

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO CEARÁ
PRÓ-REITORIA DE PESQUISA, PÓS-GRADUAÇÃO E INOVAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Alberto Régio Alves de Oliveira

**Método para identificação de danos em pás de aerogeradores com
refinamento de redes neurais convolucionais**

Fortaleza – CE

22 de abril de 2025

Alberto Régio Alves de Oliveira

**Método para identificação de danos em pás de
aerogeradores com refinamento de redes neurais
convolucionais**

Dissertação submetida ao Programa de Pós-Graduação em Ciência da Computação do Instituto Federal de Educação, Ciência e Tecnologia do Ceará, como requisito parcial para obtenção do grau de Mestre em Ciência da Computação.

Linha de Pesquisa: Inteligência Artificial

Orientador: Prof. Dr. Geraldo Luis Bezerra Ramalho

Coorientador: Prof. Dr. Cláudio Marques de Sá Medeiros

Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE)

Pró-Reitoria de Pesquisa, Pós-Graduação e Inovação (PRPI)

Programa de Pós-Graduação em Ciência da Computação (PPGCC)

Fortaleza – CE

22 de abril de 2025



Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE)
Pró-Reitoria de Pesquisa, Pós-Graduação e Inovação (PRPI)
Programa de Pós-Graduação em Ciência da Computação (PPGCC)

Alberto Régio Alves de Oliveira

Esta dissertação foi julgada adequada para a obtenção do título de Mestre em Ciência da Computação, sendo aprovada pela Coordenação do Programa de Pós-Graduação em Ciência da Computação do Instituto Federal de Educação, Ciência e Tecnologia do Ceará e pela banca examinadora:

Orientador: Prof. Dr. Geraldo Luis Bezerra Ramalho
Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE)

Coorientador: Prof. Dr. Cláudio Marques de Sá Medeiros
Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE)

Prof. Dr. Paulo Cesar Cortez
Universidade Federal do Ceará (UFC)

Prof. Dr. Marcelo Marques Simões de Souza
Universidade Federal do Ceará (UFC/Sobral)

Prof. Dr. José Renato de Brito Sousa
Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE)

Fortaleza – CE
22 de abril de 2025

Dedico esse trabalho a minha esposa, principal incentivadora e parceira.
Ao meu filho, que todos os dias ensina-me coisas novas com sua inteligência ímpar.
E aos meus pais, irmãos e sobrinhas por todo o apoio e atenção.

Agradecimentos

Primeiramente, gostaria de agradecer a toda minha família pelo apoio durante o percurso. Ao meu pequeno filho, que tantas vezes solicitou brincadeiras e atenção que momentaneamente não pude dar. A minha esposa, que me ajudou com todo o seu conhecimento em trabalhos acadêmicos. Ao meu irmão, que ajudou com a montagem e configuração da máquina que deu suporte a este trabalho. E aos meus pais e irmã. Agradeço também a todos os colegas de curso que enriqueceram as aulas com suas experiências e pontos de vista, especialmente ao Marcelo Torres e Anderson Lima. A minha equipe do Serpro, que soube gerenciar todas as demandas para liberar tempo para os meus estudos, com grande ajuda da minha líder Akemi, e em especial a Raquel Ramalho, que me mostrou a porta de entrada do mestrado no IFCE. A todos os professores que repassaram com maestria os conhecimentos necessários ao longo do curso. Ao Talisson Araujo Figueiredo, que disponibilizou tempo e esforço para ajudar a entender os assuntos relacionados aos aerogeradores. Também à empresa parceira, que disponibilizou as imagens necessárias para que ocorresse esse estudo. Ao meu coorientador, Prof. Dr. Cláudio Marques de Sá Medeiros, que disponibilizou todo o conhecimento sobre o assunto e fez revisões e considerações valiosas. Finalmente, agradeço ao Prof. Dr. Geraldo Luis Bezerra Ramalho, que orientou este estudo, dizimando todas as dúvidas que porventura surgissem ao longo do percurso, por todas as horas dedicadas, pelo apoio no congresso do Bracis 2024 e mesmo nos momentos de cobrança em busca da excelência.

*“Iluminar a inteligência com conhecimento
é uma das mais elevadas e dignas aspirações do homem.”
(Zíbia Gasparetto)*

Resumo

Os aerogeradores transformam a energia cinética produzida pelo vento, sendo suas pás o componente mais suscetível a danos. Uma vez que paradas não programadas devido aos danos resultam em prejuízos significativos, faz-se necessária a detecção precoce de falhas em pás de aerogeradores. Nessa dissertação, é analisado um método para o treinamento de um modelo de rede neural que identifica danos externos nas pás de aerogeradores a partir de fotografias coloridas. São investigadas algumas abordagens de treinamento baseadas na extração de recortes obtidos da imagem original; na aplicação de técnicas de balanceamento e aumento de dados; na técnica de separação dos recortes nos conjuntos de dados; e no ajuste fino de modelos de redes neurais convolucionais profundas pré-treinadas. Nos experimentos realizados, a melhor configuração de treinamento encontrada é utilizando o modelo EfficientNetB2 pré-treinado com o ImageNet e refinado com recortes extraídos em posições aleatórias da imagem original, aplicando técnicas de aumento de dados e balanceamento, e que são separados aleatoriamente nos conjuntos definidos. Essa configuração obtém os melhores resultados no conjunto de testes para as métricas de precisão, acurácia e sensibilidade, sendo esta a configuração selecionada para ser avaliada no conjunto de validação, obtendo uma precisão de 95%, acurácia de 92,1% e sensibilidade de 95% na classificação das imagens originais. Por fim, a classificação dos recortes permite a localização dos danos na imagem original através da exibição de um mapa de calor. Conclui-se que a configuração de treinamento utilizada contribui para uma maior diversidade das amostras de treinamento, melhorando os resultados de classificação do modelo. Os resultados indicam que o método pode ser integrado ao processo de inspeção dos aerogeradores para auxiliar no processo de identificação de danos.

Palavras-chave: Identificação de danos. Pás de aerogeradores. Redes neurais. Visão computacional.

Abstract

Wind turbines transform the kinetic energy produced by the wind, with their blades being the most susceptible component to damage. Since unplanned shutdowns can lead to significant losses, early detection of blade failures is essential. This dissertation analyzes a method for training a neural network model to identify external damage in wind turbine blades using color photographs. Several training strategies are investigated, including the extraction of patches from the original image, the application of data augmentation and class balancing techniques, the separation of patches into distinct datasets, and the fine-tuning of pre-trained deep convolutional neural network models. In the conducted experiments, the best training configuration is obtained using the EfficientNetB2 model pre-trained on ImageNet, further refined with randomly extracted patches from the original image. This configuration incorporated data augmentation and oversampling, with the extracted patches randomly assigned to the training, validation, and test sets. It achieves the best test results in terms of precision, accuracy, and sensitivity, and was subsequently selected for validation. On the validation set, it achieves 95% precision, 92.1% accuracy, and 95% sensitivity in classifying the original images. Finally, the patch classification enables damage localization in the original image through the generation of a heatmap. It is concluded that the adopted training strategy contributes to greater sample diversity, enhancing the model's classification performance. The results suggest that the proposed method can be integrated into wind turbine inspection workflows to assist in the damage identification process.

Keywords: Computer Vision. Damage Identification. Neural Network. Wind Turbine Blades.

Lista de ilustrações

Figura 1 – Componentes de um aerogerador	16
Figura 2 – Etapas genéricas da produção de uma pá eólica	17
Figura 3 – Principais partes de uma pá eólica	17
Figura 4 – Dano em uma pá eólica causado por uma descarga atmosférica	18
Figura 5 – Método proposto por Agrawal et al. (2024)	24
Figura 6 – Exemplos de danos por descargas atmosféricas	27
Figura 7 – Exemplos de danos de erosão	27
Figura 8 – Exemplos de danos de pintura	28
Figura 9 – Exemplos de danos de trincas	28
Figura 10 – Exemplos de danos por gelo	29
Figura 11 – Abordagens para visão computacional	30
Figura 12 – Arquitetura da rede LeNet-5	31
Figura 13 – Abordagens mais comuns em DTL	34
Figura 14 – Transformação geométrica de inversão horizontal e vertical	35
Figura 15 – Arquitetura da etapa de treinamento	39
Figura 16 – Exemplos das imagens originais exibindo as máscaras, com setas vermelhas indicando a localização dos danos	41
Figura 17 – Etapas de extração dos recortes baseado em janela deslizante	42
Figura 18 – Etapas da extração dos recortes baseado em posição aleatória	43
Figura 19 – Aumento da quantidade de recortes aleatórios	44
Figura 20 – Abordagens para separação dos recortes	45
Figura 21 – Configurações de treinamento utilizadas	48
Figura 22 – Arquitetura da etapa de identificação do dano	49
Figura 23 – Escala de cores do mapa de calor em danos simulados	51
Figura 24 – Resultados do experimento II (método de extração dos recortes)	54
Figura 25 – Resultados da variação da quantidade de recortes aleatórios gerados	55
Figura 26 – Resultados do experimento III (separação dos conjuntos)	56
Figura 27 – Análise através do T-SNE da distribuição das imagens nos conjuntos de treino e teste	57
Figura 28 – Análise com o T-SNE da distribuição das classes do conjunto de treinamento utilizando as abordagens de separação dos conjuntos	58
Figura 29 – Distribuição das classes no conjunto de teste utilizando a separação dos conjuntos pela imagem original, com destaque para os erros de classificação	59
Figura 30 – Exibição das miniaturas do destaque (1) da Figura 29, com borda verde para os recortes com classificação corretas e vermelho para as incorretas	59

Figura 31 – Exibição das miniaturas dos destaques (2) e (3) da Figura 29, com borda verde para os recortes com classificação corretas e vermelho para as incorretas	60
Figura 32 – Exibição das miniaturas dos destaques (4) e (5) da Figura 29, com borda verde para os recortes com classificação corretas e vermelho para as incorretas	61
Figura 33 – Resultados do experimento IV (aumento de dados e balanceamento) . .	62
Figura 34 – Matriz de confusão dos resultados no conjunto de validação, com desta- que para os casos de falso negativo	64
Figura 35 – Análise utilizando o Grad-CAM de alguns casos de recortes classificados como falso negativo	65
Figura 36 – Análise utilizando o Grad-CAM de recortes de dano de descarga e erosão que foram classificados corretamente	66
Figura 37 – Matriz de confusão dos resultados das imagens originais	67
Figura 38 – Imagem original de erosão classificada incorretamente como dano de pintura	68
Figura 39 – Imagem original de pintura classificada incorretamente como dano de trinca. Com círculos vermelhos indicando os danos de pintura e setas verdes e amarelas indicando as características que influenciam a classificação como trinca	69
Figura 40 – Imagem original de trinca classificada incorretamente como dano de pintura. Com círculos vermelhos indicando as trincas existentes	69
Figura 41 – Processo de localização do dano	70
Figura 42 – Variação do tamanho do recortes no processo de localização do dano .	71
Figura 43 – Exemplos de localização para os danos de pintura e descarga atmosférica	72
Figura 44 – Exemplos de localização para os danos de erosão e trinca	73
Figura 45 – Inclusão de caixa delimitadora	83
Figura 46 – Funcionalidade de geração da grade	85
Figura 47 – Funcionalidade de geração dos recortes aleatórios	85
Figura 48 – Funcionalidade de predição dos recortes	86
Figura 49 – Funcionalidade de análise do Grad-CAM	87
Figura 50 – Funcionalidade de segmentação SAM	88

Lista de tabelas

Tabela 1 – Comparativo dos trabalhos relacionados com abordagem CNN	22
Tabela 2 – Quantitativos de imagens disponibilizadas por tipo de dano	40
Tabela 3 – Modelos pré-treinados com o ImageNet	45
Tabela 4 – Experimentos definidos	49
Tabela 5 – Especificações do ambiente experimental	52
Tabela 6 – Resultados do experimento para definição do melhor modelo refinado .	53
Tabela 7 – Configurações do experimento para a extração dos recortes	54
Tabela 8 – Configurações do experimento para a separação dos conjuntos	56
Tabela 9 – Configurações do experimento para a uso ou não de aumento de dados e balanceamento	62
Tabela 10 – Resultados dos experimentos realizados	63
Tabela 11 – Principais parâmetros da ferramenta de rotulação e recorte	82
Tabela 12 – Tempo de classificação de todos os recortes	89
Tabela 13 – Tempo de treinamento dos modelos	90
Tabela 14 – Tempo de geração do mapa de calor do GradCam	91

Lista de Algoritmos

1	Define posicionamento das grades na imagem	84
---	--	----

Lista de abreviaturas e siglas

CNN	<i>Convolutional Neural Networks</i>
DNN	<i>Deep Neural Network</i>
DTL	<i>Deep Transfer Learning</i>
DTU	<i>Drone Inspection Dataset</i>
GAN	<i>Generative Adversarial Network</i>
GPU	<i>Graphics Processing Unit</i>
GV	Gerador de vórtice
IoT	<i>Internet of Things</i>
px	<i>Pixel</i>
RGB	<i>Red-Green-Blue</i>
SGD	<i>Stochastic Gradient Descent</i>
UAV	<i>Unmanned Aerial Vehicle</i>

Sumário

1	INTRODUÇÃO	16
1.1	Motivação para a dissertação	19
1.2	Justificativa	19
1.3	Hipóteses	20
1.4	Objetivos	21
1.4.1	Objetivo geral	21
1.4.2	Objetivos específicos	21
1.5	Trabalhos relacionados	21
1.6	Produção científica	25
1.7	Estrutura da dissertação	25
2	FUNDAMENTAÇÃO TEÓRICA	26
2.1	Danos em pás eólicas	26
2.1.1	Descargas atmosféricas	26
2.1.2	Erosão do bordo de ataque	26
2.1.3	Danos na pintura	27
2.1.4	Trincas	28
2.1.5	Outros danos	28
2.2	Visão computacional	29
2.3	Redes neurais convolucionais	30
2.3.1	VGG	31
2.3.2	ResNet	31
2.3.3	MobileNet	32
2.3.4	EfficientNet	32
2.3.5	Base de imagens ImageNet	33
2.4	Aprendizagem por transferência	33
2.5	Aumento de dados	34
2.6	Métodos de balanceamento de classes em CNNs	36
2.7	Considerações finais	37
3	MÉTODO PROPOSTO	39
3.1	Etapa de treinamento	39
3.1.1	Imagens originais	39
3.1.2	Extração dos recortes	40
3.1.2.1	Extração baseada em janela deslizante	41
3.1.2.2	Extração baseada em posição aleatória	42

3.1.3	Conjuntos de dados	43
3.1.3.1	Aumento de dados e balanceamento	44
3.1.4	Refinamento dos modelos pré-treinados	45
3.1.5	Seleção do melhor modelo	46
3.1.6	Configurações de treinamento	47
3.2	Experimentos definidos	49
3.3	Etapas de identificação de danos	49
3.3.1	Votação e classificação das imagens originais	50
3.3.2	Localização dos danos nas imagens originais	50
3.4	Considerações finais	50
4	RESULTADOS E DISCUSSÕES	52
4.1	Especificações do ambiente experimental	52
4.2	Experimento I: análise do método de refinamento dos modelos	53
4.3	Experimento II: análise do método de extração dos recortes	54
4.4	Experimento III: análise do método de separação dos conjuntos de dados	56
4.5	Experimento IV: análise do método de aumento de dados e balanceamento	62
4.6	Resultados dos experimentos	63
4.7	Resultados no conjunto de validação	63
4.8	Resultados da classificação das imagens originais	67
4.9	Resultados da localização dos danos nas imagens originais	70
4.10	Considerações finais	74
5	CONCLUSÃO E TRABALHOS FUTUROS	75
5.1	Conclusão	75
5.2	Trabalhos futuros	75
	REFERÊNCIAS	77
	APÊNDICES	81
	APÊNDICE A – FERRAMENTA PARA TRATAMENTO E ANÁLISE DAS IMAGENS ORIGINAIS	82
A.1	Módulo de rotulação	82
A.2	Módulo de recortes aleatórios	83
A.3	Módulo de predição	85
A.4	Módulo de análise Grad-CAM	87

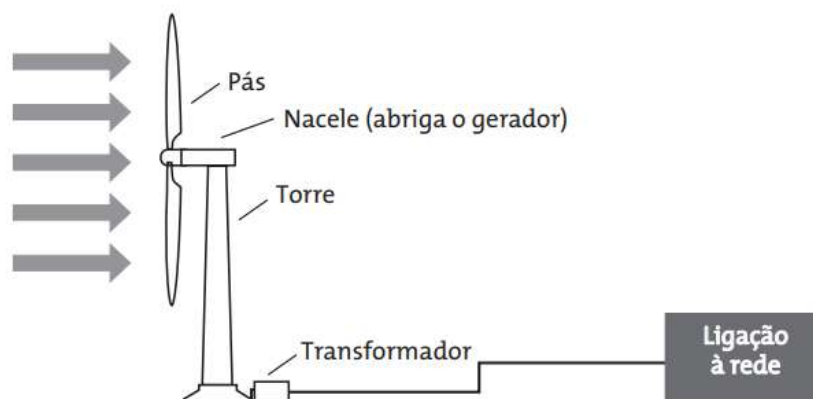
A.5	Módulo de segmentação SAM	87
A.6	Módulo de treinamento	88
	APÊNDICE B – TEMPOS DE CLASSIFICAÇÃO E TREINAMENTO	89
B.1	Tempos de classificação	89
B.2	Tempos de treinamento	89
B.3	Tempos do Grad-CAM	90

1 Introdução

As fontes de energias sustentáveis são cada vez mais importantes para o futuro da humanidade, dentre as quais destaca-se a energia eólica. Segundo o Conselho Global de Energia Eólica, para esse tipo de energia, existe uma expectativa de crescimento médio de 9% a cada ano até 2028 (Global Wind Energy Council, 2024).

A produção de energia eólica é realizada pelo aerogerador, cujos componentes estão representados na Figura 1, tendo: a turbina eólica, com seu conjunto de pás responsáveis por converter a energia cinética dos ventos em energia cinética no eixo do gerador; a nacele, que abriga o gerador e sistemas auxiliares; a torre, responsável por sustentar a turbina e a nacele; e o transformador, que adapta o nível de tensão do gerador ao da rede elétrica. Existem ainda cabines com sistemas de monitoramento, supervisão e controle eletrônicos, além de conversores eletrônicos de potência, que geralmente são instalados na base da torre (Lage; Processi, 2013).

Figura 1 – Componentes de um aerogerador

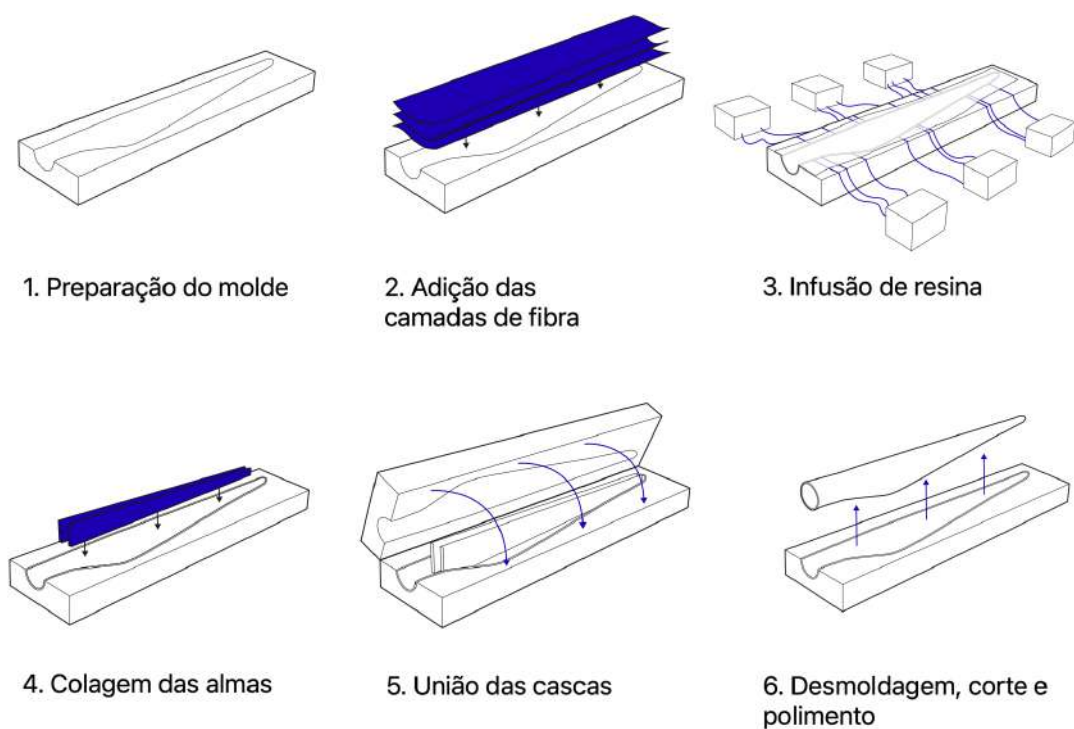


Fonte: Lage e Processi (2013).

O conjunto de pás é um importante componente do aerogerador, uma vez que é responsável pela captação da energia cinética dos ventos, além de ser o componente com o custo mais alto do aerogerador. Geralmente, são feitos de materiais compósitos (fibra de carbono, fibra de vidro, entre outros), com características de baixo peso, alta resistência e flexibilidade (Mishnaevsky et al., 2017). Na Figura 2 são ilustradas as etapas genéricas da produção de uma pá eólica. Iniciando com moldes que recebem camadas (chamadas sanduíches) de fibra de vidro ou de carbono, que são preenchidas com resina e, ao final, recebem o nome de cascas; as cascas são coladas junto a estruturas rígidas chamadas de almas; e, por fim, recebem os acabamentos finais de cortes e polimento

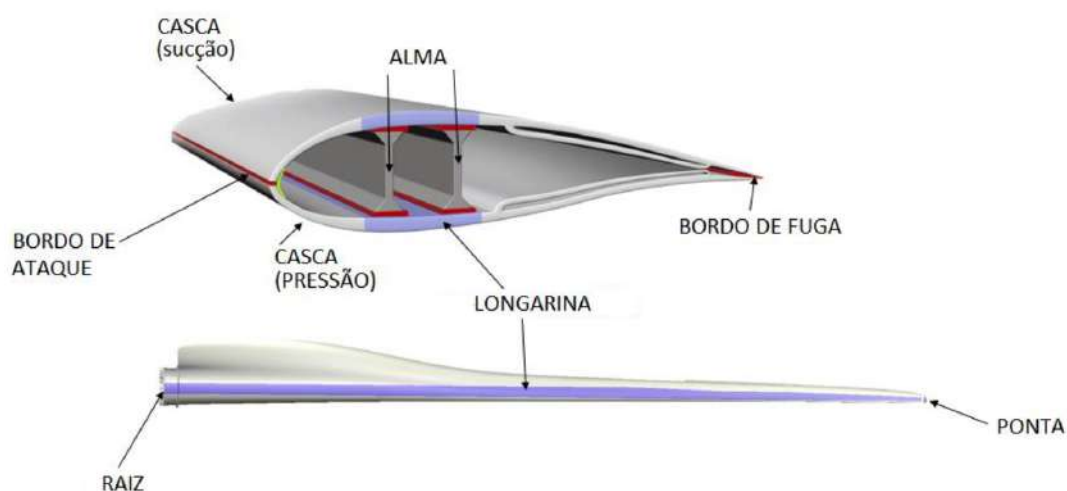
(Bladena; KirtxThomsen, 2019). O lado da pá que recebe o fluxo de vento é chamado de bordo de ataque, enquanto o lado oposto corresponde à saída do vento e é chamado de bordo de fuga. Na Figura 3 são ilustradas as principais partes de uma pá eólica.

Figura 2 – Etapas genéricas da produção de uma pá eólica



Fonte: adaptado de Bladena e KirtxThomsen (2019).

Figura 3 – Principais partes de uma pá eólica



Fonte: adaptado de Eólica (2020).

As pás são os componentes mais suscetíveis a danos (Kaewniam et al., 2022) e, segundo Scotland Against Spin (2024), são responsáveis por 9,5% (549 de 5.778) de todos os incidentes registrados mundialmente até 2024. Segundo Wang et al. (2022), os tipos de danos mais comuns são: trincas no bordo de fuga e na casca; danos por descargas atmosféricas; erosão no bordo de ataque; rotação desbalanceada ou quebra por danos de congelamento da pá; e delaminação das camadas de sanduíches. Na Figura 4 é ilustrado um dano grave em uma pá eólica causado por uma descarga atmosférica.

Figura 4 – Dano em uma pá eólica causado por uma descarga atmosférica



Fonte: Katsaprakakis, Papadakis e Ntintakis (2021).

Esses danos podem levar a paradas não programadas dos aerogeradores, gerando grandes custos de manutenção e operação, podendo até mesmo causar uma perda total do equipamento, com grandes prejuízos financeiros (Wang et al., 2022). Segundo Turnbull e Carroll (2021), o prejuízo médio (com receita perdida, custos de transporte, de pessoal e reparos) de um aerogerador é de 30,87 libras por megawatt-hora. Por isso, é muito importante identificar esses danos o quanto antes, a fim de evitar o alastramento do dano e até acidentes. Técnicas de detecção prematura de danos, como ferramentas de monitoramento, reduzem os custos de manutenção e aumentam a disponibilidade dos equipamentos. Segundo Kaewniam et al. (2022), existem técnicas baseadas em:

- medição de deformação da pá eólica: que são utilizadas para detectar fadiga, rachaduras, delaminação e outros tipos de falhas;
- emissão acústica: que detecta ondas sonoras geradas por microfissuras, delaminações ou falhas estruturais dentro da pá;

- ultrassom: que utiliza ondas sonoras de alta frequência para detectar descon- tinuidades internas nas pás, como rachaduras, bolhas de ar e delaminação de materiais compósitos;
- vibração: que monitora as mudanças nas frequências naturais da pá para identi- ficar desgastes, falhas estruturais ou desalinhamentos;
- termografia: que identifica falhas estruturais ao detectar variações térmicas na superfície das pás, como rachaduras, delaminações e regiões de adesão deficiente; e
- visão computacional: que usa imagens de alta resolução e inteligência artificial para detectar danos superficiais nas pás de turbinas eólicas.

As técnicas baseadas em visão computacional possuem alta precisão, praticidade e requerem relativamente poucos recursos materiais. (Kaewniam et al., 2022; Du et al., 2020) e, conseqüentemente, se tornaram foco de diversos estudos (Zhang; Shu, 2024).

1.1 Motivação para a dissertação

As inspeções visuais das pás de aerogeradores são geralmente realizadas com a ajuda de um Veículo Aéreo Não Tripulado, ou na língua inglesa *Unmanned Aerial Vehicle (UAV)*, e necessitam de pelo menos dois técnicos presentes no local, sendo um deles um piloto qualificado e o outro um operador de câmera de bordo. A responsabilidade desses técnicos é capturar diversas imagens de cada pá eólica e enviá-las para um especialista em danos (Cao; Shi; Zhang, 2021).

Esse especialista precisa analisar uma quantidade muito grande de imagens, sendo que apenas uma parte delas apresenta algum dano a ser identificado. Após realizada a filtragem das imagens com danos, é necessário classificar o tipo de dano e sua localização na imagem. E, finalmente, decidir o melhor planejamento para a sua correção. Esse trabalho é realizado verificando visualmente cada uma das imagens, nas quais o especialista deve avaliar características da pá eólica em busca de um dano, sendo bastante repetitivo e exaustivo, com propensão a erros de avaliação devido a fadiga visual.

