UNIVERSIDADE ESTADUAL PAULISTA "JÚLIO DE MESQUITA FILHO"

FACULDADE DE ENGENHARIA

CAMPUS DE ILHA SOLTEIRA

Bruno Soares dos Santos

Detecção de Danos em Estruturas de Concreto Utilizando Fotografias Aéreas Obtidas por VANTs

Ilha Solteira

2024

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA MECÂNICA

Bruno Soares dos Santos

Detecção de Danos em Estruturas de Concreto Utilizando Fotografias Aéreas Obtidas por VANTs

Dissertação apresentada à Faculdade de Engenharia - UNESP - Campus de Ilha Solteira, em cumprimento parcial dos requisitos para o grau de Mestre em Engenharia Mecânica.

Orientador: Prof. Dr. Douglas Domingues Bueno

Ilha Solteira

2024

FICHA CATALOGRÁFICA
Desenvolvido pelo Serviço Técnico de Biblioteca e Documentação

Santos, Bruno Soares dos.
S237d Detecção de danos em estruturas de concreto utilizando fotografias aéreas obtidas por VANTs / Bruno Soares dos Santos. -- Ilha Solteira: [s.n.], 2024
70 f. : il.

Dissertação (mestrado) - Universidade Estadual Paulista. Faculdade de Engenharia de Ilha Solteira. Área de conhecimento: Mecânica dos Sólidos, 2024

Orientador: Douglas Domingues Bueno

Inclui bibliografia

1. Monitoramento da integridade estrutural. 2. Veículo Aéro Não Tripulado. 3. Processamento de imagens. 4. Inteligência artificial.

Amanda Sertori dos Santos
Bibliotecária - CRB/8-9061
Seção Técnica de Referência, Atendimento ao Usuário e Documentação
Diretoria Técnica de Biblioteca e Documentação

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Ilha Solteira

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: DETECÇÃO DE DANOS EM ESTRUTURAS DE CONCRETO UTILIZANDO FOTOGRAFIAS AÉREAS OBTIDAS POR VANT'S

AUTOR: BRUNO SOARES DOS SANTOS
ORIENTADOR: DOUGLAS DOMINGUES BUENO

Aprovado como parte das exigências para obtenção do Título de Mestre em null, área: Mecânica dos Sólidos pela Comissão Examinadora:

Prof. Dr. DOUGLAS DOMINGUES BUENO (Participaçao Virtual)
Departamento de Engenharia Mecânica / Faculdade de Engenharia de Ilha Solteira - UNESP

Profa. Dra. THAYS APARECIDA DE ABREU SANTOS (Participaçao Virtual)
Departamento de Matemática / Faculdade de Engenharia de Ilha Solteira - UNE

gov.br
Documento assinado digitalmente
THAYS APARECIDA DE ABREU SANTOS
Data: 14/08/2024 11:33:01-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. PEDRO CHRISTIAN AYALA CASTILLO (Participaçao Virtual)
Departamento de Engenharia Mecânica / Universidade Federal do Sul e Sudeste do Pará - UNIFESSPA

Ilha Solteira, 26 de julho de 2024

À minha mãe Reinilda Soares,
pois sem o seu apoio e carinho, nada disto seria possível.

AGRADECIMENTOS

Gostaria de expressar minha profunda gratidão a todas as pessoas que, de alguma forma, contribuíram para a realização deste trabalho.

Primeiramente, agradeço à minha família, especialmente aos meus pais, Reinilda Soares e Domingos Dantas, pelo amor e apoio incondicional. Aos meus queridos irmãos, Diego e Thiago Soares, pelo incentivo e companhia em todos os momentos.

Aos meus colegas do Grupo de Materiais e Sistemas Inteligentes (GMSINT) e do Laboratório de Sistemas Avançados (Lab-SA), Renan Sanches, Bianca Carnielo, Lorena L. Dias, Renan Cavenaghi, Vítor Marchiori e Samuel Souza, pelo suporte e amizade durante esta jornada.

Aos professores Carlos R. dos Santos Jr., Camila G. Gonsalez Bueno, Rodrigo B. Santos, Leandro Salviano e Samuel da Silva, por seu apoio e colaboração durante o meu mestrado.

Um agradecimento especial à professora Ruxandra M. Botez, que me orientou durante meu estágio no Canadá, proporcionando uma experiência enriquecedora e indispensável para meu crescimento profissional.

Aos amigos que estiveram comigo no Canadá, Ana Carolina Rodrigues, Caio Marcon, Aline P. Marçal, Marco A. Assis e Elson Correia Jr. , pela amizade, apoio e momentos inesquecíveis que compartilhamos.

Na vida, há poucas pessoas que são verdadeiramente essenciais e que nos ajudam a abrir portas importantes. Entre elas, destaco o professor orientador Douglas D. Bueno. Sua orientação, amizade e paciência foram fundamentais para a concretização deste trabalho, e por isso, sou imensamente grato.

Agradeço à Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP) e à Faculdade de Engenharia de Ilha Solteira (FEIS), pelo apoio, infraestrutura e oportunidades que me foram apresentadas. A cada professor, colaborador e aluno que contribuíram para meu desenvolvimento acadêmico e profissional. Em especial, aos professores do Programa de Pós-Graduação em Engenharia Mecânica.

O presente trabalho foi realizado com o apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq. Processo número 403955/2022-0.

A todos, meu sincero agradecimento.

"Embora ninguém possa voltar atrás e fazer um novo começo, qualquer um pode começar agora e fazer um novo fim."

James R. Sherman

RESUMO

Este estudo apresenta uma abordagem para a inspeção de obras de arte especiais (OAE), como pontes, utilizando veículos aéreos não tripulados (VANTs) e técnicas avançadas de Inteligência Artificial. O objetivo é aprimorar a segurança e reduzir custos na manutenção de infraestruturas de concreto. Desenvolveu-se uma metodologia detalhada de captura de imagens, garantindo a obtenção de dados precisos e confiáveis. O método proposto emprega um banco de dados de imagens representativo, com 19.659 imagens de estruturas com e sem danos, para treinar algoritmos eficientes na detecção de fissuras e outros defeitos. A investigação dos métodos de pré-processamento revela que o filtro Gaussiano apresenta o melhor desempenho em aumentar a eficiência dos métodos de classificação. Na detecção de bordas, o método Sobel se destaca. As Redes Neurais Convolucionais (CNN) foram utilizadas na identificação de trincas, demonstrando uma acurácia de 90,31%, F1-Score de 90,50% e precisão de 91,00%, superando técnicas tradicionais de classificação, como Máquinas de Vetores de Suporte (SVM) e Regressão Logística. A análise da influência da quantidade de imagens reforçou a importância de uma distribuição equilibrada para otimizar o desempenho do modelo. Como resultado final, tem-se a elaboração de códigos computacionais para a análise automática das imagens capturadas e a criação de um protocolo de vistoria aérea otimizado, visando a manutenção eficiente e precisa de infraestruturas críticas.

Palavras-chave: monitoramento da integridade estrutural; veículo aéreo não tripulado; processamento de imagens; inteligência artificial.

ABSTRACT

This study presents an approach for the inspection of engineering structures, such as bridges, using unmanned aerial vehicles (UAVs) and advanced artificial intelligence techniques. The objective is to enhance safety and reduce costs in the maintenance of concrete structures. A detailed image capture methodology was developed, ensuring the acquisition of precise and reliable data. The proposed method employs a representative image database, with 19,659 images of structures with and without damage, to train efficient algorithms in the detection of cracks and other defects. The investigation of preprocessing methods reveals that the Gaussian filter presents the best performance in increasing the efficiency of classification methods. In edge detection, the Sobel method stands out. Convolutional Neural Networks (CNN) were used in the identification of cracks, demonstrating an accuracy of 90.31%, an F1-Score of 90.50%, and a precision of 91.00%, surpassing traditional classification techniques such as Support Vector Machines (SVM) and Logistic Regression. The analysis of the influence of the number of images reinforced the importance of a balanced distribution to optimize the model's performance. As a final result, computational codes were developed for the automatic analysis of the captured images and the creation of an optimized aerial inspection protocol, aiming at efficient and precise maintenance of critical structures.

Keywords: structural integrity monitoring; unmanned aerial vehicle; image processing; artificial intelligence.

LISTA DE FIGURAS

1 Fluxograma do processo de pré-processamento de imagens. 25
2 Modelos de bordas: degrau, rampa e telhado, da esquerda para a direita. . . . 30
3 Experimento de captura de imagem com VANT. 43
4 Imagem composta ilustrando quatro configurações diferentes: variações em altitude, presença de sombra, ângulo de inclinação da câmera e velocidade de voo. 44
5 Corpos de prova segmentados em regiões de interesse. 44
6 Fluxograma do processo de avaliação dos filtros de suavização. 45
7 Resultados das Métricas de Regressão Logística com Diferentes Filtros de Suavização 46
8 Acertos e Erros em Percentagem para Diferentes Filtros de Suavização 46
9 Acurácia Média e Acurácia Geral para Diferentes Filtros de Suavização 47
10 Curvas ROC para Regressão Logística com Diferentes Filtros de Suavização . . 47
11 Fluxograma do processo de avaliação das técnicas de detecção de bordas. . . . 48
12 Comparação de Acurácia Geral e Média para Diferentes Métodos de Detecção de Borda 50
13 Soma de Acertos e Erros dos Métodos de Detecção de Bordas 51
14 Especificidade dos Métodos de Detecção de Bordas 51
15 Sensibilidade dos Métodos de Detecção de Bordas 52
16 Comparação das Curvas ROC para Diferentes Filtros de Suavização 52
17 Comparação de métricas de avaliação entre CNN, Regressão Logística e SVM . 54
18 Acertos e Erros em Percentagem para Diferentes Métodos de Classificação . . . 55
19 Comparação das Curvas ROC para CNN, SVM e Regressão Logística 55
20 Comparação das Métricas de Desempenho entre os *Datasets* 56
21 Média de Acertos e Erros em Percentagem para Diferentes Conjuntos de Dados 58
22 Comparação da Acurácia Geral e Acurácia Média 58
23 Comparação das Curvas ROC para Diferentes *Datasets*. 59
24 Curvas de Aprendizado da Acurácia Comparativas para Três *Datasets*. 59
25 Curvas de Aprendizado da Perda Comparativas para Três *Datasets*. 60

LISTA DE TABELAS

1 Classificação das fissuras. 22
2 Matriz de Confusão 23

3 Configurações de voo utilizadas no experimento. 42
4 Matriz de Confusão em Percentagem para SVM com Diferentes Filtros de Suavização 46
5 Desempenho dos Métodos de Detecção de Bordas 49
6 Matrizes de Confusão com Percentuais para Diferentes Métodos 50
7 Detalhes da Arquitetura CNN 53
8 Comparação da Matriz de Confusão em Percentagem 54
9 Matriz de Confusão em Percentagens para Todos os Conjuntos de Dados 57

LISTA DE SÍMBOLOS

$\mathbf{\Delta}$	Número de pixeis deslocados devido ao efeito *rolling shutter*
μ_1	Média dos níveis de cinza na área de interesse
μ_2	Média dos níveis de cinza na área de fundo
μ_2	Média dos níveis de cinza na área de fundo
σ	Desvio padrão da distribuição gaussiana
σ_1	Variação padrão das intensidades de pixel na área de interesse
σ_2	Variação padrão das intensidades de pixel na área de fundo
θ	Ângulo de orientação do vetor gradiente
B	Canal de Cor Negativo
d	Diâmetro da abertura da lente em mm
E	Exposição em Joules (J) mm^{-2}
e	Base do logaritmo natural, aproximadamente igual a 2.71828
FN	Falso Negativo
FP	Falso Positivo
FR	Distância focal
F	Imagem fonte
f	Comprimento focal da lente em mm
G	Canal de Cor Verde
$\mathbf{G}_x$	Matriz gradiente com orientação horizontal
$\mathbf{G}_y$	Matriz gradiente com orientação vertical

$\mathbf{h}$	Altura do voo
$\mathbf{Im_h}$	Altura da imagem em pixeis
$\mathbf{i}$	Distância entre a lente e o plano da imagem
$\mathbf{m}_1$	Média de intensidade para os pixels do conjunto $\mathbf{G}_1$
$\mathbf{m}_2$	Média de intensidade para os pixels do conjunto $\mathbf{G}_2$
$\mathbf{o}$	Distância entre a lente e o objeto a ser fotografado
$\mathbf{p}$	Ponto específico na imagem cuja intensidade é analisada
$\mathbf{P}_1$	Probabilidade de aparição dos pixels na área de interesse
$\mathbf{P}_2$	Probabilidade de aparição dos pixels na área de fundo
$\mathbf{R}$	Canal de Cor Vermelho
$\mathbf{S_h}$	Altura do sensor
$\mathbf{s}$	Brilho da cena intrínseca em J mm^{-2} sec^{-1}
$\mathbf{T_o}$	Tempo de abertura do obturador
$\mathbf{T}$	Limiar ótimo
$\mathbf{t}$	Tempo de exposição em segundos
VN	Verdadeiro Negativo
VP	Verdadeiro Positivo
$\mathbf{V}$	Velocidade do VANT
$\mathbf{v}$	Área local em volta de um ponto na imagem

SUMÁRIO

1 INTRODUÇÃO **15**

1.1 OBJETIVOS DO TRABALHO 16

1.2 CONTRIBUIÇÕES ALCANÇADAS 17

1.3 ORGANIZAÇÃO DO TRABALHO 17

2 METODOLOGIA GERAL E FUNDAMENTOS **18**

2.1 ASPECTOS PRINCIPAIS RELATIVOS ÀS FOTOGRAFIAS AÉREAS 19

2.1.1 Fotogrametria 19

2.1.2 Exposição à Luz 19

2.1.3 Sensibilidade ISO 19

2.1.4 Distância Focal 20

2.1.5 Abertura da Lente 20

2.1.6 Velocidade do Obturador 20

2.2 DANOS EM ESTRUTURAS DE CONCRETO 21

2.3 MÉTRICAS DE AVALIAÇÃO 21

2.4 PRÉ-PROCESSAMENTO DE IMAGENS 24

2.4.1 Métodos de Conversão para Escala de Cinza 25

2.4.2 Filtros de Suavização de Ruídos 25

2.4.3 Detecção de Borda 29

3 MÉTODOS DE CLASSIFICAÇÃO UTILIZANDO INTELIGÊNCIA ARTIFICIAL **36**

3.1 REGRESSÃO LOGÍSTICA EM INTELIGÊNCIA ARTIFICIAL 36

3.2 MÉTODO *SUPPORT VECTOR MACHINE* (SVM) 37

3.3 MÉTODO DE REDES NEURAIS CONVOLUCIONAIS (CNNS) 38

3.3.1 Ativação Sigmoide 38

3.3.2 Dropout 39

4 RESULTADOS E DISCUSSÕES **40**
4.1 CRIAÇÃO DO *DATASET* COM IMAGENS CAPTURADAS POR VANTS . . 41
4.2 ANÁLISE COMPARATIVA DE FILTROS DE SUAVIZAÇÃO 45
4.3 AVALIAÇÃO DE TÉCNICAS DE DETECÇÃO DE BORDAS 48
4.4 COMPARAÇÃO DE MÉTODOS DE INTELIGÊNCIA ARTIFICIAL 53
4.5 ANÁLISE DA QUANTIDADE DE IMAGENS PARA DETECÇÃO DE TRINCAS 56

5 CONCLUSÃO **61**
5.1 SUGESTÕES PARA TRABALHOS FUTUROS 62

REFERÊNCIAS **63**

1 INTRODUÇÃO

A detecção de danos em estruturas de concrete, especialmente as de grande porte, tipicamente envolve importante esforço de engenharia, principalmente devido a dificuldades de acessar regiões específicas. Em particular, o monitoramento de prédios,viadutos e pontes, geralmente envolve avaliações visuais, realizadas por profissionais experientes, implicando em reduzida produtividade. Por outro lado, o crescente uso de VANTs (Veículos Aéreos Não-tripulados), em diversas aplicações de engenharia, tem demonstrado um importante potencial, especialmente com foco em redução de riscos relativos à segurança de equipes, bem como por redução de custos operacionais. Em particular, o uso de VANTs para detectar defeitos em estrutura de concreto oferece interessante vantagem, apesar de questões técnicas, em termos de planejamento e execução de voos para obtenção de imagens das estruturas, bem como pelo potencial de aplicação de técnicas de pós-processamento de imagens com foco em detecção de danos.

Neste contexto, Mohan e Poobal (2018) destacam a importância da detecção automática de fissuras em concreto como alternativa à inspeção manual, que é limitada pela subjetividade e experiência do inspetor. Eles realizaram uma revisão abrangente de técnicas de processamento de imagem para identificar desafios e conquistas nessa área. Flah, Suleiman e Nehdi (2020) propõem um modelo de inspeção quase automatizado para estruturas de concreto, combinando processamento de imagem e aprendizado profundo, alcançando alta precisão na classificação de defeitos. Ruiz *et al.* (2021) exploram a aplicação de técnicas de aprendizado profundo na detecção automática de fissuras em revestimentos cerâmicos, demonstrando viabilidade técnica e benefícios em termos de custo e segurança. Dais *et al.* (2021) examinam técnicas de aprendizado profundo para detecção de fissuras em muros de alvenaria, atingindo alta precisão e pontuação F1. Zhang (2014) desenvolvem um algoritmo de detecção de fissuras em superfícies de vidro baseado em processamento de imagem digital, obtendo informações de identificação precisas para avaliar a integridade estrutural.

Kim *et al.* (2018) focam na inspeção de pontes usando VANTs com sensores de visão de alta performance. Eles propuseram um método de identificação de fissuras em uma ponte de concreto envelhecida, utilizando aprendizado profundo e transferência de aprendizado, com resultados promissores. Hoang e Nguyen (2018) desenvolvem um método de processamento de imagens para detecção automática de fissuras em paredes de concreto, utilizando algoritmos de detecção de bordas e otimização meta-heurística, alcançando alta precisão. Simões, Rodrigues e Pinheiro (2021) discutem a importância dos VANTs em diversas áreas, enfatizando a necessidade de planejamento de rotas confiáveis para garantir a segurança nas operações de inspeção. Lei *et al.* (2018) abordam as limitações das técnicas tradicionais de detecção de fissuras em imagens coletadas por VANTs, propondo um novo método baseado no ponto central da fissura para identificação rápida e precisa. Reagan, Sabato e Niezrecki (2018) exploram uma abordagem inovadora que combina VANTs com correlação de imagem digital tridimensional para monitorar

a saúde de pontes, demonstrando a capacidade de detectar mudanças estruturais com alta precisão.

Lee *et al.* (2021) propõem uma metodologia para lidar com problemas de qualidade de imagem em inspeções de pontes usando VANTs, destacando técnicas para avaliação e melhoria da qualidade da imagem. Zhang *et al.* (2020) apresentam uma nova abordagem para a identificação de fissuras superficiais em áreas de mineração utilizando imagens de VANT, empregando algoritmos de aprendizado de máquina e técnicas avançadas de processamento de imagem. Simões, Rodrigues e Pinheiro (2021) oferecem uma análise crítica das metodologias de inspeção de pontes, ressaltando a necessidade de melhorias para uma avaliação mais objetiva e eficaz. Fujita, Mitani e Hamamoto (2006) e Cha, Choi e Büyüköztürk (2017) contribuem para essa discussão com métodos inovadores para detecção automática de fissuras em estruturas de concreto, empregando algoritmos avançados e redes neurais convolucionais para aumentar a precisão e eficiência das inspeções.

Neste contexto, esta dissertação apresenta uma metodologia integrada de monitoramento de integridade de estruturas de concreto utilizando imagens obtidas por VANTs e técnicas de Inteligência Artificial. A criação de um banco de dados (ou *dataset*, em inglês) de alta qualidade com 19.659 imagens, balanceando a quantidade de imagens com e sem trincas, é empregada para o treinamento dos modelos. A investigação abrangeu desde a coleta de dados com VANTs até a aplicação de técnicas de Inteligência Artificial para a detecção de trincas em estruturas de concreto. No pré-processamento, diferentes filtros são avaliados para melhorar a qualidade das imagens, enquanto diversas técnicas de detecção de bordas foram testadas para identificar as trincas de forma eficaz. A análise comparativa dos métodos de Inteligência Artificial, incluindo Redes Neurais Convolucionais (CNN), Máquinas de Vetores de Suporte (SVM) e Regressão Logística, visou determinar a técnica mais eficiente para a classificação das imagens. A influência da quantidade de imagens do conjunto de dados também foi estudada para otimizar o desempenho dos modelos. Esses esforços integrados são fundamentais para aprimorar a precisão e a eficiência na detecção de trincas, contribuindo para a segurança e durabilidade das infraestruturas, além de melhorar os métodos de monitoramento da integridade estrutural (em inglês *Structural Health Monitoring* -SHM).

1.1 OBJETIVOS DO TRABALHO

O objetivo geral desta pesquisa é desenvolver uma metodologia de monitoramento de integridade de estruturas de concreto utilizando imagens obtidas por VANTs.

Destacam-se entre os objetivos específicos:

- Determinar uma adequada sequência de pré-processamento de imagens;
- Aplicar uma técnica de inteligência artificial de segmentação de imagens para detecção de danos estruturais;
- Avaliar como se relacionam as características da câmera fotográfica embarcada com a condição de voo, visando obter imagens aéreas com qualidades adequadas.

1.2 CONTRIBUIÇÕES ALCANÇADAS

A principal contribuição desta pesquisa de mestrado é a proposição de um método vistoria e inspeção de estruturas de concreto ou obras de arte especiais (OAE).

1.3 ORGANIZAÇÃO DO TRABALHO

A presente seção apresenta a organização deste documento. Destacam-se a motivação geral, objetivos, metodologia utilizada, resultados obtidos e as considerações finais, conforme destaco a seguir:

- **Capítulo 1: Introdução**
  - Apresentação da área de estudo, importância e motivações. Definição dos objetivos da pesquisa.
- **Capítulo 2: Metodologia**
  - Descrição da metodologia de pesquisa e revisão da literatura relevante.
- **Capítulo 3: Classificação com Inteligência Artificial**
  - Análise de métodos de classificação via inteligência artificial em imagens de VANTs.
- **Capítulo 4: Resultados e Discussões**
  - Apresentação e análise dos resultados, discussões com base na metodologia e teoria.
- **Capítulo 5: Conclusão**
  - Resumo das descobertas, contribuições e limitações.

2 METODOLOGIA GERAL E FUNDAMENTOS

Este capítulo delineia a metodologia empregada no estudo, cujo objetivo central é investigar, analisar e comparar diferentes métodos de detecção de trincas em estruturas de concreto utilizando imagens capturadas por Veículos Aéreos Não Tripulados (VANTs). A pesquisa abrange desde a coleta de um vasto conjunto de imagens até a aplicação de técnicas avançadas de processamento de imagens e inteligência artificial para a análise e classificação dos dados visuais. O capítulo também descreve os fundamentos teóricos subjacentes ao estudo, proporcionando uma base sólida para as metodologias empregadas.

As principais etapas metodológicas incluem:

1. **Coleta de Dados:** Utilização de VANTs para a aquisição de imagens de estruturas de concreto, com o intuito de obter um conjunto de dados abrangente e representativo. Foram realizadas variações nas condições de captura de imagens, incluindo altitude, luz, velocidade do VANT e ângulo de inclinação da câmera.

