

Aplicação de Aprendizado Profundo na Classificação de Imagens de Patologias da Construção Civil: Análise de *Data Augmentation* para Bancos de Dados Desbalanceados

Laís Cardoso de Medeiros
RAI - Robotics and Artificial Intelligence
Centro de Ciências Exatas e Tecnológicas
Universidade Federal do Recôncavo da Bahia
Cruz das Almas, Brasil
laisdemedeiros@aluno.ufrb.edu.br

André Luiz Carvalho Ottoni
RAI - Robotics and Artificial Intelligence
Centro de Ciências Exatas e Tecnológicas
Universidade Federal do Recôncavo da Bahia
Cruz das Almas, Brasil
andre.ottoni@ufrb.edu.br

Marcela Silva Novo
Departamento de Engenharia Elétrica e Computação
Universidade Federal da Bahia
Salvador, Brasil
marcela.novo@ufba.br

Resumo — O desafio das inspeções inteligentes na construção civil tem como referência a discriminação de imagens baseada em reconhecimento de padrões. Nesse cenário, a identificação é executada por arquiteturas de aprendizado profundo como as Redes Neurais Convolucionais, através de um grande conjunto de imagens. No entanto, a captura tradicional de fotografias em alta proporção é lenta, custosa e acompanhada de erros humanos, portanto inviável em parte das aplicações. Nesse sentido, o presente trabalho tem como objetivo realizar a classificação de rachaduras em bancos desbalanceados, através de técnicas de *data augmentation* (DA). A metodologia adotada propõe a geração de imagens artificiais a partir de transformações de imagens. Em seguida, são analisados os experimentos de validação, treinamento e testes no *software* Edge Impulse. Após os experimentos baseados nas métricas de acurácia, *F1-score*, precisão e *recall*, os resultados avaliaram que os bancos compostos por balanceamento atingiram os melhores índices de classificação, em comparação ao banco desbalanceado. Além disso, a transformação “Correção Gamma” se destacou com o maior percentual de acertos para rachaduras de 97,90%, tendo também sua acurácia em 96,17%. Ante o exposto, a contribuição do presente trabalho compõe o aperfeiçoamento da pesquisa em balanceamento por DA aplicado a classificação de imagens de rachaduras.

Palavras-Chave — Banco de dados desbalanceados, Classificação de rachaduras, Data Augmentation, Edge Impulse, Redes Neurais Convolucionais.

I. INTRODUÇÃO

O desenvolvimento dos computadores propiciou a execução de tarefas com alto potencial de escalabilidade [3] [10]. Entre os avanços tecnológicos produzidos, encontra-se a ciência das máquinas inteligentes desenvolvidas primordialmente pelo matemático inglês Alan Turing [10], [23], [58]. O conceito intitulado como inteligência artificial (IA) define a habilidade de tomada de decisão por agentes robóticos [26], [45], [74]. Dessa forma, para atingir a automatização de tarefas são implementadas as estratégias de aprendizado de máquina, em inglês, *Machine Learning* (ML) [74]. Nesse contexto, com diversificada aplicabilidade social, o conceito auxilia na produção em setores como o direito [40], medicina [30], [53], administração pública [13], educação [52], mídias sociais [32], estética [32], [61], classificação imobiliária [20], segurança do trabalho [71] e na construção civil [7].

Entre as aplicações citadas, as principais áreas técnicas utilizadas são a visão computacional e a robótica móvel [58], [51], [19]. Nesse contexto, surge a necessidade do aperfeiçoamento da técnica através do aprendizado profundo, em inglês, *Deep Learning* (DL) [33], [42]. Esse modelo de aprendizado é um mecanismo capaz de efetuar detecção, classificação e reconhecimento de objetos como as detecções de incêndios florestais [1], alimentação de subestação de energia [41], acompanhamento de medicamentos ao paciente [56] e monitoramento estrutural na construção civil [60], [72] por exemplo [42]. Entre os modelos de DL que se destacam, as Redes Neurais Convolucionais, em inglês, *Convolutional Neural Networks* (CNNs) que utilizam grandes conjunto de dados, semelhantes a *Big Data* como o ImageNet, para realizar um

alto reconhecimento de padrões em imagens e a classificação discriminatória [63], [27], [37], [51].

No entanto, apesar da ampla aplicação em segmentos como danos estruturais [67], otimização na utilização de materiais [36], custo total de produção [38], índice de produtividade da construção [16] e segurança na construção de barragens [31], este tipo de treinamento das CNNs em sua melhor *performance* possui desafios para a aplicação na construção civil [43], [38], [44], [57]. Isso porque, a coleta de imagens no formato tradicional é efetuada de forma manual, dificultando assim a alta escalabilidade de fotografias necessárias [46]. Nesse sentido, a construção de uma grande base de dados torna-se custosa em tempo processamento, captura das fotografias e o potencial de armazenamento do dispositivo, além de causar erros como invariáveis no padrão de qualidade das imagens causado por falhas humanas [46], [12]. Diante das consequências apontadas, uma possível solução, é a técnica de aumento de dados, em inglês, *data augmentation*, método com capacidade de gerar imagens artificiais similares as originais configurando a quantificação e diversidade dos dados [50], [70], [69], [76]. Desse modo, com aplicações em reconhecimento de equipamentos [54], avaliação pós desastre [63] e inspeção de construções [34], [68], a técnica apresenta variedades de padrões desafiadores para as CNNs, evitando assim, casos de *overfitting*, generalização de dados e regularizações indevidas, ou seja, falhas na classificação [47], [64], [63], [65].

Sendo assim, o principal objetivo deste trabalho é classificar patologias na construção civil aplicando sistemas de aprendizado profundo baseados em validação, treinamento e testes a partir de bancos de dados desbalanceados. De maneira específica, propor algoritmo de geração de banco de dados desbalanceados aplicando técnicas de *data augmentation* a partir da utilização de filtros de transformação de imagem para geração de imagens artificiais. Além disso, visa recomendar tipos de transformações geométrica e de cor aplicado a geração de um banco de dados similar ao original para realização do treinamento do modelo de aprendizado profundo.

