

Detecção de Rachaduras em Concreto em Imagens com o uso de Redes Neurais Convolucionais

Luís Clício Carvalho Sá¹, Vinícius de Sousa Carvalho¹,
José Wanderlei Francisco de Sousa Rocha¹, Romuere Rodrigues Veloso e Silva¹

¹Curso de Sistemas de Informação, Universidade Federal do Piauí (UFPI),
Campus Senador Helvídio Nunes de Barros, Picos, PI

{luisclicio,viniciusscarvalho,jose.rocha,romuere}@ufpi.edu.br

Abstract. *Cracks in concrete infrastructure can cause serious structural damage and pose a threat to public safety, especially in areas close to these structures. Inspecting these structures is mostly difficult and expensive, in addition to being subject to the subjectivity of the inspector's assessment. A possible solution would be to use drones that employ computer vision techniques, aiming for more efficient and safer inspections. However, it is necessary to take into account the implementation cost, largely influenced by the resources available in drones. Therefore, it is appropriate to apply convolutional neural networks with architectures focused on low computational consumption to assist in detecting cracks, in order to balance cost and efficiency. In this sense, the present work aims to validate the use of the MobileNetV2 architecture for the current problem of identifying cracks in concrete structures.*

Resumo. *Rachaduras em infraestruturas de concreto podem causar danos estruturais graves e representam uma ameaça para a segurança pública, especialmente nas áreas próximas a essas estruturas. A inspeção dessas estruturas em sua maioria é difícil e dispendiosa, além de estarem sujeitas a subjetividade da avaliação do inspetor. Uma possível solução seria utilizar drones que empreguem técnicas de visão computacional, objetivando inspeções mais eficientes e seguras. No entanto, é necessário levar em consideração o custo de implementação, influenciado grandemente pelos recursos disponíveis nos drones. Desta forma, é adequado aplicar redes neurais convolucionais com arquiteturas focadas em baixo consumo computacional para auxiliar na detecção de rachaduras, de modo a balancear custo e eficiência. Nesse sentido, o presente trabalho tem o intuito de validar o uso da arquitetura MobileNetV2 para o atual problema de identificação de rachaduras em estruturas de concreto.*

1. Introdução

A identificação de rachaduras em estruturas de concreto é um problema significativo para a área de engenharia civil, pois as rachaduras podem afetar a integridade estrutural de edifícios, pontes e outras estruturas de concreto, comprometendo a região e a população que reside nas proximidades. A descoberta prematura destas pode auxiliar a elevar a longevidade de uma construção.

Entretanto, a avaliação dessas estruturas acaba por exigir altos custos financeiros e de tempo, além de colocarem em risco a segurança dos inspetores, que muitas vezes devem ir a locais de difícil acesso. Por conta disso, procedimentos auxiliados por visão computacional tem sido amplamente utilizados para automatizar a detecção e classificação de rachaduras em estruturas de concreto, uma vez que permitem o rápido levantamento e análise de informações sobre elas.

Independentemente da abordagem utilizada, a classificação assertiva de rachaduras em estruturas de concreto é crucial para garantir a segurança, estabilidade e a durabilidade dessas estruturas. Além disso, a automatização da classificação de rachaduras através da visão computacional permite a realização de inspeções mais rápidas, com menor custo, o que permite uma tomada de decisão mais rápida em caso de danos infraestruturais. Com isso, uma possível alternativa é o uso de *drones* equipados com técnicas de visão computacional. No entanto, para manter o custo baixo, os *drones* utilizados apresentariam recursos limitados, além de exigirem eficácia em sua utilização para maximizar o tempo de voo. Por conta disso, os procedimentos devem ser cuidadosamente avaliados e soluções de baixo consumo devem ser priorizadas.

