

Classificação de Rachaduras de Concreto Utilizando Rede Neural Convolutacional Profunda

Concrete Crack Classification Using Deep Convolutional Neural Network

Lucas Candido¹

 orcid.org/0009-0008-7126-4755

Leonardo Medeiros²

 orcid.org/0000-0002-5015-6957

Alexandre Machado²

 orcid.org/0009-0003-7946-6603

¹Instituto Federal de Alagoas – Campus Maceió, Maceió, Brasil.
E-mail: lucassantoscandido@gmail.com

²Instituto Federal de Alagoas – Campus Maceió, Maceió, Brasil.

DOI: 10.25286/repa.v9i3.2518

Esta obra apresenta Licença Creative Commons Atribuição-Não Comercial 4.0 Internacional.

Como citar este artigo pela NBR 6023/2018: Lucas Candido; Leonardo Medeiros; Alexandre Machado. Classificação de Rachaduras de Concreto Utilizando Rede Neural Convolutacional Profunda. Revista de Engenharia e Pesquisa Aplicada, v.9, n. 3, p. 55-69, 2024.

RESUMO

O concreto é um dos materiais mais utilizados na construção civil e como qualquer outro material está sujeito a degradação patológica, essa e outras deteriorações são comumente negligenciadas. A falta de inspeção adequada está correlacionada a 66% dos acidentes prediais ocorridos. A operação de manutenção de obras geralmente envolve uma inspeção visual para avaliação do seu estado e posteriormente a busca por soluções. Esse dano frequentemente se manifesta em forma de rachaduras, trincas ou fissuras, que caso não tratadas, podem acarretar o colapso parcial ou total da estrutura. Tendo isso em vista, o presente artigo tem como finalidade propor uma nova arquitetura de rede neural convolutacional profunda para identificar rachaduras em superfícies de concreto. O classificador proposto alcança 94,04% de assertividade no diagnóstico da patologia; sendo treinado, validado e testado em 25.000 amostras de imagens oriundas de três banco de dados.

PALAVRAS-CHAVE: Rede Neural Convolutacional; Classificação de Imagens; Concreto; Rachadura;

ABSTRACT

Concrete is one of the most used materials in civil construction, and like any other material, it is subject to pathological degradation, which is often overlooked. The lack of proper inspection is correlated with 66% of building accidents. The maintenance operation of structures typically involves a visual inspection to assess their condition and subsequently seek solutions. This damage often manifests in the form of cracks or fissures, which, if left untreated, can lead to partial or total structural collapse. This article aims to propose a new architecture of deep convolutional neural network for identifying cracks on concrete surfaces. The proposed classifier achieves 94.04% accuracy in diagnosing the pathology, having been trained, validated, and tested on 25,000 image samples from three databases.

KEY-WORDS: Convolutional Neural Network; Image Classification; Concrete; Crack;

1 INTRODUÇÃO

Historicamente na construção civil o concreto detém um papel fundamental como material de construção, tornando-se um dos mais consumidos devido sua boa resistência, baixo custo de produção e manutenção [1]. O produto fomenta a vida útil das edificações, atuando como uma substância maleável para as mais diversas aplicações que ao se enrijecer, desenvolve a resistência semelhante à de rochas naturais.

Contudo, como qualquer outro material, está sujeito a ações de agentes externos que resultam em danos estruturais e patológicos que contribuem para a degradação das construções, caso estes desgastes não sejam tratados podem ocasionar em colapso parcial ou total da estrutura [2].

Nesse contexto, os responsáveis pela edificação devem contratar profissionais técnicos para realizar a inspeção predial, que atua como uma análise visual dos danos estruturais, fornecendo uma visão ampla e detalhada do estado de conservação da edificação, conhecendo os pontos críticos a serem corrigidos em uma manutenção efetiva, estendendo a vida útil do imóvel [3].

Entretanto, Villanueva [4] explica que a prática de manutenção no Brasil é comumente negligenciada, não sendo bastante difundida na sociedade. O pesquisador afirma que a ausência do monitoramento adequado das edificações de concreto é responsável por diversas manifestações patológicas e danos significativos na estrutura.

O estudo realizado pelo IBAPE-SP [3] constata que 66% dos acidentes prediais ocorridos são relacionados a uma precária ou falta de manutenção e os outros 34% a anomalias construtivas, ocorrendo colapso parcial ou total do sistema estrutural em mais da metade dos acidentes avaliados. O presidente do CREA-PB [5], de maneira semelhante, afirma que 70% dos acidentes de edificações na Paraíba também são causados devido à falta de monitoramento adequado.

Diante das manifestações patológicas que podem resultar nas fatalidades citadas anteriormente, encontram-se as rupturas ocorridas no concreto sob ações mecânicas ou físico-químicas que são comumente classificadas como trincas, fissuras e rachaduras. Essas aberturas podem causar consequências graves na estrutura, pois servem como porta de entrada para outros agentes agressivos para o interior da estrutura, fazendo que outros fenômenos de degradação ocorram ou

auumentem exponencialmente, como a corrosão de armaduras [6].

Diante disto, o emprego de técnicas baseadas em visão computacional e aprendizado de máquina vem sendo exploradas nas últimas décadas com o intuito de monitorar e inspecionar diversos tipos de estruturas, devido a sua capacidade de processar grandes volumes de dados e analisar as informações obtidas [7]. Em especial a Convolution Neural Network (CNN), as quais tem se provado uma ferramenta poderosa capaz de compreender o conteúdo de imagens e vídeos, sendo altamente eficaz em classificação de imagens, detecção de objetos, segmentação de instâncias e outras atividades relacionadas [8].

O grande atrativo da CNN, em contrapartida aos métodos tradicionais de Machine Learning (ML), segundo Khan, A. et al. [9], seria a sua topologia dividida em vários estágios de aprendizado, cada um composto por uma combinação de camadas de processamento, proporcionando a exploração da correlação espacial e/ou temporal dos dados. Os pesquisadores complementam que a CNN, se arquitetada corretamente, consegue aprender representações internas dos pixels brutos em seus variados níveis de abstração; sem a necessidade ou com o mínimo de intervenção humana [8].

Os avanços na visão computacional viabilizam a oportunidade de expandir e aprimorar a prática da detecção de patologias através do aprendizado profundo. Tendo isso em vista, o presente estudo tem como objetivo a criação e implementação de uma nova arquitetura de Rede Neural Convolutacional (RNC) para classificação de imagens, proporcionando o diagnóstico e monitoramento de fissuras, trincas e rachaduras nas superfícies de estruturas de concreto.

2 TRABALHOS RELACIONADOS

A visão computacional aplicada a inspeção e monitoramento de infraestrutura civil vem sendo estudada extensivamente ao longo dessas últimas décadas, de tal maneira que os pesquisadores apresentaram diversas abordagens para a detecção de danos patológicos em elementos estruturais

Os primeiros estudos se baseavam em técnicas tradicionais de processamento de imagens para extração de danos em edificações [10]. Conforme Azouz et al. [11] esclarece, os filtros heurísticos utilizados focavam na detecção de bordas pelo operador de Sobel ou de Canny, também aplicando a equalização de histogramas e operações

morfológicas. Kun et al. [12] relata que a extração das características se limitava as camadas rasas e intermediárias da rede, sendo eficiente em detectar fissuras sob superfícies lisas em boas condições, o pesquisador aponta a dificuldade destes sistemas de detectar trincas em ambientes realistas.

