



UNIVERSIDADE FEDERAL DO MARANHÃO

Programa de Pós-Graduação em Ciência da Computação

João Pedro Gomes de Souza

***Detecção de Fissuras em Micro Tomografia de Concreto
Reforçado Por Fibras Utilizando Redes Neurais Profundas***

**São Luís
2025**



UNIVERSIDADE FEDERAL DO MARANHÃO
Programa de Pós-Graduação em Ciência da Computação

João Pedro Gomes de Souza

**Deteccção de Fissuras em Micro Tomografia de
Concreto Reforçado Por Fibras Utilizando Redes
Neurais Profundas**

São Luís - MA

2025

João Pedro Gomes de Souza

Detecção de Fissuras em Micro Tomografia de Concreto Reforçado Por Fibras Utilizando Redes Neurais Profundas

Dissertação apresentada como requisito parcial para obtenção do título de Mestre em Ciência da Computação, ao Programa de Pós-Graduação em Ciência da Computação, da Universidade Federal do Maranhão.

Programa de Pós-Graduação em Ciência da Computação

Universidade Federal do Maranhão

Orientador: Prof. Dr. Aristófanês Corrêa Silva

Coorientador: Prof. Dr. António Manuel Trigueiros Cunha

São Luís - MA

2025

Ficha gerada por meio do SIGAA/Biblioteca com dados fornecidos pelo(a) autor(a).
Diretoria Integrada de Bibliotecas/UFMA

Gomes de Souza, João Pedro.

Detecção de Fissuras em Micro Tomografia de Concreto
Reforçado Por Fibras Utilizando Redes Neurais Profundas /
João Pedro Gomes de Souza. - 2025.
57 f.

Coorientador(a) 1: António Manuel Trigueiros Cunha.

Orientador(a): Aristófanês Corrêa Silva.

Dissertação (Mestrado) - Programa de Pós-graduação em
Ciência da Computação/ccet, Universidade Federal do
Maranhão, São Luís, 2025.

1. Concreto. 2. Fissuras. 3. Micro-ct. 4.
Aprendizado Profundo. I. Corrêa Silva, Aristófanês. II.
Trigueiros Cunha, António Manuel. III. Título.

João Pedro Gomes de Souza

Detecção de Fissuras em Micro Tomografia de Concreto Reforçado Por Fibras Utilizando Redes Neurais Profundas

Dissertação apresentada como requisito parcial para obtenção do título de Mestre em Ciência da Computação, ao Programa de Pós-Graduação em Ciência da Computação, da Universidade Federal do Maranhão.

Trabalho aprovado. São Luís - MA, 22 de Setembro de 2025.

Prof. Dr. Aristófanês Corrêa Silva
Orientador
Universidade Federal do Maranhão

Prof. Dr. António Manuel Trigueiros Cunha
Coorientador
Universidade de Trás-os-Montes e Alto Douro

Prof. Dr. Anselmo Cardoso de Paiva
Examinador Interno
Universidade Federal do Maranhão

Prof. Dra. Aura Conci
Examinador Externo
Universidade Federal Fluminense

São Luís - MA
2025

À Maria Marly Moreira, minha bisavó. Espero ter te orgulhado.

Agradecimentos

Agradeço ao meu Senhor Jesus Cristo, por todas as bênçãos que tem depositado em minha vida e por toda força que me dá. Que toda a glória pertença a Ti.

Agradeço aos meus orientadores, Dr. Aristóфанes e Dr. António, grandes mestres, por toda sabedoria, ensinamentos e paciência. Muita paciência. Vocês enxergaram um potencial em mim que nem mesmo eu fui capaz de ver.

Agradeço a parceria do Dr. Marcello Congro e da Dra. Deane Roehl, do Instituto Tecgraf da Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), importantes colaboradores desta pesquisa, por terem fornecido a base de dados utilizada.

Agradeço à minha namorada, inesquecível Bianca, por todo o companheirismo, compreensão e carinho durante essa difícil trajetória. Você é basilar.

Agradeço à minha família, por me manter com o pé no chão e estar comigo nos momentos difíceis. Amo todos vocês.

Agradeço aos amigos Alan, Antônio e Maxwell, por dedicarem parte do seu tempo para me auxiliarem na execução de muitos experimentos necessários neste trabalho.

Agradeço a todos os meus amigos do Núcleo de Computação Aplicada, pela convivência e por terem me auxiliado mesmo que indiretamente, seja por pequenos conselhos, palavras de motivação ou pelos momentos de descontração. O apoio de todos vocês é ímpar.

Agradeço aos professores do Programa de Pós-Graduação em Ciência da Computação (PPGCC) pela capacitação na minha formação.

Agradeço à Fundação de Amparo à Pesquisa e ao Desenvolvimento Científico e Tecnológico do Maranhão (FAPEMA), à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), à Fundação para a Ciência e Tecnologia (FCT) e à Dra. Sandra Pereira Cunha (UTAD) pelo importante apoio financeiro a esta pesquisa.

A todos citados, meus mais profundos e sinceros agradecimentos. De verdade. Eu não teria conseguido sem vocês.

"Proceder pra vencer, pra crescer, prevalecer."

(Sabotage)

Resumo

O concreto reforçado por fibras é um material crucial para construção civil e monitorar sua condição é uma tarefa importante para sua preservação e prevenção de acidentes e perdas financeiras. Dentre os vários métodos não destrutivos de monitoramento da condição do concreto, destaca-se a análise por imagens de Microtomografia Computadorizada (Micro-CT), por se tratar de um método barato e livre de ruídos e de interferências externas. Contudo, a identificação manual de fissuras em imagens de Micro-CT é subjetiva e requer um esforço humano significativo. Algoritmos de aprendizado profundo vem sendo utilizados constantemente para a detecção automática de fissuras em concreto, porém apenas em regiões superficiais. Detectar fissuras em toda a estrutura do concreto reforçado por fibras, incluindo a parte interna, como em volumes de Micro-CT de alta resolução, permanece um desafio. Este estudo, portanto, propõe um método automático para detecção de fissuras especificamente para imagens de Micro-CT de concreto reforçado por fibra, integrando técnica de pré-processamento utilizando super-resolução, Detection Transformers (DETR) e pós-processamento baseadas em comitê para aprimorar a precisão e confiabilidade das detecções de fissuras. O modelo foi treinado e validado usando uma base de dados criada e publicada a partir de imagens Micro-CT de experimentos de tração fibra/matriz denominada FIRECON Dataset, composta de 4,064 imagens anotadas por especialistas. alcançando métricas de 86,098% de Interseção sobre União(IoU), 89,37% de Precisão, 83,26% de Recall, 84,99% de F1-Score e 44,69% de Precisão Média (AP). Este método portanto reduz significativamente o tempo de análise, esforço humano e melhora a consistência dos resultados em comparação com uma inspeção visual. O método proposto mostra o potencial do aprendizado profundo no auxílio do monitoramento da condição e compreensão dos mecanismos de fissuração do concreto reforçado por fibras e oferece uma ferramenta valiosa para o desenvolvimento de materiais de engenharia duráveis e de alto desempenho.

Palavras-chave: Concreto, Fissuras, Micro-CT, Aprendizado Profundo.

Abstract

Fiber-reinforced concrete is a crucial material for civil construction, and monitoring its health is an important task for preserving structures and preventing accidents and financial losses. Among the various non-destructive concrete health monitoring methods, Computed Microtomography (Micro-CT) image analysis stands out as an inexpensive method free from noise and external interference. However, manual inspection of cracks in Micro-CT images is subjective and requires significant human effort. Deep learning algorithms have been used extensively for concrete cracks automatic detection, but only in surface regions. Detecting cracks all fiber-reinforced concrete structure, including inside in high-resolution Micro-CT volumes, remains a challenge. Therefore, this work proposes a framework for automatic crack detection explicitly for fiber-reinforced concrete Micro-CT images, combining super-resolution-based preprocessing, Detection Transformers (DETR), and committee-based post-processing to improve crack detection reliability and reduce false positives. The model was trained on a new publicly available dataset, the FIRECON dataset, which consists of 4,064 images annotated by an expert, achieving metrics of 86.098% Intersection over Union (IoU), 89.37% Precision, 83.26% Recall, 84.99% F1-Score, and 44.69% Average Precision (AP). This framework, therefore, significantly reduces analysis, human effort, and improves consistency compared to manual methods used in previous studies. The results demonstrate the potential of deep learning to aid image analysis in concrete health assessments, understanding of the cracking mechanisms of fiber-reinforced concrete, and provides a valuable tool for the development of durable, high-performance engineering materials.

Keywords: Concrete, Cracks, microCT, Deep Learning.

Lista de ilustrações

Figura 1 – Exemplos de concreto reforçado por fibras.	18
Figura 2 – Alguns tipos de fissuras de concreto em macroescala: (a) química, (b) térmica, (c) por retração e (d) estrutural.	20
Figura 3 – (a) Microtomógrafo Zeiss-Xradia Versa 510 e (b) digitalização 3D de uma peça de concreto reforçado por fibras.	23
Figura 4 – Arquitetura da EfficientNetB0.	26
Figura 5 – Arquitetura da DETR.	28
Figura 6 – Arquitetura da Swim2SR.	29
Figura 7 – Etapas do método proposto.	36
Figura 8 – (a) Fatia de um volume de Micro-CT de concreto reforçado por fibras e (b) a visualização aproximada da região destacada em amarelo, mostrando os principais constituintes do concreto.	38
Figura 9 – Imagem do concreto com fibra gerado a partir de Micro-CT, antes e depois da aplicação do CLAHE.	39
Figura 10 – Fluxo da etapa de classificação de matriz cimentícia.	40
Figura 11 – Fluxo da etapa de detecção de fissuras.	41
Figura 12 – Resultado da classificação de uma fatia da Matriz B com apenas uma janela classificada como Matriz A, quantidade insuficiente para classificá-la erroneamente.	44
Figura 13 – Comparação entre a marcação do especialista(em verde) e a predição do modelo (em vermelho) para uma imagem do tipo Matriz A.	46
Figura 14 – Comparação entre a marcação do especialista(em verde) e a predição do modelo (em vermelho) para uma imagem do tipo Matriz B.	47
Figura 15 – Comparação entre a marcação do especialista (em verde) e a predição final do método após a redução de falsos positivos (em vermelho).	49

Lista de tabelas

Tabela 1 – Parâmetros de aquisição para a digitalização em Micro-CT (CONGRO et al., 2025a).	37
Tabela 2 – Desempenho	44
Tabela 3 – Matriz de Confusão da classificação binária de tipos de matriz	44
Tabela 4 – Performance of EfficientNetB0, de acordo com o tipo de imagem de entrada.	45
Tabela 5 – Métricas de detecção de fissuras para o tipo de matriz A	45
Tabela 6 – Métricas de detecção de fissuras para o tipo de matriz B	45
Tabela 7 – Comparação das métricas obtidas entre as predições de resoluções diferentes e a predição final utilizando o pós-processamento baseado em comitê, para Matriz do tipo A.	47
Tabela 8 – Comparação das métricas obtidas entre as predições de resoluções diferentes e a predição final utilizando o pós-processamento baseado em comitê, para Matriz do tipo B	48
Tabela 9 – Comparação dos modelos especializados para cada tipo de matriz com um modelo treinado com ambos os tipos.	49
Tabela 10 – Comparação dos métodos disponíveis na literatura com o método proposto, baseado nas métricas fornecidas pelos autores.	50
Tabela 11 – Produção científica diretamente ligada ao tema de pesquisa	52
Tabela 12 – Produção científica em demais áreas de pesquisa	52

Lista de abreviaturas e siglas

Micro-CT	<i>Microtomografia Computadorizada</i>
ViT	<i>Vision Transformers</i>
DETR	<i>Detection Transformer</i>
CNN	<i>Convolutional Neural Network</i>
DCNN	<i>Dilated Convolutional Neural Network</i>
IoU	<i>Intersection over Union</i>
MPa	<i>Milipascal</i>
kN	<i>quiloneutron</i>
AP	<i>Average Precision</i>
MBConv	<i>Mobile Inverted Bottleneck Convolution</i>
FFN	<i>Feed Forward Network</i>
ReLU	<i>Rectified Linear Unit</i>
CGAN	<i>Conditional Generative Adversarial Network</i>

Sumário

1	INTRODUÇÃO	14
1.1	Objetivos	16
1.1.1	Objetivos Específicos	16
1.1.2	Contribuições	16
1.2	Organização do Trabalho	17
2	FUNDAMENTAÇÃO TEÓRICA	18
2.1	O concreto reforçado por fibras	18
2.2	Fissuras e a deterioração do concreto	19
2.2.1	Métodos destrutivos	21
2.2.2	Métodos não destrutivos	21
2.3	Microtomografia Computadorizada (Micro-CT)	23
2.4	<i>Aprendizado Profundo</i>	24
2.4.1	EfficientNet	25
2.4.2	DETR	26
2.4.3	Swin2SR	29
2.5	Métricas de Avaliação	30
2.5.1	Acurácia	30
2.5.2	Precisão	30
2.5.3	<i>Recall</i>	31
2.5.4	<i>F1-Score</i>	31
2.5.5	<i>Intersection over Union</i>	31
2.5.6	<i>Average Precision</i>	31
2.6	Considerações Finais	31
3	TRABALHOS RELACIONADOS	33
3.1	Deteção utilizando deep learning em macroescala	33
3.2	Deteção de Fissuras em Micro-CT	34
4	MÉTODO PROPOSTO	36
4.1	Aquisição de Imagens	36
4.2	Marcação e Publicação da Base de Dados	38
4.3	Pré-Processamento	39
4.4	Classificação de Tipo da Matriz Cimentícia	40
4.5	Deteção de Fissuras	41
4.6	Pós-processamento	42

5	EXPERIMENTOS E RESULTADOS	43
5.1	Classificação do tipo de matriz	43
5.2	Detecção de Fissuras	45
6	CONCLUSÕES E TRABALHOS FUTUROS	51
6.1	Produções Científicas	52
	REFERÊNCIAS	53

1 Introdução

O concreto é um dos materiais mais utilizados na construção civil, desempenhando um papel fundamental no desenvolvimento urbano e industrial. Sua ampla utilização está relacionada à sua disponibilidade em larga escala, custo relativamente baixo, durabilidade e capacidade de ser moldado em diferentes formatos. Por conta disso, é a principal escolha em projetos que vão de pequenas edificações até grandes obras de infraestrutura, como pontes, túneis, barragens e viadutos (COUTO et al., 2013; MEHTA; MONTEIRO, 2006; NEVILLE et al., 1995).

Nesse sentido, o concreto reforçado por fibras trata-se de um compósito no qual fibras de diferentes naturezas são incorporadas à matriz de cimento. Essas fibras atuam como agregados descontínuos que têm a função de absorver energia e controlar a propagação de fissuras, promovendo maior resistência à tração e ao impacto. Além do tipo de fibra, seu formato também exerce papel determinante no comportamento mecânico do material, podendo ser retilíneas, em forma de gancho, frisadas, dentre outros formatos. Por essas razões, o concreto reforçado por fibras tem sido cada vez mais empregado em aplicações que demandam maior desempenho estrutural (LEITE; CASTRO, 2020; ABDUL-RAZZAK; Mohammed Ali, 2011; LEE; WON, 2014; CONGRO et al., 2025a).

Embora o concreto seja considerado um material robusto e durável, sua degradação ao longo do tempo é um inevitável. Fatores ambientais, sobrecargas não previstas em projeto, falhas de execução ou manutenção inadequada podem acelerar processos de deterioração. Entre os diversos mecanismos de degradação, a fissuração ocupa papel central, pois pode comprometer o desempenho mecânico e durabilidade da estrutura. As fissuras facilitam a penetração de agentes agressivos, como água, causando a redução de resistência e eventual colapso estrutural. Além das consequências técnicas, falhas em estruturas de concreto podem resultar em acidentes de grandes proporções, envolvendo perdas humanas, prejuízos financeiros e impactos sociais significativos (SOUZA; CONCEIÇÃO; MOREIRA, 2024). Nesse contexto, a monitoração da condição do concreto, com ênfase na detecção precoce de fissuras, é essencial para garantir a segurança estrutural e prolongar a vida útil das construções.

