

IA PARA QA/QC NA CONSTRUÇÃO: UMA SIAMESE NETWORK COM TRIPLET LOSS

Nawel Ben Haj Salah⁽¹⁾, Manuel Afonso Parente ⁽²⁾, Hélder Manuel da Silva e Sousa⁽³⁾, José Carlos Lino⁽⁴⁾, Bruno Caires⁽⁵⁾

(1) Universidade do Minho, ISISE, ARISE, Departamento de Engenharia Civil, Guimarães, Portugal.

(2) Universidade do Minho, ISISE, ARISE, Departamento de Engenharia Civil, Guimarães, Portugal, 0000-0001-5765-2622

(3) Universidade do Minho, ISISE, ARISE, Departamento de Engenharia Civil, Guimarães, Portugal, 0000-0002-4569-1090

(4) Universidade do Minho, ISISE, ARISE, Departamento de Engenharia Civil, Guimarães, Portugal, 0000-0002-1227-600X

(5) BIMMS, Porto, Portugal

Resumo

Os métodos tradicionais de garantia e controlo da qualidade (QA/QC) baseiam-se principalmente em inspeções manuais ou visuais, que são demoradas, subjetivas e propensas a erro humano. A automatização do QA/QC através da integração de Inteligência Artificial (IA), particularmente durante a fase de execução, contribui para garantir a conformidade, minimizar repetições de trabalhos e melhorar os resultados dos projetos.

Este trabalho desenvolve uma metodologia para integrar IA em QA/QC com especial enfoque em arquiteturas de modelos siameses e aplica essa metodologia num estudo de caso utilizando dados reais de um projeto de construção. O conjunto de dados é constituído por imagens “as-built” e imagens de modelos BIM “as-designed”, recolhidas através de uma plataforma de visão computacional.

O caso de estudo explora a aplicação do modelo acima descrito para comparar a similaridade entre imagens “as-built” e modelos BIM “as-designed” de painéis de parede provenientes de uma obra real. Os resultados demonstram o potencial das arquiteturas baseadas em redes siameses para automatizar a deteção de inconsistências entre o projeto e a execução, proporcionando uma base para a futura integração de métodos de IA na gestão da qualidade em construção.

1. Introdução

A garantia e controlo da qualidade (QA/QC) são processos fundamentais nos projetos de construção, assegurando que os trabalhos cumprem as especificações de projeto, as normas de segurança e os requisitos de desempenho. No entanto, em muitos estaleiros, os processos de

QA/QC ainda dependem principalmente de inspeções manuais. Estas inspeções são demoradas, exigem mão de obra intensiva e estão frequentemente sujeitas a erro humano, especialmente em projetos complexos que envolvem sistemas arquitetónicos, estruturais e MEP [1]. À medida que os projetos se tornam maiores e mais complexos, as limitações da inspeção manual tornam-se mais evidentes.

Apesar destas limitações, a investigação na área da construção digital tem crescido significativamente na Ásia, América do Norte e Europa, refletindo um interesse global em tecnologias como o Building Information Modelling (BIM), a captura da realidade e as plataformas baseadas em dados. Estas ferramentas melhoram a coordenação, a visualização e a monitorização dos processos construtivos, permitindo a comparação entre modelos BIM do estado projetado e as condições reais do estado construído, criando novas oportunidades para os processos de QA/QC [2].

Neste contexto, a Inteligência Artificial (IA), em particular a visão computacional, apresenta um forte potencial para automatizar tarefas de QA/QC. No entanto, a sua aplicação na construção permanece limitada e ainda pouco explorada em ambientes reais, devido à complexidade dos contextos de obra e à disponibilidade reduzida de conjuntos de dados estruturados e de grande dimensão [3]. Quando aplicadas ao QA/QC, a maioria das abordagens existentes baseia-se na classificação de defeitos, o que exige categorias previamente definidas e dados anotados. Contudo, os defeitos em construção são frequentemente específicos de cada projeto e visualmente diversos, dificultando a sua definição antecipada e limitando a deteção de casos não previstos [4].

