

Aprendizado de Máquina para Avaliação de Danos em Construções Históricas Utilizando Visão Computacional

Lavinia Souza Ferreira
MLBots – Machine Learning and Robotics
Escola de Minas
Universidade Federal de Ouro Preto
Ouro Preto, Minas Gerais
lavinia.ferreira@aluno.ufop.edu.br

André Luiz Carvalho Ottoni
MLBots – Machine Learning and Robotics
Departamento de Computação
Universidade Federal de Ouro Preto
Ouro Preto, Minas Gerais
andre.ottoni@ufop.edu.br

Resumo — A preservação do patrimônio histórico é um desafio crucial, exigindo métodos modernos para a identificação e monitoramento de danos em edificações antigas. O uso de inteligência artificial, especialmente *Deep Learning*, tem se mostrado uma abordagem promissora para aprimorar a precisão e a eficiência desse processo. No entanto, um dos principais desafios é desenvolver modelos de classificação de danos robustos e confiáveis, considerando a influência de diferentes hiperparâmetros. Este trabalho tem como objetivo avaliar manifestações patológicas em portas e janelas de edifícios históricos utilizando técnicas de *Deep Learning*, analisando o impacto de hiperparâmetros na performance da classificação. A metodologia envolveu a criação e rotulagem de um banco de imagens em quatro categorias: portas com danos, portas sem danos, janelas com danos e janelas sem danos. Foram testadas diferentes arquiteturas de redes neurais, variações na taxa de aprendizado e o uso de *Data Augmentation* para melhorar a generalização do modelo. Os resultados indicaram que os ajustes da taxa de aprendizado e a aplicação de *Data Augmentation* obtiveram resultados significativos de acurácia, sendo 67% para portas e 80,0% para janelas. Já com relação à variação das arquiteturas neurais, inicialmente, foi utilizada a arquitetura *MobileNetV2 96x96 0,35*. Nas etapas subsequentes, outras arquiteturas foram selecionadas para comparação como *MobileNetV1 96x96 0,1*, *MobileNetV2 160x160 0,35* e *EfficientNet B0*. Ao final dos testes, a *MobileNetV2 96x96 0,35* trouxe melhores resultados para a classe de portas enquanto para a classe de janelas, a *MobileNetV2 160x160 0,35* se destacou. Nos testes práticos, utilizando as melhores combinações entre os três hiperparâmetros, os resultados obtidos foram os seguintes: para portas com danos, a acurácia foi de 80,0%; para portas sem danos, foi alcançada uma acurácia de 100%. Em relação às janelas com danos, obteve-se uma acurácia de 100%, e, por fim, para janelas sem danos, a acurácia foi de 87,5%.

Palavras-chave: Aprendizado profundo, Classificação de danos, Patrimônio histórico, Redes Neurais Artificiais, Visão computacional.

1. INTRODUÇÃO

A conservação de edificações, especialmente as de valor histórico e cultural, representa um desafio crescente em diversas partes do mundo [23], [24], [25], [26], [27], [28], [29]. No Brasil, as edificações históricas demandam cuidados especiais, uma vez que são testemunhos de épocas passadas e parte integrante do patrimônio arquitetônico do país. A cidade de Ouro Preto, por exemplo, abriga um vasto conjunto de edificações históricas de grande relevância, sendo objeto de estudos que buscam aprimorar a preservação desse patrimônio por meio da tecnologia [1], [2].

A manutenção preventiva, com ênfase na inspeção visual, tem sido o principal meio de identificação de manifestações patológicas, como fissuras, infiltrações e corrosão, que podem comprometer a integridade dessas estruturas [16], [17], [18], [19], [20]. Entretanto, os métodos tradicionais de inspeção

apresentam limitações, tanto na capacidade humana de detecção quanto na necessidade de inspeções recorrentes, o que pode resultar em falhas na identificação de problemas em estágios iniciais [6].

Recentemente, a aplicação de tecnologias avançadas, como Inteligência Artificial (IA) e *Deep Learning*, tem mostrado grande potencial na análise e monitoramento de canteiros de obras [5], [6], [7], edificações [6], [7], [8] e identificação de manifestações patológicas [11], [15]. Essas ferramentas, ao serem aplicadas em imagens da construção, permitem a identificação de padrões de danos muitas vezes invisíveis ao olho humano, tornando o processo de detecção de manifestações patológicas mais eficiente e preciso [9], [10]. Estudos demonstram que a utilização de redes neurais convolucionais e técnicas de aumento de dados podem melhorar significativamente a detecção de rachaduras e outras manifestações patológicas em edificações [3], [4], [5], [6], [7], [8], [9], [10], [11].

A utilização de IA nesse contexto apresenta desafios técnicos, sendo um dos mais significativos a análise e ajuste de hiperparâmetros [3], [4]. O desempenho dos modelos de *Deep Learning* depende diretamente do ajuste adequado desses parâmetros, pois um ajuste inadequado pode comprometer a acurácia da detecção e aumentar os erros nos diagnósticos [12], [13]. No entanto, a literatura carece de estudos que utilizem IA para detecção de manifestações patológicas em edificações históricas de Minas Gerais, especialmente observando o ajuste de hiperparâmetros [14].

Dessa forma, o objetivo deste trabalho é explorar a aplicação de IA para a detecção de manifestações patológicas em edificações históricas de Minas Gerais, com uma abordagem focada no ajuste de hiperparâmetros. A pesquisa se baseia em abordagens recentes de aprendizado profundo e datasets já empregados na região [1], [2], considerando técnicas de otimização de modelos e melhorias na acurácia da classificação de manifestações patológicas estruturais [10].