Um modelo de visão computacional treinado para detectar danos pode agilizar a análise dos especialistas, eliminando automaticamente imagens sem danos e sinalizando a localização dos danos nas demais, tornando o processo mais rápido e confiável.

1.2 Justificativa

Nos últimos anos, diferentes trabalhos têm sido propostos para a inspeção e detecção de danos em pás eólicas utilizando aprendizado de máquina, visão computacional e Redes Neurais Convolucionais (*Convolutional Neural Network - CNN*). Um ponto em comum

entre os trabalhos é a necessidade de treinar os modelos com imagens de tamanhos menores, uma vez que as CNNs pré-treinadas têm parâmetros de entrada reduzidos. Assim, é necessário redimensionar as imagens ou extrair recortes da imagem original. Os métodos existentes realizam a “**extração dos recortes baseadas em janela deslizante**” ou extração manual da área de interesse. Esses recortes são separados nos conjuntos de treinamento, teste e validação, com todos os recortes de uma mesma imagem ficando no mesmo conjunto de dados. Neste trabalho, essa técnica é nomeada como “**separação dos conjuntos de dados pela imagem original**”.

Nesta dissertação é introduzido um método de treinamento baseado na técnica de “**extração baseada em recortes aleatórios**” e “**separação dos conjuntos de dados pelos recortes**”, onde todos os recortes são embaralhados e depois separados nos conjuntos de dados. Diferentemente das abordagens convencionais, essas técnicas permitem que a rede neural tenha como entrada amostras mais diversificadas durante a etapa de treinamento, o que contribui para:

- melhor generalização do modelo, pois os recortes extraídos capturam diferentes regiões da pá, aplicando nativamente transformações de rotação e redimensionamento;
- possibilidade da aplicação de diferentes distribuições na posição dos recortes aleatórios. Como exemplo, uma distribuição gaussiana, que permite ter um foco mais central da imagem, diminuindo assim a influência do *background*; e
- realizar a extração de uma maior quantidade de recortes de tamanho fixo pré-determinado, uma vez que a técnica baseada em janelas deslizantes possui um limitador na quantidade de recortes extraídos de mesmo tamanho.

A técnica de recortes aleatórios também pode ser aplicada em outras áreas da inspeção automatizada de estruturas, reforçando a relevância deste estudo para o avanço da visão computacional na identificação de danos.

1.3 Hipóteses

Considerando os desafios envolvidos na identificação automatizada de danos em pás de aerogeradores e nas lacunas observadas em trabalhos anteriores, são estabelecidas hipóteses que orientam a investigação conduzida nesta dissertação. Essas hipóteses buscam verificar se determinadas estratégias — como a extração de recortes aleatórios, a aplicação de técnicas de aumento de dados e o balanceamento das classes — influenciam positivamente nos resultados de modelos baseados em redes neurais convolucionais profundas. A seguir, são apresentadas as hipóteses formuladas:

- **H1** - a separação dos conjuntos de dados pelos recortes aleatórios melhora a capacidade de generalização do modelo ao aumentar a diversidade das amostras;

- **H2** - a aplicação de técnicas de aumento de dados, como rotação, espelhamento e ajuste de brilho, amplia a robustez do modelo e melhora seu desempenho na classificação de danos estruturais em diferentes condições de iluminação e perspectiva;
- **H3** - o balanceamento do conjunto de dados mitiga o impacto de classes desproporcionais, potencializando a melhora dos resultados de classificação; e
- **H4** - a classificação de recortes extraídos da imagem original permite localizar o dano na imagem original.

1.4 Objetivos

1.4.1 Objetivo geral

Propor um método para realizar o refinamento de um modelo de visão computacional, para identificar danos estruturais superficiais em pás de aerogeradores e facilitar a sua localização na imagem original através de um mapa de calor.

1.4.2 Objetivos específicos

No desenvolvimento desta dissertação, outros objetivos específicos devem ser atingidos:

- avaliar a extração dos recortes com o uso da abordagem de recortes aleatórios, analisando o impacto da variação da quantidade de recortes extraídos nos resultados;
- definir o conjunto das melhores técnicas (configuração) para o refinamento de um modelo;
- classificar a imagem original com o modelo treinado na melhor configuração encontrada; e
- localizar os danos identificados através de um mapa de calor.

1.5 Trabalhos relacionados

Os principais trabalhos publicados são detalhados a seguir e na Tabela 1 é apresentado um comparativo dos trabalhos relacionados que utilizam modelos com CNN para a identificação de danos.

Tabela 1 – Comparativo dos trabalhos relacionados com abordagem CNN

Referência	Classes identificadas	Modelos	Extração dos recortes	Tipo de separação dos conjuntos de dados	Aumento de dados
Moreno et al. (2018)	1. descargas elétricas 2. danos em pinturas 3. quebras 4. sem dano	1. CNN projetada	Não especifica a técnica de recorte	Imagem original	Não utiliza
Shihavuddin et al. (2019)	1. erosão 2. GV sem dentes 3. GV normal 4. receptor de raios 5. normal	Faster R-CNN com: 1. Inception-V2 2. ResNet-50 3. ResNet-101 4. Inception-ResNet-V2	Janela deslizante	Imagem original	Espelhamento horizontal e vertical; mudança de perspectiva; ajuste de contraste; desfoque gaussiano; e redimensionamento.
Yang et al. (2020)	1. normal 2. fundo da imagem 3. rachaduras 4. buracos 5. dano misto	1. AlexNet 2. ResNet50	Não especifica a técnica de recorte	Imagem original	Não utiliza
Foster et al. (2022)	1. dano 2. sujeira	1. Faster R-CNN com ResNet-101 2. YOLOv5	Não especifica a técnica de recorte	Pelos recortes	Aumento HSV; translação; redimensionamento; espelhamento horizontal; e mosaico.
Agrawal et al. (2024)	1. trincas	1. EfficientNetB3 2. MobileNetV3 3. ResNet-18	Janela deslizante	Imagem original	Espelhamento horizontal e vertical; ajuste de brilho, contraste, saturação e matiz; e <i>random affine transformation</i> .
Método proposto	1. descargas elétricas 2. erosão 3. danos de pintura 4. trincas 5. normal	1. EfficientNetB2 2. MobileNetV2 3. ResNet-50 4. VGG16	Janela deslizante e recortes aleatórios	Imagem original e pelos recortes	Espelhamento horizontal e vertical; ajuste de brilho, contraste, rotação e redimensionamento.

Fonte: elaborada pelo autor.

Moreno et al. (2018) propõe um processo de inspeção de pás em turbinas eólicas com um modelo de Redes Neurais Convolucionais personalizado, e previamente treinado, para a detecção e classificação de três possíveis tipos de danos: descargas elétricas, danos em pinturas e quebras. O modelo de CNN projetado pelos autores é construído com 3 conjuntos de camadas, cada uma com os seguintes componentes: camada convolucional, *Batch Normalization*, *ReLU*, *Max Pooling* e, por fim, uma camada densa de classificação com *softmax*. A métrica de validação de desempenho utilizada é a acurácia. Para o treinamento do modelo construído, os autores utilizam 78 imagens públicas coletadas da Internet. São realizadas 50 épocas de treinamento e aplicado o método *Stochastic Descent Gradient (SGD)* para a otimização. Para validar o modelo treinado, os autores constroem um protótipo de pá eólica em uma impressora 3D e os danos são emulados nesse protótipo. A acurácia final obtida é de 81,25%.

No trabalho desenvolvido por Yang et al. (2020) é aplicado um método de classificação de aprendizagem profunda utilizando como camadas de extrações de características o ResNet50 (He et al., 2016) para a classificação de danos nas pás. É utilizado um drone para fazer a aquisição das imagens, e são produzidas 1.594 imagens a partir de 20 diferentes aerogeradores. Em seguida, são feitos recortes em tamanhos menores que geram 557.900 novas imagens, das quais são selecionadas 13.200 para compor os conjuntos de dados de treino e teste. O treinamento é realizado em 40 épocas com o SGD *with momentum* para a otimização. O algoritmo demonstra melhor desempenho do que a rede *AlexNet* (Krizhevsky; Sutskever; Hinton, 2012), com acurácia final de 95,58%. O modelo é construído em MATLAB e preparado para classificar cinco tipos de imagens: normal, fundo da imagem, rachaduras, buracos e dano misto.

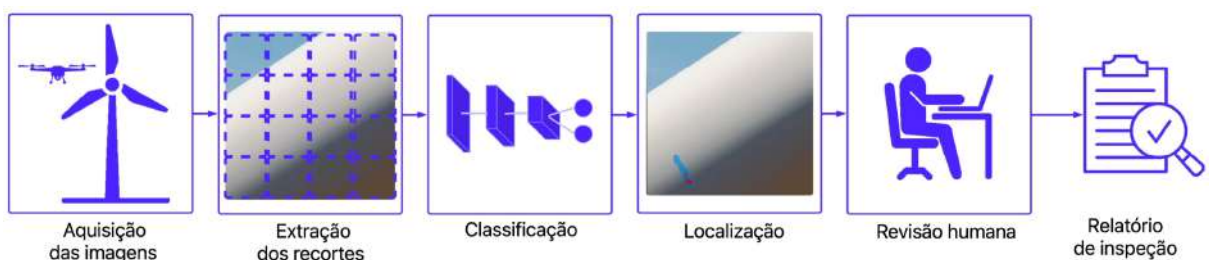
Shihavuddin et al. (2019) utilizam imagens capturadas por drones e aumentadas por técnicas de aumento de dados para treinar o modelo *Faster R-CNN* (Ren et al., 2015) com alguns extratores de características (*Inception-V2*, *ResNet-50*, *ResNet-101* e *Inception-ResNet-V2*) para detectar danos nas pás eólicas usando caixas delimitadoras. As imagens adquiridas são rotuladas manualmente por especialistas em quatro tipos de imagens: erosão, gerador de vórtice (GV) normal, GV sem dentes e receptor de raios. A métrica para avaliação do modelo é a média aritmética das precisões médias (*Mean Average Precision - mAP*). Os autores constatam que o uso de técnicas de aumento de dados e de arquiteturas mais profundas de CNN melhora o desempenho do modelo. O modelo *Faster R-CNN* com *Inception-ResNet-V2* (arquitetura mais profunda) tem o melhor resultado com um mAP de 81,10%. Uma outra contribuição do trabalho é produzir um conjunto de dados de imagens públicas de uma inspeção de turbinas eólicas com 701 imagens em alta definição, porém não rotuladas, denominado *Drone Inspection Dataset (DTU)* (Shihavuddin; Chen, 2018).

Foster et al. (2022) estendem o trabalho de Shihavuddin et al. (2019) e criam um

novo conjunto de dados de imagens públicas rotuladas derivadas do DTU. Os autores fornecem mais de 13.000 recortes das imagens originais no tamanho de 586×371 px², nos quais rotulam manualmente 3.000 instâncias das classes dano e sujeira. O objetivo dos autores também é detectar os danos nas pás eólicas usando caixas delimitadoras. A partir de seu conjunto de dados, os autores comparam o desempenho dos modelos *Faster R-CNN*, com o extrator de características ResNet-101, com o modelo YOLOv5 (Redmon et al., 2016). Embora a rotulação das imagens seja manual, os autores não deixam claro se essas rotulações são validadas por algum especialista. Nas imagens, são aplicadas técnicas de aumento de dados e utilizadas para treinar o modelo YOLOv5 em 1.000 épocas e *batch size* de 16, e o detector *Faster R-CNN* em 100.000 épocas e *batch size* de 512. O SGD é utilizado como otimizador e o modelo YOLOv5s tem o melhor resultado com 0,7937 na métrica mAP.

Por último, Agrawal et al. (2024) propõe um método para a identificação (classificação e localização) de danos focado em trincas quase invisíveis, utilizando técnicas de visão computacional e aprendizagem profunda, detalhando as etapas desde a curadoria dos dados até o treinamento do modelo. Na Figura 5 é ilustrado o método proposto por Agrawal et al. (2024). A aquisição é realizada a partir de inspeções globais e voltada especificamente para danos de trincas de alta severidade, que dá origem a um grande conjunto de dados chamado “*Zeitview Crack Detection (ZVCD) dataset*”. A partir das imagens desse conjunto são extraídos recortes no tamanho de 1024×1024 px² utilizando a técnica de janela deslizante e com aplicação de aumento de dados. Os recortes são utilizados para o refinamento de 3 modelos pré-treinados (ResNet-18, EfficientNet-B3 e MobileNetV3) no conjunto ImageNet (Deng et al., 2009). Neste caso, os modelos são utilizados para classificar os recortes e as métricas analisadas são: acurácia, precisão, sensibilidade e escore F1. A classificação dos recortes indica a posição aproximada do dano na imagem original, o que permite aplicar o Grad-CAM (Selvaraju et al., 2017) para localizar o dano nos recortes, sendo avaliada por um especialista que emite um relatório da inspeção. Os melhores resultados obtidos são: acurácia e sensibilidade de 82,20 para EfficientNet-B3; escore F1 de 74,42 para MobileNetV3-S; e precisão de 70,97 para MobileNetV3-L.

Figura 5 – Método proposto por Agrawal et al. (2024)



Fonte: adaptado de Agrawal et al. (2024).

Observa-se que os trabalhos que disponibilizam conjuntos de dados públicos de imagens ainda são muito recentes, não havendo ainda um padrão para *benchmark* estabelecido. Ademais, cada trabalho apresenta especificidades, aprofundando-se em determinados tipos específicos de danos, conforme a realidade enfrentada por cada parque eólico nas diversas localidades do mundo. O artigo de Agrawal et al. (2024), embora tenha semelhanças com o método proposto nesta dissertação (como exemplo: classificação e localização através dos recortes, métricas e modelos utilizados), não explora outras abordagens de treinamentos (como a extração baseada em recortes aleatórios ou separação dos conjuntos de dados pelos recortes) e tipos de danos (descargas atmosféricas, erosão e danos de pintura) que estão sendo explorados no método proposto.

1.6 Produção científica

Durante o desenvolvimento desta pesquisa, o seguinte trabalho científico foi aceito e publicado, a saber:

- **OLIVEIRA, A. R. A.; MEDEIROS, C. M. S.; RAMALHO, G. L. B.. Damage Identification of Wind Turbine Blades.** In: SPRINGER. *Brazilian Conference on Intelligent Systems*. Belém, PA, Brazil. 2024. p. 187–200. https://doi.org/10.1007/978-3-031-79035-5_13

1.7 Estrutura da dissertação

O restante desta dissertação está organizado da seguinte forma: no Capítulo 2 são apresentados os fundamentos teóricos para um melhor entendimento da dissertação, apresentando conceitos sobre os tipos e possíveis causas dos danos existentes nas pás de aerogeradores, além de conceitos sobre visão computacional, redes neurais profundas e técnicas de pré-processamento de dados. No Capítulo 3 é apresentada a metodologia proposta na dissertação, contendo todas as técnicas utilizadas no método proposto para classificação e localização de danos nas pás dos aerogeradores. Já no Capítulo 4, os resultados são apresentados e é realizada sua discussão. Finalmente, no Capítulo 5 estão as conclusões e propostas de trabalhos futuros.

2 Fundamentação teórica

Neste Capítulo são descritas as características dos danos em pás eólicas, abordando os danos identificados pelo método proposto; visão computacional e redes neurais convolucionais, detalhando as arquiteturas dos modelos pré-treinados utilizados; técnicas de aumento de dados, para lidar com a pequena quantidade de amostras dos conjuntos de dados; aprendizagem por transferência, que tem o objetivo de otimizar o tempo de treinamento dos modelos; e, por fim, como tratar o desbalanceamento de classes em CNNs.

2.1 Danos em pás eólicas

São dois os fatores responsáveis pelos danos nas pás eólicas: os decorrentes de falhas humanas no processo de sua fabricação e os externos, como descargas atmosféricas, acúmulo de gelo, carga irregular, umidade, rajadas de ventos fortes, entre outros. Os danos afetam o desempenho de geração de energia do aerogerador, levando-o à queda na produção de energia e, em último caso, à perda total do equipamento. Nas seções seguintes são descritas as características dos principais danos.

2.1.1 Descargas atmosféricas

A maioria das pás eólicas é equipada com dispositivos de proteção contra descargas atmosféricas. No entanto, esses sistemas não são totalmente eficazes na prevenção de todos os incidentes. Geralmente, os danos causados pelas descargas atmosféricas são buracos e queimaduras na área ao redor do ponto no qual o raio atinge a pá (Garolera et al., 2014). Esses danos apresentam características distintas, incluindo escurecimento da pintura devido às queimaduras, alteração na textura causada pelo descolamento da fibra de vidro e áreas de danos maiores. A gravidade do dano depende da extensão do impacto do raio. Na Figura 6 são ilustrados exemplos de descargas atmosféricas.

2.1.2 Erosão do bordo de ataque

As pás eólicas estão constantemente expostas a colisões repetitivas com partículas como gotas de chuva, granizo, areia, insetos, poeira, sal e óleo. Esse impacto contínuo provoca erosão e corrosão na região que recebe o vento diretamente, conhecida como bordo de ataque da pá. A erosão compromete a eficiência aerodinâmica da pá, reduzindo sua capacidade de gerar energia elétrica. Embora inicialmente o dano seja pequeno e superficial, tende a se agravar com o tempo, espalhando-se ao longo do bordo de ataque. Esse tipo de

dano apresenta uma superfície com textura rugosa, esfoliada e desgastada. Na Figura 7 são ilustrados exemplos de erosões.

Figura 6 – Exemplos de danos por descargas atmosféricas



Fonte: adaptado de Garolera et al. (2014).

Figura 7 – Exemplos de danos de erosão



Fonte: Katsaprakakis, Papadakis e Ntintakis (2021).

2.1.3 Danos na pintura

Os danos na pintura são geralmente superficiais e frequentemente classificados como “cosméticos”, pois não comprometem a estrutura da pá. Exemplos comuns incluem descascamento da pintura, manchas de óleo, entre outros. Esse tipo de dano pode evoluir para um dano de erosão, embora geralmente não precise de uma intervenção para manutenção, mas deve ser acompanhado. Na Figura 8 são ilustrados exemplos de danos de pintura.

Figura 8 – Exemplos de danos de pintura



Fonte: imagens disponibilizadas por empresa fabricante de aerogeradores.

2.1.4 Trincas

As trincas são o resultado de fadigas nos materiais que compõem as pás devido ao uso prolongado (Wang et al., 2022). Inicialmente pequenas e superficiais, essas fissuras tendem a se expandir e aprofundar com o tempo, podendo comprometer toda a estrutura da pá. Geralmente, apresentam um formato fino e longitudinal, com bordas nítidas e ásperas. Em alguns casos, observa-se ainda uma textura semelhante a uma teia de aranha. Na Figura 9 são ilustrados exemplos de trincas.

Figura 9 – Exemplos de danos de trincas



Fonte: Wang et al. (2022).

2.1.5 Outros danos

Embora os danos a seguir não sejam o escopo do trabalho, são apresentados pois existe a possibilidade de sua ocorrência e poderão ser tratados por algum trabalho futuro, caso tenham imagens disponibilizadas.

A delaminação é geralmente causada por problemas de fabricação, transporte ou instalação e caracteriza-se pelo descolamento das camadas que compõem a estrutura da

pá. Com isso, têm-se a redução da rigidez e resistência da pá, podendo até mesmo levar à quebra da pá.

Alguns parques eólicos encontram-se localizados em regiões que possuem invernos rigorosos, como: o sul do Brasil, países nórdicos ou regiões mais próximas dos polos. Os aerogeradores instalados nestes parques eólicos podem sofrer com a produção de gelo em seus componentes. O acúmulo de gelo na pá eólica é uma decorrência da convergência de fatores ambientais relacionados com temperatura, umidade, densidade do ar, velocidade do vento, entre outros. Esse gelo pode causar turbulência de ar, desbalanceamento ou vibração irregular, levando à fadiga e quebra da pá. Tudo isso reduz a produção de energia. Na Figura 10 são ilustrados exemplos de danos por gelo.

Figura 10 – Exemplos de danos por gelo

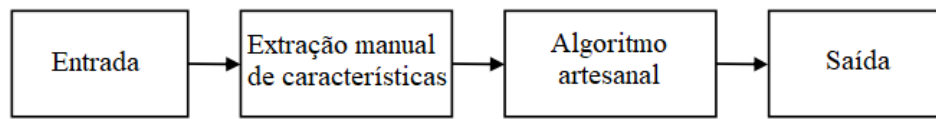


Fonte: Wang et al. (2022).

2.2 Visão computacional

Segundo Rosenfeld (1988), visão computacional é a habilidade de extrair informações de uma cena, a partir de imagens, através de um processamento computacional. Para que isso ocorra, é necessária a aplicação de técnicas computacionais consideradas de baixo-nível e de alto-nível (Dana; Christopher, 1982). As técnicas ditas de baixo-nível concentram-se em atividades tais como: reconhecimento do brilho, cores e profundidade; percepção de padrões geométricos (linhas, curvas, bordas); discriminação do fundo de uma imagem; entre outras. No entanto, a habilidade de interpretar uma cena é considerada de alto-nível. Um ponto-chave para realizar essa interpretação é entender as características contidas dentro de uma imagem, sendo necessário um modelo matemático para fazer a extração e interpretação dessas características. Na Figura 11 são ilustradas três abordagens de visão computacional: as que usam técnicas tradicionais; com aprendizagem de máquina; e com aprendizagem profunda (Szeliski, 2022). Os modelos baseados em aprendizagem profunda possuem os melhores resultados atualmente, pois integram os processos de baixo e alto nível em uma única estrutura, conforme ilustrado na Figura 11 (c).

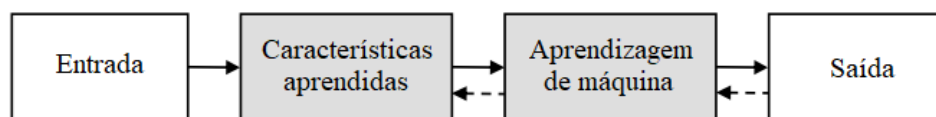
Figura 11 – Abordagens para visão computacional



(a) Fluxo tradicional da visão computacional



(b) Fluxo do aprendizado de máquina tradicional



(c) Fluxo do aprendizado profundo

Fonte: adaptado de Szeliski (2022).

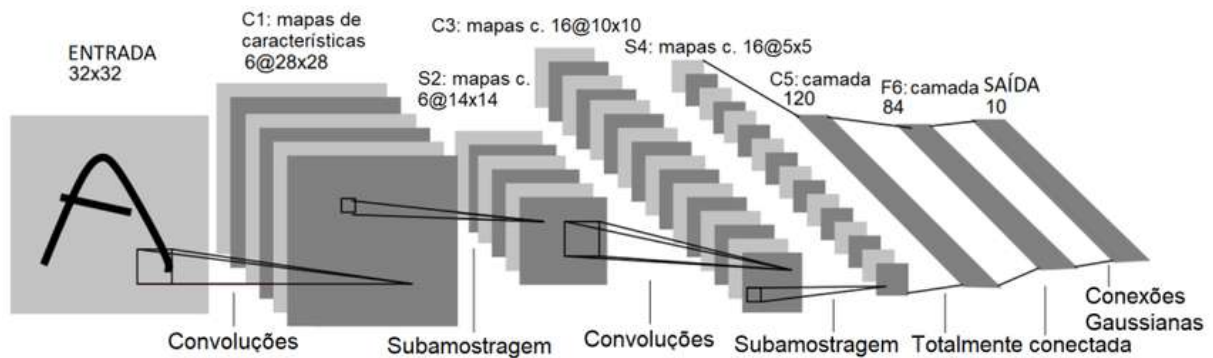
2.3 Redes neurais convolucionais

De acordo com Cichy e Kaiser (2019), as redes neurais profundas (*Deep Neural Networks – DNNs*) são modelos computacionais que abordam tarefas do mundo real nas quais os humanos se destacam, como reconhecimento de objetos, processamento de fala e planejamento cognitivo. Essas redes consistem na utilização de muitas camadas de neurônios artificiais, inspirados no funcionamento dos neurônios cerebrais humanos. A arquitetura de uma DNN é composta por três componentes: neurônios, camadas e conexões entre camadas (Khan; Laghari; Awan, 2021). Cada neurônio atribui um peso às informações recebidas dos neurônios das camadas anteriores, tendo como saída resultante a soma das entradas multiplicadas pelo seu respectivo peso, acrescido de um valor de correção (*bias*), aplicado por uma função de ativação. A grande diferença de uma rede neural padrão para uma DNN é a grande quantidade de camadas utilizadas nesta última (Cichy; Kaiser, 2019).

Um tipo particular de camada aplicada no tratamento de imagens é o operador de convolução, que origina as redes neurais convolucionais (*Convolutional Neural Networks – CNNs*). As camadas de convolução extraem características de uma imagem de entrada, repassando o resultado às camadas posteriores para realizar o aprendizado de fato. Cada camada de convolução corresponde a um conjunto de filtros responsáveis por detectar características diferentes na imagem, que são aprendidas durante o treinamento da rede. A cada época de treinamento, os filtros vão se especializando. Antes da aplicação de novos filtros na próxima camada de convolução, tipicamente é aplicada uma etapa de amostragem, chamada de *pooling*, que tem o objetivo de reduzir a resolução espacial do

mapa de atributos a ser tratado pelas camadas posteriores. As CNNs são atualmente muito utilizadas nas aplicações de visão computacional, tendo sido bem-sucedidas inicialmente no reconhecimento de dígitos na rede LeNet-5 por LeCun et al. (1998), cuja arquitetura é ilustrada na Figura 12.

Figura 12 – Arquitetura da rede LeNet-5



Fonte: adaptado de LeCun et al. (1998).

O sucesso inicial da LeNet-5 levou ao desenvolvimento de diversas outras CNNs ao longo dos últimos anos, como VGG (Simonyan; Zisserman, 2014), ResNet (He et al., 2016), MobileNet (Howard et al., 2017), EfficientNetB2 (Tan; Le, 2019), entre outras. Nas subseções que seguem são apresentadas maiores informações sobre essas redes CNNs.

2.3.1 VGG

A arquitetura VGG foi apresentada por Simonyan e Zisserman (2014) que adotou uma abordagem de utilizar camadas convolucionais com filtros pequenos (3×3) e uma maior profundidade de camadas, até então não aplicada em outros modelos existentes. Os dois melhores modelos encontrados possuem 16 (VGG-16) e 19 (VGG-19) camadas de profundidade. Essa arquitetura permitiu que os autores obtivessem destaque no *ImageNet Large Scale Visual Recognition Challenge 2014 - ILSVRC* (Russakovsky et al., 2015), que foi uma competição anual de visão computacional que ocorreu de 2010 a 2017.

À época, o modelo VGG tornou-se referência na área de redes neurais convolucionais devido à sua simplicidade estrutural e capacidade de generalização para a classificação de imagens. Porém, a principal limitação dessa arquitetura é o alto número de parâmetros, com aproximadamente 138 milhões, que aumentam os custos do tempo de treinamento e da predição.

2.3.2 ResNet

Uma problemática existente com redes muito profundas é a dissipação do gradiente, que acontece na etapa de treinamento, durante a atualização dos pesos dos neurônios

das camadas ocultas. O gradiente tende a se tornar cada vez menor à medida que é retropropagado pelas camadas ocultas da rede, o que faz com que os neurônios das camadas mais iniciais aprendam de forma significativamente mais lenta do que aqueles nas camadas posteriores.

