





Processamento digital de imagens para detecção automática de fissuras em revestimentos cerâmicos de edifícios

Digital image processing for automatic detection of cracks in buildings coatings

Ramiro Daniel Ballesteros Ruiz 
Alberto Casado Lordsleem Junior 
Arthur Flor de Sousa Neto 
Bruno José Torres Fernandes 

Resumo

Na fase de diagnóstico das manifestações patológicas em fachadas, a etapa de inspeção visual merece destacada atenção em virtude da inerente complexidade (altura, tamanho, dificuldades de acesso e condições de exposição). Nos últimos anos, o uso de técnicas de *deep learning* para detectar e classificar características específicas em imagens e vídeos vem crescendo cada vez mais e, quando combinado com o uso de veículos aéreos não tripulados (VANT) para a captura de imagens, constitui uma ferramenta que pode auxiliar e automatizar o procedimento de inspeção visual de fachadas. Este artigo teve o objetivo de realizar a análise do processamento digital de imagens para detecção automática de fissuras em revestimentos cerâmicos de edifícios, associada ao VANT ou drone, o que, potencialmente, resultaria em benefícios (prazo, custo e segurança) no que diz respeito ao diagnóstico. Assim, os resultados da pesquisa exibiram a viabilidade técnica da detecção de fissuras por técnicas de PDI. O procedimento é considerado um trabalho complexo quando há elevada variação nas imagens de estudo. No entanto, mesmo diante de um cenário limitante como a falta de *datasets* públicos para o problema, o projeto conseguiu desenvolver uma metodologia simples e eficiente para o tema para o qual foi proposto.

Palavras-chave: VANT. Drone. Inspeção predial. PDI. Inspeção visual. *Deep learning*.

Abstract

¹Ramiro Daniel Ballesteros Ruiz
¹Universidade de Pernambuco
Recife - PE - Brasil

²Alberto Casado Lordsleem
Junior
²Universidade de Pernambuco
Recife - PE - Brasil

³Arthur Flor de Sousa Neto
³Universidade de Pernambuco
Recife - PE - Brasil

⁴Bruno José Torres Fernandes
⁴Universidade de Pernambuco
Recife - PE - Brasil

In the phase of diagnosis of pathological manifestations in facades, the visual inspection stage deserves special attention due to its inherent complexity (height, size, access difficulties and exposure conditions). In recent years, there has been a growing use of deep learning techniques to detect and classify specific features in images and videos, which, when combined with the use of unmanned aerial vehicles (UAVs) for capturing images, is a potentially useful tool to assist and automate the procedures of visual inspection of facades. This paper aims to perform an analysis of digital image processing for the automatic detection of cracks in ceramic tiles, associated with UAVs or drones, which could potentially bring benefits (time, cost and safety) to the diagnostic process. The research results demonstrated the technical feasibility of detecting cracks through PDI techniques, however, the procedure is a complex process when there is high variation in the study images. At the same time, even in case of a limiting scenario such as the lack of public datasets for the problem, this research project still managed to develop a simple and efficient methodology to deal with the issue proposed.

Recebido em 02/08/19
Aceito em 11/07/20

Keywords: UAV. Drone. Building inspection. PDI. Visual inspection. *Deep learning*.

Introdução

A abordagem de inspeção visual de diferentes elementos por meio de imagens ou fotografias provou ser o recurso mais eficaz e de mais fácil acesso desde o século passado, não sendo diferente para a inspeção de manifestações patológicas em fachadas de edifícios (ADHIKARI; MOSELHI; BAGCHI, 2014).

Contudo, nos últimos anos, o uso de *deep learning* (ou aprendizado profundo – área que busca simular computacionalmente o cérebro como máquina de aprendizado) para detectar e classificar características específicas em imagens e vídeos vem crescendo cada vez mais, trazendo, em alguns casos, resultados mais precisos quando comparados à percepção humana (LIU *et al.*, 2018).

O *deep learning* também possui um desempenho de inspeção mais robusto quando comparado a técnicas tradicionais, tais como detecção de bordas através de regressão (modelo de suavização intensa de imagem, eliminando informações irrelevantes e mantendo, seletivamente, feições de interesse da área selecionada) (JAHANSHAH; MASRI, 2012).

