

UNIVERSIDADE NOVE DE JULHO – UNINOVE  
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA E GESTÃO DO  
CONHECIMENTO

GABRIEL ROSA PAZ

DETECÇÃO E MENSURAÇÃO DE FISSURAS EM EDIFICAÇÕES  
COMBINANDO O USO DE DRONES COM VISÃO COMPUTACIONAL E  
APRENDIZAGEM DE MÁQUINA

São Paulo  
2025

UNIVERSIDADE NOVE DE JULHO – UNINOVE  
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA E GESTÃO DO  
CONHECIMENTO

GABRIEL ROSA PAZ

DETECÇÃO E MENSURAÇÃO DE FISSURAS EM EDIFICAÇÕES  
COMBINANDO O USO DE DRONES COM VISÃO COMPUTACIONAL E  
APRENDIZAGEM DE MÁQUINA

Dissertação apresentada ao Programa de  
Pós-Graduação em Informática e Gestão do  
Conhecimento da UNINOVE como parte dos  
requisitos para obtenção do título de Mestre  
em Informática.

Orientador: Prof. Dr. Sidnei A. de Araújo  
Coorientador: Prof. Dr. Sergio Vicente  
Denser Pamboukian

Linha de pesquisa: LP2 – Sistemas  
Inteligentes

São Paulo

2025

Paz, Gabriel Rosa.

Detecção e mensuração de fissuras em edificações combinando o uso de drones com visão computacional e aprendizagem de máquina. / Gabriel Rosa Paz. 2026.

106 f.

Dissertação (Mestrado)- Universidade Nove de Julho - UNINOVE, São Paulo, 2026.

Orientador (a): Prof. Dr. Sidnei Alves de Araújo.

1. Construção civil. 2. Edificações. 3. Fissura. 4. Drone. 5. YOLO. 6. Visão computacional. 7. Aprendizagem de máquina.

I. Araújo, Sidnei Alves de. II. Título.

CDU 004

**ATA DE DEFESA DE DISSERTAÇÃO**

Ao décimo sexto dia do mês de dezembro de dois mil e vinte e cinco, às 14h00, do programa de Pós-Graduação, desta Universidade, reuniu-se em sessão pública a Comissão Julgadora da Dissertação de Mestrado de Gabriel Rosa Paz sob o título DETECÇÃO E MENSURAÇÃO DE FISSURAS EM EDIFICAÇÕES UTILIZANDO DRONES, VISÃO COMPUTACIONAL E APRENDIZAGEM DE MÁQUINA

Integraram a comissão os professores: Prof. Dr. Sidnei Alves de Araújo (UNINOVE), o Prof. Dr. Sergio Vicente Denser Pamboukian - Mackenzie, Prof. Dr. André Felipe Henriques Librantz - PPGI - UNINOVE, Prof. Dr. João Alexandre Paschoalin Filho - PPGCIS - UNINOVE sob a presidência do primeiro, orientador da dissertação. A banca examinadora, tendo decidido aceitar a dissertação, passou à arguição pública do candidato. Encerrados os trabalhos, os examinadores deram parecer final sobre a dissertação.

Prof. Dr. Sidnei Alves de Araújo  
Prof. Dr. Sergio Vicente Denser Pamboukian  
Prof. Dr. André Felipe Henriques Librantz  
Prof. Dr. João Alexandre Paschoalin Filho

**Parecer**

Aprovado

Aprovado

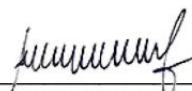
Aprovado

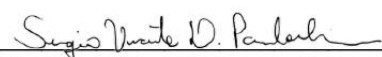
Aprovado

**Parecer:**

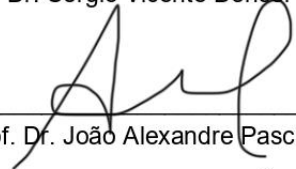
Em conclusão, o candidato foi considerado aprovado, no grau de Mestre em Informática e Gestão do Conhecimento. E, para constar, eu, Prof. Dr. André Felipe Henriques Librantz, diretor do Programa de Mestrado e Doutorado em Informática e Gestão do Conhecimento, lavrei a presente ata que assino juntamente com os membros da banca examinadora.

terça-feira, 16 de dezembro de 2025

  
Prof. Dr. Sidnei Alves de Araújo

  
Prof. Dr. Sergio Vicente Denser Pamboukian

  
Prof. Dr. André Felipe Henriques Librantz

  
Prof. Dr. João Alexandre Paschoalin Filho

  
Prof. Dr. André Felipe Henriques Librantz

## DEDICATÓRIA

A minha mãe, por me ensinar a importância da educação e por nunca me deixar desistir dos objetivos, mesmo nos momentos mais difíceis.

## **AGRADECIMENTOS**

Agradeço primeiramente ao meu Orientador o Prof. Dr. Sidnei A. de Araújo, por todo o conhecimento passado, seja durante às aulas, seja durante a orientação, agradeço também pela paciência e todo o apoio oferecido.

Ao meu Coorientador Prof. Dr. Sergio Vicente Denser Pamboukian, por todo o apoio prestado durante a pesquisa.

Ao Coordenador do PPGI o Prof. Dr. André Felipe Henriques Librantz, pelo acompanhamento próximo, incentivo constante, e as valiosas contribuições para a realização deste trabalho.

Ao PPGI e a Universidade Nove de Julho – Uninove pela oportunidade tão especial que me proporcionaram, possibilitando formação de qualidade de forma gratuita.

Aos membros da banca, pelas importantes contribuições para que o trabalho se tornasse melhor.

Ao apoio e parceria da Universidade Presbiteriana Mackenzie, ao Prof. Guilherme, a Profa. Fabíola, ao técnico Halyabner e toda a equipe do Laboratório de Geotecnologias.

Aos professores do PPGI, por compartilharem seus conhecimentos e experiências.

Ao Prof. Me. Marcos Alexandruk, ao Eng. Anderson da Silva Fatol e a todos os alunos do PPGI, em especial o grupo de pesquisa coordenado por meu orientador, por todo apoio e compartilhamento de conhecimento.

A minha família e ao meu sócio Fernando Simplício de Sousa por todo o apoio, companheirismo e paciência.

Por fim, expresso minha sincera gratidão a todas as pessoas que direta ou indiretamente colaboraram para o desenvolvimento desta pesquisa.

## RESUMO

A inspeção de estruturas e edificações é uma atividade essencial para garantir segurança, desempenho e durabilidade na construção civil. Contudo, métodos tradicionais baseados em inspeção visual apresentam limitações importantes, como dificuldade de acesso a regiões elevadas, exposição de profissionais a riscos e dependência da interpretação do inspetor. Nesse cenário, tecnologias baseadas em drones, visão computacional e aprendizagem de máquina surgem como alternativas para automatizar o processo e aumentar a confiabilidade da análise. Este trabalho propõe um método computacional para detecção e mensuração automáticas de fissuras em edificações a partir de imagens capturadas por drones. O método é composto por duas etapas: a primeira consiste na detecção automática da fissura por meio de um modelo de rede neural convolucional, treinado a partir de imagens de fissuras em obras civis; e a segunda etapa realiza a mensuração da fissura detectada (comprimento principal e angulação predominante), utilizando técnicas como *histogram of oriented gradients* (HOG), esqueletização e transformada da distância, além da conversão de medidas em pixels para unidades métricas. Os resultados obtidos nas etapas de detecção (precisão de 81,1%, recall de 69,7% e mAP@50 de 71,1%) e de mensuração (erro médio de 2,24%), juntamente com os experimentos empregando imagens em ambientes reais de edificações e obras civis, e a simulação de um cenário de monitoramento contínuo durante o voo de um drone, demonstram a viabilidade do método proposto, mesmo existindo oportunidades de aprimoramento, como a ampliação da base de treinamento e a calibração mais precisa do procedimento de mensuração geométrica das fissuras, o presente trabalho avança no desenvolvimento de um método computacional integrado que combina técnicas modernas de IA, visão computacional e o uso de drones para apoiar inspeções em edificações e obras civis.

**Palavras-chave:** Construção Civil, Edificações, Fissura, Drone, YOLO, Visão Computacional, Aprendizagem de Máquina.

## ABSTRACT

The inspection of structures and buildings is an essential activity to ensure safety, performance, and durability in civil construction. However, traditional methods based on visual inspection present significant limitations, such as difficulty accessing elevated regions, exposure of professionals to risks, and dependence on the inspector's interpretation. In this context, technologies based on drones, computer vision, and machine learning emerge as alternatives to automate the process and increase the reliability of the analysis. This work proposes a computational method for the automatic detection and measurement of cracks in buildings using images captured by drones. The method consists of two stages: the first involves automatic crack detection using a model of convolutional neural network, trained on images of cracks in civil works; the second stage performs the measurement of the detected crack (main length and predominant orientation), using techniques such as *histogram of oriented gradients* (HOG), skeletonization, and distance transform, as well as the conversion of pixel-based measurements into metric units. The results obtained in the detection stage (precision of 81.1%, recall of 69.7%, and mAP@50 of 71.1%) and in the measurement stage (average error of 2.24%), together with experiments using images from real building and civil construction environments and the simulation of a continuous monitoring scenario during drone flight, demonstrate the feasibility of the proposed method. Although opportunities for improvement were identified, such as expanding the training dataset and refining the calibration of the geometric measurement procedure, the present work advances the development of an integrated computational method that combines modern AI techniques, computer vision, and the use of drones to support inspections of buildings and civil engineering structures.

**Keywords:** Civil Construction, Buildings, Crack, Drone, YOLO, Computer Vision, Machine Learning.



## Lista de figuras

|                                                                                                 |    |
|-------------------------------------------------------------------------------------------------|----|
| Figura 1: Exemplos de anomalias em obras civis.....                                             | 27 |
| Figura 2: Fissuras, Trincas e Rachaduras .....                                                  | 28 |
| Figura 3: Régua para medir fissuras .....                                                       | 29 |
| Figura 4: Mensuração de extensão de uma fissura .....                                           | 30 |
| Figura 5: Mensuração de largura de uma fissura .....                                            | 31 |
| Figura 6: Tipos de fissuras .....                                                               | 32 |
| Figura 7: Etapas de um sistema de PID .....                                                     | 37 |
| Figura 8: Processo de aprendizado supervisionado .....                                          | 41 |
| Figura 9: Áreas de estudo IA, ML e <i>Deep Learning</i> .....                                   | 43 |
| Figura 10: Camadas de uma RNC .....                                                             | 44 |
| Figura 11: Arquitetura da rede YOLOv8 .....                                                     | 46 |
| Figura 12: Organização das palavras-chave em árvore .....                                       | 53 |
| Figura 13: Exemplos do conjunto de imagens IM_1 .....                                           | 56 |
| Figura 14: Exemplos de imagens capturadas durante a sessão 1.....                               | 57 |
| Figura 15: Exemplos de imagens capturadas durante a sessão 2 .....                              | 58 |
| Figura 16: Exemplos de imagens capturadas durante a sessão 3 .....                              | 58 |
| Figura 17: Processo de captura das imagens com drones .....                                     | 59 |
| Figura 18: Etapas do Método Proposto .....                                                      | 60 |
| Figura 19: Processos de treinamento e predição do modelo .....                                  | 60 |
| Figura 20: Exemplos de imagens anotadas do conjunto IM_1.....                                   | 61 |
| Figura 21: Segmentação e detecção de uma fissura .....                                          | 61 |
| Figura 22: Fluxo geral de execução da Etapa 2 .....                                             | 62 |
| Figura 23: Esqueletização de fissuras .....                                                     | 63 |
| Figura 24: Diferença de comprimento obtidos com soma simples e elementos 4<br>ou 8 conexo ..... | 65 |
| Figura 25: Procedimento para mensuração de fissuras .....                                       | 67 |
| Figura 26: Fissura segmentada e representação dos gradientes orientados ....                    | 68 |
| Figura 27: Histograma de angulações HOG .....                                                   | 69 |
| Figura 28: Matrizes de confusão da validação .....                                              | 72 |
| Figura 29: Matrizes de confusão do teste .....                                                  | 73 |
| Figura 30: Curvas Box-F1 e RP do Modelo treinado com 100 épocas .....                           | 74 |

|                                                                                                   |    |
|---------------------------------------------------------------------------------------------------|----|
| Figura 31: Curvas Box-F1 e RP do Modelo treinado com 300 épocas .....                             | 75 |
| Figura 32: Curvas Box-F1 e RP do Modelo treinado com 500 épocas .....                             | 75 |
| Figura 33: Imagens utilizadas para exemplificar a predição.....                                   | 77 |
| Figura 34: Anotações após predição pelo modelo treinado .....                                     | 77 |
| Figura 35: Comparativo de predições .....                                                         | 78 |
| Figura 36: Exemplos de predições corretas .....                                                   | 79 |
| Figura 37: Exemplos de falhas de predições .....                                                  | 80 |
| Figura 38: Comparação com alteração do parâmetro de confiança .....                               | 81 |
| Figura 39: Exemplos de predições corretas em um quadro do vídeo .....                             | 83 |
| Figura 40: Exemplos de predições em quadros do vídeo ilustrando casos de FP .....                 | 83 |
| Figura 41: Exemplos de predições em quadros do vídeo ilustrando casos de FN .....                 | 83 |
| Figura 42: Exemplos de mensuração de fissuras com IM_1.....                                       | 85 |
| Figura 43: Exemplos de mensuração de fissuras com IM_2.....                                       | 85 |
| Figura 44: Comparação de medição de ângulos .....                                                 | 86 |
| Figura 45: Exemplo de caso em que a abordagem por segmento de reta é melhor.....                  | 87 |
| Figura 46: Exemplo de caso em que a abordagem por HOG é melhor .....                              | 88 |
| Figura 47: Esqueletização e obtenção do comprimento .....                                         | 89 |
| Figura 48: Obtenção de comprimento por GSD e comparação com medida real (fissura simples).....    | 90 |
| Figura 49: Obtenção de comprimento por GSD e comparação com medida real (fissura ramificada)..... | 91 |
| Figura 50: Lateral da folha A4 em pixels .....                                                    | 92 |
| Figura 51: Comparação de comprimento por GSD e com medida local .....                             | 93 |
| Figura 52: Pipeline completo de detecção e mensuração .....                                       | 94 |
| Figura 53: Relatório com os resultados do pipeline .....                                          | 95 |
| Figura 54: Mensuração da largura da fissura detectada .....                                       | 97 |

## Lista de abreviaturas e siglas

|        |                                                       |
|--------|-------------------------------------------------------|
| ANAC   | Agência Nacional de Aviação Civil                     |
| ANATEL | Agência Nacional de Telecomunicações                  |
| AP     | Average Precision                                     |
| API    | Application Programming Interface                     |
| BIM    | Building Information Modeling                         |
| CLI    | Command Line Interface                                |
| CNN    | Convolutional Neural Network                          |
| CPU    | Central Processing Unit                               |
| CVPR   | Conference on Computer Vision and Pattern Recognition |
| DECEA  | Departamento de Controle do Espaço Aéreo              |
| FN     | False Negative                                        |
| FOV    | Field of View                                         |
| FP     | False Positive                                        |
| FPN    | Feature Pyramid Network                               |
| GELAN  | Generalized Efficient Layer Aggregation Network       |
| GPU    | Graphics Processing Unit                              |
| GSD    | Ground Sample Distance                                |
| HOG    | Histogram of Oriented Gradients                       |
| IA     | Inteligência Artificial                               |
| ICML   | International Conference on Machine Learning          |
| IDE    | Integrated Development Environment                    |
| IoU    | Intersection over Union                               |
| JPG    | Joint Photographic Experts Group                      |
| mAP    | mean Average Precision                                |
| ML     | Machine Learning                                      |
| MMO    | Mathematical Morphology Operators                     |
| OpenCV | Open-Source Computer Vision Library                   |
| P      | Precision                                             |

|        |                                                                       |
|--------|-----------------------------------------------------------------------|
| PAN    | Path Aggregation Networks                                             |
| PCA    | Análise de Componentes Principais                                     |
| PID    | Processamento de Imagens Digitais                                     |
| PRISMA | Preferred Reporting Items for Systematic Reviews and Meta-Analyses    |
| PVC    | POLICLORETO DE VINILA                                                 |
| R      | Recall                                                                |
| RANSAC | Random Sample Consensus                                               |
| RAM    | Random Access Memory                                                  |
| RBAC-E | Regulamento Brasileiro da Aviação Civil Especial                      |
| ReLU   | Rectified Linear Unit                                                 |
| RGB    | Red, Green and Blue                                                   |
| RNC    | Rede Neural Convolucional                                             |
| RPAS   | Remotely Piloted Aircraft System                                      |
| SARPAS | Sistema para Solicitação de Acesso de Aeronaves Remotamente Pilotadas |
| SI     | Sistema Internacional de Unidades                                     |
| SIFT   | Scale Invariant Feature Transform                                     |
| SISANT | Sistema de Aeronaves Não Tripuladas                                   |
| SURF   | Speeded-Up Robust Features                                            |
| SVM    | Support Vector Machine                                                |
| TN     | True Negative                                                         |
| TP     | True Positive                                                         |
| ULBP   | Uniform Local Binary Pattern                                          |
| VC     | Visão Computacional                                                   |
| YOLO   | You Only Look Once                                                    |

## Lista de Tabelas

|                                                                                          |    |
|------------------------------------------------------------------------------------------|----|
| Tabela 1: Técnicas Clássicas de Processamento de Imagens e Drones.....                   | 18 |
| Tabela 2: Métodos Baseados em Aprendizado Profundo.....                                  | 19 |
| Tabela 3: Métodos Híbridos e Morfológicos.....                                           | 20 |
| Tabela 4: Métricas YOLO .....                                                            | 51 |
| Tabela 5: Palavras-chaves da pesquisa bibliográfica .....                                | 53 |
| Tabela 6: Variações da expressão de busca .....                                          | 54 |
| Tabela 7: Divisão do conjunto de imagens IM_1.....                                       | 56 |
| Tabela 8: Composição do conjunto de imagens IM_2 .....                                   | 57 |
| Tabela 9: Características das GPUs utilizadas nos treinamentos e predições com YOLO..... | 62 |
| Tabela 10: Detalhes técnicos das câmeras RGB dos drones utilizados .....                 | 70 |
| Tabela 11: Resultados da etapa detecção de fissuras conjunto de validação ..             | 71 |
| Tabela 12: Resultados da etapa detecção de fissuras conjunto de teste .....              | 72 |
| Tabela 13: Comparação das medidas .....                                                  | 92 |

# Sumário

|                                                                                       |    |
|---------------------------------------------------------------------------------------|----|
| <b>1 INTRODUÇÃO</b>                                                                   | 16 |
| 1.1 CONTEXTUALIZAÇÃO DO TEMA                                                          | 16 |
| 1.2 TRABALHOS CORRELATOS E LACUNAS DE PESQUISA                                        | 18 |
| 1.3 OBJETIVOS                                                                         | 23 |
| 1.3.1 OBJETIVO GERAL                                                                  | 23 |
| 1.3.2 OBJETIVOS ESPECÍFICOS                                                           | 24 |
| 1.4 JUSTIFICATIVA                                                                     | 24 |
| 1.5 LIMITAÇÕES DA PESQUISA                                                            | 25 |
| <b>2 FUNDAMENTAÇÃO TEÓRICA</b>                                                        | 26 |
| 2.1 MANIFESTAÇÕES PATOLÓGICAS OU ANOMALIAS EM EDIFICAÇÕES E OBRAS DE CONSTRUÇÃO CIVIL | 26 |
| 2.1.1 IDENTIFICAÇÃO E MENSURAÇÃO DE FISSURAS                                          | 27 |
| 2.2 RPAS OU DRONES                                                                    | 32 |
| 2.2.1 GROUND SAMPLE DISTANCE (GSD)                                                    | 34 |
| 2.3 VISÃO COMPUTACIONAL                                                               | 36 |
| 2.3.1 HISTOGRAMA DE GRADIENTES ORIENTADOS (HOG)                                       | 38 |
| 2.3.2 ESQUELETIZAÇÃO DE OBJETOS EM IMAGENS                                            | 39 |
| 2.4 APRENDIZAGEM DE MÁQUINA                                                           | 41 |
| 2.5 APRENDIZAGEM PROFUNDA E REDES NEURAIS CONVOLUCIONAIS                              | 42 |
| 2.5.1 FRAMEWORK YOLO                                                                  | 45 |
| 2.8.1 MÉTRICAS DE AVALIAÇÃO DE DESEMPENHO                                             | 47 |
| <b>3 MATERIAIS E MÉTODOS</b>                                                          | 52 |
| 3.1 PROCEDIMENTO PARA O LEVANTAMENTO BIBLIOGRÁFICO                                    | 52 |
| 3.2 BASE DE IMAGENS EMPREGADAS NOS EXPERIMENTOS                                       | 55 |
| 3.2.1 CONJUNTO DE IMAGENS 1 (IM_1)                                                    | 55 |
| 3.2.2 CONJUNTO DE IMAGENS 2 (IM_2)                                                    | 56 |

|                                                                                                |            |
|------------------------------------------------------------------------------------------------|------------|
| 3.2.3 PROCEDIMENTO PARA CAPTURA DAS IMAGENS .....                                              | 58         |
| 3.3 MÉTODO PROPOSTO.....                                                                       | 59         |
| 3.3.1 ETAPA 1: DETECÇÃO DE FISSURAS.....                                                       | 60         |
| 3.3.2 ETAPA 2: MENSURAÇÃO DE FISSURAS .....                                                    | 62         |
| 3.3.2.1 ESQUELETIZAÇÃO.....                                                                    | 63         |
| 3.3.2.2 MENSURAÇÃO DE COMPRIMENTO.....                                                         | 63         |
| 3.3.2.3 MENSURAÇÃO DE ANGULAÇÃO.....                                                           | 66         |
| 3.3.2.4 CONVERSÃO PARA SI COM GSD .....                                                        | 69         |
| <b>4 RESULTADOS E DISCUSSÃO.....</b>                                                           | <b>70</b>  |
| 4.1 ETAPA 1: DETECAÇÃO DE FISSURAS .....                                                       | 70         |
| 4.1.1 PREDIÇÕES COM O MODELO TREINADO .....                                                    | 76         |
| 4.1.2 DETECÇÃO DE FISSURAS EM CENÁRIO DE MONITORAMENTO<br>CONTÍNUO DURANTE O VOO DO DRONE..... | 81         |
| 4.2 ETAPA 2: MENSURAÇÃO DE FISSURAS.....                                                       | 84         |
| 4.2.1 MENSURAÇÃO DO COMPRIMENTO E ANGULAÇÃO POR SEGMENTO<br>DE RETA.....                       | 84         |
| 4.2.2 MENSURAÇÃO DA ANGULAÇÃO PRINCIPAL DA FISSURA COM HOG .                                   | 86         |
| 4.2.3 MENSURAÇÃO DE COMPRIMENTO APÓS ESQUELETIZAÇÃO .....                                      | 89         |
| 4.2.4 CONVERSÃO DAS MEDIDAS DE COMPRIMENTO PARA O SI USANDO<br>GSD .....                       | 90         |
| 4.2.5 INTEGRAÇÃO DAS ETAPAS DE DETECÇÃO E MENSURAÇÃO .....                                     | 94         |
| 4.3 DISCUSSÃO DOS RESULTADOS .....                                                             | 96         |
| <b>5 CONCLUSÕES E PROPOSTAS PARA TRABALHOS FUTUROS.....</b>                                    | <b>98</b>  |
| <b>REFERÊNCIAS.....</b>                                                                        | <b>100</b> |

# 1 INTRODUÇÃO

## 1.1 CONTEXTUALIZAÇÃO DO TEMA

A inspeção sistemática de infraestruturas e edificações é uma prática essencial na engenharia civil para garantir a segurança, a durabilidade e o desempenho estrutural das construções (MOHAMMED et al., 2021). A negligência nesta etapa pode resultar na evolução das manifestações patológicas construtivas, como fissuras, que comprometem a integridade da estrutura e aumentam significativamente os custos de manutenção e recuperação (CAPORRINO, 2018).

As anomalias ou manifestações patológicas em edificações podem ser de diversos tipos, sendo as principais: fissuras, caracterizadas por aberturas lineares na superfície; fragmentação, representada pela perda de partes do concreto; eflorescência, evidenciada por manchas esbranquiçadas devido à migração de sais; manchas de corrosão, que indicam o início da oxidação das armaduras; e armadura exposta, onde o aço interno fica visível e vulnerável à corrosão, agravando os danos estruturais (DINIZ et al., 2023).

De modo geral, as manifestações patológicas em edificações podem ser atribuídas a uma variedade de fatores, como recalques diferenciais, movimentações térmicas, falhas de projeto, execução ou uso inadequado dos materiais (RODRIGUES ORTEGA; DE MARCO; FLORIAN, 2024). Dentre essas manifestações, as fissuras são as mais comuns e, também as mais negligenciadas nas etapas iniciais. A detecção precoce desse problema é fundamental para a adoção de medidas corretivas eficazes.

Entretanto, os métodos tradicionais de inspeção predial apresentam limitações consideráveis. Normalmente, essas inspeções são realizadas de forma manual, exigindo o acesso direto ao local, o que pode colocar em risco a segurança dos inspetores e, em alguns casos, inviabilizar a avaliação adequada de pontos de difícil acesso (DAVIES; KUMAR; ANTONY, 2023). Além disso, há forte dependência da experiência subjetiva dos profissionais, o que pode acarretar falhas na detecção e no diagnóstico das anomalias.



Com o avanço da tecnologia, surgem alternativas mais eficientes e seguras. Tecnologias inovadoras como os sistemas de aeronaves remotamente pilotadas (*Remotely Piloted Aircraft Systems* - RPAS) ou drones, podem ser utilizados em diversas áreas da engenharia, inclusive na inspeção de obras civis. Como abordam os trabalhos de Bravo et al. (2021) e Prabu et al. (2022) os drones permitem a captura de imagens em alta resolução de locais de difícil acesso. Dessa forma o uso desses equipamentos reduz o tempo da inspeção, minimiza riscos à integridade dos profissionais e possibilita a geração de registros mais precisos; contudo, apesar de seu potencial, o uso de drones ainda é restrito e pouco regulamentado nas práticas de inspeção rotineira. Muitos profissionais ainda desconhecem ou subutilizam esses recursos devido à ausência de normas consolidadas, altos custos iniciais de implementação ou falta de capacitação.

Neste contexto, as áreas de Visão Computacional (VC) e Aprendizagem de Máquina podem oferecer soluções promissoras para automatizar a análise de imagens obtidas pelos drones. A aplicação conjunta dessas tecnologias pode transformar completamente o paradigma atual de inspeções visuais. Por meio de algoritmos de aprendizado profundo, como as redes neurais convolucionais (RNC), é possível detectar automaticamente fissuras em estruturas de concreto com alta precisão e consistência (DAVIES; KUMAR; ANTONY, 2023).

A utilização de métodos automatizados de inspeção baseados em VC e aprendizado de máquina podem superar a acurácia de avaliações visuais humanas (CHOI et al., 2024). Além disso, esses sistemas permitem monitoramento contínuo, possibilitando ações preventivas mais eficientes, uma vez que aumentam a produtividade e otimizam recursos humanos envolvidos; e por fim evitam os altos custos dos métodos corretivos posteriores.

Dentro desse contexto a proposta desse trabalho foi desenvolver um método computacional para detecção e mensuração automáticas de fissuras em edificações e obras de construção civil a partir de imagens adquiridas por drones.

## 1.2 TRABALHOS CORRELATOS E LACUNAS DE PESQUISA

Ao realizar uma pesquisa nas bases científicas é possível encontrar na literatura estudos relacionados ao tema deste trabalho; os principais artigos selecionados durante a pesquisa bibliográfica para o presente trabalho, foram agrupados de acordo com suas abordagens em três grandes eixos:

- I) Utilização de técnicas clássicas de processamento de imagens;
- II) Aprendizagem profunda (*Deep Learning*) aplicada à detecção e segmentação automática de fissuras;
- III) Métodos híbridos e morfológicos voltados à robustez do pré-processamento e mensuração geométrica.

A Tabela 1 apresenta os estudos baseados em técnicas clássicas de processamento de imagens, os quais envolvem operações matemáticas e filtros para realçar, segmentar ou extrair informações de uma imagem.

Tabela 1: Técnicas Clássicas de Processamento de Imagens e Drones

| Trabalho              | Técnica(s) utilizada(s)                                                                   | Objetivo                                                                                             | Resultados                                                                                  | Principais Limitações                                                                                                                                                  |
|-----------------------|-------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Forcael et al. (2024) | Equalização de histograma e detecção de bordas (Canny);                                   | Avaliar fissuras em pontes via drones                                                                | Boa correlação com campo; erro < 5 mm                                                       | Detecção de bordas muito dependente da qualidade e iluminação das imagens.                                                                                             |
| Oh; Ham; Lee (2021)   | Detecção de borda (Sobel/Laplaciano/Canny)                                                | Detecção de fissuras e vazamento térmico em fachadas e otimização de rotas e uso de bateria do drone | O foco principal dos autores foi análise de eficiência (custo, rota e autonomia dos drones) | Dependência de rota/calibração do voo, uma vez que o principal objetivo do trabalho é a análise de duração da bateria e autonomia de voo.                              |
| Gao et al. (2023)     | Filtragem convencional (Canny) e adaptativa, <i>Histogram of Oriented Gradients</i> (HOG) | Detecção de fissuras em superfícies                                                                  | Acurácia ≈ 96,9%                                                                            | Sensível à condições de iluminação/ângulo de captura                                                                                                                   |
| Ko et al. (2022)      | Segmentação semântica, redes neurais convolucionais (RNC) baseada em regiões              | Detectar, segmentar e classificar angulação das fissuras                                             | IoU mediana ≈ 0,62;                                                                         | Resultados apresentam variação de acordo com o cenário (interno/externo)                                                                                               |
| Wang et al. (2024)    | Técnicas de pré-processamento, reconstrução local, euclidiana, uso de escaneamento laser  | Detecção e caracterização de fissuras (área, largura, comprimento)                                   | Recall ≈ 85%; acurácia milimétrica                                                          | Geração de Falsos Positivos (FP) em juntas de materiais, acabamentos e paredes; dependência do correto posicionamento da câmera; depende de scanner laser (alto custo) |

Fonte: Autor.

Nos trabalhos desse primeiro grupo, como os de Forcael et al. (2024) e Gao et al. (2023), as imagens obtidas por drones são processadas por algoritmos de detecção de bordas (Sobel, Laplaciano e Canny), seguidos de técnicas de

limiarização e filtragem para realce das fissuras. Em geral, esses métodos demonstram boa precisão sob iluminação controlada, mas apresentam limitações quanto à variação de perspectiva e à influência de ruídos visuais no entorno da fissura (FORCAEL et al., 2024; GAO et al., 2023). Ainda nesse eixo, Oh, Ham e Lee (2021) destacam o uso de drones em inspeções térmicas e visuais, integrando a detecção de fissuras e falhas térmicas.

O segundo grupo, apresentado na Tabela 2, compreende os trabalhos que utilizam aprendizado profundo, para detecção de fissuras em imagens.

Tabela 2: Métodos Baseados em Aprendizado Profundo

| Autor (Ano)             | Técnica Utilizada                                                                 | Objetivo                                                                             | Métricas / Resultados                                    | Principais Limitações                                                                                      |
|-------------------------|-----------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|----------------------------------------------------------|------------------------------------------------------------------------------------------------------------|
| Diniz et al. (2023)     | YOLO + RNC (multi-patologias)                                                     | Detectar várias anomalias (inclui fissuras)                                          | mAP por classe; fissura $\approx 11,8\%$ em cena ampla   | Diferença de performance entre detecção e classificação da anomalia                                        |
| Danajitha et al. (2022) | RNC + Sobel                                                                       | Fissuras em edifícios altos                                                          | Acurácia $\approx 89,8\%$                                | Generalização; calibração métrica                                                                          |
| Karmokar et al. (2020)  | RNC (AlexNet)                                                                     | Classificar fissura / sem fissura                                                    | Acurácia $\approx 98,4\%$                                | Base de imagens limitada e não detecta outras anomalias                                                    |
| Prabu et al. (2021)     | RNC + <i>Support Vector Machine</i> (SVM) (híbrido)                               | Reduzir overfitting/custo                                                            | Acurácia acima de 90%; menor custo computacional         | Dependência de base de imagens, não detecta outras anomalias além de fissuras                              |
| Choi et al. (2024)      | RNC (ResNet50) + k-means                                                          | Classificar e estimar % de área fissurada                                            | Acurácia $\approx 99,9\%$                                | Falta conversão para unidades métricas                                                                     |
| Choi et al. (2021)      | Extrator de características Scale-Invariant Feature Transform (SIFT), RNC (VGG16) | Localização espacial de fissuras dentro de um mapa de posições criado sobre a imagem | Alta precisão espacial na posição da fissura             | Custo computacional                                                                                        |
| Mohammad et al. (2019)  | RNC e descritores como HOG                                                        | Detecção de fissuras                                                                 | Acurácia acima de 87%                                    | Alto custo computacional, dependência da qualidade das imagens                                             |
| Shibano et al. (2021)   | Uso de imagens RGB e infravermelho, filtros de borda Random Forest, RNC           | Detecção automática de fissuras e eflorescência                                      | F1-score: 0,755 para fissuras e 0,803 para eflorescência | O drone precisa estar bem estabilizado para capturar as imagens                                            |
| Davies (2023)           | RNC (ResNet50) processamento em tempo real                                        | Detecção de fissuras em edifícios altos                                              | Acurácia acima de 98%                                    | Não apresenta medições, limitações de performance com baixa luminosidade                                   |
| Gan et al. (2024)       | RNC (VGG16)                                                                       | Detectar fissuras em pontes e integrar no BIM                                        | Acurácia $\approx 92\%$ ; Recall $\approx 96,5\%$        | Condições de voo; complexidade extração de informações das fissuras; etapas manuais de exportação para BIM |
| Philip et al. (2023)    | RNC (VGG16/19, ResNet50)                                                          | Classificação binária fissura / não fissura                                          | Melhor: ResNet50 (Acurácia $\approx 99,88\%$ )           | Não realiza medições e não classifica outras anomalias                                                     |

Fonte: Autor.