A complexidade da superfície de concreto dificulta o processo de segmentação do sistema, uma vez que estas variações são comumente interpretadas como aspectos da manifestação patológica e confundem o algoritmo. Dentre as principais variantes, podemos destacar a ocorrência de granulação na imagem provenientes do desgaste ou erosão do concreto, variação de iluminação em áreas com sombras e a presença de elementos físicos [12]. Diante disso, Nguyen et al. [13] esclarece que diversos pesquisadores propuseram táticas de pré-processamento de dados a fim de diminuir a influência da granulação na classificação da imagem, entretanto, o autor complementa dizendo que tais técnicas possuem limitações relacionadas à diminuição da complexidade dos elementos de cenas.

Métodos tradicionais de aprendizado de máquina, tais como Support Vector Machine (SVM), Artificial Neural Network (ANN), e Random Forest (RF) necessitam de uma fase pré-definida para extração de recursos [13]. Esta etapa tem como finalidade reduzir a complexidade de dados e facilitar o aprendizado dos padrões necessários a solução do problema [11]. Embora o tratamento da imagem melhore a precisão da classificação em determinada situação, Hu et al. [14] explica que cada técnica se aplica somente ao cenário a qual foi designada, não generalizando a ambientes com fatores divergentes. Isto é, os modelos tradicionais de ML apresentam dificuldade em lidar com rachaduras presentes em superfícies de concreto complexas e mutáveis [13] [14].

Os recentes avanços em metodologias de Deep Learning (DL) tornaram a identificação de fissuras sob ambientes de alta complexidade uma realidade [12]. A popularidade dos métodos de aprendizado profundo origina-se da sua capacidade de extrair características semânticas profundas da imagem com alta precisão. Segundo Vogado et al. [15], a CNN tem se destacado como o tipo de modelo profundo mais bem-sucedido, apresentando altos índices de acurácia para problemas conhecidos na literatura. Além disso, estas arquiteturas também são aplicáveis a tarefa de detecção de aberturas de concreto [16].

Inicialmente, Cha et al. [17] estabeleceu um marco ao empregar MatConvNet na construção de um classificador robusto aplicado na detecção das ocorrências de rachaduras em superfícies de concreto. O modelo, composto por oito camadas, (quatro convolucionais, duas de agrupamento, uma de ativação e uma de classificação com *softmax*), foi treinado e validado em um conjunto de dados próprios com 40.000 fragmentos de imagens de estruturas de concreto (32.000 para treinamento e 8.000 para validação). A CNN alcançou a acurácia de 98,22% e 97,95% durante o procedimento de treinamento e validação, respectivamente. Para avaliar a eficiência do algoritmo no ambiente de teste, os autores optaram por utilizar a técnica de *sliding window*, uma janela percorre e divide todas as regiões da superfície em fragmentos que são avaliados individualmente pelo classificador. As imagens submetidas ao teste foram capturadas em local e estrutura diferente dos dados anteriores, totalizando 55 amostras. A acurácia obtida no teste foi de 97%, valor quase idêntico ao da acurácia do processo de validação. Por fim, os pesquisadores constatam que o sistema proposto obteve um desempenho robusto em comparação a métodos tradicionais de detecção de bordas, como o filtro de Canny e Sobel.

Considerando uma visão panorâmica, Zheng et al. [18] investigou o desempenho de três modelos de CNN na detecção de rachaduras em concreto. Estas arquiteturas foram originalmente treinadas sob imagens de paisagens naturais, a estrutura dos modelos fora adaptada para solução do problema proposto. A metodologia adotada pelos autores visou comparar o desempenho de cada arquitetura sob condições semelhantes, diante disso os modelos foram submetidos ao mesmo conjunto de dados e parâmetros de treinamento. Os resultados revelaram variações na precisão e acurácia de cada rede ao decorrer do experimento, sugerindo que o aumento do número de camadas convolucionais e a profundidade gradual da rede contribuem para melhorar o desempenho na tarefa de detecção de fissuras em concreto.

Embora o aumento da profundidade de uma rede possibilite a solução de tarefas complexas [18] [19], à medida que se empilham mais camadas, problemas de degradação de gradientes, como desvanecimento e explosão, tornam-se mais proeminentes. Isso significa que o classificador pode não convergir para uma solução adequada, ou seja, em efeitos práticos a CNN terá dificuldade em detectar rachaduras finas ou sutis e classificar

erroneamente áreas não rachadas como rachadas ou não identificar rachaduras reais.

Diante disto, Gao et al. [20] propõem a aplicação de modelos da família ResNet, em conjunto com a técnica de transferência de aprendizado (*transfer learning*) para detecção de danos estruturais. A escolha da ResNet se deve ao seu aprendizado residual que atenua a degradação do gradiente, e suas conexões de atalho que capturam detalhes sutis nas imagens, melhorando a precisão no discernimento de padrões de danos estruturais. Os modelos pré-treinados no ImageNet passaram por ajuste fino (*fine-tuning*) dos parâmetros ao novo conjunto de 2.000 imagens, adaptando-os a tarefa de classificação de amostras sem deterioração, avarias estruturais por flexão e cisalhamento. Em suma, os resultados obtidos demonstram a eficácia da ResNet pré-treinada na identificação precisa de diferentes tipos de danos estruturais, por fim os autores destacam como *transfer learning* aumentou a capacidade de generalização da CNN.

O estudo conduzido por Sá et al. [21] também aplica a tática de transferência de aprendizado em modelo pré-treinado no ImageNet para classificação de fissuras de concreto. A seleção de MobileNetV2, conforme os autores explicaram, decorreu de sua eficiência e menor custo computacional comparado a outras CNN complexas (como ResNet e VGG), sendo ideal para sistemas móveis e embarcados. O modelo foi ajustado para classificação binária de manifestação patológica de fissuras, executaram o experimento em 1.000 imagens de cada categoria, fragmentadas nos subconjuntos de treinamento, validação e teste, com a proporção de 70:10:20, respectivamente. A CNN proposta atingiu 99,48% de acurácia e 99,49% de precisão, perante os dados de teste. Os autores concluem que modelos leves e eficientes, como MobileNetV2, pode ser uma opção viável para a detecção de rachaduras em imagens coletadas por veículo aéreo não tripulado (VANT).

Diniz et al. [22] desenvolvem uma metodologia para identificar de manifestações patológicas em estruturas de concreto empregando a rede neural profunda YOLOv4, devido a sua velocidade na detecção de objetos. O estudo lida com 5 tipos de danos presentes no concreto, o banco de dados com 1.052 imagens em alta resolução fora dividido em dois subconjuntos, 840 amostras usadas no treinamento e 212 no teste. A quantidade de danos presentes em cada classe corresponde a 2.507 rachaduras, 1.507 barras expostas em concreto armado, 833 eflorescência, 1.898 de fragmentação e 1.559 manchas de corrosão do aço no concreto. O

treinamento foi conduzido por 1.500 épocas e avaliado através da métrica mAP (*mean average precision*), a medida indica o quão bem o modelo de detecção de objetos opera. O algoritmo proposto conquista sua maior performance durante a 1.300 época, com mAP de 24,44% para corrosão, 27,24% em barras expostas, 19,29% para fragmentação, 11,80% para fissuras e 5,62% para eflorescência. Por fim, os pesquisadores concluem que o detector desenvolvido apresentou boas habilidades em identificar manifestações patológicas no concreto, utilizando imagens em boas condições, abrindo o questionamento de como melhorar a precisão do algoritmo na ocorrência de danos que não estejam tão bem evidenciados na imagem.