Diversos métodos têm sido empregados para detectar fissuras. Tradicionalmente, essas avaliações podem ser feitas por meio de métodos destrutivos e não destrutivos. Enquanto os primeiros envolvem a retirada de amostras ou a aplicação de esforços que danificam o corpo de prova, os métodos não destrutivos oferecem vantagens ao permitir o monitoramento sem comprometer a integridade da estrutura (NAJM; NANAYAKKARA; SABRI, 2022; AHMED; SHUKRI, 2017). Dentre as técnicas não destrutivas baseadas

em imagem, destaca-se a Microtomografia Computadorizada (Micro-CT), que tem se mostrado uma ferramenta poderosa para a análise microestrutural de materiais cimentícios. A Micro-CT permite a obtenção de imagens em microescala e possibilita não apenas a avaliação superficial, mas também a visualização interna de amostras de concreto. Essa característica confere um diferencial à técnica em relação aos demais métodos, uma vez que torna possível identificar e quantificar fissuras, porosidade, distribuição de fibras e agregados com elevado nível de detalhe ([MACHADO et al., 2015](#)).

No entanto, a análise manual de imagens de Micro-CT apresenta desafios para os avaliadores humanos. As fissuras são extremamente pequenas e difíceis de identificar visualmente, mesmo quando as imagens são ampliadas, tornando a detecção manual um processo complexo e propenso a erros. Além disso, os microtomógrafos geram uma quantidade massiva de imagens para cada amostra analisada. A reconstrução de um único volume 3D requer o processamento de centenas ou até milhares de imagens individuais. Como consequência, a avaliação manual de uma parcela significativa dessas imagens torna-se uma tarefa extremamente trabalhosa. Em casos de volumes muito grandes, essa abordagem pode até mesmo se tornar inviável na prática ([UMBACH; MIDDENDORF, 2024](#); [CONGRO, 2024](#)).

Diante dessas limitações, é pertinente considerar o uso de técnicas automáticas para auxiliar no processamento e interpretação de imagens de Micro-CT. Nesse cenário, os modelos de Aprendizado Profundo emergem como uma alternativa promissora. O Aprendizado Profundo tem se consolidado nos últimos anos como uma ferramenta capaz de realizar tarefas complexas de reconhecimento de padrões e classificação em grandes conjuntos de dados. Em aplicações relacionadas ao concreto, já existem estudos que utilizam redes neurais profundas para a detecção de fissuras a partir de imagens convencionais de superfície, apresentando resultados expressivos em termos de precisão e confiabilidade ([GOLDING et al., 2022](#)). Estes modelos auxiliam especialistas na detecção precoce de fissuras e, portanto, na monitoração da condição do concreto. Contudo, estes estudos utilizam imagens adquiridas a partir de câmeras comuns, mostrando apenas a superfície do concreto.

Aplicar técnicas de Aprendizado Profundo à análise de imagens de Micro-CT em concreto reforçado por fibras representa, portanto, uma oportunidade relevante para avançar na detecção de fissuras. Além de reduzir a carga de trabalho dos especialistas, a automação pode aumentar a consistência das detecções. Considerando a importância do concreto reforçado por fibras e a necessidade de monitorar e assegurar sua durabilidade e desempenho, a integração de métodos avançados de processamento de imagens e Aprendizado Profundo pode contribuir significativamente para o desenvolvimento de novas metodologias de avaliação não destrutiva no campo da engenharia civil e de materiais.

1.1 Objetivos

O objetivo geral deste trabalho consiste em desenvolver e implementar um método automatizado baseado em redes neurais profundas para a detecção de microfissuras em imagens de Micro-CT de concreto reforçado por fibras, visando aprimorar a análise de deterioração deste material.

1.1.1 Objetivos Específicos

Para alcançar o objetivo principal, este trabalho investigou os seguintes objetivos específicos

- Realizar uma revisão bibliográfica sobre técnicas de Aprendizado Profundo aplicadas à detecção de fissuras em materiais cimentícios e análise de imagens de Micro-CT;
- Construir e preparar uma base de dados de imagens de Micro-CT de concreto reforçado por fibras, incluindo a anotação manual das microfissuras para treinamento e validação dos modelos;
- Aplicar e avaliar arquiteturas de Aprendizado Profundo adequadas para a detecção de microfissuras em imagens de Micro-CT;
- Treinar e otimizar os modelos de Aprendizado Profundo utilizando técnicas de aumento de dados e ajuste de hiperparâmetros para melhorar a precisão da detecção;
- Avaliar o desempenho dos modelos através de métricas apropriadas como precisão, recall, F1-Score e IoU (Intersection over Union);
- Comparar os resultados obtidos com outras abordagens existentes na literatura;
- Validar a aplicabilidade do método proposto em diferentes matrizes de concreto e tipos de fibras.

1.1.2 Contribuições

Destacam-se como principais contribuições:

- Desenvolvimento de um método automático especializado para detecção de microfissuras em imagens de Micro-CT de concreto reforçado por fibras, utilizando arquiteturas de Aprendizado Profundo otimizadas para a aplicação específica;
- Aplicação inovadora da rede Swin2SR para geração de dados com super resolução em imagens de Micro-CT, aprimorando a qualidade das imagens de entrada e melhorando o desempenho de redes neurais de detecção de microfissuras;

- Proposição de uma técnica de pós-processamento baseada em comitê para redução significativa de falsos positivos na detecção de fissuras, aumentando a confiabilidade e precisão do sistema;
- Criação e publicação de uma base de dados anotada de imagens de Micro-CT de concreto reforçado por fibras com microfissuras marcadas por um especialista, contribuindo para futuras pesquisas na área;
- Avanço no estado da arte da aplicação de inteligência artificial na análise não destrutiva de materiais de construção civil.

1.2 Organização do Trabalho

Este trabalho está estruturado da seguinte forma:

- O Capítulo 2 descreve trabalhos relacionados ao tema de detecção de fissuras em concreto, aborda técnicas de monitoramento da condição do concreto tradicional em comparação com as imagens de Micro-CT e técnicas baseadas em Aprendizado Profundo, elucida o atual estado da arte e identifica lacunas presentes na literatura que este trabalho visa preencher.
- O Capítulo 3 trata da fundamentação teórica das técnicas utilizadas. São abordados conceitos sobre as redes utilizadas, técnicas de pré e pós processamento, conceitos de super-resolução, além do problema da detecção de objetos.
- O Capítulo 4 apresenta as etapas que compõem a metodologia proposta para este trabalho.
- O Capítulo 5 trata sobre os resultados obtidos e discussões em relação aos experimentos realizados.
- O Capítulo 6 apresenta as considerações finais sobre os resultados e trabalhos futuros.

2 Fundamentação Teórica

Neste capítulo apresenta-se a base teórica que fundamenta o desenvolvimento deste trabalho. Inicialmente, são expostas informações essenciais para a compreensão do concreto reforçado por fibras e dos mecanismos relacionados à sua deterioração. Na sequência, discutem-se as teorias que embasam os métodos destrutivos e não destrutivos aplicados ao monitoramento da condição do concreto e à detecção de fissuras. Posteriormente, introduzem-se informações sobre Micro-CT. Por fim, descrevem-se os fundamentos do Aprendizado Profundo, com ênfase nas arquiteturas de redes neurais empregadas no método proposto.

2.1 O concreto reforçado por fibras

O concreto é um material composto principalmente de cimento água, areia e de agregados diversos como britas. Essa mistura tem o objetivo de produzir um compósito durável e volumoso (MEHTA; MONTEIRO, 2006; NEVILLE et al., 1995).

Existem diversos tipos de concreto, dentre eles o concreto reforçado por fibras (Figura 1). Nesse tipo de concreto, as fibras atuam como agregado e podem ser feitas de materiais variados, como vidro, ferro ou fibras naturais. O formato das fibras é variável, podendo ser reta, em forma de gancho, frisada, dentre outras formas (LABIB, 2018).

Figura 1 – Exemplos de concreto reforçado por fibras.



Fonte: Daily Civil. Disponível em: <<https://dailycivil.com/fibre-reinforced-concrete/>>. Acesso em: 18 de agosto de 2025.

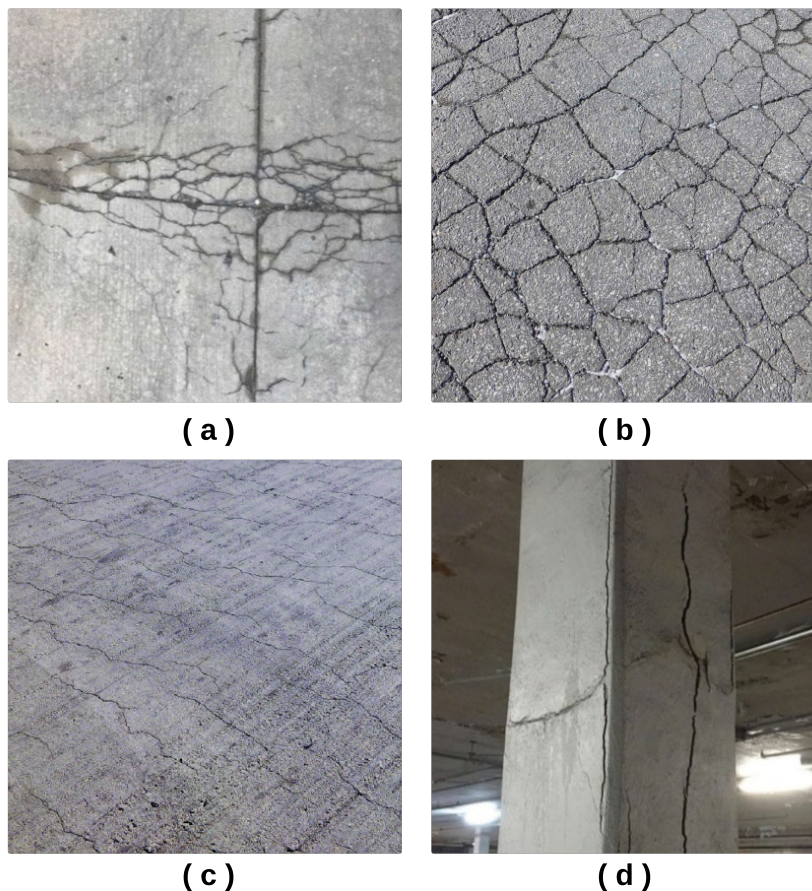
A distribuição e a quantidade de fibras dentro da matriz cimentícia é um fator variável na confecção do concreto. Cada variação proporciona distintas propriedades mecânicas ao concreto, influenciando tanto o mecanismo de transferência de tensão entre a fibra e a matriz cimentícia quanto à resistência do concreto à fissuração causada pela interação com agentes externos e a estrutura (NAAMAN; NAJM, 1991; ABDALLAH; FAN, 2017; MAČIULAITIS; VAIČIENE; ŽURAUSKIENE, 2009).

O concreto naturalmente se deteriora quando é sujeito a forças de tração. A combinação de fibra e cimento visa produzir um concreto que seja mais resistente à fissuração. O limite entre a fibra e matriz cimentícia define a capacidade do material de suportar cargas após o concreto fissurar, fazendo com que o concreto suporte uma carga maior de tração, prolongando sua durabilidade (LEITE; CASTRO, 2020; ABDUL-RAZZAK; Mohammed Ali, 2011; LEE; WON, 2014; CONGRO et al., 2025b).

2.2 Fissuras e a deterioração do concreto

As fissuras são descontinuidades presentes na matriz do concreto. Nem todas as fissuras indicam problemas estruturais (Figura 2), contudo um dos indicadores de deterioração de uma estrutura de concreto é a presença de fissuras. A deterioração do concreto reforçado por fibras ocorre quando as forças de tração superam sua resistência, causando fissuras. A detecção de fissuras em construções mostra sinais de desgaste, identificando o avanço da deterioração da estrutura e possibilitando o planejamento de reparos de forma eficaz. Dessa forma, é possível garantir a segurança e evitar acidentes que possam causar danos financeiros, físicos e até mesmo perdas de vida (LEONHARDT, 1988).

Figura 2 – Alguns tipos de fissuras de concreto em macroescala: (a) química, (b) térmica, (c) por retração e (d) estrutural.



Fonte: adaptado de Euclid Chemical. Disponível em: <<https://www.euclidchemical.com/company/blog/archive/concrete-cracks/>>. Acesso em: 18 de agosto de 2025.

O desgaste do concreto pode estar associado a fatores ambientais, climáticos ou mecânicos. Nesses contextos, o monitoramento torna-se essencial para avaliar a condição estrutural e garantir sua durabilidade. Entre as técnicas utilizadas, a detecção de fissuras desempenha papel central, pois possibilita não apenas acompanhar a evolução da deterioração, mas também analisar como diferentes composições de concreto reforçado por fibras respondem às necessidades mecânicas de uma estrutura, auxiliando no entendimento de seus mecanismos de resistência e controle da propagação de fissuras (CONGRO et al., 2025b; HASSANI; MOUSAVI; GANDOMI, 2021).

Os métodos utilizados para o monitoramento da condição do concreto podem ser classificados em destrutivos e não destrutivos, apresentando diferenças significativas quanto à forma de análise e às implicações estruturais.

2.2.1 Métodos destrutivos

Métodos destrutivos são conduzidos por meio da aplicação de uma força de tração sobre uma amostra de concreto, que geralmente é retirada de uma parte da estrutura. A amostra é então submetida a um ensaio de tração a fim de forçar a amostra à sua fissuração ou destruição. As fissuras então são observadas na própria amostra, inferindo assim a força aplicada que a amostra é capaz de suportar até a sua fissuração e, posteriormente, seu colapso. Métodos destrutivos não são adequados para realizar testes diretamente na estrutura, por deformarem a amostra, causando danos significativos à estrutura e consequentemente reduzindo sua durabilidade e resistência, ou até mesmo destruindo a própria, causando danos irreversíveis. Portanto, é necessária uma amostra à parte para realização do mesmo ([NAJM](#); [NANAYAKKARA](#); [SABRI](#), 2022).

2.2.2 Métodos não destrutivos

Métodos não destrutivos, em contraste aos anteriores, são uma maneira adequada para avaliar a deterioração de uma estrutura sem causar danos. São métodos ideais para detecção de fissuras já presentes no concreto, identificar sua resistência e julgar sua condição atual. Possibilitam a obtenção de informações mais abrangentes a respeito das características internas, propriedades em escala microscópica e indícios precoces de deterioração em estruturas. Dessa forma, destacam-se como alternativas mais eficazes para o acompanhamento do desempenho estrutural, uma vez que fornecem dados detalhados sem causar danos à estrutura analisada ([AHMED](#); [SHUKRI](#), 2017).

Diversas técnicas não destrutivas têm sido amplamente empregadas para avaliar a integridade do concreto e detectar fissuras sem causar danos à estrutura. Entre essas técnicas, destacam-se a seguir.

A emissão acústica ([DAS](#); [LEUNG](#), 2022) é amplamente empregada na verificação da integridade estrutural do concreto. Seu princípio baseia-se na detecção de ondas elásticas transitórias geradas pela liberação súbita de energia interna do material, como ocorre durante a propagação de fissuras. Por meio de sensores piezoelétricos acoplados à superfície do concreto, é possível registrar sinais acústicos que permitem identificar a ocorrência, a localização e até mesmo a intensidade de processos de deterioração em tempo real, sem a necessidade de danificar a estrutura.

A espectrometria por impedância elétrica ([DAS](#); [LEUNG](#), 2020), por sua vez, tem seu funcionamento baseado na aplicação de um sinal elétrico alternado de baixa intensidade e na medição da resposta do material em diferentes frequências, permitindo obter informações relacionadas à resistividade elétrica, à condutividade iônica e capacitância do concreto. A presença de fissuras e porosidade afeta significativamente os caminhos de condução elétrica, o que se reflete na variação dos parâmetros de impedância, possibilitando

a identificação de danos internos na estrutura do concreto.

