

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

Leonardo Adriano Vasconcelos de Oliveira

**Microstructural Classification of Non-Oriented Electrical Steels: A
Machine Learning and Computer Vision-Based Approach**

Fortaleza – CE

10 de maio de 2026

Leonardo Adriano Vasconcelos de Oliveira

Microstructural Classification of Non-Oriented Electrical Steels: A Machine Learning and Computer Vision-Based Approach

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

Linha de Pesquisa: Inteligência Artificial

Orientador: Prof. Dr. José Daniel de Alencar Santos

Coorientador: Prof. Dr. Pedro Pedrosa Rebouças Filho

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

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

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

Fortaleza – CE

10 de maio de 2026

Dados Internacionais de Catalogação na Publicação
Instituto Federal do Ceará - IFCE
Sistema de Bibliotecas - SIBI
Ficha catalográfica elaborada pelo SIBI/IFCE, com os dados fornecidos pelo(a) autor(a)

O48m Oliveira, Leonardo Adriano Vasconcelos de.

Microstructural Classification Of Non-oriented Electrical Steels: A Machine Learning And Computer Vision-based Approach / Leonardo Adriano Vasconcelos de Oliveira. - 2026.
73 f. : il. color.

Dissertação (Mestrado) – Instituto Federal do Ceará, Mestrado em Ciência da Computação, Campus Fortaleza, 2026.

Orientação: Prof. Dr. José Daniel de Alencar Santos.

Coorientação: Prof. Dr. Pedro Pedrosa Rebouças Filho.

1. Inteligência artificial. I. Título.

CDD 004



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

Leonardo Adriano Vasconcelos de Oliveira

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

**Orientador: Prof. Dr. José Daniel de Alencar
Santos**

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

**Coorientador: Prof. Dr. Pedro Pedrosa Rebouças
Filho**

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

Profa. Dra. Suane Pires Pinheiro da Silva
Instituto Federal de Educação, Ciência e Tecnologia do
Ceará (IFCE)

Prof. Dr. Aldísio Gonçalves Medeiros
Universidade Federal do Ceará (UFC)

Fortaleza – CE
10 de maio de 2026

*Dedico este trabalho à minha esposa e aos meus pets, pelo apoio constante e incondicional
ao longo de toda a minha trajetória acadêmica.*

Agradecimentos

À minha esposa e melhor amiga, Camila Angelim Ferreira, pelo companheirismo, amor e dedicação, presentes tanto nos momentos felizes quanto nos mais desafiadores.

Ao meu orientador, Prof. Dr. José Daniel de Alencar Santos, pelas valiosas orientações e pela visão singular sobre o exercício da pesquisa acadêmica e científica, além da amizade e do companheirismo ao longo desta extensa jornada acadêmica.

Ao meu coorientador, Prof. Dr. Pedro Pedrosa Rebouças Filho, pelas contribuições e pelos relevantes insights que auxiliaram no desenvolvimento deste trabalho.

Ao Prof. Dr. Francisco Nélcio Costa Freitas, pela dedicação, apoio e importância científica atribuídos a este trabalho.

À Dra. Suane Pires Pinheiro da Silva e ao Prof. Dr. Aldísio Gonçalves Medeiros, pelas incontáveis contribuições que elevaram significativamente a qualidade desta pesquisa.

Aos meus amigos dos laboratórios LAPISCO e LPC, que me acompanharam e apoiaram ao longo desta importante jornada acadêmica.

À CAPES e ao Programa de Pós-Graduação em Ciência da Computação do IFCE, pelo suporte financeiro, por meio do auxílio ao desenvolvimento deste projeto de pesquisa.

Resumo

O contínuo aumento do consumo energético global representa um dos principais desafios da atualidade, sendo a intensificação industrial um dos fatores mais relevantes nesse cenário. Os motores elétricos são responsáveis por quase metade desse consumo, o que torna fundamental a busca por melhorias na eficiência energética de máquinas e equipamentos, com vistas à redução de perdas e à mitigação de impactos em escala global. Nesse contexto, os aços elétricos de grãos não orientados (GNO) desempenham papel essencial na fabricação de rotores e estatores, uma vez que suas propriedades microestruturais influenciam diretamente o desempenho energético desses dispositivos. Entretanto, a análise detalhada do estado microestrutural desses materiais requer procedimentos laboratoriais complexos, onerosos e demorados, o que limita a quantidade de amostras disponíveis e dificulta o desenvolvimento de modelos computacionais robustos para classificação automática ou análise preditiva. Diante desse cenário, este trabalho propõe uma metodologia automatizada para a classificação dos estados microestruturais de aços elétricos GNO a partir de imagens de microscopia óptica, considerando um contexto de severo desbalanceamento entre classes. Para mitigar essa limitação, técnicas de *data augmentation* foram empregadas, incorporando imagens sintéticas geradas por um modelo de difusão (*SinDiffusion*) e combinadas às amostras reais a fim de potencializar o desempenho na classificação automática. O estudo foi conduzido a partir de amostras de aços elétricos GNO laminados a frio por meio de laminação industrial, com reduções de espessura de 50% e 70%, submetidas posteriormente a tratamentos térmicos para crescimento de grãos em temperaturas variando de 575 a 620°C. O conjunto de dados é composto por 116 imagens reais de microscopia óptica, acrescidas de 192 imagens sintéticas, totalizando 308 amostras. A extração de atributos foi realizada por meio de descritores clássicos (*Hu Moments*, *Local Binary Pattern* e *Gray-Level Co-Occurrence Matrix*) e por métodos avançados baseados em *deep learning* (*ConvNeXt*, *Swin Transformer* e *DINOv2*), os quais foram combinados com diferentes classificadores (*Random Forest*, *k-Nearest Neighbors*, *Multilayer Perceptron*, *Extreme Learning Machine*, *Support Vector Machine* e *Least Squares Support Vector Machine*) para classificar os estados microestruturais entre laminado a frio (*Cold Rolled*), recozido parcialmente (*Partial Annealed*) e recristalização primária (*Primary Recrystallization*). Os resultados demonstraram que o classificador SVM com *kernel* linear, aliado aos atributos extraídos pelo *Swin*, apresentou acurácia média de $86,52\% \pm 9,91\%$ ao longo de 50 rodadas independentes. A análise de variância (ANOVA) das características extraídas indicou que o *Swin* apresenta baixa sensibilidade à origem dos dados, evidenciando menor influência de padrões artificiais decorrentes do processo de aumento de dados, caracterizando-se como a alternativa mais equilibrada para a aplicação proposta.

Palavras-chave: Aços Elétricos de Grãos Não-Orientados (GNO). Classificação de Estado Cristalográfico. Microscopia Óptica. Aumento de dados. SinDiffusion. *Machine Learning*. Extração de Atributos com *Deep Learning*. Visão Computacional.