2. **Preparação do Conjunto de Dados:** Organização e categorização das imagens em três grupos - sem trinca, com trinca pequena e com trinca grande - para facilitar a análise subsequente e o treinamento de modelos de inteligência artificial. Técnicas de aumento de dados foram aplicadas para expandir o conjunto de dados.

3. **Pré-processamento de Imagens:** Aplicação de técnicas de processamento de imagens, como a conversão para escala de cinza e o uso de filtros de suavização e de borda, incluindo os filtros Canny, Laplacian, Log, Zero Crossing, Prewitt, Robert, Sobel, Kirsch e Marr-Hildreth.

4. **Seleção e Avaliação de Métodos de Inteligência Artificial:** Exploração e comparação de diferentes algoritmos de inteligência artificial, incluindo Redes Neurais Convolucionais (CNN), Máquinas de Vetores de Suporte (SVM) e Regressão Logística, para determinar o mais eficiente na classificação das imagens.

5. **Análise da quantidade de imagens de Dataset:** Investigação da influência da quantidade do conjunto de dados nos resultados obtidos, avaliando a precisão e a taxa de detecção de trincas em diferentes tamanhos de conjuntos de dados.

6. **Análise dos Resultados:** Avaliação crítica dos resultados obtidos para medir a efetividade das técnicas de pré-processamento e dos algoritmos de inteligência artificial na detecção de trincas. A análise inclui a comparação de métricas de desempenho, como acurácia, precisão, *recall*, F1-Score, especificidade e curvas ROC.

2.1 ASPECTOS PRINCIPAIS RELATIVOS ÀS FOTOGRAFIAS AÉREAS

Este seção discute os fatores essenciais para obter imagens de alta qualidade na detecção de trincas, incluindo fotogrametria, exposição à luz, sensibilidade ISO, distância focal, abertura da lente e velocidade do obturador.

2.1.1 Fotogrametria

A fotogrametria mede imagens para obter informações sobre a posição, tamanho e forma de objetos, ideal para reconstruções tridimensionais. A 'fotogrametria de alcance próximo' é usada quando objetos têm menos de 100 metros e câmeras estão próximas. Ela permite a criação de modelos digitais precisos e gráficos, como imagens e mapas, armazenando informações acessíveis a qualquer momento (Atkison, 1996; Luhman *et al.*, 2011).

2.1.2 Exposição à Luz

Capturar imagens precisas de trincas em estruturas com VANTs pode ser desafiador. Para obter uma imagem nítida e bem exposta, é essencial alcançar as condições ideais de exposição de luz. Segundo Afifi *et al.* (2021), a exposição da luz afeta diretamente a qualidade da imagem. Para controlar a exposição, as câmeras digitais utilizam três fatores: velocidade do obturador, valor ISO e a abertura da lente, conhecidos como o triângulo de exposição.

De acordo com Hall (2015), pequenas mudanças em qualquer um desses elementos impactam como os outros dois trabalham para expor o sensor da imagem à luz, alterando significativamente a aparência final da imagem. Conforme Lillesand, Kiefer e Chipman (2015), a exposição em um ponto específico do plano de imagem de uma câmera é determinada pelo produto da irradiância naquele ponto pela duração da exposição, expressa pela equação:

$$E = \frac{sd^2t}{4f^2} \tag{1}$$

onde E é a exposição em Joules (J) mm^{-2}, s é o brilho da cena intrínseca em J mm^{-2} sec^{-1}, d é o diâmetro da abertura da lente em mm, t é o tempo em segundos, e f é o comprimento focal da lente em mm.

2.1.3 Sensibilidade ISO

ISO, abreviação de *International Organization for Standardization*, é um indicador de sensibilidade à luz usado em filmes e sensores de câmeras digitais. Ajustar o ISO não altera a quantidade de luz que entra na câmera, mas sim a sensibilidade do sensor à luz (Judge, 2012). A luz entra pela abertura da lente, passa pelo obturador e chega ao sensor de imagem, cuja sensibilidade é controlada pelo ISO (Youn, 2016).

O valor do ISO afeta a luminosidade e a cor das imagens capturadas por VANTs. Valores altos de ISO aumentam a fotossensibilidade, mas também aumentam as manchas nas imagens. Devido às variações de iluminação durante o voo do VANT, é necessário configurar o ISO

no modo automático (Hlotov; Hunina; Siejka, 2017). Configurações ISO geralmente variam de baixo (abaixo de 100) a alto (acima de 800), com valores mais baixos indicando menor sensibilidade à luz (Vecellio; Bryant, 2018).

2.1.4 Distância Focal

A distância focal é a medida da distância entre a lente e o sensor da câmera, que pode ser ajustada para alterar o ângulo de visão e o zoom da imagem, variando desde uma lente grande angular até uma lente telefoto estreita. Essas mudanças na distância focal podem gerar diferentes graus de distorção na imagem capturada (Třebický *et al.*, 2016). Para Liu e Mason (2019), a distância focal das lentes é um fator crucial para determinar as características da distorção da projeção central em um sistema de imagem óptica. Se a área de imagem (campo de visão) for a mesma, uma distância focal menor resultará em uma maior distorção topográfica. Lillesand, Kiefer e Chipman (2015) acrescenta, existem três parâmetros fundamentais para o foco de uma câmera: a distância focal da lente, f, a distância entre a lente e o objeto a ser fotografado, o, e a distância entre a lente e o plano da imagem, i. A relação entre essas distâncias é indicada na equação 2.

$$\frac{1}{f} = \frac{1}{o} + \frac{1}{i} \tag{2}$$

Quando a câmera está corretamente focada, a relação entre a distância focal, a distância do objeto e a distância da imagem é igual a 1.

2.1.5 Abertura da Lente

A abertura da lente da câmera é um orifício ajustável que permite a passagem de luz, similar à íris do olho humano. Ela pode ser ajustada para diferentes níveis de luz, influenciando a quantidade de luz e fótons que atingem o sensor e impactando na qualidade da imagem (Youn, 2016; Ghassoun *et al.*, 2021). Aberturas pequenas podem limitar a nitidez devido à difração, enquanto aberturas grandes podem causar aberrações. Para inspeções ao ar livre, uma abertura menor é preferível, pois melhora a profundidade de campo e reduz o desfoque (Zhang *et al.*, 2020). O valor do *f-stop*, que controla a abertura, é calculado pela divisão da distância focal pelo diâmetro da abertura (Aber; Marzolff; Ries, 2010):

$$f\text{-}stop = \frac{\text{Distância focal}}{\text{Diâmetro de abertura}} = \frac{f}{d} \tag{3}$$

Comercialmente, as lentes operam em faixas de $f/1$ a $f/32$. Conforme a equação 3, a exposição diminui com o aumento do *f-stop* porque a área da abertura é proporcional ao quadrado do diâmetro, influenciando a exposição conforme a raiz quadrada do *f-stop* (Lillesand; Kiefer; Chipman, 2015).

2.1.6 Velocidade do Obturador

O obturador é um dispositivo mecânico que cobre o sensor ou o filme, exceto quando eles são expostos à luz, e a velocidade do obturador, ou *shutter speed*, determina a quantidade de tempo que o filme ou sensor é exposto à luz (Judge, 2012). Conforme Majumder e Gopi (2018), o

tempo que o obturador permanece aberto influencia linearmente a quantidade de luz que entra na câmera. Geralmente, o obturador permanece aberto por apenas uma fração de segundo, como 1/30 ou 1/60. No entanto, se algum objeto na cena se mover durante esse período, sua imagem será capturada em vários locais, resultando em um efeito conhecido como desfoque de movimento, ou efeito *rolling shutter*.

2.1.6.1 Efeito Rolling Shutter

O efeito *rolling shutter* causa distorções em imagens, já que cada linha de pixeis é exposta em momentos diferentes, resultando em artefatos visuais, especialmente em objetos em movimento (Gu *et al.*, 2010). Para evitar imagens desfocadas com VANTs, é importante ajustar a velocidade do obturador conforme a velocidade do VANT e as condições de vento (Samad *et al.*, 2013). Imagens distorcidas podem comprometer a qualidade das ortofotos geradas por softwares de fotogrametria, dificultando a correspondência de pontos-chave e resultando em deformações (Figueiredo; Figueiredo, 2018).

A obtenção de imagens de trincas pode ser afetada por vibrações e movimentos do VANT. Aumentar a velocidade do obturador pode reduzir o desfoque, mas também diminuir a exposição, exigindo uma velocidade de obturador ideal (Kim *et al.*, 2017). Antes de cada voo, deve-se estimar o efeito de desfoque de movimento com a seguinte equação adaptada (Figueiredo; Figueiredo, 2018):

$$\Delta \approx V \times h \times \mathrm{FR} \times (\frac{Im_h}{S_h}) \times T_o \tag{4}$$

onde Δ é o número de pixeis deslocados, V é a velocidade do VANT, h é a altura do voo, FR é a distância focal, Im_h é a altura da imagem em pixeis, S_h é a altura do sensor, e T_o é o tempo de abertura do obturador. Valores de Δ acima de 1,2 pixeis podem dificultar o processamento e resultar em ortofotos de qualidade inferior. Ajustar a altura e velocidade de voo pode ajudar a reduzir Δ (Figueiredo; Figueiredo, 2018).

2.2 DANOS EM ESTRUTURAS DE CONCRETO

Danos em estruturas de concreto são indicativos de falhas que podem comprometer a integridade e segurança das construções. Inspeções regulares são cruciais para a gestão eficaz de pontes, pois as medidas de gerenciamento dependem das informações coletadas (Simões; Rodrigues; Pinheiro, 2021). O avanço da construção civil trouxe consigo riscos inerentes e a necessidade de compreender a deterioração estrutural (Souza; Ripper, 2009). O estudo das patologias das estruturas é de crescente importância, diretamente ligado à segurança e desempenho das edificações (Matildes, 2022). A classificação das fissuras, como descrita na Tabela 1, é um elemento chave na avaliação da condição das estruturas de concreto.

2.3 MÉTRICAS DE AVALIAÇÃO

A avaliação de do métodos de detecção de danos requer uma série de métricas que proporcionem entendimento sobre sua eficácia. Essas métricas descrevem a habilidade do algoritmo identificar

Tabela 1 – Classificação das fissuras.

Tipo	**Abertura (mm)**
Fissura capilar	$< 0{,}2$
Fissura	0,2 - 0,5
Trinca	0,5 - 1,5
Rachadura	1,5 - 5,0
Fenda, greta ou fincha	5,0 - 10,0
Brecha	$> 10{,}0$

Fonte: Adaptada de Ibape (2021).

e classificar corretamente os de dados. Para uma compreensão abrangente, examinamos as métricas empregadas são examinadas da seguinte forma:

- **Precisão:** Indica a concordância das etiquetas de classe dos dados com aquelas de um classificador (conforme ilustrado nas equações apresentadas por Obi (2023)). A precisão quantifica o número de predições positivas corretas e indica a pureza das predições positivas. A precisão é dada por:

$$\text{Precisão} = \frac{VP}{VP + FP} \tag{5}$$

  onde VP e FP representam as contagens de verdadeiros positivos e falsos positivos, respectivamente.

- ***Recall*:** Também conhecida como sensibilidade, é a eficácia de um classificador para identificar etiquetas de classe, como deduzido pela soma de decisões por texto (Obi, 2023). A *recall* expressa a fração do total de instâncias relevantes que são recuperadas, sendo dado por:

$$\text{Recall} = \frac{VP}{VP + FN} \tag{6}$$

  onde FN representa a contagem de falsos negativos.

- ***F1 Score*:** A média harmônica de precisão e *recall*. Esta métrica é útil pois oferece uma visão mais completa do que a acurácia, especialmente quando as classes são desbalanceadas (Sokolova; Lapalme, 2009). Assim, tem se:

$$\text{F1 Score} = 2 \times \frac{\text{Precisão} \times \text{Recall}}{\text{Precisão} + \text{Recall}} \tag{7}$$

- **Suporte:** Representa o número de ocorrências reais da classe no conjunto de dados. É crucial para avaliar a relevância da métrica de desempenho em relação à frequência real da etiqueta.

- **Acurácia:** A proporção de predições que um modelo acerta, fornecendo uma eficácia média para todas as classes. Isso é matematicamente definido pela razão da observação corretamente prevista para as observações totais (Obi, 2023). Conforme Sokolova e

Lapalme (2009), a fórmula para acurácia em classificação binária é:

$$\text{Acurácia} = \frac{VP + VN}{VP + VN + FP + FN} \tag{8}$$

e para classificação multi classes, a eficácia média por classe é representada como:

$$\text{Acurácia Média} = \frac{1}{L} \sum_{i=1}^{L} \frac{VP_i}{VP_i + FN_i} \tag{9}$$

onde L é o número de classes.

- **Especificidade:** Conforme afirmado por Gorunescu (2011), a especificidade mede a proporção de 'negativos' verdadeiros que são corretamente identificados, e é definida como:

$$\text{Especificidade} = \frac{VN}{VN + FP} \tag{10}$$

onde VN representa a contagem de verdadeiros negativos.

- **Matriz de Confusão:** Uma ferramenta crucial na aprendizagem de máquina e na classificação de dados, fornecendo uma visão detalhada do desempenho do modelo ao contrastar as classificações reais e as previstas (Obi, 2023). Esta matriz é vital para a avaliação de modelos de classificação, proporcionando informações valiosas sobre suas taxas de acerto e erro. Para uma classificação binária, a matriz de confusão é matematicamente representada conforme a Tabela 2:

Tabela 2 – Matriz de Confusão

	Previsto Positivo	**Previsto Negativo**
Real Positivo	VP	FN
Real Negativo	FP	VN

Fonte: Próprio autor.

Esta matriz também pode ser aplicada em problemas de classificação multi classe, resultando em uma tabela $N \times N$, onde N é o número de classes (Deng *et al.*, 2016).

- **Análise da Curva ROC e AUC**

A análise da Curva de Característica de Operação do Receptor (ROC) e da Área Sob a Curva (AUC) é essencial para avaliar o desempenho de classificadores em problemas de classificação binária (Joshi, 2020). A curva ROC plota a Taxa de Verdadeiros Positivos (TPR) contra a Taxa de Falsos Positivos (FPR), que são definidas como:

$$TPR = \frac{TP}{TP + FN} \tag{11}$$

$$FPR = \frac{FP}{FP + TN} \tag{12}$$

Estudos de Joshi (2020), Chadha. e Juwe. (2021) e Facelli *et al.* (2011) exploram

como a variação do limiar de decisão afeta o TPR e o FPR, destacando a utilidade da curva ROC para entender o desempenho do classificador na separação das duas classes, independentemente do limiar escolhido.

A AUC, ou Área Sob a Curva ROC, mede a capacidade do classificador de discriminar entre classes positivas e negativas, variando de 0 a 1, com valores acima de 0,5 indicando um desempenho superior ao da classificação aleatória (Joshi, 2020; Chadha.; Juwe., 2021).

Gorunescu (2011) e Consoli, Recupero e Petković (2019) fornecem diretrizes para a interpretação da AUC, sublinhando sua importância como uma métrica robusta para avaliação de classificadores:

- 0.90 - 1.00 = classificação excelente;
- 0.80 - 0.90 = classificação boa;
- 0.70 - 0.80 = classificação razoável;
- 0.60 - 0.70 = classificação ruim;
- 0.50 - 0.60 = falha.

Essas diretrizes servem como um meio aproximado de avaliar a eficácia dos classificadores, com uma AUC mais alta indicando um desempenho melhor.

- **Validação Cruzada K-Fold**

  Para avaliar a capacidade de generalização do modelo CNN na detecção de fissuras em imagens, utilizamos a técnica de validação cruzada k-fold (Facelli *et al.*, 2011; Consoli; Recupero; Petković, 2019). Esta técnica divide o conjunto de dados em k subconjuntos, permitindo treinar e testar o modelo k vezes em diferentes partições (Igual; Seguí, 2017).

  Escolhemos k=5 para equilibrar a precisão da avaliação e o custo computacional, uma vez que valores como 5 ou 10 são frequentemente utilizados (Nwanganga; Chapple, 2020). Em cada iteração, um subconjunto é usado para teste e os outros k-1 para treinamento, garantindo que cada subconjunto seja usado uma vez como conjunto de teste.

  Essa abordagem proporciona uma avaliação robusta do desempenho do modelo, mitigando a aleatoriedade na partição dos dados (Consoli; Recupero; Petković, 2019; Plaue, 2023).

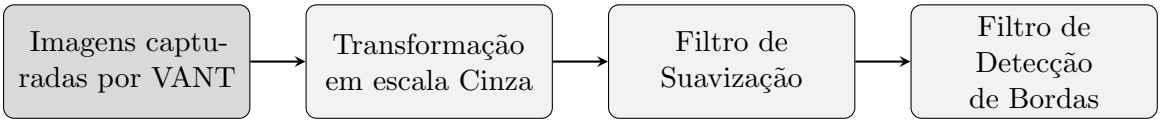
2.4 PRÉ-PROCESSAMENTO DE IMAGENS

O pré-processamento desempenha um papel fundamental na preparação das imagens para análise, assegurando que as características relevantes sejam realçadas e que as imagens estejam em um formato adequado para aplicação de algoritmos de inteligência artificial. Este método compreende a alteração e o tratamento das imagens iniciais, visando aprimorar a clareza e a exatidão das informações, preparando-as para a subsequente etapa de identificação de defeitos.

As etapas principais do pré-processamento incluem a conversão das imagens para escala de cinza, a aplicação de filtros de suavização para reduzir ruídos e irregularidades, e a utilização de filtros de borda para destacar as características estruturais essenciais. Estas etapas são cruciais para garantir que os algoritmos de aprendizado de máquina, como a regressão logística

e redes neurais profundas, possam operar com a máxima eficiência e precisão. O fluxograma a seguir ilustra o processo completo de pré-processamento.

Figura 1 – Fluxograma do processo de pré-processamento de imagens.



Fonte: Próprio autor.

As seções subsequentes detalham cada etapa do pré-processamento, destacando sua importância e metodologia.

2.4.1 Métodos de Conversão para Escala de Cinza

Existem várias técnicas para a transformação de imagens coloridas em escala de cinza, cada uma destacando aspectos distintos da percepção visual e das propriedades técnicas das imagens (Kanan; Cottrell, 2012; Saravanan, 2010; Biswas *et al.*, 2011). Embora haja uma ampla gama de abordagens, este estudo analisará os seguintes exemplos: ***Intensity***, ***Luminance***, ***Luster***, e ***Glightness***. Cada um desses processos proporciona uma abordagem singular na conversão de imagens para escala de cinza.

O método *Intensity* considera a média dos canais RGB (Kanan; Cottrell, 2012):

$$Y_{\text{Intensity}} = \frac{1}{3}(R + G + B) \tag{13}$$

O método *Luminance* utiliza uma combinação ponderada dos canais RGB para alinhar-se com a percepção humana de luminosidade (Kanan; Cottrell, 2012):

$$Y_{\text{Luminance}} = 0{,}3R + 0{,}59G + 0{,}11B \tag{14}$$

Luster é baseado na média dos valores máximo e mínimo dos canais RGB (Kanan; Cottrell, 2012):

$$Y_{\text{Luster}} = \frac{1}{2}(\max(RGB) + \min(RGB)) \tag{15}$$

Neste trabalho utiliza-se o método *Glightness*, seguindo o estudo de Ahmad, Moon e Shin (2018), que indica ser este método como sendo superior na preservação das bordas das imagens:

$$Y_{\text{Glightness}} = 0{,}21R + 0{,}71G + 0{,}07B \tag{16}$$

2.4.2 Filtros de Suavização de Ruídos

Filtros de suavização são frequentemente empregados para desfocar e reduzir ruídos no processamento de imagens. Eles auxiliam em tarefas de pré-processamento, como a remoção de detalhes minúsculos e a conexão de descontinuidades em linhas e curvas (Gonzalez; Woods, 2010). Filtros lineares de suavização podem ser aplicados no domínio espacial, definindo pesos apropriados na máscara de convolução, ou no domínio da frequência, atenuando frequências

altas. Estes filtros desempenham um papel crucial na obtenção de imagens mais suaves e na redução de ruídos (Distante; Distante, 2020).

2.4.2.1 Filtro da Média

O filtro da média é reconhecido por sua simplicidade e eficácia no processamento de imagens. Conforme Solomon e Breckon (2011), este método de processamento espacial consiste em atribuir a cada pixel o valor médio dos níveis de intensidade de seus pixels adjacentes, atribuindo uma ponderação uniforme para cada pixel na área circunvizinha. Pedrini e Schwartz (2007) destaca que tal técnica contribui para a manutenção do valor médio dos níveis de cinza na imagem e promove a uniformidade nas regiões onde os níveis de cinza são estáveis.

No entanto, Zhang (2014) destaca que este método pode resultar em perda de nitidez da imagem, particularmente em suas bordas e detalhes, uma desvantagem compensada pela sua simplicidade e rapidez de cálculo. Além disso, Gonzalez e Woods (2010) enfatizam que, ao reduzir as transições abruptas de intensidades, o filtro da média serve como um método eficaz para a supressão de ruído.

Seja uma imagem $f[x, y]$ de entrada representada pela matriz $M \times N$, e um núcleo ou máscara de suavização, representado pela matriz $m \times n$, $w[i, j]$. O objetivo é aplicar a máscara $w[i, j]$ sobre a imagem de entrada $f[x, y]$, visando gerar uma versão suavizada da mesma. Esta imagem resultante é expressa pela matriz $M \times N$, indicada como $g[x, y]$, tal que

$$g[x,y] = \frac{\sum_{s=-a}^{a} \sum_{t=-b}^{b} w[s+i,t+j] \times f[x+s, y+t]}{\sum_{s=-a}^{a} \sum_{t=-b}^{b} w[s+i,t+j]} \tag{17}$$

onde para a máscara w, uma matriz $m \times n$ é considerada $m = 2a + 1$ e $n = 2b + 1$. Além disso, $g[x, y]$, $f[x, y]$ e $w[i, j]$ correspondem aos elementos individuais de suas respectivas matrizes.

Considerando que as matrizes de suavização espaciais são intrinsecamente simples, é conveniente e esclarecedor apresentar um exemplo numérico do filtro da média em processamento de imagens. Para isso, utilizaremos matrizes simples, representando uma imagem, e uma máscara de média correspondente. Este método não só simplifica o entendimento, mas também ilustra claramente o conceito.

A imagem de entrada $f[x, y]$ é dada por:

$$f[x, y] = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

A máscara de média aplicada é:

$$w[x, y] = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

A convolução é realizada pela equação 17, considerando bordas preenchidas com zeros, um processo chamado de *padding*. A imagem para convolução $f[x, y]$ passa a ser representada por:

$$f[x,y] = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 & 0 \\ 0 & 4 & 5 & 6 & 0 \\ 0 & 7 & 8 & 9 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

A matriz resultante $g[x, y]$ após a convolução é:

$$g[x,y] = \begin{bmatrix} 1.\overline{3} & 2.\overline{3} & 1.\overline{7} \\ 3 & 5 & 3.\overline{6} \\ 2.\overline{6} & 4.\overline{3} & 3.\overline{1} \end{bmatrix}$$

2.4.2.2 Filtro Gaussiano

O filtro gaussiano é caracterizado por ser um filtro linear e utilizar pesos derivados da distribuição gaussiana. Esta técnica é reconhecida por sua capacidade de minimizar ruídos que seguem uma distribuição normal, atuando eficazmente em ambos os domínios, espacial e a frequência (Wang *et al.*, 2014). Devido à capacidade de realizar convoluções bidimensionais de maneira rápida e eficaz, este filtro é extensivamente utilizado (Smith, 2003). A formação de seu núcleo é influenciada pelo parâmetro do desvio padrão da distribuição gaussiana (Jain; Gupta, 2015), proporcionando características como simetria rotacional, separabilidade e excelente localização tanto no tempo quanto na frequência (Charalampidis, 2016).

