

METODOLOGIA DE DEEP LEARNING E RAY CASTING APLICADA À INSPEÇÃO REMOTA DE PONTES

João Fonseca ⁽¹⁾, Tomás Jorge ⁽¹⁾, Eduardo Soares ⁽¹⁾, Pedro Oliveira ⁽¹⁾, Rafael Cabral ⁽¹⁾⁽²⁾, Bernardino Lima ⁽³⁾, Eduardo Dias ⁽³⁾, Daniel Monteiro ⁽⁴⁾, Ricardo Santos ⁽¹⁾⁽²⁾, Diogo Ribeiro ⁽¹⁾⁽²⁾

(1) iBuilt, ISEP, Politécnico do Porto, Porto, Portugal

(2) CONSTRUCT, FEUP, Universidade do Porto, Porto, Portugal

(3) TECHONIS – Global Consulting Services, Porto, Portugal

(4) Câmara Municipal de Gondomar, Porto, Portugal

Resumo

A crescente necessidade de metodologias de inspeção estrutural mais eficientes e objetivas tem impulsionado o desenvolvimento de abordagens baseadas em aquisição remota e processamento automatizado. Este trabalho apresenta uma metodologia integrada para a deteção e o mapeamento tridimensional de anomalias superficiais em pontes de betão armado, estabelecendo um pipeline completo desde a aquisição de dados até à extração de métricas geométricas. A abordagem combina imagens adquiridas por UAV, reconstrução tridimensional por Fotogrametria e *Laser Scanning*, segmentação por instâncias baseada em *Deep Learning* e projeção geométrica 3D das anomalias. Foi desenvolvido e ajustado um modelo de segmentação YOLO11-seg para a identificação de fissuras, eflorescências e armaduras expostas, que atingiu um desempenho de mAP@50 de 61%. A metodologia foi aplicada a um caso de estudo real, demonstrando a sua capacidade de integrar dados heterogéneos, assegurar coerência geométrica e produzir resultados interpretáveis para suporte a fluxos de inspeção semi-automatizados.

1. Introdução

As pontes de betão armado constituem infraestruturas essenciais, amplamente utilizadas pela sua durabilidade e relação custo-eficiência [1], mas encontram-se particularmente expostas a processos de degradação ao longo do seu ciclo de vida, como fissuração, eflorescências e a exposição e corrosão das armaduras. Estes mecanismos têm vindo a ser acelerados pelas alterações climáticas aumentando a necessidade de inspeções mais frequentes e proativas [2]. No entanto, os métodos tradicionais de inspeção continuam a ser maioritariamente baseados em observação visual, dependente das condições de acesso e equipamentos especializados, sendo frequentemente dispendiosos, demorados, potencialmente perigosos e marcados pela elevada subjetividade na avaliação do estado de conservação [3].

Neste contexto, a transformação digital tem impulsionado a adoção de tecnologias como UAV, reconstrução tridimensional e *Deep Learning*, promovendo o desenvolvimento de metodologias de inspeção remota mais eficientes, seguras e objetivas [4]. O presente trabalho propõe uma metodologia integrada que unifica estas tecnologias, permitindo transformar dados, como imagens, em resultados tridimensionais objetivos, precisos e facilmente interpretáveis.

2. Metodologia de inspeção remota

A metodologia de inspeção remota proposta, ilustrada na Figura 1, está focada em quatro tópicos principais: Aquisição, Reconstrução, Inferência e Mapeamento. Cada uma destas atividades é composta por vários passos intermédios que, quando executados de forma sequencial, permitem gerar modelos tridimensionais e garantem que as respetivas anomalias são devidamente identificadas, localizadas e destacadas na estrutura.