Para resolver esse problema, He et al. (2016) desenvolveu a arquitetura ResNet que é composta por blocos residuais, que permitem a passagem direta de informações entre camadas distantes, facilitando a aprendizagem de redes profundas e evitando a dissipação do gradiente. As implementações mais comuns da arquitetura ResNet são: ResNet18, ResNet34, ResNet50, ResNet101 e ResNet152, com os números indicando a quantidade de camadas treináveis.

2.3.3 MobileNet

Com a dissipação do gradiente já tendo sido resolvida por outras arquiteturas, Howard et al. (2017) estudou como reduzir a quantidade de parâmetros dos modelos e apresentou a arquitetura MobileNet, que visa aplicações em dispositivos móveis e sistemas embarcados. Com a adoção das convoluções *depthwise* separáveis, a MobileNet consegue reduzir significativamente o número de operações, tornando-a mais leve.

Sandler et al. (2018) evoluiu a primeira versão e apresentou a MobileNetV2, que trouxe como novidade os blocos de convolução invertidos com conexões residuais (*Inverted Residual Blocks*), que primeiro expandem a dimensionalidade dos canais, aplicam a convolução *depthwise* e depois reduzem novamente a dimensionalidade, trazendo mais generalização e tornando a rede mais eficiente.

A MobileNet é leve e bastante eficiente em termos computacionais, permitindo execução em smartphones e IoT. Porém, pode ter desempenho inferior em comparação com modelos com mais parâmetros em tarefas complexas.

2.3.4 EfficientNet

Tan e Le (2019) propuseram uma arquitetura que equilibrasse a profundidade, largura e resolução da entrada para maximizar o desempenho com menor custo computacional. Para isso, utiliza o método de *compound scaling*. Também é utilizado o bloco *MBConv*, uma variante dos *Inverted Residual Blocks* da MobileNetV2. Essa arquitetura é chamada de EfficientNet-B0 e serve como base para outras variações (EfficientNet-B1 a B7), que escalam progressivamente os parâmetros com base nessa estrutura inicial. A variante EfficientNet-B2 oferece um bom *trade-off* entre eficiência e acurácia, com melhores resultados do que o EfficientNet-B0 e B1, enquanto mantém um custo computacional significativamente menor do que variantes mais profundas como EfficientNet-B3 a B7.

2.3.5 Base de imagens ImageNet

Todas as arquiteturas apresentadas nas seções anteriores são treinadas e avaliadas no conjunto de imagens chamado ImageNet. Essa é uma base de imagens apresentada por Deng et al. (2009), que contém mais de 14 milhões de imagens rotuladas, distribuídas em mais de 21.000 classes. É importante destacar que não existe nenhuma categoria para danos, de qualquer tipo, ou de aerogeradores e pás eólicas. Para que essas arquiteturas identifiquem outras categorias que não estão presentes no ImageNet, é necessário realizar o refinamento dos modelos através do processo de transferência de conhecimento.

2.4 Aprendizagem por transferência

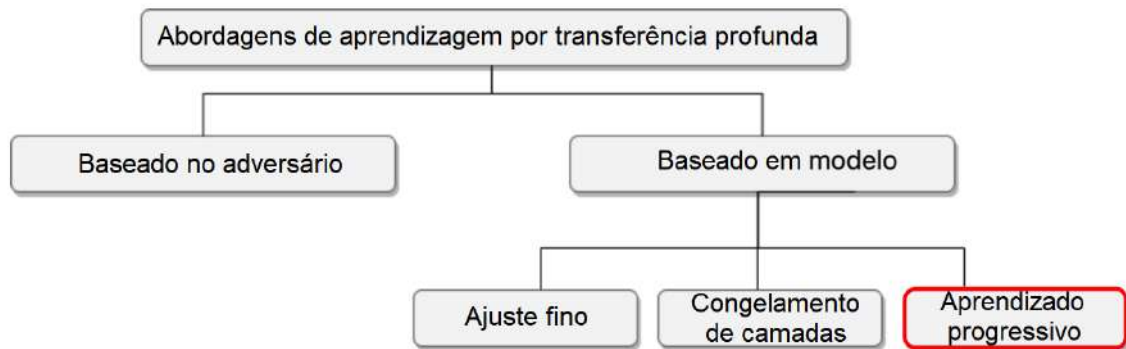
O treinamento de redes profundas pode demandar muitas horas, dias ou semanas, dependendo da arquitetura do modelo treinado e do volume de dados necessário para obter resultados satisfatórios. Para mitigar esse problema, uma técnica utilizada é a aprendizagem por transferência (*Deep Transfer Learning - DTL*). Segundo Iman, Arabnia e Rasheed (2023), essa técnica permite reduzir tanto o tempo, quanto o custo computacional ao aproveitar o conhecimento prévio de um modelo já treinado em um determinado domínio e refiná-lo com novos dados para uma nova tarefa.

No contexto da análise de pás eólicas, o DTL pode ser aplicado utilizando os modelos previamente treinados no conjunto de imagens ImageNet. Esses modelos já possuem aprendizado consolidado em estruturas básicas (baixo nível) e no reconhecimento de padrões mais complexos (alto nível). Assim, esse conhecimento prévio pode ser adaptado e aprimorado com imagens específicas de danos em pás eólicas, acelerando o processo de treinamento e melhorando a precisão dos resultados.

São apresentadas na Figura 13 as abordagens mais comuns para a aprendizagem por transferência em aprendizado profundo. Segundo Iman, Arabnia e Rasheed (2023), as abordagens mais eficientes e efetivas são baseadas em modelo (*Model-based*). Essa abordagem é detalhada nesta dissertação; para detalhes sobre a abordagem baseada em adversários (*Adversarial-based*), consultar o estudo de Iman, Arabnia e Rasheed (2023).

A abordagem baseada em modelo possui três sub-abordagens: ajuste fino das camadas de classificação (*Fine-tuning*), congelamento das camadas (*Frozen CNN layers*) e aprendizado progressivo (*Progressive learning*). Essas sub-abordagens podem ser utilizadas isoladamente ou em conjunto. O ajuste fino consiste em carregar os valores de pesos dos modelos já treinados e ajustar todas as camadas de extração de características e de classificação, realizando um novo treinamento com dados do novo domínio, assim, necessitando de menos tempo para o treinamento, uma vez que todo o conhecimento do modelo anterior é reaproveitado. No congelamento das camadas, alguma parte do conhecimento adquirido pelo modelo anterior não é ajustado em um novo treinamento.

Figura 13 – Abordagens mais comuns em DTL



Fonte: adaptado de Iman, Arabnia e Rasheed (2023).

Nessa abordagem, algumas camadas não têm os pesos atualizados, enquanto as outras camadas têm o ajuste dos pesos realizado. O aprendizado progressivo, destacado em vermelho na Figura 13, consiste no congelamento de camadas pré-treinadas e na adição, ou substituição, de camadas para aprender novas características dos dados, utilizando os pontos positivos das duas abordagens anteriores. Essa última abordagem tenta imitar as características do aprendizado humano que aprende novas habilidades com a ajuda de habilidades já adquiridas.

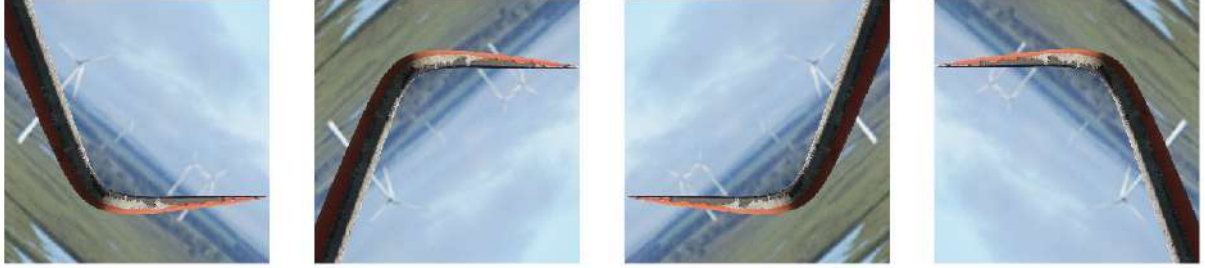
2.5 Aumento de dados

O treinamento de uma CNN para um determinado contexto necessita de um grande volume de dados, às vezes na casa de milhões de imagens. Obter essa grande quantidade de imagens é um grande desafio, sendo o aumento de dados ou em inglês *data augmentation* uma estratégia para lidar com esse problema, que consiste em gerar novas imagens com algumas pequenas alterações de modo que fiquem diferentes da imagem original. Isso traz benefícios como a maior diversidade das imagens, ampliação da quantidade de amostras e redução do *overfitting*. Possíveis problemas são: introdução de ruídos ou informações irrelevantes nas imagens, aumento do tempo de treinamento dos modelos e dificuldade para a definição das transformações e identificação dos parâmetros adequados.

De acordo com Xu et al. (2023), os algoritmos de aumento de dados da abordagem livre de modelo encontram-se entre as principais categorias existentes. Essa abordagem é aplicada a uma única ou múltiplas imagens, que consiste na aplicação de transformações na imagem, como geométricas (rotação, inversão, alteração de escala, etc.), cor (contraste, brilho, etc.), intensidade (ruídos, oclusões, etc.) ou até a mesclagem de várias imagens. Na Figura 14 é ilustrado um conjunto de imagens resultantes das transformações geométricas de inversão horizontal e vertical da abordagem livre de modelo. Na imagem original à esquerda, são aplicadas as transformações geométricas de inversão vertical, horizontal e

ambas ao mesmo tempo.

Figura 14 – Transformação geométrica de inversão horizontal e vertical



Fonte: Khalifa, Loey e Mirjalili (2022).

As transformações geométricas aplicadas no contexto de pás eólicas ajudam a simular as várias posições das pás no momento de aquisição das imagens, uma vez que o processo de rotação ocasionado pelos ventos pode deixá-las em posições verticais, horizontais ou inclinadas. As transformações de cor ajudam a simular os vários aspectos de luminosidade ocasionados pelo clima do dia ou horário de aquisição das imagens. Esses dois tipos de transformações permitem aplicar pequenas alterações nas imagens das pás eólicas sem, no entanto, gerar problemas de introdução de informações irrelevantes nas imagens. As equações dessas transformações são descritas abaixo:

1. **Inversão vertical**, Equação (2.1):

$$I_{\text{inv_vertical}}(i, j) = I(H - 1 - i, j) \quad (2.1)$$

em que:

- $I_{\text{inv_vertical}}(i, j)$ é o pixel na posição (i, j) da imagem invertida verticalmente;
- $I(H - 1 - i, j)$ é o pixel correspondente na imagem original;
- H é a altura total da imagem;
- $i \in [0, H - 1]$, $j \in [0, W - 1]$.

2. **Inversão horizontal**, Equação (2.2):

$$I_{\text{inv_horizontal}}(i, j) = I(i, W - 1 - j) \quad (2.2)$$

em que:

- $I_{\text{inv_horizontal}}(i, j)$ representa o pixel na posição (i, j) da imagem invertida horizontalmente;
- $I(i, W - 1 - j)$ é o pixel correspondente da imagem original;
- W é a largura total da imagem;
- $i \in [0, H - 1]$, $j \in [0, W - 1]$.

3. Ajuste de brilho, Equação (2.3):

$$I_{\text{brilho}}(i, j) = \min(255, \max(0, I(i, j) + \beta)) \quad (2.3)$$

em que:

- $I_{\text{brilho}}(i, j)$ é o valor do pixel após o ajuste de brilho;
- $I(i, j)$ é o valor original do pixel na posição (i, j) ;
- β é o valor de brilho adicionado (positivo para clarear, negativo para escurecer);
- $\min$ e $\max$ garantem que o valor resultante esteja no intervalo permitido de 0 a 255.

4. Ajuste de contraste, Equação (2.4):

$$I_{\text{contraste}}(i, j) = \min(255, \max(0, \alpha \cdot I(i, j))) \quad (2.4)$$

em que:

- $I_{\text{contraste}}(i, j)$ é o valor do pixel após o ajuste de contraste;
- $I(i, j)$ é o valor original do pixel na posição (i, j) ;
- α é o fator de contraste ($\alpha > 1$ aumenta o contraste, $0 < \alpha < 1$ reduz o contraste);
- $\min$ e $\max$ asseguram que o resultado esteja entre 0 e 255.

2.6 Métodos de balanceamento de classes em CNNs

A maioria dos problemas do mundo real possui mais dados de determinadas classes do que de outras, o que se denomina desbalanceamento de classes. No contexto de danos em pás eólicas também possui esse padrão: existem bem menos imagens com danos do que imagens consideradas normais. Desta forma, os modelos de aprendizagem de máquina são afetados por esses desbalanceamentos, pois, com poucas amostras de determinada classe, não é possível aprender todas as características necessárias para a identificação dessa classe. Além disso, amostras da classe dominante acabam tendo um peso maior durante o treinamento.

Uma forma de contornar esse problema é aplicando métodos de balanceamento de classes, que são divididos em dois grandes grupos: métodos em nível de dados e em nível de classificador. No primeiro grupo, o conjunto de dados é modificado para que as classes apresentem quantidades de amostras similares. Já o método em nível de classificador, não modifica a quantidade das amostras, porém os modelos são modificados para aprender a lidar com o desbalanceamento dos dados (Buda; Maki; Mazurowski, 2018).

Dentro dos métodos em nível de dados, o *oversampling* é um método que inclui novas amostras para as classes minoritárias para que estas passem a possuir uma quantidade de amostras próximas da classe majoritária. Na versão básica, as novas amostras podem ser aleatoriamente replicadas das amostras originais. Apenas a replicação simples das amostras pode levar o modelo ao *overfitting*. Para evitar isso, é possível aplicar as técnicas de aumento de dados para diferenciar as imagens ou utilizar técnicas mais sofisticadas como o SMOTE (Chawla et al., 2002) ou Redes Adversárias Generativas — *Generative Adversarial Network* - GANs (Goodfellow et al., 2014) que criam amostras sintéticas. Já o *undersampling* consiste em remover aleatoriamente amostras das classes majoritárias para igualar a quantidade de amostras da classe minoritária. Sua desvantagem é que parte dos dados que seriam usados para o treinamento são descartados.

Nos métodos em nível de classificador, o *thresholding* ajusta a probabilidade da classe prevista baseada na quantidade de amostras de cada classe, sendo uma previsão ponderada pelas classes. Em um exemplo hipotético, dado um problema de classificação binária com classes bastante desbalanceadas, é possível ter um caso em que a classe majoritária somente seria escolhida vencedora com uma probabilidade acima de 90%.

Já o método *cost sensitive learning* atua no momento de treinamento do modelo, ajustando os custos dos erros de predição de acordo com o desbalanceamento das classes. A título de exemplo, em um contexto de predição de casos de doenças graves, geralmente existem muitas amostras de pacientes normais e pouquíssimos casos de pacientes graves, o que torna o conjunto de dados altamente desbalanceado. A técnica de *cost sensitive learning* consiste em aumentar o custo do erro quando o modelo previr incorretamente as classes minoritárias, assim o otimizador do modelo vai, a cada iteração de treinamento, buscar errar menos a predição dessas classes.

Buda, Maki e Mazurowski (2018) avaliaram vários modelos aplicados a três conjuntos de dados de imagem (MNIST, CIFAR-10 e ImageNet) e concluíram que: i) classes desbalanceadas prejudicam o desempenho dos modelos; ii) o método de melhor desempenho, na maioria dos casos, para modelos CNNs multi-classe é o *oversampling*; e iii) para desbalanceamentos extremos ou muitas classes desbalanceadas, o melhor método é o *undersampling*.

2.7 Considerações finais

Neste Capítulo, foram apresentados os principais fundamentos teóricos que sustentam o desenvolvimento deste trabalho. Inicialmente, foram descritos os principais tipos de danos abordados nesta dissertação, enfatizando os aspectos estruturais e visuais relacionados às descargas atmosféricas, erosão, degradação da pintura e trincas. Em seguida, foram abordados os conceitos de visão computacional e redes neurais convolucionais, com ênfase

nos modelos utilizados ao longo dos experimentos: VGG-16, ResNet-50, MobileNetV2 e EfficientNet-B2, pré-treinados com o ImageNet.

Foram também discutidas técnicas complementares que se mostram essenciais para o aprimoramento do desempenho dos modelos, como o aumento de dados baseados na abordagem livre de modelo, o balanceamento de classes por *oversampling* aplicado a modelos CNNs multi-classe, e aprendizagem por transferência, com ênfase na abordagem de aprendizado progressivo.

Esses conceitos formam a base teórica para o método proposto no Capítulo 3, onde são detalhadas as etapas do processo de treinamento e validação do modelo, bem como os experimentos realizados para avaliar o impacto das técnicas exploradas neste Capítulo.

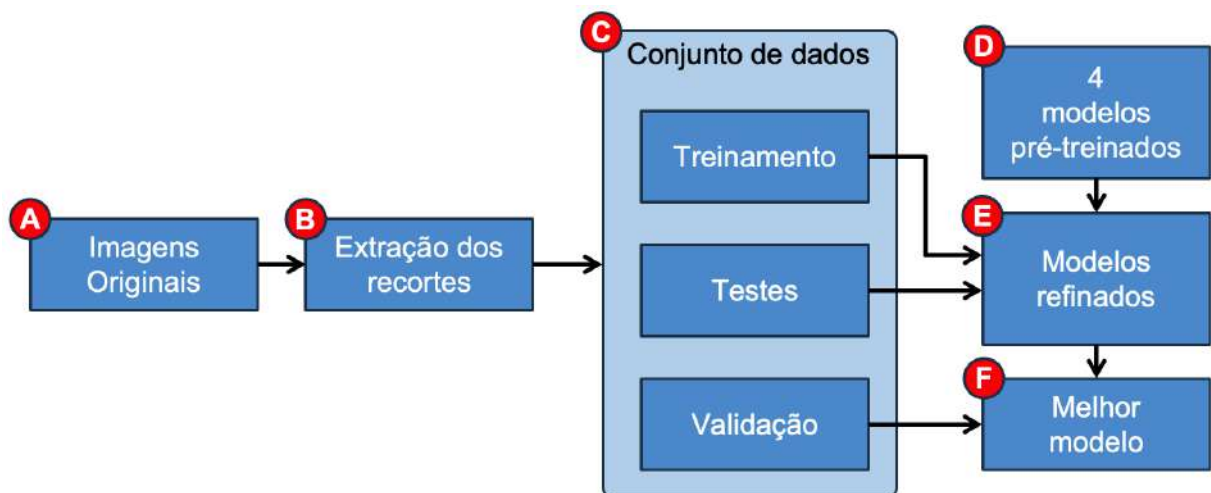
3 Método Proposto

Neste Capítulo é apresentado o método proposto para a identificação (classificação e localização) dos danos em pás eólicas. A metodologia proposta é subdividida em duas etapas: a etapa de treinamento dos modelos e a etapa de identificação dos danos. A primeira tem o objetivo de executar os passos necessários para a obtenção de um modelo treinado para a etapa subsequente. Nas seções seguintes são apresentados detalhes dessas etapas.

3.1 Etapa de treinamento

A arquitetura da etapa de treinamento é ilustrada na Figura 15. **Imagens originais** (A) indica as imagens disponibilizadas com os 4 possíveis tipos de danos. **Extração dos recortes** (B) é onde são extraídos os vários recortes menores da imagem original que, posteriormente, são separados nos **Conjuntos de dados** (C) utilizados para o refinamento e avaliação dos modelos. Já os **4 modelos pré-treinados** (D) são usados como base para fazer o refinamento nos **Modelos refinados** (E) com os conjuntos de dados de treinamento e testes. Por fim, o **Melhor modelo** (F) treinado é avaliado no conjunto de validação. E todos esses passos são detalhados nas subseções seguintes.

Figura 15 – Arquitetura da etapa de treinamento



Fonte: elaborada pelo autor.

3.1.1 Imagens originais

As imagens de pás eólicas com danos são disponibilizadas por uma filial de uma empresa de aerogeradores de grande porte, que possui uma parceria com o IFCE. Por

questões estratégicas, o nome da empresa é mantido em sigilo. Segundo a empresa parceira, as imagens são capturadas pelos técnicos da empresa durante campanhas de inspeção das pás dos aerogeradores, que ocorrem anualmente ou a cada dois anos, e são utilizadas câmeras profissionais de alta capacidade de *zoom*, com uso de tripé em solo e também com UAV voando. Estes aerogeradores estão localizados no Nordeste e Sul do Brasil, e as inspeções não ocorrem à noite e nem em dias de chuva, embora possam ocorrer em dias nublados. Na Tabela 2 é apresentado um resumo da quantidade, tamanho máximo e mínimo em *pixel* das imagens disponibilizadas por tipo de dano. São fornecidas 101 imagens com 4 tipos de danos e imagens sem danos não são disponibilizadas.

Tabela 2 – Quantitativos de imagens disponibilizadas por tipo de dano

Tipos de danos	Imagens		
	Quantidade	Tam. máximo (px ²)	Tam. mínimo (px ²)
Descargas atmosféricas	27	1632 × 1224	640 × 480
Erosão	27	4608 × 3456	
Danos na pintura	12	1024 × 768	
Trincas	35	4608 × 3456	
Total	101	-	-

Fonte: elaborada pelo autor.

Um trabalho prévio de localização do dano através de uma máscara foi realizado e validado por um especialista em danos de pás eólicas utilizando uma ferramenta de rotulação. Cada imagem possui apenas um único dano, tamanhos variados e áreas de imagem não relacionadas aos danos, como exemplo: o céu, praias, árvores e outras áreas da pá eólica sem danos; algumas possuem indicações textuais da data e hora da obtenção do registro fotográfico. Na Figura 16 são ilustrados exemplos de imagens com danos e suas máscaras, assim como áreas nas imagens que não possuem danos.

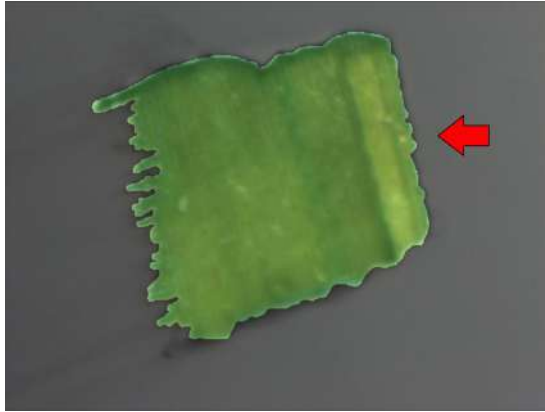
Para auxiliar todo o processo de rotulação das imagens originais, extração e classificação dos recortes, uma ferramenta auxiliar para esse trabalho foi desenvolvida e usada juntamente com o especialista. Detalhes dessa ferramenta encontram-se no Apêndice A.

3.1.2 Extração dos recortes

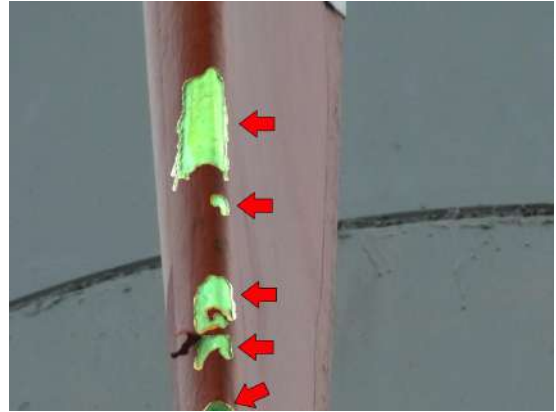
Os modelos pré-treinados necessitam de imagens com tamanhos menores do que o tamanho das imagens originais disponibilizadas. Porém, o redimensionamento das imagens torna imperceptíveis alguns dos danos menores. Assim, optou-se por extrair recortes menores das imagens originais e realizar a tarefa de classificação de danos em cada um dos recortes.

Nesse passo, as imagens originais juntamente com as máscaras dos rótulos são utilizadas como entrada. Em seguida, recortes são delimitados ao longo da imagem,

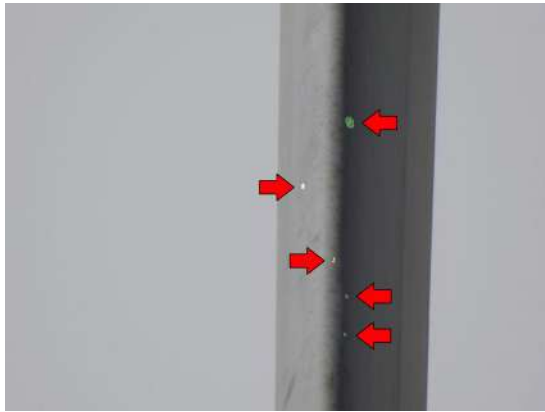
Figura 16 – Exemplos das imagens originais exibindo as máscaras, com setas vermelhas indicando a localização dos danos



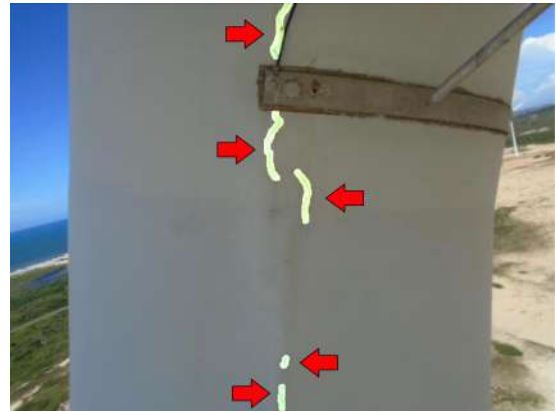
(a) Dano por descarga



(b) Erosão



(c) Dano na pintura



(d) Trinca

Fonte: imagens disponibilizadas por empresa de aerogeradores parceira do trabalho.

extraídos e rotulados nas classes Normal, Descarga, Erosão, Pintura e Trinca. O rótulo Normal é atribuído quando um recorte não possui nenhum dano indicado pela máscara. Para a delimitação e extração dos recortes, duas abordagens podem ser utilizadas: extração baseada em janela deslizante e baseada em posição aleatória.

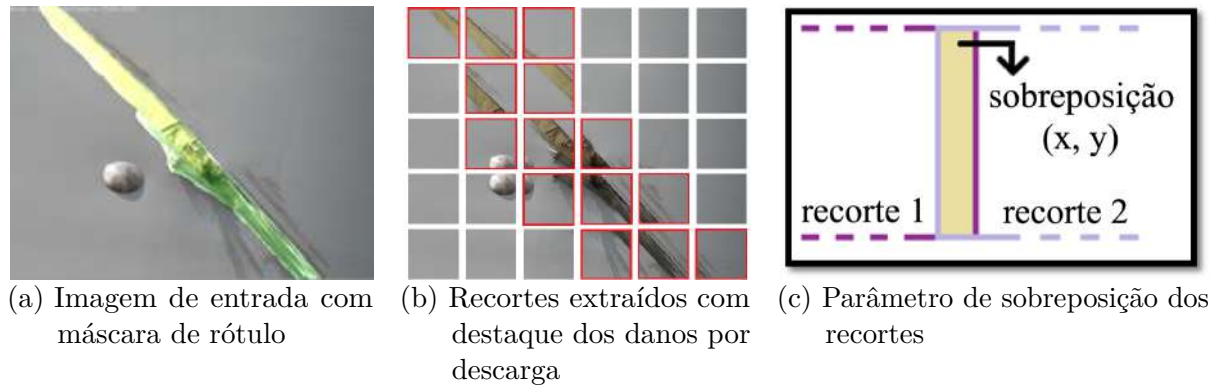
3.1.2.1 Extração baseada em janela deslizante

Essa abordagem utiliza o auxílio da técnica da janela deslizante para fazer a extração dos recortes de tamanhos fixos pré-definidos. Com o intuito de evitar que os danos sejam prejudicados pela divisão da imagem original nos recortes, é possível definir uma sobreposição com um valor mínimo de *pixels* a ser considerado. No caso do tamanho da imagem original não ser divisível pelo tamanho dos recortes com a sobreposição, o tamanho da sobreposição é aumentado até conseguir um ajuste perfeito.