Assim, para equilibrar custo e eficiência, este trabalho propõe a utilização de redes neurais convolucionais (CNNs) com arquiteturas de baixo custo computacional, em particular a arquitetura MobileNetV2 [Sandler et al. 2018], para o desenvolvimento de um sistema de detecção de rachaduras em concretos. Por fim, também é esperado apresentar de modo claro os procedimentos e resultados obtidos a partir dos experimentos realizados.

2. Trabalhos relacionados

O trabalho proposto por [Flah et al. 2020], descreve um modelo de inspeção automatizado baseado em processamento de imagem e aprendizado profundo para detectar defeitos em áreas normalmente inacessíveis de estruturas de concreto. As imagens foram selecionadas manualmente, sendo redimensionadas para uma resolução de 227 x 227 *pixels*. O classificador foi treinado usando 10.000 imagens, 5.000 para imagens com rachaduras e 5000 para imagens sem rachaduras, aplicando uma divisão de 60:20:20 para treinamento, validação e teste. A precisão obtida no teste foi de 98,25%, com uma perda de 0,057, demonstrando que o classificador não foi influenciado por problemas de *overfitting*. Em resumo, o modelo proposto pode ser uma ferramenta eficaz para inspeção automatizada de estruturas de concreto.

O estudo conduzido por [Arafin et al. 2022] focou na avaliação comparativa do desempenho de modelos para a detecção de rachaduras e descamação em superfícies de concreto. Nesse contexto, foram empregadas redes neurais convolucionais (CNNs) pré-treinadas, incluindo VGG-19, ResNet-50, InceptionV3, Xception e MobileNetV2, para a classificação das rachaduras. O conjunto de dados utilizado abrangeu um total de 4087 imagens de rachaduras e 1100 imagens de descamação, todas com resolução de 224 x 224 *pixels*, divididas de maneira aleatória na proporção de 70% para treinamento, 20% para validação e 10% para teste. Com base nos resultados obtidos, a arquitetura InceptionV3 superou os outros modelos, obtendo acurácia, precisão e recall de 91%, 82% e 100%, respectivamente.

[Kim et al. 2020] apresenta uma técnica de análise de imagens usando *deep learning* desenvolvida para detectar rachaduras e analisar características (por exemplo, comprimento e largura) em imagens para instalações de pequena escala. o trabalho foi dividido em 3 estágios, onde o primeiro e segundo compreende detecção de imagem (por exemplo, classificação e segmentação). Para a detecção, foram usadas 40.000 imagens, divididas em duas classes, cada uma com 20.000 imagens, sendo elas, com e sem rachaduras. Após a detecção baseada em aprendizado profundo, no terceiro estágio, algoritmos de desbaste e rastreamento são aplicados para analisar o comprimento e a largura da trinca na imagem. Os resultados obtidos apresentaram 99% de acurácia na detecção das rachaduras. Embora o estudo tenha demonstrado grande potencial na detecção e medição de rachaduras de concreto, o método proposto deve ser posteriormente verificado com um conjunto de dados maior, a fim de ser utilizado no campo prático.

[Dung and Anh 2019] propõe a utilização de uma rede totalmente convolucional profunda (FCN) para a detecção de fissuras em imagens de concreto. O modelo foi treinado em um conjunto de dados público com 40.000 imagens de fissuras de concreto de 227x227 pixels, divididos em em uma proporção de 80:10:10 para, respectivamente, treinamento, validação e teste. Para o experimento, foram utilizadas as arquiteturas de rede pré-treinadas VGG16, ResNet e InceptionV3. Os resultados indicaram que o modelo baseado na VGG16 obteve melhores resultados, com precisão obtida no teste de 89,3% e F1 de 89,3%.

A Tabela 1 apresenta um resumo das principais características dos trabalhos relacionados.

Tabela 1. Resumo das características dos trabalhos relacionados.