Em uma direção complementar, Kao et al. [23] combinam YOLOv4 com técnicas de processamento de imagens para identificar e mensurar trincas em pontes por intermédio de VANT. A metodologia da pesquisa consistiu em instalar marcadores de plano em locais de interesse da ponte em conjunto a uma estação total para medição da manifestação patológica. Em seguida o VANT coleta as imagens através de uma câmera fotográfica, as amostras passam por correção de distorção e são enviadas ao detector de objetos. O YOLOv4 detecta e recorta as regiões da cena que possuem rachaduras, os fragmentos passam por limiarização da imagem a fim de estimar as dimensões da trinca através de escala. Os resultados obtidos demonstram que o sistema alcançou uma precisão (mAP) de 92% na identificação da manifestação patológica, sendo capaz de medir rachaduras com 0,22 milímetros de largura. Embora o algoritmo tenha alcançado bons números, a taxa de erro da medição oscilou entre 0,31% e 38,29% de acordo com a amostra usada, validando a necessidade de aperfeiçoamento da técnica de medição proposta na pesquisa.

Com isso, observamos que modelos de análise preditiva possuem uma vasta aplicação, podendo ser incorporado em conjunto com outras tecnologias para aumentar a sua eficácia. Sendo assim, neste artigo exploramos uma nova abordagem para classificação de rachaduras em imagens de concreto usando redes neurais convolucionais.

3 MATERIAIS E MÉTODOS

O vigente projeto fundamenta-se nos conceitos metodológicos definidos por Gil [24], atuando como uma pesquisa exploratória em seu objetivo de estudo, com o intuito de aprofundar o

conhecimento na área de aprendizado profundo para classificação de imagens. Dito isto, a pesquisa classifica-se como de natureza aplicada, pois tem como propósito dirigido à solução do problema de pesquisa, havendo abordagem quantitativa para a análise dos resultados.

3.1 BANCO DE DADOS

A eficiência da construção de uma arquitetura de rede neural limita-se ao banco de dados utilizado, seja à quantidade de amostras ou devido a qualidade dos dados [25]. Além disto, a relevância desta problemática intensifica-se ao analisarmos que a manifestação patológica de fendas e aberturas em estruturas e superfícies de concreto apresenta diversas variações de tamanho e formato.

Portanto, optou-se para o desenvolvimento da pesquisa a criação de um conjunto de dados a partir da coleta de amostras de imagens provenientes de três estudos relacionados a classificação de trincas, fissuras e rachaduras de concreto [26] [27] [28].

A base de dados contém 25.000 imagens em sua totalidade, sendo composta por 12.500 amostras positivas para a ocorrência de rachaduras e outras 12.500 sem a manifestação patológica. O conjunto abrange fragmentos de estruturas de concreto, formado por edifícios, pontes, decks, paredes e pisos. A Figura 1 demonstra como cada superfície apresenta variação na condição física do material e

servindo informações para a aprendizagem e extração dos padrões das imagens. O *dataset* de teste representa um ambiente experimental para avaliar o desempenho da rede neural em classificar dados desconhecidos.

3.1.2 Pré-Processamento dos Dados

Atualmente as imagens são obtidas por diversos equipamentos fotográficos e fontes, ocasionando variações no formato do material que afetam o desempenho da classificação. O pré-processamento atua como solução a esta divergência dos dados, a técnica selecionada redimensiona as imagens para a dimensão de 224 x 224 x 3, e reescala os valores entre o intervalo de [0, 255].

O método de *Data Augmentation* gera novos exemplares durante o processo de treinamento, as amostras passam por transformações geométricas e/ou fotométricas com finalidade de melhorar a generalidade do modelo [25]. As técnicas de aumento de dados utilizadas foram: inversão vertical, inversão horizontal, rotação, alteração do brilho, ruído gaussiano e variação no zoom da imagem.

3.2 CONSTRUÇÃO DA ARQUITETURA

As redes neurais convolucionais são designadas para processar múltiplos dados matriciais através de camadas empilhadas. A arquitetura básica de uma CNN é estruturada por uma série de estágios, as primeiras camadas são para a extração de recursos e redimensionamento dos dados, as últimas camadas analisam as características extraídas baseando-se em padrões descobertos previamente na finalidade de classificar os elementos.

Neste presente trabalho é proposto uma nova arquitetura para classificação de imagens de superfícies de concreto com rachaduras. O modelo foi construído a partir do zero, selecionando diferentes tipos de camadas e hiperparâmetros que as compõem, a fim de obter boa precisão em detectar rachaduras e melhor custo computacional. Esta seleção foi realizada manualmente por tentativa e erro, tendo em vista que não há formulação matemática para o cálculo de camadas e parâmetros adequados para um conjunto de dados específico.

Como demonstrado na Figura 2, o modelo recebe de entrada imagens em 224 x 224 x 3 pixels, os

Figura 1 – Amostras do conjunto de dados.



Fonte: Adaptado de [26] [27] [28] .

em sua área de iluminação.

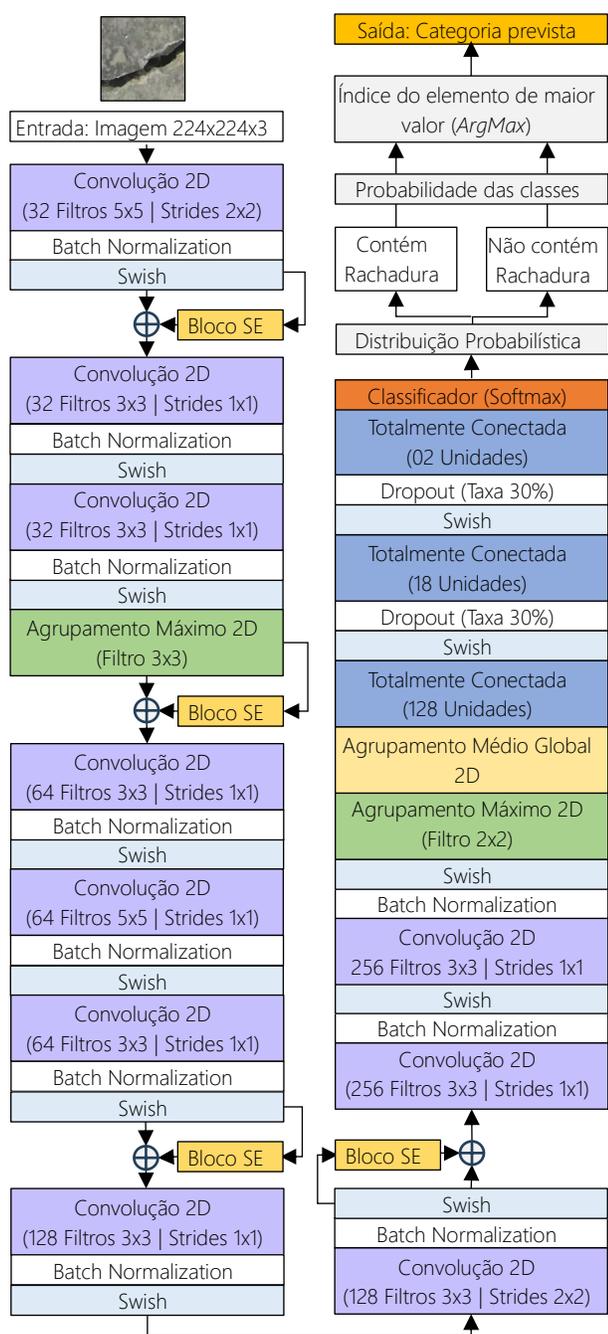
3.1.1 Preparação dos Dados

As imagens são categorizadas em duas classes, conforme indicado na Figura 1, fracionadas nos subconjuntos de treinamento com 17.500 imagens (70%), validação contendo 5.000 (20%) e teste com 2.500 (10%), mantendo a proporção de classes.