Neste segundo grupo Diniz et al. (2023) exploram o uso do YOLO para identificar diferentes manifestações patológicas em concreto, demonstrando a robustez do modelo na classificação de múltiplas classes. De modo semelhante, Danajitha et al. (2022) e Prabu et al. (2021) empregam RNCs e classificadores híbridos RNC-SVM em imagens capturadas por drones, obtendo acurácia superior a 90% na detecção de fissuras. Outros autores, como Karmokar et al. (2020) e Alheeti et al. (2022), também exploram RNCs e SVMs, enfatizando o potencial de generalização dessas técnicas, embora ressaltem a necessidade de conjuntos de dados amplos e diversificados para evitar sobreajuste (*overfitting*). Complementarmente, Choi et al. (2024) propõem um método de classificação e segmentação baseado em ResNet50 e k-means, que, além de identificar fissuras, estima o percentual de área fissurada, possibilitando análises quantitativas de severidade. Os resultados apresentados pelos trabalhos desse grupo reforçam a vantagem das abordagens de aprendizado profundo quanto a robustez à variação de textura e iluminação, embora sua aplicação prática ainda dependa de maior padronização dos conjuntos de imagens.

No terceiro eixo, apresentado na Tabela 3, encontram-se os métodos morfolologicamente orientados, que utilizam operadores matemáticos para reduzir ruídos e melhorar a consistência das detecções.

Tabela 3: Métodos Híbridos e Morfológicos

| Autor (Ano)                  | Técnica Utilizada                                                 | Objetivo                                                                                                                                 | Métricas / Resultados                                        | Principais Limitações                                                                                                      |
|------------------------------|-------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------|
| Abrudan et al. (2024)        | Morfologia + RNC                                                  | Reduzir ruído e custo computacional na detecção de fissuras                                                                              | Acurácia de 80,93%                                           | Depende da qualidade da segmentação                                                                                        |
| Alheeti et al. (2022)        | SVM                                                               | Classificar: fissura / não fissura                                                                                                       | Acurácia: 83 – 100%                                          | Utilização apenas de imagens de superfícies de concreto, não faz segmentação                                               |
| Ma; Li; Zhao (2020)          | RNC (YOLOv5) + segmentação adaptativa                             | Detecção de fissuras milimétricas                                                                                                        | mAP@50: 0,826                                                | Alto custo computacional                                                                                                   |
| Kim et al. (2021)            | Utilização de câmera (RGB-D), que traz informação de profundidade | Medição da largura da fissura independente do ângulo de aquisição da imagem                                                              | Média de erro de 6,56%                                       | Custo/complexidade das câmeras, não utilizou drones                                                                        |
| Flah; Suleiman; Nehdi (2020) | RNC e processamento de imagens (Otsu)                             | Classificar quanto a angulação (vertical, horizontal e diagonal) e mensurar fissuras (Comprimento - L, Largura - W e ângulo - $\theta$ ) | Acurácia acima de 96%; erro: L (1,5%), W (5%), $\theta$ (2%) | Necessita de imagens capturadas com distância de no máximo 50cm entre a câmera e a superfície onde a fissura está presente |
| Bolaybolay et al. (2022)     | Segmentação, morfologia + mensuração geométrica                   | Medir comprimento em máscaras                                                                                                            | Erro médio $\approx$ 3 – 4 mm                                | Discretização; calibração métrica                                                                                          |

Fonte: Autor.

Neste terceiro grupo Abrudan, Drăgulescu e Vizireanu (2024) apresentam uma aplicação de operações morfológicas em conjunto com redes neurais para inspeção de infraestruturas de obras civis, demonstrando ganhos de desempenho e eficiência computacional. Já Bolaybolay et al. (2022) focam na mensuração direta de fissuras segmentadas, utilizando operações morfológicas para suavização das bordas antes da extração do comprimento, obtendo erro médio inferior a 5 mm em comparação com medições de referência. Esses estudos indicam que o uso de descritores geométricos e morfologia matemática contribui para análises mais estáveis e quantitativamente consistentes.

De modo geral, observa-se que os métodos clássicos apresentam simplicidade e redução de esforço computacional, os baseados em aprendizagem profunda exibem maior robustez e adaptabilidade para classificação, mas requerem treinamento supervisionado com grandes volumes de dados e maior capacidade de processamento. Os métodos híbridos e morfológicos, por sua vez, surgem como alternativas intermediárias, conciliando desempenho e interpretabilidade. Assim, as tendências indicam uma evolução em direção a sistemas integrados que combinem detecção automática, mensuração geométrica precisa e calibração fotogramétrica, promovendo inspeções mais rápidas, seguras e quantitativamente confiáveis no contexto da engenharia civil (DINIZ et al., 2023; ABRUDAN et al., 2024; CHOI et al., 2024; FORCAEL et al., 2024).

A análise dos trabalhos selecionados nas bases de pesquisa evidencia que o tema é relevante e vem sendo amplamente explorado por pesquisadores de diversos países. No entanto observa-se que a maioria dos estudos se concentra na detecção automática das fissuras e na aplicação de RNC para este propósito (SHIBANO et al., 2024; MA; LI; ZHAO, 2024; DAVIES et al., 2023; Gao et al., 2023).

Uma lacuna da literatura diz respeito à mensuração das fissuras, que são anomalias recorrentes em obras civis. Nesse contexto, foram identificados apenas os trabalhos de Bolaybolay et al. (2022), Forcael et al. (2024), Wang et al. (2024), Gan et al. (2024) e Flah, Suleiman e Nehdi (2020), que se propuseram a realizar mensurações de características geométricas. O trabalho de Diniz et al.

(2023), embora não adentre na mensuração de fissuras também foi selecionado como uma referência importante, pelo fato de utilizar uma abordagem próxima a utilizada neste trabalho para a detecção de fissuras.

Bolaybolay et al. (2022) propuseram um método para calcular automaticamente o comprimento de fissuras em imagens segmentadas. A proposta visa automatizar a medição da extensão dessas anomalias com base em imagens binárias segmentadas, permitindo ao usuário selecionar os pontos extremos da fissura. O algoritmo emprega dilatação, afinamento e identificação de curvaturas para garantir maior precisão na contagem dos pixels pertencentes à fissura. Apesar da eficácia, o método apresenta limitações relacionadas à dependência da segmentação prévia e da calibração manual para conversão métrica (BOLAYBOLAY et al., 2022).

O trabalho de Forcael et al. (2024) foca na inspeção de estruturas de pontes, utilizando drones e técnicas clássicas de processamento de imagem para detectar fissuras, e mensurar extensão e largura dessas anomalias fazendo uso de técnicas de detecção de borda.

Wang et al. (2024) apresentam uma abordagem baseada em nuvens de pontos para detecção de fissuras e extração de atributos geométricos em edifícios históricos, composta por etapas de limpeza, segmentação, filtragem, reconstrução local de superfície e distância Euclidiana, seguida de mensuração. Os resultados de campo indicam acurácia milimétrica nas medidas e recall em torno de 85%. As limitações salientadas incluem a confusão com juntas de alvenaria e a necessidade de dados de boa qualidade para robustez em geometrias mais complexas.

Gan et al. (2024) apresentam um método para inspeção de estrutura de pontes, no qual uma RNC identifica fissuras e extrai atributos, posteriormente mapeados para um modelo BIM para gestão visual e avaliação de estado. Os autores relatam desempenho elevado na detecção (Acurácia  $\approx 92\%$ ; Recall  $\approx 96,5\%$ ), demonstrando viabilidade prática para inspeções remotas. As limitações reconhecidas incluem a dependência de condições operacionais do drone e complexidade elevada das etapas para modelar e extrair informações das fissuras.

Flah, Suleiman e Nehdi (2020) apresentam um *pipeline* que integra aprendizagem profunda para detecção e classificação (direção da fissura) e processamento de imagens para quantificação geométrica (comprimento, largura). Os autores relatam elevada acurácia de classificação (entre 96% e 97%) e baixos erros de quantificação (1,5% em comprimento, 5% em largura e 2% em ângulo), evidenciando a viabilidade de um fluxo quase automatizado para avaliação de danos em concreto, desde que haja controle de qualidade do conjunto de imagens e condições adequadas para mitigar ruídos e variações de iluminação.

Diniz et al. (2023) propuseram um método baseado em redes neurais profundas para detectar e classificar as manifestações patológicas (fragmentação, fissura, eflorescência, barras expostas e corrosão) em estruturas de concreto. A pesquisa foi estruturada em duas abordagens: uma voltada à classificação de imagens recortadas com presença de fissuras e outra focada na detecção de múltiplas manifestações patológicas (como corrosão, eflorescência e fragmentação) em imagens amplas, utilizando a arquitetura YOLOv4. Os resultados para detecção mostraram variação de desempenho entre as classes, com destaque para a detecção de barras expostas e corrosão. O estudo ressalta a importância da qualidade das imagens e da anotação das bases para melhorar os resultados e sugere futuras integrações entre os módulos de detecção e classificação para reduzir falsos positivos.

Por fim, observa-se uma lacuna quanto à proposição de métodos integrados de detecção e mensuração automatizada de atributos geométricos das fissuras, foco do presente trabalho.

### 1.3 OBJETIVOS

#### 1.3.1 OBJETIVO GERAL

Desenvolver um método computacional capaz de detectar automaticamente fissuras em obras e edificações a partir de imagens capturadas por drones e, em seguida, mensurar suas principais características geométricas, como comprimento e angulação predominante.

### 1.3.2 OBJETIVOS ESPECÍFICOS

1.3.2.1 Analisar a viabilidade, limitações e o potencial de adoção de um método computacional para detecção e mensuração automática de fissuras;

1.3.2.2 Desenvolver e treinar um modelo de RNC (YOLOv8) para detectar automaticamente fissuras em imagens de edificações e obras de construção civil.

1.3.2.3 Desenvolver um procedimento que possibilite mensuração de comprimento e ângulo de propagação de fissuras.

1.3.2.4 Desenvolver um método computacional integrando os procedimentos de detecção e mensuração de fissuras descritos nos itens 1.3.2.2 e 1.3.2.3.

### 1.4 JUSTIFICATIVA

Como já mencionado nas seções anteriores, a inspeção de estruturas e edificações na construção civil é uma atividade crítica para garantir a segurança, a durabilidade e o desempenho das obras. No entanto, os métodos tradicionais de inspeção, geralmente manuais e baseados na avaliação visual direta, apresentam limitações significativas, como a exposição de profissionais a riscos, dificuldade de acesso a pontos elevados ou de difícil alcance e forte dependência da experiência subjetiva do inspetor (DAVIES; KUMAR; ANTONY, 2023). Nesse cenário, manifestações patológicas recorrentes, como fissuras, acabam sendo frequentemente negligenciadas em suas fases iniciais, dificultando intervenções preventivas mais eficazes (CAPORRINO, 2018).

Diante deste contexto, tecnologias emergentes como os drones, integradas a métodos de VC e aprendizagem de máquina, têm se mostrado promissoras para modernizar os processos de inspeção (PRABU et al., 2022; MOHAMMED et al., 2021). Os drones permitem a obtenção de imagens aéreas em alta resolução, mesmo em locais de difícil acesso, enquanto as RNC, como YOLO, vêm sendo amplamente utilizadas para a detecção automática de fissuras e outras anomalias com alta acurácia (SHIBANO et al., 2024; MA; LI; ZHAO, 2024). Ainda assim, a aplicação prática e integrada dessas soluções encontra desafios relacionados à ausência de normatização, custos de implementação, e à limitação de muitos estudos que focam exclusivamente na



detecção binária, sem avançar para a classificação detalhada ou medição das anomalias.

A literatura atual, embora conte com propostas para a detecção automática com uso de RNC, apresenta lacunas relevantes quanto à caracterização e mensuração das anomalias. Poucos estudos, como os de Diniz et al. (2023), Bolaybolay et al. (2022), Forcael et al. (2024), Wang et al. (2024), Gan et al. (2024) e Flah, Suleiman e Nehdi (2020) abordam a classificação de diferentes tipos de anomalias ou a mensuração de dimensões e angulação de fissuras com base em atributos geométricos. Ainda mais raro é encontrar abordagens que integram detecção e mensuração em um único método automatizado e aplicável a diferentes tipologias de obras e estruturas.

Neste cenário, o presente trabalho de pesquisa justifica-se por propor uma abordagem que responde diretamente as lacunas identificadas. A proposta visa contribuir com um modelo mais completo, com potencial para apoiar a tomada de decisão técnica em inspeções prediais e estruturais, por exemplo se comparado como trabalho de Diniz et al. (2023), o presente trabalho aperfeiçoa a detecção de fissuras, implementa a segmentação automática e agrega mensuração de características geométricas das fissuras.

## 1.5 LIMITAÇÕES DA PESQUISA

Conforme observado na revisão bibliográfica apresentada na seção 1.2 a área de estudo deste trabalho é vasta e vem sendo constantemente explorada, justamente por isso não é possível exaurir o tema, dessa forma, este estudo apresenta algumas limitações que apontam oportunidades para investigações futuras, sendo as principais descritas a seguir.

- Diferentemente de Diniz et al. (2023), que abordou vários tipos de manifestações patológicas em obras de construção civil, este trabalho se concentrará na detecção e mensuração de fissuras.
- Embora este trabalho utilize imagens capturadas por drones, o foco será apenas o método computacional baseado em VC e aprendizagem de máquina para detecção e mensuração, não faz parte deste estudo

informar a localização da anomalia na edificação como fizeram Choi et al. (2021).

- A qualidade das imagens e os processos de segmentação e conversão empregados impossibilitam a medição das larguras das fissuras, uma vez que em abordagens normativas é necessário medição de valores inferiores a 1 mm.

## **2 FUNDAMENTAÇÃO TEÓRICA**

### **2.1 MANIFESTAÇÕES PATOLÓGICAS OU ANOMALIAS EM EDIFICAÇÕES E OBRAS DE CONSTRUÇÃO CIVIL**

Patologia das edificações é um campo de pesquisa da engenharia civil voltado ao estudo das anomalias construtivas, englobando suas causas, manifestações, consequências e métodos de prevenção e correção. Seu objetivo principal é compreender os mecanismos que levam ao comprometimento do desempenho dos elementos construtivos, assegurando que cada componente da edificação mantenha sua funcionalidade conforme os requisitos para os quais foi projetado. Essas anomalias podem ser decorrentes de diversos fatores, como deficiências no projeto, falhas na execução, uso inadequado de materiais ou condições ambientais adversas (CAPORRINO, 2018).

As manifestações patológicas em edificações podem ser classificadas segundo diversos critérios, tais como sua origem: física, química, mecânica ou biológica; natureza: estrutural ou não estrutural; grau de severidade: leve, moderado ou grave; bem como a fase em que ocorrem, podendo surgir ainda no projeto, durante a execução, no uso ou na manutenção da edificação. Conforme estabelece a ABNT NBR 15575:2013, conhecida como Norma de Desempenho, os sistemas construtivos das edificações devem atender a requisitos mínimos relacionados à segurança estrutural, durabilidade, estanqueidade, manutenibilidade, entre outros, de modo a assegurar as condições adequadas de conforto, saúde e segurança aos usuários ao longo da vida útil da construção (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2013).

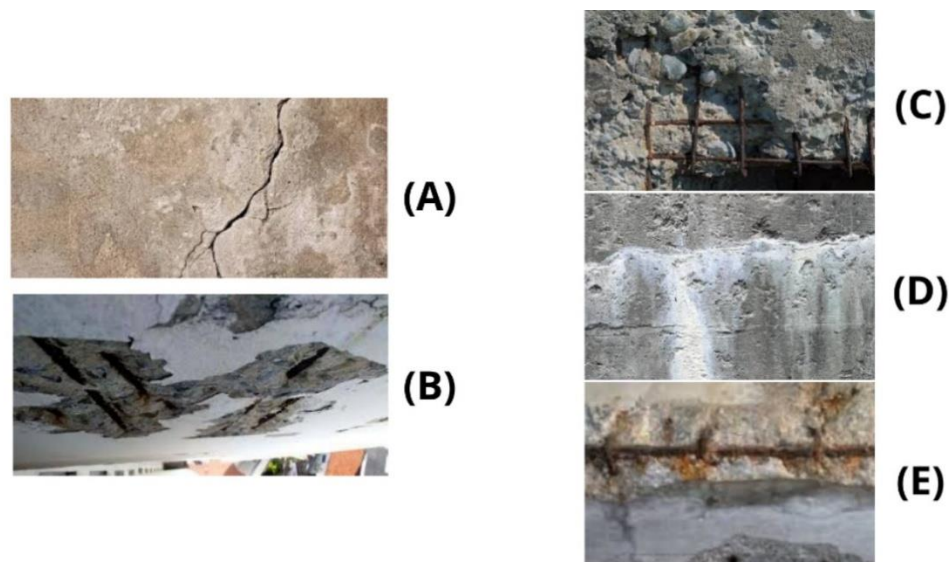
As manifestações patológicas correspondem a anomalias que reduzem o desempenho esperado das construções, podendo comprometer segurança, durabilidade e estética. Seu estudo envolve identificar causas, mecanismos de

degradação e formas de mitigação. Fissuras são recorrentes em estruturas de concreto e alvenaria, resultando, entre outros fatores, de recalques diferenciais, variações térmicas e falhas de execução (THOMAZ, 2020).

Nas fundações, problemas como recalque, erosão ou dimensionamento inadequado também se refletem em fissuras e deformações na superestrutura, exigindo diagnóstico geotécnico criterioso (ORTEGA; DE MARCO; FLORIAN, 2024).

A Figura 1 apresenta exemplos de anomalias frequentes em edificações e obras de construção civil; além das fissuras (A), outras manifestações patológicas frequentemente observadas em superfícies de concreto incluem fragmentação (B), evidenciada com a perda de partes do concreto; eflorescência (D), caracterizada por manchas esbranquiçadas devido à migração de sais; corrosão de armaduras (E) e exposição de barras de aço (C) como apontam Diniz et al. (2023). Tais defeitos podem ser detectados visualmente, mas a inspeção manual é onerosa, sujeita a erros e depende fortemente da experiência do avaliador.

Figura 1: Exemplos de anomalias em obras civis



Fonte: Autor.

### 2.1.1 IDENTIFICAÇÃO E MENSURAÇÃO DE FISSURAS

As fissuras, também chamadas de trincas ou rachaduras estão entre as manifestações patológicas mais comuns nas edificações e estruturas de concreto, podendo afetar tanto o desempenho quanto a durabilidade da construção. A ABNT NBR 15575-2:2013 define “fissura de componente

estrutural” como um seccionamento na superfície ou em toda a seção transversal de um componente, com abertura capilar, provocado por tensões normais ou tangenciais. Essas fissuras podem ser classificadas como ativas, quando sua abertura varia em função de movimentações higrotérmicas ou estruturais; ou passivas, quando permanecem com abertura constante (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2013, NBR 15575-2).

No mesmo documento normativo, o termo “trinca” é tratado como uma expressão coloquial qualitativa aplicável à fissura, ou seja, sem definição técnica precisa (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2013, NBR 15575-2). Contudo, há tentativas de caracterizar essas manifestações com base na largura de abertura. Segundo a ABNT NBR 6118:2023, para que a durabilidade da armadura não seja comprometida, a abertura máxima característica das fissuras, sob ações frequentes, não deve exceder entre 0,2 mm e 0,4 mm, dependendo da exposição (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2023, NBR 6118).

Autores como Thomaz (2020) reforçam que não há valores padronizados na literatura para a diferenciação entre fissuras, trincas e rachaduras. No entanto, em termos práticos, fissuras são aberturas que variam desde microfissuras capilares até cerca de 0,5 mm, trincas variam de 1 mm a 3 mm, e rachaduras são aberturas mais pronunciadas, geralmente acima de 3 mm, a Figura 2 fornece um exemplo de diferenciação visual de fissuras, trincas e rachaduras.

Figura 2: Fissuras, Trincas e Rachaduras



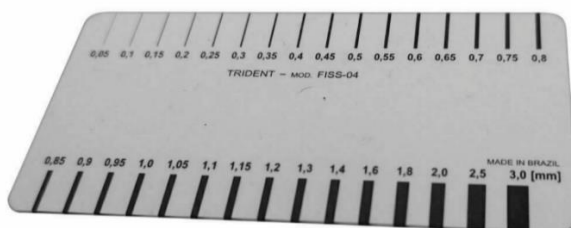
Fonte: Kasai (2021).

Em estudos internacionais, como o de Forcael et al. (2024), é adotado como limite técnico a abertura de 5 mm para diferenciar fissuras de trincas (*cracks*), conforme diretrizes do Ministério de Obras Públicas do Chile. Essa abordagem revela a existência de critérios distintos conforme o contexto técnico e geográfico, reforçando a ausência de consenso universal.

Essa variabilidade nos critérios de classificação evidencia a complexidade em adotar limites fixos para diferenciar essas anomalias em termos de nomenclatura. Dessa forma, adotou-se neste trabalho o uso do termo “fissura” de forma generalista, englobando quaisquer aberturas visíveis que indiquem perda de integridade superficial ou estrutural de componentes construtivos, independentemente de sua medida de largura exata. Essa escolha visa padronizar a análise e facilitar a aplicação de técnicas automatizadas de detecção e mensuração por meio de VC e aprendizagem de máquina.

Apenas a título de entendimento do processo de inspeção visual de fissuras em edificações, cabe informar que este geralmente utiliza como instrumento para a medição da abertura, o chamado “fissurômetro”, também conhecido como gabarito de fissuras ou calibre para medir fissuras. Trata-se de uma régua transparente, geralmente de POLICLORETO DE VINILA (PVC), com escala milimétrica graduada de 0,05 mm a 3,0 mm. Seu uso consiste na sobreposição direta sobre a fissura identificada, permitindo estimar sua abertura com base na escala impressa no instrumento. Esse procedimento é simples, de baixo custo e bastante eficaz para diagnósticos preliminares, especialmente em avaliações no local (*in loco*). A medição da largura da fissura serve de base para determinar sua gravidade e a necessidade de intervenções. A Figura 3 ilustra a régua para medir fissuras (“Fissurômetro”) modelo FISS-04 do fabricante Trident®, cujo as dimensões são 85 mm x 45 mm e 0,75 mm de largura.

Figura 3: Régua para medir fissuras



Fonte: Adaptado de TRIDENT ([s.d.]).

A mensuração das características geométricas das fissuras, como o comprimento e o ângulo de propagação, pode ser utilizada no diagnóstico de anomalias em edificações. O comprimento está diretamente relacionado à extensão do dano e à perda de integridade do elemento estrutural, podendo indicar a progressão da deterioração. Já o ângulo de propagação auxilia na identificação dos mecanismos de falha, como cisalhamento ou flexão, e influencia o comportamento estrutural, revelando se a ruptura tende a ser gradual ou súbita (THOMAZ, 2020).

O trabalho de Sahade (2005) aborda de forma abrangente o fenômeno das fissuras em alvenarias de vedação e seus revestimentos, destacando a relevância desse tipo de manifestação patológica por comprometer a durabilidade, estanqueidade e estética das edificações. As fissuras podem ser geométricas (isoladas, geralmente lineares) ou mapeadas (superficiais e em formato de rede), e ainda classificadas quanto à atividade em ativas (com variação de abertura) e passivas (estabilizadas). Essa tipologia auxilia no entendimento de suas causas, que podem estar associadas a retrações, movimentações térmicas e higroscópicas, recalques de fundação, deformações estruturais ou alterações químicas nos materiais.

O diagnóstico enfatiza a necessidade de uma análise prévia das causas e da severidade das fissuras, observando sua forma, extensão, orientação e variação de abertura ao longo do tempo. Para fissuras ativas, recomenda-se o monitoramento instrumental, avaliando a evolução da abertura e a influência de variações térmicas e estruturais, a Figura 4 exemplifica a mensuração de extensão de uma fissura por meio da medição de seu comprimento principal.

Figura 4: Mensuração de extensão de uma fissura



Fonte: Autor.



A severidade é determinada pela largura da abertura, parâmetro crítico para estimar o grau de comprometimento da vedação, sendo que aberturas acima de 0,5 mm tendem a afetar a estanqueidade e requerem tratamento específico. Sahade (2005) destaca que o sucesso na recuperação depende da correta identificação das causas e da seleção de sistemas de reparo compatíveis com o tipo e a atividade da fissura, evitando reincidências e garantindo o desempenho e a durabilidade das alvenarias, a Figura 5 apresenta a mensuração de largura de uma fissura.

Figura 5: Mensuração de largura de uma fissura



Fonte: CHERPINSKI ENGENHARIA ([s.d.]).

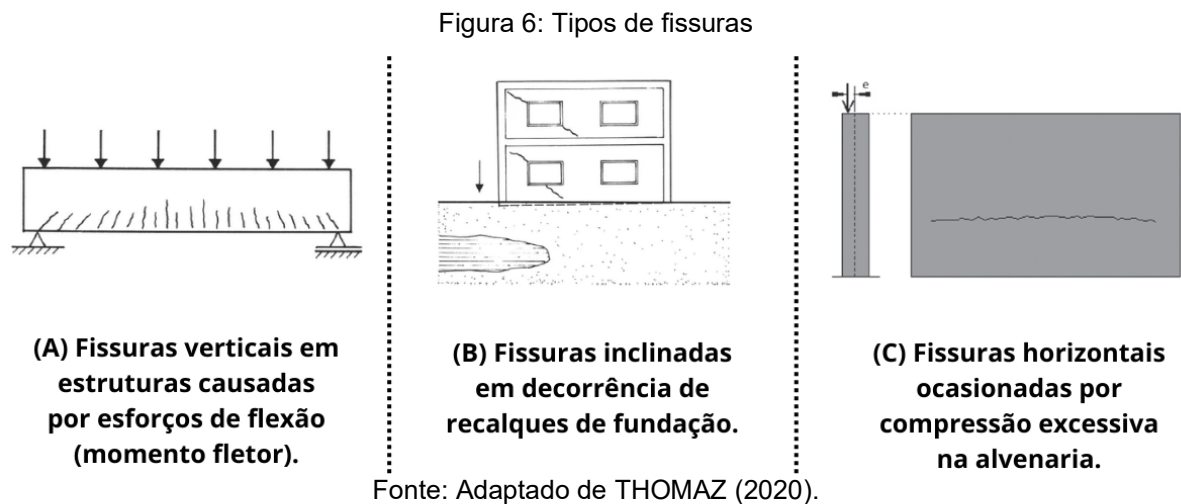
A direção e o sentido das fissuras em edificações constituem importantes indicadores para a determinação de suas causas prováveis, pois tais manifestações patológicas possuem padrões geométricos diretamente associados aos esforços mecânicos, movimentos estruturais e interações físico-químicas atuantes nos elementos construtivos; fissuras verticais tendem a estar relacionadas a sobrecargas e esforços de flexão nos elementos estruturais, sendo típicas no meio dos vãos das vigas, onde o momento fletor gera tração nas fibras inferiores e compressão nas superiores (THOMAZ, 2020)

Já as fissuras inclinadas normalmente indicam esforços de cisalhamento, deslocamentos diferenciais ou recalques de fundação, revelando movimentações estruturais incompatíveis com a capacidade de absorção de deformações dos materiais (THOMAZ, 2020)

As fissuras horizontais, por sua vez, estão frequentemente associadas a compressão excessiva na alvenaria ou a tensões de flexocompressão, podendo

também decorrer de variações térmicas e dilatações diferenciais entre materiais com coeficientes de dilatação distintos (THOMAZ, 2020).

A Figura 6 exemplifica alguns tipos comuns de fissuras e suas causas, de qualquer forma é importante ressaltar a necessidade de análise por parte de profissionais especializados e experientes, uma vez que fissuras de formatos semelhantes podem ter origens distintas, exigindo diferentes abordagens de intervenção e profilaxia (THOMAZ, 2020).



## 2.2 RPAS OU DRONES

Os RPAS, popularmente conhecidos como drones, representam uma inovação tecnológica com crescente aplicação na engenharia civil, especialmente na inspeção de estruturas e edificações. Essas aeronaves não tripuladas são capazes de realizar voos autônomos ou controlados remotamente, equipadas com câmeras de alta resolução e sensores inteligentes que permitem a captura de imagens e vídeos em áreas de difícil acesso ou que ofereçam riscos à integridade de inspetores humanos (ALHEETI et al., 2022; GAO et al., 2023).

A utilização de drones na inspeção de edificações permite uma abordagem não destrutiva, segura, rápida e precisa, otimizando o tempo de coleta de dados e reduzindo custos operacionais. Essa tecnologia é especialmente eficaz para a detecção de fissuras em fachadas, lajes, pontes e outros elementos estruturais, mesmo em locais elevados ou de difícil visualização (KARMOKAR et al., 2020).



Esses sistemas podem ser associados a técnicas de VC e aprendizagem de máquina, que processam as imagens capturadas por meio de algoritmos como RNC, SVM, ou métodos de segmentação para aumentar significativamente a acurácia da detecção e classificação de anomalias, mesmo em condições adversas, como superfícies com sombras, sujeira ou manchas (GAO et al., 2023; KARMOKAR et al., 2020; ALHEETI et al., 2022).

Os benefícios do uso de drones incluem:

- Alta resolução nas imagens capturadas, favorecendo análises detalhadas das superfícies;
- Redução do risco humano em inspeções manuais em locais perigosos;
- Rapidez na coleta de dados, especialmente em grandes estruturas;
- Viabilidade econômica, especialmente em comparação com métodos manuais ou uso de andaimes e plataformas elevatórias (GAO et al., 2023).

Com isso, a combinação entre drones, processamento de imagens digitais (PID) e IA consolida-se como uma tendência promissora para inspeções técnicas na construção civil, contribuindo significativamente para o monitoramento de anomalias e para a manutenção preventiva de estruturas.

Por fim, cabe ressaltar que no Brasil, o uso de RPAS ou drones, é regulamentado por diferentes órgãos federais, cada um responsável por aspectos específicos da operação. A Agência Nacional de Telecomunicações (ANATEL) trata da homologação do equipamento quanto às frequências de radiocomunicação. A Agência Nacional de Aviação Civil (ANAC), por sua vez, regula o uso e o registro das aeronaves por meio do Sistema de Aeronaves Não Tripuladas (SISANT), com base no Regulamento Brasileiro da Aviação Civil Especial (RBAC-E) nº 94, exigindo registro para drones com peso superior a 250g e classificando-os em três categorias conforme seu peso e risco operacional. Já o Departamento de Controle do Espaço Aéreo (DECEA) é responsável pela autorização do uso do espaço aéreo, por meio do Sistema para Solicitação de Acesso de Aeronaves Remotamente Pilotadas (SARPAS), e o Ministério da Defesa regulamenta as atividades de aerolevanteamento para fins técnicos e cartográficos. A operação legal de drones exige ainda a observância de critérios como distância mínima de pessoas não envolvidas, avaliação de risco operacional e, em alguns casos, contratação de seguro com cobertura contra danos a terceiros (ANAC, 2017; DECEA, 2016; BRAVO et al., 2021).

### 2.2.1 GROUND SAMPLE DISTANCE (GSD)

O *Ground Sample Distance* (GSD) é a distância no terreno correspondente a um pixel da imagem capturada pelo drone, também chamada de resolução espacial ao nível do solo. Em termos práticos, se  $GSD = 2,5$  cm/pixel cada pixel da imagem representa aproximadamente 2,5 cm no terreno. O GSD é um parâmetro central para planejamento de voo, dimensionamento de acurácia e mensuração em produtos fotogramétricos (WOLF; DEWITT; WILKINSON, 2014; MIKHAIL; BETHEL; MCGLONE, 2001).

Nos sistemas fotogramétricos convencionais, a relação de semelhança entre triângulos fornece a razão de ampliação entre o objeto no terreno e sua representação na imagem. Para esse modelo, utiliza-se a distância de voo, a distância focal e as características físicas do sensor. Para uma imagem nadiral (o ponto diretamente abaixo do observador) sobre terreno plano, o GSD é dado pela Equação 1.

$$GSD = \frac{H \cdot p}{f} \left[ \frac{\text{m}}{\text{pixel}} \right] \quad (1)$$

Onde:

$H$ : altura de voo sobre o terreno, em metros;

$f$ : distância focal efetiva da câmera (em mesmas unidades de  $H$ );

$p$ : tamanho físico do pixel no sensor (passo do pixel).

Uma forma alternativa de determinar o GSD consiste em substituir o tamanho do pixel pela razão entre a largura física do sensor e o número total de pixels nessa dimensão, conforme Equação 2:

$$GSD = \frac{H \cdot S_w}{f \cdot N_w} \left[ \frac{\text{m}}{\text{pixel}} \right] \quad (2)$$

Onde:

$S_w$ : largura física do sensor (m);

$N_w$ : número de pixels na largura da imagem (px).

A página web permite o cálculo do GSD.

Para fins de verificação e comparação dos valores teóricos obtidos, é possível utilizar ferramentas de cálculo disponíveis online. Entre elas, destaca-se a calculadora de GSD fornecida pelo site Handal Selaras, que permite estimar a resolução espacial a partir da informação do modelo do drone, ou dos parâmetros da câmera, e da distância de captura (HANDAL SELARAS, [s.d.]).

Quando nem todos os parâmetros necessários da câmera estão disponíveis na documentação técnica do fabricante, é possível estimar valores como a largura do sensor ou a distância focal utilizando relações geométricas do campo de visão ou *Field of View* (FOV). Uma forma comum de realizar essa estimativa é por meio da Equação 3.

$$S = 2f \cdot \tan\left(\frac{FOV}{2}\right) \quad (3)$$

Onde:

$S$ : representa a largura física do sensor (mm ou cm),

$f$ : é a distância focal real da câmera (na mesma unidade de  $S$ ),

$FOV$ : corresponde ao campo de visão, geralmente fornecido pelo fabricante em graus.

Assim, caso o valor de  $f$  seja conhecido, pode-se calcular a largura do sensor  $S$ ; da mesma forma, se  $S$  for conhecido, é possível isolar a equação para determinar a distância focal  $f$ . Essa abordagem é amplamente utilizada quando informações completas sobre o conjunto óptico não são fornecidas pelo fabricante, permitindo reconstruir as dimensões geométricas fundamentais necessárias para cálculos como o do GSD.

