

**UNIVERSIDADE FEDERAL DE SÃO CARLOS**  
**CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA URBANA**

**Identificação de Imperfeições em Pavimentos Flexíveis de  
Rodovias por Meio de Imagens de Vídeo e Aprendizado de  
Máquina Profundo**

Diego Eduardo de Vita

São Carlos – SP

2024

**UNIVERSIDADE FEDERAL DE SÃO CARLOS**  
**CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA URBANA**

**Identificação de Imperfeições em Pavimentos Flexíveis de  
Rodovias por Meio de Imagens de Vídeo e Aprendizado de  
Máquina Profundo**

Diego Eduardo de Vita

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Urbana da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de Mestre em Engenharia Urbana.

Orientação Prof. Dr. Edson Augusto Melanda

São Carlos – SP

2024

*"Dedico este trabalho a Deus, para que eu nunca me esqueça de Te agradecer por tudo,  
Senhor."*

# Agradecimentos

Ao Prof. Dr. Edson Augusto Melanda pela aceitação do desafio, confiança, orientação ao longo do caminho e pela amizade.

Aos meus pais, Marcos e Maria, e meus irmãos, Marcos e Maryana, por serem essenciais em minha vida e entenderem minhas preocupações e aflições.

À Clarissa pelo companherismo de todos os dias e pelo incentivo e apoio em todos os momentos.

Ao CPGEU pela compreensão e a todos os professores e colaboradores do PPGEU por todo conhecimento passado.

Aos membros da banca examinadora, Prof. Dr. Abimael Cereda Jr. e Prof. Dr. Diego Oliveira Martins, que tão gentilmente aceitaram participar e colaborar com esta dissertação, bem como ao membro da banca de qualificação Prof. Dr. Marcelo Monari pelas sugestões apresentadas na ocasião.

Às diversas amizades do meio técnico público e privado que fiz ou cultivei durante essa pesquisa e que de alguma maneira colaboraram com ela.

*“Porque sou eu que conheço os planos que tenho para vocês, diz o Senhor, planos de fazê-los prosperar e não de causar dano, planos de dar a vocês esperança e um futuro.”*  
*(Jeremias 29:11)*

# Resumo

Segundo dados da Confederação Nacional dos Transportes de 2023, quase 57% das rodovias brasileiras apresentam algum tipo de problema relacionado ao pavimento, isso após análise das condições de 43.843 km de rodovias estaduais e 67.659 km de rodovias federais. Em 2022, cerca de 60% de toda carga transportada em território nacional usava o sistema rodoviário, sendo que, somente entre 2019 e 2021, ocorreram 195.326 acidentes. Isso mostra a importância do sistema rodoviário, mas também, de certa forma, adverte para as condições de conservação que as rodovias devem possuir, necessitando também de celeridade e assertividade na manutenção, que está diretamente relacionada com a eficiência na avaliação das condições do pavimento. Compreende-se, portanto, o impacto da morosidade na avaliação do pavimento das estradas brasileiras. Nesse sentido, grandes esforços têm sido realizados com o objetivo de realizar uma avaliação automatizada, conferindo a celeridade que o processo necessita. Em virtude disso, técnicas computacionais de aprendizado de máquina têm sido empregadas, com treinamento de algoritmos de inteligência artificial capazes de identificar padrões através de imagens e, portanto, diagnosticar a situação encontrada nas estradas por meio das imagens analisadas. Assim, esse trabalho apresenta os conceitos ligados à estrutura, à patologias e ao gerenciamento de pavimentos. Também traz as principais literaturas à cerca das redes neurais profundas, redes convolucionais e detecção de objetos, para, então, propor a inspeção automatizada em pavimentos rodoviários utilizando algoritmos de aprendizagem profunda, por meio de imagens de vídeo. Os resultados desse trabalho mostram a identificação de patologias em 98% das imagens selecionadas para um teste de detecção real. O método escolhido para composição do *dataset* possibilitou ao modelo processar a etapa de treinamento e posteriormente ser aplicado ao teste. Além disso, a formulação do *dataset* ocorreu após a seção de detecção inicial, a qual conduziu à melhor estratégia de rotulação para o caso de detecção das panelas e fissuras. Por fim, o resultado médio de confiança baseado nas imagens detectadas mostrou-se promissor, quando é feita uma análise da dimensão do banco de imagens empregado no treinamento.

**Palavras-chave:** Análise de Imagens Baseada em Objetos; Aprendizado Profundo; Detecção Automatizada de Objetos; Pavimentos Flexíveis; Gerência de Pavimentos.

# Abstract

According to data from the National Transportation Confederation in 2023, almost 57% of Brazilian highways have some kind of pavement-related problem, after analyzing the 43.843 km of state highways and 67.659 km of federal highways. By 2022, around 60% of all cargo transported in the country used the road system. and between 2019 and 2021 alone, there were 195,326 accidents. This shows the importance of the road system, but also, in a way, warns of the conservation conditions that roads must be in, and they also need to be maintained quickly and assertively. and assertiveness in maintenance, which is directly related to efficiency in assessment of sidewalk conditions. We can therefore understand the impact of the slowness in assessing the sidewalk of Brazilian roads. In this sense, great efforts have been made to carry out an automated assessment, providing the speed that the process needs. As a result, computational machine learning techniques have been used to train artificial intelligence algorithms capable of identifying patterns through images and, therefore, diagnosing the situation found on the roads using the images analyzed. This paper therefore presents structure, pathologies and sidewalk management. It also the main literature on deep neural networks, convolutional networks and object detection, and then proposes automated inspection of road sidewalks using deep learning algorithms, by means of video images. The results of this work show the identification of pathologies in 98% of the images selected for a real detection test. The method chosen for dataset enabled the model to process the training stage and then be applied to the test. In addition, the dataset was formulated after the initial detection section, which led to the best section, which led to the best labeling strategy for the case of detecting the pans and cracks. Finally, the average confidence result based on the detected images was promising when the size of the image bank used for training is analyzed.

**Keywords:** Object-Based Image Analysis; Deep Learning; Automated Object Detection; Flexible Pavements; Pavement Management.

# Lista de ilustrações

Figura 1 – Estrutura dos pavimentos rígidos e flexíveis . . . . .	18
Figura 2 – Preenchimento de dados no MEDINA . . . . .	20
Figura 3 – Terminologia e tipo de defeitos - Fendas: Fissuras ou trincas . . . . .	21
Figura 4 – Terminologia e tipo de defeitos . . . . .	23
Figura 5 – Estrutura de um SGP para Jain et al. 2003 . . . . .	24
Figura 6 – Estrutura do SGP para o DNIT . . . . .	25
Figura 7 – Estrutura do SGP para Haas, Hudson e Zanieski . . . . .	26
Figura 8 – Processo decisório para Haas, Hudson e Zanieski . . . . .	27
Figura 9 – Faixas de variação do IRI em função do pavimento e da velocidade . . . . .	29
Figura 10 – Coeficiente de impacto em função do perfil do pavimento . . . . .	30
Figura 11 – Coeficiente de impacto em função da patologia . . . . .	30
Figura 12 – Classificação da serventia . . . . .	32
Figura 13 – Ficha de avaliação de serventia . . . . .	33
Figura 14 – Modelo do neurônio artificial . . . . .	34
Figura 15 – Operações de Max Pooling e Average Pooling . . . . .	40
Figura 16 – Arquitetura de uma rede neural convolucional . . . . .	41
Figura 17 – Filtro Kernel e Operação de Max Pooling . . . . .	41
Figura 18 – Ilustração das operações de uma rede CNN . . . . .	42
Figura 19 – Métodos de interpretação de imagens utilizados pela CV . . . . .	43
Figura 20 – Segmentação semântica e Segmentação de instâncias . . . . .	43
Figura 21 – Exemplos reais de detecção do modelo AttentionNet . . . . .	45
Figura 22 – Funcionamento do modelo G-CNN . . . . .	46
Figura 23 – Funcionamento do modelo YOLO . . . . .	47
Figura 24 – Cálculo de IoU . . . . .	48
Figura 25 – Valores de IoU . . . . .	48
Figura 26 – Fluxograma descrevendo as etapas dos materiais e métodos . . . . .	51
Figura 27 – Câmera de ação GoPro Hero 12 Black . . . . .	52
Figura 28 – Configuração adotada para coleta de dados . . . . .	53
Figura 29 – Caixa estanque gp-wtp-901 e suporte tipo tripé utilizados . . . . .	54
Figura 30 – Instalação do tripé na estrutura veicular . . . . .	54
Figura 31 – Instalação do cabo de fitilho de polipropileno . . . . .	55
Figura 32 – Trajeto sendo realizado apenas com a caixa estanque, sem a câmera . . . . .	55
Figura 33 – Trajeto realizado e pontos em que foi avaliada a estabilidade . . . . .	56
Figura 34 – Trajeto realizado e pontos de referência do trajeto . . . . .	56
Figura 35 – Coleta de imagens do dataset de treinamento . . . . .	57
Figura 36 – Estação de controle automático A701 . . . . .	57

Figura 37 – Gráfico de umidade e chuva acumulada no dia 10/05/2024 . . . . .	58
Figura 38 – Banco de imagens . . . . .	59
Figura 39 – Exemplificação do agrupamento de imagens selecionadas para rotulação	59
Figura 40 – Exemplo de rotulação para os treinos 2 e 3 . . . . .	61
Figura 41 – Arquivo do tipo txt gerado para cada imagem . . . . .	62
Figura 42 – Configuração de processamento escolhida . . . . .	63
Figura 43 – Rede YOLO V8 e suas camadas . . . . .	64
Figura 44 – Treino 3 - processamento . . . . .	64
Figura 45 – Exemplo de detecção inicial alcançada . . . . .	65
Figura 46 – Perda das caixas delimitadoras para os treinos 2, 3 e 4 . . . . .	66
Figura 47 – Probabilidade predita de classe para os treinos 2, 3 e 4 . . . . .	66
Figura 48 – Precisão geral e com 50% de confiança para os treinos 2, 3 e 4 . . . . .	67
Figura 49 – Recall e precisão com confiança de 95% para os treinos 2, 3 e 4 . . . . .	67
Figura 50 – Detecção para os treinos 2, 3 e 4 em imagens específicas . . . . .	68
Figura 51 – Matriz de confusão para o treino 2 . . . . .	68
Figura 52 – Matriz de confusão para o treino 3 . . . . .	69
Figura 53 – Matriz de confusão para o treino 4 . . . . .	69
Figura 54 – Configuração adotada para detecção . . . . .	71
Figura 55 – Verificação de estabilidade e instalação da caixa estanque e da câmera de ação . . . . .	72
Figura 56 – Trajeto entre as cidades de Votorantim e Piedade em São Paulo . . . . .	73
Figura 57 – Locais de início (a), término (b) e parada para resfriamento (b) e (c) .	73
Figura 58 – Condições climáticas em Votorantim, SP . . . . .	74
Figura 59 – Banco de imagens da 2ª etapa . . . . .	75
Figura 60 – Banco de imagens selecionadas . . . . .	75
Figura 61 – Exemplo de imagens selecionadas . . . . .	76
Figura 62 – Configuração da ferramenta CVAT . . . . .	76
Figura 63 – Rotulação para o treino . . . . .	77
Figura 64 – Arquivos do tipo txt para cada patologia . . . . .	77
Figura 65 – Nomeação das pastas . . . . .	78
Figura 66 – Nomeação dos arquivos . . . . .	79
Figura 67 – Nomeação das pastas no pré-processamento . . . . .	79
Figura 68 – Exemplos dos efeitos Blur e Brightness . . . . .	80
Figura 69 – Processamento do treino . . . . .	81
Figura 70 – Trajeto para coleta de dados da detecção . . . . .	82
Figura 71 – Condições climáticas em Campinas, SP . . . . .	82
Figura 72 – Registro do percurso de detecção na Rodovia Adalberto Pazam . . . . .	83
Figura 73 – Processamento da detecção . . . . .	85
Figura 74 – Perda das caixas delimitadoras para o treinamento da seção 3.2.5 . . . . .	87

Figura 75 – Probabilidade predita de classe para o treinamento da seção 3.2.5 . . .	88
Figura 76 – Precisão geral e com confiança de 50% para o treinamento da seção 3.2.5	88
Figura 77 – Recall para o treinamento da seção 3.2.5 . . . . .	89
Figura 78 – Perda de distribuição focalizada para o treinamento da seção 3.2.5 . . .	89
Figura 79 – Matriz de confusão do treinamento para o treinamento da seção 3.2.5 .	90
Figura 80 – Exemplo de detecção gerada pelo modelo treinado . . . . .	93

# Lista de tabelas

Tabela 1 – Anexo A do DNIT - resumo dos defeitos nos pavimentos . . . . .	22
Tabela 2 – Formatação e composição dos minidatasets . . . . .	63
Tabela 3 – Avaliação do treinamento . . . . .	70
Tabela 4 – Composição do dataset . . . . .	81
Tabela 5 – Identificação das patologias encontradas nas imagens do teste . . . . .	84
Tabela 6 – Resumo dos resultados da fase de pós processamento . . . . .	91

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>13</b>
<b>1.1</b>	<b>Justificativa</b>	<b>14</b>
<b>1.2</b>	<b>Objetivo</b>	<b>15</b>
<b>2</b>	<b>REVISÃO DA LITERATURA</b>	<b>17</b>
<b>2.1</b>	<b>Pavimentos</b>	<b>17</b>
2.1.1	Estrutura dos pavimentos	17
2.1.2	Dimensionamento da estrutura dos pavimentos	18
2.1.3	Patologias encontradas nos pavimentos flexíveis e semi-flexíveis	20
<b>2.2</b>	<b>Sistema de Gerenciamento de Pavimentos</b>	<b>24</b>
2.2.1	Avaliação de pavimentos	28
2.2.1.1	Avaliação do desempenho funcional	28
2.2.1.2	Levantamento da irregularidade longitudinal	28
2.2.1.3	Levantamento das patologias de superfície	31
<b>2.3</b>	<b>Identificação automática de objetos</b>	<b>33</b>
2.3.1	Redes neurais artificiais	34
2.3.2	Tipos de aprendizado	36
2.3.3	Aprendizado de máquina	37
2.3.4	Aprendizado profundo	39
2.3.5	Redes convolucionais	39
2.3.6	Visão Computacional	42
2.3.7	Detecção de Objetos	44
2.3.7.1	YOLO	47
<b>3</b>	<b>MATERIAIS E MÉTODOS</b>	<b>50</b>
<b>3.1</b>	<b>Detecção inicial</b>	<b>52</b>
3.1.1	Configuração e instalação dos equipamentos	52
3.1.2	Coleta de dados para o treinamento	56
3.1.3	Preparação do banco de imagens do treinamento	58
3.1.4	Seleção de parte das imagens do banco de imagens	60
3.1.5	Rotulação	60
3.1.6	Preparação do minidataset	62
3.1.7	Treinamento	63
3.1.8	Avaliação do treinamento	65
<b>3.2</b>	<b>Detecção final</b>	<b>71</b>
3.2.1	Configuração, instalação e coleta de dados para treinamento	71

3.2.2	Preparação do banco de imagens e rotulação . . . . .	74
3.2.3	Pré-processamento . . . . .	77
3.2.4	Preparação do dataset . . . . .	80
3.2.5	Treinamento . . . . .	81
3.2.6	Coleta de dados para detecção . . . . .	82
3.2.7	Detecção . . . . .	84
<b>4</b>	<b>RESULTADOS . . . . .</b>	<b>87</b>
<b>4.1</b>	<b>Avaliação do treinamento . . . . .</b>	<b>87</b>
<b>4.2</b>	<b>Avaliação da detecção . . . . .</b>	<b>90</b>
<b>5</b>	<b>CONSIDERAÇÕES FINAIS . . . . .</b>	<b>94</b>
<b>5.1</b>	<b>Principais Resultados . . . . .</b>	<b>94</b>
<b>5.2</b>	<b>Trabalhos futuros . . . . .</b>	<b>95</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>96</b>

# 1 Introdução

As rodovias são componentes essenciais da infraestrutura de transportes no Brasil, desempenhando um papel crucial na competitividade do país. Em 2018, a Confederação Nacional da Indústria (CNI) enfatizou, em seu relatório anual, Proposta da Indústria para as Eleições: Grandes obras paradas: como enfrentar o problema?, a necessidade de investir em infraestrutura, por exemplo de rodovias, como um dos desafios para aumentar a competitividade brasileira. Esse investimento, à ocasião, foi visto como parte primordial de um conjunto de obras de infraestrutura que poderiam impulsionar o desenvolvimento econômico efetivo do Brasil (CNI, 2018).

Pouco antes, em 2017, com o mesmo propósito, foi lançado pelo Governo Federal Brasileiro, como parte da estratégia para melhoria da infraestrutura brasileira, o programa "Agora, é Avançar".

Nesse programa foram listadas 7.439 obras de infraestrutura, incluindo as obras rodoviárias necessárias ao desenvolvimento brasileiro. Assim, também faziam parte do escopo as obras atrasadas ou as paralisadas, as quais representavam 18,5% do total, conforme dados do Ministério do Planejamento Desenvolvimento e Gestão (MPDG, 2017).

Diretamente relacionados à questão competitiva, estão os custos de transporte, que são diretamente afetados pelo estado de conservação das rodovias e impactam os gastos com manutenção de veículos, consumo de combustível e tempo de deslocamento.

Muito antes, em 2001, uma avaliação, composta pelos pareceres do Departamento Nacional de Estradas de Rodagem (DNER) e pela Empresa Brasileira de Planejamento de Transportes (Gepoit), mostrou os impactos que uma estrada degradada poderia trazer. Por exemplo, aumentar em até 58% o consumo de combustíveis, 38% os gastos de manutenção dos veículos, 50% o índice de acidentes e 100% o tempo gasto nas viagens (CNT, 2001 apud BARTHOLOMEU; CAIXETA, 2008).

Em 2021, uma pesquisa elaborada pela CNT analisou o aumento de custo operacional para o transporte rodoviário de carga, para o caso do transporte em pavimento rodoviário classificado como péssimo, chegando ao acréscimo de até 91,5% no custo de operação (CNT, 2021).

As causas de deterioração das rodovias são multifacetadas, podendo incluir erros no dimensionamento das camadas, falhas durante a construção, falta de manutenção preventiva e fiscalização inadequada do tráfego de veículos (CNT, 2021).

Tão importante quanto à questão econômica encontra-se a questão de segurança das estradas brasileiras. Segundo CNT (2021), somente entre 2019 e 2021, ocorreram 195.326

acidentes nas rodovias federais, dos quais 160.383 resultaram em vítimas, envolvendo 222.120 feridos e 16.010 óbitos em geral. Os acidentes nas rodovias podem estar associados a vários fatores, como aos comportamentais, às especificidades dos veículos, às condições climáticas, mas também a características e condições das estradas, nos quesitos geometria, pavimentação e sinalização.

À vista disso, os Sistemas de Gerenciamento de Pavimentos (SGP's) tem como principal finalidade fornecer ao usuário, através das condições do pavimento, um transporte seguro, compatível e econômico.

Logo os SGP's, conceituados como um ferramenta de gestão, ganham mais espaço, disponibilizando conjunto de informações assertivas para tomada de decisão sobre as condições das malhas viárias, as metodologias construtivas empregadas, os materiais e projetos utilizados na construção, as manutenções e reabilitações dos pavimentos, as informações sobre o tráfego de veículos, as projeções referentes à durabilidade de todos os trechos, às especificações quanto a um plano prioritário de manutenção, a lista de empresas capacitadas para manutenções, etc (SHAHIN, 2005).

## 1.1 Justificativa

Bertollo (1997) destaca que embora os SGP's objetivem formar um sistema de gerenciamento que proporcione um transporte seguro, confortável e econômico ao usuário, almejam, adicionalmente, alcançar a melhoria das condições do pavimento e a redução dos custos de manutenção, de reabilitação e de operação dos veículos.

As atividades relacionadas a um SGP's estão comumente relacionadas à área de planejamento, podendo ser divididas em: sistemas de referência, avaliação do trecho pavimentado, estabelecimento de prioridades e elaboração de um plano plurianual de investimento, sendo a atividade de avaliação de alta complexidade (DNIT, 2011).

Ainda nesse mesmo sentido, Almeida (2018) destacou que a avaliação de pavimentos no Brasil ainda não utiliza técnicas avançadas e não possibilita um diagnóstico eficiente, por consequência, esse trabalho ainda possui alto nível de empirismo no país.

Para Branco (2016), a inspeção visual em rodovias, realizada por técnicos, gera um trabalho acentuado e até mesmo tedioso, sujeito a alta variabilidade de diagnósticos e perda de informações importantes, além disso, a utilização do método tradicional de inspeção é perigoso e caro.

Da mesma maneira, diversas tarefas, que demandam esforço interpretativo, alto índices de assertividades e eficiência executiva, vêm obtendo apoio computacional para sua execução.

Assim, o progresso tecnológico tem assistido à diversas tarefas, por meio do de-

envolvimento de poderosas ferramentas de automatização, através da introdução do aprendizado de máquina às etapas dos processos escolhidos, acelerando e sistematizando os resultados, por exemplo, nas tarefas de classificação do uso do solo e mapeamento de locais de acidentes de trânsito (WANLI et al., 2018).

Nesse contexto, os métodos tradicionais de inspeção em pavimento são atividades que podem ser aperfeiçoadas, pois são demoradas, perigosas e dispendiosas, além de muitas vezes resultarem em análises subjetivas e com grande variabilidade. Diante disso, a adoção de novas tecnologias, como a análise de imagens por meio de aprendizado de máquina, é apresentada como uma alternativa promissora para superar esses desafios e impulsionar os SGP's na etapa de avaliação.

Para Hilt (2017), os progressos alcançados na criação dos Big Datas em consonância com as Inteligências Artificiais (IA's) formaram uma conjectura que viabiliza o estudo de novos cenários, substancialmente em atividades de assitência. Nesse sentido o aprendizado de máquina (AM) que é uma das dimensões da IA, possibilita que os algoritmos melhorarem sem intervenção humana (HILT, 2017).

Portanto, o desenvolvimento de um método de identificação de imperfeições em pavimentos por meio do apredinzagem de máquina, pode contribuir na avaliação de pavimentos e ajudar no aperfeiçoamento dos SGP's, auxiliando para a redução de acidentes, dos custos com transportes e dos impactos ambientais, pelo fato de poderem gerar levantamentos das condições das estradas mais assertivas e eficientes, contribuindo para o desenvolvimento efetivo brasileiro.

## 1.2 Objetivo

O objetivo geral desse trabalho é treinar um algoritmo para identificação automatizada de fissuras e painelas em pavimentos flexíveis, por meio de imagens de vídeo, utilizando técnicas de Aprendizado de Máquina.

Os objetivos específicos deste trabalho são:

- Montar um dataset para realizar o treinamento do algoritmo responsável pela identificação das imperfeições;
- Indicar por meio de uma etapa de detecção inicial a forma de rotulação mais eficaz para detecção de fissuras e painelas;
- Quantificar a detecção alcançada do modelo para um banco de imagens previamente escolhido;
- Apresentar a métrica de confiança alcançada pelo modelo treinado na detecção final;

- Disponibilizar a base de dados rotulados publicamente.

## 2 Revisão da Literatura

### 2.1 Pavimentos

O pavimento é a conformação ordenada de camadas de materiais compactados, com a finalidade de atender estruturalmente e operacionalmente aos esforços providos do tráfego, assistindo aos usuários, através das melhorias das condições de rolamento, com conforto, economia e segurança, tendo também como critérios a durabilidade e o custo. Seu dimensionamento busca projetar uma estrutura na qual as camadas repassem esforços dissipados ao terreno natural, impossibilitando problemas intrínsecos ou na base que a apoia (BALBO, 2007).

#### 2.1.1 Estrutura dos pavimentos

A estrutura do pavimento é apoiada sobre a superfície final de terraplenagem e visa resistir, mecanicamente, aos esforços providos do tráfego, do clima e fornecer conforto, segurança e economia aos usuários (BRANCO; PEREIRA; SANTOS, 2006).

A última camada do pavimento indica o tipo de rigidez dessa, podendo ser conhecida como asfáltica/flexível (em que foram utilizados componentes derivados do petróleo) ou cimentícia/rígida (indicando a utilização de concreto de cimento).

Na estruturação das tipologias mencionadas, conforme pode ser visto na Figura 1, para pavimentação flexível, objeto de estudo desse trabalho, observa-se tradicionalmente a seguinte subdivisão: revestimento asfáltico, base, sub-base e reforço do sub-leito, enquanto na tipologia rígida, placas de concreto, armadas ou não, apoiadas sobre uma sub-base e sobre o sub-leito, inexistindo, entretanto, para ambos os casos, a obrigatoriedade de todos os níveis (BERNUCCI et al., 2006).

O revestimento é a parte que possui contato direto com as rodas dos veículos e suporta os esforços diretamente aplicados, transmitindo-os com perda gradual de intensidade às camadas inferiores. Além disso, pode ser composto por camadas de rolamento e pela camada de ligação, chamada *binder* (BERNUCCI et al., 2006).

A estruturação subjacente, base, sub-base e reforço do sub-leito, por meio de seus materiais componentes, das espessuras constituintes e da metodologia executiva fornecem as tensões e deformações máximas que podem ser suportadas pela estrutura (MEDINA, 1997).

A base poderá ser composta por misturas de brita graduada tratada ou não, por misturas de solos com agregados e solos estabilizados quimicamente ou naturalmente,

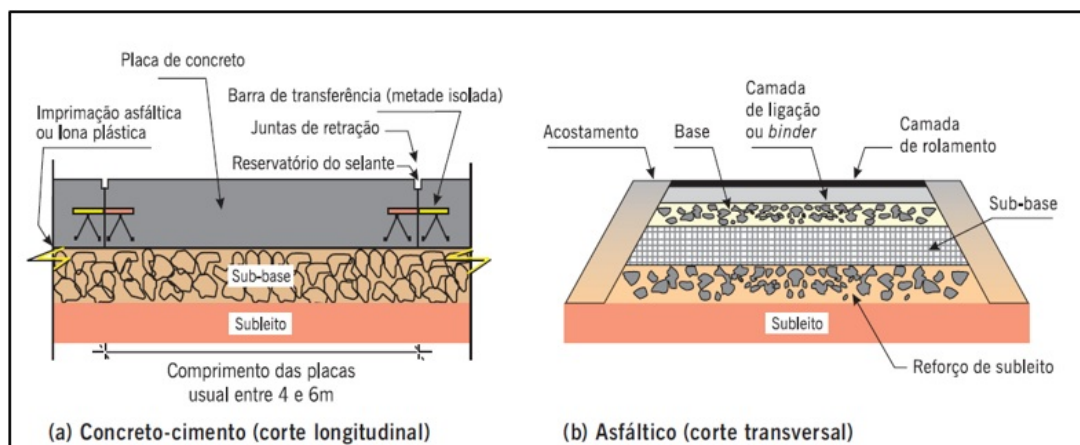


Figura 1 – Estrutura dos pavimentos rígidos e flexíveis

Fonte: Extraído de [Bernucci et al. \(2006\)](#)

possuindo a finalidade de suportar as solicitações transmitidas pelo revestimento, bem como atuar na drenagem do pavimento ([BALBO, 2007](#); [BRANCO](#); [PEREIRA](#); [SANTOS, 2006](#)).

Já a sub-base possui função de suportar os esforços oriundos da base, atuando diretamente na redução da espessura dessa. Há possibilidade de utilização dos mesmos materiais provindos da base para sua execução, com a vantagem de redução no consumo dos aglomerantes ([BRANCO](#); [PEREIRA](#); [SANTOS, 2006](#); [BALBO, 2007](#)).

Nos casos em que o sub-leito possui baixa capacidade de suporte, existe a necessidade de ser executado seu reforço, aumentando sua capacidade de resistência e com isso a redução da espessura das camadas superiores, modificando os custos de execução ([BRANCO](#); [PEREIRA](#); [SANTOS, 2006](#); [BALBO, 2007](#)).

### 2.1.2 Dimensionamento da estrutura dos pavimentos

À determinação dos materiais das camadas e das suas respectivas espessuras em função do número de solicitações atribuiu-se o nome de dimensionamento do pavimento asfáltico, pela qual se buscou produzir uma estrutura que suporte aos esforços resultantes do tráfego e a mantenha com as exigências funcionais de vida útil ([FRANCO, 2007](#)).

No Brasil, por muito anos, foi a metodologia empírica, atualmente em desuso, a principal técnica para dimensionamento de pavimentos. Essa, apesar de muito utilizada, é considerada insatisfatória ao tentar-se provar o comportamento estrutural, principalmente porque seus resultados não possibilitam, com fidedignidade, ser estendidos para diferentes tipos de materiais, locais e solicitações de tráfego ([FRANCO, 2014](#)).

A metodologia empírica foi desenvolvida na década de 60 e introduzida pelo Departamento Nacional de Estradas de Rodagem (DNER) em 1966, passando pela alteração do número N (que representa o tráfego) em 1981 ([MOTTA](#); [MEDINA, 2006](#)).

Como primeira etapa, são determinados os valores de capacidade de suporte, através do ensaio *California Bearing Ratio* (CBR), ou em português, Índice de Suporte Califórnia (ISC), e do número equivalente de operações de um eixo, número N, durante o período de projeto admitido de operação para via, sendo aplicado sobre esse, um fator climático regional, o qual considera a variação da capacidade de suporte dos materiais devido a inconstância de umidade (DNER, 1981).

A determinação da altura total do pavimento, de acordo com o método, é realizada por meio da utilização de um ábaco que correlaciona a capacidade de suporte com o número N, sendo que a determinação da altura das outras camadas constituintes deve obedecer um processo iterativo de resolução de inequações, até que se obtenha a altura de cada camada da estrutura do pavimento. Ressalta-se que existem diversas advertências de acordo com essa técnica, como altura mínima das camadas e a capacidade de suporte mínima para o sub leito.

Outra forma para realização do dimensionamento da estrutura do pavimento é a que ocorre por meio da análise mecanicista empírica, a qual pode utilizar diferentes teorias estruturais, como a Teoria das Multicamadas Elásticas (TME), ou, algum outro tipo, como o Método de Elementos Finitos (MEF) por exemplo, permitindo sua aplicação a diferentes tipos de carregamentos, materiais e condições ambientais (TORQUATO E SILVA et al., 2021).

Esse método analisa questões ligadas à mecânica comportamental das camadas que integram o pavimento, correlacionando os esforços provindos do tráfego às espessuras das camadas e às propriedades dos materiais utilizados, verificando as tensões e deformações, as quais a estrutura ficará submetida (SANTOS et al., 2020). Em vista disso, essa análise tem por fim dimensionar a estrutura observando o equilíbrio dos esforços imputados pelo tráfego, considerando também o número de repetições e a capacidade de suporte dos materiais.

Um dos principais parâmetros da análise mecanicista empírica é o módulo de resiliência dos materiais, ou, módulo resiliente, que é a relação entre a tensão aplicada e a deformação elástica no material observada (VENESCAU, 2020).

Assim, com o auxílio computacional e o conhecimento matemático do estado de tensão e de deformação de locais da estrutura, é possível realizar simulações de desempenho da estrutura que se pretende construir (SANTOS et al., 2020).

No Brasil, durante os últimos anos, foi implementado o novo Método de Dimensionamento de Pavimentos (MEDINA), que utiliza uma abordagem mecanicista empírica para análise do comportamento estrutural (SANTIAGO et al., 2018), o MEDINA foi desenvolvido como software computacional e realiza a verificação da estrutura se baseando nos conceitos da TME (VENESCAU, 2020).

Complementa esse software o programa AEMC, que faz a análise das camadas elásticas, utilizando para isso a estrutura de integração de cálculo de Gauss-Laguerre para cálculo das equações integrais (CALÇAS et al., 2021).

Para utilização do MEDINA é necessário determinar a quantidade de camadas que será composta a estrutura e indicar seus respectivos materiais, sendo que o software disponibiliza uma biblioteca de materiais que podem ser utilizados, além de indicar o tipo de revestimento apropriado em função do número de solicitações (N) o qual será submetido. Outra possibilidade permitida é a escolha do tipo de via, alteração do número N e período de projeto (CALÇAS et al., 2021), conforme apresentado na Figura 2.

CAMADA	DESCRIÇÃO DO MATERIAL	TIPO	ESPESSURA (cm)	MÓDULO (MPa)	COEFICIENTE DE POISSON
>> 1 <<	CONCRETO ASFÁLTICO	RJ CAP 30/45 #12,5mm Sepebe	10,0	9000	0,30
2	MATERIAL GRANULAR	Brita Graduada - Gnaisse C3	20,0	381	0,35
3	SOLO FINO, SILTOSO OU ARGILOSO	Solo Argiloso LG(1)	20,0	250	0,45
SL	SUBLEITO	Solo Silteco NS'	8,0	189	0,45

**EIXO PADRÃO RODOVIÁRIO**

**DADOS DO TRÁFEGO**

Tipo de Via: Sistema Arterial Primário

VMD (1º ano): 1370

FV: 1,000

N anual (1º ano): 5,00e+05

% Veículos na faixa de projeto: 100

N Anual da faixa: 5,00e+05

Taxa de crescimento (%): 0,0

Período de projeto (anos): 10

N Total: 5,00e+05

Figura 2 – Preenchimento de dados no MEDINA

Fonte: Extraído de Calças et al. (2021)

Posto isso, no Brasil, o MEDINA é o atual método de dimensionamento, sendo amplamente utilizado, tanto pelos projetista como por outras partes interessadas.

### 2.1.3 Patologias encontradas nos pavimentos flexíveis e semi-flexíveis

A norma número 005 de 2003 do Departamento Nacional de Infraestrutura de Transportes (DNIT, 2003) discorre detalhadamente a respeito da identificação visual de defeitos na superfície dos pavimentos, uma vez que cada tipo concorre para estabelecer

sua parcela quanto ao indicador da qualidade da superfície avaliada, através do Índice de Gravidade Global (IGG).

Dito isso, a Tabela 1 apresenta os tipos de fendas que são encontrados nas superfícies dos pavimentos. Dessa forma, as fendas são apresentadas como aberturas nos pavimentos e subdivididas em: fissuras, quando se fazem perceptíveis a distância menor que 1,5 m; ou trincas que são maiores que as fissuras. Quanto à gravidade, são classificadas como do tipo 1 (com abertura de até 1 mm); tipo 2 (superiores a 1 mm) e tipo 3 (superiores a 1 mm com desagregação ou superiores a 1 mm com erosão das bordas) (BERNUCCI et al., 2006).

Incrementalmente, as trincas podem se apresentar isoladas como transversais curtas (TTC), transversais longas (TTL), longitudinais curtas (TLC), longitudinais longas (TLL) e de retração (TRR), ou se apresentarem agrupadas, em blocos, quando possuem uniformidade geométrica (TB), em blocos com erosão das bordas (TBE), em couro de jacaré, quando não possuem uniformidade geométrica (J), ou em couro de jacaré com erosão das bordas (JE), conforme apresentado na Figura 3 .