O presente artigo está organizado em cinco seções. Na Seção II, são apresentados os conceitos e fundamentos teóricos base para o desenvolvimento do trabalho. Na sequência, a Seção III descreve a metodologia dos experimentos produzidos. Em seguida, a Seção IV apresenta os resultados analisados nas etapas de validação, treinamento e teste em cada experimento. Por fim, na Seção V são apresentadas as contribuições construídas e propostas de trabalhos futuros baseados nos experimentos realizados.

II. FUNDAMENTAÇÃO TEÓRICA

O ramo das Redes Neurais Artificiais (RNAs) se destaca entre as aplicações de aprendizado de máquina em virtude da alta precisão obtida no reconhecimento de padrões [2], [43], [55]. Diante disso, os seus resultados são a consequência da arquitetura de aprendizagem e processamento representada pela Figura 1 baseado na similaridade com a estrutura de neurônios biológicos [64], [15].

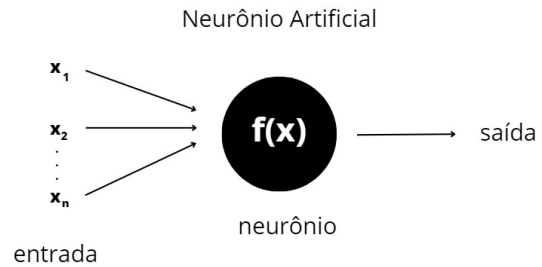


Figura 1: Estrutura do neurônio artificial.

Fonte: Baseado em [21].

Nessa perspectiva, há analogias entre a função dos dendritos e as entradas, a função de atividade do neurônio e o corpo celular, e por fim, a função das sinapses com as saídas [22]. A apontada definição das RNAs tinha como pressuposto inicial a existência de apenas um neurônio, o Perceptron, modelo desenvolvido por Franklin Rosenblatt (1958), sendo considerado o primeiro neurônio artificial de aprendizagem. Todavia, houve no decorrer na história o desenvolvimento deste conceito, ocorrendo o aperfeiçoamento da técnica de reconhecimento de padrões, emergindo no âmbito das CNNs [64], [22], [39].

O surgimento das CNNs apresenta no desenvolvimento do DL um alto potencial na extração de características de imagens processadas [29], [11]. A mencionada extração ocorre a partir dos filtros convolucionais, em inglês, *feature detector*, ou *kernels* composto pela multiplicação entre *pixels* da imagem (z) e os pesos sinápticos (w) representados na Equação 1 [43], [35], [42]. Importa salientar que este procedimento acontece na camada convolucional responsável pela aplicação da convolução, gerando novas matrizes e construindo o mapeamento de características, em inglês, *feature map*, de cada imagem [22], [55].

$$R = \sum_{i=1}^9 w_i \cdot z_i \quad (1)$$

III. METODOLOGIA

Esta seção descreve a metodologia para a execução dos experimentos, fragmentadas em: (1) Materiais de *Hardware* e *Software*, (2) Arquitetura da Rede Neural Convolucional, (3) Base de dados e Transformações de imagens, (4) Planejamento dos experimentos, (5) Validação, treinamento e testes.

A. Materiais de Hardware e Software

Para a execução do experimento foi utilizada a linguagem R na versão 4.2.2 e o ambiente de desenvolvimento integrado (IDE) RStudio na versão 2022.12.0+353 [3]. A preferência para escolha desta linguagem de programação se deu pela especialidade em projetos estatísticos e visualização gráfica de transformações de imagens, como na biblioteca "EBImage" aplicada na versão 4.42.0 [4], [73].

Outra tecnologia essencial para o processo de composição dos experimentos, foi a plataforma de Inteligência Artificial e Aprendizado de Máquina, utilizada em sua versão gratuita, Edge impulse¹ [62]. O *Software* é responsável por efetuar etapas de treinamento e teste, com grande detalhamento de informações, permitindo interpretações em diferentes perspectivas [62]. Nesse sistema, os processos executados utilizaram a arquitetura neural MobileNet [24]. Além disso, para execução do experimento foi utilizado o sistema operacional *Microsoft Windows 10* de *x64 bits* em uma máquina Intel(R) Core(TM) i3-7020U e 8GB de memória RAM. Ademais, o *hardware* conta com placa integrada Intel Graphics.

B. Arquitetura da Rede Neural Convolutional

Nesse trabalho, foi adotada a Rede Neural Convolutional MobileNet. Essa arquitetura está disponível na plataforma Edge Impulse com pesos pré-treinados no *dataset* ImageNet. [59]. A ferramenta possui estrutura baseada em convoluções fatoriais, além de 28 camadas e 3.732.2893 parâmetros. No contexto da classificação de imagens e visão computacional, o modelo se destaca por sua capacidade de atingir resultados de classificação satisfatórios, com poder de processamento computacional limitado e execução em menor tempo [43].

Este fato ocorre devido a arquitetura efetuar através de convoluções separáveis em profundidade, ou seja, nas primeiras camadas acontece a queda do volume das restrições, reduzindo o processamento computacional e armazenamento. Assim, objetiva-se atingir a rápida execução com menor custo de armazenamento [8]. Dessa forma, o modelo descrito se configura adequado para o presente projeto, visto que, a aplicação web Edge impulse oferece de forma gratuita, a realização de experimentos com tempo de tratamento de dados em até vinte minutos.

C. Base de dados e transformação de Imagens

O experimento foi realizado a partir da base de dados [14] devido a sua composição por imagens binárias, como mostra a Figura 2, devido a menor exigência de armazenamento e tempo de processamento [9]. Em resumo, ao realizar o *download* do banco de imagens, são carregadas 40 mil imagens com ocupando no total 230MB de memória, em dois diretórios. Dessa forma, o banco de fotografias possui a pasta com nome *"Positive"* que contém 20 mil imagens da classe de rachaduras e 20 mil imagens de paredes na pasta *"Negative"*. Além disso, são disponibilizados o documento do artigos construído a partir da composição do banco de dados. É importante citar que o conjunto de imagens foi escolhido devido aos bons resultados de classificação apresentados e o grande volume de imagens, representando assim elementos que contribuem para a identificação do rótulo de rachaduras e menor incidência de generalização dos treinamentos [14].