Embora seja altamente eficaz na eliminação de ruídos gaussianos e mantenha as mesmas características independentemente da orientação (Das *et al.*, 2015), o filtro gaussiano pode ser computacionalmente intensivo e tender a reduzir o nível de detalhes da imagem, afetando a nitidez e as minúcias (Seddik; Tebbini; Braiek, 2014). A largura do filtro, controlada pelo parâmetro σ, é crucial para o equilíbrio entre suavização excessiva e insuficiente (Wang *et al.*, 2014).

O filtro gaussiano apresenta o mesmo grau de suavidade em todas as direções, o que é uma vantagem ao lidar com imagens com direção de borda desconhecida (Fude *et al.*, 2008). No entanto, ele processa a imagem inteira uniformemente, o que pode criar disparidades devido à distribuição não linear das intensidades da imagem (Seddik; Tebbini; Braiek, 2014). Neste trabalho, adota-se uma função gaussiana discreta bidimensional com valor médio zero como filtro de suavização. Conforme Gonzalez e Woods (2010), a equação adaptada do filtro gaussiano em duas dimensões é dada por:

$$w[i,j] = e^{-\frac{i^2+j^2}{2\sigma^2}} \tag{18}$$

onde $w[i, j]$ é o valor do núcleo em uma dada posição (i,j) e σ é o desvio padrão da distribuição gaussiana.

E para exemplificar a aplicação do filtro gaussiano, consideramos a imagem $f[x, y]$ e a

máscara $w[x, y]$, onde cada termo da matriz w é obtido pela equação 18. A imagem suavizada $g[x, y]$ é calculada pela operação de convolução após a normalização da máscara.

Temos a matriz de entrada (imagem) $f[x, y]$:

$$f[x, y] = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

Seja a máscara Gaussiana Normalizada $w[x, y]$:

$$w[x, y] = \begin{bmatrix} 0.058 & 0.096 & 0.058 \\ 0.096 & 0.159 & 0.096 \\ 0.058 & 0.096 & 0.058 \end{bmatrix}$$

Assim temos a imagem suavizada:

$$g[x,y] = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 3 \\ 2 & 3 & 3 \end{bmatrix}$$

2.4.2.3 Filtro de Mediana

O filtro de mediana é comumente empregado no tratamento de imagens, destacando-se por sua habilidade em atenuar distúrbios visuais, notavelmente o ruído do tipo sal e pimenta, ao mesmo tempo em que preserva as bordas e detalhamentos sutis da imagem (Solomon; Breckon, 2011; Aranda; Reviriego; Maestro, 2017). O termo mediana, no contexto do filtro de mediana, refere-se ao valor que está no ponto médio em uma sequência ordenada de valores dos pixels (Zhang, 2014). A aplicação do filtro de mediana envolve deslizar uma janela pela imagem, e a intensidade mediana dos pixels dentro da janela é selecionada como a nova intensidade do pixel processado (Malothu Nagu, 2014).

O processo de cálculo do filtro de mediana consiste em ordenar os valores dos pixels em uma vizinhança e escolher o valor mediano como o novo valor do pixel (Szeliski, 2021; Chang; Hsiao; Hsieh, 2008). Quando o conjunto de números possui uma quantidade ímpar de elementos, a mediana corresponde ao elemento central após a ordenação dos dados. Em contraste, se o conjunto contém uma quantidade par de elementos, a mediana é calculada como a média aritmética dos dois números que se encontram no meio da sequência ordenada (Kumar; Singh, 2020). Os filtros de mediana são populares devido à sua notável capacidade de redução de ruído com menos borrão em comparação a filtros de suavização linear de tamanho similar (Gupta, 2011). Contudo, o custo computacional dos filtros de mediana pode ser maior que o dos filtros de média, o que pode restringir seu uso (Solomon; Breckon, 2011; Szeliski, 2021). A equação do filtro de mediana é a seguinte:

$$g[x, y] = \text{median}\left(\{f[x + i, y + j] : \quad (i, j) \in w\}\right) \tag{19}$$

onde $g[x, y]$ representa o valor resultante do pixel, $f[x + i, y + j]$ são os valores dos pixels adjacentes em torno de (x, y) e w engloba todas as posições dos pixels na área coberta pelo mascara *median* () é a função mediana.

Para demonstrar o filtro da mediana em processamento de imagens, começamos com uma matriz de entrada específica. Seja a matriz de entrada $f[x,y]$ A matriz de entrada é:

$$f[x,y] = \begin{bmatrix} 0 & 0 & 0 \\ 25 & 62 & 30 \\ 95 & 125 & 255 \end{bmatrix}$$

Com *padding* espelhado, a matriz de entrada se expande para:

$$f[x,y] = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 25 & 25 & 62 & 30 & 30 \\ 95 & 95 & 125 & 255 & 255 \\ 95 & 95 & 125 & 255 & 255 \end{bmatrix}$$

A mascara utilizada é:

$$w[i,j] = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Para o pixel g[2,2], considera-se os termos do kernel em ordem crescente:

$$\text{Mediana de } [0, 0, 0, 25, 30, 62, 95, 125, 255]$$

A mediana é 30, então substituímos o valor original do pixel central (62) por 30. Aplicando este processo para todos os pixels, obtemos a matriz de saída $g[x,y]$:

$$g[x,y] = \begin{bmatrix} 0 & 0 & 0 \\ 25 & 30 & 30 \\ 95 & 95 & 95 \end{bmatrix}$$

2.4.3 Detecção de Borda

A detecção de bordas é fundamental no processamento de imagens, identificando limites entre áreas distintas através de descontinuidades nos níveis de cinza (Pedrini; Schwartz, 2007). Essas bordas, características primitivas na imagem, indicam a extensão física de objetos (Pratt, 2007).

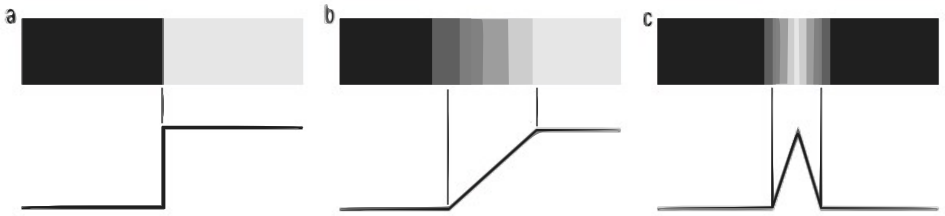
Operadores sensíveis às mudanças são utilizados na detecção de bordas, minimizando áreas de valores de cinza constantes (Jähne, 2005). No entanto, descontinuidades abruptas são raras, e bordas podem ser borradas por suavização e ruído (Gonzalez; Woods, 2010; Pedrini; Schwartz, 2007).

2.4.3.1 Modelos de Borda

A segmentação de imagens geralmente utiliza a identificação de bordas, onde mudanças bruscas na intensidade indicam pontos críticos (Gonzalez; Woods, 2010; Pratt, 2007). As bordas são divididas em três categorias: borda tipo degrau, rampa e telhado, como ilustrado na Figura 2. A borda tipo degrau é caracterizada por uma alteração súbita entre dois níveis de intensidade, separados por um pixel. Esse padrão é comum em imagens sintéticas, como modelagem tridimensional (Gonzalez; Woods, 2010). A borda tipo rampa, típica em imagens digitais do cotidiano, apresenta bordas menos definidas e mais ruidosas devido às limitações do mecanismo de foco. Este modelo possui uma série de pontos dispersos ao longo de uma inclinação, formando um segmento contínuo (Gonzalez; Woods, 2010; Pratt, 2007).

A borda tipo telhado é definida por linhas que cruzam uma área, com a base determinada pela espessura e definição da linha. Este tipo é frequente em imagens que capturam detalhes finos, como dados de profundidade ou imagens de satélite (Gonzalez; Woods, 2010; Majumder; Gopi, 2018). Diversas técnicas foram desenvolvidas para refinar a detecção de bordas, resultando em abordagens variadas que produzem resultados distintos, conforme a natureza da imagem (Leta, 2014).

Figura 2 – Modelos de bordas: degrau, rampa e telhado, da esquerda para a direita.



Fonte: Reproduzido da referência (Gonzalez; Woods, 2010).

2.4.3.2 Definição Matemática, Método das Diferenças Finitas e Principais Etapas do Processo de Detecção de Bordas

Nesta seção, são apresentados os fundamentos matemáticos subjacentes à detecção de bordas, com foco no conceito de gradientes e no método das diferenças finitas para computá-los. Além disso, descrevem-se as principais etapas do processo de detecção de bordas, incluindo suavização, cálculo do gradiente, não-maxima supressão e limiarização para a identificação precisa das bordas.

A importância do gradiente na detecção de bordas é amplamente reconhecida na literatura de processamento de imagens (Hsu; CHen, 2008; Pedrini; Schwartz, 2007; Gonzalez; Woods, 2010).

Segundo Hsu e CHen (2008), o gradiente de uma imagem é uma função que, em cada pixel, o vetor gradiente aponta na direção do maior possível aumento de intensidade. Este é definido como o vetor:

$$\vec{\nabla} F = \frac{\partial F}{\partial x}\hat{i} + \frac{\partial F}{\partial y}\hat{j} \tag{20}$$

Nessa equação, F é a imagem fonte, e $\frac{\partial F}{\partial x}$ e $\frac{\partial F}{\partial y}$ são imagens separadas, onde cada pixel contém os respectivos valores de derivada.

O gradiente eficientemente identifica a intensidade e a orientação de bordas em um ponto (x, y) da imagem, destacando-se na direção de maior variação, conforme descrito por Gonzalez e Woods (2010). Ele calcula a magnitude do gradiente através da equação:

$$M[x,y] = \sqrt{\left(\frac{\partial F}{\partial x}\right)^2 + \left(\frac{\partial F}{\partial y}\right)^2} \tag{21}$$

A magnitude do gradiente, que também pode ser entendida como sua intensidade, corresponde ao grau máximo de variação na direção do vetor gradiente. Adicionalmente, a orientação deste vetor é definida pelo ângulo que ele forma com o eixo x, conforme demonstrado na equação a seguir:

$$\theta = \arctan\left(\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial x}}\right) \tag{22}$$

Pedrini e Schwartz (2007) destacam a relevância da detecção de bordas para revelar mudanças notáveis nos níveis de cinza de uma imagem. Utilizando derivadas parciais para descrever essas variações em duas dimensões, o gradiente serve como um operador essencial, sinalizando áreas com maior alteração nos níveis de cinza. Complementando a descrição matemática, Gonzalez e Woods (2010) destacam a forma matricial do gradiente de uma imagem como $\vec{\nabla} F = \left\{\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}\right\}$.

Essa abordagem permite a identificação precisa de regiões onde ocorrem mudanças significativas nos níveis de cinza, fundamental para a análise de bordas em processamento de imagens.

Hsu e CHen (2008) notam que a magnitude e a direção do gradiente podem ser determinadas pela combinação de derivadas. Em harmonia com as observações de Pedrini e Schwartz (2007), a relevância do gradiente na detecção de bordas é ressaltada, permitindo identificar variações de intensidade na imagem e facilitando o reconhecimento de bordas verticais e horizontais pela análise das diferenças entre pixels adjacentes (Pedrini; Schwartz, 2007).

Hsu e CHen (2008) apontam que o histograma do gradiente de uma imagem nítida é densamente povoado por valores baixos, com algum ruído em valores altos, enquanto uma imagem desfocada mostra pouca presença em valores altos. Isso indica a capacidade do histograma do gradiente em distinguir entre imagens focadas e desfocadas. Por outro lado, Gonzalez e Woods (2010) e Pedrini e Schwartz (2007) concordam que a direção da borda em um ponto $[x, y]$ é perpendicular ao gradiente nesse local, evidenciando a utilidade do gradiente na detecção precisa de bordas. Estes conceitos fundamentam a capacidade de criar algoritmos robustos para a detecção efetiva de bordas em imagens.

2.4.3.3 Supressão Não-Máxima

A Supressão Não-Máxima é essencial em algoritmos de detecção de bordas, como exemplo o Filtro Canny, pois considera apenas os máximos locais para melhorar a precisão. Segundo Gonzalez e Woods (2010), a imagem gradiente $M[x, y]$ frequentemente apresenta regiões largas

em torno dos máximos locais. A técnica usa a magnitude do gradiente (equação 21) e o ângulo do gradiente (equação 22), definindo orientações discretas do vetor gradiente em quatro direções: horizontal, vertical, +45° e -45°.

O processo envolve: identificar picos na imagem gradiente $M[x, y]$, definir orientações do vetor gradiente em uma região 3×3, comparar a magnitude de cada pixel com seus vizinhos na direção do gradiente, e suprimir pixels não máximos, resultando em bordas mais finas e definidas.

Pedrini e Schwartz (2007) destaca que selecionar pontos onde a magnitude do gradiente é máxima localmente afina as bordas e remove artefatos causados por ruído ou texturas sutis. Vale e Poz (2002) observa que altos valores na matriz de magnitude $M[i, j]$ indicam gradientes significativos, mas não definem bordas com precisão. Artero e Tommaselli (2009) descreve a exclusão de pixels com magnitudes inferiores, perpendiculares à direção da borda, retendo apenas aqueles de maior magnitude e categorizando as direções das bordas em oito ângulos diferentes.

2.4.3.4 Limiarização

A técnica de limiarização é fundamental no tratamento de imagens, utilizada para segmentar pixels de acordo com um limiar estabelecido (Filho; Neto, 1999). Seu principal intuito é distinguir regiões de uma imagem que diferem entre si, tipicamente o plano de fundo e o elemento de interesse.

A abordagem matemática para limiarização transforma uma imagem de entrada $F[x,y]$, que possui N níveis de tonalidade, em uma imagem de saída $G[x,y]$, que frequentemente é convertida para uma forma binária (Filho; Neto, 1999; Solomon; Breckon, 2011):

$$G[x,y] = \begin{cases} 1 & \text{se } F[x,y] \geq T \\ 0 & \text{se } F[x,y] < T \end{cases} \tag{23}$$

Em certos contextos, a limiarização pode ser estendida para alocar pixels em mais de duas categorias (Gonzalez; Woods, 2010; Sundararajan, 2017):

$$G[x,y] = \begin{cases} a & \text{se } F[x,y] > T_2 \\ b & \text{se } T_1 < F[x,y] \leq T_2 \\ c & \text{se } F[x,y] < T_1 \end{cases} \tag{24}$$

Onde a representa a categoria atribuída aos pixels com intensidade maior que o limiar superior T_2; b representa a categoria atribuída aos pixels com intensidade entre os limiares T_1 e T_2; e c representa a categoria atribuída aos pixels com intensidade menor que o limiar inferior T_1.

2.4.3.5 Limiarização Conduzida por Propriedades Estatísticas

Como indicado por Filho e Neto (1999), a definição do limiar ótimo deve considerar as propriedades estatísticas inerentes à imagem. Presumindo que os níveis de cinza sigam uma distribuição de probabilidade normal ou gaussiana, os parâmetros chave envolvem:

1. μ_1: Média dos níveis de cinza na área de interesse.
2. μ_2: Média dos níveis de cinza na área de fundo.
3. σ_1 e σ_2: Variação padrão das intensidades de pixel nas áreas de interesse e fundo, respectivamente.
4. $P1$ e $P2$: Probabilidades de aparição dos pixels nas áreas de interesse e fundo, respectivamente.

Com estas informações, o limiar ótimo T é calculado através da equação:

$$AT^2 + BT + C = 0, \tag{25}$$

onde

$$\begin{aligned} A &= \sigma_1^2 - \sigma_2^2, \\ B &= 2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2), \\ C &= \mu_2\sigma_1^2 - \mu_1\sigma_2^2 + 2\sigma_1^2\sigma_2^2 \ln\left(\frac{\sigma_2 P_1}{\sigma_1 P_2}\right). \end{aligned}$$

Quando as variações são equivalentes e as classes são igualmente prováveis, a equação de limiar se resume a:

$$T = \frac{\mu_1 + \mu_2}{2}. \tag{26}$$

2.4.3.6 Processamento de Limiar Global

Em cenários onde a separação de intensidade entre fundo e objetos é marcante, a aplicação de um limiar único em toda a imagem é suficiente (Gonzalez; Woods, 2010). No entanto, para imagens com transições de intensidade menos definidas, Gonzalez e Woods (2010) propõe um método iterativo:

1. Definir um valor inicial para o limiar global, T.
2. Utilizar T para segmentar a imagem.
3. Determinar as médias de intensidade $m1$ e $m2$ para os pixels dos conjuntos G1 e G2, respectivamente.
4. Atualizar T com base na Equação (26).
5. Continuar o processo das etapas 2 a 4 até que a variação de T seja inferior a um limite estabelecido ΔT.

2.4.3.7 Processamento de Limiar Local

Quando a aplicação de um limiar único em toda a imagem não é eficaz, o método de limiarização local se apresenta como uma solução mais adequada (Pedrini; Schwartz, 2007). Existem várias abordagens possíveis, entre elas:

A média de intensidade em uma área circunvizinha local pode ser calculada como:

$$T = \text{média}_v(p) \tag{27}$$

Para a mediana dos valores na área local, tem-se:

$$T = \text{mediana}_v(p) \tag{28}$$

E para o cálculo baseado nos valores extremos da vizinhança:

$$T = \frac{\min_v(p) + \max_v(p)}{2} \tag{29}$$

Definições relevantes:

- v: representa a área local em volta de um ponto na imagem, como um bloco de tamanho $m \times n$.
- p: é um ponto específico na imagem cuja intensidade é analisada.

2.4.3.8 Filtro de Borda Sobel

O filtro de Sobel, formulado por Sobel (1970), atua como um operador de gradiente discreto para a identificação de bordas em imagens (Ntalianis; Croitoru, 2018; Oskoei; Hu, 2010). A técnica emprega um máscara que acentua o pixel central, conseguindo atenuar o ruído mais eficientemente do que métodos análogos (Ansari; Kurchaniya; Dixit, 2017), contudo este processo envolve o cálculo do gradiente de intensidade para cada pixel, permitindo assim determinar a orientação e o ritmo da variação de intensidade (Ntalianis; Croitoru, 2018).

Utilizando pares de matrizes de convolução de dimensão 3x3, o Sobel estima o gradiente espacial em duas orientações distintas, horizontal (G_x) e vertical (G_y), realçando as bordas detectadas (Hussain; Agarwal, 2015). Essas matrizes são aplicadas de maneira individualizada para se extrair os componentes do gradiente em ambas as orientações (Gonzalez; Woods, 2010). O método também é capaz de inferir o ângulo das arestas na imagem (Sluder; Wolf, 2003). Contudo, apesar da sua eficiência, ele apresenta sensibilidade a interferências aleatórias e pode gerar bordas mais grossas que o desejado (Oskoei; Hu, 2010; Distante; Distante, 2020).

As matrizes do filtro Sobel são definidas como:

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \quad G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}.$$

A utilização dessas matrizes gera dois conjuntos de dados de gradientes, um para cada direção espacial. Os valores de magnitude e direção das bordas são determinados por meio de fórmulas que se valem do método de diferenças finitas (Pedrini; Schwartz, 2007).

2.4.3.9 Filtro de bordas Canny

Desenvolvido por John Canny em 1986, o filtro Canny é reconhecido por sua eficácia na detecção de bordas em imagens digitais (Canny, 1986). O algoritmo de Canny é estruturado em torno de três critérios essenciais para otimizar a detecção de bordas, conforme discutido por Sekehravani, Babulak e Masoodi (2020), Hussain e Agarwal (2015) e Pedrini e Schwartz (2007):

1. Precisão: A técnica se destina a identificar corretamente todas as bordas relevantes, reduzindo detecções incorretas e mantendo dados valiosos da imagem.

2. Precisão de localização: Os pontos identificados como bordas devem ser localizados tão perto quanto possível do centro exato da borda real para reduzir a discrepância entre a localização detectada e a real.

3. Singularidade de resposta: O algoritmo visa assegurar que cada borda autêntica seja representada por um único ponto, prevenindo respostas redundantes para a mesma borda.

Canny (1986) concebeu uma formulação matemática desses princípios e buscou soluções ideais. Por meio de otimização numérica, descrito por Distante e Distante (2020), eles propuseram operadores ótimos para diferentes tipos de arestas, elucidando um princípio de incerteza para a detecção e localização de bordas em meio a ruído e estabelecendo um equilíbrio entre esses fatores.

O processo do algoritmo de Canny ocorre em múltiplas fases, como indicado por Sekehravani, Babulak e Masoodi (2020) e Hussain e Agarwal (2015):

1. **Suavização**: O processo inicia com o amaciamento da imagem através do filtro gaussiano.

2. **Extração de gradientes**: A etapa seguinte envolve calcular os gradientes da imagem ao longo das direções x e y, seguido pela definição da magnitude e orientação dos gradientes.

3. **Supressão de não-máximos**: Esta fase refina as bordas, resultando numa imagem binária onde apenas os valores máximos são retidos (definidos como 1), enquanto outros são atenuados (definidos como 0).

4. **Limiarização progressiva**: Esta etapa diferencia os pixels de borda em fortes e fracos, baseando-se em dois níveis de limiar. pixels que excedem o limiar superior são classificados como bordas fortes, e aqueles que se situam entre os dois limiares são identificados como bordas frágeis. Um pixel de borda frágil é considerado parte da borda se estiver ligado a um pixel de borda forte.

Segundo Distante e Distante (2020), a metodologia de Canny é notavelmente eficaz para a delimitação precisa de bordas e na supressão do ruído, estabelecendo-se como referência para detecção de bordas em contextos digitais.

3 MÉTODOS DE CLASSIFICAÇÃO UTILIZANDO INTELIGÊNCIA ARTIFICIAL

Inteligência artificial é um ramo da ciência computacional que ensina os sistemas a adaptarem suas ações para melhorar a precisão, um conceito cunhado na era moderna da computação (Marsland, 2009). Baseia-se no princípio de que máquinas podem aprender a partir de dados e *feedback* do ambiente, refinando comportamentos para alcançar resultados desejados, uma ideia que tem suas raízes matemáticas e abstratas bem estabelecidas (Joshi, 2020). Apresenta-se a seguir os métodos avaliados neste trabalho:

3.1 REGRESSÃO LOGÍSTICA EM INTELIGÊNCIA ARTIFICIAL

A regressão logística é um método de classificação fundamental no campo da inteligência artificial, sendo crucial para entender eventos binários. Diferente da regressão linear, que prevê resultados contínuos, a regressão logística é projetada para problemas de classificação binária (Izbicki; Santos, 2020; Zomchak; Starchevska, 2023).