Figura 3 – Terminologia e tipo de defeitos - Fendas: Fissuras ou trincas  
Fonte: Adaptado de DNIT (2003a apud BERNUCCI et al., 2006)

Tabela 1 – Anexo A do DNIT - resumo dos defeitos nos pavimentos

Fendas				Codificação	Classe das Fendas		
Fissuras				FI	-	-	-
Trincas no revestimento geradas por deformação permanente excessiva e/ou decorrentes do fenômeno de fadiga	Trincas Isoladas	Transversais	Curtas	TTC	FC-1	FC-2	FC-3
			Longas	TTL	FC-1	FC-2	FC-3
		Longitudinais	Curtas	TLC	FC-1	FC-2	FC-3
			Longas	TLL	FC-1	FC-2	FC-3
	Trincas Interligadas	“Jacaré”	Sem erosão acentuada nas bordas das trincas	J	-	FC-2	-
			Com erosão acentuada nas bordas das trincas	JE	-	-	FC-3
Trincas no revestimento não atribuídas ao fenômeno de fadiga	Trincas Isoladas	Devido à retração térmica ou dissecação da base (solo-cimento) ou do revestimento		TRR	FC-1	FC-2	FC-3
	Trincas Interligadas	“Bloco”	Sem erosão acentuada nas bordas das trincas	TB	FC-1	FC-2	FC-3
			Com erosão acentuada nas bordas das trincas	TBE	-	-	FC-3
Outros Defeitos					Codificação		
Afundamento	Plástico	Local	Devido à fluência plástica de uma ou mais camadas do pavimento ou do subleito	ALP			
		da Trilha	Devido à fluência plástica de uma ou mais camadas do pavimento ou do subleito	ATP			
	De Consolidação	Local	Devido à consolidação diferencial ocorrente em camadas do pavimento ou do subleito	ALC			
		da Trilha	Devido à consolidação diferencial ocorrente em camadas do pavimento ou do subleito	ATC			
<b>Ondulação/Corrugação</b> - Ondulações transversais causadas por instabilidade da mistura betuminosa constituinte do revestimento ou da base					O		
<b>Escorregamento</b> (do revestimento betuminoso)					E		
<b>Exsudação</b> do ligante betuminoso no revestimento					EX		
<b>Desgaste</b> acentuado na superfície do revestimento					D		
<b>“Painéis”</b> ou buracos decorrentes da desagregação do revestimento e às vezes de camadas inferiores					P		
<b>Remendos</b>				Remendo Superficial		RS	
				Remendo Profundo		RP	
<b>Nota 1:</b> Classe das trincas isoladas							
As trincas interligadas são classificadas como FC-3 e FC-2 caso apresentem ou não erosão nas bordas.							
As trincas interligadas são classificadas como FC-3 e FC-2 caso apresentem ou não erosão nas bordas.							
<b>Nota 2:</b> Classe das trincas interligadas							
As trincas interligadas são classificadas como FC-3 e FC-2 caso apresentem ou não erosão nas bordas.							

Já na Figura 4, são apresentados outros defeitos, que são: os afundamentos (A); as corrugações e ondulações (O); as exsudações (EX); os desgastes ou desagregações (D); as panelas ou buracos (P); e os remendos (R).

Os afundamentos subdividem-se em plásticos ou por consolidação (vide Figura 4). Plásticos são aqueles em que ocorreu a fluência plástica de uma ou mais camadas do pavimento, sendo locais (ALP) ou da trilha de rodas (ATP). Os por consolidação são defeitos causados pela consolidação diferencial de uma ou mais camadas do pavimento ou subleito sem erguimento, podendo ser locais (ALC), quando registrados em trechos menores que 6 metros de comprimento; e afundamento de consolidação de trilho de roda (ATC), quando maiores em comprimento e na região longitudinal das rodas (DNIT, 2003a).



Figura 4 – Terminologia e tipo de defeitos

Fonte: Adaptado DNIT (2003a apud BERNUCCI et al., 2006)

Quanto às corrugações e ondulações, representadas pela letra O, ambas aparecem no sentido transversal ao da pista (DNIT, 2003a), sendo as corrugações defeitos da ordem de dezenas de centímetros entre cristas, enquanto as ondulações são defeitos da ordem de metros e se sucedem pela consolidação diferencial do subleito (BERNUCCI et al., 2006).

A exsudação (EX) é a problemática relacionada à migração do ligante betuminoso pelo revestimento, com aparecimento do excesso migrante, através de manchas escuras, à superfície do pavimento; enquanto o desgaste (D), por sua vez, é a retirada violenta de agregados da superfície devido aos esforços transmitidos pelo tráfego, em que se pode observar aspereza superficial; e o escorregamento (E) é o deslocamento do pavimento em relação a camada inferior com surgimento de forma "meia-lua" (vide Figura 4) (DNIT, 2003a).

O deslocamento das camadas com surgimento de uma cavidade é a patologia conhecida como panela (P) (vide Figura 4), a qual pode possuir inúmeras causas, inclusive a falta de aderência entre camadas subjacentes.

Já o remendo, que está ligado à atividade de manutenção das vias para correção de panelas, orifícios ou quaisquer outras depressões, gera irregularidades longitudinais com conseqüente desconforto ao usuário, sendo, portanto, considerado uma patologia (ADLINGE; GUPTA, 2013).

## 2.2 Sistema de Gerenciamento de Pavimentos

Na opinião de Vasquez (2011), os administradores de estradas podem economizar uma grande quantidade de dinheiro ao implementarem os SGP's, concluindo que as boas estradas custam menos ao longo do tempo.

Jain et al. (2003) menciona que um SGP demanda um sistema que forneça apoio à tomada de decisão, citando, por exemplo, softwares capazes de facilitar a preparação, análise, exibição e gerenciamento de dados geográficos conforme a Figura 5.

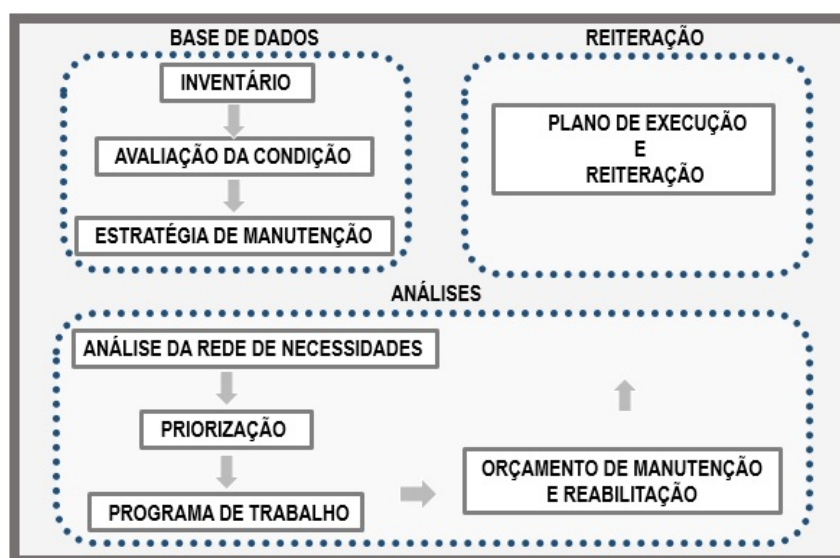


Figura 5 – Estrutura de um SGP para Jain et al. 2003  
Fonte: Adaptado de Jain et al. (2003)

Para Ragnoli, Blasiis e Benedetto (2018), os SGP's podem informar sobre trabalhos específicos, por exemplo, à respeito de manutenções prioritárias, informando o local a ser reparado, como deverá ser realizada a manutenção e quando essa deverá ocorrer (RAGNOLI; BLASIIS; BENEDETTO, 2018).

Para Ahmed (2013), a maximização do benefício para a sociedade ou maximização dos recursos disponíveis advêm da utilização de um sistema de gerenciamento e manutenção do pavimento.

Segundo o manual de gerência de pavimento, os SGP's são constituídos por elementos internos e externos. Os elementos internos se referem ao planejamento, ao projeto, à construção e à manutenção, enquanto os externos, às diretrizes políticas e administrativas, aos recursos orçamentários e à coleta e inserção de dados necessários à alimentação e retroalimentação do sistema de acordo com a Figura 6 (DNIT, 2011).

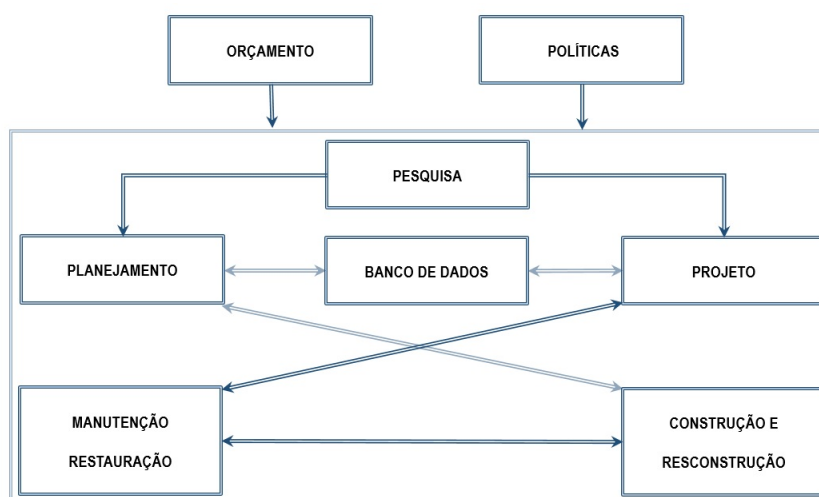


Figura 6 – Estrutura do SGP para o DNIT

Fonte: Adaptado de DNIT (2011)

As atividades essenciais de um SGP são: sistemas de referência, avaliação do trecho pavimentado, estabelecimento de prioridades e elaboração de um plano plurianual de investimento.

Pelo sistema de referência, entende-se estabelecer fisicamente a nomenclatura em relação aos trechos da rede rodoviária, podendo ainda ser subdivididos. Na atividade de avaliação, os dados devem ser coletados periodicamente, permitindo que as condições estruturais e operacionais sejam sempre verificadas (DNIT, 2011).

A avaliação, atividade de alta complexidade, deve ser subdividida no histórico em: implantação, manutenção e melhoramentos, orografia da região, características regionais, condições estruturais, condições funcionais e do tráfego (DNIT, 2011).

Ainda nesse sentido, toda implementação do SGP tem por finalidade direcionar os gestores às prioridades que causem intervenções na malha rodoviária, ou seja, o estabeleci-

mento racional de prioridades e, além disso, um plano de orçamentário customizado de longo prazo (plano plurianual de investimento).

Haas, Hudson e Zaniewski (1994 apud BERTOLLO, 1997) também informaram que os componentes de um SGP são formados pelos componentes externos e internos, embora, aos internos, também exista o componente de pesquisa, Figura 7.

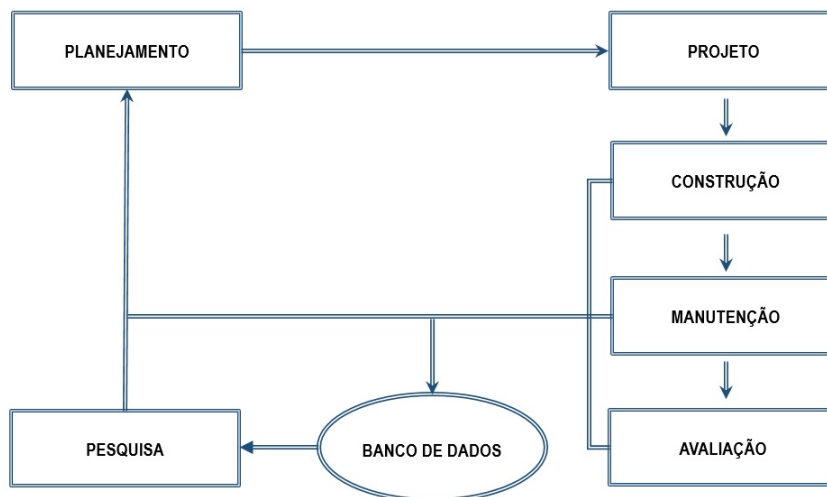


Figura 7 – Estrutura do SGP para Haas, Hudson e Zanieski

Fonte: Adaptado de Haas, Hudson e Zaniewski (1994 apud BERTOLLO, 1997)

Quanto ao processo de tomada de decisão, os SGPs podem ser subdivididos em dois níveis: rede e projeto. Nos processos de rede estão as atividades de planejamento, programação e orçamento, enquanto nas atividades de projeto estão as de dimensionamento, construção e manutenção (BERTOLLO, 1997).

Oliveira (2013) descreveu que via de regra, as informações em nível de rede são superficiais e pobres em detalhes, sendo de toda a malha rodoviária, contudo, é o nível em que as decisões administrativas são tomadas.

Na atividade de planejamento uma investigação pormenorizada é realizada e tudo deverá ser preparado, o que implicará de um banco de dados com as características básicas da estrutura do pavimento, do monitoramento contínuo, do modelo de previsão de durabilidade e da filosofia estratégica de investimento em função do custo da manutenção e reabilitação (WILHEIM, 1978).

A atividade de programação é tida como o conjunto de ações preliminares que serão executadas pela equipe gestora, através de uma análise preliminar dos dados encontrados no banco de dados (OLIVEIRA, 2013), enquanto na de orçamento, segundo Gadelha (2011), todas as obras devam ser estudadas como um todo, com levantamentos detalhados, com estudos de soluções, nas quais os problemas deverão ser criticados, solucionados e equalizados sobre a ótica do custo, da quantidade e da qualidade, com projetos e especificações, inclusive com o estudo de alternativas relacionadas aos processos construtivos

e com uma programação inicial de execução.

A atividade de dimensionamento, referente à tomada de decisão quanto nível de projeto, é a elaboração dos memoriais de cálculo e todo o detalhamento necessário para cada região à qual será realizada a construção, manutenção ou reabilitação (OLIVEIRA, 2013). As atividades de manutenção são divididas em preventivas e corretivas, sendo essas para corrigir um defeito que já apareceu, enquanto as de manutenção para reduzir a taxa de deterioração que foi iniciada. Adicionalmente, o pavimento também necessita de atividades que prolonguem sua vida útil, atividade de reabilitação, a qual tem a finalidade de correção mais efetiva, afim de fornecer reforço, adaptação ou recuperação (ALVES, 2016).

Na Figura 8, é apresentado o processo de tomada de decisão dos SGPs baseado no nível de rede e de projeto (BERTOLLO, 1997).

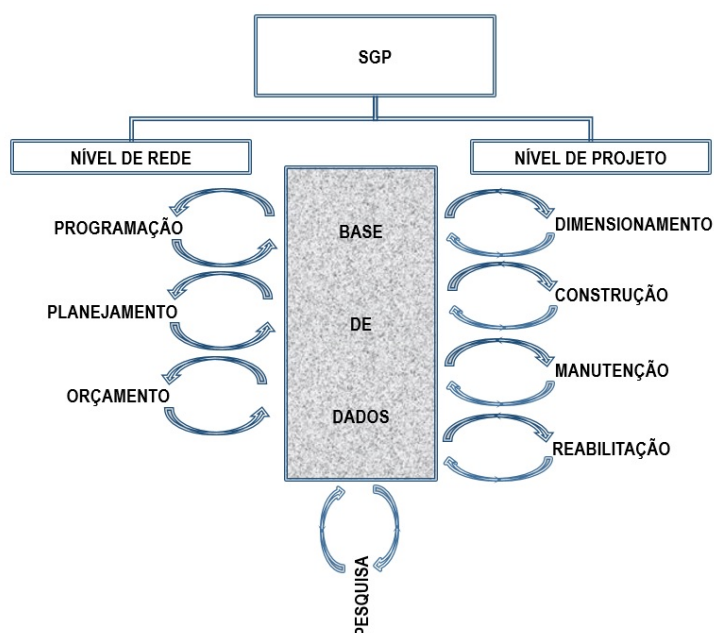


Figura 8 – Processo de tomada de decisão de um SGP segundo Haas, Hudson e Zanieski  
Fonte: Adaptado de Haas, Hudson e Zaniewski (1994 apud OLIVEIRA, 2013)

Para Fernandes Júnior, Oda e Zerbini (2006), a aptidão de um determinado segmento do pavimento de servir ao tráfego, em seu arranjo de cargas, velocidades e volumes, à época da observação, recebe o nome de serventia. Enquanto o desempenho é a variação da serventia ao longo do tempo.

Por fim, a norma DNIT (2003b), estabelece que “ a serventia atual é a capacidade de um trecho específico de pavimento de proporcionar, na opinião do usuário, rolamento suave e confortável em determinado momento, para quaisquer condições de tráfego ”. Já Bernucci et al. (2006), definem serventia como o aspecto para o qual foi produzido, levando em consideração o conforto ao rolamento que a via deve oferecer e a segurança ao tráfego, para isso estipulam um período aconselhável para manutenção corretiva.

## 2.2.1 Avaliação de pavimentos

Bertollo (1997) dispõe que o sucesso na implementação do SGP depende não só do arquivamento, mas também, da recuperação das informações. Além disso, o autor esclarece que inicialmente pressupõem-se maior tempo na busca dessas, devido à necessidade de levantamentos precisos, os quais, geralmente, não estão cadastrados nas prefeituras, como por exemplo o histórico da estrutura, o de recuperações, o de materiais utilizados, surgindo, portanto, há necessidade de retirada de corpos de provas e consulta a funcionários antigos.

Quanto à avaliação das condições do pavimento, três desempenhos devem ser considerados: funcional, estrutural, operacional. O primeiro expressa, principalmente, a característica do pavimento quanto sua serventia adequada, ou seja, a qualidade de rolamento e, portanto, está atrelado à irregularidade longitudinal e aos defeitos de superfície. Já o estrutural reflete a capacidade da pista de se manter íntegra, protegida de grandes patologias. Por fim, o desempenho operacional está relacionado às questões de geometria da via, à possibilidade de hidroplanagem, às questões de sinalização viária, à derrapagem, entre outras DNIT (2011).

### 2.2.1.1 Avaliação do desempenho funcional

Na visão dos usuários, a característica que mais afeta o desempenho funcional é irregularidade longitudinal, a qual pode ser avaliada mediante o índice de irregularidade internacional, qualificado como IRI - *International Roughness Index* - (AASHTO, 1993), sendo apropriado incluir o levantamento das patologias de superfície para obtenção de um diagnóstico consistente da funcionalidade (DNIT, 2011).

### 2.2.1.2 Levantamento da irregularidade longitudinal

Domingues (2004 apud BARELLA, 2008) explicou que a irregularidade longitudinal é o conjunto dos desvios indesejáveis, na trilha das rodas do caminho de rolamento, em relação ao plano do projeto geométrico, plano ideal, os quais podem causar perturbação ao rolamento rápido e confortável.

Ademais, além da qualidade horizontal, a irregularidade longitudinal afeta o efeito dinâmico das cargas, a drenagem da pista e a própria dinâmica dos veículos. O efeito dinâmico, por si, possibilita a intensificação dos esforços solicitantes sobre a estrutura, causando a progressividade dos defeitos e ciclicamente o aumento da irregularidade (BERNUCCI et al., 2006).

Em vista dos efeitos causados pela irregularidade horizontal, inúmeros sistemas de medição, classificados consoante a literatura como: diretos, que utilizam nível e mira para medição; indiretos – os quais utilizam perfilômetro de superfície para o mesmo fim (CHLOE, GMR, Merlin TRR, etc.); do tipo resposta, que utiliza equações, as quais correlacionam a

irregularidade com o perfil observado, fornecendo uma estimativa de variação (rugosímetro, Bump Integrator, Mayseter, Integrador IPR/USP, etc.); do tipo sonda sem contato – que utiliza perfilômetro a laser ou acústico, podem ser utilizados (DNIT, 2011).

Em 1982, um experimento internacional para determinação de medidas de irregularidade foi conduzido no Brasil, a fim de obter medidas de irregularidade comparáveis em todo mundo, já que os Sistemas de Medição de Irregularidade de Rodovia Tipo Resposta, conhecidos como RTRMMS (Response-Type Road Roughness Measurement Systems), haviam se difundido, e, embora precisos, ainda não forneciam valores adequados para a maioria das aplicações de engenharia à ocasião, o que impulsionou esse desenvolvimento em escala estatística, a qual, após o experimento, foi notabilizada com o nome IRI (GILLESPIE, 1982).

O IRI (Índice de Irregularidade Internacional) é uma métrica de rugosidade que reflete a resposta de um veículo genérico às irregularidades da superfície da estrada. Seu valor é obtido por meio da medição precisa do perfil da via, que é então processado por um algoritmo capaz de simular a reação de um veículo de referência às variações do pavimento, acumulando o deslocamento da suspensão. Dessa forma, o IRI funciona como uma replicação matemática de um roadmeter (GILLESPIE, 1982).

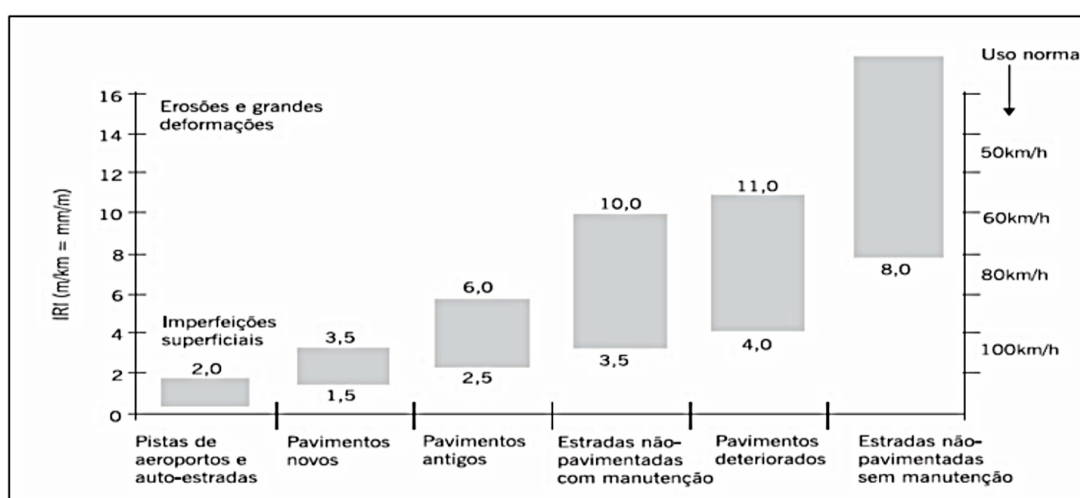


Figura 9 – Faixas de variação do IRI em função do pavimento e da velocidade  
 Fonte: Sayers e Karamihas (1998 apud BERNUCCI et al., 2006)

Júnior e Barbosa (2000) revelaram o quanto as irregularidades horizontais influenciam na intensificação dos esforços sobre a estrutura do pavimento, para isso, apresentaram o coeficiente de impacto para pavimentos e pontes devido às irregularidades.

Algebricamente, o coeficiente foi representado pela letra  $f$ , sendo estipulado como o resultado da divisão da carga dinâmica pela carga estática, o qual exprimiu o número de vezes que a carga dinâmica foi multiplicada. Os autores fizeram o cálculo para diferentes níveis de IRI, diferentes perfis de pista, determinando que  $f = 1$  representava o  $IRI = 0$ , situação hipotética ideal, com pista de rolamento sem imperfeições.

Os resultados mostraram para irregularidades horizontais isoladas, para o caso de pistas com IRI = 4,4 (onduladas e de asfalto), quando submetidas às cargas de 60 kN, 170 kN e 230 kN, empreendidas por caminhão trucado a velocidade de 90 km/h, os acréscimos de 30%, 34% e 27% da carga inicial pré-estabelecida, representados pelos coeficientes de 1,3, 1,34 e 1,27, respectivamente, além de picos de sobrecarga e alívio constantes, como pode ser visto na Figura 10.

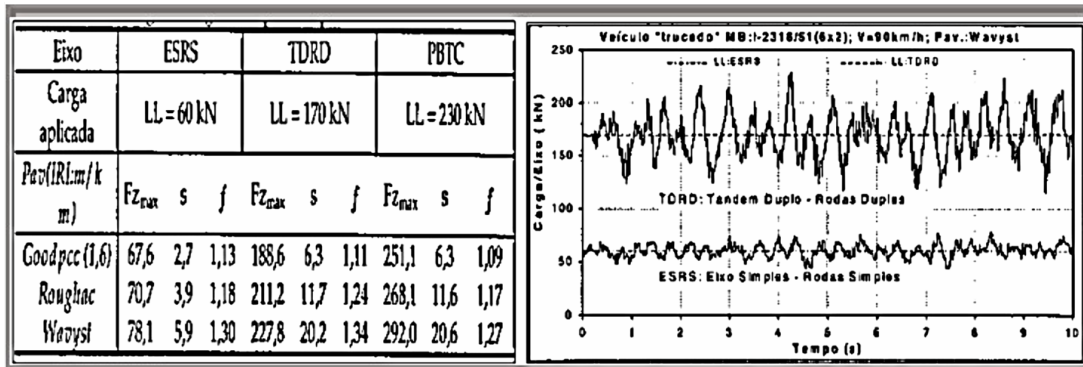


Figura 10 – À esquerda: coeficiente de impacto de veículo trucado a 90 km/h, com limite legal de cargas, sobre os perfis de pavimento; À direita: simulação de veículo trucado a 90 km/h sobre o pavimento asfáltico ondulado

Fonte: Adaptado de Júnior e Barbosa (2000)

Para as depressões com profundidade de 16 cm, tipo 3, considerando IRIs diferentes, observou-se, para as cargas estipuladas de 60 kN, 100 kN e 160 kN, os acréscimos de 117%, 362% e 245%, apresentando coeficientes de 2,17, 4,62 e 3,45, respectivamente, quando submetidas à utilização de caminhão de eixo simples a uma velocidade de 90 Km/h, o que pode ser visualizado na Figura 11.

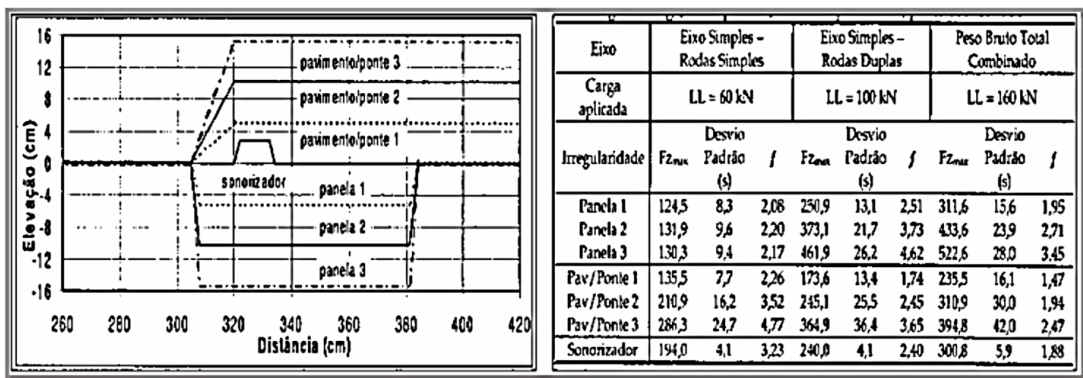


Figura 11 – À esquerda: saliência ou depressão pela extensão, para painéis, pontes e sonozador; À direita: coeficientes de impacto (f) para saliências e depressões no pavimento, representados por painéis, pontes e sonozador, utilizando veículo caminhão toco a 90 km/h com limite legal de carga

Fonte: Adaptado de Júnior e Barbosa (2000)

Portanto, o IRI pode contribuir para a majoração do aumento das cargas dos veículos sobre a estrutura do pavimento, assim como, determinadas patologias, o que

colabora para a celeridade do estado de deteriorização, à medida que ocorre o agravamento das condições de superfície.

### 2.2.1.3 Levantamento das patologias de superfície

O processo de avaliação envolve o avaliador, o veículo e o perfil longitudinal da superfície, essa interação dinâmica é responsável pelo envio de estímulos e a geração de sensações. Contudo, as sensações podem ser interpretadas de maneira diferente, mesmo quando estímulos iguais são enviados, dando-lhe caráter subjetivo (NAIR; HUDSON, 1986).

Nesse sentido, foi convencionada como subjetiva a avaliação das condições de superfície do pavimento segundo o ponto de vista dos usuários, a qual foi normalizada no Brasil pelo DNER (Departamento Nacional de Estrada de Rodagem) através da norma PRO-007/94, Valor de Serventia Atual (FONTENELE; JUNIOR; CORREIO, 2007).

O primeiro método de avaliação sistemática da serventia foi desenvolvido em 1960 por Carey e Irick, nas pistas experimentais da AASHTO, antigamente AASHO (American Association of State Highway Officials), em que o valor de serventia, o qual ficou foi denominado de Present Serviceability Rating (PSR), era estipulado pela média aritmética dos valores individuais atribuídos por cada avaliador ao trafegar em um mesmo subtrecho (BERNUCCI et al., 2006).

Carey e Irick (1960 apud BARELLA, 2008) explanaram brevemente quanto ao procedimento para avaliação de pavimento pela sistemática do PSR, informando da quantidade de avaliadores necessária, no caso 12, e da escala gradativa de notas em função da situação, de 0 a 5. Além disso, apontou para a necessidade dos avaliadores de responderem a seguinte questão: O pavimento tem condição aceitável, caso fosse necessário trafegar sobre ele por um longo período? No Brasil, a norma DNER PRO-007/94 representou a avaliação subjetiva do pavimento em um primeiro momento, posteriormente sendo substituída pela norma do DNIT 009/2003 PRO - Avaliação subjetiva da superfície de pavimentos (DNIT, 2003b).

Pela norma nacional, a avaliação subjetiva do pavimento deve fornecer o cálculo do Valor de Serventia Atual (VSA), correspondente ao método PSR da AASHTO (BERNUCCI et al., 2006).

Salviatto (2021) dissertou sobre o procedimento estabelecido pela normativa brasileira, esclarecendo aspectos importantes, como: a avaliação dos trechos de forma independente, o número mínimo de avaliadores, no caso 5, o intervalo de notas estabelecido, de 0 a 5, a máxima extensão dos trechos avaliados, 2 km, a necessidade das condições climáticas de estarem favoráveis na ocasião da avaliação, a exclusão da avaliação dos aspectos geométricos, de resistência a derrapagem, do desconforto do veículo ao passar

sobre linhas férreas, bueiros e trechos de acesso a pontes e viadutos.

O procedimento de avaliação, ainda, impõe que os avaliadores não devem comentar uns com outros sobre qualquer aspecto, nem para o caso de auxílio técnico. Também precisam considerar qualquer trecho avaliado como uma rodovia de tráfego intenso, por onde passam veículos comerciais e de passageiros, além de examinar principalmente as panelas, as irregularidades e as saliências, ignorando grandes depressões causadas por recalques de aterro DNIT (2003b).

Substancialmente, as avaliações são influenciadas pelas condições de rolamento, mas também, por desníveis (por exemplos as panelas), remendos e trincas. Dessa forma, quem avalia deve somente considerar de que forma o pavimento se comporta no momento da avaliação DNIT (2003b).

Pelo procedimento de avaliação, cada avaliador deve registrar sua nota numa ficha de avaliação, que vai de 0 a 5, correspondendo respectivamente de péssima a ótima. Assim, notas, entre o intervalo de 0 e 1, avaliam a condição superficial do pavimento como péssima, entre 1 e 2, como ruim, entre 2 e 3, como regular, entre 3 e 4, como boa, entre 4 e 5, como ótima, como pode ser visto na Figura 12. Durante a avaliação, cada responsável pela análise necessita se questionar como se comportaria o trecho em análise durante um período de percurso de 24 horas, para o caso de ser o principal trecho rodoviário, e como seria a sensação de dirigir por 800 quilômetros sobre essa superfície DNIT (2003b).

VSA	Conceito
0-1	Péssimo
1-2	Ruim
2-3	Regular
3-4	Bom
4-5	Ótimo

Figura 12 – Classificação da serventia

Fonte: Adaptado de DNIT (2003c apud SALVIATTO, 2021)

Logo após a análise, a nota estipulada deve ser marcada na escala vertical da ficha de avaliação, que é apresentado na Figura 13, DNIT (2003b). Assim, os resultados parciais de cada trecho do percurso são relacionados individualmente, correspondendo ao somatório da nota de cada avaliador dividido pelo número de avaliadores.

Em síntese, o levantamento das patologias de superfície, como parte da avaliação do desempenho funcional do pavimento, é considerado como um procedimento subjetivo,

NORMA DNIT 009/2003-PRO

**Anexo A (normativo)**  
Ficha de avaliação de serventia

VSA – Valor de Serventia Atual	5	ÓTIMO	Conceito
	4	BOM	
	3	REGULAR	
	2	RUIM	
	1	PÉSSIMO	
	0		

Rodovia: \_\_\_\_\_

Observações: \_\_\_\_\_

Nº do Avaliador: \_\_\_\_\_

Data: \_\_\_/\_\_\_/\_\_\_

\_\_\_\_\_ / Índice geral

Figura 13 – Ficha de avaliação de serventia  
Fonte: Adaptado de DNIT (2003b)

que submete-se ao parecer de um grupo de avaliadores, que normativamente devem desconsiderar aspectos relevantes observados.

### 2.3 Identificação automática de objetos

Esta seção explicará os principais conceitos utilizados na elaboração da técnica de detecção de panelas e fissuras, se fazendo imprescindível ao pleno entendimento e desenvolvimento do estudo proposto.

Sendo assim, serão esclarecidos os seguintes itens : Redes Neurais Artificiais, Tipos de Aprendizado, Aprendizado de Máquina, Aprendizagem Profundo, Redes Neurais Profundas,

Redes Convolucionais, Técnica de Entrada de Dados, Visão Computacional, Detecção de Objetos e YOLO V8.

### 2.3.1 Redes neurais artificiais

As Redes Neurais Artificiais (RNAs) são compostas por modelos matemáticos influenciados pela estrutura neural dos organismos inteligentes, capazes, por meio da tentativa, de atingir conhecimento. Essas redes são formadas por unidades de processamento chamadas de neurônios artificiais, possuindo duas características fundamentais para seu desempenho; a arquitetura e o algoritmo de treinamento (HAYKIN, 1994 apud FERNANDES, 1999).

Para McCulloch e Pitts (1990), o neurônio  $k$ , pertencente a uma determinada rede, recebe os sinais de entrada  $x_1; x_2; \dots; x_p$ , que, ao serem ponderados pelos pesos  $w_1; w_2; \dots; w_p$  e somados ao limiar de ativação inerente (conhecido como *bias*) e com a função de ajuste da saída do neurônio, que introduz mais um grau de liberdade, fornecerão a saída  $v_k$ . O sinal egresso, conforme Figura 14, é gerado após a saída  $v_k$  passar pela função de ativação  $\varphi(\cdot)$ .