Nesse sentido, as principais contribuições deste trabalho são:

- Avaliação de manifestações patológicas em portas e janelas de edifícios históricos;
- Aprendizado de máquina utilizando variação de hiperparâmetros para o reconhecimento de duas classes binárias, sendo elas: Portas com danos, portas sem danos, janelas com danos e janelas sem danos;
- Experimentos simulados e aplicações reais utilizando visão computacional e dispositivos móveis (smartphones) para a detecção de manifestações patológicas em portas e janelas de edifícios religiosos históricos.

Este artigo está organizado em quatro seções. A *Seção II* apresenta a metodologia proposta para a aplicação de Inteligência Artificial na detecção de manifestações patológicas. A *Seção III* descreve os resultados obtidos a partir da implementação da metodologia proposta, apresentando análises e comparações de desempenho. Por fim, a *Seção IV* discute as conclusões do estudo.

II. METODOLOGIA

Esta seção descreve a metodologia proposta neste estudo, dividida nas seguintes etapas (*Fig. 1*): (A) Descrição do Processo de Rotulagem, (B) Dataset de Manutenções Patológicas, (C) Experimentos no Edge Impulse e Ajuste de Hiperparâmetros e (D) Testes Práticos.

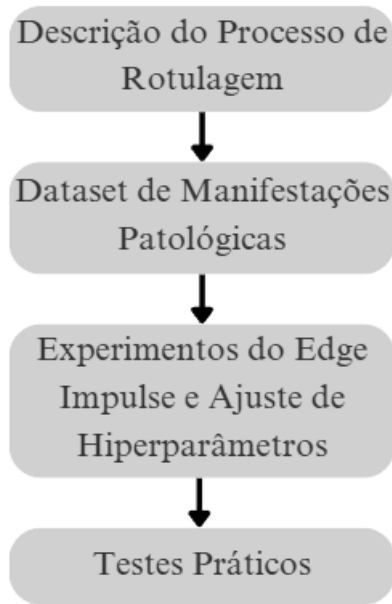


Fig. 1 - Sequência da Metodologia.

A. Descrição do processo de rotulagem

Os datasets escolhidos para o aprendizado de máquina foram o *ImageOP* [1] e o *ImageMG* [2]. O *ImageOP* foi desenvolvido com o objetivo de fornecer uma base de imagens voltada para o treinamento de modelos de *Deep Learning*,

com foco na identificação e classificação de elementos arquitetônicos presentes em edificações religiosas brasileiras, com ênfase na cidade de Ouro Preto, em Minas Gerais. Já o *ImageMG* engloba imagens de edificações de diversas cidades do estado de Minas Gerais, permitindo uma análise mais abrangente das manifestações patológicas e dos elementos arquitetônicos presentes nessas construções históricas.

O processo de rotulagem das imagens de janelas e portas no banco de dados foi desenvolvido com o objetivo de classificar as imagens de edifícios históricos em dois conjuntos e dois subconjuntos binários: portas com danos, portas sem danos, janelas com danos e janelas sem danos [16-20]. No entanto, realizar essa classificação e avaliar a presença ou ausência de danos em cada imagem representa um grande desafio, especialmente devido à presença de outros elementos visuais nas fotos.

A primeira categoria, denominada *alto nível de danos*, envolve imagens de elementos estruturais que apresentam manifestações patológicas graves, enquanto a segunda, *baixo nível de danos* ou *bem conservado*, é destinada às imagens de elementos que estão em boas condições ou com danos mínimos [16], [17], [18], [19].

Para uma imagem ser classificada na categoria de *alto nível de danos*, tanto o elemento (janela ou porta) quanto a parede à qual ele está inserido devem exibir manifestações patológicas significativas que exigem intervenção imediata. As manifestações patológicas observadas nessa categoria incluem desagregação de materiais, infiltração de água, danos estruturais graves, fissuras superficiais, desgaste de revestimentos e umidade leve.

Por outro lado, uma imagem é classificada na categoria de *baixo nível de danos* ou *bem conservado* quando os elementos presentes não apresentam manifestações patológicas graves ou quando as condições de conservação são adequadas. Para essa classificação, verifica-se a ausência de fissuras nas paredes, com acabamento intacto e sem sinais de umidade ou degradação. Além disso, os acabamentos, como pinturas, azulejos e detalhes arquitetônicos, devem estar em bom estado de conservação, sem danos visíveis ou alterações que comprometam a estética ou a funcionalidade do ambiente. Embora possa haver pequenas manifestações patológicas ou sinais de desgaste, elas não representam risco imediato e podem ser resolvidas em um médio ou longo prazo. A *Fig. 2* ilustra exemplos das duas classificações.



Fig. 2 - Exemplos de Imagens Classificadas Após a Rotulagem: (a, b) Portas Com Danos; (c, d) Portas Sem Danos; (e, f) Janelas Com Danos e (g, h) Janelas Sem Danos. Fonte: Imagens do Banco de Dados [1]

B. Dataset de Manifestações Patológicas

O banco de dados inicialmente possuía um total de 1318 imagens de portas e janelas de edifícios históricos de algumas cidades e distritos do estado de Minas Gerais, incluindo Acuruí, Itatiaia, Lagoa Dourada, São Brás do Suaçuí, Ouro Preto, São João del-Rei, Tiradentes, Marilândia, Oliveira e Itapeçerica. Estas imagens foram coletadas com o objetivo de realizar a rotulagem e classificação das condições de conservação desses elementos arquitetônicos.

Após o processo de rotulagem, as imagens foram divididas em duas categorias principais: imagens com danos e imagens sem danos. O conjunto final de dados resultou em 795 imagens com danos (391 de portas e 404 de janelas) e 523 imagens sem danos (234 de portas e 289 de janelas).