Algumas etapas do processo de extração dos recortes baseado em janela deslizante

são ilustradas na Figura 17. Mais precisamente, na Figura 17 (a) é exibida a imagem original juntamente com a máscara de rótulo que serve como entrada para a etapa de extração. Já na Figura 17 (b), vários recortes são extraídos a partir da aplicação de uma janela deslizante com sobreposição, com destaque em vermelho para os recortes com rótulos “Descarga” e sem destaque os recortes com rótulos “Normal”. Finalmente, na Figura 17 (c) é ilustrada a sobreposição dos recortes.

Figura 17 – Etapas de extração dos recortes baseado em janela deslizante



Fonte: elaborado pelo autor.

A abordagem baseada em janela deslizante tem a vantagem de extrair recortes com a cobertura de toda a área da imagem original, não ficando nenhum *pixel* de fora dos recortes, porém a desvantagem é a limitação da quantidade de recortes extraídos para um recorte de tamanho fixo. Como exemplo disto, na Figura 17 (b) foi utilizado um recorte de tamanho $320 \times 320 \text{ px}^2$, o que extraiu recortes similares a uma grade 5×6 , fixando em 30 recortes extraídos. Só é possível extrair outros recortes da imagem, caso seja alterado o tamanho do recorte, mas ainda assim esse novo tamanho deve ter um limite da quantidade extraída. Outro aspecto importante é que é possível aplicar rotações na imagem original, mas todos os recortes são afetados da mesma maneira.

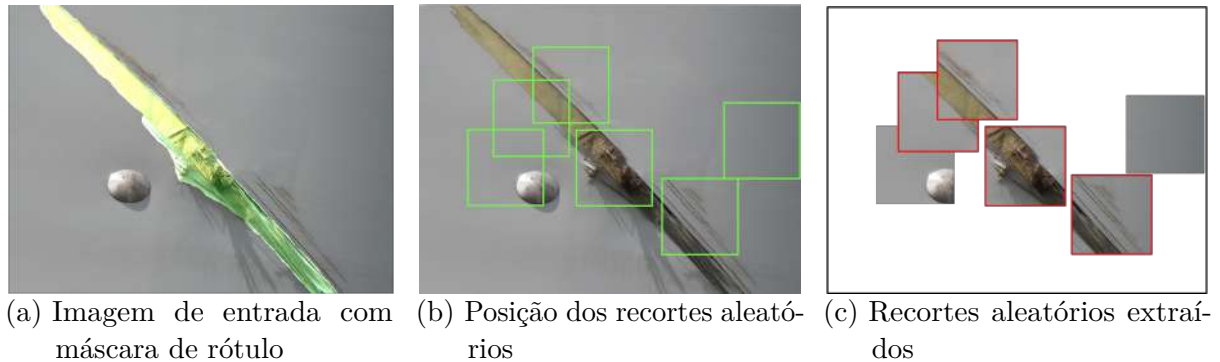
3.1.2.2 Extração baseada em posição aleatória

Na abordagem baseada em posição aleatória, os recortes ficam posicionados de forma livre e aleatória na área da imagem original. O ponto central do recorte (x, y) é escolhido de forma aleatória, dentro das dimensões da imagem original. Também é possível aplicar, para cada posição, transformações de redimensionamento e rotação, extraíndo assim recortes de tamanhos variados e angulações diferentes.

Algumas etapas da extração dos recortes baseados em posição aleatória são ilustradas na Figura 18, sem a aplicação do redimensionamento e rotação. Na Figura 18 (a) é exibida a imagem original juntamente com a máscara de rótulo que serve como entrada para a extração dos recortes. Já na Figura 18 (b) são apresentados os recortes aleatórios

posicionados na imagem original. E, por fim, na Figura 18 (c), são ilustrados os recortes aleatórios extraídos, com destaque em vermelho para os recortes com rótulos “Descarga” e, sem destaque, os recortes com rótulos “Normal”.

Figura 18 – Etapas da extração dos recortes baseado em posição aleatória



Fonte: elaborado pelo autor.

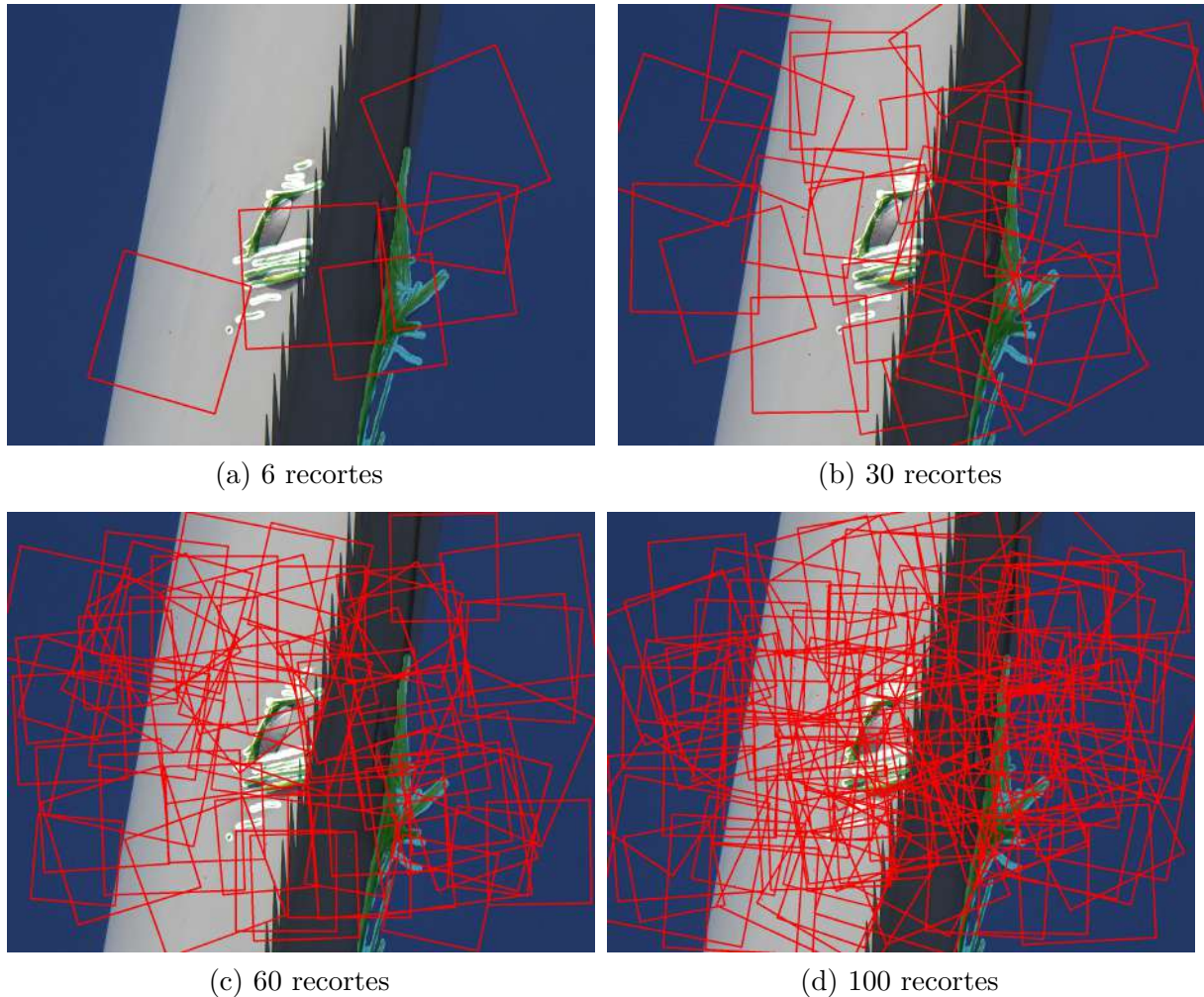
A abordagem baseada em posição aleatória tem a desvantagem de não cobrir toda a área da imagem original, tendo a possibilidade de não extrair um dano encontrado em uma área não coberta. Porém, essa abordagem possibilita a utilização do redimensionamento e rotação para cada recorte. Também é possível a aplicação de métodos de distribuição dos pontos, por exemplo: uniforme ou gaussiana (que permite um viés de centro, diminuindo a influência do fundo). Como também, o aumento da quantidade de recortes, de forma fixa ou dinâmica (baseado no tamanho da imagem de entrada). Na Figura 19 é ilustrada a utilização de 6, 30, 60 e 100 recortes, com a aplicação do redimensionamento e rotação.

3.1.3 Conjuntos de dados

Os recortes extraídos são separados nos conjuntos de dados utilizados no projeto dos modelos de classificação. Quando utilizados dois conjuntos de dados, a proporção mais comum utilizada é de 2/3 para treino e 1/3 para teste, com outras variações comuns de treino/teste em 60/40, 70/30 ou 80/20 (Raschka, 2018). Para a separação entre os conjuntos de treinamento, teste e validação, as proporções de 60%, 15% e 25% são adotadas, respectivamente, de forma a ter um bom equilíbrio entre os conjuntos para evitar viés e garantir uma estimativa confiável do desempenho do modelo (Xu; Goodacre, 2018). Os recortes são separados de forma estratificada pelas classes disponíveis e, nesse passo, também existem duas possíveis abordagens a serem adotadas: separação dos recortes pela imagem original ou pelos recortes.

A **abordagem de separação dos conjuntos de dados pela imagem original** consiste em manter todos os recortes de uma mesma imagem no mesmo conjunto de dados. Ou seja, primeiramente são definidas quais imagens originais compõem cada conjunto

Figura 19 – Aumento da quantidade de recortes aleatórios



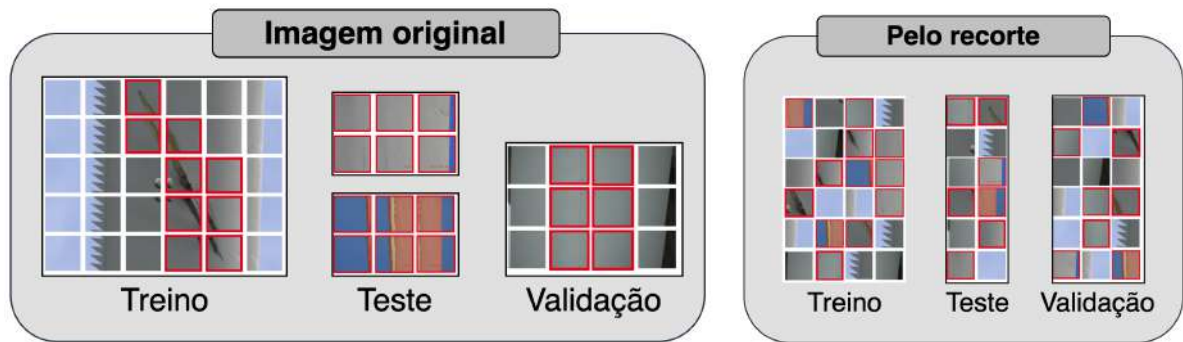
Fonte: elaborado pelo autor.

de dados, respeitando a proporção e a estratificação, para, em seguida, separar todos os recortes da imagem no conjunto de dados correspondente. Já na **abordagem de separação dos conjuntos de dados pelos recortes**, independentemente da imagem original, os recortes são separados aleatoriamente nos conjuntos disponíveis. Na Figura 20 é ilustrado um exemplo da separação dos recortes de 4 imagens originais utilizando as duas abordagens.

3.1.3.1 Aumento de dados e balanceamento

Os recortes extraídos possuem um grande desbalanceamento das classes, com imagens da classe normal dominando o quantitativo de amostras. Optou-se por fazer o treinamento dos modelos com outras duas abordagens: **sem** aumento de dados ou balanceamento e **com** o aumento de dados e balanceamento, somente aplicados ao conjunto de treinamento. Neste caso, o método de balanceamento utilizado é o *oversampling* com

Figura 20 – Abordagens para separação dos recortes



(a) separação pela imagem original

(b) separação pelos recortes

Fonte: elaborado pelo autor.

métodos a nível de dados.

Já o processo de aumento dos dados aplica métodos da abordagem livre de modelo em única imagem, empregando as transformações: inversão vertical, inversão horizontal, ajuste de brilho e ajuste de contraste. Conforme detalhado na Seção 2.5, para o ajuste de brilho, Equação (2.3), é utilizado um β aleatório entre $[-0,10, +0,10]$ e para o ajuste de contraste, Equação (2.4), um α aleatório entre $[+0,7, +1,3]$. Cada uma dessas transformações tem uma probabilidade de 30% de ser aplicada, sendo que, pelo menos, uma função de transformação é sempre aplicada à imagem.

3.1.4 Refinamento dos modelos pré-treinados

O processo de aprendizagem por transferência baseado no aprendizado progressivo é utilizado para otimizar o tempo de treinamento dos modelos. Na Tabela 3 são detalhados os 4 modelos pré-treinados com o ImageNet.

Tabela 3 – Modelos pré-treinados com o ImageNet

Modelos	Referência	Tam. (MB)	Camadas	Total Parâmetros (congelados)	Parâmetros treináveis (descongelados)
EfficientNetB2	(Tan; Le, 2019)	35	227	9,2M	7.045
MobileNetV2	(Sandler et al., 2018)	14	53	3,4M	6.405
ResNet50	(He et al., 2016)	98	50	25,6M	10.245
VGG16	(Simonyan; Zisserman, 2014)	528	16	138,4M	2.565

Fonte: elaborada pelo autor.

As camadas pré-treinadas desses modelos são congeladas e a última camada de classificação é substituída por uma nova camada de saída para classificação das cinco classes definidas, com o uso da função *Softmax* (Goodfellow; Bengio; Courville, 2016) como ativação. Essa substituição se deve ao fato de que a camada original foi treinada para classificar as 21.000 classes do ImageNet. Esta função *Softmax*, que é uma generalização da função sigmoide para múltiplas classes, recebe um vetor de valores reais (*logits*) e os transforma em uma distribuição de probabilidade, em que a soma dos valores de saída é sempre 1. Sua equação matemática é

$$\text{softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^n e^{z_j}}, \quad (3.1)$$

em que z_i é o *logit* de entrada para a classe i ; e^{z_i} é o exponencial do *logit*; e $\sum_{j=1}^n e^{z_j}$ é a soma das exponenciações de todos os *logits*. Os modelos são otimizados pelo algoritmo Adam (Diederik, 2014) para minimizar a perda da entropia cruzada.

Esses modelos pré-treinados e com a última camada modificada são os modelos-base utilizados no processo de refinamento. Cada modelo-base passa por cinco repetições de refinamento para avaliar sua robustez à inicialização dos pesos da última camada, com base na média e no desvio padrão das métricas adotadas. Assim, a cada repetição do refinamento, o modelo-base é reutilizado, e os conjuntos de treinamento e teste são novamente divididos, garantindo que cada repetição utilize conjuntos de dados de treinamento diferentes. Já a avaliação é realizada por métricas que são calculadas a partir do conjunto de teste, considerando a classificação das imagens entre Normal ou com Dano (Descarga, Erosão, Pintura ou Trinca).

3.1.5 Seleção do melhor modelo

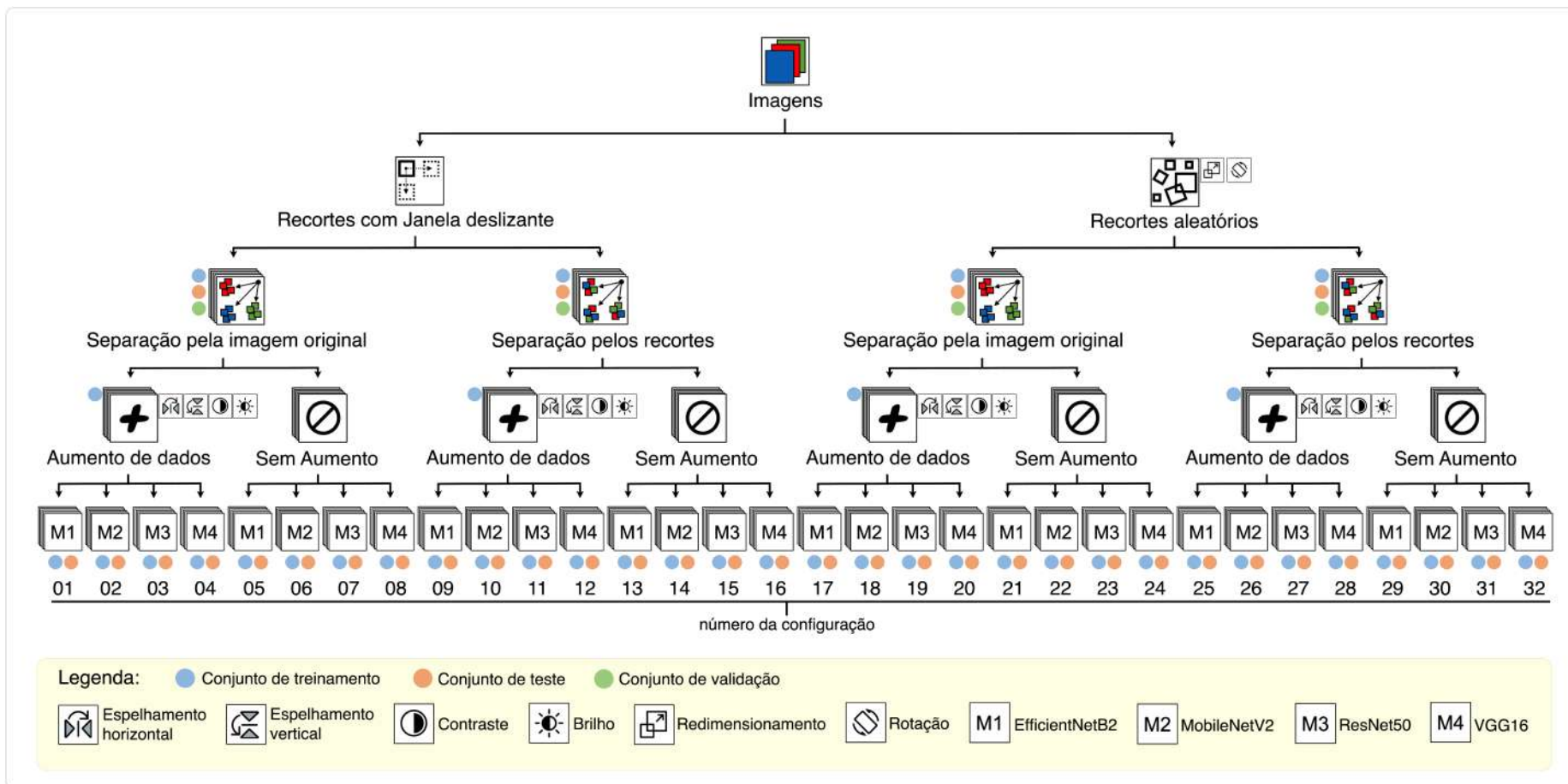
No processo de construção do melhor modelo, as principais métricas analisadas são precisão, acurácia e sensibilidade. Neste processo, a escolha do melhor modelo é baseada na média da precisão obtida ao longo de cinco repetições de treinamento. Isto porque a precisão é adotada como métrica principal devido à sua relevância na seleção de imagens com danos, pois uma alta precisão ajuda a minimizar a ocorrência de falsos negativos, que correspondem a casos em que um dano é incorretamente classificado como uma imagem normal.

Para desempate entre modelos, considera-se inicialmente a acurácia e, caso necessário, a sensibilidade. Após a seleção, o melhor modelo é reavaliado utilizando o conjunto de validação. Além disso, as métricas são analisadas de forma detalhada para cada tipo de dano, abrangendo as classes Normal, Descarga, Erosão, Dano de Pintura e Trinca.

3.1.6 Configurações de treinamento

A partir das abordagens apresentadas, são definidas 32 configurações de treinamento diferentes, um exemplo é: modelo ResNet50, com extração de recortes baseados em janela deslizante, utilizando separação dos conjuntos pelos recortes e sem a aplicação de aumento de dados e balanceamento. Todas as configurações utilizadas estão ilustradas na Figura 21. Como cada modelo é treinado em 5 conjuntos de dados diferentes, ao final do processo, são obtidos um total de 160 modelos (32 configurações x 5 repetições).

Figura 21 – Configurações de treinamento utilizadas



Fonte: elaborada pelo autor.

3.2 Experimentos definidos

Para determinar a melhor configuração de treinamento, são definidos quatro experimentos. Na Tabela 4 é apresentada a descrição de cada experimento, bem como as questões que são respondidas ao seu término. Em todos os casos, os resultados são expressos pelos valores de precisão média, calculados a partir de cinco repetições para cada configuração testada. O modelo treinado nessa configuração é considerado o melhor modelo encontrado.

Tabela 4 – Experimentos definidos

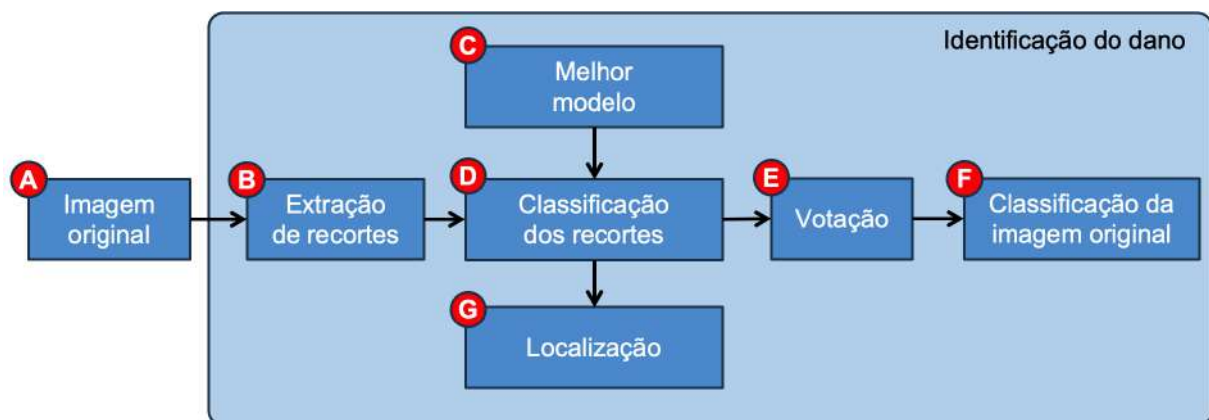
Experimento	Pergunta a ser respondida
I. Refinamento dos modelos	Qual modelo refinado é o melhor para classificação de danos em pás eólicas? EfficientNetB2, MobileNetV2, ResNet50 ou VGG16?
II. Extração dos recortes	Qual o melhor método de extração dos recortes? Janela deslizante ou posição aleatória?
III. Separação dos conjuntos	Qual o melhor método de separação dos conjuntos de dados? Pela imagem original ou recortes?
IV. Aumento de dados e balanceamento	É melhor treinar sem aumento de dados e balanceamento ou aplicar o aumento de dados e balanceamento nos recortes?

Fonte: elaborada pelo autor.

3.3 Etapa de identificação de danos

Uma vez que o melhor modelo é definido, é possível utilizá-lo na etapa de identificação de danos. Essa etapa é ilustrada na Figura 22.

Figura 22 – Arquitetura da etapa de identificação do dano



Fonte: elaborada pelo autor.

A **Imagem original** (A) é usada como entrada para a etapa de **Identificação do dano**. A partir da imagem original, é realizada a **Extração de recortes** (B) e

com o **Melhor modelo** (C) é realizada a **Classificação dos recortes** (D), definindo uma única classe para cada recorte extraído no passo anterior. A seguir, é feita uma **Votação** (E) das classificações dos recortes para decidir a classe final da imagem e, com isso, ter a **Classificação da imagem original** (F). Também existe a opção de realizar a **Localização** (G) do dano, utilizando a posição dos recortes classificados. Essa localização é exibida projetando um mapa de calor na imagem original. Nessa etapa, é utilizada a extração de recortes baseada na abordagem de janela deslizante.

3.3.1 Votação e classificação das imagens originais

Os modelos de classificação são treinados e avaliados utilizando os diversos recortes extraídos das imagens originais. Neste passo, é realizado o processo inverso: o mapeamento dos recortes de volta à imagem original, cujo melhor modelo é, então, aplicado para classificar todos os recortes extraídos. Em seguida, utiliza-se um método de votação por maioria para determinar a classe final da imagem original com base nos resultados dos recortes. O dano mais votado define o rótulo atribuído à imagem. Caso nenhum recorte seja classificado como dano, a imagem é considerada normal.

3.3.2 Localização dos danos nas imagens originais

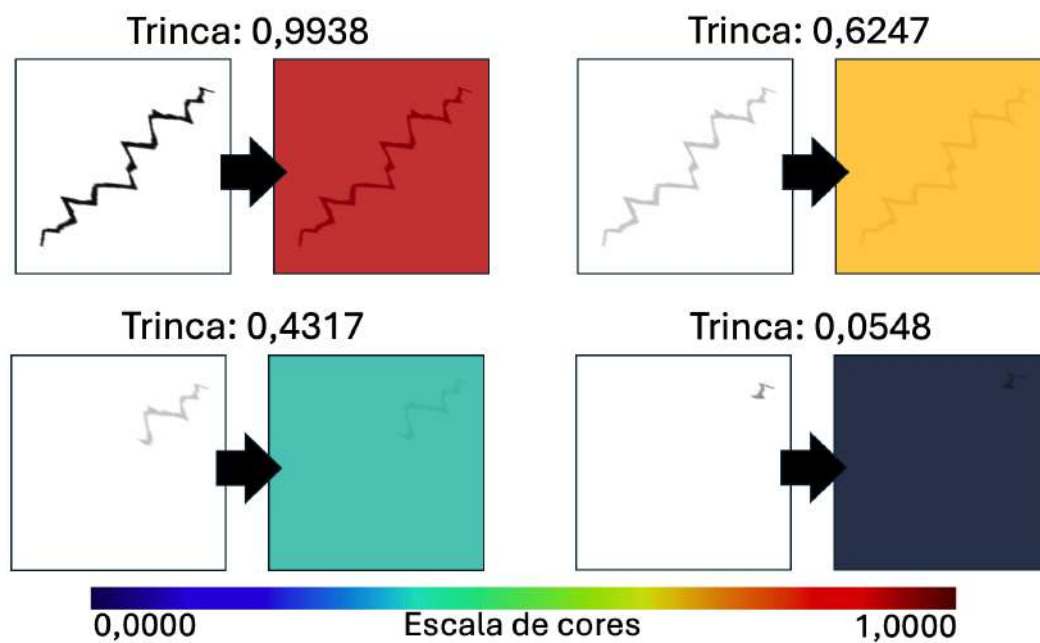
A classificação de todos os recortes permite mapear as posições dos danos na imagem original. No processo de classificação, a camada de saída do melhor modelo gera valores normalizados por meio da função *Softmax*, garantindo que a soma das probabilidades seja sempre igual a 1. Com base na classificação dos recortes, em suas posições e nos valores normalizados das classes, é possível localizar os danos na imagem original por meio de um mapa de calor. Quanto mais próximo de 1 for o valor predito, mais vermelho é o recorte, indicando maior probabilidade de dano. Por outro lado, valores próximos de 0 resultam em recortes exibidos em azul, representando menor probabilidade de dano. Na Figura 23 é ilustrada a aplicação da escala de cores em ilustrações de recortes de trinca, com o número representando o valor predito da classificação para a classe trinca.