Trabalho	Arquitetura utilizada	Qtd. de imagens	Acurácia
[Flah et al. 2020]	Experimental	10.000	98.25%
[Arafin et al. 2022]	InceptionV3	5187	91%
[Kim et al. 2020]	Experimental	40.000	99%
[Dung and Anh 2019]	VGG16	40.000	-

3. Metodologia

Esta seção apresenta a metodologia empregada no trabalho, bem como conceitos importantes para o entendimento dos processos e resultados. Nesse sentido, a abordagem adotada para o experimento consiste em 7 etapas: aquisição e seleção da base de imagens, seleção do modelo pré-treinado, pré-processamento das imagens, transferência de aprendizagem, treinamento do modelo e obtenção das métricas de desempenho. A figura 1 apresenta as etapas realizadas neste trabalhos, como também sua sequência de execução.

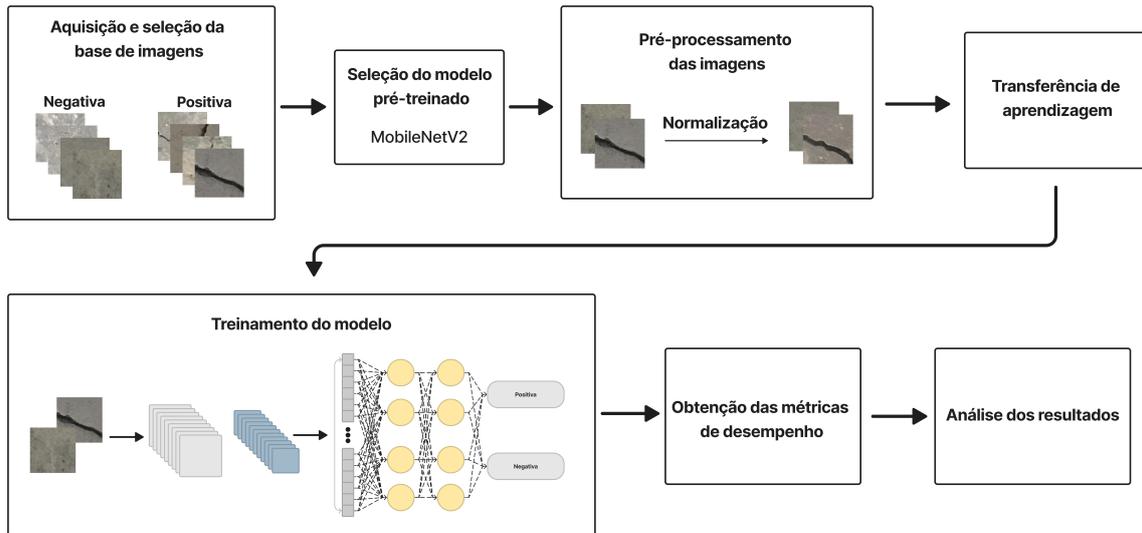


Figura 1. Procedimentos adotados neste trabalho.

3.1. Aquisição e seleção da base de imagens

Para a realização do experimento, foi utilizada uma base com imagens de concreto, dividida em imagens de concreto sem rachaduras (*Negativa*) e com rachaduras (*Positiva*) [Özgenel 2018]. Todas as imagens da base selecionada possuem as dimensões de 227×227 pixels no sistema de cores *RGB*. O conjunto completo é composto por um total de 40 mil imagens de concretos, sendo 20 mil imagens para cada classe. A Figura 2 apresenta exemplos de cada uma das classes.

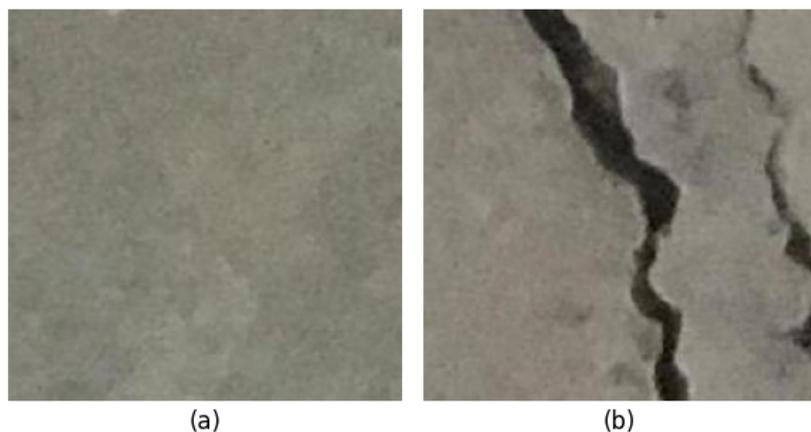


Figura 2. Amostras de ambas as classes. (a) Classe Negativa; (b) Classe Positiva.