Os dados de treino e validação atuam paralelamente durante o processo de treinamento,

dados passam sequencialmente por diversas camadas a fim de obter a classificação.

Figura 2 – Diagrama da arquitetura proposta.

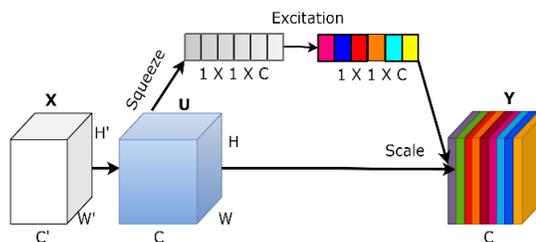


Fonte: Os autores.

2.1 Camada Convolutiva

A camada convolutiva é o bloco de construção principal de uma rede neural convolutiva. O seu papel fundamental é a extração de recursos através da aplicação de filtros nas imagens de entrada, utilizando a operação de convolução. Cada filtro

Figura 3 – Módulo de Squeeze-and-Excitation.



Fonte: Os autores.

convolutiva consiste em uma matriz de pesos específica que atua na vizinhança local de pixels da amostra. O uso de múltiplos filtros geram um número arbitrário de feature maps, os quais representam as diferentes características extraídas. A operação pode ser expressa pela Equação 1.

$$Y_i = B_i + \sum_{j=1}^n X_j * K_{i,j} \quad (1)$$

Onde a saída, Y_i , consiste em i -ésimos mapas de características. X_j e B_i representam a entrada da camada e a matriz de bias, respectivamente. $K_{i,j}$ denomina o filtro que conecta o i -ésimo e o j -ésimo feature map de saída e entrada, respectivamente.

3.2.2 Camada de Batch Normalization

Proposta por Ioffe et. al. [29], *Batch Normalization* (BN) funciona como uma camada suplementar que regulariza os dados de entrada, transformando todos os valores de saída para que mantenham a sua média próxima a 0 e o desvio padrão próximo a 1.

O procedimento de normalização com BN melhora a eficiência e velocidade de treinamento de uma rede neural, diminui os riscos de *overfitting* e permite altas taxas de aprendizado, sendo geralmente aplicada entre a camada de convolução e ativação.

3.2.2 Camada de Ativação

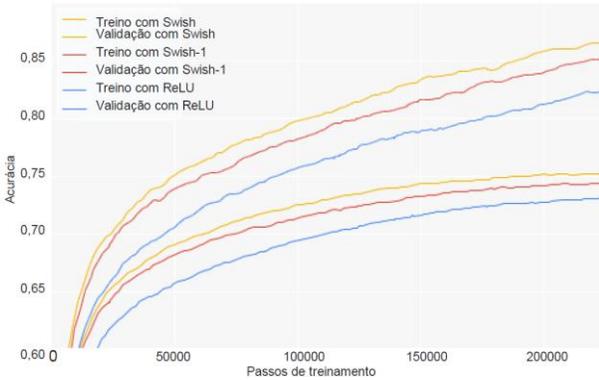
A escolha da função de ativação em uma rede neural profunda impacta de forma significativa o desempenho do modelo durante o seu treinamento e classificação, como explica Ramachandran et. al. [30]. O pesquisador apresenta a função de ativação swish como substituição ao ReLU, embasando-se na superioridade do swish ao utilizá-lo em arquiteturas originalmente projetadas com ReLU, como validado na Figura 4.

Swish atua como uma função não monotônica e de curva suave, definida pela Equação 2. Em

virtude disso, a função permite a propagação de valores negativos, ao contrário da função ReLU, aproveitando os pequenos números negativos para capturar padrões subjacentes e relevantes dos dados. Em contrapartida, a função demanda maior custo computacional.

$$f(x) = x * sigmoid(x) \tag{2}$$

Figura 4 – Curvas de treinamento da rede Mobile NASNet-A no conjunto de dados ImageNet.



Fonte: Adaptado de [30].

3.2.2 Camada de Agrupamento

A operação de *pooling* reduz a dimensão dos mapas, preservando as características principais e descartando as irrelevantes, consequentemente reduzindo o tempo de computação e melhorando a capacidade de generalização do modelo.

A camada de agrupamento divide os mapas em sub-regiões e executam sua operação de acordo com a estratégia de agrupamento. Max pooling descarta todos os valores mínimos de uma determinada área e retorna seu valor máximo, enquanto Average pooling calcula a média dos valores de cada vizinhança.

3.2.4 Camada Totalmente Conectada

Também conhecida como camada densa, funciona de forma semelhante a redes neurais *Multilayer Perceptron* (MLP). Atua ajustando os parâmetros do vetor unidimensional de entrada através da utilização de pesos para criar uma representação de probabilidade estocástica de cada classe, a última camada densa do modelo fornece as informações à camada de *softmax* para realizar a classificação da imagem. O resultado da operação é definido conforme Eq. 3.

$$Y_i = B_i + \sum_{j=i}^n W_{i,j} X_j \tag{3}$$

Em que Y_i representa a saída, B_i denota o bias, X_j equivale ao vetor de entrada, $W_{i,j}$ corresponde a

conexão ponderada entre a j -ésima unidade de X_j e o i -ésimo elemento de Y_i .

3.2.5 Camada Dropout

Dropout regulariza a problemática do overfitting enquanto amplia a precisão da arquitetura. A camada opera durante o treinamento escolhendo aleatoriamente um grupo de neurônios que serão descartados durante alguma etapa do aprendizado. O processo assegura melhor processamento dos dados e generalização da rede, prevenindo que os neurônios de uma determinada camada sejam sempre influenciados pelas unidades anteriores.

3.2.6 Camada de Classificação

A classificação é realizada através da última camada da arquitetura, a *Softmax*, que calcula a distribuição das probabilidades para cada classe de imagem existente no modelo.

A saída de *Softmax* (Equação 4) fornece a probabilidade de uma imagem pertencer a uma determinada classe, através de uma distribuição categórica entre o intervalo de 0 e 1, em que a soma dos valores será igual a 1. A categoria com o maior escore probabilístico será selecionada como a classe prevista.

$$Y(X_i) = \frac{e^{X_i}}{\sum_{j=i}^n e^{X_j}} \tag{4}$$

O i -ésimo elemento da entrada, X_i , na saída, Y , representa a previsão probabilística de X pertencer a classe i .

3.2.6 Módulo de Squeeze-and-Excitation

O bloco de Squeeze-and-Excitation (SE) [31] funciona como uma unidade arquitetural de channel-wise attention para redes neurais convolucionais, ou seja, o mecanismo busca obter as representações das propriedades mais relevantes de uma imagem, enquanto suprime os recursos desprezíveis através do ajuste de pesos.