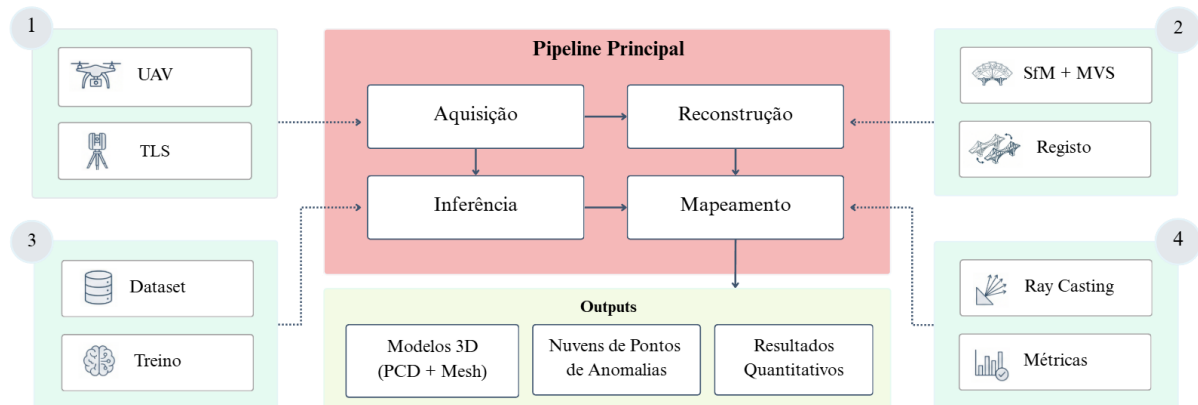


Figura 1: Fluxograma da metodologia de inspeção remota proposta

2.1 Aquisição de dados

A aquisição de dados baseia-se numa estratégia multisensorial, combinando dados visuais (imagens) e geométricos (nuvens de pontos), recolhidos através de câmaras *Red-Green-Blue* (RGB) e sensores *Light Detection and Ranging* (LiDAR). Embora a metodologia possa ser aplicada recorrendo exclusivamente a dados visuais, a adoção desta estratégia permite tirar partido das vantagens específicas de cada tecnologia e mitigar algumas das suas limitações inerentes. As imagens RGB são adquiridas por veículos aéreos não tripulados (*Unmanned Aerial Vehicles* - UAV), permitindo a captura remota em zonas de difícil acesso, enquanto os dados tridimensionais podem ser obtidos por *Terrestrial Laser Scanning* (TLS), assegurando a disponibilidade dos diferentes tipos de dados necessários às etapas subseqüentes de processamento, inferência e mapeamento.

O planeamento da aquisição de imagens por UAV é integrado na metodologia de forma a responder simultaneamente às necessidades de contextualização geométrica da estrutura e à identificação detalhada de anomalias. Para esse efeito, a estratégia de voo combina levantamentos a maior distância, destinados à captura de imagens de contexto geral da estrutura, com levantamentos de proximidade, realizados a menor distância da superfície, com o objetivo de aumentar a resolução espacial e facilitar a identificação de anomalias superficiais. Por sua vez, o posicionamento das estações de TLS assume igualmente um papel relevante, uma vez

que um planeamento adequado da localização de cada estação permite reduzir o tempo de trabalho em campo, garantir a sobreposição necessária entre levantamentos e assegurar a integridade da representação geométrica da estrutura.

2.2 Processamento e reconstrução

Esta etapa tem como objetivo transformar os dados brutos recolhidos na fase de aquisição em representações geométricas coerentes e compatíveis com as fases subsequentes de inferência e de mapeamento tridimensional de anomalias. A Figura 2 apresenta o fluxo de processamento dos dados.