¹<https://www.edgeimpulse.com/>



Figura 2: Imagem de rachadura.

Fonte: Imagem do banco de dados [14].

Em seguida, foram aplicados os filtros da biblioteca "EBImage" em imagens selecionadas, apresentadas na Figura 3. Nesse momento, o objetivo principal é observar quais as transformações geométricas e de cor aplicadas preservam e destacam os rótulos de cada imagem, visando desenvolver desafios na aprendizagem do reconhecimento de padrões para a arquitetura durante o treinamento com alta relevância para a identificação das rachaduras [63], [42].

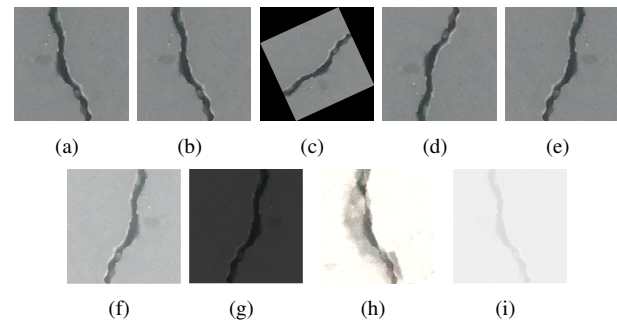


Figura 3: Transformações de imagem: base [14].

(a) Foto original; (b) *Resize*; (c) *Rotate*; (d) *Flip*; (e) *Flop*; (f) *Brilho*; (g) *Contraste*; (h) *Combinação*; (i) *Correção Gamma*; Fonte: Rstudio.

Ao final do experimento foram selecionados uma quantidade limitada de filtros destinados para a execução dos experimentos devido ao custo computacional dos processos. Dessa forma, para o campo das transformações de geométricas foram utilizados os filtros de "Rotação", "Flip/Flop" e "Rotação/FlipFlop", enquanto para as transformações de cor a "Correção Gamma", "Brilho/Contraste" e "Brilho/Contraste/CombinaçãoGamma". Por fim, também para reduzir consumo de processamento, as figuras serão redimensionadas para tamanhos menores através do parâmetro *resize*.

D. Algoritmo DAGenerator

Nesta seção, é apresentado o escopo do algoritmo proposto para geração de imagens artificiais. Nesse contexto, desenvolvido na linguagem R e em sua *Integrated Development Environment* (IDE), o Rstudio, o código do DAGenerator apresentado nas Figuras 4 e 5 tem como objetivo gerar novas imagens, a partir de transformações geométricas e de cor. Isto

é, construir um grande banco de imagens a partir de uma pequena quantidade de fotografias originais.

A primeira parte da descrição do algoritmo apontado na Figura 4, inicia nas linhas 1 e 2 com a leitura da biblioteca e a listagem de imagens na pasta do computador em processamento. Em seguida, a linha 4 até 16 representa um comando de repetição "for", responsável por gerar as imagens originais em menor resolução, diminuindo assim, o custo computacional, porém sem transformações aparentes nas imagens.

```

1 library(EBImage)
2 nomes=list.files(pattern = ".jpg")
3
4 for(i in 1:length(nomes)){
5   im=readImage(nomes[i])
6   im2 = resize(im, 500) #6 MB
7   plot(im2)
8
9   randomValue= runif(1, min=0, max=1000);
10
11  png(paste("transform/rotate/rotate",randomValue,nomes[i],".png"))
12
13  plot(im2)
14  dev.off()
15
16 }
17

```

Figura 4: Primeira etapa - Algoritmo DAgenerator.

Fonte: Rstudio.

Sendo assim, o código segue com a segunda etapa exibida na Figura 5, definindo na linha 18 a variável "TOTAL" responsável por indicar a quantidade final de imagens geradas, conforme a Equação 2

$$TOTAL = nImage * fotografias \quad (2)$$

em que, "nImage" é o valor que ao ser multiplicado com o número de fotografias produz a quantidade total desejada e "fotografias" é a quantidade de fotografias originais.

A partir da linha 20 a 37, são implementados dois comandos de repetição que implementam através de código a multiplicação indicada na Equação 2. Por fim, o código segue com os métodos para rotacionar a imagem. Vale ressaltar que este código, apenas realiza a transformação de imagem de rotação. Isto significa que para cada transformação, as linhas 26 e 27 eram preenchidas por códigos de cada transformação.

```

18 nImage = 99
19
20 for(j in 1:nImage){
21   for(i in 1:length(nomes)){
22     im=readImage(nomes[i])
23     im2 = resize(im, 500) #6 MB
24     plot(im2)
25
26     randomValueRotate= runif(1, min=0, max=360);
27     im3 = rotate(im2,randomValueRotate);
28
29     randomValue= runif(1, min=0, max=1000);
30
31     png(paste("transform/rotate/rotate",randomValue,nomes[i],".png"))
32
33     plot(im3)
34     dev.off()
35
36   }
37 }

```

Figura 5: Segunda etapa - Algoritmo DAgenerator.

Fonte: Rstudio.

E. Planejamento dos experimentos

Nessa etapa do experimento, são planejados os formatos dos três tipos de bancos: (i) Banco original, (ii) Banco desbalanceado e (iii) Banco com *Data Augmentation*. Em primeiro plano, o banco original é composto por 40 mil imagens originais, divididas em 20 mil para as rachaduras, sendo 16 mil para treinamento e 4 mil para teste, seguindo o mesmo padrão para a classe parede. Em seguida para o banco desbalanceado, são selecionadas para o treinamento 160 imagens originais de rachaduras e 16 mil imagens originais de paredes, enquanto para o teste 4 mil imagens originais de rachaduras e paredes.

