



UNIVERSIDADE FEDERAL DE SERGIPE
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**Classificação e segmentação de danos em edifícios após
desastres naturais utilizando dados tabulares e imagens
registradas pré e pós-desastre**

Dissertação de Mestrado

Antônio dos Santos Ramos Neto



São Cristóvão – Sergipe

2024

UNIVERSIDADE FEDERAL DE SERGIPE
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Antônio dos Santos Ramos Neto

**Classificação e segmentação de danos em edifícios após
desastres naturais utilizando dados tabulares e imagens
registradas pré e pós-desastre**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Sergipe como requisito parcial para a obtenção do título de mestre em Ciência da Computação.

Orientador(a): Prof. Daniel Dantas

São Cristóvão – Sergipe

2024

**FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL
UNIVERSIDADE FEDERAL DE SERGIPE**

R175c Ramos Neto, Antônio dos Santos
Classificação e segmentação de danos em edifícios após
desastres naturais utilizando dados tabulares e imagens
registradas pré e pós-desastre / Antônio dos Santos Ramos Neto ;
orientador Daniel Dantas. - São Cristóvão, 2024.
57 f. : il.

Dissertação (mestrado em Ciência da Computação) –
Universidade Federal de Sergipe, 2024.

1. Inteligência artificial. 2. Redes neurais (Computação). 3.
Edifícios - Deterioração. I. Dantas, Daniel orient. II. Título.

CDU 004.8



UNIVERSIDADE FEDERAL DE SERGIPE
PRÓ-REITORIA DE PÓS-GRADUAÇÃO E PESQUISA
COORDENAÇÃO DE PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Ata da Sessão Solene de Defesa da Dissertação do
Curso de Mestrado em Ciência da Computação-UFS.
Candidato: ANTONIO DOS SANTOS RAMOS NETO

Em 29 dias do mês de agosto do ano de dois mil e vinte quatro, com início às 13h, realizou-se na Sala de Seminários do PROCC da Universidade Federal de Sergipe, na Cidade Universitária Prof. José Aloísio de Campos, a Sessão Pública de Defesa de Dissertação de Mestrado do candidato **Antônio dos Santos Ramos Neto**, que desenvolveu o trabalho intitulado: ***“Classificação e segmentação de danos em edifícios após desastres naturais utilizando dados tabulares e imagens registradas pré e pós-desastre”***, sob a orientação do Prof. Dr. Daniel Oliveira Dantas. A Sessão foi presidida pelo Prof. Dr. Daniel Oliveira Dantas (PROCC/UFS), que após a apresentação da dissertação passou a palavra aos outros membros da Banca Examinadora, Prof. Dr. Wonder Alexandre Luiz Alves (UNINOVE) e, em seguida, o Prof. Dr. Leonardo Nogueira Matos (Procc/UFS). Após as discussões, a Banca Examinadora reuniu-se e considerou o mestrando (a) aprovado. Atendidas as exigências da Instrução Normativa 05/2019/PROCC, do Regimento Interno do PROCC (Resolução 67/2014/CONEPE), e da Resolução no 04/2021/CONEPE que regulamentam a Apresentação e Defesa de Dissertação, e nada mais havendo a tratar, a Banca Examinadora elaborou esta Ata que será assinada pelos seus membros e pelo mestrando.

Cidade Universitária “Prof. José Aloísio de Campos”, 29 de agosto de 2024.

Documento assinado digitalmente
 DANIEL OLIVEIRA DANTAS
Data: 29/08/2024 16:51:53-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Daniel Oliveira Dantas
(PROCC/UFS)
Presidente

Documento assinado digitalmente
 LEONARDO NOGUEIRA MATOS
Data: 01/09/2024 08:17:15-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Leonardo Nogueira Matos
(PROCC/UFS)
Examinador Interno

Documento assinado digitalmente
 WONDER ALEXANDRE LUZ ALVES
Data: 30/08/2024 05:58:39-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Wonder Alexandre Luiz Alves
(Uninove)
Examinador Externo

Documento assinado digitalmente
 ANTONIO DOS SANTOS RAMOS NETO
Data: 02/09/2024 07:56:51-0300
Verifique em <https://validar.iti.gov.br>

Antônio dos Santos Ramos Neto
Candidato

Agradecimentos

Em primeiro lugar, agradeço a Deus por me conceder forças, coragem e sabedoria ao longo de toda a minha jornada. O caminho não foi fácil, mas todos esses sonhos começaram quando meus pais e um dos meus professores do ensino médio, mesmo com muita dificuldade em passar essa mensagem a um menino inquieto, me mostraram que a educação é capaz de transformar vidas. Foi nesse momento que comecei a sonhar e acreditar que poderia alcançar voos mais altos.

Em 2017, deixei minha cidade natal, Canindé do São Francisco, com pouca bagagem, mas com uma enorme vontade de vencer. Enfrentei dificuldades e desafios, mas, com persistência, consegui concluir a faculdade em 2021. Esse foi um passo gigantesco para a minha vida pessoal e profissional, e senti que precisava buscar ainda mais conhecimento, inspirado por diversas pessoas.

No mesmo ano, com o apoio inestimável de Thaís, a mulher da minha vida, e após algumas tentativas, fui aprovado no mestrado. Agora, após dois anos e meio, concluo esta etapa que me faz lembrar ainda mais daquela criança sonhadora que sempre fui e sempre serei.

Agradeço profundamente ao meu orientador e professor, Daniel Oliveira Dantas, por toda a dedicação e apoio ao longo deste período. Sua orientação foi fundamental para o desenvolvimento deste trabalho e para minha formação profissional. Não tenho palavras para expressar minha gratidão.

Minha imensa admiração e agradecimento vão para Thaís, que esteve ao meu lado em toda essa trajetória e na vida. Você me inspira a ser melhor a cada dia.

Agradeço enormemente à minha família, que me apoiou em todas as escolhas que fiz. Aos meus pais, Rosa e Paulo, que são meus maiores tesouros, e às minhas irmãs, Paula e Kellyane, assim como aos meus sobrinhos, Pablo, Guilherme, Alany e Alice. Sem vocês, eu não estaria aqui; vocês são a minha fortaleza.

Minha eterna gratidão ao meu professor do ensino médio, Robenilson, por ter me mostrado que a educação pode transformar vidas. Suas palavras me motivaram a alcançar todos os meus objetivos.

Por fim, agradeço a todos os meus professores, que compartilharam seus conhecimentos comigo, e à Universidade Federal de Sergipe. Obrigado por tudo.

Se vi mais longe, foi por estar de pé sobre ombros de gigantes.
(Isaac Newton)

Resumo

Desastres naturais que causam grandes danos às estruturas urbanas, representando riscos significativos para a vida humana e perdas financeiras consideráveis. A rápida identificação de edifícios com níveis críticos de danos é essencial para mitigar riscos adicionais e acelerar os esforços de recuperação. Este trabalho desenvolveu modelos para classificar e segmentar danos em edifícios, utilizando dados tabulares para classificação e imagens registradas para segmentação.

Na abordagem de classificação, utilizamos dados tabulares do terremoto de Gorkha, aplicando técnicas de engenharia de características para aprimorar o desempenho dos modelos de aprendizado de máquina. O modelo XGBoost, treinado com essas técnicas, obteve um F1-score de 0,7542, demonstrando sua eficácia na identificação dos níveis de danos em edifícios. Este resultado destaca a importância da preparação e otimização dos dados para melhorar a performance dos modelos de classificação em cenários de desastre.

Para a segmentação, exploramos várias arquiteturas de redes neurais profundas, focando na Unet com *backbone* BDANet, além de ResNet18, VGG16 e ResNet50. Utilizando imagens de satélite, aplicamos operações de morfologia matemática para refinar a segmentação dos danos. Esta abordagem permitiu a segmentação dos níveis de danos, incluindo *no damage*, *minor damage* e *major damage*. O modelo Unet com BDANet, combinado com técnicas de morfologia matemática, alcançou um F1-score de 0,7999, evidenciando seu superior desempenho na segmentação de danos em imagens pós-desastre.

As abordagens de classificação com dados tabulares e segmentação com imagens registradas mostrou-se eficaz para a avaliação dos danos em edifícios após desastres naturais. Cada metodologia aprimora a precisão das avaliações em seu respectivo domínio, oferecendo ferramentas valiosas para autoridades e equipes de resgate. Esses modelos avançados auxiliam na priorização de esforços e na alocação de recursos de maneira mais eficiente em situações de desastres, contribuindo significativamente para a eficácia das respostas em emergências.

Palavras-chave: Danos em edifícios, engenharia de características, redes neurais para feature engineering, otimização de hiperparâmetros, XGBoost, operações morfológicas, imagens registradas, classificação e segmentação de nível de dano, BDANet, CutMix.

Abstract

Natural disasters that cause significant damage to urban structures, representing considerable risks to human life and substantial financial losses. The rapid identification of buildings with critical levels of damage is essential to mitigate further risks and accelerate recovery efforts. This work developed models to classify and segment building damage, using tabular data for classification and recorded images for segmentation.

In the classification approach, we used tabular data from the Gorkha earthquake, applying feature engineering techniques to enhance the performance of machine learning models. The XGBoost model, trained with these techniques, achieved an F1-score of 0.7542, demonstrating its effectiveness in identifying building damage levels. This result highlights the importance of data preparation and optimization to improve the performance of classification models in disaster scenarios.

For segmentation, we explored various deep neural network architectures, focusing on Unet with a BDANet *backbone*, in addition to ResNet18, VGG16, and ResNet50. Using satellite images, we applied mathematical morphology operations to refine the damage segmentation. This approach allowed for the segmentation of damage levels, including no damage, *minor damage*, and *major damage*. The Unet model with BDANet, combined with mathematical morphology techniques, achieved an F1-score of 0.7999, evidencing its superior performance in post-disaster damage segmentation.

The classification approaches using tabular data and segmentation using recorded images proved effective in assessing building damage after natural disasters. Each methodology enhances the accuracy of assessments in its respective domain, providing valuable tools for authorities and rescue teams. These advanced models assist in prioritizing efforts and allocating resources more efficiently in disaster situations, significantly contributing to the effectiveness of emergency responses.

Keywords: Building damage, feature engineering, neural networks for feature engineering, hyperparameter optimization, XGBoost, morphological operations, recorded images, damage classification and segmentation, BDANet, CutMix.

Lista de ilustrações

Figura 1 – Etapas da metodologia	31
Figura 2 – Engenharia de características de dados de geolocalização	33
Figura 3 – Rede neural de redução dimensional para dados de geolocalização	35
Figura 4 – Exemplo de aplicação CutMix, que está representado nos retângulos vermelhos, nas imagens.	38
Figura 5 – Arquitetura da rede BDANet resumida com foco na camada morfológica de dilatação e na saída final da rede.	40
Figura 6 – Comparação do F1-score antes e depois da otimização e com engenharia de características	43
Figura 7 – Matriz de confusão do XGBoost	44
Figura 8 – Posição no desafio	45
Figura 9 – Perda de treinamento e score de teste por épocas do melhor modelo	46
Figura 10 – F1-score por classe em cada época do melhor modelo	47
Figura 11 – Segmentação do melhor modelo em um exemplo de imagem.	49

Lista de tabelas

Tabela 1 – Modelo usados.	20
Tabela 2 – Parâmetros testados por modelo. Os selecionados estão destacados em negrito.	32
Tabela 3 – Valores únicos da coluna <code>geo_level_3</code> relacionados às colunas <code>geo_level_1</code> e <code>geo_level_2</code> iguais a 6 e 487.	34
Tabela 4 – Escala de danos observada em edifícios. Adaptado de Defense Innovation Unit, DoD (2023).	36
Tabela 5 – F1-score dos modelos propostos	42
Tabela 6 – Comparação dos resultados do modelo com otimização de hiperparâmetros.	42
Tabela 7 – Comparação de modelos entre estudos	45
Tabela 8 – Resultados dos modelos no dataset de teste (5 maiores F1-score)	45
Tabela 9 – Resultados da aplicação de camadas morfológicas no melhor modelo.	48
Tabela 10 – Comparação do F1-score entre outros estudos. A tabela está em ordem decrescente pelo F1-score.	48

Lista de códigos

Código 1 – Obtendo características a partir das probabilidades condicionais das geolocalizações	33
---	----

Lista de abreviaturas e siglas

AdamW	Adaptive Moment Estimation with Weight Decay
ADASYN	Adaptive Synthetic Sampling
BDANet	Multiscale Convolutional Neural Network with Cross-directional Attention for Building Damage Assessment
CNN	Convolutional Neural Network
CTGAN	Conditional Tabular Generative Adversarial Network
FPN	Feature Pyramid Network
GPU	Graphics Processing Unit
kNN	k-Nearest Neighbors
LGBM	Light Gradient Boosting Machine
LSTM	Long Short-Term Memory
MM	Morfologia Matemática
PCA	Principal Component Analysis
PSPNet	Pyramid Scene Parsing Network
RF	Random Forest
RGB	Red, Green, Blue
SMOTE	Synthetic Minority Over-sampling Technique
UNet	Fully Convolutional Network for Image Segmentation
XGBoost	Extreme Gradient Boosting

Sumário

1	Introdução	14
1.1	Objetivo geral	15
1.2	Objetivos específicos	16
1.3	Organização do trabalho	16
2	Fundamentação teórica	18
2.1	Classificação	18
2.1.1	Aprendizado supervisionado	18
2.1.2	Técnicas de <i>ensemble: bootstrap e boosting</i>	19
2.1.3	Modelos de classificação	19
2.1.4	Otimização de modelos de classificação	21
2.1.5	Engenharia de características	21
2.2	Segmentação	22
2.2.1	Segmentação de imagens	22
2.2.2	Preparação dos dados para segmentação de imagens	23
2.2.3	Abordagens para imagens registradas	23
2.2.4	Arquiteturas de redes neurais na segmentação de imagens	24
2.2.5	Morfologia matemática	24
2.2.6	Visualização e avaliação	25
3	Trabalhos relacionados	26
3.1	Classificação	26
3.2	Segmentação	27
4	Metodologia	30
4.1	Classificação de nível de dano em dados tabulares	30
4.2	Segmentação de edifícios em imagens aéreas de desastres	35
5	Resultados	41
5.1	Classificação de nível de dano em dados tabulares	41
5.1.1	Treinamento dos modelos	41
5.1.2	Otimização de hiperparâmetros	42
5.1.3	Engenharia de características	42
5.1.4	Avaliação do modelo	43
5.1.5	Posição no desafio e comparação com outros estudos	44
5.2	Segmentação de edifícios em imagens aéreas de desastres	45

5.2.1	Experimentos	45
5.2.2	Comparação com outros estudos	47
6	Conclusões	50
	Referências	53

1

Introdução

Terremotos são desastres naturais que causam danos em grande escala. [Girgin \(2011\)](#) descreve o terremoto de Kocaeli em 1999 na Turquia. Este terremoto foi um dos maiores desastres naturais da história, causando quase 18.000 mortes e mais de US\$15 bilhões em danos materiais.

[Krausmann e Cruz \(2013\)](#) descrevem outro grande terremoto, o de Tohoku em 2011 no Japão. O terremoto causou grande devastação, deixando mais de 16.000 vítimas e danos econômicos próximos a US\$200 bilhões. Esses desastres destacam as destruições que os terremotos podem causar.

Esses eventos podem causar uma série de consequências negativas para as sociedades afetadas. Ao discutir o risco de vida em um terremoto, uma das primeiras coisas a se considerar são as estruturas dos edifícios.

Segundo [Peña \(2012\)](#), em regiões onde a atividade sísmica é alta e os edifícios não são projetados para resistir a ações sísmicas, a fragilidade e vulnerabilidade das estruturas se tornam evidentes. Mesmo em áreas de baixa atividade sísmica, terremotos podem causar danos consideráveis às estruturas, pois a maioria das edificações não construídas de acordo com normas sísmicas possui baixa ductilidade.

Devido à conveniência e ao baixo preço dos materiais de construção, muitas das estruturas construídas no Vale de Kathmandu, Nepal, durante o terremoto de Gorkha em 2015 eram feitas de alvenaria ou estruturas de concreto não reforçadas. Infelizmente, essas estruturas foram destruídas em graus variados durante o terremoto, algumas impedindo seu uso futuro ([LIU; FANG; ZHAO, 2021](#)).