Nas tarefas de QA/QC, o principal objetivo não é apenas classificar defeitos, mas verificar a correspondência entre o estado construído e a intenção de projeto. Assim, abordagens de IA baseadas em similaridade constituem uma alternativa que permite a comparação direta entre imagens do estaleiro e modelos derivados do BIM. As Siamese Networks são amplamente utilizadas na visão computacional para tarefas como correspondência de imagens [5] e rastreamento de objetos [6]. Estes modelos aprendem a medir a similaridade entre representações de imagens e a determinar se duas entradas correspondem ou não [7]. No entanto, a sua aplicação no QA/QC na construção, especialmente para a comparação direta entre modelos BIM e imagens do estaleiro, permanece ainda pouco explorada.

Este trabalho propõe um enquadramento de QA/QC baseado em Inteligência Artificial, utilizando modelos siameses com triplet loss para comparar imagens do estado construído com imagens derivadas do modelo BIM do estado projetado. Ao focar-se na similaridade em vez de classes de defeitos previamente definidas, o método proposto reduz a necessidade de anotação extensiva e aumenta a robustez face a desvios novos ou inesperados [8].

As principais contribuições deste artigo são:

- (i) a aplicação de técnicas de aprendizagem métrica ao QA/QC baseado em BIM;
- (ii) a utilização de dados reais de um projeto de construção;
- (iii) a avaliação experimental do desempenho do modelo utilizando conjuntos de dados de diferentes dimensões.

2. Metodologia

2.1 Contexto da Investigação e Fonte de Dados

Esta investigação é desenvolvida em colaboração com a BIMMS, uma empresa especializada em construção digital. A empresa utiliza uma plataforma de visão computacional (OpenSpace), que disponibiliza imagens captadas por drones, câmaras 360°, dispositivos móveis e scanners laser. A plataforma permite a inspeção remota de estaleiros de construção e o alinhamento das condições reais com os modelos BIM, através de registo espacial e estimativa da posição da câmara [9]. Este processo permite a comparação sincronizada entre o estado construído e o modelo do estado projetado, como ilustrado na Figura 1.

Apesar destas funcionalidades, o processo de comparação continua a ser maioritariamente manual, demorado e sujeito a erro humano. Assim, este trabalho tem como objetivo automatizar este processo através de técnicas de aprendizagem baseada em similaridade.



Figura 1: Interface da plataforma OpenSpace com comparação entre imagem do estaleiro (as-built) e modelo BIM (as-designed).

2.2 Preparação do Conjunto de Dados

O conjunto de dados é composto por pares de imagens que representam as condições do estado construído e as correspondentes vistas do modelo BIM do estado projetado. As imagens foram extraídas da plataforma OpenSpace, que assegura o alinhamento automático entre as imagens do estaleiro e os modelos BIM através de registo espacial e estimativa da posição da câmara, garantindo pontos de vista semelhantes. Este conjunto inclui aproximadamente 100 pares de imagens de paredes provenientes de um projeto real, abrangendo diferentes níveis de complexidade, desde paredes arquitetónicas simples até paredes que integram sistemas MEP visíveis (e.g., tubagens, condutas e cabos), como ilustrado nas Figuras 2 e 3.

As imagens foram selecionadas em condições adequadas de iluminação e visibilidade, evitando oclusões severas. No entanto, foram mantidas pequenas variações de perspetiva, iluminação e oclusão parcial, típicas do ambiente de obra. Antes do treino, todas as imagens foram redimensionadas e normalizadas. Apesar do alinhamento automático, podem existir pequenas diferenças geométricas e de perspetiva, que contribuem para a robustez do modelo.