O experimento segue com a composição dos bancos balanceados por DA. Dessa forma, para as paredes 16 mil imagens de treinamento e 4 mil imagens de testes, sendo estas imagens originais. Todavia, para a classe de rachaduras são utilizadas as 160 imagens separadas no treinamento para o banco desbalanceado e a partir dessas imagens originais são geradas, pelo algoritmo "DAgenerator", as 15.840 imagens artificiais, totalizando 16 mil imagens de treinamento e 4 mil imagens originais para o teste. É importante ressaltar que, o 1% de imagens originais selecionadas tem como objetivo desafiar experimentalmente as possibilidades dos resultados que serão obtidos, ou seja, em caso de bons resultados demonstrar que mesmo no "pior caso" a técnica de DA pode ser uma solução interessante e robusta.

F. Validação, treinamento e testes

Nesta etapa do experimento são executadas as fases de treinamento, validação e testes realizado no *software* de aplicações em IA, o Edge impulse, tais fases foram realizadas em oito bancos no formato apresentados na subseção C. Em resumo, inicialmente no *software* foi realizado a configuração dos dados de entrada e saída, ou seja, imagens e parede/rachadura. Assim como, o *upload* das imagens com inserção direcionada para treinamento ou teste, conforme o formato planejado. Em seguida, definição de parâmetros como o tamanho das imagens e a escala de cor definida como RGB. Além disso, a configuração do treinamento da CNNs com a arquitetura neural MobileNetV1 96x96 0.25, o número de ciclos/épocas sendo 20 e o percentual de imagens de treinamento destinadas para validação definida em 20%. Sendo estas baseadas na análise tempo de processamento, devido ao *software* Edge impulse disponibilizar apenas 20 minutos de processamento gratuito. Por fim, são analisados os resultados da matriz de confusão para treinamento e testes de cada banco de dados.

Nesse contexto, para melhor interpretação dos resultados é apresentada a matriz de confusão de cada classificação de treinamento e testes, representada na Figura 6 e os parâmetros de avaliação métrica do classificador [6].

	PAREDE	RACHADURA
PAREDE	96.2%	3.8%
RACHADURA	1.3%	98.7%
F1 SCORE	0.97	0.98

Figura 6: Matriz de confusão.

Fonte: *Edge Impulse*.

Dessa forma, em todo o trabalho são determinados o padrão de rachadura como classe positiva e parede classe negativa em que, Verdadeiro negativo (VN) representa acerto para paredes, Falso positivo (FP) representa erro para paredes, Verdadeiro positivo (VP) representa acerto para rachaduras e Falso Negativo (FN) representa erro para rachaduras. Ademais, o parâmetro estatístico da acurácia representada através da Equação 3, indica a fração de previsões corretas sobre cada imagem e a definição de cada classe categorizada, como mostra a equação abaixo:

$$acuracia = \frac{VP + VN}{VP + FN + VN + FP} \quad (3)$$

No contexto das rachaduras, além da acurácia, outras métricas de avaliação complementam as interpretações. Entre elas, a precisão que possui melhor valor quando há a diminuição dos falsos positivos, logo para o experimento entre paredes e rachaduras, são apontadas rachaduras quando na verdade são paredes, definida então pela Equação 4:

$$precisao = \frac{VP}{VP + FP} \quad (4)$$

Outra métrica importante é o *recall* que tem melhores valores quando são diminuídos falsos negativos, isto porque, contém grande número de indicações de parede quando na verdade são rachaduras, esse parâmetro apresenta relevância para o presente trabalho em bancos balanceados, a Equação 5 é representada por:

$$recall = \frac{VP}{VP + FN} \quad (5)$$

Somando a precisão e o *recall*, a Equação 6 nomeado como "F1Score", representa a média harmônica entre as Equações 4 e 5 determinado por,

$$F1Score = \frac{2 * precisao * recall}{precisao + recall} \quad (6)$$

IV. RESULTADOS

A seção apresenta os principais resultados obtidos durante os experimentos. Sendo estes, os (1) Resultados Gerais e de forma detalhada os (2) Resultados para banco de dados desbalanceados e OS (3) Resultados para banco de dados com Correção Gamma, apresentando assim a importância do balanceamento por DA no reconhecimento de padrões das CNNs.

A. Resultados Gerais

Para iniciar as discussões sobre os resultados apresentados são destacados os Verdadeiros Positivos (VP) que representam o percentual de acertos da classe rachaduras em cada classificação. Durante a análise, esse percentual é muito importante, porque na situação de classificação de rachaduras em uma construção mesmo com um menor percentual de acurácia, é necessário que as rachaduras sejam indicadas. Afinal, entre os objetivos da classificação estão indicar precisamente o rótulo das rachaduras.

Dessa forma, o percentual de acertos ou VP para os testes são apresentados na Tabela I. Nesse contexto, o banco original atinge 97,60% de acurácia e no banco desbalanceado a taxa tem valor de apenas 76,20%. O resultado indica a importância de um maior conjunto de dados para treinamento, impactando diretamente a classificação de testes. Por outro lado, os bancos balanceados por DA apresentaram a maior sendo 97,90% para a "Correção Gamma". É importante ressaltar, que o maior percentual dos verdadeiros positivos para banco balanceado por DA foi superior ao banco original, indicando o destaque ao rótulo das fendas realizado através dos filtros de transformações de imagens.

Tabela I: Resultado geral: Verdadeiros positivos para testes.