De acordo com [Miranda \(2010\)](#), esses danos estão relacionados a estribos insuficientes e baixa capacidade de resistência das paredes de alvenaria. Com todos esses danos que as estruturas dos edifícios podem sofrer, avaliar essas estruturas após desastres em grandes áreas urbanas é uma tarefa complexa e pode ser perigosa para os avaliadores.

Um aspecto crítico nos cenários pós-terremoto é a avaliação dos danos estruturais aos edifícios. Para abordar a complexidade da avaliação de tais danos e acelerar os esforços de recuperação, várias soluções foram propostas. Uma solução possível envolve a utilização de modelos de classificação ou segmentação para identificar estruturas com níveis críticos de danos.

Esses modelos aproveitam tecnologias avançadas, como imagens térmicas infravermelhas, aprendizado de máquina e dados de sensoriamento remoto para melhorar as métricas de avaliação de danos em edifícios pós-terremoto (ZHANG et al., 2020; LIU, 2024; WANG et al., 2022; MING, 2024).

A classificação e a segmentação podem ser soluções para detecção do nível de danos após desastres naturais. Dessa forma, este trabalho concentra-se na aplicação de aprendizado de máquina para a classificação de danos e no uso de redes neurais para segmentação de imagens de edifícios após desastres naturais. Utilizamos dados tabulares, que oferecem informações estruturadas sobre as construções e o impacto dos desastres, e imagens de satélite pré e pós-evento para identificar níveis de dano.

A avaliação manual de danos em grandes áreas urbanas é um processo demorado, arriscado e complexo. Neste contexto, o problema central abordado é o desenvolvimento de modelos capazes de automatizar a classificação e segmentação de danos em edifícios, proporcionando uma resposta mais rápida e segura em cenários de desastres, com um alto F1-score.

A relevância desta pesquisa está em sua contribuição potencial para melhorar as métricas de avaliação de danos pós-desastres. Embora estudos anteriores, como os de Zhang et al. (2020) e Gupta et al. (2019), tenham destacado a importância de modelos automáticos para essa tarefa, ainda há uma lacuna no uso de técnicas mais avançadas, como feature engineering para classificação e operações morfológicas em redes neurais para segmentação. Este estudo visa preencher essa lacuna, oferecendo modelos mais abrangentes e escaláveis para a avaliação de danos.

Nossa abordagem de classificação de dados tabulares, com a adição de técnicas de engenharia de características, utilizando dados tabulares do terremoto de Gorkha (DRIVENDATA, 2022), obteve o F1-score de 0,7542.

Por outro lado, utilizando a abordagem de segmentação com a adição de operações morfológicas, aplicadas ao dataset de imagens registradas pré e pós-desastre xBD (GUPTA et al., 2019), obtivemos um modelo com F1-score de 0,7999.

1.1 Objetivo geral

Este estudo tem como objetivo desenvolver, implementar e avaliar modelos de classificação e segmentação para a avaliação de danos em edifícios após desastres naturais, utilizando dados tabulares e imagens registradas pré e pós-desastre.

1.2 Objetivos específicos

1. Desenvolver um modelo de classificação de danos em edifícios após terremotos utilizando técnicas de aprendizado de máquina em dados tabulares.
2. Desenvolver um modelo de segmentação de danos em edifícios após desastres naturais utilizando técnicas de *deep learning* em imagens de satélite.
3. Validar os modelos utilizando datasets de desastres naturais reais e avaliar seu desempenho utilizando o F1-score.
4. Propor uso de técnicas como engenharia de características, para classificação, e morfologia matemática, para segmentação, para melhorar o F1-score dos modelos.
5. Obter F1-score superior aos algoritmos de classificação e segmentação da literatura.

1.3 Organização do trabalho

Este trabalho está organizado da seguinte forma:

No Capítulo 1, apresentamos a introdução e a contextualização do problema, destacando a importância da avaliação de danos em edifícios após desastres naturais. Este capítulo também inclui os objetivos geral e específicos do trabalho, além de uma visão geral sobre a estrutura da dissertação.

No Capítulo 2, discutimos a fundamentação teórica, abordando as principais técnicas e metodologias utilizadas na classificação e segmentação de danos em edifícios. Este capítulo é dividido em duas seções principais:

- Classificação: inclui modelos de classificação para dados tabulares, otimização de hiperparâmetros dos modelos de classificação e engenharia de características.
- Segmentação: abrange a preparação dos dados para segmentação de imagens, abordagens para imagens registradas, arquiteturas de redes neurais na segmentação de imagens, morfologia matemática e técnicas de visualização e avaliação.

No Capítulo 3, discutimos os trabalhos relacionados, apresentando estudos e pesquisas relevantes que abordam a classificação e segmentação de danos em edifícios, tecnologias e metodologias aplicadas, e os avanços mais recentes na área.

No Capítulo 4, detalhamos a metodologia adotada para o desenvolvimento dos modelos de classificação e segmentação. Este capítulo descreve o processo de coleta e preparação dos dados, a escolha das arquiteturas dos modelos, os procedimentos de treinamento, otimização de hiperparâmetros, engenharia de características e os critérios de avaliação dos modelos.

- Classificação de nível de dano em dados tabulares: detalha o processo de coleta de dados, a preparação dos dados tabulares, as técnicas de engenharia de características aplicadas e a avaliação do desempenho do modelo.
- Segmentação de edifícios em imagens aéreas de desastres: descreve o uso de imagens pré e pós-desastre para segmentação, as arquiteturas de redes neurais empregadas e o uso de operações morfológicas para aprimorar a segmentação.

No Capítulo 5, apresentamos os resultados obtidos com a aplicação dos modelos em datasets reais. Este capítulo é dividido em duas seções principais:

- Classificação de nível de dano em dados tabulares: inclui o treinamento dos modelos, otimização de hiperparâmetros, engenharia de características, avaliação dos modelos e comparação com outros estudos.
- Segmentação de edifícios em imagens aéreas de desastres: descreve os experimentos realizados, os resultados obtidos e a comparação com outros estudos.

No Capítulo 6, apresentamos as conclusões do trabalho, destacando as contribuições, limitações e sugestões para trabalhos futuros. Este capítulo resume os principais achados e discute as implicações práticas e teóricas das técnicas de classificação e segmentação aplicadas à avaliação de danos em edifícios após desastres naturais.

2

Fundamentação teórica

Este capítulo apresenta os principais conceitos teóricos e técnicas fundamentais utilizadas no desenvolvimento deste trabalho.

Inicialmente, será discutido o aprendizado supervisionado, as técnicas de *ensemble*, os modelos de classificação empregados, bem como suas características e métodos de otimização. Em seguida, serão abordadas as técnicas de engenharia de características, que desempenham um papel crucial na melhoria do desempenho dos modelos de aprendizado de máquina.

Por fim, será discutido o processo de segmentação de imagens, explorando arquiteturas de redes neurais, abordagens para segmentação e visualização dos resultados. A fundamentação teórica estabelecida aqui serve como base para a análise e desenvolvimento dos métodos propostos neste estudo.

2.1 Classificação

2.1.1 Aprendizado supervisionado

Aprendizado supervisionado é caracterizado pelo uso de dados rotulados para treinar modelos que realizam tarefas como classificação, previsão ou segmentação. Essa abordagem é amplamente aplicada em áreas como reconhecimento de padrões em imagens, processamento de linguagem natural e diagnósticos médicos.

O aprendizado supervisionado depende da relação entre entradas e saídas rotuladas, o que permite ao modelo aprender associações específicas entre os dados de entrada e seus respectivos rótulos. Durante o treinamento, o modelo ajusta seus parâmetros com base na minimização de uma função de erro, de modo a reduzir a diferença entre as previsões feitas e os resultados reais. Após o treinamento, o modelo é capaz de realizar previsões para novos dados, aplicando as regras que aprendeu a partir do conjunto de treinamento.

Embora o objetivo seja que o modelo generalize para novos dados, isso nem sempre acontece, e há o risco de *overfitting*, quando o modelo se ajusta excessivamente aos dados de treinamento e perde a capacidade de fazer boas previsões em dados não vistos durante o treinamento.

A classificação é uma das tarefas mais comuns no aprendizado supervisionado, onde o objetivo é atribuir rótulos a novos dados com base em um conjunto de treinamento previamente rotulado.

Técnicas como a classificação por *K-means*, discutida por Yang (2010), embora geralmente associadas ao aprendizado não supervisionado, também podem ser aplicadas no contexto supervisionado, utilizando funções de distância adequadas para agrupar dados rotulados.

Embora o aprendizado supervisionado ofereça uma base sólida para a resolução de problemas de classificação, técnicas mais avançadas, como o *ensemble learning*, como *bagging* e *boosting*, têm mostrado ser eficazes para melhorar ainda mais o desempenho preditivo.

2.1.2 Técnicas de *ensemble*: *bootstrap* e *boosting*

Ensemble learning é uma abordagem poderosa que combina previsões de múltiplos modelos para formar um único modelo mais robusto e preciso. Essa estratégia é particularmente útil em cenários onde os modelos individuais sofrem de alta variância ou viés.

Bagging, ou *bootstrap aggregating*, é uma técnica que treina vários modelos em diferentes subconjuntos dos datasets de treinamento. Fazer isso reduz a variabilidade e melhora a estabilidade das previsões. Um dos exemplos mais conhecidos dessa técnica é o *random forest*, que combina múltiplas árvores de decisão para gerar previsões mais precisas (LIAW; WIENER, 2002).

Boosting, por outro lado, foca em transformar classificadores fracos em fortes. Esse método envolve o treinamento sequencial de modelos, onde cada novo modelo corrige os erros cometidos pelos anteriores. O algoritmo AdaBoost é um exemplo notável de *boosting*, atribuindo pesos maiores a instâncias mal classificadas e ajustando esses pesos em iterações subsequentes para melhorar a precisão (LI; WANG; SUNG, 2008).

Essas técnicas têm se mostrado altamente eficazes em uma ampla gama de aplicações, incluindo classificação de dados tabulares. Ao combinar múltiplos modelos, *bagging* e *boosting* mitigam problemas de variância e viés, resultando em uma melhoria significativa no desempenho preditivo. A seguir, apresentamos os modelos de classificação que foram utilizados neste estudo.

2.1.3 Modelos de classificação

Existem vários modelos de aprendizado de máquina para resolver problemas de classificação. Neste estudo, usaremos os modelos mostrados na Tabela 1.

O *random forest* é um modelo de aprendizado supervisionado que é amplamente

Tabela 1 – Modelo usados.

Nº	Modelos
1	Extra trees classifier
2	Gradient boosting classifier
3	Light gradient boosting machine
4	Random forest classifier
5	XGBoost classifier
6	CatBoost classifier
7	AdaBoost classifier
8	AsymBoost classifier

reconhecido por sua precisão e eficiência no treinamento (LIAW; WIENER, 2002). Ele é robusto ao ruído nos dados e apresenta boa capacidade de generalização, combinando várias árvores de decisão para gerar um modelo mais estável.

O *extra trees classifier*, é semelhante ao *random forest* por usar a técnica de *boosting* (ABHISHEK, 2020), mas difere na escolha das divisões das árvores, que são feitas de forma totalmente aleatória. Essa abordagem pode reduzir a variabilidade do modelo (GEURTS; ERNST; WEHENKEL, 2006).

De acordo com Li (2016), como nos outros modelos de *boosting*, o *gradient boosting* combina diversos modelos com alto erro de classificação em um modelo com menor erro de classificação de maneira iterativa. O modelo geralmente tem uma árvore de decisão de tamanho fixo como modelo principal.

O *light gradient boosting machine* (LGBM) é um modelo de *gradient boosting* que usa modelos baseados em árvore. Ke et al. (2017) apontam que este modelo pode superar outros modelos de classificação em acurácia e precisão, com consumo de memória consideravelmente menor em experimentos realizados em dados públicos. Além disso, é possível reduzir o tempo de treinamento usando processamento distribuído (KE et al., 2017).

O modelo *random forest* é um algoritmo de aprendizado supervisionado baseado em árvores de decisão (BREIMAN, 2001). Uma das principais vantagens do modelo de *random forest* é sua capacidade de lidar com dados de alta dimensionalidade e grande escala (BIAU; CADRE; MEVEL, 2008).

Da maneira semelhante à forma que o LGBM funciona, Chen e Guestrin (2016) descrevem o XGBoost como um modelo distribuído de *gradient boosting* criado para ser altamente eficiente, flexível e portátil. Ele cria múltiplas *decision trees* em paralelo para aumentar a eficiência do treinamento (KE et al., 2017).

Prokhorenkova et al. (2018) descrevem CatBoost como uma implementação de *gradient boosting* que usa árvores de decisão binárias como preditores de *baseline*. Samat et al. (2020) mostram que CatBoost pode evitar *overfitting* por meio de estimativas de gradiente e demonstrou

precisões superiores às implementações com XGBoost e LGBM.

O modelo AdaBoost cria um conjunto de classificadores com baixa precisão que possuem pesos e são treinados em diferentes amostras dos dados de treinamento com ajuste adaptativo dos pesos após cada iteração de treinamento (LI; WANG; SUNG, 2008).

O AsymBoost (VIOLA; JONES, 2001) é uma variante do AdaBoost com um método *cost-sensitive boost*. Ele usa *asymmetric misclassification costs* para fazer a atualização do erro nas épocas de treinamento. É um classificador de alto desempenho quando o banco de dados possui classes desbalanceadas

2.1.4 Otimização de modelos de classificação

Ataei e Osanloo (2004) e Alvarenga-Júnior (2018) apresentam algumas maneiras de realizar a otimização de hiperparâmetros com foco na melhoria da saída do modelo. Um dos métodos apresentados é o *grid search*. O *grid search* realiza uma busca exaustiva em um espaço de hiperparâmetros de tamanho pequeno. A técnica fornece uma maneira simples de encontrar os melhores hiperparâmetros e é fácil de implementar e paralelizar.

Os autores afirmam ainda que neste método, cada hiperparâmetro é definido em torno de um intervalo de pesquisa específico que se acredita ter um potencial local para varredura. Para cada hiperparâmetro, uma resolução de grade determinará o número de candidatos a serem considerados.

2.1.5 Engenharia de características

De acordo com Khurana et al. (2016), a engenharia de características consiste em extrair características de dados brutos e transformá-los para melhorar a precisão dos modelos de aprendizado de máquina. O autor reforça que a maior parte do tempo na construção de modelos de classificação é gasto em engenharia de características e limpeza de dados.

De acordo com Khalid, Khalil e Nasreen (2014) e Deng (2024), a engenharia de características é um processo crítico para melhorar o desempenho de modelos de aprendizado de máquina, particularmente em tarefas de classificação. Envolve a conversão de dados brutos em características significativas que podem melhorar a precisão do modelo.

As técnicas de engenharia de características para dados tabulares são essenciais para melhorar o desempenho dos modelos de aprendizado de máquina. A lista abaixo mostra alguns exemplos de métodos de engenharia de características que podem melhorar o desempenho do modelo.

1. Criação de novas características a partir de combinações lineares de características existentes, como multiplicar duas características ou dividir uma característica por outra;

2. Transformação de características, por exemplo, normalização de características numéricas ou codificação de características categóricas;
3. Criação de características agregadas, por exemplo a média ou soma de características relacionadas;
4. Redução da dimensionalidade, como o uso de técnicas de *principal component analysis* (PCA) para selecionar características relevantes.

Além disso, técnicas como *synthetic minority oversampling technique* (SMOTE) e *conditional tabular generative adversarial network* (CTGAN) são comumente usadas para aumentar dados tabulares e melhorar o desempenho do modelo (SATHIANARAYANAN et al., 2022).

Também métodos avançados, como redes *long short-term memory* (LSTM), se destacaram na engenharia de características para tarefas de classificação. As redes LSTM são excelentes na captura de dependências de longo prazo em dados sequenciais, tornando-as adequadas para tarefas que envolvem análise de séries temporais e extração de características (XU et al., 2023).