Para a realização da inspeção, é necessário o completo entendimento de uma imagem. Segundo Zhao *et al.* (2019), deve-se não somente classificar várias imagens diferentes, como também estimar com precisão as localizações dos objetos contidos nos cenários.

O processamento digital de imagens, dentro do contexto apresentado, configura uma alternativa promissora para o diagnóstico de manifestações patológicas em edificações, em particular auxiliando na detecção de fissuras em revestimentos cerâmicos de fachadas.

As manifestações patológicas das fachadas sob a forma de fissuras assumem especial importância em função da preocupação quanto ao risco à segurança e à perda de desempenho do edifício, já que respondem também pelos aspectos relativos ao conforto, à higiene e à saúde (CEOTTO; BONDUK; NAKAKURA, 2005; SAHADE, 2005; COSTA E SILVA, 2008; RIBEIRO *et al.*, 2013).

Na fase de diagnóstico das manifestações patológicas em fachadas, a etapa de inspeção visual (vistoria) merece destacada atenção em virtude da inerente complexidade (altura, tamanho, dificuldades de acesso e condições exposição).

Atualmente, no Brasil, a inspeção visual das fachadas de edificações é realizada, em sua maioria, pelo método tradicional, no qual o profissional (alpinista industrial) realiza a verificação visual e o registro fotográfico dos principais pontos a serem analisados (AGUILAR, 2018).

Por outro lado, como resultado do desenvolvimento e aprimoramento de novas ferramentas tecnológicas na indústria da construção civil, percebe-se que o procedimento de inspeção visual e, particularmente, os registros fotográficos podem ser auxiliados e otimizados usando veículo aéreo não tripulado (VANT), comumente conhecido como drone. O drone é utilizado como uma ferramenta para a captura de imagens aéreas (YANG *et al.*, 2018; IRIZARRY; COSTA, 2016; MADER *et al.*, 2016; REAGAN; SABATO; NIEZRECKI, 2017; HUNG *et al.*, 2018).

Associado ao VANT ou drone, o processamento digital de imagens para detecção automática de fissuras em revestimentos cerâmicos de edifícios, foco deste trabalho, resultaria em benefícios (prazo, custo e segurança) no que diz respeito ao diagnóstico, auxiliando no processo de resolução da manifestação patológica.

No entanto, não existe uma base de dados (*dataset*) disponível publicamente de manifestações patológicas na construção civil, como, por exemplo, fissuras em estrutura de concreto ou revestimentos de argamassa e cerâmicos. Isso evidencia a ausência de imagens pré-rotuladas para a realização da detecção e segmentação automatizada das manifestações patológicas (YANG *et al.*, 2018).

Dessa forma, este trabalho teve como objetivo principal realizar estudo de detecção e segmentação de patologias na construção civil através de *deep learning*, com foco em fissuras em cerâmicas.

Considerando a limitação encontrada, foi desenvolvido um sucinto *dataset* preliminar para a realização do estudo de inspeção de fissuras em cerâmicas utilizando *deep learning*, apresentado nas seções seguintes.

Referencial teórico

Com a proposta de detectar e segmentar as fissuras, vários estudos foram realizados tendo como desafio comum a criação de seus próprios *datasets*.

Entre os métodos adotados, Coudray *et al.* (2018) destaca quatro:

- (a) *Threshold*, bastante populares e simples, mas ineficientes;
- (b) *Morphology*, que combina *Threshold* com morfologia matemática, apresentando forte dependência dos parâmetros estáticos iniciais;
- (c) *Multi-Scale*, que, mesmo tendo melhor desempenho que os anteriores, obtém mais falsos negativos; e
- (d) *Neural Network*, que é mais eficiente, mas pode se tornar de elevado custo com a etapa de treinamento.

Seguindo o método de *neural network*, Cha, Choi e Büyüköztürk (2017) apresentaram a arquitetura de modelo *Convolutional Neural Network* (CNN) com foco na detecção de fissuras em concreto, que demonstrou bons resultados a partir de 10 mil imagens no *dataset*, conforme ilustra a Figura 1.

Com uma metodologia um pouco diferente, Silva e Lucena (2018) introduziram a implementação de um modelo de *Machine Learning* para classificar imagens em uma inspeção de infraestrutura. Para isso, foi utilizado um modelo pré-treinado (VGG16) e a técnica de *transferlearning*, objetivando possibilitar a classificação de imagens com e sem fissuras, como ilustra a Figura 2.