BANCO	ACURÁCIA
Correção Gamma	97,90%
Balanceado	97,60%
Brilho/Contraste/CorreçãoGamma	96,60%
Brilho/Contraste	92,80%
Rotate	88,30%
FlipFlop	87,20%
Rotate/Flip/Flop	86,80%
Desbalanceado	76,20%

Conforme os resultados de acurácia expostos nas Tabelas II e a III, é verificado que altos valores de acurácia na validação do treinamento em alguns casos indicam resultados de generalização. Dessa forma, ao comparar as tabelas, a análise verifica uma diferença entre os valores de treinamento e teste, isto porque somente a análise do treinamento não determina a avaliação do classificador. Portanto, a Tabela III representa a sequência de resultados de teste, que apontam no conjunto de fotografias originais a melhor percentagem de teste, 97,78%, enquanto o banco desbalanceado apresenta a pior percentagem de acurácia de testes, 87,95%. Esse resultado ratifica as afirmações do Seção II, sobre a discrepante perda de precisão ao diminuir a quantidade de informações para treinamento. Nesse caso se aproximando de 10% de diferença, entre o banco original e o banco desbalanceado.

Além disso, os bancos balanceados por DA demonstraram os valores da acurácia de classificação durante os testes foram superiores ao banco desbalanceado. Dessa forma, validando a hipótese das gerações das imagens artificiais para o balanceamento. Contudo, entre os filtros utilizados a transformação de cor "Correção Gamma" obteve valor de acurácia em quase 2% menor que a acurácia das fotografias originais. Diante desse fato, fica claro como o estilo da transformação "Correção

Gamma”ressalta a rachadura no fundo da imagem produzindo uma maior possibilidade de reconhecimento do rótulo de rachadura.

Tabela II: Resultado geral: Treinamento e validação.

BANCO	ACURÁCIA
FlipFlop	99,90%
Rotate/Flip/Flop	99,90%
Desbalanceado	99,50%
Balanceado	97,50%
Correção Gamma	97,50%
Rotate	97,50%
Brilho/Contraste	97,50%
Brilho/Contraste/CorreçãoGamma	93,50%

Tabela III: Resultado geral: Teste.

	BANCO	ACURÁCIA
1°	Balanceado	97,78%
2°	Correção Gamma	96,17%
3°	Brilho/Contraste	95,54%
4°	Brilho/Contraste/CorreçãoGamma	95,52%
5°	Rotate	93,97%
6°	Rotate/Flip/Flop	93,20%
7°	FlipFlop	93,12%
8°	Desbalanceado	87,95%

B. Resultados para banco de dados desbalanceados

A Tabela IV evidencia os resultados de treinamento para bancos de dados desbalanceados, ou seja, 16 mil imagens para a classe parede e com apenas 160 para a classe de rachadura. Na verdade, reconhecidos pela aplicação, 22.870 imagens no total de imagens, esse valor total é o menor entre as tabelas devido ao desbalanceamento. Portanto, para o treinamento 14.929 fotografias para treinamento, em que 14.769 para parede e 160 para rachaduras. Nesse quadro, a Tabela IV expõe como o desbalanceamento entre as classes reduz o potencial de treinamento das CNN.

Isto porque, a tabela IV apresenta no treinamento sendo altos valores de classificação da classe parede 99,9% (2.948) de VN e 0,1% (2) de FP, no entanto baixos valores de classificação de rachaduras 36,1% (13) de FN e 63,9% (23) de VP para rachaduras. Concluindo com a métrica de *F1-Score* de 100% para as paredes, acompanhado dos mesmos valores de *recall* e precisão, e somente 75% para rachaduras para o *F1-score*, composto por um valor ainda mais baixo de 65% de *recall* e precisão de 92%, demonstrando uma possível generalização no treinamento.

Tabela IV: Resultados do treinamento e validação: Banco desbalanceado.

	PAREDE	RACHADURA
PAREDE	99,9%	0,1%
RACHADURA	36,1%	63,9%
F1 SCORE	100,0%	75,0%

O treinamento de testes foi realizado com bancos de 4 mil fotografias em cada classe. Nesse sentido, são reconhecidos

7.941, sendo 3.991 para rachaduras e 3.950 para parede. Nesse caso, a partir da Tabela V que apresenta, diferente da etapa de treinamento, uma acurácia de 87,95% para os testes, valor quase 10% menor que o apresentado nos bancos de dados original.

Em outros parâmetros, 99,8% (3.944) de VN e 0,1% (3) de FP para imagens de parede, e 18,8% (750) de FN e 76,2% (3.040) de VP para as rachaduras. Na métrica do *F1-Score*, para rachadura 86%, seguido de 100% de precisão e 86% de *recall*, enquanto para a parede o *F1-score* é de 0,91, associado de 100% de *recall* e 84% de precisão. Concluindo de fato, a previsão feita durante o treinamento de que nesta etapa ocorre a generalização, o *overfitting*.

Tabela V: Resultados de teste: Banco desbalanceado.

	PAREDE	RACHADURA
PAREDE	99,8%	0,1%
RACHADURA	18,8%	76,2%
F1 SCORE	91,0%	86,0%

C. Resultados para banco de dados com Correção Gamma

Em seguida, são apresentados na Tabela VI os valores da validação do treinamento para o conjunto de dados balanceado com a transformação de cor ”Correção Gamma”. Este experimento é composto por 38.712 imagens, destinadas a 20% de 30.769 para validação de treinamento, em que 16.000 são de rachaduras e 14.679 são de paredes. Esses dados possuem acurácia na validação de 97,4% e perda de 0,11. Além disso, para a matriz de confusão, para este parâmetro a classe parede contém 96,8% (2.846) de VN e 3,2% (94) de FP, enquanto para a classe rachadura 97,9% (3145) de VP e somente 2,1% (69) de FN. Para ambos rótulos, o *F1-Score* possui a taxa de 97%, em que nas rachaduras a precisão é de 97% e *recall* de 98%, no caso das paredes esses valores são trocados. Sendo a classe rachadura com o maior valor de *recall*.

Tabela VI: Resultados do treinamento e validação: Correção Gamma.