Por fim, quando não se dispõe de todos os parâmetros internos da câmera como distância focal real, tamanho do pixel ou dimensões completas do sensor, é possível determinar o GSD por meio de um procedimento empírico baseado em um objeto de dimensões previamente conhecidas. Essa abordagem permite estimar o GSD real da imagem sem depender integralmente das informações fornecidas pelo fabricante.

O procedimento consiste em posicionar no campo de visão da câmera um alvo de referência contendo uma medida física conhecida, como uma régua, placa graduada, padrão quadriculado ou objeto com dimensões certificadas. Em seguida, realiza-se a captura de uma fotografia mantendo o sensor o mais

perpendicular possível ao plano do alvo, reduzindo efeitos de perspectiva. Após a captura, o comprimento do objeto conhecido é medido diretamente em pixels, utilizando software de edição ou ferramentas de inspeção digital da imagem.

A partir dessas informações, o GSD empírico é obtido pela relação entre a dimensão real do objeto e seu tamanho em pixels na imagem, conforme a Equação 4.

$$GSD = \frac{D_{real}}{D_{px}} \quad (4)$$

Onde:

GSD é o Ground Sample Distance (cm/pixel),

$D_{real}$  é a dimensão real do objeto de calibração (cm),

$D_{px}$  é a medida do mesmo objeto na imagem, expressa em pixels.

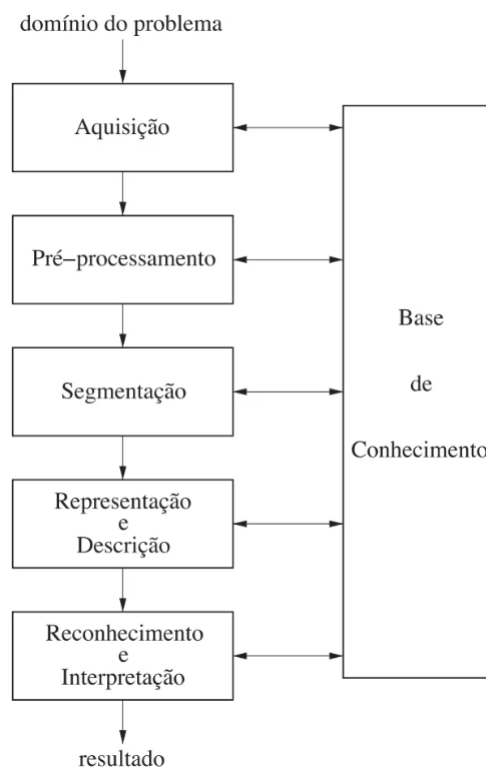
O valor resultante representa a resolução espacial real da imagem, sem a necessidade de conhecer parâmetros ópticos avançados da câmera. Esse procedimento é particularmente eficaz quando a captura ocorre em distâncias reduzidas, situação em que pequenas variações de distância, inclinação da câmera e distorções ópticas podem introduzir erros no cálculo teórico do GSD. Assim, a calibração empírica fornece uma estimativa mais realista, pois incorpora as características reais do sistema óptico, as condições de captura e eventuais distorções da lente (HONKAVAARA et al., 2007).

## 2.3 VISÃO COMPUTACIONAL

Visão Computacional (VC) é uma área interdisciplinar da computação que visa capacitar os computadores a interpretar e processar imagens e vídeos de maneira semelhante à percepção visual humana. Trata-se de uma área ampla e complexa, que integra conceitos de Inteligência Artificial (IA), aprendizagem de máquina, matemática e processamento de sinais. Os algoritmos de VC operam a partir de imagens digitais de entrada, extraindo delas características visuais, como contornos, regiões de interesse, texturas, e objetos, que são posteriormente analisadas ou classificadas. Um dos principais objetivos dessa tecnologia é automatizar a interpretação visual de forma confiável e eficiente, sendo essencial para aplicações como inspeção de estruturas, monitoramento de segurança, robótica e veículos autônomos (SPIZHEVOY; RYBNIKOV, 2018).

No contexto da engenharia civil, a VC tem se mostrado uma ferramenta poderosa na automação da inspeção de estruturas. Estudos recentes demonstram a aplicação bem-sucedida de algoritmos de VC combinados a drones para a identificação de fissuras, utilizando desde técnicas clássicas de segmentação até o uso de técnicas modernas utilizando RNC como YOLOv8. O Processamento de Imagens Digitais (PID), por sua vez, é o conjunto de técnicas e operações aplicadas a imagens digitais com o propósito de melhorar sua qualidade, extrair informações úteis ou prepará-las para análise posterior. Este campo inclui etapas como aquisição, pré-processamento, segmentação, extração de características, classificação e pós-processamento. Segundo Pedrini e Schwartz (2008), o processamento digital é essencial para lidar com problemas relacionados a ruídos, iluminação, contraste e resolução das imagens, garantindo resultados mais precisos e robustos em aplicações práticas. Técnicas como filtros espaciais, transformadas (Fourier, Wavelet), realce de bordas e limiarização são amplamente utilizadas para preparar as imagens para a etapa de análise computacional. A Figura 7 apresenta as etapas de um sistema de processamento de imagens digitais.

Figura 7: Etapas de um sistema de PID



Fonte: Pedrini e Schwartz (2008).

Na prática, bibliotecas como *Open-Source Computer Vision Library* (OpenCV) têm se consolidado como ferramentas fundamentais para o desenvolvimento de soluções baseadas em VC. Essa biblioteca oferece uma ampla gama de funções para o processamento de imagens, incluindo leitura, manipulação, visualização, detecção de bordas (algoritmo de Canny), reconhecimento de padrões e integração com algoritmos de aprendizagem de máquina. O uso de frameworks como o OpenCV permite não apenas a prototipação rápida, mas também o desenvolvimento de sistemas robustos e em tempo real para tarefas de VC (SPIZHEVOY; RYBNIKOV, 2018).

Há inúmeras técnicas de VC e PID aplicáveis à análise de anomalias estruturais, variando desde métodos clássicos baseados em filtros e gradientes até abordagens mais avançadas de extração de características. No desenvolvimento do método proposto neste trabalho, empregam-se duas dessas técnicas: o descritor *Histogram of Oriented Gradients* (HOG), utilizado para representar padrões relacionados à orientação, e o processo de esqueletização, aplicado para simplificar a forma das fissuras e permitir a mensuração de suas características geométricas. Ambas as técnicas são descritas nas seções seguintes.

### 2.3.1 HISTOGRAMA DE GRADIENTES ORIENTADOS (HOG)

Descritores de imagem são representações vetoriais ou estruturadas que codificam, de forma compacta e discriminativa, propriedades visuais relevantes de um sinal bidimensional (ou tridimensional), com o objetivo de possibilitar tarefas como reconhecimento, classificação, detecção e recuperação de padrões (GONZALEZ; WOODS, 2018; SZELISKI, 2010). Em geral, entende-se “descritor” como a saída de um processo de caracterização: a partir de uma imagem (ou de regiões/pontos de interesse), extrai-se um conjunto de características (*features*) projetadas para serem informativas e, idealmente, relativamente invariantes a certas transformações (iluminação, escala, rotação, pequenas deformações) (LOWE, 2004; SZELISKI, 2010).

Na literatura, é comum distinguir: (i) descritores globais, que resumem a imagem inteira (histogramas globais de cor/gradiente); e (ii) descritores locais, associados a pontos, regiões ou janelas que “varrem” a imagem (*Scale Invariant Feature Transform* (SIFT), *Speeded-Up Robust Features* (SURF), HOG em

janela deslizante) (LOWE, 2004; SZELISKI, 2010). Outra distinção relevante é entre descritores projetados (“*hand-crafted*”), como HOG, e descritores aprendidos por redes neurais profundas, onde a própria rede aprende representações hierárquicas a partir dos dados. Apesar do avanço dos métodos aprendidos, descritores projetados continuam úteis pela interpretabilidade, baixo custo computacional e facilidade de integração em *pipelines* clássicos (DALAL; TRIGGS, 2005; GONZALEZ; WOODS, 2018).

O HOG, proposto por Dalal e Triggs (2005), descreve uma imagem (ou região) por meio da distribuição de orientações de gradiente locais, acumuladas em histogramas. A intuição é que a geometria local de bordas e arestas — capturada por direções e magnitudes do gradiente — é altamente discriminativa para formas e aparências, enquanto variações globais de iluminação podem ser atenuadas por normalização apropriada (DALAL; TRIGGS, 2005; SZELISKI, 2010).

A magnitude e a orientação do gradiente são definidas pela Equação 5.

$$M(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2}, \quad \theta(x, y) = \arctan\left(\frac{G_y(x, y)}{G_x(x, y)}\right) \quad (5)$$

Dada uma imagem em tons de cinza  $I(x, y)$ , os gradientes de primeira ordem  $G_x$  e  $G_y$  são obtidos, tipicamente, por filtros de Sobel ou diferenças centrais, assim é possível se obter a magnitude  $M$  e a orientação do gradiente.

Para uso em HOG, emprega-se comumente a orientação não sinalizada (*unsigned*), isto é,  $\theta \in [0, 180^\circ)$ , que identifica direções opostas como equivalentes (DALAL; TRIGGS, 2005).

### 2.3.2 ESQUELETIZAÇÃO DE OBJETOS EM IMAGENS

A esqueletização é uma operação fundamental em processamento de imagens, cujo objetivo é reduzir uma forma bidimensional a uma representação unidimensional que preserve sua estrutura topológica e geométrica essencial. Essa operação permite descrever o eixo medial de objetos segmentados, mantendo propriedades como conectividade e orientação, o que é particularmente útil em análises morfométricas, como a extração de comprimento e direção de objetos segmentados de uma imagem (GONZALEZ; WOODS, 2018).

Um dos métodos mais utilizados para a obtenção do esqueleto é o baseado na transformada da distância (*Distance Transform*). Essa técnica consiste em atribuir a cada pixel do objeto um valor correspondente à menor distância em relação ao fundo. O esqueleto é então definido pelos pontos locais máximos dessa transformada, ou seja, pelos pixels equidistantes das bordas do objeto resultando em uma linha central que representa sua geometria principal (SOILLE, 2004). A formalização matemática básica pode ser descrita pela Equação 6,

$$D(p) = \min_{q \in B} d(p, q) \quad (6)$$

onde  $D(p)$  é a distância do ponto  $p$  ao conjunto de pixels de borda  $B$ , e  $d(p, q)$  representa uma métrica de distância, usualmente euclidiana. Os pontos que satisfazem condições de máxima distância local em  $D(p)$  compõem o esqueleto  $S$ .

A morfologia matemática, por sua vez, fornece um conjunto de operações que permitem refinar e simplificar o esqueleto obtido. Utilizando operadores como erosão e dilatação, é possível gerar esqueletos progressivamente mais simples e eliminar ramificações espúrias (SOILLE, 2004; SERRA, 1982). A operação de esqueletização morfológica pode ser expressa iterativamente pela Equação 7.

$$S(A) = \bigcup_{k=0}^n [(A \ominus kB) - ((A \ominus kB) \circ B)] \quad (7)$$

Em que  $A$  representa o conjunto de pixels do objeto,  $B$  é o elemento estruturante, " $\ominus$ " indica a erosão morfológica e " $\circ$ " a abertura morfológica. Essa formulação remove camadas externas do objeto até atingir o eixo central, preservando a conectividade topológica.

Combinadas, a transformada da distância e a morfologia matemática fornecem um arcabouço robusto para extração de esqueletos precisos, que podem ser aplicados em análises de fissuras. A combinação desses métodos garante redução de ruídos, continuidade do eixo medial e preservação de propriedades geométricas, características essenciais para a mensuração confiável de parâmetros estruturais em imagens digitais (VINCENT, 1993; SOILLE, 2004).

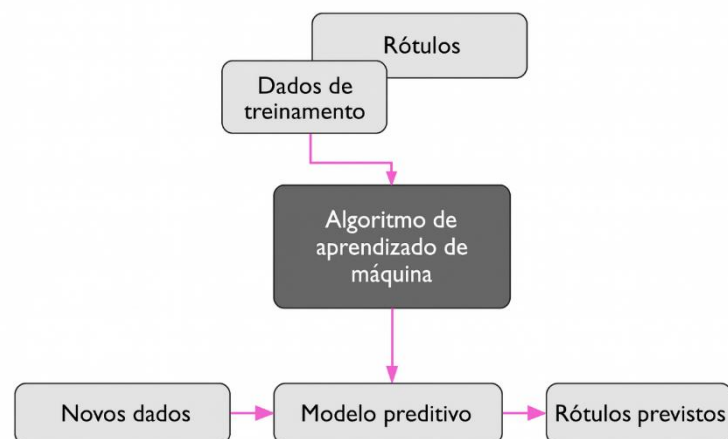


## 2.4 APRENDIZAGEM DE MÁQUINA

Aprendizagem de Máquina ou *Machine Learning* é uma subárea da IA que se dedica ao desenvolvimento de algoritmos capazes de aprender padrões a partir de dados e tomar decisões ou realizar previsões sem serem explicitamente programados para tarefas específicas. De forma geral, um modelo de aprendizagem de máquina é construído a partir da análise de dados históricos, com o objetivo de reconhecer padrões e generalizá-los para novos dados (RASCHKA; LIU; MIRJALILI, 2022).

Existem três principais paradigmas de aprendizagem de máquina: aprendizagem supervisionada, não supervisionada e por reforço. Na aprendizagem supervisionada, os modelos são treinados a partir de dados rotulados, ou seja, cada exemplo do conjunto de treinamento possui uma entrada associada a uma saída desejada. Este tipo de aprendizagem é amplamente utilizado em tarefas de classificação e regressão. Por outro lado, na aprendizagem não supervisionada, os dados não possuem rótulos, e o objetivo é encontrar estruturas, padrões ou agrupamentos nos dados, como ocorre em algoritmos de agrupamento (*clustering*) e redução de dimensionalidade. Já a aprendizagem por reforço envolve a interação de um agente com um ambiente, onde ele aprende por meio de recompensas ou punições obtidas ao executar ações, buscando maximizar a recompensa acumulada ao longo do tempo (RASCHKA; LIU; MIRJALILI, 2022). A Figura 8 apresenta de forma simplificada um diagrama do processo de aprendizado supervisionado.

Figura 8: Processo de aprendizado supervisionado



Fonte: Raschka; Liu; Mirjalili (2022).

Os modelos de aprendizagem de máquina são compostos por dois elementos principais: os parâmetros, ajustados durante o treinamento, e os hiperparâmetros, que são definidos antes do processo de treinamento e impactam diretamente o desempenho do modelo. Outro conceito fundamental é o de *overfitting* e *underfitting*. O *overfitting* ocorre quando o modelo aprende tão bem os dados de treinamento que não consegue generalizar para novos dados, enquanto o *underfitting* acontece quando o modelo não consegue capturar os padrões dos dados, resultando em baixo desempenho tanto no treinamento quanto na validação (RASCHKA; LIU; MIRJALILI, 2022).

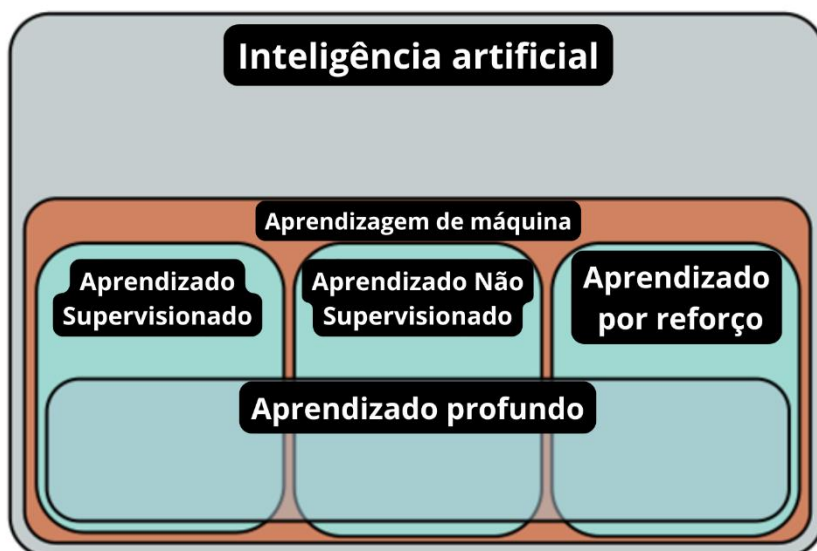
A construção de modelos envolve etapas como pré-processamento de dados, seleção de características (*feature selection*), escolha do algoritmo, treinamento, validação e teste. Além disso, é fundamental utilizar técnicas de avaliação como validação cruzada, bem como métricas como acurácia, precisão, *recall* e F1-score para mensurar o desempenho dos modelos. Atualmente, a integração entre algoritmos tradicionais de aprendizagem de máquina, e técnicas avançadas de aprendizado profundo (*Deep Learning*), tem possibilitado avanços significativos em diversas aplicações (RASCHKA; LIU; MIRJALILI, 2022).

## 2.5 APRENDIZAGEM PROFUNDA E REDES NEURAIIS CONVOLUCIONAIS

Aprendizagem profunda ou (*Deep Learning*) é um subcampo de aprendizagem de máquina, que, por sua vez, faz parte do amplo universo da IA, essa abordagem surge como uma das técnicas mais avançadas e eficazes da atualidade, sendo baseado no uso de redes neurais profundas, que são estruturas compostas por múltiplas camadas que permitem modelar relações complexas entre entradas e saídas. Essas redes são altamente versáteis, capazes de processar dados de grande dimensão, comprimento variável e estruturas internas complexas, produzindo saídas que podem variar de valores únicos a distribuições de probabilidades para múltiplas classes. Os modelos de aprendizagem profunda estão no centro das aplicações modernas, como tradução automática, reconhecimento de fala, VC e sistemas de assistentes digitais, consolidando-se como a base das soluções mais avançadas em IA (PRINCE, 2023). A Figura 9 apresenta de forma simplificada a relação das áreas

de estudo: IA, aprendizagem de máquina ou *Machine Learning* (ML) e aprendizagem profunda ou *Deep Learning*.

Figura 9: Áreas de estudo IA, ML e *Deep Learning*



Fonte: Prince (2023).

Redes Neurais Convolucionais (RNC), são um modelo de *deep learning*, constituindo uma classe especializada de redes neurais artificiais projetadas para o processamento de dados com estrutura em grade, como imagens. Diferentemente das redes neurais tradicionais, as RNC exploram as propriedades espaciais dos dados visuais por meio da aplicação de filtros convolucionais, que detectam padrões locais como bordas, texturas, formas e objetos inteiros. Isso torna as RNCs particularmente eficazes em tarefas de classificação, segmentação e detecção de objetos em imagens, sendo amplamente utilizadas em aplicações de VC (PRINCE, 2023).

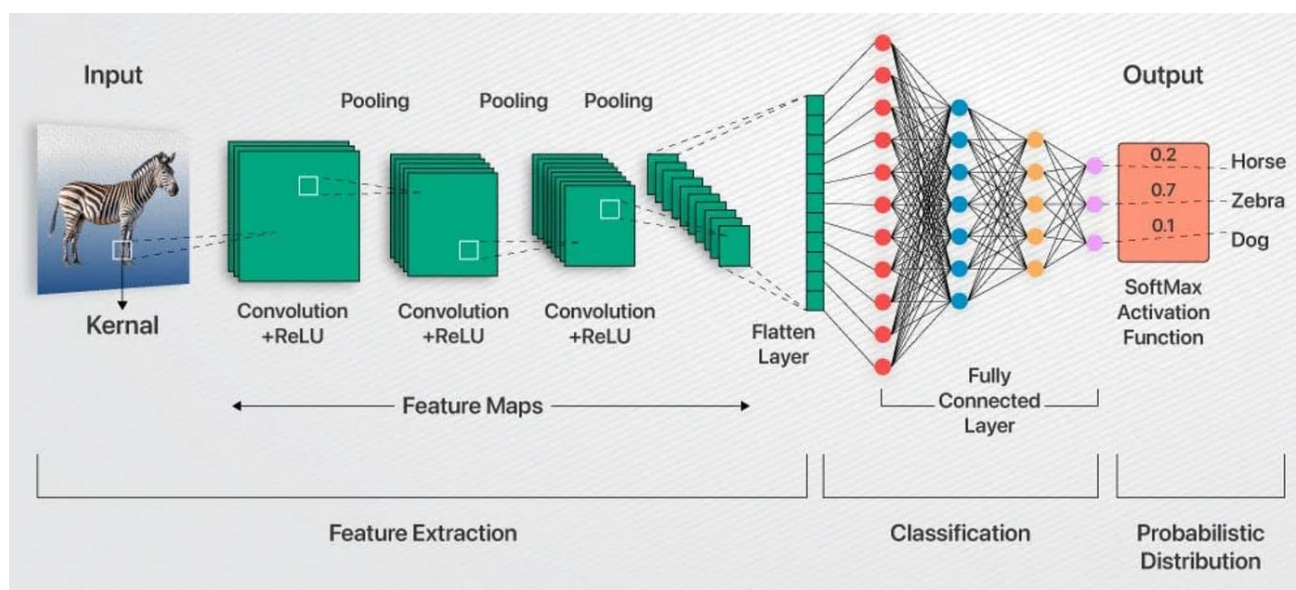
A arquitetura típica de uma RNC é composta por três tipos principais de camadas: camadas convolucionais, responsáveis pela extração de características locais; camadas de *pooling*, que reduzem a dimensionalidade dos dados preservando as informações mais relevantes; e camadas totalmente conectadas (*fully connected*), que realizam a classificação com base nas características extraídas. O uso de funções de ativação não lineares, como *Rectified Linear Unit* (ReLU), permite à rede aprender representações complexas, enquanto técnicas como *dropout*, que é a desativação aleatória de um certo número de neurônios de uma camada de rede neural, e normalização

são empregadas para melhorar a generalização e evitar *overfitting*, que é quando um modelo treinado fica viciado no conjunto de dados de treinamento e perde a capacidade de generalização (ARTASANCHEZ; JOSHI, 2020).

As RNCs têm revolucionado a área de inspeção automatizada de estruturas de concreto, especialmente na detecção de fissuras, corrosão, eflorescências e outras manifestações patológicas visuais. De acordo com Diniz et al. (2023), modelos baseados em RNC são capazes de atingir altos valores de acurácia na classificação de fissuras em imagens recortadas de estruturas de concreto, além de realizar detecção em imagens amplas utilizando arquiteturas como YOLO com resultados promissores para múltiplas classes de anomalias.

Essas redes também se beneficiam do uso de aprendizado transferido (*transfer learning*), que aproveita o conhecimento de modelos previamente treinados em grandes bases de dados, como o ImageNet, e os adapta para tarefas específicas. Isso reduz significativamente o tempo de treinamento e melhora a performance, especialmente em cenários com conjuntos de dados limitados (ARTASANCHEZ; JOSHI, 2020). A Figura 10 apresenta as camadas de uma rede RNC.

Figura 10: Camadas de uma RNC



Fonte: Singh (2024).

Existem inúmeros modelos de RNC, variando desde arquiteturas clássicas até modelos avançados de detecção em tempo real. Entre essas

alternativas, o modelo YOLOv8 foi escolhido para compor a etapa de detecção de fissuras no método proposto, por apresentar elevada eficiência computacional e capacidade de detecção rápida e robusta, características essenciais para aplicações de monitoramento estrutural próximas ao tempo real.

### 2.5.1 FRAMEWORK YOLO

O algoritmo *You Only Look Once* (YOLO) representa um dos avanços mais significativos em detecção de objetos em tempo real dentro do campo da VC. Diferentemente de abordagens clássicas que realizam a detecção por meio de múltiplas etapas (como extração de regiões de interesse, classificação e refinamento), o YOLO realiza todas essas tarefas em uma única passagem da imagem pela rede neural, o que permite alta velocidade de processamento com desempenho competitivo em termos de acurácia (ARTASANCHEZ; JOSHI, 2020).

A arquitetura do YOLO divide a imagem de entrada em uma grade e, para cada célula, prevê simultaneamente as classes dos objetos e os parâmetros das caixas delimitadoras (*bounding boxes*). Essa abordagem unificada reduz significativamente o tempo de inferência e torna o algoritmo ideal para aplicações em tempo real (PRINCE, 2023).

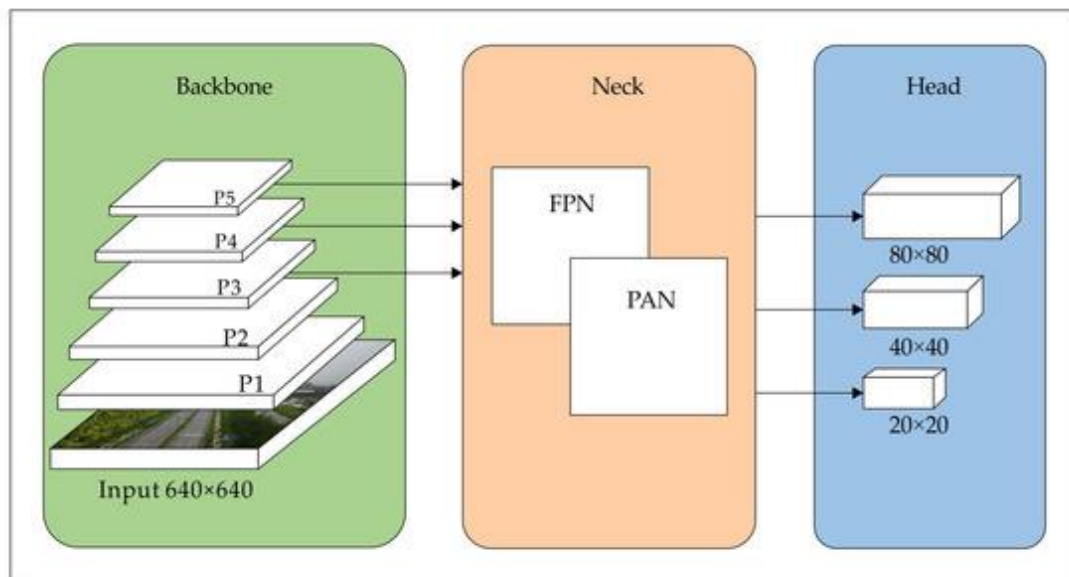
O YOLOv8, lançado em janeiro de 2023 pela Ultralytics, representa um salto evolutivo na família YOLO, consolidando-se como uma escolha pertinente para tarefas de detecção de objetos, classificação e segmentação em tempo real. Entre os principais avanços arquitetônicos destacam-se:

1. *Backbone* atualizado: essa atualização da arquitetura melhora a extração de características sem aumentar significativamente o custo computacional.
2. *Anchor-free head*: ao eliminar âncoras pré-definidas, o modelo simplifica o processo de previsão de caixas delimitadoras e reduz a necessidade de ajustes específicos para diferentes domínios.
3. *Neck* otimizado: combina múltiplas resoluções por meio de *Feature Pyramid Network* (FPN) + *Path Aggregation Networks* (PAN), melhorando a detecção de objetos em diferentes escalas.

- Ferramentas voltadas a desenvolvedores: o pacote ultralytics oferece *Command Line Interface* (CLI) unificada e *Application Programming Interface* (API) PyTorch, facilitando tarefas de treinamento, inferência e implantação.

A Figura 11 apresenta de forma simplificada a arquitetura da rede YOLOv8.

Figura 11: Arquitetura da rede YOLOv8



Fonte: LI et al. (2025).

A escolha do framework YOLO para aplicações de detecção de objetos justifica-se por sua eficiência computacional, alta precisão e arquitetura unificada de inferência, que o tornam particularmente adequado para cenários que exigem processamento em tempo real e robustez em ambientes complexos.

Diferentemente dos métodos tradicionais de visão computacional, baseados em extração de descritores (como SIFT, SURF ou HOG) e posterior classificação com modelos supervisionados, o YOLO executa a detecção de maneira integrada e holística, tratando o problema como uma única tarefa de regressão espacial. Essa abordagem permite que a rede estime simultaneamente as coordenadas das caixas delimitadoras (*bounding boxes*) e as probabilidades de classe, em um único passo de inferência (REDMON et al., 2016). Além disso, comparado a arquiteturas de múltiplos estágios, como R-

CNN, Fast R-CNN e Faster R-CNN, o YOLO apresenta desempenho significativamente superior em termos de velocidade de processamento, mantendo níveis competitivos de *mAP* (*mean Average Precision*). A evolução para versões mais recentes, como o YOLOv8, incorporou avanços estruturais como backbone mais leve, detecção ancorada em *features* multiescala e treinamento otimizado, que ampliam a capacidade de generalização do modelo mesmo em bases de imagens limitadas ou não padronizadas. Assim, o YOLO destaca-se como uma solução eficiente e precisa para a detecção automática de fissuras em imagens de inspeção, conciliando desempenho e confiabilidade em aplicações de engenharia civil.

### 2.8.1 MÉTRICAS DE AVALIAÇÃO DE DESEMPENHO

Para que seja possível avaliar o modelo treinado da primeira etapa do método computacional proposto neste trabalho, a qual é responsável pela detecção automática de fissuras, e sua capacidade de inferência é necessária a utilização de métricas padronizadas de desempenho, as principais métricas disponibilizadas pelo modelo YOLO são divididas em duas categorias: *Box*, que se refere à detecção por caixas delimitadoras (*bounding boxes*), e *Mask*, que se refere à segmentação da forma da fissura por meio de máscaras, para ambas as categorias as métricas fornecidas são precisão ou *precision* (*P*), *recall* (*R*) e *mean Average Precision* (*mAP@0,50* e *mAP@0,50:0,95*).

A Equação 8 apresenta a expressão matemática da precisão (*P*).

$$P = \frac{TP}{TP+FP} \quad (8)$$

A precisão (*P*) é a razão entre os verdadeiros positivos (*TP*) e todos os itens classificados como positivos, ou seja, todos os positivos preditos pelo sistema incluindo os falsos positivos (*FP*); ela indica a proporção de predições corretas entre todas as que o modelo fez. Quanto mais próxima de 1,0, melhor. Valores abaixo de 0,5 geralmente indicam que o modelo está fazendo muitas predições incorretas (falsos positivos).

Na Equação 9 a expressão matemática da sensibilidade ou *Recall* (*R*) é apresentada.

$$R = \frac{TP}{TP+FN} \quad (9)$$

O *Recall* ( $R$ ) é a razão entre os verdadeiros positivos ( $TP$ ) e todos os itens realmente positivos, ou seja, as verdadeiras instâncias, incluindo os falsos negativos ( $FN$ ); ela mostra a capacidade do modelo de encontrar fissuras que realmente existem na imagem. Também deve estar próximo de 1,0 para ser considerado bom. Valores abaixo de 0,5 mostra que o modelo está deixando de detectar parcela significativa de objetos (ou seja, muitos falsos negativos), o que é um problema sério em aplicações onde a detecção completa é importante.

A acurácia é uma métrica fundamental utilizada para avaliar o desempenho de modelos de classificação, expressando a proporção de previsões corretas em relação ao total de previsões realizadas. Em termos gerais, ela quantifica o quão próximo o modelo está de reproduzir corretamente as classes reais, englobando tanto os casos positivos quanto negativos identificados de forma correta (REZENDE, 2022). Sua formulação matemática básica é apresentada na Equação 10.

$$Acurácia = \frac{TP+TN}{TP+TN+FP+FN} \quad (10)$$

Onde:

- $TP$  (*True Positives*) representa o número de amostras corretamente classificadas como pertencentes à classe positiva;
- $TN$  (*True Negatives*) indica o número de amostras corretamente classificadas como pertencentes à classe negativa;
- $FP$  (*False Positives*) corresponde às amostras incorretamente classificadas como positivas;
- $FN$  (*False Negatives*) refere-se às amostras incorretamente classificadas como negativas.

Assim, a acurácia fornece uma visão geral da capacidade global do modelo em classificar corretamente os dados. Entretanto, em bases desbalanceadas, nas quais há discrepância entre o número de amostras das classes, a métrica pode apresentar viés, motivo pelo qual recomenda-se o uso conjunto de outras métricas, como *Precisão*, *Recall* e *F1-score* (REZENDE, 2022).



O *F1-score* é uma métrica de avaliação que combina de forma harmônica as medidas de Precisão (*P*) e *Recall* (*R*), buscando um equilíbrio entre ambas. É especialmente útil em problemas de classificação nos quais há desequilíbrio entre classes, pois fornece uma visão mais representativa do desempenho do modelo do que a acurácia isoladamente (REZENDE, 2022). Sua expressão matemática básica é dada pela Equação 11.

$$F1 = 2 \times \frac{(P \times R)}{(P + R)} \quad (11)$$

Onde:

*P* (Precisão) indica a proporção de amostras corretamente identificadas como positivas entre todas as amostras classificadas como positivas;

*Recall* (Sensibilidade) representa a proporção de amostras positivas corretamente identificadas entre todas as amostras realmente positivas.

O *F1-score* varia de 0 a 1, sendo que valores próximos a 1 indicam melhor desempenho do modelo, ao passo que valores próximos de 0 indicam baixa capacidade de identificar corretamente as instâncias positivas. Assim, o *F1-score* é amplamente empregado em tarefas de visão computacional e aprendizado de máquina, por equilibrar a penalização tanto de falsos positivos quanto de falsos negativos.

A *mAP* é uma métrica bastante utilizada para avaliação de sistemas de recomendação e de modelos de segmentação de imagens, onde é necessário detectar e segmentar objetos; e ela é obtida pela média aritmética das *Average Precisions* (*AP*) (FILHO, 2023). Para segmentação de imagem, o *AP* pode ser calculado pela Equação 12, dependendo do limiar de *IoU* e do método de interpolação (ZHAO et al., 2022).

$$AP = \int_0^1 P(R) dR \quad (12)$$

A *AP* de uma classe é uma média das precisões em diferentes pontos de *R* (sensibilidade) ao longo de uma curva *P* x *R*.