3.4 Considerações finais

Este Capítulo apresentou o método proposto para a identificação e localização de danos em pás de aerogeradores por meio do refinamento de modelos de redes neurais convolucionais. Foram detalhadas as etapas de pré-processamento, extração e separação dos recortes, estratégias de aumento de dados e balanceamento, além da configuração dos experimentos e do ajuste fino dos modelos pré-treinados. As abordagens adotadas visam aprimorar a capacidade de generalização e a acurácia dos modelos. No capítulo seguinte,

são descritos os resultados obtidos e a análise do desempenho das diferentes configurações experimentais avaliadas.

Figura 23 – Escala de cores do mapa de calor em danos simulados



Fonte: elaborada pelo autor.

4 Resultados e Discussões

Neste Capítulo são apresentados e analisados os resultados dos experimentos definidos, com o objetivo de identificar a melhor configuração de treinamento. O modelo treinado nesta configuração é avaliado no conjunto de dados de validação e, em seguida, são avaliadas a classificação das imagens originais e a localização dos danos através da utilização de um mapa de calor.

4.1 Especificações do ambiente experimental

A configuração da máquina utilizada para o refinamento dos modelos é apresentada na Tabela 5.

Tabela 5 – Especificações do ambiente experimental

Configuração	Especificação
CPU	AMD Athlon 3000G with <i>Radeon Vega Graphics</i> 3.50GHz
Mémoria RAM	16 GB
Disco rígido	SSD Haste, 960gb, 2,5", 550mb/s <i>read</i> , 420mb/s <i>write</i>
GPU	<i>NVidia GeForce GTX 1660 (2x) Super</i>
Sistema operacional	Linux Ubuntu 22.04.2 LTS
<i>Framework</i>	<i>Tensorflow 2.12.1</i> configurado para GPUs em paralelo
Outros <i>softwares</i>	Cuda <i>Toolkit</i> 11.8.0; e cuDNN 8.6.0.163

Fonte: elaborada pelo autor.

Os modelos são construídos utilizando o *framework TensorFlow 2.12.1* que é uma ferramenta de *software* livre e de código aberto para aprendizagem de máquina e inteligência artificial focada principalmente em aprendizagem profunda. Todos os modelos são refinados utilizando os seguintes parâmetros: 0,001 de taxa de aprendizagem, 200 épocas, otimizador Adam, processamento em GPU, 5 repetições de treinamento, tamanho de *batch* de 20 imagens e funções de eventos (*callbacks*). Os *callbacks* utilizados são os seguintes: redução de 20% da taxa de aprendizagem ao atingir um platô em 5 épocas na métrica de perda da entropia cruzada do conjunto de testes, com limite mínimo da taxa de aprendizagem em 0,0001; e parada antecipada do treinamento quando a métrica de perda do conjunto de validação não reduz após 20 épocas, sendo restaurados os melhores pesos encontrados.

4.2 Experimento I: análise do método de refinamento dos modelo

O objetivo desse experimento é entender qual o modelo pré-treinado obtém os melhores resultados após o refinamento com a utilização das seguintes variações de configurações: abordagem de extração dos recortes, tipo de separação e uso ou não do aumento de dados e balanceamento. Na Tabela 6 são apresentados os resultados de cada configuração para a métrica de precisão, considerando a média das cinco repetições realizadas.

Tabela 6 – Resultados do experimento para definição do melhor modelo refinado

Métrica: Precisão (média)			Modelos				
Extração dos recortes	Tipo de separação	Aumento e Balanc.	M1	M2	M3	M4	Melhor
Janela deslizante	Imagem original	Com	0,927 ±0,055	0,905 ±0,084	0,924 ±0,063	0,898 ±0,078	M1
		Sem	0,859 ±0,098	0,836 ±0,756	0,893 ±0,072	0,824 ±0,150	M3
	Recorte	Com	0,970 ±0,006	0,953 ±0,015	0,967 ±0,015	0,966 ±0,017	M1
		Sem	0,951 ±0,013	0,945 ±0,019	0,948 ±0,005	0,938 ±0,013	M1
Aleatório	Imagem original	Com	0,889 ±0,106	0,939 ±0,030	0,918 ±0,062	0,924 ±0,058	M2
		Sem	0,887 ±0,080	0,891 ±0,126	0,877 ±0,064	0,823 ±0,143	M2
	Recorte	Com	0,986 ±0,011	0,967 ±0,015	0,972 ±0,024	0,959 ±0,014	M1
		Sem	0,952 ±0,014	0,943 ±0,015	0,964 ±0,011	0,941 ±0,018	M3
Melhor geral:							M1

Legenda: M1: EfficientNetB2, M2: MobileNetV2, M3: ResNet50 e M4: VGG16.

Fonte: elaborada pelo autor.

Para cada configuração, é analisado qual modelo obtém o melhor resultado. Por fim, uma votação simples dos melhores modelos é realizada. Com isso, tem-se que o modelo EfficientNetB2 (M1) obtém o melhor resultado em 4 configurações. Já os modelos MobileNetV2 (M2) e ResNet50 (M3) obtêm, cada um, 2 melhores resultados. E o modelo VGG16 (M4) não obtém o melhor resultado em nenhuma das configurações. Os melhores resultados apresentados pelo modelo M1 podem ser justificados por esse modelo possuir uma arquitetura mais recente (EfficientNetB2) em relação aos outros modelos, incorporando avanços tecnológicos que tornam sua estrutura mais eficiente e adaptável a diferentes cenários de visão computacional. Além disso, possui a rede mais profunda, o que possibilita obter características mais detalhadas das imagens.

De forma a reduzir a apresentação e análise dos resultados, os próximos experimentos apresentam apenas os valores para o modelo EfficientNetB2.

4.3 Experimento II: análise do método de extração dos recortes

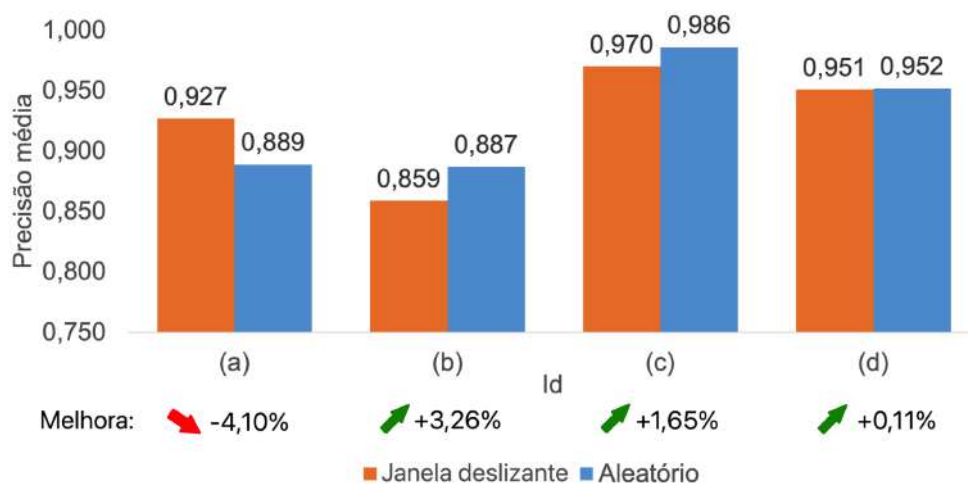
O objetivo desse experimento é entender a influência das diferentes abordagens da extração dos recortes, baseada em janela deslizante ou posição aleatória, a partir da imagem original. Para isso, as duas abordagens extraem, cada uma, os mesmos números de recortes: 1.828, com um tamanho de recorte padrão de 320×320 px². E no caso da abordagem baseada em janela deslizante, uma sobreposição de 32 px. A abordagem de posição aleatória permite a aplicação de transformações de redimensionamento do tamanho e rotação individualmente para cada recorte. Sobre o tamanho padrão dos recortes, é aplicado aleatoriamente um redimensionamento com um fator de escala entre a faixa de $[+0,7, +1,3]$ e para a rotação entre $[-45, +45]$ graus. Na Tabela 7 são apresentadas as configurações utilizadas para o experimento de geração de recortes. Um identificador (Id) é atribuído a cada configuração e este é utilizado na apresentação dos resultados contidos na Figura 24.

Tabela 7 – Configurações do experimento para a extração dos recortes

Id	Tipo de separação	Aumento de dados	Modelo
(a)	Imagem original	Com aumento e balanceamento	EfficientNetB2
(b)	Imagem original	Sem aumento e balanceamento	
(c)	Recorte	Com aumento e balanceamento	
(d)	Recorte	Sem aumento e balanceamento	

Fonte: elaborada pelo autor.

Figura 24 – Resultados do experimento II (método de extração dos recortes)

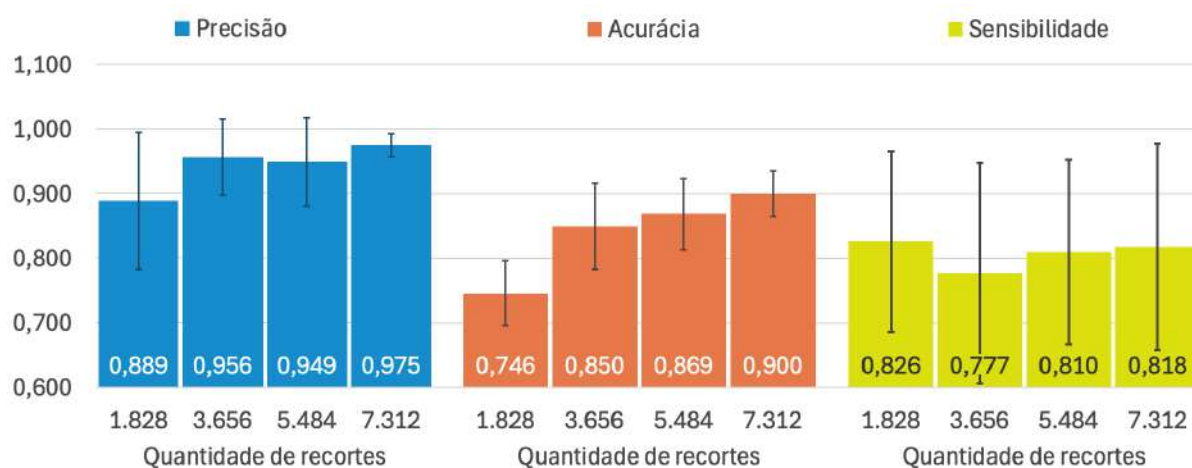


Fonte: elaborada pelo autor.

Analisando os resultados das 4 configurações, observa-se que a melhor abordagem para extração dos recortes é obtida pela abordagem de **recorte aleatório**. Embora, na configuração (a) o resultado seja menor em relação à abordagem de janela deslizante, nas configurações (b) e (c) os resultados são melhores e praticamente os mesmos na configuração (d). Quando analisados apenas os resultados para o modelo EfficientNetB2, a melhora média do recorte aleatório sobre o recorte baseado em janela deslizante é de **0,23%**. E quando se considera todos os modelos, a melhora média é de **1,06%**.

A razão pela qual as duas abordagens apresentam resultados próximos se deve ao fato de ambas extraírem a mesma quantidade de recortes. Entretanto, a abordagem de recortes aleatórios tem como vantagem a possibilidade do aumento da quantidade de recortes extraídos. A diferença dos resultados é acentuada quando ocorre o aumento da quantidade de recortes. Na Figura 25 são ilustrados os resultados da variação da quantidade de recortes extraídos, aplicada à configuração (a), que apresenta resultados piores comparados à abordagem baseada em grade.

Figura 25 – Resultados da variação da quantidade de recortes aleatórios gerados



Fonte: elaborado pelo autor.

O eixo vertical apresenta os resultados obtidos pelas métricas precisão, acurácia e sensibilidade. O eixo horizontal exibe a quantidade de recortes extraídos, com os valores de 1.828, 3.656, 5.484 e 7.312 recortes, sendo respectivamente 1x, 2x, 3x e 4x a quantidade de recortes extraídos no experimento original. Nota-se que os resultados melhoram para precisão (+9,67%) e acurácia (+20,64%) com o aumento da quantidade de recortes (de 1.828 para 7.312 recortes) e exibem resultados mais consistentes com a diminuição do desvio padrão, já ultrapassando o resultado baseado em janela deslizante quando utilizados 3.656 recortes, enquanto a sensibilidade apresenta resultados estáveis. Essa melhora justifica-se pela maior diversidade das imagens extraídas e maior cobertura da imagem original.

4.4 Experimento III: análise do método de separação dos conjuntos de dados

O objetivo desse experimento é avaliar o impacto do tipo de separação dos conjuntos de dados nos resultados avaliados. Na abordagem baseada na imagem original, a separação dos recortes é realizada com base nos recortes das imagens originais, enquanto que na abordagem pelos recortes, a separação é realizada levando-se em consideração apenas os recortes. Na prática, é como se na primeira abordagem fosse realizada a separação das imagens e depois fosse feita a extração dos recortes, enquanto que na segunda abordagem, ocorresse a extração dos recortes e depois a separação.

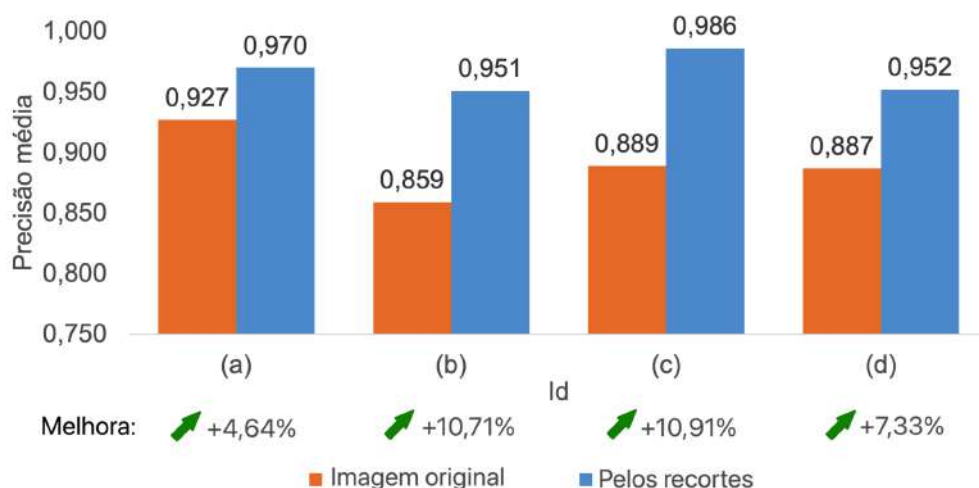
As configurações utilizadas para o experimento de separação dos conjuntos são apresentadas na Tabela 8. Um identificador (Id) é atribuído a cada configuração e este é utilizado na apresentação dos resultados contidos na Figura 26.

Tabela 8 – Configurações do experimento para a separação dos conjuntos

Id	Extração dos recortes	Aumento de dados	Modelo
(a)	Janela deslizante	Com aumento e balanceamento	EfficientNetB2
(b)	Janela deslizante	Sem aumento e balanceamento	
(c)	Aleatório	Com aumento e balanceamento	
(d)	Aleatório	Sem aumento e balanceamento	

Fonte: elaborada pelo autor.

Figura 26 – Resultados do experimento III (separação dos conjuntos)



Fonte: elaborada pelo autor.

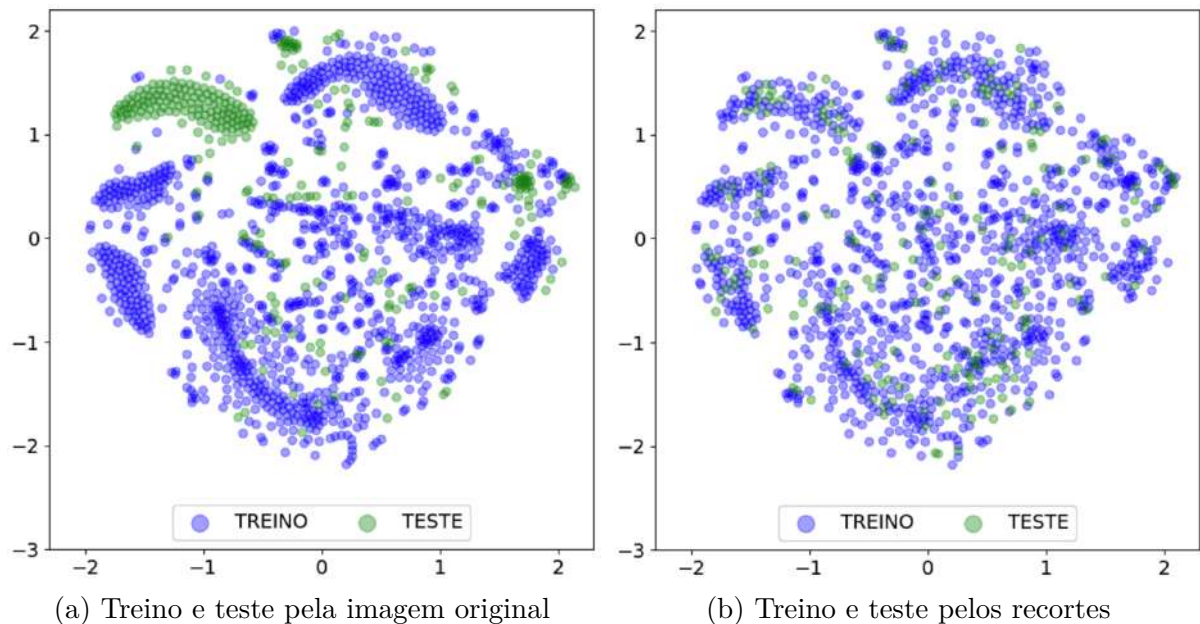
Analisando os resultados das 4 configurações, o melhor resultado é obtido pela separação dos conjuntos com abordagem baseada **pelos recortes**, com uma média de melhora sobre a abordagem pela imagem original de **8,40%** para o modelo EfficientNetB2.

Essa técnica é a melhor em todas as configurações e todos os modelos, apresentando uma melhora média de **7,46%**.

Um aspecto importante a ressaltar é que na abordagem pela imagem original, a separação nos conjuntos de treinamento, testes e validação respeitam as proporções estabelecidas na Seção 3.1.3 de 60%, 15% e 25%, quando é verificada no nível das imagens originais. Porém, quando é verificada a separação a nível dos recortes, esses valores podem mudar. Uma vez que o tamanho das imagens originais pode variar e com a separação sendo feita de forma aleatória e estratificada, pode acontecer, por exemplo, de imagens com tamanhos grandes ficarem no conjunto de testes ou validação e tamanhos pequenos no conjunto de treinamento, o que leva a ter a quantidade de recortes em proporções diferentes. Isso não ocorre na abordagem pelos recortes.

Uma análise das distribuições das imagens nos conjuntos de dados de treino e teste utilizando o T-SNE é ilustrada na Figura 27 para a configuração (b), conforme descrito na Tabela 8. Na Figura 27 (a) é exibida a distribuição **pela imagem original** e a Figura 27 (b) a distribuição **pelos recortes**. Observa-se que os conjuntos separados pela imagem original possuem uma densidade maior de recortes, enquanto que os conjuntos separados pelos recortes possuem amostras mais distribuídas.

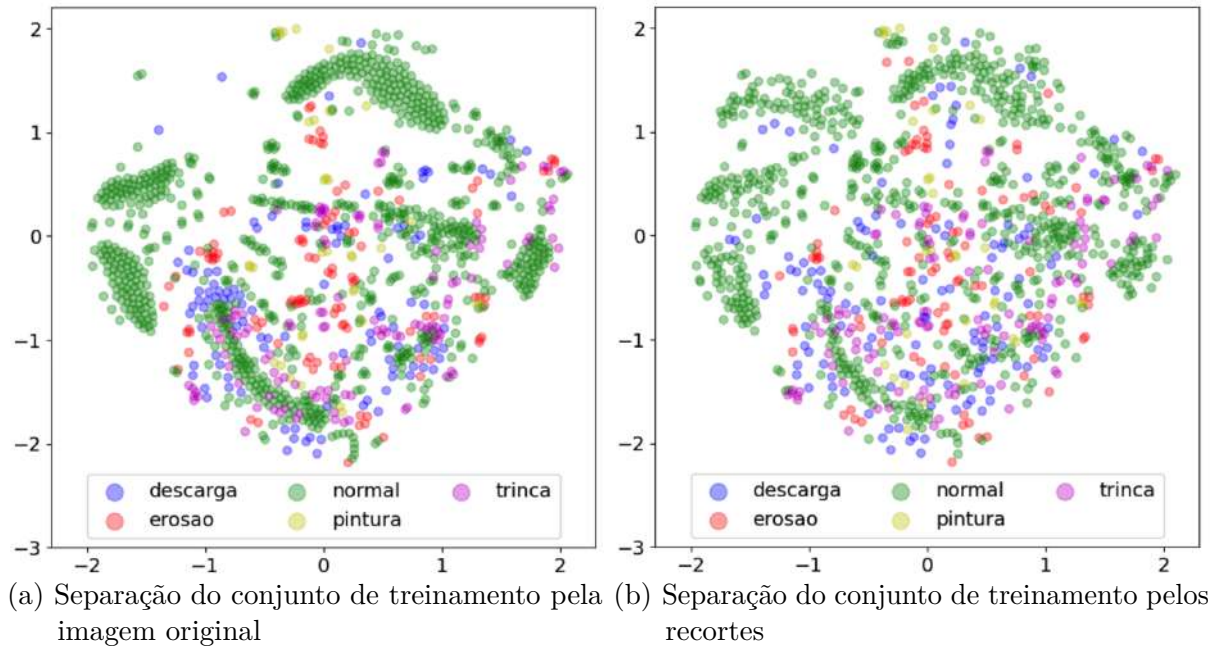
Figura 27 – Análise através do T-SNE da distribuição das imagens nos conjuntos de treino e teste



Fonte: elaborado pelo autor.

Na Figura 28 é ilustrado o detalhamento da distribuição por classe no conjunto de treinamento utilizando as duas abordagens de separação dos conjuntos. Observa-se na Figura 28 (a) que existe uma concentração de imagens normais que não ocorre na Figura 28 (b); isso acaba por impactar os resultados dos modelos treinados com a abordagem de separação pela imagem original.

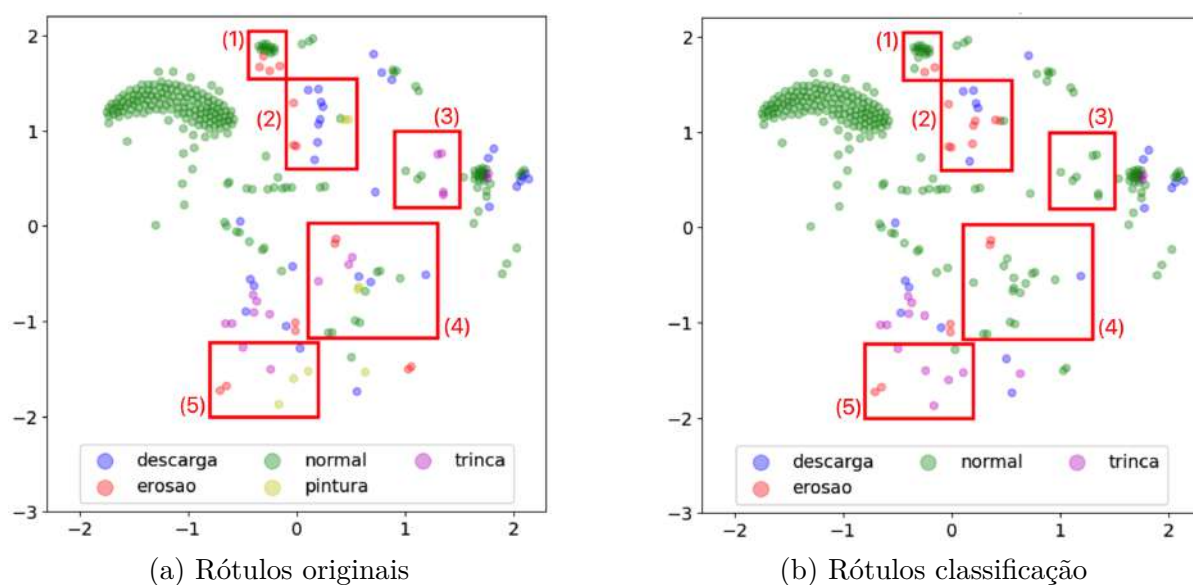
Figura 28 – Análise com o T-SNE da distribuição das classes do conjunto de treinamento utilizando as abordagens de separação dos conjuntos



Fonte: elaborado pelo autor.

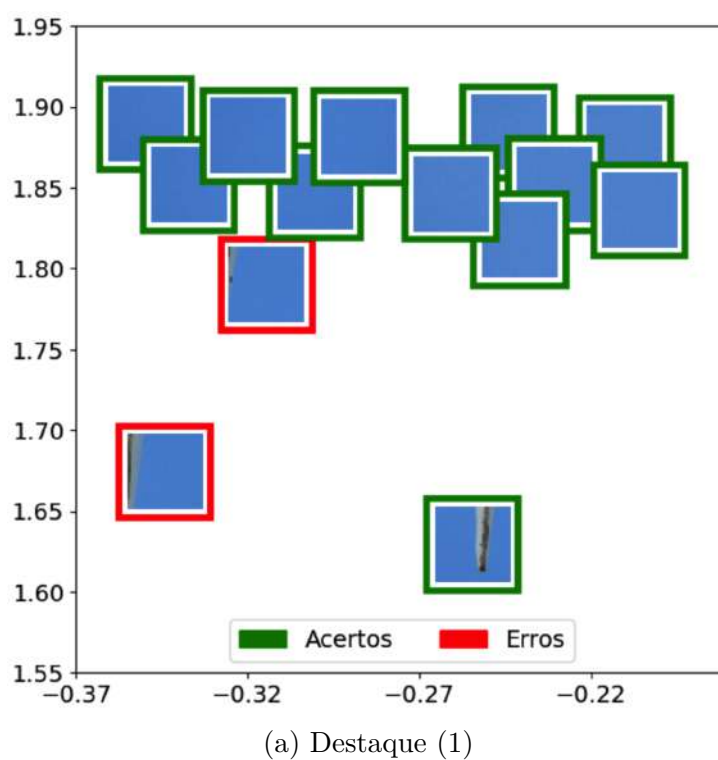
Na Figura 29 é ilustrada a distribuição das classes no conjunto de teste utilizando a separação dos conjuntos pela imagem original, mostrando os rótulos originais e os rótulos das classificações realizadas pelo modelo, sendo mostradas em destaque amostras às quais o modelo atribui o rótulo incorreto. As miniaturas desses recortes são ilustradas nas Figuras 30, 31 e 32. Essa análise indica que a diversificação das imagens é um fator importante para o treinamento de modelos baseados em pequenos conjuntos de dados, sendo esse o motivo dos melhores resultados da abordagem de separação dos conjuntos pelos recortes.