Figura 2: Site/BIM de uma parede arquitetónica com porta, janela, radiador e tubagens



Figura 3: Site/BIM de uma parede que integra elementos MEP ao nível do teto (condutas e tubagens)

2.3 Arquitetura dos modelos siameses

Os modelos siameses são arquiteturas utilizadas para tarefas de comparação baseada em similaridade. São compostos por sub-redes idênticas que partilham os mesmos pesos e extraem características visuais das imagens de entrada através de redes neuronais convolucionais (*Convolutional Neural Networks* – CNN).

Neste estudo, foi implementado um modelo siamês com três sub-redes, correspondentes às imagens âncora, positiva e negativa. Cada sub-rede processa uma imagem e gera um vetor de características (*embedding*) num espaço comum. Estes vetores são posteriormente comparados para avaliar a similaridade entre imagens do estado construído e do modelo BIM do estado projetado, como ilustrado na Figura 4.

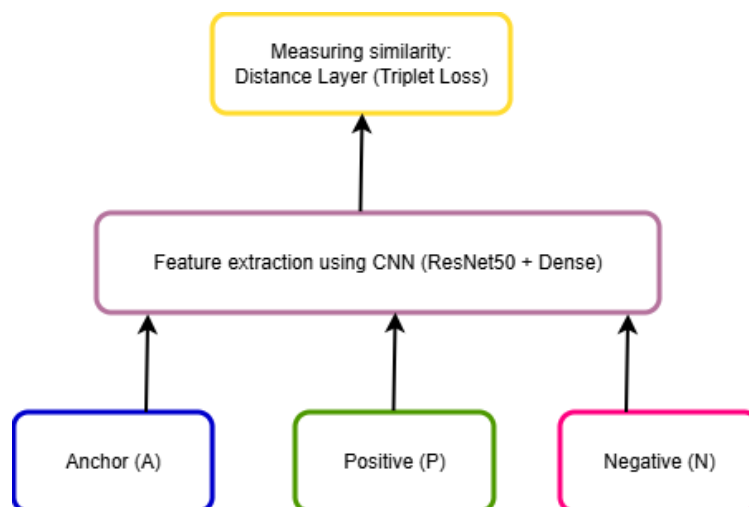


Figura 4: Arquitetura de modelos siameses com *triplet loss*

2.4 Estratégia de Triplet Loss

Os modelos siameses são treinados utilizando uma estratégia de triplet loss, baseada em triplos compostos por uma imagem âncora (imagem do estaleiro), uma imagem positiva (imagem correspondente do modelo BIM) e uma imagem negativa (imagem não correspondente), como ilustrado na Figura 5.

O objetivo da função de *triplet loss* é minimizar a distância entre os vetores de características da imagem âncora e da imagem positiva, e maximizar a distância em relação à imagem negativa. Desta forma, o modelo aprende relações de similaridade e distingue entre pares de imagens correspondentes e não correspondentes.