Cabe ressaltar que nas imagens da base não foram aplicadas técnicas de aumento de dados, tais como rotações. No entanto, para verificar se era possível obter resultados satisfatórios na classificação com o uso de uma quantidade menor de imagens, o presente trabalho utilizou apenas 5% da base de imagens original, divididas igualmente entre as classes existentes, ou seja, 1 mil imagens de ambas as classes.

3.2. Seleção do modelo pré-treinado

A seleção da rede neural para este estudo envolveu uma análise criteriosa dos modelos disponíveis no framework Keras. Os critérios considerados incluíram a acurácia dos modelos, com uma ênfase na escolha dos cinco melhores classificados na plataforma, além da busca por modelos com um número mínimo de parâmetros para otimizar a eficiência computacional. Outro fator crítico foi a adequação dos modelos para dispositivos com altos requisitos de eficiência energética e computacional.

Após uma análise criteriosa, optamos pelo modelo MobileNetV2 devido à sua eficiência e tamanho compacto, em contraste com redes mais complexas como a VGG ou a ResNet. Essa escolha é especialmente relevante para dispositivos móveis e sistemas embarcados, onde a otimização de recursos computacionais e eficiência energética é fundamental. Além disso, a qualidade das imagens no conjunto de dados utilizado aumenta nossa confiança de que o MobileNetV2 entregará resultados precisos na tarefa de classificação de imagens.

3.3. Pré-processamento das imagens

A MobileNetV2 no Keras inicia com uma camada de pré-processamento que normaliza e centraliza os valores dos pixels das imagens, garantindo que estejam dentro de uma faixa específica. Este modelo pode ser inicializado com os pesos da base de imagens ImageNet, que contém milhões de imagens classificadas em diversas categorias.

Para aumentar a capacidade de aprendizado, camadas adicionais são incorporadas à rede e treinadas usando 70% do conjunto de dados de treinamento. Os hiperparâmetros são otimizados utilizando 10% do conjunto de validação, e a capacidade de generalização é avaliada com os 20% restantes do conjunto de teste. A última camada da rede utiliza a função de ativação *sigmoid*, que mapeia os valores de entrada em uma escala de 0 a 1, fornecendo a probabilidade de cada classe ser a correta para a imagem. Isso capacita a rede a realizar classificação multiclasse com alta precisão.

3.4. Transferência de aprendizagem

A MobileNetV2 é carregada sem as camadas do topo. As camadas base são congeladas, o que significa que seus pesos e parâmetros são mantidos fixos durante o treinamento. Isso é feito para evitar que as camadas base sejam ajustadas e para preservar as características já aprendidas durante o treinamento no conjunto de dados ImageNet.

Em sequência, novas camadas são adicionadas à rede para realizar a identificação de rachaduras em concreto. Essas camadas são treinadas no conjunto menor de dados para a classificação correta, enquanto as camadas base são mantidas fixas. O objetivo é ajustar apenas as camadas adicionadas para que elas possam aprender a classificar as imagens de acordo com a tarefa específica.

3.5. Treinamento do modelo

O treinamento da rede MobileNetV2 segue uma abordagem de transferência de aprendizado, utilizando uma rede pré-treinada na ImageNet como base para a iden-

tificação de rachaduras em estruturas de concreto. Camadas adicionais são incorporadas à rede e treinadas com o conjunto de dados de treinamento. Ajustes nos hiperparâmetros, como a taxa de aprendizagem e a arquitetura da rede, são feitos usando o conjunto de validação para melhorar o desempenho.