Figura 29 – Distribuição das classes no conjunto de teste utilizando a separação dos conjuntos pela imagem original, com destaque para os erros de classificação



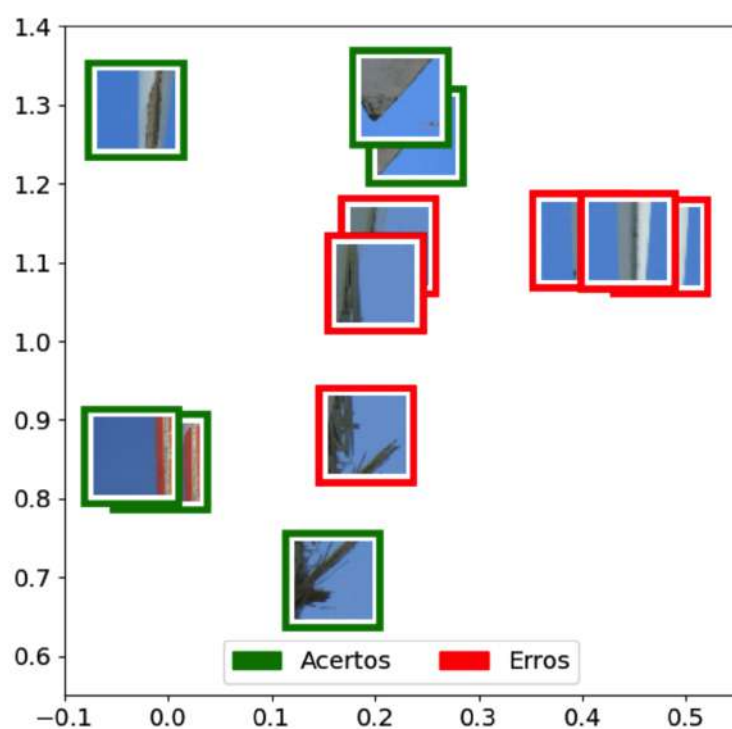
Fonte: elaborado pelo autor.

Figura 30 – Exibição das miniaturas do destaque (1) da Figura 29, com borda verde para os recortes com classificação corretas e vermelho para as incorretas

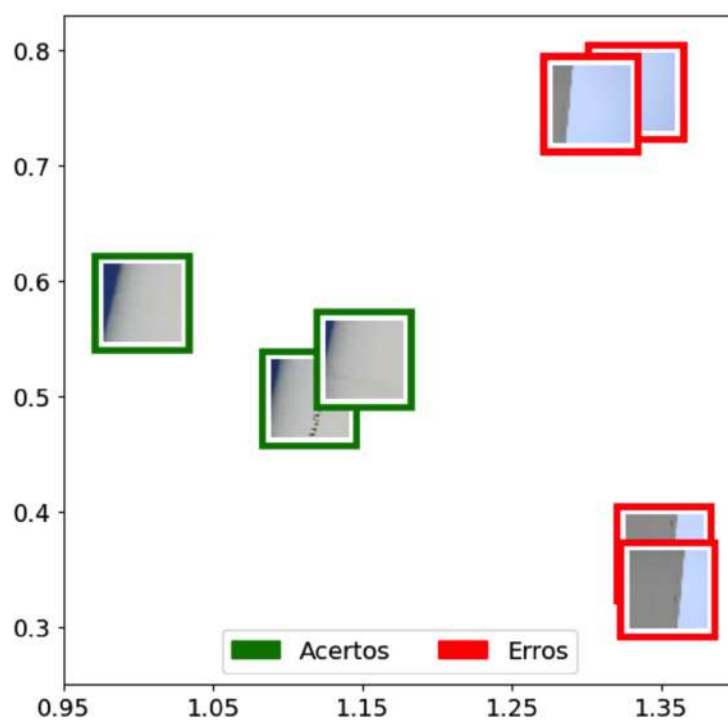


Fonte: elaborado pelo autor.

Figura 31 – Exibição das miniaturas dos destaques (2) e (3) da Figura 29, com borda verde para os recortes com classificação corretas e vermelho para as incorretas



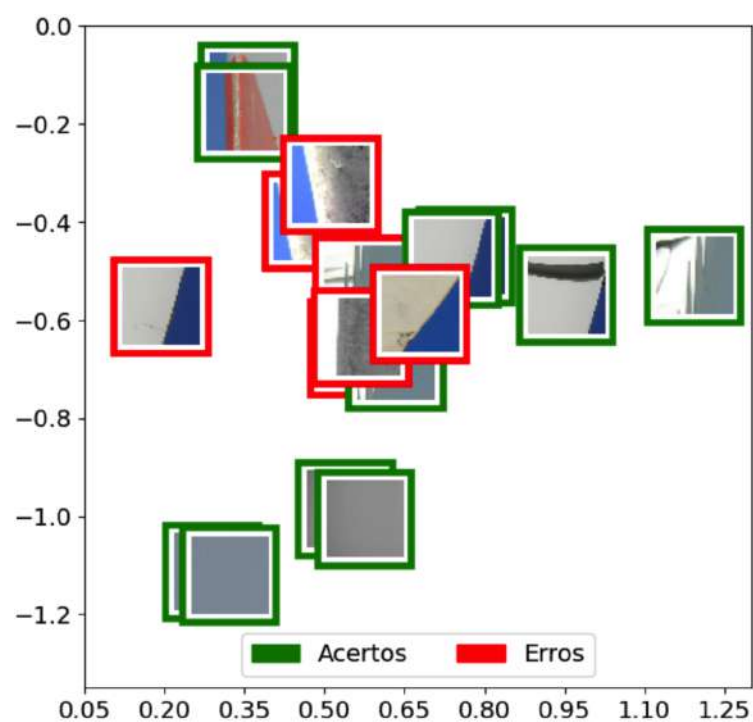
(a) Destaque (2)



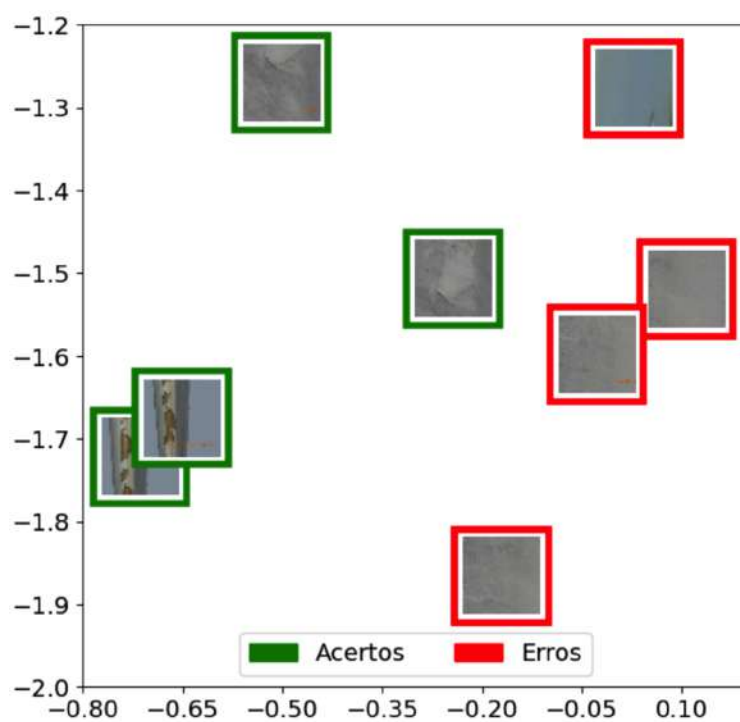
(b) Destaque (3)

Fonte: elaborado pelo autor.

Figura 32 – Exibição das miniaturas dos destaques (4) e (5) da Figura 29, com borda verde para os recortes com classificação corretas e vermelho para as incorretas



(a) Destaque (4)



(b) Destaque (5)

Fonte: elaborado pelo autor.

4.5 Experimento IV: análise do método de aumento de dados e balanceamento

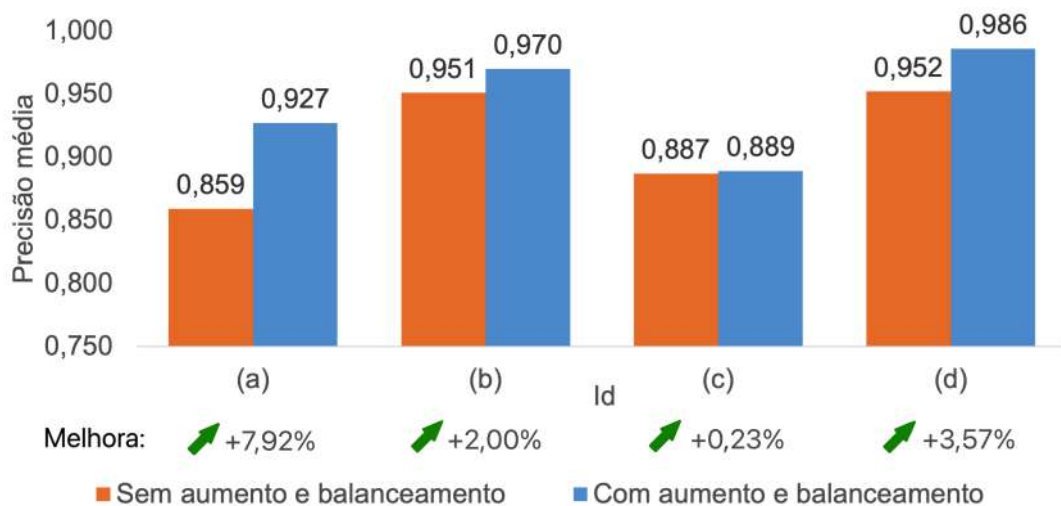
O objetivo desse experimento é verificar a influência do uso ou não do aumento de dados e balanceamento nos recortes gerados. Na abordagem sem o uso do aumento de dados e balanceamento, os recortes não têm nenhuma alteração em relação à extração e separação originais. Quando é aplicado o aumento de dados e balanceamento, os recortes são aleatoriamente replicados, até o limite máximo da classe com maior quantidade, e são aplicadas técnicas de transformações, conforme descritas na Seção 3.1.3.1. As configurações utilizadas para o experimento do uso ou não do aumento de dados e balanceamento são apresentadas na Tabela 9 e os resultados são apresentados na Figura 33.

Tabela 9 – Configurações do experimento para a uso ou não de aumento de dados e balanceamento

Id	Extração dos recortes	Tipo de separação	Modelo
(a)	Janela deslizante	Imagem original	EfficientNetB2
(b)	Janela deslizante	Recorte	
(c)	Aleatório	Imagem original	
(d)	Aleatório	Recorte	

Fonte: elaborada pelo autor.

Figura 33 – Resultados do experimento IV (aumento de dados e balanceamento)



Fonte: elaborada pelo autor.

Analisando as 4 configurações, a técnica baseada no **uso de aumento de dados e balanceamento** traz melhores resultados, com uma média de melhora sobre a não utilização de **3,43%** para o modelo EfficientNetB2. Essa técnica é a melhor em todas as configurações e todos os modelos, apresentando uma melhora média de **4,24%**.

4.6 Resultados dos experimentos

Como resultado final dos experimentos, é encontrada a melhor configuração para a realização do treinamento. Na Tabela 10 é apresentado um resumo da configuração encontrada. O modelo treinado com essa configuração é escolhido como o melhor modelo e avaliado no conjunto de validação, entre as 5 repetições de treinamento, é selecionada a que apresenta a precisão mais próxima da média.

Tabela 10 – Resultados dos experimentos realizados

Experimento	Resposta
Melhor modelo	EfficientNetB2
Extração dos recortes	Posição aleatória
Separação dos conjuntos	Pelos recortes
Aumento de dados e balanceamento	Com aumento de dados e balanceamento

Fonte: elaborada pelo autor.

Os tempos de treinamento, classificação e geração do mapa de calor não são aspectos estudados nesta pesquisa, porém esses tempos podem ser consultados no Apêndice B.

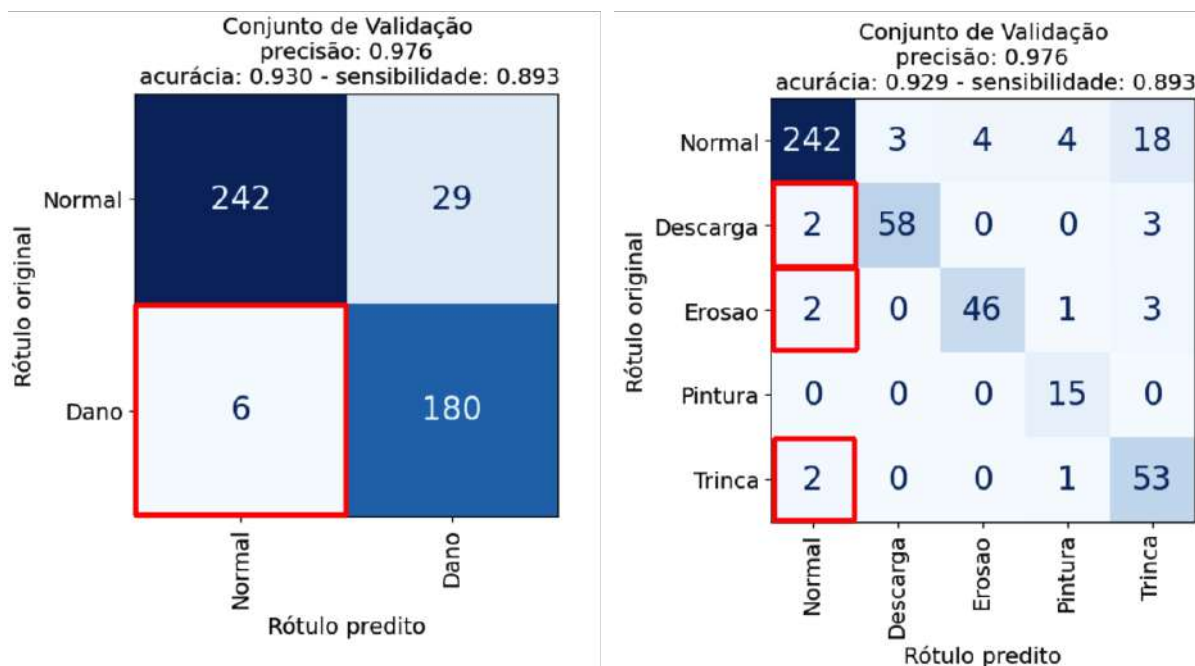
4.7 Resultados no conjunto de validação

O melhor modelo encontrado é avaliado com o conjunto de validação. Na Figura 34 são ilustradas as matrizes de confusão da classificação dos recortes no conjunto de validação. São avaliadas a ocorrência ou não de danos nos recortes na Figura 34 (a), como também de forma mais detalhada para as classes existentes na Figura 34 (b).

Os resultados mostram que o modelo obtém uma precisão de 97,60%, acurácia de 93,00% e sensibilidade de 89,30%. Observa-se que a matriz de confusão resumida mostra 6 recortes com danos que são classificados incorretamente com o rótulo normal, esses são os falsos negativos. A matriz detalhada indica os rótulos corretos desses danos. São eles: descarga, erosão e trinca, cada um com 2 recortes classificados incorretamente.

Fazendo um estudo mais aprofundado do que leva esses danos a serem classificados incorretamente, é possível atribuir que a grande área não relacionada ao dano na imagem influencia no resultado da classificação do modelo. Uma análise utilizando o Grad-CAM (Selvaraju et al., 2017), que é uma técnica para destacar visualmente na imagem as áreas mais relevantes para os pesos de uma rede CNN, é apresentada na Figura 35 e ilustra os resultados das predições, da aplicação do Grad-CAM para o rótulo normal e o Grad-CAM do rótulo verdadeiro.

Figura 34 – Matriz de confusão dos resultados no conjunto de validação, com destaque para os casos de falso negativo



(a) Avaliação resumida (dano e não dano)

(b) Avaliação detalhada (por classe)

Fonte: elaborada pelo autor.

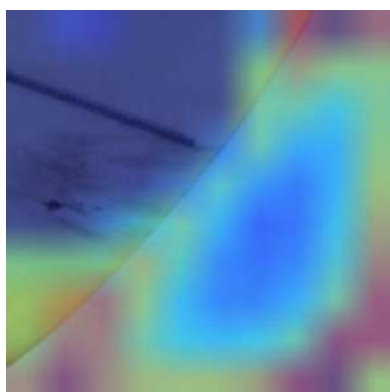
São apresentados um erro de cada classe, tendo: descarga, em (a), (b) e (c); erosão, em (d), (e) e (f); e trinca, em (g), (h) e (i). As figuras possuem tamanhos diferentes devido à aplicação da transformação de redimensionamento dos recortes aleatórios. As classificações incorretas dos recortes estão ilustradas em (a), (d) e (g). As áreas das imagens que sensibilizaram os neurônios do modelo para a classe normal estão destacadas em (b), (e) e (h). E os destaques para a classe verdadeira estão ilustrados em (c), (f) e (i).

Observa-se que os recortes classificados como normal possuem uma vasta área uniforme e sem danos que se mostra importante para os neurônios responsáveis pela identificação das características dessa classe. Em contrapartida, os danos contidos nos recortes são bastante pequenos ou localizados às margens da imagem. Embora os neurônios responsáveis pela identificação desses danos os tenham localizado, não é relevante o suficiente para apresentar um resultado confiável. Esse resultado pode indicar que o método proposto tenha limitações para identificar danos nos estágios iniciais, que geralmente são pequenos e de difícil identificação, até mesmo para um especialista treinado. Já a Figura 36 ilustra a análise do Grad-CAM de recortes classificados corretamente.

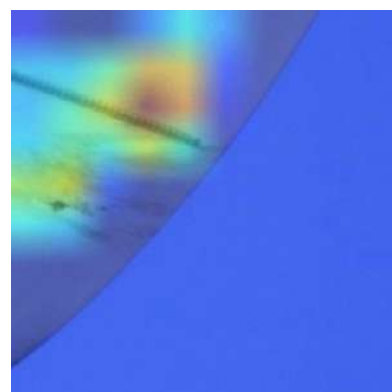
Figura 35 – Análise utilizando o Grad-CAM de alguns casos de recortes classificados como falso negativo



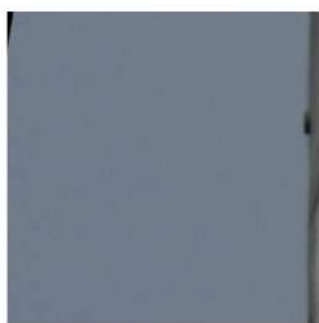
(a) Dano de descarga classificado como normal



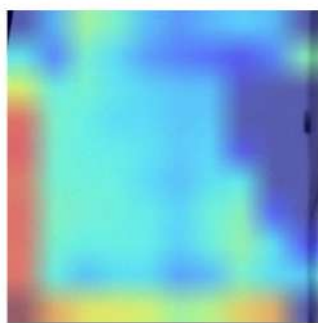
(b) Gradcam classe normal



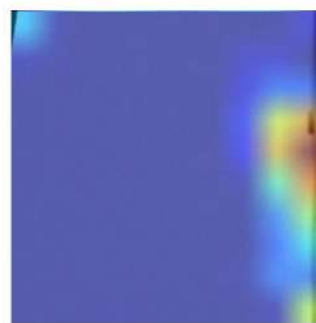
(c) Gradcam classe descarga



(d) Dano de erosão classificado como normal



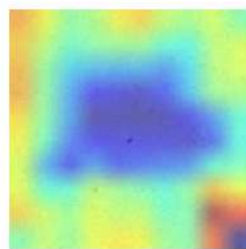
(e) Gradcam classe normal



(f) Gradcam classe erosão



(g) Dano de trinca classificado como normal



(h) Gradcam classe normal



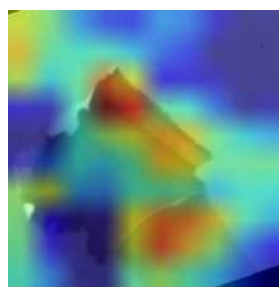
(i) Gradcam classe trinca

Fonte: elaborado pelo autor.

Figura 36 – Análise utilizando o Grad-CAM de recortes de dano de descarga e erosão que foram classificados corretamente



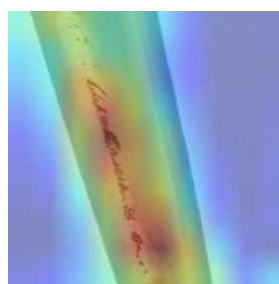
(a) Dano de descarga classificado corretamente



(b) Gradcam classe descarga



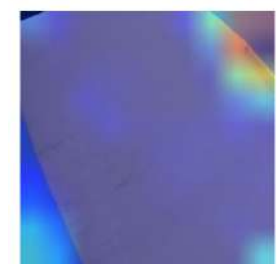
(c) Dano de erosão classificado corretamente



(d) Gradcam classe erosão



(e) Dano de pintura classificado corretamente



(f) Gradcam classe pintura



(g) Dano de trinca classificado corretamente



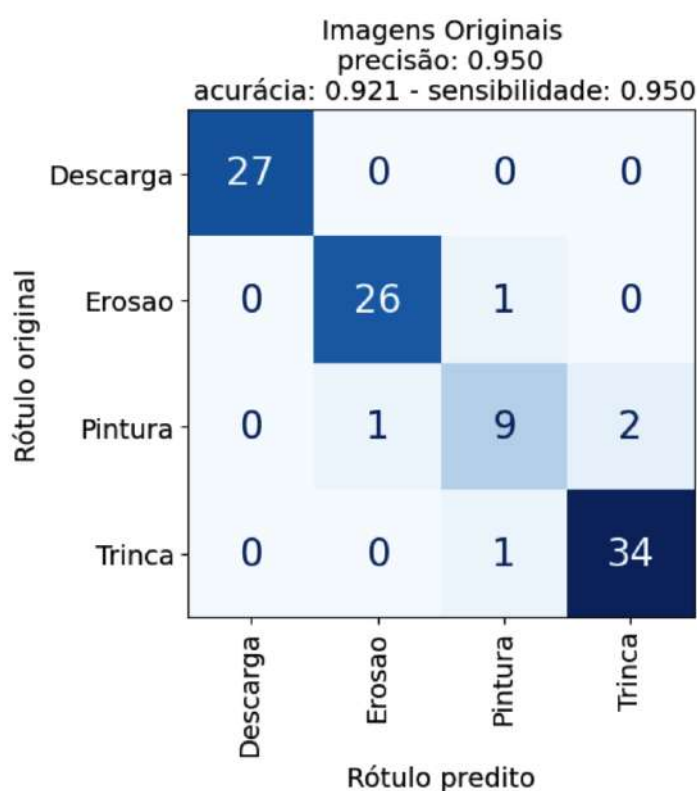
(h) Gradcam classe trinca

Fonte: elaborado pelo autor.

4.8 Resultados da classificação das imagens originais

O melhor modelo também é utilizado para classificar todos os recortes extraídos e, através de um processo de votação, definir uma classe única para cada imagem original, conforme especificado na Seção 3.3.1. Na Figura 37 é ilustrada a matriz de confusão com o resultado final da classificação das imagens originais, obtendo uma precisão de 95,00%, acurácia de 92,10% e sensibilidade de 95,00%.

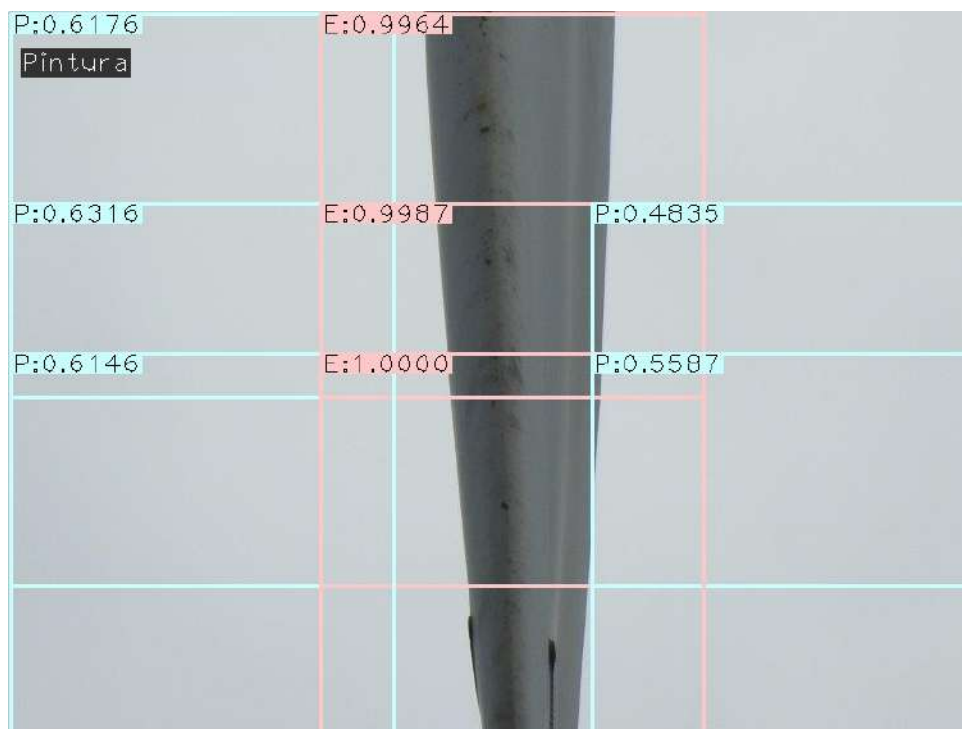
Figura 37 – Matriz de confusão dos resultados das imagens originais



Fonte: elaborada pelo autor.

Mesmo o modelo errando a classificação de alguns recortes, o processo de votação consegue produzir bons resultados com a previsão correta da maioria dos recortes. A Figura 38 é ilustrado o caso da previsão incorreta do dano de erosão, que é classificado como pintura. A imagem ilustra uma pá eólica com o céu nublado; essa característica faz com que o céu possa ser confundido com uma pá eólica. Os recortes com predominância do fundo são classificados como dano de pintura com uma baixa confiabilidade (entre 0,4835 e 0,6316), provavelmente por possuírem alguma inconformidade em alguns *pixels*. Já os danos reais são classificados corretamente para o dano de erosão com uma alta confiabilidade (entre 0,9964 e 1,0000). O processo de votação classifica a imagem original como pintura, devido a 5 recortes terem sido classificados como dano de pintura, enquanto somente 3 recortes são classificados como dano de erosão. Isso indica que o processo de votação pode ser evoluído com outras técnicas, por exemplo, votação ponderada.

Figura 38 – Imagem original de erosão classificada incorretamente como dano de pintura



Fonte: elaborada pelo autor.

Outros casos de classificações incorretas são ilustrados nas Figuras 39 e 40. Na Figura 39, os dois recortes destacados em azul claro são classificados corretamente como dano de pintura, embora tenham resultados baixos (0,5013 e 0,4639). Isso ocorre devido a imagem possuir danos de pintura de tamanho pequeno (círculos vermelhos destacados) em relação ao tamanho do recorte, levando a uma baixa confiança na classificação. Os recortes destacados em azul escuro são classificados incorretamente como dano de trinca. Isso se deve a uma mancha na pintura (seta amarela) e a linha vertical (setas verdes) inerente a estrutura que confundem com uma trinca, já que ambas possuem características de um dano de trinca. Na Figura 40, os três recortes destacados em azul escuro são classificados corretamente como dano de trinca. Os recortes destacados em azul claro são classificados incorretamente como pintura. Observa-se que as classificações incorretas das imagens originais estão relacionadas à classe Pintura, seja como rótulo original ou previsto, conforme detalhado na matriz de confusão (Figura 37). Isso indica que essa classe necessita de um melhor refinamento, justificando-se pela quantidade reduzida de imagens disponibilizadas para essa classe, sendo necessário uma maior quantidade e diversidade de imagens de danos de pintura.

Figura 39 – Imagem original de pintura classificada incorretamente como dano de trinca. Com círculos vermelhos indicando os danos de pintura e setas verdes e amarelas indicando as características que influenciam a classificação como trinca



Fonte: elaborada pelo autor.