O estudo de Yang *et al.* (2018) apresentou o modelo *InspectionNet*, inspirado nos modelos HED (XIE; TU, 2015) e *U-net* (RONNEBERGER; FISCHER; BROX, 2015). Esses três modelos têm como propósito a detecção e a segmentação ao nível de pixel. Desse modo, o estudo focou na detecção de fragmentação do concreto, mostrado na Figura 3.

Diferentemente dos estudos citados anteriormente, o presente trabalho foi realizado com o método de *neural network*, cujo modelo de *deep learning* utilizado foi o *U-net* (RONNEBERGER; FISCHER; BROX, 2015), visto que necessita de poucas imagens para um bom desempenho.

A seguir, detalha-se a metodologia, assim como as etapas do trabalho, como o *dataset* desenvolvido, o pré-processamento das imagens e o modelo escolhido.

Método

Para a construção do projeto, foi necessário separar todo o processo em módulos, como mostra a Figura 4. Iniciando-se com a criação do *dataset* (base de dados), os dados são pré-processados com técnicas de processamento digital de imagens (PDI), e, por fim, a etapa de treinamento e teste da *deep learning*.

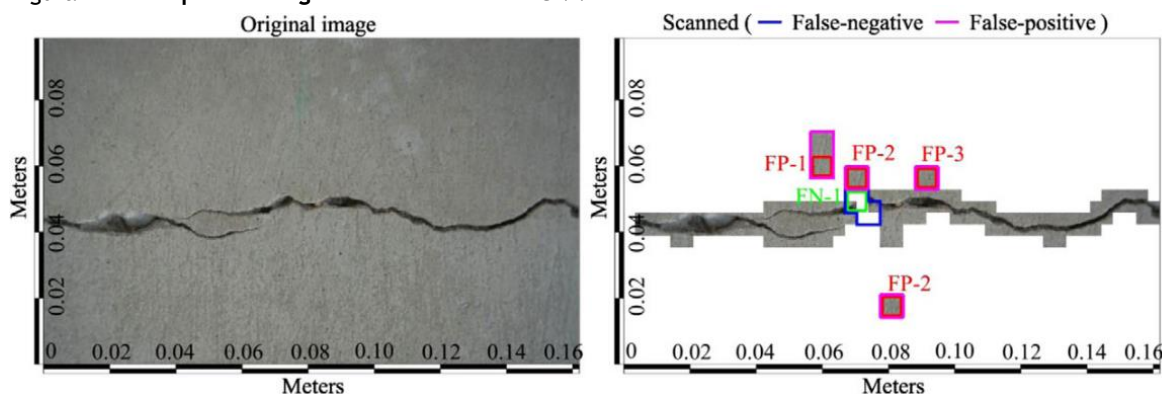
Construção do Dataset

Com a inexistência de um *dataset* público para o problema, foi necessária a criação de um pequeno conjunto de dados, buscando atender aos requisitos mínimos da *deep learning*, mesmo sabendo que para bons resultados seriam necessárias milhares de imagens para o treinamento (aprendizado).

O modelo adotado, *U-net* (RONNEBERGER; FISCHER; BROX, 2015), trouxe grande versatilidade nessa limitação, já que conseguiu lidar bem com pouca massa de dados, quando comparado a outros modelos.

Desse modo, foram coletadas 167 imagens diferentes, que apresentavam cerâmicas com fissuras, e, assim, foram criados rótulos (*labels*) para a *deep learning* interpretar como “respostas corretas”, conforme ilustra a Figura 5.

Figura 1 - Exemplo de imagem testada em uma CNN



Fonte: adaptada de Cha, Choi e Büyüköztürk (2017).

Figura 2 - Exemplo das imagens com e sem fissuras ou trincas que compõem o dataset



Fonte: Silva e Lucena (2018).

Figura 3 - Imagem original, seguida por seu rótulo (label), e a detecção da fragmentação



Fonte: Yang et al. (2018).