As LSTMs podem, por exemplo, capturar a evolução temporal de certos parâmetros, como variações nas medições de estabilidade estrutural ou relação entre as regiões que tiveram maior impacto, permitindo que o modelo identifique padrões que indicam deterioração progressiva. Dessa forma, ela consegue associar essas mudanças temporais aos níveis de dano, auxiliando na classificação de gravidade das áreas afetadas.

Segundo Casari e Zheng (2018), a engenharia de características é importante, mas pouco discutida em estudos de aprendizado de máquina. Em muitos casos, a técnica é específica para cada dataset e difícil de generalizar, mas é crucial para aprender modelos preditivos de alto desempenho.

2.2 Segmentação

2.2.1 Segmentação de imagens

A segmentação de imagens é um processo fundamental em visão computacional, cujo objetivo é particionar uma imagem em várias regiões ou segmentos, facilitando a análise e interpretação da imagem. Essa técnica é amplamente utilizada em diversas áreas, como detecção de objetos, reconhecimento de imagens, diagnóstico médico e veículos autônomos (DING; WU, 2019; ADHVARYU, 2012).

Os principais tipos de segmentação incluem a segmentação semântica, que rotula cada pixel com uma classe, como “carro” ou “pedestre”, e a segmentação de instâncias, que distingue diferentes instâncias de objetos da mesma classe. Há também a segmentação panóptica, que

combina essas duas abordagens para uma compreensão mais completa da cena (DESIANI et al., 2023).

Com o avanço das redes neurais, métodos de segmentação baseados em aprendizado profundo, como U-Net e Mask R-CNN, têm se destacado por sua precisão em tarefas de segmentação complexas. Essas abordagens têm sido aplicadas com sucesso em diversas áreas, incluindo diagnóstico médico e agricultura de precisão, oferecendo resultados cada vez mais acurados e eficientes (ZHAO et al., 2021).

Antes de aplicar essas técnicas de segmentação, porém, é fundamental garantir que os dados estejam devidamente preparados para atender aos requisitos de cada método. A seguir, discutiremos as etapas de preparação dos dados para segmentação de imagens.

2.2.2 Preparação dos dados para segmentação de imagens

A preparação de dados é uma etapa crucial na segmentação de imagens, que envolve o ajuste das condições das imagens para atender aos requisitos de segmentação.

1. Recorte: seleciona uma parte da imagem, faz um recorte e a utiliza como entrada para o modelo de segmentação. Esta técnica ajuda a aumentar a diversidade dos dados de treinamento e evitar *overfitting* (HE et al., 2022)
2. Transformações para data augmentation: aplicação de diversas transformações à imagem de entrada para criar novos exemplos de treinamento. As transformações mais comuns incluem rotação, escala, espelhamento, cisalhamento, translação e normalização de contraste. Essas transformações ajudam a aumentar o tamanho do dataset de treinamento e a melhorar a robustez do modelo (HE et al., 2022).
3. Redimensionamento: alterar o tamanho da imagem de entrada para corresponder ao tamanho de entrada do modelo de segmentação. Esta técnica é útil quando o tamanho da imagem de entrada é diferente do tamanho de entrada do modelo (ALAMIN; JEBERSON; BAJAJ, 2016).
4. CutMix: É uma estratégia de data augmentation proposta por Yun et al. (2019) que envolve cortar e colar *patches* entre imagens de treinamento, onde os rótulos do *groundtruth* também são misturados proporcionalmente à área dos *patches*.

2.2.3 Abordagens para imagens registradas

Existem diversas abordagens possíveis para segmentar imagens registradas, como imagens de desastre onde temos uma imagem pré e uma pós-desastre.

Uma abordagem possível para segmentar imagens registradas, como imagens pré e pós-desastre, é concatenar as duas imagens de entrada na dimensão dos canais, ou seja, empilhá-las

gerando uma imagem com seis canais. Esta abordagem foi usada por [Li, Jafari e Rother \(2019\)](#) e [Muhadi et al. \(2020\)](#). Nesta abordagem, a imagem concatenada é alimentada em um modelo de segmentação para obter a saída segmentada.

Outra abordagem possível é alimentar as imagens separadas antes e depois do desastre em uma rede siamesa. Esta abordagem foi usada por [Da, Ji e Zhou \(2022\)](#) e [Chowdhury e Rahnemoonfar \(2021\)](#). Nessa abordagem, a rede siamesa aprende uma métrica de similaridade entre as imagens pré e pós-desastre, que pode ser usada para identificar mudanças na cena.

Uma terceira abordagem possível é calcular a diferença entre as imagens pré e pós-desastre. Esta abordagem foi usada por [Rudner et al. \(2019\)](#) e [Yang et al. \(2020\)](#). Nesta abordagem, a imagem diferencial é alimentada em uma rede de segmentação para obter a saída segmentada.

2.2.4 Arquiteturas de redes neurais na segmentação de imagens

UNet é uma arquitetura de rede *encoder-decoder* totalmente convolucional com conexões de salto entre módulos *encoder-decoder* ([RONNEBERGER; FISCHER; BROX, 2015](#)).

Feature pyramid network (FPN), é uma arquitetura de rede neural que foi originalmente proposta para detecção de objetos, mas também foi aplicada a tarefas de segmentação ([MINAEI et al., 2021](#)).

LinkNet é uma arquitetura de rede neural que foi proposta para tarefas de segmentação semântica ([BASU, 2023](#)). É uma rede totalmente convolucional que utiliza uma combinação de módulos *encoder-decoder* para extrair características da imagem de entrada e gerar o mapa de segmentação.

Pyramid scene parsing network (PSPNet) é uma arquitetura de rede neural que foi proposta para tarefas de segmentação semântica e é baseado em um *pyramid pooling module* (PPM) que captura informações contextuais multiescala ([LI, 2023](#)).

BDAnet é uma arquitetura de rede neural que foi proposta para tarefas de segmentação de imagens ([MOCHALOV; MOCHALOVA, 2019](#)). É uma rede totalmente convolucional que usa uma combinação de convoluções dilatadas e *atrous spatial pyramid pooling* (ASPP) para capturar informações contextuais em multiescala. O ASPP aplica convoluções com diferentes taxas de amostragem, permitindo que o modelo capture tanto detalhes locais quanto informações de contexto mais amplas, sem aumentar a quantidade de parâmetros ([CHEN, 2017](#)).

2.2.5 Morfologia matemática

A morfologia matemática é um campo do processamento de imagens que trata da análise e manipulação de estruturas geométricas em imagens. As duas operações primitivas da morfologia matemática são dilatação e erosão. Além disso, existem as operações de abertura e fechamento, que são derivadas ([CHEN et al., 2021](#)).

A dilatação é uma operação que expande os limites de uma região que possui valores de pixel maiores que 0 em uma imagem, enquanto a erosão os reduz. A abertura é uma erosão seguida de uma dilatação, que pode ser usada para remover pequenos objetos de uma imagem. O fechamento é uma dilatação seguida de uma erosão, que pode ser usada para preencher pequenos buracos em uma imagem (BANON; BARRERA, 1994).

2.2.6 Visualização e avaliação

Existem diversas técnicas de visualização e métricas de erro que podem ser usadas para avaliar os resultados dos modelos de segmentação de imagens. Para visualização utilizamos segmentação semântica visual (BENKHOUI; EL-KORCHI; REINHOLD, 2021) e para avaliação dos resultados utilizamos o F1-score (LAINE et al., 2021). As métricas de erro são calculadas usando os valores de *true positives* (TP), *false positives* (FP), *true negatives* (TN) e *false negatives* (FN).

Precisão e sensibilidade são métricas fundamentais para avaliar o desempenho de modelos de segmentação e classificação. A precisão é definida como a razão entre as previsões verdadeiramente positivas e o total de previsões positivas feitas pelo modelo:

$$\text{Precisão} = \frac{TP}{TP + FP}$$

A sensibilidade mede a capacidade do modelo de identificar todas as instâncias positivas, sendo a razão entre TP e o total de instâncias realmente positivas:

$$\text{Sensibilidade} = \frac{TP}{TP + FN}$$

Com base nessas duas métricas, o F1-score é utilizado como uma métrica combinada que representa a média harmônica entre precisão e sensibilidade, e é dado pela fórmula:

$$\text{F1-score} = 2 \times \frac{\text{Precisão} \times \text{Sensibilidade}}{\text{Precisão} + \text{Sensibilidade}}$$

Para calcular o F1-score, as máscaras de segmentação geradas pelos modelos foram comparadas com as máscaras *groundtruth*.

Além disso, a segmentação semântica é uma técnica de processamento de imagens que envolve a classificação de cada pixel em uma imagem em uma categoria semântica específica. Esta técnica visa atribuir um rótulo a cada pixel, permitindo a identificação e distinção de diferentes objetos ou regiões dentro de uma imagem.

3

Trabalhos relacionados

Este capítulo apresenta os trabalhos relacionados às duas áreas centrais deste trabalho: a classificação de danos em edifícios após desastres naturais e a segmentação de imagens para a detecção de áreas afetadas.

Inicialmente, discutimos abordagens de classificação aplicadas à avaliação de danos estruturais em dados tabulares, com ênfase nos métodos baseados em aprendizado de máquina e nas técnicas de engenharia de características.

Em seguida, abordamos as metodologias de segmentação de imagens utilizadas para a análise de imagens pré e pós-desastre, explorando arquiteturas de redes neurais voltadas para a segmentação de danos em edifícios.

Embora tratem de métodos diferentes e utilizem diferentes tipos de dados, ambos os problemas visam o mesmo objetivo de avaliar os danos causados aos edifícios. Dessa forma, cada uma dessas abordagens oferece vantagens específicas e pode ser aplicada em diferentes contextos, dependendo dos dados disponíveis e das necessidades da análise.

3.1 Classificação

Os trabalhos de [Kumaraswamy, Reddy e Kolla \(2020\)](#), [Guyette, Lewis e Stone \(2021\)](#) e [Finotti et al. \(2018\)](#) apresentam abordagens distintas, mas complementares, para a classificação de níveis de dano em estruturas após desastres naturais. Todos os estudos utilizaram dados fornecidos pela plataforma [DrivenData \(2022\)](#), destacando a relevância e a consistência dos dados dessa fonte para a tarefa de classificação de danos estruturais.

[Kumaraswamy, Reddy e Kolla \(2020\)](#) utilizaram uma combinação de quatro modelos de classificação, explorando a força de um classificador *ensemble* para integrar as previsões dos modelos individuais e melhorar o desempenho geral, alcançando um F1-score de 0,7459.

A abordagem de *ensemble* sugere que a combinação de diferentes modelos pode capturar de maneira mais robusta os padrões nos dados, mitigando a limitação de um único modelo em situações complexas como a classificação de danos estruturais.

Por outro lado, [Guyette, Lewis e Stone \(2021\)](#) focaram em uma combinação de métodos tradicionais de aprendizado de máquina, como o *k-nearest neighbors* (kNN) e o *random forest*, que também são amplamente utilizados para tarefas de classificação. O resultado foi um F1-score de 0,7260, ligeiramente inferior ao de [Kumaraswamy, Reddy e Kolla \(2020\)](#), mas ainda assim destacando a viabilidade dessas técnicas. Este estudo, comparado ao anterior, mostra que a simplicidade de métodos como kNN pode ser compensada pela robustez de abordagens baseadas em árvores, como o *random forest*, na presença de dados estruturados.

[Finotti et al. \(2018\)](#) exploraram a eficácia de modelos como redes neurais, *support vector machines* e *decision trees*, destacando a importância de ajustes em hiperparâmetros e técnicas complementares, como a engenharia de características, para melhorar a precisão de predições. Diferente dos trabalhos anteriores, este enfatiza a necessidade de otimizações e ajustes para obter um desempenho superior, sugerindo que o sucesso na classificação de danos depende não apenas da escolha do modelo, mas também da personalização do processo de treinamento.

Em todos esses trabalhos, observa-se a importância crescente da engenharia de características para potencializar a capacidade preditiva dos modelos. Técnicas como *gradient boosting* ([NABI; SAEED; HARRON, 2020](#)) e modelos *ensemble* são apontadas como soluções efetivas para melhorar o desempenho dos classificadores.

Além disso, abordagens mais recentes, como o uso de redes neurais profundas, como CNN e LSTM, para extração automática de características, são promissoras na combinação com classificadores *ensemble*, conforme discutido em [Shi, Suganthan e Katuwal \(2022\)](#), [Nanni, Lumini e Brahnam \(2022\)](#). Esses trabalhos sublinham que a sofisticação na extração de características e o uso de modelos combinados, como *deep random vector functional link* (dRVFL), são estratégias eficazes para aprimorar a classificação de dados tabulares e complexos.

3.2 Segmentação

A segmentação de imagens é uma técnica fundamental no campo do processamento de imagens e visão computacional que envolve dividir uma imagem em regiões significativas e não sobrepostas. O processo de segmentação de imagens é essencial para a compreensão de cenas naturais, pois permite a identificação de objetos e seus limites dentro de uma imagem ([YU, 2023](#)).

A importância da segmentação de imagens reside em sua capacidade de extrair informações relevantes de uma imagem, que podem ser usadas para várias aplicações, como reconhecimento de objetos, compressão de imagens e aprimoramento de imagens ([LI et al., 2020](#)).

Uma das aplicações da segmentação de imagens é na avaliação de danos a edifícios e na identificação de mudanças na paisagem antes e depois de desastres naturais. Isso permite a localização e identificação de edifícios e outras estruturas, que podem então ser analisados para danos.

Uma rede neural pode ser usada para localizar as áreas mais afetadas e segmentar danos em edifícios (DA; JI; ZHOU, 2022). Da mesma forma, Wang e Li (2022) utilizaram segmentação automática de imagens para simular e avaliar riscos de inundações e deslizamentos em áreas de *hotspots*.

No entanto, para realizar essa tarefa de segmentação de imagens após desastres naturais, a disponibilidade de datasets que tenham anotações de qualidade pré e pós-desastre é crucial. O dataset xBD é amplamente utilizado para avaliar a segmentação de danos a edifícios a partir de imagens de satélite (GUPTA et al., 2019).

Este dataset inclui imagens de satélite de alta resolução de 19 eventos de desastres naturais, como furacões, terremotos, inundações, incêndios florestais, erupções vulcânicas e tsunamis, cobrindo mais de 45.000 quilômetros quadrados de vários desastres naturais que ocorreram ao redor do mundo.

Várias arquiteturas de redes neurais têm sido usadas na segmentação de imagens de satélite após desastres naturais. UNet é uma arquitetura popular que usa um *encoder-decoder* para capturar o contexto e expandir gradualmente as características, permitindo uma segmentação precisa (MA et al., 2020).

Multiscale convolutional neural network with cross-directional attention (BDAnet) é uma arquitetura que usa atenção multidirecional e fusão de características multiescala para melhorar a avaliação de danos a edifícios a partir de imagens de satélite (SHEN et al., 2021).

Assim como a BDAnet, a *siamese hierarchical transformer framework* é outra arquitetura que usa uma rede neural para segmentar imagens (DA; JI; ZHOU, 2022). A aplicação bem-sucedida das arquiteturas dos modelos para segmentação de imagens pré e pós-desastre depende do uso de imagens registradas. Isso apresenta desafios substanciais, como a identificação de características estáveis em imagens multitemporais (LYU; JIANG, 2017) e a segmentação e agrupamento de imagens planas, ambos problemas críticos no campo da visão computacional (BURDESCU et al., 2014). Também é crucial manter a relação espacial entre essas imagens para uma segmentação precisa (DA; JI; ZHOU, 2022).

Os estudos de Shen et al. (2021) e Weber e Kan (2020) propõem novas arquiteturas de redes neurais convolucionais para avaliação de danos a edifícios a partir de imagens de satélite após desastres naturais. Shen et al. (2021) propõem a BDAnet, uma CNN de duas etapas que utiliza mecanismos de atenção multidimensional e multidirecional para melhorar o F1-score da segmentação de imagens. Weber e Kan (2020) propõem uma arquitetura CNN Inception V3 aprimorada que combina imagens de sensoriamento remoto e dados vetoriais de blocos para

avaliar o grau de danos de grupos de edifícios.