Figura 40 – Imagem original de trinca classificada incorretamente como dano de pintura. Com círculos vermelhos indicando as trincas existentes

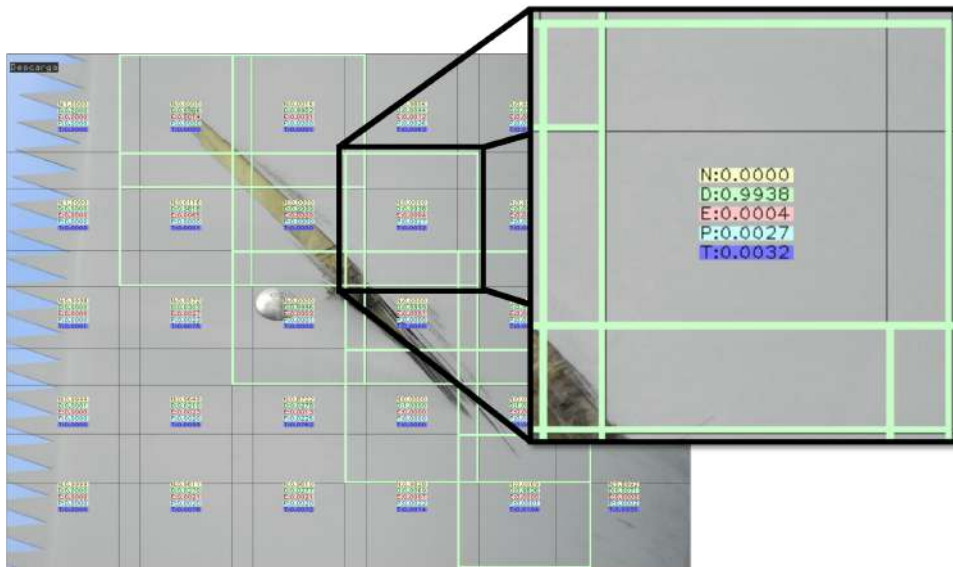


Fonte: elaborada pelo autor.

4.9 Resultados da localização dos danos nas imagens originais

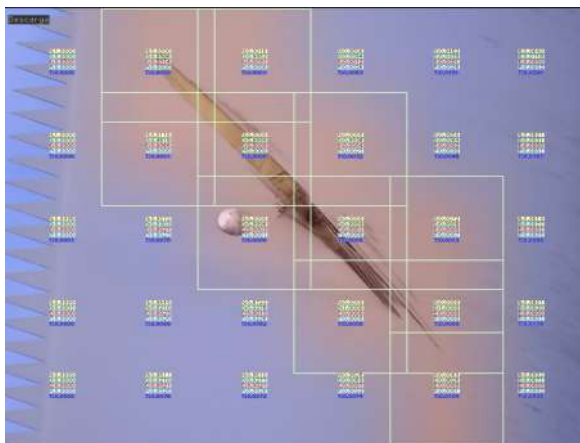
Com base nos resultados das classificações dos recortes um mapa de calor é exibido para destacar a localização do dano. Na Figura 41 é ilustrado o resultado da localização do dano. Na Figura 41 (a) é ilustrada uma imagem de dano por descarga atmosférica, com a exibição da grade guia dos recortes, com os destaques na cor verde para os recortes classificados como descarga e com um dos recortes ampliados, mostrando um valor normalizado de 0,9938 para a classe descarga. Na Figura 41 (b) é exibido um mapa de calor utilizando o valor normalizado para a classe descarga de cada recorte. Na Figura 41 (c) é apresentado o resultado final da localização do dano, sem a exibição das grades.

Figura 41 – Processo de localização do dano



(a) Resultado da classificação dos recortes.

Legenda: N: Normal, D: Descarga, E: Erosão, P: Pintura e T: Trinca



(b) Localização baseado no resultado dos recortes

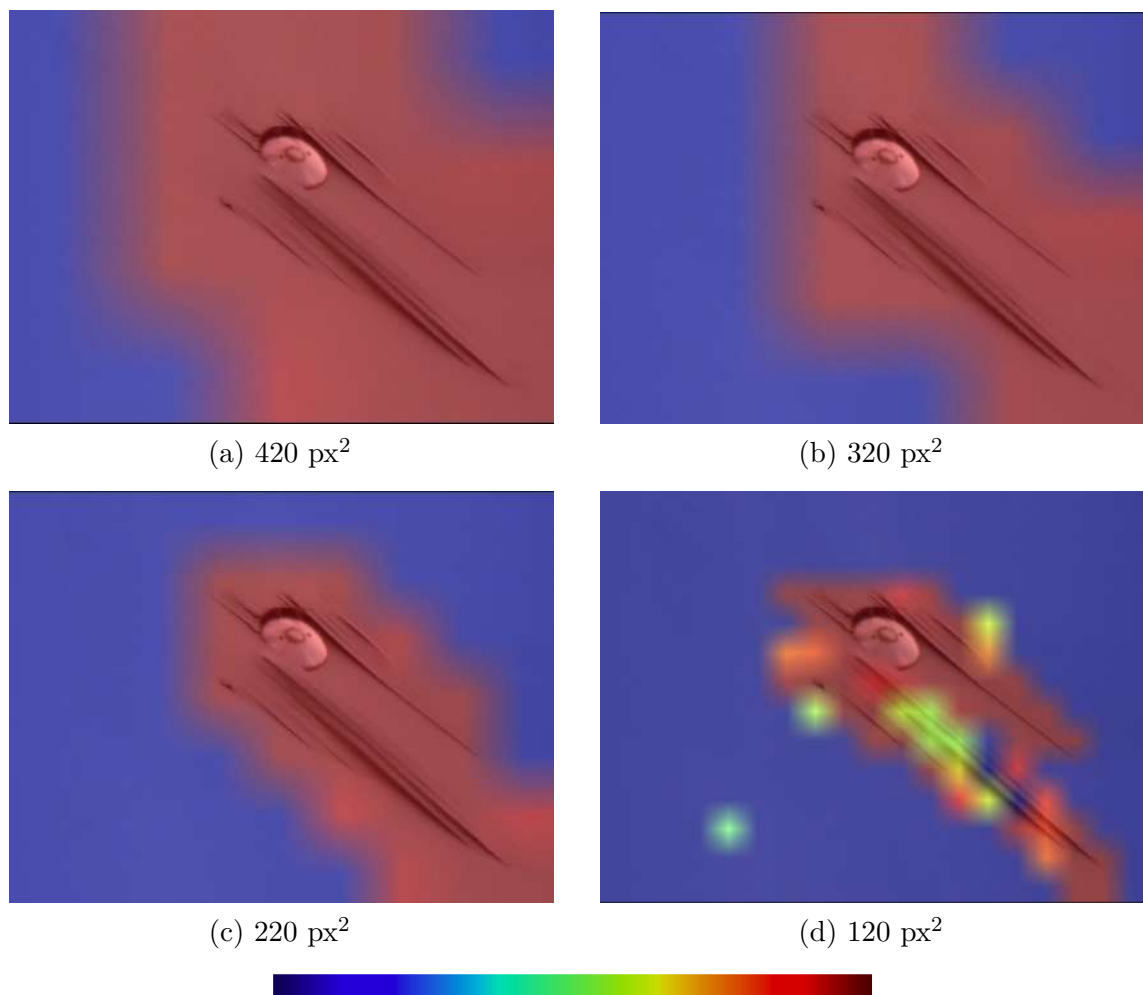


(c) Resultado final da localização

Fonte: elaborado pelo autor.

Observa-se que o mapa de calor não fica exatamente ajustado ao dano, isso deve-se ao fato dos recortes terem a resolução de $320 \times 320 \text{ px}^2$ e alguns recortes classificados como dano terem bastante área não correspondente ao dano em si. Na Figura 42 é apresentada uma avaliação do impacto da variação do tamanho dos recortes no processo de localização durante a etapa de identificação do dano. São exibidos os resultados para os tamanhos de 420 px^2 , 320 px^2 , 220 px^2 e 120 px^2 (Figuras 42 (a), (b), (c) e (d), respectivamente). É possível notar que, com um tamanho maior de recorte, muita área não relativa ao dano é destacada em vermelho. Quando o tamanho do recorte diminui, o mapa de calor vai se ajustando mais ao dano. Porém, em recortes pequenos, caso da Figura 42 (d), os valores dos resultados da classificação também começam a ficar mais baixos, chegando até a exibir falsos positivos e impactando na exibição do mapa de calor, que deixa de exibir algumas áreas em vermelho e começa a exibi-las em amarelo. Uma perspectiva futura é estudar o melhor tamanho de recortes a serem utilizados, ou até mesmo ir ajustando dinamicamente esse tamanho apenas para os recortes classificados como dano.

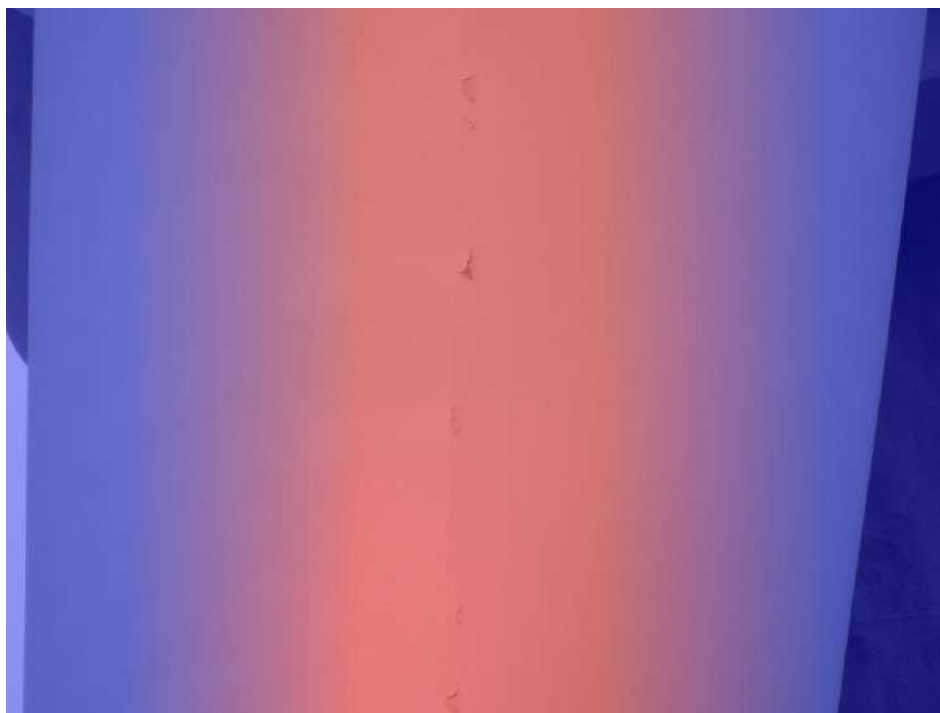
Figura 42 – Variação do tamanho do recortes no processo de localização do dano



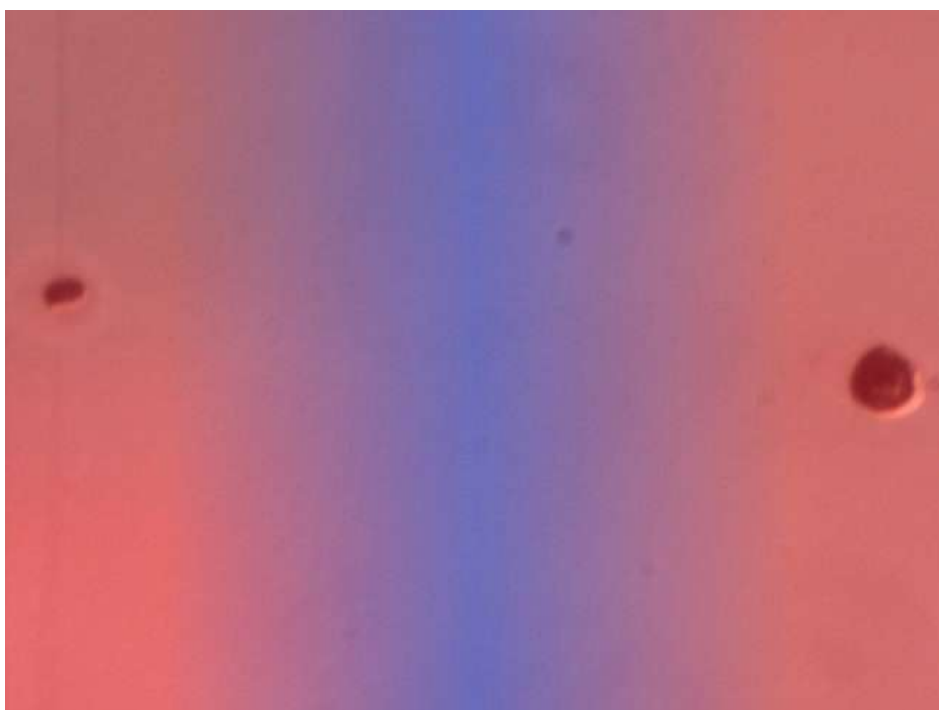
Fonte: elaborado pelo autor.

Outros exemplos da localização dos danos através do mapa de calor são ilustrados nas Figuras 43 e 44. Os danos ilustrados são: dano de pintura (a) e descarga atmosférica (b) na Figura 43; e erosão (a) e trinca (b) na Figura 44.

Figura 43 – Exemplos de localização para os danos de pintura e descarga atmosférica



(a) Dano de pintura



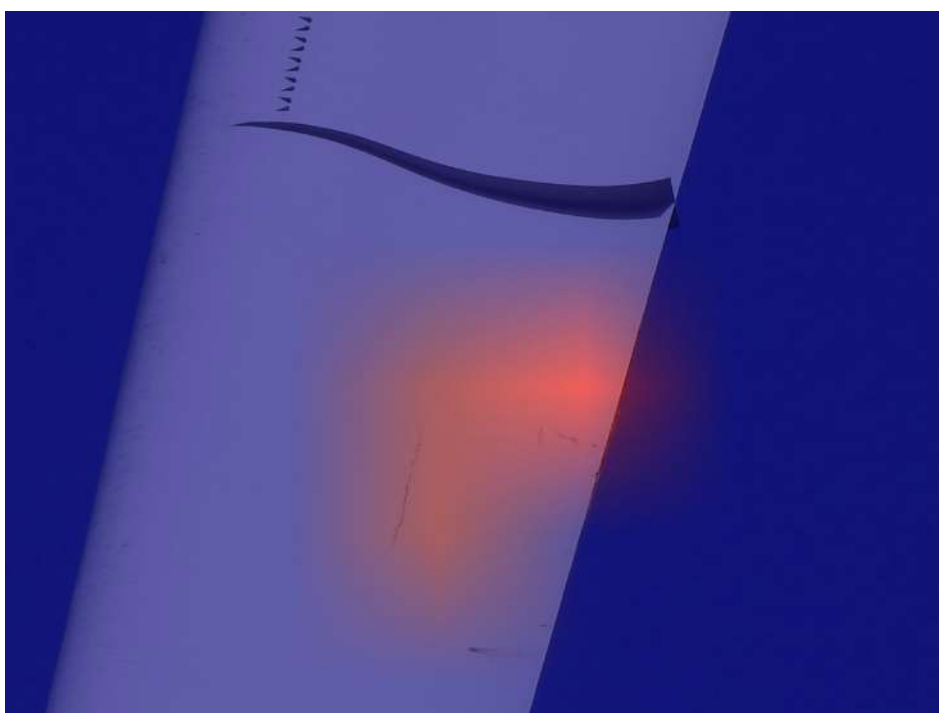
(b) Descarga atmosférica

Fonte: elaborado pelo autor.

Figura 44 – Exemplos de localização para os danos de erosão e trinca



(a) Erosão



(b) Trinca

Fonte: elaborado pelo autor.

4.10 Considerações finais

Neste Capítulo, foram apresentados e analisados os resultados obtidos a partir dos experimentos realizados com diferentes configurações de extração e separação dos recortes, aumento de dados e balanceamento das amostras, e refinamento dos modelos pré-treinados para a classificação de danos em pás de aerogeradores. Os testes demonstraram que a abordagem baseada na extração de recortes aleatórios, aliada à separação dos conjuntos pelos recortes, o uso de aumento de dados e balanceamento de classes, resultou em melhorias expressivas nas métricas de avaliação. Dentre os modelos avaliados, o EfficientNet-B2 obteve o melhor desempenho, destacando-se como a arquitetura mais eficaz para o contexto proposto.

Além da tarefa de classificação, os resultados demonstraram a viabilidade da localização dos danos por meio de mapas de calor gerados com base nas previsões dos recortes. Essa estratégia possibilita interpretar visualmente a região da imagem afetada. No Capítulo seguinte, são apresentadas as conclusões finais da pesquisa, bem como as perspectivas para trabalhos futuros.

5 Conclusão e Trabalhos Futuros

Neste Capítulo são apresentadas as conclusões e planos para evoluir o método proposto.

5.1 Conclusão

Nesta dissertação é apresentado um método para a identificação e localização de danos estruturais em pás de aerogeradores, utilizando técnicas de visão computacional e aprendizado profundo. Este método utiliza abordagens de extração e separação dos recortes nos conjuntos de dados, aumento de dados e balanceamento de classes, e refinamento de modelos pré-treinados para classificação e localização dos danos.

Os experimentos demonstram que, na melhor configuração encontrada, o modelo refinado alcança 97,6% de precisão, 93,0% de acurácia e 89,3% de sensibilidade na classificação dos recortes extraídos, e 90,6% de precisão, 92,9% de acurácia e 90,6% de sensibilidade na classificação das imagens originais. A melhor configuração encontrada utiliza recortes aleatórios, com separação dos conjuntos de dados pelos recortes, com a aplicação de aumento de dados e balanceamento para o refinamento do modelo EfficientNetB2.

A extração de recortes utilizando a abordagem dos recortes aleatórios e a separação dos conjuntos de dados pelos recortes contribuem para uma maior diversidade das amostras de treinamento, possibilitando melhorar os resultados encontrados. Além disso, a extração de um número maior de recortes aleatórios amplia a cobertura da imagem original, aumentando a diversidade das imagens.

Vale destacar que o método proposto também permite a classificação e localização dos danos com o suporte visual de mapas de calor, facilitando a interpretação dos resultados.

Os resultados encontrados indicam que a metodologia proposta tem potencial para ser integrada ao fluxo de trabalho de manutenção preditiva de aerogeradores, contribuindo para a identificação de danos em pás eólicas, podendo otimizar o tempo de análise de imagens por especialistas e reduzir custos operacionais.

5.2 Trabalhos futuros

Como ponto de melhoria, é importante avaliar a eficácia do método com novas imagens, uma vez que o conjunto utilizado foi limitado tanto em volume quanto em variedade, especialmente para a classe de danos de pintura, que teve baixa representatividade, além de não incluir imagens sem danos. Também é possível explorar uma fusão das abordagens

de extração de recortes, combinando as técnicas de janela deslizante e recortes aleatórios, de forma a aproveitar os pontos fortes de ambas e ampliar a cobertura e diversidade das imagens analisadas.

Outra possível melhoria é a análise e otimização do tempo de classificação, com o objetivo de viabilizar sua implementação embarcada em drones, permitindo a identificação de danos em tempo real durante as inspeções. Além disso, o método também pode evoluir com a exploração de outros algoritmos de votação, como a votação ponderada pelo resultado da classificação ou pela distância entre os recortes, buscando melhorar a acurácia da classificação das imagens originais.

Outros aprimoramentos incluem a adaptação do método para classificar imagens que contenham múltiplos tipos de danos, uma vez que, nesta pesquisa, todas as imagens analisadas apresentavam apenas um único dano, o que limitou a classificação a uma única classe por imagem. Também a ampliação da capacidade de localização para identificar não apenas a classe vencedora, mas todas as classes detectadas em uma mesma imagem.

Por fim, o processo de localização pode ser melhorado por meio de uma exibição mais precisa dos danos identificados, o que envolve tanto o ajuste do tamanho ideal dos recortes quanto a substituição do atual mapa de calor por um mapa de saliência, oferecendo uma representação mais detalhada e interpretável das regiões afetadas.

Referências

- AGRAWAL, S. et al. Barely-visible surface crack detection for wind turbine sustainability. In: IEEE. *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Abu Dhabi, Emirados Árabes Unidos, 2024. p. 5933–5939. Citado 4 vezes nas páginas 8, 22, 24 e 25.
- BLADENA; KIRTXTHOMSEN. *Wind Turbine Blades Handbook*. 2019. Acessado em: 25 fev. 2025. Disponível em: <https://www.bladena.com/uploads/8/7/3/7/87379536/cortir_handbook_2019.pdf>. Citado na página 17.
- BUDA, M.; MAKI, A.; MAZUROWSKI, M. A. A systematic study of the class imbalance problem in convolutional neural networks. *Neural networks*, Elsevier, v. 106, p. 249–259, 2018. Citado 2 vezes nas páginas 36 e 37.
- CAO, J.; SHI, Y.; ZHANG, H. Wind energy: History, current design principles, market penetration, environmental impacts. In: IEEE. *2021 International Conference on Control Science and Electric Power Systems (CSEPS)*. Shanghai, China, 2021. p. 398–405. Citado na página 19.
- CHAWLA, N. V. et al. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, v. 16, p. 321–357, 2002. Citado na página 37.
- CICHY, R. M.; KAISER, D. Deep neural networks as scientific models. *Trends in cognitive sciences*, Elsevier, v. 23, n. 4, p. 305–317, 2019. Citado na página 30.
- DANA, H. B.; CHRISTOPHER, M. B. *Computer Vision*. Englewood Cliffs, New Jersey: Prentice Hall, Inc., 1982. Citado na página 29.
- DENG, J. et al. Imagenet: A large-scale hierarchical image database. In: IEEE. *2009 IEEE Conference on Computer Vision and Pattern Recognition*. Miami, FL, USA, 2009. p. 248–255. Citado 2 vezes nas páginas 24 e 33.
- DIEDERIK, K. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. Citado na página 46.
- DU, Y. et al. Damage detection techniques for wind turbine blades: A review. *Mechanical Systems and Signal Processing*, Elsevier, v. 141, p. 106445, 2020. Citado na página 19.
- EÓLICA, C. T. S. de E. *Terminologia Básica Aplicável aos Componentes do Aerogerador*. Associação Brasileira de Ensaios Não Destrutivos e Inspeção (ABENDI), 2020. Prática Recomendada PRe-004, Revisão 00. Disponível em: <https://www.abendi.org.br/abendi/Upload/file/consulta_publica/pre-004_rev00_terminologia-setor-eolico.pdf>. Citado na página 17.
- FOSTER, A. et al. Drone footage wind turbine surface damage detection. In: *2022 IEEE 14th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP)*. Piscataway, NJ, USA: IEEE, 2022. p. 1–5. Citado 2 vezes nas páginas 22 e 23.

GAROLERA, A. C. et al. Lightning damage to wind turbine blades from wind farms in the us. *IEEE Transactions on Power Delivery*, IEEE, v. 31, n. 3, p. 1043–1049, 2014. Citado 2 vezes nas páginas 26 e 27.

Global Wind Energy Council. *GWEC Global Wind Report 2024*. 2024. Acessado em: 31 jan. 2025. Disponível em: <https://gwec.net/wp-content/uploads/2024/05/GWR-2024_digital-version_final-2.pdf>. Citado na página 16.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. Cambridge, Massachusetts, EUA: MIT Press, 2016. <<http://www.deeplearningbook.org>>. Citado na página 46.

GOODFELLOW, I. et al. Generative adversarial nets. In: GHAHRAMANI, Z. et al. (Ed.). *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2014. v. 27. Disponível em: <https://proceedings.neurips.cc/paper_files/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf>. Citado na página 37.

HE, K. et al. Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, NV, USA: IEEE, 2016. p. 770–778. Citado 4 vezes nas páginas 23, 31, 32 e 45.

HOWARD, A. G. et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*, 2017. Citado 2 vezes nas páginas 31 e 32.

IMAN, M.; ARABNIA, H. R.; RASHEED, K. A review of deep transfer learning and recent advancements. *Technologies*, MDPI, v. 11, n. 2, p. 40, 2023. Citado 2 vezes nas páginas 33 e 34.

KAEWNIAM, P. et al. Recent advances in damage detection of wind turbine blades: A state-of-the-art review. *Renewable and Sustainable Energy Reviews*, Elsevier, v. 167, p. 112723, 2022. Citado 2 vezes nas páginas 18 e 19.

KATSAPRAKAKIS, D. A.; PAPADAKIS, N.; NTINTAKIS, I. A comprehensive analysis of wind turbine blade damage. *Energies*, MDPI, v. 14, n. 18, p. 5974, 2021. Citado 2 vezes nas páginas 18 e 27.

KHALIFA, N. E.; LOEY, M.; MIRJALILI, S. A comprehensive survey of recent trends in deep learning for digital images augmentation. *Artificial Intelligence Review*, Springer, p. 1–27, 2022. Citado na página 35.

KHAN, A. A.; LAGHARI, A. A.; AWAN, S. A. Machine learning in computer vision: a review. *EAI Endorsed Transactions on Scalable Information Systems*, v. 8, n. 32, p. e4–e4, 2021. Citado na página 30.

KIRILLOV, A. et al. Segment anything. *arXiv preprint arXiv:2304.02643*, 2023. Disponível em: <<https://arxiv.org/abs/2304.02643>>. Citado na página 87.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, v. 25, 2012. Citado na página 23.

LAGE, E. S.; PROCESSI, L. D. Panorama do setor de energia eólica. Banco Nacional de Desenvolvimento Econômico e Social, 2013. Citado na página 16.

- LECUN, Y. et al. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, Ieee, v. 86, n. 11, p. 2278–2324, 1998. Citado na página 31.
- MISHNAEVSKY, L. et al. Materials for wind turbine blades: An overview. *Materials*, MDPI, v. 10, n. 11, p. 1285, 2017. Citado na página 16.
- MORENO, S. et al. A new vision-based method using deep learning for damage inspection in wind turbine blades. In: IEEE. *2018 15th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE)*. Mexico City, Mexico: IEEE, 2018. p. 1–5. Citado 2 vezes nas páginas 22 e 23.
- RASCHKA, S. Model evaluation, model selection, and algorithm selection in machine learning. *arXiv preprint arXiv:1811.12808*, 2018. Citado na página 43.
- REDMON, J. et al. You only look once: Unified, real-time object detection. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. Las Vegas, NV, USA: IEEE, 2016. p. 779–788. Citado na página 24.
- REN, S. et al. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, v. 28, 2015. Citado na página 23.
- ROSENFELD, A. Computer vision: basic principles. *Proceedings of the IEEE*, IEEE, v. 76, n. 8, p. 863–868, 1988. Citado na página 29.
- RUSSAKOVSKY, O. et al. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)*, v. 115, n. 3, p. 211–252, 2015. Citado na página 31.
- SANDLER, M. et al. Mobilenetv2: Inverted residuals and linear bottlenecks. In: IEEE. *Proceedings of the IEEE conference on computer vision and pattern recognition*. Salt Lake City, UT, USA, 2018. p. 4510–4520. Citado 2 vezes nas páginas 32 e 45.
- Scotland Against Spin. *Turbine Accident Summary to 31 December 2024*. 2024. Acessado em: 31 jan. 2025. Disponível em: <<https://scotlandagainstspin.org/wp-content/uploads/2025/01/Turbine-Accident-Summary-to-31-December-2024.pdf>>. Citado na página 18.
- SELVARAJU, R. R. et al. Grad-cam: Visual explanations from deep networks via gradient-based localization. In: IEEE. *2017 IEEE International Conference on Computer Vision (ICCV)*. Venice, Italy, 2017. p. 618–626. Citado 2 vezes nas páginas 24 e 63.
- SHIHAVUDDIN, A.; CHEN, X. Dtu - drone inspection images of wind turbine. Mendeley Data, 2018. Disponível em: <<https://data.mendeley.com/datasets/hd96prn3nc/1>>. Citado na página 23.
- SHIHAVUDDIN, A. et al. Wind turbine surface damage detection by deep learning aided drone inspection analysis. *Energies*, MDPI, v. 12, n. 4, p. 676, 2019. Citado 2 vezes nas páginas 22 e 23.
- SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014. Citado 2 vezes nas páginas 31 e 45.
- SZELISKI, R. *Computer vision: algorithms and applications*. Cham, Switzerland: Springer Nature, 2022. Citado 2 vezes nas páginas 29 e 30.