Para o treinamento do modelo de aprendizado de máquina, as imagens foram separadas em dois conjuntos: um conjunto de treinamento e um conjunto de teste. As imagens de janelas e portas, tanto com danos quanto sem danos, provenientes das localidades de Marilândia, Oliveira e Itapeçerica foram destinadas ao conjunto de teste, enquanto as imagens das demais localidades foram utilizadas para o treinamento. Essa divisão geográfica do *dataset* com cidades para treinamento e teste segue a abordagem proposta pelo ImageOP [1]. A *Tabela I* apresenta a distribuição detalhada das imagens conforme a rotulação realizada.

TABELA I. CONJUNTO DE DADOS

Imagens de treinamento				
Região	Porta		Janela	
	Com danos	Sem danos	Com danos	Sem danos
Acuruí	33	7	35	0
Itatiaia	14	2	5	0
Lagoa Dourada e São Brás do Suaçuí	3	9	5	18
Ouro preto	194	98	203	138
São João del-Rei	17	32	21	27
Tiradentes	63	43	92	55
Amostras Treinamento	324	191	361	238
Imagens de teste				
Marilândia, Oliveira e Itapeçerica	67	43	43	51

C. Experimentos do Edge Impulse e Ajuste de Hiperparâmetros

Utilizou-se a plataforma *Edge Impulse* para realizar o *upload* e a organização das imagens destinadas ao treinamento e teste do modelo. O *Edge Impulse* é uma plataforma de desenvolvimento voltada à construção, treinamento e implantação de modelos de aprendizado de máquina em dispositivos embarcados, sendo amplamente utilizada em aplicações de *TinyML* (*Machine Learning* em dispositivos com recursos computacionais limitados, como microcontroladores, sensores inteligentes e sistemas IoT). A plataforma facilita o pré-processamento de dados, a definição de arquiteturas de rede, a validação dos modelos e a exportação final para diferentes hardwares. Um dos grandes

diferenciais da ferramenta é a possibilidade de capturar e enviar dados diretamente por meio de um *smartphone*, utilizando *QR Code*, sem a necessidade de equipamentos especializados.

Antes de iniciar o processo de treinamento, foi necessário definir alguns parâmetros essenciais para a configuração do modelo. No que se refere às imagens, optou-se por mantê-las no modelo de cor RGB (cores reais), sem conversão para escala de cinza, a fim de preservar informações cromáticas importantes para a etapa de classificação. Para *Data Augmentation*, utilizou-se a configuração padrão do Edge Impulse. A arquitetura de rede neural selecionada foi a MobileNetV2, com dimensões de entrada de 96×96 pixels e fator de profundidade 0,35 — uma configuração leve e eficiente, otimizada para dispositivos móveis e adequada a tarefas de classificação de imagens em ambientes com restrições de processamento e memória.

Outros parâmetros foram definidos, como o número de Número de Épocas (*Training cycles*) em 20 e o Tamanho do Lote (*Batch size*) em 32. Durante toda a utilização do Edge Impulse, tanto o treinamento quanto os testes do modelo foram realizados na nuvem.

Conforme a *Tabela II*, com relação à taxa de aprendizado, foram realizados os experimentos de 1 a 6, que tiveram a taxa de aprendizado de 0,0005; 0,005 e 0,05, aplicados tanto para a classe de janelas quanto para a de portas. Esses valores foram selecionados para explorar o impacto de diferentes taxas de aprendizado (*Learning Rate*) sobre o desempenho do modelo.

TABELA II. RESUMO DOS EXPERIMENTOS

Elemento	Experimento	Taxa de Aprendizado	Aumento de Dados
Porta	1	0,0005	-
	2	0,005	-
	3	0,05	-
Janela	4	0,0005	-
	5	0,005	-
	6	0,05	-
Porta	7	0,0005	✓
	8	0,005	✓
	9	0,05	✓
Janela	10	0,0005	✓
	11	0,005	✓
	12	0,05	✓

Diante desses resultados, foi realizado a aplicação de outro hiperparâmetro, que são os experimentos de 7 a 12, onde além da variação da taxa de aprendizado também implantamos o aumento de dados (*Data Augmentation*). Esses valores foram selecionados para explorar o impacto de diferentes taxas de aprendizado e como aumento de dados interferiria nesses resultados.

Inicialmente, foi utilizada a arquitetura MobileNetV2 96×96 0,35 [30]. Essa arquitetura é conhecida por sua eficiência e leveza, sendo otimizada para dispositivos com recursos computacionais limitados. A MobileNetV2 utiliza blocos inverted residual com conexões lineares, o que

permite manter uma boa acurácia mesmo em modelos compactos. Nas etapas seguintes, outras arquiteturas foram testadas para fins comparativos:

A MobileNetV1 (96x96, 0,1) que é uma versão ainda mais leve da arquitetura MobileNet, que utiliza convoluções separáveis em profundidade (depthwise separable convolutions), reduzindo significativamente o número de parâmetros e operações. Apesar de ser mais eficiente, pode apresentar perda de acurácia em cenários mais complexos [30].

A MobileNetV2 (160x160, 0,35) que é uma variação da primeira arquitetura, com resolução de entrada maior (160x160 pixels), permitindo capturar mais detalhes das imagens. Essa mudança, embora aumente o custo computacional, pode contribuir para uma melhor generalização do modelo [30].

E por fim, a EfficientNet B0 que é uma arquitetura mais recente que aplica uma estratégia de dimensionamento composta (compound scaling) para ajustar simultaneamente a profundidade, largura e resolução da rede. É reconhecida por alcançar altos níveis de acurácia com uma eficiência computacional superior em comparação com outras redes de mesma categoria [31].