As técnicas baseadas em vibração (IMAI et al., 2010) constituem métodos cujo princípio fundamental reside no fato de que a presença de descontinuidades altera as propriedades dinâmicas do material, como frequência natural, modos de vibração e amortecimento. Por meio da aplicação de excitações externas e do registro das respostas vibracionais com acelerômetros ou sensores ópticos, é possível identificar mudanças significativas no comportamento estrutural. Fissuras e danos internos tendem a reduzir a rigidez e, conseqüentemente, deslocar as frequências naturais do sistema, permitindo a detecção de fissuração na estrutura.

Os pulsos ultrassônicos guiados (ZHAO et al., 2019) baseiam-se na propagação de ondas elásticas ao longo da estrutura. Os pulsos guiados permitem que as ondas se propaguem por longas distâncias, contornando geometrias complexas e cobrindo áreas grandes com número pequeno de sensores. A presença de fissuras, vazios ou regiões degradadas altera a propagação das ondas, seja pela dispersão, reflexão, atenuação ou mudança na velocidade, o que possibilita a identificação e localização de danos estruturais.

As técnicas baseadas em sensores de fibra ótica baseiam-se em uma rede de sensores de fibra ótica que ficam inseridos dentro da matriz do concreto. Quando uma força de tração forte o suficiente para fissurar o concreto ocorre e causa a ruptura de um dos sensores, ocorre a interrupção do sinal de luz do sensor. Essa interrupção sinaliza uma região com fissura. Com uma quantidade suficiente de sensores é possível monitorar toda uma estrutura em busca de regiões danificadas (SCHÖTTL et al., 2020). Este método, entretanto, não é capaz de detectar regiões danificadas antes da instalação dos sensores. Outro ponto é a necessidade dos sensores dentro da matriz do concreto, o que altera as propriedades de resistência à tração da estrutura.

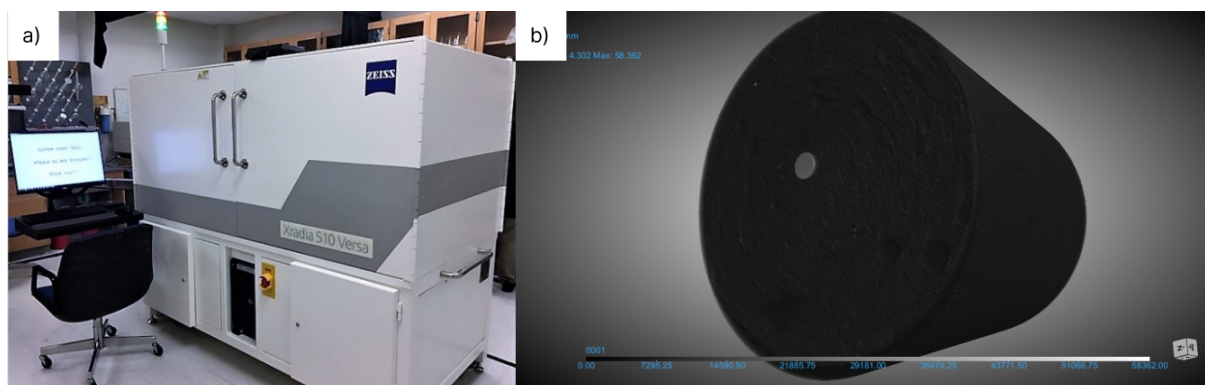
Embora essas técnicas sejam eficazes, elas apresentam algumas limitações importantes quando aplicadas ao concreto reforçado por fibras. Do ponto de vista econômico e operacional, observa-se o alto custo dos equipamentos necessários para implementação e a necessidade de instalar uma rede extensa de sensores por toda a estrutura. Quanto às limitações técnicas, destaca-se a interferência de ruídos externos que podem comprometer a precisão dos resultados, a dispersão irregular das ondas dentro da estrutura quando as fibras não estão distribuídas uniformemente, e a dificuldade de interpretação dos dados, podendo haver confusão entre as fissuras reais e outros elementos do concreto, como vazios e agregados. Essas limitações podem dificultar ou até mesmo inviabilizar o uso dessas técnicas para a detecção eficaz de problemas em estruturas de concreto reforçado por fibras.

2.3 Microtomografia Computadorizada (Micro-CT)

As tomografias computadorizadas destacam-se dos demais métodos de aquisição de imagens combinando informações de uma série de imagens de absorção de raios X 2D registradas à medida que a fonte de raios X e o detector são girados em torno do objeto. Usando princípios matemáticos de tomografia, essa série de imagens é reconstruída para produzir uma imagem digital tridimensional em que cada voxel (elemento de volume ou pixel 3D) representa a absorção de raios X nesse ponto. Devido à relação entre a absorção de raios X e a densidade do material, a estrutura interna 3D pode ser inferida a partir das imagens, e as características internas podem ser posicionadas de forma única. As imagens 3D resultantes são normalmente exibidas como uma série de “fatias” 2D ([MACHADO et al., 2015](#); [LANDIS](#); [KEANE, 2010](#)).

A Figura 3 exibe um exemplo de tomógrafo e visualização 3D de uma amostra que foi digitalizada.

Figura 3 – (a) Microtomógrafo Zeiss-Xradia Versa 510 e (b) digitalização 3D de uma peça de concreto reforçado por fibras.



Fonte: Elaborado pelo autor.

Um exemplo do uso de imagens de Micro-CT na identificação de fissuras em concreto é o estudo de Congro et al. ([CONGRO et al., 2025b](#)), que consiste na criação de espécimes de concreto de diversas configurações de fibra e matriz cimentícia, que são submetidos a um ensaio de tração e em seguida são analisados via imagens de Micro-CT, onde o especialista avalia imagem por imagem em busca das regiões fissuradas. A aquisição de imagens por Micro-CT permite ao especialista a visualização de peças de dimensões muito reduzidas, fato que não seria viável por meio de câmeras convencionais. Além disso, essa técnica oferece a vantagem de possibilitar a visualização da estrutura interna do concreto, e não apenas da superfície, de maneira nítida e identificando as formas e limites de cada elemento do concreto ([SADOWSKI; STEFANIUK, 2017](#)). Em casos como o estudo de Congro et al., onde busca-se entender a fissuração nas heterogeneidades do concreto, onde ocorre os estágios iniciais da fissuração na escala dos milímetros, a utilização de

imagens de Micro-CT torna-se uma alternativa promissora (SILVA, 2020).

No entanto, a análise manual de imagens de Micro-CT apresenta desafios significativos para os avaliadores humanos. As fissuras são extremamente pequenas e difíceis de identificar visualmente, mesmo quando as imagens são ampliadas, tornando a detecção um processo manual, complexo e propenso a erros humanos. Além disso, os microtomógrafos geram uma quantidade massiva de imagens para cada amostra analisada. A reconstrução de um único volume 3D requer o processamento de centenas ou até milhares de imagens individuais. Como consequência, a avaliação manual de uma parcela significativa dessas imagens torna-se uma tarefa extremamente trabalhosa. Em casos de volumes muito grandes, essa abordagem pode até mesmo se tornar inviável na prática (UMBACH; MIDDENDORF, 2024).

2.4 Aprendizado Profundo

O aprendizado profundo é um subcampo do aprendizado de máquina que se fundamenta no uso de redes neurais artificiais com múltiplas camadas para a extração hierárquica de representações dos dados.

Métodos tradicionais de *machine learning* exigem a definição manual de atributos para a análise, conhecidos como características, ou *features*. O desempenho desses modelos, embora eficazes em diversas situações, estava limitado à qualidade de representação das *features* criadas, que podiam não representar em sua totalidade a complexidade das informações contidas nos dados (SHINDE; SHAH, 2018).

Os modelos de aprendizado profundo, por sua vez, são capazes de aprender automaticamente características relevantes a partir de grandes volumes de dados, organizando-as em diferentes níveis de abstração. Tratam-se de modelos que possuem camadas neurais profundas conectadas, de forma que as características extraídas não são definidas por um especialista, mas sim inferidas de acordo com os dados. Dessa forma, o aprendizado de modelos de aprendizado profundo possuem um treinamento voltado aos dados que, graças a enorme disponibilidade de dados, é capaz de aprender características muito mais subjetivas sobre os dados de entrada. Essa abordagem possibilita resolver problemas complexos em domínios como visão computacional, reconhecimento de fala, processamento de linguagem natural e análise de sinais, alcançando desempenho superior ao de técnicas convencionais em diversas tarefas (LECUN; BENGIO; HINTON, 2015).

Existem diversas redes de aprendizado profundo, cada uma projetada com propósitos e funcionamentos específicos. A seguir, serão apresentadas as redes utilizadas neste trabalho.

2.4.1 EfficientNet

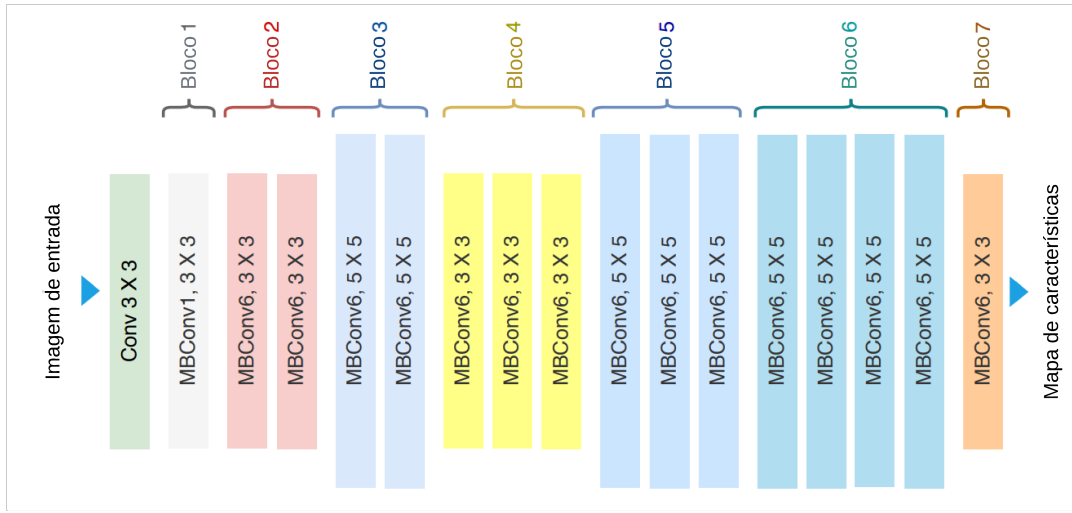
Redes Neurais Convolucionais (CNN) são um tipo de rede de *deep learning* projetada para o processamento de dados espaciais, como imagens e vídeos. Consistem em camadas de convolução, cujos filtros são aprendidos automaticamente e processam os dados de entrada aprendendo padrões espaciais complexos como bordas e texturas. Utilizam frequentemente camadas adicionais entre as convoluções, como camadas de normalização, subamostragem e regularização, que aprimoram a extração de características e filtro de informações dos modelos, contribuindo para a generalização dos modelos. Nos últimos anos, têm sido amplamente usadas para tarefas de visão computacional, como classificação, detecção e segmentação (LECUN; BENGIO; HINTON, 2015; KRIZHEVSKY; SUTSKEVER; HINTON, 2012).

Embora redes básicas de CNN fossem ideais para lidar com imagens, percebeu-se que não são necessariamente eficientes. As CNN são projetadas para uma determinada quantidade de dados e são expandidas em termos de arquitetura de acordo com a disponibilidade de dados. Quanto mais dados e recursos computacionais são disponibilizados, maiores as dimensões das redes em termos de profundidade (quantidade de camadas), largura (quantidade de neurônios presentes em cada camada) e resolução (tamanho dos dados de entrada de cada camada). As EfficientNets, por sua vez, tratam-se de uma família de redes, que vai do modelo B0 até o B7, cujo objetivo é tornar uma rede eficiente por meio de uma estratégia de escalonamento, onde as dimensões de profundidade, largura e resolução são definidas de acordo com um coeficiente de dimensionamento composto, otimizando a rede, fazendo com que cada arquitetura obtenha a melhor precisão com o menor custo computacional possível. Essa estratégia possibilitou que a EfficientNet alcançasse desempenho de estado da arte em tarefas de classificação de imagens em base de dados de *benchmarking* como a ImageNet, ao mesmo tempo em que reduziu significativamente o número de parâmetros e operações em comparação com modelos anteriores à sua publicação. Além disso, a arquitetura tornou-se referência em aplicações práticas devido à sua eficiência energética e escalabilidade, características fundamentais para o treinamento em larga escala e para a implantação em dispositivos com recursos limitados (TAN; LE, 2019).

A rede inicial, EfficientNetB0, é formada a partir da arquitetura exposta na Figura 4. A rede é formada principalmente de camadas *Mobile Inverted Bottleneck Convolution* (MBConv) propostas em (SANDLER et al., 2018). O bloco MBConv é adota uma estrutura invertida em relação aos blocos de redução de dimensionalidade tradicionais: inicialmente expande os canais por meio de uma convolução 1×1 , em seguida aplica uma convolução separável em profundidade (*depthwise convolution*) para processar cada canal de forma independente, e por fim projeta os mapas de características de volta a uma dimensionalidade reduzida com outra convolução 1×1 , esta última sem função de ativação, a fim de evitar

perda de informação em baixa dimensão. Além disso, incorpora conexões residuais quando há compatibilidade entre entrada e saída, o que favorece a propagação do gradiente, e integra mecanismos *Squeeze-and-Excitation*, responsáveis por atribuir maior peso aos canais que são mais relevantes para a tarefa e reduzindo a influência dos canais menos informativos. Sucessivos blocos com filtros 3x3 são aplicados a fim de reduzir a dimensão das imagens ao mesmo tempo que aumenta a quantidade de canais. Estes blocos são intercalados com blocos que usam filtros 5x5 a fim de capturar características globais. Essa arquitetura inicial estabelece uma base um modelo base que busca reduzir a quantidade de parâmetros do modelo sem reduzir sua precisão. Em seguida, as dimensões dos modelos seguintes são definidas de acordo com um coeficiente de dimensionamento.

Figura 4 – Arquitetura da EfficientNetB0.



Fonte: Adaptado de (AHMED; SABAB, 2022).

A Equação 2.1 mostra a definição das dimensões, onde d representa a profundidade da rede, w corresponde à largura da rede, r denota a resolução da entrada, α, β, γ são constantes que podem ser determinadas por uma pequena busca em grade, e ϕ é o coeficiente de dimensionamento, definido pelo usuário que controla quantos recursos estão disponíveis para o redimensionamento do modelo.

$$\begin{aligned} d &= \alpha^\phi, \\ w &= \beta^\phi, \quad \text{onde: } \alpha \cdot \beta^2 \cdot \gamma^2 \approx 2, \quad \alpha, \beta, \gamma \geq 1 \\ r &= \gamma^\phi, \end{aligned} \tag{2.1}$$

2.4.2 DETR

As redes baseadas em *Transformers* foram originalmente propostas para tarefas de processamento de linguagem natural, que até então eram tarefas onde modelos de

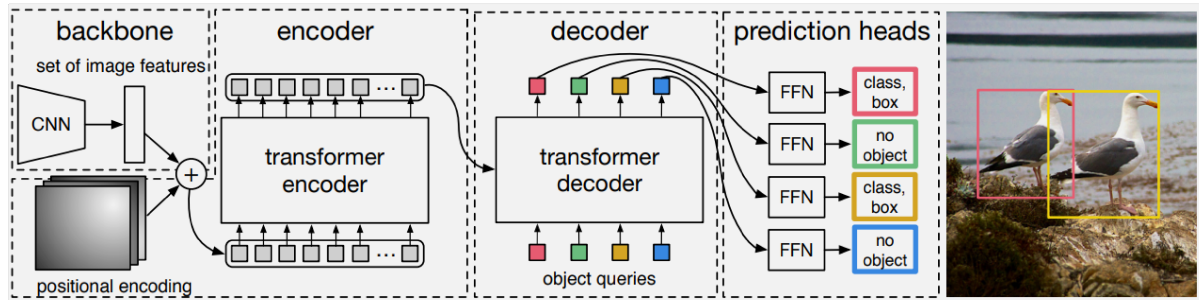
recorrência e CNNs. Contudo, Vaswani et al. (VASWANI et al., 2017) propuseram um novo tipo de rede neural baseado inteiramente em mecanismos de atenção, que permite modelar dependências de longo alcance entre elementos de uma sequência de forma mais eficiente do que arquiteturas recorrentes ou convolucionais. Em seus testes iniciais, as *Transformers* se destacaram por superar o desempenho de todos os modelos anteriores, além de serem mais paralelizáveis e necessitarem de menos tempo para treinar.