- TAN, M.; LE, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In: PMLR. *International conference on machine learning*. Long Beach, California, USA, 2019. p. 6105–6114. Citado 3 vezes nas páginas 31, 32 e 45.
- TURNBULL, A.; CARROLL, J. Cost benefit of implementing advanced monitoring and predictive maintenance strategies for offshore wind farms. *Energies*, MDPI, v. 14, n. 16, p. 4922, 2021. Citado na página 18.
- WANG, W. et al. Review of the typical damage and damage-detection methods of large wind turbine blades. *Energies*, MDPI, v. 15, n. 15, p. 5672, 2022. Citado 3 vezes nas páginas 18, 28 e 29.
- XU, M. et al. A comprehensive survey of image augmentation techniques for deep learning. *Pattern Recognition*, Elsevier, p. 109347, 2023. Citado na página 34.
- XU, Y.; GOODACRE, R. On splitting training and validation set: a comparative study of cross-validation, bootstrap and systematic sampling for estimating the generalization performance of supervised learning. *Journal of analysis and testing*, Springer, v. 2, n. 3, p. 249–262, 2018. Citado na página 43.
- YANG, P. et al. The surface damage identifications of wind turbine blades based on resnet50 algorithm. In: IEEE. *2020 39th Chinese Control Conference (CCC)*. Shenyang, China, 2020. p. 6340–6344. Citado 2 vezes nas páginas 22 e 23.
- ZHANG, Z.; SHU, Z. Unmanned aerial vehicle (uav)-assisted damage detection of wind turbine blades: A review. *Energies*, MDPI, v. 17, n. 15, p. 3731, 2024. Citado na página 19.

Apêndices

APÊNDICE A – Ferramenta para tratamento e análise das imagens originais

Uma ferramenta específica foi desenvolvida nesse trabalho para auxiliar nas tarefas de rotulação da imagem original, extração dos recortes em grade e aleatórios, treinamento do modelo, classificação dos recortes e localização dos danos, análise com o Grad-CAM, além de testes de segmentação. Na Tabela 11 é mostrado os parâmetros de entrada da ferramenta. O único parâmetro obrigatório é o caminho das imagens originais (`--folder`).

Tabela 11 – Principais parâmetros da ferramenta de rotulação e recorte

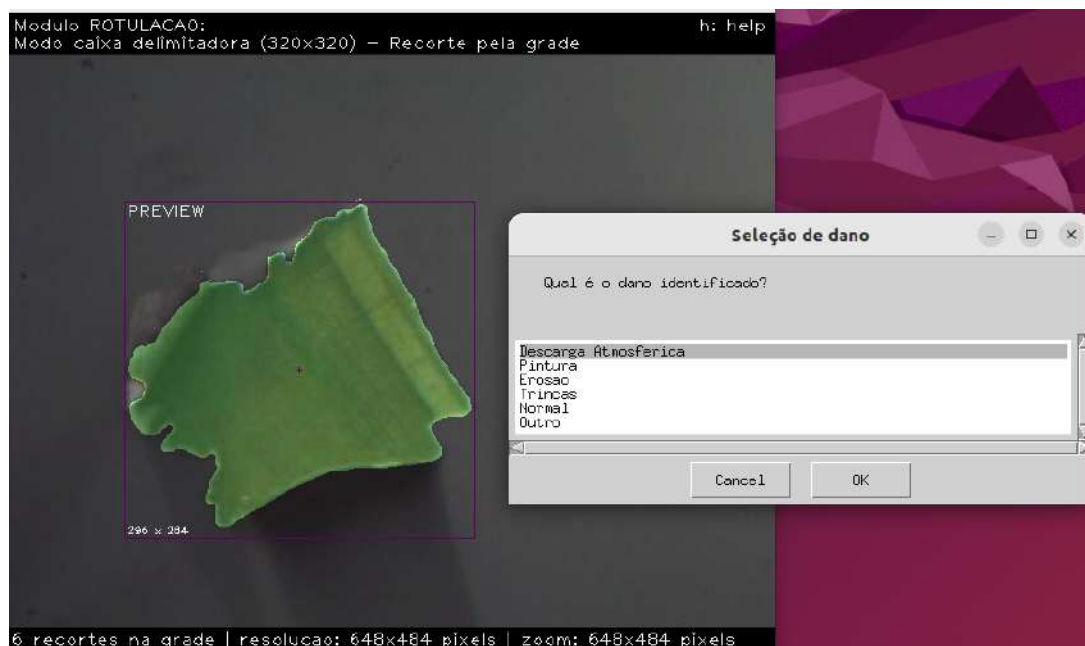
Parâmetros	Descrição
<code>--folder</code>	Caminho das imagens originais.
<code>--folder_mask</code>	Caminho das máscaras com a localização dos danos.
<code>--folder_parameter</code>	Caminho dos parâmetros das caixas delimitadoras.
<code>--folder_parameter_sam</code>	Caminho dos parâmetros da configuração do SAM
<code>--bb_x_size</code>	Tamanho horizontal do bounding box
<code>--bb_y_size</code>	Tamanho vertical do bounding box
<code>--extract</code>	Indica que os recortes extrados deverão ser salvos no caminho especificado.
<code>--extract_sam</code>	Extraí o SAM das imagens para a pasta indicada
<code>--extract_x_size</code>	Tamanho horizontal da imagem extraída
<code>--extract_y_size</code>	Tamanho vertical da imagem extraída
<code>--overlap</code>	Indica o tamanho mínimo de sobreposição de uma grade na outra.
<code>--read_only</code>	Não salva as informações alteradas nos arquivos de parâmetros
<code>--force_size</code>	Força um tamanho do bounding box, substituindo o valor encontrado no arquivo de parâmetros
<code>--quiet</code>	Executa a extração de forma silenciosa. Não exibe componentes visuais.

Fonte: elaborada pelo autor.

A.1 Módulo rotulação

O módulo de **Rotulação** permite que o especialista rotule manualmente o dano para posterior realização do treinamento. Para isso, é possível realizar a segmentação manual dos *pixels* referentes aos danos com uma opção de “pintura” sobre a imagem ou adicionar uma caixa delimitadora sobre uma determinada área. Na Figura 45 é ilustrada a inclusão de uma caixa delimitadora em um dano que já foi segmentado.

Figura 45 – Inclusão de caixa delimitadora



Fonte: elaborado pelo autor.

A funcionalidade de segmentação foi utilizada juntamente com um especialista em danos de pás eólicas, indicado pela empresa fabricante de aerogeradores, para fazer a demarcação das áreas com danos. Ao final, todas as segmentações foram revisadas e validadas por este especialista.

A funcionalidade de visualização das grades foi implementada para definir o posicionamento das mesmas. A ferramenta realiza um cálculo a partir de três parâmetros: o tamanho da imagem, o tamanho da grade e o tamanho da sobreposição (*overlap*). O parâmetro do tamanho da sobreposição define o valor mínimo a ser considerado. No caso de o tamanho da imagem não ser divisível pelo tamanho das grades com sobreposição, a ferramenta adiciona mais um conjunto de grades (horizontal ou vertical) e aumenta o tamanho da sobreposição para conseguir um ajuste mais adequado das grades. A implementação dessa funcionalidade é demonstrada no Algoritmo 1 e seu resultado é ilustrado na Figura 46.

A.2 Módulo recortes aleatórios

O módulo de **Recortes Aleatórios** permite que os recortes sejam extraídos utilizando uma outra abordagem. A posição central dos recortes é definida aleatoriamente e é possível habilitar separadamente as transformações de rotação e redimensionamento. A máscara do dano é utilizada para atribuir o rótulo ao recorte gerado. Na Figura 47 é ilustrado a utilização dessa funcionalidade.

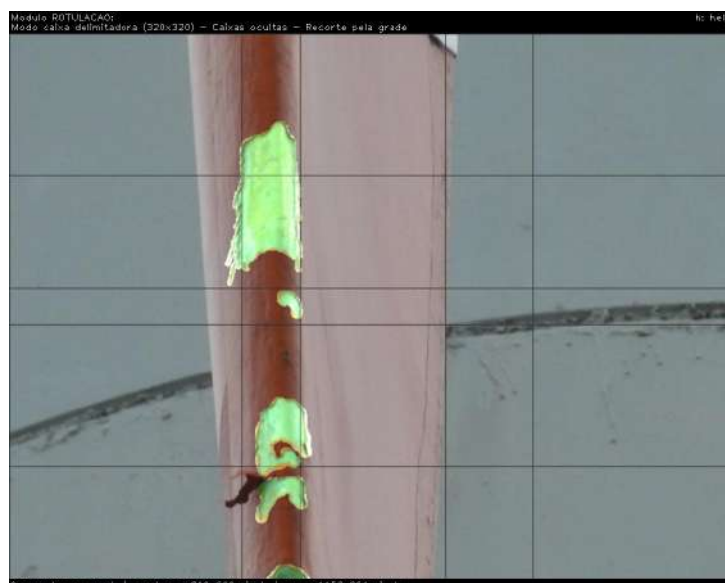
Algoritmo 1 Define posicionamento das grades na imagem

```

1: entradas:
2:   imagem, tamanho da imagem original no formato (x,y)
3:   grade, tamanho da grade no formato (x,y)
4:   sobreposição, tamanho no formato (x,y) que uma grade irá se sobrepor a outra
5: início
6:   vetor_coordenadas  $\leftarrow$  [ ] {Vetor com as coordenadas das grades}
7:   deslocamento  $\leftarrow$  grade – sobreposicao {Calcula deslocamento das grades}
8:   qtde_grades  $\leftarrow$  imagem div deslocamento {Calcula quantas grades cabem}
9:
10:  {Calcula sobra de pixels na coordenada x, caso não seja divisível}
11:  se (imagemx mod deslocamentox) > 0 faça
12:    qtde_gradesx  $\leftarrow$  qtde_gradesx + 1
13:  fim se
14:  sobrax  $\leftarrow$  ((qtde_gradesx * deslocamentox) – imagex) div (qtde_gradesx – 1)
15:
16:  {Calcula sobra de pixels na coordenada y, caso não seja divisível}
17:  se (imagemy mod deslocamentoy) > 0 faça
18:    qtde_gradesy  $\leftarrow$  qtde_gradesy + 1
19:  fim se
20:  sobray  $\leftarrow$  ((qtde_gradesy * deslocamentoy) – imagey) div (qtde_gradesy – 1)
21:
22:  deslocamento  $\leftarrow$  grade – sobreposicao – sobra {Define novo deslocamento considerando as sobras}
23:
24:  {Calcula coordenadas das grades}
25:  para i2 até qtde_gradesy faça
26:    para i1 até qtde_gradesx faça
27:
28:      {Calcula coordenadas x}
29:      x1  $\leftarrow$  i1 * deslocamentox
30:      x2  $\leftarrow$  x1 + gradex
31:      se x2 > imagemx faça
32:        x1  $\leftarrow$  x1 – (x2 – imagemx)
33:        x2  $\leftarrow$  x2 – (x2 – imagemx)
34:      fim se
35:
36:      {Calcula coordenadas y}
37:      y1  $\leftarrow$  i2 * deslocamentoy
38:      y2  $\leftarrow$  y1 + gradey
39:      se y2 > imagemy faça
40:        y1  $\leftarrow$  y1 – (y2 – imagemy)
41:        y2  $\leftarrow$  y2 – (y2 – imagemy)
42:      fim se
43:
44:      vetor_coordenadas  $\leftarrow$  INSIRA(x1, y1, x2, y2, vetor_coordenadas)
45:    fim para
46:  fim para
47:  retorna vetor_coordenadas

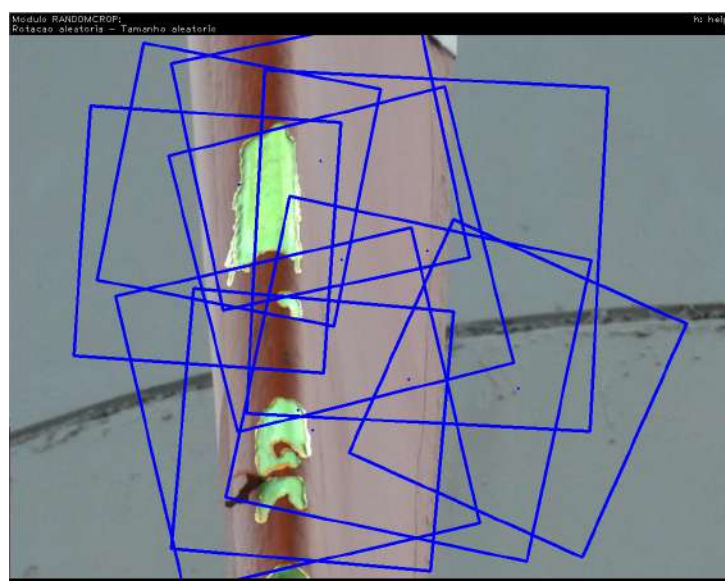
```

Figura 46 – Funcionalidade de geração da grade



Fonte: elaborado pelo autor.

Figura 47 – Funcionalidade de geração dos recortes aleatórios



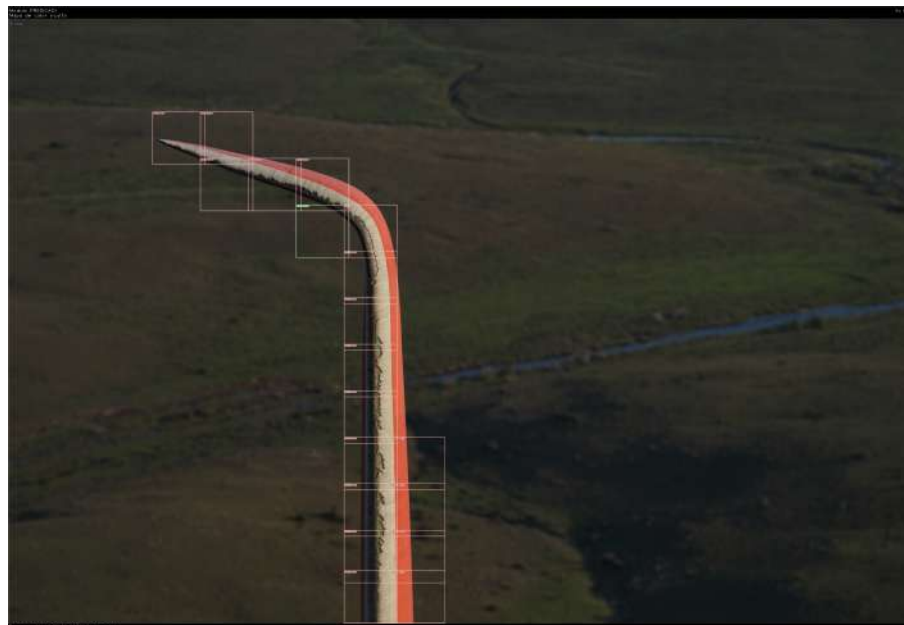
Fonte: elaborado pelo autor.

A.3 Módulo de predição

O módulo de **Predição** permite visualizar o resultado da utilização do modelo treinado para classificar os recortes definidos pelo Algoritmo 1. É possível habilitar a exibição da previsão com caixas delimitadoras, da previsão com texto e do mapa de calor; mudar o *layout* dos textos para exibir somente o rótulo vencedor e mostrar em formato percentual ou numérico. Na Figura 48 (a) é ilustrada a utilização da funcionalidade de

predição com somente as caixas delimitadoras habilitadas. E na Figura 48 (b) são ilustradas as caixas delimitadoras e o mapa de calor habilitados.

Figura 48 – Funcionalidade de predição dos recortes



(a) Somente as caixas delimitadoras habilitadas



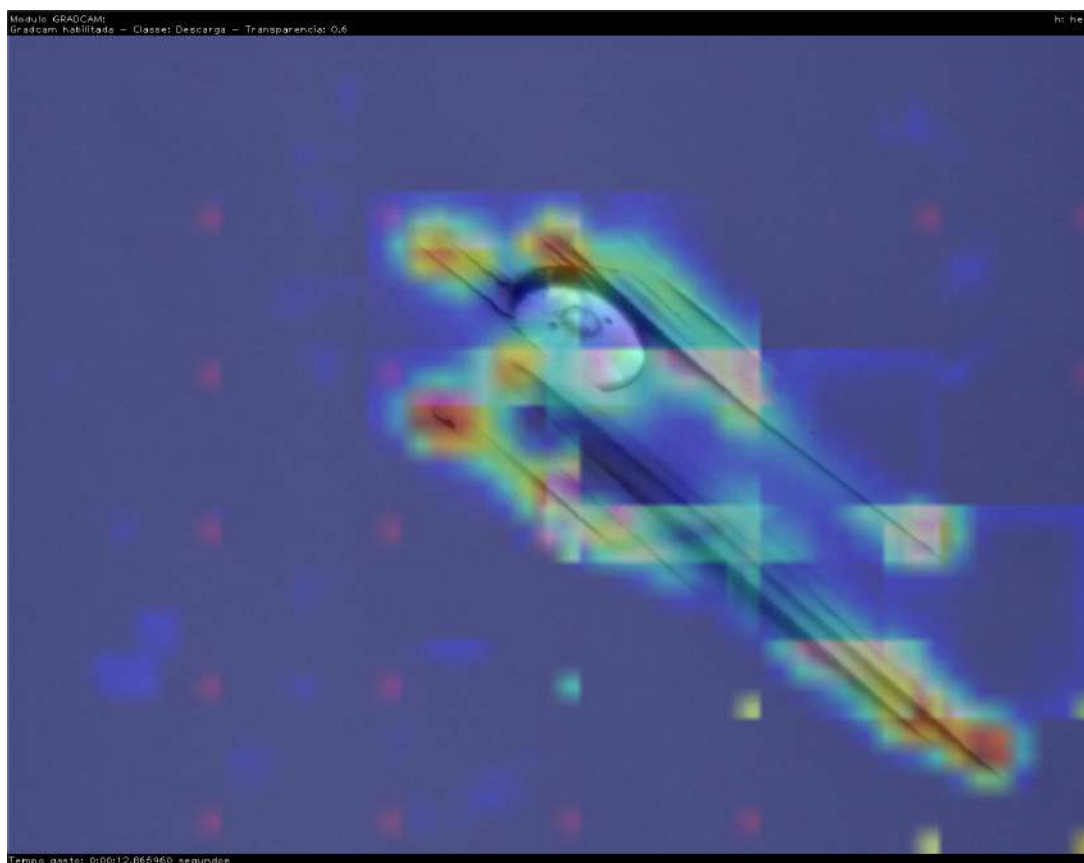
(b) Caixas delimitadoras e o mapa de calor habilitados

Fonte: elaborado pelo autor.

A.4 Módulo de análise Grad-CAM

O módulo de **Grad-CAM** permite analisar as áreas da imagem que mais sensibilizaram os pesos da última camada convolucional. É possível verificar o resultado do Grad-CAM para cada classe individualmente. Na Figura 49 é ilustrada a utilização da funcionalidade de análise do Grad-CAM.

Figura 49 – Funcionalidade de análise do Grad-CAM



Fonte: elaborado pelo autor.

A.5 Módulo de segmentação SAM

O módulo de **Segmentação** permite aplicar na imagem pontos que serviram como base para o algoritmo *Segment Anything Meta* (Kirillov et al., 2023) utilizar para segmentar a imagem. Essa funcionalidade foi utilizada em alguns experimentos que não obtiveram bons resultados. Primeiramente, devido a incluir no método a responsabilidade de ter que definir a quantidade e localização dos pontos para conseguir uma segmentação correta, o que não é uma tarefa fácil. E, em segundo lugar, mesmo conseguindo isolar somente a pá eólica do fundo, a borda da segmentação apresenta muito serrilhado em muitas imagens, o

que acaba confundindo com o dano de trinca, por exemplo. Na Figura 50 é ilustrada a utilização da funcionalidade de segmentação com o SAM.

Figura 50 – Funcionalidade de segmentação SAM



Fonte: elaborado pelo autor.

A.6 Módulo de treinamento

O módulo de **Treinamento** está internamente dentro do código-fonte da ferramenta desenvolvida. Ele permite que outros códigos desenvolvidos em Python, como exemplo, notebooks Jupyter ou Colab, possam utilizar as classes e funções para realizar treinamentos de modelos.

APÊNDICE B – Tempos de classificação e treinamento

Embora esse estudo não tenha como foco a análise dos tempos de execução e treinamento dos modelos, as seções a seguir exibem alguns exemplos dos tempos que foram registrados.

B.1 Tempos de classificação

Na Tabela 12 é detalhado alguns tempos de classificação coletados para imagens em diversas resoluções. São mostradas informações do nome do arquivo, sua resolução, a quantidade de recortes que foram gerados e classificados e os tempos sem e com o uso de GPU.

Tabela 12 – Tempo de classificação de todos os recortes

Arquivo	Resolução	Recortes	Sem GPU	Com GPU
descarga_01.jpg	648 × 484	6	00:00:01.642576	00:00:00.426646
descarga_05.jpg	800 × 600	9	00:00:02.519167	00:00:00.598430
erosao_08.jpg	1024 × 768	12	00:00:03.212247	00:00:00.771084
descarga_06.jpg	1632 × 1224	30	00:00:09.238286	00:00:01.893359
trinca_35.jpg	4032 × 3024	154	00:00:42.916935	00:00:09.817154
erosao_27.jpg	4608 × 3456	192	00:00:59.895024	00:00:12.103420
1-0-1688597565697-DSC06707-4812033.jpg	5456 × 3632	247	00:01:06.106782	00:00:15.847702
ffb8bba9-e065-4cd1-a73a-2254250db3c2_1067.jpg	5280 × 3956	266	00:01:10.346666	00:00:17.188168

Fonte: elaborada pelo autor.

B.2 Tempos de treinamento

Na Tabela 13 é detalhado alguns tempos de treinamento dos modelos. A coluna de configuração refere-se ao modelo que foi treinado de acordo com a numeração contida na Figura 21. Alguns modelos possuem informações entre parênteses que indicam os testes que foram realizados na geração de recortes aleatórios com a variação da quantidade de recortes gerados, conforme ilustrado na Figura 25. Em seguida, estão os tempos de cada repetição e a média desses tempos.

Tabela 13 – Tempo de treinamento dos modelos

Configuração	Repetições					Média
	1	2	3	4	5	
01	0:08:33	0:09:08	0:06:57	0:07:40	0:08:49	00:08:13
02	0:06:27	0:05:26	0:04:16	0:05:02	0:05:32	00:05:21
03	0:12:15	0:11:17	0:08:35	0:08:52	0:11:06	00:10:25
04	0:29:47	0:23:16	0:21:05	0:21:02	0:26:53	00:24:25
05	0:04:25	0:03:51	0:04:15	0:04:21	0:05:12	00:04:25
06	0:02:54	0:01:48	0:02:07	0:02:30	0:02:10	00:02:18
07	0:04:33	0:03:28	0:03:44	0:03:34	0:03:40	00:03:48
08	0:17:52	0:07:46	0:11:09	0:08:41	0:19:33	00:13:00
09	0:12:24	0:13:19	0:09:41	0:09:32	0:13:29	00:11:41
10	0:08:32	0:06:25	0:08:58	0:07:20	0:07:41	00:07:47
11	0:14:43	0:10:38	0:13:27	0:11:06	0:15:29	00:13:05
12	0:39:47	0:33:38	0:53:58	0:42:48	0:38:03	00:41:39
13	0:07:00	0:09:24	0:06:07	0:06:39	0:08:23	00:07:31
14	0:05:10	0:04:16	0:08:55	0:05:37	0:04:31	00:05:42
15	0:05:50	0:06:47	0:08:55	0:08:24	0:12:10	00:08:25
16	0:17:19	0:18:51	0:16:06	0:23:35	0:25:28	00:20:16
17	0:08:01	0:13:21	0:06:05	0:07:22	0:11:53	00:09:20
17 (2x)	0:13:22	0:13:18	0:12:46	0:12:51	0:13:44	00:13:12
17 (3x)	0:19:05	0:19:52	0:18:51	0:18:48	0:19:12	00:19:10
17 (4x)	0:24:35	0:24:37	0:24:42	0:24:42	0:24:45	00:24:40
18	0:03:47	0:03:36	0:02:26	0:03:20	0:04:59	00:03:38
19	0:08:02	0:09:03	0:05:45	0:07:11	0:08:40	00:07:44
20	0:18:32	0:20:48	0:14:24	0:28:56	0:20:47	00:20:41
21	0:05:49	0:03:35	0:11:22	0:04:56	0:05:06	00:06:10
22	0:02:40	0:01:55	0:06:48	0:02:36	0:02:12	00:03:14
23	0:05:34	0:03:40	0:04:44	0:03:24	0:04:12	00:04:19
24	0:15:58	0:07:38	0:18:25	0:11:04	0:08:23	00:12:18
25	0:11:09	0:09:16	0:09:21	0:12:49	0:11:39	00:10:51
26	0:04:49	0:03:42	0:05:43	0:06:14	0:05:18	00:05:09
27	0:10:05	0:08:35	0:08:00	0:13:43	0:12:35	00:10:36
28	0:43:35	0:30:41	0:24:52	0:27:02	0:31:35	00:31:33
29	0:08:15	0:05:33	0:06:24	0:08:09	0:10:44	00:07:49
29 (3x)	0:28:20	0:24:26	0:38:34	0:17:40	0:13:23	00:24:29
30	0:04:46	0:05:30	0:03:28	0:04:27	0:03:18	00:04:18
31	0:13:25	0:06:30	0:08:22	0:10:29	0:10:54	00:09:56
32	0:23:46	0:13:04	0:41:18	0:21:40	0:27:11	00:25:24

Fonte: elaborada pelo autor.

B.3 Tempos do Grad-CAM

Na Tabela 14 é detalhado alguns tempos para a geração do mapa de calor do Grad-CAM para a análise de uma única classe. Para a avaliação das classes definidas no

método, é necessária a multiplicação dos tempos por 5. São mostradas informações do nome do arquivo, sua resolução, a quantidade de recortes que foram gerados e os tempos sem e com o uso de GPU para a análise de uma única classe.

Tabela 14 – Tempo de geração do mapa de calor do GradCam

Arquivo	Resolução	Recortes	Sem GPU	Com GPU
descarga_01.jpg	648×484	6	00:00:11.614508	00:00:03.059431
descarga_05.jpg	800×600	9	00:00:17.249171	00:00:03.955823
erosao_08.jpg	1024×768	12	00:00:23.983076	00:00:05.151997
descarga_06.jpg	1632×1224	30	00:00:59.103169	00:00:12.935715
trinca_35.jpg	4032×3024	154	00:04:59.589517	00:01:06.957830
erosao_27.jpg	4608×3456	192	00:06:04.434580	00:01:24.465940
1-0-1688597565697-DSC06707-4812033.jpg	5456×3632	247	00:07:52.451017	00:01:50.089444
ffb8bba9-e065-4cd1-a73a-2254250db3c2_1067.jpg	5280×3956	266	00:08:31.944988	00:01:55.231119

Fonte: elaborada pelo autor.