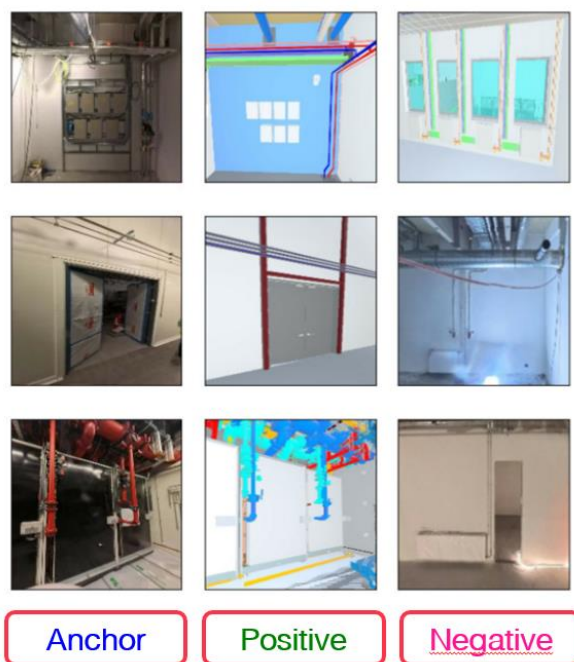


Figura 5: Exemplo de triplo: âncora, positiva e negativa

2.5 Procedimento de Treino

As bibliotecas Keras e TensorFlow foram utilizadas em Python para o cálculo da função de perda e para o treino do modelo. Os modelos siamês são então treinados com estes triplos, de forma a aprender representações de similaridade entre imagens do estado construído e do modelo BIM do estado projetado. O modelo foi treinado durante 10 *epochs*, com um tempo total de aproximadamente 13 minutos, utilizando uma estação de trabalho equipada com um processador Intel Core i7-6700HQ, 16 GB de RAM e uma GPU com 4 GB de memória. Por fim, o desempenho do modelo é avaliado através de métricas de precisão, de modo a verificar a sua capacidade de distinguir corretamente pares de imagens correspondentes e não correspondentes. O processo descrito é sintetizado no fluxo de trabalho apresentado na Figura 6.

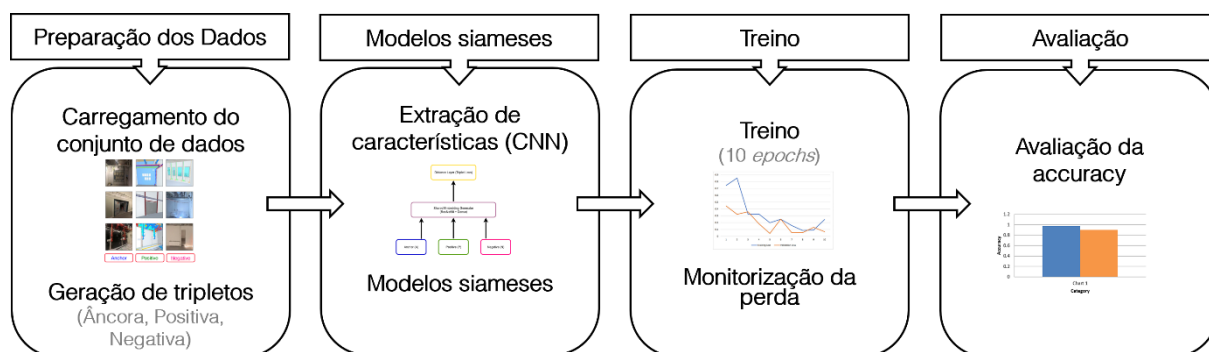


Figura 6: Fluxo de trabalho simplificado do método proposto baseado em modelos siameses

3. Resultados e Discussão

O modelo é inicialmente treinado com um conjunto de dados composto por 50 triplos de imagens. O treino é realizado ao longo de 10 *epochs*, sendo registadas as perdas de treino (*training loss*) e de validação (*validation loss*). Alguns triplos são também visualizados para verificar o correto alinhamento e rotulagem das imagens.

As perdas de treino e de validação são monitorizadas para avaliar o desempenho do processo de aprendizagem e detetar possíveis situações de sobreajustamento (*overfitting*). Após esta fase inicial, o conjunto de dados é expandido para 100 triplos, e o modelo é treinado novamente sob a mesma configuração.

Esta secção apresenta e discute os resultados das perdas de treino e de validação obtidos em ambas as experiências, seguida de uma análise dos resultados de precisão correspondentes.

3.1 Perda de Treino e de Validação do Modelo

3.1.1 Conjunto de Dados com 50 triplos de Imagens

Tabela 1: Perda de treino e de validação do modelo antes do aumento do conjunto de dados

Epoch	Training loss	Validation loss
Epoch 1/10	0.7447	0.4439
Epoch 2/10	0.8525	0.3218
Epoch 3/10	0.3256	0.358
Epoch 4/10	0.325	0.1825
Epoch 5/10	0.198	0.0418
Epoch 6/10	0.2494	0.2508
Epoch 7/10	0.1566	0.0523
Epoch 8/10	0.0849	0.0529
Epoch 9/10	0.0906	0.1311
Epoch 10/10	0.2446	0.0667