Figura 4 - Etapas do processo, separado por módulos

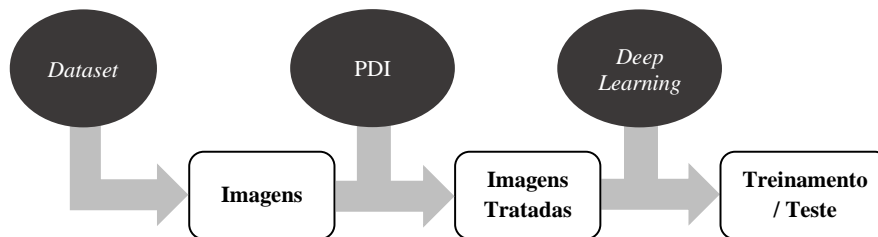


Figura 5 - Exemplo das imagens coletadas e seus rótulos



Do conjunto de dados coletados, 67 imagens foram separadas para a etapa de teste, e 100 para treinamento. Entre as imagens de treinamento, foi realizado o processo de *data augmentation*, o qual aplica técnicas de PDI em cada imagem, variando alguns de seus aspectos, como angulação, *zoom* e deslocamento. Com essa técnica, foi possível criar 10 mil novas imagens e incrementar a base de treinamento, o que resultou em 10.100 imagens.

Pré-processamento

Nesta etapa, as imagens foram tratadas para que o resultado fosse o mais próximo dos rótulos, ou seja, binarizadas, com fundo preto e traços brancos.

Em processamento digital de imagens, uma técnica de binarização bastante utilizada é a *Otsuthresholding* (OTSU, 1979). Antes de realizá-la, foi necessária a aplicação de algumas outras técnicas, envolvendo iluminação, morfologia e filtros de ruído. Essa série de técnicas possibilitou visualizar o objeto de estudo (as fissuras) na grande parte das imagens, que antes apresentavam problemas, conforme ilustra a Figura 6.

No entanto, as imagens resultantes apresentaram grande variedade de cenários, iluminação, angulação e distância. Isso fez com que o pré-processamento não atendesse a todas as imagens, como mostra a Figura 7.

Aplicação do *Deep learning*

Como mencionado, o *U-net* (RONNEBERGER; FISCHER; BROX, 2015) foi o modelo adotado, criado para a área de biomedicina com a finalidade de segmentar imagens de microestrutura de amostras de alta complexidade e resolução, como membranas e vasos sanguíneos.

Figura 6 - Exemplo de imagens somente com *Otsuthresholding* (a) e com a série de técnicas aplicadas (b)

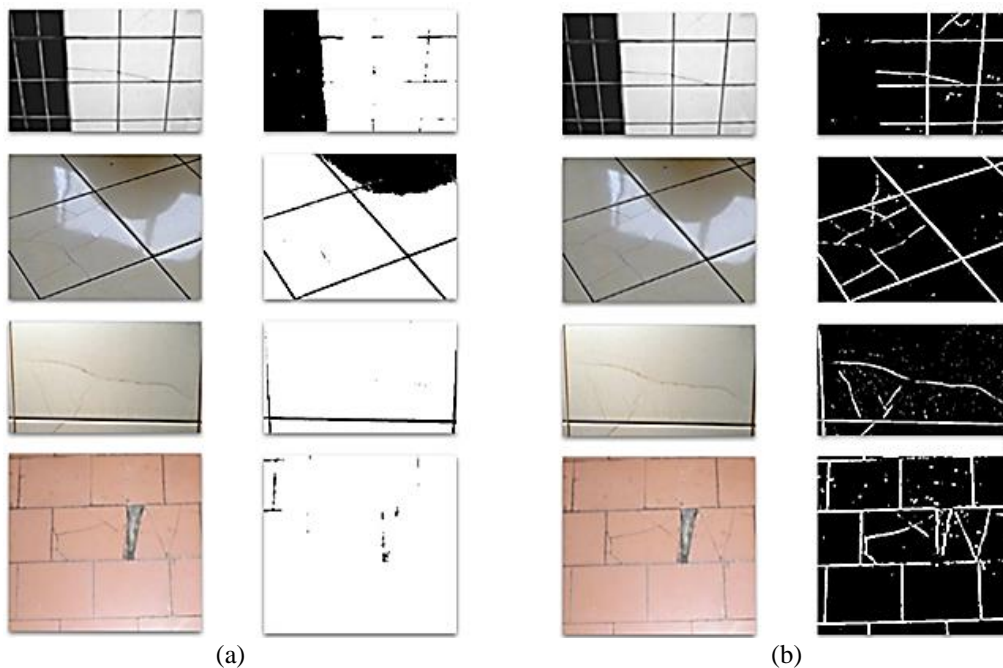
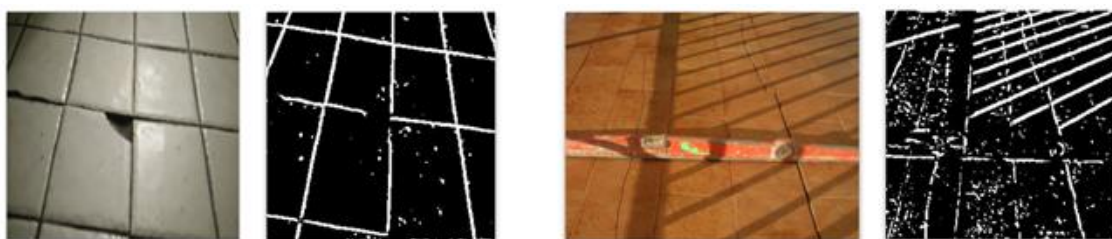