Por fim O *mAP* é a média *AP* em diferentes limiares de Interseção sobre União (*IoU*), e a expressão matemática do *mAP* varia um pouco dependendo de como o *AP* é calculado, mas geralmente é dada pela equação 13.

$$mAP = \frac{1}{C} \sum_{c=1}^C AP_c \quad (13)$$

Onde  $C$  é o número total de classes (por exemplo, diferentes tipos de objetos ou categorias) e  $AP_c$  é a precisão média para cada classe  $c$ .

A  $mAP@0.50$  (com  $IoU \geq 50\%$ ) é uma métrica mais tolerante, e valores acima de 0,5 já indicam bom desempenho, ou seja, é considerado acerto do modelo se ele localiza o objeto com pelo menos 50% de sobreposição com a posição real (ULTRALYTICS, 2023<sup>a</sup>).

A Equação 14 apresenta a expressão da *Intersection over Union* ( $IoU$ ), que é uma métrica que mede a sobreposição entre a caixa predita pelo modelo (*bounding box* da predição), e a caixa real (*bounding box* da anotação manual) (ULTRALYTICS, 2023<sup>a</sup>).

$$IoU = \frac{\text{Área da interseção}}{\text{Área da união}} \quad (14)$$

|  $IoU = 1,0$  | → A caixa predita é idêntica à real (perfeita)

|  $IoU = 0,0$  | → Nenhuma sobreposição

|  $IoU \geq 0,5$  | → Muitas vezes considerada “acerto” na detecção

A  $mAP@0,50:0,95$ , representada pelas Equações 15 e 16, por sua vez considera a média do  $AP$  calculado em dez diferentes limiares de  $IoU$ , variando de 0,50 até 0,95, em passos de 0,05:

$$IoU = \{0,50; 0,55; 0,60; 0,65; 0,70; 0,75; 0,80; 0,85; 0,90; 0,95\} \quad (15)$$

Assim:

$$mAP@0,50:0,95 = \frac{1}{10} \sum_{t=0,50}^{0,95} AP@t \quad (16)$$

Portanto, para ser considerada um “acerto” em cada limiar, a detecção precisa ter  $IoU \geq$  limiar atual. (ULTRALYTICS, 2023<sup>a</sup>). Valores abaixo de 0,3 são considerados fracos.

A Tabela 4 resume os intervalos de valores das métricas.

Tabela 4: Métricas YOLO

| Métrica                       | Valor ideal | Desempenho moderado | Desempenho fraco |
|-------------------------------|-------------|---------------------|------------------|
| Precisão ( <i>Precision</i> ) | $\geq 0.8$  | 0.5 – 0.79          | $< 0.5$          |
| Recall                        | $\geq 0.8$  | 0.5 – 0.79          | $< 0.5$          |
| mAP@0.50                      | $\geq 0.5$  | 0.3 – 0.49          | $< 0.3$          |
| mAP@0.50:0.95                 | $\geq 0.3$  | 0.15 – 0.29         | $< 0.15$         |

Fonte: Autor.

No modelo YOLO, o valor de confiança (*51onfidence score*) associado a cada *bounding box* é definido a partir da probabilidade de existência de um objeto naquela região, denominada *objectness score*. Esse valor é obtido pela aplicação de uma função sigmoide na saída da rede, garantindo que o resultado esteja entre 0 e 1. Durante o treinamento, a rede aprende a estimar corretamente essa probabilidade a partir das anotações reais, sendo penalizada quando atribui alta probabilidade a regiões sem objeto ou baixa probabilidade a regiões que contêm objeto. Em cenários com múltiplas classes, esse *objectness score* é multiplicado pela probabilidade condicional de classe; entretanto, no caso de apenas uma classe, o valor de confiança corresponde diretamente ao *objectness score*.

Quando o modelo YOLO é treinado para detectar mais de uma classe (caso de múltiplas classes), o *51onfidence score* de uma *bounding box* para uma classe  $c$  é definido pela expressão da Equação 17.

$$Confidence_c = P_{(objeto)} \times P_{(classe = c \mid objeto)} \quad (17)$$

Onde:

$P_{(objeto)}$  é o *objectness score*, ou seja, a probabilidade de existir um objeto dentro da *bounding box*.

$P_{(classe = c \mid objeto)}$  é a probabilidade condicional de que o objeto pertença à classe  $c$ .

Quando o modelo é treinado para detectar somente uma classe (por exemplo, fissura), não há necessidade de estimar probabilidades condicionais de classe. Nesse caso, o *confidence score* reduz-se a equação 18.

$$Confidence = P_{(objeto)} \quad (18)$$

Ou seja, o valor de confiança corresponde diretamente ao *objectness score*.

### 3 MATERIAIS E MÉTODOS

Neste capítulo são apresentados os materiais e métodos empregados na pesquisa. Primeiro são apresentados os procedimentos adotados no levantamento bibliográfico na seção 3.1. Na sequência, a base de imagens empregada nos experimentos conduzidos é apresentada na seção 3.2. Por fim, o método proposto é descrito em detalhes na seção 3.3.

#### 3.1 PROCEDIMENTO PARA O LEVANTAMENTO BIBLIOGRÁFICO

O levantamento bibliográfico foi conduzido a partir das seguintes etapas: seleção das bases de dados científicas; escolha das palavras-chave e formulação das estratégias de busca; definição dos critérios de elegibilidade (inclusão e exclusão); e análise das publicações selecionadas. As duas últimas etapas foram orientadas pela metodologia *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA), conforme proposta por Moher et al. (2010).

A seleção de bases científicas de alta qualidade é crucial para assegurar a robustez e relevância de uma pesquisa. Neste estudo, foram escolhidas as bases Scopus, Compendex e Web of Science, devido à sua abrangência, credibilidade e rigor na indexação de publicações. A Compendex, voltada à engenharia e ciências aplicadas, oferece mais de 20 milhões de registros especializados, sendo essencial para pesquisas técnicas. A Scopus, uma das maiores bases multidisciplinares, cobre mais de 70 milhões de documentos, destacando-se pelas ferramentas de análise de citações e pela ampla cobertura de periódicos e conferências importantes, como *International Conference on Machine Learning* (ICML) e *Conference on Computer Vision and Pattern*

*Recognition* (CVPR), especialmente relevantes na ciência da computação. Já a Web of Science se destaca pela seleção criteriosa de publicações e pelas ferramentas avançadas de análise de impacto e tendências em áreas como física, matemática, química e ciências ambientais. A combinação dessas três bases garante uma investigação abrangente, atualizada e alinhada com o estado da arte nas ciências exatas.

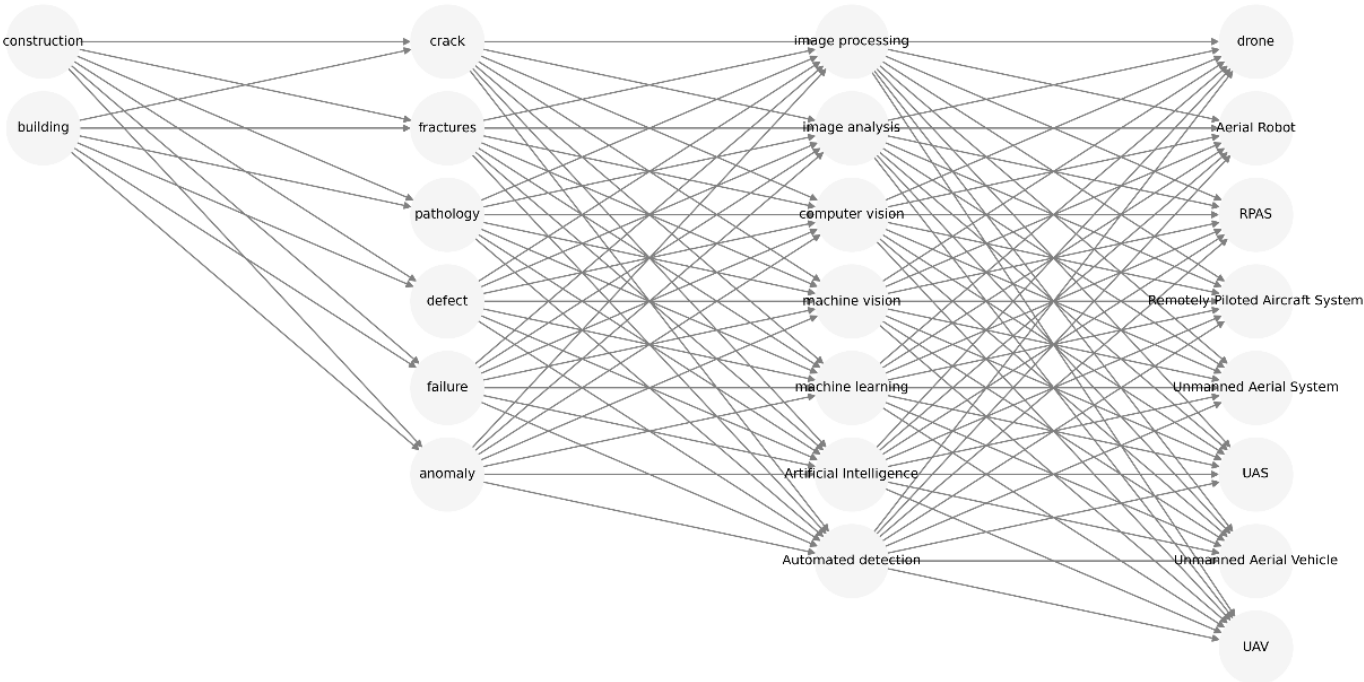
Para a realização da pesquisa sobre o tema “detecção e mensuração automática de fissuras em edificações a partir de imagens de drones” foram definidas as palavras-chave apresentadas na Tabela 5 e na Figura 12 (em formato de árvore), no idioma inglês.

Tabela 5: Palavras-chaves da pesquisa bibliográfica

| Grupo 1      | Grupo 2   | Grupo 3                 | Grupo 4                          |
|--------------|-----------|-------------------------|----------------------------------|
| construction | crack     | image processing        | drone                            |
| building     | fractures | image analysis          | Aerial Robot                     |
|              | pathology | computer vision         | RPAS                             |
|              | defect    | machine vision          | Remotely Piloted Aircraft System |
|              | failure   | machine learning        | Unmanned Aerial System           |
|              | anomaly   | image processing        | UAS                              |
|              |           | Artificial Intelligence | Unmanned Aerial Vehicle          |
|              |           | Automated detection     | UAV                              |

Fonte: Autor.

Figura 12: Organização das palavras-chave em árvore



Fonte: Autor.

A partir das palavras-chave e do período de publicação dos artigos, definido entre janeiro de 2020 e janeiro de 2025, foi elaborada a expressão de busca, e suas variações para cada uma das bases conforme Tabela 6.

Tabela 6: Variações da expressão de busca

|                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|-----------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Expressão Base</b> | (construction OR building) AND (crack OR fractures OR pathology OR defect OR failure OR anomaly) AND ("image processing" OR "image analysis" OR "computer vision" OR "machine vision" OR "machine learning" OR "Artificial Intelligence" OR "Automated detection") AND (drone OR "Aerial Robot" OR RPAS OR "Remotely Piloted Aircraft System" OR "Unmanned Aerial System" OR UAS OR "Unmanned Aerial Vehicle" OR UAV)                                                                                                                                                                       |
| <b>Scopus</b>         | TITLE-ABS-KEY((construction OR building) AND (crack OR fractures OR pathology OR defect OR failure OR anomaly) AND ("image processing" OR "image analysis" OR "computer vision" OR "machine vision" OR "machine learning" OR "Artificial Intelligence" OR "Automated detection") AND (drone OR "Aerial Robot" OR RPAS OR "Remotely Piloted Aircraft System" OR "Unmanned Aerial System" OR UAS OR "Unmanned Aerial Vehicle" OR UAV)) AND (LANGUAGE(english) OR LANGUAGE(portuguese)) AND (PUBYEAR > 2019)                                                                                   |
| <b>Web Of Science</b> | TS=((construction OR building) AND (crack OR fractures OR pathology OR defect OR failure OR anomaly) AND ("image processing" OR "image analysis" OR "computer vision" OR "machine vision" OR "machine learning" OR "Artificial Intelligence" OR "Automated detection") AND (drone OR "Aerial Robot" OR RPAS OR "Remotely Piloted Aircraft System" OR "Unmanned Aerial System" OR UAS OR "Unmanned Aerial Vehicle" OR UAV)) AND PY=(2020-2025)                                                                                                                                               |
| <b>Compendex</b>      | ((((construction OR building) AND (crack OR fractures OR pathology OR defect OR failure OR anomaly) AND ("image processing" OR "image analysis" OR "computer vision" OR "machine vision" OR "machine learning" OR "Artificial Intelligence" OR "Automated detection") AND (drone OR "Aerial Robot" OR RPAS OR "Remotely Piloted Aircraft System" OR "Unmanned Aerial System" OR UAS OR "Unmanned Aerial Vehicle" OR UAV)) AND ((LANGUAGE:english) OR (LANGUAGE:portuguese)) AND (PUBLISHED:2020 OR PUBLISHED:2021 OR PUBLISHED:2022 OR PUBLISHED:2023 OR PUBLISHED:2024 OR PUBLISHED:2025)) |

Fonte: Autor.

É importante esclarecer que não foram utilizadas palavras-chave em português, uma vez que todos os artigos nas bases pesquisadas independente da língua original possuem o resumo e as palavras-chave em inglês (*abstract* e *key words*) permitindo localizar trabalhos de autores brasileiros, como o trabalho de Diniz et al. (2023).

## 3.2 BASE DE IMAGENS EMPREGADAS NOS EXPERIMENTOS

Para a condução de experimentos com o método proposto neste trabalho foi composta uma base contendo conjuntos de imagens, o primeiro (IM\_1) utilizado para treinamento, validação e testes do modelo de detecção das fissuras, e o segundo (IM\_2) utilizado para submeter ambas as etapas do método proposto às imagens capturadas por drones em ambientes de obras civis.

### 3.2.1 CONJUNTO DE IMAGENS 1 (IM\_1)

Para o treinamento do modelo YOLOv8, responsável por detectar e segmentar automaticamente fissuras em imagens de edificações, foi utilizada a base de imagens RGB denominada *crack*<sup>1</sup> (*Instance Segmentation*), disponível no repositório Roboflow<sup>2</sup> Universe, um projeto de segmentação de instâncias para visualizar fissuras detectadas em concreto. Para melhor entendimento dos procedimentos experimentais e dos resultados, denominamos essa base de imagens de IM\_1.

A motivação principal para a escolha desse conjunto foi o fato delas já estarem devidamente rotuladas (marcadas com a máscara que indica o local da fissura). As imagens desse conjunto resultaram de um processo de *data augmentation* realizado sobre uma base inicial de 1.551 imagens, utilizando os seguintes parâmetros da plataforma Roboflow:

- Espelhamento (*Flip*): Horizontal
- Rotações de 90°: Sentido Horário, Sentido Anti-horário
- Recortes (*Crop*): Zoom mínimo de 0%, Zoom máximo de 30%
- Rotações: Entre -15° e +15°
- Alteração de cores (Hue): Entre -10° e +10°
- Saturação (*Saturation*): Entre -25% e +25%
- Brilho (*Brightness*): Entre -20% e +20%
- Desfoque (*Blur*): Até 1px

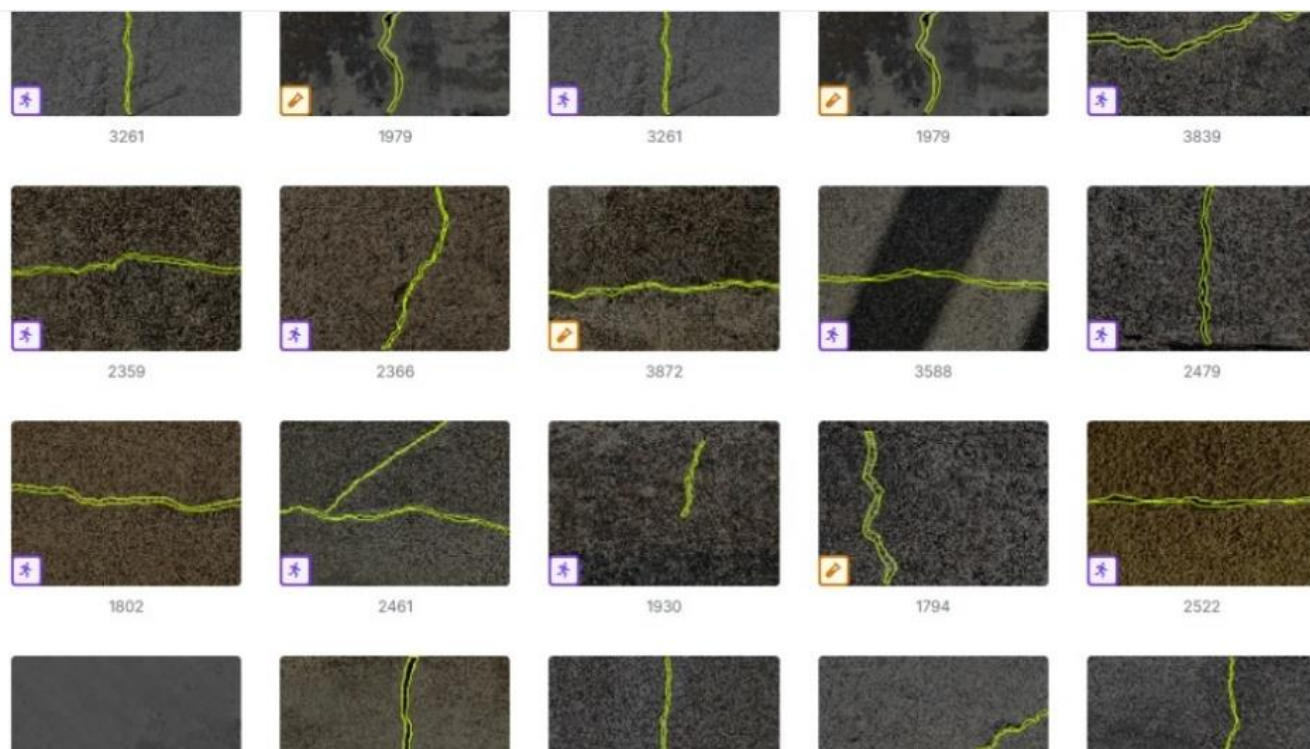
---

<sup>1</sup> Base crack presente no repositório Roboflow Universe neste link: <https://universe.roboflow.com/university-bswxt/crack-bphdr> (acesso em 20/06/2025)

<sup>2</sup> Roboflow é uma plataforma de software como serviço (SaaS) que ajuda desenvolvedores a construir e implantar aplicações de VC (<https://roboflow.com/>)

As imagens também foram pré-processadas e padronizadas, sendo todas escalonadas para dimensão  $416 \times 416$  pixels. Alguns exemplos de imagens da base IM\_1 são ilustrados na Figura 13.

Figura 13: Exemplos do conjunto de imagens IM\_1



Fonte: Autor.

As 4.029 imagens do conjunto IM\_1 foram divididas para treinamento, validação e testes conforme apresentado na Tabela 7.

Tabela 7: Divisão do conjunto de imagens IM\_1

| Finalidade: | Quantidade: | Percentual: |
|-------------|-------------|-------------|
| Treinamento | 2.821       | 70%         |
| Validação   | 808         | 20%         |
| Testes      | 400         | 10%         |

Fonte: Autor.

### 3.2.2 CONJUNTO DE IMAGENS 2 (IM\_2)

O conjunto IM\_2 é composto por imagens adquiridas em 3 momentos distintos (sessões). Na primeira sessão, a aquisição das imagens foi realizada pelo autor deste trabalho na Universidade Presbiteriana Mackenzie com equipamento e apoio da equipe técnica do Laboratório de Geotecnologias da instituição. A segunda sessão ocorreu em um canteiro de obras localizado no



município de Itapetininga, e contou com o auxílio de alunos de mestrado do PPGI da Universidade Nove de Julho. Por fim, a terceira sessão foi realizada de forma independente pelo autor deste trabalho. Os detalhes das imagens do conjunto IM\_2, como quantidade, resolução, tipo e formato dos arquivos estão informados na Tabela 8.

Tabela 8: Composição do conjunto de imagens IM\_2

| Sessão: | Drone Utilizado:  | Qtde.: | Resolução (px): | Formato: |
|---------|-------------------|--------|-----------------|----------|
| 1       | DJI Phantom 4 Pro | 17     | 5.472x3.648     | RGB/JPG  |
| 2       | DJI Mavic 2 Pro   | 36     | 5.472x3.648     | RGB/JPG  |
| 3       | DJI Mini 3        | 16     | 1.600x900       | RGB/JPG  |

Fonte: Autor.

Esse segundo conjunto de imagens (IM\_2), foi utilizado para testar o método proposto com imagens de fissuras obtidas por drones em ambientes reais de edificações e obras civis.

A Figuras 14 apresenta exemplos das imagens que compõem o conjunto IM\_2, adquiridas na sessão 1.

Figura 14: Exemplos de imagens capturadas durante a sessão 1



Fonte: Autor.

A sessão 1 de captura de imagens foi realizada nas dependências do Campus Higienópolis da Universidade Presbiteriana Mackenzie. Próximo as fissuras fotografadas foram inseridas marcações visuais com graduação de 1 cm visando facilitar a comparação com a medição realizada pela segunda etapa do método proposto, que é baseada no cômputo do GSD.

A segunda sessão de fotos foi realizada em um canteiro de obras na cidade de Itapetininga, tendo como alvo uma estrutura de contenção do tipo Solo

Grampeado. A Figura 15 apresenta exemplos das imagens capturadas nessa sessão.

Figura 15: Exemplos de imagens capturadas durante a sessão 2



Fonte: Autor.

A terceira sessão de fotos foi realizada uma construção anexa à residência do autor deste trabalho. A Figura 16 apresenta exemplos de imagens capturadas durante nessa seção.

Figura 16: Exemplos de imagens capturadas durante a sessão 3



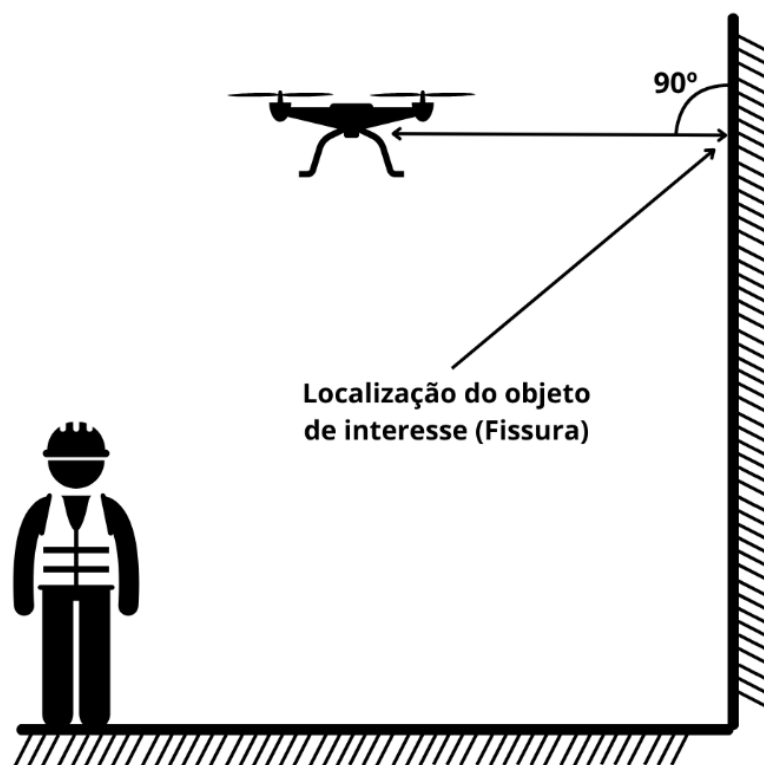
Fonte: Autor.

### 3.2.3 PROCEDIMENTO PARA CAPTURA DAS IMAGENS

A captura das imagens seguiu rigorosamente as normas e procedimentos da ANAC para operação segura. Para assegurar condições adequadas à análise por VC, priorizou-se a obtenção das fotografias em ambientes bem iluminados por luz natural, garantindo maior contraste entre a superfície da alvenaria e a fissura. O drone foi aproximado o máximo possível da região de interesse, mantendo sempre condições seguras de voo, com distância entre a câmera e a fissura variando entre 0,6 m e 5 m. Durante os registros, a câmera foi mantida perpendicular ao plano da parede e a aeronave estabilizada antes do disparo,

como ilustrado na Figura 17, a fim de minimizar distorções, desfoques e artefatos que pudessem comprometer a detecção e mensuração das fissuras.

Figura 17: Processo de captura das imagens com o uso de drones



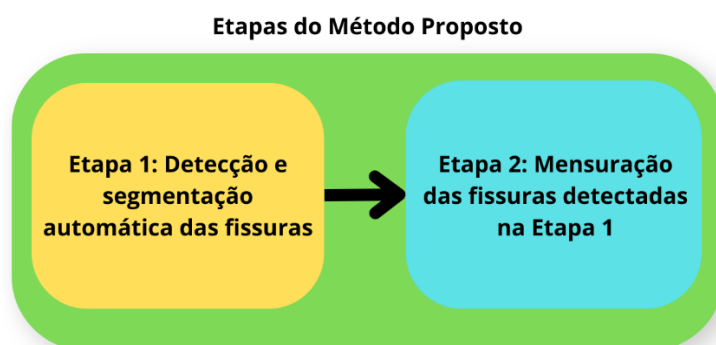
Fonte: Autor.

### 3.3 MÉTODO PROPOSTO

O método computacional proposto neste trabalho contempla duas etapas, como ilustrado na Figura 18. A primeira utiliza um modelo de RNC do framework YOLO, denominado YOLOv8, para detecção e segmentação automática de fissuras, e que foi treinado e validado com o conjunto de imagens IM\_1. A segunda etapa emprega técnicas de VC e PID, como HOG, esqueletização e transformada da distância para extrair e mensurar características geométricas da fissura detectada, como comprimento do eixo principal e ângulo de propagação predominante. A implementação dos algoritmos que compõem essas etapas foi realizada utilizando linguagem Python 3, no ambiente Google Colab<sup>3</sup>.

<sup>3</sup> O Google Colab é uma ferramenta web para estudo e pesquisa de projetos de aprendizagem de máquina e análise de dados (<https://colab.research.google.com/>)

Figura 18: Etapas do Método Proposto

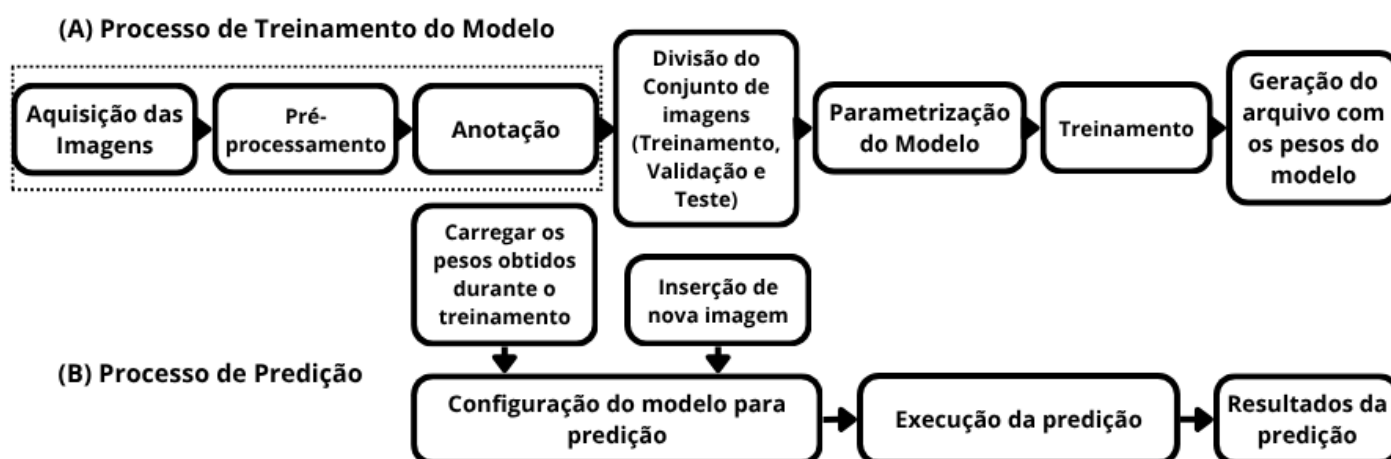


Fonte: Autor.

### 3.3.1 ETAPA 1: DETECÇÃO DE FISSURAS

Na primeira etapa do método, as imagens do conjunto IM\_1, com as fissuras devidamente demarcadas (anotações), foram utilizadas como conjunto de treinamento do modelo YOLOv8. Assim, o modelo aprendeu padrões visuais associados à presença de fissuras, de forma a identificar automaticamente instâncias semelhantes em novas imagens. O diagrama dos processos de treinamento e predição do modelo de detecção automática de fissuras está representado na Figura 19.

Figura 19: Processos de treinamento e predição do modelo



Fonte: Autor.

A região pontilhada da Figura 19 representa os passos envolvidos no treinamento do modelo que foram previamente realizados pela Ultralytics<sup>4</sup> na plataforma Robofow, pois o conjunto de imagens IM\_1 é disponibilizado pré-

<sup>4</sup> A Ultralytics é uma empresa de IA e VC, que atua no desenvolvimento de modelos de código aberto YOLO (<https://www.ultralytics.com/>)



processado e anotado, a Figura 20 apresenta alguns exemplos de imagens anotadas do conjunto de imagens IM\_1.

Figura 20: Exemplos de imagens anotadas do conjunto IM\_1

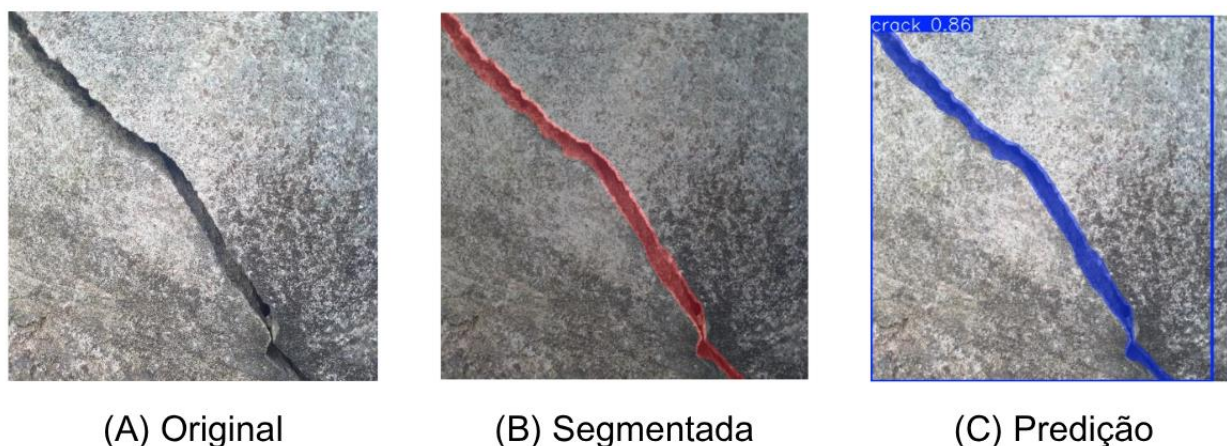


Fonte: Autor.

Quanto à RNC YOLO utilizada como algoritmo principal do modelo, os experimentos foram conduzidos com a versão Ultralytics 8.3.205, empregando o modelo YOLOv8s (Small), que possui 11,2 milhões de parâmetros. A escolha dessa versão teve como objetivo um melhor equilíbrio entre performance e exigência de poder computacional, permitindo inclusive implementação do modelo em sistemas de predição local, sem a necessidade de conexão com ambientes de computação em nuvem. Além disso, foi utilizada a variação de modelo especializado YOLOv8s-seg, que têm como objetivo segmentação de instâncias, ou seja, faz detecção e segmentação com contorno do objeto alvo da detecção (ULTRALYTICS, 2023a).

A segmentação da fissura detectada nesta primeira etapa é utilizada na segunda etapa do método, responsável pela mensuração de sua geometria. A Figura 21 apresenta uma imagem de fissura original (A), segmentada (B) e com a caixa de detecção (*bounding box*) na saída do processo de predição (C).

Figura 21: Segmentação e detecção de uma fissura



Fonte: Autor.

O mesmo modelo treinado pode tanto realizar predições em imagens individualizadas quanto em sequências de imagens (ou frames), ou seja, a mesma técnica pode ser utilizada para detecção automática em transmissões ao vivo ou vídeos gravados previamente, como demonstrado nos experimentos descritos na seção 4.1.2.

O modelo de detecção de fissuras foi treinado no ambiente de desenvolvimento Google Colab, utilizando Unidades de Processamento Gráfico ou *Graphics Processing Unit* (GPU) disponíveis na plataforma (T4, L4 e A100, detalhados na Tabela 9), de acordo com a disponibilidade das mesmas no momento de execução dos experimentos.

Tabela 9: Características das GPUs utilizadas nos treinamentos e predições com YOLO

| Característica                 | NVIDIA T4                                          | NVIDIA L4                                              | NVIDIA A100                                           |
|--------------------------------|----------------------------------------------------|--------------------------------------------------------|-------------------------------------------------------|
| <b>Finalidade</b>              | Uso geral, aprendizado básico                      | Inferência de IA, modelos complexos                    | Treinamento de IA em larga escala, HPC                |
| <b>Performance</b>             | Moderada (base de comparação)                      | Aprox. 2,2x mais rápida que T4                         | Aprox. 6x a 13x mais rápida que T4                    |
| <b>VRAM (Memória de Vídeo)</b> | ~15 GB                                             | ~22 GB                                                 | 40 GB (ou 80 GB em versões específicas)               |
| <b>Custo/Benefício</b>         | Excelente para prototipagem e orçamentos limitados | Bom equilíbrio entre custo e performance superior à T4 | Alto desempenho com custo mais elevado                |
| <b>Disponibilidade</b>         | Disponível no plano gratuito e pagos               | Disponível em planos pagos (Pro/Pro+)                  | Disponível principalmente no plano Pro+ (prioritário) |
| <b>Arquitetura</b>             | Turing                                             | Ada Lovelace                                           | Ampere                                                |

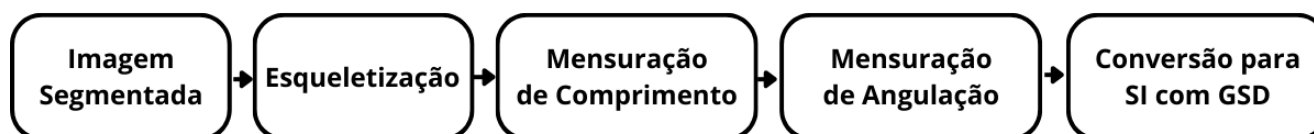
Fonte: Autor.