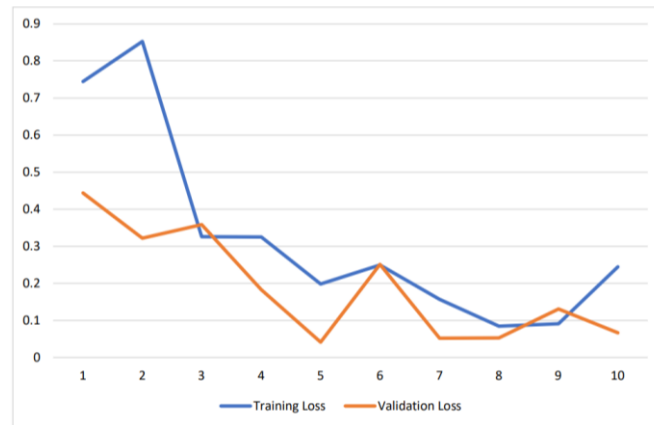


Figura 7: Perda de treino e de validação ao longo de 10 *epochs* (antes do aumento do conjunto de dados)

Os resultados obtidos são apresentados na Tabela 1 e na Figura 7. A perda de treino inicia-se em aproximadamente 0,75, atinge um pico no epoch 2 ($\approx 0,85$) e diminui rapidamente para cerca de 0,3 no epoch 3, indicando alguma instabilidade inicial. Após o epoch 3, a perda de treino estabiliza entre 0,1 e 0,2, com um ligeiro aumento no epoch 10.

Em contraste, a perda de validação inicia-se em aproximadamente 0,45 e diminui progressivamente até ao epoch 5, aproximando-se de zero. No entanto, observa-se um pico no epoch 6, indicando alguma instabilidade no desempenho de validação. Estas oscilações sugerem que o modelo ainda não atingiu uma representação totalmente estável ou generalizável, mantendo-se algum risco de overfitting.

De forma geral, o modelo demonstra uma aprendizagem rápida, mas um desempenho instável na validação, indicando que a reduzida dimensão do conjunto de dados constitui um fator limitante. Para ultrapassar esta limitação, o modelo foi posteriormente treinado com um conjunto de dados expandido para 100 triplos.

3.1.2 Conjunto de Dados com 100 triplos de Imagens

Tabela 2: Perda de treino e de validação do modelo após o aumento do conjunto de dados

Epoch	Training loss	Validation loss
Epoch 1/10	0.6853	0.4264
Epoch 2/10	0.4675	0.2609
Epoch 3/10	0.3274	0.1191
Epoch 4/10	0.1615	0.1135
Epoch 5/10	0.1293	0.1507
Epoch 6/10	0.1162	0.0517
Epoch 7/10	0.1308	0.1346
Epoch 8/10	0.1320	0.2421
Epoch 9/10	0.0309	0.0436
Epoch 10/10	0.0583	0.0297

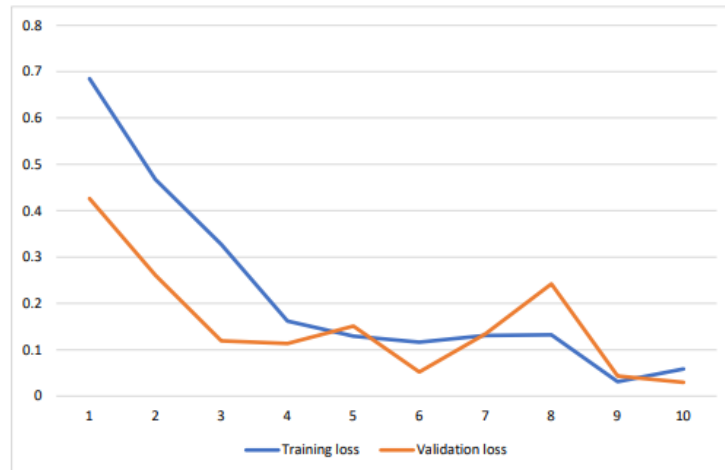


Figura 8: Perda de treino e de validação ao longo de 10 *epochs* (após o aumento do conjunto de dados)