O módulo é composto por três operações com funcionalidades distintas, denominadas de squeeze, excitation e scale, como demonstrado na Figura 3. A operação de squeeze tem como função principal a extração de informações globais de cada canal dos feature maps. Excitation manuseia camadas totalmente conectadas para gerar os pesos dos canais e obter sua correlação com os recursos. A etapa de scale redimensiona os canais com os pesos das informações adquiridas, calibrando as dimensões do resultado com o tamanho original da amostra de entrada.

Em suma, as operações do módulo de Squeeze-and-Excitation redimensionam as informações dos

feature maps, com finalidade de manter elementos pertinentes a sua atividade e reduzir a influência de dados irrelevantes, amplificando a competência do reconhecimento da rede.

3.3 DEFINIÇÕES DE APRENDIZADO

O procedimento de treinamento da rede neural sucede de forma supervisionada. O aprendizado supervisionado utiliza conjuntos de dados previamente rotulados, no qual para cada amostra haverá um vetor com a verdadeira classe da imagem e outro com o valor previsto pelo modelo. Os resultados serão comparados e analisados através do supervisor externo da rede, o qual fornecerá as relações entre os parâmetros de aprendizado utilizados para a solução da tarefa e o direcionamento do reajuste de seus pesos.

A chave para obter a performance ideal durante este processo condiz na configuração dos parâmetros de treino, os quais influenciam de forma significativa o comportamento do modelo, descritos na Tabela 1.

A quantidade de épocas (*epochs*) define quantas vezes o conjunto de dados de treino passa pelo modelo. O tamanho do lote (*batch size*) delimita o número de amostras utilizadas em cada iteração do treinamento. A taxa de aprendizado (*learning rate*) indica a magnitude de como os pesos do modelo serão atualizados durante o processo de backpropagation do treinamento.

Tabela 1 – Definição dos parâmetros de treinamento.

PARÂMETRO	VALOR
ÉPOCAS	50
TAMANHO DO LOTE	128
OTIMIZADOR	ADAM
FUNÇÃO DE PERDA	CROSS-ENTROPY
TAXA DE APRENDIZADO INICIAL	1×10^{-3}

Fonte: Os autores.

3.3.1 Função de Perda

A Função de Perda *Cross-Entropy* é utilizada para quantificar a discrepância entre o resultado correto esperado e o obtido pela arquitetura, este erro é usado para atualizar os parâmetros internos e pesos da rede [32]. O algoritmo mede o desempenho do modelo de classificação através de uma distribuição probabilística cuja saída é um escore entre 0 e 1.

3.3.2 Otimizador

O Otimizador *Adaptive Moment Estimation* (ADAM) [33] é um método empregado para melhorar a precisão da rede neural com a otimização estocástica eficiente. O algoritmo utiliza taxas de aprendizado adaptáveis para encontrar as taxas de aprendizado individuais dos diferentes hiperparâmetros que compõe a arquitetura durante o treinamento, encontrando valores que minimizem o erro do modelo.

3.4 MEDIDAS DE DESEMPENHO

A avaliação do modelo ocorre por meio de diferentes expressões matemáticas que medem a performance da rede, as quatro métricas empregadas neste estudo são acurácia (*accuracy*), precisão (*precision*), sensibilidade (*recall*) e pontuação F1 (*F1-Score*).

- **Acurácia** define a proporção de classificações corretas para a quantidade total de amostras.

$$Acurácia = \frac{(VP + VN)}{(VP + FP + VN + FN)} \quad (5)$$

- **Precisão** mede a exatidão do número de imagens com fissuras que foram previstas corretamente.

$$Precisão = \frac{VP}{(VP + FP)} \quad (6)$$

- **Sensibilidade** ou simplesmente revocação, calcula a porcentagem da quantidade de superfícies de concreto com rachaduras previstas corretamente em relação ao número total de imagens com a patologia.

$$Sensibilidade = \frac{VP}{(VP + FN)} \quad (7)$$

- **Pontuação F1** é a média harmônica entre precisão e sensibilidade, a expressão foca na análise de classes positivas, seu alto valor indica que o modelo performa melhor em classes positivas.

$$Pontuação F1 = 2 \times \left(\frac{Precisão \times Sensibilidade}{Precisão + Sensibilidade} \right) \quad (8)$$

Onde, nas Equações (5-7), VP, VN, FP e FN, representam a quantidade de amostras classificadas como verdadeira positiva, verdadeira negativa, falsa positiva e falsa negativa, respectivamente.

4 ANÁLISE E DISCUSSÃO DOS RESULTADOS

Os experimentos neste estudo foram conduzidos no Google Colaboratory (Colab) [34], serviço na nuvem de máquinas virtuais designadas para aprendizado de máquina [35]. Operando sob uma CPU Intel(R) Xeon(R) @ 2.00GHz e 13,6GB RAM com GPU Tesla V100 SXM2 contendo 16GB de VRAM. Os códigos foram implementados através da linguagem Python em conjunto com as bibliotecas: Tensorflow, Keras, Scikit-learn, Numpy, Pandas, Matplotlib e Seaborn.

4.1 TREINAMENTO E VALIDAÇÃO

Em ordem de treinar o classificador proposto nesta pesquisa para detecção de rachaduras em superfícies de concreto, treinamos a arquitetura desde o início com as amostras de imagens de treino e validação, não empregamos pesos de modelos pré-treinados através da transferência de aprendizado.

O procedimento de treinamento durou 30,45 minutos para treinar a arquitetura em 50 épocas, cada época levou em média 36,54 segundos para conduzir 137 iterações com a aceleração da GPU, enquanto o mesmo cálculo computacional na CPU foi estimado o tempo médio de 72,67 minutos por época. Monitoramos o cálculo da soma de erros nos dados de validação (perda de validação), adotando o critério de salvar o modelo com os pesos que obteve o melhor valor.

O resultado do treinamento e validação da rede neural sob a taxa de aprendizado da Tabela 2 alcançou a melhor solução na 50ª época, com um valor da perda de 16,19% e um valor da acurácia de 94,54%, conforme demonstrado na Figura 5.

Tabela 2 – Relação de taxa de aprendizado por época.

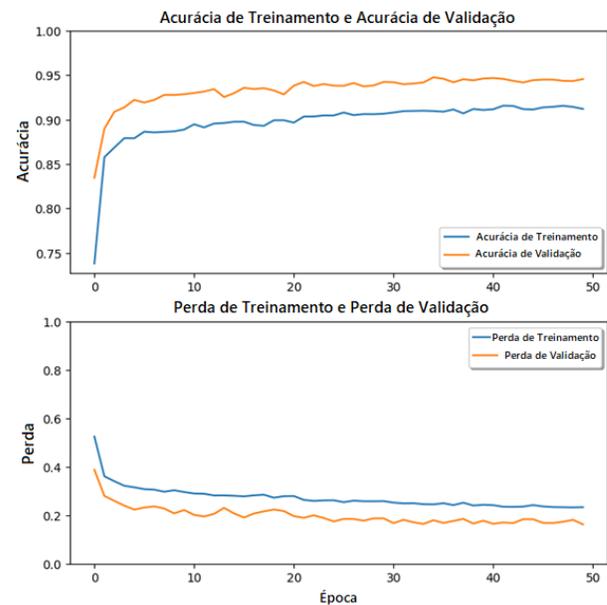
INTERVALO DE ÉPOCAS	TAXA DE APRENDIZADO
01 - 21	1×10^{-3}
22 - 30	5×10^{-4}
31 - 39	$2,5 \times 10^{-4}$
40 - 44	$1,25 \times 10^{-4}$
45 - 49	$6,25 \times 10^{-5}$
50	$3,125 \times 10^{-5}$

Fonte: Os autores.