	PAREDE	RACHADURA
PAREDE	96,8%	3,2%
RACHADURA	2,1%	97,9%
F1 SCORE	97,0%	97,0%

Nessa perspectiva, os resultados de testes exibidos na Tabela VII, foram construídos por 7.943 imagens, com 3.993 rachaduras e 3.950 paredes. Com resultado relevante, a acurácia de 96,12%, representa a maior entre os bancos balanceados com DA. Para a matriz de confusão, na classe parede possui VN de 98,2% (3.879) e um baixo FP de somente 1,2% (48). Da mesma forma, as rachaduras com 94,2% (3756) de VP e 4,6% (135) de FN. Portanto, em métricas detalhadas o *F1-score* as paredes contém 97%, com 95% de precisão e 98% de *recall*. Dessa forma para a classe rachadura ,a pontuação é de 96%, acompanhada de 99% de precisão e 94% *recall*. Concluindo assim, que as métricas indicam uma excelente classificação para os testes.

Tabela VII: Resultados de teste: Correção Gamma.

	PAREDE	RACHADURA
PAREDE	98,2%	1,2%
RACHADURA	4,6%	94,2%
F1 SCORE	97,0%	96,0%

V. CONCLUSÃO

Ante o exposto, o presente trabalho tem como objetivo principal a classificação de patologias de rachaduras aplicando sistemas de aprendizado profundo através de técnicas de DA. Dessa forma, com intuito de balancear o conjunto de fotografias são geradas imagens artificiais a partir de experimentos executados com transformações geométricas e de cor. Nesse sentido, baseado nos resultados de testes são recomendadas as melhores transformações para a classe rachadura. Em síntese, a fundamentação teórica revisou bibliograficamente conceitos que contemplam a construção do experimento de classificação. Em seguida, a metodologia descreveu a execução do experimento através das etapas de (1) Materiais de *Hardware* e *Software*, (2) Arquitetura da Rede Neural Convolutiva, (3) Base de dados e Transformações de imagens, (4) Planejamento dos experimentos e (5) Validação, treinamento e testes.

Diante da metodologia aplicada os resultados apontaram relevantes dados de análise da acurácia de teste, que alcançou em todos banco de dados balanceados por DA percentuais superiores quando comparados com o banco desbalanceado. Dessa forma, tem-se que o balanceamento por DA produz um percentual de acurácia entre 5,0% até aproximadamente 10% maior para acertos quando comparados aos bancos desbalanceados. Dado o exposto, entre as transformações aplicadas, o filtro de "Correção Gamma" obteve o maior percentual de acertos ou VP para rachaduras. Ademais, ficou evidente que a transformação mencionada atingiu o melhor resultado entre as bases de dados artificiais 96,17% de acurácia, sendo este valor apenas 2,0% inferior a base de dados com maior resultado de acurácia, o banco de dados original, composto por 100% de fotografias originais.

Em síntese, os conhecimentos construídos durante os experimentos representam uma contribuição para o avanço nos estudos do balanceamento por DA e para as inspeções inteligentes na construção civil. Por fim, para a continuação do desenvolvimento nos estudos realizados são propostos a realização dos experimentos a partir da metodologia apresentada com as bases [12] e [75] e o desenvolvimento do aplicativo para inspeções inteligentes na construção civil, composto por geração de imagens artificiais para balanceamento de bases de dados e classificação de imagens de rachaduras.

AGRADECIMENTOS

Os autores agradecem ao apoio da FAPESB (Pedido N° 4153/2022), CAPES, CNPq, UFRB e UFBA.

REFERÊNCIAS

[1] ALVES, J. M. R. Detecção de incêndios florestais com recurso a Deep Learning e Visão Computacional. Tese (Doutorado) — Universidade Fernando Pessoa (Portugal), 2018.

[2] BALESTRIERO, R.; BOTTOU, L.; LECUN, Y. The effects of regularization and data augmentation are class dependent. arXiv preprint arXiv:2204.03632, 2022.

[3] Bates, Douglas . The R Project for Statistical Computing. 2023. Disponível em: <https://www.r-project.org/other-docs.html>. Acesso em: 12 de maio 2023.

[4] Carey, Vince. EBImage. 2023. Disponível em: <https://www.edgeimpulse.com/>. Acesso em: 12 de maio 2023.

[5] CASTLEMAN, K. R. Digital image processing. [S.l.]: Prentice Hall Press, 1996.

[6] CASTRO, L. N. de; FERRARI, D. G. Introdução à mineração de dados. [S.l.]: Saraiva Educação SA, 2017..

[7] CAVALCANTI, V. Y. S. de L. et al. Indústria 4.0: desafios e perspectivas na construção civil. Revista Campo do Saber, v. 4, n. 4, 2018.

[8] Chollet, François. Keras. 2023. Disponível em: <https://keras.io/>. Acesso em: 21 de maio 2023.

[9] CONCI, A.; AZEVEDO, E.; LETA, F. R. Computação gráfica—teoria e prática, [v. 2]. Rio de Janeiro, 2008.

[10] COPELAND, B. J. The modern history of computing. 2000

[11] CUN, Y. L. et al. Handwritten digit recognition: Applications of neural network chips and automatic learning. IEEE Communications Magazine, IEEE, v. 27, n. 11, p. 41–46, 1989.

[12] DAIS, D. et al. Automatic crack classification and segmentation on masonry surfaces using convolutional neural networks and transfer learning. Automation in Construction, Elsevier, v. 125, p. 103606, 2021.

[13] DESORDI, D.; BONA, C. D. A inteligência artificial e a eficiência na administração pública. Revista de Direito, Universidade Federal Viçosa, v. 12, n. 2, p. 1–22, 2020.

[14] 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, Elsevier, v. 186, p. 1031–1045, 2018.

[15] ELGENDY, M. Deep learning for vision systems. [S.l.]: Simon and Schuster, 2020.

[16] ERPEN, M. L. et al. Análise de sucesso na indústria da construção civil utilizando redes neurais. [S.l.]: Bookerfield Editora, 2021.

[17] FORSYTH, D. A.; PONCE, J. Computer vision: a modern approach. [S.l.]: prentice hall professional technical reference, 2002.

[18] GOMES, J.; VELHO, L. Computação gráfica: imagem. [S.l.]: IMPA, 2002.