Ambos os estudos visaram melhorar o F1-score dos modelos de segmentação de danos a edifícios a partir de imagens de satélite usando técnicas de aprendizado profundo. Apesar das contribuições dos estudos mencionados, ainda há uma lacuna em relação à exploração da morfologia matemática e diferentes abordagens para imagens registradas na avaliação de danos a edifícios em imagens de satélite pós-desastre.

A morfologia matemática, com suas operações de dilatação, erosão, abertura e fechamento, permite a melhoria da máscara prevista pelo modelo, removendo ruídos e imperfeições gerados pelo modelo, facilitando assim a identificação de edifícios danificados.

4

Metodologia

Este capítulo apresenta as metodologias adotadas para resolver os dois principais problemas deste trabalho, a classificação dos níveis de danos em edifícios utilizando dados tabulares e a segmentação de imagens aéreas para identificar visualmente os danos.

Para a classificação, utilizamos um dataset tabular contendo informações estruturais e geográficas sobre os edifícios afetados por desastres, como o tipo de material, número de andares e localização geográfica. Esses dados foram processados e transformados por meio de técnicas de engenharia de características, como a substituição de categorias por probabilidades condicionais e a extração de características relacionadas à geolocalização. Diversos modelos de aprendizado de máquina foram treinados e avaliados, buscando maximizar o desempenho por meio da otimização de hiperparâmetros e ajustes no modelo.

Na segmentação, trabalhamos com imagens de satélite de alta resolução, que capturam edifícios antes e após o desastre. Aplicamos redes neurais convolucionais, como UNet e FPN, utilizando abordagens distintas de entrada, como imagens empilhadas, diferenças de imagens e redes siamesas. Essas redes foram treinadas para segmentar os danos visíveis nos edifícios, utilizando técnicas de transformações e *CutMix* para lidar com o desbalanceamento entre as diferentes classes. O processo de segmentação é crucial para identificar visualmente as áreas afetadas e fornecer uma análise espacial detalhada dos danos.

4.1 Classificação de nível de dano em dados tabulares

O dataset utilizado contém dados sobre a estrutura dos edifícios. O dataset possui 260.601 amostras, cada uma representando um edifício na região de Gorha atingida pelo terremoto em 2015.

São 39 características, uma com um identificador único do edifício e outras 38 com dados das estruturas do edifício, como tipo de estrutura, número de famílias que moravam no edifício,

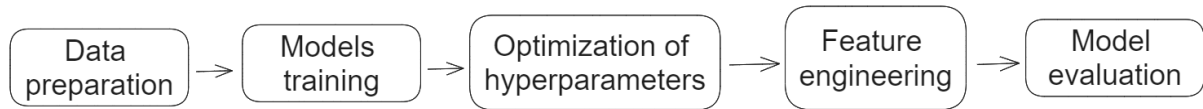


Figura 1 – Etapas da metodologia

se era utilizado para fins industriais, e assim por diante (DRIVENDATA, 2022).

Além dos atributos estruturais, o dataset inclui informação geográfica como o distrito e aldeia onde cada edifício está localizado. Esses dados geográficos ajudam a compreender as variações regionais nos níveis de danos.

O dataset também apresenta detalhes do material de construção, incluindo o tipo de fundação, telhado e materiais de piso. Além disso, indicadores socioeconômicos, como o número de famílias que residem em cada edifício e se este serve fins industriais, fornecem informações sobre a utilização do edifício.

O dataset possui três classes: a *classe 0* representa dano baixo, a *classe 1* representa dano médio e a *classe 2* representa destruição completa. Este dataset é considerado um dos maiores já coletados após uma catástrofe natural, contendo informações valiosas sobre os efeitos do terremoto, a condição das famílias e estatísticas socioeconômicas e demográficas.

Dividimos as 260.601 amostras rotuladas em dois subconjuntos: treino (70%) e teste (30%). Essa divisão foi realizada para garantir uma avaliação justa do modelo. Além disso, foram disponibilizadas 86.868 amostras sem rótulos, que foram utilizadas para gerar as previsões finais enviadas como parte do desafio. As previsões para essas amostras seguiram o formato exigido para submissão, onde cada *building_id* foi associado ao respectivo *damage_grade* previsto.

Os modelos foram treinados na linguagem Python, versão 3.7.13. Utilizamos as bibliotecas pandas e NumPy para manipulação e processamento dos datasets. Também usamos o scikit-learn, que é uma biblioteca de aprendizado de máquina de código aberto, para treinamento e otimização.

A Figura 1 mostra as cinco etapas do método proposto.

A transformação dos dados categóricos em variáveis foi feita aplicando a técnica de *one-hot encoding*, que transforma cada categoria em uma nova coluna com valor 0 ou 1.

Na etapa de treinamento foram testados oito modelos de classificação. A Tabela 1 mostra os oito modelos testados na etapa de treinamento.

A otimização de hiperparâmetros refere-se ao processo de ajuste dos hiperparâmetros para obter melhores resultados de F1-score. Essa otimização foi feita usando a técnica de *grid search* com conjuntos de parâmetros de grade baseados em experimentação. A Tabela 2 mostra os parâmetros de cada modelo testado no *grid search*. Os valores em negrito foram os selecionados.

Após a otimização dos hiperparâmetros, técnicas de engenharia de características foram aplicadas para avaliar se os algoritmos poderiam aumentar o F1-score. Foram aplicadas duas

Tabela 2 – Parâmetros testados por modelo. Os selecionados estão destacados em negrito.

Modelos	Parâmetros	Valores testados
Extra trees classifier	max depth	{4,5,6,7,8, 25 ,30,32}
	n estimators	{360, 500 }
Gradient boosting classifier	learning rate	{0.01, 0.025 ,0.15,0.2}
	loss	{ deviance }
	max depth	{3,5, 8 }
	max features	{ log2 ,sqrt}
	subsample	{0.5,0.618,0.7, 0.8 }
Light gradient boosting machine	n estimators	{200, 400 }
	colsample bytree	{ 0.7 ,0.8}
	max depth	{15, 20 ,25}
	min split gain	{ 0.3 ,0.4}
	n estimators	{ 100 }
	num leaves	{50, 100 ,200}
	reg alpha	{1.1, 1.2 ,1.3}
	reg lambda	{1.1, 1.2 ,1.3}
	subsample	{0.7,0.8, 0.9 }
Random forest classifier	subsample freq	{ 20 }
	criterion	{ gini ,entropy}
	max depth	{4,5,6,7, 8 }
	max features	{ auto ,sqrt,log2}
XGBoost classifier	n estimators	{200, 500 }
	objective	{ multi:softmax }
	gamma	{0.8, 0.85, 0.9 , 0.95, 1.0}
	eta	{0.022, 0.023, 0.025 , 0.027, 0.028}
	max depth	{9, 10, 11 , 12, 13}
	min child weight	{5.5, 5.8, 6.0 , 6.4, 6.7}
CatBoost classifier	subsample	{0.76, 0.78, 0.79, 0.8 , 0.82}
	colsample bytree	{0.66, 0.67, 0.69, 0.70 , 0.73}
AdaBoost classifier	border count	{2,3,6,8,10, 13 ,15}
	depth	{2,3,6,8, 10 ,13,15}
	iterations	{100,300,500,1000,3000, 5000 }
	l2 leaf reg	{1, 3 ,5,7,10}
	learning rate	{0.01,0.025,0.15,0.2, 0.5 ,1}
Asym boost classifier	algorithm	{SAMME, SAMME.R }
	learning rate	{0.01,0.025, 0.15 ,0.2,0.5,1}
	n estimators	{50,100,200,400, 500 ,1000}
Asym boost classifier	algorithm	{SAMME, SAMME.R }
	learning rate	{0.01,0.025,0.15,0.2,0.5, 1 }
	n estimators	{ 50 ,100,200,400,500,1000}



Figura 2 – Engenharia de características de dados de geolocalização

Código 1 – Obtendo características a partir das probabilidades condicionais das geolocalizações

```

1 damage1 = {}
2 damage2 = {}
3 damage3 = {}
4 for i, j in (data['geo_level_1_id'].value_counts()).iteritems():
5     n1 =
6         len(feat[feat['damage_grade']==0][feat['geo_level_1_id']==i])
7     n2 =
8         len(feat[feat['damage_grade']==1][feat['geo_level_1_id']==i])
9     n3 =
10         len(feat[feat['damage_grade']==2][feat['geo_level_1_id']==i])
11     damage1[i] = n1/j
12     damage2[i] = n2/j
13     damage3[i] = n3/j

```

técnicas de engenharia de características: substituição das variáveis geográficas por probabilidades condicionais e uma rede neural para redução da dimensionalidade das características geolocalizacionais.

A primeira técnica de engenharia de características envolve substituir os valores das colunas de geolocalização, representadas por x , por suas probabilidades condicionais em relação a cada possível classe de decisão na coluna de dano, denominada *damage_grade* (representada por y).

Assim, cada coluna de geolocalização (*geo_level*) é expandida em três novas colunas, cada uma representando a probabilidade condicional de uma classe de dano específica. Por exemplo, para cada valor em uma coluna *geo_level_1*, calculamos as probabilidades condicionais $P(y = 1|x = x_i)$, $P(y = 2|x = x_i)$ e $P(y = 3|x = x_i)$, correspondentes às classes de dano.

Como mostrado na Figura 2, essa transformação resulta em três colunas adicionais para cada nível de geolocalização, capturando a relação probabilística entre a localização geográfica e o nível de dano.

O Código 1 mostra como foi feito o processo de extração das probabilidades condicionais relacionadas a geolocalizações.

Observamos que os modelos de classificação apresentam dificuldades de classificação

Tabela 3 – Valores únicos da coluna `geo_level_3` relacionados às colunas `geo_level_1` e `geo_level_2` iguais a 6 e 487.

Geo_level_3_id	Frequência
139	11
12198	11
1755	7
7439	7
382	4
835	3
6845	3
2908	2
6936	2

quando as características categóricas possuem muitos valores únicos. Observamos que as colunas `geo_level_1`, `geo_level_2` e `geo_level_3` possuíam, além de muitos valores únicos, valores coincidentes que mostravam uma relação entre as colunas.

Um exemplo dessa relação é quando na coluna `geo_level_1` o valor é igual a 6, e ao mesmo tempo, na coluna `geo_level_2` o valor é igual a 487. Quando duas colunas ocorrem nessas condições, a coluna `geo_level_3` tem uma frequência maior de alguns valores específicos, como 139 e 12198. A Tabela 3 indica como esse relacionamento pode existir.

A partir desta observação, foi criada uma rede neural com camadas de embeddings seguidas de camadas densas para realizar a redução dimensional das características geolocalizacionais. Essa arquitetura foi escolhida porque as colunas de geolocalização possuem uma relação hierárquica: o nível de geolocalização um é superior ao nível dois e o nível dois é superior ao três. A Figura 3 mostra a arquitetura da rede.

A arquitetura da rede neural é composta por três camadas de embeddings que mapeiam as características categóricas das colunas geográficas para representações vetoriais contínuas. Cada uma dessas camadas de embedding tem diferentes tamanhos de vocabulário e dimensões.

A primeira camada, que processa a coluna `geo_level_1`, possui um vocabulário de 31 possíveis valores e uma dimensão de saída de 16. A segunda camada de embedding, responsável por processar a coluna `geo_level_2`, tem um vocabulário de 1414 valores e uma dimensão de 128. A terceira camada, que processa a coluna `geo_level_3`, tem um vocabulário ainda maior, com 11861 possíveis valores, e também gera uma representação vetorial de 128 dimensões.

Após o processamento pelas camadas de embeddings, as saídas resultantes das três camadas são concatenadas em um único vetor, conforme mostrado na operação de concatenação na Figura 3. Esse vetor concatenado tem uma dimensão total de 272 (16 da primeira camada, mais 128 da segunda e mais 128 da terceira). Essa concatenação combina as informações de cada coluna `geo_level` em uma única representação vetorial, capturando as relações hierárquicas entre as colunas `geo_level_1`, `geo_level_2` e `geo_level_3`.

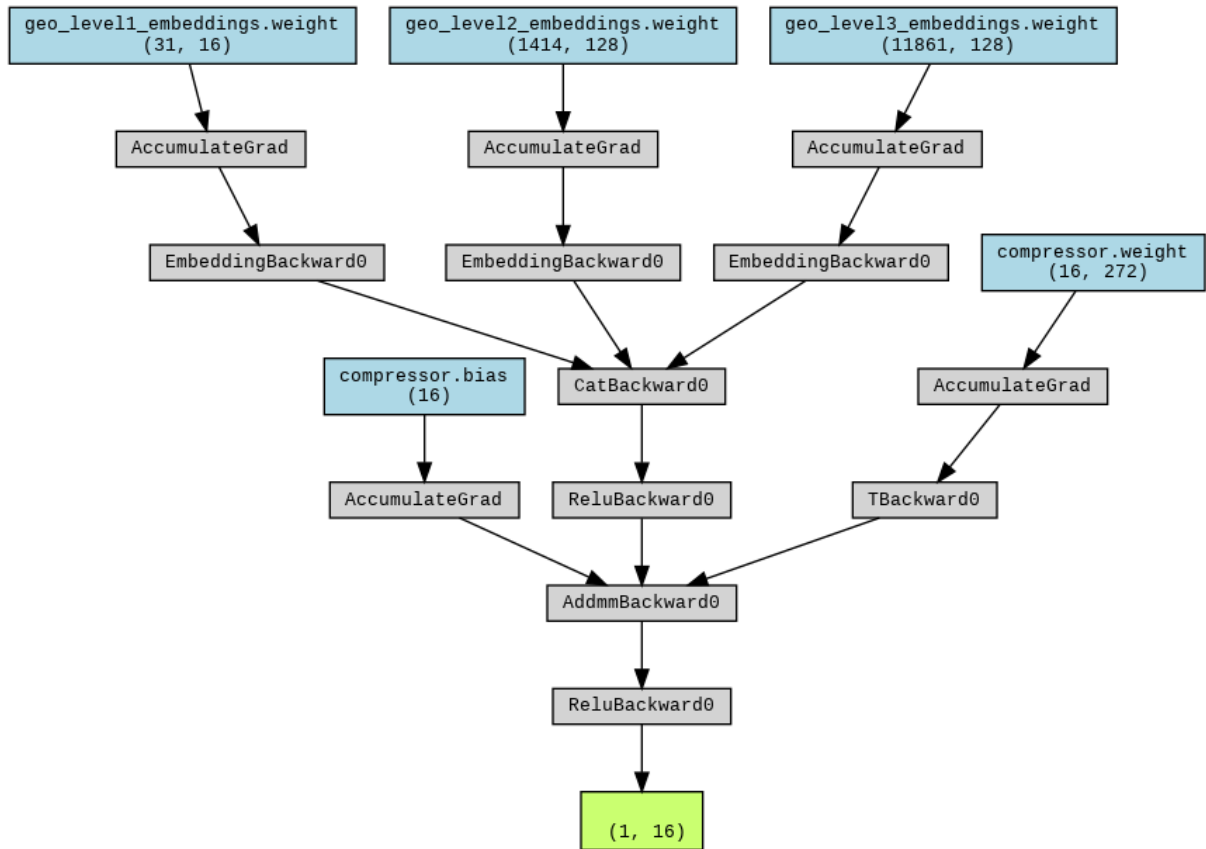


Figura 3 – Rede neural de redução dimensional para dados de geolocalização

Em seguida, o vetor concatenado passa por uma camada densa de compressão, que reduz a dimensionalidade de 272 para um espaço latente de 16 dimensões. Essa camada é responsável por comprimir a informação geográfica em uma representação compacta que retém as principais características. A função de ativação *ReLU* é aplicada após a concatenação e também após a compressão para introduzir não linearidade ao modelo, permitindo que ele capture interações complexas entre as variáveis geográficas. O resultado final da rede é um vetor de 16 dimensões que pode ser utilizado como uma entrada para outras tarefas, como classificação ou regressão.