No contexto dos resultados obtidos por meio da matriz de confusão de cada classificação, serão apresentados as métricas estatísticas utilizadas para a interpretação desses resultados, tais como Acurácia, Precisão, Recall e F1-Score.

Em relação à matriz de confusão, o Verdadeiro Negativo (VN) refere-se ao número de acertos na classificação da classe “sem danos”, ou seja, quando o modelo corretamente identifica a ausência de danos. O Falso Positivo (FP) representa um erro de classificação, ocorrendo quando o modelo incorretamente classifica um caso “sem danos” como “com danos”. O Verdadeiro Positivo (VP) indica o número de acertos na classificação da classe “com danos”, ou seja, quando o modelo corretamente identifica a presença de danos. Por fim, o Falso Negativo (FN) refere-se ao erro de classificação em que o modelo falha ao identificar um caso “com danos”, classificando-o incorretamente como “sem danos”.

O parâmetro estatístico da acurácia representada através da Equação (1), indica a fração de previsões corretas sobre cada imagem e a definição de cada classe categorizada, como mostra a equação abaixo:

$$Acuracia = \frac{VP + VN}{VP + FN + VN + FP} \quad (1)$$

No contexto dos danos em portas e janelas, além da acurácia, outras métricas de avaliação complementam a 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 (2):

$$Precisão = \frac{VP}{VP + FP} \quad (2)$$

Outra métrica importante é o Recall que tem melhores valores quando são diminuídos falsos negativos. esse parâmetro apresenta relevância para o presente trabalho em bancos balanceados, a Equação (3) é representada por:

$$Recall = \frac{VP}{VP + FN} \quad (3)$$

Somando a Precisão e o Recall, a Equação (4) nomeado como “F1-Score”, representa a média harmônica entre as Equações 2 e 3 determinado por:

$$F1 - Score = \frac{2 \cdot Precisão \cdot Recall}{Precisão + Recall} \quad (4)$$

Na análise da matriz de confusão referente ao último experimento, foram avaliadas as melhores combinações de hiperparâmetros previamente ajustados, como a taxa de aprendizado e as técnicas de aumento de dados (*data augmentation*), com o objetivo de identificar o desempenho mais eficaz até aquele momento. Com base nos resultados positivos obtidos na primeira fase de ajustes, a segunda etapa do experimento concentrou-se na variação dessas arquiteturas de redes neurais convolucionais, com o intuito de verificar o impacto da complexidade da arquitetura no desempenho do modelo.

D. Testes Práticos

Dando continuidade ao processo de avaliação das arquiteturas, as que apresentaram o melhor desempenho nas etapas anteriores foram selecionadas para a realização dos testes em campo. Nesta fase, experimentos em mundo real foram conduzidos para validar o conjunto de dados proposto para classificação por aprendizado profundo. Os testes foram realizados na cidade de Catas Altas, conforme a Fig. 3.

Catas Altas, localizada no estado de Minas Gerais, é uma cidade de grande relevância histórica e cultural, inserida na região da Estrada Real, um dos mais importantes roteiros do período colonial brasileiro. Fundada no século XVIII, sua origem está diretamente ligada à exploração do ouro, atividade que impulsionou seu crescimento e deixou marcas profundas em sua arquitetura. Essa herança ainda é visível nas igrejas, casarões e ruas de pedra que preservam características autênticas da época colonial.

Esses elementos tornam Catas Altas um local de grande valor para estudos voltados à conservação do patrimônio edificado e à aplicação de novas tecnologias na preservação de edificações históricas. Em reconhecimento à sua importância, o centro histórico da cidade foi tombado pelo Instituto Estadual do Patrimônio Histórico e Artístico de Minas Gerais (IEPHA), conforme o Decreto Nº 29.399, de 21 de abril de 1989, estando inscrito no Livro do Tombo Arqueológico, Etnográfico e Paisagístico, bem como no Livro do Tombo Histórico, das Obras de Arte Históricas e dos Documentos Paleográficos ou Bibliográficos.



Fig. 3 - Processo de visão computacional utilizando dispositivo móvel na cidade histórica de Catas Altas.

A escolha de Catas Altas como área de estudo para a presente pesquisa justifica-se por seu elevado grau de preservação histórica e pelo desafio constante de manter suas edificações em bom estado de conservação frente às ações do tempo e à necessidade de intervenções técnicas cuidadosas.

Para esta fase, um dispositivo móvel (smartphone Iphone 11) foi empregado para a detecção em tempo real de elementos de construção usando visão computacional, conforme a Fig. 3. O acesso ao modelo treinado no software Edge Impulse foi obtido por meio do smartphone. Vale ressaltar que, nesta fase, seis edifícios religiosos foram selecionados, conforme a Fig. 4. O objetivo foi avaliar o

desempenho do conjunto de dados proposto e do modelo treinado em ambientes reais não incluídos nas imagens do conjunto de dados.

Os modelos de aprendizado profundo foram treinados na plataforma Edge Impulse. Em seguida, o software Edge Impulse disponibiliza a aplicação do sistema de inteligência artificial em dispositivos móveis por meio da leitura de um código QR. No smartphone, o aplicativo é aberto em um navegador web. Assim, é possível direcionar a câmera do dispositivo para a igreja onde se deseja detectar os componentes do edifício. Os resultados do reconhecimento em tempo real são exibidos na interface do smartphone.