[19] GONZALES, R. C.; WINTZ, P. Digital image processing. [S.l.]: Addison-Wesley Longman Publishing Co., Inc., 1987.

[20] GONZÁLEZ, M. A. S. Aplicação de técnicas de descobrimento de conhecimento em bases de dados e de inteligência artificial em avaliação de imóveis. 2002.

[21] HAMDI, M.; ALOUI, C. et al. Forecasting crude oil price using artificial neural networks: a literature survey. Econ. Bull, v. 35, n. 2, p. 1339–1359, 2015.

[22] HAYKIN, S. Neural networks: a comprehensive foundation. [S.l.]: Prentice Hall PTR, 1998

[23] HODGES, A. Alan turing: the enigma. In: Alan Turing: The Enigma. [S.l.]: Princeton University Press, 2014.

[24] HOWARD, A. G. et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861, 2017.

[25] JAYARAMAN, S.; ESAKKIRAJAN, S.; VEERAKUMAR, T. Digital image processing. [S.l.]: Tata McGraw Hill Education New Delhi, 2009. v. 7014.

[26] KONAR, A. Artificial intelligence and soft computing: behavioral and cognitive modeling of the human brain. [S.l.]: CRC press, 2018.

[27] KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. Communications of the ACM, AcM New York, NY, USA, v. 60, n. 6, p. 84–90, 2017.

[28] LEACH, S. et al. Data augmentation for improving deep learning models in building inspections or postdisaster evaluation. Journal of Performance of Constructed Facilities, American Society of Civil Engineers, v. 35, n. 4, p. 04021029, 2021.

[29] LECUN, Y. et al. Gradient-based learning applied to document recognition. Proceedings of the IEEE, Ieee, v. 86, n. 11, p. 2278–2324, 1998.

[30] LOBO, L. C. Inteligência artificial e medicina. Revista Brasileira de Educação Médica, SciELO Brasil, v. 41, p. 185–193, 2017.

[31] LOPES, R. A. et al. Emprego de modelo regressivo linear e rede neural artificial no controle da segurança de barragens de concreto. Revista Sul-Americana de Engenharia Estrutural, v. 5, n. 3, 2011.