Ao perceber o potencial das *Transformers*, Dosovitskiy et al. (DOSOVITSKIY et al., 2020) introduziram os Vision Transformers (ViT). O ViT processa uma imagem como uma sequência de pequenos blocos (*patches*), geralmente de tamanho 16×16 pixels. Cada *patch* é achatado em um vetor e projetado linearmente para formar *embeddings* de dimensão fixa, semelhantes aos *embeddings* de palavras em tarefas de processamento de linguagem natural. Em seguida, é adicionado a cada *embedding* um *positional encoding*, que preserva a informação de posição de cada *patch* na imagem. Todos os *embeddings* resultantes formam uma sequência de entrada que é processada pelo mecanismo de atenção do Transformer, permitindo que o modelo capture interações globais entre diferentes regiões da imagem.

Baseado nisso, Carion et al. (CARION et al., 2020) propuseram o *Detection Transformer* (DETR), uma abordagem que reformula a tarefa de detecção de objetos como um problema de predição direta por meio de um *Transformer*. O DETR elimina a necessidade de componentes tradicionais em detecção, como âncoras e propostas de região, ao utilizar uma arquitetura codificador-decodificador baseada em atenção para mapear diretamente imagens em conjuntos de caixas delimitadoras e rótulos de classe. Essa formulação simplificou o processado de detecção de objetos, alcançou um novo estado-da-arte em base de dados conhecidos pela comunidade científica e reduziu o tempo de processamento necessário para o treinamento de um modelo de detecção de objetos.

A arquitetura da DETR pode ser visualizada na Figura 5 e é composta de 4 partes principais: a espinha dorsal (*backbone*), o codificador (*transformer encoder*), o decodificador (*transformer decoder*) e as cabeças de predição (*prediction heads*). A espinha dorsal da rede consiste em uma CNN, tipicamente a ResNet-50 (HE et al., 2016), que recebe a imagem de entrada e resulta, como saída, em um conjunto de características extraídas da imagem, com 2048 canais e redução espacial de $32 \times$ em relação à entrada.

Figura 5 – Arquitetura da DETR.



Fonte: Adaptado de (CARION et al., 2020).

Esse conjunto bidimensional resultante é então "achato", formando uma sequência unidimensional de características que são tratadas como sequências de caracteres em um problema de processamento de texto. Em uma imagem, a ordem das partes da imagem que compõem a sequência é tão importante quanto a ordem das palavras em um texto. Portanto, para tornar a DETR invariante à permutação, a sequência de características é somada a vetores posicionais.

Essa sequência de características de entrada, agora somadas aos vetores posicionais são utilizadas como entrada para o codificador. Este, por sua vez, possui várias camadas que consistem de um módulo de autoatenção de múltiplas cabeças (*multi-head self-attention*) e uma rede neural de avanço ou *feed forward network* (FFN). A saída final do codificador é um conjunto de vetores de alta dimensão, chamados *object queries*, contextualizados e cheios de informações sobre as características e suas relações espaciais na imagem original.

O decodificador, por sua vez, recebe como entrada as *object queries* geradas pelo codificador. Em seguida, as *object queries* são decodificadas individualmente em coordenadas de caixas delimitadoras (*bounding boxes*) e rótulos de classe (*labels*), resultando em N previsões finais. Usando uma autoatenção codificador-decodificador sobre essas características, o modelo usa a imagem inteira como contexto e raciocina globalmente sobre todos os objetos. Por fim, o decodificador resulta em um conjunto de previsões compostas de caixas delimitadores e rótulos de classes.

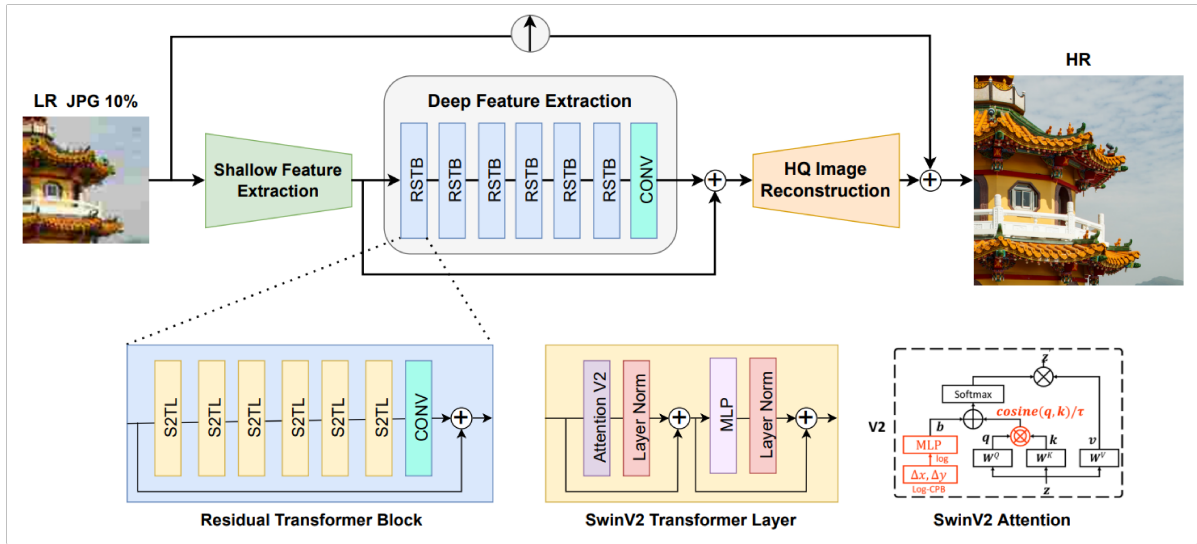
Finalmente, as cabeças de predição recebem as previsões finais geradas pelo decodificador, com camadas compostas de FFN com função de ativação *Rectified Linear Unit* (ReLU), e uma camada de projeção linear. A FFN prevê as coordenadas centrais normalizadas, altura e largura da caixa delimitadora em relação à imagem de entrada, e a camada linear prevê o rótulo da classe usando uma função *softmax*. É previsto um conjunto de tamanho fixo N, parametrizável, de caixas delimitadoras, onde N é geralmente muito maior do que o número real de objetos de interesse em uma imagem. Portanto, um rótulo de classe especial adicional é usado para representar que nenhum objeto foi

detectado, significando o fundo da imagem.

2.4.3 Swin2SR

A rede Swin2SR é concebida para aprimorar a super-resolução e a restauração de imagens. Proposta por Conde et al. (CONDE et al., 2022), sua arquitetura, que pode ser visualizada na Figura 6, é composta por três módulos principais: a extração de características rasas (*shallow feature extraction*), a extração de características profundas (*deep feature extraction*), e a reconstrução da imagem de alta qualidade (*high-quality image reconstruction*).

Figura 6 – Arquitetura da Swin2SR.



Fonte: Adaptado de (CONDE et al., 2022).

O processo se inicia com o módulo de extração de características rasas, que utiliza uma camada de convolução para processar a imagem de entrada de baixa resolução e extrair características gerais sobre a imagem. Essas características são então transmitidas diretamente ao módulo de reconstrução final para preservar os detalhes básicos da imagem.

A saída deste primeiro módulo serve como entrada para o estágio central de extração de características profundas, que é o coração da rede. Este estágio é formado por uma sequência de Blocos Residuais de SwinV2 ou *Residual SwinV2 Transformer Blocks (RSTB)*, que consistem de várias camadas de redes SwinV2 Transformer (LIU et al., 2022) seguidas de uma camada convolucional, que empregam atenção baseada em janela (*window-based attention*) para modelar dependências globais e relações entre *pixels* distantes.

Por fim, no módulo de reconstrução, as características rasas e profundas são unidas para obter uma reconstrução da imagem em alta qualidade. Para aumentar a resolução da imagem, utiliza-se uma operação de embaralhamento de pixels (*pixel shuffle*). Uma

interpolação bicúbica é aplicada na imagem de entrada e adicionada ao resultado final, a fim de complementar a recuperação de informações básicas na imagem. O resultado de todo esse processo é uma imagem de alta resolução, com artefatos de compressão significativamente reduzidos.

2.5 Métricas de Avaliação

Este trabalho se desenvolve em dois problemas específicos de visão computacional, sendo eles textcolorreda classificação de tipos de matriz cimentícia e detecção de fissuras. Portanto há algumas diferenças entre a forma como são avaliados. Esta seção apresenta as métricas que serão utilizadas para verificar o desempenho dos modelos.

Para a tarefa de Classificação serão utilizadas as seguintes métricas: Acurácia, Precisão, Recall e F1-Score. Para avaliar os modelos de detecção de objetos, serão utilizadas as métricas de Precisão, Recall, F1-Score, Average Precision (AP) e Intersection over Union (IoU). Embora a Precisão, o Recall e o F1-Score sejam usados em ambos os modelos, a forma de cálculo difere para cada um.

Para modelos de classificação binária, onde só há uma classe e o modelo precisa prever se uma determinada entrada pertence ou não à classe, o modelo retorna uma probabilidade, que é processada por uma binarização através de um limiar de confiança, a fim de converter as probabilidades em rótulos de classe. Nesse caso, cada exemplo tem uma única previsão, o que simplifica o cálculo.

Modelos de Detecção, por outro lado, fornece em sua saída rótulos, acompanhados de níveis de confiança e caixas delimitadoras. A partir disso, para calcular Verdadeiros Positivos (VP), Falsos Positivos (FP), Falsos Negativos (FN) e Verdadeiros Negativos (VN) é necessário checar se cada predição possui colisão com algum objeto real marcado na imagem.

2.5.1 Acurácia

Mede a proporção de previsões corretas em relação ao total de amostras, dada por:

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

2.5.2 Precisão

Indica a proporção de verdadeiros positivos em relação a todos os positivos previstos, definida como:

$$\text{Precisão} = \frac{\text{VP}}{\text{VP} + \text{FP}},$$

2.5.3 Recall

Também chamada de Sensibilidade, mede a proporção de verdadeiros positivos em relação a todos os positivos reais, expresso por:

$$Recall = \frac{VP}{VP + FN},$$

2.5.4 F1-Score

Média harmônica entre Precisão e Recall, dada por:

$$F1-Score = 2 \times \frac{Precisão \times Recall}{Precisão + Recall}.$$

2.5.5 Intersection over Union

Significa a qualidade de detecção de objetos, comparando a região prevista pelo modelo com a região real. Quanto mais próximo de 100, significa que mais próxima está da marcação real do especialista. Dada pela fórmula:

$$IoU = \frac{\text{Área de Interseção}}{\text{Área de União}} = \frac{|A \cap B|}{|A \cup B|}$$

2.5.6 Average Precision

Avalia a soma das precisões à medida que o Recall aumenta. Quanto maior for a métrica AP, mais o modelo poderá classificar corretamente sem gerar novos casos de falsos positivos. A equação a seguir fornece esse resultado:

$$AP = \int_0^1 p(r) dr$$

2.6 Considerações Finais

Inicialmente, neste capítulo foi abordada a composição do concreto reforçado por fibras, suas características e a importância da detecção de fissuras para estudar o concreto, bem como prevenir colapsos estruturais.

Em seguida, foram abordadas as técnicas para detecção de fissuras, com ênfase nas imagens adquiridas por Micro-CT, imagens utilizadas neste trabalho, bem como suas características e vantagens em relação às demais técnicas, o que justifica sua escolha.

Por fim, foram abordados os conceitos de *deep learning* bem como os modelos utilizados neste trabalho, além de abordar as métricas de avaliação para os mesmos.

A seguir, será mostrado os trabalhos relacionados que foram conduzidos até o momento pela comunidade acadêmica, bem como a identificação de lacunas que este trabalho busca sanar.

3 Trabalhos Relacionados

Estudos sobre detecção automática de fissuras em concreto já vem sendo realizados, mostrando que é um tema de interesse da comunidade acadêmica. Primeiro, elicitam-se técnicas tradicionais sem o uso de aprendizado profundo. Posteriormente, serão vistos trabalhos que utilizam modelos de aprendizado profundo para detecção automática de microfissuras utilizando Micro-CT disponíveis na literatura.

3.1 Detecção utilizando deep learning em macroescala

Nos últimos anos, a aplicação de modelos de *deep learning* tem se consolidado como uma abordagem promissora para a automação de tarefas que seriam exaustivas e suscetíveis a erros quando realizadas manualmente por seres humanos. A detecção de fissuras em concreto constitui um dos problemas recorrentes na literatura, contando com diversas soluções baseadas em técnicas de aprendizado profundo (GOLDING et al., 2022; DAHLHOFF; RAUPACH, 2023; ZHOU et al., 2022).

Diniz et al. (DINIZ et al., 2023) propuseram um método automatizado baseado na rede neural YOLOv4 para detectar patologias visuais em superfícies de estruturas de concreto; fissuras, fragmentação, corrosão e barras de aço expostas diretamente a partir de imagens feitas por câmeras convencionais. A base de dados utilizada foi a CODEBRIM, onde cada imagem possui especificamente uma patologia, resultando em 11,80% de mAP na detecção de fissuras.

Sohaib et al. (SOHAIB; JAMIL; KIM, 2024), por sua vez, desenvolveram uma abordagem baseada em *ensemble* de modelos para realizar uma segmentação final precisa usando a base de dados pública *Dataset for Drone-based Inspection of Road Pavement Structures for Cracks* composta de 4215 imagens de fissuras adquiridas a partir de diversos modelos de câmeras a bordo de drones. Utilizaram-se das diversas versões da rede YOLOv8, onde cada versão realizou uma segmentação individual. A segmentação final, portanto, se tratou da interseção das máscaras dos três modelos treinados, resultando em uma segmentação final mais apurada. O método de *ensemble* foi testado em três datasets disponíveis publicamente, sendo eles: *Heterogenous Dataset*, *Pavement Cracks Dataset* e *Deepcrack Dataset*, alcançando a métrica 89,15% de F1-Score médio, superando outros resultados da literatura.

Park et al. (PARK; EEM; JEON, 2020) desenvolveram um estudo para detecção de fissuras em estruturas de concreto baseado em um drone equipado com uma câmera e *lasers*. As imagens das câmeras são utilizadas como entrada para uma CNN Yolo (REDMON

et al., 2016), treinada para detectar fissuras em concreto em uma base de dados privada criada pelos autores. Após detectá-las, o drone utilizava os pontos de lasers projetados para mensurar o tamanho das fissuras. O método então seria equipado em um drone que faria o papel de um inspetor ao monitorar estruturas de concreto que necessitam de manutenções preventivas, identificando pontos danificados com antecedência e sem a necessidade de alocar um trabalhador humano.

Há vários outros estudos lidando com detecção de fissuras em concreto usando algoritmos de Deep Learning. Lee et al. (LEE; KIM; LEE, 2019) propôs um método robusto de detecção de fissuras por meio de segmentação usando uma rede CNN, capaz de segmentar corretamente até mesmo imagens completas com a presença de objetos como vegetação, que dificultam a delimitação da trinca. O método proposto alcançou um F1-Score próximo a 80% utilizando uma base de dados privada composta de 542 imagens.

Guo et al. (GUO; MENG; BAO, 2024) estudaram o uso de *Conditional Generative Adversarial Networks* (CGAN) em uma rede U-Net com mecanismos de atenção para melhorar a detecção de pequenas fissuras e escala macro. Este método utilizou uma base de dados privada de 2200 imagens de dimensões 512x512 *pixels* alcançou métricas de O F1-Score e IOU de 98,2% e 98%, respectivamente.

Os resultados de todos esses estudos são satisfatórios no que diz respeito à automação da detecção de fissuras em macroescala. No entanto, nenhum deles apresentou dados em microescala (nanômetros), nem em imagens de Micro-CT. As imagens utilizadas foram fotos em escala comum. Portanto, os métodos apresentados não são capazes de identificar o estado interno das estruturas de concreto, nem foi documentada a detecção de fissuras em microescala.