Figura 7 - Exemplo de imagens prejudicadas no pré-processamento



Desse modo, a arquitetura foi planejada para oferecer um bom desempenho mesmo com um pequeno conjunto de imagens (comparado aos outros modelos de *deep learning*).

A arquitetura da *U-net*, mostrada na Figura 8, consiste em duas atividades. A primeira atividade corresponde à contração, aplicando repetidas convoluções, cada uma seguida por uma *Rectified Linear Units* (ReLU – retificador de redes neurais) e uma operação de agrupamento. A segunda atividade corresponde à expansão, através de convoluções e concatenações com recursos de alta resolução (RONNEBERGER; FISCHER; BROX, 2015).

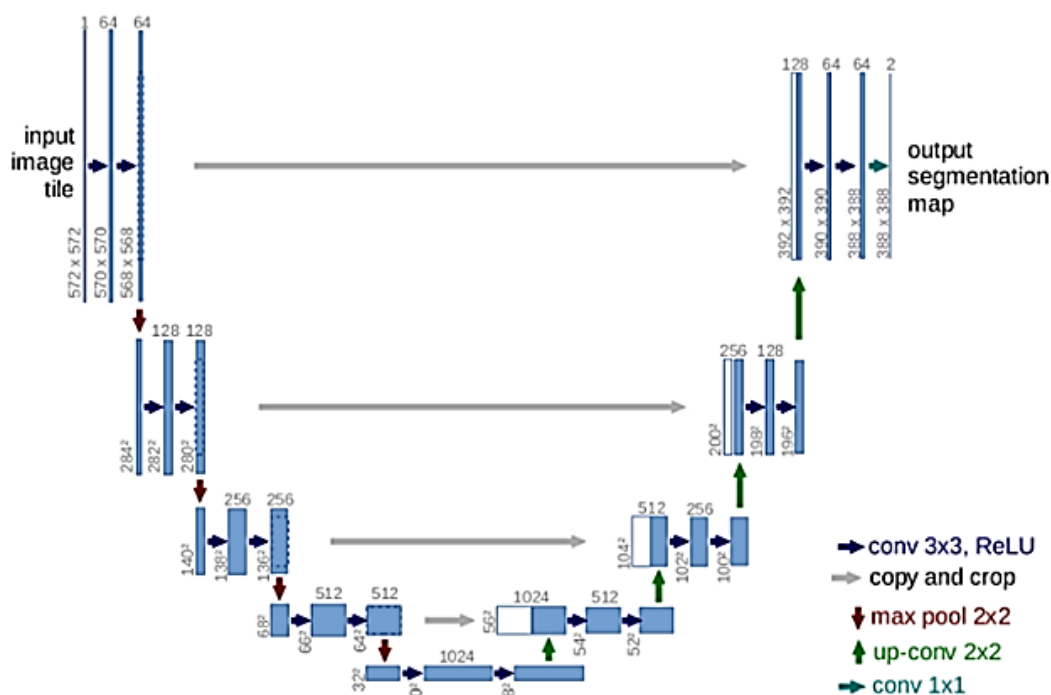
Em outras palavras, o *U-net* é uma arquitetura de rede *convolucional* para segmentação rápida e, para isso, precisa de imagens. A estratégia usada pelo *U-net*, sucintamente, é a classificação do *pixel*, considerando o contexto de toda a imagem. Quanto à convolução, é a operação de filtragem no domínio da intensidade (como a imagem se apresenta na prática, com os valores de cada *pixel*) (RONNEBERGER; FISCHER; BROX, 2015).