O modelo matemático da regressão logística é representado pela função logística, que modela a probabilidade de uma variável de resposta binária como uma função de variáveis preditoras:

$$P(Y = 1) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X)}} \tag{30}$$

onde $P(Y = 1)$ é a probabilidade do resultado binário ser 1, β_0 e β_1 são os parâmetros do modelo, e X é a variável preditora. Para múltiplas variáveis, a forma paramétrica é (Izbicki; Santos, 2020):

$$P(Y = 1|x) = \frac{e^{\beta_0 + \sum_{i=1}^{n} \beta_i x_i}}{1 + e^{\beta_0 + \sum_{i=1}^{n} \beta_i x_i}} \tag{31}$$

Os coeficientes são estimados pelo método de máxima verossimilhança, maximizando a função de verossimilhança condicional.

A regressão logística é amplamente utilizada em diversos domínios, incluindo saúde e finanças. Na saúde, prevê a probabilidade de doenças específicas (Thabtah; Abdelhamid; Peebles, 2019), enquanto nas finanças, detecta fraudes em transações com cartão de crédito (Schein; Ungar, 2007).

Extensões incluem a regressão logística multinomial e ordinal, adequadas para problemas de

classificação multi-classe (Gladence *et al.*, 2015). A regressão logística ordinal modela variáveis de resposta ordinais.

Apesar de sua ampla aplicação, a regressão logística tem limitações, como a suposição de linearidade entre a variável dependente e as variáveis independentes, além de ser sensível ao sobreajuste em espaços de alta dimensão (Musa, 2013; Saberioon *et al.*, 2018).

3.2 MÉTODO *SUPPORT VECTOR MACHINE* (SVM)

O método Máquina de Vetores de Suporte (em inglês *Support Vector Machine* - SVM) emergiu como uma técnica robusta para classificação e regressão, com fortes aplicações em Inteligência Artificial (IA). A abordagem da SVM é baseada na Minimização de Risco Estrutural (SRM), buscando um hiperplano ótimo em um espaço de características elevado (Mavroforakis; Theodoridis, 2006). O princípio matemático é expresso por:

$$f(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x} + b \tag{32}$$

onde $f(\mathbf{x})$ é a função de decisão, $\mathbf{w}$ é o vetor de pesos, $\mathbf{x}$ é o vetor de entrada, e b é o viés (Mavroforakis; Theodoridis, 2006). Para dados não linearmente separáveis, a SVM emprega um *kernel* para transformar o espaço de características (Wu *et al.*, 2023). A seleção criteriosa de parâmetros, incluindo o fator de penalidade C e a função de kernel, é crucial (Wu *et al.*, 2023).

SVMs são métodos não paramétricos, supervisionados e baseados em *kernel*, que utilizam um mapeamento implícito dos dados de entrada em um espaço de características de alta dimensão (Saberioon *et al.*, 2018). Este método usa um hiperplano linear como função de decisão para problemas não lineares e aplica uma transformação reversa no espaço não linear. A SVM utiliza os multiplicadores de Lagrange para resolver a otimização do problema, permitindo encontrar a solução ótima para a função de decisão. O classificador é dado por:

$$y(x) = \sigma \left(\sum_{k=1}^{N} \alpha_k y_k \phi(x, x_k) + b \right) \tag{33}$$

onde $y(x)$ é a saída do classificador, α_k são constantes reais positivas, y_k é a saída correspondente ao ponto x_k, $\phi(x, x_k)$ é a função *kernel*, e b é uma constante real. A função de base radial (*Radial Basis Function* - RBF) é um *kernel* popular, calculada como:

$$\phi(x, x_k) = \exp\left(-\frac{\|x - x_k\|^2}{2\sigma^2} \right) \tag{34}$$

onde σ é a largura da função de base radial (Saberioon *et al.*, 2018). A divisão de amostra adotada é a validação cruzada *leave-one-out*, que proporciona uma acurácia diagnóstica estável e evita o sobreajuste (Wu *et al.*, 2023).

A SVM também se aplica à detecção de fissuras em estruturas de concreto, mostrando eficácia na identificação de fissuras com o uso de filtros de suavização (Ding *et al.*, 2023). O método busca um hiperplano ótimo no novo espaço de características, maximizando o intervalo entre o conjunto de amostras e o hiperplano classificador.

SVMs são introduzidas como uma solução para problemas de duas classes, diferenciando entre dois grupos com um hiperplano ótimo que maximiza a margem de separação (Brereton; Lloyd, 2009). A formulação SVM incorpora o princípio de minimização de risco estrutural (em inglês, *Structural Risk Minimizatio* - SRM), demonstrando superioridade em relação ao princípio de minimização de risco empírico (em inglês, *Empirical Risk Minimization* - ERM) usado por outras técnicas (Dibike *et al.*, 2001).

3.3 MÉTODO DE REDES NEURAIS CONVOLUCIONAIS (CNNS)

Redes Neurais Convolucionais (CNNs) são categorizadas como redes neurais especializadas, apresentando alto desempenho em tarefas de reconhecimento e classificação de imagens (Santos; Araújo, 2023). As CNNs evoluíram para uma ampla gama de aplicações desde seu desenvolvimento inicial para o reconhecimento de dígitos manuscritos (Tuba *et al.*, 2021).

Uma CNN típica é estruturada da seguinte maneira:

1. **Camada Convolucional:** Esta camada utiliza filtros para processar a entrada, gerando mapas de ativação que capturam características importantes da imagem (Santos; Araújo, 2023).

2. **Camada ReLU:** Implementa um limiar para introduzir não-linearidade ao processo, convertendo todos os valores negativos em zero (Santos; Araújo, 2023).

3. **Camada de Pooling:** Trabalha na redução da dimensionalidade dos mapas de ativação, o que diminui os parâmetros da rede e consolida informações importantes (Santos; Araújo, 2023).

4. **Camadas Totalmente Conectadas:** Estas camadas integram todos os recursos aprendidos anteriormente para realizar a classificação final da imagem (Santos; Araújo, 2023).

3.3.1 Ativação Sigmoide

A função sigmoide, com sua forma em S, é usada em redes neurais para limitar a saída entre 0 e 1 e fornecer resultados probabilísticos. Waoo e Soni (2021) destaca sua utilidade em tarefas de classificação binária. Segundo Nwankpa *et al.* (2018), a função sigmoide é importante em aprendizado profundo devido à sua natureza não linear, expressa pela equação:

$$f(\mathbf{X}) = \frac{1}{1 + e^{-\mathbf{X}}} \tag{35}$$

onde $f(\mathbf{X})$ representa a saída da função sigmoide aplicada a uma matriz $\mathbf{X}$, e $\mathbf{X}$ é a matriz de entrada. Esta fórmula transforma qualquer valor real em um valor entre 0 e 1, tornando-se útil para modelar probabilidades.

A função sigmoide é fácil de entender e usada em redes simples, mas tem desvantagens em redes profundas, como saturação do gradiente e convergência lenta (Waoo; Soni, 2021; Nwankpa *et al.*, 2018). Ela é eficaz para valores de entrada entre 0 e 1, mas pode sofrer com

gradientes reduzidos e saída não centralizada em zero, o que pode afetar o desempenho da rede.

3.3.2 Dropout

Dropout é uma técnica para evitar *overfitting* em redes neurais. Srivastava *et al.* (2014) descreve o *dropout* como a remoção aleatória de neurônios durante o treinamento. Cada neurônio é mantido com uma probabilidade p, geralmente definida como 0,5 para unidades ocultas.

Hinton *et al.* (2012) explica que o *dropout* impede que os neurônios dependam uns dos outros, o que melhora a generalização do modelo. Durante o treinamento, cada neurônio é desativado com uma probabilidade de 0,5, criando várias sub-redes. Rosa, Papa e Yang (2018) descreve que, durante o teste, os pesos da rede são ajustados pela razão p para refletir a média das sub-redes treinadas. O *dropout* é realizado através da seguinte equação:

$$\mathbf{W}_{\text{test}}^{(L)} = p\mathbf{W}^{(L)} \tag{36}$$

onde $\mathbf{W}_{\text{test}}^{(L)}$ é a matriz de pesos ajustada para o teste, $\mathbf{W}^{(L)}$ é a matriz de pesos original e p é a probabilidade de retenção dos neurônios.

4 RESULTADOS E DISCUSSÕES

Neste capítulo são apresentados os resultados obtidos na pesquisa, organizados em três partes principais. A primeira parte discute os resultados das imagens obtidas pelos VANTs e a criação do *dataset* personalizado. A segunda parte investiga os métodos de pré-processamento, incluindo a análise de filtros de suavização e filtros de bordas. A terceira parte foca na análise dos métodos de Inteligência Artificial (IA) utilizados para a detecção de trincas, abrangendo CNN, SVM e regressão logística, além de avaliar a influência da quantidade de imagens do *dataset* nos resultados.

Os corpos de de provas com contem danos foram categorizados como trincas, de acordo com a tabela 1 e as imagens capturadas são organizadas em um *dataset* personalizado, categorizado em três grupos: sem trinca, trinca pequena e trinca grande. Técnicas de aumento da quantidade de imagens foram aplicadas para expandir o *dataset*, resultando em um total de 19.659 imagens. Todas as simulações utilizaram a seguinte distribuição de imagens, exceto no estudo da quantidade de imagens do *dataset* na secção 4.5, onde a quantidade de imagens foram alteradas propositalmente para estudar sua influência: um total de 9859 imagens com trincas, 9800 imagens sem trincas, com 13761 (70.00%) das imagens destinadas ao treinamento, 3932 (20.00%) para teste e 1966 (10.00%) para validação.

Estudar a influência da quantidade de imagens do *dataset* é fundamental, pois em trabalhos de campo é comum encontrar *datasets* desequilibrados, com uma quantidade significativamente maior de imagens sem danos. Essas imagens são mais fáceis de obter e, portanto, mais disponíveis para pesquisas. Para avaliar a influência da quantidade de imagens de imagens na eficácia do modelo CNN, o *dataset* foi ampliado através de técnicas de aumento de dados e organizado em três diferentes proporções de imagens.

Diversos métodos de pré-processamento foram investigados para melhorar a qualidade das imagens e a detecção de trincas. Foram analisados vários filtros de suavização para reduzir o ruído das imagens e filtros de bordas para realçar as trincas e defeitos estruturais. A eficácia de cada filtro foi avaliada com base na clareza das trincas detectadas. Para a avaliação dos filtros de bordas, utilizamos a regressão logística, enquanto o método SVM (Máquinas de Vetores de Suporte) foi empregado para analisar os filtros de suavização. A escolha desses métodos se deve à sua simplicidade e ampla utilização na literatura para detecção de danos de forma binária. A SVM tem sido amplamente utilizada em estudos de SHM (em inglês, *Structural Health Monitoring*) e detecção de trincas, como demonstrado por Haoran *et al.* (2020), Chaiyasarn *et al.* (2018), Hoshyar, Kharkovsky e Samali (2018). Por outro lado, a regressão logística também se mostrou eficaz nesses contextos, como evidenciado nos trabalhos de Li *et al.* (2019) e Ali *et al.* (2021). A regressão logística e o SVM são métodos robustos e

bem estabelecidos, frequentemente utilizados devido à sua capacidade de lidar com problemas de classificação binária e sua eficiência em detectar padrões específicos em imagens. Além disso, durante o processo de aprendizagem foi necessário começar com métodos mais simples para construir uma base sólida de conhecimento.

Na análise de pré-processamento, descobriu-se que o filtro de suavização Gaussiano apresentou o melhor desempenho, seguido pelo filtro de borda Sobel. No entanto, na análise do filtro de borda, utilizou-se o filtro de suavização mediana. Nas demais análises (comparação de métodos de IA e diferentes volumes de *dataset*), empregou-se a suavização mediana e o filtro de borda Canny. Essa escolha deve-se ao fato de que a primeira versão da análise foi realizada com o *dataset* descrito em Özgenel (2018), no qual os melhores filtros foram a suavização Mediana e o filtro de borda Canny. Com o aprimoramento do código e a utilização do *dataset* produzido pelo autor, identificou-se que os filtros Gaussiano e Sobel eram mais eficazes. No entanto, para manter a consistência e comparabilidade dos resultados, optou-se por continuar utilizando os filtros Mediana e Canny nas análises subsequentes.

A investigação sobre os métodos de IA focou na comparação de três técnicas principais: Redes Neurais Convolucionais (CNN), Máquinas de Vetores de Suporte (SVM) e Regressão Logística. Embora seja sabido que as CNNs geralmente apresentam um desempenho superior em tarefas de reconhecimento de padrões em imagens, optou-se por incluir a SVM e a Regressão Logística para fornecer uma análise comparativa abrangente. Essas técnicas tradicionais são frequentemente usadas como referências básicas e ajudam a contextualizar os ganhos de desempenho alcançados com as CNNs. Além disso, devido ao maior grau de complexidade do código das CNNs, foi importante avaliar se o uso dessa técnica avançada era realmente necessário para obter resultados satisfatórios. Cada método foi treinado e testado utilizando o *dataset* pré-processado. Além disso, foi investigada a influência da quantidade de imagens do *dataset* nos resultados obtidos, avaliando-se a precisão e a taxa de detecção de trincas.

Para todas as análises computacionais, foi utilizado o *framework* Python 3 no ambiente de desenvolvimento colaborativo online Google Colab, aproveitando o acelerador de hardware V100 GPU. Esta configuração, suportada por 51 GB de RAM do sistema, 16 GB de RAM da GPU e 200 GB de espaço em disco, foi crucial para processar eficientemente a grande quantidade de imagens em um período de tempo adequado. Devido à variabilidade dos recursos disponibilizados pelo Google Colab em diferentes momentos, a análise de tempo de execução de cada código foi deixada de lado, já que a mesma tarefa poderia ser executada mais rapidamente ou mais lentamente dependendo da disponibilidade dos recursos no momento da execução.

Os resultados obtidos fornecem uma base sólida para futuras pesquisas, destacando a eficácia dos métodos de captura e análise de imagens na detecção de trincas em estruturas de concreto, e oferecendo valiosas perspectivas para aprimoramentos e desenvolvimentos subsequentes.

4.1 CRIAÇÃO DO *DATASET* COM IMAGENS CAPTURADAS POR VANTS

Com o objetivo de obter um dataset de imagens, foi utilizado um Veículo Aéreo Não Tripulado (VANT) para capturar imagens de trincas em estruturas de concreto. O modelo escolhido foi

um quadrotor comercial DJI Phantom 3 Standard foi utilizado para coletar imagens aéreas em diferentes períodos e configurações de voo. O Phantom 3 possuía uma câmera integrada com estabilizador tipo *gimbal*, sensor CMOS de 1/2.3 polegadas, que capturava vídeos até 4096x2160p a 24 fps e fotos de 12 megapixels. A câmera tinha uma lente de 20 mm, abertura f/2.8, velocidade do obturador de 8 a 1/8000 s, resolução de 4000x3000 e ISO variando de 100 a 1600. Essas características foram cruciais para obter imagens nítidas, necessárias para a detecção precisa de trincas.

Para simular cenários reais de inspeção de estruturas, foram realizadas variações nas condições de captura de imagens, incluindo altitude, luz, velocidade do VANT e ângulo de inclinação da câmera. As alturas de voo aproximadas foram de 1 metro (baixa), 2 metros (média) e 4 metros (alta). As configurações variáveis são detalhadas na Tabela 3. Foram fotografados três corpos de prova de concreto (um sem trinca, um com trinca grande e outro com trinca pequena) em dez configurações diferentes de voo, gerando um total de 100 imagens.

Tabela 3 – Configurações de voo utilizadas no experimento.

Configuração	**Altitude**	**Luz**	**Velocidade**	**Ângulo de Inclinação**
1	Baixa	Alta	Estacionário	90 graus
2	Média	Alta	Estacionário	90 graus
3	Alta	Alta	Estacionário	90 graus
4	Baixa	Baixa	Estacionário	90 graus
5	Média	Baixa	Estacionário	90 graus
6	Alta	Baixa	Estacionário	90 graus
7	Baixa	Alta	Estacionário	45 graus
8	Baixa	Baixa	Estacionário	45 graus
9	Baixa	Alta	1 m/s	90 graus
10	Média	Alta	1 m/s	90 graus

Fonte: Próprio autor.

A captura de imagens foi realizada com o VANT DJI Phantom 3 Standard, cujas características técnicas permitiram a captura de imagens de alta resolução, essenciais para a detecção precisa de trincas. As configurações do VANT, como ISO, velocidade do obturador e abertura, foram ajustadas automaticamente para se adaptar às variações de iluminação e condições ambientais durante os experimentos ao ar livre. A Figura 3 ilustra o experimento de captura de imagem com o VANT em voo, preparado para fotografar os corpos de prova de concreto.

Figura 3 – Experimento de captura de imagem com VANT.

Fonte: Próprio autor.

As diferentes configurações de captura de imagem foram cruciais para avaliar a eficácia na detecção de trincas e defeitos estruturais em estruturas de concreto. Estas configurações incluíram variações na altitude do VANT, condições de iluminação (como a presença de sombra), o ângulo de inclinação da câmera e a velocidade de voo. Cada uma dessas variáveis influenciou significativamente a clareza e visibilidade dos detalhes nas estruturas.

Para ilustrar o impacto dessas variáveis, a Figura 4 apresenta quatro diferentes configurações de captura. Esta abordagem permitiu uma comparação direta entre os cenários e destacou como cada configuração afetou a qualidade final da imagem, facilitando a identificação das condições mais favoráveis para a detecção precisa de trincas e defeitos.

Para aumentar a robustez e a generalização do modelo, foi implementada uma técnica de aumento de dados em 300 imagens de Regiões de Interesse (ROI) derivadas de 10 imagens de cada configuração, categorizadas em três grupos: sem trinca, trinca pequena e trinca grande, conforme ilustrado na Figura 5. Técnicas de aumento de dados, incluindo *flips* horizontais e verticais, rotação, mudanças de contraste, brilho, desfoque, ruído e recorte, foram empregadas, resultando em um total de 9.859 imagens com trincas e 9.800 sem trincas, totalizando 19.659 imagens.

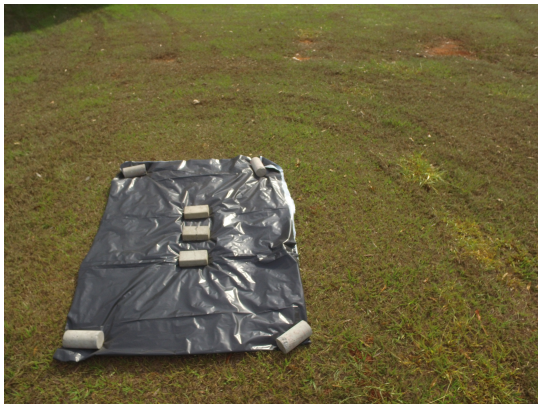
As imagens de 4000x3000 pixels capturadas pelo VANT incluíam três corpos de prova. Enquanto o ROI pode ser identificado automaticamente através de técnicas de processamento de imagem, como detecção de bordas, segmentação ou algoritmos de aprendizado de máquina, neste experimento foi empregado um método assistido, onde a seleção das regiões de interesse é feita manualmente por um operador. Para identificar as regiões de interesse (ROIs), foi utilizado o

Figura 4 – Imagem composta ilustrando quatro configurações diferentes: variações em altitude, presença de sombra, ângulo de inclinação da câmera e velocidade de voo.

a) Configuração 1

b) Configuração 3

c) Configuração 4



d) Configuração 7

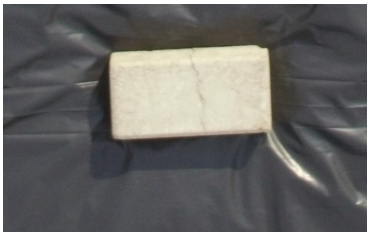
Fonte: Próprio autor.

software livre Inkscape. Neste experimento, cada corpo de prova foi cuidadosamente selecionado e extraído das imagens originais. Esses recortes foram posteriormente redimensionados para 64x64 pixels, padronizando o tamanho para uma análise consistente pelos algoritmos de IA, assegurando que variações no tamanho das imagens não afetassem a detecção de trincas.

Figura 5 – Corpos de prova segmentados em regiões de interesse.

a) Sem Trinca (ST)

b) Trinca Grande (TG)



c) Trinca Pequena (TP)

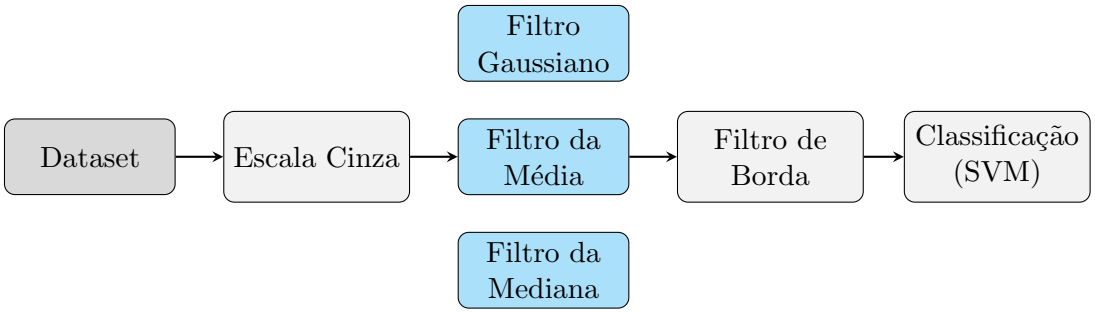
Fonte: Próprio autor.

4.2 ANÁLISE COMPARATIVA DE FILTROS DE SUAVIZAÇÃO

Nesta seção, apresentam-se os resultados da avaliação comparativa dos filtros de suavização aplicados às imagens de estruturas de concreto. Utiliza-se o filtro de borda Canny, reconhecido por sua eficácia em destacar as bordas importantes em imagens, para determinar qual filtro de suavização maximiza o desempenho de um algoritmo de *Support Vector Machine* (SVM) na detecção de trincas. Os filtros testados incluem o filtro Gaussiano, filtro de Média e filtro de Mediana. As imagens foram inicialmente convertidas para escala de cinza utilizando o método *Gligtness* antes da aplicação dos filtros de suavização. O objetivo desta investigação é analisar a influência dos diferentes filtros de suavização na detecção de trincas e avaliar como cada filtro impacta o desempenho do algoritmo SVM.

Para ilustrar o processo de avaliação, o fluxograma representado na Figura 6 detalha as etapas desde a seleção das imagens até a aplicação dos filtros de suavização e a análise dos resultados.

Figura 6 – Fluxograma do processo de avaliação dos filtros de suavização.



Fonte: Próprio autor.