(a)



(b)



(c)



(d)



(e)



(f)

Fig. 4: Edificações Históricas utilizadas para realização dos testes práticos na cidade de Catas Altas – Minas Gerais, Brasil: (a) Igreja Nossa Senhora da Conceição; (b) Igreja Nossa Senhora do Rosário; (c) Igreja Nosso Senhor do Bonfim, (d) Secretaria Municipal de Educação; (e) Prefeitura Municipal de Catas Altas e (f) Câmara Municipal de Catas Altas.

III. RESULTADOS

Os resultados obtidos no desenvolvimento deste trabalho estão organizados em duas seções principais, conforme descrito a seguir:

- Avaliação da Variação de Taxa de Aprendizado e Aumento de Dados: onde serão apresentados os impactos da variação da taxa de aprendizado e da aplicação de técnicas de aumento de dados no desempenho do modelo.

- Avaliação da Variação de Arquiteturas Neurais: onde serão comparadas diferentes arquiteturas de redes neurais convolucionais (CNNs), observando o desempenho individual de cada uma frente ao mesmo conjunto de dados.

A. Avaliação da Variação de Taxa de Aprendizizado e Aumento de Dados

Na classe de portas, conforme a *Tabela III*, percebe-se que dentre as taxas de aprendizado avaliados a que obteve melhores resultados foi o valor médio. Além disso, após introduzir o hiperparâmetro de aumento de dados, observou-se uma melhoria na Acurácia do modelo. A melhor taxa de aprendizado continuou sendo 0,005, porém houve uma redução na Precisão enquanto Recall e F1-Score se mantiveram.

TABELA III. RESULTADOS DA VALIDAÇÃO PARA OS SUBCONJUNTOS DE PORTAS. EM NEGRITO OS MELHORES RESULTADOS POR HIPERPARÂMETRO.

Taxa de aprendizado	Acurácia	Precisão	Recall	F1-Score
<i>Sem aumento de dados</i>				
0,0005	63,1%	72,0%	60,0%	60,0%
0,005	66,0%	76,0%	70,0%	70,0%
0,05	62,1%	70,0%	60,0%	60,0%
<i>Com aumento de dados</i>				
0,0005	65,0%	69,0%	60,0%	60,0%
0,005	67,0%	73,0%	67,0%	67,0%
0,05	48,5%	62,0%	50,0%	40,0%

Na classe de janelas, conforme os dados apresentados na *Tabela IV*, observa-se que menores taxas de aprendizado resultaram em melhores desempenhos dos parâmetros avaliados. Além disso, após a introdução do hiperparâmetro de aumento de dados (*Data Augmentation*), a taxa de aprendizado que apresentou melhor desempenho manteve-se inalterada, porém houve uma melhora na acurácia do modelo. Por outro lado, os valores de Precisão, Recall e F1-Score permaneceram estáveis.

TABELA IV. RESULTADOS DA VALIDAÇÃO PARA OS SUBCONJUNTOS DE JANELAS. EM NEGRITO OS MELHORES RESULTADOS POR HIPERPARAMETRO.

Taxa de aprendizado	Acurácia	Precisão	Recall	F1-Score
<i>Sem aumento de dados</i>				
0,0005	78,3%	80,0%	80,0%	80,0%
0,005	73,3%	80,0%	70,0%	70,0%
0,05	61,7%	40,0%	60,0%	50,0%
<i>Com aumento de dados</i>				
0,0005	80,0%	80,0%	80,0%	80,0%
0,005	74,2%	80,0%	70,0%	70,0%
0,05	61,7%	40,0%	60,0%	50,0%

Diante dos resultados obtidos, foram selecionadas, para as classes de portas e janelas, as melhores combinações de hiperparâmetros, bem como suas respectivas matrizes de confusão. Conforme apresentado na *Tabela V*, observa-se que, para as imagens de portas, os valores registrados foram: 54,1% de verdadeiros positivos (VP), 14,3% de falsos positivos (FP), 45,9% de falsos negativos (FN) e 85,7% de verdadeiros negativos (VN).

TABELA V. MATRIZ DE CONFUSÃO PARA TAXA DE APRENDIZADO DE 0,005 E AUMENTO DE DADOS PARA PORTAS

	<i>Porta com danos</i>	<i>Porta sem danos</i>
<i>Porta com danos</i>	54,1%	45,9%
<i>Porta sem danos</i>	14,3%	85,7%
F1-Score	66,0%	68,0%

Já em relação às imagens de janelas, foram obtidos 81,1% de VP, 21,7% de FP, 18,9% de FN e 78,3% de VN, conforme observa-se na *Tabela VI*.

TABELA VI. MATRIZ DE CONFUSÃO PARA TAXA DE APRENDIZADO DE 0,0005 E AUMENTO DE DADOS PARA JANELAS

	<i>Porta com danos</i>	<i>Porta sem danos</i>
<i>Porta com danos</i>	81,1%	18,9%
<i>Porta sem danos</i>	21,7%	78,3%
F1-Score	83,0%	75,0%

A. Avaliação da Variação de Arquiteturas Neurais

A análise das arquiteturas aplicadas à classe de portas foi conduzida com base nos resultados obtidos nos testes realizados sobre o subconjunto de dados, conforme apresentado na *Tabela VII*. Dentre as arquiteturas avaliadas, a MobileNetV2 com resolução 96×96 e largura 0.35 destacou-se por apresentar o melhor desempenho. Essa configuração obteve elevada precisão na classificação de imagens tanto de portas com danos quanto de portas sem danos, superando as demais alternativas testadas.

TABELA VII. RESULTADOS DOS TESTES PARA O SUBCONJUNTO DE PORTAS PARA CADA ARQUITETURA NEURAL. EM NEGRITO OS MELHORES RESULTADOS POR HIPERPARAMETRO.