Abstract

The continuous increase in global energy consumption represents one of the major challenges of modern society, with industrial intensification being one of the most relevant contributing factors in this context. Electric motors account for nearly half of this consumption, making the pursuit of improvements in the energy efficiency of machines and equipment essential to reduce losses and mitigate impacts on a global scale. In this scenario, non-oriented grain electrical steels (NOES) play a crucial role in the manufacturing of rotors and stators, as their microstructural properties directly influence the energy performance of these devices. However, a detailed analysis of the microstructural state of these materials requires complex, costly, and time-consuming laboratory procedures, which limits the number of available samples and hinders the development of robust computational models for automatic classification or predictive analysis. To address this challenge, this work proposes an automated methodology for classifying the microstructural states of NOES based on optical microscopy images, considering a scenario of severe class imbalance. To mitigate this limitation, data augmentation techniques were employed by incorporating synthetic images generated using a diffusion model (SinDiffusion) and combining them with real samples to enhance the performance of automatic classification. The study was conducted using cold-rolled NOES samples produced through industrial rolling with thickness reductions of 50% and 70%, which were subsequently subjected to heat treatments for grain growth at temperatures ranging from 575 to 620°C. The dataset consists of 116 real optical microscopy images complemented by 192 synthetic images, totaling 308 samples. Feature extraction was performed using classical descriptors (Hu Moments, Local Binary Pattern, and Gray-Level Co-Occurrence Matrix) and advanced deep learning-based methods (ConvNeXt, Swin Transformer, and DINOv2), which were combined with different classifiers (Random Forest, k-Nearest Neighbors, Multilayer Perceptron, Extreme Learning Machine, Support Vector Machine, and Least Squares Support Vector Machine) to classify the microstructural states into cold rolled, partial annealed, and primary recrystallization. The results demonstrated that the SVM classifier with a linear kernel, combined with features extracted by the Swin Transformer, achieved an average accuracy of $86.52\% \pm 9.91\%$ over 50 independent runs. The analysis of variance (ANOVA) of the extracted features indicated that the Swin Transformer exhibits low sensitivity to data origin, reflecting reduced influence of artificial patterns introduced by data augmentation, characterizing it as the most balanced alternative for the proposed application.

Keywords: Non-Oriented Electrical Steels (NOES). Crystallographic State Classification. Optical Microscopy. Data Augmentation. Diffusion Models. SinDiffusion. Machine Learning. Deep Learning Feature Extraction. Computer Vision.

Lista de ilustrações

Figura 1 – Exemplo do fluxo de treinamento dos modelos de difusão, adicionando ruído e retirando ruído ao longo de várias etapas. Adaptado de (Croitoru <i>et al.</i> , 2023).	23
Figura 2 – Comparação do modelo de difusão e o <i>SinDiffusion</i> , onde o campo de recepção do modelo clássico cobre toda a imagem, enquanto o do <i>SinDiffusion</i> aprende a estimar o ruído em <i>patches</i> , gerando maior variações de imagens. Adaptado de (Wang <i>et al.</i> , 2025).	24
Figura 3 – Imagens de microscopia óptica de aços elétricos GNO laminados a frio, com redução de espessura de 50% e ampliação de 500×.	27
Figura 4 – Metodologia da abordagem proposta.	30
Figura 5 – Microscópio Óptico utilizado para aquisição das imagens digitais através da fotomicrografia, destacando o material de fixação e a amostra a ela acoplada.	31
Figura 6 – Análise comparativa da estrutura cristalográfica das 12 imagens sintéticas geradas pelo <i>SinDiffusion</i> a partir do treinamento da imagem digital de microscopia óptica da amostra laminada a frio do tipo de laminação 50% com zoom de 500X.	33
Figura 7 – Análise comparativa da estrutura cristalográfica das 12 imagens sintéticas geradas pelo <i>SinDiffusion</i> a partir do treinamento da imagem digital de microscopia óptica da amostra da classe CR (Laminado a frio) do tipo de laminação 50% com zoom de 500X, após a aplicação da técnica de <i>histogram matching</i> .	35
Figura 8 – Análise comparativa da estrutura cristalográfica das 12 imagens sintéticas geradas pelo <i>SinDiffusion</i> a partir do treinamento da imagem digital de microscopia óptica da amostra da classe CR (Laminado a frio) do tipo de laminação 50% com zoom de 500X, após a aplicação da técnica de <i>histogram matching</i> e <i>data augmentation</i> clássica.	37
Figura 9 – Representação bidimensional por t-SNE dos atributos extraídos por técnicas clássicas para os dados reais (Figuras. 9a, 9d e 9g), após o aumento de dados sintéticos utilizando o <i>SinDiffusion</i> com ajuste de histograma (Figuras. 9b, 9e e 9h) e, por fim, após a aplicação das técnicas clássicas de aumento de dados (Figuras. 9c, 9f e 9i).	49

Figura 10 – Representação bidimensional por t-SNE dos atributos extraídos por técnicas baseadas em <i>deep learning</i> para os dados reais (Figuras. 10a, 10d e 10g), após o aumento de dados sintéticos utilizando o <i>SinDiffusion</i> com ajuste de histograma (Figuras. 10b, 10e e 10h) e, por fim, após a aplicação das técnicas clássicas de aumento de dados (Figuras. 10c, 10f e 10i)..	50
Figura 11 – Análise ANOVA em histograma dos atributos de cada extrator para distribuição estatística entre classes minoritárias (CR e PR), entre origem (real e artificial) e, entre classe e origem.	52
Figura 12 – Comparação, por meio de boxplots, da acurácia dos classificadores na fase de teste antes (boxplots verdes) e após (boxplots laranja) a aplicação de técnicas padrão de aumento de dados, ao longo de 50 execuções independentes para cada um dos extratores de características.	54
Figura 13 – Matrizes de confusão exibidas como <i>heatmaps</i> para as amostras de teste, agregando os resultados de 50 execuções independentes dos classificadores de melhor desempenho, utilizando os extratores de atributos LBP, ConvNeXt, Swin e DINOv2.	63

Lista de tabelas

Tabela 1 – Características visuais e magnéticas de cada estado microestrutural dos aços elétricos GNO.	28
Tabela 2 – Composição química do aço elétrico GNO utilizada.	31
Tabela 3 – Quantidade de amostras de imagens digitais de microscopia por temperatura e classe.	32
Tabela 4 – Média e desvio padrão das fases de treinamento e teste para cada estimador, bem como os valores de PPV, sensibilidade (Se) e F-score referentes às amostras de teste, calculados ao longo de 50 execuções independentes para os extratores de atributos clássicos.	57
Tabela 5 – Média e desvio padrão das fases de treinamento e teste para cada estimador, bem como os valores de PPV, sensibilidade (Se) e F-score referentes às amostras de teste, calculados ao longo de 50 execuções independentes para os extratores de atributos baseados em <i>deep learning</i>	58
Tabela 6 – Escalabilidade computacional dos extratores de atributos avaliados.	60

Lista de abreviaturas e siglas

Acc	<i>Accuracy</i>
CNN	<i>Convolutional Neural Networks</i>
CR	<i>Cold Rolled</i>
DINO	<i>Self-Distillation with no labels</i>
DL	Direção de Laminação
DM	<i>Diffusion Model</i>
ELM	<i>Extreme Learning Machine</i>
GLCM	<i>Gray-Level Co-Occurrence Matrix</i>
GNO	Grão Não Orientado
IA	Inteligência Artificial
KNN	<i>k-Nearest Neighbors</i>
LBP	<i>Local Binary Pattern</i>
LSSVM	<i>Least-Squares Support Vector Machine</i>
MLP	<i>Multilayer Perceptron</i>
PA	<i>Partial Annealed</i>
PR	<i>Primary Recrystallization</i>
RBF	<i>Radial Basis Function</i>
RF	<i>Random Forest</i>
Swin	<i>Shifted Window</i>
t-SNE	<i>(t-Distributed Stochastic Neighbor Embedding</i>
ViT	<i>Vision Transformer</i>