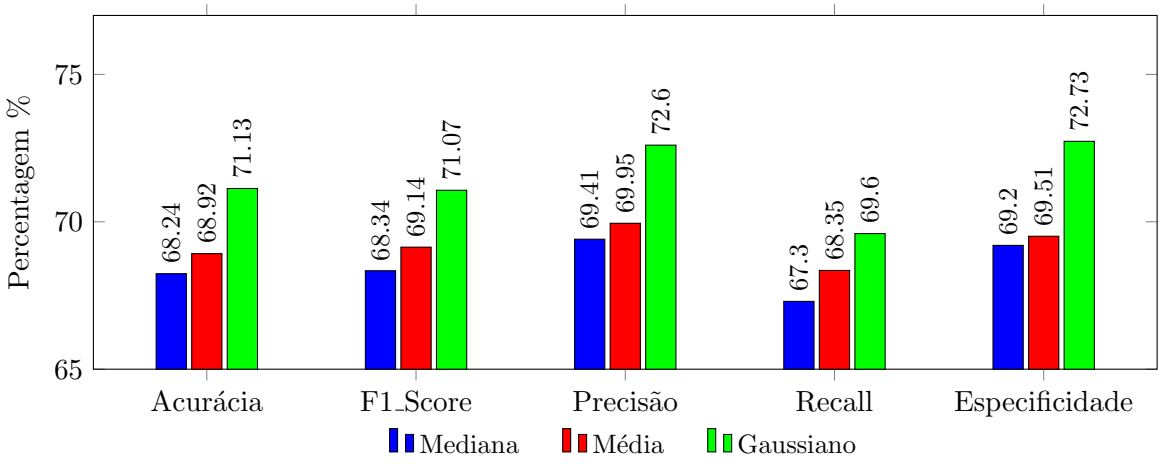
Os resultados das métricas principais apresentados na Figura 7 mostram que o filtro Gaussiano apresentou a melhor performance em termos de acurácia (71,13%), F1-Score (71,07%) e precisão (72,60%). Comparativamente, o filtro de Média obteve 68,92% de acurácia, 69,14% de F1-Score e 69,95% de precisão, enquanto o filtro de Mediana apresentou 68,24% de acurácia, 68,34% de F1-Score e 69,41% de precisão. Esses resultados, sugerem que o filtro Gaussiano é mais eficaz para a detecção de falhas em imagens, possivelmente devido à sua capacidade de suavizar ruídos mantendo as características essenciais das bordas.

A especificidade e sensibilidade, que podem ser analisadas na Figura 7, destacam que o filtro Gaussiano tem a maior especificidade (72,73%) e a maior sensibilidade (*recall*) (69,60%). Isso indica que o filtro Gaussiano é mais eficaz tanto em evitar falsos positivos quanto na detecção de verdadeiros positivos, superando os outros filtros em ambas as métricas.

A análise da matriz de confusão (Tabela 4) revela que o filtro Gaussiano também obteve os menores valores de Falsos Positivos (13,38%) e Falsos Negativos (15,49%). Este desempenho superior sugere que o filtro Gaussiano tem uma melhor capacidade de discriminar entre as classes, resultando em menos erros de classificação.

A média simples dos erros e acertos ilustrada na Figura 8 mostra que o filtro Gaussiano apresentou a maior percentagem de acertos (35,57%), seguido pelo filtro de Média (34,46%)

Figura 7 – Resultados das Métricas de Regressão Logística com Diferentes Filtros de Suavização



Fonte: Próprio autor.

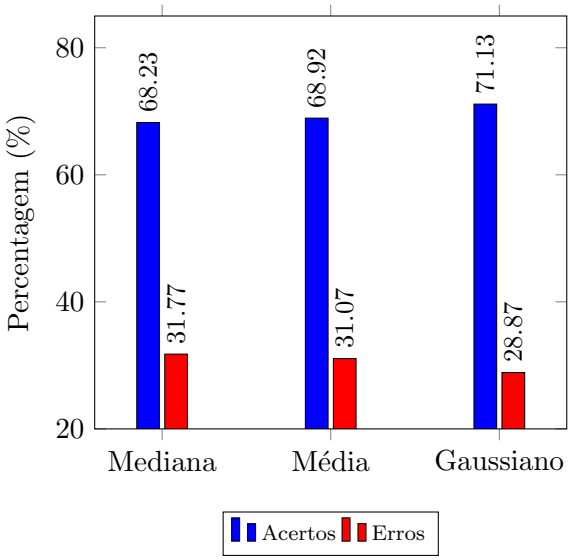
Tabela 4 – Matriz de Confusão em Percentagem para SVM com Diferentes Filtros de Suavização

Filtro	VP (%)	FP (%)	VN (%)	FN (%)
Mediana	34,28	15,11	33,95	16,66
Média	34,82	14,95	34,10	16,12
Gaussiano	35,45	13,38	35,68	15,49

Fonte: Próprio autor.

e pelo filtro de Mediana (34,12%). Quanto aos erros, o filtro Gaussiano teve o menor percentual (14,435%), enquanto os filtros de Média e Mediana apresentaram 15,54% e 15,89% respectivamente.

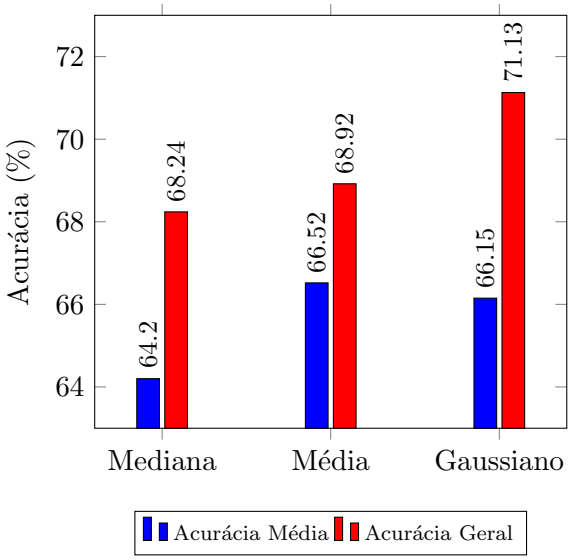
Figura 8 – Acertos e Erros em Percentagem para Diferentes Filtros de Suavização



Fonte: Próprio autor.

Observou-se que, embora o filtro Gaussiano tenha mostrado os melhores resultados gerais nas métricas principais, na validação cruzada, o filtro de Média mostrou uma acurácia ligeiramente superior (66,52%) em comparação ao filtro Gaussiano (66,15%), como mostrado na Figura 9. Essa discrepância pode ser atribuída à variação nos subconjuntos de dados usados em diferentes iterações da validação cruzada, o que indica que, apesar do filtro Gaussiano ter um desempenho superior nas métricas gerais, o filtro de Média pode ser mais robusto em cenários variados.

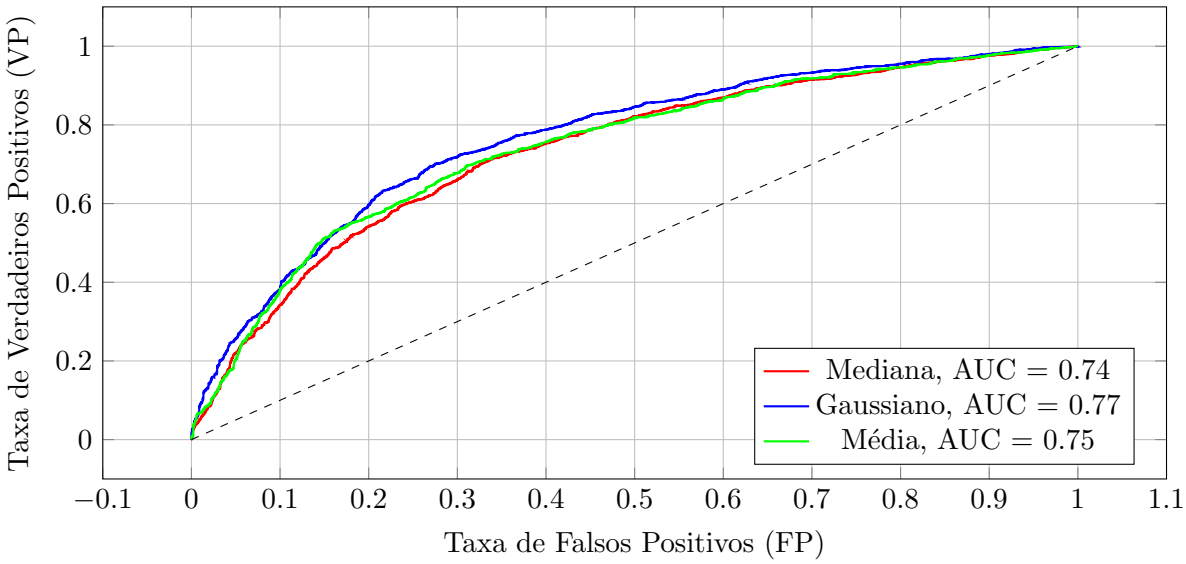
Figura 9 – Acurácia Média e Acurácia Geral para Diferentes Filtros de Suavização



Fonte: Próprio autor.

As curvas ROC (Figura 10) revela que o filtro Gaussiano fica acima das curvas dos filtros de Mediana e Média, com uma AUC de 0.77, indicando uma melhor taxa de verdadeiros positivos em relação à taxa de falsos positivos. As curvas ROC dos filtros de Mediana e Média também mostraram boas performances, com AUCs de 0.72 e 0.74, respectivamente.

Figura 10 – Curvas ROC para Regressão Logística com Diferentes Filtros de Suavização



Fonte: Próprio autor.

Em resumo, os resultados indicam que o filtro Gaussiano é a escolha mais robusta para a detecção de falhas em imagens, apresentando melhor desempenho em quase todas as métricas analisadas. Embora o filtro de Média tenha mostrado uma acurácia ligeiramente melhor, o filtro Gaussiano oferece um equilíbrio superior entre precisão, *recall*, e especificidade, tornando-o mais confiável para aplicações práticas de detecção de falhas.

Também é necessário citar que os resultados desta pesquisa, particularmente no que diz respeito à análise comparativa dos filtros de suavização para a detecção de trincas em estruturas de concreto, foram publicados no *27th COBEM – International Congress of Mechanical Engineering* (Santos *et al.*, 2023). No entanto, no artigo do COBEM foi utilizado o *dataset* disponibilizado por Özgenel (2018), onde as imagens são mais centralizadas e uniformes, e um menor número de imagens foi usado. Nesse estudo, o filtro de suavização que apresentou melhores resultados foi o filtro da Mediana. A diferença nos resultados observados entre os dois estudos pode ser atribuída às variações nos *datasets* e na quantidade de imagens.

O código utilizado para a aplicação de suavização e detecção de bordas na identificação de trincas em imagens via Máquina de Vetores de Suporte (SVM) foi registrado no INPI. A patente, intitulada "Aplicação de Suavização e Detecção de Bordas para a Identificação de Trincas em Imagens via Máquina de Vetores de Suporte (SVM)", foi registrada como Programa de Computador sob o número BR512023002986-6, com data de registro em 20/08/2023, na Instituição de registro: INPI - Instituto Nacional da Propriedade Industrial.

4.3 AVALIAÇÃO DE TÉCNICAS DE DETECÇÃO DE BORDAS

Na busca pela técnica mais eficaz para detecção de trincas em estruturas de concreto, empregou-se o filtro de mediana como método de suavização nas imagens. Esta escolha se baseou no desempenho do filtro de mediana. Além disso, a regressão logística foi empregada como o método de inteligência artificial para analisar os resultados da aplicação dos diferentes filtros de borda. Esta combinação de suavização e análise foi crucial para avaliar com precisão a eficácia das técnicas de detecção de bordas.

Para compreender melhor a metodologia aplicada na avaliação das diferentes técnicas de detecção de bordas, elaboramos um fluxograma do processo. Este fluxograma detalha cada etapa do processo, desde a seleção das imagens até a aplicação das técnicas de detecção de bordas e a análise dos resultados utilizando a regressão logística.

Figura 11 – Fluxograma do processo de avaliação das técnicas de detecção de bordas.

Fonte: Próprio autor.

O fluxograma apresentado na Figura 11 ilustra as etapas sequenciais adotadas na avaliação. As imagens foram inicialmente coletadas e preparadas para o processamento, utilizando o filtro

de mediana para suavização. Em seguida, cada técnica de detecção de bordas foi aplicada, e o desempenho de cada uma foi avaliado utilizando o método de regressão logística.

Foram testadas diversas técnicas de detecção de bordas, incluindo os filtros Canny, Laplacian, Log, Zero Crossing, Prewitt, Robert, Sobel, Kirsch e Marr-Hildreth. Cada uma dessas técnicas possui características específicas que podem influenciar a eficácia na detecção de trincas em imagens de concreto.

Para cada técnica de detecção de bordas, as imagens foram suavizadas utilizando o filtro de mediana antes da aplicação da técnica. As imagens resultantes foram então analisadas utilizando regressão logística para determinar a eficácia de cada técnica na detecção de trincas.

Tabela 5 – Desempenho dos Métodos de Detecção de Bordas

Método	Acurácia (%)	F1_Score (%)	Precisão (%)	Recall (%)	Especificidade (%)
Canny	68.24	68.34	69.41	67.30	69.21
Laplacian	70.60	70.53	72.07	69.05	72.21
Log	68.57	68.65	69.78	67.55	69.62
Zero_Crossing	60.30	60.85	61.14	60.56	60.03
Prewitt	69.20	69.02	70.78	67.35	71.12
Robert	69.84	70.10	70.81	69.40	70.30
Sobel	71.03	71.53	71.62	71.44	70.61
Kirsch	68.26	68.55	69.21	67.90	68.64
Marr_Hildreth	61.88	61.65	63.22	60.16	63.66

Fonte: Próprio autor.

Conforme mostrado na Tabela 5, o método Sobel apresentou os melhores resultados gerais, com uma acurácia de 71.03%, um F1-Score de 71.53% e valores altos de precisão e recall, ambos em torno de 71.44%. Isso indica que o método Sobel é eficaz na detecção de trincas, equilibrando bem a precisão e o recall.

Os métodos Laplacian e Robert também demonstraram um bom desempenho, com acurácias de 70.60% e 69.84%, respectivamente. O método Laplacian teve um bom equilíbrio entre precisão e recall, enquanto o método Robert teve um F1-Score ligeiramente superior (70.10%).

As matrizes de confusão apresentadas na Tabela 6 fornecem uma visão detalhada dos verdadeiros negativos (VN), falsos positivos (FP), falsos negativos (FN) e verdadeiros positivos (VP) para cada método. O método Sobel novamente se destacou, com uma proporção de 34.61% de VN, 14.45% de FP, 14.45% de FN e 36.50% de VP, indicando um bom equilíbrio entre a detecção de trincas e a minimização de falsos alarmes.

Os resultados da validação cruzada de acurácia para os diferentes métodos de detecção de bordas estão apresentados na Figura 12. Observa-se que a acurácia média na validação cruzada é ligeiramente inferior à acurácia observada nas métricas principais para todos os métodos, o que pode ser atribuído à variabilidade dos dados utilizados em diferentes subconjuntos durante o processo de validação. O método Sobel apresentou a maior acurácia média (69.74%), enquanto o método Zero Crossing teve a menor (58.68%).

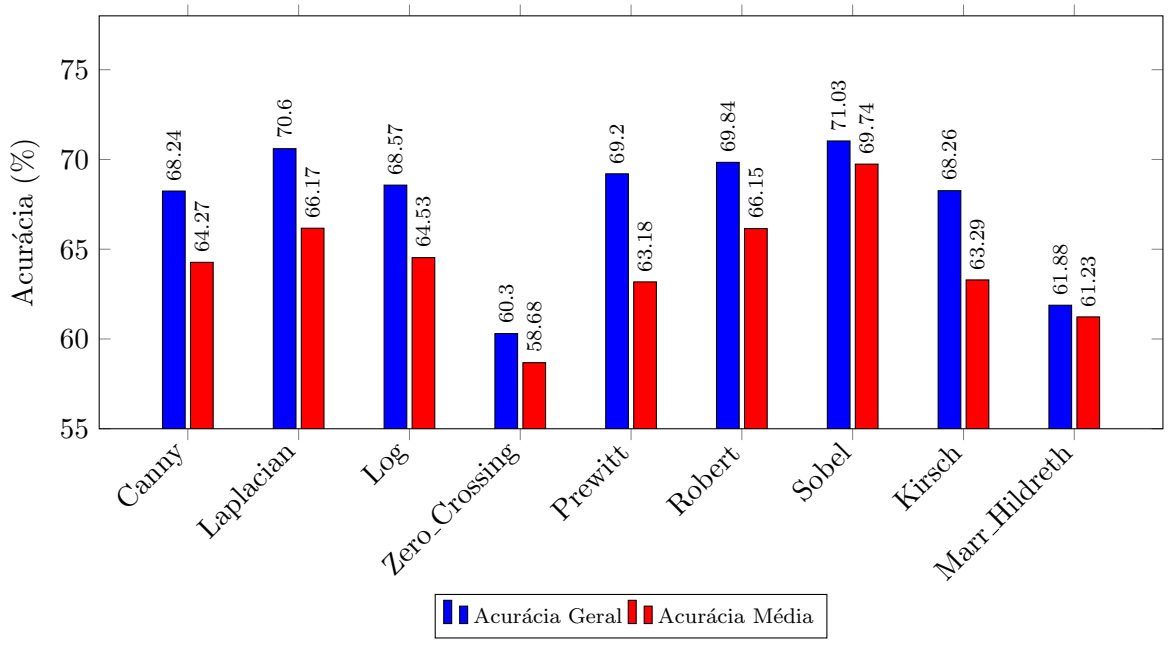
O gráfico de acertos e erros, apresentados nas Figuras 13, mostra que o método Sobel teve o maior número de acertos (71.11%) e o menor número de erros (28.89%). O método Laplacian seguiu com 70.60% de acertos e 29.40% de erros, enquanto o método Zero Crossing apresentou

Tabela 6 – Matrizes de Confusão com Percentuais para Diferentes Métodos

Método	VN (%)	FP (%)	FN (%)	VP (%)
Canny	33.90	15.16	16.63	34.31
Laplacian	35.40	13.66	15.77	35.17
Log	34.16	14.90	16.53	34.41
Zero_Crossing	29.45	19.61	20.09	30.85
Prewitt	34.92	14.14	16.61	34.33
Robert	34.49	14.57	15.59	35.35
Sobel	34.61	14.45	14.45	36.50
Kirsch	33.70	15.36	16.38	34.56
Marr_Hildreth	31.23	17.83	20.27	30.67

Fonte: Próprio autor.

Figura 12 – Comparação de Acurácia Geral e Média para Diferentes Métodos de Detecção de Borda



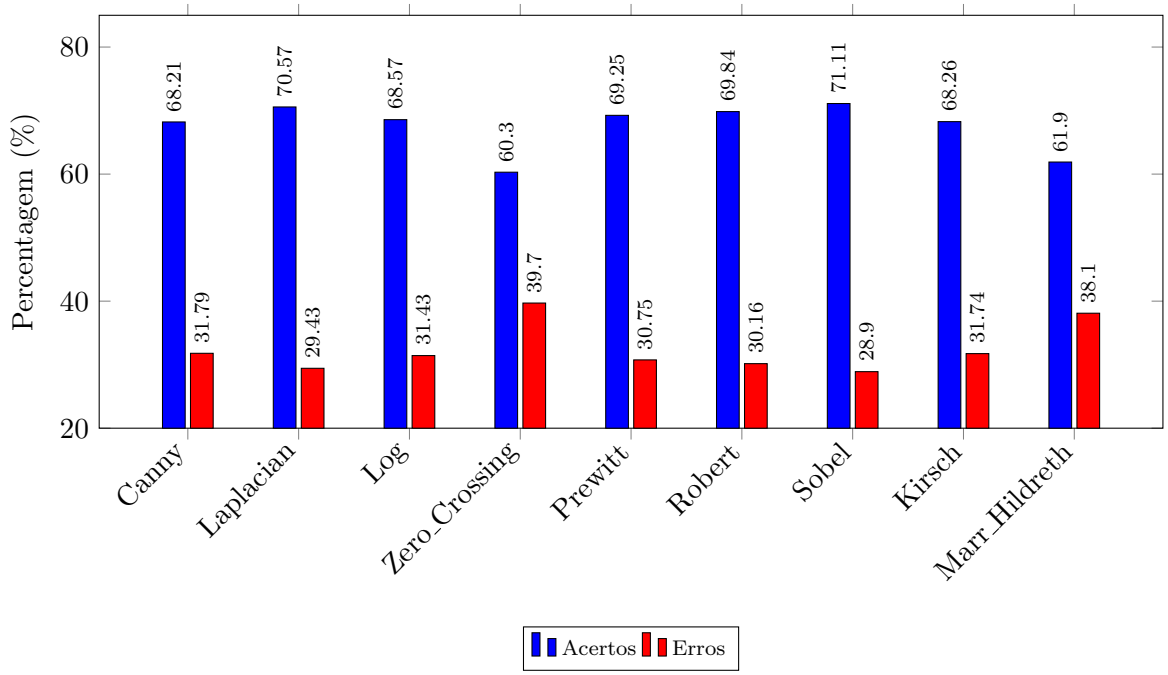
Fonte: Próprio autor.

os piores resultados com 60.30% de acertos e 39.70% de erros.

A análise da sensibilidade (recall) e especificidade, conforme mostrado nas Figuras 15 e 14, revela que o método Sobel também teve a maior sensibilidade (71.44%) e especificidade (70.61%). O método Laplacian seguiu com uma sensibilidade de 69.05% e especificidade de 72.21%, enquanto o método Zero Crossing teve a menor sensibilidade (60.56%) e especificidade (60.03%).

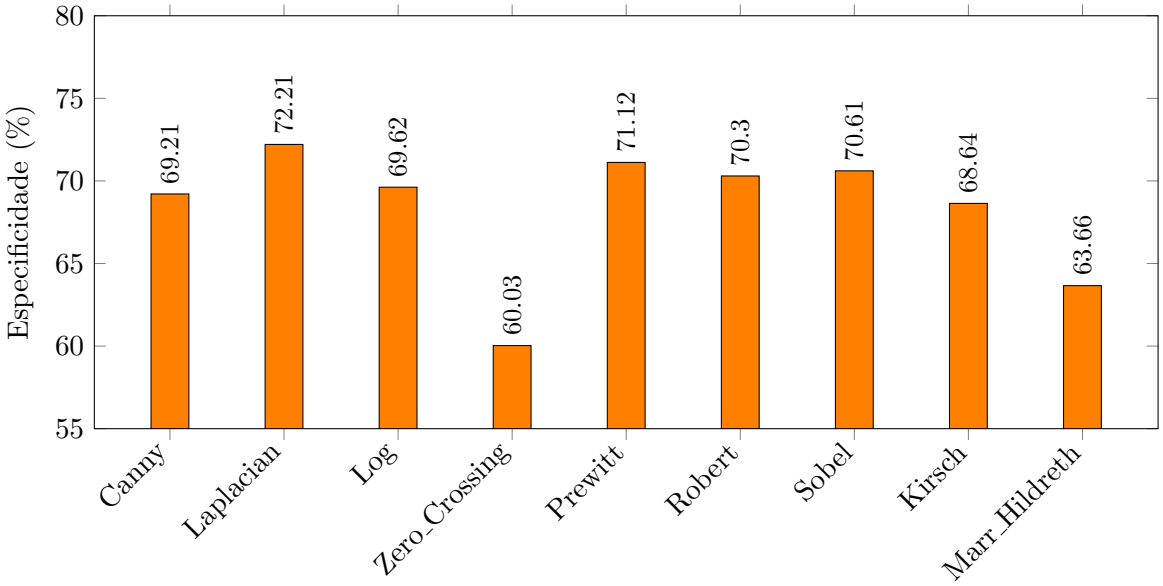
A análise das curvas ROC, apresentada na Figura 16, fornece informações adicionais sobre a performance dos métodos de detecção de bordas. O método Sobel novamente se destaca, com a maior área sob a curva (AUC) de 0.77, indicando um desempenho superior na separação entre as classes de trincas e não trincas. Os métodos Laplacian e Robert também apresentaram bons resultados, ambos com uma AUC de 0.75. Esses resultados reforçam as observações anteriores de que esses métodos são eficazes na detecção de trincas em imagens de concreto. Por outro

Figura 13 – Soma de Acertos e Erros dos Métodos de Detecção de Bordas



Fonte: Próprio autor.