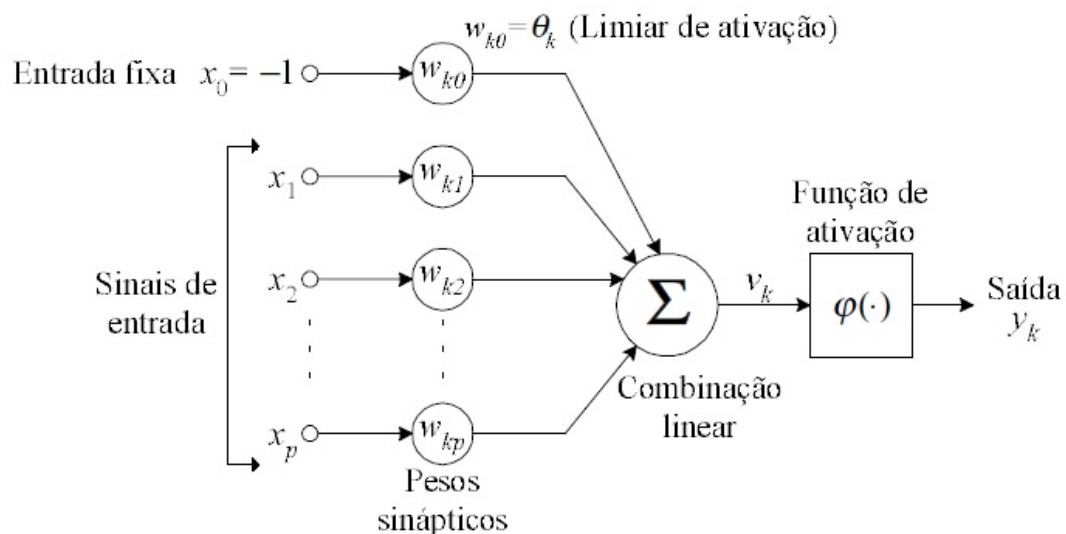


Figura 14 – Modelo do neurônio artificial

Fonte: Extraído de McCulloch e Pitts (1990 apud FERNANDES, 1999)

Fernandes (1999) explicou que a inteligência integral da rede é estabelecida através do potencial de processamento paralelo, que ocorre quando os neurônios são interligados. Assim um neurônio recebe um estímulo, o processa e emite um sinal atual, o qual é recebido por outros da rede. As redes, que o fluxo de propagação do sinal opera unidirecionalmente, recebem o nome de redes de propagação diretas (ou do inglês *feedforwards*); Já para aquelas em que não existe limitação quanto à interligação entre os neurônios e, assim, não

há um sentido único de propagação, recebem o nome de redes neurais recorrentes (RNNs) (ou do inglês *recurrent*).

Inicialmente, nas RNAs, os pesos eram valores fixos e, portanto, não podiam se auto-ajustarem. Com a evolução, surgiram as redes com múltiplas camadas, originando as camadas intermediárias, também chamadas de escondidas, as quais oportunizaram o ajuste dos pesos, atribuindo flexibilidade em aplicações que requeriam maior adaptação (FRANCO, 2014 apud OLIVATTO, 2021).

Na década de 1950, Frank Rosenblatt desenvolveu a rede *Perceptron*, composta por neurônios de multicamadas, conhecidos como MPL (*Percéptrons* de Múltiplas Camadas), o que desenvolveu os trabalhos com redes neurais por determinado tempo (ROSENBLATT, 1958 apud FLECK et al., 2016).

Rauber (2014) explicou que a modelagem *Perceptron* foi capaz de classificar dados que são linearmente separáveis. A maior importância disso está na apresentação de uma rede com habilidade de moldar os pesos internos do neurônio, dando a capacidade de resolver problemas de classificação linear.

Apesar do desenvolvimento, a rede *Perceptron* possuía limitações, como é o caso do problema XOR ( $f(0,0) = f(1,1) = 0$ ,  $f(1,0) = f(0,1) = 1$ ), que a rede não conseguiu resolver, mesmo sendo uma função simples, o que levou parte da comunidade científica a deixar o estudo das RNAs por determinado tempo (MINSKY; PAPERT, 1969 apud RAUBER, 2014).

O acréscimo de camadas intermediárias poderia possibilitar o ajuste dos pesos através de treinamento e classificar corretamente problemas mais complexos, solução que já era conhecida naquela ocasião, inclusive para o problema XOR, mas que somente foi divulgada no surgimento da rede de retropropagação de erro (ou do inglês *error backpropagation*), a qual trouxe uma solução geral para esse desafio. Após isso, vários modelos de redes foram desenvolvidos (RAUBER, 2014).

Haykin (2001) explicou que para utilização de uma RNA, a partir de seu novo desenvolvimento, ou seja da introdução de camadas intermediárias, era necessário uma pressuposição inicial dos parâmetros de entrada (para os pesos e a bias), sendo essa entrada conhecida como treinamento, que é um processo iterativo em que os padrões iniciais são utilizados até a convergência.

Assim, as RNAs tornam-se cada vez mais atraentes, à medida que ocorre o crescimento da capacidade computacional (FLECK et al., 2016). Fora isso, os recentes avanços das técnicas de inteligência artificial permitiram o desenvolvimento de redes para problemas específicos, representando uma alternativa aos algoritmos tradicionais (OLESKOVICZ; COURY; AGGARWAL, 2003 apud FLECK et al., 2016).

O processo de ajuste sobreposto aos pesos sinápticos conduz ao aprendizado da

rede, o qual somente ocorre quando um problema é resolvido através de uma solução universalizada. Em resumo, o treinamento da rede coordena o aprendizado, de modo a ajustar a matriz de pesos sinápticos ao vetor de saída que corresponda ao estimado para cada de entrada (MIRANDA; FREITAS; FAGGION, 2009 apud FLECK et al., 2016).

Eyng (2008) explicou que as RNAs extraem informações que elas consideram importantes durante o treinamento, traduzindo-as em uma representação personalíssima. Dessa forma os parâmetros são ajustados pelos pesos das conexões entre os neurônios e, ao fim, esses parâmetros são armazenados.

Para Mienye, Swart e Obaido (2024), as redes neurais têm sido aplicáveis a diferentes ambientes, alcançando performances de última geração, principalmente em séries que são obtidas ao longo do tempo. Apesar do feito, persevera um constante aperfeiçoamento dessas redes.

Logo, a crescente complexidade dos dados advindos de novas tarefas conduziu para melhoria contínua das arquiteturas e das redes variantes, ocorrendo uma expansão da aplicação para tarefas cada vez mais complexas, por exemplo, para sistemas de tomada de decisão em tempo real (MIENYE; SWART; OBAIDO, 2024)

### 2.3.2 Tipos de aprendizado

Fernandes (1999) explicou que, para as redes neurais, a propriedade mais relevante é a habilidade de aprender e melhorar seu desempenho, sendo o aprendizado, conforme mencionado anteriormente, um processo iterativo de ajustes dos parâmetros livres.

Precedentemente, é necessário entender que o ser humano adquire conhecimento por meio da indução lógica. Essa o permite adquirir conclusões genéricas sobre amostragens, amparando-se em conceitos peculiares e, posteriormente, generalizando-os (MONARD; BARANAUSKAS, 2003)

Para Monard e Baranauskas (2003), um dos principais conceitos ligados à indução é a inferência indutiva sobre exemplares. Logo, as hipóteses criadas podem ou não preservar a verdade a depender da inferência estabelecida. Nesse sentido, a inferência indutiva pode derivar conhecimento e até predizer eventos futuros, no entanto, deve existir cautela em sua utilização, pois um número insuficiente de exemplos, ou mesmo exemplos mal escolhidos, podem levar a hipóteses de pouco valor.

A literatura divide o aprendizado indutivo em aprendizado supervisionado e não supervisionado. Braga, Ludermir e Carvalho (2000) explicaram que o aprendizado supervisionado se faz quando as entradas e saídas são providas às redes, essas adquirem *know-how* para provisionarem saídas não existentes, ou seja, que estão fora do conjunto de treino, tendo como finalidade encontrar a correlação dos parâmetros providos.

Às não supervisionadas, são fornecidos apenas os dados de entrada. Assim a rede responsabiliza-se por reconhecer os padrões possíveis, sejam ocultos ou com ruídos, para disponibilizar as saídas (FRANCO, 2014). Para esse tipo, o *know-how* acontece através da redundância e da correlação das entradas (BRAGA; LUDERMIR; CARVALHO, 2000).

Mohri, Rostamizadeh e Talwalkar (2018) explicaram que as tarefas de classificação, que conferem ordem para determinados itens, e as de regressão, que realizam prognóstico de resultados, pertencem ao aprendizado supervisionado. Já as de agrupamento, que dividem um conjunto em subconjuntos, e tarefas de redução dimensional, que mantendo as características fundamentais, convertem a representação inicial em representação simplificada, pertencem ao aprendizado não supervisionado.

Existem ainda outras classificações. Nilsson (1998), por exemplo, discorreu sobre o aprendizado por reforço e o semisupervisionado.

O aprendizado por reforço simula de certa forma o comportamento do ser humano, buscando estimular o comportamento esperado para correta convergência dos resultados. Nesse sentido, existe uma realimentação (*feedback loop*) dos resultados, em que ocorre uma constante reavaliação do aprendizado no ambiente, sendo necessária a avaliação dos dados de entrada e saída ciclicamente (KAELBLING; LITTMAN; MOORE, 1996).

À combinação dos aprendizados se nomeou de semisupervisionado. Nele os dados de entrada podem estar parcialmente rotulados (que é identificação do dado bruto e sua marcação com um ou mais rótulos, especificando o contexto ao modelo, permitindo as previsões futuras) ou com rótulos faltando. Esse conjunto será utilizado para gerar mais rótulos pela parte não supervisionada. Após isso, o produto será gerado pela parte supervisionada (BOCHIE et al., 2020).

### 2.3.3 Aprendizado de máquina

Mueller e Massaron (2018 apud NEVES, 2020) relataram que apenas parte das inteligências artificiais (IAs) são modeladas para simular a inteligência humana, enfatizando que, ao se pensar em IA, deve-se atentar ao estudo entre a interação da busca do objetivo, do processamento de dados e do método de entrada, destacando que parte dos algoritmos que fundamentam as IAs não têm por finalidade simular a inteligência humana.

Inicialmente a pesquisa sobre IA tinha como prioridade a resolução de argumentos complexos através de regras de inferência lógica, o que era conhecido como abordagem através da base de conhecimento (GOODFELLOW; BENGIO; COURVILLE, 2016 apud JANIESCH; ZSCHECH; HEINRICH, 2021), entretanto, esse tipo de abordagem se deparou frente com um grande desafio, o de resolver questões que exigissem conhecimento tácito, uma vez que era indecifrável como repassa-lo, sendo fundamental na resolução de diversas questões (BRYNJOLFSSON; MCAFEE, 2017 apud JANIESCH; ZSCHECH; HEINRICH,

2021).

Elbeck (2018) dispôs que o AM simula o cérebro humano objetivando a melhoria da decisão, enquanto Wang, Zhang e Kinsner (2010) esclareceu que o aprendizado de máquina, também conhecido como computação cognitiva ou processamento de linguagem natural, é um campo que abrange a ciência da computação e informação, a área de cognição e inteligência, procurando entender os arranjos e os processos relacionados ao encadeamento e funcionamento da inteligência originária.

Ainda nesse sentido, Jordan e Mitchell (2015 apud JANIESCH; ZSCHECH; HEINRICH, 2021) discorreram que a melhoria da eficiência de um programa através da experiência, quando relacionada a tarefa, no que tange a um determinado desempenho, é o conceito de AM. Sholevar, Golroo e Esfahani (2022) explicaram que além de uma melhoria incremental através do aprendizado, os modelos AM fornecem uma melhor abstração e, portanto, resultados mais assertivos, quando dados ocultos, ou mesmo com ruídos, são analisados.

Em vista disso, o conjunto de modelos matemáticos, algoritmos, ou outras técnicas, que visam advinhar a saída de informações somente com utilização dos dados de entrada para treinamento, recebeu o nome de Aprendizado de Máquina (GRUS, 2015 apud MAGALHÃES; OLIVEIRA, 2020).

Para Bochie et al. (2020), a mais importante premissa no AM é a competência para generalização adequada dos resultados, nos casos que os modelos não passaram pelo aprendizado. A avaliação dessa capacidade é realizada através do treinamento (em que os dados brutos são utilizados para ajustar os parâmetros internos à cada camada percorrida), da validação (na qual o erro é medido para avaliação do aprendizado, sendo verificada a possibilidade de melhora ou encerramento do ciclo) e do teste (que corresponde à fase final, em que se verificará o desempenho do modelo como um todo, sendo necessária a utilização de amostras inéditas).

Na verificação do desempenho, algumas situações podem ser constatadas, dentre as quais estão o subajuste (ou *underfitting*) e o sobreajuste (ou *overfitting*). No subajuste a inexatidão do modelo ocorre por falha no ajuste dos parâmetros. Um exemplo seria quando o modelo tenta representar uma curva de segundo grau através de uma reta, nesse caso a exposição do modelo a mais dados de entrada melhoraria o resultado da representação.

Já no sobreajuste o treinamento é tido como demasiado e prejudica o resultado. Um exemplo seria a representação de um polinômio de grau 2 por meio de um de grau 8, mostrando uma grande discrepância quando as amostras diferem infinitesimalmente.

Nessa lógica, o AM é uma das dimensões da IA, que tem por finalidade buscar modelos que representem soluções, para os casos que somente se disponha dos elementos de entrada, gerando resultados aperfeiçoados.

### 2.3.4 Aprendizado profundo

LeCun e Bengio (1995) explicaram que o aprendizado de máquina convencional, por vezes, necessita de um pré-tratamento dos dados de entrada, exigindo conhecimento antecipado do programador. Isso é percebido na representação linear de uma função não linear, por exemplo, em que a solução ocorre através do treinamento da rede pela representação gráfica projetada, ao invés, apenas, do treinamento com os dados de entrada previamente tratados (BOCHIE et al., 2020).

No entanto, quando a representação gráfica é tão difícil quanto a determinação da função, tal como para conceitos abstratos, é necessária a utilização do aprendizado profundo, o qual fará a abstração das representações complexas em representações mais simples, que serão organizadas hierarquicamente, se complementando ao final (GOODFELLOW; BENGIO; COURVILLE, 2016).

Bochie et al. (2020) descreveram que as seções dos algoritmos, para o caso de reconhecimento de imagens, podem especializar-se, de modo que as bordas, por exemplo, sejam identificadas à parte e outras progressivamente, até que os arranjos das seções estejam todos especializados, a ponto de identificarem a imagem como um todo. Nesse sentido as camadas são estruturadas paralelamente àquelas de reconhecimento menos complexo, num sentido progressivo de dificuldade de extração.

O aprendizado profundo descreve de forma ampla qualquer modelo de aprendizado que o implemente de forma hierárquica. No entanto, o termo via de regra é associado às Redes Neurais Profundas (*Deep Neural Networks* - DNNs), que são revestidas dessa finalidade.

Por último, as DNNs são compostas por diversas camadas ocultas, as quais são responsáveis pela abstração das informações de forma progressiva de complexidade. A especialização ocorre à medida em que o número de camadas é introduzido na estrutura das redes, sendo que as iniciais, que estão próximas ao começo do extração, responsáveis pela abstração de atributos simples, enquanto aquelas próximas ao fim pelos complexos (BOCHIE et al., 2020).

### 2.3.5 Redes convolucionais

As Redes Neurais Convolucionais (Convolutional Neural Networks – CNNs) são estruturas que solucionam o problema das RNAs, nas quais o processamento pode ser inviabilizado devido à necessidade de grande quantidade de camadas ou mesmo da grande quantidade amostral (LECUN; BENGIO, 1995).

Essa necessidade ocorre porquê nas RNAs acontece uma grande interação para os ajustes de pesos, logo, durante o processamento, é realizada uma multiplicação geral das matrizes de dados (GOODFELLOW; BENGIO; COURVILLE, 2016; FARIA, 2018).

No decorrer do processamento de uma RNA, todos os neurônios se conectam, ocorrendo um aumento demasiado do tempo para o completo treinamento da rede (SHARMA et al., 2019).

Em uma CNN, existem as camadas de operações especializadas (camadas convolucionais), as de ativação não lineares (Rectified Linear Unit - ReLu) e as camadas de subamostragens (Poolings) tornando-as mais otimizadas (GOODFELLOW; BENGIO; COURVILLE, 2016; FARIA, 2018), como será explicado à frente.

A maior otimização das CNNs advém de possuírem maior resistência a variação dos dados. Essa resistência decorre da camada de Pooling. A função Pooling é a responsável por realizar o cálculo de submatrizes por meio da média de entrada e do valor máximo das camadas anteriores, como ser visto na Figura 15.

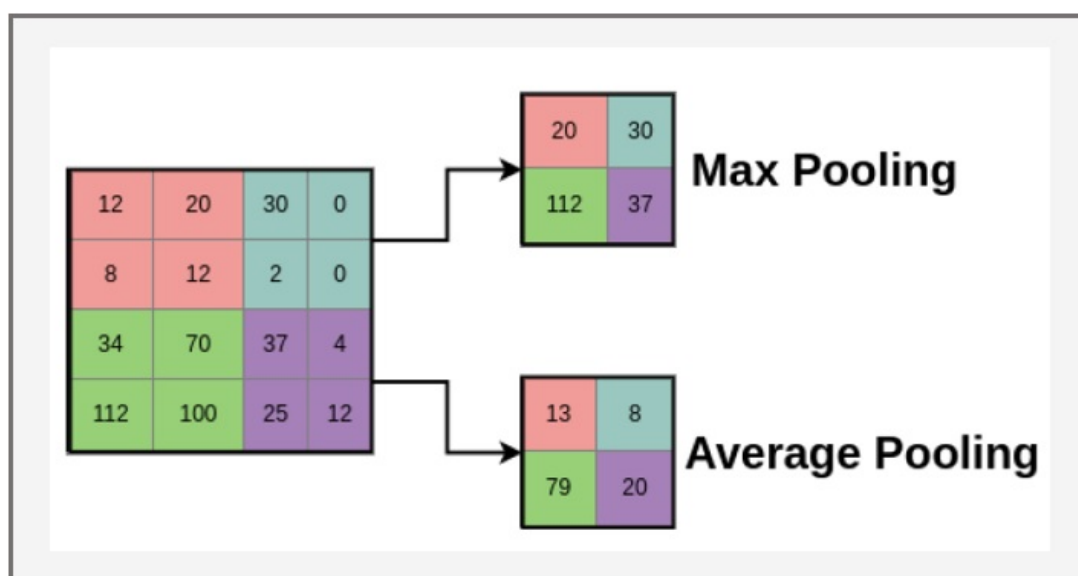


Figura 15 – Operações de Max Pooling e Average Pooling  
Fonte: Extraído de Bochie et al. (2020)

Ademais, a resistência também pode ser realizada pela função *zero padding*, a qual pode incluir zeros aos dados de entrada de uma determinada camada, evitando que a rede, ou o seu núcleo, tenha de ser reduzido, fornecendo adaptabilidade as CNNs (GOODFELLOW; BENGIO; COURVILLE, 2016).

Outra característica dessa função é de realizar a redução da influência dos dados tratados através da análise das dimensões envolvidas e necessárias, já que a pooling atua nas saídas dos neurônios em regiões específicas (BOCHIE et al., 2020), como pode ser visto na Figura 16.

As CNNs possuem outra grande vantagem quando comparadas às RNAs: a de trabalhar com uma quantidade reduzida de parâmetros. Assim, para essa tipologia, existe a esparsialidade entre os neurônios de camadas adjacentes e o compartilhamento de parâmetros (BOCHIE et al., 2020).

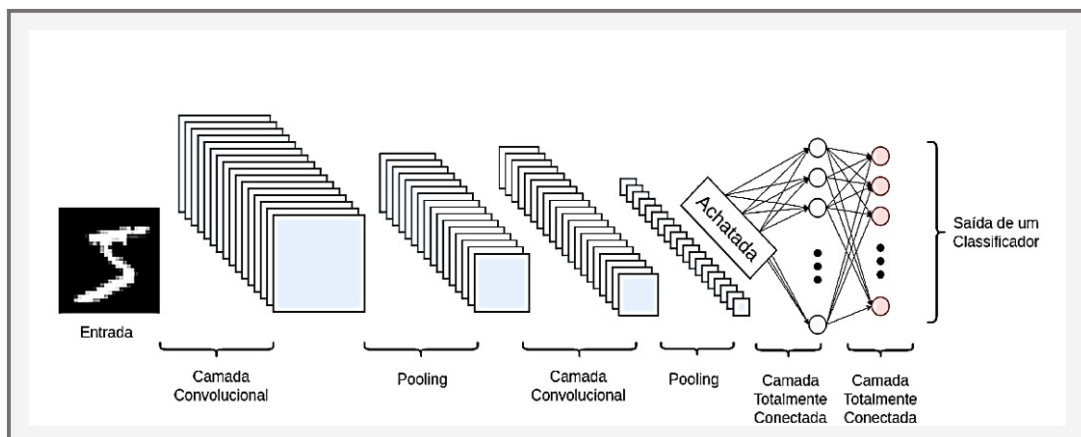


Figura 16 – Arquitetura de uma rede neural convolucional  
 Fonte: Extraído de [Bochie et al. \(2020\)](#)

A esparsialidade ocorre pela limitação de conexões entre os neurônios de uma mesma região, evitando que algumas regiões não sejam utilizadas plenamente ([GOODFELLOW; BENGIO; COURVILLE, 2016](#)).

Quanto ao compartilhamento, somente a um conjunto específico de pesos é permitida a replicação. Logo, toda a estrutura é submetida aos núcleos de valores conforme é apresentado na Figura 17, também conhecido como filtro de Kernel, que é estabelecido pela operação do produto da camada anterior com a sua matriz de valores. Desse modo, cada célula da matriz produto fica representada por um valor que a rede deverá apreender ([KHAN et al., 2018](#)).

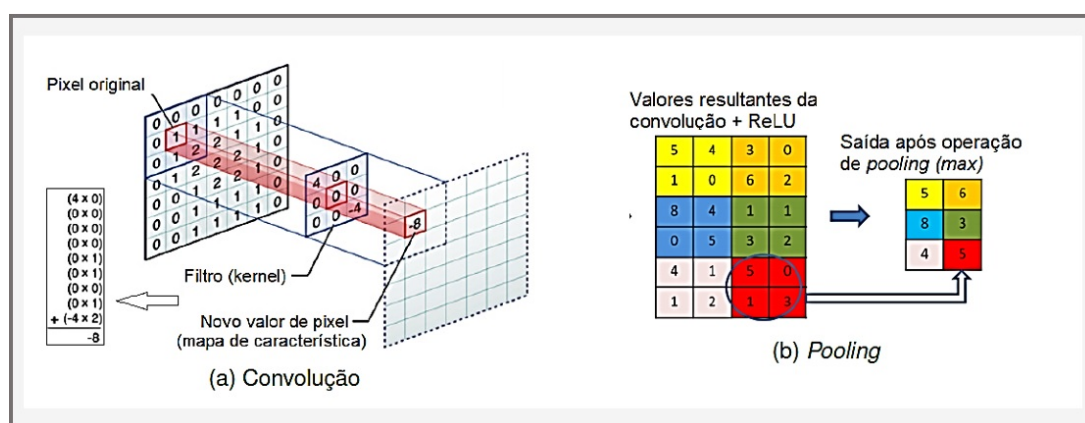


Figura 17 – Filtro Kernel (a) e Operação de Max Pooling (b)  
 Fonte: Adaptado de (a) [APPLE \(2016\)](#) e (b) [Faria \(2018\)](#)

De um modo geral, as CNNs também são conhecidas por conseguirem processar dados padronizados disponíveis em matrizes múltiplas, por exemplo matrizes 2D (arrays), que contenham pixel nos três canais de cores das camadas RGB. Logo, essa estrutura possibilita, por exemplo, o processamento de imagens e vídeos (espectogramas), mas também o processamento das matrizes 3D, com imagens volumétricas e vídeos ([LECUN; BENGIO; HINTON, 2015; GOODFELLOW; BENGIO; COURVILLE, 2016](#)).

Em síntese, as camadas convolucionais são formadas por 3 estágios: o primeiro estágio, formado pelas camadas especializadas, as quais produzem saídas lineares através de diversas convoluções, no segundo, formado pela função ReLu, também conhecido como estágio de detecção, em que ocorre a ativação e no terceiro, estágio de Pooling, em que é feita a computação dos valores de cada região (BOCHIE et al., 2020).

Na Figura 18 (a), são apresentadas as operações que repetidamente são aplicadas por uma rede CNN. Assim, a camada convolucional é concebida com seus filtros lineares. A camada ReLu e a Pooling também são representadas. Os mapas de recursos  $M$  são submetidos aos filtros  $N$ . Os  $N$  mapas resultantes desse processo são submetidos a função ReLu e agrupados pela função pooling, fornecendo um mapa de recursos com função reduzida.

Na Figura 18 (b), é ilustrada uma CNN com 11 camadas e uma imagem com três canais de cores. A CNN possui 8 camadas convolucionais conjuntamente a ReLu, sendo 3 totalmente conectadas (as três últimas), e as 5 camadas de Pooling.

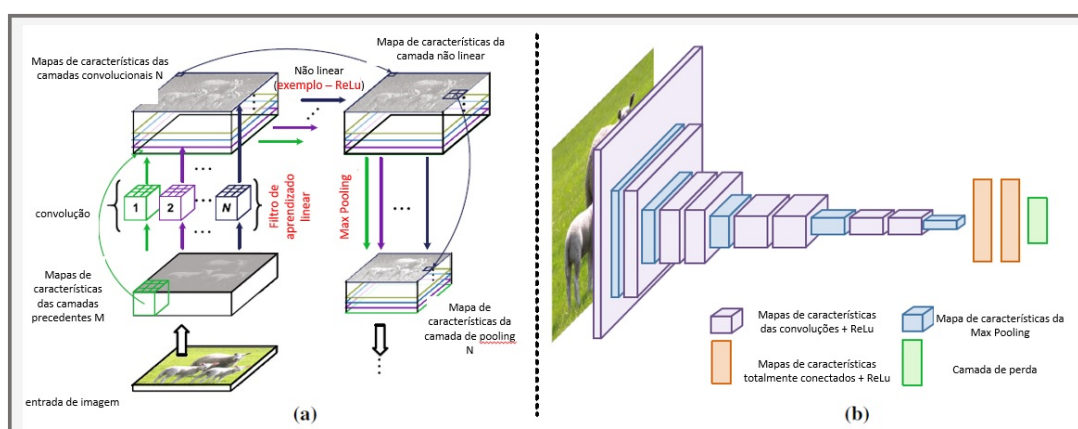


Figura 18 – Ilustração das operações de uma rede CNN

Fonte: Adaptado de Liu et al. (2020)

Nesse sentido, as CNNs são redes que obtêm melhores resultados quando existe uma grande quantidade de dados a serem abstraídos, obtendo, assim, um ajuste mais refinado dos modelos treinados, uma vez que possui funções especializadas, as quais são capazes de conter as variações de dados não significativas nas partes específicas do processo.

### 2.3.6 Visão Computacional

A Visão Computacional (Computer Vision - CV) é uma área do processamento de imagens que compreende diversos métodos e técnicas para interpretação de imagens (GOODFELLOW; BENGIO; COURVILLE, 2016).

Dentre os principais métodos de interpretação de imagens estão a classificação, a classificação e localização conjuntamente, a segmentação e a detecção (GOODFELLOW; BENGIO; COURVILLE, 2016).

Para Li, Johnson e Yeung (2017), a classificação é a concessão de um rótulo a um determinado objeto, já a classificação conjuntamente a localização estão relacionados à atribuição de um rótulo, mas também a determinação da localização. A segmentação está associada ao correlacionamento dos pixels das classes dos objetos. Por fim, a detecção implementa a rotulação na coleção de objetos que foi classificada e localizada.

Na Figura 19, é possível visualizar caixas circunscrevendo os objetos. Essas caixas são comumente nomeadas de *bounding boxes*, e definem a localização espacial dos objetos nas imagens (LIU et al., 2020).



Figura 19 – Métodos de interpretação de imagens utilizados pela CV  
Fonte: Adaptado de Liu et al. (2020)

Há também, segundo Li, Johnson e Yeung (2017), dois tipos de segmentação, a semântica e a de instâncias. A diferença entre ambas é que na segmentação de instâncias, os pixels da mesma instância de objetos são indenficados individualmente, como pode ser visto na Figura 20.



Figura 20 – (a) Segmentação semântica e (b) Segmentação de instâncias  
Fonte: Adaptado de Liu et al. (2020)

Por fim, parte dos modelos de detecção, utiliza os *bounding boxes* para, então, efetuarem a classificação conjuntamente a localização, ou, a detecção de objetos de interesse. Já, a segmentação, também muito utilizada, atua na identificação de classes individuais ou de forma coletiva.

### 2.3.7 Detecção de Objetos

Para [Szeliski \(2022\)](#), examinar detalhadamente uma imagem com o objetivo de identificar uma equivalência é o fundamento da detecção.

Em suma, detectar objetos é identificar a que classe pertence e exibir sua posição através de um *bounding box* ([LIU et al., 2020](#); [RUSSAKOVSKY et al., 2015](#)).

Nesse sentido, a classificação pode ser interpretada como a categorização, enquanto a localização é a regressão ([OLIVATTO, 2021](#)).

Além disso, na detecção de objetos, existe o objetivo de identificar todas as instâncias de objetos que estejam em uma determinada imagem, tornando o desafio ainda maior. A fim de conseguir alcançar esse objetivo, foram desenvolvidas 2 estruturas de CNNs, a *Region Proposal Based Framework* e a *Regression/Classification Based Framework Based* ([WANGENHEIM, 2018](#)).

#### *Region Proposal Based Framework*

Essa estrutura foi concebida no formato multi-etapas. Logo, a detecção é realizada através de segmentação das regiões. Dessa maneira, existe uma similiaridade com o funcionamento do cérebro humano, em que é realizada uma varredura integral do espaço e, a partir disso, as regiões de interesse são evidenciadas ([ZHAO et al., 2019](#)).

Para [Girshick et al. \(2013\)](#), *Region Proposal Based Framework* gera, em um primeiro momento, as regiões candidatas a classificação, as quais vieram de uma busca seletiva inicial. Na sequência, essas regiões são agrupadas em regiões maiores, devido a estrutura ter detectado similaridade de características. No fim, as regiões maiores são classificadas.

Um dos modelos que representa a aplicação dessa estrutura é o *Region-Based Convolutional Neural Networks* (R-CNN). Apesar do modelo possuir essa estrutura, ele requisita um tamanho definido para imagem de entrada, com isso a imagem fica submetida ao redimensionamento, o que resulta em perda de conteúdo ou comprometimento da exatidão de detecção. O redimensionamento da imagem de entrada pode ocorrer por corte ou distorção, o que explica respectivamente os efeitos mencionados. ([ZHAO et al., 2019](#)).

O aprimoramento da R-CNN se desenvolveu através da adição da camada *Spatial Pyramid Pooling* (SPP), a qual possibilitou que as imagens fossem carregadas integralmente, antes de ser subdividas. O primeiro algoritmo a receber-lá foi o SPP-NET, desenvolvido por Kaiming He, Xiangyu Zhang, Shaoting Ren e Jian Sun ([OLIVATTO, 2021](#)).

Embora a SPP-NET permitisse que as imagens fossem carregadas integralmente, sua estrutura realizava a extração das características preliminarmente, para posteriormente classifica-lás e, por fim, definir as suas localizações ([GIRSHICK, 2015](#)).

A revolução adveio da criação de uma camada de *pooling*, em que todas as etapas foram substituídas em um único algoritmo, que foi chamado de *Fast R-CNN*. Nesse modelo,

os produtos de cada camada são classificados por outra, prognosticando a categoria e seu *bounding box* (GIRSHICK, 2015).

Para Zhao et al. (2019), o progresso das novas estruturas de detecção tiveram o objetivo de melhorar o desempenho do processamento das camadas de *pooling* das regiões de interesse, da produção de mapas de características, além do desenvolvimento da possibilidade de segmentação de instâncias (ZHAO et al., 2019).

#### *Regression/Classification Based Framework Based*

Para essa estrutura, a imagem é considerada como um todo, logo os pixels da imagem e os *bounding boxes* são de uma em vez só detectados, o que reduz o tempo gasto, já que não há produção de sub-regiões de interesse. Além disso, a probabilidade da classe do objeto, que foi identificado, já é realizada na mesma etapa. Os modelos pioneiros dessa estrutura são o *You Look Once* (YOLO) e o *Single Shot MultiBox Detector* (SSD) (ZHAO et al., 2019).

Szegedy, Toshev e Erhan (2013) desenvolveram outro modelo utilizando a estrutura *Regression/Classification Based Framework Based*, o qual foi chamado de *DNN Based*. O modelo em questão produz um máscara binária para etapa de teste e para extração das características.

Embora apresentasse essa inovação, havia dificuldade do modelo no manuseio de objetos jtapostos e *bounding boxes* advindos de sobreamostragem (SZEGEDY; TOSHEV; ERHAN, 2013).

Yoo et al. (2015) apresentaram o modelo de CNN chamado de *AttentionNet*, que alcançou a marca de 65% de reconhecimento, utilizando uma única rede convolucional de 8 camadas de arquitetura. Nele, a detecção ocorre por meio da classificação iterativa. O *AttentionNet* produz direções fracas de limite do objeto (*bounding boxes* em azul) em direção ao objeto que se pretende identificar, convergindo para um *bounding box* preciso (*bounding boxes* em vermelho) do que se deseja detectar, de acordo com a Figura 21.

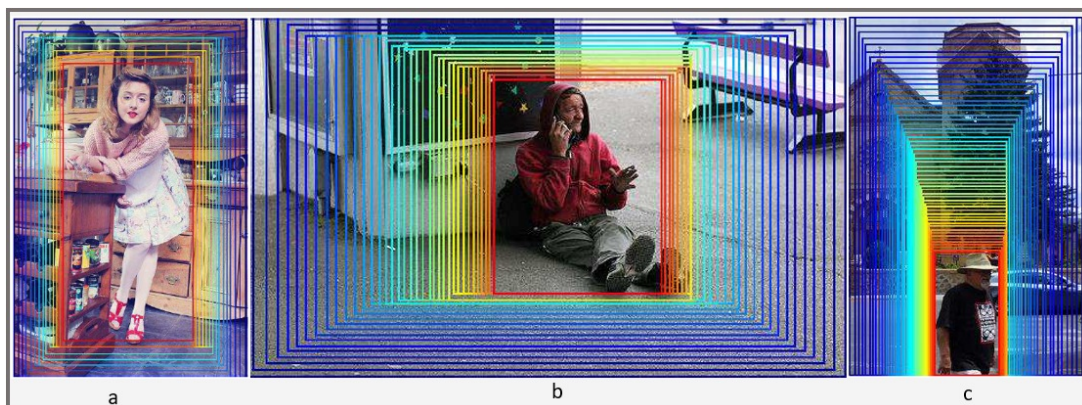


Figura 21 – Imagens (a), (b) e (c) Exemplos reais de detecção do modelo AttentionNet  
Fonte: Adaptado de Yoo et al. (2015)

O *AttentionNet* foi capaz de produzir os resultados dos modelos mais sofisticados naquela ocasião, sem incluir quaisquer modelos separados, por meio da soma da previsão de identificação dos *bounding boxes* (YOO et al., 2015).