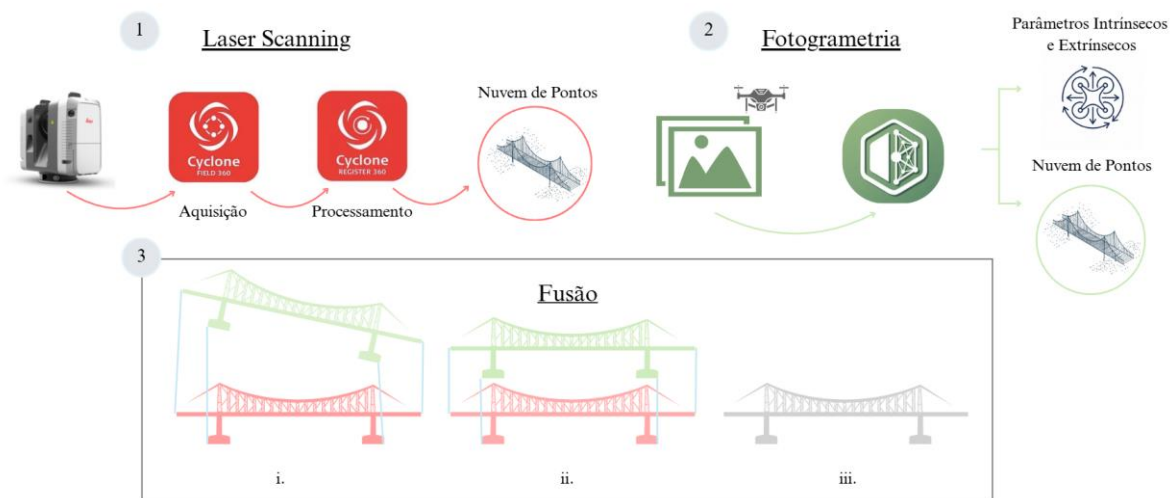


Figura 2: Fluxograma do processamento dos dados

No caso da fotogrametria, o processamento é realizado a partir de imagens RGB, seguindo um fluxo de trabalho típico baseado em técnicas de *Structure from Motion* (SfM) e *Multi-View Stereo* (MVS), que incluem a calibração da câmara, o registo das imagens e a reconstrução e densificação geométrica da cena. Nesta metodologia, recorreu-se ao Bentley iTwin Modeler [5] pela sua capacidade de processamento de conjuntos de dados extensos e pela geração de modelos tridimensionais de elevada definição. O processo conduz à estimação das poses das câmaras e à geração de uma representação 3D da estrutura sob a forma de nuvem de pontos.

A extração dos parâmetros intrínsecos e extrínsecos das câmaras constitui um passo fundamental, uma vez que estes parâmetros são posteriormente utilizados na projeção das anomalias detetadas.

Em relação aos dados de TLS, o processamento incide no registo e validação das nuvens de pontos adquiridas a partir de diferentes estações. Neste trabalho, a aquisição e verificação em campo foi suportada pelo Cyclone Field [6], enquanto o registo e o ajuste global foram efetuados no Cyclone Register [7], permitindo controlar a qualidade das ligações entre estações e obter uma nuvem de pontos consistente.

De forma a assegurar a correspondência espacial entre os diferentes conjuntos de dados e as várias etapas do processamento, é adotada uma estratégia de registo multifase. O processo inicia-se com um alinhamento global, realizado no CloudCompare [8], destinado a corrigir discrepâncias de maior magnitude entre conjuntos de dados, sendo depois refinado através de fases sucessivas de registo mais preciso.

O refinamento geométrico é efetuado com recurso ao algoritmo *Iterative Closest Point* (ICP), aplicado de forma progressiva (3 iterações), permitindo minimizar gradualmente a distância entre correspondências espaciais e reduzir erros residuais. O processo é finalizado com a transformação da nuvem de pontos numa malha triangular texturizada.

2.3 Estratégia de Treino e Inferência para identificação de anomalias

Para a identificação automática de anomalias estruturais foi adotado o modelo *You Only Look Once* 11 (YOLO11-seg), uma arquitetura de deteção e segmentação por instâncias baseada em *Deep Learning*, pré-treinada no dataset *Common Objects in Context* (COCO). O modelo foi ajustado através de um processo de *fine-tuning* para permitir a deteção e segmentação de fissuras, eflorescências e armaduras expostas. A adaptação do modelo a este domínio específico exigiu a construção de um dataset de treino dedicado, composto por imagens provenientes de datasets públicos, como SDNET 2018 [9] e DACL 10k [10], e por dados recolhidos em campanhas fotográficas dedicadas, que incluíram a inspeção de diversas pontes e viadutos localizados na Área Metropolitana do Porto. No total o dataset resultante integra cerca de 7500 imagens, as quais foram compiladas, validadas e, quando necessário, anotadas.