3.2 Detecção de Fissuras em Micro-CT

Alguns estudos foram identificados quando se diz respeito à utilização de Micro-CT para encontrar microfissuras em concreto.

Li et al. (LI et al., 2023) realizaram uma análise da progressão da fissuração em microescala usando uma base de dados privada de 20 volumes de imagens de Micro-CT, cada volume contando com 454 fatias. Com uma CNN, conseguiram identificar fissuras no corpo de amostras de concreto por meio de sucessivos ensaios de tração. No entanto, esse é um estudo sobre a progressão e a morfologia da fissuração do concreto, sem detalhes adequados sobre o número de imagens e a métrica de detecção do modelo usado em relação à marcação do especialista.

(TIAN et al., 2021) desenvolveram um método para detectar fissuras e vazios em Micro-CT de concreto com base em um método de segmentação seguido de classificação. A

base de dados utilizada privada é composta de 10 volumes, totalizando 2400 imagens, onde os espécimes foram submetidos a testes de carga consecutivos para aumentar a fissuração. A base de dados foi aumentada utilizando operações de rotação e espelhamento, gerando um conjunto de dados de 5.000 imagens de Micro-CT de concreto em micro escala. O método proposto foi capaz de identificar a maioria das fissuras e vazios nos espécimes, alcançando métricas de 94,2% de IoU. Porém, trata-se de um estudo de segmentação e as demais métricas de avaliação não são fornecidas.

Por fim, (DONG et al., 2020) desenvolveram uma CNN Dilatada (DCNN) baseada em convolução atroz que foi treinada separadamente para detectar vazios, agregados e fissuras em imagens de Micro-CT de concreto em microescala. A base de dados utilizada também é privada e composta de 3 volumes, totalizando 1223 imagens. Se tratando de fissuras, o modelo conseguiu atingir 78,5% de F1-Score, sendo capaz de detectar a maioria das fissuras presentes na base de dados utilizada na pesquisa.

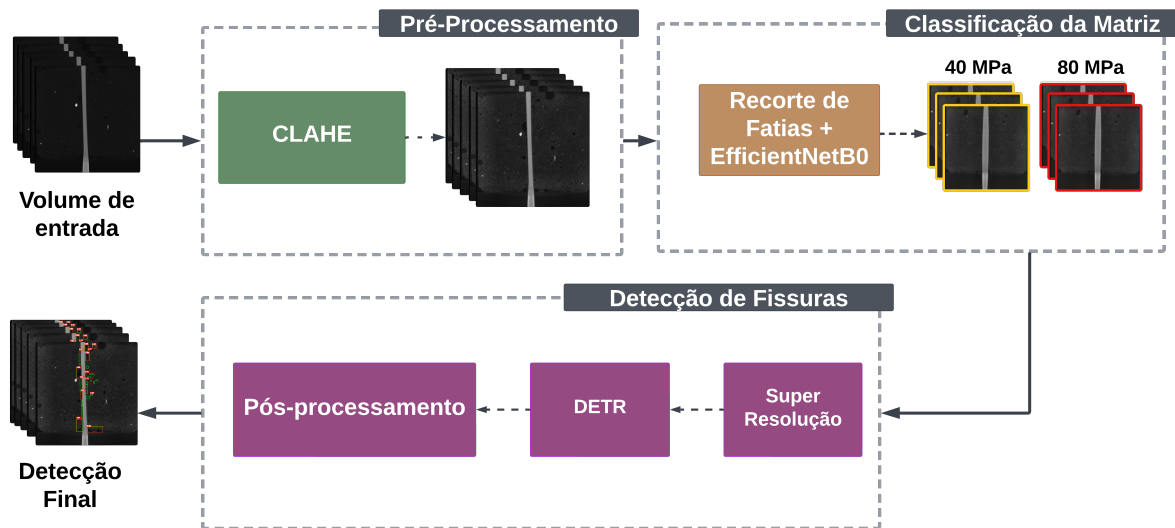
Com isso, podemos perceber que há vários estudos no campo da macroescala, todos atingindo resultados satisfatórios na detecção de fissuras. Contudo, quando se trata do campo da microescala, percebe-se que, apesar de haver estudos com resultados bons, nenhum estudo sobre concreto reforçado por fibras foi identificado. Em adicional, também não foi identificado nenhum estudo onde a base de dados é pública, de forma a contribuir com outras pesquisas no avanço da detecção de fissuras em imagens de Micro-CT. Outro ponto é a ausência de base de dados que possuem marcações para a tarefa de detecção de objetos, como caixas delimitadoras. Todas as bases dos trabalhos relacionados utilizam-se de máscaras de segmentação, que foram geradas ou por especialistas ou por software. Como a base de dados não é fornecida, também não é possível extrair as caixas delimitadoras a partir das máscaras de segmentação.

Por fim, é perceptível a discrepância entre os métodos disponíveis para macroescala e microescala. Portanto, torna-se importante um estudo aprofundado sobre detecção de microfissuras em imagens de Micro-CT de concreto reforçado por fibras, por se tratar de um concreto com um tipo de agregado diferente, e que investigue e documente adequadamente os dados obtidos como um problema de detecção de objetos, para fornecer informações para estudos futuros.

4 Método Proposto

Este trabalho propõe um método para detecção de fissuras em imagens de Micro-CT de concreto reforçado por fibras utilizando modelos de *deep learning*. O método consiste nas etapas mostradas na Figura 7.

Figura 7 – Etapas do método proposto.



Fonte: Elaborado pelo autor.

Serão explicados os detalhes de aquisição das imagens que compõem a base de dados utilizada neste trabalho, bem como de marcação por parte de um especialista na Seção 4.1. Neste trabalho está sendo disponibilizada para a comunidade uma nova base de dados de imagens de Micro-CT de concreto reforçado por fibras. De acordo com nossa revisão bibliográfica, esta é a primeira base pública para este fim.

Em seguida, seguem as etapas de pré-processamento para ajustar as imagens adquiridas para as entradas de seus respectivos modelos. Em seguida, ocorre a separação entre as matrizes presentes na base, com o intuito de produzir modelos especializados para cada matriz. Por fim, o método proposto realiza a detecção das fissuras automaticamente. Mais detalhes sobre cada etapa do método estão descritos nas seções que se seguem.

4.1 Aquisição de Imagens

A preparação dos corpos de prova, os procedimentos de ensaios de tração, os protocolos de aquisição de imagens foram realizados por Congro et al. (CONGRO et al., 2025a). Duas matrizes distintas de cimento foram preparadas: (i) Matriz A, com uma

resistência à tração de 40 MPa (milipascals), formulada com uma relação água/cimento de 0,5, incorporando cimento ASTM tipo IL - Cimento Portland-Calciário e (ii) Matriz B, com resistência à compressão de 80 MPa, formulada com uma relação água/cimento de 0,4 e incorporando cimento ASTM C150 Tipo III. As fibras de aço utilizadas nos ensaios de tração foram do tipo DRAMIX 3D 80/60, com um comprimento de 60 mm e um diâmetro de 0,75 mm. Foram consideradas duas geometrias de fibra: fibras de aço retas e fibras de aço com extremidades em gancho. O comprimento de encaixe foi fixado em 40 mm para todas as amostras. Os corpos de prova foram projetados com uma geometria cilíndrica. Cada corpo de prova possui 40 mm de altura e 26 mm de diâmetro, que consiste em: Uma parte maior com 25 mm de altura e 26 mm de diâmetro, localizada na parte de baixo do corpo de prova, utilizada para fixação na máquina de ensaio de tração; Uma parte menor, localizada na parte de cima do corpo de prova, com altura de 15 mm e diâmetro de 11 mm, com a fibra localizada no meio, que é a região de interesse do corpo de prova tomografada pelo microtomógrafo.

Os ensaios de tração foram conduzidos usando uma máquina de ensaios mecânicos universal EMIC® DL 3000 com uma capacidade de carga máxima de 30 kN (quiloNewtons). Uma célula de carga de 2 kN foi utilizada para garantir a medição precisa da força. Os testes foram realizados sob controle de deslocamento a uma taxa de 0,5 mm/min.

A digitalização de Micro-CT foi gerada a partir de um microtomógrafo Zeiss-Xradia Versa 510 com um tamanho de voxel de 5 micrômetros para garantir imagens de alta resolução da interface fibra/matriz. Cada corpo de prova foi digitalizado após o ensaio de tração para gerar volumes com fissuras. Os parâmetros da digitalização seguem como apresentado em Congro et al. e são mostrados na Tabela 1

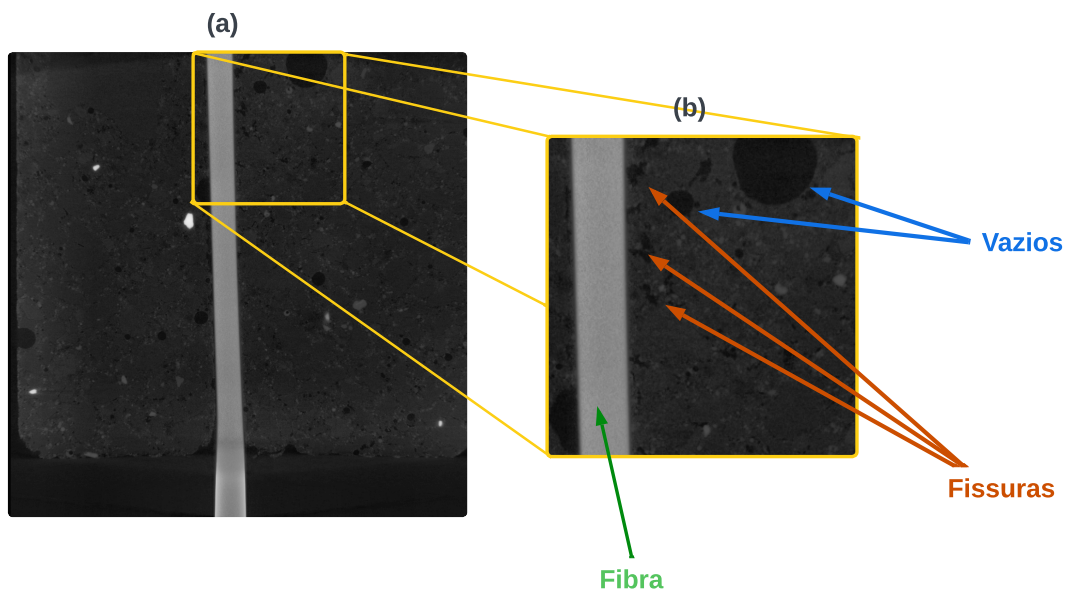
Tabela 1 – Parâmetros de aquisição para a digitalização em Micro-CT (CONGRO et al., 2025a).

Informação	Valor
Lente óptica	0.4x
Camera binning	1
Campo de visão	2048 x 2048 pixels
Tamanho de voxel	5 μ m
Voltagem	80 kV
Potência	7 W
Tempo de exposição	6.4 - 8 sec

No final desta etapa, a região de cada corpo de prova de dimensões 11 mm x 15 mm gera um volume 3D de dimensão $1000 \times 1000 \times 1000$ *pixels*. A Figura 8 mostra uma fatia central de um volume gerado a partir de um dos corpos de prova. Na seção destacada em amarelo, as heterogeneidades dentro do concreto são visíveis. A fibra aparece proeminentemente em uma intensidade maior, seguida dos agregados, que também são

distinguíveis. As regiões mais escuras correspondem às discontinuidades presentes no concreto, incluindo vazios, que geralmente são causados por bolhas de ar durante a preparação do concreto, e fissuras apresentando estágio inicial de deterioração do concreto. As fissuras identificadas na imagem são bem pequenas, na dimensão dos microns, de forma que sua visualização é um desafio mesmo com o aumento de contraste apresentado na figura.

Figura 8 – (a) Fatia de um volume de Micro-CT de concreto reforçado por fibras e (b) a visualização aproximada da região destacada em amarelo, mostrando os principais constituintes do concreto.



Fonte: Elaborado pelo autor.

4.2 Marcação e Publicação da Base de Dados

Para todos os dez volumes na base de dados, fatias do plano XZ foram usadas, totalizando 4064 imagens contendo fissuras, com 1901 imagens pertencendo à Matriz A (40 MPa) e 2163 imagens pertencendo à Matriz B (80 MPa).

O processo de marcação foi conduzido pelo software LabelImg (versão 1.8.1 para Windows 11)¹. Para garantir a precisão e confiabilidade, todas as marcações foram conduzidas por um engenheiro civil com expertise em compósitos cimentícios e análise de imagens de Micro-CT de concreto. As marcações foram salvas em arquivos de formato

¹ Disponível em: <https://github.com/HumanSignal/labelImg>

.txt contendo marcações seguindo o formato das caixas delimitadores utilizadas na YOLO (REDMON et al., 2016).

A base de dados marcada foi nomeada "*FIRECON - FIber REinforced CONcrete Cracks Detection Dataset*" e está disponível publicamente em: <<https://kaggle.com/datasets/a56850cac6de7ecd9ca296d0a02cc664ec0971d1182863bb06e1e1d97b98de78>>.

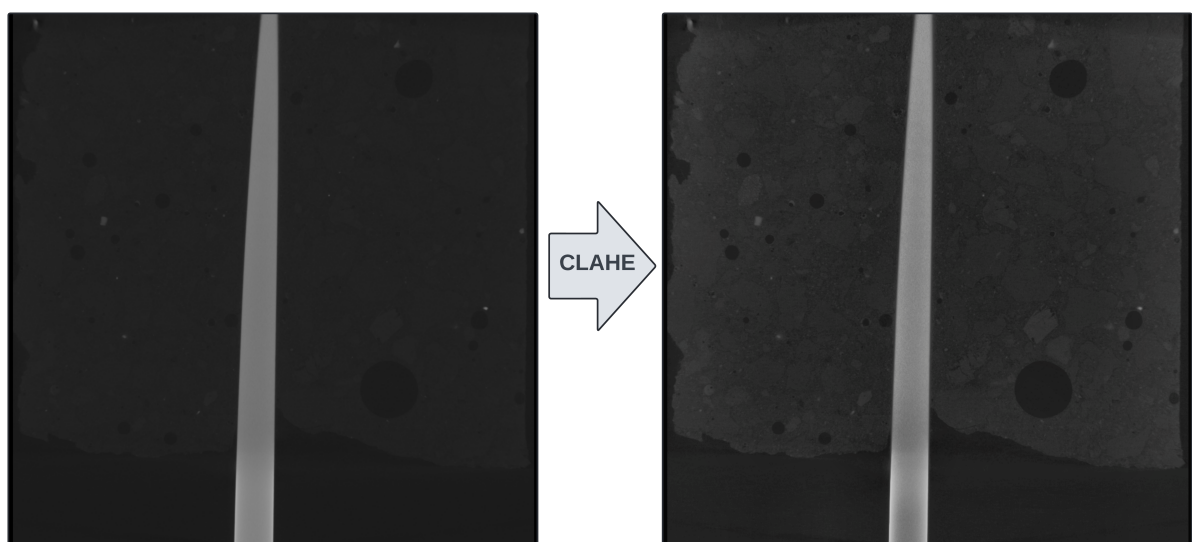
4.3 Pré-Processamento

A fim aumentar o contraste das imagens da base e melhorar a visualização dos detalhes, bem como realçar as bordas das heterogeneidades do compósito, como pré-processamento foi aplicada uma Equalização do Histograma Adaptativa Limitada ao Contraste ou *Contrast-Limited Adaptive Histogram Equalization (CLAHE)* (MUSA; RAFI; LAMSANI, 2018).

Como parâmetros para a aplicação do CLAHE, foi estabelecido que o limite dos picos de intensidade não poderiam exceder o dobro do valor médio, além de *kernel* de dimensão 8×8 pixels. O resultado pode ser visualizado na Figura 9.

As imagens resultantes demonstram uma distinção mais aparente entre os elementos do concreto. Melhorar o contraste das imagens facilita a identificação de características menos discerníveis nas imagens originais, como os limites das fibras e dos vazios, agregados e potenciais fissuras. O pré-processamento se mostrou crucial para melhorar a performance das etapas subsequentes deste método. Sem a utilização de um ajuste de contraste, é muito difícil diferenciar os elementos presentes na imagem a olho nu.