Apesar dos resultados promissores, o modelo se torna bastante ineficiente ao proceder com a detecção de várias categorias ao mesmo tempo, utilizando duas instâncias do modelo consecutivas (ZHAO et al., 2019).

Najibi, Rastegari e Davis (2016) desenvolveram uma nova técnica de detecção que foi chamada de G-CNN. A G-CNN é capaz de treinar uma rede convolucional para direcionar e estimar uma grade apoiada no tamanho dos objetos. Após o mapeamento, a camada de *pooling* do modelo realiza o cálculo das regiões de interesse (*region of interest* - ROI) e um regressor linear realiza a definição da localização, além de estimar o tamanho necessário do *bounding box*, vide Figura 22 (NAJIBI; RASTEGARI; DAVIS, 2016).

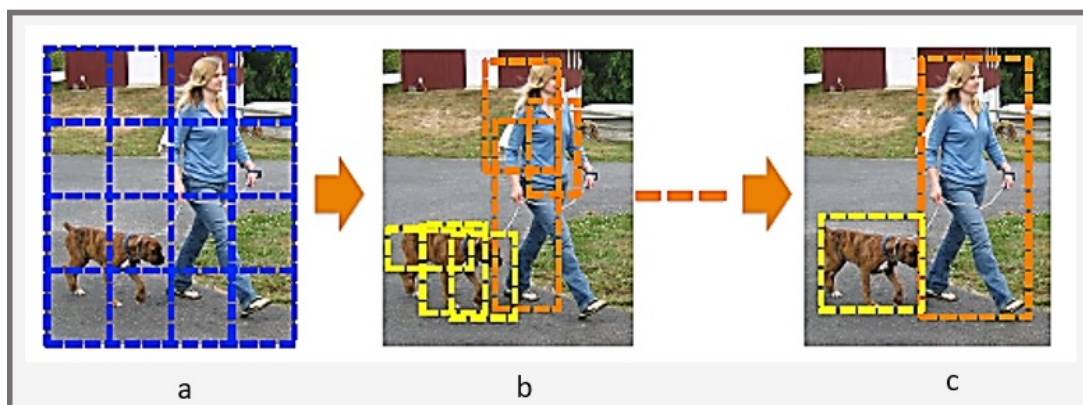


Figura 22 – Modelo G-CNN: (a) Grade retangular multiescala, (b) *bounding boxes* deslocados para os objetos e (c) *bounding boxes* de delimitação dos objetos

Fonte: Adaptado de Najibi, Rastegari e Davis (2016)

O modelo G-CNN possui semelhanças à arquitetura do modelo *Fast R-CNN*, no entanto, o G-CNN atua na detecção dos objetos nos limites das caixas, iniciando pela grade multiescalar gerada, o que se traduz no principal objetivo dessa estrutura. O aprendizado é feito a fim que a estrutura consiga mover os *bounding boxes* em direção aos objetos em apenas algumas iterações (NAJIBI; RASTEGARI; DAVIS, 2016).

Apesar do G-CNN possuir o desempenho de identificação dos modelos em vigor naquela ocasião, o mesmo apresenta dificuldade na detecção de pequenos objetos e na detecção de objetos com alta sobreposição (ZHAO et al., 2019).

Redmon et al. (2016) propuseram o *YOLO* (*You Look Once*). Essa estrutura, assim como a G-CNN, divide a imagem em uma grade de tamanho variável. Com isso, cada célula fica encarregada de identificar o objeto que está centralizado nela, conforme apresentado na Figura 23.

Para Jiang et al. (2022), o algoritmo *YOLO* é de estrutura simples, possuindo

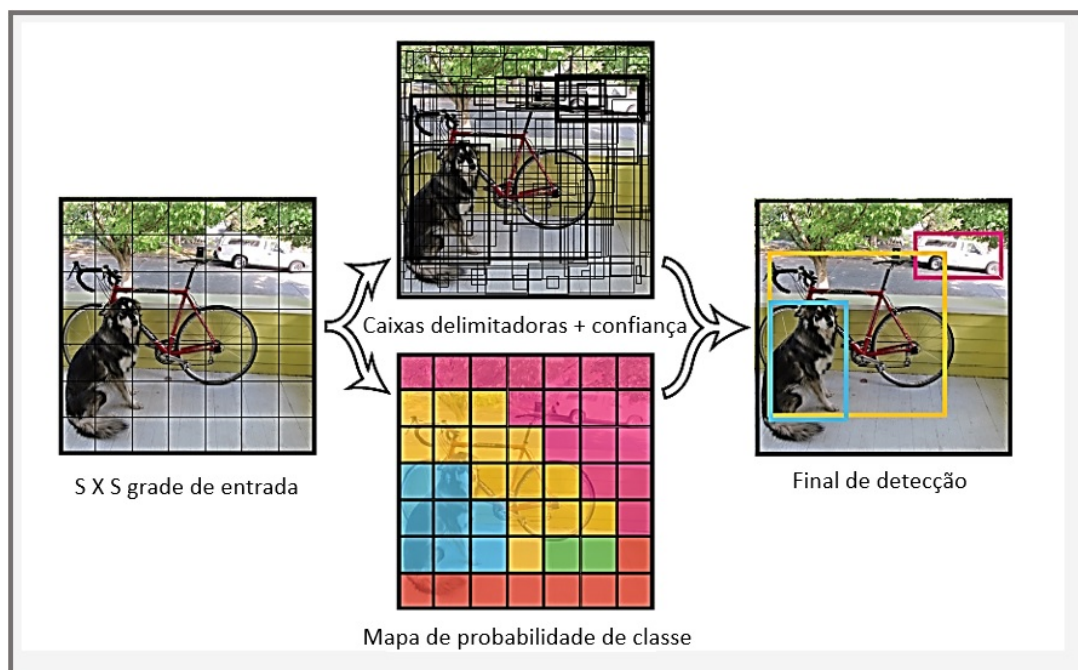


Figura 23 – Funcionamento do modelo YOLO

Fonte: Adaptado de [Redmon et al. \(2016\)](#)

como principal característica seu tamanho pequeno e a alta velocidade na identificação de imagens, sendo que essa última característica está alicerçada, principalmente, na colocação da imagem global em rede, a qual pode ser codificada com informações gerais, diminuindo o erro de identificar o plano de fundo como sendo o objeto a ser detectado.

Outrossim, se refere à grande capacidade de generalização, podendo aprender recursos generalizados. Para tanto, o algoritmo transforma detecção de imagens em questões de regressão linear.

### 2.3.7.1 YOLO

Muito rápido na detecção de objetos, o modelo YOLO não necessita de uma estrutura complexa para essa finalidade. Em vista disso, já na sua primeira versão, produziu detecções a uma taxa de 150 quadros por segundo, ou 150 FPS (do inglês *Frames per Second*), quando executado na versão *fast*, denotando a forte capacidade de detecções em tempo real. Além disso, o modelo alcançou mais que o dobro de precisão média (mAP) de detecção em tempo real do que qualquer outro modelo àquela ocasião ([REDMON et al., 2016](#)).

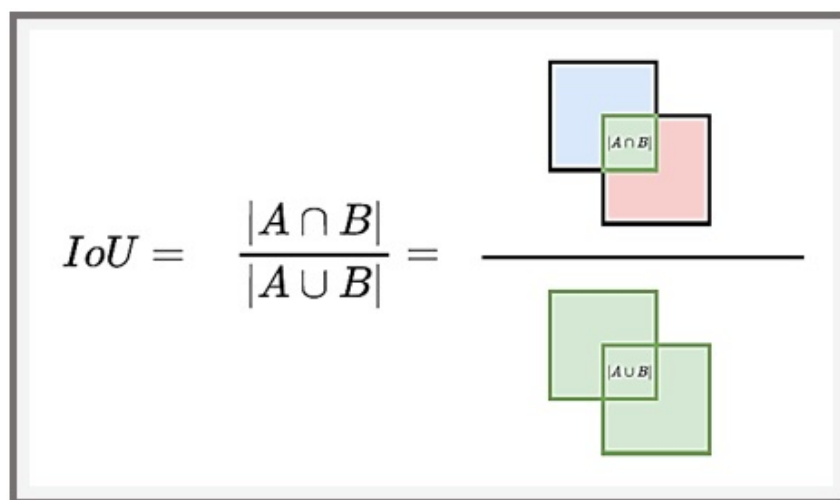
[Redmon et al. \(2016\)](#) explicaram que diferentemente do modelo *Fast R-CNN*, o YOLO observa a imagem integralmente, codificando informações implícitas do ambiente das classes e da sua aparência. Isso, de modo geral, leva o modelo a um desempenho superior na identificação, sendo também, amplamente aplicável a inúmeras classes.

As grades das células preveem um número  $N$  de *bounding boxes* e apresentam a

confiança do modelo quanto a esses possuírem um objeto e a sua precisão na detecção. Para isso, os *bounding boxes* produzem as previsões  $x$ ,  $y$ ,  $w$ ,  $h$  e a confiança.  $X$  e  $y$  representam as coordenadas de largura e altura do centro do *bounding box*, enquanto  $w$  e  $h$ , as medidas de largura e altura da imagem integralmente (REDMON et al., 2016).

Para entender o indicador de confiança é necessário compreender o conceito de *Intersection-over-Union* - IoU (Intersecção sobre União).

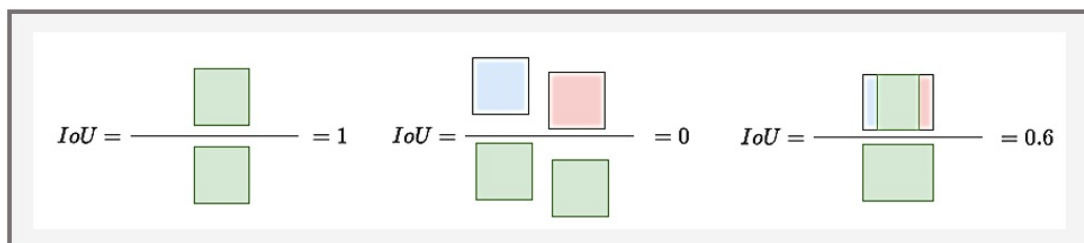
A IoU representa a média da divisão entre a área do *bounding box* correspondente a intersecção pela área de união, como pode ser visto na Figura 24 (WU; SAHOO; HOI, 2020).



$$IoU = \frac{|A \cap B|}{|A \cup B|}$$

Figura 24 – Cálculo de IoU  
Fonte: Extraído de (SAVINAINEN, 2024)

Nesse sentido, a IoU informa quanto o *bounding box* previsto corresponde ao *bounding box* de verdadeiro. Assim, quando resultar em 1, a sobreposição prevista fica perfeitamente coincidente. Quando igual a zero, não existe sobreposição e, quando entre 0 e 1, existe uma sobreposição parcial, ver Figura 25.



$$IoU = \frac{\text{Area of intersection}}{\text{Area of union}} = 1 \quad IoU = \frac{\text{Area of intersection}}{\text{Area of union}} = 0 \quad IoU = \frac{\text{Area of intersection}}{\text{Area of union}} = 0.6$$

Figura 25 – Valores de IoU  
Fonte: Extraído de (SAVINAINEN, 2024)

A confiança é a multiplicação entre o resultado da probabilidade de que o objeto esteja alocado dentro do limite do *bounding box* pelo IoU do *boundig box*.

Além de desempenhar papel fundamental no cálculo da confiança, a IoU fornece outras informações importantes, que são as classificações em Verdadeiros Positivos (VP), Falsos Positivos (FP) e Falsos Negativos (FN).

- Verdadeiro Positivo (VP): significa uma detecção assertiva do objeto (quando existe a detecção e a mesma ocorreu corretamente);
- Falso Positivo (FP): significa uma detecção não assertiva do objeto (quando existe a detecção, mas não deveria ter ocorrido);
- Falso Negativo (FN): representa uma detecção não assertiva para o que é o objeto (quando não existe a detecção, onde deveria ter ocorrido).

A partir do enquadramento nessas classes é possível calcular a *Precision* e o *Recall*, conforme mostrado nas Equações 2.1 e 2.2 respectivamente:

$$Precision = \frac{VP}{VP + FP} \quad (2.1)$$

$$Recall = \frac{VP}{VP + FN} \quad (2.2)$$

VP = Verdadeiro Positivo

FP = Falso Positivo

FN = Falso Negativo

### 3 Materiais e Métodos

Esta pesquisa assume uma abordagem quantitativa, uma vez que as amostras são grandes e representam a população, sendo assim, constituem um retrato real dessa (FONSECA, 2002 apud GERHARDT; SILVEIRA, 2009); e, quanto ao procedimento, experimental, em que as etapas iniciam pela formulação exata do problema e das hipóteses, as quais circunscrevem as variáveis necessárias e controladas que estão envolvidas no evento estudado (TRIVIÑOS, 1987 apud GERHARDT; SILVEIRA, 2009).

No quesito objetivo, trata-se de estudo exploratório, a fim de construir novas hipóteses e tornar o problema mais explícito (GIL, 2002). Por fim, quanto a sua natureza, classifica-se como aplicada, já que se dirige a objetivos específicos, tendo a gerar aplicação prática (GERHARDT; SILVEIRA, 2009).

Preliminarmente foi realizada uma pesquisa com a finalidade de entender a representatividade das estradas brasileiras no escoamento de cargas, através de pesquisas realizadas nos repositórios de diversas universidades nacionais e nos sites especializados em artigos acadêmicos. Com isso, foi possível entender a importância da manutenção dessa infraestrutura para o desenvolvimento brasileiro em nível nacional. Contudo, através de pesquisas sobre indicadores qualitativos das rodovias brasileiras, desenvolvidos principalmente por entidades de classe, foi possível aferir a nota fornecida pelos usuários ao sistema viário atual.

A revisão bibliográfica, já numa nova fase, abarcando principalmente sites internacionais especializados na produção acadêmica voltada à transportes, à infraestrutura e à inteligência artificial, permitiu identificar novas tecnologias para detecção de imperfeições em pavimentos, com potencial de empregabilidade nos sistemas atuais de gerenciamento de pavimentos.

Para o desenvolvimento desse trabalho, como forma de melhoria contínua, foi proposta uma etapa preliminar, intitulada de detecção inicial, que visou planejar, corrigir e melhorar todo processo final de detecção, por meio da análise e implementação de melhorias, a qual se materializou como uma etapa independente da detecção final.

Referente ao nome utilizado durante o procedimento experimental, para o caso das patologias que serão objeto de identificação automatizada, foi feita a opção pela terminologia de panelas e fissuras, uma vez que esse trabalho tem aspecto quantitativo em sua essência, não cabendo para esse estágio de desenvolvimento a diferenciação entre trincas e fissuras.

Diante desse contexto, a implementação ocorreu nas seguintes subetapas: Para

etapa de detecção inicial em: configuração e instalação dos equipamentos, coleta de dados para o treinamento, preparação do banco de imagens de treinamento, rotulação, preparação do minidataset, treinamento e avaliação do treinamento.

Para etapa detecção final em: configuração, instalação e coleta de dados para treinamento, preparação do banco de imagens e rotulação, pré-processamento, preparação do dataset, treinamento, coleta de dados para detecção, detecção, conforme apresentado na Figura 26.

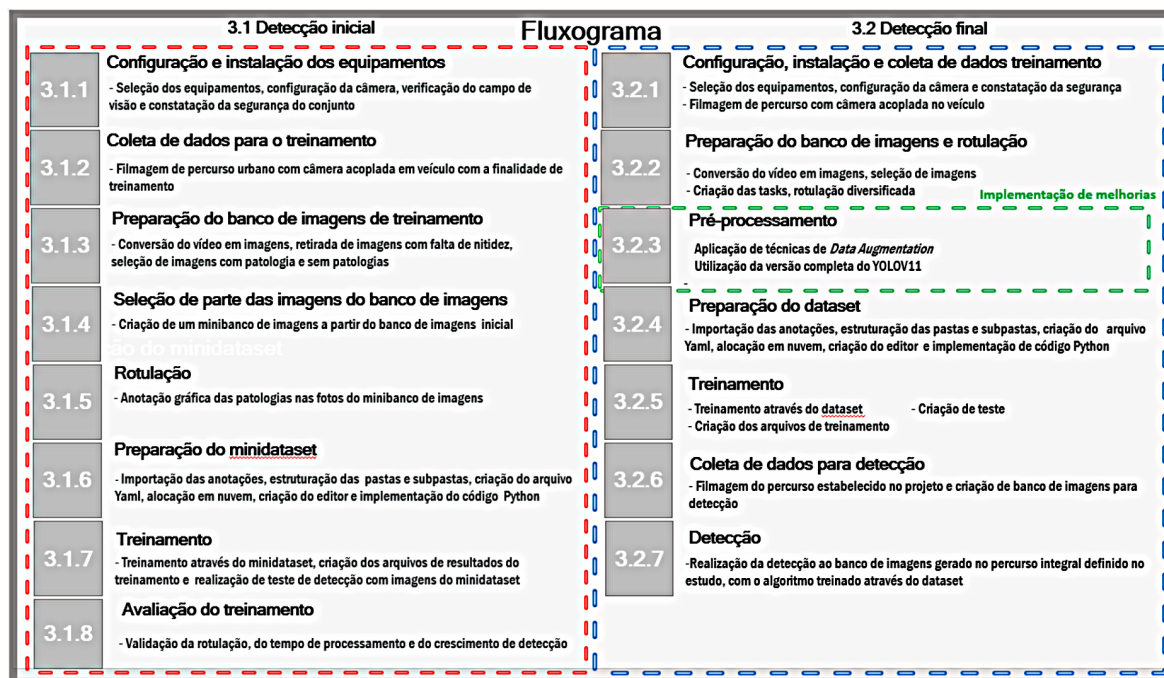


Figura 26 – Fluxograma descrevendo as etapas dos materiais e métodos

Fonte: Elaborado pelo autor

Por fim, para o desenvolvimento também foram utilizados os seguintes recursos:

- Google: para pesquisa de material acadêmico de fundamentação;
- CVAT: para rotulação das patologias nas imagens;
- Conversor Python de "\*.xml" para "\*.txt": para a conversão dos arquivos de rótulos para o formato compatível com YOLO;
- Roboflow: para implementação de técnicas de pré-processamento (redimensionamento e Data Augmentation);
- Ferramenta Python de Tiling: para implementação da técnica de Tiling;
- Google Drive: para armazenamento de arquivos dos treinamentos de CNNs (entrada, backup e saída), inclusive arquivos utilizados nos testes;
- Google Colab (Python Notebook com GPU NVÍDEA SMI 535.104.05 A 100 SX-M4-40 GB, CUDA VERSION 12.2, cuDNN 7.6.5 e OpenCV 3.2.0): para execução dos

treinamentos e testes das CNNs;

- YOLOv8s e YOLOv11x (com Darknet AlexeyAB): para estruturação e execução dos treinamentos e testes das CNNs.

### 3.1 Detecção inicial

A indispensabilidade de entender e confirmar o processo como um todo, com o objetivo de orientar, planejar, corrigir e melhorar a detecção final, tornando-a mais eficaz, oportunizou a criação da etapa de detecção inicial (uma forma de prototipagem). Em vista disso, as subetapas de seleção de imagens para treinamento, para rotulação, para preparação dos minidatasets e para treinamento foram elaboradas de modo que 4 treinos fossem realizados. Já, na subetapa de avaliação do treinamento foram considerados apenas 3 treinos, conforme será explicado detalhadamente mais à frente.

#### 3.1.1 Configuração e instalação dos equipamentos

A coleta de dados foi realizada através da utilização de uma câmera de ação GoPro Hero 12 Black, Figura 27 (vistas de A a E), a qual foi configurada para um padrão específico de gravação, que foi chamado de Customizado 2, Figura 28.



Figura 27 – Câmera de ação GoPro Hero 12 Black  
Fonte: Elaborado pelo autor

Para essa configuração, a resolução escolhida foi de 3840 pixels na horizontal por 2160 pixels na vertical (4K), e o perfil de iluminação, que faz o ajuste da iluminação, foi o de modo natural. Além do mais, a proporção de aspecto foi ajustada para 16 por 9, permitindo, posteriormente, que as imagens do banco de dados ficassem no formato retangular, conforme Figura 28 vistas A e B respectivamente.

A taxa de quadros por segundo (do inglês *Frames per Second*- FPS) foi ajustada para 60 quadros e o campo de visão selecionado foi o com lente de ajuste para o horizonte, o qual mantém a linha de horizonte estabilizada e evita distorções da borda, vide Figura 28, vista B.

Quanto à estabilização digital do vídeo, foi utilizada a função *AutoBoost*, que permite um reforço da estabilização do vídeo de forma automática, e que tem por objetivo preservar o campo de visão assentado com o mais alto nível de estabilização suportado, apresentada na Figura 28, vista D.

O tempo de abertura e fechamento do obturador do equipamento, designado na configuração como *shutter*, foi configurado no modo automático, o qual está em consonância com a taxa de *frames*, ou seja, ocorrendo de acordo com o tempo de luz necessário para formação da foto, com a quantidade e qualidade previamente estabelecidas, havendo, dessa maneira, um equilíbrio de entrada de luz. A seleção da opção auto na função *shutter*, apresentada na Figura 28, vista C, demonstra a escolha dessa configuração.



Figura 28 – Configuração adotada para coleta de dados  
Fonte: Elaborado pelo autor

Após a configuração, uma caixa estanque, referência Telesin modelo gp-wtp-901 (ver Figura 29, vistas D e E), que tem a finalidade de proteger a câmera contra água, sujeira, arranhões e pequenas quedas, foi acoplada num suporte tipo tripé com ventosa, consoante à Figura 29, vistas A, B e C.



Figura 29 – Caixa estanque gp-wtp-901 e suporte tipo tripé utilizados  
Fonte: Elaborado pelo autor

O tripé também possibilitou a fixação da câmera à carroceria do veículo Citroen, modelo C3, consoante ao que pode ser verificado nas Figuras 30.



Figura 30 – Instalação do tripé na estrutura veicular  
Fonte: Elaborado pelo autor

Por segurança, foi colocado um cabo de fitilho de polipropileno na cor cinza, que amarrou o tripé ao limpador de para-brisa do veículo, proporcionando um travamento anti-quedas, apresentado na Figura 31.

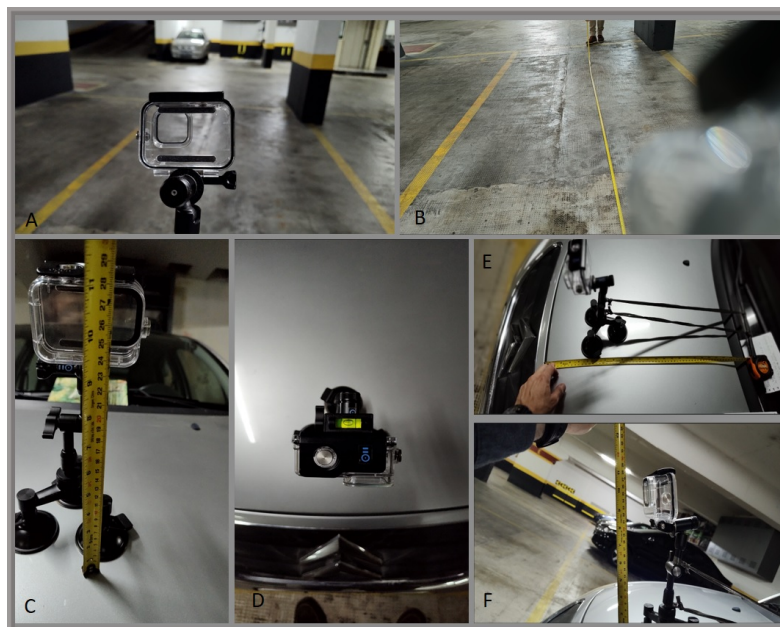


Figura 31 – Instalação do cabo de fitilho de polipropileno  
 Fonte: Elaborado pelo autor

Realizados os procedimentos descritos e antes da incorporação da câmera ao sistema, com o objetivo de verificar a estabilidade, o percurso foi realizado conforme apresentado nas Figuras 32 e 33, o que permitiu observar a eficácia da fixação e assim submeter a câmera também ao trajeto proposto.

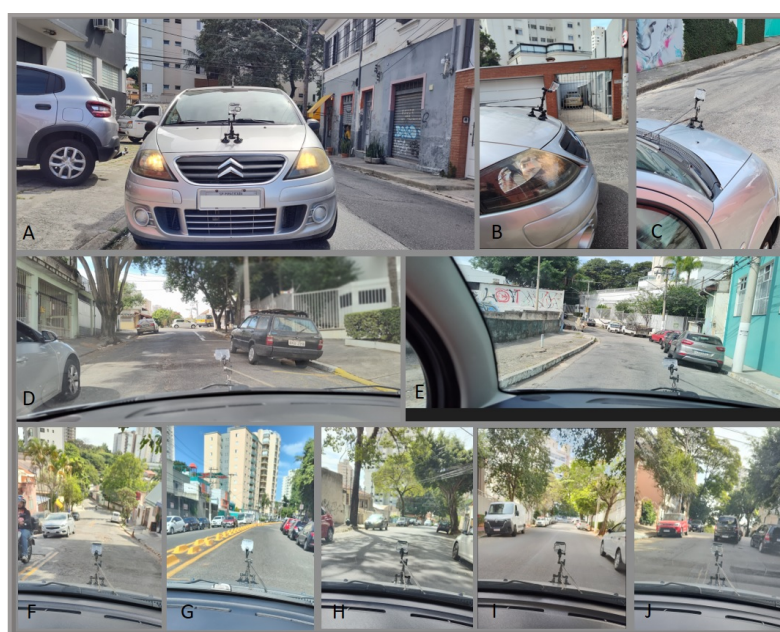


Figura 32 – Trajeto sendo realizado apenas com a caixa estanque, sem a câmera  
 Fonte: Elaborado pelo autor



Figura 33 – Trajeto realizado e pontos em que foi avaliada a estabilidade  
 Fonte: Adaptado de [GOOGLE MAPS \(2024\)](#)

### 3.1.2 Coleta de dados para o treinamento

O trajeto realizado foi de 2,8 Km e ocorreu entre os bairros Vila Anglo e Pompéia, situados na cidade de São Paulo, SP, em conformidade com o que está apresentado na Figura 34. Esse percurso foi escolhido por possuir um grande número de patologias.

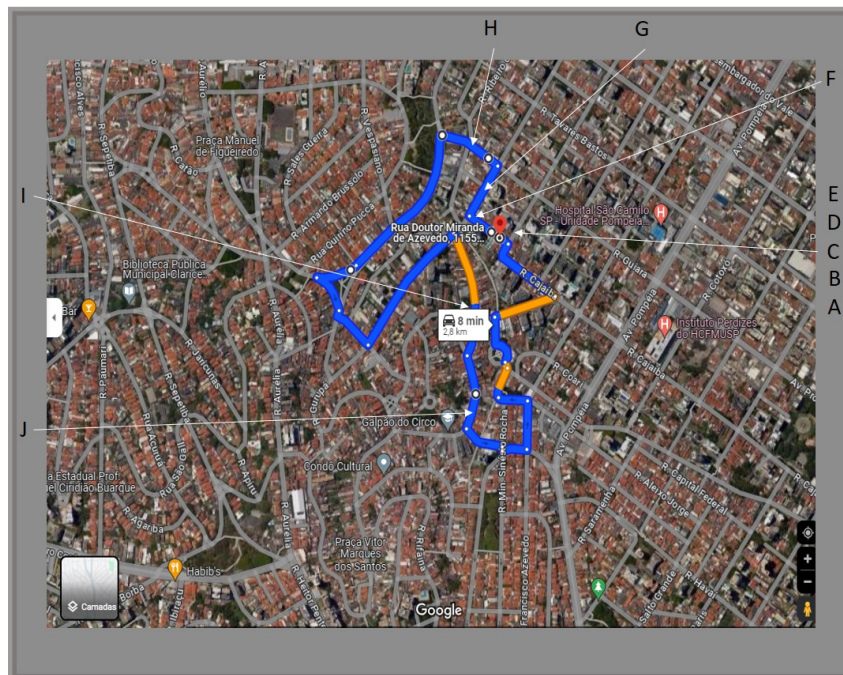


Figura 34 – Trajeto realizado e pontos de referência do trajeto  
 Fonte: Adaptado de [GOOGLE MAPS \(2024\)](#)

A câmera GoPro foi paramentada com a caixa estanque e acoplada ao tripé, que já

se encontrava instalado no veículo, vide Figura 35. A coleta de dados para o treinamento ocorreu no dia 10/05/2024, entre as 8h00 e 9h00.



Figura 35 – Coleta de imagens do dataset de treinamento  
Fonte: Elaborado pelo autor

Nesse período, as condições climáticas estavam estáveis e a temperatura oscilou entre 20° e 22°C para o horário estabelecido. A Figura 36 apresenta a localização relativa da estação de controle mais próxima, A701, do Instituto Nacional de Meteorologia (INMET), em relação ao local de coleta dos dados e o gráfico de temperatura em função do horário.

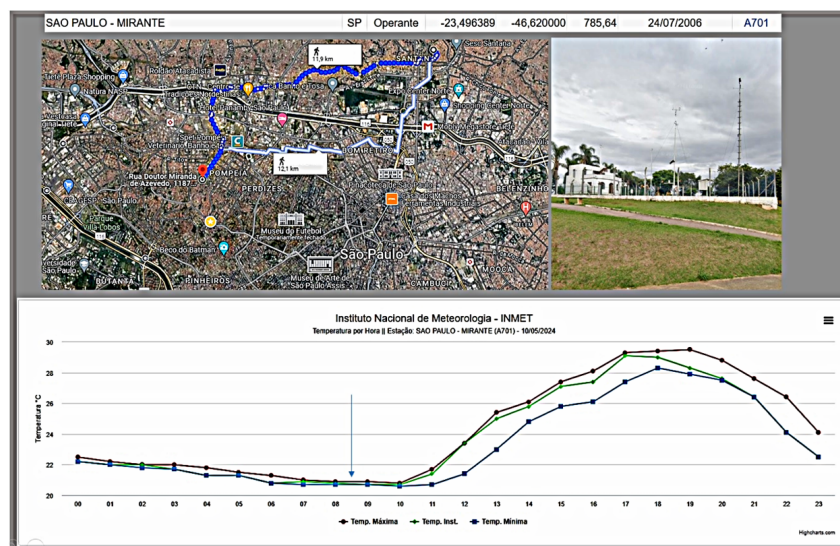


Figura 36 – Estação de controle automático A701  
Fonte: Adaptado de INMET (2024)

A umidade do ar oscilou entre 70% e 80 %, não sendo observada a ocorrência de chuva. Desse modo, o pavimento se encontrava livre de empoçamentos. A Figura 37 mostra a umidade relativa do ar em função do horário e a chuva acumulada em mm.

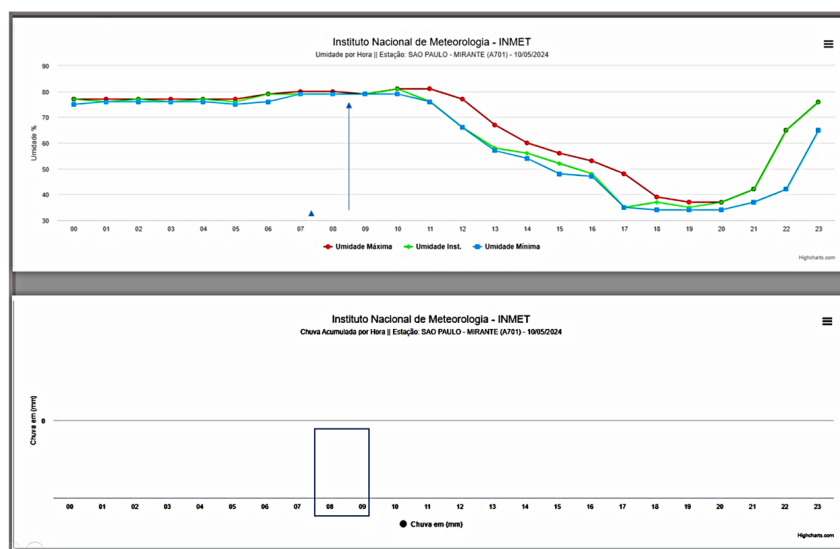


Figura 37 – Gráfico de umidade e chuva acumulada no dia 10/05/2024 entre 8h00 e 9h00 na cidade de São Paulo

Fonte: Adaptado de [INMET \(2024\)](#)

A velocidade do veículo oscilou entre 0 e 40 km/h, permitindo a captura de imagens por cerca de 11 minutos e 41 segundos ininterruptos.

### 3.1.3 Preparação do banco de imagens do treinamento

A preparação do banco de imagens adveio em parte da conversão do vídeo, a qual foi realizada com o software Free Video to JPG Converter, versão 5.1.1. Para tal propósito, optou-se pela configuração de conversão por FPS, o que gerou um total de 42.035 imagens. Essas foram analisadas visualmente e foi constatada a necessidade de exclusão de 2446 imagens, as quais apresentavam problemas de foco e nitidez, restando um total de 39.589 imagens, correspondente ao armazenamento de de 78,2 GB, Figura 38.

Ademais, foram separadas 2121 fotos com a patologia de fissuras, 2339 com as panelas e 1093 sem patologias, Figura 38.

O principal objetivo dessa etapa é prover ao software de imagens (CVAT) as imagens com as patologias escolhidas para anotação. Para isso, foram selecionadas fotos, em que a mesma patologia pudesse ser observada várias vezes e, preferencialmente, permitissem a circunscrição integral da patologia, de acordo com a evolução do vídeo. A Figura 39 exemplifica as fotos selecionadas.

A seleção, o armazenamento e a separação foram realizados no ambiente Windows, para, posteriormente, serem transferidos para o armazenamento em nuvem do *Google*.