Arquitetura Neural	Acurácia	Precisão	Recall	F1-Score
MobileNetV2 96x96 0,35	67,3%	71,0%	71,0%	71,0%
MobileNetV1 96x96 0,1	55,4%	68,0%	68,0%	68,0%
MobileNetV2 160x160 0,35	55,4%	63,0%	58,0%	58,0%
EfficientNet B0	66,4%	75,0%	68,0%	68,0%

Para a classe de janelas, a arquitetura MobileNetV2 com resolução 160×160 e largura 0.35 manteve-se como a melhor opção, apresentando resultados semelhantes aos obtidos para as portas, conforme detalhado na *Tabela VIII*.

TABELA VIII. RESULTADOS DOS TESTES PARA O SUBCONJUNTO DE JANELAS PARA CADA ARQUITETURA NEURAL. EM NEGRITO OS MELHORES RESULTADOS POR HIPERPARAMETRO.

Arquitetura Neural	Acurácia	Precisão	Recall	F1-Score
MobileNetV2 96x96 0,35	56,4%	65,0%	64,0%	64,0%
MobileNetV1 96x96 0,1	45,7%	62,0%	61,0%	60,0%
MobileNetV2 160x160 0,35	63,8%	69,0%	69,0%	69,0%
EfficientNet B0	61,0%	69,0%	69,0%	69,0%

B. Avaliação dos Testes Práticos

Os testes foram avaliados com base nos resultados apresentados na Fig. 5. Para as portas com danos, foram analisadas 10 imagens, das quais o modelo classificou corretamente 8, resultando em uma acurácia de 80,0%. Em relação às portas sem danos, 7 imagens foram avaliadas e

todas foram corretamente classificadas, alcançando uma acurácia de 100%. Para as janelas com danos, o modelo obteve acerto total nas 12 imagens analisadas, também com 100% de acurácia. Por fim, no caso das janelas sem danos, 7 das 8 imagens foram corretamente classificadas, resultando em uma acurácia de 87,5%. Os resultados completos estão sintetizados na Tabela IX.

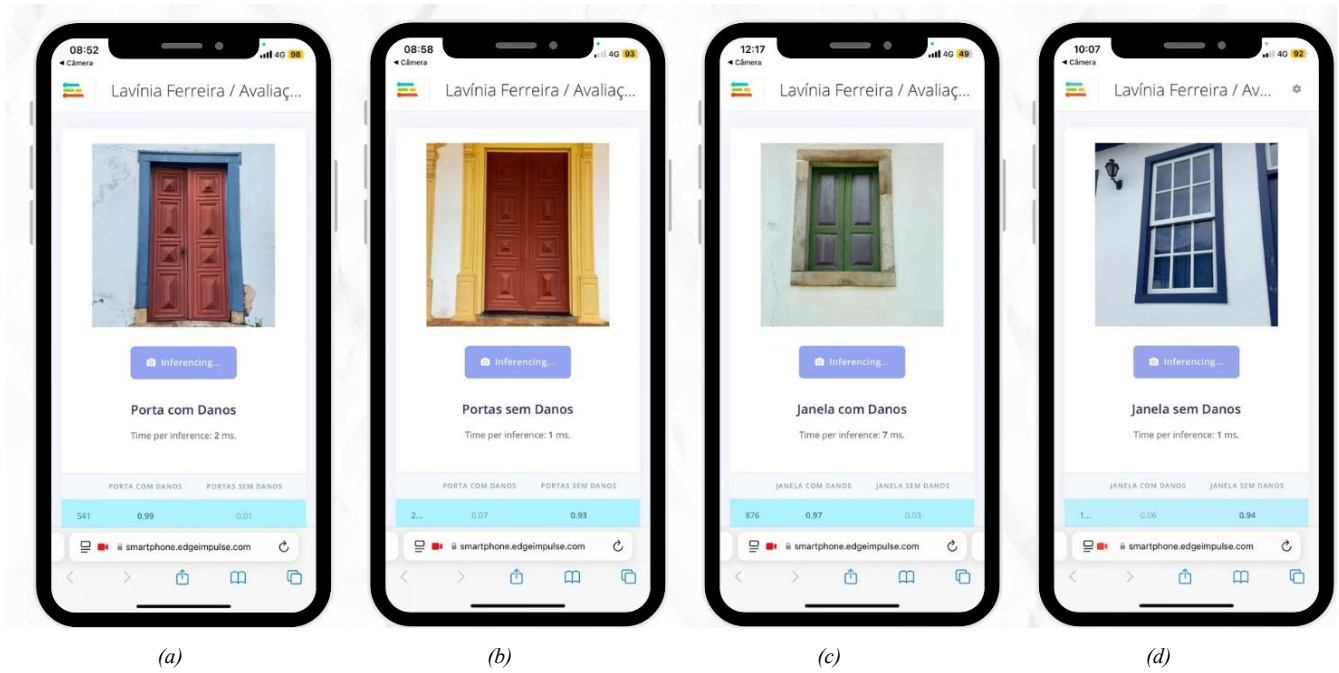


Fig. 5: Exemplos de classificação em tempo real a partir de capturas de tela da interface gráfica do Edge Impulse acessadas no dispositivo móvel.

A Fig. 5 apresenta exemplos do processo de visão computacional utilizando um dispositivo móvel na cidade histórica de Catas Altas. Observa-se que, ao apontar a câmera do smartphone para elementos dos edifícios religiosos, as inferências são exibidas em tempo real na interface web do Edge Impulse. A figura mostra capturas de tela da interface gráfica acessada no dispositivo, ilustrando classificações realizadas corretamente para as quatro classes abordadas neste artigo.