Após comparar o F1-score dos modelos antes e depois da aplicação da engenharia de características, a matriz de confusão do modelo com o F1-score mais alto foi utilizada para analisar o desempenho de classificação do modelo.

Todo o código da solução está disponível no GitHub para referência futura ¹.

4.2 Segmentação de edifícios em imagens aéreas de desastres

O dataset xBD do desafio xview2 (Defense Innovation Unit, DoD, 2023) contém imagens de satélite de alta resolução de seis tipos diferentes de desastres naturais em todo o mundo, cobrindo uma área total de mais de 45000 quilômetros quadrados (GUPTA et al., 2019).

¹ https://github.com/tonylincon1/neural_network_feature_engineering

Tabela 4 – Escala de danos observada em edifícios. Adaptado de [Defense Innovation Unit, DoD \(2023\)](#).

Classe	Rótulo	Descrição visual da estrutura
0	<i>Background</i>	Região de fundo das imagens.
1	<i>No damage</i>	Imperturbável. Nenhum sinal de água, danos estruturais, danos nas telhas ou máscaras de queimadura.
2	<i>Minor damage</i>	Edifício parcialmente queimado, água ao redor da estrutura, fluxo vulcânico próximo, elementos do telhado faltando ou rachaduras visíveis.
3	<i>Major damage</i>	Colapso parcial da parede ou do telhado, invasão do fluxo vulcânico ou a estrutura está cercada por água ou lama.
4	<i>Destroyed</i>	A estrutura está queimada, completamente desmoronada, parcialmente ou completamente coberta com água ou lama, ou não está mais presente.

Cada pixel da imagem possui um valor que corresponde a um dos cinco rótulos definidos para classificar o estado de cada edifício. Esses rótulos foram definidos de acordo com o grau de dano apresentado pelo edifício e são os seguintes: *background*, *no damage*, *minor damage*, *major damage*, e *destroyed*. A Tabela 4 mostra os valores correspondentes a cada nível de dano. Os níveis variam de 1, *no damage*, a 4, *destroyed*. O 0 corresponde ao *background*.

O dataset xBD foi dividido em 9162 imagens de treinamento, 906 imagens de teste e 906 imagens de validação, totalizando 10974 imagens. Esta divisão permite o aprendizado, avaliação e validação do modelo, evitando *overfitting* e melhorando a generalização.

A etapa de preparação de dados desempenha um papel crucial na construção de modelos robustos de segmentação de imagens, permitindo que o modelo capture a diversidade das imagens de entrada e minimize o risco de *overfitting*.

As imagens e as máscaras possuem tamanho original de $1024 \times 1024 \times 3$ pixels, para reduzir o consumo de memória da GPU, as imagens e as máscaras foram cortadas para o tamanho de $512 \times 512 \times 3$ pixels, reduzir a carga computacional, facilitar a normalização dos dados e focar nos detalhes mais relevantes das imagens. Isso auxilia as redes a processar os dados de forma rápida e precisa, melhorando a segmentação e avaliação dos danos aos edifícios. Após o corte, usamos três abordagens para preparar a entrada de dados:

1. Imagens empilhadas: concatena duas imagens de entrada RGB em uma única com seis canais, sendo três canais de cada imagem de entrada.
2. Imagens separadas: passar imagens pré e pós-desastre separadamente para alimentar uma rede siamesa.
3. Diferença entre imagens: diferença entre as duas imagens de entrada, formando uma única

imagem de entrada com três canais correspondente à diferença de intensidade entre as imagens.

Na etapa de treinamento do modelo, foram testados quatro modelos de redes neurais, disponíveis nas bibliotecas Python *segmentation_models* e *PyTorch*: UNet, FPN, Linknet e PSPNet. Cada um desses modelos foi implementado usando sete *backbones*: BDANet, VGG16, ResNet18, ResNet50, ResNet34, ResNeXt50 e SENet154. Um total de 28 combinações foram treinadas. Todas as arquiteturas testadas recebem como entrada imagens com $512 \times 512 \times 3$ pixels e retornam como saída cinco imagens, uma para cada rótulo, com a mesma dimensão.

Usamos como rótulo final o que tiver a maior probabilidade retornada pela camada *res* que é uma camada convolucional final que reduz as saídas da rede para cinco mapas de pontuação, correspondendo a cada rótulo de classe. Essa camada recebe o resultado da última camada de decodificação e aplica uma convolução com um filtro de dimensão (1×1) , sem stride ou padding, o que permite gerar uma saída com o mesmo tamanho espacial da entrada, mas com cinco canais, um para cada classe. Esses valores servem como pontuações de probabilidade para cada classe em cada pixel da imagem de saída.

Para treinar os modelos, foi utilizada a métrica *dice loss*, que é uma métrica comumente usada para avaliar a sobreposição entre a máscara prevista e a máscara de verdade. Também utilizamos a função *focal loss*, que visa resolver o problema de desbalanceamento de classes.

Além disso, foi aplicado um peso à função de perda para lidar com o desbalanceamento de classes. Esses pesos foram definidos com base na frequência das classes nas imagens de treinamento. O otimizador usado foi AdamW com uma taxa de aprendizado de 10^{-4} e uma *weight decay* de 10^{-6} . A métrica utilizada para avaliar o modelo foi o F1-score.

Em cada época do treinamento, o modelo recebe todas as imagens do dataset de treinamento. As transformações são aplicadas aleatoriamente a cada imagem para aumentar a variabilidade dos dados. Aumentar a variabilidade dos dados permite que o modelo generalize melhor para dados novos e desconhecidos. As seguintes transformações foram aplicadas.

1. Espelhamento horizontal: espelha a imagem ao longo do eixo vertical.
2. Espelhamento vertical: espelha a imagem ao longo do eixo horizontal.
3. Rotação: gira a imagem em um ângulo aleatório entre -10 e 10 graus.
4. Escala: aplica uma escala aleatória à imagem entre 0,8 e 1,2.
5. Ruído gaussiano aditivo: adiciona ruído gaussiano à imagem com uma escala aleatória entre 0 e 0,05.
6. Normalização de contraste: ajusta o contraste da imagem para um valor aleatório entre 0,8 e 1,2.

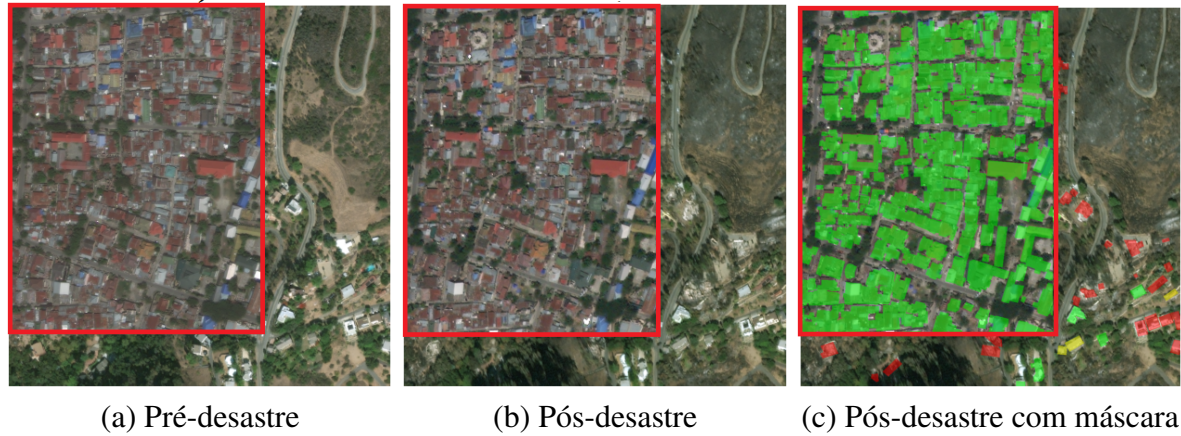


Figura 4 – Exemplo de aplicação CutMix, que está representado nos retângulos vermelhos, nas imagens.

7. Transformação elástica: aplica uma deformação elástica à imagem.

Para aumentar a quantidade de exemplos por classe, foi implementada a técnica CutMix. Esta técnica consiste em selecionar uma parte retangular aleatória, com probabilidade de 87%, de uma imagem e substituí-la por uma parte correspondente de outra imagem, com as respectivas máscaras de anotação ajustadas de forma semelhante ao mostrado na Figura 4.

A aplicação de operações de morfologia matemática (MM) foi realizada com o objetivo de melhorar o F1-score do melhor modelo de segmentação obtido no treinamento. A camada morfológica consiste em aplicar uma operação morfológica com elemento estruturante quadrado de tamanho 3×3 pixels.

Testamos quatro operações morfológicas tradicionais: erosão, dilatação, abertura e fechamento. Estas operações são amplamente utilizadas no processamento de imagens, especialmente em tarefas de segmentação, devido à sua capacidade de ajustar a estrutura das regiões segmentadas. As camadas morfológicas foram inseridas ao final do modelo Unet com o *backbone*, logo após a camada *res*, com o intuito de refinar as previsões de rótulos de segmentação. A camada *res* é uma convolução que reduz a dimensionalidade das saídas dos ramos de decodificação, gerando mapas de características finais que representam *logits* para cada classe. As saídas da *res* apresentam cinco canais, representando diferentes classes de rótulos na imagem de dimensão $512 \times 512 \times 5$.

A Figura 5 mostra parte da arquitetura onde a camada morfológica foi adicionada. Nesse exemplo em especial, mostramos a dilatação morfológica. A rede foi estruturada com as seguintes etapas:

1. BDANet: a rede começa processando as imagens de entrada (512×512) por meio do bloco BDANet, responsável pela segmentação inicial. Essa parte da rede é representada de forma abstrata como um bloco único na figura, uma vez que nosso foco está no refinamento morfológico subsequente.

2. Camada de predição (*res layer*): após o processamento pelo bloco BDANet, a camada *res* gera uma predição de segmentação com 5 canais, cada um representando uma classe de rótulo diferente. A convolução é realizada com um kernel de 1×1 , ajustando as predições para as 5 classes de saída. Na Figura 5, é possível visualizar os pesos da camada *res.weight* e o viés *res.bias*, que são ajustados durante o treinamento.
3. Camada morfológica: após a predição inicial, aplicamos a operação morfológica, nesse exemplo, estamos exibindo a operação de dilatação através da camada *self.dilation*. Essa operação utiliza um kernel de 3×3 , que expande as áreas segmentadas, suavizando os limites e garantindo uma melhor captura das regiões de interesse. A Figura 5 mostra os parâmetros da camada de dilatação *dilation.weight* e *dilation.bias*, que são treináveis.
4. Saída final: após a aplicação da dilatação, a rede gera uma saída final de dimensão (1, 5, 512, 512), com 5 canais representando as classes de rótulo refinadas. A operação morfológica contribui para o refinamento das predições, melhorando a qualidade das segmentações e eliminando pequenos ruídos ou artefatos.

Cada uma das operações morfológicas foram aplicadas individualmente após a predição de segmentação para avaliar seu impacto na qualidade final das predições.

Utilizamos o F1-score para avaliar o desempenho dos modelos de segmentação utilizados no problema de segmentação de imagens pré e pós-desastre.

O F1-score é definido como a média harmônica entre precisão e a sensibilidade e pode variar de 0 a 1, com valores próximos de 1 indicando um modelo com melhor desempenho. Para calcular o F1-score, as máscaras de segmentação produzidas pelos modelos foram comparadas com as máscaras de segmentação verdadeiras.

Além disso, para visualizar os resultados do modelo, exibimos as suas previsões para verificar se estava gerando as máscaras corretamente.

O código da solução está disponível no GitHub para referência futura ².

² https://github.com/ddantas-ufs/2024_building

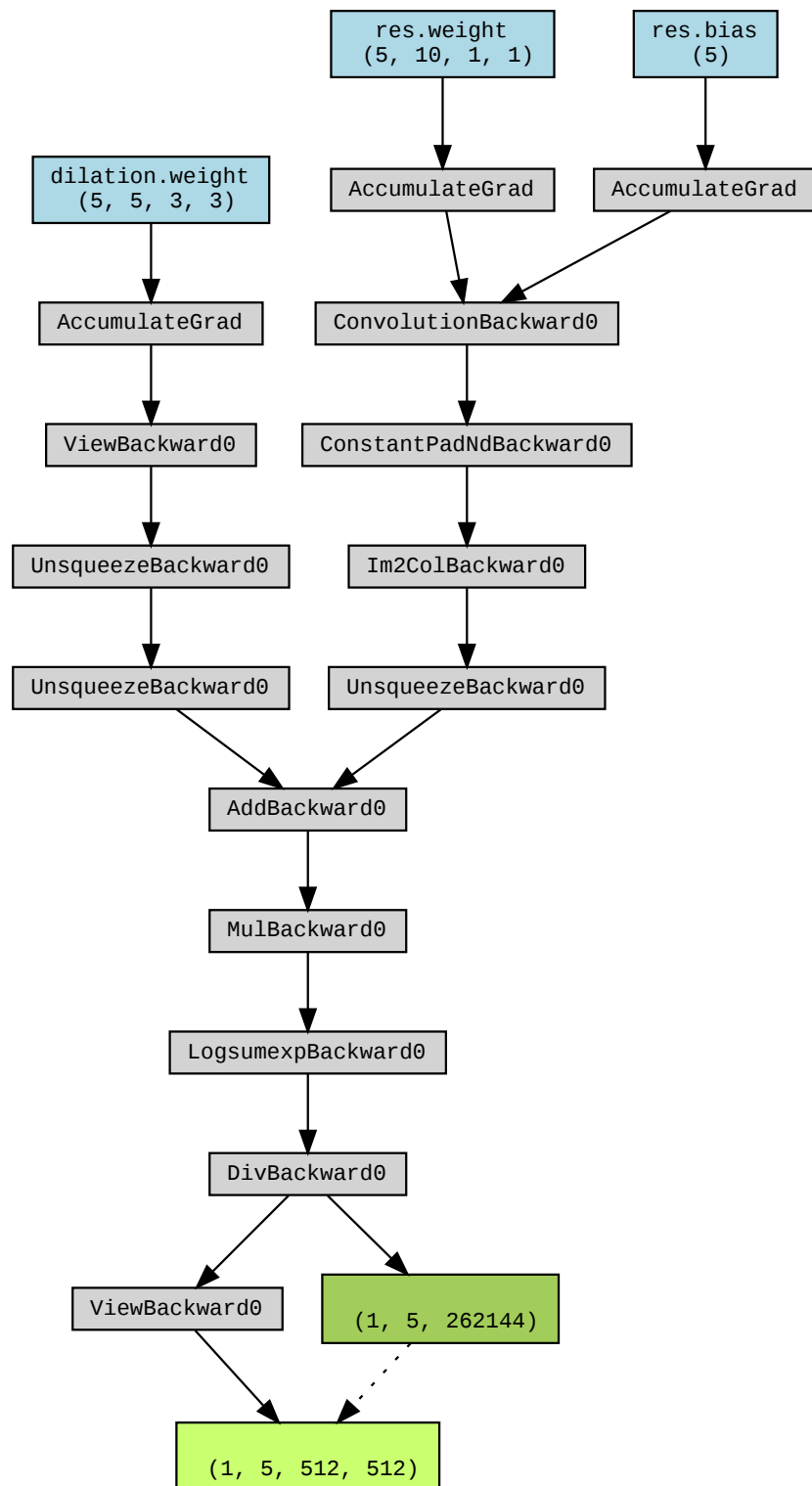


Figura 5 – Arquitetura da rede BDANet resumida com foco na camada morfológica de dilatação e na saída final da rede.

5

Resultados

Este capítulo apresenta os resultados obtidos para os dois problemas abordados neste trabalho: a classificação dos níveis de dano em edifícios utilizando dados tabulares e a segmentação de imagens aéreas para a identificação de danos visuais.

Para a classificação, foram testados diversos modelos de aprendizado de máquina com diferentes técnicas de otimização de hiperparâmetros e engenharia de características, com o objetivo de maximizar o F1-score. Para a segmentação, foram implementados e avaliados modelos de redes neurais convolucionais, utilizando imagens pré e pós-desastre para segmentar as áreas danificadas.