- [32] MANOVICH, L. Automatizando a estética: inteligência artificial e cultura das imagens. *Esferas*, n. 11, 2017.
- [33] MEDEIROS, C. P. Processamento de imagem e aprendizagem de máquina para classificação de anomalias em vias públicas. 2018.
- [34] MELO, R. R. S. de; COSTA, D. B.; ÁLVARES, J. S. Diretrizes para o uso de veículo aéreo não tripulado (vant) para inspeção de segurança em canteiros de obra. *Tecnologia para canteiro de obras sustentável*, Editora Scienza, 2017.
- [35] MITCHELL, T. *Machine learning* (wbc/mcgraw-hill, boston). MA, 1997.
- [36] MORETTI, J. F. Sistema inteligente baseado nas redes neurais artificiais para dosagem do concreto. Universidade Estadual Paulista (Unesp), 2010.
- [37] NEGASSI, M.; WAGNER, D.; REITERER, A. Smart (sampling) augment: Optimal and efficient data augmentation for semantic segmentation. *Algorithms*, MDPI, v. 15, n. 5, p. 165, 2022.
- [38] NETO, L. B. et al. Estimativa de valores ausentes com redes neurais artificiais: o caso dos custos de construção civil. *Revista de Economia e Administração*, v. 3, n. 2, 2004.
- [39] NIXON, M.; AGUADO, A. Feature extraction and image processing for computer vision. [S.l.]: Academic press, 2019.
- [40] NUNES, D.; MARQUES, A. L. P. C. Inteligência artificial e direito processual: vieses algorítmicos e os riscos de atribuição de função decisória às máquinas. In: *Revista de Processo*. [S.l.: s.n.], 2018. v. 285, n. 2018, p. 421–447.
- [41] OLIVEIRA, B. A. S. et al. Deep learning para detecção de componentes em 69 alimentadores de subestações. *Simpósio Brasileiro de Telecomunicações e Processamento de Sinais-SBRt*, v. 1, 2020.
- [42] OTTONI, A. L.; NOVO, M. S.; COSTA, D. B. Deep learning for vision systems in construction 4.0: a systematic review. *Signal, Image and Video Processing*, Springer, p. 1821-1829, 2023.
- [43] OTTONI, A. L. C. et al. Tuning of data augmentation hyperparameters in deep learning to building construction image classification with small datasets. *International Journal of Machine Learning and Cybernetics*, Springer, v. 14, n. 1, p. 171–186, 2023.
- [44] OTTONI, A. L. C.; NOVO, M. S. A deep learning approach to vegetation images recognition in buildings: a hyperparameter tuning case study. *IEEE Latin America Transactions*, IEEE, v. 19, n. 12, p. 2062–2070, 2021.
- [45] OTTONI, A. L. C. et al. Métodos para recomendação de hiperparâmetros de aprendizado de máquina na classificação de imagens da construção civil. Universidade Federal da Bahia, 2022.
- [46] ÖZGENEL, Ç. F.; SORGUÇ, A. G. Performance comparison of pre-trained convolutional neural networks on crack detection in buildings. In: *IAARC PUBLICATIONS*. Isarc. proceedings of the international symposium on automation and robotics in construction. [S.l.], 2018. v. 35, p. 1–8.
- [47] PARK, D. S. et al. Specaugment: A simple data augmentation method for automatic speech recognition. *arXiv preprint arXiv:1904.08779*, 2019.
- [48] PARKER, J. R. *Algorithms for image processing and computer vision*. [S.l.]: John Wiley Sons, 2010.
- [49] PAWARA, P. et al. Data augmentation for plant classification. In: *SPRINGER. Advanced Concepts for Intelligent Vision Systems: 18th International Conference, ACIVS 2017, Antwerp, Belgium, September 18-21, 2017, Proceedings 18*. [S.l.], 2017. p. 615–626.
- [50] PEREZ, F. et al. Data augmentation for skin lesion analysis. In: *SPRINGER. OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis: First International Workshop, OR 2.0 2018, 5th International Workshop, CARE 2018, 7th International Workshop, CLIP 2018, Third International Workshop, ISIC 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16 and 20, 2018, Proceedings 5*. [S.l.], 2018. p. 303–311.
- [51] POGGIO, T.; TORRE, V.; KOCH, C. Computational vision and regularization theory. *Readings in computer vision*, Elsevier, p. 638–643, 1987.
- [52] POZZEBON, E.; FRIGO, L. B.; BITTENCOURT, G. Inteligência artificial na educação universitária: quais as contribuições. *Campinas: Revista CCEI*, v. 8, n. 13, p. 34–41, 2004.
- [53] RAMESH, A. et al. Artificial intelligence in medicine. *Annals of the Royal College of Surgeons of England, Royal College of Surgeons of England*, v. 86, n. 5, p. 334, 2004.
- [54] RASHID, K. M.; LOUIS, J. Times-series data augmentation and deep learning for construction equipment activity recognition. *Advanced Engineering Informatics*, Elsevier, v. 42, p. 100944, 2019.
- [55] RICI, P.; SANTOS, S. O. S.; OTTONI, A. L. C. Tuning of data augmentation hyperparameters to covid-19 detection in x-ray images with deep learning. 2022.
- [56] RODRIGUES, J. SISAMED: sistema de baixo custo de apoio e alerta à toma de medicação em casa recorrendo a uma visão computacional e modelos deep learning. Tese (Doutorado), 2022.
- [57] RUIZ, R. D. B. et al. Processamento digital de imagens para detecção automática de fissuras em revestimentos cerâmicos de edifícios. *Ambiente Construído, SciELO Brasil*, v. 21, p. 139–147, 2020.
- [58] RUSSELL, S. J.; NORVIG, P. *Artificial intelligence: a modern approach*. [S.l.: s.n.], 2013.
- [59] SAMSURYADI, S.; ERMATITA, E. Efficient mobilenet architecture as image recognition on mobile and embedded devices. *Indonesian Journal of Electrical Engineering and Computer Science, Institute of Advanced Engineering and Science (IAES)*, v. 22, n. 1, p. 389–394, 2019.
- [60] SANTOS, L. d. et al. Inteligência artificial aplicada ao monitoramento da saúde estrutural. Universidade Federal de Campina Grande, 2021.
- [61] SASAKI, K.; HASHIMOTO, M.; NAGATA, N. Person invariant classification of subtle facial expressions using coded movement direction of keypoints. In: *SPRINGER. Video Analytics. Face and Facial Expression Recognition and Audience Measurement: Third International Workshop, VAAM 2016, and Second International Workshop, FFER 2016, Cancun, Mexico, December 4, 2016, Revised Selected Papers 2*. [S.l.], 2017. p. 61–72.
- [62] Shelby, Zach. Optimize AI for the edge. 2023. Disponível em: <https://bioconductor.org/packages/release/bioc/html/EBImage.html>. Acesso em: 12 de maio 2023.
- [63] SHORTEN, C.; KHOSHGOFTAAR, T. M. A survey on image data augmentation for deep learning. *Journal of big data, SpringerOpen*, v. 6, n. 1, p. 1–48, 2019.
- [64] SILVA, I. N. d.; SPATTI, D. H.; FLAUZINO, R. A. Redes neurais artificiais para engenharia e ciências aplicadas. 2010.
- [65] SONG, C. et al. Analysis on the impact of data augmentation on target recognition for uav-based transmission line inspection. *Complexity, Hindawi Limited*, v. 2020, p. 1–11, 2020.
- [66] SONKA, M.; HLAVAC, V.; BOYLE, R. *Image processing, analysis, and machine vision*. [S.l.]: Cengage Learning, 2014.
- [67] SOUZA, D. A. de et al. Redes neurais artificiais: Identificação de danos em elementos estruturais. *Anais do SIMPÓSIO NACIONAL DE CIÊNCIAS E ENGENHARIAS (SINACEN)*, v. 6, n. 1, p. 109–119, 2021.
- [68] STAFFA, L. et al. Data augmentation approach in detecting roof pathologies with uass images. In: *IOP PUBLISHING. IOP Conference Series: Earth and Environmental Science*. [S.l.], 2022. v. 1101, n. 9, p. 092002.
- [69] TRAN, N.-T. et al. On data augmentation for gan training. *IEEE Transactions on Image Processing*, IEEE, v. 30, p. 1882–1897, 2021.
- [70] UM, T. T. et al. Data augmentation of wearable sensor data for parkinson’s disease monitoring using convolutional neural networks. In: *Proceedings of the 19th ACM international conference on multimodal interaction*. [S.l.: s.n.], 2017. p. 216–220.
- [71] VEIGA, R.; PIRES, C. C.; ASCENSO, R. O impacto da inteligência artificial na gestão de riscos. *ISLA Multidisciplinary e-Journal*, v. 3, n. 1, p. 19–34, 2020.
- [72] WANG, Z. et al. Cnn training with twenty samples for crack detection via data augmentation. *Sensors*, MDPI, v. 20, n. 17, p. 4849, 2020.
- [73] Wickham, Hadley. *RStudio Desktop*. 2023. Disponível em: <https://posit.co/download/rstudio-desktop/>. Acesso em: 17 de maio 2023.
- [74] WINSTON, P. H. *Artificial intelligence*. [S.l.]: Addison-Wesley Longman Publishing Co., Inc., 1984.
- [75] YANG, X. et al. Automatic pixel-level crack detection and measurement using fully convolutional network. *Computer-Aided Civil and Infrastructure Engineering*, Wiley Online Library, v. 33, n. 12, p. 1090–1109, 2018.
- [76] ZOPH, B. et al. Learning data augmentation strategies for object detection. In: *SPRINGER. Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXVII 16*. [S.l.], 2020. p. 566–583.