Figura 9 – Imagem do concreto com fibra gerado a partir de Micro-CT, antes e depois da aplicação do CLAHE.



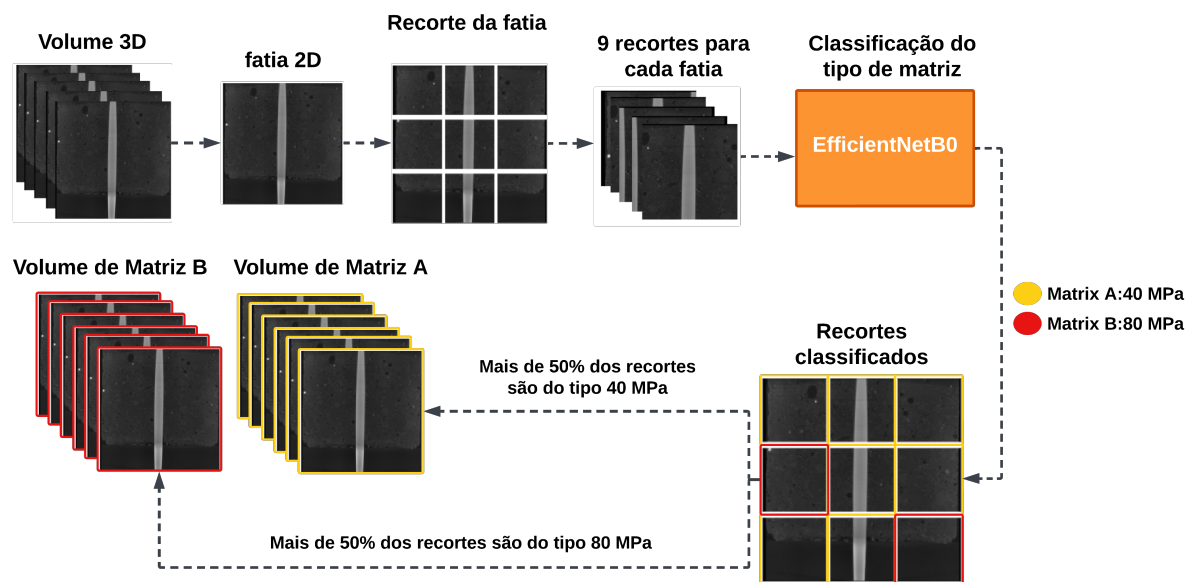
Fonte: Elaborado pelo autor.

4.4 Classificação de Tipo da Matriz Cimentícia

Esta etapa consiste em classificar as imagens presentes na base de dados entre os dois tipos de matriz cimentícia: Matriz A (40 MPa) e Matriz B (80 MPa). Esta etapa utiliza a arquitetura EfficientNetB0 para classificar as imagens de Micro-CT, como é ilustrado na Figura 10.

O objetivo dessa classificação é separar os dados de acordo com o tipo de matriz para a etapa posterior. Como será demonstrado nos experimentos no Capítulo 5, a separação entre os tipos de matrizes melhora o resultado final. Modelos de detecção treinados especificamente para um tipo de matriz possuem melhor desempenho que treinar utilizando ambos, pois desta forma os estágios posteriores podem focar em aprender e extrair características de um único tipo de matriz, ao invés de dois.

Figura 10 – Fluxo da etapa de classificação de matriz cimentícia.



Fonte: Elaborado pelo autor.

Cada fatia é dividida entre nove partes iguais para um determinado volume, e então são utilizadas como a entrada de um classificador binário com a arquitetura da EfficientNetB0. Como decodificador da rede, utiliza-se ao final uma camada completamente conectada que recebe 1280 características do codificador, responsável por aplicar uma transformação linear para reduzir a saída do modelo para duas classes que significam, respectivamente, Matriz A e Matriz B.

Dividir cada imagem em nove recortes reduz a complexidade visual, fazendo com que a rede adquira características mais relevantes sobre a matriz, além de funcionar como um aumento de dados, visto que a quantidade de imagens de entrada se multiplica.

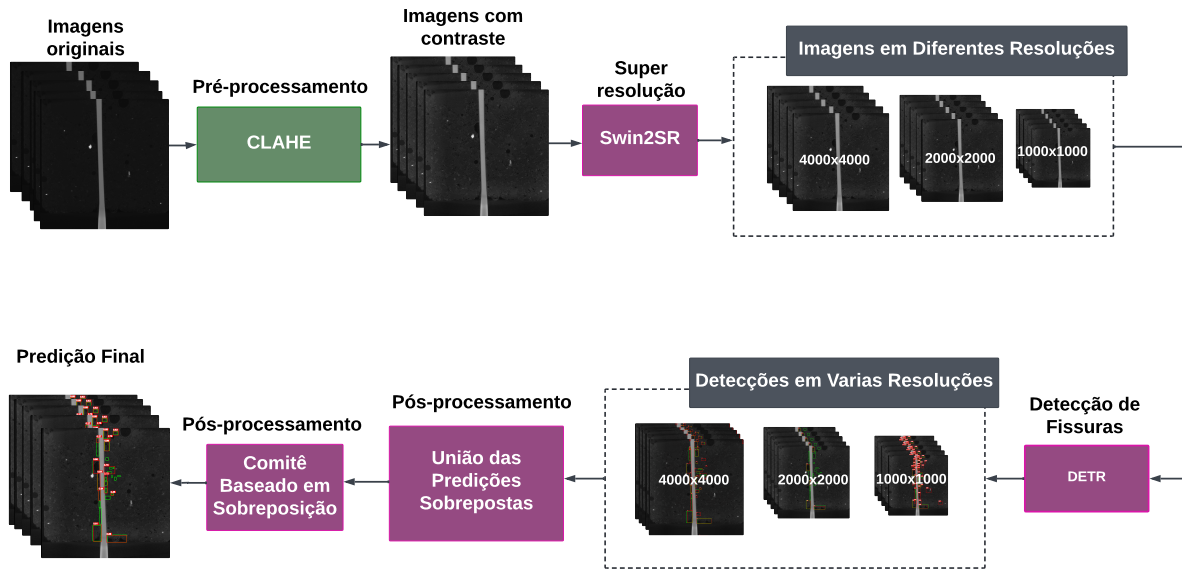
O classificador binário prediz o tipo da matriz de cada uma dessas nove imagens.

Um limiar então é aplicado para determinar a classificação final de cada fatia: Se mais da metade (ou seja, pelo menos 5 recortes) forem classificados como uma única classe, então a fatia inteira é classificada com a mesma classe. Da mesma forma, se mais da metade das fatias presentes em um volume (precisamente, quinhentas e treze fatias) são classificadas com a mesma matriz, então o volume inteiro é classificado com o mesmo tipo de matriz. Isso reduz a possibilidade de erros ou de que fatias de um mesmo volume sejam classificadas em tipos de matrizes diferentes.

4.5 Detecção de Fissuras

Após a classificação do tipo de matriz, cada imagem é direcionada a um processo de detecção de fissuras que se utiliza de um pré-processamento baseado em super-resolução para aumentar a confiabilidade e precisão da detecção, assim como ilustrado na Figura 11.

Figura 11 – Fluxo da etapa de detecção de fissuras.



Fonte: Elaborado pelo autor.

O primeiro passo da detecção é utilizar o modelo Swin2SR (CONDE et al., 2022) para melhorar a visibilidade de detalhes pequenos e ampliar as características da imagem. Trata-se de um modelo de estado-da-arte para gerar imagens em super resolução. O modelo gera versões de uma mesma imagem de alta resolução nas resoluções 2000 x 2000 e 4000 x 4000 pixels, além da imagem original nas dimensões 1000 x 1000 pixels.

Gerando duas amostras em super resolução da mesma imagem, o método expande suas informações visuais disponíveis, permitindo que o modelo de detecção capture detalhes mais sutis em diferentes escalas que talvez são imperceptíveis na dimensão original. Essa detecção em múltiplas resoluções tem uma função crucial em melhorar a detecção de

fissuras pequenas e melhorar a confiabilidade da predição final aliada ao pós-processamento que será apresentado.

Em seguida, a rede Detection Transformer (DETR) (CARION et al., 2020) é utilizada para detecção de fissuras. Três modelos são treinados, um para cada resolução, gerando 3 predições para cada imagem. Essa abordagem de detecção utiliza informações complementares em cada uma das resoluções.

4.6 Pós-processamento

Após a etapa de detecção, três predições diferentes são geradas para uma mesma imagem. Uma etapa de pós-processamento de dois estágios é aplicada para redução de falsos positivos e melhoria da detecção final. Primeiramente, as predições sobrepostas em uma mesma resolução são unidas, de forma que dimensão da caixa delimitadora final englobe todas as caixas delimitadores sobrepostas e a confiança final se torna a média das confianças das caixas delimitadores envolvidas na união.

Em seguida, as predições de todas as resoluções são unidas em uma única precisão final, por meio de uma abordagem baseada em comitê. Para cada predição, o algoritmo checa se a mesma fissura foi predita em uma outra resolução. Para isso, é necessário que haja uma interseção sobre união superior 80% entre as predições de resoluções diferentes. Caso haja alguma interseção, é verificado se a soma das confianças das predições em questão é superior ao limiar de 50%. Em caso positivo, a predição final será a caixa delimitadora da sobreposição das predições envolvidas na interseção e a confiança da predição final será igual a maior confiança entre as predições em interseção.

Este método garante que a predição final inclua somente predições confirmadas por pelo menos dois dos modelos de resoluções diferentes e que essas predições tenham uma alta confiança. Predições entre dois ou mais modelos cuja confiança é inferior a 50% ou então que somente um modelo prediz são considerados falsos positivos e portanto são descartados. Desta forma, a predição final conta somente com caixas delimitadoras de confiança alta e sem sobreposições.

5 Experimentos e Resultados

Esse capítulo apresenta os resultados obtidos dos experimentos realizados com o método proposto. Também é discutido o impacto de cada etapa do método.

5.1 Classificação do tipo de matriz

Para criar a base de dados para treinar e avaliar o classificador de tipo de matriz, fatias centrais foram selecionadas manualmente dentre os volumes. Como o formato de cada corpo de prova é cilíndrico, as fatias próximas às margens do volume são compostas predominantemente do fundo da imagem (*background*), ilustrados como *pixels* de cor preta na Micro-CT, de forma que não possuem informação relevante o suficiente na imagem toda para contribuir com o treinamento. Essas fatias portanto não são utilizadas.

As fatias centrais, por outro lado, possuem pouca ou nenhuma predominância de fundo, pois quanto mais próximo do centro do cilindro, mais informação a imagem contém, se assemelhando a fatias de um cubo. Em seguida, cada fatia é recortada em 9 janelas iguais, resultando em uma base de dados final consistindo de 16.281 imagens da Matriz A e 26.838 imagens da Matriz B.

Como pré-processamento, cada imagem é redimensionada para as dimensões 224 x 224 pixels, para equivaler às dimensões de entrada da EfficientNetB0. Esta, por sua vez, foi treinada utilizando os seguintes hiperparâmetros: tamanho de batch 32, otimizador Adam, Binary Cross-Entropy como função de perda, 100 épocas, decaimento de taxa de aprendizagem de 0,0001 e 10 épocas de paciência, ou seja, a cada 10 épocas a taxa de aprendizado irá decair a uma taxa de 0,0001 caso não haja melhora nas métricas.

A rede foi treinada e avaliada usando hold-out cross-validation com 5 iterações, onde 80% dos dados foram usados para treinamento e 20% para teste, garantindo que não houvessem dados repetidos em cada divisão.

Um estudo de ablação foi realizado para comparar a performance da EfficientNetB0, enquanto classificador de tipo de matriz de concreto, com outras arquiteturas. Todas as redes foram treinadas seguindo os mesmos hiperparâmetros. Os resultados obtidos estão na Tabela 2.

A EfficientNetB0 desempenhou-se satisfatoriamente melhor que as demais, e foi capaz de diferenciar corretamente entre os tipos de matriz de concreto presentes na base. O desvio padrão sugere que o modelo fez uma boa generalização do problema, classificando corretamente as imagens da base de teste. A Tabela 3 mostra que, de acordo com os experimentos conduzidos, o número de erros para cada tipo de matriz é insuficiente para

Modelo	Acurácia	Precisão	Recall	F1-Score
EfficientNetB0	0,9986 $\pm$ 0,0005	0,9974 $\pm$ 0,0016	0,9989 $\pm$ 0,0012	0,9982 $\pm$ 0,0006
AlexNet	0,8509 $\pm$ 0,0071	0,8346 $\pm$ 0,0065	0,8721 $\pm$ 0,0073	0,8518 $\pm$ 0,0069
DenseNet121	0,9204 $\pm$ 0,0043	0,9022 $\pm$ 0,0039	0,9416 $\pm$ 0,0047	0,9227 $\pm$ 0,0041
ResNet101	0,9316 $\pm$ 0,0037	0,9235 $\pm$ 0,0035	0,9543 $\pm$ 0,0042	0,9348 $\pm$ 0,0038
ViT_b_16	0,9432 $\pm$ 0,0031	0,9378 $\pm$ 0,0028	0,9654 $\pm$ 0,0035	0,9449 $\pm$ 0,0032

Tabela 2 – Desempenho

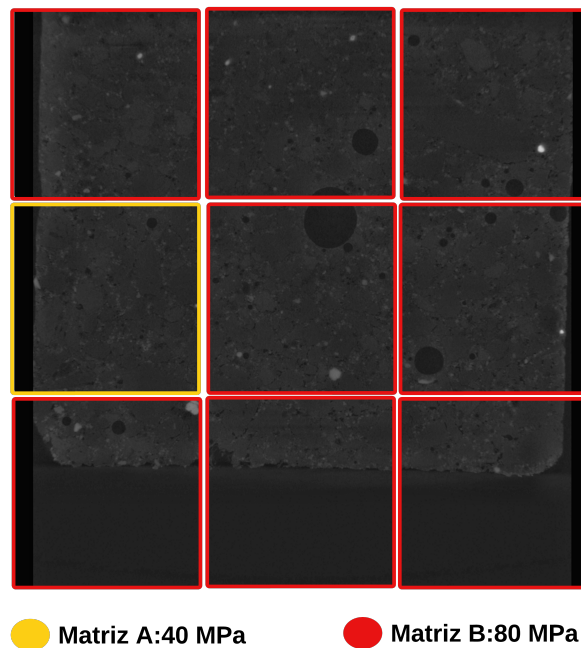
dos modelos de classificação de tipos de matriz, com média e desvio padrão

classificar um volume de maneira incorreta, pois a quantidade de erros é menor do que a quantidade de imagens necessárias para classificar uma imagem erroneamente e, por consequência, um volume. Isso também pode ser visto na Figura 12, que mostra um exemplo de classificação contendo uma janela predita na classe errada. Portanto, de acordo com os experimentos conduzidos e resultados obtidos, há evidências de que o modelo foi capaz de classificar corretamente os tipos de matriz de concreto presentes na base de dados.

Class	Matriz A	Matriz B
Matriz A	3254	3
Matriz B	3	5364

Tabela 3 – Matriz de Confusão da classificação binária de tipos de matriz

Figura 12 – Resultado da classificação de uma fatia da Matriz B com apenas uma janela classificada como Matriz A, quantidade insuficiente para classificá-la erroneamente.



Fonte: Elaborado pelo autor.

A Tabela 4 mostra que, de acordo com os experimentos executados, a EfficientNetB0 melhora a performance se comparada a um classificador treinado com os mesmos hiperparâmetros, mas com a imagem inteira ao invés de janelas. Para imagens de entrada com a resolução menor, os resultados evidenciam que o modelo pode extrair características mais significantes para a classificação da matriz.

Tipo de entrada	Acurácia	Precisão	Recall	F1-Score
Imagem Inteira	0,7195 $\pm$ 0,0020	0,8470 $\pm$ 0,1827	0,1533 $\pm$ 0,2268	0,1721 $\pm$ 0,1922
Janelas	0,9986 $\pm$ 0,0005	0,9974 $\pm$ 0,0016	0,9989 $\pm$ 0,0012	0,9982 $\pm$ 0,0006

Tabela 4 – Performance of EfficientNetB0, de acordo com o tipo de imagem de entrada.