Para a realização do treinamento, foi monitorado o aprendizado do modelo, evitando situações de *overfitting* (sobreajuste ou superajuste – quando um modelo estatístico se ajusta muito bem ao conjunto de dados anteriormente observado, mas se mostra ineficaz para prever novos resultados) e *underfitting* (subajuste – quando o modelo não se adapta bem sequer aos dados com os quais foi treinado), além da implementação de *EarlyStopping* (parada antecipada ou parada precoce), com tolerância de até 5 épocas sem uma taxa de aprendizagem significativa.

Esta etapa foi realizada sob o sistema operacional Arch Linux, com processador Intel i7-7500U (4) 3.500 GHz, memória de 16 GB e uma unidade de processamento gráfico GPU (*graphics processing unit*) Nvidia GeForce 940M de 4 GB. Realizada com 10.100 imagens, a etapa consumiu cerca de 10 h para obter resultados mais precisos nas 67 imagens de teste.

De certo modo, o resultado do modelo tende a melhorar à medida que o volume de dados é aumentado, possibilitando passar mais tempo na etapa de treinamento e sofrer menos problemas de *overfitting*.

Figura 8 - Arquitetura da U-net



Fonte: adaptada de Ronneberger, Fischer e Brox (2015).

Resultados

O resultado do estudo proporcionou, mesmo sob as limitações do *dataset* e com imagens prejudicadas no pré-processamento, uma boa análise do cenário, alcançando 71,64% de acurácia nas segmentações das imagens, conforme ilustram a Figura 9 e o Quadro 1. Tanto com imagens com bastante ruído, quanto com fissuras discretas, a inspeção através de *deep learning* demonstrou ser adaptável a situações complexas e flexível aos obstáculos.

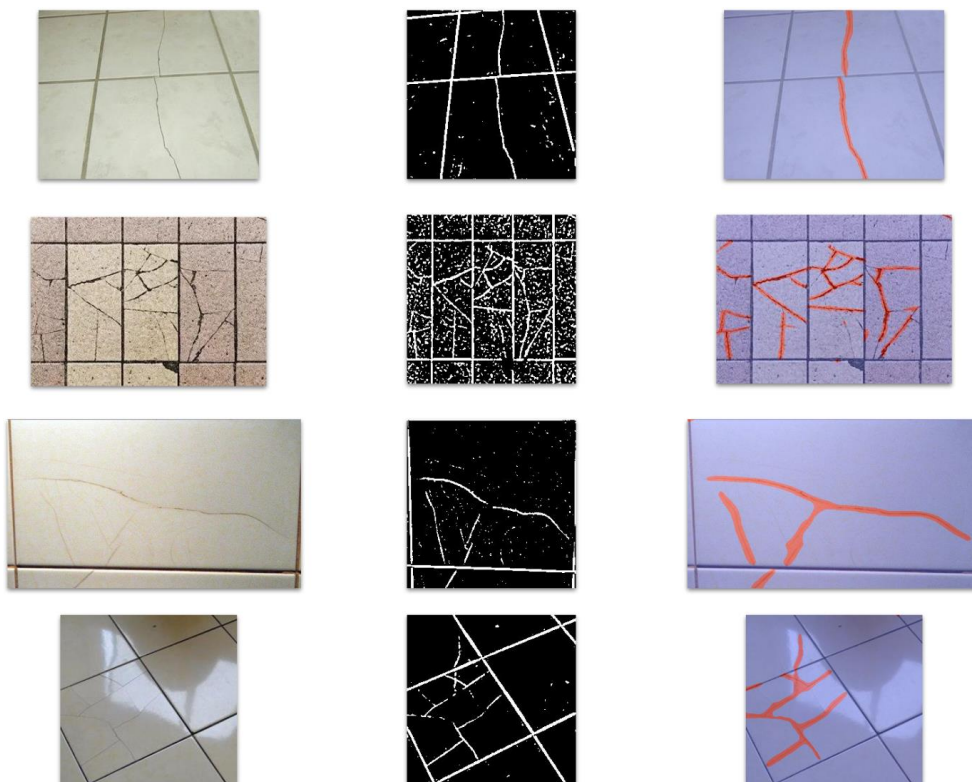
Entretanto, a inexistência de um *dataset* público disponível para esse tipo de problema influenciou diretamente no desempenho da *deep learning*, já que foi necessário buscar imagens que representassem o objeto de estudo a ser segmentado (fissuras em cerâmicas).