TABELA IX. RESULTADOS DE ACURÁCIA PARA OS TESTES PRÁTICOS

Elemento	Com danos	Sem danos
Porta	80,0%	100,0%
Janela	100,0%	87,5%

A leve variação nas taxas de acerto pode ser atribuída a fatores como a qualidade das imagens ou a complexidade das cenas, que podem impactar a capacidade do modelo de generalizar para todos os casos. No entanto, a alta taxa de acerto nas classes de portas e janelas com danos, bem como a precisão nas portas sem danos, sugerem que a arquitetura selecionada tem grande potencial para aplicações práticas na detecção de danos em estruturas, especialmente em contextos urbanos históricos, como os de Catas Altas.

IV. CONCLUSÃO

O presente trabalho teve como objetivo avaliar manifestações patológicas em portas e janelas de edifícios históricos utilizando técnicas de *Deep Learning*, com foco

na influência de diferentes hiperparâmetros na classificação de danos. A principal contribuição deste estudo foi a análise detalhada, por meio do *dataset ImageOP*, de como variáveis como a taxa de aprendizado, o uso de *Data Augmentation* e a escolha da arquitetura têm influência na acurácia do modelo.

Os resultados mostraram que a escolha adequada do *Learning Rate* teve um impacto significativo no desempenho da rede, evidenciando a importância de um ajuste criterioso desse parâmetro. Além disso, o uso de *Data Augmentation*, especialmente com transformações como rotações, espelhamentos e ajustes de brilho, melhorou a generalização do modelo e reduziu a ocorrência de *Overfitting*. A análise comparativa de diferentes arquiteturas, como a *MobileNetV2 160x160 0.35* e a *MobileNetV2 96x96 0.35*, indicou que algumas configurações foram mais eficazes em termos de precisão e eficiência computacional, enquanto outras se destacaram na redução de falsos positivos e falsos negativos. Com relação aos testes em campo, os resultados obtidos foram satisfatórios, uma vez que a acurácia variou entre 80,0% e 100,0%. Assim, conclui-se que as combinações entre taxa de aprendizado, aumento de dados e escolha da arquitetura neural foram bastante eficazes.

Como trabalhos futuros, sugere-se a ampliação do conjunto de dados, incorporando imagens de diferentes regiões geográficas, de modo a aumentar a diversidade e a representatividade dos exemplos. Outra possibilidade seria o desenvolvimento de um sistema automatizado capaz de classificar não apenas a presença de danos, mas também o tipo específico de patologia identificada, contribuindo para diagnósticos mais precisos e intervenções mais eficazes na conservação do patrimônio histórico. Além disso, propõe-se estender a classificação para danos presentes em outros elementos arquitetônicos de construções históricas, como

frontões, torres e ornamentos, auxiliando de forma ainda mais precisa na conservação do patrimônio histórico.

AGRADECIMENTOS

Os autores agradecem ao apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq (Edital PIBIC/CNPq/UFOP Nº 03/2024) e da Universidade Federal de Ouro Preto – UFOP (Edital PROPPI 18/2024).