O conjunto de dados foi organizado em subconjuntos de treino, validação e teste, segundo uma divisão 80/10/10. Inicialmente, foi realizado um treino base recorrendo às configurações padrão disponibilizadas pela Ultralytics. Para garantir que o conjunto de validação representava adequadamente a complexidade dos dados, foi aplicada validação cruzada do tipo K-Fold. Posteriormente, a otimização dos hiperparâmetros de treino foi realizada com recurso à biblioteca Optuna [11], sendo aplicada especificamente a uma versão final refinada do dataset que integra imagens adquiridas por UAV. Esta etapa teve como objetivo especializar o modelo para o contexto de inspeção remota, ajustando os parâmetros de treino de forma a melhorar o desempenho e a estabilidade do modelo de forma iterativa (40 *trials*).

A métrica utilizada para avaliar quantitativamente o desempenho do modelo foi uma variante do *mean Average Precision* (mAP), nomeadamente a mAP@50, descrita na expressão (1).

$$mAP_{(IoU=50\%)} = \frac{1}{n} \sum_{i=0}^n AP_i \quad (1)$$

Esta métrica foi selecionada não apenas por ser amplamente adotada na literatura para a avaliação de modelos de deteção e segmentação por instâncias, mas também por ser independente do limiar de confiança, permitindo uma melhor comparabilidade entre modelos e configurações distintas.

Numa fase inicial do processo de *fine-tuning*, as imagens que compunham o conjunto de dados de treino incluíam apenas exemplos de fissuras grosseiras. Durante os levantamentos fotográficos verificou-se uma maior presença de microfissuras. A inclusão destas instâncias no dataset, resultante da aquisição de novas imagens em contexto real, refletiu um aumento da complexidade da classe “fissuras”, traduzindo-se num decréscimo das métricas associadas a essa categoria (Tabela 1). Importa referir que a diminuição das métricas não se refletiu numa degradação significativa da qualidade das inferências, uma vez que o modelo manteve um bom desempenho na identificação das fissuras mais evidentes, sendo o impacto mais relevante na deteção de microfissuras.

Relativamente às restantes classes, as eflorescências, apesar da melhoria das métricas ao longo das iterações, continuam a apresentar maior complexidade, sobretudo devido à elevada

propensão para falsos positivos, frequentemente induzidos por variações de iluminação e padrões superficiais semelhantes.

Por outro lado, a classe de armaduras expostas apresentou um desempenho consistente desde as fases iniciais, registando melhorias progressivas ao longo do processo de otimização, tanto ao nível das métricas como da qualidade das inferências, o que poderá estar relacionado com a maior definição geométrica e contraste desta anomalia.

Tabela 1: Resultados de treino

Nome	Épocas	Geral	Fissuras	Eflorescências	Arm. Expostas
Base	120	51.4%	58.2%	39.4%	56.5%
K-Fold	5 x 120	51.8%	56.8%	39.8%	59.1%
Optuna	(40 x 40) + 120	61.1%	44.9%	62.0%	76.3%

Antes da implementação do modelo de segmentação desenvolvido e da projeção das anomalias no modelo 3D, as imagens provenientes da fase de aquisição são corrigidas quanto à distorção ótica com base nos parâmetros intrínsecos da câmara, assegurando consistência geométrica com a reconstrução fotogramétrica. A inferência é realizada sobre mosaicos (*tiles*) de 640×640 pixels, cujos resultados são posteriormente recombinaados na resolução original da imagem. As máscaras de segmentação seguem para o mapeamento com uma cor atribuída por classe (verde-eflorescências, azul-armaduras expostas, vermelho-fissuras) e podem ainda ser validadas e ajustadas manualmente.