É importante ressaltar a importância de um amplo *dataset* para se trabalhar com *deep learning*, com imagens de boa qualidade e rotuladas, para, conseqüentemente, se obterem melhores resultados.

Mesmo após selecionar os dados mais relevantes para compor o *dataset*, ainda assim, em muitos casos, houve grande variação nas imagens, como angulação, distância e qualidade.

Além disso, o pré-processamento dos dados também impactou em alguns dos resultados, já que não atendeu a todas as imagens do *dataset*, principalmente pela variação da iluminação.

Figura 9 - Exemplo das imagens originais, pré-processadas, e do resultado da segmentação sobreposta na imagem original



Quadro 1 - Acurácia nas segmentações das imagens

Descrição	Número de imagens	Porcentagem
Imagens processadas no teste	67	100,00
Imagens segmentadas	59	88,06
Segmentação imprecisa	11	16,42
Acurácia	48	71,64
Taxa de acerto	71,64%	

Conclusão

O presente estudo teve como proposta a utilização de *deep learning* no processo de inspeção de patologias na construção civil, especificamente na detecção e na segmentação de fissuras em cerâmicas.

A detecção das fissuras por PDI é viável do ponto de vista técnico, no entanto é considerado um trabalho complexo quando há elevada variação nas imagens de estudo. Diante de um cenário limitante, tal como a falta de *datasets* públicos para o problema, o projeto, ainda assim, conseguiu desenvolver uma metodologia simples e eficiente para o que foi proposto.

Além disso, o projeto proposto se mostrou mais eficiente a problemas mais comuns e com *dataset* já bem definido, chegando a 95% de acurácia na segmentação de rachaduras em concreto, por exemplo.

Por fim, há um caminho promissor para o aperfeiçoamento do projeto, buscando simplificar mais a abordagem (minimizando a etapa de pré-processamento, por exemplo) e melhorar a precisão dos resultados.

Percebeu-se também a ausência de *datasets* públicos referentes a patologias na construção civil, de modo geral, abrindo espaço para novos estudos que avaliem e reforcem a criação desses conjuntos de dados conforme a necessidade de inspeção automatizada para cada problema da área.

Referências

- ADHIKARI, R.; MOSELHI, O.; BAGCHI, A. Image-based retrieval of concrete crack properties for bridge inspection. **Automation in Construction**, v. 39, p. 180-194, 2014.
- AGUILAR, G. **Inspeção Predial de Fachadas com Drones à luz da ABNT NBR 5674:2012**: manutenção de edificações. 2018. Disponível em: <https://www.linkedin.com/pulse/inspe%C3%A7%C3%A3o-predial-de-fachadas-com-drones-%C3%A0-luz-da-abnt-gabriel-aguiar/?published=t>. Acesso em: 19 maio 2018.
- CEOTTO, L. H.; BONDUK, R. C.; NAKAKURA, E. H. **Revestimentos de argamassa**: boas práticas em projeto, execução e avaliação. Porto Alegre: ANTAC, 2005. (Recomendações Técnicas Habitar, v. 1).
- CHA, Y.-J.; CHOI, W.; BÜYÜKÖZTÜRK, O. Deep learning based crack damage detection using convolutional neural networks. **Computer Aided Civil and Infrastructure Engineering**, v. 32, p. 361-378, 2017.
- COSTA E SILVA, A. J. **Descolamento nos revestimentos cerâmicos de fachada na cidade do Recife**. São Paulo, 2008. 221 f. Dissertação (Mestrado em Engenharia Civil) – Escola Politécnica, Universidade de São Paulo, São Paulo, 2008.
- COUDRAY, N. *et al.* Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. **Nature Medicine**, v. 24, p. 1559–1567, 2018.
- HUNG, M. N. W. B. *et al.* Levantamento com veículo aéreo não tripulado para geração de modelo digital do terreno em bacia experimental com vegetação florestal esparsa. **Revista Ra'e Ga Espaço Geográfico em Análise**, v. 43, p. 215-231, 2018.
- IRIZARRY, J.; COSTA, D. B. Exploratory study of potential applications of unmanned aerial systems for construction management tasks. **Journal of Management in Engineering**, v. 32, n. 3, p. 05016001, 2016.
- JAHANSHAH, M.; MASRI, S. Adaptive vision-based crack detection using 3D scene reconstruction for condition assessment of structures. **Automation in Construction**, v. 22, p. 567-576, 2012.
- LIU, Y. *et al.* Artificial intelligence-based breast cancer nodal metastasis detection. **Archives of Pathology & Laboratory Medicine**, v. 7, p. 859-868, 2018.
- MADER, D. *et al.* Potential of UAV-Based laser scanner and multispectral camera data in building inspection. In: THE INTERNATIONAL ARCHIVES OF THE PHOTOGRAMMETRY, REMOTE SENSING AND SPATIAL INFORMATION SCIENCES, 41., Prague, 2016. **Anais [...]** Prague, 2016.
- OTSU, N. A threshold selection method from gray-level histograms. **IEEE Transactions on Systems, Man and Cybernetics**, v. 9, p. 62-66, 1979.
- REAGAN, D.; SABATO, A.; NIEZRECKI, C. Feasibility of using digital image correlation for unmanned aerial vehicle structural health monitoring of bridges structural health monitoring. **SAGE Publishing**, v. 17, n. 3, p. 289-302, 2017.