Observa-se a ocorrência de oscilações durante o aprendizado, tendo como origem a divergência das condições das imagens no *dataset*, incluso a

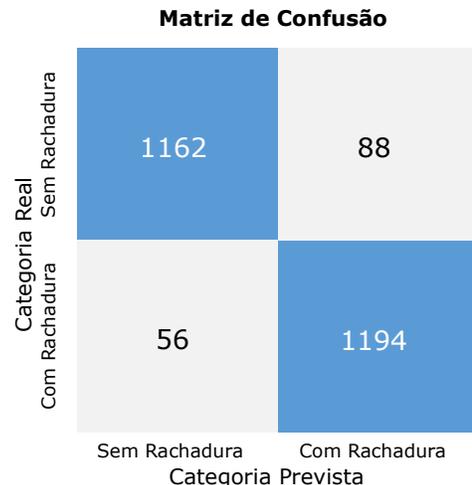
existência de amostras com informações gráficas que inviabilizam a classificação correta. Ademais, é importante evidenciar que a acurácia de validação está ligeiramente superior a acurácia de treinamento, assim como a perda de treinamento se manifesta acima de perda de validação. Uma das razões deste fenômeno se deve a camada de *dropout* descartar uma fração dos neurônios classificadores durante procedimento de treino, reduzindo os recursos que influenciam a análise dos dados. Por outro lado, *dropout* não opera durante o tempo de inferência, permitindo a livre circulação das informações para a previsão na validação.

Figura 5 – Histórico do treinamento e validação.



Fonte: Os autores.

Figura 6 – Matriz de confusão da classificação binária.



Fonte: Os autores.

4.2 AVALIAÇÃO DO MODELO

Tendo em vista medir a eficiência da rede neural proposta neste artigo, utilizamos as 2.500 imagens do subconjunto de teste, nunca vistas antes, com a finalidade de simular um ambiente experimental. A distribuição de classes nas imagens se mantém proporcional, tendo em vista que na ocasião de haver desequilíbrio das classes o resultado tende a ser favorável a categoria que apresentar a maior quantidade de dados, uma vez que existe mais informações sobre um determinado rótulo.

A matriz de confusão da Figura 6 demonstra a distribuição das previsões do modelo em relação às classes verdadeiras. O experimento resultou na classificação correta de 95,52% das superfícies com rachaduras e 92,56% de amostras sem patologia, classificando erroneamente apenas 149 das 2500 imagens. Calculamos as métricas de desempenho com base nos valores da matriz de confusão, a arquitetura conquistou em sua totalidade 94,04% de acurácia e 92,77% de precisão, conforme exposto na Tabela 3.

Tabela 3 - Resultado da avaliação da performance.

MÉTRICA	RESULTADO	RESULTADO (%)
ACURÁCIA	0,9404	94,04%
PRECISÃO	0,9277	92,77%
SENSIBILIDADE	0,9552	95,52%
PONTUAÇÃO F1	0,9413	94,13%

Fonte: Os autores.

4.3 VISUALIZAÇÃO DA CLASSIFICAÇÃO

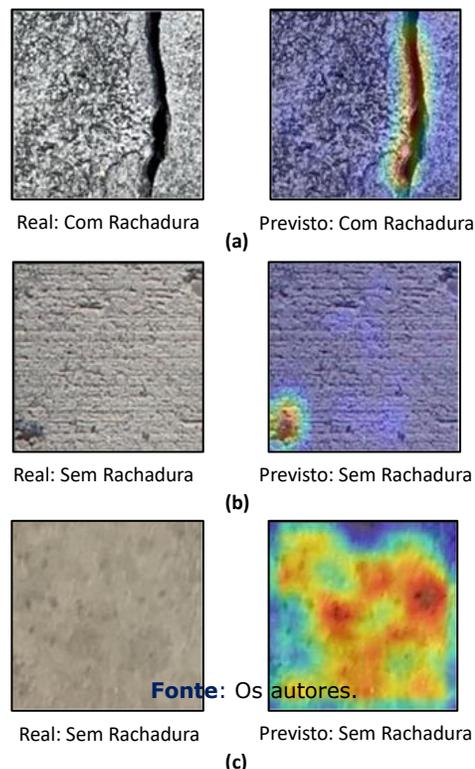
O processo de aprendizado das redes neurais artificiais profundas é comumente considerado uma caixa-preta, considerando-se o fato de que não existe uma simples explicação de seu funcionamento ou como o algoritmo determina as informações relevantes para realizar uma decisão. Diante dessa lacuna, o *Gradient-weighted Class Activation Mapping* (Grad-CAM) [36] nos permite desmistificar o funcionamento da CNN ao produzir um mapa de calor indicando visualmente a importância de cada parte da imagem para a classificação.

Como resultado, podemos visualizar de forma intuitiva se o nosso modelo foi capaz de aprender a prever corretamente levando em consideração os recursos relevantes de cada categoria e descobrir se a rede está focando na região de interesse correta.

Ao analisarmos o mapa de calor gerado a partir

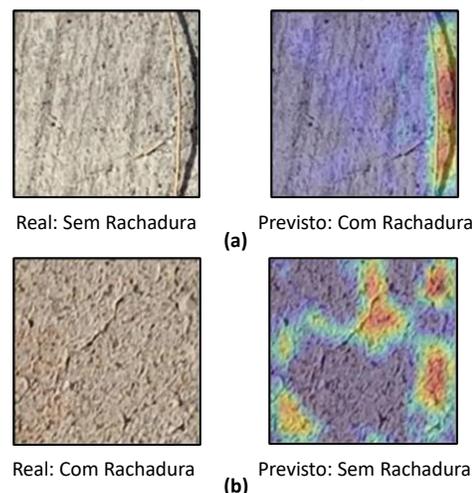
das previsões corretas da avaliação de teste (Figura 7), observamos que o modelo identifica áreas com aberturas em superfícies complexas, se

Figura 7 – Mapa de calor Grad-CAM para amostras classificadas corretamente na avaliação do modelo.



Fonte: Os autores.

Figura 8 – Mapa de calor Grad-CAM para amostras classificadas incorretamente na avaliação do modelo.



Fonte: Os autores.

baseando no formato da região localizada para classificar a possível existência de fissuras (a-b). No entanto, em imagens com superfícies planas (c) a arquitetura foca em identificar regiões com poucas variações e fronteiras visuais.

A Figura 8 ilustra as previsões incorretas de amostras de teste durante o experimento de avaliação. Observamos que superfícies complexas impactam diretamente as tomadas de decisões, a ocorrência de variação na iluminação da superfície confunde a área de foco do modelo, possibilitando a interpretação de sombras como evidências de classes, conforme o demonstrado na imagem (a). Superfícies granuladas com alta variação na textura dificultam a concentração da rede nos pixels que identificam as rachaduras de concreto, conforme o exemplo exposto da imagem (b).