Após a otimização dos hiperparâmetros, o conjunto de teste é usado para avaliar quão bem a rede consegue aplicar seu aprendizado. A precisão é medida comparando suas previsões com os valores reais. O treinamento ocorre ao longo de 10 épocas, onde uma época representa uma passagem completa pelos dados de treinamento. Durante esse processo, os pesos das camadas adicionais são ajustados para minimizar a função de perda, que avalia a diferença entre as previsões da rede e os valores reais dos dados de treinamento. A taxa de aprendizagem é fixada em 0,0001, controlando o tamanho dos ajustes de peso em cada iteração de treinamento.

3.6. Obtenção das métricas de desempenho

A MobileNetV2 é uma rede neural convolucional de aprendizado profundo que pode ser configurada para retornar diversas métricas de desempenho. Para a realização deste trabalho foram selecionadas as métricas de acurácia, precisão, *recall* e *F1-score*. A obtenção dessas métricas envolve a avaliação da rede em um conjunto de dados de teste, onde a saída prevista da rede é comparada com a saída real para cada imagem.

A acurácia é uma métrica que representa a proporção de classificações corretas pelo total de classificações do domínio [Rizk et al. 2019]. Já a precisão, entrega uma informação que contrasta com a acurácia, pois apresenta a visualização de quantos resultados positivos obtidos são verdadeiramente positivos [Rizk et al. 2019]. O F1 Score mede a precisão geral de um modelo de classificação, levando em consideração tanto a precisão quanto o *recall*. O *recall*, é a relação entre os resultados previstos e os resultados previstos corretamente [Antelo 2022].

4. Resultados e Discussão

Nesta seção, apresentaremos os resultados obtidos em nossos experimentos com a técnica de transferência de aprendizagem na arquitetura MobileNetV2 utilizando os pesos da ImageNet. Com isso, foi possível alcançar resultados expressivos e de alta qualidade.

A Tabela 2 apresenta uma comparação dos resultados de desempenho obtidos neste trabalho com os resultados dos outros trabalhos relacionados. Nela há as métricas de acurácia, precisão, *F1*, *recall* e perda (*loss*) obtidos por cada trabalho, quando disponíveis. A acurácia foi comum a todos os trabalhos listados, sendo esta a principal métrica utilizada para comparação.

Os resultados deste trabalho foram obtidos após o treinamento do modelo com a MobileNetV2 e sua avaliação perante o conjunto de imagens de teste, que contém cerca de 20% das 2.000 imagens utilizadas. Como visto na Tabela 2, foi possível obter acurácia de 99,48% e precisão de 99,49%, além de uma taxa de perda de 0,057. Isso mostra a capacidade do modelo em convergir para resultados robustos mesmo com uma quantidade reduzida de imagens.

Tabela 2. Comparação dos resultados de desempenho entre esse trabalho e os relacionados.

Trabalho	Acurácia	Precisão	F1	Recall	Loss
[Flah et al. 2020]	0,9825	-	-	-	0,0570
[Arafin et al. 2022]	0,9100	0,8200	-	0,100	-
[Kim et al. 2020]	0,9900	-	-	-	-
[Dung and Anh 2019]	-	0,989	0,988	-	0.0420
Este	0,9948	0,9949	0,9948	0,9949	0,0715

A Figura 3 apresenta as curvas de aprendizagem do modelo, onde o primeiro gráfico representa o treino e validação na métrica de acurácia durante as épocas, mostrando que já na 4^a época o modelo se aproximou dos resultados finais, e o segundo gráfico mostra a taxa de perda durante o treino e validação.