5.2 Detecção de Fissuras

Os dados para treinamento do modelo de detecção de fissuras são menores. Ao todo 4064 fatias foram anotadas, sendo 1901 para a Matriz A e 2163 imagens para Matriz B. A divisão entre treino, validação e teste foi feita de acordo com o seguinte critério: Para cada volume 80% das fatias são seleccionadas aleatoriamente para treino, 10% para treino e 10% para teste, de forma que não se repitam entre treino, teste e validação.

Os hiperparâmetros para o treinamento da DETR seguem: tamanho de batch 4, taxa de aprendizagem 0,00001, Otimizador AdamW, Focal Loss como função de perda, 100 épocas e decaimento de taxa de aprendizagem de 0,00001 a cada 10 épocas.

Um estudo de ablação também foi conduzido para os modelos de detecção de fissuras, comparando a DETR com outras duas versões da arquitetura YOLO (v8 e v10), utilizando os mesmos hiperparâmetros. As métricas alcançadas são mostradas nas Tabelas 5 e 6.

Modelo	IoU	Precisão	Recall	F1-Score	AP
Yolov8	0,1421	0,1820	0,0550	0,0850	0,0750
Yolov10	0,2394	0,2680	0,0910	0,1364	0,1260
DETR	0,6353	0,6831	0,5363	0,5701	0,3416

Tabela 5 – Métricas de detecção de fissuras para o tipo de matriz A

Modelo	IoU	Precisão	Recall	F1-Score	AP
Yolov8	0,1325	0,1502	0,0764	0,1003	0,0857
Yolov10	0,2264	0,2518	0,1149	0,1573	0,1410
DETR	0,8609	0,8937	0,8326	0,8499	0,4469

Tabela 6 – Métricas de detecção de fissuras para o tipo de matriz B

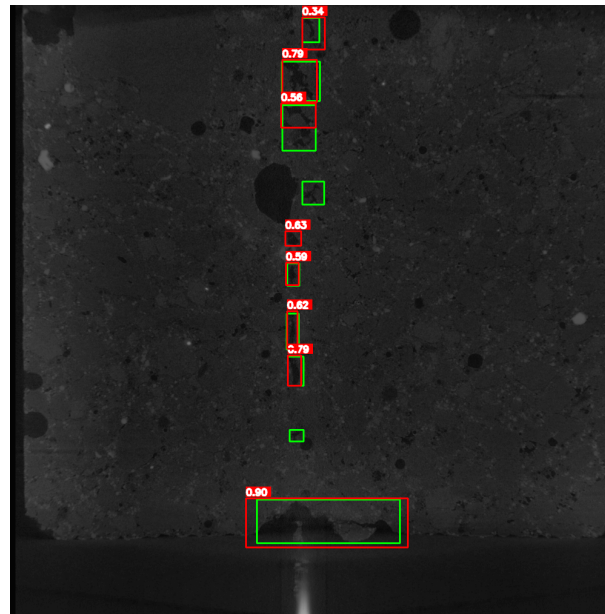
Como é perceptível nas métricas, ambos os modelos foram capazes de detectar fissuras automaticamente. As métricas de *Recall* e AP indicam que o modelo treinado

para a Matriz A não foi capaz de reconhecer algumas fissuras. Por outro lado, o modelo foi capaz de aprender características para a Matriz B, alcançando métricas superiores a 80%, exceto AP, o que indica que, o modelo pode gerar falsos positivos a medida em que se utiliza limiares menores de confiança.

A rede DETR superou a performance das duas versões da YOLO utilizadas, sugerindo que, dada a base de dados utilizada, o modelo é capaz de aprender características profundas relacionadas às fissuras que os demais modelos não foram. Entretanto, até mesmo utilizando a DETR, o método proposto assumiu uma postura conservativa em relação à predição, detectando apenas parte das marcações do especialista.

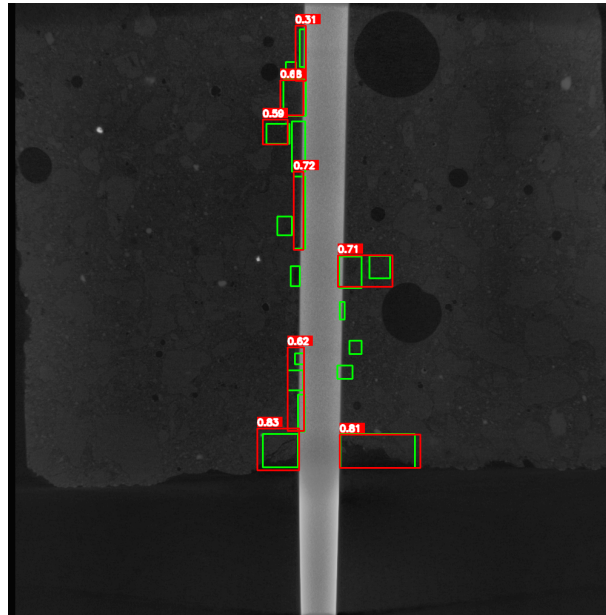
As Figuras 13 e 14 mostram um exemplo para cada tipo de matriz da detecção final do método proposto. Mesmo que o método seja capaz de identificar algumas fissuras com precisão, é notável que muitas fissuras não são detectadas, principalmente as menores. Em adicional, nota-se que algumas predições circundam algumas fissuras menores, mas com forma e tamanho diferentes, de forma que o modelo é capaz de identificar uma região fissurada, mas não detecta as fissuras em si com precisão. Como nesses casos o IoU entre a predição e a marcação do especialista é inferior a 50%. Consequentemente, a predição não conta como um acerto, uma vez que está distante demais da forma da marcação do especialista.

Figura 13 – Comparação entre a marcação do especialista(em verde) e a predição do modelo (em vermelho) para uma imagem do tipo Matriz A.



Fonte: Elaborado pelo autor.

Figura 14 – Comparação entre a marcação do especialista(em verde) e a predição do modelo (em vermelho) para uma imagem do tipo Matriz B.



Fonte: Elaborado pelo autor.

É importante notar o fator de redução de falsos positivos do método proposto. As Tabelas 7 e 8 mostram as métricas que o método alcançou utilizando a DETR treinada com cada resolução, comparada com o resultado final contendo a etapa de redução de falsos positivos. É notável que algumas predições que não são relevantes são descartadas porque não há a interseção com predições de modelos treinados com resoluções diferentes ou porque não há confiança suficiente para a mesma ser mantida no resultado final. Em outros casos, a união entre duas predições resultou em um acerto com maior IoU entre a predição final e uma marcação do especialista após o pós-processamento. Esse pós-processamento melhora as métricas do modelo e promove uma detecção final de maior confiabilidade.

Modelo	IoU	Precisão	Recall	F1-Score	AP
DETR - 1x	0,5109	0,1776	0,4585	0,2168	0,0888
DETR - 4x	0,3545	0,3835	0,3681	0,3411	0,1917
DETR - 2x	0,4609	0,6149	0,2447	0,3506	0,1505
Comitê	0,6353	0,6831	0,5363	0,5701	0,3416

Tabela 7 – Comparação das métricas obtidas entre as predições de resoluções diferentes e a predição final utilizando o pós-processamento baseado em comitê, para Matriz do tipo A.

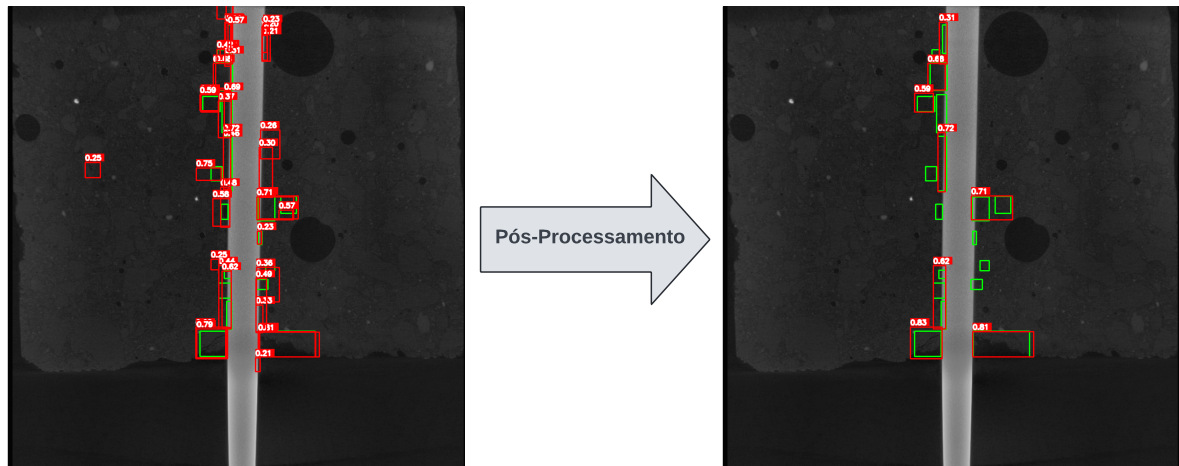
Modelo	IoU	Precisão	Recall	F1-Score	AP
DETR - 1x	0,4805	0,7645	0,3210	0,4481	0,5428
DETR - 2x	0,4073	0,5751	0,1699	0,2612	0,3725
DETR - 4x	0,4136	0,6687	0,5856	0,6245	0,6272
Comitê	0,8609	0,8937	0,8326	0,8499	0,4469

Tabela 8 – Comparação das métricas obtidas entre as predições de resoluções diferentes e a predição final utilizando o pós-processamento baseado em comitê, para Matriz do tipo B

A Figura 15 mostra a representação visual da redução de falsos positivos após o pós-processamento baseado em comitê. Muitas predições sem interseção entre as resoluções foram descartadas. Em adicional, aquelas que foram preditas em mais de uma resolução e tinha confiança o suficiente foram mantidas e consideradas como uma predição única no final.

Os resultados mostram que predições feitas por modelos das três resoluções diferem entre si na base de dados utilizada, tanto em tamanho quanto em confiança. A simples sobreposição entre as predições mantém falsos positivos além de múltiplas detecções a um mesmo objeto, o que impacta negativamente no desempenho do modelo. A união de predições do método proposto baseada na interseção e nos valores de confiança entre as resoluções preserva predições que são importantes para o resultado final ao mesmo tempo que descarta predições isoladas feitas somente por um dos modelos, que são potenciais falsos positivos. Desta forma, o resultado final carrega detecções baseados nas características aprendidas pelos modelos baseados nas resoluções original, 2x e 4x, que são complementares. Um ponto negativo é o descarte de uma predição correta caso apenas um dos modelos tenha predito corretamente, pois o resultado final se trata de um comitê.

Figura 15 – Comparação entre a marcação do especialista (em verde) e a predição final do método após a redução de falsos positivos (em vermelho).



Fonte: Elaborado pelo autor.

A Tabela 9 exibe a melhora nas métricas alcançadas ao usar modelos especializados para cada tipo de matriz ao invés de um modelo para ambos os tipos. Usando a base de dados proposta neste trabalho, quando o modelo é treinado com ambos os tipos de matriz, a mistura de características possivelmente causa redução no aprendizado dos modelos. Dessa forma, os resultados obtidos evidenciam a necessidade da etapa de classificação do tipo de matriz anterior à detecção de fissuras.

Modelo	IoU	Precisão	Recall	F1-Score	AP
Ambas as matrizes	0,3912	0,3272	0,0571	0,0944	0,1636
Apenas Matriz A	0,6353	0,6831	0,5363	0,5701	0,3416
Apenas Matriz B	0,8609	0,8937	0,8326	0,8499	0,4469

Tabela 9 – Comparação dos modelos especializados para cada tipo de matriz com um modelo treinado com ambos os tipos.

Por fim, a Tabela 10 mostra o comparativo entre as métricas alcançadas pelo método proposto e os estudos encontrados na literatura abordados no Capítulo 3.2. Os métodos e métricas fornecidos pelos respectivos autores não são adequados para uma justa comparação com um trabalho voltado para detecção de objetos, uma vez que não disponibilizaram a base de dados e respectivas predições para extração das caixas delimitadoras, a fim de reproduzir os cálculos das métricas adequados, uma vez que se tratam de trabalhos de segmentação semântica e produzem uma saída de formato diferente.

Outro fator que impossibilita uma boa comparação é a disponibilidade de métricas. Enquanto o trabalho de Li et al. (LI et al., 2023) não relatou quaisquer métricas alcançadas pelos modelos, o trabalho de Tian et al. (TIAN et al., 2021) disponibilizou apenas IoU, e o trabalho de Dong et al. (DONG et al., 2020) apenas F1-Score, o que não somente

inviabiliza uma comparação justa entre os métodos entre si, como também com o método proposto.

Embora não supere as métricas dos outros modelos discutidos na literatura, o método proposto compartilha semelhanças com eles. Além de se destacar como o único método explicitamente projetado para concreto reforçado com fibra, é também o único que disponibiliza uma base de dados e uma avaliação adequada para modelos de detecção de objetos, permitindo que métodos propostos posteriormente tenham uma avaliação mais abrangente utilizando a mesma base, o que pode possibilitar uma comparação mais justa.

Método	Tarefa	Base de Dados	Marcações	Métricas
Método proposto	Detecção	Pública	caixas delimitadoras	0,8499 (F1-Score) e 0,8609 (IoU)
Li et al. (LI et al., 2023)	Segmentação	Privada	máscaras	-
Tian et al. (TIAN et al., 2021)	Segmentação	Privada	máscaras	0,942 (IoU)
Dong et al. (DONG et al., 2020)	Segmentação	Privada	máscaras	0,785 (F1-Score)

Tabela 10 – Comparação dos métodos disponíveis na literatura com o método proposto, baseado nas métricas fornecidas pelos autores.

6 Conclusões e Trabalhos Futuros

Este trabalho apresentou o estudo sobre um método para detecção automática de fissuras em microtomografias de concreto com fibras. Como visto, a detecção de fissuras em microescala é um desafio, uma vez que a visualização das fissuras a olho nu é muito difícil e a detecção manual das fissuras por parte de um especialista é um trabalho árduo e demorado para o ser humano realizar.

Até onde abrangeu a revisão bibliográfica, existem uma série de métodos para monitorar a condição do concreto e detectar fissuras, contudo estes métodos possuem diversas limitações como a necessidade de uma extensa rede de sensores, a possibilidade de fatores externos à amostra influenciarem as medições realizadas, bem como as limitações técnicas que muitos métodos têm em relação à morfologia das fissuras encontradas. diversos estudos utilizando modelos de aprendizado profundo para automação da tarefa de detecção de fissuras foram identificados, porém apenas em macroescala, que são muito maiores e mais fáceis de identificar.

Por conta destes fatores, foi proposto a utilização de modelos de aprendizado profundo para detecção automática de fissuras em microescala, utilizando imagens de Micro-CT como dados de entrada. De acordo com a literatura, há poucos estudos voltados para detecção de fissuras utilizando imagens de Micro-CT, que é um método adequado pois não é influenciável por interferências externas e fornece visualização da parte interna da estrutura de concreto. Dentre os métodos abordados, as principais limitações encontradas foi a falta de uma avaliação adequada dos métodos, a indisponibilidade da base de dados e a falta de métodos voltados especificamente para concreto reforçado por fibras.

Portanto, foi proposto um método baseado em aprendizado profundo utilizando a rede EfficientNetB0 para classificação do tipo de matriz, pré-processamento baseado em super-resolução, a rede DETR para detecção de fissuras, além de um pós-processamento baseado em comitê para redução de falsos positivos e aumento na confiabilidade do resultado final. o método proposto não foi capaz de identificar todas as microfissuras presentes na base de dados, alcançando métricas de 57,01% e 84% de F1-Score para imagens pertencentes à Matriz A e B, respectivamente. Em termos de semelhança com a marcação do especialista, o método proposto alcançou métricas para a Matriz A e B, respectivamente, de 63,53% e 86,09% de IoU, evidenciando que uma parte das predições possuem áreas que ou perdem parte da fissura, ou abrangem partes do concreto além da mesma. Contudo os resultados mostram que ele é capaz de reduzir significativamente o esforço empregado pelos especialistas ao analisar o concreto reforçado por fibras. Este resultado, aliado à publicação da base de dados representa uma importante contribuição

no problema da detecção de fissuras em Micro-CT.