5 CONCLUSÕES E TRABALHOS FUTUROS

Diante dos resultados adquiridos ao longo da pesquisa, apresentamos um novo algoritmo para identificação de patologias de concreto baseado em redes neurais convolucionais. O experimento foi realizado a partir da geração de um novo banco de dados por meio da mesclagem das amostras de outros três datasets. O desenvolvimento do modelo foi conduzido através de ensaios com variação dos fatores, como na quantidade e classe das camadas de processamento, verificando-se a configuração ideal dos hiper parâmetros, tais como a definição dos neurônios de saídas, dimensão dos filtros, taxa de aprendizado, número de épocas, tamanho de lote, otimizador, função de perda, complexidade da rede e outras variáveis. Desta forma, investigamos a correlação entre estes elementos e seu efeito na performance do modelo, ordenando o ajuste mais adequado para a tarefa.

O modelo proposto tem alta aplicabilidade na detecção de aberturas em superfícies de concreto, demonstrando boa performance e capacidade de generalização ao alcançar 94,04% de acurácia no subconjunto de teste.

O propósito principal deste estudo foi ofertar um método eficiente para a detecção de rachaduras de concreto, possibilitando o uso da rede neural em conjunto com outros dispositivos com finalidade de efetuar o levantamento do estado de conservação das estruturas de forma rápida e segura, poupando tempo em comparação a métodos de inspeção convencionais.

Apesar dos resultados expostos, é importante observar que o modelo sugerido se limita a tarefa de analisar uma única patologia, havendo a necessidade de expansão para detecção de múltiplas degradações. Além disto, outra grande limitação ocorre resultante do classificador de

imagens não poder delimitar visualmente a região da patologia analisada na superfície de concreto, impossibilitando segmentação da área afetada.

Portanto, em trabalhos futuros visamos ampliar o algoritmo com o intuito de classificar diversas patologias, juntamente com a capacidade de identificar visualmente as regiões danificadas. Vale ressaltar que não foi utilizado nenhum método capaz de obter ou analisar dimensões físicas, como comprimento, largura, altura e área. Desta forma, nos próximos estudos pretendemos aprimorar o modelo para capturar estes dados, possibilitando a percepção precisa de danos estruturais.

6 RECONHECIMENTOS

O presente trabalho foi realizado através do apoio da Pró-reitoria de Pesquisa, Pós-graduação e Inovação (PRPPI) do Instituto Federal de Alagoas (IFAL), por meio do Programa Institucional de Bolsa de Iniciação Científica (PIBIC) diante do edital de nº 13/2021/PRPPI/IFAL, sob concessão da bolsa de estudo/pesquisa do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq); projeto posteriormente contemplado por intermédio do edital de nº 16/2022/PRPPI/IFAL, com financiamento da PRPPI.

REFERÊNCIAS

- [1] LIMA, S. F. E. A. CONCRETO E SUAS INOVAÇÕES. **Caderno de Graduação-Ciências Exatas e Tecnológicas-UNIT-ALAGOAS**, 09 Maio 2014. 31-40.
- [2] BAUER, L. A. F. **Materiais de Construção Vol. 1**. 5. ed. [S.l.]: LTC, 2000.
- [3] IBAPE-SP, INSTITUTO BRASILEIRO DE AVALIAÇÕES E PERÍCIAS DE ENGENHARIA DE SÃO PAULO. **Inspeção Predial: a Saúde dos Edifícios**. 2. ed. São Paulo: [s.n.], 2015.
- [4] VILLANUEVA,. **A importância da manutenção preventiva para o bom desempenho da edificação**. Rio de Janeiro: Universidade Federal do Rio de Janeiro, 2015.
- [5] CREA-PB, C. R. D. E. E. A. D. P. Conselho Regional de Engenharia e Agronomia da Paraíba, 16 out. 2019. Disponível em: <<https://creapb.org.br/noticias/crea->

- pb-alerta-sobre-necessidade-de-manutencao-predial-periodica/>. Acesso em: 13 jun. 2022.
- [6] SILVA, L. K. D.; CABRAL, A. E. B. **LEVANTAMENTO DE MANIFESTAÇÕES PATOLÓGICAS EM ESTRUTURAS DE CONCRETO ARMADO NO ESTADO DO CEARÁ**. ENCONTRO NACIONAL DE TECNOLOGIA DO AMBIENTE CONSTRUÍDO. Maceió: [s.n.]. 2014. p. 1626-1634.
- [7] BILLIE F. SPENCER, V. H. Y. N. Advances in Computer Vision-Based Civil Infrastructure Inspection and Monitoring. **Engineering**, 2019. 199-222.
- [8] GUO , et al. Deep learning for visual understanding: A review. **Neurocomputing**, 26 Novembro 2015.
- [9] KHAN, A.; SOHAIL, A.; , Z. U. A survey of the recent architectures of deep convolutional neural networks. **Artificial Intelligence Review**, 21 Abril 2020. 5455-5516.
- [10] JAHANSHAHI, R. M.; AL., E. A survey and evaluation of promising approaches for automatic image-based defect detection of bridge structures. **Structure and Infrastructure Engineering**, 2009. 455-486.
- [11] AZOUZ, Z.; HONARVAR SHAKIBAEI ASLI, B.; KHAN, M. Evolution of Crack Analysis in Structures Using Image Processing Technique: A Review. **Electronics**, 12 Setembro 2023. 3862. Disponível em: <<https://www.mdpi.com/2079-9292/12/18/3862>>.
- [12] KUN, J. et al. A deep learning-based method for pixel-level crack detection on concrete bridges. **IET Image Process**, 16, n. 10, 6 Abril 2022. 2609–2622. Disponível em: <<https://ietresearch.onlinelibrary.wiley.com/doi/epdf/10.1049/ipr2.12512>>. Acesso em: 23 Janeiro 2024.
- [13] NGUYEN, S. D. et al. Deep Learning-Based Crack Detection: A Survey. **International Journal of Pavement Research and Technology**, 16, n. 4, 1 Julho 2023. 943–967. Disponível em: <<https://link.springer.com/article/10.1007/s42947-022-00172-z>>. Acesso em: 18 Janeiro 2024.
- [14] HU, G. X. . E. A. Pavement Crack Detection Method Based on Deep Learning Models. **Wireless Communications and Mobile Computing**, 2021, 17 Maio 2021. 1-13.
- [15] VOGADO, L. H. E. A. **Rede neural convolucional para o diagnóstico de leucemia**. Anais do XIX Simpósio Brasileiro de Computação Aplicada à Saúde. [S.l.]: [s.n.]. 2019. p. 46-57.
- [16] ALI, L.; AL., E. Performance Evaluation of Deep CNN-Based Crack Detection and Localization Techniques for Concrete Structures. **Sensors**, 21, n. 5, 01 Março 2021. n.p. Disponível em: <<https://www.mdpi.com/1424-8220/21/5/1688>>. Acesso em: 07 Fevereiro 2024.
- [17] CHA, Y.; CHOI, W.; BÜYÜKÖZTÜRK, O. Deep learning-based crack damage detection using convolutional neural networks. **Computer-Aided Civil and Infrastructure Engineering**, 32, n. 5, 24 Março 2017. 361-378.
- [18] ZHENG , H.; MOHAMMED , M. ; YANGE, L. Exploring the Detection Accuracy of Concrete Cracks Using. **Advances in Materials Science and Engineering**, 2021.
- [19] PAULY, L. E. A. **Detection, Deeper Networks for Pavement Crack**. Proceedings of the 34th ISARC. Taiwan: IAARC. 2017. p. 479-485.
- [20] GAO, Y. E. A. **Deep Residual Network with Transfer Learning for Image-based Structural Damage Recognition**. 11th National Conference on Earthquake Engineering. Los Angeles: [s.n.]. 2018. p. n.p.
- [21] SÁ, L. C. C. E. A. **Detecção de Rachaduras em Concreto em Imagens com o uso de Redes Neurais Convolucionais**. Anais do