2.4 Mapeamento

A etapa final da metodologia consiste no mapeamento tridimensional (3D) das anomalias detetadas nas imagens, através da projeção das máscaras de segmentação no modelo 3D da estrutura e do cálculo da sua interseção. Este processo permite localizar espacialmente as anomalias diretamente na geometria reconstruída, podendo ser utilizada, para o efeito, a nuvem fotogramétrica ou qualquer modelo no mesmo referencial. A informação resultante é passível de integração em ambientes BIM, levando, por exemplo, à associação de cada anomalia a um elemento estrutural específico.

A projeção é realizada recorrendo a técnicas de *Ray Casting*, nas quais pixéis das máscaras de segmentação são convertidos em raios tridimensionais definidos a partir do centro ótico da câmara e orientados segundo os parâmetros intrínsecos e extrínsecos estimados durante a fotogrametria. Os raios gerados são posteriormente intersetados com a malha tridimensional da estrutura, permitindo determinar as coordenadas espaciais dos pontos correspondentes às anomalias detetadas, resultando numa nuvem de pontos tridimensional que preserva a informação associada a cada deteção, nomeadamente a classe e o respetivo nível de confiança. A partir desta representação tridimensional, são extraídas métricas geométricas relevantes para a caracterização das anomalias, incluindo área, extensão e percentagem de superfície danificada, de acordo com a tipologia de anomalia considerada. Os resultados são disponibilizados sob a forma de uma nuvem de pontos colorida, que se sobrepõe ao modelo tridimensional da estrutura, e de um relatório quantitativo que sintetiza as métricas calculadas.

3. Caso de estudo

3.1 Descrição

O caso de estudo selecionado foi a ponte rodoviária de Santa Helena, uma estrutura em betão armado localizada no concelho de Gondomar, Portugal, ilustrada na Figura 3. A ponte é composta por quatro vãos, um sobre o rio Ferreira e três sobre o correspondente leito de cheia. Do ponto de vista estrutural, é constituída por dois encontros e três apoios intermédios, sendo cada apoio formado por uma viga de coroamento e por dois pilares.

Numa avaliação visual preliminar foram identificadas as três classes de anomalias consideradas neste trabalho, com predominância de armaduras expostas na zona das cornijas do tabuleiro.



Figura 3: Ponte de Santa Helena: vista geral da estrutura (superior) e esquema simplificado da geometria longitudinal (inferior) [12].

3.2 Implementação da metodologia de inspeção remota

A metodologia de inspeção remota foi aplicada à ponte de Santa Helena recorrendo a uma estratégia combinada de aquisição por Fotogrametria e *Laser Scanning*. A aquisição fotogramétrica foi realizada com recurso a dois UAVs, um DJI Mavic 3 Enterprise e um DJI Mini 4 Pro, permitindo assim a recolha de imagens de contexto geral e levantamentos de proximidade, particularmente relevantes nas zonas inferiores do tabuleiro e nos pilares.

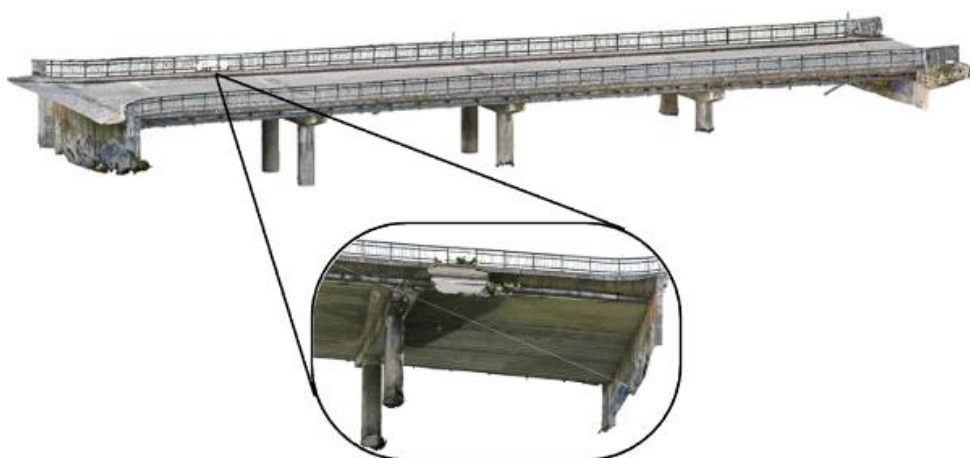
Em complemento, foi realizada aquisição TLS utilizando o equipamento Leica RTC360. Foram adquiridas cerca 1850 imagens e realizadas 47 estações TLS.