Os resultados obtidos são apresentados na Tabela 2 e na Figura 8. Tanto a perda de treino como a perda de validação diminuem rapidamente nos primeiros epochs, indicando uma aprendizagem rápida na fase inicial. Entre os epochs 1 e 4, ambas as perdas apresentam uma redução significativa, sugerindo que o modelo aprende eficazmente a distinguir pares de imagens.

Observa-se alguma instabilidade na perda de validação após o epoch 4, particularmente com o pico no epoch 8, possivelmente devido à reduzida dimensão do conjunto de validação, à presença de ruído nos dados ou a uma taxa de aprendizagem relativamente elevada. No epoch 10, ambas as perdas atingem valores reduzidos (perda de treino $\approx 0,058$ e perda de validação $\approx 0,030$), indicando um bom ajustamento aos dados, sem sinais claros de *overfitting*.

Em síntese, o modelo apresenta uma convergência rápida nos primeiros epochs, uma flutuação pontual na validação e uma boa capacidade de generalização.

3.2 Avaliação da Acuidade

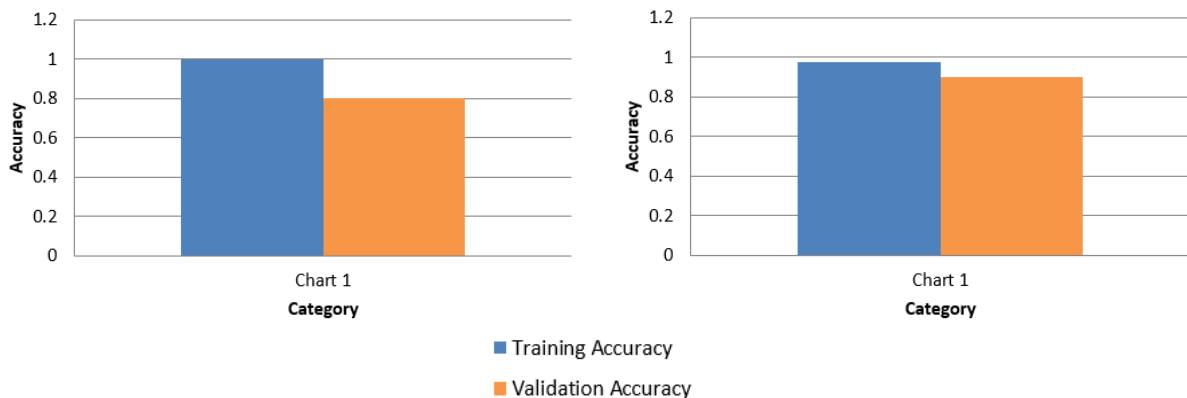


Figura 9 e 10: Acuidade antes e após o aumento do conjunto de dados

Os resultados da acuidade antes e após o aumento do conjunto de dados são apresentados nas Figuras 9 e 10. A acuidade é utilizada para avaliar o desempenho do modelo, medindo a proporção de previsões corretas. É definida como:

$$\text{Acuidade} = \frac{\text{Número de previsões corretas}}{\text{Número total de previsões}}$$

Antes do aumento do conjunto de dados, o modelo apresentou uma acuidade de treino elevada (100%), mas uma acuidade de validação mais baixa (80%), indicando possível *overfitting* e uma capacidade limitada de generalização devido à reduzida dimensão do conjunto de dados. Após o aumento do conjunto de dados, a acuidade de treino manteve-se elevada ($\approx 97,5\%$), enquanto a acuidade de validação ($\approx 90\%$) melhorou consideravelmente. Esta redução da diferença entre os resultados de treino e validação sugere uma melhor capacidade de generalização e um desempenho mais fiável em dados não vistos.