Sumário

1	INTRODUÇÃO	15
1.1	Objetivo Geral	16
1.2	Objetivos específicos	17
1.3	Organização do Texto	17
2	REVISÃO DE LITERATURA	19
2.1	Aplicações de Aprendizado de Máquina na Ciência dos Materiais e em Aços Elétricos GNO	19
2.2	Geração de Dados Sintéticos e Técnicas de Data Augmentation	20
2.2.1	Data augmentation clássico	21
2.2.2	Modelos generativos aplicados à visão computacional	21
2.2.3	Modelos de difusão	22
2.2.4	SinDiffusion	22
3	FUNDAMENTOS DOS AÇOS ELÉTRICOS DE GRÃOS NÃO ORIENTADOS E METALOGRAFIA	25
3.1	Aços Elétricos de Grão Não Orientado	25
3.2	Metalografia	28
4	METODOLOGIA PROPOSTA	30
4.1	Aquisição de Dados	31
4.2	Definição de Classe	32
4.3	Aumento de Dados Sintéticos	32
4.4	Harmonização Visual	34
4.5	<i>Augmentation</i> Clássico	35
4.6	Extratores de Atributos	37
4.6.1	Extratores de Atributos Clássicos	38
4.6.1.1	Hu Moments	38
4.6.1.2	Local Binary Pattern	38
4.6.1.3	Gray-Level Co-Occurrence Matrix	39
4.6.2	Extratores de Atributos baseados em <i>Deep Learning</i>	40
4.6.2.1	ConvNeXt	40
4.6.2.2	Swin Transformer	41
4.6.2.3	DINOv2	41
4.7	Modelos para Classificação	42
4.7.1	Support Vector Machine	42

4.7.2	Least Squares Support Vector Machine	43
4.7.3	Multilayer Perceptron	43
4.7.4	Extreme Learning Machine	43
4.7.5	k-Nearest Neighbors	44
4.7.6	Random Forest	44
5	RESULTADOS E DISCUSSÃO	46
5.1	Data Augmentation	46
5.1.1	Cenário de Criação de Dados Sintéticos	46
5.1.2	Configuração de Simulação dos Extratores de Atributos	47
5.1.3	Representação Bidimensional dos Atributos	48
5.1.4	Análise Estatística dos Atributos	51
5.1.5	Resultados de Simulação	53
5.2	Simulações dos Classificadores	55
5.2.1	Matrizes de Confusão	61
6	CONCLUSÃO	64
6.1	Limitações do Trabalho	65
	REFERÊNCIAS	67

1 Introdução

O crescimento contínuo do consumo global de energia representa um dos desafios mais críticos da era contemporânea, com projeções que indicam um aumento de aproximadamente 50% até o ano de 2050 (Ghasemian *et al.*, 2024). Essa tendência ascendente está diretamente associada à intensificação das emissões de gases de efeito estufa (Abbass *et al.*, 2022), bem como ao agravamento de impactos socioeconômicos, tais como o aumento dos custos de produção e o maior risco de ampliação da pobreza energética em escala global (Guan *et al.*, 2023; Simões; Leder, 2022).

A expansão da automação industrial e a ampla adoção de máquinas elétricas, sistemas robóticos e tecnologias digitais têm intensificado a demanda por energia, gerando um efeito de rebote que acaba por compensar parte dos ganhos de produtividade (Deng; Zou; Guang, 2025). Como consequência, o consumo energético passou a representar um dos principais custos operacionais no setor industrial (Yang; Zhou; Gao, 2025; Opoku *et al.*, 2022), o que reforça a necessidade de políticas e tecnologias voltadas à melhoria da eficiência energética e à garantia da sustentabilidade dos processos produtivos (Dimian *et al.*, 2025).

Nesse contexto, os aços elétricos desempenham um papel central, uma vez que são materiais ferromagnéticos amplamente utilizados na fabricação de núcleos de motores elétricos e transformadores (Li *et al.*, 2021; Karami; Butler; Tamimi, 2024). Esses materiais apresentam uma microestrutura, isto é, a organização interna dos grãos e das fases metálicas observável em escala microscópica, a qual governa suas propriedades magnéticas e mecânicas. Os aços elétricos são geralmente divididos em dois grupos principais: os aços de grão orientado (GO), cuja textura cristalográfica do tipo Goss proporciona elevada permeabilidade magnética e menores perdas energéticas (Karami; Butler; Tamimi, 2024; Hayakawa, 2022; Jain *et al.*, 2022), e os aços de grão não orientado (GNO), que apresentam uma microestrutura isotrópica¹ e são amplamente empregados em máquinas rotativas, como motores e geradores (Ahn *et al.*, 2020).

A microestrutura dos aços elétricos, moldada pelos processos de laminação e recozimento, exerce influência direta sobre suas propriedades magnéticas e, consequentemente, sobre sua eficiência em aplicações industriais (Li; Wang; Liu, 2023). Em particular, o estado cristalográfico descreve o grau de orientação e a distribuição dos grãos, refletindo as etapas metalúrgicas que determinam o desempenho magnético final do material. Dessa forma, a caracterização precisa dessas estruturas é fundamental para a otimização da textura cristalográfica e para a minimização de perdas nos processos de conversão eletromecânica

¹ Microestrutura isotrópica indica distribuição aproximadamente aleatória das orientações dos grãos, sem textura cristalográfica predominante.

de energia, contribuindo, em última instância, para o aumento da eficiência energética e para a sustentabilidade dos sistemas de manufatura (Tamimi *et al.*, 2022).

Entre as técnicas mais comumente empregadas para essa caracterização, destaca-se a microscopia óptica, uma vez que permite a observação, com elevada precisão, dos grãos e das orientações cristalográficas responsáveis pelo desempenho magnético dos aços elétricos (Li; Wang; Liu, 2023; Li *et al.*, 2021; Jain *et al.*, 2022). No entanto, apesar de sua relevância científica e industrial, esse processo ainda se mostra lento, oneroso e fortemente dependente da avaliação humana, o que limita a padronização e a reprodutibilidade das análises (Ivo *et al.*, 2020; Gola *et al.*, 2018).

Essas restrições tornam a avaliação microestrutural particularmente desafiadora em escala industrial, uma vez que cada amostra deve ser submetida a múltiplas etapas laboratoriais, tais como corte, polimento e ataque químico, o que acarreta aumento do tempo de processamento e dos custos associados (Rebouças Filho *et al.*, 2019). Além disso, a obtenção de imagens microscópicas de alta qualidade em quantidade suficiente é dificultada pelo desgaste dos equipamentos e pela variação microestrutural ao longo da espessura do material, o que limita a representatividade dos conjuntos de dados experimentais (Durmaz *et al.*, 2021; Cui *et al.*, 2023).