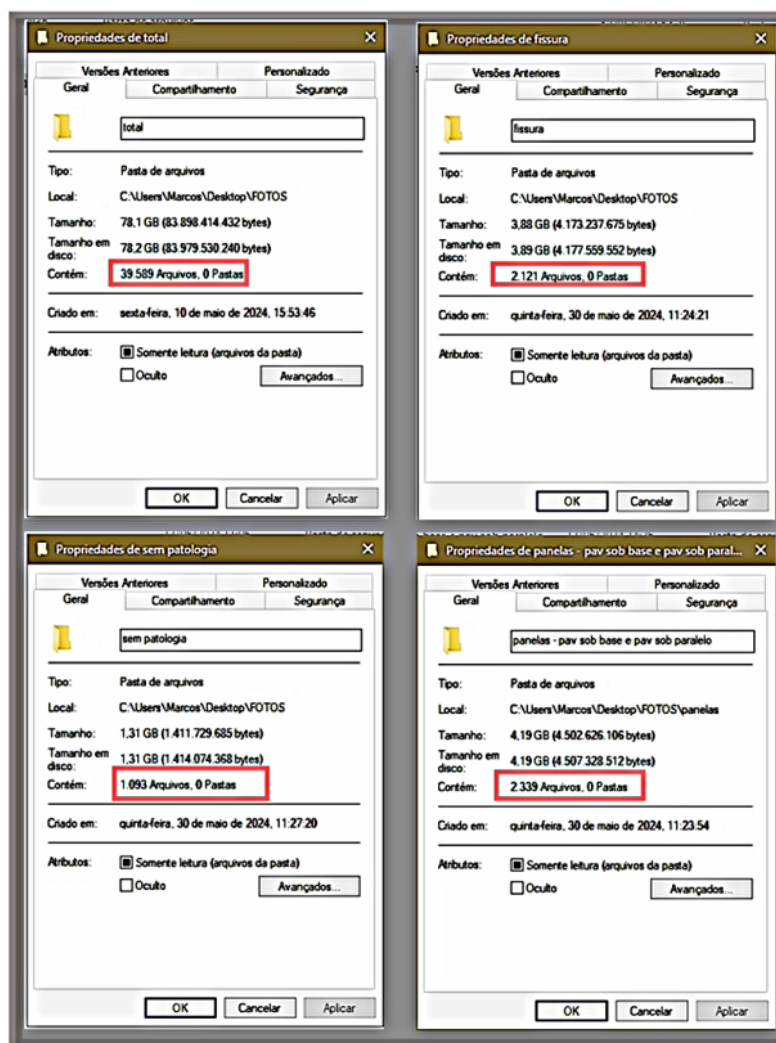


Figura 38 – Banco de imagens  
 Fonte: Elaborado pelo autor



Figura 39 – Exemplificação do agrupamento de imagens selecionadas para rotulação  
 Fonte: Elaborado pelo autor

### 3.1.4 Seleção de parte das imagens do banco de imagens

Consoante com Yao et al. (2019) e Russakovsky et al. (2015), a construção do banco de imagens deve envolver a diversibilidade, que é a variedade de visões do objeto e dos planos de fundo, assim como a escalabilidade, significando o equilíbrio de classes ao qual se almeja identificar.

Referente ao primeiro requisito, as imagens coletadas apresentaram as patologias de diferentes ângulos de visão, mas também, continham diferentes fundos de coleta, nesse sentido, essa propriedade foi obdecida.

Quanto ao segundo, os minibancos receberam uma quantidade equilibrada de imagens de cada patologia.

À vista disso, foram criados minibancos de imagens com diferentes quantidades. Logo, os minibancos de imagens 1, 2, 3 e 4 receberam respectivamente 22, 66, 240 e 306 imagens.

Por último, destaca-se que o treino 1, em todas suas subetapas, teve caráter exploratório, a fim de entender todo processo e assegurar o processamento do treino e da detecção, por meio da leitura correta do minidataset, da formulação da rotulação e da implementação do código de programação.

### 3.1.5 Rotulação

A rotulação é a etapa em que o objeto de interesse recebe sua demarcação ou delimitação. Para isso, existem algumas ferramentas que permitem realizar esse tipo de anotação.

Na internet é possível encontrar uma variedade de datasets preparados para detecção de objetos, representando um banco de imagens que já foi rotulado, como por exemplo o dataset COCO, que possui 91 categorias de objetos rotulados (LIN et al., 2014).

Apesar dessa facilidade, fissuras e panelas não pertencem à classe de objetos rotulados do COCO, ou mesmo, do *Google* Imagens, outro dataset pronto. Portanto, essas classes precisam ser delimitadas manualmente.

Russakovsky et al. (2015) discorreu sobre acurácia do rótulo, que é a habilidade humana de constatar e circunscrever o objeto de interesse. Logo, a rotulação foi realizada para os treinos 1,2,3 e 4, mas de 3 maneiras distintas, já que o objetivo era analisar a detecção obtida pelo modelo treinado em função do tipo de rotulação elaborada e, com isso, decidir quanto à forma mais eficaz.

Enfatiza-se que na rotulação do treino 1, o objetivo foi compreender como a ferramenta CVAT operava. Assim, foram verificados e constatados os aspectos relacionados às aberturas de projetos, às criações de tarefas (do inglês *tasks*), às adições de categorias de

rótulos, aos procedimentos de exportações das imagens e aos de importações das anotações geradas. Em vista disso, essa rotulação não teve a finalidade de analisar a detecção, mas sim compreender o processo.

Quanto ao treino 2, a rotulação buscou anotar as fissuras em pequenas áreas, em muitos casos até de forma individualizada, embora se apresentassem agrupadas. Em vista disso, foi evitada a sobreposição de categorias. Dessa forma, as patologias (panelas e fissuras) foram marcadas separadamente, havendo sobreposição em poucos casos, vide Figura 40 (à esquerda).

Para o 3, fez-se opção de agrupar as fissuras em blocos maiores e permitir a sobreposição de marcação. Assim, as panelas poderiam representar uma área dentro de uma região maior que estivesse com delimitação de fissuras, vide Figura 40 (à direita).

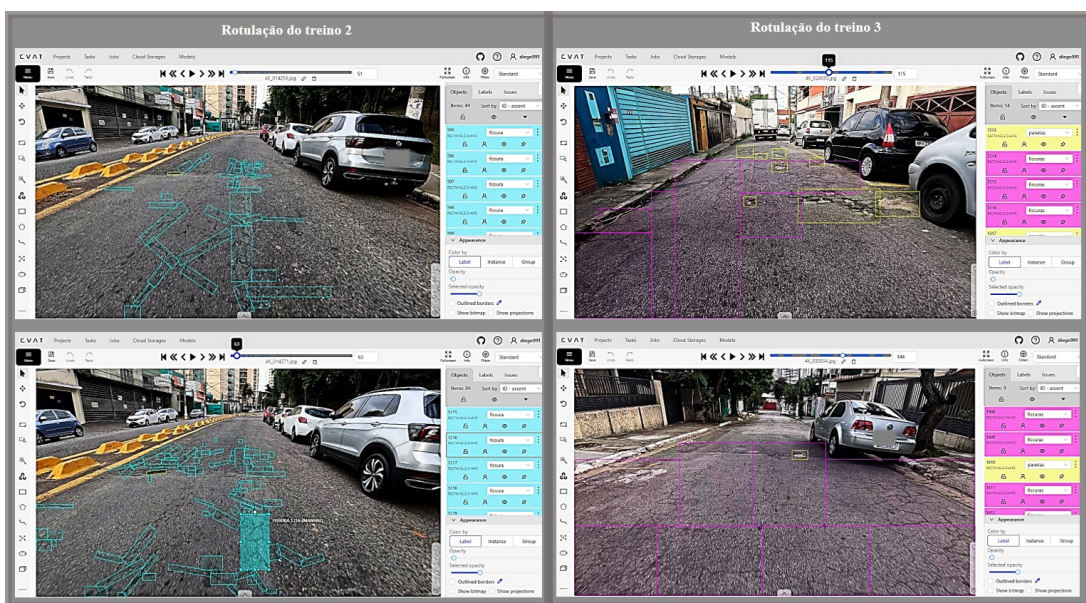


Figura 40 – Exemplo de rotulação para os treinos 2 e 3

Fonte: Adaptado de CVAT (2024)

Para o treino 4, as rotulações não foram propriamente realizadas: os arquivos gerados para a rotulação 2 e 3 foram unidos, como será apresentado na subetapa de preparação do minidataset.

Após término de um projeto, aqui entendido como a rotulação para um dos treinos, enviava-se as informações de marcação (categoria e as coordenadas dos boxes, ou, coordenadas dos bounding boxes em formato txt) para armazenamento no servidor local.

Para cada imagem, foi gerado um arquivo do tipo txt, contendo a lista de todas as rotulações realizadas, Figura 41

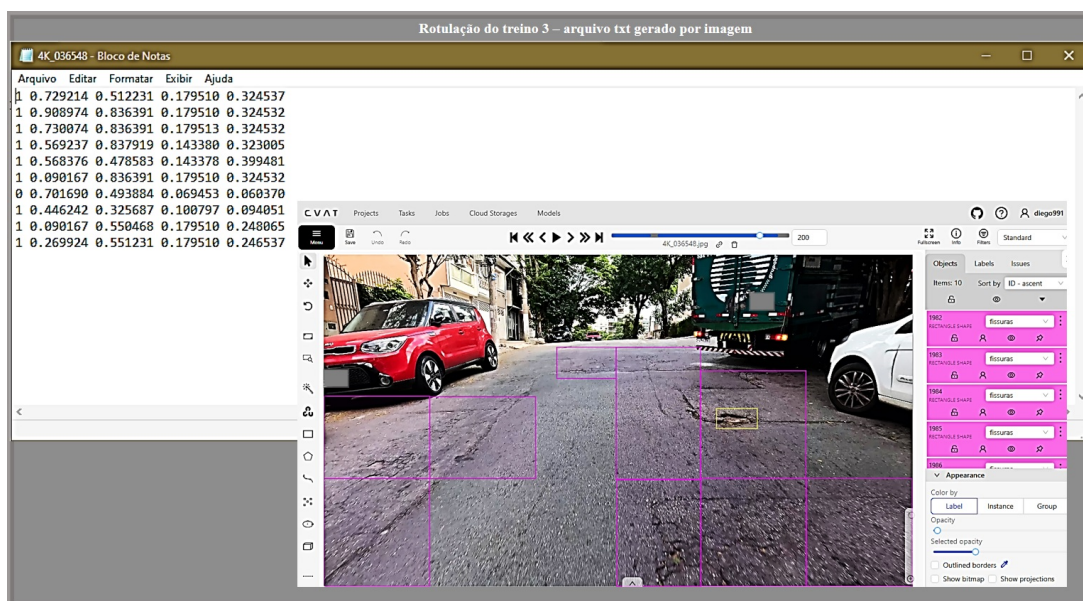


Figura 41 – Arquivo do tipo txt gerado para cada imagem  
 Fonte: Adaptado de CVAT (2024)

### 3.1.6 Preparação do minidataset

A preparação do minidataset sobreveio pela criação de uma estrutura de pastas, as quais dispuseram de todos os arquivos necessários à atividade de detecção. Portanto, inicialmente, ainda no ambiente Windows, foi criada uma pasta com o nome Dataset. Dentro dessa, foram geradas outras duas pastas com as nomenclaturas Main e Teste.

Além disso, a pasta Main recebeu outras duas pastas, com os respectivos nomes: Images e Labels, ambas também receberam outras duas, com os nomes de: Train e Val.

Após montada essa estrutura, partiu-se para distribuição das imagens do minibanco e dos arquivos txt baixados na etapa anterior, conforme apresentado na Tabela 2.

Para o treino 1, a Train, pertencente a Images, recebeu 9 imagens com patologias do tipo panelas e 9 imagens com patologias do tipo fissuras; já a Val recebeu 4 imagens, sendo 2 de fissuras e 2 de panelas.

A Train, pertencente a Label, recebeu os arquivos do tipo txt referentes às imagens que foram colocadas na Train de Images; da mesma forma, a Val recebeu os arquivos do tipo txt das imagens que foram colocadas na Val de Images.

A pasta Teste recebeu 15 imagens, sendo : 5 de fissuras, 5 de panelas e 5 sem patologias.

Após realizada essa etapa, a estrutura foi transferida para o ambiente de armazenamento em nuvem do *Google* (*Google Drive*).

Para cada treino essa montagem foi realizada da mesma forma, apenas mudando a quantidade de imagens que cada subpasta recebeu, conforme apresentado na Tabela 2:

Tabela 2 – Formatação e composição dos minidatasets

Train	Pasta	Subpasta	Subpasta	Qtde.	Subpasta	Qtde.
1	Dataset	Images	Train	18	Val	4
1	Dataset	Labels	Train	18	Labels	4
1	Teste			15		
2	Dataset	Images	Train	55	Val	11
2	Dataset	Labels	Train	55	Labels	11
2	Testes			15		
2	Dataset	Images	Train	200	Val	40
2	Dataset	Labels	Train	200	Labels	40
2	Testes			15		
2	Dataset	Images	Train	255	Val	51
2	Dataset	Labels	Train	255	Labels	51
2	Teste			15		

Fonte: Elaborado pelo autor

Dentro da pasta dataset também foi criado um arquivo do tipo txt, com o nome "Config.yaml", o qual recebeu o caminho de origem da pasta dataset, da pasta de imagens train, da pasta de imagens val e as classes de patologias existentes.

No arquivo config. yaml, a classe de patologia de painelas recebeu o código 0, enquanto o de fissuras o 1.

### 3.1.7 Treinamento

O processamento ocorreu no ambiente virtual do *Google (Colab)* com utilização da linguagem de programação Python.

Devido ao nível de processamento, optou-se pela utilização da unidade de processamento gráfico (do inglês *Graphics Processing Units- GPU*) da NVÍDEA na configuração A100, vide Figura 42, com utilização da arquitetura de dispositivos unificados de computação (do inglês *Compute Unified Device Architecture- CUDA*) na versão 12.2.



Figura 42 – Configuração de processamento escolhida  
Fonte: Adaptado de [GOOGLE COLAB \(2024\)](#)

Após elaborada a programação de acesso ao dataset, foi realizada a instalação da biblioteca da Ultralytics para utilização da rede AP YOLO V8 na versão short, conforme Figura 43, a qual operou com utilização de 22 camadas de processamento e 50 épocas de treinamento, conforme Figura 44.

```

from ultralytics import YOLO
model=YOLO("yolov8s.pt")
model.train(data="/content/drive/MyDrive/dataset_1/config.yaml", epochs=50, imgsz=640)

...
Downloading https://github.com/ultralytics/assets/releases/download/y8.2.8/yolov8s.pt to 'yolov8s.pt'...
100% [██████████] 21.50/21.50 [00:00:00:00, 43998B/s]
Ultralytics YOLOv8.2.38 Python-3.10.12 torch-2.3.8+cu121 CUDA:0 (NVIDIA A100-SXM4-40GB, 40514MiB)
engine/trainer: task=detect, mode=train, model=yolov8s.pt, data=/content/drive/MyDrive/dataset_1/config.yaml, epochs=50, time=None, patience=100, batch=16, imgsz=640, save=True, save_dir=runs/detect/train
Downloading https://ultralytics.com/assets/Arial.ttf to '/root/.config/Ultralytics/Arial.ttf'...
100% [██████████] 755k/755k [00:00:00:00, 99.0MB/s]
Overriding model.yaml nc=80 with nc=2

   from  n  params module  arguments
   ---  --  -
0      -1  1     928 ultralytics.nn.modules.conv.Conv [3, 32, 3, 2]
1      -1  1    1856 ultralytics.nn.modules.conv.Conv [32, 64, 3, 2]
2      -1  1   29056 ultralytics.nn.modules.block.C2F [64, 64, 1, True]
3      -1  1   73984 ultralytics.nn.modules.conv.Conv [64, 128, 3, 2]
4      -1  2  197632 ultralytics.nn.modules.block.C2F [128, 128, 2, True]
5      -1  1  295424 ultralytics.nn.modules.conv.Conv [128, 256, 3, 2]
6      -1  2  788480 ultralytics.nn.modules.block.C2F [256, 256, 2, True]
7      -1  1 1180672 ultralytics.nn.modules.conv.Conv [256, 512, 3, 2]
8      -1  1 1838080 ultralytics.nn.modules.block.C2F [512, 512, 1, True]
9      -1  1  656096 ultralytics.nn.modules.block.SPPF [512, 512, 5]
10     -1  1     0 torch.nn.modules.upsampling.Upsample [None, 2, 'nearest']
11     [-1, 6] 1     0 ultralytics.nn.modules.conv.Concat [1]
12     -1  1   591360 ultralytics.nn.modules.block.C2F [768, 256, 1]
13     -1  0     0 torch.nn.modules.upsampling.Upsample [None, 2, 'nearest']
14     [-1, 4] 1     0 ultralytics.nn.modules.conv.Concat [1]
15     -1  1  148224 ultralytics.nn.modules.block.C2F [384, 128, 1]
16     -1  1  147712 ultralytics.nn.modules.conv.Conv [128, 128, 3, 2]
17     [-1, 12] 1     0 ultralytics.nn.modules.conv.Concat [1]
18     -1  1  493056 ultralytics.nn.modules.block.C2F [384, 256, 1]
19     -1  1  590336 ultralytics.nn.modules.conv.Conv [256, 256, 3, 2]
20     [-1, 9] 1     0 ultralytics.nn.modules.conv.Concat [1]
21     -1  1 1969152 ultralytics.nn.modules.block.C2F [768, 512, 1]
22     [15, 18, 21] 1 2116822 ultralytics.nn.modules.head.Detect [2, [128, 256, 512]]
    
```

Figura 43 – Rede YOLO V8 e suas camadas  
 Fonte: Adaptado de GOOGLE COLAB (2024)

```

Image sizes b40 train, b40 val
Using 8 dataloader workers
Logging results to runs/detect/train
Starting training for 50 epochs...
    
```

Epoch	GPU_mem	box_loss	cls_loss	df1_loss	Instances	Size	
1/50	4.05G	2.802	3.962	2.499	42	640: 100% ██████████ 13/13 [00:08:00:00, 1.45it/s]	
	Class	Images	Instances	Box(P)	R	mAP50 mAP50-95): 100% ██████████ 2/2 [00:08:00:00, 4.03s/it] all	
2/50	3.9G	2.422	2.682	2.193	34	640: 100% ██████████ 13/13 [00:01:00:00, 6.83it/s]	
	Class	Images	Instances	Box(P)	R	mAP50 mAP50-95): 100% ██████████ 2/2 [00:00:00:00, 4.26it/s]	
		all	40	281	0.614	0.269	0.118 0.0472
3/50	4.8G	2.286	2.389	2.022	24	640: 100% ██████████ 13/13 [00:03:00:00, 3.41it/s]	
	Class	Images	Instances	Box(P)	R	mAP50 mAP50-95): 100% ██████████ 2/2 [00:00:00:00, 7.25it/s] all	
4/50	4.2G	2.138	2.215	1.931	38	640: 100% ██████████ 13/13 [00:02:00:00, 5.14it/s]	
	Class	Images	Instances	Box(P)	R	mAP50 mAP50-95): 100% ██████████ 2/2 [00:00:00:00, 6.64it/s] all	

Figura 44 – Treino 3 - processamento  
 Fonte: Adaptado de GOOGLE COLAB (2024)

Ainda no Colab, foi executada a plotagem da detecção inicial alcançada pelo modelo treinado, quando aplicado às imagens da pasta teste, como pode ser visualizado na Figura 45. Posteriormente, as principais métricas de desempenho do modelo treinado

foram avaliadas para as fases train e val. Após isso, todos os arquivos do *Google Colab* foram transferidos para o armazenamento virtual do *Google*.

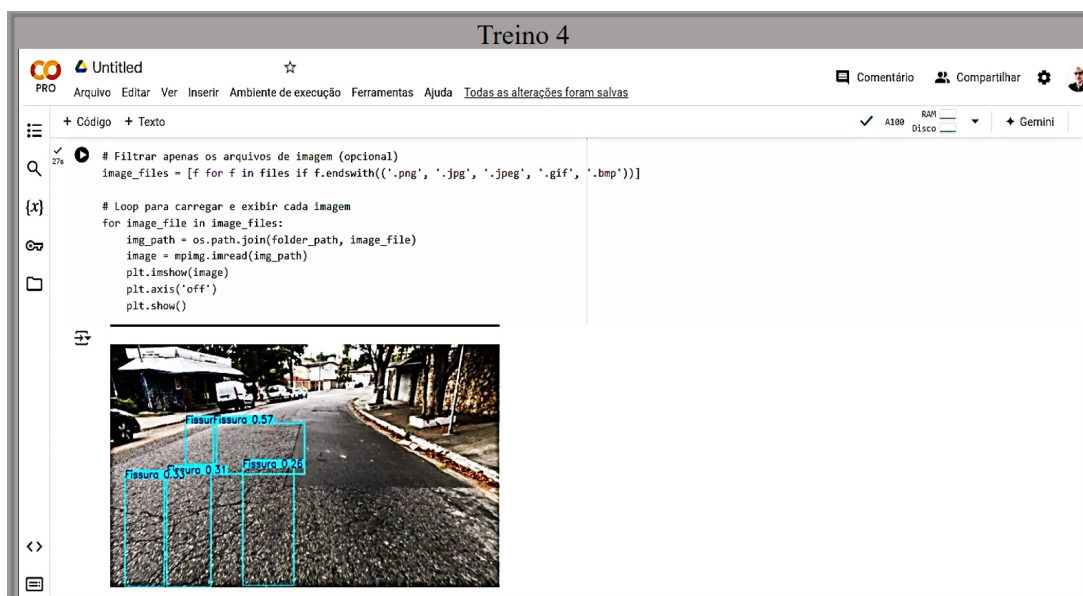


Figura 45 – Exemplo de detecção inicial alcançada

Fonte: Adaptado de [GOOGLE COLAB \(2024\)](#)

### 3.1.8 Avaliação do treinamento

Nessa etapa foi analisada a performance das fases train e val, que conjuntamente fazem parte dos treinos. A fase train é utilizada para treinar o modelo através do conjunto de dados disponibilizado, ou seja, o modelo aprende utilizando o conjunto de dados fornecidos.

A fase val é utilizada para validar o modelo após o treino. Logo, o modelo é medido em relação à sua precisão e desempenho geral, podendo ser ajustado para melhorar o seu desempenho.

Em vista disso, foram analisadas as seguintes métricas: perda das caixas delimitadoras e probabilidade predita de classe, para as fases de train e val.

Singularmente para fase val, foram analisadas as seguintes métricas: precisão geral, recall, precisão com confiança de 50%, precisão com confiança de 95%, teste de detecção e matriz de confusão.

O gráfico da Figura 46, indica a métrica de perda das caixas delimitadoras, representando o somatório do desvio da confiança da classe, com a soma do desvio das coordenadas das caixas delimitadoras com maior índice IOU para os treinos 2, 3 e 4.

Assim, para a fase train, foi possível verificar uma tendência de redução em função das épocas para todos os treinos, sendo também possível observar a menor perda para o treino 3. Já na fase val, os treinos 3 e 4 forneceram as menores perdas.

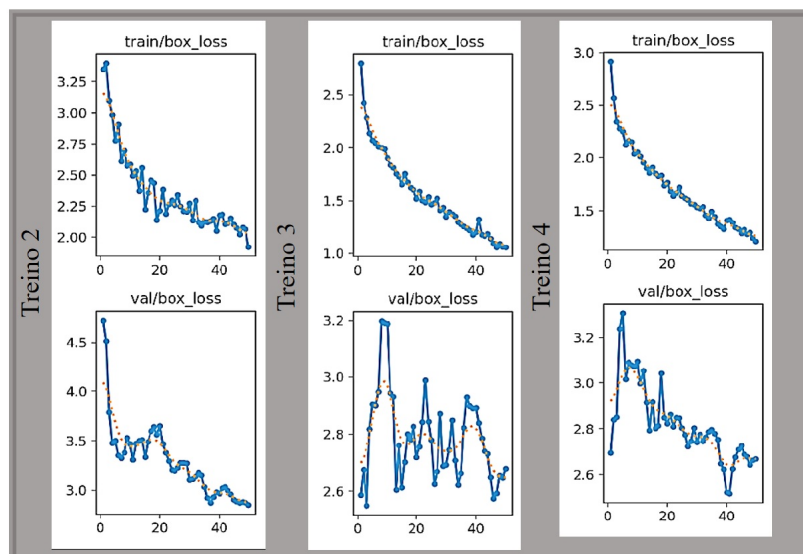


Figura 46 – Perda das caixas delimitadoras para os treinos 2, 3 e 4  
 Fonte: Extraído pelo autor do YOLO V8

A Figura 47 apresentou o gráfico de probabilidade predita de classe para os treinos 2, 3 e 4, nas fases de train e val, também para esse caso, quanto mais próximo de zero, menor o desvio do indicador.

A análise desse parâmetro, para a fase train, permitiu constatar que os treinos 3 e 4 obtiveram os melhores desempenhos, abaixo de 1. No que se refere à fase val, o treino 2 obteve o melhor índice, ficando próximo a 2.

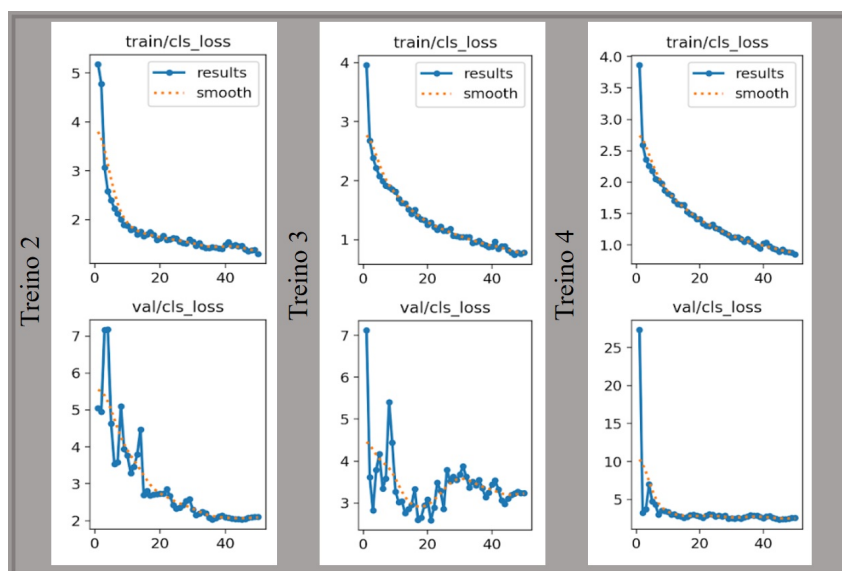


Figura 47 – Probabilidade predita de classe para os treinos 2, 3 e 4  
 Fonte: Extraído pelo autor do YOLO V8

Com relação à precisão geral (*Precision*) e à precisão com confiança de 50% (mAP50), Figura 48, foi possível notar que os treinos 3 e 4 obtiveram os melhores valores finais para ambos indicadores, permanecendo entre 0,6 e 0,8 para precisão geral e entre

0,4 e 0,5 para precisão com confiança de 50%.

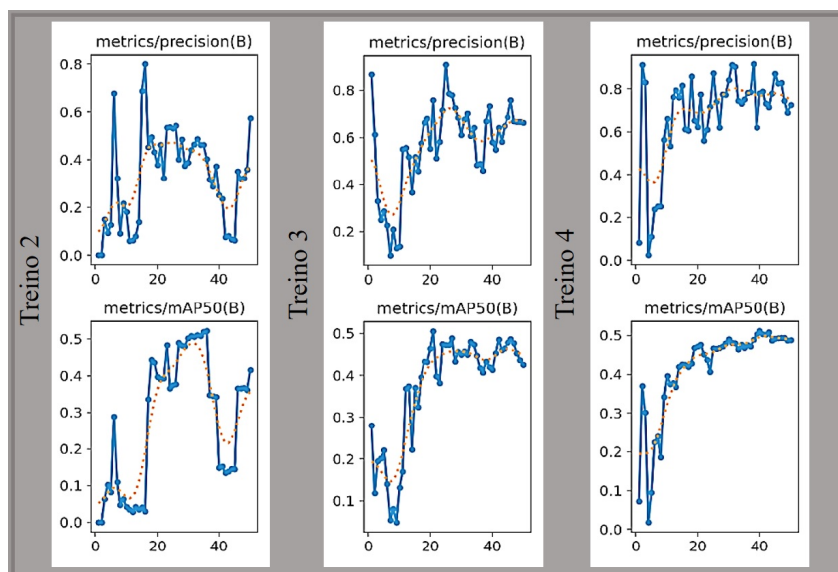


Figura 48 – Precisão geral e com 50% de confiança para os treinos 2, 3 e 4  
Fonte: Extraído pelo autor do YOLO V8

No gráfico da Figura 49, que apresenta a capacidade de identificação da rede quando submetida à detecção de vários objetos diferentes (recall), o treino 4 obteve o melhor resultado, finalizando entre 0,5 e 0,45. Em relação à precisão com confiança de 95% (mAP50-95), foi possível observar que o treino 4 também alcançou o melhor índice, que terminou entre 0,25 e 0,2.

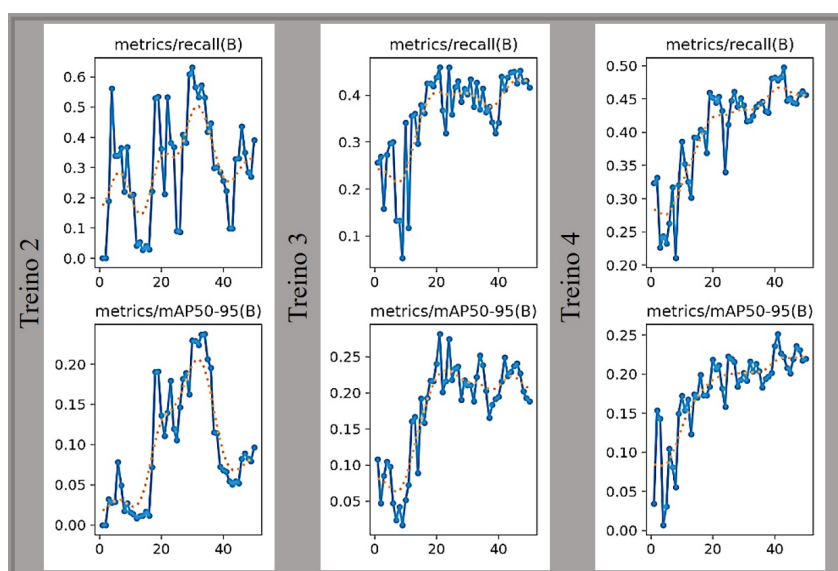


Figura 49 – Recall e precisão com confiança de 95% para os treinos 2, 3 e 4  
Fonte: Extraído pelo autor do YOLO V8

A detecção inicial, que mostra a eficácia dos modelos treinados quando aplicados às imagens de testes, Figura 50, mostrou que o treino 3 identificou o maior número de patologias.

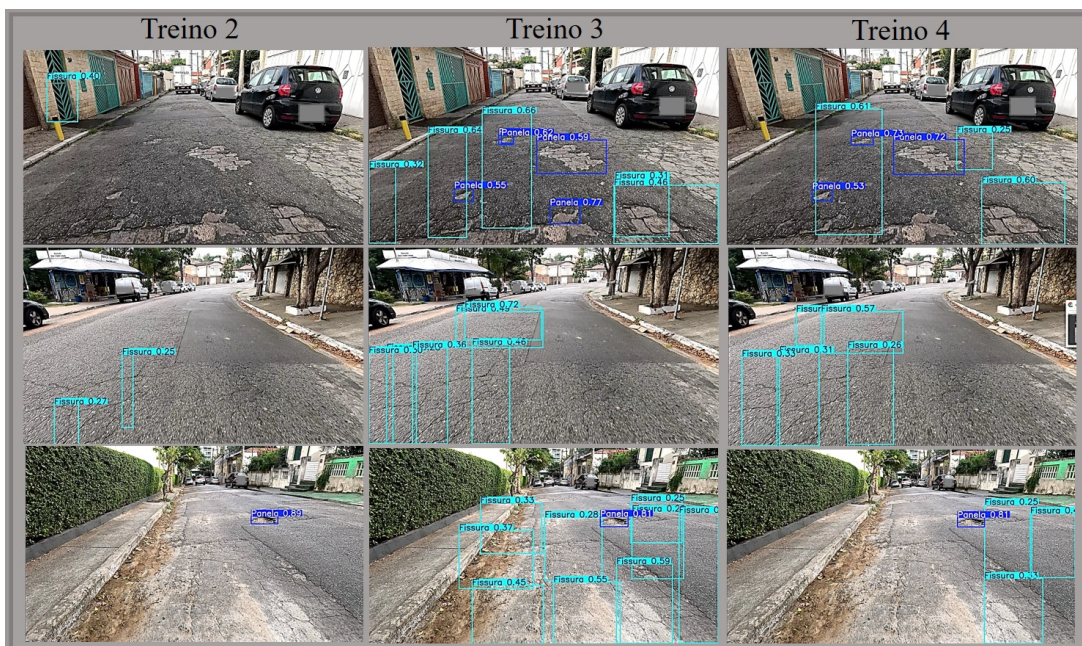


Figura 50 – Detecção para os treinos 2, 3 e 4 em imagens específicas  
 Fonte: Extraído pelo autor do YOLO V8

Referente a matriz de confusão do treino 2, Figura 51, foi possível verificar que 2 painelas e 25 fissuras foram identificadas corretamente, enquanto 34 objetos foram identificados erroneamente como fissuras . Além disso, 1 painela e 194 fissuras não foram identificadas por esse modelo.

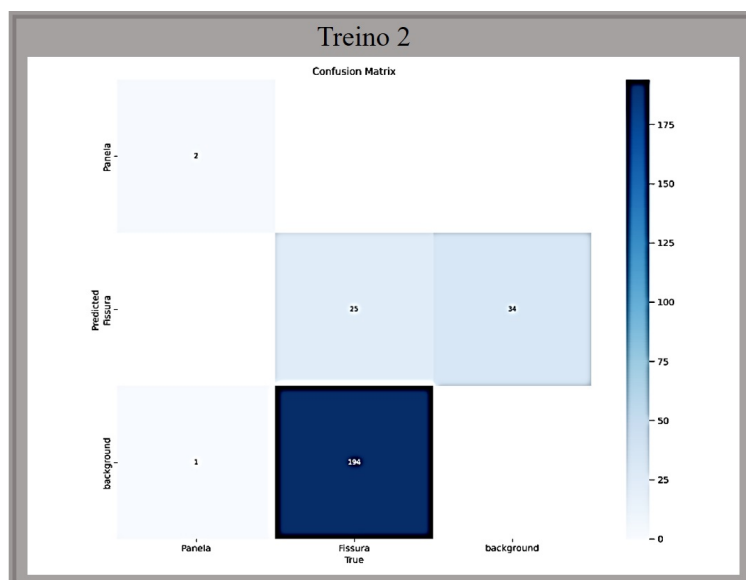


Figura 51 – Matriz de confusão para o treino 2  
 Fonte: Extraído pelo autor do YOLO V8

Para a matriz de confusão do treino 3, Figura 52, notou-se que 31 painelas e 53 fissuras foram identificadas corretamente, enquanto 7 painelas e 35 fissuras foram identificadas incorretamente. Ademais, 8 painelas e 189 fissuras não foram indentificadas.

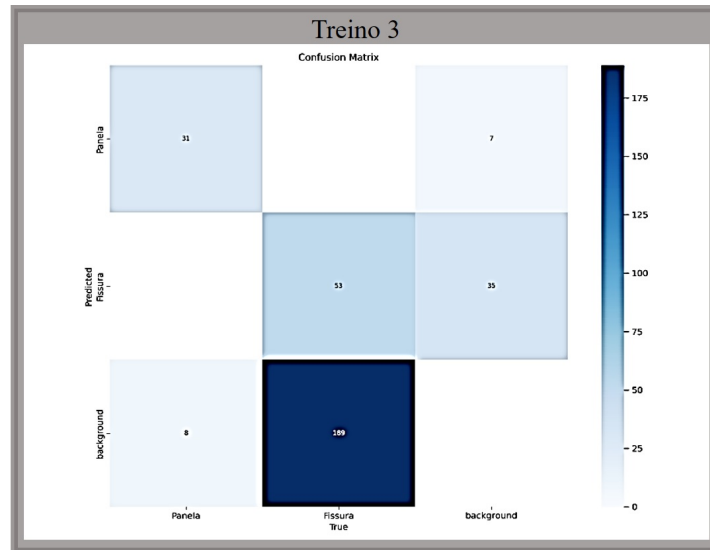


Figura 52 – Matriz de confusão para o treino 3  
Fonte: Extraído pelo autor do YOLO V8

Referente ao treino 4, Figura 53, foi constatado que 34 painelas e 64 fissuras foram identificadas corretamente, enquanto 30 fissuras foram identificadas incorretamente, além de que 8 painelas e 397 fissuras não foram identificadas.