A reconstrução 3D conduziu ao desenvolvimento de uma nuvem de pontos densa a partir dos dados fotogramétricos, bem como de uma nuvem de pontos proveniente do TLS registada e georreferenciada.

Os dois conjuntos de dados foram posteriormente fundidos, resultando na geração de uma malha triangular representativa da geometria da estrutura. Durante o processamento fotogramétrico identificaram-se dificuldades localizadas, nomeadamente associadas à perda de sinal GNSS, o que obrigou a uma maior intervenção manual na seleção de keypoints para

garantir o correto registo das imagens. Adicionalmente, registaram-se limitações decorrentes da presença de reflexos na superfície da água sob o tabuleiro, bem como de oclusões parciais da estrutura (Figura 5 a)), resultantes da presença de vegetação.

Antes da execução das inferências, procedeu-se à correção da distorção geométrica das imagens. No caso do DJI Mini 4 Pro, esta correção é aplicada internamente pelo equipamento, sendo, assim, esta etapa apenas aplicável às imagens provenientes do DJI Mavic 3 Enterprise. Após a correção geométrica, as imagens foram processadas pelo modelo de segmentação treinado, utilizando um limiar de confiança de 70%, definido de forma a reduzir a ocorrência de falsos positivos num cenário caracterizado pela complexidade geométrica e operacional, bem como pela presença de elementos não representados no conjunto de dados de treino, como condutas e respetivos apoios metálicos (Figura 5 b)).



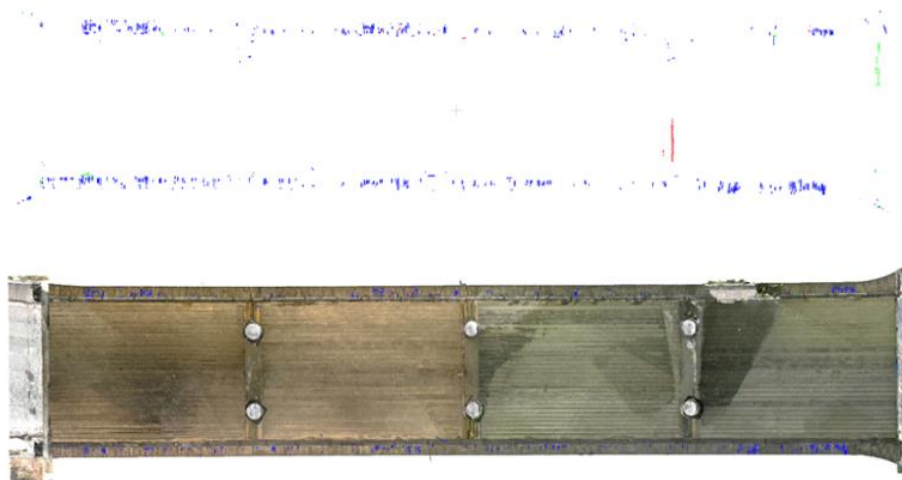
a)



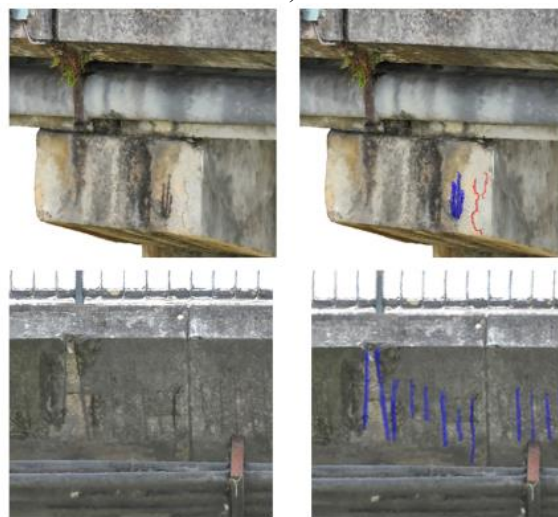
b)

Figura 5: Resultados intermédios: a) malha triangular reconstruída; b) resultado de inferência das anomalias [12].