Nesse contexto de restrições operacionais, elevado custo e limitada escalabilidade das análises convencionais, este trabalho propõe uma abordagem automatizada para a classificação do estado cristalográfico de aços elétricos GNO a partir de imagens de microscopia óptica, por meio da aplicação de técnicas de aprendizado de máquina voltadas à caracterização microestrutural. A utilização de ferramentas computacionais sobre fotomicrografias possibilita a discriminação entre diferentes estados microestruturais, os quais influenciam diretamente o desempenho eletromagnético do material e sua adequação a distintas aplicações industriais. Dessa forma, busca-se mitigar a subjetividade inerente à análise microestrutural convencional, reduzindo o tempo de avaliação e a ambiguidade entre especialistas.

1.1 Objetivo Geral

Desenvolver e avaliar uma metodologia automatizada para a classificação do estado cristalográfico de aços elétricos GNO entre laminado a frio, recristalização parcial e recristalização primária, a partir de imagens de microscopia óptica, empregando técnicas de aprendizado de máquina.

A abordagem proposta visa discriminar diferentes estados microestruturais, incluindo condições parcialmente recristalizadas, considerando cenários caracterizados por desbalanceamento de classes e limitada disponibilidade de dados experimentais.

1.2 Objetivos específicos

De forma mais específica, a dissertação possui os seguintes objetivos:

- a) Aplicar modelo de difusão especializado em imagens para a geração de imagens sintéticas representativas das classes minoritárias, visando mitigar o desbalanceamento severo entre classes;
- b) Investigar técnicas clássicas de extração de atributos aplicadas a imagens de microscopia óptica de aços elétricos GNO;
- c) Avaliar técnicas de extração de atributos baseadas em *deep learning*, utilizando *transfer learning*, a partir de imagens de microscopia óptica de aços GNO;
- d) Avaliar o desempenho de diferentes classificadores combinados com os extratores de atributos na classificação dos estados microestruturais dos aços GNO entre laminado a frio, recristalização parcial e primária, antes e após a aplicação de técnicas de *data augmentation*;
- e) Comparar o desempenho de abordagens baseadas em extratores clássicos e em *deep learning*, em conjunto com os classificadores, para identificar a metodologia mais adequada à classificação automática do estado microestrutural dos aços GNO;

1.3 Organização do Texto

Esta dissertação está organizada em oito capítulos, além das referências bibliográficas e apêndices, conforme a seguinte estrutura:

O Capítulo 2 apresenta a revisão da literatura, abordando os principais estudos que investigam a aplicação de técnicas de *machine learning* na classificação automática de aços elétricos GNO, com ênfase em métodos, abordagens e desafios associados a esse problema.

O Capítulo 3 apresenta os fundamentos dos aços elétricos GNO e os princípios da metalografia. Nesse capítulo, são discutidas as características dos diferentes estados microestruturais dos aços GNO, bem como as etapas do processo metalográfico que conduzem à aquisição das imagens de microscopia óptica utilizadas neste trabalho.

O Capítulo 4 apresenta a metodologia proposta neste trabalho, descrevendo de forma detalhada todas as etapas desenvolvidas, desde a aquisição e preparação dos dados até os procedimentos de extração de atributos e classificação das microestruturas dos aços elétricos GNO. Inicialmente, são abordadas a definição das classes, a geração de imagens sintéticas por meio de modelos de difusão, a harmonização visual dos dados e a aplicação de técnicas clássicas de *data augmentation*. Em seguida, são apresentadas

as técnicas empregadas para a extração de atributos das imagens de microscopia óptica, contemplando tanto métodos clássicos, como *Hu Moments*, *Local Binary Pattern* (LBP) e *Gray-Level Co-Occurrence Matrix* (GLCM), quanto arquiteturas de *deep learning* baseadas em redes neurais convolucionais e *Transformers*, incluindo *ConvNeXt*, *Swin Transformer* e *DINOv2*. Por fim, são descritos os modelos utilizados para a classificação dos aços elétricos GNO, abrangendo diferentes abordagens de *machine learning*, como *Multilayer Perceptron* (MLP), *Extreme Learning Machine* (ELM), *Support Vector Machine* (SVM), *Least Squares Support Vector Machine* (LSSVM), *k-Nearest Neighbors* (KNN) e *Random Forest* (RF), permitindo uma análise comparativa entre técnicas de distintas naturezas metodológicas.

No Capítulo 5, são apresentados e discutidos os resultados das simulações realizadas. Inicialmente, são analisados os diferentes cenários de simulação relacionados aos extratores de atributos, incluindo uma avaliação preliminar do desempenho da classificação após a incorporação de dados sintéticos gerados pelo modelo *SinDiffusion* e, posteriormente, após a aplicação de técnicas clássicas de *data augmentation*. O objetivo dessa análise é investigar a variabilidade das amostras sintéticas e seu impacto no desempenho dos modelos de classificação. Em seguida, são apresentados e discutidos os resultados das simulações dos classificadores, avaliando seus desempenhos em diferentes cenários, tanto em relação aos extratores de atributos clássicos quanto aos métodos avançados baseados em *deep learning*.

Por fim, o Capítulo 6 apresenta as conclusões do trabalho, destacando os principais resultados obtidos quanto ao desempenho da metodologia proposta, bem como uma discussão sobre perspectivas de trabalhos futuros com o objetivo de aprimorar a classificação automática dos estados microestruturais dos aços elétricos GNO.

2 Revisão de literatura

Este capítulo apresenta uma revisão da literatura sobre a aplicação de técnicas de aprendizado de máquina e de visão computacional na análise microestrutural de aços elétricos GNO. Apesar do crescimento significativo do uso de métodos baseados em inteligência artificial na ciência dos materiais, observa-se que o número de trabalhos que abordam diretamente a classificação microestrutural desses aços com base em dados experimentais é relativamente limitado. Dessa forma, a revisão é direcionada a estudos que exploram imagens de microscopia óptica, dados cristalográficos e propriedades magnéticas como fontes de informação para a caracterização microestrutural. Posteriormente, é abordada uma breve revisão sobre metodologias de *data augmentation* clássicas e baseadas em modelos generativos e de difusão, de modo a contextualizar o estado da arte em relação à geração de amostras sintéticas.

O capítulo está organizado de modo a apresentar inicialmente as principais abordagens baseadas em aprendizado de máquina aplicadas a aços GNO, seguido da discussão de diferentes fontes de dados utilizadas na literatura, culminando na identificação das limitações e lacunas que motivam o desenvolvimento da metodologia proposta neste trabalho. Por fim, o capítulo aborda a geração de dados sintéticos, descrevendo as principais técnicas clássicas e baseadas em *deep learning*, bem como a adoção da técnica de geração de amostras sintéticas para mitigar o problema de desbalanceamento das classes.

2.1 Aplicações de Aprendizado de Máquina na Ciência dos Materiais e em Aços Elétricos GNO

Nos últimos anos, os avanços em visão computacional e inteligência artificial têm impulsionado a adoção de técnicas de aprendizado de máquina e aprendizado profundo na área de ciência dos materiais (Cooper *et al.*, 2025; Guo *et al.*, 2022; Azimi *et al.*, 2018; Mao *et al.*, 2025; Hiremath *et al.*, 2025). No entanto, apenas um número limitado de estudos tem explorado essas abordagens especificamente para a análise e classificação de aços elétricos GNO. Rebouças Filho *et al.* (Rebouças Filho *et al.*, 2019) propuseram uma metodologia que combina extração de características com algoritmos de aprendizado de máquina para a classificação de aços GNO a partir de imagens de microscopia óptica. De forma complementar, Ivo *et al.* (Ivo *et al.*, 2020) empregaram diferentes modelos de redes neurais convolucionais previamente treinados para a extração de características dessas imagens, as quais foram posteriormente integradas a técnicas de aprendizado de máquina com o objetivo de classificar estados microestruturais, incluindo condições de material

laminado a frio, crescimento normal de grão e crescimento anormal de grão.