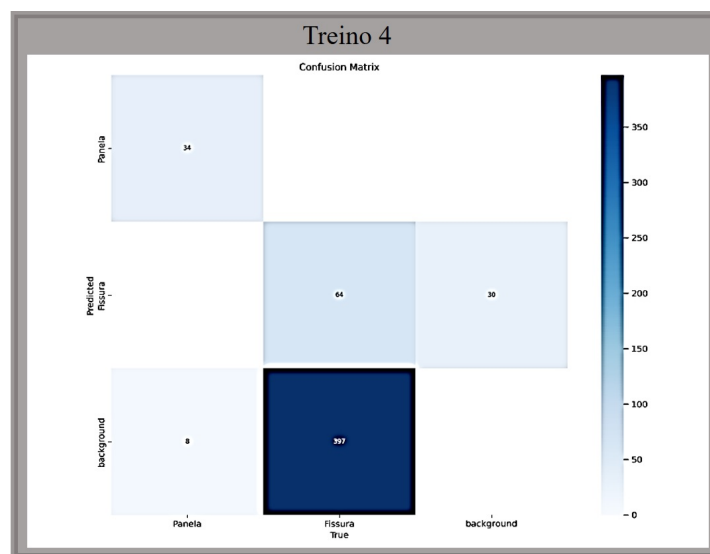


Figura 53 – Matriz de confusão para o treino 4  
Fonte: Extraído pelo autor do YOLO V8

Nesse sentido, notou-se que quando comparados os treinos 2 e 3, a identificação incorreta de painelas é 0 para o treino 2 e 7 para o treino 3, enquanto a de fissuras é 34 para o treino 2 e 35 para o treino 3. Além disso, a não identificação de painelas é de 1 para o treino 2 e de 8 para o treino 3, ao mesmo tempo que o de fissuras é de 194 para o treino 2 e de 189 para o treino 3.

Ainda no tocante das matrizes de confusão, foi possível verificar que a identificação incorreta de painelas do modelo 4 superou ao somatório dos treinos 2 e 3, assim como, a

não identificação de fissuras. Além disso, para o treino 4, a indentificação adequada de fissuras ficou abaixo do somatório alcançado pelos treinos 2 e 3.

Por último, a matriz de confusão do treino 3, apresenta o maior número de identificações corretas em termos proporcionais.

A avaliação do treinamento inicial permitiu constatar que o indicador relacionado à perda das caixas delimitadoras foi melhor sucedido para o treino 3, quando observada a fase train. Atinente à fase val, houve o mesmo desempenho para os treinos 3 e 4.

No que se refere ao indicador de probabilidade predita de classe, observou-se que os treinos 3 e 4 produziram os melhores resultados para fase train. Para a fase val, o treino 2 teve o melhor resultado.

No quesito precisão geral e precisão com confiança de 50%, os treinos 3 e 4 produziram os melhores resultados. Já, quanto ao recall e à precisão com confiança de 95%, o treino 4 apresentou o melhor desempenho.

Quando analisados o teste de detecção inicial e a matriz de confusão, o modelo estruturado no treino 3 atingiu o melhor desempenho.

Por conseguinte, o padrão de rotulação para a detecção final utilizará a forma de rotulação produzida no treino 3, o qual obteve a melhor performance, quando analisados todos os parâmetros conjuntamente. Na Tabela 3, as células das fases train e val, foram preenchidas com um "X" quando o treino avaliado obteve o melhor desempenho.

Tabela 3 – Avaliação do treinamento

Treino	2		3		4	
	train	val	train	val	train	val
Perda das caixas delimitadoras			x	x		x
Probabilidade predita de classe		x	x		x	
Precisão geral				x		x
Precisão com confiança de 50%				x		x
Recall						x
Precisão com confiança de 95%						x
Teste de detecção				x		
Matriz de confusão				x		
<b>Total</b>	<b>1</b>		<b>7</b>		<b>6</b>	

Fonte: Elaborado pelo autor

Logo, cumpri-se nessa seção o objetivo específico de indicar por meio de uma etapa de detecção inicial a forma de rotulação mais eficaz para a detecção de painéis e fissuras, a qual será utilizada para rotulação da detecção final.

## 3.2 Detecção final

Essa etapa tem por finalidade apresentar todos materiais e métodos utilizados na etapa de detecção final, considerando a avaliação realizada na etapa inicial.

### 3.2.1 Configuração, instalação e coleta de dados para treinamento

A coleta de dados para o treinamento foi realizada com o mesmo modelo de câmera utilizada na seção 3.1.1 e 3.1.2 . Desse jeito, foi utilizada uma câmera de ação GoPro Hero 12 Black.

Também foi mantida a configuração utilizada na seção 3.1.1, conforme Figura 54. Logo, a resolução selecionada foi de 3840 pixels na horizontal por 2160 pixels na vertical (4K), sendo o perfil de iluminação em modo natural. Ademais, a proporção de aspecto foi de 16 por 9.

A taxa de quadros por segundo foi de 60 fps e campo de visão foi ajustado para horizonte. A estabilização foi a de formato digital, na função *AutoBoost*.

Foi selecionado o modo *Shutter* para o tempo de abertura e fechamento do obturador do equipamento, permanecendo em consonância com a taxa de frames e o tempo de luz necessário para formação da foto.

Diferente do teste inicial, foi selecionada a opção de alta nitidez, para configuração de filmagem, sendo mais uma melhoria implementada após a fase de detecção inicial, como pode ser visto na Figura 54, na parte inferior à direita.



Figura 54 – Configuração adotada para detecção  
Fonte: Elaborado pelo autor

Após a configuração, o suporte tipo tripé com ventosa, foi fixado na parte da carroceria do veículo Citroen C3. Por segurança, um cabo de fitilho de polipropileno na cor

cinza, foi amarrado do tripé à antena do veículo, proporcionando um mecanismo contra quedas.

Realizados os procedimentos descritos, antes da montagem da caixa estanque e da câmera à carroceria, com o objetivo de verificar a estabilidade, o percurso foi realizado até a chegada da Estrada Vicinal da Represa de Itupararanga, situada na cidade de Votorantim no estado de São Paulo. Durante o trajeto até a estrada, foi possível observar a segurança e a estabilidade do suporte para as velocidades que, posteriormente, seriam utilizadas para a coleta de dados. conforme apresentado na Figura 55 (a).

A caixa estanque e a câmera foram fixadas ao tripé no início do percurso, conforme registro apresentado na Figura 55 (b).



Figura 55 – Verificação de estabilidade e instalação da caixa estanque e da câmera de ação  
Fonte: Elaborado pelo autor

O percurso, conforme apresentado na Figura 56, ocorreu entre as cidades de Votorantim (SP), indicada com a letra (a), e Piedade (SP), indicada com a letra (d), no dia 12 de novembro de 2024, entre as 8h22 e as 9h32, em que foram percorridos de 28,3 Km.

O trajeto iniciou pela Estrada Vicinal da Represa de Itupararanga e passou pela Estrada Vicinal Carolina Paes Granjero.

Durante o percurso foi necessária a realização de duas entrepausas, correspondentes aos locais (b) e (c) também indicados na Figura 56, para resfriamento da câmera de ação. Esses momentos, assim como os de início (a) e fim (d), foram registrados e estão apresentados na Figura 57.

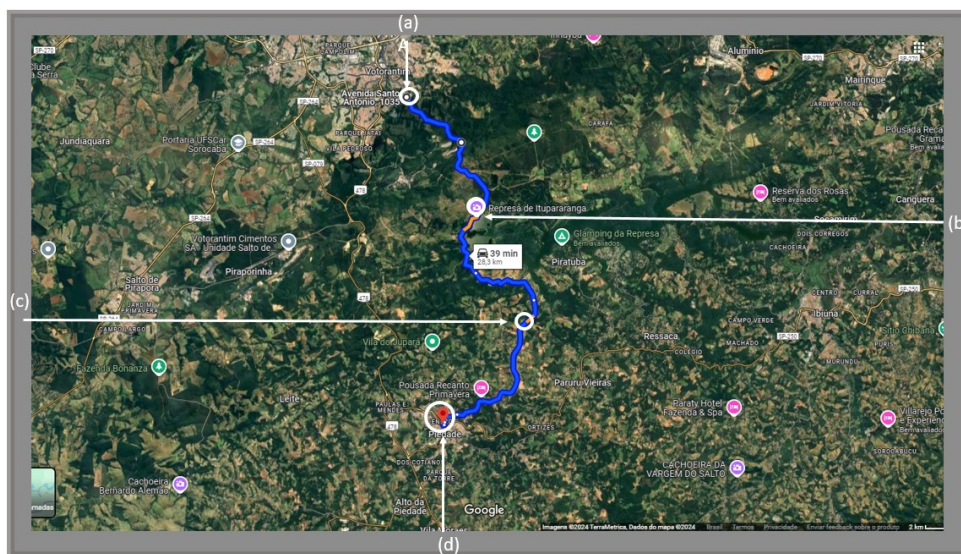


Figura 56 – Trajeto entre as cidades de Votorantim e Piedade em São Paulo  
 Fonte: Adaptado de [GOOGLE MAPS \(2024\)](#)



Figura 57 – Locais de início (a), término (b) e parada para resfriamento (b) e (c)  
 Fonte: Elaborado pelo autor

Na ocasião da coleta dos dados, as condições climáticas estavam estáveis, não sendo observada a ocorrência de chuva. Desse modo, o pavimento se encontrava livre de empoçamentos. O dia permaneceu ensolarado com temperatura média próxima de 22°C para o dia, ver Figura 58 (d); já a umidade média do ar foi 75 %, ver Figura 58 (c); a quantidade de chuva acumulada diária foi igual a zero, conforme Figura 58 (b). Todos os dados se referem à estação de controle climático ETEC V0455 pertencente ao CIIAGRO, cuja localização segue indicada pela letra (a), na Figura 58.

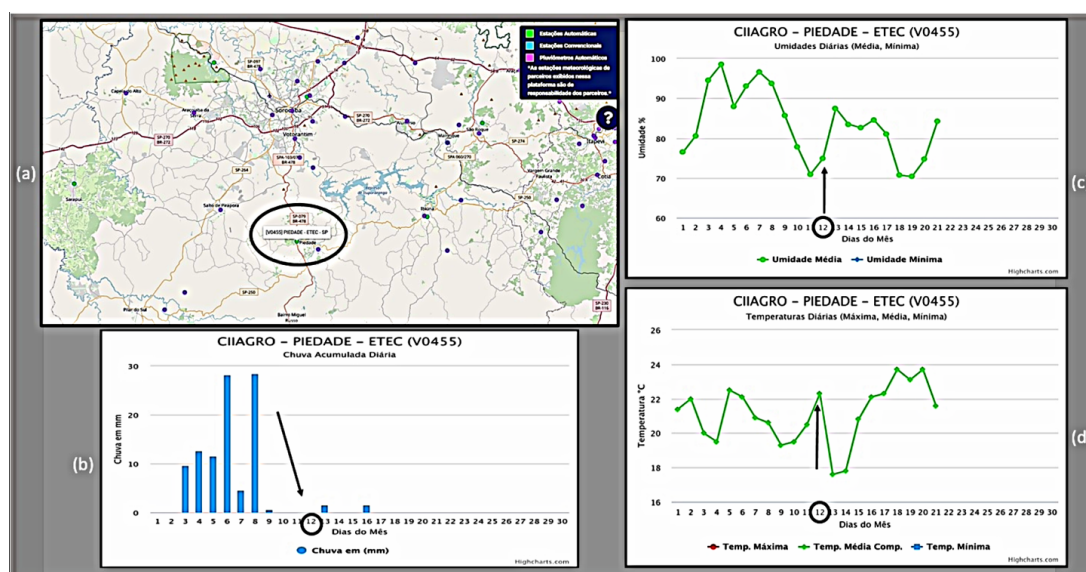


Figura 58 – Condições climáticas em Votorantim, SP

Fonte: Adaptado de CIIAGRO (2024)

O vídeo foi gravado a uma velocidade variável devido às condições de tráfego. A velocidade do veículo oscilou entre 0 e 80 km/h, permitindo a captura de imagens por cerca de 50 minutos e 18 segundos. O primeiro trecho a ser filmado, entre os locais a e b, conforme Figura 56, teve duração de 16 minutos e 17 segundos. O subsequente, entre b e c, teve duração de 17 minutos, enquanto o trecho de c e d durou 17 minutos e 1 segundo.

### 3.2.2 Preparação do banco de imagens e rotulação

Para a preparação do banco de imagem, os vídeos foram convertidos em fotos através do software Free Video to JPG Converter, versão 5.1.1, conforme pode ser visto na Figura 59 (a). A conversão de todos os vídeos gerou um total de 119.759 imagens, proporcionando 21,5 GB de armazenamento para o trecho AB, 10,6 GB para o trecho BC e 10 GB para o trecho CD, Figura 59 (b).

Todas as imagens foram analisadas visualmente. Assim, foram selecionadas 300 imagens contendo fissuras e panelas ao mesmo tempo, sendo essas imagens copiadas numa pasta com o mesmo nome. O mesmo foi feito com a patologia fissura e com a patologia panela, que receberam individualmente 100 imagens, cada qual com sua patologia.

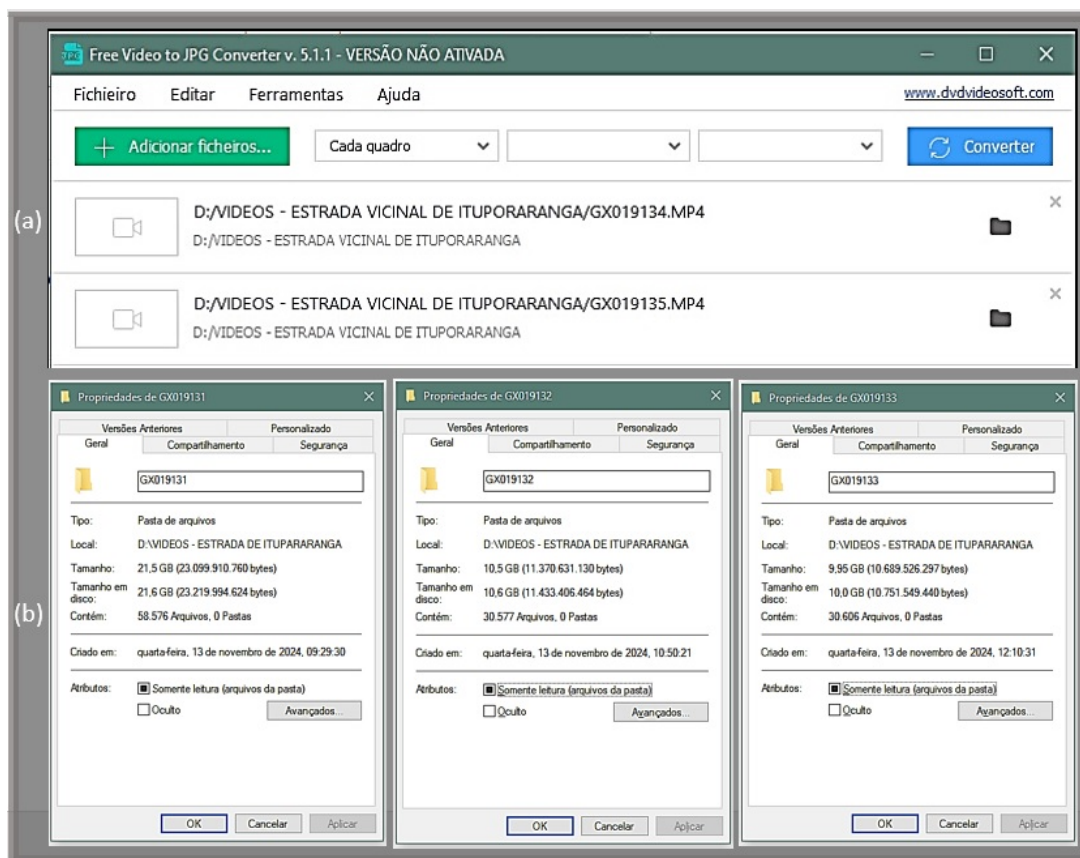


Figura 59 – Banco de imagens da 2ª etapa  
 Fonte: Elaborado pelo autor

Além disso, como forma de melhoria do processo de treinamento, foram separadas 100 imagens sem patologias, as quais funcionaram como background do treinamento. Todo esse procedimento de armazenamento está registrado na Figura 60.

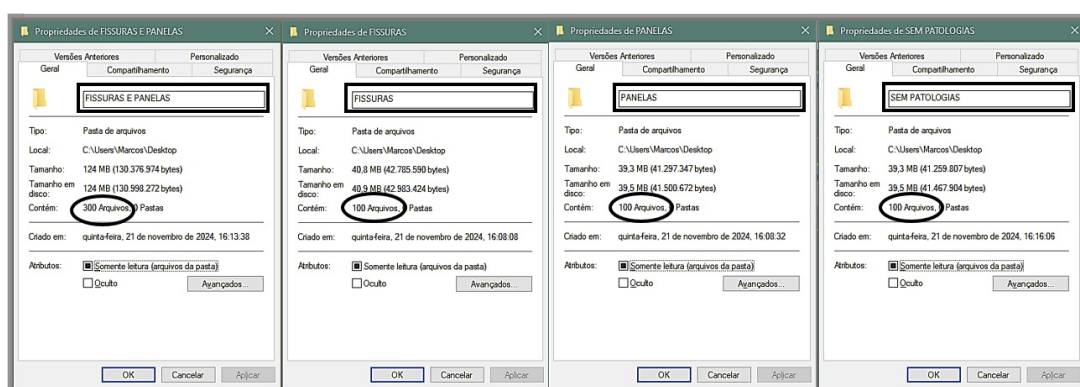


Figura 60 – Banco de imagens selecionadas  
 Fonte: Elaborado pelo autor

A análise das imagens teve por finalidade selecionar aquelas que possuísem painéis e fissuras com boa nitidez, preferencialmente sem ocorrência de sombras, sujeiras, marcas ou sinalizações no asfalto. Também teve por objetivo escolher imagens, em que as patologias pudessem ser rotuladas mais eficazmente. Além disso, evitou-se a seleção de imagens com

excesso de patologias, uma vez que essas não podem ser anotadas com precisão devido sua dispersão, vide exemplo 61.

A seleção, o armazenamento e a separação foram realizados no ambiente *Windows*.



Figura 61 – Exemplo de imagens selecionadas  
Fonte: Elaborado pelo autor

Após montagem do banco de imagens, foi criado na ferramenta CVAT o projeto mestrado. Nele, foram abertas 3 *tasks* (tarefas em português). Essas tarefas foram nomeadas de acordo com as pastas, ou seja, fissuras e painelas, fissuras e painelas. Dentro de cada tarefa foram adicionados 2 layers, os quais foram nomeados de painelas (em amarelo) e fissuras (em rosa), ver Figura 62.

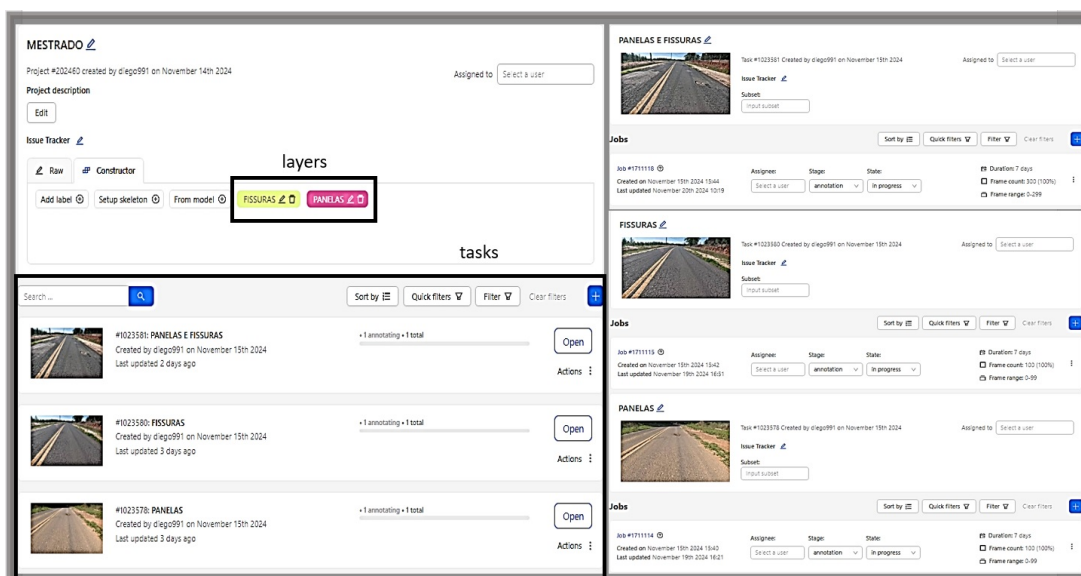


Figura 62 – Configuração da ferramenta CVAT  
Fonte: Adaptado de CVAT (2024)

Após realizadas essas atividades, foi feito o upload de todas as pastas para sua correspondente task e iniciada a rotulação das 500 imagens.

Como forma de melhoria desse trabalho, a rotulação obedeceu às orientações alcançadas na seção 3.1.8, pertencente a etapa de detecção inicial. A figura 63 apresenta alguns exemplos de rotulações realizadas.



Figura 63 – Rotulação para o treino

Fonte: Elaborado pelo autor

Após término das *tasks*, foi efetuado o download das informações de marcação, para armazenamento no servidor local.

Em consequência, para cada imagem, foi criado um arquivo do tipo txt, contendo a lista de todas as rotulações realizadas por tipo de *task*, Figura 64.

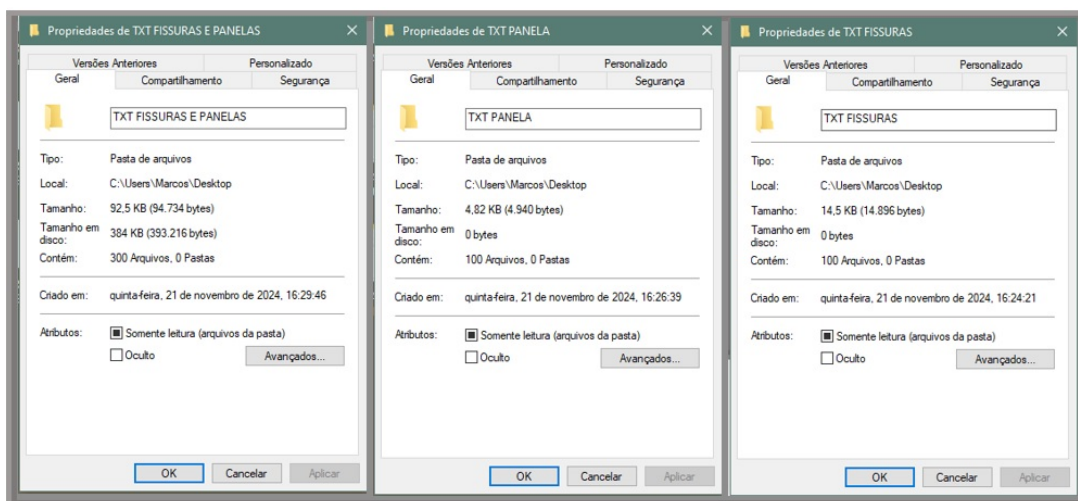


Figura 64 – Arquivos do tipo txt para cada patologia

Fonte: Elaborado pelo autor

### 3.2.3 Pré-processamento

A etapa de pré-processamento foi julgada como necessária após a análise realizada na seção de detecção inicial. À vista disso, essa etapa objetiva tornar o modelo

mais generalizável através da aplicação de técnicas de *Data Augmentation* (SHORTEN; KHOSHGOFTAAR, 2019).

Dentre as técnicas implementadas, está a técnica *brightness*, que adiciona variação do brilho nas imagens. Assim, o modelo tende a ser mais eficaz para detectar objetos com luminosidades diferentes. Observando os testes realizados por (ZAWORSKI, 2018), adotou-se um intervalo de aplicação, que considerou a aplicação da variação de 0,8 a 1,2 na escala de brilho em todas as imagens.

Outra técnica aplicada nessa seção foi a do efeito *blur*. Ela é útil quando se deseja identificar objetos que estão sujeitos a movimento (NELSON, 2020). No caso dessa ferramenta, optou-se por aplicar um desfoque leve em todas as imagens.

Para aplicação dessas ferramentas, foram feitas 2 cópias das pastas de fissuras e painelas, fissuras, painelas e sem patologias. A primeira cópia de cada pasta foi renomeada mantendo o descritivo do tipo da patologia mais o nome *blur*. A segunda também manteve seu nome, com o acréscimo da palavra *brightness*. No total, cada pasta foi duplicada. Na Figura 65 (a), pode ser vista a nomeação aplicada à pasta de fissuras e painelas.

Esse mesmo procedimento foi realizado para as pastas do tipo txt, nesse sentido, cada uma também foi duplicada, vide Figura 65 (b).

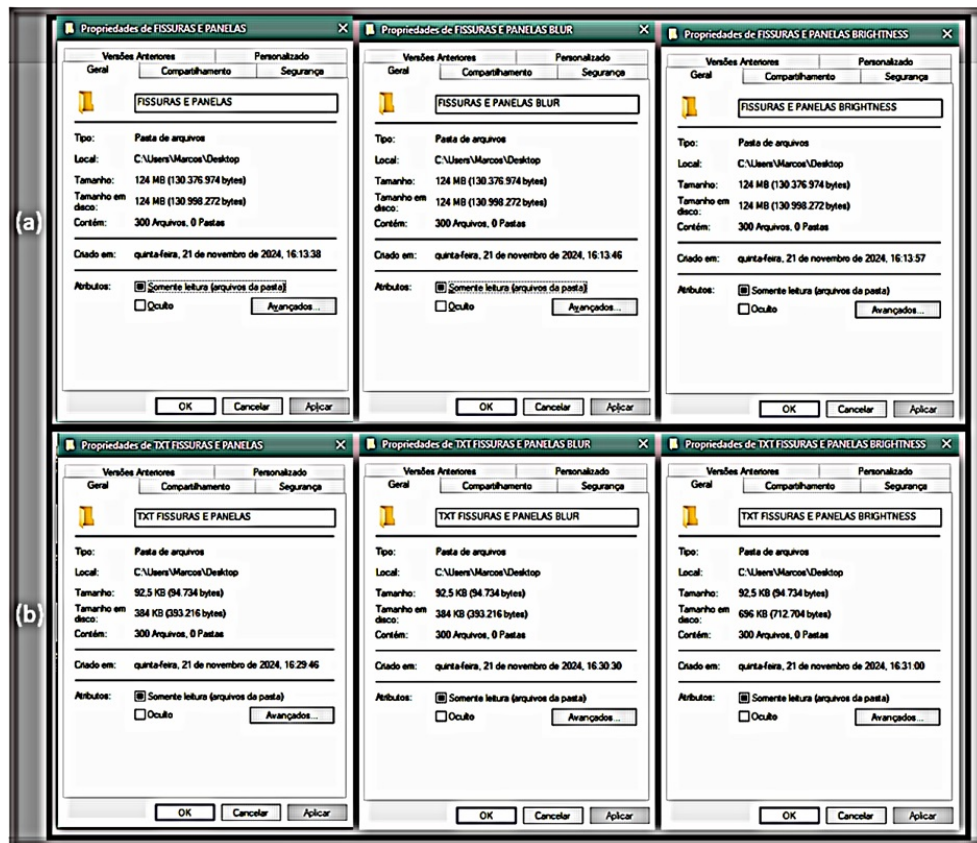


Figura 65 – Nomeação das pastas  
 Fonte: Elaborado pelo autor

Utilizando a linguagem de programação python, versão 3.10.4, todas as extensões dos arquivos txt, para as nomenclaturas *blur* e *brightness*, foram alteradas para *blur* e *brightness*, de acordo com sua patologia, conforme Figura 66.

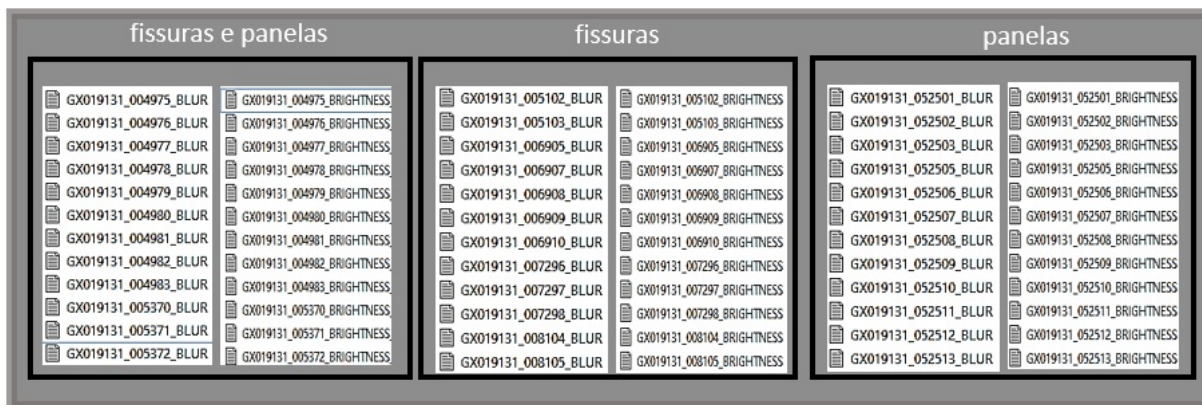


Figura 66 – Nomeação dos arquivos  
 Fonte: Elaborado pelo autor

Esses arquivos foram transferidos do servidor local para armazenamento em nuvem da *Google Drive*.

O pré-processamento ocorreu no ambiente virtual do *Google (Colab)* com utilização da linguagem de programação *Python*. Além disso, foram instaladas os códigos *Augmenters* e *Imageio V2* da biblioteca *Imgaug*, que é uma biblioteca para acréscimo de imagens quando se trabalha com AM. Ademais, para verificação dos resultados no ambiente *Colab*, optou-se por utilizar as bibliotecas *Numpy* e *Matplotlib*.

Na aplicação dessas técnicas, a fim de não haver desordem de identificação, antes do início do pré-processamento, foram criadas as pastas que receberam o *out put* do pré-processamento, dessa forma foram adicionadas as letras AP às pastas de imagens que contiveram os arquivos, conforme apresentado na Figura 67.

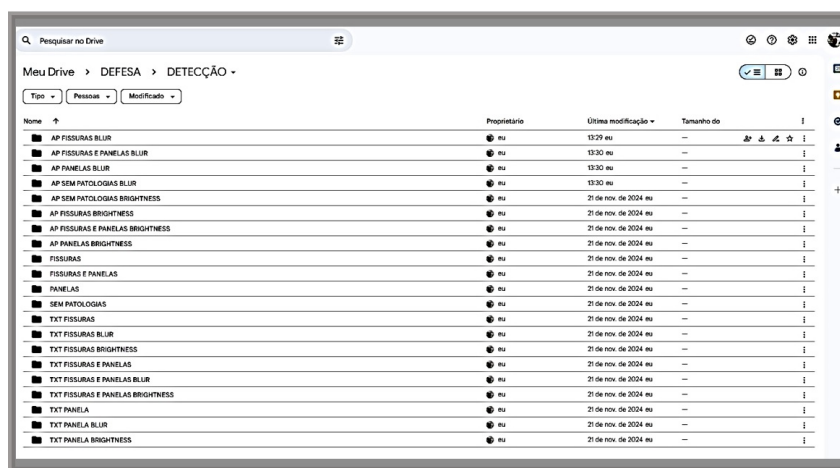


Figura 67 – Nomeação das pastas no pré-processamento  
 Fonte: Adaptado de [GOOGLE DRIVE \(2024\)](#)

A aplicação dos efeitos *Blur* e *Brightness* triplicou o número de imagens de entrada para o treinamento e validação. Os efeitos aplicados às imagens podem ser visualizados na Figura 68.



Figura 68 – Exemplos dos efeitos Blur e Brightness  
Fonte: Elaborado pelo autor

### 3.2.4 Preparação do dataset

O dataset foi criado diretamente no ambiente do *Google Drive*, diferente da etapa de prototipagem, uma vez que todos os arquivos do pré-processamento já haviam sido manuseados no ambiente de armazenamento em nuvem do *Google*. Posto isso, foi criada a pasta Main.

A pasta Main recebeu outras duas pastas, com os respectivos nomes: Images e Labels, ambas também receberam outras duas de nomes: Train e Val.

Com a estrutura montada, partiu-se para distribuição das imagens e dos arquivos txt baixados, conforme apresentado na Tabela 4.

A pasta Train, pertencente a Images, recebeu 720 imagens com patologias do tipo fissuras e panelas, 240 imagens com patologias do tipo fissuras, 240 imagens do tipo panelas e 240 imagens sem patologias, totalizando 1440 imagens já com as técnicas de pré-processamento aplicadas; já a Val recebeu 180 de fissuras e panelas, 60 de fissuras, 60 de panelas e 60 sem patologias, totalizando 360 imagens, já com as técnicas de pré-processamento aplicadas.

A Train, pertencente a Label, recebeu os arquivos do tipo txt referentes às imagens que foram colocadas na Train de Images; da mesma forma, a Val recebeu os arquivos do tipo txt das imagens que foram colocadas na Val de Images.

Dentro da pasta dataset foi criado um arquivo do tipo txt, com o nome "Config.yaml", o qual recebeu o caminho de origem da pasta dataset, da pasta de imagens do train, da pasta de imagens da pasta val e das classes de patologias existentes.

Tabela 4 – Composição do dataset

Pasta	Subpasta	Subpasta	Qtde.	Subpasta	Qtde.
Dataset	Images	Train	1440	Val	360
Dataset	Labels	Train	1440	Labels	360

Fonte: Elaborado pelo autor

No arquivo config. yaml, a classe de patologia de placas recebeu o código 0, enquanto o de fissuras o 1.

### 3.2.5 Treinamento

O processamento foi realizado no ambiente virtual do *Google* com utilização da linguagem de programação *Phyton* na versão 3.10.12. Além disso, foi utilizado a unidade de processamento gráfico da *NVÍDEA* na configuração A100. As configurações seguem apresentadas na Figura 69.

Após elaborada a programação de acesso ao dataset, foi realizada a instalação da biblioteca da *Ultralytics* para utilização da rede *AP YOLO V11* na versão X, a qual foi configurada para 200 épocas de treinamento.