Durante a fase de mapeamento 3D, foram projetados 12,356,036 raios, correspondentes a 2451 máscaras de anomalias, obtendo-se uma taxa de interseção com a geometria da ponte de 91,1%. Este valor indica uma elevada correspondência entre as deteções em imagem e a superfície da estrutura, sugerindo que o modelo de deteção é capaz de distinguir adequadamente a estrutura de elementos de fundo. O resultado do mapeamento foi uma nuvem de pontos de anomalias apresentada na Figura 6.



a)



b)

Figura 6: Resultados do mapeamento: a) Nuvem de pontos de anomalias e sobreposição no intradorso do tabuleiro; b) Pormenores de projeção [12].

Após a conclusão e validação da projeção, procedeu-se à extração das métricas. Os resultados apresentados na Tabela 2 confirmam que as armaduras expostas constituem o tipo de anomalia mais prevalente, enquanto as fissuras representam a classe menos representada em termos de superfície danificada, verificando-se, contudo, valores elevados de extensão máxima em ambas as classes, associados a uma fissura no pavimento rodoviário e a um erro de projeção. Este tipo de dados permite definir um valor que poderá ser comparado em inspeções futuras determinando assim evolução.

Tabela 2: Resultados quantitativos do mapeamento

Classe	Área (m ²)	Superfície Danificada (%)	Extensão Média (m)	Extensão Máxima (m)
Fissuras	0.219	0.01	0,492	2,785
Eflorescências	0.516	0.03	0.189	1,093
Arm. Expostas	6,16	0,31	0.18	2,684

4. Conclusões

O presente trabalho demonstrou a viabilidade e o potencial da integração de aquisição remota, reconstrução 3D, identificação automática de anomalias e projeção geométrica num único fluxo metodológico aplicado à inspeção de pontes de betão armado. A metodologia proposta combina fotogrametria e *Laser Scanning* como tecnologias complementares para a geração de modelos 3D, aplica técnicas de *Deep Learning* para a deteção de fissuras, eflorescências e armaduras expostas, e utiliza uma abordagem baseada em *Ray Casting* para o mapeamento 3D consistente das anomalias e a extração de métricas geométricas.

Os resultados obtidos evidenciam ganhos claros em termos de eficiência, segurança e objetividade face a abordagens tradicionais, ainda que revelem limitações associadas à variabilidade das condições de aquisição, à presença de oclusões e à heterogeneidade visual de algumas classes de anomalias. Estes fatores confirmam que, embora a automatização total ainda não seja alcançável, a adoção de fluxos semi-automatizados constitui uma solução tecnicamente robusta e alinhada com as exigências atuais de monitorização estrutural e gestão baseada em dados. Embora as métricas obtidas não permitam, isoladamente, definir um estado global de conservação, a sua aplicação sistemática a diferentes estruturas viabiliza análises comparativas e a definição de escalas relativas de deterioração, podendo estas ser utilizadas para atribuir níveis de importância ou relevância às anomalias detetadas com base na sua extensão e características geométricas. Adicionalmente, este tipo de informação quantitativa permite estabelecer valores de referência passíveis de comparação em campanhas de inspeção futuras, possibilitando a avaliação da evolução temporal das anomalias e apoiando estratégias de manutenção baseadas em dados.