Outros autores ampliaram essas investigações ao explorar diferentes fontes de dados. Ivo *et al.* (Ivo *et al.*, 2019) utilizaram informações provenientes da Função de Distribuição de Orientações Cristalográficas (Crystallographic Orientation Distribution Function – CODF), combinadas com extratores clássicos de características e algoritmos de aprendizado de máquina, com o objetivo de distinguir diferentes estados microestruturais. De forma complementar, Duarte *et al.* (Duarte *et al.*, 2021) correlacionaram propriedades magnéticas obtidas a partir de curvas de histerese com as respectivas condições microestruturais, aplicando o método SMOTE para aumentar o volume de amostras experimentais. Apesar desses avanços, nenhum dos estudos existentes abordou a classificação de estados microestruturais intermediários, em particular aqueles parcialmente recristalizados, nem tratou de maneira adequada o severo desbalanceamento de classes comumente encontrado em conjuntos de dados industriais reais.

Apesar dos avanços apresentados, observa-se que os estudos existentes concentram-se majoritariamente na classificação de estados microestruturais bem definidos, desconsiderando a presença de estados intermediários, em particular aqueles parcialmente recristalizados, os quais são frequentes em ambientes industriais reais. Além disso, as abordagens propostas não tratam do severo desbalanceamento entre classes, característica comum em conjuntos de dados industriais, o que compromete sua aplicabilidade prática, uma vez que frequentemente demandam grandes volumes de dados para o adequado treinamento dos modelos desenvolvidos. Essas limitações evidenciam uma lacuna relevante na literatura, indicando a necessidade de metodologias capazes de lidar simultaneamente com a complexidade microestrutural dos aços elétricos GNO e com as restrições impostas por dados oriundos de processos industriais reais.

2.2 Geração de Dados Sintéticos e Técnicas de Data Augmentation

Diante de cenários eminentes de baixa quantidade de dados, principalmente relacionada ao custo da aquisição de grandes quantidades de amostras, a área de geração de dados sintéticos juntamente com técnicas de *data augmentation* emerge como alternativa para contornar tais problemas. Tais metodologias foram aplicadas neste trabalho, utilizando uma combinação de técnicas de *data augmentation* clássicas com um modelo generativo baseado em difusão. Portanto, esta seção apresenta uma breve descrição das técnicas utilizadas para gerar amostras sintéticas dos aços elétricos GNO e mitigar o problema de desbalanceamento severo entre as classes.

2.2.1 Data augmentation clássico

As técnicas clássicas de *data augmentation* consistem em aplicações de transformações simples aplicadas diretamente às imagens originais com o objetivo de aumentar a variabilidade dos dados, buscando melhorar a capacidade de generalização dos algoritmos de machine learning em diferentes tarefas, sejam em classificação, segmentação e detecção de objetos. Essas transformações normalmente abrangem alterações geométricas, como *flipping*, rotação e recorte, bem como ajustes fotométricos como mudança de espaço de cor e aplicação de filtros para produzir efeitos de *blurring*, que preservam a imagem sem gerar uma nova classe de dados (Shorten; Khoshgoftaar, 2019; Yang *et al.*, 2023).

Em estudos de revisão, técnicas como essas têm sido amplamente empregadas para ampliar conjuntos de dados de imagem e reduzir o risco de *overfitting* em tarefas de reconhecimento e classificação visual, ao gerar versões ligeiramente modificadas dos exemplos originais, sem a necessidade de aquisição de novas amostras (Shorten; Khoshgoftaar, 2019; Yang *et al.*, 2023).

Em contrapartida, embora tenham ampla aplicação e baixo custo computacional, as técnicas de *data augmentation* clássicas apresentam limitações. As transformações geométricas e fotométricas não são capazes de introduzir novas variações semânticas ou estruturais nos dados. Consequentemente, tais amostras sintéticas tendem a ter uma alta correlação com as imagens originais, o que pode restringir a variabilidade diante de tarefas como classificação, limitando a capacidade de generalização dos estimadores (Kumar *et al.*, 2024; Mumuni; Mumuni, 2022).

2.2.2 Modelos generativos aplicados à visão computacional

Diante das limitações das técnicas de *data augmentation* clássicas, ao longo dos últimos anos surgiram técnicas baseadas em *deep learning* capazes de gerar amostras sintéticas com variações semânticas e estruturais, mitigando as limitações inerentes aos métodos clássicos (Mumuni; Mumuni, 2022; Shorten; Khoshgoftaar, 2019).

Modelos generativos baseados em *deep learning* têm como objetivo aprender a distribuição estatística subjacente às imagens treinadas, permitindo uma síntese de novas amostras que preservam características visuais, estruturais e semânticas do conjunto inicial. Tais modelos buscam capturar padrões complexos de textura, forma e organização espacial das imagens, indo além de simples transformações geométricas e fotométricas. Assim, diferentemente das técnicas clássicas de *data augmentation*, os modelos generativos possuem a capacidade de capturar relações não lineares de alto nível, o que os permite gerar imagens sintéticas que apresentam variações estruturais plausíveis e semanticamente consistentes (Gao *et al.*, 2023).

Diante da variedade de modelos generativos desenvolvidos até então, dentre as

abordagens mais exploradas, destacam-se os *autoencoders* variacionais (VAEs - *Variational Autoencoders*), as redes adversárias generativas (GANs - *Generative Adversarial Networks*) e, mais recentemente, os modelos de difusão, que têm demonstrado elevado potencial na geração de dados realistas para tarefas de classificação, segmentação e análise microestrutural (Yazdani *et al.*, 2025).

2.2.3 Modelos de difusão

Diante dos modelos mais utilizados no estado da arte para geração de imagens sintéticas com *deep learning*, as GANs apresentam instabilidades durante o treinamento, frequentemente resultando na geração de amostras com baixa diversidade (Gui *et al.*, 2021). Em contrapartida, os VAEs tendem a produzir imagens com perda de nitidez e excesso de suavização (Zhao; Song; Ermon, 2017).

Mais recentemente, os modelos de difusão têm se destacado por sua maior capacidade de diversidade amostral, aliada a uma estrutura de treinamento mais estável e à elevada fidelidade em relação às amostras originais (Croitoru *et al.*, 2023). Com base nessas características, o presente trabalho adotou o modelo de difusão para a geração de dados sintéticos a partir de imagens de microscopia óptica.

Os modelos de difusão tratam as imagens de treinamento por meio de um processo sequencial de adição gradual de ruído. Nesse contexto, uma rede neural profunda é treinada para aprender a reverter tal processo, removendo iterativamente o ruído inserido e reconstruindo as estruturas latentes dos dados.