### 3.3.2 ETAPA 2: MENSURAÇÃO DE FISSURAS

A segunda etapa do método, cujo *pipeline* de funcionamento é ilustrado na Figura 22, tem como objetivo a mensuração do comprimento e angulação predominante das fissuras.

Figura 22: Fluxo geral de execução da Etapa 2

#### Fluxo de trabalho (*pipeline*) da Etapa 2



Fonte: Autor.

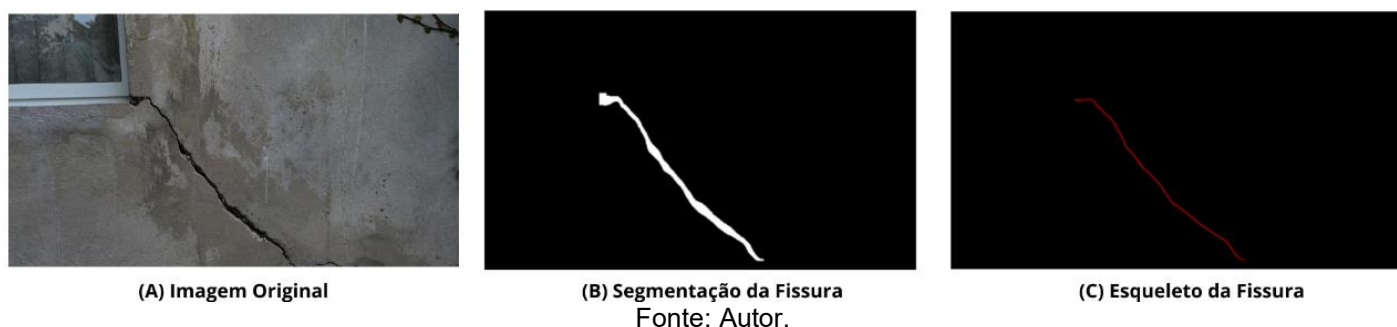
Esta etapa tem como entrada a imagem com as fissuras segmentadas pela etapa anterior, ou seja, a área de interesse já está em destaque, o que

dispensa pré-processamentos ou segmentação utilizando técnicas clássicas de VC e PID.

#### 3.3.2.1 ESQUELETIZAÇÃO

Na região segmentada é aplicado o processo de esqueletização, por meio da transformada da distância, ilustrado na Figura 23, para reduzir a fissura a uma linha de um pixel de espessura, mantendo seu formato e extensão. Operações de morfologia matemática podem ser utilizadas neste passo para eliminar ramificações espúrias.

Figura 23: Esqueletização de fissuras



#### 3.3.2.2 MENSURAÇÃO DE COMPRIMENTO

Foram realizados experimentos com a mensuração do comprimento da extensão da fissura sendo obtida por meio de duas abordagens:

- I. Medição do comprimento a partir de um segmento de reta interligando os pontos de extremidade do esqueleto da fissura
- II. Medição do comprimento a partir da análise dos pixels que compõem o esqueleto da fissura.

Na primeira abordagem (I), algoritmos de varredura percorrem os pixels da fissura, calculando distâncias entre eles através da técnica da distância Euclidiana (BOLAYBOLAY et al., 2022). Para possibilitar o cálculo do comprimento principal da fissura detectada é traçado um segmento de reta unindo os pixels das extremidades mais distantes da fissura, do resultado da esqueletização, de onde são extraídas as coordenadas de todos os pixels que compõem a fissura esqueletizada. A partir dessas coordenadas, calcula-se a maior distância euclidiana possível entre dois pontos pertencentes à fissura, definindo assim as duas extremidades  $A$  e  $B$ , sendo a posição de cada ponto representada por um par de coordenadas, como segue:  $A(x_A, y_A)$  e  $B(x_B, y_B)$ . A partir desses dois pontos extremos, traça-se um segmento de reta que

representa o comprimento principal da fissura (em pixels). A Equação 19 apresenta a expressão matemática utilizada para calcular esse comprimento, correspondente à extensão principal da fissura.

$$D(A, B) = \sqrt{(x_A - x_B)^2 + (y_A - y_B)^2} \quad (19)$$

Já na segunda abordagem (II), estima-se o comprimento da fissura contando-se o número total de pixels pertencentes ao esqueleto, por meio da Equação 20.

$$L_{\text{px, simples}} = \sum_{i,j} S(i, j) \quad (20)$$

Na qual  $S$  é o conjunto de pixels do esqueleto, com  $S(i, j) = 1$  se o pixel  $(i, j)$  pertence ao esqueleto, e 0 caso contrário.

Nesse caso, cada pixel do esqueleto contribui com uma unidade de comprimento. Essa abordagem é computacionalmente simples e direta, e funciona bem quando o esqueleto é predominantemente formado por segmentos horizontais e verticais. No entanto, em trechos diagonais, a contagem pura de pixels tende a subestimar o comprimento real, uma vez que um deslocamento de um pixel na diagonal representa, geometricamente, uma distância maior que um deslocamento puramente horizontal ou vertical.

Assim, para obter uma estimativa mais fiel ao comprimento da fissura, foi adotada uma sofisticação dessa segunda abordagem baseada na soma das arestas entre pixels vizinhos conectados no esqueleto, distinguindo deslocamentos ortogonais e diagonais. Dessa forma, considera-se o esqueleto como um grafo em que cada pixel é um nó, e que há uma aresta entre dois nós se eles são vizinhos 4 ou 8 conexos (acima, abaixo, esquerda, direita e diagonais).

O comprimento total é calculado percorrendo o esqueleto e somando:

- 1 unidade de comprimento para cada par de pixels vizinhos em 4-conexos (horizontal ou vertical);
- $\sqrt{2}$  unidades de comprimento para cada par de pixels vizinhos em 8-conexos diagonal.

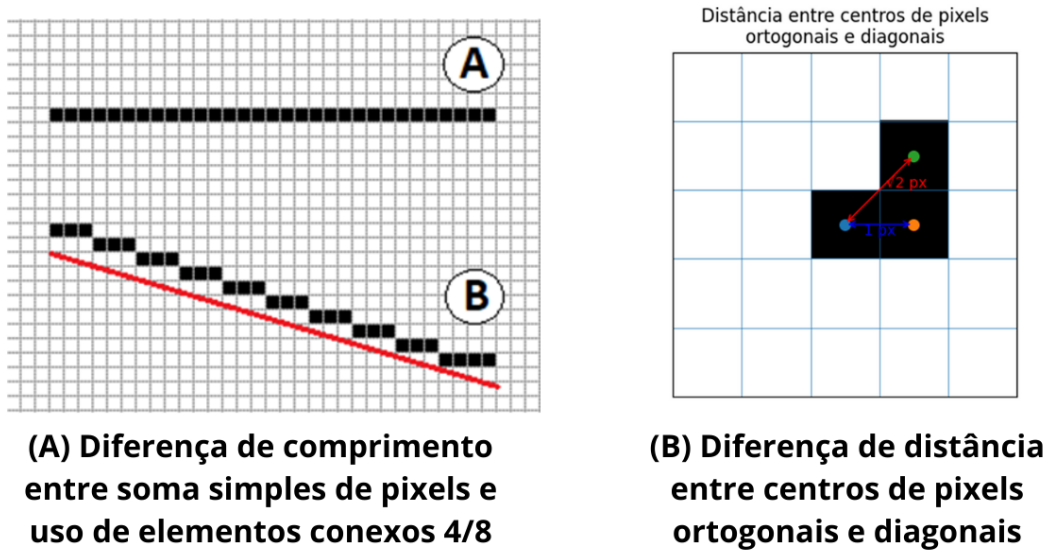


Seja  $E_{\text{ort}}$  o conjunto de pares de pixels adjacentes ortogonais (4-conexos) e  $E_{\text{diag}}$  o conjunto de pares adjacentes diagonais, o comprimento aproximado em pixels é dado pela Equação 21.

$$L_{\text{px, conexo}} = \sum_{(p,q) \in E_{\text{ort}}} 1 + \sum_{(p,q) \in E_{\text{diag}}} \sqrt{2} \quad (21)$$

Na implementação, essa soma é realizada de forma vetorizada, criando imagens deslocadas em direções pré-definidas (direita, baixo, diagonal baixo-direita e diagonal baixo-esquerda) e contando quantas vezes um pixel do esqueleto possui um vizinho também pertencente ao esqueleto naquela direção. Dessa forma, evita-se contagem duplicada de arestas e obtém-se um valor de comprimento que considera explicitamente a geometria (horizontal, vertical ou diagonal) do traçado. A Figura 24 ilustra a diferença de comprimento obtida entre soma simples de pixels (Equação 20) e a abordagem considerando os 4 ou 8 elementos conexos e o uso das diagonais ( $\sqrt{2}$ ) para o cálculo do comprimento (Equação 21).

Figura 24: Diferença de comprimento obtidos com soma simples e elementos 4 ou 8 conexo



Fonte: Autor.

Essa abordagem por elementos conexos apresenta duas vantagens principais:

- Fornece valores mais próximos do comprimento euclidiano da fissura, especialmente quando há trechos significativos em diagonal;
- Mantém baixo custo computacional, pois opera sobre vizinhanças locais e utiliza operações matriciais simples.

Em ambos os casos, quando conhecido o tamanho do pixel (por exemplo, a partir do GSD da imagem), o comprimento em pixels pode ser convertido para uma unidade física (metro, centímetros, milímetros, etc.), multiplicando-se o valor obtido pelo fator de escala correspondente.

### 3.3.2.3 MENSURAÇÃO DE ANGULAÇÃO

De forma semelhante à medição de comprimento, também foram consideradas duas abordagens distintas para a mensuração do ângulo predominante da fissura detectada, quais sejam:

- I. Obtenção do ângulo predominante a partir do segmento de reta traçado entre os pontos das extremidades da fissura.
- II. Obtenção do ângulo predominante da fissura por meio do descritor HOG.

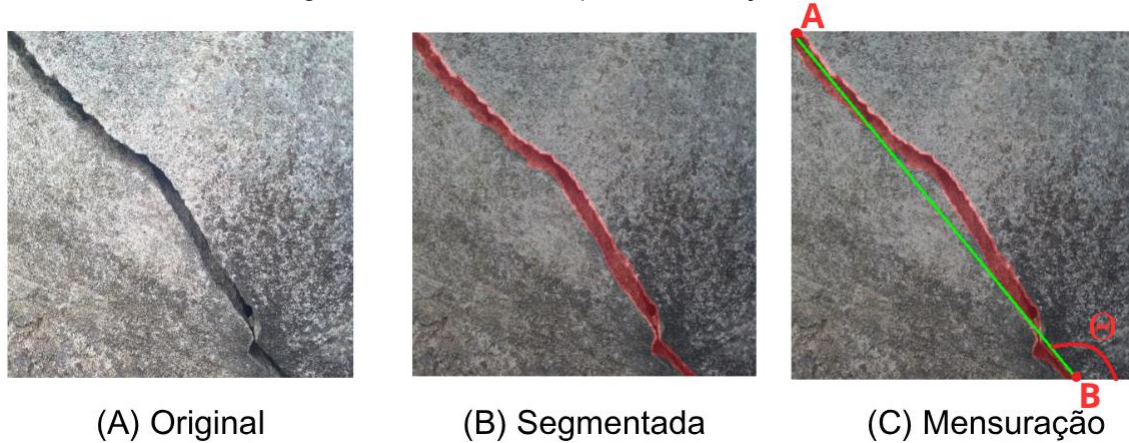
Utilizando a primeira abordagem (I), o cálculo da angulação predominante da fissura detectada consiste em determinar o ângulo formado pelo segmento de reta traçado entre as extremidades da fissura, em relação a um eixo ou plano de referência, conforme a Equação 22.

$$\theta = \tan^{-1}\left(\frac{y_B - y_A}{x_B - x_A}\right) \quad (22)$$

O ângulo ( $\theta$ ) é calculado em relação ao eixo horizontal da imagem, e  $(x_A, y_A)$  e  $(x_B, y_B)$  representam as coordenadas dos pontos que determinam as extremidades do segmento de reta traçado.

O ângulo predominante e o comprimento principal permitem a caracterização geométrica da fissura, contribuindo para uma análise mais detalhada. A Figura 25 ilustra os passos da etapa de mensuração de comprimento e ângulo principal de propagação de uma fissura: imagem original (A), imagem segmentada (B) e (C) imagem com os pontos de extremidades definidos, o segmento de reta traçado e o ângulo ( $\theta$ ) destacado.

Figura 25: Procedimento para mensuração de fissuras



Fonte: Autor.

Este procedimento permite quantificar, de forma simples e rápida, a extensão e orientação da fissura, sendo um parâmetro relevante para análise estrutural e diagnóstico de manifestações patológicas (BOLAYBOLAY et al., 2022; WANG et al., 2024).

A segunda abordagem (II) para determinação do ângulo predominante de propagação da fissura utiliza o descritor HOG. A técnica baseia-se na extração dos vetores de gradiente de intensidade de cada pixel da região segmentada, permitindo quantificar as orientações locais das variações de brilho que caracterizam as bordas da fissura. A partir desses gradientes, é gerado um histograma de orientações no intervalo de  $0^\circ$  a  $180^\circ$ , no qual cada barra ou *bin* representa a frequência relativa de direções predominantes dos gradientes.

O HOG é aplicado sobre fissura segmentada, obtida pelo modelo YOLOv8. A partir da qual extrai-se o campo de gradientes de intensidade, e, em seguida, constroi-se o histograma de orientações dos gradientes, ponderado pela magnitude  $M(x,y)$ . O pico do histograma, correspondente ao ângulo de maior frequência ponderada, indica a orientação predominante da fissura.

Entretanto, é importante ressaltar que o gradiente de intensidade é, por definição, perpendicular à direção da borda, logo, os vetores do HOG representam a orientação normal à fissura e não sua direção tangente de propagação. Assim, para converter a orientação dominante do HOG ( $\theta_H$ ) em orientação tangente da fissura ( $\theta_T$ ), aplica-se uma rotação de  $90^\circ$  no espaço angular conforme a Equação 23.

$$\theta_T = (\theta_H + 90^\circ) \bmod 180^\circ \quad (23)$$

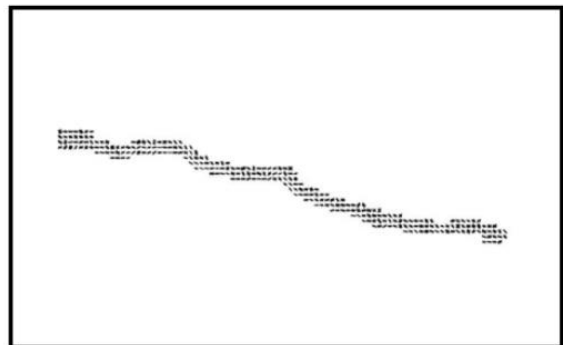
Essa transformação assegura que o ângulo resultante esteja expresso na convenção geométrica ( $0^\circ$  a  $180^\circ$ , sentido anti-horário), permitindo a comparação com a outra abordagem de estimação angular (segmento entre extremidades).

O uso do HOG nesse contexto apresenta características interessantes: o método é robusto a ruídos pontuais, já que o histograma agrega a informação direcional de uma região inteira; a representação angular é independente da posição e escala da fissura; o descritor fornece uma medida estatística global da orientação predominante, útil para complementar métodos puramente geométricos. Assim, a orientação principal de propagação derivada do HOG foi utilizada como parâmetro de comparação e validação das orientações obtidas pelo segmento de reta de extremidades, permitindo uma interpretação mais completa da direção e padrão de propagação das fissuras detectadas. A Figura 26 exemplifica uma fissura segmentada e uma representação visual de orientação de gradientes da segmentação.

Figura 26: Fissura segmentada e representação dos gradientes orientados



**(A) Fissura com a máscara de segmentação**

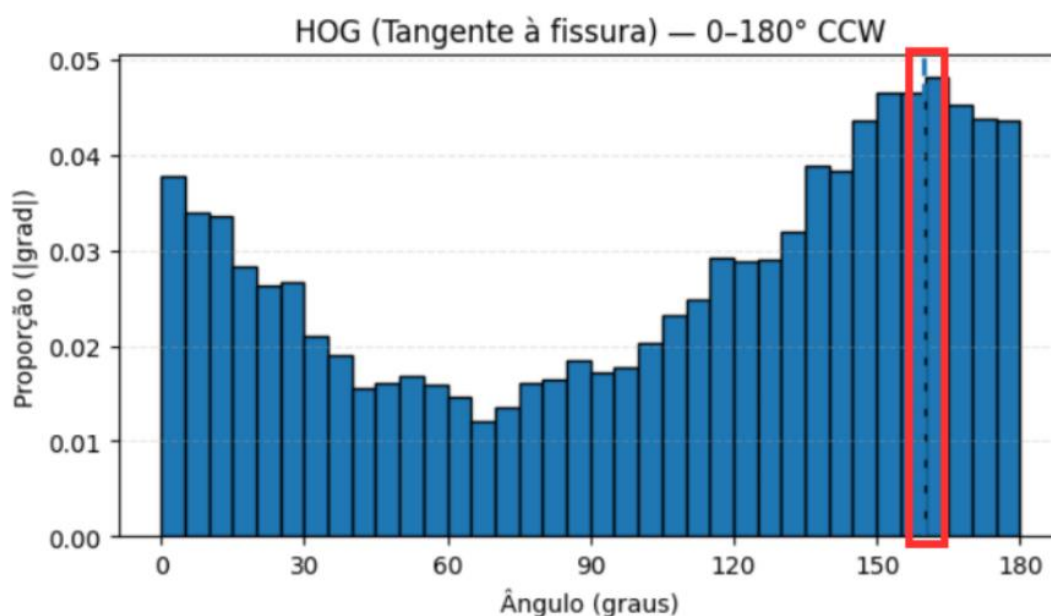


**(B) Representação visual do HOG da segmentação**

Fonte: Autor.

A Figura 27 apresenta um histograma das angulações, já normalizadas, dos gradientes extraídos por meio do HOG, a partir do qual é possível determinar a angulação predominante da fissura.

Figura 27: Histograma de angulações HOG



Fonte: Autor.

### 3.3.2.4 CONVERSÃO PARA SI COM GSD

Uma vez obtido o comprimento em pixels da fissura, é possível utilizar o conceito de GSD para estimar as características geométricas reais da fissura em unidades do SI. Para cada experimento realizado, foram definidos manualmente três parâmetros fundamentais:

- I. A distância entre a câmera e o plano da superfície onde se situava a fissura;
- II. A largura física do sensor da câmera, expressa em milímetros;
- III. A distância focal real da câmera, também expressa em milímetros.

Esses valores permitem estimativas precisas em função das condições reais de captura. O resultado do GSD é inicialmente calculado em metros por pixel (m/px), sendo posteriormente convertido para milímetros por pixel (mm/px) para facilitar a análise dos comprimentos das fissuras. Assim, uma vez determinado o GSD, o comprimento real  $L_{real}$  de cada fissura é obtido multiplicando-se o comprimento medido em pixels ( $L_{px}$ ) pelo GSD, conforme apresentado na Equação 24.

$$L_{real} = L_{px} \cdot GSD \quad (24)$$

A Tabela 10 apresenta as especificações técnicas das câmeras RGB dos 3 drones utilizados para captura das imagens do conjunto IM\_2. Nela estão informados os dados de largura e altura do sensor, distância focal e campo e visão (FOV).

Tabela 10: Detalhes técnicos das câmeras RGB dos drones utilizados

| Sessão: | Drone Utilizado:        | Resolução (px): | Largura (W): | Altura (H): | Distância Focal (f) | Campo de Visão (FOV): |
|---------|-------------------------|-----------------|--------------|-------------|---------------------|-----------------------|
| 1       | DJI Phantom 4 Pro       | 5.472x3.648     | 13,2mm       | 8mm         | 8,8mm               | 84°                   |
| 2       | DJI Mavic 2 Pro         | 5.472x3.648     | 12,8mm       | 9,6mm       | 10,35mm             | 77°                   |
| 3       | DJI Mini 3 <sup>5</sup> | 1.600x900       | 9,7mm        | 7,3mm       | 6,72mm              | 82,1°                 |

Fonte: Autor.

Após o detalhamento dos processos e algoritmos que compõem o método, e dos conjuntos de imagens utilizadas nos experimentos, procede-se à avaliação e discussão dos resultados obtidos, conforme apresentado na seção seguinte.

## 4 RESULTADOS E DISCUSSÃO

Este capítulo apresenta os resultados e as discussões resultantes dos experimentos computacionais com o método proposto. Nas Seções 4.1 e 4.2 são apresentados, respectivamente, os resultados referentes à detecção e à mensuração, enquanto a Seção 4.3 apresenta as discussões.

### 4.1 ETAPA 1: DETECAÇÃO DE FISSURAS

Foram realizados diversos ciclos de treinamentos do modelo utilizando o conjunto de imagens IM\_1. Ao longo dos ciclos foram sendo implementadas alterações nos modelos de GPU disponíveis na plataforma do Google Colab e ajustes nas configurações e hiperparâmetros, objetivando melhorar a performance do modelo treinado. Como referência de performance, foram consideradas as métricas disponibilizadas pelo framework YOLO, que são divididas em duas categorias: *Box*, detecção dos objetos-alvo (fissura presente

<sup>5</sup> Para o drone DJI Mini 3, não foram encontrados valores oficiais explícitos de W, H e f.

na imagem), assinaladas por caixas delimitadoras (*bounding boxes*), e *Mask*, que se refere à segmentação da forma da fissura por meio de máscaras.

A Tabela 11 apresenta as configurações, os hiperparâmetros e os valores das métricas correspondentes aos melhores resultados obtidos nos ciclos de treinamento, considerando-se as 808 imagens do conjunto IM\_1 destinadas à validação, conforme Tabela 7. Devido ao longo tempo de cada ciclo de treinamento, adotou-se três valores fixos de épocas: 100, 300 e 500.

Tabela 11: Resultados da etapa detecção de fissuras conjunto de validação

|                                  | Modelo YOLOv8:                 | yolov8s-seg   |                 |               |
|----------------------------------|--------------------------------|---------------|-----------------|---------------|
|                                  | GPU:                           | T4            | L4              | L4/A100       |
|                                  | Tempo de Treinamento:          | ≈2:30h        | ≈4:30h          | ≈6:00h        |
| Ajuste de hiperparâmetros        | Num. Épocas ( <i>epochs</i> ): | 100           | 300             | 500           |
|                                  | Tam. imgs ( <i>imgsz</i> ):    | 640 pixels    | 640 pixels      | 768 pixels    |
|                                  | Batch size ( <i>batch</i> ):   | 16            | -1 <sup>6</sup> | -1            |
|                                  | Tx. Ap. In. ( <i>lr0</i> ):    | 0.001         | 0.001           | 0.001         |
|                                  | Fat. Dec. ( <i>lrf</i> ):      | 0.1           | 0.1             | 0.01          |
|                                  | Reg. ( <i>weight_decay</i> ):  | 0.0005        | 0.0005          | 0.0005        |
| Resultados de Detecção (BOX)     | Precisão ( <i>P</i> ):         | 0,819 (81,9%) | 0,800 (80,0%)   | 0,871 (87,1%) |
|                                  | Recall ( <i>R</i> ):           | 0,727 (72,7%) | 0,787 (78,7%)   | 0,775 (77,5%) |
|                                  | mAP@0,50:                      | 0,784 (78,4%) | 0,777 (77,7%)   | 0,820 (82,0%) |
| Resultados de Segmentação (MASK) | Precisão ( <i>P</i> ):         | 0,729 (72,9%) | 0,772 (77,2%)   | 0,741 (74,1%) |
|                                  | Recall ( <i>R</i> ):           | 0,622 (62,2%) | 0,679 (67,9%)   | 0,703 (70,3%) |
|                                  | mAP@0,50:                      | 0,641 (64,1%) | 0,669 (66,9%)   | 0,659 (65,9%) |

Fonte: Autor.

A Tabela 12, por sua vez, apresenta os valores das métricas obtidas para o conjunto de imagens IM\_1, considerando a porção reservada para teste, composta por 400 imagens, o que permite avaliar a capacidade real de generalização do modelo treinado.

<sup>6</sup> Com o intuito de otimizar o tempo de treinamento para número de épocas maiores foi utilizado o -1 na configuração de *batch size* para treinamento do modelo, dessa forma o sistema de maneira dinâmica tenta alocar o maior *batch* possível na memória RAM da GPU

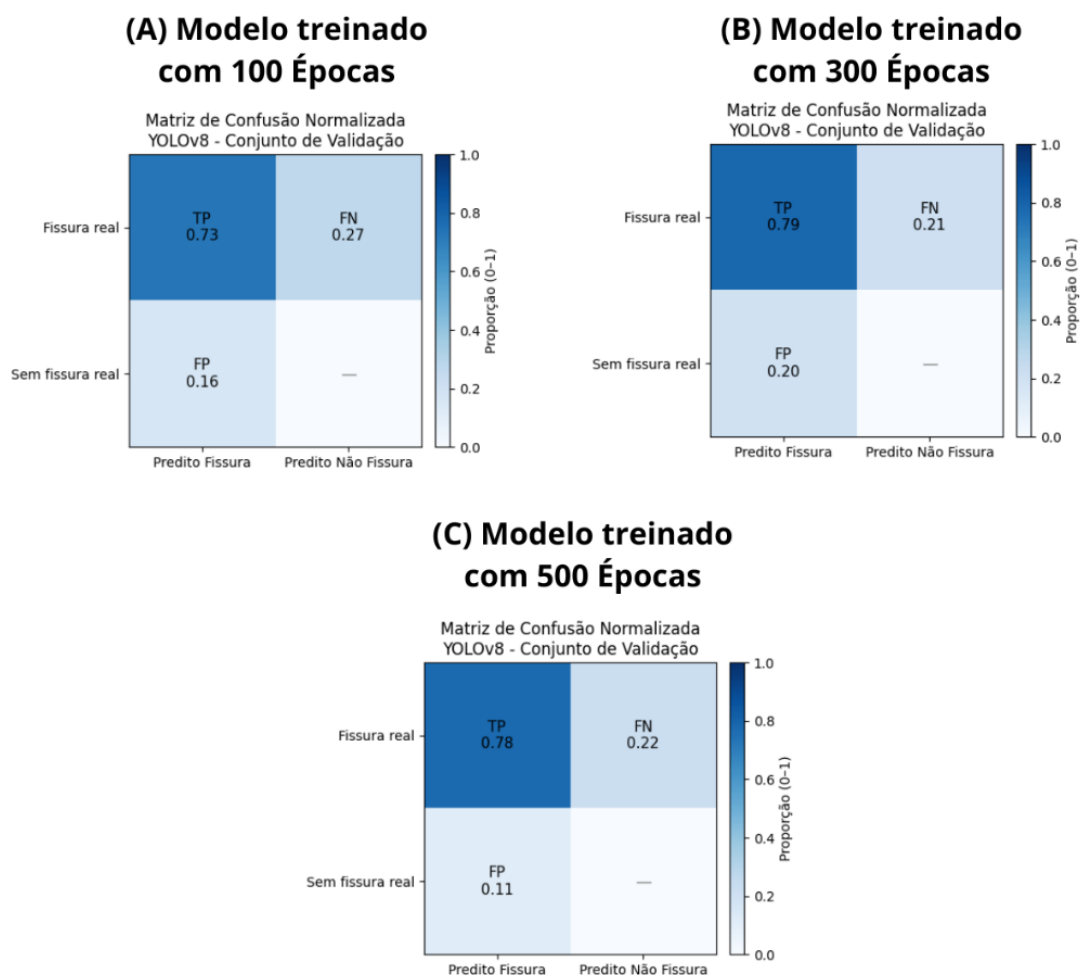
Tabela 12: Resultados da etapa detecção de fissuras conjunto de teste

|                                  | Modelo YOLOv8:    | yolov8s-seg   |               |               |
|----------------------------------|-------------------|---------------|---------------|---------------|
|                                  | GPU:              | T4            | T4            | T4            |
|                                  | Duração do Teste: | ≈15s          | ≈15s          | ≈15s          |
| Resultados de Detecção (BOX)     | Precisão (P):     | 0,771 (77,1%) | 0,811 (81,1%) | 0,794 (79,4%) |
|                                  | Recall (R):       | 0,696 (69,6%) | 0,697 (69,7%) | 0,649 (64,9%) |
|                                  | mAP@0,50:         | 0,756 (75,6%) | 0,711 (71,1%) | 0,704 (70,4%) |
| Resultados de Segmentação (MASK) | Precisão (P):     | 0,710 (71,0%) | 0,711 (71,1%) | 0,752 (75,2%) |
|                                  | Recall (R):       | 0,596 (59,6%) | 0,614 (61,4%) | 0,616 (61,6%) |
|                                  | mAP@0,50:         | 0,593 (59,3%) | 0,550 (55,0%) | 0,599 (59,9%) |

Fonte: Autor.

A Figura 28 apresenta as matrizes de confusão normalizadas obtidas na etapa de validação do modelo durante o treinamento.

Figura 28: Matrizes de confusão da validação

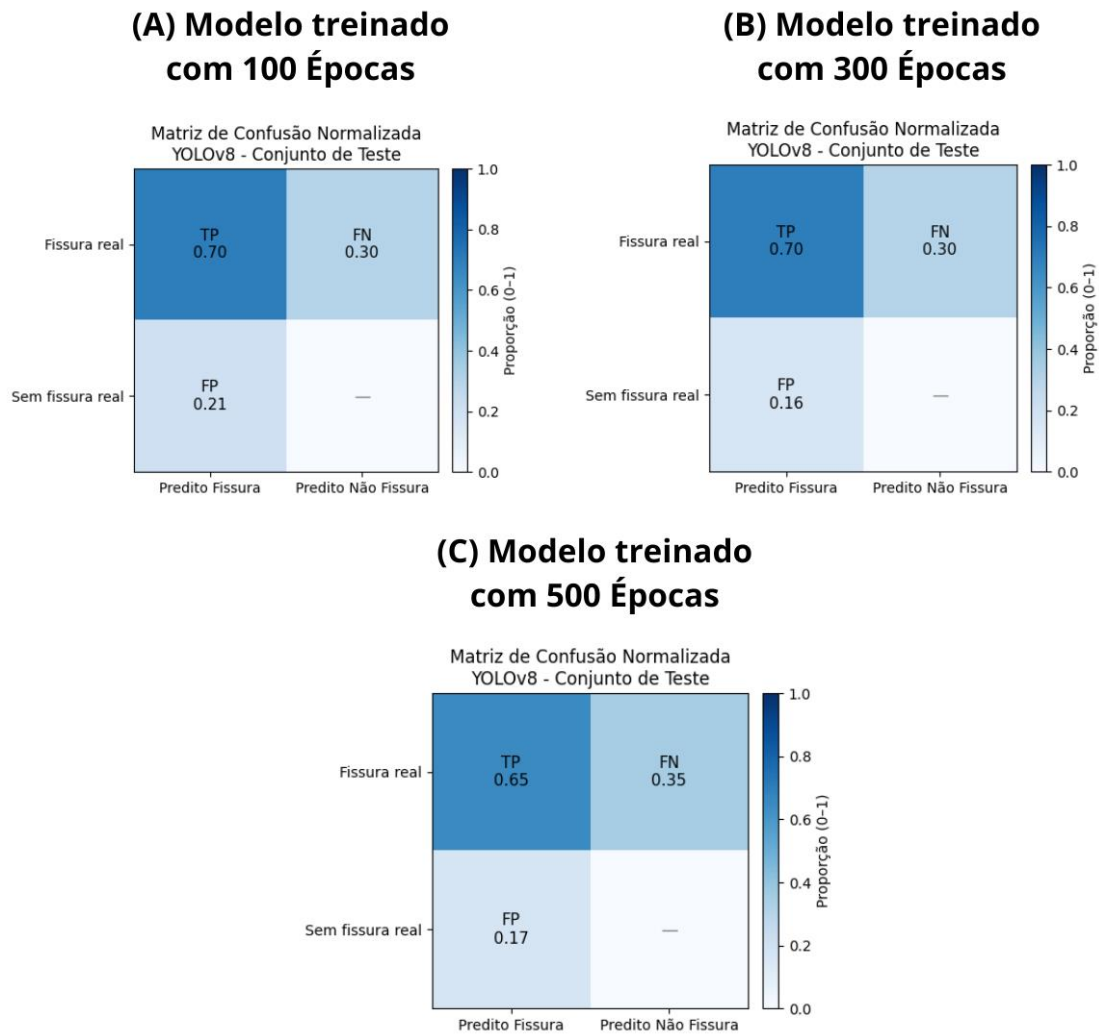


Fonte: Autor.



A Figura 29 apresenta as matrizes de confusão normalizadas obtidas nos testes com o modelo treinado.

Figura 29: Matrizes de confusão do teste



Fonte: Autor.