Antes do início do treinamento do modelo, também como forma de melhoria do método, foram configuradas as funções *fliplr*, que realiza o espelhamento das imagens na horizontal, e *flipud*, que faz essa mesma argumentação, contudo, na direção vertical. Assim, para ambas as técnicas, foram atribuídos os valores integrais 1.0, os quais possibilitaram um acréscimo de imagens utilizando um espelhamento de até 180 °.

Para mais, foi modelada a técnica de processamento *shear*, que acrescenta ao banco de imagens um corte em grau, para o valor de até 30 graus, sendo mais uma melhoria do procedimento.

```

Processamento
Ultralytics 8.3.40 Python-3.10.12 torch-2.5.1+cu121 CUDA:0 (NVIDIA A100-50M-40GB, 40514MiB)
engine/trainer: task=detect, mode=train, model=yolo11x.pt, data=/content/drive/MyDrive/DEFESA/dataset/config.yaml, epochs=200
Overriding model.yaml nc=88 with nc=2

   from  n  params module                                arguments
   ---  --  -
0      -1  1    2794  ultralytics.nn.modules.conv.Conv                    [3, 96, 3, 2]
1      -1  1  166272  ultralytics.nn.modules.conv.Conv                    [96, 192, 3, 2]
2      -1  2   389760  ultralytics.nn.modules.block.C3k2                  [192, 384, 2, True, 0.25]
3      -1  1  1327872  ultralytics.nn.modules.conv.Conv                    [384, 384, 3, 2]
4      -1  2  1553664  ultralytics.nn.modules.block.C3k2                  [384, 768, 2, True, 0.25]
5      -1  1  5389952  ultralytics.nn.modules.conv.Conv                    [768, 768, 3, 2]
6      -1  2  5022720  ultralytics.nn.modules.block.C3k2                  [768, 768, 2, True]
7      -1  1  5389952  ultralytics.nn.modules.conv.Conv                    [768, 768, 3, 2]
8      -1  2  5022720  ultralytics.nn.modules.block.C3k2                  [768, 768, 2, True]
9      -1  1  1476864  ultralytics.nn.modules.block.SPPF                  [768, 768, 5]
10     -1  2  3264768  ultralytics.nn.modules.block.C2P5A                  [768, 768, 2]
11     -1  1      0  torch.nn.modules.upsampling.Upsample                [None, 2, 'nearest']
12     [-1, 6] 1      0  ultralytics.nn.modules.conv.Concat                  [1]
13     -1  2  5612544  ultralytics.nn.modules.block.C3k2                  [1536, 768, 2, True]
14     -1  1      0  torch.nn.modules.upsampling.Upsample                [None, 2, 'nearest']
15     [-1, 4] 1      0  ultralytics.nn.modules.conv.Concat                  [1]
16     -1  2  1700352  ultralytics.nn.modules.block.C3k2                  [1536, 384, 2, True]
17     -1  1  1327872  ultralytics.nn.modules.conv.Conv                    [384, 384, 3, 2]
18     [-1, 13] 1      0  ultralytics.nn.modules.conv.Concat                  [1]
19     -1  2  5317632  ultralytics.nn.modules.block.C3k2                  [1152, 768, 2, True]
20     -1  1  5389952  ultralytics.nn.modules.conv.Conv                    [768, 768, 3, 2]
21     [-1, 10] 1      0  ultralytics.nn.modules.conv.Concat                  [1]
22     -1  2  5612544  ultralytics.nn.modules.block.C3k2                  [1536, 768, 2, True]
23     [16, 19, 22] 1  3147862  ultralytics.nn.modules.head.Detect                  [2, [384, 768, 768]]

YOLO11x summary: 631 layers, 56,876,086 parameters, 56,876,070 gradients, 195.5 GFLOPS
    
```

Figura 69 – Processamento do treino  
 Fonte: Adaptado de [GOOGLE COLAB \(2024\)](#)

### 3.2.6 Coleta de dados para detecção

A coleta de dados obedeceu aos mesmos procedimentos executados na instalação, montagem e configuração dos equipamentos apresentados nas seções 3.2.1. À vista disso, foi utilizado o mesmo modelo de câmera apresentado nessas seções.

O percurso foi realizado na cidade de Campinas (SP), no dia 12 de novembro de 2024, entre as 12h52 e as 12h57, em que foram percorridos cerca de 5 Km.

O trajeto envolveu a Rodovia Adalberto Panzam, iniciando próximo à intersecção com a Rodovia Anhaguera (a) e terminando na junção com a Rodovia dos Bandeirantes (b), que segue apresentado na Figura 70.

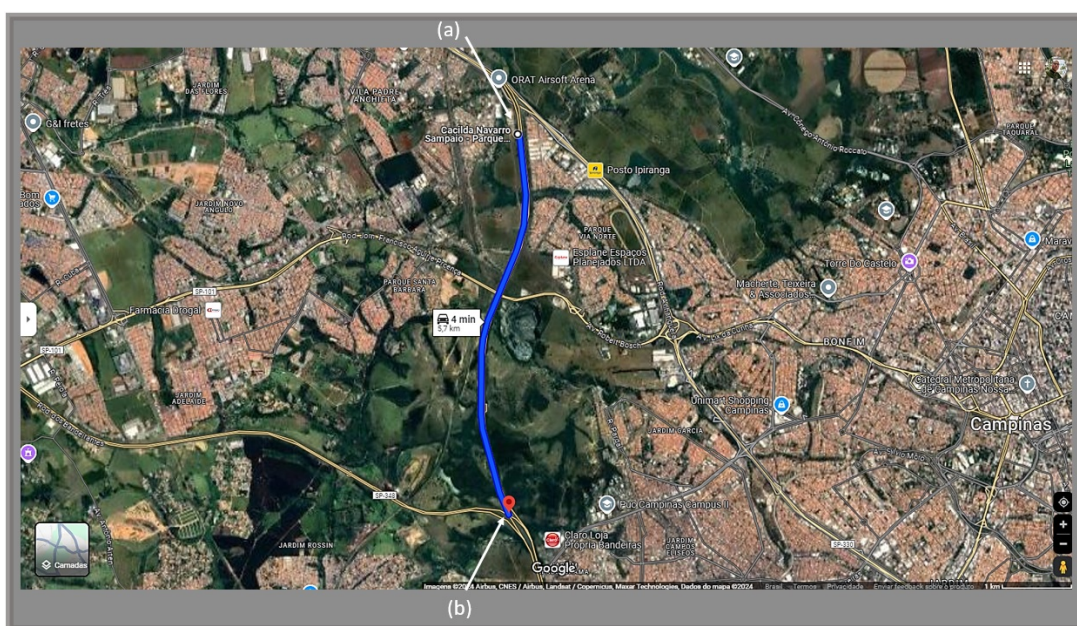


Figura 70 – Trajeto para coleta de dados da detecção  
 Fonte: Adaptado de [GOOGLE MAPS \(2024\)](#)

Durante a coleta dos dados, foram observadas condições climáticas estáveis. Assim, não foi observada precipitação pluviométrica. A temperatura estava próxima aos 32°C e a umidade do ar média foi de 63,59 mm ([CIIAGRO, 2024](#)). Os dados do clima seguem apresentados na Figura 71.

Detalhamento dos período - Período da consulta: 12-11-2024 a 12-11-2024								
Local	Data	Temperatura (°C)			Umidade do ar (mm)			Precipitação (mm)
		Mínima	Média	Máxima	Mínima	Média	Máxima	
Campinas	12-11-2024	17.97	25.09	32.21	39.28	63.59	87.9	0

Figura 71 – Condições climáticas em Campinas, SP  
 Fonte: Adaptado de [CIIAGRO \(2024\)](#)

A Coleta de dados foi realizada a uma velocidade que variou entre 0 e 70 km/h, permitindo a captura de imagens por cerca de 5 minutos. Durante todo o trajeto, o veículo

permaneceu na faixa da direita. Na Figura 72 (a), são apresentados alguns momentos do percurso pela câmera interna do veículo, apenas utilizada para registro do horário e velocidade do veículo, enquanto na 72 (b), são apresentados os registros da câmera de ação GoPro.



Figura 72 – Registro do percurso de detecção na Rodovia Adalberto Pazam  
Fonte: Elaborado pelo autor

Após a coleta de dados, o vídeo foi convertido em fotos através do software Free Video to JPG Converter, versão 5.1.1. A conversão gerou um total de 18.460 imagens, proporcionando 4,47 GB de armazenamento .

Todas as imagens foram analisadas visualmente. Por conseguinte, foram selecionadas 50 imagens contendo as patologias fissuras e panelas, fissuras e panelas.

Todas as imagens selecionadas para o teste receberam suas classificações de acordo com a patologia encontrada. Na Tabela 5 é apresentada a classificação para cada imagem.

Assim, as iniciais F e P foram atribuídas às imagens que continham a patologia conjunta de fissuras e panelas. Para patologia fissura, foi atribuída com a letra F.

Tabela 5 – Identificação das patologias encontradas nas imagens do teste

<b>Imagem</b>	<b>Tipo</b>	<b>Imagem</b>	<b>Tipo</b>
2133	F	8574	F E P
2269	F	9888	F
2365	F	9940	F
2366	F	9947	F
2556	F	10141	F
2759	F	11976	F
2760	F	12396	F
2820	F	12404	F
2821	F	12464	F
3060	F	12544	F
3061	F	12568	F
3108	F	14018	F E P
3121	F	15293	F
3122	F	15759	F
3135	F	15768	F
3342	F	16121	F
3516	F	16182	F
4029	F	16246	F
4030	F	16273	F E P
4410	F E P	16362	F E P
4777	F	16410	F E P
4962	F	17210	F E P
5447	F	18048	F
7776	F E P	19600	F
7777	F E P	19639	F

Fonte: Elaborado pelo autor

Terminada a classificação, as imagens escolhidas foram separadas em uma pasta do servidor local, a qual foi nomeada de Teste. Na sequência, foi realizado o *upload* dessa pasta para o *Google Drive*.

### 3.2.7 Detecção

Assim como no treinamento, o processamento da detecção foi realizado no *Google Colab* em linguagem de programação *Python* na versão 3.10.12. Para isso, foi utilizado a unidade de processamento gráfico da *NVIDIA* na configuração L4.



específico de montar um dataset para realizar o treinamento do algoritmo responsável pela identificação das imperfeições.

## 4 Resultados

### 4.1 Avaliação do treinamento

Nessa seção estão apresentados os resultados alcançados na etapa de treinamento da seção 3.2.5. Logo, foram analisadas as seguintes métricas: perda das caixas delimitadoras, probabilidade predita de classe, precisão geral, precisão com confiança de 50%, recall, matrizes de confusão e perda de distribuição focalizada.

A Figura 74 indica a métrica de perda das caixas delimitadoras para as fases train e val. À vista disso, é possível verificar uma tendência de redução, em função das épocas, mais acentuada na fase val. Contudo, após esse acentuado declínio inicial, inicia-se uma tendência de aumento da perda das caixas.

Os parâmetros finais para a perda das caixas delimitadoras na fase train, que prosseguiram reduzindo, obtêm o resultado entre o intervalo de 1 e 0,5, enquanto na fase val, é observado o resultado entre o intervalo de 3 e 2,75.

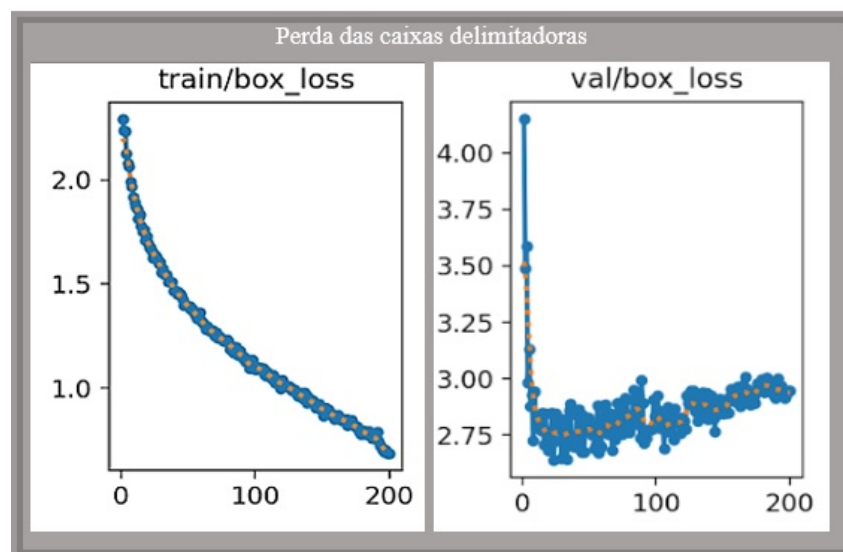


Figura 74 – Perda das caixas delimitadoras para o treinamento da seção 3.2.5

Fonte: Extraído pelo autor do YOLO V11

Com relação à probabilidade predita de classe, que segue apresentada na Figura 75, verifica-se, assim como para o indicador de perda das caixas delimitadoras, uma evidente diminuição inicial da perda de classe na fase val, enquanto para fase train, esse declínio possui menor intensidade.

Assim, para fase train, essa métrica fica entre os valores 0,5 e 0. Já para val, entre os valores 4 e 2.

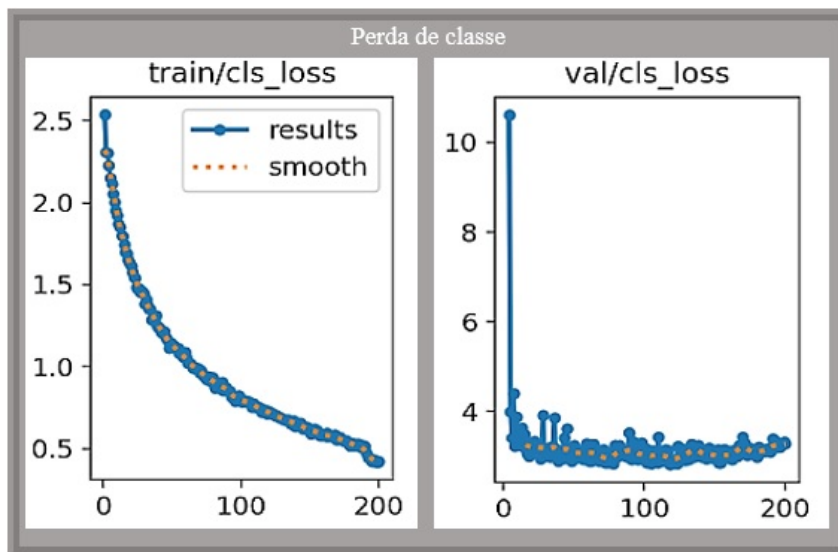


Figura 75 – Probabilidade predita de classe para o treinamento da seção 3.2.5  
Fonte: Extraído pelo autor do YOLO V11

No quesito precisão geral, Figura 76, é possível verificar que o modelo atinge um valor entre o intervalo de 0,4 e 0,2. Já em relação à precisão com confiança de 50%, o resultado se aproxima de 0,2, ambas relacionadas à fase val.

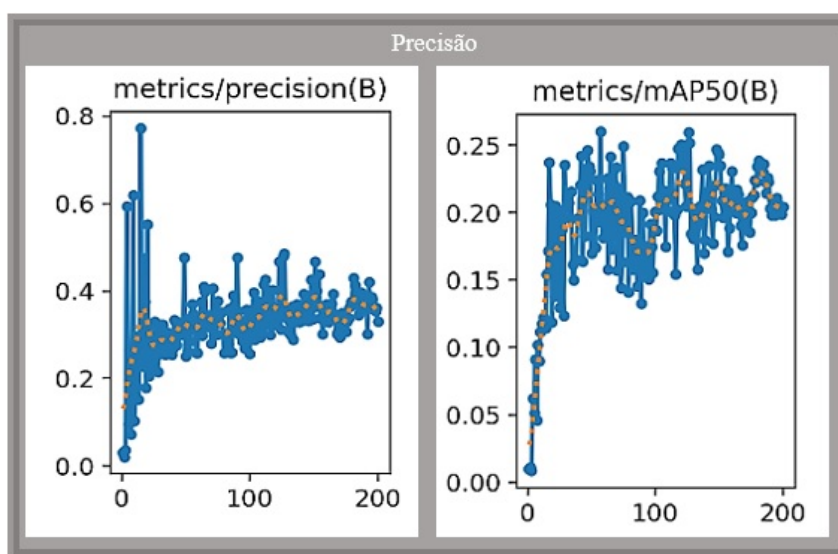


Figura 76 – Precisão geral e com confiança de 50% para o treinamento da seção 3.2.5  
Fonte: Extraído pelo autor do YOLO V11

Na Figura 77, que apresenta o parâmetro de recall, foi possível observar uma tendência de crescimento mais acentuada nas primeiras épocas de treinamento. Posteriormente, nota-se uma estabilização entre o intervalo de 0,25 e 0,20.

Na Figura 78, perda de distribuição focalizada, é possível verificar um acentuado decaimento inicial na fase de val, enquanto na fase train, esse declínio ocorre durante as 200 épocas.

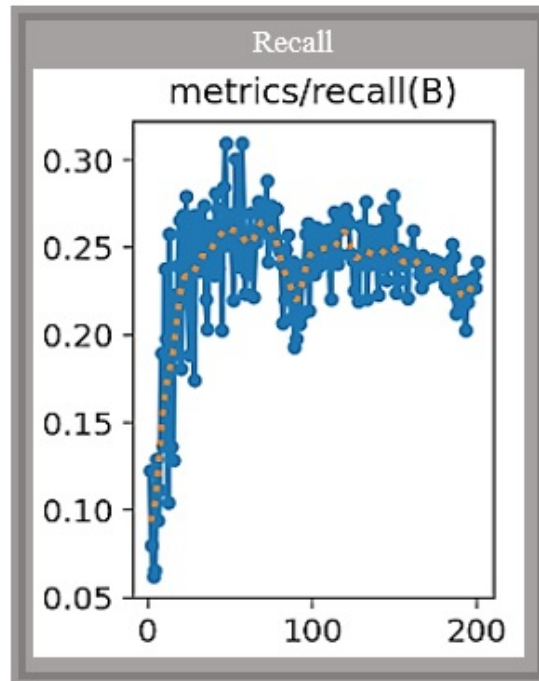


Figura 77 – Recall para o treinamento da seção 3.2.5  
Fonte: Extraído pelo autor do YOLO V11

Logo, para fase val, esse índice varia entre o intervalo de 4 e 2. Porém, para a fase train, varia entre 1,25 e 1,00.

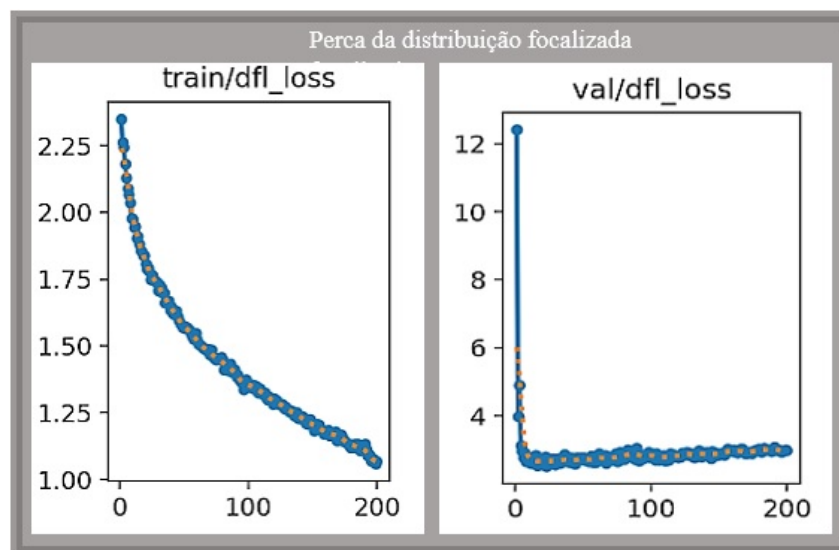


Figura 78 – Perda de distribuição focalizada para o treinamento da seção 3.2.5  
Fonte: Extraído do YOLO v11 pelo autor

Referente a matriz de confusão, Figura 79, é possível constatar que 223 fissuras e 109 fissuras foram identificadas corretamente, enquanto 237 fissuras e 80 painéis foram identificados equivocadamente. Além disso, 1630 fissuras e 179 painéis não foram identificadas pelo modelo.

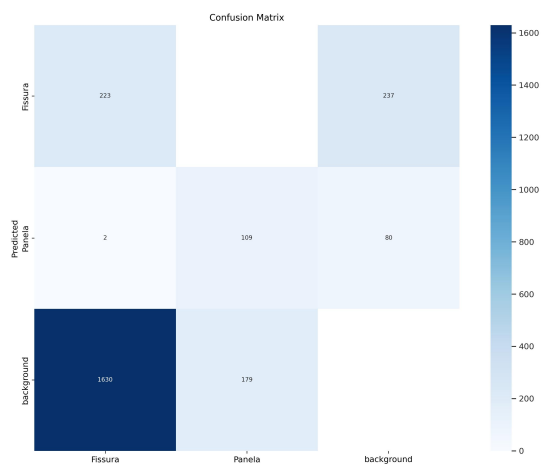


Figura 79 – Matriz de confusão do treinamento para o treinamento da seção 3.2.5  
 Fonte: Extraído pelo autor do YOLO V11

## 4.2 Avaliação da detecção

Nessa seção são descritos e quantificados os resultados do modelo após submetido à detecção, referindo-se às informações da seção 3.2.7.

Dessa forma, é possível constatar que o modelo treinado encontrou algum tipo de patologia em 49 das 50 imagens processadas.

Para todos os casos de imagens previamente classificadas com a patologia de fissura, às quais correspondem a um total de 41 imagens, mesmo que parcialmente, é possível ver que o modelo encontrou parte das patologias que haviam sido observadas.

Para o caso de fissuras e panelas concomitantemente, correspondente aos 9 casos restantes, verifica-se que o modelo identificou em uma 1 imagem a presença de ambas patologias. Entretanto, foram identificadas panelas em 4 imagens e fissuras em 3 imagens. Em apenas 1 imagem, não foi identificada nenhuma patologia.

Na Tabela 6, está descrito para cada imagem a saída observada pós processamento da Seção 3.2.7. Logo, estão apresentadas as confianças para cada bounding box elaborado pelo modelo no processo de detecção.

Além disso, em verde é apresentada a patologia que foi plenamente detectada, em amarelo que foi parcialmente detectada e as que não foram em vermelho.

Também é possível observar que a média de confiança do modelo para detecção de fissuras é de 0,39 e 0,57 para panelas.

Tabela 6 – Resumo dos resultados da fase de pós processamento da Seção 3.2.7

Imagem	Tipo	Identificação	Confiança			
			Bounding boxes			
			Fissura		Panela	
2133	F	id parcial	0,43			
2269	F	id parcial	0,27	0,36		
2365	F	id parcial	0,37	0,59	0,34	
2366	F	id parcial	0,27	0,38		
2556	F	id parcial	0,29			
2759	F	id parcial	0,48			
2760	F	id parcial	0,37			
2820	F	id parcial	0,31	0,39		
2821	F	id parcial	0,36	0,26		
3060	F	id parcial	0,33			
3061	F	id parcial	0,26			
3108	F	id parcial	0,34			
3121	F	id parcial	0,45	0,38	0,38	
3122	F	id parcial	0,64			
3135	F	id parcial	0,56			
3342	F	id parcial	0,37			
3516	F	id parcial	0,30	0,42		
4029	F	id parcial	0,72			
4030	F	id parcial	0,52			
4410	F	P				0,52
4777	F		0,47			
4962	F		0,52	0,32		
5447	F		0,52			
7776	F	P	0,62			
7777	F	P	0,4			0,52
8574	F	P				0,67
9888	F		0,38			
9940	F		0,45			
9947	F		0,36			id falsa
10141	F		0,28			
11976	F		0,28			
12396	F		0,68			
12404	F		0,38			
12464	F		0,27			

Tabela 6 - continuação da página anterior

			Confiança			
12544	F	id parcial	0,66			
12568	F	id parcial	0,29	0,28		
14018	F P	id parcial	0,29			
15293	F	id parcial	0,34	0,38		
15759	F	id parcial	0,29	0,36		
15768	F	id parcial	0,59	0,30		
16121	F	id parcial	0,47			id falsa
16182	F	id parcial	0,41			
16246	F	id parcial	0,26			
16273	F P	id parcial				
16362	F P	id parcial				0,76
16410	F P	id parcial				0,38
17210	F P	id parcial	0,26			
18048	F	id parcial	0,30			
19600	F	id parcial	0,29			
19639	F	id parcial	0,52	0,27		
<b>Confiança média</b>			<b>0,39</b>		<b>0,57</b>	

id parcial = identificação de parte das patologias da imagem

id falsa = identificação que não corresponde a uma forma de patologia

	Identificação parcial da patologia discutida na imagem
	Identificação total da patologia discutida na imagem
	Patologia não identificada

Fonte: Elaborado pelo autor

Na Figura 80 estão apresentadas algumas das detecções geradas pelo modelo treinado. Nas imagens, é possível encontrar os bounding boxes, a classificação da patologia e a confiança do modelo.

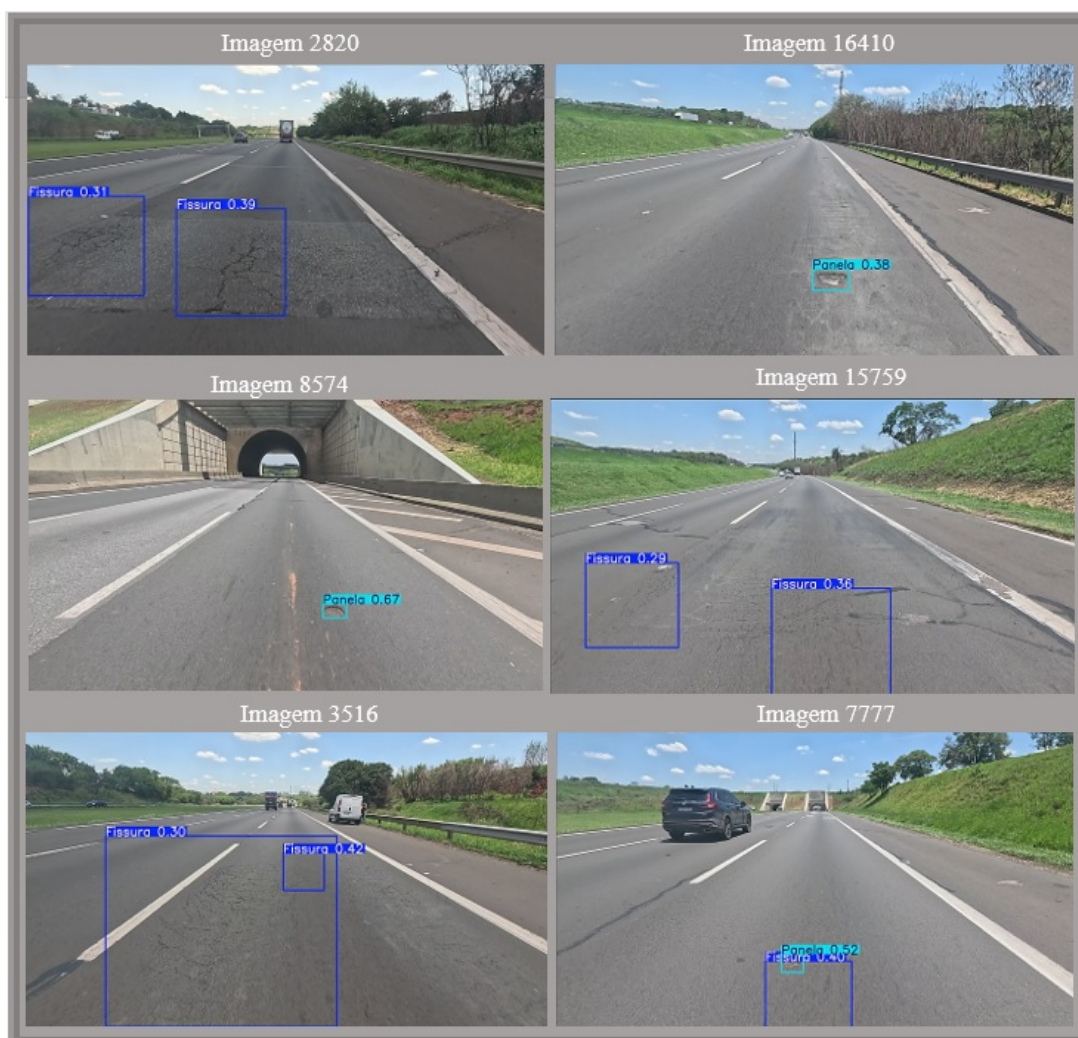


Figura 80 – Exemplo de detecção gerada pelo modelo treinado

Fonte: Extraído pelo autor do YOLO V11

Almejando tornar a base de imagens pública, para que outros usuários possam utilizá-la e cumprindo o objetivo específico de disponibilizar a base de dados rotulados publicamente, o dataset que contempla as fotos e os arquivos do tipo txt, utilizados na seção 3.2.5, foram indexadas nas plataformas Github e Zenodo, sob os weblinks:

[https://github.com/estudentediegoufscar/txt\\_fissuras\\_panelas.git](https://github.com/estudentediegoufscar/txt_fissuras_panelas.git)

<https://zenodo.org/uploads/14425458>

# 5 Considerações Finais

## 5.1 Principais Resultados

Essa seção resume as principais conclusões, além de dispor quanto aos aspectos relevantes à continuidade do progresso desse estudo, que busca sofisticar os SGP's por meio do diagnóstico de patologias em pavimentos flexíveis através do treinamento de um algoritmo utilizando aprendizado de máquina.

Ressalta-se que as conclusões obtidas estão relacionadas às peculiaridades de todo um processo elaborativo contido nessa dissertação, mas que permite deduzir suas potencialidades e limitações.

Para atualidade é atípico desenvolver soluções em métodos ou processos relacionados à área de engenharia sem aplicação de IA. Nesse sentido, significativos progressos em automação da detecção de objetos estão sendo produzidos nos últimos anos, inclusive o crescente desenvolvimento métodos e técnicas para interpretação de imagens.

Contudo, detectar imperfeições em pavimentos flexíveis é considerado um grande desafio. Para mais, no caso brasileiro, verifica-se que a maior parte das estradas é considerada ruim ou péssima e que os métodos utilizados para inspeção em pavimentos flexíveis são vagarosos, duvidosos e onerosos.

Frente à isso, dada a complexidade da detecção, foi imprescindível compreender os principais conceitos relacionados à Detecção Automática de Objetos, passando pelos conceitos relacionados às Redes Neurais Artificiais, Tipos de Aprendizado, Aprendizado de Máquina, Aprendizado Profundo, Redes Neurais Profundas, Redes Convolucionais, Visão Computacional e Detecção Automatizada de Objetos.

O método escolhido para composição do dataset permitiu ao modelo processar a etapa de treinamento, atendendo todas especificidades, inclusive as relacionadas à diversibilidade e à escalabilidade de sua composição. Além disso, foi implementado ao dataset as técnicas *brightness* e *blur*, as quais são técnicas de Data Augmentation, que aumentaram qualitativamente e quantitativamente os bancos de imagens associados, possibilitando a realização do treinamento do algoritmo. Logo, o objetivo de montar um dataset para realizar o treinamento do algoritmo responsável pela identificação das imperfeições foi satisfeito.

A análise dos indicadores de perda das caixas delimitadoras, perda de classe, precisão média, precisão com confiança de 50%, recall, detecção inicial e matriz de confusão, da etapa de detecção inicial, conforme apresentado na seção 3.1.8, mostra que a estratégia de

rotulação do treino 3 se fez mais adequada para detecção de fissuras e trincas. Assim, a rotulação para a detecção final obedece essa propositura. Por conseguinte, entende-se que o direcionamento ao objetivo geral foi realizado, sendo cumprido o objetivo de indicar por meio de uma etapa de detecção inicial a forma de rotulação mais eficaz para detecção de fissuras e panelas;

Os resultados do teste de detecção mostram a identificação de patologias em 98% das imagens selecionadas. Ainda que de forma parcial, o modelo, mesmo composto por um dataset pequeno, produz resultados assertivos, na identificação de panelas e fissuras. Em vista disso, constata-se que a quantificação da detecção alcançada pelo modelo para um banco de imagens previamente escolhido foi concluída, satisfazendo mais uma especificidade do escopo. Todavia, evidencia-se que o sucesso dessa aplicação fundamenta-se em todo um procedimento, que abrange desde a coleta de imagens até a implementação de técnicas no YOLO.

Além disso, o resultado médio de confiança baseado nas imagens detectadas, apesar de reduzido, se mostra promitente, quando é analisada a dimensão do dataset construído. Assim, a confiança média conjunta chegou a 48%.

Cumprindo ao último objetivo específico, objetivando dar continuidade na evolução de modelos que visem à identificação automatizada de patologias em asfalto flexível com utilização de aprendizado de máquina, o dataset desse estudo foi disponibilizado publicamente e pode ser facilmente localizado na seção 4.2.

Em síntese, esse estudo desenvolve o treinamento de um algoritmo para identificação automatizada de fissuras e panelas em pavimentos flexíveis, por meio de imagens de vídeo, utilizando técnicas de Aprendizado de Máquina.

## 5.2 Trabalhos futuros

Para trabalhos futuros sugere-se a aplicação de outras redes neurais convolucionais de aprendizado não supervisionado. Também pode ser verificada a possibilidade de mudança de estratégia de rotulação a depender da patologia escolhida. Infere-se, com base nos resultados observados, que o modelo pode ser utilizado para outras detecções importantes nas estradas, como sinalização viária por exemplo.

A utilização de um banco de imagens pronto, que também possibilitasse a rotulação, poderia otimizar significativamente o processo de treinamento.

Por fim, a otimização do método de rotulação também poderia trazer maior desempenho ao processo de treinamento.