A seguir, detalhamos os resultados para cada uma das abordagens, com comparações do F1-score, para identificar o desempenho em diferentes classes de dano.

5.1 Classificação de nível de dano em dados tabulares

5.1.1 Treinamento dos modelos

Após a preparação dos dados foi possível comparar o desempenho dos diversos modelos treinados. A Tabela 5 mostra os resultados em ordem decrescente de F1-score.

O F1-score mostra que alguns modelos tiveram melhor desempenho na classificação dos danos, principalmente os dois primeiros modelos, CatBoost e XGBoost, que obtiveram o maior F1-score.

Os modelos treinados também foram comparados com o modelo de classificação *baseline* que teve um F1-score de 0,5248. O modelo *baseline* serve de base de comparação com modelos mais complexos, pois é um modelo que faz previsões sem tentar encontrar padrões nos dados, o que equivale à classificação aleatória. A diferença de F1-score entre o melhor modelo e o *baseline* é de 27,9%, o que evidencia o bom resultado do modelo.

Tabela 5 – F1-score dos modelos propostos

Nº	Modelos	F1-score
1	CatBoost classifier	0,7279
2	XGBoost classifier	0,7248
3	Random forest classifier	0,7187
4	Light gradient boosting machine	0,7155
5	AsymBoost classifier	0,7015
6	Gradient boosting classifier	0,6866
7	AdaBoost classifier	0,6686
8	Extra trees classifier	0,6404

Tabela 6 – Comparação dos resultados do modelo com otimização de hiperparâmetros.

Nº	Modelos	F1-score Antes da Otimização	F1-score Depois da Otimização	Diferença
1	XGBoost classifier	0,7248	0,7419	0,0171
2	Light gradient boosting machine	0,7155	0,7324	0,0169
3	Gradient boosting classifier	0,6866	0,7038	0,0172
4	Extra trees classifier	0,6404	0,6993	0,0589
5	CatBoost classifier	0,7279	0,6973	-0,0305
6	AdaBoost classifier	0,6686	0,6684	-0,0002
7	AsymBoost classifier	0,7015	0,6526	-0,0489
8	Random forest classifier	0,7187	0,6502	-0,0685

5.1.2 Otimização de hiperparâmetros

Na tentativa de melhorar o F1-score, foi aplicada a técnica de *grid search* e *cross validation* para otimizar os hiperparâmetros. A Tabela 6 mostra a diferença no F1-score dos modelos antes e depois da otimização.

Houve melhoria no F1-score em quatro modelos: XGBoost, LGBM, *gradient boosting* e *extra trees*. Pode-se observar, por exemplo, que o F1-score aumentou no modelo *extra trees* em 0,0589. Por outro lado, os modelos CatBoost, AsymBoost e *random forest* apresentaram uma diminuição no seu F1-score. O modelo AdaBoost permaneceu quase inalterado.

5.1.3 Engenharia de características

Após a otimização dos hiperparâmetros, técnicas de engenharia de características foram aplicadas na tentativa de aumentar o F1-score dos modelos. A Figura 6 mostra os resultados dos modelos antes e depois da otimização, bem como após a aplicação da engenharia de características.

O F1-score de alguns dos modelos aumentou após a aplicação das técnicas de engenharia de características. O classificador XGBoost mostrou um aumento de 0,7419 após otimização para

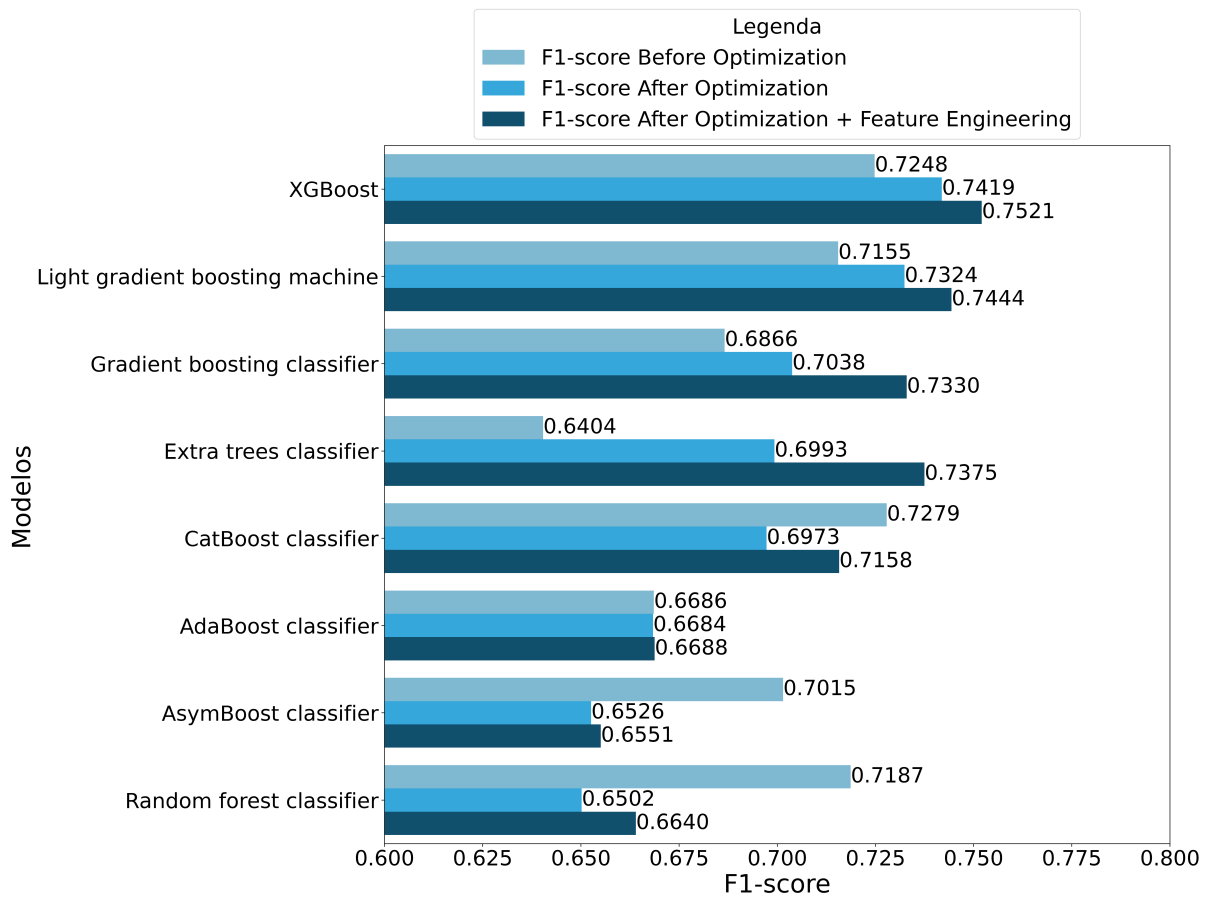


Figura 6 – Comparação do F1-score antes e depois da otimização e com engenharia de características

0,7521 com engenharia de características. Da mesma forma, os classificadores LGBM e *gradient boosting* exibiram aumentos no F1-score, aumentando para 0,7444 e 0,7330, respectivamente.

No entanto, alguns modelos tiveram poucas melhorias ou alterações insignificantes. Por exemplo, o *extra trees* melhorou de 0,6993 para 0,7375. O random forest, no entanto, mostrou uma diminuição no F1-score após a otimização dos hiperparâmetros, mas viu uma recuperação com engenharia de características, melhorando para 0,6640.

Esses resultados sugerem que a engenharia de características, particularmente técnicas como a substituição de categorias de dados geográficos por probabilidades condicionais e o uso de redes neurais que reduzem a dimensionalidade das características de geolocalização, melhoram o F1-score. As melhorias observadas em múltiplos modelos demonstram o potencial de melhoria dos modelos de classificação.

5.1.4 Avaliação do modelo

Após treinamento, otimização do modelo e implementação das técnicas de engenharia de características, o XGBoost foi escolhido como o melhor modelo com base no F1-score. Para melhor analisar os resultados, obtivemos a matriz de confusão para o XGBoost, conforme

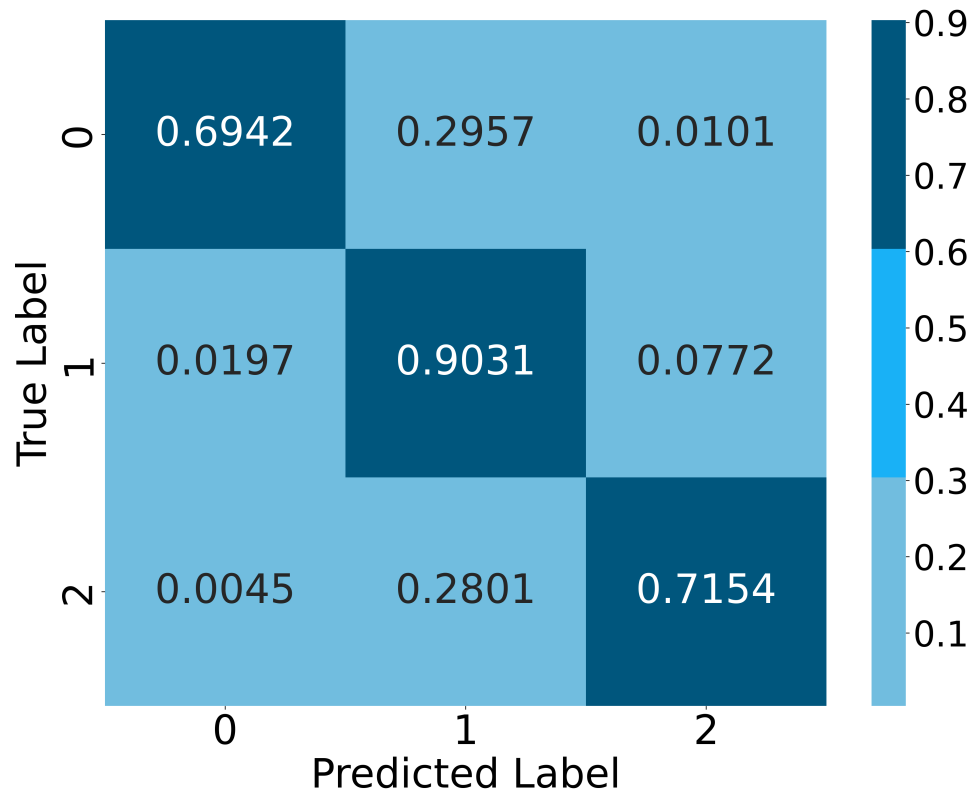


Figura 7 – Matriz de confusão do XGBoost

mostrado na Figura 7.

A matriz de confusão revela diferentes acurácias para as classes dos dados de teste. A acurácia para a classe 0 (dano baixo) é de 69,42%, para a classe 1 (dano médio) é de 90,31% e para a classe 2 (destruição completa) é de 71,54%. Estes resultados indicam que o modelo tem melhor desempenho na classificação de edifícios com níveis de danos médios, o que pode ser atribuído ao desbalanceamento nos dados de treinamento.

5.1.5 Posição no desafio e comparação com outros estudos

Treinamento, otimização e teste foram feitos com os dados que foram disponibilizados pela plataforma, dados rotulados, mas o desafio da plataforma Data Driven ([DRIVENDATA, 2022](#)) forneceu um dataset não rotulado a ser utilizado e enviados como forma de validação do modelo criado. Após treinar o modelo, fizemos as predições para o dataset não rotulado e enviamos os resultados a plataforma do desafio. Neste dataset não rotulado, alcançamos um F1-score de 0,7542. Essa pontuação nos posicionou em 4^a entre os 7621 participantes do desafio, conforme mostrado na Figura 8.

O modelo XGBoost, em comparação com outros estudos, teve o F1-score mais alto no dataset não rotulado. A Tabela 7 compara o F1-score dos estudos publicados pelos participantes do desafio.

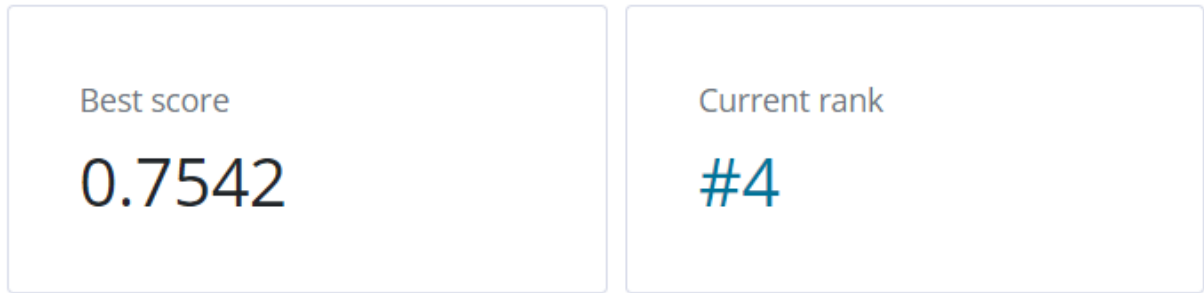


Figura 8 – Posição no desafio

Tabela 7 – Comparação de modelos entre estudos

Nº	Modelos	F1-score	Posição no desafio
1	Modelo proposto	0,7542	4
2	Kumaraswamy, Reddy e Kolla (2020)	0,7459	631
3	Guyette, Lewis e Stone (2021)	0,7260	1410
4	Liang (2021)	0,7150	1642
5	Mangalathu et al. (2020)	0,6600	2020

Tabela 8 – Resultados dos modelos no dataset de teste (5 maiores F1-score)

Nº	Abordagens diferentes para imagens registradas	Modelo	Backbone	F1-score geral
1	Imagens separadas	Unet	BDANet	0,7613
2	Imagens empilhadas	Linknet	ResNet18	0,4332
3	Imagens empilhadas	Linknet	VGG16	0,4012
4	Imagens empilhadas	Linknet	ResNet50	0,3687
5	Diferença entre imagens	Linknet	ResNet50	0,2031

Comparando a versão atual do modelo proposto com o trabalho de [Kumaraswamy, Reddy e Kolla \(2020\)](#) temos uma diferença de 0,0083 no F1-score e há uma diferença de 627 posições.

5.2 Segmentação de edifícios em imagens aéreas de desastres

5.2.1 Experimentos

Realizamos 28 experimentos para identificar quais modelos possuem maior F1-score. A Tabela 8 mostra o F1-score dos cinco melhores modelos para segmentação de imagens pré e pós-desastre. Os modelos foram avaliados utilizando o F1-score com 25 épocas de treinamento.

O modelo Unet com *backbone* BDANet, que tratou as imagens separadas, teve o melhor desempenho, com F1-score de 0,7613. Este resultado destaca a eficácia da abordagem de imagens

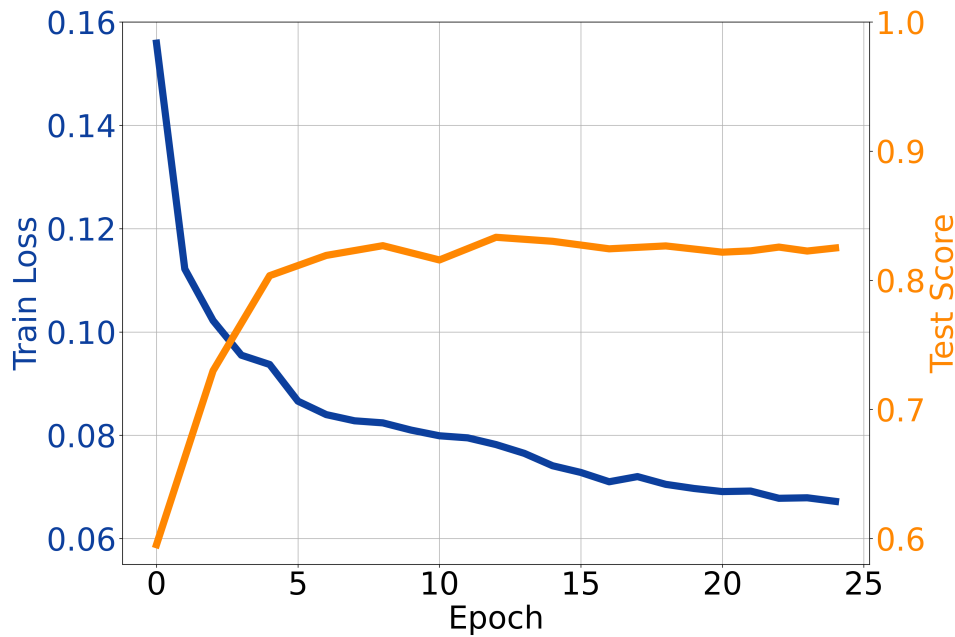


Figura 9 – Perda de treinamento e score de teste por épocas do melhor modelo

separadas, sugerindo que a manutenção da individualidade da imagem pode reter características críticas que podem ser perdidas em outras abordagens.