Nas matrizes de confusão, o valor de TN não é aplicável. Isso ocorre porque o treinamento foi realizado com apenas uma classe de interesse, e todas as imagens de IM\_1 contêm ao menos uma fissura rotulada, inexistindo amostras negativas que permitam quantificar verdadeiros negativos. Considerar o fundo da imagem como classe negativa geraria forte desequilíbrio, pois a grande maioria dos pixels seriam classificados como TN, também tornando essa medida irrelevante para a avaliação. Assim, o desempenho do modelo é analisado apenas quanto à capacidade de localizar e segmentar fissuras. Além

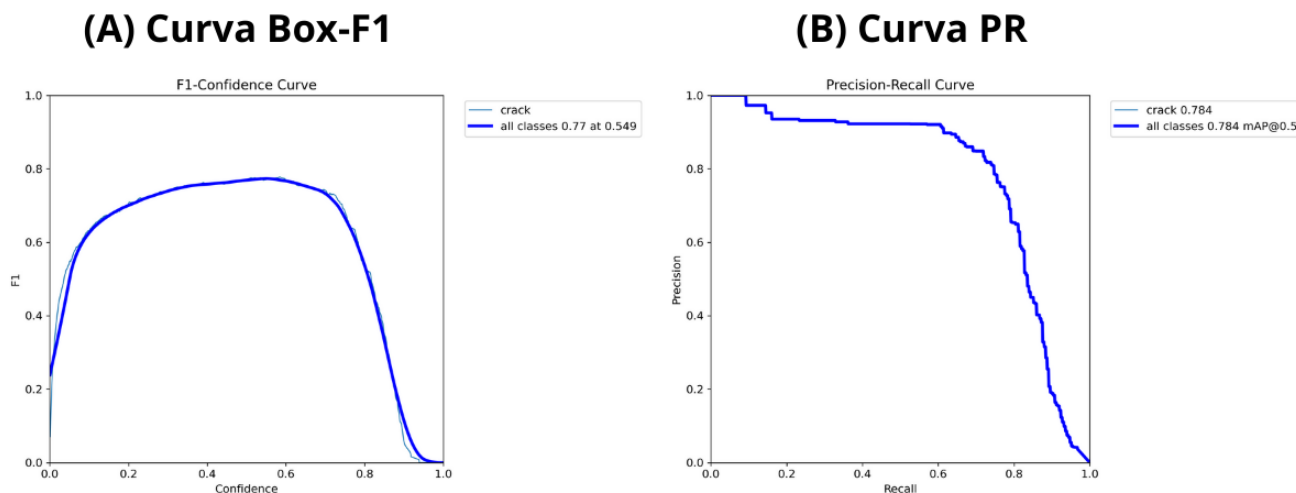
disso, as matrizes foram normalizadas por classe, de modo que apenas a linha correspondente à classe “fissura” possui soma igual a 1.

A apresentação de 2 curvas geradas a partir das métricas obtidas durante os treinamentos e testes podem auxiliar em uma análise mais generalista dos resultados, são elas:

- Curva Box-F1, que representa a variação do F1-score em função do limiar de confiança (*confidence threshold*) durante a avaliação das caixas delimitadoras (*bounding boxes*) que representam as predições do modelo. O ponto de máximo da curva indica o melhor equilíbrio entre precisão e recall, ou seja, define o limiar de confiança ideal para as predições.
- Curva PR: Plota a Precisão (*Precision*) versus o *Recall*. Para tarefas de detecção de objetos, onde muitas vezes há um grande desequilíbrio de classes (muito mais fundo do que objetos de interesse), a curva PR oferece uma visão mais realista e informativa do desempenho do modelo, focando na performance da classe minoritária (os objetos que queremos detectar).

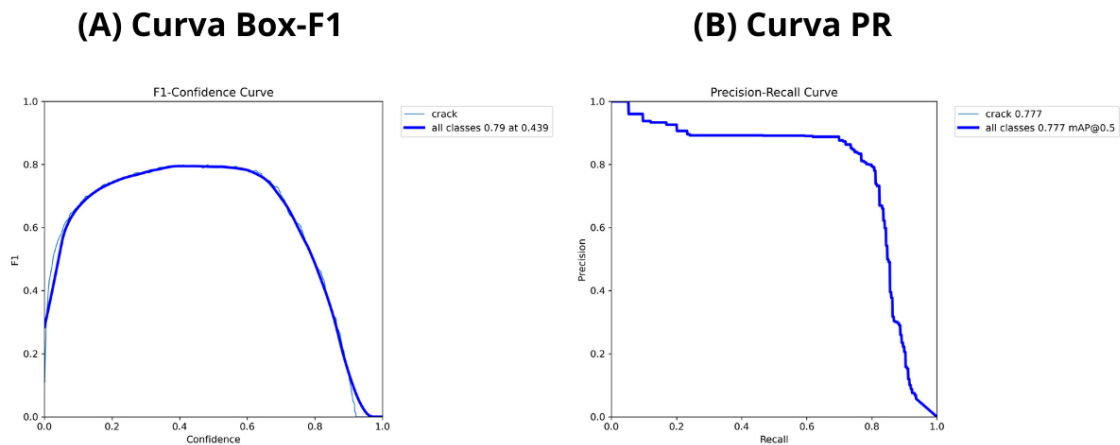
As imagens apresentadas nas figuras 30 a 32 apresentam as duas curvas para cada um dos modelos treinados, considerando os três valores adotados para o número de épocas.

Figura 30: Curvas Box-F1 e RP do Modelo treinado com 100 épocas



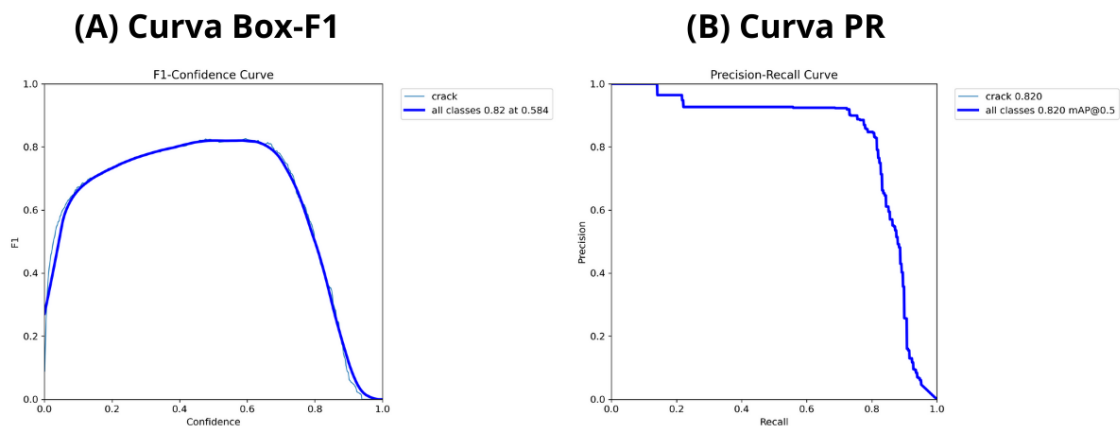
Fonte: Autor.

Figura 31: Curvas Box-F1 e RP do Modelo treinado com 300 épocas



Fonte: Autor.

Figura 32: Curvas Box-F1 e RP do Modelo treinado com 500 épocas



Fonte: Autor.

A etapa 1 do modelo proposto, responsável pela detecção de fissuras nas imagens, alcançou valores de precisão ( $P$ ) em torno de 80% na métrica de detecção por caixas delimitadoras (*bounding boxes*), tanto na validação quanto no teste. Entretanto, os valores de *recall* ( $R$ ) e mAP@0,5 foram inferiores na fase de teste (entre 64% e 69%) em comparação à validação. Esses resultados indicam que o modelo apresenta desempenho promissor, embora ainda exista margem para aprimorar sua capacidade de generalização.

Uma vez que diferentes configurações e hiperparâmetros foram explorados durante o treinamento dos modelos, acarretando apenas em variações insignificantes dos resultados, recomenda-se a ampliação do conjunto de treinamento, acrescentando imagens com diferentes tipos de acabamento

(não apenas concreto) como forma de melhorar a capacidade de generalização do modelo. Outra possibilidade seria a alteração na estratégia de treinamento acrescentando uma nova classe com imagens que não contenham fissuras, de modo a minimizar a ocorrência de FP.

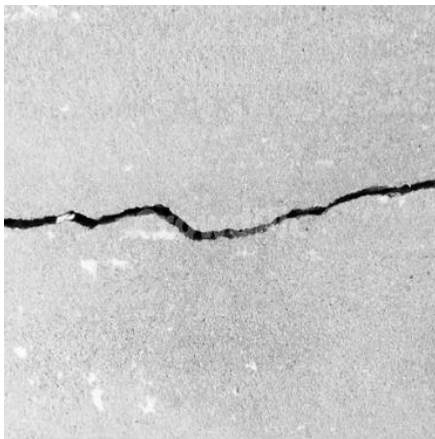
Para fins de comparações, embora boa parte da literatura pesquisada tenha dado ênfase a acurácia como principal métrica, para o presente trabalho, que envolve apenas uma classe de instâncias (fissuras), a acurácia não é a melhor forma de avaliar o modelo treinado; assim, adotou-se o mAP@0,50 como métrica principal, por integrar em um único indicador o equilíbrio entre precisão e recall em diferentes limiares de confiança, oferecendo uma visão global do desempenho do modelo. Essa métrica é amplamente utilizada nos principais benchmarks da área, como PASCAL VOC, COCO e a família YOLO, e consolidou-se como referência em sistemas de detecção de objetos (EVERINGHAM et al., 2010; REDMON et al., 2016; REDMON; FARHADI, 2018; LIN et al., 2014).

#### 4.1.1 PREDIÇÕES COM O MODELO TREINADO

Nesta seção são apresentadas algumas imagens resultantes do processo de predição, visando ilustrar os resultados da etapa de detecção. Para isso, foram selecionadas algumas imagens do conjunto IM\_1 separadas para os testes contendo fissuras com diferentes níveis de dificuldade para detecção e imagens do conjunto IM\_2 para analisar a generalização do modelo.

A Figura 33 apresenta três exemplos distintos: (A) uma fissura linear e bem destacada em relação aos demais elementos da cena; (B) uma fissura mais tênue; e (C) uma fissura com múltiplas ramificações. Para cada caso, são exibidas a imagem original e o resultado após a predição, com as anotações automáticas geradas pelo modelo. A Figura 34 mostra os resultados das predições, incluindo segmentações e caixas delimitadoras. Ressalta-se que os três modelos treinados apresentaram resultados semelhantes para esses três exemplos.

Figura 33: Imagens utilizadas para exemplificar a predição



(A)



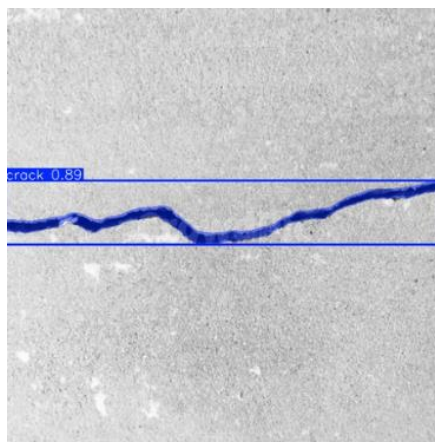
(B)

Fonte: Autor.

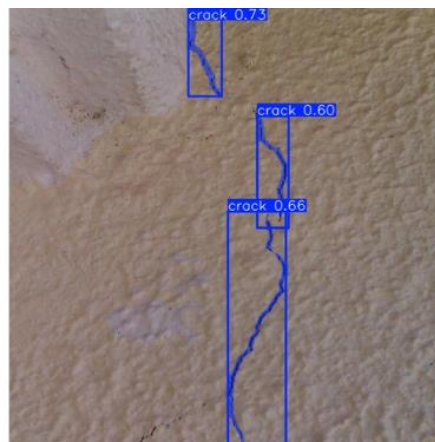


(C)

Figura 34: Anotações após predição pelo modelo treinado

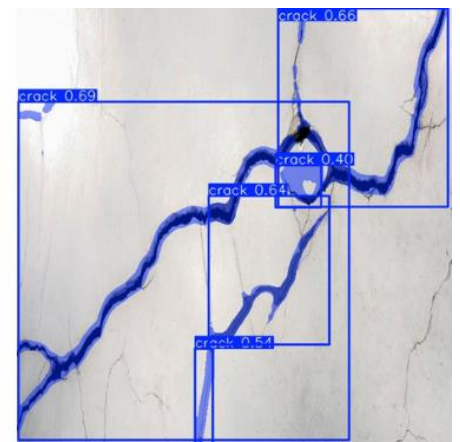


(A)



(B)

Fonte: Autor.

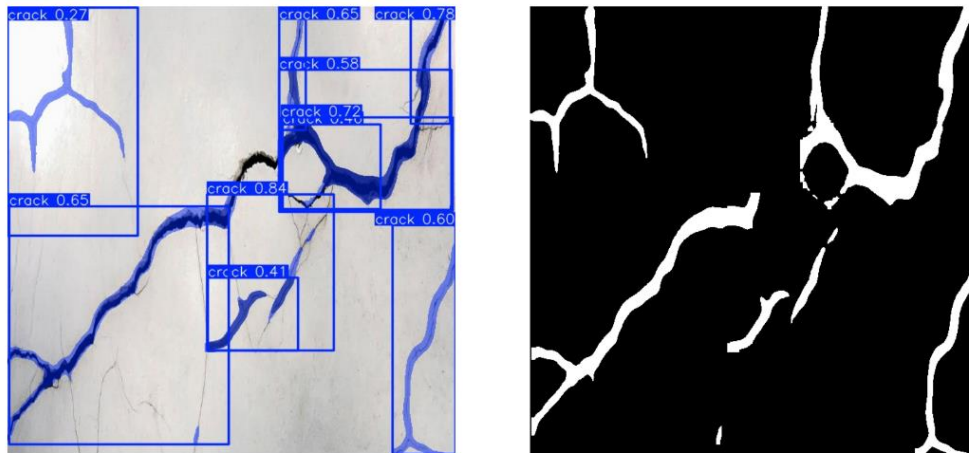


(C)

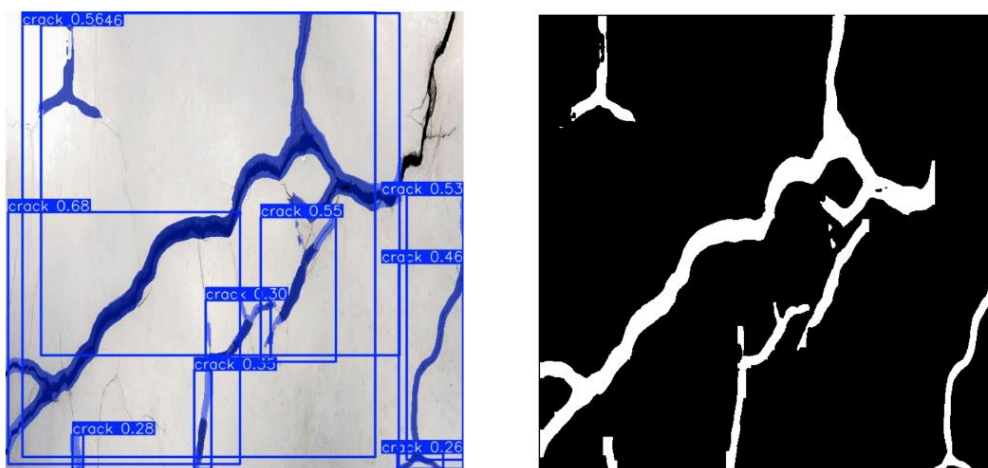
A Figura 35 apresenta um comparativo dos três modelos treinados com 100 (Modelo 1), 300 (Modelo 2) e 500 (Modelo 3) épocas, mostrando a similaridade dos resultados de detecção e segmentação aplicados à imagem C da Figura 33.

Figura 35: Comparativo de predições

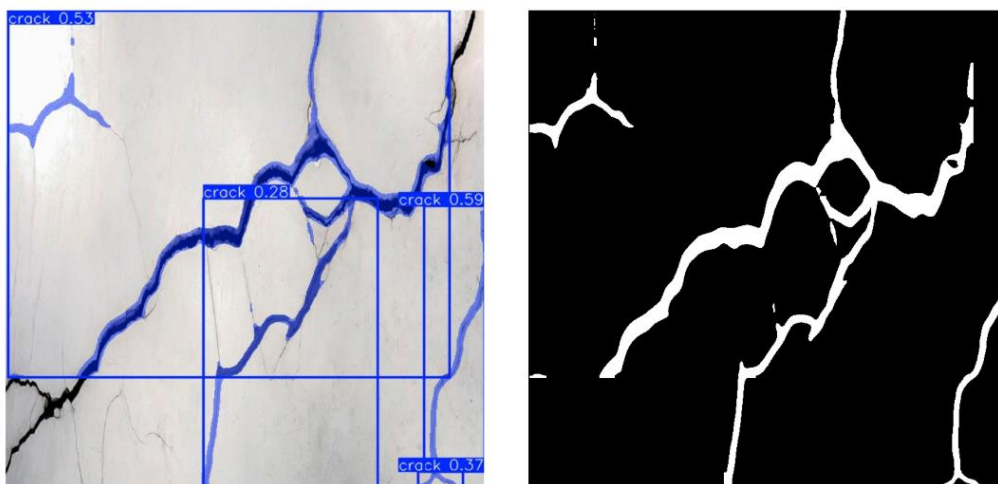
**(A) Predição com Modelo 1**



**(B) Predição com Modelo 2**



**(C) Predição com Modelo 3**

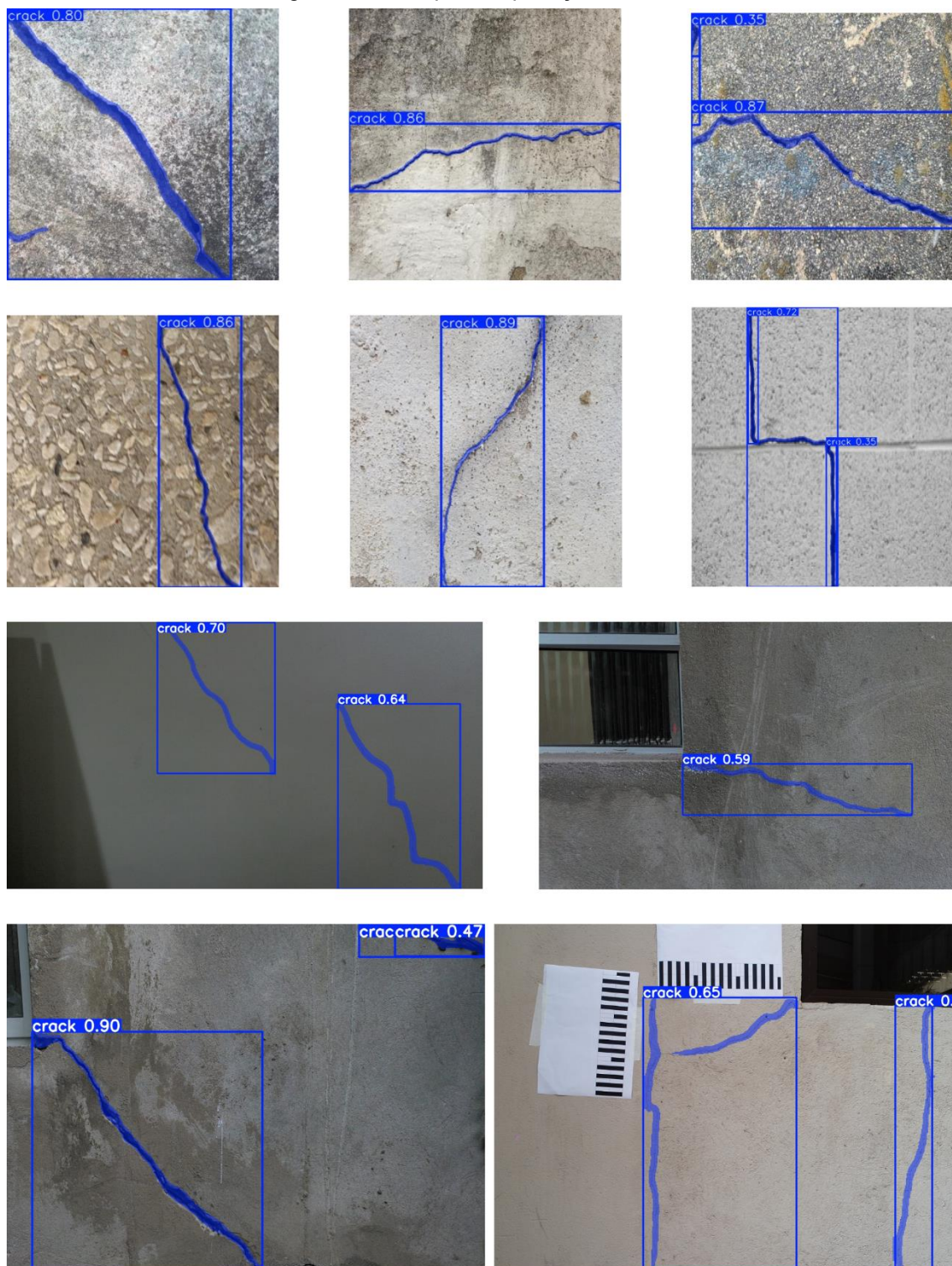


Fonte: Autor.

Os modelos obtiveram bom desempenho de detecção e segmentação em fissuras com aspecto bem definido e sem muitas ramificações, conforme exemplos apresentados na Figura 36.



Figura 36: Exemplos de predições corretas

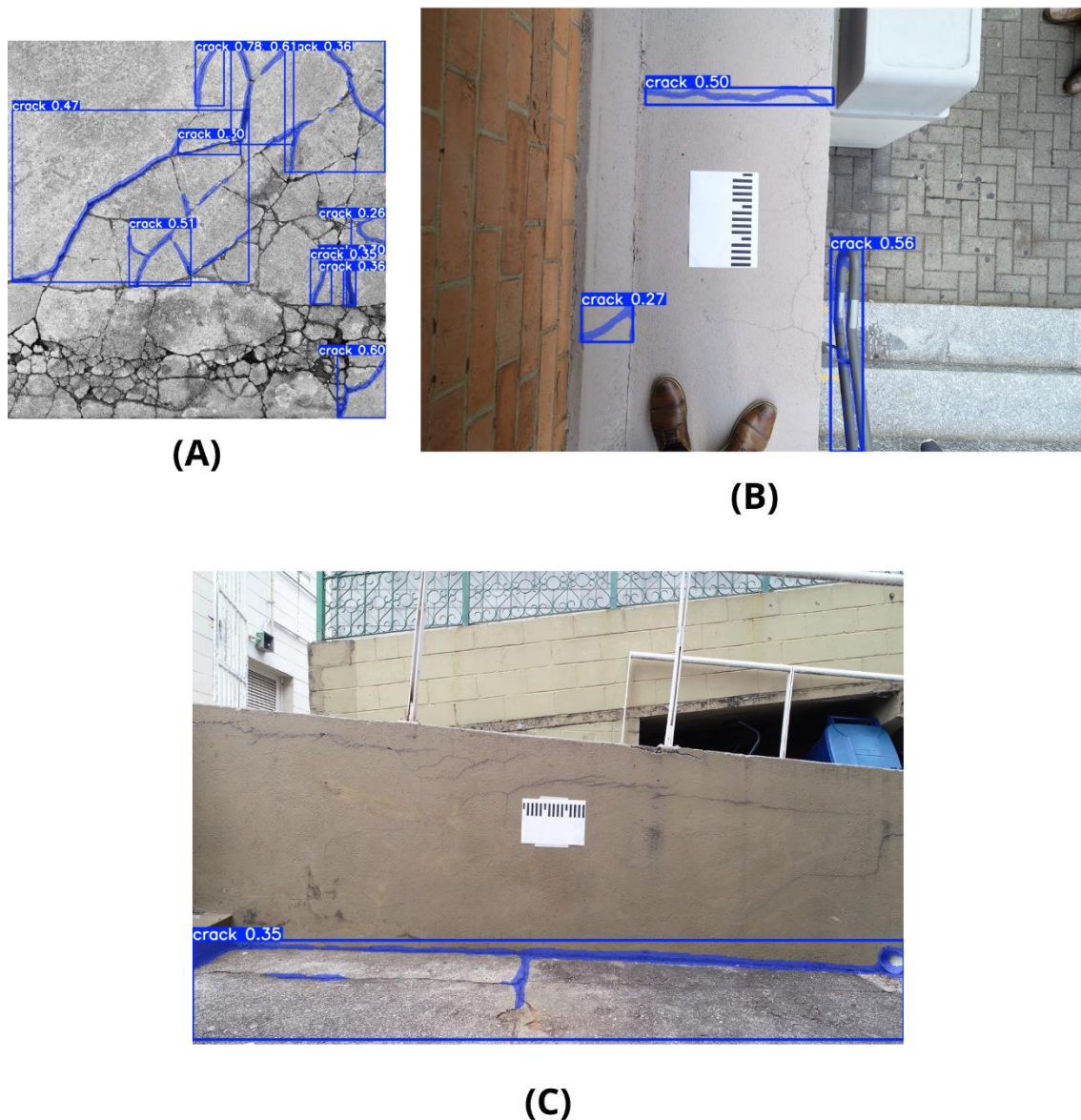


Fonte: Autor.

No entanto, foi observada dificuldade dos modelos em detectar fissuras muito tênues ou finas, fissuras muito ramificadas, bem como dificuldade em distinguir elementos como arestas de fissuras, que neste último caso gerou falsos positivos. A Figura 37 apresenta imagens exemplificando algumas das

dificuldades observadas. A Figura 37 (A) evidencia dificuldades na detecção e segmentação de fissuras com múltiplas ramificações. Na Figura 37 (B) o corrimão foi confundido com uma fissura, e algumas fissuras mais tênues não foram detectadas. Por fim, na Figura 37 (C), a junção da parede com o piso foi considerada uma fissura, e fissuras tênues na parede não foram detectadas.

Figura 37: Exemplos de falhas de predições



Fonte: Autor.

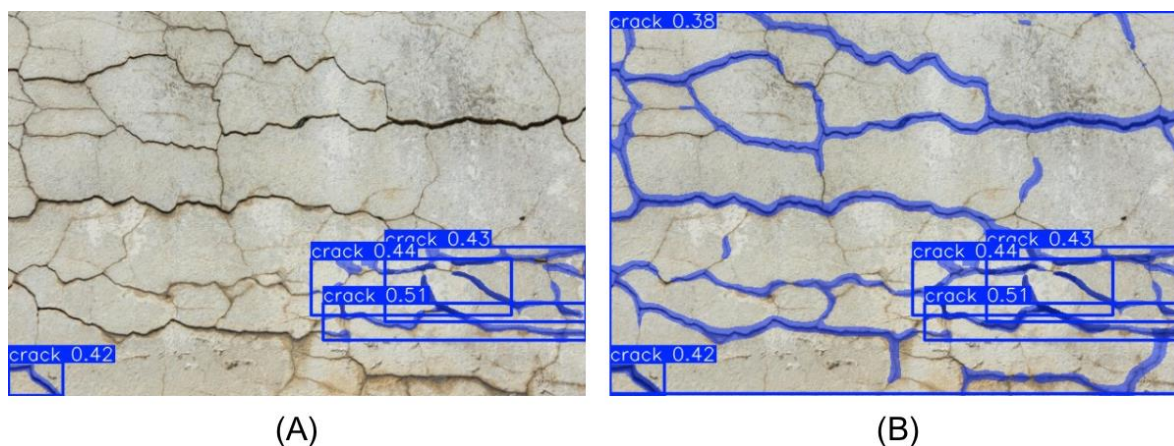
O YOLOv8 dispõe de um parâmetro de confiança (*conf*), que permite calibrar a partir de qual grau de confiança a detecção é considerada verdadeira. Esse parâmetro é configurável e aceita valores de 0 a 1, sendo 0,25 o valor padrão. A calibração correta desse parâmetro pode aumentar a sensibilidade da



detecção quando reduzido, porém o aumento da sensibilidade pode acarretar maior ocorrência de FP.

A Figura 38 apresenta um comparativo envolvendo um tipo de fissura particularmente desafiador para os modelos treinados: fissuras altamente ramificadas. Na Figura 38 (A), as predições foram obtidas com limiar de confiança equivalente à 40% ( $\text{conf} = 0,40$ ), enquanto na Figura 38 (B) o limiar foi reduzido para 25% ( $\text{conf} = 0,25$ ).

Figura 38: Comparação com alteração do parâmetro de confiança



Fonte: Autor.

Este último caso evidencia que a calibração conveniente do modelo preditivo representa um desafio para melhorar o desempenho das predições sem, no entanto, aumentar o número de FP.

#### 4.1.2 DETECÇÃO DE FISSURAS EM CENÁRIO DE MONITORAMENTO CONTÍNUO DURANTE O VOO DO DRONE

Com o objetivo de avaliar o desempenho do modelo YOLOv8-seg na detecção de fissuras em condições próximas ao tempo real, foi realizado um experimento utilizando quadros de um vídeo produzido pelo Laboratório de Geotecnologias da Universidade Mackenzie. Esse vídeo foi obtido por meio de um drone DJI Mavic 3 Thermal (equipado com câmeras RGB e térmica), registrando imagens das superfícies externas das paredes de um shopping center. Nos experimentos, cada quadro do vídeo, com duração de 1 min e 34 s

e resolução de  $1920 \times 1080$  pixels, foi processado individualmente pelo modelo, permitindo simular o cenário de monitoramento contínuo durante um voo real.

Utilizando um computador com processador Intel i7-1355U, 32 GB de RAM, observou-se que o modelo YOLOv8s-seg, quando executado exclusivamente em CPU, alcançou entre 4 e 6 frames por segundo (FPS) após a redução da resolução da imagem para  $640 \times 480$  pixels. Esses valores mostram que o processamento apenas em CPU é insuficiente para aplicações próximas ao tempo real. Para avaliar alternativas, também foram realizados experimentos utilizando uma unidade de processamento gráfico (*Graphics Processing Unit – GPU*) NVIDIA MX550 2 GB, nos quais o desempenho foi significativamente superior, acelerando o processamento dos quadros em aproximadamente dez vezes, chegando a atingir 40 FPS. Esse ganho reforça a necessidade de GPU para viabilizar aplicações com baixa latência e capacidade de processamento compatível com o fluxo contínuo de quadros transmitidos pelo drone. Ressalta-se, entretanto, que no caso de inspeções baseadas em imagens adquiridas por drones, o desempenho final também depende do atraso (*delay*) associado ao sistema de transmissão de imagens. Assim, além do hardware de processamento, uma infraestrutura de transmissão com alta taxa de transferência é igualmente essencial para o uso prático dessa solução em campo.

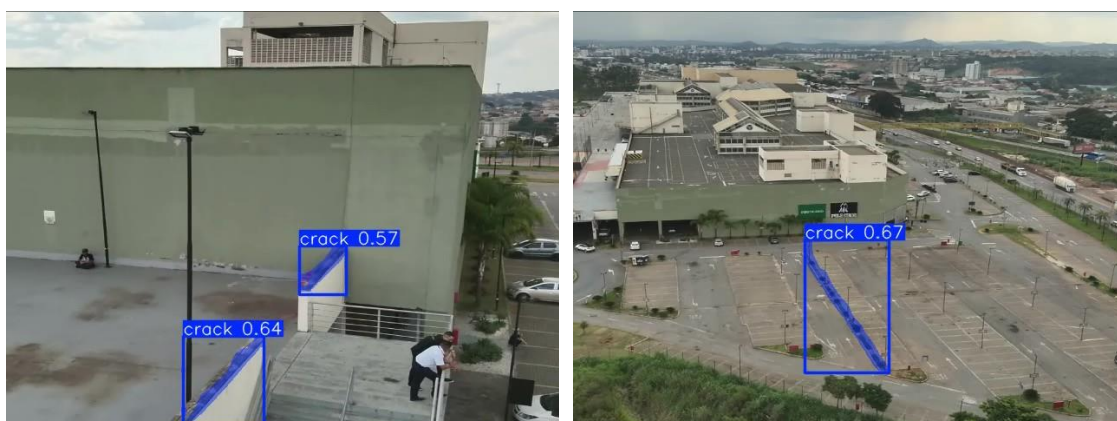
Apesar da viabilidade observada, conforme ilustrado na Figura 39, que apresenta o resultado das predições do modelo para um dos quadros do vídeo, alguns desafios foram identificados nesse experimento. Por exemplo, houve casos frequentes de FP, especialmente associados a feições estruturais como quinas de beirais, cantos de muros e bordas de canteiros, que foram erroneamente interpretadas como fissuras. Como pode ser visto na Figura 40, esse problema se intensifica quando os objetos são capturados a longas distâncias (superiores a 30 metros), levando a uma redução da resolução espacial, que dificulta a distinção entre fissuras reais e padrões geométricos lineares presentes na cena.

Figura 39: Exemplo de predições corretas em um quadro do vídeo



Fonte: Autor.

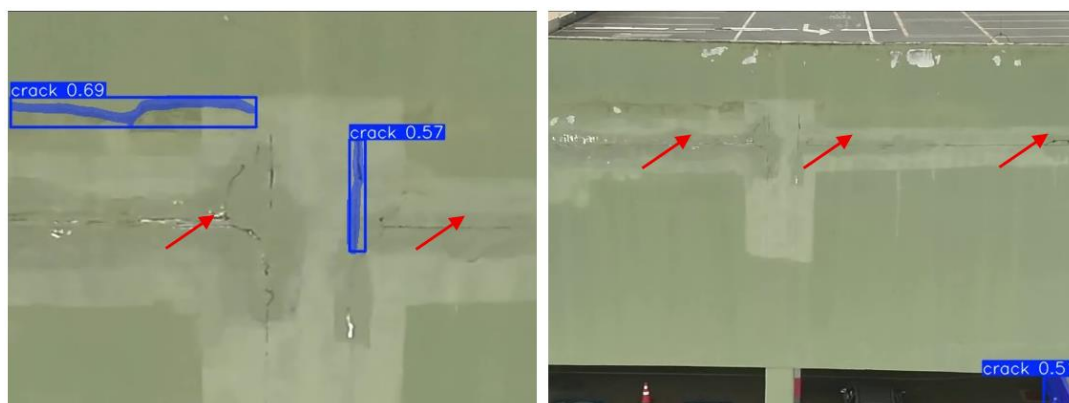
Figura 40: Exemplos de predições em quadros do vídeo ilustrando casos de FP



Fonte: Autor.

Também foram notados casos de FN, principalmente decorrentes da distância entre o drone e a superfície inspecionada, da falta de foco em determinados quadros, e dos efeitos de movimento do drone, que prejudicam a nitidez das anomalias. Alguns desses problemas estão ilustrados na Figura 41.