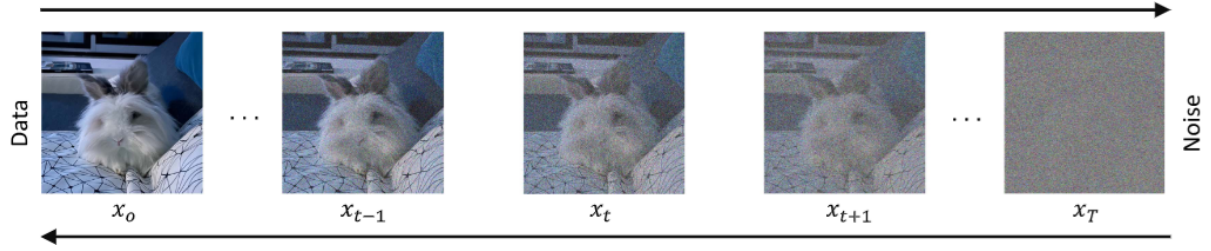
O funcionamento desses modelos baseia-se em uma Cadeia de Markov, caracterizada como um processo estocástico no qual o estado futuro depende apenas do estado atual. Assim, cada imagem ruidosa em um passo t depende exclusivamente da imagem obtida no passo imediatamente anterior $t - 1$. Esse mecanismo é construído a partir de uma sequência de pequenas corrupções gaussianas, aplicadas de forma incremental.

No processo de *forward*, adiciona-se ruído gaussiano gradualmente à imagem ao longo de T passos, conduzindo a distribuição dos dados a um ruído gaussiano isotrópico. Em contrapartida, o processo reverso, comumente denominado *denoising*, realiza a remoção progressiva desse ruído por meio de uma rede neural treinada para estimar as componentes estocásticas adicionadas, possibilitando a reconstrução, ou geração, de amostras realistas a partir de ruído puro. A Figura 1 apresenta este processo gradual de adição de ruído na fase *forward* e remoção do mesmo no sentido *reverse*.

2.2.4 SinDiffusion

Os modelos de difusão e demais citados precisam de uma quantidade significativa de dados para o seu processo de treinamento ser eficiente, de modo a gerar imagens sintéticas

Figura 1 – Exemplo do fluxo de treinamento dos modelos de difusão, adicionando ruído e retirando ruído ao longo de várias etapas. Adaptado de (Croitoru *et al.*, 2023).



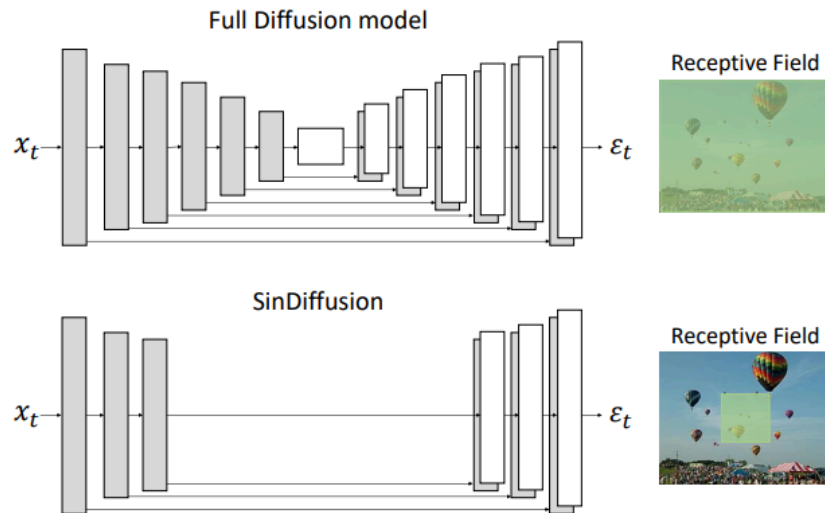
minimamente representativas do problema. Em contrapartida, o presente trabalho possui um cenário totalmente contrário à tais condições, com poucas amostras de imagens de microscopia óptica. Diante deste cenário, buscou-se modelos de difusão capazes de treinar com poucas amostras, chegando à adoção do modelo recentemente desenvolvido, chamado de *SinDiffusion*.

Proposto por Weilun Wang *et al.* (Wang *et al.*, 2025), o *SinDiffusion* é uma abordagem recente e inovadora em *deep learning* que emprega modelos de difusão para capturar as distribuições internas de *patches* provenientes de uma única imagem original. Essa estratégia demonstrou desempenho superior às abordagens anteriores baseadas em GANs, tanto em termos de qualidade quanto de diversidade das amostras geradas (Wang *et al.*, 2025).

As principais inovações do *SinDiffusion* concentram-se no treinamento em uma única escala e no uso de um campo receptivo em nível de *patch*. Essa abordagem elimina a necessidade de pirâmides de imagens e múltiplos estágios, evitando a propagação de erros sutis ao longo das escalas. Diferentemente dos modelos de difusão tradicionais, cujo campo receptivo abrange a imagem inteira, favorecendo a memorização exata da amostra de treino, o *SinDiffusion* restringe o campo receptivo para que o modelo aprenda distribuições estatísticas locais de *patches*, em vez do conteúdo global. Como resultado, aumenta-se a diversidade das amostras geradas, preservando ao mesmo tempo a estrutura visual da imagem. A Figura 2 mostra um comparativo entre os modelos de difusão clássicos e o *SinDiffusion*, ressaltando o campo receptivo reduzido.

Até o momento da elaboração deste trabalho, não foram identificados na literatura estudos que empreguem o modelo de difusão *SinDiffusion* para a geração de amostras sintéticas de imagens de microscopia óptica de aços elétricos de grão não orientado (GNO), com o objetivo de mitigar problemas severos de desbalanceamento de dados.

Figura 2 – Comparação do modelo de difusão e o *SinDiffusion*, onde o campo de recepção do modelo clássico cobre toda a imagem, enquanto o do *SinDiffusion* aprende a estimar o ruído em *patches*, gerando maior variações de imagens. Adaptado de (Wang *et al.*, 2025).



3 Fundamentos dos Aços Elétricos de Grãos Não Orientados e Metalografia

Neste capítulo são apresentados os fundamentos teóricos relacionados aos aços elétricos GNO e à metalografia necessários para a compreensão do desenvolvimento deste trabalho. Inicialmente, são discutidos os princípios básicos da microestrutura dos aços elétricos, com ênfase nos estágios da microestrutura e suas características eletromecânicas. Em seguida, são abordados os conceitos fundamentais do ensaio metalográfico, incluindo as etapas de preparação das amostras e os aspectos relevantes da microscopia óptica, estabelecendo a base conceitual para as análises microestruturais e para a aplicação das metodologias propostas nos capítulos subsequentes.

Nesse contexto, os aços elétricos, em especial os aços elétricos GNO, apresentam propriedades magnéticas fortemente dependentes de sua microestrutura e das etapas de processamento às quais são submetidos (Chu; Zhou; Mao, 2025; Winter *et al.*, 2024). A compreensão da influência de processos como a laminação a frio, a recristalização parcial e a recristalização primária sobre o comportamento elétrico e mecânico desses materiais requer a revisão dos fundamentos associados à sua fabricação e à evolução de sua estrutura interna. Nesse contexto, este capítulo aborda os principais estados microestruturais observados nesses aços, bem como os fundamentos do ensaio metalográfico, técnica amplamente empregada na análise e caracterização de ligas metálicas (Scott, 2019). Esses conceitos fornecem a base necessária para a interpretação das análises microestruturais e para o entendimento das abordagens computacionais desenvolvidas ao longo deste trabalho.