Além disso, o F1-score obtido também pode ser atribuído ao uso do *backbone* BDANet, uma vez que o modelo apresentou F1-score mais alto do que os outros modelos testados. A BDANet foi proposta por Shen et al. (2021) e obteve F1-score superior a outras arquiteturas na tarefa de segmentação de danos.

Após a etapa de treinamento, o modelo com maior F1-score foi escolhido para analisar os resultados e encontrar oportunidades de melhorar o F1-score.

A Figura 9 mostra a *dice loss* e o score de teste ao longo das épocas de treinamento. O score de teste foi calculado como uma média ponderada dessas duas métricas: 0,3 vezes o *dice loss* mais 0,7 vezes o F1-score.

A *dice loss* está diminuindo consistentemente ao longo das épocas, sugerindo que o modelo está aprendendo e melhorando continuamente. Este é um sinal positivo de que o treinamento está progredindo conforme o esperado.

O score de teste geralmente aumenta com o tempo. Isto indica que o desempenho do modelo no dataset está melhorando.

A Figura 10 mostra a evolução do F1-score por classe, ao longo de cada época, no dataset de teste do melhor modelo.

O F1-score *overall* teve uma tendência crescente no dataset de teste, embora tenha havido algumas oscilações. Isto sugere que o modelo melhorou ao longo das épocas.

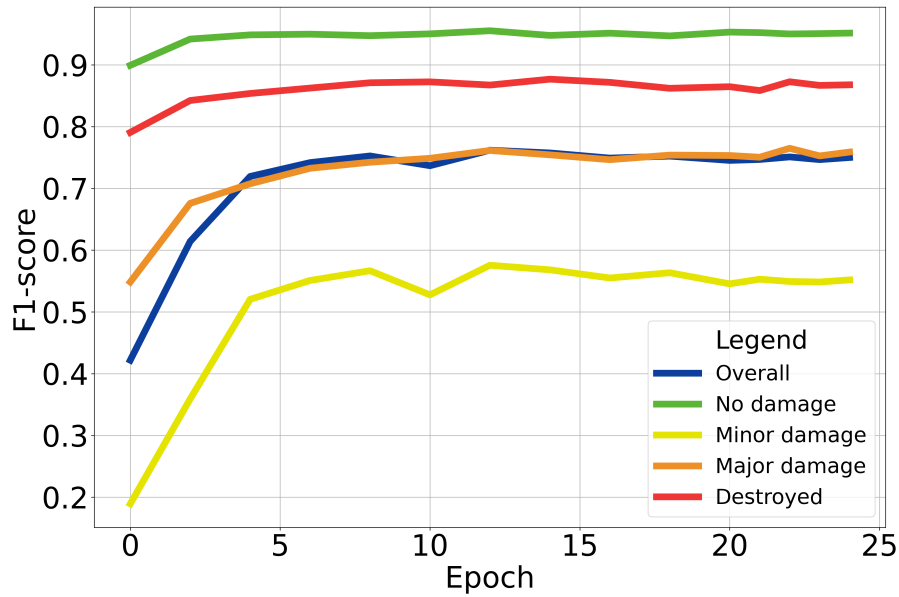


Figura 10 – F1-score por classe em cada época do melhor modelo

Observamos que a classe *no damage* consistentemente tem o F1-score mais alto, sugerindo que o modelo é mais eficaz na previsão desta classe.

Em contraste, as classes *minor damage* e *major damage* têm os F1-scores mais baixos, sugerindo que o modelo tem mais dificuldade em prever essas classes ou que pode haver poucos exemplos dessas classes em comparação com as demais classes.

Além de observar a tendência por classe do F1-score *overall* ao longo das épocas, os resultados visuais das previsões também foram avaliados, conforme mostrado na Figura 11.

Apesar do F1-score do modelo ter tido um resultado satisfatório, as previsões do modelo apresentaram pequenos erros na completude dos polígonos. Um desses erros pode ser visto claramente na segunda linha da Figura 11, onde temos alguns retângulos previstos com pequenos furos ou preenchimentos inadequados.

A partir dessas observações, foi realizado um experimento com aplicação de camadas com operações morfológicas para melhorar o F1-score. A Tabela 9 mostra o F1-score *overall* dos experimentos depois da adição de camadas morfológicas ao melhor modelo.

A última coluna da Tabela 9 mostra a melhoria do F1-score obtido com a adição das camadas morfológicas. Quando a camada morfológica é aplicada após a camada de decisão, há um aumento do F1-score, levando a uma melhoria de 0,0386 em relação ao melhor modelo, atingindo um escore F1 de 0,7999.

5.2.2 Comparação com outros estudos

A Tabela 10 mostra como o modelo proposto neste estudo se compara a outros modelos de estudos relacionados.

Tabela 9 – Resultados da aplicação de camadas morfológicas no melhor modelo.

Operação	F1-score	Aumento F1-score
Dilatação	0,7999	0.0386
Erosão	0,7813	0.0198
Fechamento	0,7766	0.0151
Abertura	0,7615	0,0000

Tabela 10 – Comparação do F1-score entre outros estudos. A tabela está em ordem decrescente pelo F1-score.

Models	<i>Overall</i>	<i>No damage</i>	<i>Minor damage</i>	<i>Major damage</i>	<i>Destroyed</i>
Shen et al. (2021)	0,806	0,925	0,616	0,788	0,876
Modelo proposto	0,7999	0,9543	0,6010	0,7621	0,8792
Long, Shelhamer e Darrell (2015)	0,765	0,919	0,532	0,708	0,861
Weber e Kan (2020)	0,741	0,906	0,493	0,722	0,837
Hou et al. (2019)	0,737	0,884	0,518	0,684	0,855
Gupta et al. (2019)	0,265	0,663	0,143	0,009	0,465

O modelo proposto obteve F1-score de 0,7999. Este modelo superou o BDANet nas classes *no damage* e *destroyed* com um F1-score de 0,9543 e 0,8792 respectivamente. Contudo, o modelo proposto teve desempenho inferior em todas as outras classes quando comparado ao BDANet. Podemos tentar outras técnicas para melhorar o F1-score, como métodos *ensemble*, CutMix concentrado nas classes que tiveram F1-score mais baixas ou ajustar os hiperparâmetros com um *grid search*.

Os modelos Long, Shelhamer e Darrell (2015), Weber e Kan (2020), Hou et al. (2019) e Gupta et al. (2019) alcançaram F1-scores mais baixos do que o modelo proposto e o BDANet.

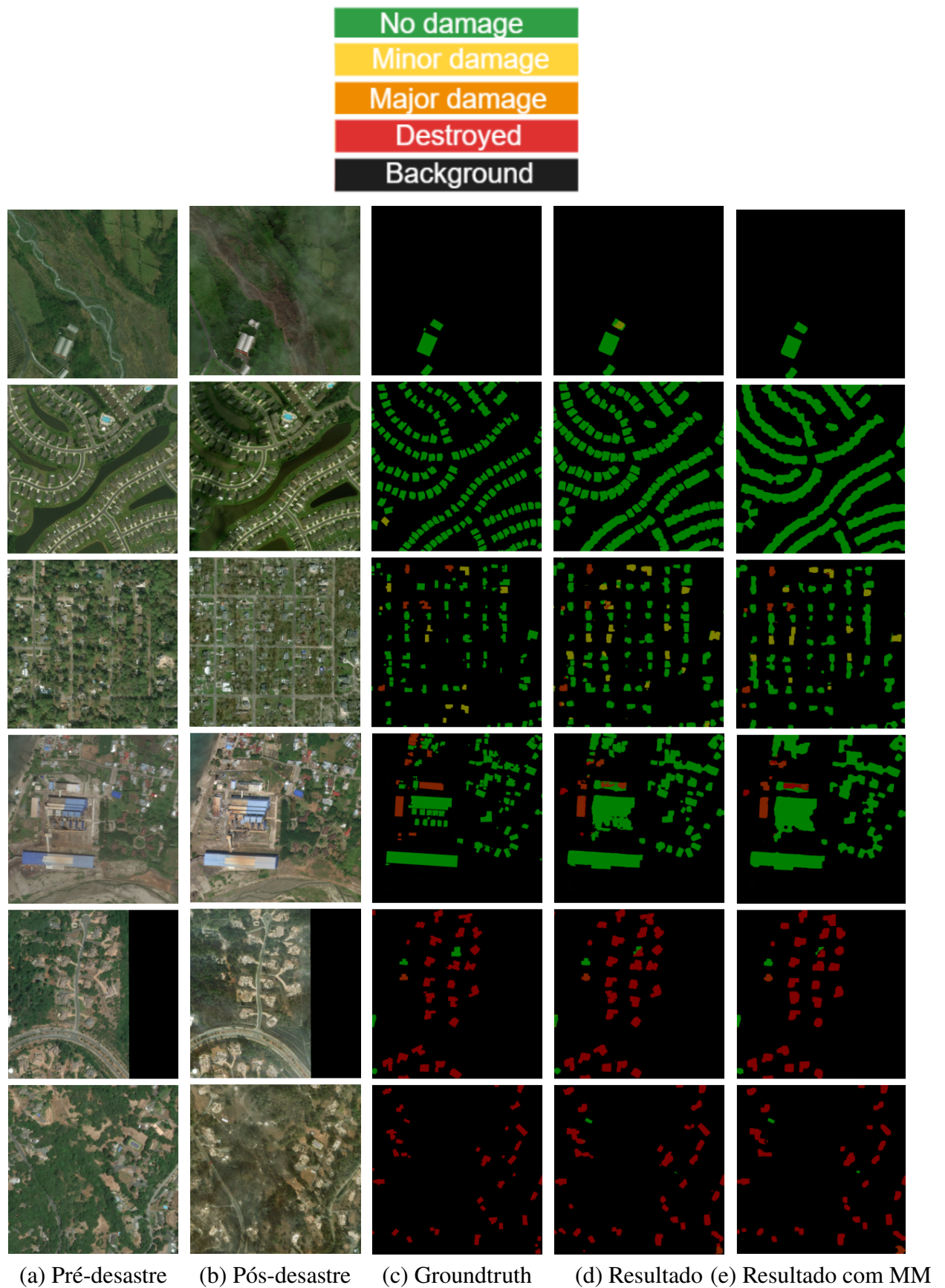


Figura 11 – Segmentação do melhor modelo em um exemplo de imagem.

6

Conclusões

Neste estudo, foram desenvolvidos, implementados e avaliados modelos para classificar e segmentar níveis de danos em edifícios após desastres naturais, utilizando tanto dados tabulares quanto imagens de satélite pré e pós-desastre registradas. A pesquisa baseou-se na criação de modelos de classificação e segmentação com otimização de hiperparâmetros e engenharia de características.

Para a criação do modelo de classificação que prevê o nível de danos para o desafio Richter's Predictor ([DRIVENDATA, 2022](#)), utilizamos um dataset com informações sobre os níveis de danos ocorridos durante o terremoto de Gorkha no Vale de Kathmandu. A linguagem de programação utilizada foi Python. Diversos modelos de classificação foram treinados, otimizados e, finalmente, utilizamos engenharia de características para maximizar o F1-score.

O modelo XGBoost de classificação, utilizando dados tabulares, demonstrou um desempenho satisfatório, posicionando-se em 4ª colocação no desafio e superando outros estudos sobre o tema. Com a otimização de hiperparâmetros usando *grid search* e o uso de engenharia de características, o F1-score do modelo aumentou de 0,7248 para 0,7521 nos dados de teste e obteve um F1-score de 0,7542 após a submissão da classificação dos dados não rotulados ao site do desafio.

As técnicas de engenharia de características, especialmente a substituição das categorias de dados geográficos por probabilidades condicionais e o uso de redes neurais que reduzem a dimensionalidade das características de geolocalização, contribuíram consideravelmente para o aumento do F1-score.

Para a segmentação de imagens pré e pós-desastre, vários modelos foram avaliados quanto ao seu F1-score. O modelo Unet com a arquitetura BDANet obteve o melhor desempenho, alcançando um F1-score de 0,7613, indicando a eficácia da abordagem de imagens separadas, preservando suas características individuais que podem ser perdidas na abordagem de diferença entre imagens.

A arquitetura BDANet pode ter contribuído para o alto desempenho deste modelo na extração de características relevantes das imagens, conforme evidenciado no estudo de Shen (SHEN et al., 2021). A abordagem de imagens empilhadas obteve um desempenho inferior, com F1-scores mais baixos que os outros modelos testados. Observou-se que a arquitetura ResNet18 obteve um F1-score mais alto que VGG16 e ResNet50. Com a aplicação de operações de morfologia matemática, o F1-score global aumentou de 0,7613 para 0,7999.

Comparando o desempenho do modelo proposto com outros estudos, observou-se que o modelo proposto superou o BDANet nas classes *no damage* e *destroyed*. Além disso, o modelo proposto obteve maior F1-score em todas as classes comparado aos modelos FCN, MTF, WNet e ao modelo base. No entanto, o modelo proposto teve desempenho inferior em todas as outras classes quando comparado ao BDANet, sugerindo que ainda há espaço para melhorar o F1-score.

Observou-se que, tanto nos modelos de classificação com dados tabulares quanto nos modelos de segmentação de imagens, o uso de técnicas de engenharia de características, otimização de hiperparâmetros e morfologia matemática proporcionou melhorias consideráveis do F1-score.

Essa diferença pode ser atribuída à natureza dos dados e às técnicas empregadas. Após desastres naturais, a disponibilidade de dados tabulares detalhados pode ser limitada, enquanto imagens de satélite são mais acessíveis e oferecem uma cobertura abrangente das áreas afetadas. Isso favorece o uso de modelos de segmentação, que podem explorar melhor essas imagens para identificar e avaliar os danos. Assim, a capacidade de segmentar imagens com precisão torna-se crucial para avaliações rápidas e eficientes, destacando a importância de investir em modelos que possam aproveitar dos vastos datasets de imagens disponíveis após desastres.

Trabalhos futuros podem focar em melhorar tanto a parte de classificação como segmentação desenvolvidos por esse trabalho. Sobre a classificação de dados tabulares, trabalhos futuros podem explorar o balanceamento do dataset, pois dados balanceados tendem a melhorar o desempenho dos modelos.

Dado o desbalanceamento das classes observadas nos dados de treinamento, técnicas como *synthetic minority oversampling technique* (SMOTE), *adaptive synthetic sampling* (ADASYN) e *conditional tabular generative adversarial networks* (CTGAN) poderiam ser empregadas para gerar dados sintéticos. Essas técnicas podem aumentar o dataset criando novas instâncias de classes sub-representadas, ajudando os modelos a generalizarem melhor e alcançarem um F1-score mais alto.

Além disso, outras técnicas de pré-processamento, treinamento e otimização podem ser exploradas para melhorar ainda mais o desempenho na classificação dos níveis de danos. A análise exploratória de dados, quando combinada com engenharia de características, pode ajudar a identificar e mitigar problemas associados ao uso de modelos de classificação aplicado a dados de alta dimensionalidade.

Por outro lado, trabalhos futuros podem melhorar os modelos de segmentação com a implementação de métodos *ensemble* e aplicação de CutMix concentrado nas classes com F1-scores mais baixos e o ajuste fino dos hiperparâmetros com *grid search* ou *random search* são estratégias adicionais que podem ser exploradas para melhorar o desempenho dos modelos.