Figura 41: Exemplos de predições em quadros do vídeo ilustrando casos de FN



Fonte: Autor.

Assim como nos experimentos com as imagens do conjunto de teste, os casos de FP e FN na detecção em tempo real podem ser diminuídos por meio da ampliação do conjunto de imagens de treinamento, incorporando maior variabilidade de cenários, de distâncias e das condições gerais de aquisição. Além disso, recomenda-se o treinamento do modelo utilizando duas classes: “fissura” e “não fissura”. Essa estratégia permitiria que o modelo aprendesse explicitamente padrões associados a regiões que não representam anomalias, reduzindo a taxa de falsos positivos. Nesse caso, durante a etapa de predição, o sistema deve ser configurado para reportar apenas ocorrências da classe “fissura”.

Em resumo, o experimento descrito nesta seção sinaliza o potencial do método proposto para aplicações reais de detecção de fissuras com o uso de drones, ao mesmo tempo em que evidencia oportunidades de aprimoramentos visando aumentar a confiabilidade das predições em situações de monitoramento contínuo durante o voo do drone.

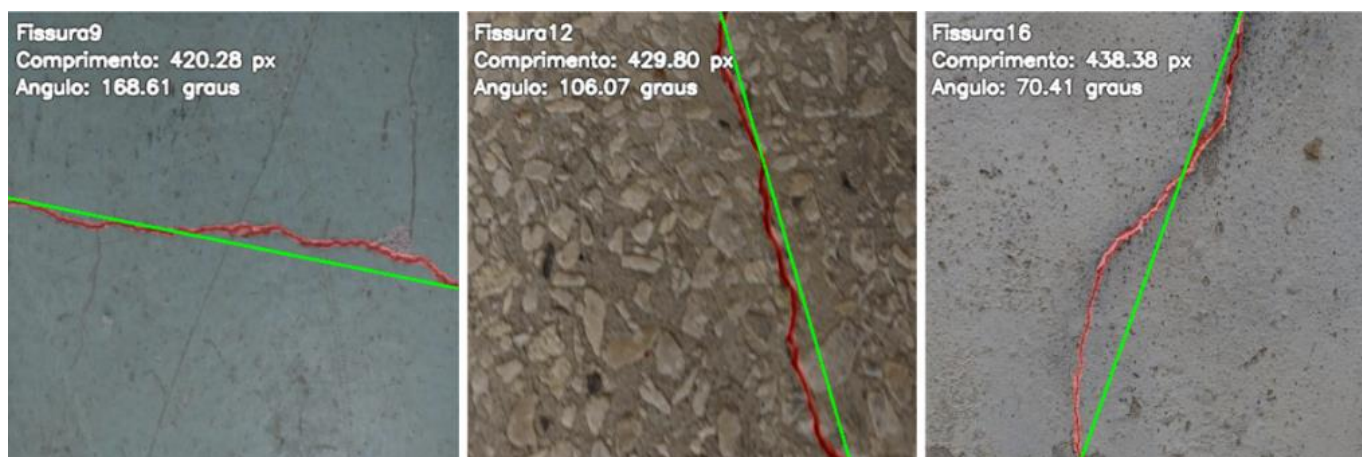
## 4.2 ETAPA 2: MENSURAÇÃO DE FISSURAS

Após a detecção automática, a máscara de segmentação é utilizada na etapa de mensuração, para extrair as características geométricas da fissura, como comprimento e angulação. Contudo, para que essa mensuração seja confiável, é fundamental seguir o procedimento de aquisição apresentado na seção 3.2.3. Qualquer falha na captura, como ângulo inadequado, distorções ou iluminação incorreta, pode comprometer toda a análise e gerar valores imprecisos.

### 4.2.1 MENSURAÇÃO DO COMPRIMENTO E ANGULAÇÃO POR SEGMENTO DE RETA

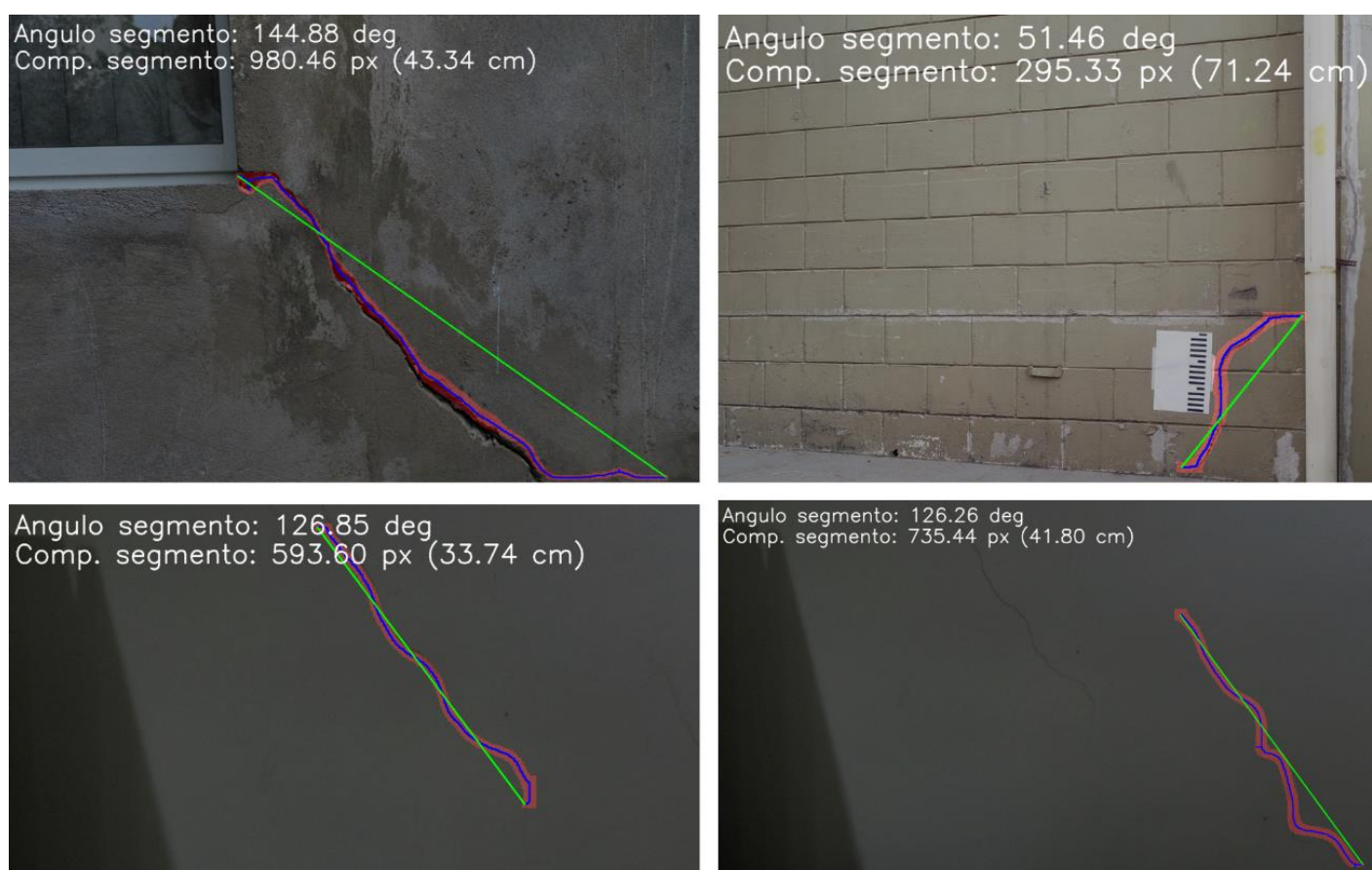
A Figura 42 apresenta alguns resultados de mensuração das características geométricas das fissuras detectadas nas imagens de teste do conjunto IM\_1, enquanto a Figura 43 exemplifica os resultados obtidos para imagens do conjunto IM\_2.

Figura 42: Exemplos de mensuração de fissuras com IM\_1



Fonte: Autor.

Figura 43: Exemplo de mensuração de fissura com IM\_2



Fonte: Autor.

Nas Figuras 42 e 43 é possível verificar que o método proposto foi capaz de traçar o segmento de reta entre as extremidades das fissuras detectadas e segmentadas, dessa forma foi possível obter a angulação predominante em relação ao eixo horizontal, e medir o a extensão da fissura por meio do comprimento do segmento de reta em pixels (Figuras 42 e 43); para as imagens

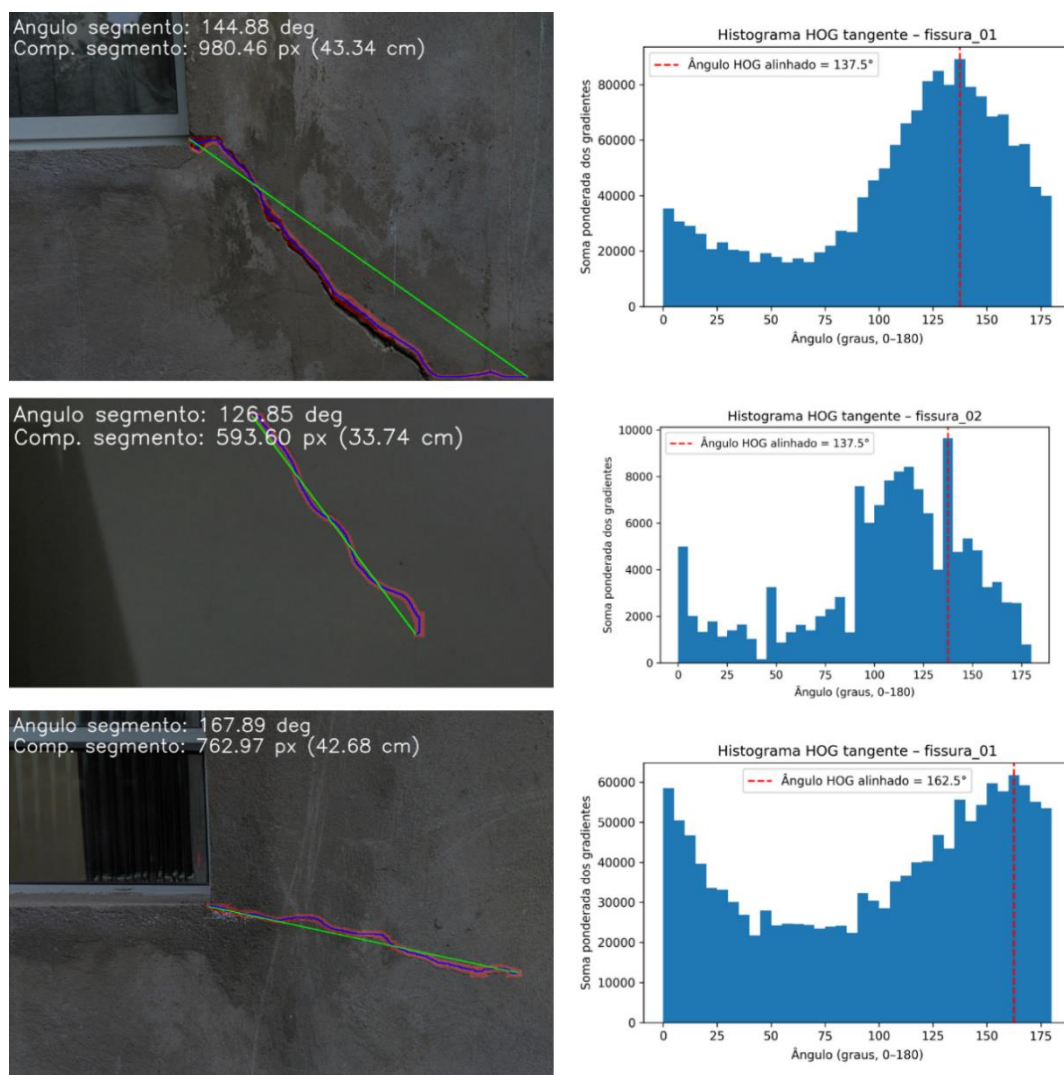


do conjunto IM\_2 (Figura 43) das quais dispomos das informações técnicas da câmera e da distância entre a câmera e a superfície que apresenta a fissura, também foi possível converter a medida de comprimento de pixels para unidades do SI, na seção 4.2.4 é abordada a análise da precisão da conversão de pixels para SI utilizando o GSD.

#### 4.2.2 MENSURAÇÃO DA ANGULAÇÃO PRINCIPAL DA FISSURA COM HOG

A Figura 44 apresenta algumas imagens com comparação entre a mensuração de angulação principal da fissura pela abordagem do segmento de reta e pela abordagem usando HOG.

Figure 44: Comparação da medição de ângulos

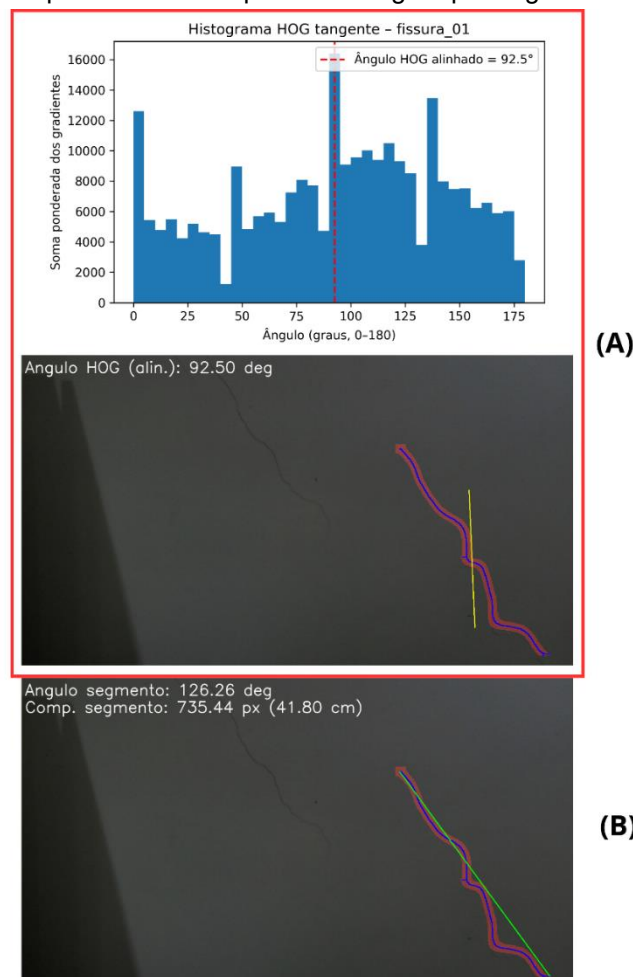


Fonte: Autor.

As angulações obtidas nos exemplos da Figura 44 nas duas abordagens foram próximas, satisfazendo o objetivo de determinar a angulação predominante da fissura. Com essa informação é possível para o profissional de inspeção saber se a fissura tem direção predominante horizontal, vertical ou diagonal, e dessa forma utilizar essa informação em conjunto com outras para realizar o diagnóstico.

Ambas as abordagens utilizadas para obtenção da angulação predominante das fissuras se mostraram promissoras, a abordagem puramente geométrica (segmento de reta) obteve melhores resultados em fissuras lineares sem ramificações, pois em alguns casos apenas a avaliação utilizando HOG pode levar a conclusões equivocadas, como apresentado na Figura 45 (A), onde a angulação principal obtida via HOG indica uma fissura vertical e não diagonal. Já na Figura 45 (B), a abordagem via segmento de reta descreve melhor a angulação predominante da fissura.

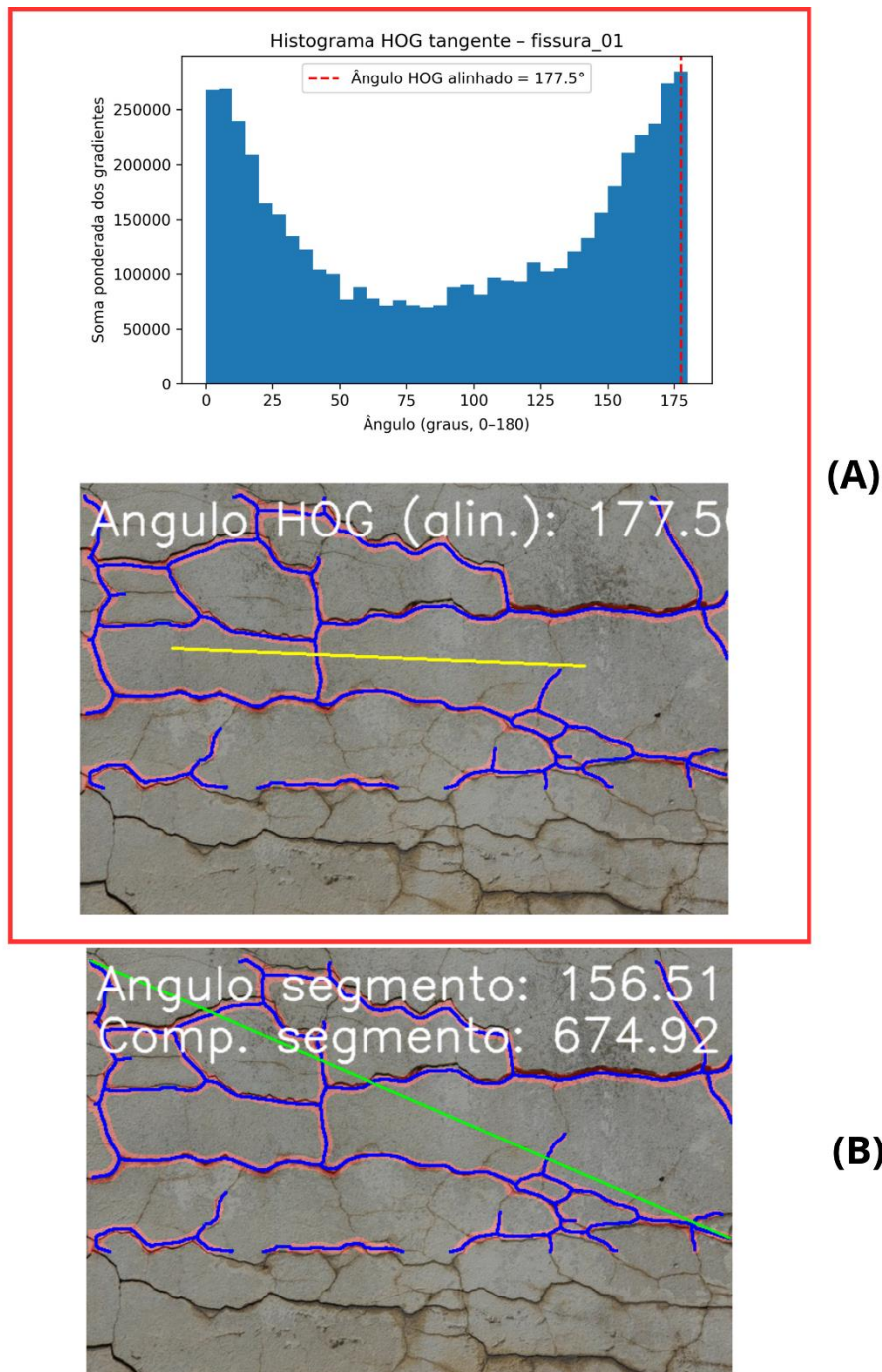
Figure 45: Exemplo de caso em que a abordagem por segmento de reta é melhor



Fonte: Autor.

No entanto, a abordagem com uso do HOG apresenta vantagem em fissuras com múltiplas ramificações, conforme exemplifica a Figura 46 (A), já que o histograma agrega informação direcional de uma região inteira, permitindo uma medida estatística global da orientação predominante. Para este tipo de fissura, a abordagem por segmento de reta não descreve convenientemente a angulação predominante como mostra a Figura 46 (B).

Figure 46: Exemplo de caso em que a abordagem por HOG é melhor



Fonte: Autor.

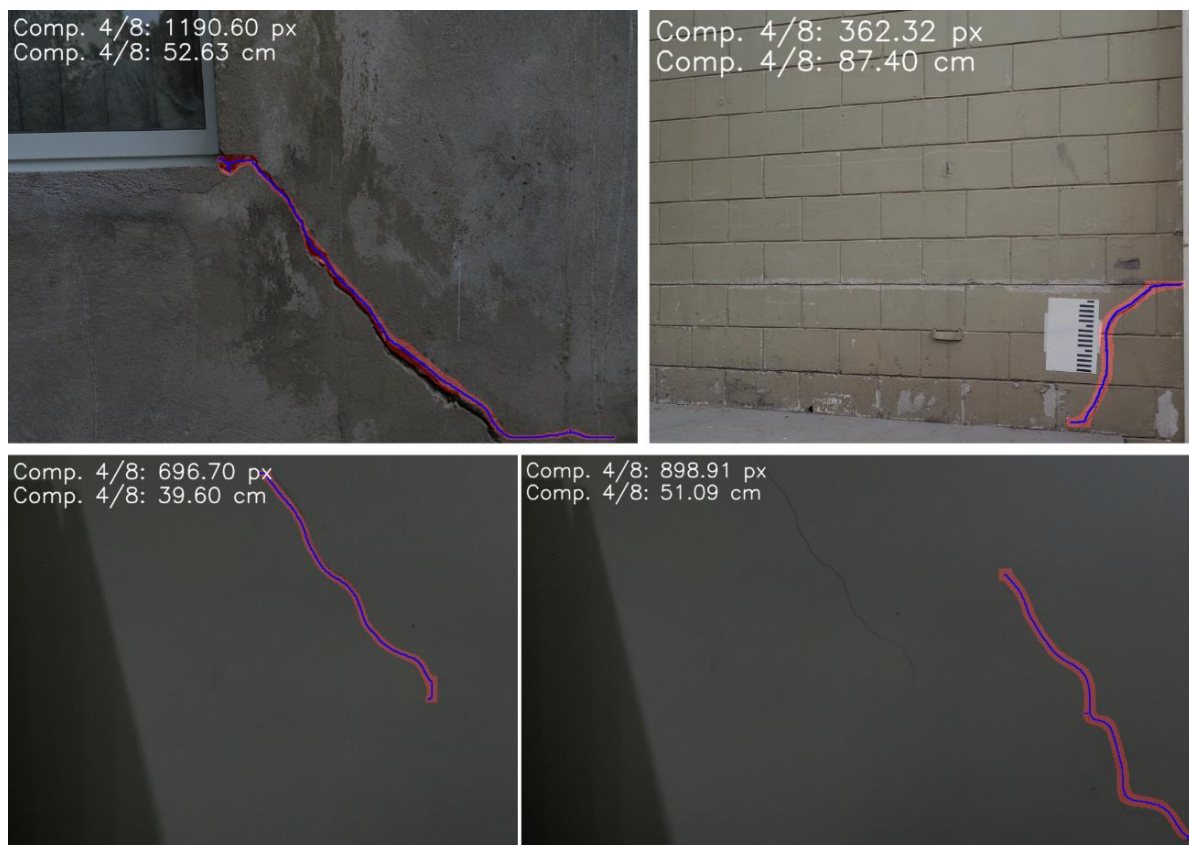


Por fim, cabe ressaltar que outras abordagens como análise de componentes principais (PCA) também podem ser utilizados para obtenção de angulação predominante das fissuras, por PCA a partir da distribuição dos pixels da segmentação, é possível identificar a direção de maior variância dos pontos. O primeiro componente principal define o eixo predominante do traçado, cuja orientação em relação ao eixo horizontal da imagem corresponde à angulação principal da fissura, fornecendo uma estimativa global e robusta de sua direção de propagação.

#### 4.2.3 MENSURAÇÃO DE COMPRIMENTO APÓS ESQUELETIZAÇÃO

Para possibilitar a mensuração de toda a extensão de uma fissura, e não apenas de seu comprimento principal, é conveniente aplicar a esqueletização da fissura segmentada, permitindo determinar seu comprimento por meio da soma dos pixels pertencentes ao esqueleto e pelo método de elementos conexos, conforme ilustrado na Figura 47.

Figure 47: Esqueletização e obtenção do comprimento



Fonte: Autor.

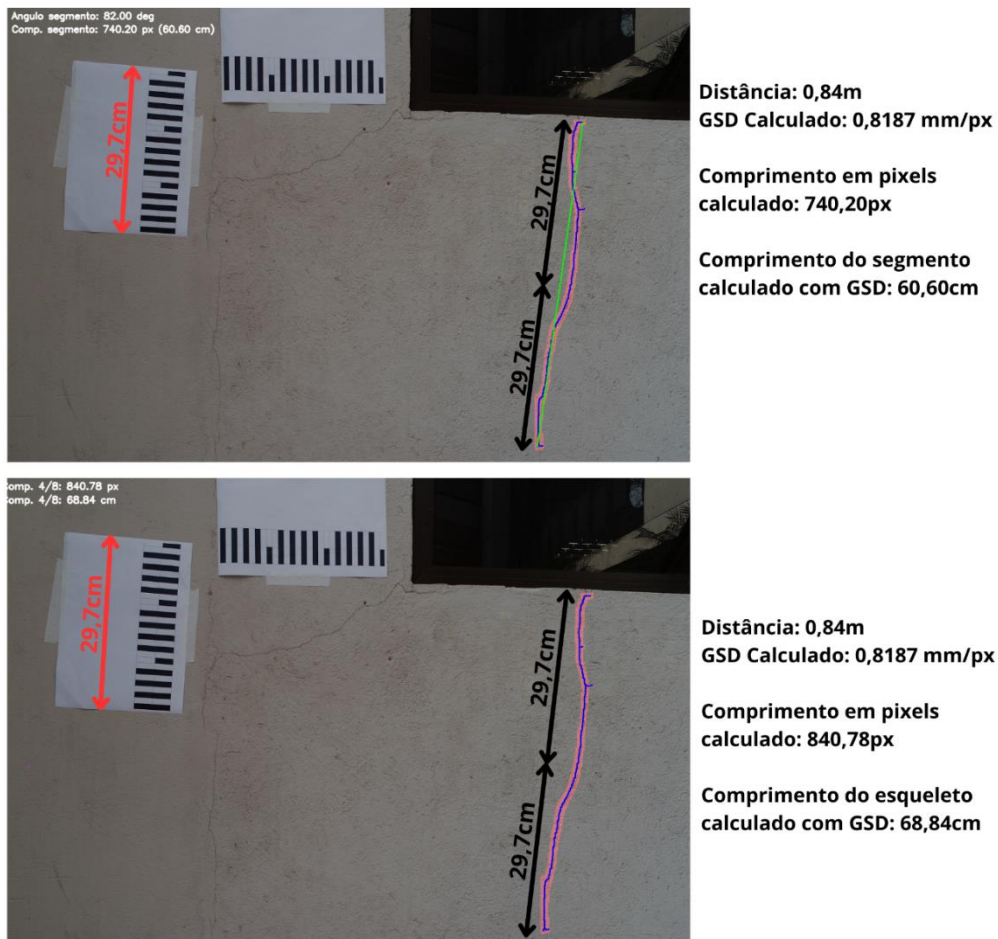
Na Figura 47, além da segmentação e esqueletização, foram apresentados os valores obtidos de comprimento com base nos pixels que formam o esqueleto da fissura, bem como os valores convertidos em unidades de comprimento do SI. Essa conversão foi realizada utilizando o GSD, como explicado na seção 3.3.2.4.

#### 4.2.4 CONVERSÃO DAS MEDIDAS DE COMPRIMENTO PARA O SI USANDO GSD

Inicialmente a etapa de mensuração do método proposto obtém as medidas de comprimento em pixels, sendo necessária uma calibração para permitir a conversão para uma unidade do SI, como m, cm ou mm, de modo a possibilitar a comparação com a medição realizada *in loco*. Para tanto, foi utilizado o GSD como anteriormente explicado.

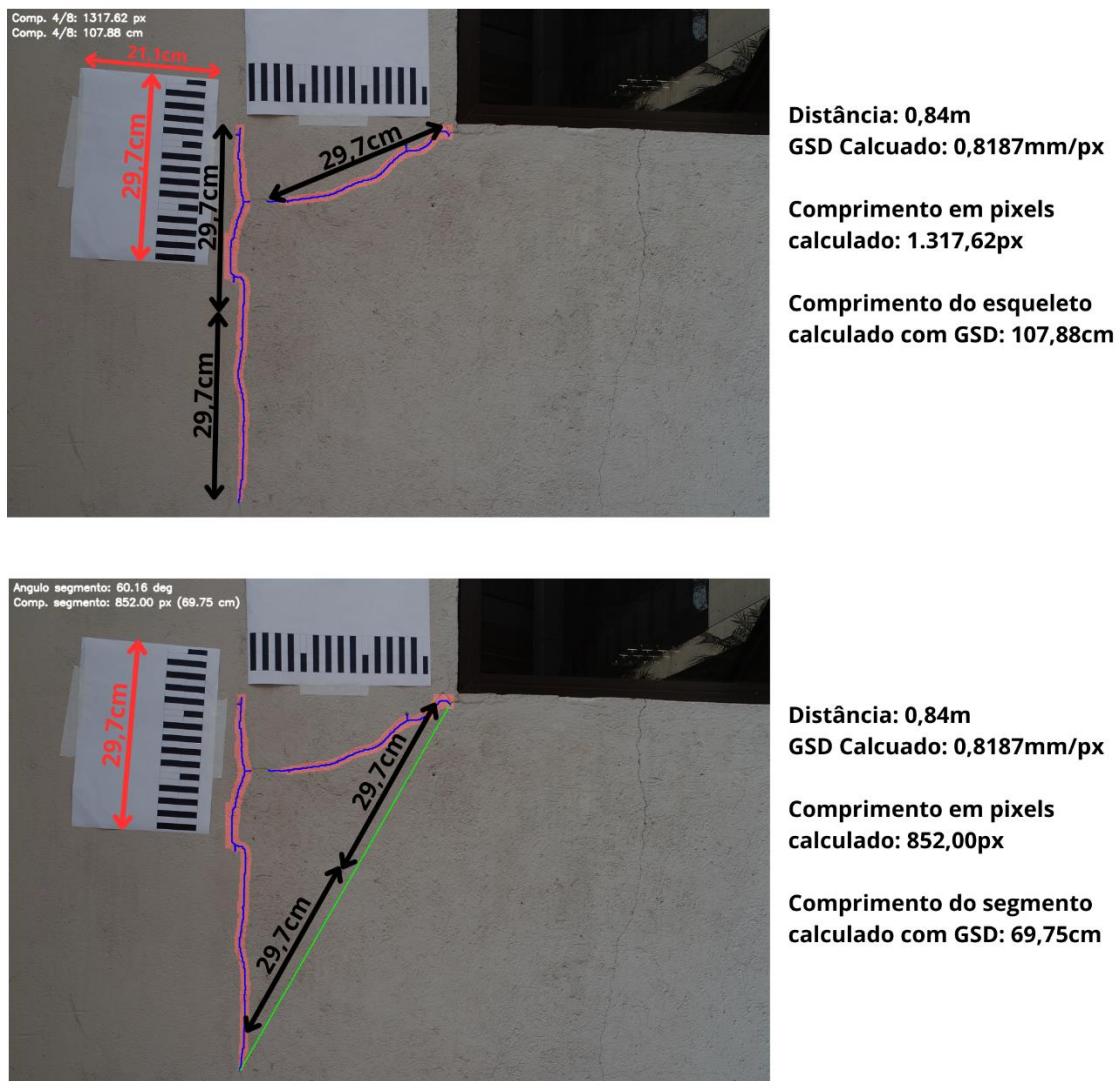
As figuras 48 e 49 apresentam os valores de comprimento calculados utilizando GSD e a comparação com dimensões conhecidas.

Figure 48: Obtenção de comprimento por GSD e comparação com medida real (fissura simples)



Fonte: Autor.

Figure 49: Obtenção de comprimento por GSD e comparação com medida real (fissura ramificada)

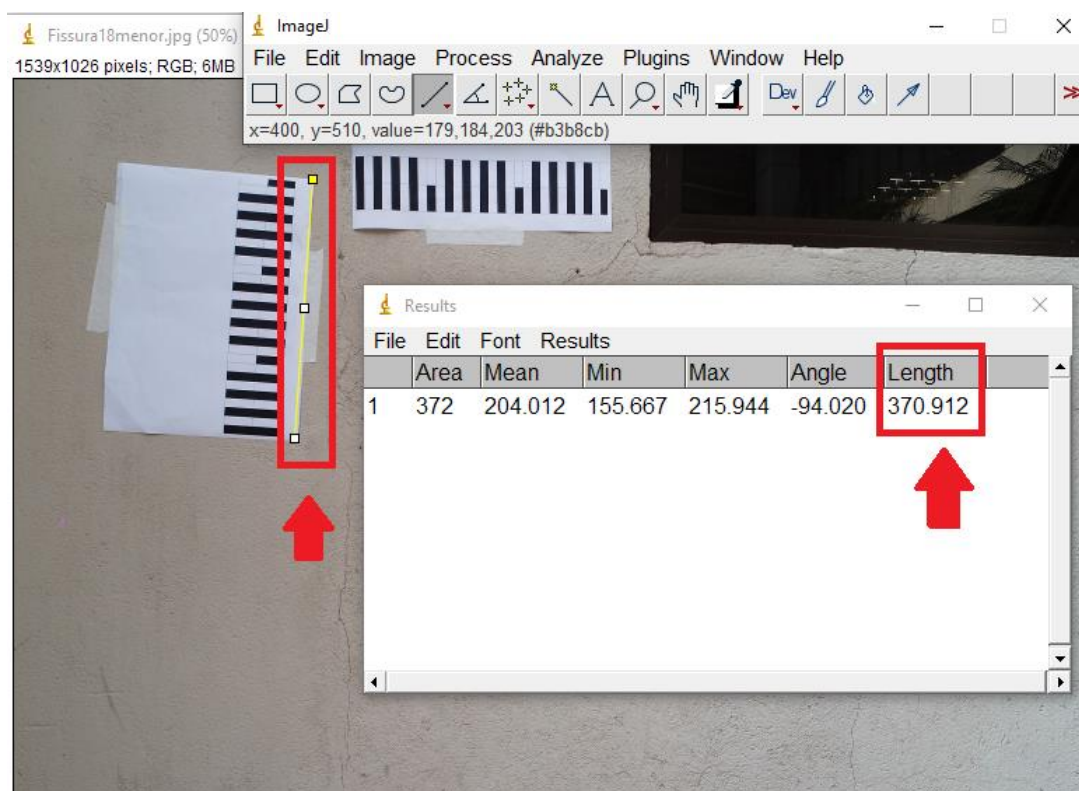


Fonte: Autor.