4. Conclusão

Este trabalho demonstra o potencial da utilização de aprendizagem baseada em similaridade para comparar imagens do estado construído com modelos BIM do estado projetado no contexto de QA/QC. Os resultados mostram que a abordagem proposta pode apoiar a automatização destes processos, reduzindo o esforço associado à inspeção visual manual e minimizando erros e atrasos. No entanto, persistem algumas limitações, nomeadamente o tempo necessário para a preparação do conjunto de dados, a dependência da qualidade e do alinhamento das imagens, e a influência significativa da dimensão do conjunto de dados no desempenho e na capacidade de generalização do modelo. Como trabalho futuro, recomenda-se a expansão do conjunto de dados, incluindo cenários mais complexos, como sistemas MEP ao nível do teto, bem como a melhoria da qualidade dos dados. Adicionalmente, a integração do modelo em workflows de QA/QC baseados em BIM e *digital twins* poderá aumentar a sua aplicabilidade prática. Por fim, o desenvolvimento de modelos de IA para deteção e comparação automática de defeitos poderá contribuir para melhorar a eficiência e a fiabilidade na identificação de discrepâncias entre o modelo BIM do estado projetado e as condições reais em obra.

Agradecimentos

Este trabalho foi financiado por PRR – Plano de Recuperação e Resiliência no âmbito do projeto “R2U Technologies | modular systems” com a referência C644876810-00000019. Este trabalho foi também apoiado pela FCT/MCTES no âmbito da Unidade de I&D Instituto para a Sustentabilidade e Inovação em Estruturas de Engenharia (ISISE), com as referências UID/4029/2025 (<https://doi.org/10.54499/UID/04029/2025>), UID/PRR/04029/2025 (<https://doi.org/10.54499/UID/PRR/04029/2025>) e UID/PRR2/04029/2025 (<https://doi.org/10.54499/UID/PRR2/04029/2025>), e no âmbito do Laboratório Associado Produção Avançada e Sistemas Inteligentes ARISE com a referência LA/P/0112/2020 (<https://doi.org/10.54499/LA/P/0112/2020>).

Referências

- [1] X. Tang, Q. Zhou, M. Yu and J. Wu, "Benefits of terrestrial laser scanning for construction QA/QC: A time and cost analysis," *Journal of Management in Engineering*, vol. 38, no. 1, 2022.
- [2] J. O. Toyin and A. Sattineni, "Innovative approaches to QA/QC in construction: Current state," in *Proceedings of the 23rd CIB World Building Congress*, West Lafayette, USA, 2025.
- [3] Y. Xu et al., "Machine learning in construction: From shallow to deep learning," *Developments in the Built Environment*, vol. 6, p. 100045, 2021.
- [4] M. De Filippo et al., "AI-powered inspections of facades in reinforced concrete buildings," *HKIE Transactions*, vol. 30, no. 1, pp. 1–14, 2023.
- [5] I. Melekhov, J. Kannala and E. Rahtu, "Siamese network features for image matching," in *Proceedings of the 23rd International Conference on Pattern Recognition (ICPR)*, 2016, pp. 378–383.
- [6] Y. Zhang et al., "Structured Siamese Network for Real-Time Visual Tracking," in *Computer Vision – ECCV 2018*, pp. 355–370, 2018.
- [7] Y. Li, C. L. P. Chen and T. Zhang, "A Survey on Siamese Network: Methodologies, Applications, and Opportunities," *IEEE Transactions on Artificial Intelligence*, vol. 3, no. 6, pp. 994–1014, 2022.
- [8] E. Hoffer and N. Ailon, "Deep metric learning using triplet network," in *International Conference on Learning Representations (ICLR)*, 2015.
- [9] OpenSpace, "OpenSpace reality capture platform." [Online]. Available: <https://www.openspace.ai>. [Accessed: Feb. 2026].