V. REFERÊNCIAS

- [1] A. L. C. Ottoni, L. T. C. Ottoni, “ImageOP: The Image Dataset with Religious Buildings in The World Heritage Town of Ouro Preto for Deep Learning Classification” *Heritage* 2024, 7(11), 6499–6525.
- [2] A. L. C. Ottoni, L. T. C. Ottoni, “ImageMG. Mendeley Data”, V1, 2025. DOI: 10.17632/nhfp5yd3jn.1.
- [3] A. L. C. Ottoni, A. M. Arcanjo, “Análise de modelos de aprendizado profundo para sistemas embarcados aplicados à classificação de rachaduras em construções”. In: XVI Brazilian Conference on Computational Intelligence (CBIC 2023), Salvador, 2023. Anais... Salvador: CBIC, 2023.
- [4] A. L. C. Ottoni, L. C. Medeiros, M. S. Novo, “Aplicação de aprendizado profundo na classificação de imagens de patologias da construção civil: análise de data augmentation para bancos de dados desbalanceados”. In: XVI Brazilian Conference on Computational Intelligence (CBIC 2023), Salvador, 2023. Anais... Salvador: CBIC, 2023.
- [5] A. L. C. Ottoni, H. S. Peinado, C. A. Oliveira, R. R. S. Melo, D. B. Costa, and M. S. Novo, “Análise de técnicas de data augmentation para aperfeiçoamento da detecção de sistemas de guarda-corpo e rodapés em canteiros de obras com inteligência artificial,” *Proc. 13º Simpósio Brasileiro de Gestão e Economia da Construção e 4º Simpósio Brasileiro de TIC na Construção*, Aracaju, 2023.
- [6] S. Silva et al., “Modelo de aprendizado de máquina para inspeção automatizada de fachadas de paredes de concreto,” *Proc. 13º Simpósio Brasileiro de TIC na Construção*, Aracaju, 2023.
- [7] M. I. S. C. Lima, R. O. Rey, L. V. S. Sá, D. B. Costa, R. R. S. Melo, J. L. T. Nogueira, and L. B. S. Junior, “Plataforma web para inspeção da segurança em canteiros de obra apoiado por VANT e dispositivos móveis,” *Proc. 3º Simpósio Brasileiro de TIC na Construção*, Uberlândia, pp. 1–11, 2021.
- [8] H. S. Peinado, C. A. Oliveira, A. L. C. Ottoni, R. R. S. Melo, D. B. Costa, and M. S. Novo, “Análise de técnicas de data augmentation para aperfeiçoamento da detecção de sistemas de guarda-corpo e rodapés em canteiros de obras com inteligência artificial,” *Proc. 13º Simpósio Brasileiro de Gestão e Economia da Construção*, Aracaju, 2023.
- [9] M. Mishra and P. B. Lourenço, “Artificial intelligence-assisted visual inspection for cultural heritage: state-of-the-art review,” *Journal of Cultural Heritage*, vol. 66, pp. 536–550, 2024.
- [10] L. A. S. Oliveira, M. V. F. Nunes, and A. L. C. Ottoni, “Desenvolvimento de um sistema otimizado para a classificação de manifestações patológicas em estradas utilizando redes neurais convolucionais,” *Proc. XXVII ENMC, XV ECTM*, Porto Alegre, 2024.
- [11] R. C. Gomes, A. L. C. Ottoni, and R. C. Barros, “Detecção de rachaduras em edificações por meio da análise de imagens aéreas com VANT,” *Proc. XXVII ENMC, XV ECTM*, Porto Alegre, 2024.
- [12] G. Silva and A. L. C. Ottoni, “Análise de técnicas de aumento de dados para processamento de linguagem natural,” *Proc. XVI Brazilian Conference on Computational Intelligence (CBIC 2023)*, Salvador, 2023.
- [13] A. L. C. Ottoni, A. M. Souza, and M. S. Novo, “Automated hyperparameter tuning for crack image classification with deep learning,” *Neural Networks*, 2023.
- [14] A. L. C. Ottoni, M. S. Novo, and D. B. Costa, “Deep learning for vision systems in construction 4.0: A systematic review,” *Original Paper*, 2022.
- [15] M. O. Rebouças, “Aplicação de aprendizado de máquina e data Augmentation para detecção de rachaduras em edificações na construção civil: uma abordagem prática utilizando smartphone,” *TCC na Universidade Federal Do Recôncavo Da Bahia - UFRB*, 2023.
- [16] M. T. G. Barbosa, A. E. Polisseni, and F. M. Tavares, “Análise E Representação Em Contextos Diversos: Projeto, Técnica E Gestão Do Ambiente Construído - Patologias De Edifícios Históricos Tombados”, *Encontro Nacional da Associação Nacional de Pesquisa e Pós-Graduação em Arquitetura e Urbanismo, ENANPARQ*, 2010.
- [17] S. Gomes, R. M. N. Lima, E. C. L. Bezerra, and H. C. Santos, “Identificação de grau de risco em patologias prediais e priorização de suas ações reparadoras através da matriz GUT: estudo em um prédio comercial e um prédio residencial na cidade de Mossoró/RN,” 2022.
- [18] M. X. O. Sá and M. R. M. Pedreiro, “Patologias mais comuns dentro da engenharia civil,” *Revista Ibero-Americana de Humanidades, Ciências e Educação - Rease*, 2023.
- [19] B. Souza Filho, H. O. O. Miranda, and J. A. G. Souza, “Patologias da construção civil,” *Faculdade Ages*.
- [20] T. F. B. Ramos, L. A. Araújo, R. T. Alencar, and E. P. M. Guerra, “Manifestações patológicas em edificação centenária: um estudo de caso,” *Congresso Técnico Científico da Engenharia e da Agronomia - CONTECC*, 2017.
- [21] Instituto Estadual do Patrimônio Histórico e Artístico de Minas Gerais (IEPHA-MG), “Centro histórico de Catas Altas.”
- [22] Silva and L. Oliveira, “Artificial intelligence at the interface between cultural heritage and photography: a systematic literature review,” *Heritage*, vol. 7, pp. 3799–3820, 2024.
- [23] M. Samhour, L. Al-Arabi, and F. Al-Atrash, “Prediction and measurement of damage to architectural heritages facades using convolutional neural networks,” *Neural Computing and Applications*, vol. 34, pp. 18125–18141, 2022.
- [24] N. Karimi, M. Mishra, and P. B. Lourenço, “Automated surface crack detection in historical constructions with various materials using deep learning-based YOLO network,” *International Journal of Architectural Heritage*, Jul. 2024, DOI: 10.1080/15583058.2024.2376177.
- [25] N. Karimi, N. Valibeig, and H. R. Rabiee, “Deterioration detection in historical buildings with different materials based on novel deep learning methods with focusing on Isfahan historical bridges,” *International Journal of Architectural Heritage*, vol. 18, no. 6, pp. 981–993, 2024, doi: 10.1080/15583058.2023.2201576.
- [26] N. Karimi, M. Mishra, and P. B. Lourenço, “Deep learning-based automated tile defect detection system for Portuguese cultural heritage buildings,” *Journal of Cultural Heritage*, vol. 68, pp. 86–98, Jul./Aug. 2024.
- [27] M. Mishra and P. B. Lourenço, “Artificial intelligence-assisted visual inspection for cultural heritage: state-of-the-art review,” *Journal of Cultural Heritage*, vol. 66, pp. 536–550, 2024.
- [28] M. D’Orazio, A. Gianangeli, F. Monni, and E. Quagliarini, “Automatic monitoring of the bio-colonisation of historical building’s facades through convolutional neural networks (CNN),” *Journal of Cultural Heritage*, vol. 70, pp. 80–89, Nov. 2024.
- [29] M. Mishra, “Machine learning techniques for structural health monitoring of heritage buildings: a state-of-the-art review and case studies,” *Journal of Cultural Heritage*, vol. 47, pp. 227–245, 2021.
- [30] Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.C. Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 18–23 June 2018; pp. 4510–4520.
- [31] Tan, M. Efficientnet: Rethinking model scaling for convolutional neural networks. *arXiv* 2019, arXiv:1905.11946.