Figura 14 – Especificidade dos Métodos de Detecção de Bordas

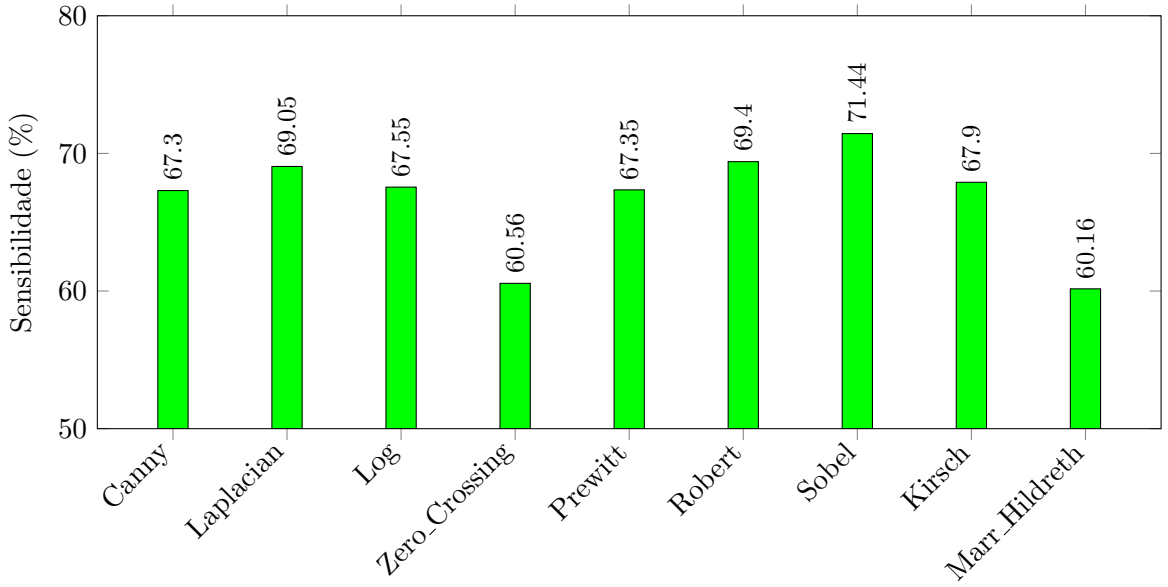


Fonte: Próprio autor.

lado, o método Zero Crossing teve o pior desempenho, com uma AUC de 0.64, evidenciando sua menor eficácia comparativa. Os demais métodos, como Canny, Log, Prewitt e Kirsch, apresentaram AUCs variando entre 0.73 e 0.74, indicando desempenhos moderados.

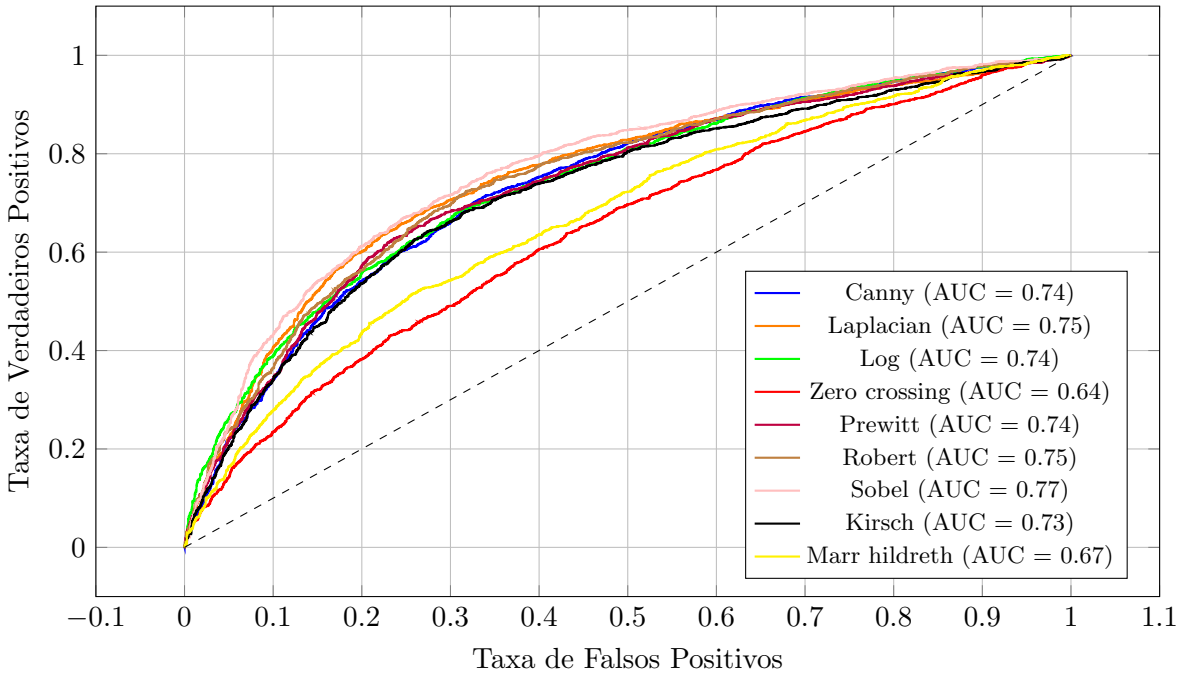
Os dados indicam que o método Sobel é o mais eficaz para a detecção de trincas em imagens de estruturas de concreto, apresentando as melhores métricas de acurácia, F1-Score e AUC. O método Laplacian também se mostrou uma alternativa viável, combinando um bom desempenho com uma boa acurácia. Métodos como Robert e Prewitt demonstraram resultados

Figura 15 – Sensibilidade dos Métodos de Detecção de Bordas



Fonte: Próprio autor.

Figura 16 – Comparação das Curvas ROC para Diferentes Filtros de Suavização



Fonte: Próprio autor.

aceitáveis, enquanto o método Zero Crossing foi o menos eficaz. Essas descobertas destacam a importância da escolha do método de detecção de bordas no pré-processamento de imagens para maximizar a precisão e a eficiência na identificação de trincas em concreto.

O código utilizado para a comparação de técnicas de detecção de bordas na identificação de trincas em imagens via Regressão Logística foi registrado no INPI. A patente, intitulada "Comparação de Técnicas de Detecção de Bordas na Identificação de Trincas em Imagens via Regressão Logística", foi registrada como Programa de Computador sob o número BR512023003631-5,

com data de registro em 20/11/2023, na Instituição de registro: INPI - Instituto Nacional da Propriedade Industrial.

4.4 COMPARAÇÃO DE MÉTODOS DE INTELIGÊNCIA ARTIFICIAL

Nesta seção, são apresentados os resultados da comparação entre diferentes métodos de Inteligência Artificial (IA) utilizados para a detecção de trincas em imagens de estruturas de concreto. Foram analisados três métodos principais: Redes Neurais Convolucionais (CNN), Máquinas de Vetores de Suporte (SVM) e Regressão Logística.

A CNN utilizada foi inspirada na arquitetura LeNet-5 (LeCun *et al.*, 1998). A arquitetura detalhada é apresentada na Tabela 7. Esta CNN foi projetada para extrair automaticamente as características das imagens, facilitando a detecção precisa de trincas em estruturas de concreto. A arquitetura da CNN é composta por três camadas convolucionais seguidas de camadas de *max pooling*, uma camada de achatamento, uma camada densa com função de ativação ReLU, uma camada de *dropout* para evitar *overfitting*, e uma camada densa final com função de ativação Sigmoid.

Tabela 7 – Detalhes da Arquitetura CNN

Camada	**Configuração**	**Saída**	**Parâmetros**
Conv 1	32 filtros, kernel 3x3, ReLU	62x62x32	320
Max Pooling 1	Tamanho de pool 2x2	31x31x32	0
Conv 2	64 filtros, kernel 3x3, ReLU	29x29x64	18,496
Max Pooling 2	Tamanho de pool 2x2	14x14x64	0
Conv 3	128 filtros, kernel 3x3, ReLU	12x12x128	73,856
Max Pooling 3	Tamanho de pool 2x2	6x6x128	0
Flatten	-	4608	0
Densa	512 unidades, ReLU	512	2,359,808
Dropout	Taxa de 0.5	512	0
Densa	1 unidade, Sigmoid	1	513

Fonte: Próprio autor.

A Regressão Logística, por sua vez, é um modelo linear amplamente utilizado para tarefas de classificação binária. No contexto deste estudo, a Regressão Logística foi implementada com a biblioteca *scikit-learn* em Python, configurada para realizar até 20.000 iterações, garantindo a convergência do modelo.

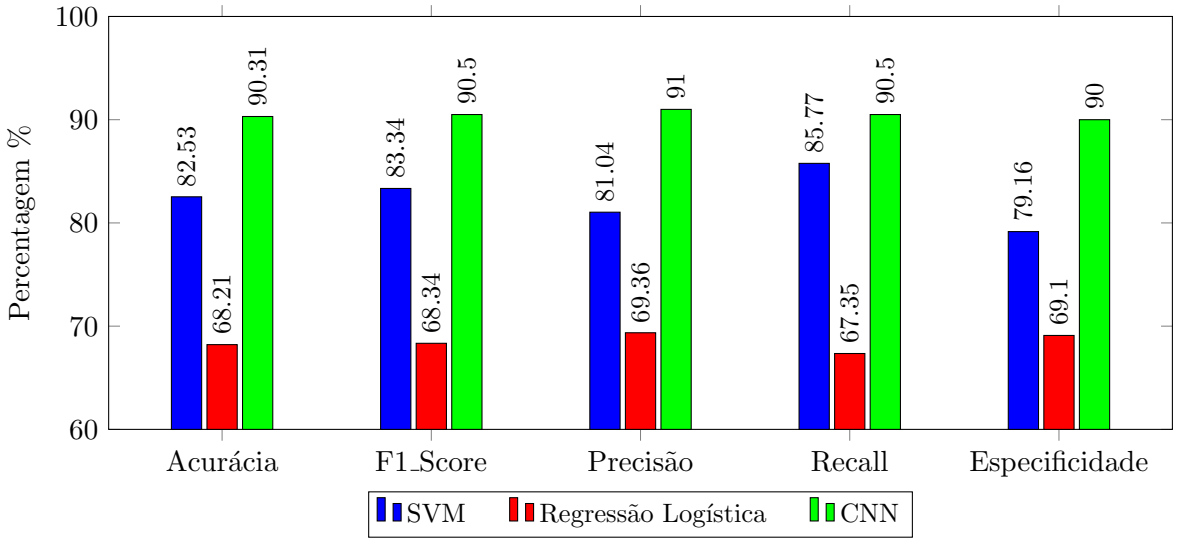
As Máquinas de Vetores de Suporte (SVM) são conhecidas por sua eficácia em espaços de alta dimensionalidade. Para este estudo, a SVM foi implementada utilizando um *pipeline* da biblioteca *scikit-learn*, que incluía a padronização dos dados seguida pela aplicação do classificador SVM.

Os filtros de pré-processamento foram aplicados uniformemente nas imagens antes de serem alimentadas nos modelos de IA. As imagens foram inicialmente convertidas para escala de cinza utilizando o método *Gligtness*, seguidas pela aplicação do filtro de suavização mediana e, em seguida, pelo filtro de borda Canny. Esta padronização permitiu uma comparação justa

entre os diferentes métodos de IA. A seguir, são apresentados os resultados detalhados da comparação de desempenho entre CNN, SVM e Regressão Logística.

A análise das métricas principais, mostrada na Figura 17, confirma que a CNN superou os outros métodos em termos de acurácia, F1-Score e precisão. A acurácia da CNN foi de 90.31%, significativamente maior que a da SVM (82.53%) e da Regressão Logística (68.21%). O F1-Score da CNN foi de 90.50%, enquanto o da SVM foi de 83.34% e o da Regressão Logística foi de 68.34%. Esses resultados sugerem que a CNN não apenas identificou com mais precisão as trincas, mas também foi mais eficaz em manter um equilíbrio entre precisão e *recall*.

Figura 17 – Comparação de métricas de avaliação entre CNN, Regressão Logística e SVM



Fonte: Próprio autor.

A Figura 17 revela que a CNN apresentou a melhor performance em termos de sensibilidade (*recall*) e especificidade. A CNN alcançou uma sensibilidade de 90.50%, superando a SVM (85.77%) e a Regressão Logística (67.35%). Em relação à especificidade, a CNN obteve 90.00%, enquanto a SVM teve 79.16% e a Regressão Logística 69.10%. Esses resultados indicam que a CNN é mais eficaz tanto em identificar corretamente as trincas quanto em evitar falsos positivos.

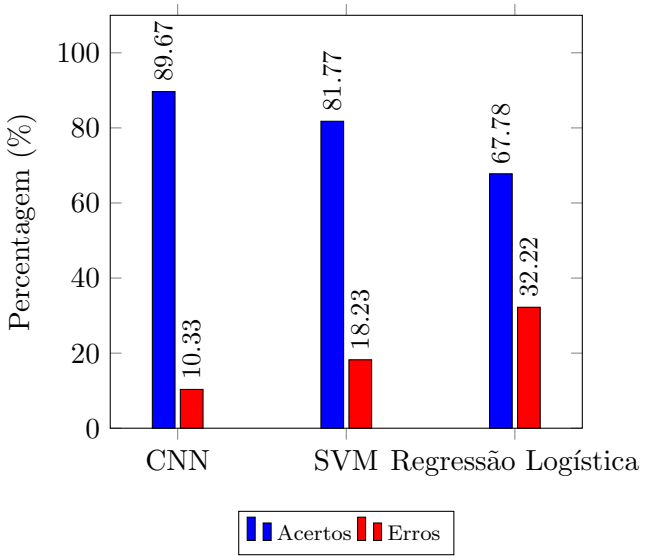
As matrizes de confusão na Tabela 8 revelam que a CNN foi mais precisa tanto na identificação de imagens com trincas (VP = 85.00%) quanto na exclusão de falsos casos (VN = 93.43%), enquanto a SVM e a Regressão Logística tiveram mais dificuldades, evidenciado pelo número maior de falsos negativos e, no caso da SVM, falsos positivos também. A Tabela 8 compara as porcentagens de verdadeiros positivos (VP), falsos positivos (FP), verdadeiros negativos (VN) e falsos negativos (FN) para os três métodos.

Tabela 8 – Comparação da Matriz de Confusão em Percentagem

Método	**VP (%)**	**FP (%)**	**VN (%)**	**FN (%)**
CNN	85.00	5.00	93.43	6.57
SVM	82.53	18.96	81.04	14.23
Regressão Logística	68.21	15.00	67.35	32.65

Fonte: Próprio autor.

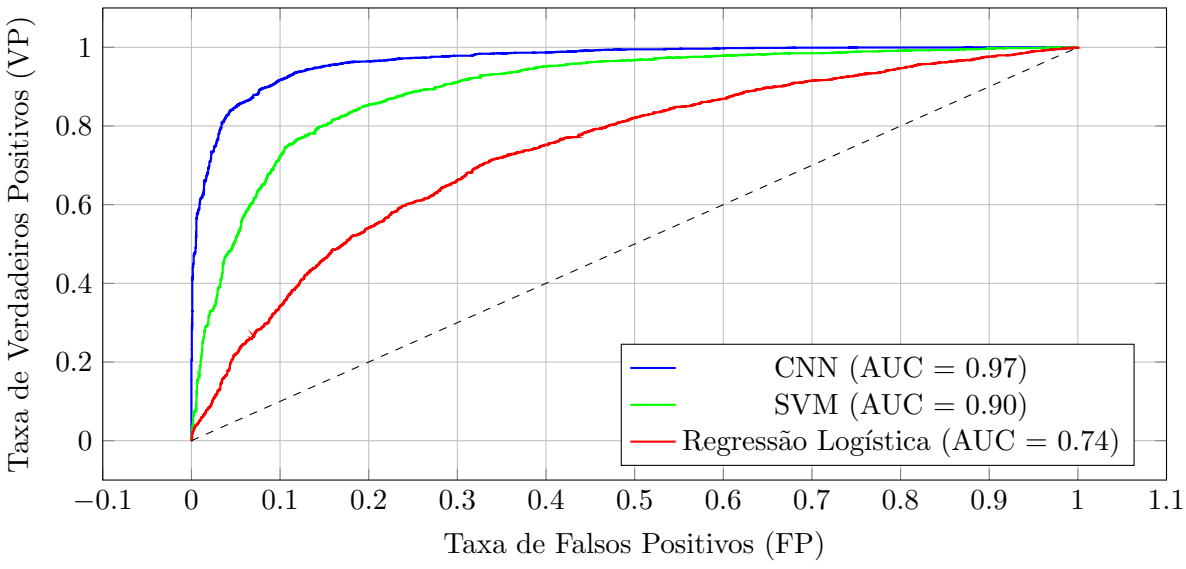
Figura 18 – Acertos e Erros em Percentagem para Diferentes Métodos de Classificação



Fonte: Próprio autor.

A Figura 18 mostra a comparação dos acertos e erros em percentagem para cada método de classificação. Os resultados indicam que a CNN apresentou a maior percentagem de acertos (89.67%) e a menor percentagem de erros (10.33%), seguida pela SVM com 81.77% de acertos e 18.23% de erros, e pela Regressão Logística com 67.78% de acertos e 32.22% de erros. Isso evidencia a superioridade da CNN na classificação correta das imagens.

Figura 19 – Comparação das Curvas ROC para CNN, SVM e Regressão Logística



Fonte: Próprio autor.

As curvas ROC dos métodos mostradas na Figura 19 fornecem uma visualização clara da performance dos métodos. A curva ROC da CNN, com uma AUC de 0.97, indica uma excelente taxa de verdadeiros positivos em relação à taxa de falsos positivos. A SVM apresentou uma AUC de 0.90, enquanto a Regressão Logística obteve uma AUC de 0.74. De acordo com Gorunescu (2011) e Consoli, Recupero e Petković (2019), a AUC da Regressão Logística

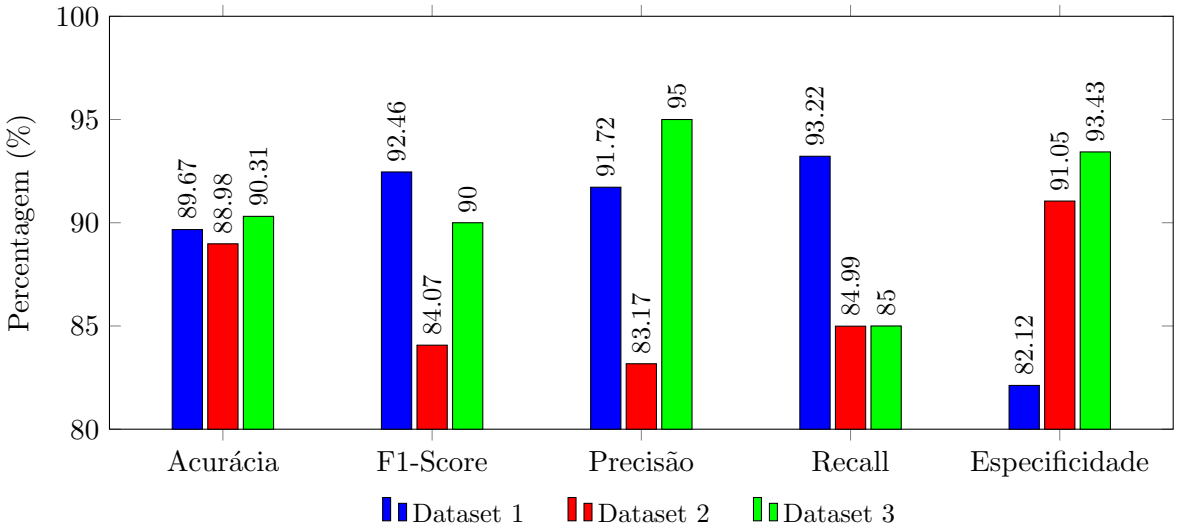
é classificada como razoável, enquanto as AUCs da CNN e da SVM são classificadas como excelentes. Portanto, todas as simulações com os diferentes métodos de IA demonstraram um desempenho satisfatório, destacando a eficácia geral dos métodos utilizados neste estudo. A superioridade da CNN em minimizar falsos positivos e negativos a estabelece como a técnica mais confiável para a detecção de trincas em imagens.

Em resumo, os resultados indicam que a CNN é a escolha mais robusta para a detecção de trincas em imagens, apresentando melhor desempenho em quase todas as métricas analisadas. A superioridade da CNN é evidenciada tanto nas matrizes de confusão quanto nas curvas ROC, além dos gráficos de acertos e erros.

4.5 ANÁLISE DA QUANTIDADE DE IMAGENS PARA DETECÇÃO DE TRINCAS

A avaliação de desempenho do modelo de Rede Neural Convolucional (CNN), como refletido na Figura 20, demonstra como diferentes composições de conjuntos de dados influenciam a capacidade do modelo em detectar trincas em estruturas de concreto. Para investigar essa influência, três configurações distintas de conjuntos de dados foram utilizadas, denominadas Dataset 1, Dataset 2 e Dataset 3.

Figura 20 – Comparação das Métricas de Desempenho entre os *Datasets*



Fonte: Próprio autor.

Na Configuração do Dataset 1, havia um total de 9859 imagens positivas (com trincas) e 4900 imagens negativas (sem trincas). A divisão dos dados foi feita da seguinte forma: para treinamento, foram utilizadas 10331 imagens, das quais 6872 eram positivas (66.52%) e 3459 negativas (33.48%); para teste, foram usadas 2952 imagens, sendo 2007 positivas (67.99%) e 945 negativas (32.01%); e para validação, foram utilizadas 1476 imagens, com 980 positivas (66.40%) e 496 negativas (33.60%).

A Configuração do Dataset 2 inverteu a proporção de imagens positivas e negativas em relação ao Dataset 1. Nesta configuração, havia 4900 imagens positivas e 9800 imagens

negativas. A divisão dos dados seguiu o mesmo padrão do Dataset 1, mas com as proporções invertidas.

Na Configuração do Dataset 3, a distribuição das imagens foi equilibrada, totalizando 9859 imagens positivas e 9800 imagens negativas. A proporção de imagens para cada conjunto foi a seguinte: 13761 imagens (70.00%) para treinamento, 3932 imagens (20.00%) para teste e 1966 imagens (10.00%) para validação.

Na Configuração do Conjunto de Dados 1, a acurácia geral alcançou 89.67%, com uma precisão de 91.72% e um *recall* de 93.22%. A Configuração do Conjunto de Dados 2 apresentou uma acurácia ligeiramente menor, em 88.98%, com uma precisão de 83.17% e um *recall* de 84.99%. Já a Configuração do Conjunto de Dados 3 obteve a maior acurácia geral de 90.31%, com uma precisão de 95.00% e um *recall* de 85.00%.