- XVI Encontro Unificado de Computação do Piauí. Piauí: SBC. 2023. p. 49-56.
- [22] DINIZ, J. D. C. N.; AL., E. A Method for Detecting Pathologies in Concrete Structures Using Deep Neural Networks. **Applied Sciences**, 9, 07 Maio 2023. n.p. Disponível em: <<https://www.mdpi.com/2076-3417/13/9/5763>>. Acesso em: 19 fev. 2024.
- [23] KAO, S. P.; AL., E. Combining the YOLOv4 Deep Learning Model with UAV Imagery Processing Technology in the Extraction and Quantization of Cracks in Bridges. **Sensors**, 23, 25 fev. 2023. n.p. Disponível em: <<https://www.mdpi.com/1424-8220/23/5/2572>>.
- [24] GIL, A. C. **Como Elaborar Projetos de Pesquisa**. 4º. ed. São Paulo: Atlas, 2002.
- [25] KUHN, M.; JOHNSON,. **Applied Predictive Modeling**. 1. ed. New York: Springer New York, NY, v. 1, 2013.
- [26] XU, H.; WANG, Y. Automatic Bridge Crack Detection Using a Convolutional Neural Network. **Applied Sciences**, 18 Julho 2019. n.p.
- [27] DORAFSHAN, ; THOMAS, R. ; MAGUIRE,. SDNET2018: An annotated image dataset for non-contact concrete crack detection using deep convolutional neural networks. **Data in brief**, 21, 2018. 1664-1668. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2352340918314082>>.
- [28] ÖZGENEL, Ç. F. Concrete Crack Images for Classification. **Mendeley Data**, 23 Julho 2019. Disponível em: <<https://data.mendeley.com/dataset/s/5y9wdsg2zt>>. Acesso em: 01 jul. 2021.
- [29] IOFFE, S.; SZEGEDY, C. **Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift**. International conference on machine learning. Lille: [s.n.]. 2015. p. 448-456.
- [30] RAMACHANDRAN, PRAJIT ; ZOPH, BARRET ; LE, QUOC V. SEARCHING FOR ACTIVATION FUNCTIONS. **arXiv**, 27 Outubro 2017. 1-13.
- [31] HU, J.; SHEN, L.; SUN, G. **Squeeze-and-excitation networks**. Proceedings of the IEEE conference on computer vision and pattern recognition. [S.l.]: [s.n.]. 2018. p. 7132-7141.
- [32] JINDAL, ; SARDANA, ; MEHTA,. Analyzing Performance of Deep Learning Techniques for Web Navigation Prediction. **Procedia Computer Science**, v. 167, p. 1739-1748, 2020.
- [33] KINGMA, ; BA, L. **ADAM: A METHOD FOR STOCHASTIC OPTIMIZATION**. ICLR. [S.l.]: [s.n.]. 2015. p. n.p.
- [34] GOOGLE RESEARCH. Colaboratory. **Colaboratory**. Disponível em: <<https://research.google.com/colaboratory/intl/pt-BR/faq.html>>. Acesso em: 18 maio 2023.
- [35] CARNEIRO, T. et al. Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications. **IEEE Access**, 8 out. 2018.
- [36] SELVARAJU, R. R. et al. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. **International Journal of Computer Vision**, 128, 11 out. 2019. 336-359.
- [37] HAFIZ SULIMAN MUNAWAR, F. U. A. H. Inspecting Buildings Using Drones and Computer Vision: A Machine Learning Approach to Detect Cracks and Damages. **Drones**, 24 Dezembro 2021.
- [38] GONÇALVES, A. B. **ESTUDO DE PATOLOGIAS E SUAS CAUSAS NAS ESTRUTURAS DE CONCRETO ARMADO DE OBRAS DE EDIFICAÇÕES**. Rio de Janeiro: UFRJ / ESCOLA POLITÉCNICA, 2015.
- [39] FERREIRA, ; LOBÃO, W. N. MANIFESTAÇÕES PATOLÓGICAS NA CONSTRUÇÃO CIVIL. **Cadernos de Graduação: Ciências Exatas e Tecnológicas**, Outubro 2018. 71-80.

- [40] ULIANA, J. G. et al. **ESTUDO DA INCIDÊNCIA DE MANIFESTAÇÕES PATOLÓGICAS EM FACHADAS DE EDIFÍCIOS LOCALIZADOS EM REGIÃO LITORÂNEA.** 1º CONGRESSO BRASILEIRO DE PATOLOGIAS DA CONSTRUÇÃO. Foz do Iguaçu: [s.n.]. 2014.
- [41] SIMONYAN, K.; ZISSERMAN, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. **arXiv**, 4 Setembro 2014. Disponível em: <<https://arxiv.org/abs/1409.1556>>.
- [42] GU, J. et al. Recent Advances in Convolutional Neural Networks. **Pattern Recognition**, v. 77, p. 354-377, Maio 2018. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0031320317304120>>.
- [43] STUTZ, D. **Understanding Convolutional Neural Networks.** [S.l.]: [s.n.]. 2014.
- [44] RUIZ, R. D. B. E. A. Processamento digital de imagens para detecção automática de fissuras em revestimentos cerâmicos de edifícios. **Ambiente Construído**, Porto Alegre, 21, n. 1, Março 2021. 139-147.
- [45] PANTA, M. E. A. IterLUNet: Deep Learning Architecture for Pixel-Wise Crack Detection in Levee Systems. **IEEE Access**, 11, 8 Fevereiro 2023. 12249-12262. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/10035954>>. Acesso em: 29 Janeiro 2024.
- [46] YOKOYAMA, S.; MATSUMOTO, T. Development of an automatic detector of cracks in concrete using machine learning. **Procedia engineering**, 171, 2017. 1250-1255.
- [47] DORAFSHAN, S.; THOMAS, R. J.; MAGUIRE, M. Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete.. **Construction and Building Materials**, 03 Agosto 2018. 1031-1045.
- [48] WANG, Z. et al. CNN Training with Twenty Samples for Crack Detection via Data Augmentation. **Sensors**, 27 Agosto 2020. n.p.
- [49] PRASHANT KUMAR, S. B. . N. S. S. . S. R. K. Real-Time Concrete Damage Detection Using Deep Learning for High Rise Structures. **IEEE Access**, 9, 5 Agosto 2021. 112312-112331.
- [50] NAIR, V.; HINTON, G. E. **Rectified Linear Units Improve Restricted Boltzmann Machines.** Proceedings of the 27th international conference on machine learning. [S.l.]: [s.n.]. 2010. p. 807-814.
- [51] LI, S.; ZHAO, X. Image-Based Concrete Crack Detection Using Convolutional Neural Network and Exhaustive Search Technique. **Advances in Civil Engineering**, 2019, 30 Abril 2019. n.p.