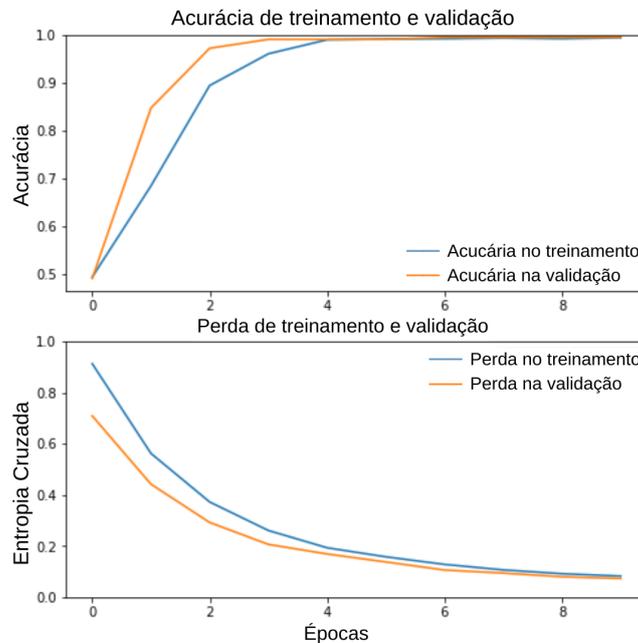


Figura 3. Curvas de aprendizagem da acurácia/perda de treinamento e validação.

5. Conclusão

Este estudo teve como objetivo treinar um modelo de rede neural convolucional capaz de identificar rachaduras em estruturas de concreto a partir de imagens, visando uma avaliação mais precisa e menos custosa do que as tradicionais inspeções humanas em áreas de risco. Além disso, buscou-se validar a usabilidade de arquiteturas leves para esse problema, permitindo uma avaliação mais rápida, precisa e menos custosa em relação à subjetividade da avaliação de inspetores e ao deslocamento para áreas de risco.

O modelo produzido apresentou através da métrica de acurácia um desempenho de 99,48%, um valor próximo aos resultados obtidos pelos trabalhos relacionados

listado, utilizando 5% da quantidade de imagens do *dataset* sendo esta uma quantidade inferior a quantidade utilizada nos demais trabalhos. Vale ressaltar que a utilização de uma arquitetura de rede neural convolucional adequada para dispositivos móveis, como a MobileNetV2, possibilitou a obtenção desses resultados com quantidades menores de imagens. Isso sugere que o uso de modelos leves e eficientes pode ser uma opção viável para a identificação de rachaduras em estruturas de concreto por meio de imagens coletadas por *drones*.

Para trabalhos futuros, pode-se explorar ainda mais a utilização de arquiteturas leves, a fim de tornar o processo mais eficiente e menos custoso. Além disso, é viável testar o modelo produzido em um ambiente real, implantando-o em um *drone* para identificar rachaduras em tempo real.

Referências

- Antelo, A. C. L. (2022). *Visão computacional para detecção de hábitos alimentares*. PhD thesis.
- Arafin, P., Issa, A., and Billah, A. H. M. M. (2022). Performance comparison of multiple convolutional neural networks for concrete defects classification. *Sensors*, 22(22).
- Dung, C. V. and Anh, L. D. (2019). Autonomous concrete crack detection using deep fully convolutional neural network. *Automation in Construction*, 99:52–58.
- Flah, M., Suleiman, A. R., and Nehdi, M. L. (2020). Classification and quantification of cracks in concrete structures using deep learning image-based techniques. *Cement and Concrete Composites*, 114:103781.
- Kim, J. J., Kim, A.-R., and Lee, S.-W. (2020). Artificial neural network-based automated crack detection and analysis for the inspection of concrete structures. *Applied Sciences*, 10(22):8105.
- Özgenel, Ç. F. (2018). Concrete crack images for classification. *Mendeley Data*, 1(1).
- Rizk, Y., Hajj, N., Mitri, N., and Awad, M. (2019). Deep belief networks and cortical algorithms: A comparative study for supervised classification. *Applied computing and informatics*, 15(2):81–93.
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., and Chen, L.-C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4510–4520.