RIBEIRO, D. V. *et al.* **Corrosão em estruturas de concreto armado**: teoria, controle e métodos de análise. Rio de Janeiro: Elsevier Brasil, 2013.

RONNEBERGER, O.; FISCHER, P.; BROX, T. U-Net: convolutional networks for biomedical image segmentation. **Medical Image Computing and Computer-Assisted Intervention**, v. 9351, p. 234-241, 2015.

SAHADE, R. F. **Avaliação de sistemas de recuperação de fissuras em alvenaria de vedação**. São Paulo, 2005. 169 f. Dissertação (Mestrado em Habitação) – Instituto de Pesquisas Tecnológicas do Estado de São Paulo, São Paulo, 2005.

SILVA, W. R. L.; LUCENA, D. S. Concrete cracks detection based on deep learning image classification. In: INTERNATIONAL CONFERENCE OF EXPERIMENTAL MECHANICS, 18., Brussels, 2018. **Proceedings [...]** Brussels, 2018.

XIE, S.; TU, Z. Holistically-nested edge detection. **Computer Vision Foundation**, v. 1, p. 1395-1403, 2015.

YANG, L. *et al.* Semantic metric 3D reconstruction for concrete inspection. In: CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION WORKSHOPS, Salt Lake City, 2018. **Proceedings [...]** Salt Lake City, 2018.

ZHAO, Z.-Q. *et al.* Object detection with deep learning: a review. **IEEE Transactions on Neural Networks and Learning Systems**, v. 1, p. 1-21, 2019.

Ramiro Daniel Ballesteros Ruiz

Departamento Engenharia Civil, Escola Politécnica | Universidade de Pernambuco | Rua Benfica, 455, Madalena | Recife - PE - Brasil | CEP 50720-001 | Tel.: (11) 95913-0030 | E-mail: rdb@poli.br

Alberto Casado Lordsleem Junior

Departamento Engenharia Civil, Escola Politécnica | Universidade de Pernambuco | Tel.: (81) 3184-7566 | E-mail: acasado@poli.br

Arthur Flor de Sousa Neto

Departamento Engenharia Civil, Escola Politécnica | Universidade de Pernambuco | Tel.: (81) 3184-7548 | E-mail: afsn@ecomp.poli.br

Bruno José Torres Fernandes

Departamento Engenharia Civil, Escola Politécnica | Universidade de Pernambuco | E-mail: bjtf@ecomp.poli.br

Ambiente Construído

Revista da Associação Nacional de Tecnologia do Ambiente Construído

Av. Osvaldo Aranha, 99 - 3º andar, Centro

Porto Alegre - RS - Brasil

CEP 90035-190

Telefone: +55 (51) 3308-4084

Fax: +55 (51) 3308-4054

www.seer.ufrgs.br/ambienteconstruido

E-mail: ambienteconstruido@ufrgs.br



This is an open-access article distributed under the terms of the Creative Commons Attribution License.