Em conclusão, embora os modelos propostos tenham alcançado um alto F1-score em relação a outros trabalhos, esforços contínuos para balancear o dataset e experimentar diferentes modelos e técnicas de otimização podem contribuir para alcançar ainda maior precisão e robustez na classificação e segmentação de danos em edifícios após desastres naturais. O uso combinado de *grid search* e engenharia de características para dados tabulares demonstrou-se eficaz na melhoria do F1-score dos modelos de classificação. Já na segmentação, a combinação de transformações, *Curmix* e operações de morfologia matemática aumentou o F1-score, reforçando a relevância dessas técnicas no treinamento de modelos de aprendizado de máquina.

Referências

- ABHISHEK, L. Optical character recognition using ensemble of SVM, MLP and extra trees classifier. In: IEEE. *2020 International Conference for Emerging Technology (INCET)*. [S.l.], 2020. p. 1–4. Citado na página 20.
- ADHVARYU, P. S. A review on diverse ensemble methods for classification. *IOSR Journal of Computer Engineering*, v. 1, p. 27–32, 2012. Citado na página 22.
- ALAMIN, I. M. J.; JEBERSON, W.; BAJAJ, H. Improved region growing based breast cancer image segmentation. *International Journal of Computer Applications*, v. 975, p. 888, 2016. Citado na página 23.
- ALVARENGA-JÚNIOR, W. J. de. *Métodos de otimização hiperparamétrica: um estudo comparativo utilizando árvores de decisão e florestas aleatórias na classificação binária*. Dissertação (Master's Thesis) — Universidade Federal de Minas Gerais, 2018. Citado na página 21.
- ATAEI, M.; OSANLOO, M. Using a combination of genetic algorithm and the grid search method to determine optimum cutoff grades of multiple metal deposits. *International Journal of Surface Mining, Reclamation and Environment*, Taylor & Francis, v. 18, n. 1, p. 60–78, 2004. Citado na página 21.
- BANON, G. J. F.; BARRERA, J. *Bases da Morfologia Matemática para a análise de imagens binárias*. [S.l.]: UFPE-DI, 1994. Citado na página 25.
- BASU, A. A survey on recent trends in deep learning for nucleus segmentation from histopathology images. *Evolving Systems*, p. 1–46, 2023. Citado na página 24.
- BENKHOUI, Y.; EL-KORCHI, T.; REINHOLD, L. Effective pavement crack delineation using a cascaded dilation module and fully convolutional networks. In: *Geometry and Vision: First International Symposium, ISGV 2021, Auckland, New Zealand, January 28-29, 2021, Revised Selected Papers 1*. [S.l.: s.n.], 2021. Citado na página 25.
- BIAU, G.; CADRE, B.; MEVEL, L. Consistency of random forests. *Journal of Machine Learning Research*, v. 9, n. Oct, p. 1925–1947, 2008. Citado na página 20.
- BREIMAN, L. Random forests. *Machine learning*, Springer, v. 45, n. 1, p. 5–32, 2001. Citado na página 20.
- BURDESCU, D. D. et al. Efficient volumetric segmentation method. In: *2014 Federated Conference on Computer Science and Information Systems*. [S.l.: s.n.], 2014. p. 659–668. Citado na página 28.
- CASARI, A.; ZHENG, A. Feature engineering for machine learning. *O'Reilly Media, Inc.*, p. 218, 2018. Citado na página 22.
- CHEN, L.-C. Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587*, 2017. Citado na página 24.

- CHEN, Q. et al. Identification and evaluation of urban construction waste with vhr remote sensing using multi-feature analysis and a hierarchical segmentation method. *Remote Sensing*, v. 13, n. 1, p. 158, 2021. Citado na página 24.
- CHEN, T.; GUESTRIN, C. XGBoost: A scalable tree boosting system. In: *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. [S.l.: s.n.], 2016. p. 785–794. Citado na página 20.
- CHOWDHURY, T.; RAHNEMOONFAR, M. Self attention based semantic segmentation on a natural disaster dataset. In: *2021 IEEE International Conference on Image Processing (ICIP)*. [S.l.: s.n.], 2021. p. 2798–2802. Citado na página 24.
- DA, Y.; JI, Z.; ZHOU, Y. Building damage assessment based on siamese hierarchical transformer framework. *Mathematics*, v. 10, n. 11, p. 1898, 2022. Citado 2 vezes nas páginas 24 e 28.
- Defense Innovation Unit, DoD. *xView2: Assess building damage*. 2023. <<https://xview2.org/dataset>>. Accessed: 30/04/2023. Citado 3 vezes nas páginas 9, 35 e 36.
- DENG, X. Fault prediction of complete electrical engineering equipment based on deep learning technology. 2024. Citado na página 21.
- DESIANI, A. et al. Majority voting as ensemble classifier for cervical cancer classification. *Science and Technology Indonesia*, v. 8, p. 84–92, 2023. Citado na página 23.
- DING, W.; WU, S. Abc-based stacking method for multilabel classification. *Turkish Journal of Electrical Engineering & Computer Sciences*, v. 27, p. 4231–4245, 2019. Citado na página 22.
- DRIVENDATA. *Driven Data - Richter's Predictor: Modeling Earthquake Damage*. 2022. <<https://www.drivendata.org/competitions/57/nepal-earthquake/page/134/>>. Accessed: 14/05/2022. Citado 5 vezes nas páginas 15, 26, 31, 44 e 50.
- FINOTTI, R. P. et al. Detecção de danos estruturais utilizando análise estatística e auto-codificador esparso. *Mecânica Computacional*, v. 36, n. 19, p. 881–890, 2018. Citado 2 vezes nas páginas 26 e 27.
- GEURTS, P.; ERNST, D.; WEHENKEL, L. Extremely randomized trees. *Machine learning*, Springer, v. 63, n. 1, p. 3–42, 2006. Citado na página 20.
- GIRGIN, S. The natech events during the 17 August 1999 Kocaeli earthquake: aftermath and lessons learned. *Natural Hazards and Earth System Sciences*, Copernicus GmbH, v. 11, n. 4, p. 1129–1140, 2011. Citado na página 14.
- GUPTA, R. et al. Creating xBD: A dataset for assessing building damage from satellite imagery. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*. [S.l.: s.n.], 2019. p. 10–17. Citado 4 vezes nas páginas 15, 28, 35 e 48.
- GUYETTE, S.; LEWIS, W.; STONE, N. Predicting damage based on Gorkha earthquake dataset. *Proceedings of Student Research and Creative Inquiry Day*, v. 5, 2021. Citado 3 vezes nas páginas 26, 27 e 45.
- HE, W. et al. Differentiable automatic data augmentation by proximal update for medical image segmentation. *IEEE/CAA Journal of Automatica Sinica*, v. 9, n. 7, p. 1315–1318, 2022. Citado na página 23.

- HOU, B. et al. From W-Net to CDGAN: Bitemporal change detection via deep learning techniques. *IEEE Transactions on Geoscience and Remote Sensing*, IEEE, v. 58, n. 3, p. 1790–1802, 2019. Citado na página 48.
- KE, G. et al. LightGBM: a highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, v. 30, 2017. Citado na página 20.
- KHALID, S.; KHALIL, T.; NASREEN, S. A survey of feature selection and feature extraction techniques in machine learning. *2014 Science and Information Conference*, 2014. Citado na página 21.
- KHURANA, U. et al. Cognito: Automated feature engineering for supervised learning. In: *IEEE. 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW)*. [S.l.], 2016. p. 1304–1307. Citado na página 21.
- KRAUSMANN, E.; CRUZ, A. M. Impact of the 11 March 2011, Great East Japan earthquake and tsunami on the chemical industry. *Natural hazards*, Springer, v. 67, n. 2, p. 811–828, 2013. Citado na página 14.
- KUMARASWAMY, A.; REDDY, B. N.; KOLLA, R. Richter’s predictor: Modelling earthquake damage using multi-class classification models. In: *IEEE. 2020 Third International Conference on Advances in Electronics, Computers and Communications (ICAEECC)*. [S.l.], 2020. p. 1–5. Citado 3 vezes nas páginas 26, 27 e 45.
- LAINE, R. F. et al. Avoiding a replication crisis in deep-learning-based bioimage analysis. *Nature Chemical Biology*, v. 18, n. 10, p. 1136–1144, 2021. Citado na página 25.
- LI, C. *A gentle introduction to gradient boosting*. 2016. [Online]. Available: <http://www.ccs.neu.edu/home/vip/teach/MLcourse/4_boosting/slides/gradient_boosting.pdf>. Last accessed: 2016. Citado na página 20.
- LI, K. Upernet optimisation and application to mousehole segmentation. In: *International Conference on Artificial Intelligence, Virtual Reality, and Visualization (AIVRV 2022)*. [S.l.: s.n.], 2023. v. 12588, p. 157–163. Citado na página 24.
- LI, T.-T. et al. Semantic segmentation of urban street scene based on convolutional neural network. *Journal of Physics Conference Series*, v. 1682, n. 1, p. 012077, 2020. Citado na página 27.
- LI, W.; JAFARI, O.; ROTHER, C. Deep object co-segmentation. In: *Computer Vision–ACCV 2018: 14th Asian Conference on Computer Vision, Perth, Australia, December 2–6, 2018, Revised Selected Papers, Part III 14*. [S.l.: s.n.], 2019. p. 638–653. Citado na página 24.
- LI, X.; WANG, L.; SUNG, E. AdaBoost with SVM-based component classifiers. *Engineering Applications of Artificial Intelligence*, Elsevier, v. 21, n. 5, p. 785–795, 2008. Citado 2 vezes nas páginas 19 e 21.
- LIANG, Y. Application of machine learning methods to predict the level of buildings damage in the Nepal earthquake. *Academia.edu*, p. 1–7, 2021. Citado na página 45.
- LIAW, A.; WIENER, M. Classification and regression by RandomForest. *R News*, v. 2, n. 3, p. 18–22, 2002. Citado 2 vezes nas páginas 19 e 20.

- LIU, C.; FANG, D.; ZHAO, L. Reflection on earthquake damage of buildings in 2015 Nepal earthquake and seismic measures for post-earthquake reconstruction. In: ELSEVIER. *Structures*. [S.l.], 2021. v. 30, p. 647–658. Citado na página 14.
- LIU, J. Bdhe-net: A novel building damage heterogeneity enhancement network for accurate and efficient post-earthquake assessment using aerial and remote sensing data. *Applied Sciences*, 2024. Citado na página 15.
- LONG, J.; SHELHAMER, E.; DARRELL, T. Fully convolutional networks for semantic segmentation. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2015. p. 3431–3440. Citado na página 48.
- LYU, C.; JIANG, J. Remote sensing image registration with line segments and their intersections. *Remote Sensing*, v. 9, n. 5, p. 439, 2017. Citado na página 28.
- MA, H. et al. Improved CNN classification method for groups of buildings damaged by earthquake, based on high resolution remote sensing images. *Remote Sensing*, v. 12, n. 2, p. 260, 2020. Citado na página 28.
- MANGALATHU, S. et al. Classifying earthquake damage to buildings using machine learning. *Earthquake Spectra*, SAGE Publications Sage UK: London, England, v. 36, n. 1, p. 183–208, 2020. Citado na página 45.
- MINAEE, S. et al. Image segmentation using deep learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 44, n. 7, p. 3523–3542, 2021. Citado na página 24.
- MING, D. Cascaded framework for earthquake building damage detection combining spatial and frequency domain feature integration. *The International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences*, 2024. Citado na página 15.
- MIRANDA, P. d. S. T. *Avaliação da vulnerabilidade sísmica na realidade predial brasileira*. Dissertação (Post-graduation) — Universidade Federal do Ceará, 2010. Citado na página 14.
- MOCHALOV, V.; MOCHALOVA, A. Extraction of ionosphere parameters in ionograms using deep learning. *E3s Web of Conferences*, v. 127, p. 01004, 2019. Citado na página 24.
- MUHADI, N. A. et al. Image segmentation methods for flood monitoring system. *Water*, v. 12, n. 6, p. 1825, 2020. Citado na página 24.
- NABI, R. M.; SAEED, S.; HARRON, H. A novel approach for stock price prediction using gradient boosting machine with feature engineering GBM-wFE. *Kurdistan Journal of Applied Research*, v. 5, p. 28–48, 2020. Citado na página 27.
- NANNI, L.; LUMINI, A.; BRAHNAM, S. Neural networks for anatomical therapeutic chemical ATC classification. *Applied Computing and Informatics*, 2022. Citado na página 27.
- PEÑA, L. A. P. *Análise dos efeitos provocados por abalos sísmicos em estruturas irregulares*. Dissertação (Master's Thesis) — Universidade de Brasília, 2012. Citado na página 14.
- PROKHORENKOVA, L. et al. CatBoost: unbiased boosting with categorical features. *Advances in neural information processing systems*, v. 31, 2018. Citado na página 20.

- RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. In: SPRINGER. *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III* 18. [S.l.], 2015. p. 234–241. Citado na página 24.
- RUDNER, T. G. J. et al. Multi3net: Segmenting flooded buildings via fusion of multiresolution, multisensor, and multitemporal satellite imagery. *Proceedings of the Aaai Conference on Artificial Intelligence*, v. 33, n. 1, p. 702–709, 2019. Citado na página 24.
- SAMAT, A. et al. CatBoost for RS image classification with pseudo label support from neighbor patches-based clustering. *IEEE Geoscience and Remote Sensing Letters*, IEEE, 2020. Citado na página 20.
- SATHIANARAYANAN, B. et al. Feature-based augmentation and classification for tabular data. *CAAI Transactions on Intelligence Technology*, v. 7, p. 481–491, 2022. Citado na página 22.
- SHEN, Y. et al. Bdanet: Multiscale convolutional neural network with cross-directional attention for building damage assessment from satellite images. *IEEE Transactions on Geoscience and Remote Sensing*, IEEE, v. 60, p. 1–14, 2021. Citado 4 vezes nas páginas 28, 46, 48 e 51.
- SHI, Q.; SUGANTHAN, P. N.; KATUWAL, R. Weighting and pruning based ensemble deep random vector functional link network for tabular data classification. *Pattern Recognition*, v. 132, p. 108879, 2022. Citado na página 27.
- VIOLA, P.; JONES, M. Fast and robust classification using asymmetric adaboost and a detector cascade. *Advances in neural information processing systems*, v. 14, 2001. Citado na página 21.
- WANG, J.; LI, M. Evaluation model of regional comprehensive disaster reduction capacity under complex environment. *Journal of Environmental and Public Health*, v. 2022, 2022. Citado na página 28.
- WANG, Y. et al. A two-stage seismic damage assessment method for small, dense, and imbalanced buildings in remote sensing images. *Remote Sensing*, 2022. Citado na página 15.
- WEBER, E.; KAN, H. Building disaster damage assessment in satellite imagery with multi-temporal fusion. *arXiv preprint arXiv:2004.05525*, 2020. Citado 2 vezes nas páginas 28 e 48.
- XU, F. et al. High accuracy key feature extraction approach for the non-stationary signals measurement based on NGO-VMD noise reduction and CNN-LSTM. *Measurement Science and Technology*, v. 35, p. 015031, 2023. Citado na página 22.
- YANG, M.-D. et al. Semantic segmentation using deep learning with vegetation indices for rice lodging identification in multi-date uav visible images. *Remote Sensing*, v. 12, n. 4, p. 633, 2020. Citado na página 24.
- YANG, Z. Introduction to supervised learning. p. 92–103, 2010. Citado na página 19.
- YU, Y. Techniques and challenges of image segmentation: A review. *Electronics*, v. 12, n. 5, p. 1199, 2023. Citado na página 27.
- YUN, S. et al. CutMix: Regularization strategy to train strong classifiers with localizable features. In: *Proceedings of the IEEE/CVF international conference on computer vision*. [S.l.: s.n.], 2019. Citado na página 23.

- ZHANG, R. et al. Automatic detection of earthquake-damaged buildings by integrating uav oblique photography and infrared thermal imaging. *Remote Sensing*, 2020. Citado na página [15](#).
- ZHAO, D. et al. Experimental study and comparison of imbalance ensemble classifiers with dynamic selection strategy. *Entropy*, v. 23, p. 822, 2021. Citado na página [23](#).