Como trabalhos futuros, espera-se ampliar a base de dados com mais amostras, envolvendo outros tipos de matrizes e diferentes formatos, quantidades e disposição de fibras, a fim de produzir modelos mais generalizados. Sugere-se a investigação de outros modelos de detecção, como o as arquiteturas Mamba (LI et al., 2024). Outro fator que não foi abordado nesta pesquisa é a utilização de modelos de aprendizado profundo voltados para aprendizado com menos dados, como abordagens de *few-shot learning*.

6.1 Produções Científicas

Nesta seção são apresentadas as informações sobre a produção científica produzida durante este programa de mestrado, conforme mostram as Tabelas 11 e 12.

Tabela 11 – Produção científica diretamente ligada ao tema de pesquisa

Título	Local de publicação	Ano	Status	Qualis
<i>Automated Crack Detection in Micro-CT Scanning for Fiber-Reinforced Concrete Using Super-Resolution and Deep Learning</i>	<i>Electronics</i>	2025	Aceito	A4

Tabela 12 – Produção científica em demais áreas de pesquisa

Título	Local de publicação	Ano	Status	Qualis
Método de Recuperação de Imagens Baseado em Conteúdo de Tomografia de Rocha utilizando Triplet Loss	SBAI	2025	Publicado	B4
<i>A Deep Metric Learning Approach for Content Based Image Retrieval in Rock Tomography: A Triplet Loss Study</i>	SIBGRAPI	2025	Aceito	A3
<i>Digital Rock Sample Retrieval Using Variational Autoencoder and Clustering</i>	<i>Computational Geosciences</i>	2025	Submetido	A2
<i>Rock Tomography Image Retrieval: A Deep Metric Learning Approach with N-pair Loss</i>	<i>Multimedia Tools and Applications</i>	2025	Submetido	A2

Referências

ABDALLAH, S.; FAN, M. Anchorage mechanisms of novel geometrical hooked-end steel fibres. *Materials and Structures*, v. 50, 2017. Citado na página 19.

ABDUL-RAZZAK, A. A.; Mohammed Ali, A. A. Influence of cracked concrete models on the nonlinear analysis of high strength steel fibre reinforced concrete corbels. *Composite Structures*, v. 93, n. 9, p. 2277–2287, 2011. ISSN 0263-8223. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0263822311001000>>. Citado 2 vezes nas páginas 14 e 19.

AHMED, J. K.; SHUKRI, M. T. Comparison between destructive and non-destructive test on concrete. *Eurasian Journal of Science and Engineering*, v. 3, n. 2, p. 215–223, 2017. Citado 2 vezes nas páginas 14 e 21.

AHMED, T.; SABAB, N. H. N. Classification and understanding of cloud structures via satellite images with efficientunet. *SN Computer Science*, Springer, v. 3, n. 1, p. 99, 2022. Citado na página 26.

CARION, N.; MASSA, F.; SYNNAEVE, G.; USUNIER, N.; KIRILLOV, A.; ZAGORUYKO, S. End-to-end object detection with transformers. In: SPRINGER. *European conference on computer vision*. [S.l.], 2020. p. 213–229. Citado 3 vezes nas páginas 27, 28 e 42.

CONDE, M. V.; CHOI, U.-J.; BURCHI, M.; TIMOFTE, R. Swin2sr: Swinv2 transformer for compressed image super-resolution and restoration. In: SPRINGER. *European Conference on Computer Vision*. [S.l.], 2022. p. 669–687. Citado 2 vezes nas páginas 29 e 41.

CONGRO, M. *Experimental and Numerical Investigation of Damage and Stress Transfer Mechanisms in Cement Materials*. 190 p. Tese (Tese de Doutorado) — Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Civil e Ambiental, 2024. Citado na página 15.

CONGRO, M.; VIEIRA, J. D.; SILVA, F. de A.; ROEHL, D. Integrated workflow for experimental and numerical analysis of damage mechanisms at the fiber/matrix interface in cement composite materials. *Construction and Building Materials*, Elsevier, v. 479, p. 141493, 2025. Citado 4 vezes nas páginas 10, 14, 36 e 37.

CONGRO, M.; VIEIRA, J. D.; SILVA, F. de A.; ROEHL, D. Integrated workflow for experimental and numerical analysis of damage mechanisms at the fiber/matrix interface in cement composite materials. 2025. Citado 3 vezes nas páginas 19, 20 e 23.

COUTO, J. A. S.; CARMINATTI, R. L.; NUNES, R. R. A.; MOURA, R. C. A. et al. O concreto como material de construção. *Caderno de Graduação-Ciências Exatas e Tecnológicas-UNIT-SERGIPE*, v. 1, n. 3, p. 49–58, 2013. Citado na página 14.

DAHLHOFF, A.; RAUPACH, M. Crack analysis of textile reinforced concrete using automated crack evaluation via digital image correlation. *Buildings*, MDPI, v. 13, n. 12, p. 2984, 2023. Citado na página 33.

- DAS, A. K.; LEUNG, C. K. A strategy for in situ determination of self-healing state for strain hardening cementitious composites. *Cement and Concrete Composites*, Elsevier, v. 112, p. 103641, 2020. Citado na página [21](#).
- DAS, A. K.; LEUNG, C. K. Fast tomography: A greedy, heuristic, mesh size-independent methodology for local velocity reconstruction for ae waves in distance decaying environment in semi real-time. *Structural Health Monitoring*, SAGE Publications Sage UK: London, England, v. 21, n. 4, p. 1555–1573, 2022. Citado na página [21](#).
- DINIZ, J. d. C. N.; PAIVA, A. C. d.; JUNIOR, G. B.; ALMEIDA, J. D. S. de; SILVA, A. C.; CUNHA, A. M. T. d. S.; CUNHA, S. C. A. P. d. S. A method for detecting pathologies in concrete structures using deep neural networks. *Applied Sciences*, MDPI, v. 13, n. 9, p. 5763, 2023. Citado na página [33](#).
- DONG, Y.; SU, C.; QIAO, P.; SUN, L. Microstructural crack segmentation of three-dimensional concrete images based on deep convolutional neural networks. *Construction and Building Materials*, Elsevier, v. 253, p. 119185, 2020. Citado 3 vezes nas páginas [35](#), [49](#) e [50](#).
- DOSOVITSKIY, A.; BEYER, L.; KOLESNIKOV, A.; WEISSENBORN, D.; ZHAI, X.; UNTERTHINER, T.; DEGHANI, M.; MINDERER, M.; HEIGOLD, G.; GELLY, S. et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020. Citado na página [27](#).
- GOLDING, V. P.; GHARINEIAT, Z.; MUNAWAR, H. S.; ULLAH, F. Crack detection in concrete structures using deep learning. *Sustainability*, MDPI, v. 14, n. 13, p. 8117, 2022. Citado 2 vezes nas páginas [15](#) e [33](#).
- GUO, P.; MENG, W.; BAO, Y. Intelligent characterization of complex cracks in strain-hardening cementitious composites based on generative computer vision. *Construction and Building Materials*, Elsevier, v. 411, p. 134812, 2024. Citado na página [34](#).
- HASSANI, S.; MOUSAVI, M.; GANDOMI, A. H. Structural health monitoring in composite structures: A comprehensive review. *Sensors*, MDPI, v. 22, n. 1, p. 153, 2021. Citado na página [20](#).
- HE, K.; ZHANG, X.; REN, S.; SUN, J. Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2016. p. 770–778. Citado na página [27](#).
- IMAI, M.; NAKANO, R.; KONO, T.; ICHINOMIYA, T.; MIURA, S.; MURE, M. Crack detection application for fiber reinforced concrete using bocda-based optical fiber strain sensor. *Journal of structural engineering*, American Society of Civil Engineers, v. 136, n. 8, p. 1001–1008, 2010. Citado na página [22](#).
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, v. 25, 2012. Citado na página [25](#).
- LABIB, W. A. Fibre reinforced cement composites. *Cement Based Materials*, InTech London, UK, v. 31, 2018. Citado na página [18](#).

- LANDIS, E. N.; KEANE, D. T. X-ray microtomography. *Materials characterization*, Elsevier, v. 61, n. 12, p. 1305–1316, 2010. Citado na página 23.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *nature*, Nature Publishing Group UK London, v. 521, n. 7553, p. 436–444, 2015. Citado 2 vezes nas páginas 24 e 25.
- LEE, D.; KIM, J.; LEE, D. Robust concrete crack detection using deep learning-based semantic segmentation. *International Journal of Aeronautical and Space Sciences*, Springer, v. 20, p. 287–299, 2019. Citado na página 34.
- LEE, S.-J.; WON, J.-P. Flexural behavior of precast reinforced concrete composite members reinforced with structural nano-synthetic and steel fibers. *Composite Structures*, v. 118, p. 571–579, 2014. ISSN 0263-8223. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0263822314003651>>. Citado 2 vezes nas páginas 14 e 19.
- LEITE, A.; CASTRO, A. de. Influência da matriz cimentícia no comportamento de concretos reforçados com fibras. *Revista IBRACON de Estruturas e Materiais*, SciELO Brasil, v. 13, p. 543–562, 2020. Citado 2 vezes nas páginas 14 e 19.
- LEONHARDT, F. Cracks and crack control in concrete structures. *Pci Journal*, v. 33, n. 4, p. 124–145, 1988. Citado na página 19.
- LI, M.; YUAN, J.; CHEN, S.; ZHANG, L.; ZHU, A.; CHEN, X.; CHEN, T. 3det-mamba: Causal sequence modelling for end-to-end 3d object detection. *Advances in Neural Information Processing Systems*, v. 37, p. 47242–47260, 2024. Citado na página 52.
- LI, N.; ZHAO, Y.; XING, Y.; HE, X.; LI, H. Meso-damage analysis of concrete based on x-ray ct in-situ compression and using deep learning method. *Case Studies in Construction Materials*, Elsevier, v. 18, p. e02118, 2023. Citado 3 vezes nas páginas 34, 49 e 50.
- LIU, Z.; HU, H.; LIN, Y.; YAO, Z.; XIE, Z.; WEI, Y.; NING, J.; CAO, Y.; ZHANG, Z.; DONG, L. et al. Swin transformer v2: Scaling up capacity and resolution. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. [S.l.: s.n.], 2022. p. 12009–12019. Citado na página 29.
- MACHADO, A.; SILVA, M.; FILHO, R.; PFEIL, M.; LIMA, I.; LOPES, R. 3d investigation of steel fiber distribution in reinforced concrete by x-ray microtomography. *Revista Ibracon de Estruturas e Materiais*, SciELO Brasil, v. 8, n. 5, p. 707–720, 2015. Citado 2 vezes nas páginas 15 e 23.
- MAČIULAITIS, R.; VAIČIENE, M.; ŽURAUSKIENE, R. The effect of concrete composition and aggregates properties on performance of concrete. *Journal of civil engineering and management*, Taylor & Francis, v. 15, n. 3, p. 317–324, 2009. Citado na página 19.
- MEHTA, P. K.; MONTEIRO, P. J. *Concrete microstructure, properties, and materials*. [S.l.]: McGraw-hill, 2006. Citado 2 vezes nas páginas 14 e 18.
- MUSA, P.; RAFI, F. A.; LAMSANI, M. A review: Contrast-limited adaptive histogram equalization (clahe) methods to help the application of face recognition. In: *IEEE. 2018 third international conference on informatics and computing (ICIC)*. [S.l.], 2018. p. 1–6. Citado na página 39.

- NAAMAN, A. E.; NAJM, H. Bond-slip mechanisms of steel fibers in concrete. *Aci Materials Journal*, v. 88, p. 135–145, 1991. Disponível em: <<https://api.semanticscholar.org/CorpusID:138860949>>. Citado na página 19.
- NAJM, H. M.; NANAYAKKARA, O.; SABRI, M. M. S. Destructive and non-destructive evaluation of fibre-reinforced concrete: A comprehensive study of mechanical properties. *Materials*, MDPI, v. 15, n. 13, p. 4432, 2022. Citado 2 vezes nas páginas 14 e 21.
- NEVILLE, A. M. et al. *Properties of concrete*. [S.l.]: Longman London, 1995. v. 4. Citado 2 vezes nas páginas 14 e 18.
- PARK, S. E.; EEM, S.-H.; JEON, H. Concrete crack detection and quantification using deep learning and structured light. *Construction and Building Materials*, Elsevier, v. 252, p. 119096, 2020. Citado na página 33.
- REDMON, J.; DIVVALA, S.; GIRSHICK, R.; FARHADI, A. You only look once: Unified, real-time object detection. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2016. p. 779–788. Citado 2 vezes nas páginas 34 e 39.
- SADOWSKI, Ł.; STEFANIUK, D. Microstructural evolution within the interphase between hardening overlay and existing concrete substrates. *Applied Sciences*, MDPI, v. 7, n. 2, p. 123, 2017. Citado na página 23.
- SANDLER, M.; HOWARD, A.; ZHU, M.; ZHMOGINOV, A.; CHEN, L.-C. Mobilenetv2: Inverted residuals and linear bottlenecks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2018. p. 4510–4520. Citado na página 25.
- SCHÖTTL, L.; KOLB, P.; LIEBIG, W. V.; WEIDENMANN, K. A.; INAL, K.; ELSNER, P. Crack characterization of discontinuous fiber-reinforced composites by using micro-computed tomography: Cyclic in-situ testing, crack segmentation and crack volume fraction. *Composites Communications*, Elsevier, v. 21, p. 100384, 2020. Citado na página 22.
- SHINDE, P. P.; SHAH, S. A review of machine learning and deep learning applications. In: IEEE. *2018 Fourth international conference on computing communication control and automation (ICCCUBEA)*. [S.l.], 2018. p. 1–6. Citado na página 24.
- SILVA, M. C. D. da. *Modelagem Numérica do Comportamento Mecânico de Materiais Compósitos Cimentícios em uma Abordagem Multiescala*. Tese (Doutorado) — PUC-Rio, 2020. Citado na página 24.
- SOHAIB, M.; JAMIL, S.; KIM, J.-M. An ensemble approach for robust automated crack detection and segmentation in concrete structures. *Sensors*, MDPI, v. 24, n. 1, p. 257, 2024. Citado na página 33.
- SOUZA, L. M. de; CONCEIÇÃO, W. B. da; MOREIRA, K. C. B. Análise de acidentes e colapsos estruturais: investigação das causas e lições aprendidas. *Revista Novos Desafios*, v. 4, n. 2, p. 201–215, 2024. Citado na página 14.
- TAN, M.; LE, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In: PMLR. *International conference on machine learning*. [S.l.], 2019. p. 6105–6114. Citado na página 25.

- TIAN, W.; CHENG, X.; LIU, Q.; YU, C.; GAO, F.; CHI, Y. Meso-structure segmentation of concrete ct image based on mask and regional convolution neural network. *Materials & Design*, Elsevier, v. 208, p. 109919, 2021. Citado 3 vezes nas páginas 34, 49 e 50.
- UMBACH, C.; MIDDENDORF, B. Investigations on lightweight concrete by in situ compression tests using high-resolution computed tomography (μ -ct). *Journal of Microscopy*, Wiley Online Library, v. 294, n. 2, p. 177–190, 2024. Citado 2 vezes nas páginas 15 e 24.
- VASWANI, A.; SHAZEER, N.; PARMAR, N.; USZKOREIT, J.; JONES, L.; GOMEZ, A. N.; KAISER, Ł.; POLOSUKHIN, I. Attention is all you need. *Advances in neural information processing systems*, v. 30, 2017. Citado na página 27.
- ZHAO, G.; WANG, B.; WANG, T.; HAO, W.; LUO, Y. Detection and monitoring of delamination in composite laminates using ultrasonic guided wave. *composite Structures*, Elsevier, v. 225, p. 111161, 2019. Citado na página 22.
- ZHOU, S.; PAN, Y.; HUANG, X.; YANG, D.; DING, Y.; DUAN, R. Crack texture feature identification of fiber reinforced concrete based on deep learning. *Materials*, MDPI, v. 15, n. 11, p. 3940, 2022. Citado na página 33.