Sendo assim, como trabalho futuro, prevê-se a exploração de técnicas emergentes de reconstrução tridimensional, como o *Gaussian Splatting*, bem como de arquiteturas avançadas de *Deep Learning*, nomeadamente *Vision Transformers*, com vista ao reforço da qualidade da representação geométrica e do desempenho na identificação de anomalias. Adicionalmente, pretende-se reforçar a integração da metodologia no fluxo de *Building Information Modeling* (BIM), explorando a incorporação automática das anomalias detetadas em modelos BIM através de esquemas normalizados, como *Industry Foundation Classes* (IFC), permitindo a sua associação a elementos estruturais específicos e a definição de propriedades para apoio à manutenção. Por fim, será relevante avaliar a robustez da metodologia em múltiplos casos de estudo, promovendo uma aplicação mais consistente e escalável em contextos reais.

Agradecimentos

Este trabalho foi desenvolvido no âmbito do projeto AIDETECT – " Identificação automática de anomalias em pontes de betão armado com recurso a inteligência artificial aplicada à visão computacional financiado pela empresa Techonis. Os autores agradecem igualmente o apoio da unidade de investigação UIDB/04708/2020 (DOI: 10.54499/UIDB/04708/2020) e do financiamento programático - UIDP/04708/2020 (DOI: 10.54499/UIDP/04708/2020), atribuídos ao CONSTRUCT - Instituto de I&D em Estruturas e Construções – financiados por fundos nacionais através da FCT/MCTES (PIDDAC).

Referências

- [1] V. Prakash *et al.*, “Structural Health Monitoring of Concrete Bridges Through Artificial Intelligence: A Narrative Review,” *Applied Sciences*, vol. 15, no. 9, 2025, doi: 10.3390/app15094855.
- [2] M. Zucca *et al.*, “Climate Change Impact on Corrosion of Reinforced Concrete Bridges and Their Seismic Performance,” *Applied Sciences*, vol. 14, p. 60, Sep. 2023, doi: 10.3390/app14010060.
- [3] B. G. Pantoja-Rosero, R. Achanta, and K. Beyer, “Damage-augmented digital twins towards the automated inspection of buildings,” *Autom. Constr.*, vol. 150, p. 104842, Jun. 2023, doi: 10.1016/J.AUTCON.2023.104842.
- [4] A. M. Rakoczy *et al.*, “Technologies and Platforms for Remote and Autonomous Bridge Inspection – Review,” *Structural Engineering International*, vol. 35, no. 3, pp. 354–376, Dec. 2025, doi: 10.1080/10168664.2024.2368220.
- [5] Bentley Systems, “iTwin Capture Modeler.” [Online]. Available: <https://www.bentley.com/software/itwin-capture-modeler/>
- [6] Leica Geosystems AG, “Leica Cyclone FIELD 360.” [Online]. Available: <https://leica-geosystems.com/products/laser-scanners/software/leica-cyclone/leica-cyclone-field-360>
- [7] Leica Geosystems AG, “Leica Cyclone REGISTER 360.” [Online]. Available: <https://leica-geosystems.com/products/laser-scanners/software/leica-cyclone/leica-cyclone-register-360>
- [8] CloudCompare Development Team, “CloudCompare – 3D Point Cloud and Mesh Processing Software,” 2024.
- [9] S. Dorafshan, R. J. Thomas, and M. Maguire, “SDNET2018: An annotated image dataset for non-contact concrete crack detection using deep convolutional neural networks,” *Data Brief*, vol. 21, pp. 1664–1668, Dec. 2018, doi: 10.1016/J.DIB.2018.11.015.
- [10] J. Flotzinger, P. J. Rösch, and T. Braml, “dacl10k: Benchmark for Semantic Bridge Damage Segmentation,” Sep. 2023, [Online]. Available: <http://arxiv.org/abs/2309.00460>
- [11] T. Akiba *et al.*, “Optuna: A Next-generation Hyperparameter Optimization Framework,” *CoRR*, vol. abs/1907.10902, 2019, [Online]. Available: <http://arxiv.org/abs/1907.10902>
- [12] J. Fonseca, “Deep Learning-based Damage Identification for the Remote Inspection of Bridges,” ISEP, Politécnico do Porto, Porto, 2025.