A Configuração do Conjunto de Dados 3 emergiu como a mais eficiente, apresentando a maior acurácia geral de 90.31% e as menores taxas de erro, sendo 5.00% de falsos positivos e 4.00% de falsos negativos. A análise enfatiza a importância de uma distribuição equilibrada de imagens no conjunto de dados, destacando que uma composição balanceada contribui para a melhoria da acurácia e confiabilidade do modelo.

Tabela 9 – Matriz de Confusão em Percentagens para Todos os Conjuntos de Dados

Conjunto de Dados	**VP (%)**	**FP (%)**	**VN (%)**	**FN (%)**
Dataset 1	93.22	17.88	82.12	6.78
Dataset 2	84.99	8.95	91.05	15.01
Dataset 3	95.00	5.00	96.00	4.00

Fonte: Próprio autor.

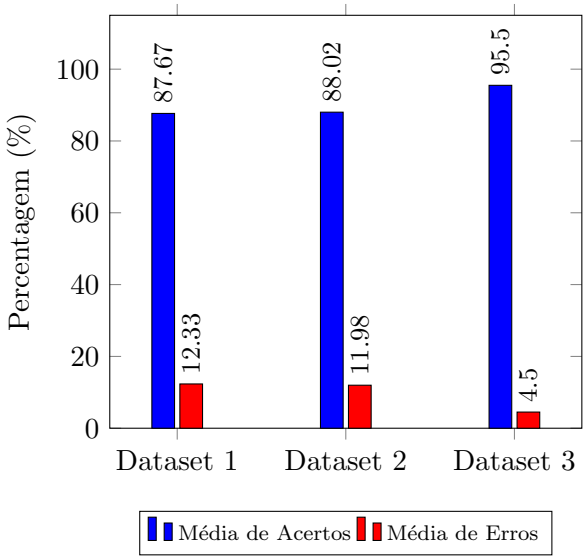
A média de acertos e erros, conforme mostrado na Figura 21, destaca a superioridade do Conjunto de Dados 3, com 95.50% de acertos e 4.50% de erros. Comparativamente, o Conjunto de Dados 1 teve 87.67% de acertos e 12.33% de erros, e o Conjunto de Dados 2 apresentou 88.02% de acertos e 11.98% de erros. Esses resultados confirmam a eficiência e precisão do Conjunto de Dados 3 na detecção de trincas.

Para a análise de sensibilidade e especificidade, o Conjunto de Dados 1 alcançou uma sensibilidade de 93.22% e uma especificidade de 82.12%, indicando forte detecção de trincas, mas com mais falsos positivos. O Conjunto de Dados 2 teve uma sensibilidade de 84.99% e uma alta especificidade de 91.05%. O Conjunto de Dados 3 apresentou uma sensibilidade de 85.00% e a maior especificidade de 93.43%.

A acurácia das configurações dos conjuntos de dados também foi avaliada por meio de validação cruzada, que poder ser observado na Figura 22, onde se observou um aumento progressivo na acurácia, com a Configuração 1 alcançando 91.49%, a Configuração 2 atingindo 90.91% e a Configuração 3 obtendo 92.48%. Esses incrementos sucessivos confirmam que ajustes estratégicos na composição do conjunto de dados podem otimizar significativamente o desempenho da CNN, o que é vital para a classificação precisa em aplicações de monitoramento da integridade estrutural.

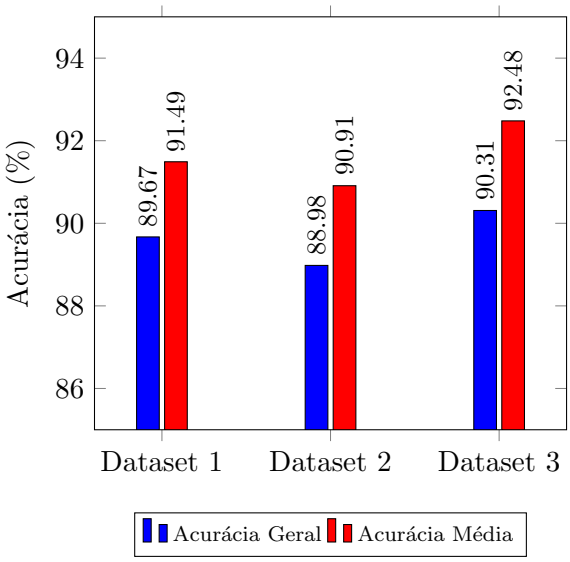
A análise da curva ROC para as três configurações de conjuntos de dados, ilustrada na Figura 23, mostra que a capacidade do modelo CNN de discriminar entre imagens com e sem trincas varia ligeiramente com a composição do conjunto de dados. Os valores de AUC foram

Figura 21 – Média de Acertos e Erros em Percentagem para Diferentes Conjuntos de Dados



Fonte: Próprio autor.

Figura 22 – Comparação da Acurácia Geral e Acurácia Média

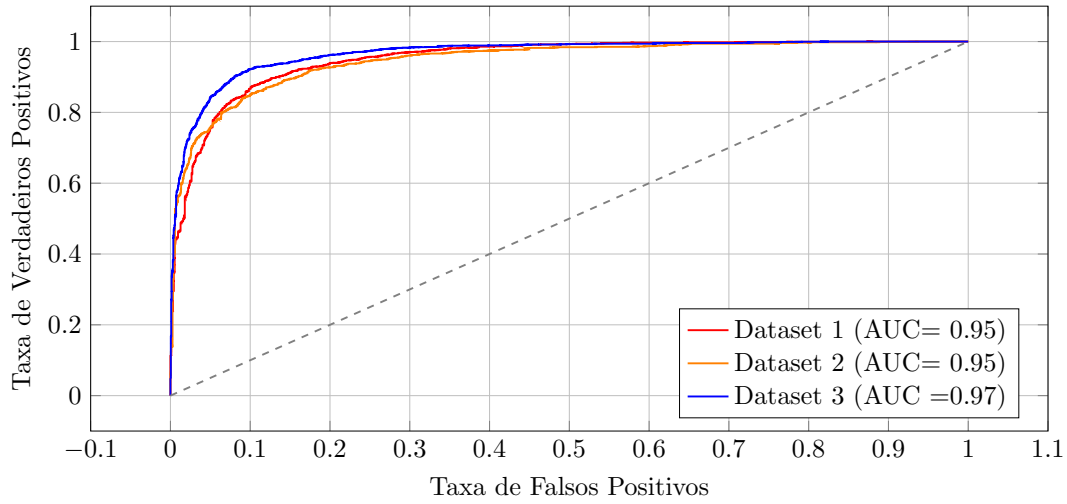


Fonte: Próprio autor.

0.95 para as duas primeiras configurações, que predominantemente apresentam imagens "Sem Trinca"e "Com Trinca", respectivamente. No entanto, a terceira configuração, que apresenta uma distribuição equilibrada de imagens, alcançou um valor AUC superior de 0.97. Este resultado destaca a influência positiva de um equilíbrio adequado de classes no conjunto de dados, demonstrando que uma proporção equilibrada melhora significativamente a precisão diagnóstica do modelo, tornando-o mais eficaz para aplicações práticas em monitoramento da saúde estrutural.

As curvas de aprendizado nas Figuras 24 e 25 mostram a generalização do modelo CNN com diferentes configurações de conjuntos de dados. A Figura 24 indica um aumento consistente na acurácia de treinamento e validação, especialmente na Configuração 3, onde as curvas

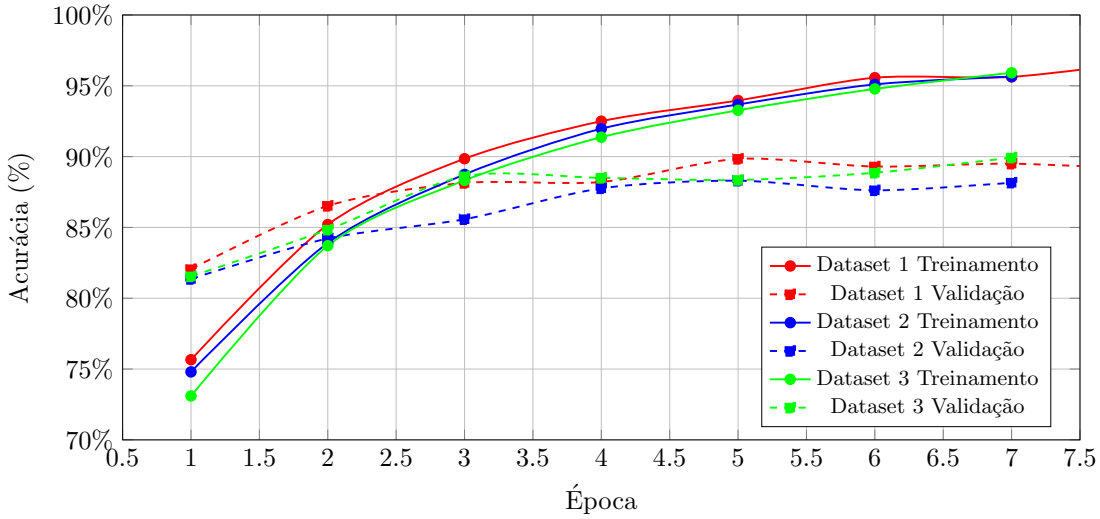
Figura 23 – Comparação das Curvas ROC para Diferentes *Datasets*.



Fonte: Próprio autor.

permanecem próximas, sinalizando equilíbrio entre aprendizado e generalização.

Figura 24 – Curvas de Aprendizado da Acurácia Comparativas para Três *Datasets*.

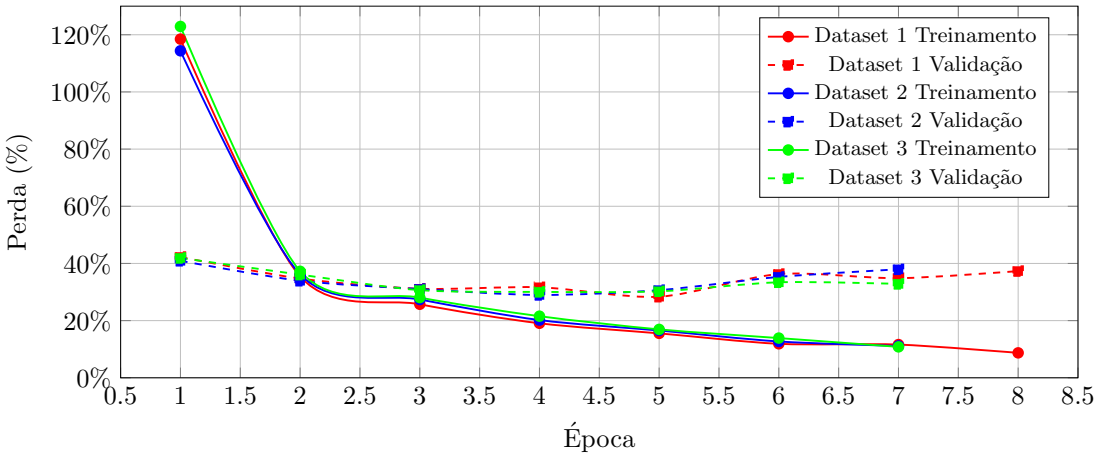


Fonte: Próprio autor.

As curvas de perda na Figura 25 mostram uma rápida diminuição na perda de treinamento, estabilizando em um valor baixo. Na Configuração 3, o alinhamento entre a perda de treinamento e validação sugere menor tendência ao *overfitting*, indicando que o modelo está aprendendo de forma eficaz, capturando padrões generalizáveis sem reter excessivamente as peculiaridades do conjunto de treinamento.

O modelo CNN demonstra alta eficiência na detecção de trincas, com a Configuração do Conjunto de Dados 3 provando ser a configuração mais equilibrada e eficaz. A melhoria progressiva no desempenho ao longo dos conjuntos de dados destaca a importância da composição do conjunto de dados. Futuras melhorias podem se concentrar em refinar ainda mais o equilíbrio do conjunto de dados e explorar técnicas adicionais de pré-processamento de imagens para aumentar a acurácia e robustez do modelo.

Figura 25 – Curvas de Aprendizado da Perda Comparativas para Três *Datasets*.



Fonte: Próprio autor.

Adicionalmente, parte dos resultados desta secção foi apresentada no VIII Encontro Regional de Matemática Aplicada e Computacional da Regional 9 (ERMAC) (Santos *et al.*, 2024), promovido pela Sociedade Brasileira de Matemática Aplicada e Computacional (SBMAC), demonstrando a aplicabilidade do modelo para o monitoramento da integridade estrutural.

5 CONCLUSÃO

A detecção de trincas em estruturas de concreto é crucial para o monitoramento de integridade estrutural (SHM), garantindo a segurança e durabilidade de infraestruturas essenciais como pontes, túneis e edifícios. Se não detectadas precocemente, trincas podem levar a falhas estruturais significativas, resultando em altos custos de reparo e riscos à segurança pública. Este estudo propôs uma abordagem combinando o uso de veículos aéreos não tripulados (VANTs) com técnicas avançadas de Inteligência Artificial para aprimorar a precisão e eficiência na detecção de trincas.

Os principais objetivos desta pesquisa foram determinar uma sequência adequada de pré-processamento de imagens e aplicar uma técnica de inteligência artificial eficaz para a segmentação de imagens visando à detecção de danos estruturais. Todos os objetivos foram alcançados, resultando em contribuições significativas para a área de SHM.

A coleta de imagens aéreas foi realizada utilizando o VANT DJI Phantom 3 Standard, que se mostrou eficaz na criação de um *dataset* diversificado e abrangente. A criação de um *dataset* de qualidade é fundamental para o treinamento e validação de algoritmos de detecção de trincas. Com a aplicação de técnicas de aumento de dados, foi gerado um total de 19.600 imagens, equilibrando a quantidade de imagens com e sem trincas, proporcionando uma base sólida para o desenvolvimento de algoritmos de detecção.

Na etapa de pré-processamento, a investigação dos filtros de suavização revelou que o filtro Gaussiano foi o mais eficaz, superando os filtros de Média e Mediana. O filtro Gaussiano apresentou melhor desempenho em métricas-chave como acurácia (71,13%), F1-Score (71,07%), precisão (72,60%) e especificidade (73,00%). Além disso, a taxa de acertos foi a maior (35,57%), enquanto a taxa de erros foi a menor (14,43%), destacando-se como a escolha mais eficaz para melhorar a qualidade das imagens e facilitar a detecção de trincas.

A análise das técnicas de detecção de bordas indicou que o método Sobel era o mais eficaz para identificar trincas em imagens de concreto. O método Sobel apresentou as melhores métricas de acurácia (71,03%), F1-Score (71,53%) e AUC (0,77), além de uma alta sensibilidade (71,44%) e especificidade (70,61%). Este método também obteve a maior taxa de acertos (71,11%) e a menor taxa de erros (28,89%), tornando-se a opção mais confiável para o pré-processamento das imagens.

Na comparação das técnicas de Inteligência Artificial, as Redes Neurais Convolucionais (CNN) superaram significativamente as Máquinas de Vetores de Suporte (SVM) e a Regressão Logística. A CNN alcançou a maior acurácia (90,31%), F1-Score (90,50%) e precisão (91,00%), destacando-se como a técnica mais eficaz e confiável para a detecção de trincas. As curvas ROC e as matrizes de confusão corroboraram esses resultados, com a CNN apresentando uma

sensibilidade de 90,50% e especificidade de 90,00%. A taxa de acertos foi a mais alta, enquanto a taxa de erros foi a menor, evidenciando a capacidade da CNN em minimizar falsos positivos e falsos negativos.

Finalmente, a análise da influência da quantidade de imagens do *dataset* destacou a importância de uma distribuição equilibrada para otimizar o desempenho da CNN. A Configuração do Conjunto de Dados 3, com uma distribuição balanceada, obteve a maior acurácia (90,31%), maior precisão (95,00%) e uma sensibilidade de 85,00%, além de apresentar a menor média de erros (4,50%) e a maior média de acertos (95,50%). Essas descobertas reforçam a eficácia de uma composição balanceada do *dataset* para melhorar a precisão e a robustez do modelo.

Em particular, este estudo a viabilidade e a eficácia do uso de VANTs na inspeção de estruturas de concreto, e também forneceu uma metodologia detalhada e comprovada para o desenvolvimento de algoritmos de detecção de trincas. A abordagem integrada de coleta de dados, pré-processamento de imagens e aplicação de técnicas de Inteligência Artificial resultou em um sistema de monitoramento de alta precisão, eficiência e custo-benefício. Essas contribuições representam um avanço significativo na modernização dos métodos de SHM, promovendo maior segurança e durabilidade das infraestruturas.

5.1 SUGESTÕES PARA TRABALHOS FUTUROS

Nesta seção são apresentadas sugestões para trabalhos futuros, visando a continuidade da pesquisa:

- **Desenvolvimento de Modelos 3D:** Implementar técnicas de reconstrução 3D para criar modelos detalhados das estruturas de concreto, permitindo uma visualização e análise mais precisa das trincas e outros defeitos estruturais.
- **Classificação de Tipos de Danos:** Desenvolver algoritmos que não apenas detectem a presença de danos, mas também classifiquem os tipos de defeitos, como vergalhão exposto, trincas, infiltrações por umidade, entre outros.
- **Análise com Diferentes Modelos de VANTs e Câmeras:** Realizar estudos utilizando diversos modelos de VANTs e tipos de câmeras para avaliar a influência das especificações técnicas dos equipamentos na qualidade da detecção e classificação de trincas.
- **Treinamento com Múltiplos *dataset*s:** Ampliar o treinamento dos modelos de inteligência artificial com diferentes *dataset*s, incluindo imagens de asfalto, trilhos de ferrovias, entre outros, para verificar a robustez e a generalização dos algoritmos em diversas aplicações.
- **Avaliação de Técnicas de Fusão de Dados:** Investigar técnicas de fusão de dados para combinar informações de imagens aéreas com dados de outras fontes, como sensores de vibração ou medições geotécnicas, a fim de proporcionar uma análise mais completa e precisa da saúde estrutural.

REFERÊNCIAS

ABER, J. S.; MARZOLFF, I.; RIES, J. B. **Small Format Aerial Photography. Principles, Techniques and Geoscience Applications**. 1. ed. Emporia, KS, USA: Elsevier Science, 2010. 261 p. ISSN 2060-3274. ISBN 9780444532602.

AFIFI, M. *et al.* Learning multi-scale photo exposure correction. In: **Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)**. Nashville, TN, USA: IEEE, 2021. p. 9153–9163. ISBN 978-1-6654-4509-2. ISSN 1063-6919.

AHMAD, I.; MOON, I.; SHIN, S. J. Color-to-grayscale algorithms effect on edge detection - a comparative study. In: **2018 International Conference on Electronics, Information, and Communication (ICEIC)**. Honolulu, HI, USA: IEEE, 2018. DOI: 10.23919/ELINFOCOM.2018.8330719.

ALI, L. *et al.* Performance evaluation of different algorithms for crack detection in concrete structures. In: **Proceedings of 2nd International Conference on Computation, Automation and Knowledge Management, ICCAKM 2021**. Abu Dhabi, UAE: Institute of Electrical and Electronics Engineers Inc., 2021. p. 53–58. ISBN 9781728194912. DOI: 10.1109/ICCAKM50778.2021.9357717.

ANSARI, M. A.; KURCHANIYA, D.; DIXIT, M. A Comprehensive Analysis of Image Edge Detection Techniques. **International Journal of Multimedia and Ubiquitous Engineering**, Yogjakarta, Indonesia, v. 12, n. 11, p. 1–12, 2017. ISSN 19750080.

ARANDA, L. A.; REVIRIEGO, P.; MAESTRO, J. A. Error detection technique for a median filter using denoising algorithm. **Transactions on Nuclear Science**, IEEE, Madrid, Spain, v. 30, p. 2219–2226, 8 2017. DOI: 10.1109/TNS.2017.2666843.

ARTERO, A. O.; TOMMASELLI, A. M. G. Detecção e afinamento de bordas em direções previamente conhecidas. **Boletim de Ciencias Geodesicas**, Curitiba, Paraná, Brazil, v. 15, p. 155–177, 6 2009. ISSN 14134853.

ATKISON, K. B. **Close Range Photogrammetry and Machine Vision**. 1. ed. Caithness, Scotland: Whittles Publishing Services, 1996. 382 p. ISBN 187032546X.

BISWAS, D. *et al.* Novel gray scale conversion techniques based on pixel depth. **Journal of Global Research in Computer Science**, West Bengal, India, v. 2, 1 2011. DOI: 10.23919/ELINFOCOM.2018.8330719.

BRERETON, R. G.; LLOYD, G. R. Support vector machines for classification and regression. **Analyst**, Royal Society of Chemistry, Bristol, UK, v. 135, p. 230–267, 2009. ISSN 1364-5528.

CANNY, J. A Computational Approach to Edge Detection. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Cambridge, MA, USA, PAMI-8, n. 6, p. 679–698, 1986. ISSN 01628828.

CHA, Y. J.; CHOI, W.; BüYüKöZTüRK, O. Deep learning-based crack damage detection using convolutional neural networks. **Computer-Aided Civil and Infrastructure Engineering**, Blackwell Publishing Inc., Hoboken, NJ, United States, v. 32, p. 361–378, 5 2017. ISSN 14678667.

CHADHA., B.; JUWE., S. **Agile Machine Learning with DataRobot: Automate Each Step of the Machine Learning Life Cycle, From Understanding Problems to Delivering Value**. 1. ed. Birmingham,UK: Packt Publishing, 2021. v. 1. -511 p. ISBN 9781801076807.

CHAIYASARN, K. *et al.* Crack detection in masonry structures using convolutional neural networks and support vector machines. In: **35th International Symposium on Automation and Robotics in Construction (ISARC 2018)**. Berlin, Germany: The International Association for Automation and Robotics in Construction, 2018. ISBN 978-3-00-060855-1.

CHANG, C. C.; HSIAO, J. Y.; HSIEH, C. P. An adaptive median filter for image denoising. In: **Proceedings - 2008 2nd International Symposium on Intelligent Information Technology Application, IITA 2008**. Xi'an, China: IEEE, 2008. v. 2, p. 346–350. ISBN 9780769534978.

CHARALAMPIDIS, D. Recursive implementation of the Gaussian filter using truncated cosine functions. **IEEE Transactions on Signal Processing**, Athens, Greece, v. 64, n. 14, p. 3554–3565, 2016. ISSN 1053587X.

CONSOLI, S.; RECUPERO, D. R.; PETKOVIć, M. **Data Science for Healthcare: Methodologies and Applications**. Eindhoven, The Netherlands: Springer International Publishing, 2019. 1-367 p. ISBN 9783030052492.

DAIS, D. *et al.* Automatic crack classification and segmentation on masonry surfaces using convolutional neural networks and transfer learning. **Automation in Construction**, Leeds, UK, v. 125, 2021. DOI: 10.1016/j.autcon.2021.103606.

DAS, S. *et al.* A comparative study of different noise filtering techniques in digital images. **International Journal of Engineering Research and General Science**, Dimapur, India, v. 3, n. 5, p. 180–191, 2015. ISSN 2091-2730.

DENG, X. *et al.* An improved method to construct basic probability assignment based on the confusion matrix for classification problem. **Information Sciences**, Elsevier Inc., Chongqing, China, v. 340-341, p. 250–261, 2016. ISSN 00200255.