3.1 Aços Elétricos de Grão Não Orientado

Os métodos de produção dos aços elétricos seguem, em linhas gerais, a rota convencional de processamento de chapas de aço (Moses *et al.*, 2019). Inicialmente, o aço líquido é obtido na etapa de aciaria, por meio de fornos a oxigênio básico ou fornos elétricos a arco. Nessa fase, a composição química do material é cuidadosamente ajustada, com especial atenção ao teor de silício e à adição controlada de elementos de liga, como manganês, alumínio e fósforo, visando ao aumento da resistividade elétrica e à redução das perdas magnéticas (Moses *et al.*, 2019).

Após o ajuste da composição química, o aço líquido passa pelo processo de solidificação, seguido do lingotamento contínuo. Nessa etapa, o material é resfriado e solidificado na forma de placas, tarugos ou blocos, dando origem à microestrutura inicial do material (Moses *et al.*, 2019). Essa microestrutura serve como base para as etapas subsequentes

de laminação e recozimento, as quais desempenham papel fundamental na definição da microestrutura final e da textura cristalográfica dos aços elétricos.

Após as etapas iniciais de produção e solidificação, a laminação a frio (cold rolling – CR) constitui o primeiro processo responsável por modificar significativamente a microestrutura dos aços elétricos. Esse processo introduz deformação plástica severa, promovendo o alongamento dos grãos e o aumento substancial da densidade de discordâncias, fatores que influenciam diretamente a textura cristalográfica e as propriedades mecânicas do material (Neundlinger *et al.*, 2024; Sistaninia *et al.*, 2025). Durante a laminação a frio, a chapa de aço é deformada plasticamente ao passar entre dois cilindros metálicos que giram em sentidos opostos, à temperatura ambiente, resultando em uma redução significativa de espessura. A microestrutura deformada obtida nessa etapa atua como um estado precursor essencial para os tratamentos térmicos subsequentes, nos quais os fenômenos de recuperação e recristalização governam o desenvolvimento do tamanho de grão e da textura cristalográfica, parâmetros decisivos para o desempenho magnético dos aços elétricos empregados em dispositivos eletromagnéticos, como transformadores e motores elétricos.

Embora a laminação a frio proporcione a redução da espessura da chapa, esse processo também promove a subdivisão dos grãos e a introdução de tensões residuais, modificando as curvas de histerese e aumentando as perdas magnéticas do material (Cui *et al.*, 2018; Karami; Butler; Tamimi, 2024). Esses efeitos são indesejáveis para diversas aplicações, especialmente em máquinas elétricas que demandam elevada eficiência energética. Para mitigar tais efeitos, aplica-se subsequentemente um tratamento térmico conhecido como recristalização primária (ou recozimento), cujo objetivo é eliminar os defeitos introduzidos pela deformação plástica, reduzir as tensões internas e minimizar as perdas por histerese.¹

Durante a recristalização primária dos aços elétricos, ocorrem três processos principais: o rearranjo e a aniquilação parcial das discordâncias, o crescimento de grãos recristalizados no interior da matriz deformada e a competição entre os grãos recém-formados pelo crescimento (Ayad *et al.*, 2021). A microestrutura resultante é composta por grãos equiaxiais, livres de deformação, o que torna o material menos suscetível à fratura, reduz as perdas magnéticas e contribui para o aumento da eficiência energética do aço.

A Figura 3 apresenta imagens de microscopia óptica com ampliação de 500×, ilustrando as estruturas de grãos que compõem a microestrutura de um aço elétrico não orientado (GNO) laminado a frio, com redução de espessura de 50%, observadas antes (Figura 3a), durante (Figura 3b) e após a recristalização primária (Figura 3c). Na imagem

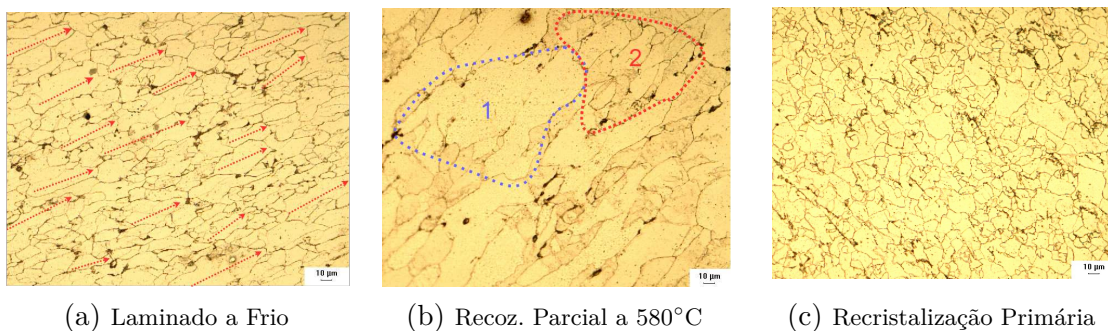
¹ A curva de histerese representa a relação entre o campo magnético aplicado e a densidade de fluxo resultante em um material ferromagnético, evidenciando perdas energéticas durante ciclos de magnetização e desmagnetização.

do material laminado a frio (Figura 3a), setas vermelhas destacam o padrão de grãos alongados orientados preferencialmente em uma direção específica, como consequência direta do processo de laminação.

Na etapa de recristalização parcial a 580°C , ilustrada na Figura 3b, inicia-se a reorganização parcial da microestrutura deformada. Nesse estágio, as discordâncias passam por um rearranjo parcial, resultando em uma microestrutura mista, composta tanto por grãos alongados remanescentes do processo de laminação a frio quanto por novos grãos recristalizados, caracterizando o início da recristalização primária. Na Figura 3b, duas regiões são destacadas pelas cores azul e vermelha. A região em azul (1) corresponde a um grão completamente recristalizado, enquanto a região em vermelho (2) indica grãos que permanecem associados à estrutura deformada proveniente da laminação a frio.

Por fim, a Figura 3c apresenta a microestrutura totalmente recristalizada. Após o recozimento completo, todos os grãos deformados são transformados em grãos equiaxiais, livres de tensões internas. Nesse estágio, a força motriz para o crescimento dos grãos é reduzida, iniciando-se um processo de competição espacial entre eles, o que limita seu crescimento excessivo e resulta em tamanhos de grão menores e mais uniformes. Como consequência, os aços elétricos no estado de recristalização primária exibem propriedades mecânicas e magnéticas mais homogêneas em todas as direções, tornando-se adequados para aplicações como máquinas elétricas, núcleos de estatores de turbogeradores, geradores e máquinas síncronas de ímãs permanentes (Karami; Butler; Tamimi, 2024).

Figura 3 – Imagens de microscopia óptica de aços elétricos GNO laminados a frio, com redução de espessura de 50% e ampliação de $500\times$.



Fonte: Adaptado de Freitas (Freitas, 2011)

A Tabela 1 apresenta um resumo das principais características visuais e das propriedades magnéticas típicas associadas aos três estados microestruturais dos aços elétricos GNO abordados neste trabalho.