Com o intuito de estabelecer maior rigor e precisão na comparação entre as medidas reais e as medidas calculas pelo método proposto por meio do GSD, foi utilizado um software open source de processamento e análise científica de imagens chamado ImageJ (disponível em <https://imagej.net/ij/index.html>), este mesmo software foi utilizado em outros trabalhos como em Forcael et al (2024), seu uso no presente trabalho foi feito, conforme apresentado na Figura 50, para medir em pixels o comprimento do maior lado da folha A4 de referência, essa medida em pixels e medida real da folha A4 (297 mm) foram utilizadas para calcular o comprimento no SI dos segmentos de reta traçados, através de cálculo baseado em proporções, possibilitando comparação numérica e estimativa do

erro percentual da medida obtida por meio do GSD, esses valores encontram-se na Tabela 13.

Figure 50: Lateral da folha A4 em pixels



Fonte: Autor.

Tabela 13: Comparação das medidas

| Item                  | Comprimento (px) | Comprimento por GSD (cm) | Comprimento por proporção (cm) | Erro (%) |
|-----------------------|------------------|--------------------------|--------------------------------|----------|
| Segmento<br>Figura 48 | 740,20           | 60,60                    | 59,27                          | 2,24%    |
| Segmento<br>Figura 49 | 852,00           | 69,75                    | 68,22                          | 2,24%    |

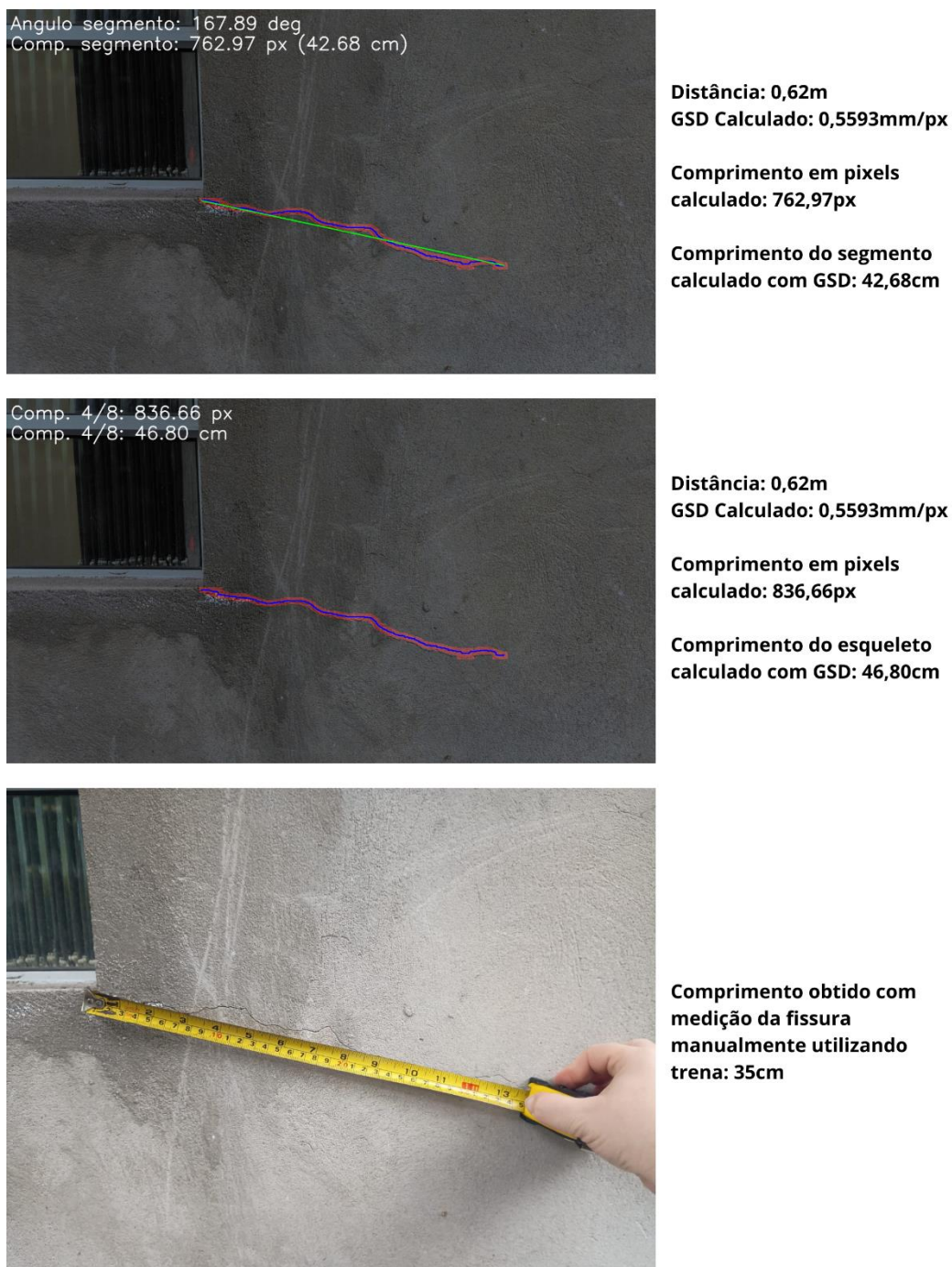
Fonte: Autor.

É importante ressaltar que essas imagens foram capturadas com o drone DJI Phantom 4 Pro, cujos parâmetros da câmera: largura ( $w = 13,2\text{mm}$ ), altura ( $h = 8\text{mm}$ ) e distância focal ( $f = 8,8\text{mm}$ ), estão disponíveis na calculadora de GSD do site: <https://www.handelselaras.com/calculator/>, o que torna o cálculo do comprimento mais coerente com a medida real.



A Figura 51 apresenta comparação entre o comprimento calculado por meio do GSD e medição manual utilizando trena no local da fissura.

Figure 51: Comparação de comprimento por GSD e com medida local



Fonte: Autor.

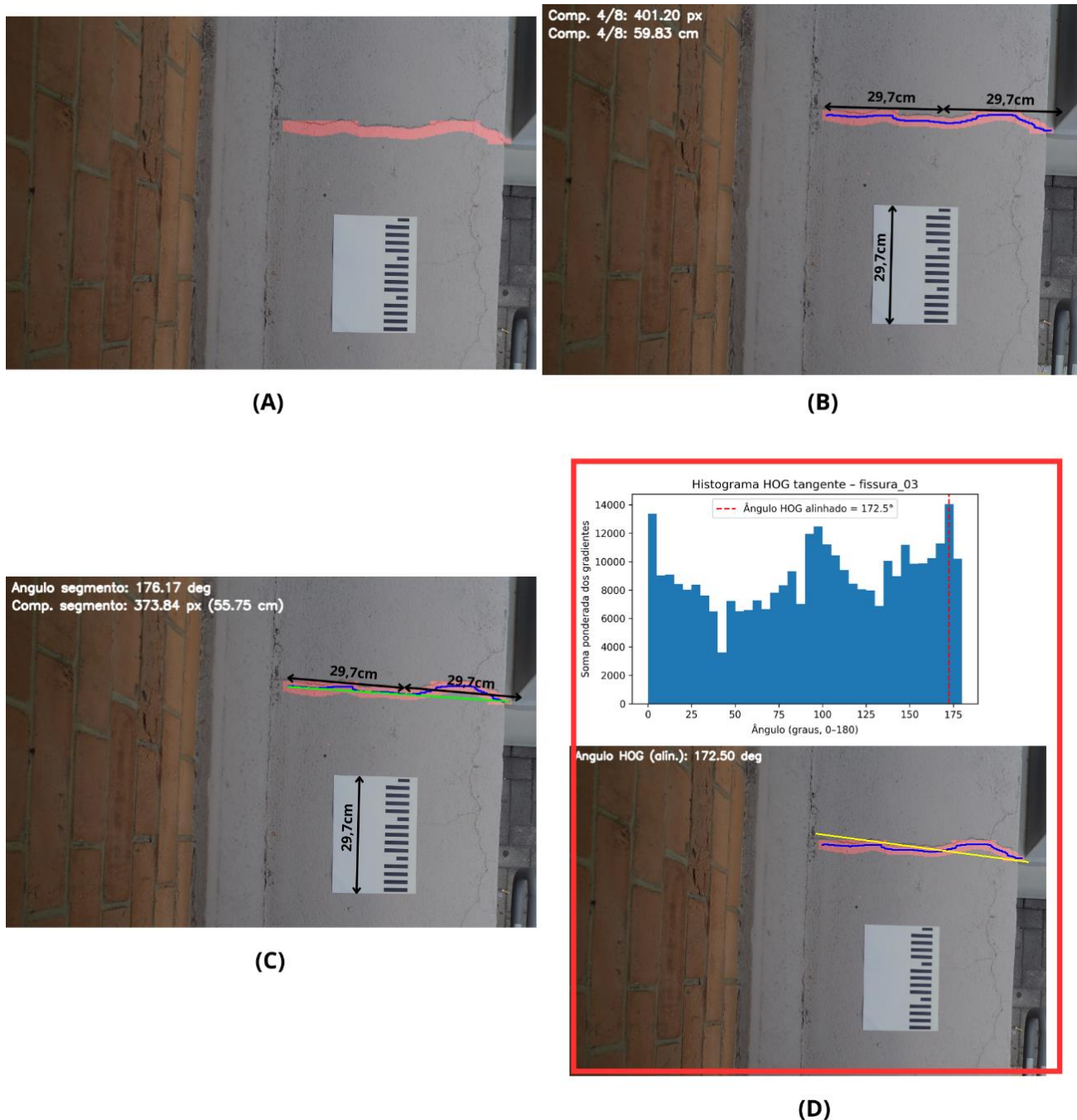
A imagem mostrada na Figura 51 foi capturada utilizando o drone DJI Mini 3, cujo a única informação precisa encontrada foi o FOV (82,1°), sendo as outras características estimadas, o que acarretou substancial diferença entre o valor calculado por meio do GSD do segmento de reta e o valor medido manualmente

no local (erro de 21,94%), evidenciando que calibrações imperfeitas inviabilizam a mensuração precisa por meio do GSD.

#### 4.2.5 INTEGRAÇÃO DAS ETAPAS DE DETECÇÃO E MENSURAÇÃO

Conforme já explicado, o método computacional proposto neste trabalho foi implementado de modo a executar o *pipeline* de detecção, segmentação e mensuração de fissuras em imagens. A Figura 52 apresenta as saídas do *pipeline*.

Figure 52: *Pipeline* completo de detecção e mensuração



Fonte: Autor.

Inicialmente, o usuário deve informar o local onde a imagem com as fissuras se encontra, além dos seguintes parâmetros da câmera empregada na aquisição da imagem: largura do sensor (mm), distância focal (mm) e distância (m) entre a lente da câmera e a fissura alvo. Com essas informações, o *pipeline* é executado, produzindo os resultados das duas etapas que compõem o método, a Figura 51 (A) apresenta a saída da etapa 1 do método responsável pela detecção e segmentação da fissura, realizada pelo modelo YOLO treinado; na sequência inicia-se a segunda etapa do método, responsável pela mensuração da fissura detectada, a Figura 51 (B) apresenta a saída da esqueletização da segmentação e a mensuração da extensão da fissura pela abordagem da contagem de pixels (4/8 conexos) com resultado em pixels e unidade de comprimento do SI; na sequência a Figura 51 (C) apresenta a saída da mensuração de comprimento e ângulo pela abordagem do segmento de reta o ângulo predominante é exibido em graus e o comprimento do segmento de reta é exibido em pixels e em unidade de comprimento do SI; por fim a Figura 51 (D) apresenta a angulação predominante utilizando HOG, com saída é apresentado um segmento de reta representando a angulação predominante obtida e o histograma de angulações tangenciais dos gradientes. Além das saídas em arquivos de imagem o código gera um relatório em arquivo de texto (txt) com os resultados numéricos do *pipeline*, a Figura 53 apresenta o relatório.

Figure 53: Relatório com os resultados do *pipeline*

```
Fissura: fissura_03
Índice da máscara no YOLO: 1
Área da máscara (antes da limpeza): 1854.00 px^2

Comprimento (contagem simples):      379.00 px
Comprimento (arestas 4/8-conexas):    401.20 px

Extremidades do esqueleto: p1=(473, 185), p2=(846, 210)
Comprimento do segmento de extremidades: 373.84 px

Largura da imagem: 1368 px
Tamanho de pixel estimado no sensor: 0.009649 mm/px
GSD ≈ 0.001491 m/px  (≈ 1.4912 mm/px)

Comprimento estimado da fissura (arestas 4/8):
  ≈ 401.20 px
  ≈ 598.27 mm
  ≈ 59.83 cm
  ≈ 0.5983 m

Comprimento estimado da fissura (segmento de extremidades):
  ≈ 373.84 px
  ≈ 557.48 mm
  ≈ 55.75 cm
  ≈ 0.5575 m

Ângulo da fissura pelo segmento de extremidades: 176.17° (0-180°, sistema geométrico)
Ângulo principal pelo HOG tangente (bruto): 172.50° (0-180°)
Ângulo principal pelo HOG tangente (alinhado ao segmento): 172.50° (0-180°)
```

Fonte: Autor.

### 4.3 DISCUSSÃO DOS RESULTADOS

Os resultados obtidos nos testes relativos a primeira etapa do método proposto foram: precisão de 81,1%, *recall* de 69,7% e *mAP@0,50* de 71,1% para a detecção das fissuras, conforme explicado a métrica principal para avaliação da etapa de detecção é a *mAP@0,50*, onde foi obtido resultado superior ao de Diniz et al. (2023) que alcançaram *mAP@0,50* de 11,80% para detecção de fissuras; porém inferior ao obtido por Ma; Li; Zhao (2020) que atingiram *mAP@0,50* de 82,6%; a comparação com outros trabalho não é direta, pois a maioria utilizou conjuntos de imagens com pelo menos 2 classes (fissura e não fissura) e consideraram como principal métrica a Acurácia. Esta primeira etapa do método proposto se mostrou promissora e pode ser melhorada principalmente com o aumento e maior diversificação do conjunto de imagens utilizado para treinamento, essa ação pode aperfeiçoar a generalização do método melhorando a performance de detecção e segmentação para superfícies com acabamentos diferentes de alvenaria, o modelo atual foi treinado basicamente com imagens de fissuras em estruturas de concreto, além disso mudanças na estratégia de treinamento utilizando mais classes, como por exemplo imagens que não tenham fissuras, ou imagens que tenham instâncias que parecem fissuras (mas não são), também pode aprimorar o modelo.

A segunda etapa do método proposto, responsável pela mensuração das fissuras segmentadas na etapa anterior, também se mostrou promissora, satisfazendo o objetivo de determinar a angulação predominante e o comprimento tanto da direção principal da fissura, quanto de toda sua extensão atingindo erro de 2,24% entre a medida real e a calculada por meio do GSD. Bolaybolay et al. (2022) atingiram erro percentual médio de 5,56% na mensuração de comprimento e Flah, Suleiman e Nehdi (2020) obtiveram 1,5% na mensuração de comprimento, Forcael et al. (2024) utilizaram como métrica a significância estatística ( $p\text{-valor} - \alpha = 0,05$ ) o que faz com que a comparação dos resultados não seja direta; embora todos tenham alcançado resultados expressivos, também se depararam com o desafio da necessidade de calibração para possibilitar a conversão da medição de pixels para unidades de medida de comprimento do SI.



Além disso, observou-se que o procedimento de aquisição das imagens é fundamental, e deve ser seguido de forma criteriosa, principalmente para melhorar a eficiência da etapa de mensuração do método proposto.

Embora não faça parte do escopo principal deste trabalho, o instrumental desenvolvido permite estimar a largura máxima das fissuras detectadas, em pixels, por meio da segmentação, esqueletização e transformada da distância, conforme ilustrado na Figura 54.

Figure 54: Mensuração da largura da fissura detectada

Largura máxima (px): 26.00  
Ponto de máxima largura (x, y): (682, 495)

Largura máxima estimada: 26.00 px



Largura máxima (px): 28.00  
Ponto de máxima largura (x, y): (1290, 460)

Largura máxima estimada: 28.00 px



Fonte: Autor.

Contudo, em virtude de as fissuras apresentarem larguras muito finas, frequentemente da ordem de milímetros ou inferiores, tanto a calibração para conversão em unidades do SI via GSD quanto a própria segmentação tornam-se passos críticos, por exemplo, caso a máscara de segmentação não fique bem

delineada sobre a fissura acarretará imprecisão para determinação das bordas da fissura.

## 5 CONCLUSÕES E PROPOSTAS PARA TRABALHOS FUTUROS

O método proposto neste trabalho se mostrou promissor para as tarefas de detecção e mensuração de fissuras a partir de imagens de edificações e obras de construção civil capturadas por drones. Na primeira etapa do método, responsável pela detecção automática das fissuras, foram alcançadas precisão de 81,1%, *recall* de 69,7% e *mAP@50* de 71,1%, mas foi constatado espaço para aperfeiçoar a generalização do modelo YOLO, a partir de melhorias no conjunto de treinamento. A segunda etapa do método, responsável pela mensuração das fissuras segmentadas na etapa anterior, também se mostrou promissora, satisfazendo o objetivo de determinar a angulação predominante e o comprimento da fissura, inclusive convertendo o valor em pixels para unidade de comprimento do SI por meio do GSD, atingindo erro de 2,24% na comparação da medida real com a calculada por meio do GSD. No entanto, a disponibilidade das características técnicas da câmera e a correta calibração do sistema são essenciais para garantir a precisão dessa conversão.

O presente trabalho buscou avançar tanto no desenvolvimento de um método computacional integrado para detecção e mensuração automática de fissuras em edificações e obras civis, que representa uma importante contribuição tecnológica, quanto na consolidação do uso combinado de técnicas de IA, VC e drones como ferramentas de apoio a inspeções estruturais. As contribuições abrangem desde a aplicação de modelos modernos de detecção, como o YOLO, até a utilização de descritores e operações, como HOG, transformada da distância, esqueletização e cálculo do GSD, demonstrando o potencial dessas tecnologias para automatizar processos, reconhecer padrões estruturais e ampliar a precisão e eficiência das análises na engenharia civil. Do ponto de vista da Computação e Informática, o trabalho contribui ao propor e validar um *pipeline* computacional integrado que combina aprendizagem profunda, visão computacional clássica e procedimentos geométricos para solucionar um problema real da engenharia. Além disso, a avaliação empírica

com imagens reais e simulação de condições de voo demonstra a aplicabilidade prática do *pipeline*, contribuindo para o avanço de sistemas inteligentes embarcados, processamento de imagens em tempo real, e aplicações de IA orientadas à inspeção automatizada.

Como sugestões para trabalhos futuros propõe-se: a criação de uma base de imagens aberta de grande escala e diversificada, com vários tipos de acabamentos sobre a alvenaria, além de imagens de paredes e estruturas de concreto; implementação de detecção de outros tipos de manifestações patológicas, como fragmentação, corrosão e eflorescência, incluindo a utilização de termografia (multiespectral); implementação de novas técnicas para a etapa de mensuração, como PCA e transformação de coordenadas, de modo a tornar os resultados menos dependentes da posição do drone e da angulação da câmera; uma investigação aprofundada sobre as formas de mensuração da largura das fissuras; e por fim, implementação de localização espacial das fissuras e integração com BIM.

## REFERÊNCIAS

ALHEETI, K. M. A.; AI-ANI, M. S.; AI-ALOOSY, A. K. N.; ALZHRANI, A.; RUKAN, D. A. S. Intelligent mobile detection of cracks in concrete utilising an unmanned aerial vehicle. **Bulletin of Electrical Engineering and Informatics**, v. 11, n. 1, p. 176–184, 1 fev. 2022.

ANAC – AGÊNCIA NACIONAL DE AVIAÇÃO CIVIL. **Regulamento Brasileiro de Aviação Civil Especial nº 94/2017**. Disponível em: <https://www.anac.gov.br/assuntos/legislacao/legislacao-1/rbha-e-rbac/rbac/rbac-e-94>. Acesso em: 12 de maio de 2025.

ARTASANCHEZ, Alberto; JOSHI, Prateek. **Artificial Intelligence with Python: Your complete guide to building intelligent apps using Python 3.x**. 2. ed. Birmingham: Packt Publishing, 2020.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. ABNT NBR 15575:2013 – **Edificações habitacionais – Desempenho**. Rio de Janeiro: ABNT, 2013.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. ABNT NBR 6118:2023 – **Projeto de estruturas de concreto — Procedimento**. Rio de Janeiro: ABNT, 2023.

BOLAYBOLAY, J. M. A.; ALELUYA, E. R. M.; CLAR, S. E.; SALAAN, C. J. O. Surface Cracks Length Calculation in Segmented Images using Image Processing Techniques. **2022 IEEE 14th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)**, p. 1–6, 1 dez. 2022.

BOSCH, A.; ZISSERMAN, A.; MUÑOZ, X. Representing Shape with a Spatial Pyramid Kernel. In: **Proceedings of the 6th ACM International Conference on Image and Video Retrieval (CIVR)**. Amsterdam: ACM, 2007. p. 401–408.

BRAVO, D. T.; LIMA, G. A.; ALVES, W. A. L.; COLOMBO, V. P.; DJOGBENOU, L.; PAMBOUKIAN, S. V. D.; QUARESMA, C. C.; ARAÚJO, S. A. Automatic detection of potential mosquito breeding sites from aerial images acquired by unmanned aerial vehicles. **COMPUTERS ENVIRONMENT AND URBAN SYSTEMS**, v. 90, p. 101692, 2021.

CAPORRINO, Cristiana Furlan. **Patologias em alvenarias**. 2. ed. São Paulo: Oficina de Textos, 2018.

CHERPINSKI ENGENHARIA. **Você conhece um fissurômetro?** Cherpinski Engenharia, [s.d.]. Disponível em: <https://cherpinskiengenharia.com.br/avaliacao-de-imoveis/voce-conhece-um-fissurometro/>. Acesso em: 04 dez. 2025.

CHOI, D.; BELL, W.; KIM, D.; KIM, J. UAV-Driven Structural Crack Detection and Location Determination Using Convolutional Neural Networks. **Sensors**, v. 21, n. 8, p. 2650, 9 abr. 2021.

CHOI, Y.; PARK, H. W.; MI, Y.; SONG, S. Crack detection and analysis of concrete structures based on neural network and clustering. **Sensors**, v. 24, n. 6, p. 1725, 2024.

DALAL, N.; TRIGGS, B. Histograms of Oriented Gradients for Human Detection. In: **IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)**. San Diego: IEEE, 2005. p. 886–893.

DANAJITHA, K. K.; SHEEBA, A.; SOPHIYA, P.; VC, M. V. Detection of Cracks in High Rise Buildings using Drones. **2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC)**, 17 ago. 2022.

DAVIES, T.; KUMAR, R. S.; ANTONY, A. M. A Novel Method for Concrete Crack Detection Using Image Processing Technique. **2023 Advanced Computing and Communication Technologies for High Performance Applications (ACCTHPA)**, p. 1–8, 20 jan. 2023.

DUMITRU A.; DRĂGULINESCU A.; VIZIREANU N. Applications of mathematical morphology operators in civil infrastructures. **Earth Science Informatics**, v. 17, n. 5, p. 4027–4033, 24 jun. 2024.

EVERINGHAM, M.; VAN G., L.; WILLIAMS, C. K.; WINN, J.; Zisserman, A. The pascal visual object classes (voc) challenge. **International journal of computer vision**, v. 88, n. 2, p. 303-338, 2010.

FELZENSZWALB, P. F. et al. Object Detection with Discriminatively Trained Part-Based Models. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 32, n. 9, p. 1627–1645, 2010.

FILHO, Mario. **Mean Average Precision (MAP) em Machine Learning**. Mario Filho | Machine Learning, 17 jan. 2023. Disponível em: <https://mariofilho.com/mean-average-precision-map-em-machine-learning/>. Acesso em: 20 jun. 2025.

FLAH, M.; SULEIMAN, A. R.; NEHDI, M. L. Classification and quantification of cracks in concrete structures using deep learning image-based techniques. **Cement and Concrete Composites**, v. 114, art. 103781, nov. 2020.

FORCAEL, E.; ROMÁN, O.; STUARDO, H.; HERRERA, R. F.; SOTO-MUÑOZ, J. Evaluation of Fissures and Cracks in Bridges by Applying Digital Image Capture Techniques Using an Unmanned Aerial Vehicle. **Drones**, v. 8, n. 1, p. 8–8, 30 dez. 2023.

GAN, L.; Liu, H.; YAN, Y.; CHEN, A. Bridge bottom crack detection and modeling based on faster R-CNN and BIM. **IET Image Processing**, v. 18, n. 3, p. 664–677, 2 nov. 2023.

GAO, Y.; LIU, Y.; GUO, F.; JIA, X. UAV image and intelligent detection of building surface cracks. Paladyn, **Journal of Robotics, Intelligent Agents, and Artificial Intelligence**, v. 14, n. 1, p. 20220122, 2023.

GONZALEZ, R. C.; WOODS, R. E. **Digital Image Processing**. 4. ed. New York: Pearson, 2018.

HONKAVAARA, E.; MARKELIN, L.; AHOKAS, E.; KUITTINEN, R.; PELTONIEMI, J. Calibrating digital photogrammetric airborne imaging systems: test field calibration. **International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. XXXVI, B1, p. 1-8, 2007.

KARMOKAR, A.; JANI, N.; KALLA, A.; HARLALKA, H.; Sonar, P. Inspection of Concrete Structures by a Computer Vision Technique and an Unmanned Aerial Vehicle. 2020 **International Conference on Computational Performance Evaluation (ComPE)**, p. 338–343, jul. 2020.

KASAI, Murilo Shoji. **Por que surgem as fissuras?** Alicerce Engenharia Júnior, 2021. Disponível em: <https://www.alicerceejr.com/post/por-que-surgem-as-fissuras>. Acesso em: 12 jun. 2025.

KIM, H.; LEE, S.; AHN, E.; SHIN, M.; SIM, S. H. Crack identification method for concrete structures considering angle of view using RGB-D camera-based sensor fusion. **Structural Health Monitoring**, v. 20, n. 2, p. 500–512, 15 jul. 2020.

KO, P.; PRIETO, S. A.; BORJA. Developing a Free and Open-Source Semi-Automated Building Exterior Crack Inspection Software for Construction and Facility Managers. **IEEE Access**, v. 11, p. 77099–77116, 1 jan. 2023.

KO, P.; PRIETO, S. A.; DE SOTO, B. García. ABECIS: An automated building exterior crack inspection system using UAVs, open-source deep learning and photogrammetry. In: **ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction. IAARC Publications**, 2021. p. 637-644.

LI, H.; LI, Y.; XIAO, L.; ZHANG, Y.; CAO, L.; WU, D. RLRD-YOLO: An Improved YOLOv8 Algorithm for Small Object Detection from an Unmanned Aerial Vehicle (UAV) Perspective. **Drones**, v. 9, n. 4, p. 293, 10 abr. 2025.

LOWE, D. G. Distinctive Image Features from Scale-Invariant Keypoints. **International Journal of Computer Vision**, v. 60, n. 2, p. 91–110, 2004.

MA, X.; LI, Y.; ZHAO, W. Millimeter-level concrete crack detection network with layer aggregation capability. **Fourth International Conference on Image Processing and Intelligent Control (IPIC 2024)**, p. 36, 23 ago. 2024.

MIKHAIL, E. M.; BETHEL, J. S.; MCGLONE, J. C. **Introduction to Modern Photogrammetry**. New York: Wiley, 2001.

MOHAMMAD, F.; ALZOUBI, A.; DU, H.; JASSIM, S. A generic approach for automatic crack recognition in buildings glass facade and concrete structures. **Thirteenth International Conference on Digital Image Processing (ICDIP 2021)**, p. 70, 30 jun. 2021.

MOHER, D.; LIBERATI, A.; TETZLAFF, J.; ALTMAN, D. G.; PRISMA GROUP\*, T. Preferred reporting items for systematic reviews and meta-analyses: **The**

**PRISMA statement.** *International Journal of Surgery*, v. 8, n. 5, p. 336–341, 2010.

NOGUEIRA DINIZ, J. D. C.; PAIVA, A. C. D.; JUNIOR, G. B.; De ALMEIDA, J. D. S., SILVA, A. C.; CUNHA, A. M. T. D. S.; CUNHA, S. C. A. P. D. S. A Method for Detecting Pathologies in Concrete Structures Using Deep Neural Networks. *Applied Sciences*, v. 13, n. 9, p. 5763–5763, 7 maio 2023.

OH, S.; HAM, S.; LEE, S. Drone-Assisted Image Processing Scheme using Frame-Based Location Identification for Crack and Energy Loss Detection in Building Envelopes. *Energies*, v. 14, n. 19, p. 6359, 5 out. 2021.

PEDRINI, H.; SCHWARTZ, W. R. **Análise de imagens digitais: princípios, algoritmos e aplicações.** São Paulo: Cengage Learning, 2008.

PRABU, R. T.; ANITHA, G.; MOHANAVEL, V.; TAMILSELVI, M.; RAMKUMAR, G. Automated Crack and Damage Identification in Premises using Aerial Images based on Machine Learning Techniques. **2022 Sixth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)**, 10 nov. 2022.

PHILIP, R. E.; ANDRUSHIA, A. D.; NAMMALVAR, A.; GURUPATHAM, B. G. A.; ROY, K. A Comparative Study on Crack Detection in Concrete Walls Using Transfer Learning Techniques. *Journal of Composites Science*, v. 7, art. 169, 2023.

PRINCE, Simon. **Understanding Deep Learning.** Cambridge: The MIT Press, 2023. Disponível em: <http://udlbook.com>. Acesso em: 12 jun. 2025.

RASCHKA, S.; LIU, Y. (Hayden); MIRJALILI, Vahid. **Machine Learning with PyTorch and Scikit-Learn.** Birmingham: Packt Publishing, 2022.

REDMON, J.; DIVVALA, S.; GIRSHICK, R.; FARHADI, A. You Only Look Once: Unified, Real-Time Object Detection. In: **Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).** Las Vegas, NV: IEEE, 2016. p. 779–788.

REDMON, J.; FARHADI, A. **YOLOv3: An Incremental Improvement.** Disponível em: <<https://arxiv.org/abs/1804.02767>>.



REZENDE, S. O. **Sistemas Inteligentes: Fundamentos e Aplicações**. 4. ed. São Paulo: Manole, 2022.

RODRIGUES ORTEGA, M.; DE MARCO, G.; FLORIAN, F. ANÁLISE DAS PATOLOGIAS ESTRUTURAIS EM FUNDAÇÕES: CAUSAS, DIAGNÓSTICOS E SOLUÇÕES. **Revista Foco (Interdisciplinary Studies Journal)**, v. 17, n. 12, 2024.

SAHADE, R. F. **Estudo das fissuras em alvenarias de vedação e seus revestimentos**. 2005. 174 f. Dissertação (Mestrado em Engenharia Civil) – Escola Politécnica, Universidade de São Paulo, São Paulo, 2005.

SHIBANO, K.; MOROZOVA, N.; SHIMAMOTO, Y.; ALVER, N.; SUZUKI, T. Improvement of crack detectivity for noisy concrete surface by machine learning methods and infrared images. **Case Studies in Construction Materials**, v. 20, p. e02984, 2024.

SINGH, R. **Decoding CNNs: a beginner's guide to Convolutional Neural Networks and their applications**. Medium, 2024. Disponível em: <https://ravjot03.medium.com/decoding-cnns-a-beginners-guide-to-convolutional-neural-networks-and-their-applications-1a8806cbf536>. Acesso em: 12 jun. 2025.

SPIZHEVOY, A.; RYBNIKOV, A. **OpenCV 3 Computer Vision with Python Cookbook: Leverage the power of OpenCV 3 and Python to build computer vision applications**. Birmingham: Packt Publishing, 2018.

SZELISKI, Richard. **Computer Vision: Algorithms and Applications**. London: Springer, 2010.

THOMAZ, Ercio. **Trincas em edifícios: causas, prevenção e recuperação**. 2. ed. rev. e ampl. São Paulo: Oficina de Textos, 2020.

TRIDENT. **Calibre para medir fissuras 85 x 45 mm em PVC transparente – mod. FISS-04**. [S.l.]: Trident, [s.d]. Disponível em: <https://trident.com.br/produto/calibre-para-medir-fissuras-85-x-45-mm-em-pvc-transparente-mod-fiss-04/>. Acesso em: 12 jun. 2025.

ULTRALYTICS. **YOLOv8: Performance Metrics**. Ultralytics YOLO Docs, 2023a. Disponível em: <https://docs.ultralytics.com/pt/models/yolov8/#performance-metrics>. Acesso em: 13 jun. 2025.

ULTRALYTICS. **YOLOv8: supported tasks and modes**. **Ultralytics YOLO Docs**, 2023. Disponível em: <https://docs.ultralytics.com/pt/models/yolov8/#supported-tasks-and-modes>. Acesso em: 12 jun. 2025.

WANG, H.; SHI, Y.; YUAN, Q.; LI, M. Crack Detection and Feature Extraction of Heritage Buildings via Point Clouds: A Case Study of Zhonghua Gate Castle in Nanjing. **Buildings**, v. 14, n. 8, p. 2278, 2024.

WOLF, Paul R.; DEWITT, Bon A.; WILKINSON, Benjamin E. **Elements of photogrammetry with applications in GIS**. 4. ed. New York: McGraw-Hill Education, 2014.

ZHAO, H.; LI, J.; NIE, J.; GE, J.; YANG, S.; YU, L.; WANG, K. Identification Method for Cone Yarn Based on the Improved Faster R-CNN Model. **Processes**, v. 10, n. 4, p. 634, 24 mar. 2022.