DIBIKE, Y. B. *et al.* Model induction with support vector machines: Introduction and applications. **Journal of Computing in Civil Engineering**, ASCE - American Society of Civil Engineers, Reston, Virginia, EUA, v. 15, n. 3, p. 208–216, 6 2001. DOI: 10.1061/ASCE0887-3801(2001)15:3(208).

DING, C. *et al.* Performance prediction for a fuel cell air compressor based on the combination of backpropagation neural network optimized by genetic algorithm (GA-BP) and support vector machine (SVM) algorithms. **Thermal Science and Engineering Progress**, Elsevier Ltd, Changsha, China, v. 44, n. August, p. 102070, 2023. ISSN 24519049.

DISTANTE, A.; DISTANTE, C. **Handbook of image processing and computer vision: Volume 2: From image to pattern**. Lecce, Italy: Springer, 2020. v. 2. 1–431 p. ISBN 9783030423742.

FACELLI, K. *et al.* **Inteligência Artificial: Uma Abordagem de Aprendizado de Máquina**. 1. ed. Rio de Janeiro, Brasil: LTC - Livros Técnios e Cientficicos Editora Ltda., 2011. 396 p. ISBN 978-85-216-1880-5.

FIGUEIREDO, E. O.; FIGUEIREDO, S. M. d. M. **Planos de voo semiautônomos para fotogrametria com aeronaves remotamente pilotadas de classe 3**. Rio Branco, Brasil: Embrapa, 2018. 56 p. ISSN 0100-9915.

FILHO, O. M.; NETO, H. V. **Processamento Digital de Imagens**. 1ª. ed. Rio de Janeiro, Brasil: Brasport, 1999. v. 66. 331 p. ISSN 17730627. ISBN 8574520098.

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**, Elsevier Ltd, London, Canada, v. 114, 11 2020. ISSN 09589465.

FUDE, G. *et al.* Edge detection based on wavelet analysis with gaussian filter. In: **Proceedings - 1st International Congress on Image and Signal Processing, CISP 2008**. Hainan, China: IEEE, 2008. v. 2, p. 724–728. ISBN 9780769531199.

FUJITA, Y.; MITANI, Y.; HAMAMOTO, Y. A method for crack detection on a concrete structure. In: **18th International Conference on Pattern Recognition (ICPR'06)**. Hong Kong, China: IEEE, 2006. ISBN 0-7695-2521-0.

GHASSOUN, Y. *et al.* Implementation and validation of a high accuracy uav-photogrammetry based rail track inspection system. **Remote Sensing**, Braunschweig, Germany, v. 13, p. 1–22, 2021. ISSN 20724292.

GLADENCE, M. *et al.* A statistical comparison of logistic regression and different bayes classification methods for machine learning. **ARPN Journal of Engineering and Applied Sciences**, Tamil Nadu, India, v. 10, 2015. ISSN 1819-6608.

GONZALEZ, R. C.; WOODS, R. E. **Processamento Digital De Imagens**. 3. ed. Knoxville, EUA: Pearson, 2010. 624 p. ISBN 9788581435862.

GORUNESCU, F. **Data Mining: Concepts, Models and Techniques**. 1. ed. Craiova, Romania: Springer, 2011. v. 12. -364 p. ISSN 1868-4394. ISBN 9783642017988.

GU, J. *et al.* Coded rolling shutter photography: Flexible space-time sampling. In: **2010 IEEE International Conference on Computational Photography, ICCP 2010**. Cambridge, MA, USA: IEEE, 2010. p. 0–7. ISBN 9781424470235.

GUPTA, G. Algorithm for image processing using improved median filter and comparison of mean, median and improved median filter. In: **International Journal of Soft Computing**. Jaipur, India: Blue Eyes Intelligence Engineering and Sciences Publication, 2011. v. 1, p. 304–311. ISSN 2231-2307.

HALL, M. **Digital Photography - Mastering Aperture, Shutter Speed, Iso and Exposure**. 1. ed. Penrith Cumbria, UK: JVzon Studio, 2015. 43 p. ISBN 9780760768587.

HAORAN, C. *et al.* Egg crack detection based on support vector machine. In: **Proceedings - 2020 International Conference on Intelligent Computing and Human-Computer Interaction, ICHCI 2020**. [S.l.]: Institute of Electrical and Electronics Engineers Inc., 2020. p. 80–83. ISBN 9781665423168.

HINTON, G. E. *et al.* Improving neural networks by preventing co-adaptation of feature detectors. **arXiv preprint arXiv:1207.0580**, Toronto, Canada, p. 1–18, 2012. DOI: https://doi.org/10.48550/arXiv.1207.0580.

HLOTOV, V.; HUNINA, A.; SIEJKA, Z. Accuracy Investigation of Creating Orthophotomaps Based on Images Obtained by Applying Trimble-UX5 UAV. **Reports on Geodesy and Geoinformatics**, Warsaw, Poland, v. 103, n. 1, p. 106–118, 2017. DOI: 10.1515/rgg-2017-0009.

HOANG, N. D.; NGUYEN, Q. L. Metaheuristic optimized edge detection for recognition of concrete wall cracks: A comparative study on the performances of roberts, prewitt, canny, and sobel algorithms. **Advances in Civil Engineering**, Da Nang, Vietnam, v. 2018, 2018. ISSN 16878094. DOI: 10.1155/2018/7163580.

HOSHYAR, A. N.; KHARKOVSKY, S.; SAMALI, B. Statistical features and traditional sa-svm classification algorithm for crack detection. **Journal of Signal and Information Processing**, Scientific Research Publishing, v. 09, p. 111–121, 2018. ISSN 2159-4465.

HSU, P.; CHEN, B.-Y. Blurred image detection and classification. In: **14th International Multimedia Modeling Conference, MMM 2008**. Kyoto, Japan: Springer, 2008. p. 277–286. ISBN 0302-9743.

HUSSAIN, Z.; AGARWAL, D. A Comparative Analysis of Edge Detection Techniques Used in Flame Image Processing. **International Journal of Advance Research in Science and Engineering**, Mathura, India, v. 4, n. 4, p. 1335–1343, 2015. ISSN 2319-8354(E).

IGUAL, L.; SEGUí, S. **Introduction to Data Science: A Python Approach to Concepts, Techniques and Applications**. Cham, Switzerland: Springer, 2017. v. 1. -227 p. ISSN 1863-7310. ISBN 978-3-319-50017-1.

IZBICKI, R.; SANTOS, T. M. dos. **Aprendizado de máquina: uma abordagem estatística**. 1. ed. São Carlos, Brasil: Rafael Izbicki, 2020. v. 1. -272 p. ISBN 978-65-00-02410-4.

JÄHNE, B. **Jähne2005.** 6. ed. Heidelberg, Germany: Springer, 2005. 654 p. ISBN 3-540-24035-7.

JAIN, A.; GUPTA, R. Gaussian filter threshold modulation for filtering flat and texture area of an image. In: **Conference Proceeding - 2015 International Conference on Advances in Computer Engineering and Applications, ICACEA 2015**. Ghaziabad, India: IEEE, 2015. p. 760–763. ISBN 9781467369114.

JOSHI, A. V. **Machine Learning and Artificial Intelligence**. 1. ed. Redmond, USA: Springer International Publishing, 2020. 262 p. ISBN 978-3-030-26621-9.

JUDGE, A. **Mastering Aperture, Shutter Speed, ISO and Exposure**. 1. ed. Sedona, Arizona, USA: CCreateSpace Independent Publishing Platform, 2012. 102 p. ISBN 978-1482314458.

KANAN, C.; COTTRELL, G. W. Color-to-grayscale: Does the method matter in image recognition? **PLoS ONE**, La Jolla, California, United States of America, v. 7, 1 2012. ISSN 19326203.

KIM, H. *et al.* Concrete crack identification using a UAV incorporating hybrid image processing. **Sensors**, Ulsan, Korea;, p. 1–14, 2017. ISSN 14248220.

KIM, I. H. *et al.* Application of crack identification techniques for an aging concrete bridge inspection using an unmanned aerial vehicle. **Sensors**, Daejeon, Korea, v. 18, p. 1–14, 2018. ISSN 14248220.

KUMAR, A.; SINGH, S. Comparative analysis of gaussian filter, median filter and denoise autoenocoder. In: **2020 7th International Conference on Computing for Sustainable Global Development**. Delhi, India: IEEE, 2020. v. 6, p. 45–51. DOI: 10.23919/INDIACom49435.2020.9083712.

LECUN, Y. *et al.* Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, v. 86, n. 11, p. 2278–2323, 1998. ISSN 00189219.

LEE, J. H. *et al.* A new image-quality evaluating and enhancing methodology for bridge inspection using an unmanned aerial vehicle. **Smart Structures and Systems**, Daejeon, South Korea, v. 27, n. 2, p. 209–226, 2021. ISSN 17381991.

LEI, B. *et al.* New crack detection method for bridge inspection using uav incorporating image processing. **Journal of Aerospace Engineering**, American Society of Civil Engineers (ASCE), Hubei, China, v. 31, 9 2018. ISSN 0893-1321.

LETA, F. R. **Visual Computing- Scientific Visualization and Imaging Systems**. 4. ed. Rio de Janeiro, Brasil: Springer, 2014. 256 p. ISSN 2190-5924. ISBN 978-3-642-55131-4.

LI, C. *et al.* Tunnel crack detection using coarse-to-fine region localization and edge detection. **Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery**, Wiley-Blackwell, v. 9, 9 2019. ISSN 19424795.

LILLESAND, T. M.; KIEFER, R. W.; CHIPMAN, J. W. **Remote Sensing and Image Interpretation**. 7. ed. Madison, USA: Wiley, 2015. 770 p. ISBN 9781118343289.

LIU, J. G.; MASON, P. J. **Image processing and GIS for remote Sensing**. 2. ed. London, UK: Wiley, 2019. 472 p. ISBN 9781118724200.

LUHMAN, T. *et al.* **Close Range Photogrammetry - Principles, techniques and applications**. 1. ed. Oldenburg, Germany: [s.n.], 2011. v. 30. 510 p. ISSN 14779730. ISBN 9781849950572.

MAJUMDER, A.; GOPI, M. **Introduction Visual Computing: Core Concepts in Computer Vision, Graphics, and Image Processing**. 1. ed. Irvine,USA: CRC Press, 2018. 393 p. ISBN 2013206534.

Malothu Nagu, N. V. S. Image De-Noising By Using Median Filter and Weiner Filter. **International Journal of Innovative Research in Computer and Communication Engineering**, Andhra Pradesh, India, v. 2, n. 9, p. 5641–5649, 2014.

MARSLAND, S. **Machine Learning - An Algorithmic Perspective**. 1. ed. Boca Raton , USA: CRC PRESS, 2009. v. 1. ISBN 13981420067187.

MATILDES, C. M. Concreto armado e suas patologias. **Revista Científica Semana Acadêmica**, Revista Científica Semana Acadêmica, São Paulo, Brasil, v. 10, p. 1–31, sep 2022. DOI: 10.35265/2236-6717-225-12187.

MAVROFORAKIS, M. E.; THEODORIDIS, S. A geometric approach to support vector machine (SVM) classification. **IEEE Transactions on Neural Networks**, v. 17, n. 3, p. 671–682, 2006. ISSN 10459227.

MOHAN, A.; POOBAL, S. Crack detection using image processing: A critical review and analysis. **Alexandria Engineering Journal**, Elsevier, Kottayam, Kerala, v. 57, p. 787–798, 6 2018. ISSN 1110-0168.

MUSA, A. B. Comparative study on classification performance between support vector machine and logistic regression. **International Journal of Machine Learning and Cybernetics**, Springer Verlag, Wad Madani Sudan, v. 4, p. 13–24, 2 2013. ISSN 1868808X.

NTALIANIS, K.; CROITORU, A. Applied physics, system science and computers. In: **Applied Physics, System Science and Computers - Proceedings of the 1st International Conference on Applied Physics**. Dubrovnik, Croatia: Springer, 2018. v. 428, p. 245–251. ISBN 978-3-319-53933-1. DOI: 10.1007/978-3-319-53934-8.

NWANGANGA, F.; CHAPPLE, M. **Practical Machine Learning in R**. 1. ed. Notre Dame, USA: Wiley, 2020. v. 1. 464 p. ISBN 978-1-5231-3319-2.

NWANKPA, C. E. *et al.* Activation functions: Comparison of trends in practice and research for deep learning. **arXiv preprint arXiv:1811.03378**, Glasgow, UK, p. 1–20, 2018. DOI: https://doi.org/10.48550/arXiv.1811.03378.

OBI, J. C. A comparative study of several classification metrics and their performances on data. **World Journal of Advanced Engineering Technology and Sciences**, GSC Online Press, v. 8, p. 308–314, 2 2023. DOI: 10.30574/wjaets.2023.8.1.0054.

OSKOEI, M. A.; HU, H. A Survey on Edge Detection Methods. **Evaluation**, Colchester, United Kingdom, n. February, p. 1–36, 2010. ISSN 1744-8050.

ÖZGENEL, Ç. F. **Concrete Crack Images for Classification**. *s.l.*: Mendeley Data, 2018. DOI: 10.17632/5y9wdsg2zt.1.

PEDRINI, H.; SCHWARTZ, W. R. **Análise de imagens digitais : princípios, algoritmos e aplicações**. 1ª. ed. São Paulo, Brasil: Cengage Learning, 2007. 529 p. ISBN 8522105952.

PLAUE, M. **Data Science: An Introduction to Statistics and Machine Learning**. Berlin, Germany: Springer Berlin Heidelberg, 2023. v. 1. –372 p. DOI: 10.1007/978-3-662-67882-4. ISBN 978-3-662-67882-4.

PRATT, W. K. **Digital Image Processing**. 4. ed. Los Altos, USA: Wiley, 2007. 807 p. ISBN 9780471767770.

REAGAN, D.; SABATO, A.; NIEZRECKI, C. Feasibility of using digital image correlation for unmanned aerial vehicle structural health monitoring of bridges. **Structural Health Monitoring**, SAGE Publications Ltd, Lowell, MA, USA, v. 17, p. 1056–1072, 9 2018. ISSN 17413168. DOI: 10.1177/1475921717735326.

ROSA, G. H. de; PAPA, J. P.; YANG, X. S. Handling dropout probability estimation in convolution neural networks using meta-heuristics. **Soft Computing**, Springer Berlin Heidelberg, v. 22, n. 18, p. 6147–6156, 2018. ISSN 14337479. DOI: 10.1007/s00500-017-2678-4.

RUIZ, R. D. B. *et al.* Processamento digital de imagens para detecção automática de fissuras em revestimentos cerâmicos de edifícios. **Ambiente Construído**, Recife, Brasil, v. 21, p. 139–147, 2021. ISSN 1415-8876. DOI: 10.1590/s1678-86212021000100498.

SABERIOON, M. *et al.* Comparative performance analysis of support vector machine, random forest, logistic regression and k-nearest neighbours in rainbow trout (oncorhynchus mykiss) classification using image-based features. **Sensors**, MDPI AG, Nové Hrady, Czech Republic, v. 18, 4 2018. ISSN 14248220. DOI: 10.3390/s18041027.

SAMAD, A. M. *et al.* The potential of unmanned aerial vehicle (uav) for civilian and mapping application. In: **Proceedings - 2013 IEEE 3rd International Conference on System Engineering and Technology, ICSET 2013**. Shah Alam, Malaysia: IEEE, 2013. p. 313–318. ISBN 9781479910304. DOI: 10.1109/ICSEngT.2013.6650191.

SANTOS, B. S. D. *et al.* Fotografias aéreas para detecção de trincas em estruturas de concreto usando redes neurais convolucionais. In: **VIII Encontro Regional de Matemática Aplicada e Computacional da Regional 9**. Ilha Solteira, SP, Brasil: [s.n.], 2024. Apresentação na modalidade pôster. Disponível em: https://www.even3.com.br/viiiermacr9/.Acessoem:18set.2024.

SANTOS, B. S. dos *et al.* On the use of different image filters for damage detection in concrete-like structures using aerial. In: **27th ABCM Internacional Congress of Mechanical Engineering**. Florianopolis, Brazil: ABCM, 2023. p. 10. DOI: 10.26678/ABCM.COBEM2023.COB2023-1825.

SANTOS, W. F. M. D.; ARAúJO, F. M. U. D. Detecção e diagnóstico de falhas em rolamentos, sob diferentes cargas e velocidades, utilizando redes neurais convolucionais. **Brazilian Journal of Development**, South Florida Publishing LLC, Lagoa Nova, Natal, v. 9, p. 24682–24700, 8 2023. DOI: 10.34117/bjdv9n8-098.

SARAVANAN, C. Color image to grayscale image conversion. In: **2010 2nd International Conference on Computer Engineering and Applications, ICCEA 2010**. Bali, Indonesia: IEEE, 2010. v. 2, p. 196–199. ISBN 9780769539829. DOI: 10.1109/ICCEA.2010.192.

SCHEIN, A. I.; UNGAR, L. H. Active learning for logistic regression: An evaluation. **Machine Learning**, Philadelphia, USA, v. 68, p. 235–265, 10 2007. ISSN 08856125. DOI: 10.1007/s10994-007-5019-5.

SEDDIK, H.; TEBBINI, S.; BRAIEK, E. B. **Smart Real Time Adaptive Gaussian Filter Supervised Neural Network for Efficient Gray Scale and RGB Image De-noising**. Tunis, Tunisia: Taylor and Francis, 2014. 347–364 p. DOI: 10.1080/10798587.2014.888242.

SEKEHRAVANI, E. A.; BABULAK, E.; MASOODI, M. Implementing canny edge detection algorithm for noisy image. **Bulletin of Electrical Engineering and Informatics**, Lynchburg, USA, v. 9, n. 4, p. 1404–1410, 2020. ISSN 23029285. DOI: 10.11591/eei.v9i4.1837.

SIMõES, L. F.; RODRIGUES, P. F.; PINHEIRO, D. C. C. Análise crítica dos métodos de inspeção da abnt nbr 9452:2019. In: **XII Congresso Brasileiro de Pontes e Estruturas**. Belo Horizonte, Minas Gerais, Brasil: Associação Brasileira de Pontes e Estruturas (ABPE), 2021. v. 9452. ISSN 0102-8878.

SLUDER, G.; WOLF, D. E. **Methods in Cell Biology - Digital Microscopy: A Second Edition of Video Microscopy**. 2. ed. Santa Barbara, California: Elsevier, 2003. 544 p. ISBN 0-12-649161-5.

SMITH, S. W. **Digital Signal Processing: A Practical Guide for Engineers and Scientists**. Burlington, UK: Newnes, 2003. v. 17. 650 p. ISBN 075067444X.

SOBEL, I. E. **Camera Models and Machine Perception**. -110 p. Tese (Doutorado) — Stanford University, Sacramento, USA, 1970. ISBN: 9798661468925.

SOKOLOVA, M.; LAPALME, G. A systematic analysis of performance measures for classification tasks. **Information Processing and Management**, Elsevier Ltd, Ontario, Ottawa, Canada, v. 45, n. 4, p. 427–437, 2009. ISSN 03064573. DOI: 10.1016/j.ipm.2009.03.002.

SOLOMON, C.; BRECKON, T. **Fundamentals of Digital Image Processing- A Practical Approach with Examples in Matlab**. 1. ed. Canterbury, Kent, UK: Wiley-Blackwell, 2011. 320 p. ISBN 978-0-470-84472-4.

SOUZA, V. C. M. D.; RIPPER, T. **Patologia, Recuperação e Reforço de Estruturas de Concreto**. 1. ed. São Paulo, Brasil: Pini, 2009. 259 p. ISBN 8572660968.

SRIVASTAVA, N. *et al.* Dropout: A simple way to prevent neural networks from overfitting. **Journal of Machine Learning Research**, Toronto, Ontario, Canada., v. 15, p. 1929–1958, 2014. ISSN 15337928.

SUNDARARAJAN, D. **Digital image processing: A signal processing and algorithmic approach**. 1. ed. Montreal, Canada: Springer, 2017. 468 p. ISBN 978-981-10-6113-4.

SZELISKI, R. **Computer Vision : Algorithms and Applications 2nd Edition**. Seattle, WA, USA: Springer, 2021. 1196 p. ISSN 1668-0941. ISBN 978-3-030-34371-2.

THABTAH, F.; ABDELHAMID, N.; PEEBLES, D. A machine learning autism classification based on logistic regression analysis. **Health Information Science and Systems**, Springer International Publishing, Auckland, New Zealand, v. 7, 12 2019. ISSN 20472501.

TUBA, E. *et al.* **Convolutional Neural Networks Hyperparameters Tuning**. 1. ed. Warsaw, Poland: Springer, 2021. v. 1. 65-84 p. ISBN 978-3-030-72711-6.

TřEBICKÝ, V. *et al.* Focal length affects depicted shape and perception of facial images. **PLoS ONE**, Klecany, Czech Republic, v. 11, p. 1–14, 2016. ISSN 19326203.

VALE, G. M. D.; POZ, A. P. D. O processo de detecção de bordas de canny: Fundamentos, algoritmos e avaliação experimental. **I Simpósio Brasileiro de Geomática**, Curitiba, Brasil, p. 292–303, 7 2002. ISSN 1982-2170.

VECELLIO, M. E.; BRYANT, E. P. **Pocket Guide to Crime Scene Photography**. 1. ed. Boca Raton, USA: CRC Press, 2018. 202 p. ISBN 9788527729833.

WANG, M. *et al.* A new image denoising method based on gaussian filter. In: **Proceedings - 2014 International Conference on Information Science, Electronics and Electrical Engineering**. Sapporo, Japan: IEEE, 2014. v. 1, p. 163–167. ISBN 9781479931965. DOI: 10.1109/InfoSEEE.2014.6948089.

WAOO, A. A.; SONI, B. K. Performance analysis of sigmoid and relu activation functions in deep neural network. In: **Procedia Computer Science**. Pradesh, India: Springer, 2021. v. 171, p. 39–52. ISBN 9789811622472. ISSN 18770509.

WU, G. *et al.* Compared between support vector machine (svm) and deep belief network (dbn) for multi-classification of raman spectroscopy for cervical diseases. **Photodiagnosis and Photodynamic Therapy**, Elsevier B.V., Pune, India, v. 42, n. February, p. 103340, 2023. ISSN 18731597. DOI: 10.1016/j.pdpdt.2023.103340.

YOUN, J. **Mastering Digital Photography: Understanding the Art and Science of Aperture, Shutter, Exposure, Light, and Composition**. 2. ed. *[s.i.]*: CreateSpace Independent Publishing Platform, 2016. 207 p. ISBN 2013206534.

ZHANG, D. *et al.* inspection using unmanned aerial vehicles. **Engineering Structures**, Elsevier, Glasgow, UK, v. 209, n. November 2019, p. 109940, 2020. ISSN 18737323.

ZHANG, Y. The design of glass crack detection system based on image preprocessing technology. In: **2014 IEEE 7th Joint International Information Technology and Artificial Intelligence Conference, ITAIC 2014**. Chongqing, China: IEEE, 2014. p. 72. ISBN 9781479944200. DOI: 10.1109/ITAIC.2014.7065001.

ZOMCHAK, L.; STARCHEVSKA, I. Macroeconomic determinants of economic development and growth in ukraine: Logistic regression analysis. In: HU, Z.; WANG, Y.; HE, M. (Ed.). **Advances in Intelligent Systems, Computer Science and Digital Economics IV**. Barcelona, Spain: Springer, 2023. p. 358–368.