# Referências

- AASHTO, 1993 AMERICAN ASSOCIATION AND TRANSPORTANTION OFFICIALS. *Guide for Design for Pavement Structures*. Washington, D.C. EUA, 1993. Citado na página 28.
- ADLINGE, S. S.; GUPTA, A. K. Pavement deterioration and its causes. *International journal of innovative research and development*, v. 2, n. 4, p. 9–15, 2013. Citado na página 24.
- AHMED, B. Q. *Developing of Pavement Management System (PMS) for EMU Campus Pavement in GIS Environment*. Dissertação (Mestrado) — Eastern Mediterranean University, Gazimağusa, North Cyprus, 2013. Citado na página 25.
- ALMEIDA, L. C. d. A. *Aplicativo para Smartphone destinado à Medição da Irregularidade Longitudinal em Rodovias*. Dissertação (Mestrado) — Centro de Tecnologia, Departamento de Engenharia de Transportes, Universidade Federal do Ceará, 2018. Citado na página 14.
- ALVES, K. d. R. *A gestão de pavimentos aplicada à manutenção rodoviária*. Dissertação (Mestrado) — Universidade Federal de Santa Maria, Porto Alegre, 2016. Citado na página 27.
- BALBO, J. T. *Pavimentação asfáltica: materiais, projeto, e restauração*. 1. ed. São Paulo: Oficina de textos, 2007. 560 p. Citado 2 vezes nas páginas 17 e 18.
- BARELLA, R. M. *Contribuição para a avaliação da irregularidade longitudinal de pavimentos com perfilômetros inerciais*. Tese (Doutorado) — Universidade de São Paulo, Escola Politécnica da Universidade de São Paulo, São Paulo, 2008. Citado 2 vezes nas páginas 28 e 31.
- BARTHOLOMEU, D. B.; CAIXETA, J. V. Impactos econômicos e ambientais decorrentes do estado de conservação das rodovias brasileiras: um estudo de caso. *Revista de Economia e Sociologia Rural*, SciELO Brasil, v. 46, p. 703–738, 2008. Disponível em: <<https://www.scielo.br/j/resr/a/zWSrV5qVwGvBXzXqF6BMBtd/>>. Citado na página 13.
- BERNUCCI, L. B. et al. *Pavimentação asfáltica: Formação básica para engenheiros*. 3. ed. Rio de Janeiro: PETROBRAS: ABEDA, 2006. 504 p. Citado 8 vezes nas páginas 17, 18, 21, 23, 27, 28, 29 e 31.
- BERTOLLO, S. A. M. *Considerações sobre a gerência de pavimentos urbanos em nível de rede*. Dissertação (Mestrado) — Universidade de São Paulo, Escola de Engenharia de São Carlos, São Carlos, 1997. Citado 4 vezes nas páginas 14, 26, 27 e 28.
- BOCHIE, K. et al. *Aprendizado Profundo em Redes Desafiadoras: Conceitos e Aplicações*. In: *SBRC 2020, GOMES, D. G.; MORAES I. M.; CAMPISTA M. E. M.;*(Org.). *Minicursos do XXXVIII Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*. SBC, 2020. 140–189 p. Disponível em: <<https://doi.org/10.5753/sbc.5033.7>>. Acesso em: 9 jun.2024. Citado 6 vezes nas páginas 37, 38, 39, 40, 41 e 42.

- BRAGA, A. d. P.; LUDERMIR, T. B.; CARVALHO, A. *Redes neurais artificiais: teoria e aplicações*. [S.l.]: LTC, 2000. 262 p. Citado 2 vezes nas páginas 36 e 37.
- BRANCO, F.; PEREIRA, P.; SANTOS, L. P. *Pavimentos rodoviários*. 1. ed. São Paulo: Coimbra Almedina, 2006. 388 p. Citado 2 vezes nas páginas 17 e 18.
- BRANCO, L. H. C. *MANIAC: Uma metodologia para o monitoramento automatizado das condições dos pavimentos utilizando VANTs*. Tese (Doutorado) — Universidade de São Paulo, Escola de Engenharia de São Carlos, São Carlos, 2016. Citado na página 14.
- BRASIL. MINISTÉRIO DO PLANEJAMENTO, DESENVOLVIMENTO E GESTÃO. OLIVEIRA, DYOGO HENRIQUE DE. *Agora, é avançar*. [S.l.], 2017. (Apresentação de Slides). Disponível em: <<https://www.gov.br/economia/pt-br/centrais-de-conteudo/apresentacoes/planejamento/2017/projeto-avancar-apresentacao-ministro-dyogo-projecao.pdf/view>>. Citado na página 13.
- BRYNJOLFSSON, E.; MCAFEE, A. Artificial intelligence, for real. *Harvard business review*, v. 1, p. 1–31, jul 2017. Disponível em: <<https://starlab-alliance.com/wp-content/uploads/2017/09/AI-Article.pdf>>. Acesso em: 11 jun.2024. Citado na página 37.
- CALÇAS, A. S. P. et al. Comparativo de dimensionamento de pavimentos pelo método medina e pelo método semiempírico. *Revista Mackenzie*, 2021. Disponível em: <<https://adelfa-api.mackenzie.br/server/api/core/bitstreams/61a7377f-1379-4228-9442-78366c09316b/content>>. Acesso em: 09 out.2024. Citado na página 20.
- CAREY, W. N.; IRICK, P. E. The pavement serviceability performance concept. *Highway Research Board, HRB Bulletin*, v. 250, 1960. Citado na página 31.
- CENTRO INTEGRADO DE INFORMAÇÕES AGROMETEOROLÓGICAS. *Portal Agrometeorológico e Hidrológico do Estado de São Paulo*. São Paulo, 2024. Disponível em: <<http://www.ciiagro.org.br/diario/periodo>>. Citado 2 vezes nas páginas 74 e 82.
- CNI. CONFEDERAÇÃO NACIONAL DA INDÚSTRIA. *Proposta da Indústria para as Eleições: Grandes obras paradas: como enfrentar o problema?* Brasília, 2018. 124 p. Disponível em: <[https://static.portaldaindustria.com.br/media/filer\\_public/26/96/2696958e-7e26-47f4-b3a4-261a26742592/25\\_-\\_grandes\\_obras\\_paradas\\_resumo.pdf](https://static.portaldaindustria.com.br/media/filer_public/26/96/2696958e-7e26-47f4-b3a4-261a26742592/25_-_grandes_obras_paradas_resumo.pdf)>. Citado na página 13.
- CNT. CONFEDERAÇÃO NACIONAL DO TRANSPORTE. *Revista CNT*. Brasília, 2001. 40 p. Disponível em: <<https://repositorio.itl.org.br/jspui/handle/123456789/132>>. Citado na página 13.
- CNT. CONFEDERAÇÃO NACIONAL DO TRANSPORTE. *Pesquisa CNT de Rodovias 2021*. Brasília, 2021. 231 p. Disponível em: <<https://static.poder360.com.br/2021/12/pesquisa-cnt-rodovias.pdf>>. Citado na página 13.
- CVAT. *Open Data Annotation Platform*. [S.l.], 2024. Disponível em: <<https://www.cvat.ai/>>. Citado 3 vezes nas páginas 61, 62 e 76.

DEPARTAMENTO NACIONAL DE ESTRADAS DE RODAGEM - INSTITUTO DE PESQUISAS RODOVIÁRIAS - DIVISÃO DE INFORMÁTICA TÉCNICO-CIENTÍFICA. DNER 667-22/1981. *Método de projeto de pavimentos flexíveis*. 3. ed. [S.l.], 1981.

Disponível em: <[https://www.gov.br/dnit/pt-br/assuntos/planejamento-e-pesquisa/ipr/coletanea-de-manuais/vigentes/667\\_metodo\\_de\\_projeto\\_de\\_pavimentos\\_flexiveis.pdf](https://www.gov.br/dnit/pt-br/assuntos/planejamento-e-pesquisa/ipr/coletanea-de-manuais/vigentes/667_metodo_de_projeto_de_pavimentos_flexiveis.pdf)>. Citado na página 19.

DEPARTAMENTO NACIONAL DE INFRAESTRUTURA DE TRANSPORTES - INSTITUTO DE PESQUISAS RODOVIÁRIAS. DNIT 005/2003 -TER. *Defeitos nos pavimentos flexíveis e semirrígidos : Terminologia*. Rio de Janeiro, 2003a.

Disponível em: <[https://www.gov.br/dnit/pt-br/assuntos/planejamento-e-pesquisa/ipr/coletanea-de-normas/coletanea-de-normas/terminologia-ter/dnit\\_005\\_2003\\_ter-1.pdf](https://www.gov.br/dnit/pt-br/assuntos/planejamento-e-pesquisa/ipr/coletanea-de-normas/coletanea-de-normas/terminologia-ter/dnit_005_2003_ter-1.pdf)>. Citado 4 vezes nas páginas 21, 22, 23 e 24.

DEPARTAMENTO NACIONAL DE INFRAESTRUTURA DE TRANSPORTES - INSTITUTO DE PESQUISAS RODOVIÁRIAS. DNIT 008/2003 -PRO. *Levantamento visual contínuo para Avaliação da superfície de pavimentos flexíveis e semi-rígidos : Procedimento*. Rio de Janeiro, 2003c.

Disponível em: <[https://www.gov.br/dnit/pt-br/assuntos/planejamento-e-pesquisa/ipr/coletanea-de-normas/coletanea-de-normas/procedimento-pro/DNIT\\_008\\_2003\\_PRO](https://www.gov.br/dnit/pt-br/assuntos/planejamento-e-pesquisa/ipr/coletanea-de-normas/coletanea-de-normas/procedimento-pro/DNIT_008_2003_PRO)>. Citado na página 32.

DEPARTAMENTO NACIONAL DE INFRAESTRUTURA DE TRANSPORTES - INSTITUTO DE PESQUISAS RODOVIÁRIAS. DNIT 009/2003 -PRO. *Avaliação subjetiva da superfície de pavimentos flexíveis e semi-rígidos : Procedimento*. Rio de Janeiro, 2003b.

Disponível em: <[https://www.gov.br/dnit/pt-br/assuntos/planejamento-e-pesquisa/ipr/coletanea-de-normas/coletanea-de-normas/procedimento-pro/DNIT\\_009\\_2003\\_PRO](https://www.gov.br/dnit/pt-br/assuntos/planejamento-e-pesquisa/ipr/coletanea-de-normas/coletanea-de-normas/procedimento-pro/DNIT_009_2003_PRO)>. Citado 4 vezes nas páginas 27, 31, 32 e 33.

DEPARTAMENTO NACIONAL DE INFRAESTRUTURA DE TRANSPORTES - INSTITUTO DE PESQUISAS RODOVIÁRIAS. DNIT-IPR 745/2011. *Manual de gerência de pavimentos - Procedimento*. [S.l.], 2011.

Disponível em: <[https://www.gov.br/dnit/pt-br/assuntos/planejamento-e-pesquisa/ipr/coletanea-de-manuais/vigentes/745\\_manual\\_de\\_gerencia\\_de\\_pavimentos.pdf](https://www.gov.br/dnit/pt-br/assuntos/planejamento-e-pesquisa/ipr/coletanea-de-manuais/vigentes/745_manual_de_gerencia_de_pavimentos.pdf)>. Citado 4 vezes nas páginas 14, 25, 28 e 29.

DOMINGUES, F. A. A. Notas de aula da disciplina: Irregularidade longitudinal dos pavimentos-ptr-5008. *Escola Politécnica da Universidade de São Paulo. Curso de pós-graduação em engenharia de transportes*. São Paulo, 2004. Citado na página 28.

ELBECK, M. The fourth industrial revolution's potential influence on marketing education. *E-Journal of Business Education and Scholarship of Teaching*, v. 12, n. 1, p. 112–119, 2018. Disponível em: <<https://files.eric.ed.gov/fulltext/EJ1183303.pdf>>. Acesso em: 15 jun.2024. Citado na página 38.

EYNG, E. *Controle feedforward-feedback aplicado as colunas de absorção do processo de produção de etanol por fermentação*. Tese (Doutorado) — Faculdade de Engenharia Química, Universidade Estadual de Campinas, 2008. Citado na página 36.

FARIA, E. L. d. *Redes neurais convolucionais e máquinas de aprendizado extremo aplicadas ao mercado financeiro brasileiro*. Tese (Doutorado) — COOPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2018. Citado 3 vezes nas páginas 39, 40 e 41.

FERNANDES JÚNIOR, J. L.; ODA, S.; ZERBINI, L. F. *Defeitos e atividades de manutenção e reabilitação em pavimentos asfálticos*. 2006. Reimpressão. Disponível em: <<http://repositorio.eesc.usp.br/handle/RIEESC/7448>>. Citado na página 27.

FERNANDES, M. A. c. *Redes Neurais Artificiais Aplicadas à Detecção Inteligente de Sinais*. Dissertação (Mestrado) — Centro de Tecnologia. Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal do Rio Grande do Norte, 1999. Citado 2 vezes nas páginas 34 e 36.

FLECK, L. et al. Redes neurais artificiais: Princípios básicos. *Revista Eletrônica Científica Inovação e Tecnologia*, v. 1, n. 13, p. 47–57, jun 2016. Disponível em: <<https://periodicos.utfpr.edu.br/recit/article/viewFile/4330/Leandro>>. Acesso em: 9 jun.2024. Citado 2 vezes nas páginas 35 e 36.

FONSECA, J. J. S. d. *Metodologia da pesquisa científica*. [S.l.]: UECE, 2002. 127 p. Citado na página 50.

FONTENELE, H. B.; JUNIOR, C. A. P. da S.; CORREIO, C. L. P. Análise de ocorrência de avaliações subjetivas. *Synergismus scyentifica UTFPR*, v. 2, n. 1, nov 2007. Disponível em: <<https://typeset.io/papers/analise-de-ocorrencia-de-avaliacoes-subjetivas-2bl87e6rpb>>. Acesso em: 10-maio-2024. Citado na página 31.

FRANCO, C. R. Inteligência artificial. *Uniasselvi*, 2014. Disponível em: <[ISBN978-85-68075-77-7](https://doi.org/10.1007/978-85-68075-77-7)>. Citado 3 vezes nas páginas 18, 35 e 37.

FRANCO, F. A. C. d. P. *Método de dimensionamento mecânico empírico de pavimentos asfálticos*. Tese (Doutorado) — Curso de Engenharia Civil, COOPE/UFRJ, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2007. Citado na página 18.

GADELHA, L. G. C. *Orçamento na Construção Pesada – Orçamento e programação de Barragens de Terra*. 2. ed. Recife: do Autor, 2011. 504 p. Citado na página 26.

GERHARDT, T. E.; SILVEIRA, D. T. *Métodos de Pesquisa: Série Educação a Distância*. 1. ed. [S.l.]: Editora da UFRGS, 2009. 120 p. Citado na página 50.

GIL, A. C. *Como elaborar projetos de pesquisa*. 4. ed. [S.l.]: Atlas S A, 2002. 177 p. Citado na página 50.

GILLESPIE, T. D. Everything you always wanted to know about the iri, but were afraid to ask! The University of Michigan Transportation Research Institute, Lincoln, Nebraska, september 1982. Citado na página 29.

GIRSHICK, R. et al. Rich feature hierarchies for accurate object detection and semantic segmentation. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [s.n.], 2013. p. 580–587. Disponível em: <[https://openaccess.thecvf.com/content\\_cvpr\\_2014/html/Girshick\\_Rich\\_Feature\\_Hierarchies\\_2014\\_CVPR\\_paper.html](https://openaccess.thecvf.com/content_cvpr_2014/html/Girshick_Rich_Feature_Hierarchies_2014_CVPR_paper.html)>. Acesso em: 28 out.2024. Citado na página 44.

GIRSHICK, R. B. Fast r-cnn. corr abs/1504.08083. *arXiv preprint arXiv:1504.08083*, 2015. Disponível em: <<http://arxiv.org/abs/1504.08083>>. Acesso em: 28 out.2024. Citado 2 vezes nas páginas 44 e 45.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep learning*. [S.l.]: MIT Press, 2016. 775 p. Citado 5 vezes nas páginas 37, 39, 40, 41 e 42.

GOOGLE INC. *Google Colab*. [S.l.], 2024. Disponível em: <<https://colab.research.google.com/>>. Citado 5 vezes nas páginas 63, 64, 65, 81 e 85.

GOOGLE INC. *Google Drive*. [S.l.], 2024. Disponível em: <<https://drive.google.com/pt-br>>. Citado na página 79.

GOOGLE INC. *Google Maps*. [S.l.], 2024. Disponível em: <<https://www.google.com.br/maps/preview>>. Citado 3 vezes nas páginas 56, 73 e 82.

GRUS, J. *Data Science from scratch: first principles from Python*. 1. ed. [S.l.]: O'Reilly Media, 2015. 330 p. Citado na página 38.

HAAS, R.; HUDSON, W. R.; ZANIEWSKI, J. P. *Modern pavement management*. 1. ed. Flórida: Krieger Publishing Company, 1994. 604 p. Citado 2 vezes nas páginas 26 e 27.

HAYKIN, S. *Neural networks: a comprehensive foundation*. [S.l.]: Prentice Hall PTR, 1994. Citado na página 34.

HAYKIN, S. *Redes neurais: princípios e prática*. 2. ed. [S.l.]: Art Med Editora SA, 2001. 262 p. Citado na página 35.

HILT, K. What does the future hold for the law librarian in the advent of artificial intelligence? *Canadian Journal of Information and Library Science*, v. 41, n. 3, p. 211–227, Sept 2017. Disponível em: <<https://muse.jhu.edu/pub/50/article/686190/pdf>>. Acesso em: 15 jun.2024. Citado na página 15.

INMET. *Instituto Nacional de Meteorologia: Gráfico Diário de Estações*. São Paulo, 2024. Disponível em: <<https://tempo.inmet.gov.br/Graficos/A001>>. Citado 2 vezes nas páginas 57 e 58.

JAIN, N. et al. Geographical information system for pavement management system. *Transportation*, nov 2003. Disponível em: <<https://www.yumpu.com/en/document/read/21524673/map-asia-2003-transportation-geographical-information-system-for->>. Acesso em: 5-agosto-2023. Citado na página 24.

JANIESCH, C.; ZSCHECH, P.; HEINRICH, K. Machine learning and deep learning. *Electronic Markets*, Springer, v. 31, n. 3, p. 685–695, apr 2021. Disponível em: <<https://doi.org/10.1007/s12525-021-00475-2>>. Acesso em: 21 jun.2024. Citado 2 vezes nas páginas 37 e 38.

JIANG, P. et al. A review of yolo algorithm developments. *Procedia computer science*, v. 199, p. 1066–1073, 2022. Disponível em: <<https://doi.org/10.1016/j.procs.2022.01.135>>. Acesso em: 25 maio.2022. Citado na página 46.

JORDAN, M. I.; MITCHELL, T. M. Machine learning: Trends, perspectives and prospects. *Science*, v. 349, n. 6245, p. 255–260, 2015. Disponível em: <<https://dl.acm.org/doi/abs/10.1145/3345120.3345147>>. Acesso em: 27 jun.2024. Citado na página 38.

JÚNIOR, J. L. F.; BARBOSA, R. E. Estimativa das cargas dinâmicas dos veículos rodoviários que atuam sobre os pavimentos e pontes. *Transportes*, v. 8, n. 2, nov 2000. Disponível em: <[DOI:https://doi.org/10.14295/transportes.v8i2.187](https://doi.org/10.14295/transportes.v8i2.187)>. Acesso em: 11-maio-2024. Citado 2 vezes nas páginas 29 e 30.

- KAEHLING, L. P.; LITTMAN, M. L.; MOORE, A. W. Reinforcement learning: A survey. *Journal of artificial intelligence research*, v. 4, p. 237–285, 1996. Disponível em: <<https://doi.org/10.1613/jair.301>>. Acesso em: 9 jun.2024. Citado na página 37.
- KHAN, S. et al. *A guide to convolutional neural networks for computer vision*. [S.l.]: Springer, 2018. 187 p. Citado na página 41.
- LECUN, Y.; BENGIO, Y. Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, v. 3361, n. 10, jan 1995. Disponível em: <<https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=e26cc4a1c717653f323715d751c8dea7461aa105>>. Acesso em: 15 jun.2024. Citado na página 39.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *Nature*, Nature Publishing Group UK London, v. 521, n. 7553, p. 436–444, 2015. Disponível em: <<https://www.nature.com/articles/nature14539>>. Acesso em: 20 out.2024. Citado na página 41.
- LI, F.-F.; JOHNSON, J.; YEUNG, S. *Lecture 11: Detection and segmentation*. [S.l.]: Stanford university, 2017. Citado na página 43.
- LIN, T.-Y. et al. *Microsoft coco: common objects in context*. In: *13th European Conference, FLEET, D. et al.(Org.). Computer Vision ECCV 2014*. Springer, 2014. 740–755 p. Disponível em: <<https://arxiv.org/pdf/1405.0312>>. Acesso em: 10 fev.2023. Citado na página 60.
- LIU, L. et al. Deep learning for generic object detection: A survey. *International journal of computer vision*, Springer, v. 128, n. 2, p. 261–318, 2020. Disponível em: <<https://link.springer.com/article/10.1007/s11263-019-01247-4>>. Acesso em: 24 out.2024. Citado 3 vezes nas páginas 42, 43 e 44.
- MAGALHÃES, L. M.; OLIVEIRA, F. H. L. d. Identificação de condições funcionais em pavimentos urbanos auxiliada por machine learning. In: *CONGRESSO DE PESQUISA E ENSINO EM TRANSPORTES. 34º Congresso de Pesquisa e Ensino em Transporte da ANPET*. [S.l.]: Editora: Associação Nacional de Pesquisa e Ensino em Transportes, 2020. p. 1070–1080. Anais [...]. Citado na página 38.
- MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *Bulletin of mathematical biology*, v. 52, p. 115–133, 1990. Disponível em: <<https://www.cs.cmu.edu/~epxing/Class/10715/reading/McCulloch.and.Pitts.pdf>>. Acesso em: 25 jun.2024. Citado na página 34.
- MEDINA, J. *Mecânica dos Pavimentos*. 1. ed. Rio de Janeiro: COPPE/UFRJ, 1997. 380 p. Citado na página 17.
- MIENYE, I. D.; SWART, T. G.; OBAIDO, G. Recurrent neural networks: A comprehensive review of architectures, variants, and applications. *Information*, MDPI, v. 15, n. 9, p. 517, 2024. Citado na página 36.
- MINSKY, M.; PAPER, S. *Perceptrons: An Introduction to Computational Geometry*. [S.l.]: MIT Press, 1969. 268 p. Citado na página 35.

- MIRANDA, F. A.; FREITAS, S. R. C. D.; FAGGION, P. L. Integração e interpolação de dados de anomalias ar livre utilizando-se a técnica de rna e krigagem. *Boletim de Ciências Geodésicas*, v. 15, n. 3, p. 428–443, set 2009. Disponível em: <<https://www.redalyc.org/pdf/3939/393937709008.pdf>>. Acesso em: 13 jun.2024. Citado na página 36.
- MOHRI, M.; ROSTAMIZADEH, A.; TALWALKAR, A. *Foundations of machine learning*. 2. ed. [S.l.]: MIT press, 2018. 504 p. Citado na página 37.
- MONARD, M. C.; BARANAUSKAS, J. A. Conceitos sobre aprendizado de máquina. *Sistemas inteligentes-Fundamentos e aplicações*, v. 1, n. 1, p. 32, 2003. Citado na página 36.
- MOTTA, L. M. G.; MEDINA, J. *Investigação e desenvolvimento em mecânica dos pavimentos na COPPE-Brasil*. [S.l.], 2006. Disponível em: <[https://www.civil.uminho.pt/revista/artigos/Num26/n\\_26\\_pag\\_85-97.pdf](https://www.civil.uminho.pt/revista/artigos/Num26/n_26_pag_85-97.pdf)>. Acesso em: 07 out.2024. Citado na página 18.
- MUELLER, J. P.; MASSARON, L. *Artificial intelligence*. 1. ed. [S.l.]: For Dummies, 2018. 336 p. Citado na página 37.
- NAIR, S. K.; HUDSON, W. R. Serviceability prediction from user-based evaluations of pavement ride quality. *Transportation Research Record*, v. 1084, p. 66–75, jan 1986. Disponível em: <<http://onlinepubs.trb.org/Onlinepubs/trr/1986/1084/1084-009.pdf>>. Acesso em: 12-maio-2024. Citado na página 31.
- NAJIBI, M.; RASTEGARI, M.; DAVIS, L. S. G-cnn: an iterative grid based object detector. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [s.n.], 2016. p. 2369–2377. Disponível em: <<https://ieeexplore.ieee.org/document/7780629>>. Acesso em: 30 out.2024. Citado na página 46.
- NELSON, J. The importance of blur as an image augmentation technique. *Roboflow*, 2020. Disponível em: <<https://blog.roboflow.com/using-blur-in-computer-vision-preprocessing/>>. Acesso em: 25 ago.2024. Citado na página 78.
- NEVES, B. C. Inteligência artificial e computação cognitiva em unidades de informação: conceitos e experiências. *Logeion: filosofia da informação*, v. 7, n. 1, p. 186–205, 2020. Disponível em: <<https://doi.org/10.21728/logeion.2020v7n1.p186-205>>. Acesso em: 25 jun.2024. Citado na página 37.
- NILSSON, N. J. *Introduction to machine learning: An early draft of a proposed textbook*. 2. ed. [S.l.]: s. n., 1998. 188 p. Citado na página 37.
- OLESKOVICZ, M.; COURY, D. V.; AGGARWAL, R. K. O emprego de redes neurais artificiais na detecção, classificação e localização de faltas em linhas de transmissão. *Revista Controle & Automação*, v. 14, n. 2, p. 138–150, jun 2003. Disponível em: <<https://www.scielo.br/j/ca/a/4ML9FmdfChGjZSPPMGsvZDk/?format=pdf&lang=pt>>. Acesso em: 10 jun.2024. Citado na página 35.
- OLIVATTO, T. F. *Identificação automática de rampas de acessibilidade apoiada por visão computacional a partir de imagens panorâmicas street-level*. Dissertação (Mestrado) — Centro de Ciências Exatas e de Tecnologia, Programa de Pós Graduação em Engenharia Urbana, Universidade Federal de São Carlos, 2021. Citado 2 vezes nas páginas 35 e 44.

- OLIVEIRA, J. J. *Experiência de implantação de sistema de gerência de pavimentos em cidade de médio porte-estudo de caso: Anápolis-GO*. Dissertação (Mestrado) — Universidade de São Paulo, Escola de Engenharia de São Carlos, São Carlos, 2013. Citado 2 vezes nas páginas 26 e 27.
- PERFORMING Convolution Operations. Programming Guide. [S.l.], 2016. Disponível em: <<<https://developer.apple.com/library/archive/documentation/Performance/Conceptual/vImage/ConvolutionOperations/ConvolutionOperations.html>>>. Acesso em: 22 out.2024. Citado na página 41.
- RAGNOLI, A.; BLASIIS, M. R.; BENEDETTO, A. Pavement distress detection methods: A review. *Infrastructures*, v. 3, n. 58, p. 1–19, set 2018. Disponível em: <<https://doi.org/10.3390/infrastructures3040058>>. Citado na página 25.
- RAUBER, T. W. Redes neurais artificiais. *Universidade Federal do Espírito Santo*, p. 1–29, may 2014. Disponível em: <[https://www.researchgate.net/profile/ThomasRauber2/publication/228686464\\_Rednes\\_neurais\\_artificiais/links/02e7e521381602f2bd000000/Redes-neurais-artificiais.pdf](https://www.researchgate.net/profile/ThomasRauber2/publication/228686464_Rednes_neurais_artificiais/links/02e7e521381602f2bd000000/Redes-neurais-artificiais.pdf)>. Acesso em: 24 jun.2024. Citado na página 35.
- REDMON, J. et al. You only look once: Unified, real-time object detection. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [s.n.], 2016. p. 779–788. Disponível em: <[https://www.cv-foundation.org/openaccess/content\\_cvpr\\_2016/papers/Redmon\\_You\\_Only\\_Look\\_CVPR\\_2016\\_paper.pdf](https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/Redmon_You_Only_Look_CVPR_2016_paper.pdf)>. Acesso em: 06 nov.2024. Citado 3 vezes nas páginas 46, 47 e 48.
- ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, American Psychological Association, v. 65, n. 6, p. 386–408, 1958. Citado na página 35.
- RUSSAKOVSKY, O. et al. Imagenet large scale visual recognition challenge. *International journal of computer vision*, v. 115, p. 211–252, Apr 2015. Disponível em: <<https://link.springer.com/article/10.1007/s11263-015-0816-y>>. Acesso em: 05 abr.2022. Citado 2 vezes nas páginas 44 e 60.
- SALVIATTO, V. H. *Ferramenta de avaliação da condição de pavimentos flexíveis urbanos baseada em uma análise multicritério*. Dissertação (Mestrado) — Universidade Estadual de Londrina, Londrina, 2021. Citado 2 vezes nas páginas 31 e 32.
- SANTIAGO, L. S. et al. Determinação do dano em pavimentos asfálticos por meio da combinação do modelo s-vecd com análises elásticas. *Revista Transporte*, v. 26, n. 2, p. 31–43, 2018. Disponível em: <<https://www.revistatransportes.org.br/anpet/article/view/1446>>. Acesso em: 15 set.2024. Citado na página 19.
- SANTOS, T. A. et al. Análise do desempenho de subleitos rodoviários à luz de conceitos da mecânica dos pavimentos. *Revista Matéria*, SciELO Brasil, v. 25, n. 03, p. E–12817, 2020. Disponível em: <<https://www.scielo.br/j/rmat/a/QbBXS LHkrbCRkJ63p86wZff/?lang=pt>>. Acesso em: 09 out.2024. Citado na página 19.
- SAVINAINEN, O. *Uncertainty Estimation and Confidence Calibration in YOLO5Face*. Dissertação (Mestrado) — Department of Electrical Engineering, Linköping University, Sweden, 2024. Citado na página 48.

- SAYERS, M. W.; KARAMIHAS, S. M. *The little book of profiling: basic information about measuring and interpreting road profiles*. 1. ed. Michigan: The University of Michigan, 1998. 100 p. Citado na página 29.
- SHAHIN, M. Y. *Pavement management for airports, roads, and parking lots*. 2. ed. New York: Springer, 2005. 571 p. Citado na página 14.
- SHARMA, A. et al. Deepinsight: A methodology to transform a non-image data to an image for convolution neural network architecture. *Scientific reports*, Nature Publishing Group UK London, v. 9, n. 1, p. 11399, 2019. Citado na página 40.
- SHOLEVAR, N.; GOLROO, A.; ESFAHANI, S. R. Machine learning techniques for pavement condition evaluation. *Automation in Construction*, v. 136, n. 104190, mar 2022. Disponível em: <<https://doi.org/10.1016/j.autcon.2022.104190>>. Acesso em: 05 abr.2022. Citado na página 38.
- SHORTEN, C.; KHOSHGOFTAAR, T. M. A survey on image data augmentation for deep learning. *Journal of big data*, Springer, v. 6, n. 1, p. 1–48, 2019. Disponível em: <<https://journalofbigdata.springeropen.com/articles/10.1186/s40537-019-0197-0#Sec3>>. Acesso em: 05 nov.2024. Citado na página 78.
- SZEGEDY, C.; TOSHEV, A.; ERHAN, D. Deep neural networks for object detection. In: *Advances in neural information processing systems*. [S.l.]: NeurIPS Proceedings, 2013. Anais [...]. Citado na página 45.
- SZELISKI, R. *Computer vision: algorithms and applications*. Springer Nature Switzerland, 2022. 273–331 p. Disponível em: <<https://link.springer.com/book/10.1007/978-3-030-34372-9>>. Acesso em: 24 out.2024. Citado na página 44.
- TORQUATO E SILVA, S. A. et al. Utilização do método dos elementos finitos em retroanálise de pavimentos flexíveis. *Researchgate*, Universidade Federal do Ceará - Departamento de Engenharia de Transportes - Programa de Pós-Graduação em Engenharia de Transportes, out 2021. Disponível em: <[https://www.researchgate.net/publication/355712806\\_UTILIZACAO\\_DO\\_METODO\\_DOS\\_ELEMENTOS\\_FINITOS\\_EM\\_RETROANALISE\\_DE\\_PAVIMENTOS\\_FLEXIVEIS](https://www.researchgate.net/publication/355712806_UTILIZACAO_DO_METODO_DOS_ELEMENTOS_FINITOS_EM_RETROANALISE_DE_PAVIMENTOS_FLEXIVEIS)>. Acesso em: 11 set.2024. Citado na página 19.
- TRIVIÑOS, A. N. S. *Introdução à pesquisa em ciências sociais: a pesquisa qualitativa em educação*. 1. ed. [S.l.]: Atlas S A, 1987. 176 p. Citado na página 50.
- VASQUEZ, C. A. *Pavement management systems on a local level*. Dissertação (Mestrado) — Utah State University, Logan, Utah, 2011. Citado na página 24.
- VENESCAU, R. Conheça o método de dimensionamento nacional. *Engenharia Rodoviária - Teoria e Aplicação*, 2020. Disponível em: <<https://engenhariarodoviaria.com.br/conheca-o-novo-metodo-de-dimensionamento-nacional-medina/>>. Acesso em: 30 set.2024. Citado na página 19.
- WANG, Y.; ZHANG, D.; KINSNER, W. *Advances in cognitive informatics and cognitive computing*. [S.l.]: Springer, 2010. v. 323. 297 p. Citado na página 38.

- WANGENHEIM, A. V. Deep learning: Detecção de objetos em imagens. *LAPiX-Image Processing and Computer Graphics Lab - INCoD - Instituto Nacional para Convergência Digital*, 2018. Disponível em: <<https://lapix.ufsc.br/ensino/visao/visao-computacionaldeep-learning/deteccao-de-objetos-em-imagens/>>. Acesso em: 24 out.2024. Citado na página 44.
- WANLI, M. et al. People logistics in smart cities. *Computer*, ACM New York, NY, USA, v. 61, n. 11, p. 54–59, 2018. Disponível em: <<https://dl.acm.org/doi/pdf/10.1145/3239546>>. Acesso em: 05 nov.2024. Citado na página 15.
- WILHEIM, J. *Planejando para o desenvolvimento*. São Paulo: Governo do Estado de São Paulo, Secretaria de Economia e Planejamento, 1978. v. 15. 220 p. Citado na página 26.
- WU, X.; SAHOO, D.; HOI, S. C. Recent advances in deep learning for object detection. *Neurocomputing*, Elsevier, v. 396, p. 39–64, 2020. Disponível em: <<https://doi.org/10.1016/j.neucom.2020.01.085>>. Acesso em: 07 nov.2024. Citado na página 48.
- YAO, Y. et al. Towards automatic construction of diverse, high-quality image datasets. *IEEE Transactions on Knowledge and Data Engineering*, v. 32, n. 6, p. 1199–1211, 2019. Disponível em: <<https://arxiv.org/pdf/1708.06495>>. Acesso em: 6 fev.2024. Citado na página 60.
- YOO, D. et al. Attentionnet: Aggregating weak directions for accurate object detection. In: *Proceedings of the IEEE international conference on computer vision*. [s.n.], 2015. p. 2659–2667. Anais [...]. Disponível em: <[https://www.cv-foundation.org/openaccess/content\\_iccv\\_2015/html/Yoo\\_AttentionNet\\_Aggregating\\_Weak\\_ICCV\\_2015\\_paper.html](https://www.cv-foundation.org/openaccess/content_iccv_2015/html/Yoo_AttentionNet_Aggregating_Weak_ICCV_2015_paper.html)>. Acesso em: 31 out.2024. Citado 2 vezes nas páginas 45 e 46.
- ZAWORSKI, R. Data augmentation techniques and pitfalls for small datasets. *Snowdog*, 2018. Disponível em: <<https://snow.dog/blog/data-augmentation-for-small-datasets>>. Acesso em: 25 ago.2024. Citado na página 78.
- ZHAO, Z.-Q. et al. Object detection with deep learning: A review. *IEEE transactions on neural networks and learning systems*, IEEE, v. 30, n. 11, p. 3212–3232, 2019. Disponível em: <<https://ieeexplore.ieee.org/document/8627998>>. Acesso em: 28 out.2024. Citado 3 vezes nas páginas 44, 45 e 46.