Tabela 1 – Características visuais e magnéticas de cada estado microestrutural dos aços elétricos GNO.

Estado Microestrutural	Características Visuais	Características Magnéticas
Laminado a frio	Grãos achatados e alongados, orientados preferencialmente em uma direção específica, com contornos altamente irregulares	Aumento das perdas magnéticas por histerese
Recristalização parcial	Grãos alongados e achatados coexistindo com alguns grãos equiaxiais recém-formados	Redução parcial das perdas por histerese; leve aumento da permeabilidade magnética
Recristalização primária	Grãos totalmente equiaxiais e uniformes, com contornos suaves e regulares	Minimização das perdas magnéticas e máxima permeabilidade

3.2 Metalografia

As imagens de microscopia apresentadas na Figura 3 foram obtidas a partir da aplicação de práticas e técnicas metalográficas consolidadas. De acordo com Rusanovsky *et al.* (Rusanovsky; Beerli; Oren, 2022), a metalografia consiste no estudo da microestrutura de ligas metálicas. Esse campo desenvolve e aplica métodos que possibilitam a extração de informações relacionadas às propriedades elétricas, magnéticas e mecânicas dos aços elétricos. De modo geral, a análise microestrutural de um material é realizada por meio da interação entre uma sonda e uma amostra cuidadosamente preparada (Brandon; Kaplan, 2013), destacando-se o exame metalográfico como uma das principais metodologias empregadas nesse contexto (Scott, 2019).

Os exames metalográficos seguem etapas bem definidas com o objetivo de investigar a estrutura interna de um material metálico e correlacionar suas características microestruturais com o desempenho em serviço. Essas etapas compreendem, de modo geral, a preparação da amostra, o embutimento em resina sintética, como a baquelite, o lixamento e o polimento da superfície. Na etapa final, realiza-se o ataque químico da superfície da amostra, procedimento que permite evidenciar as diferenças entre grãos e regiões do material. Todos esses processos são indispensáveis para a obtenção de imagens microestruturais por meio da microscopia óptica, resultando nas imagens apresentadas na Figura 3.

Entretanto, embora a microscopia óptica seja amplamente empregada, sua aplicação na análise microestrutural de aços elétricos apresenta limitações relevantes. Além de se tratar de um processo demorado, que exige esforço significativo do operador (Bulgarevich *et al.*, 2018), sua avaliação é fortemente dependente da experiência do observador e está

sujeita à variabilidade inerente entre diferentes especialistas (Hiremath *et al.*, 2025).

As limitações da análise microestrutural por microscopia óptica tornam-se ainda mais pronunciadas quando se trata da classificação de diferentes estados microestruturais. Isso se deve, principalmente, à ambiguidade inerente às transições entre os estados laminado a frio, parcialmente recristalizado e totalmente recristalizado (recristalização primária), o que torna a avaliação manual suscetível a inconsistências (Sun *et al.*, 2020; Chen *et al.*, 2025). Diante desse cenário, abordagens computacionais, incluindo algoritmos de aprendizado de máquina e técnicas de visão computacional, podem ser empregadas para automatizar a classificação microestrutural, promovendo a padronização do processo, a redução de vieses subjetivos e o aumento da reprodutibilidade das análises.

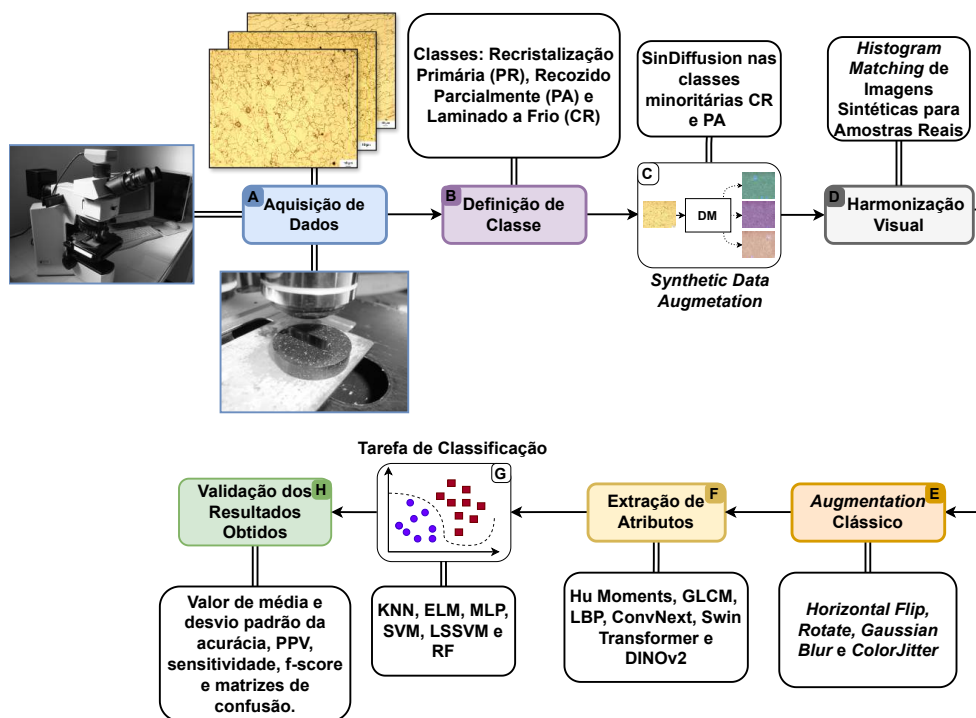
Em síntese, embora a microscopia óptica seja fundamental para a observação de microestruturas, suas limitações no processo de classificação manual, tais como a variabilidade entre operadores e a necessidade de elevado grau de especialização, evidenciam a demanda por abordagens computacionais. Nesse contexto, este trabalho emprega técnicas de visão computacional e aprendizado de máquina para automatizar e padronizar a classificação microestrutural de aços elétricos não orientados a partir de imagens de microscopia óptica, integrando de forma eficaz a análise em ciência dos materiais com métodos computacionais.

4 Metodologia Proposta

Neste trabalho, propõe-se uma metodologia baseada em técnicas de aprendizado de máquina e visão computacional com o objetivo de otimizar a análise microestrutural. Essa abordagem possibilita a classificação dos estados microestruturais de aços elétricos não orientados a partir de imagens de microscopia óptica, reduzindo tanto o tempo de análise quanto a subjetividade do processo. Adicionalmente, a metodologia integra técnicas clássicas de aumento de dados com abordagens baseadas em modelos de difusão, de modo a contornar a limitação de conjuntos de dados reduzidos e ampliar sua aplicabilidade em diferentes contextos experimentais.

Portanto, este capítulo descreve as etapas da metodologia proposta. Inicialmente, é apresentada a aquisição das imagens de microscopia óptica dos aços elétricos não orientados, bem como a definição de suas respectivas classes. Em seguida, são detalhados os procedimentos para a geração de dados sintéticos destinados às classes minoritárias, incluindo o processo de harmonização visual dessas imagens artificiais. Por fim, são descritas as técnicas clássicas de aumento de dados aplicadas às imagens sintéticas, com o objetivo de ampliar sua variabilidade. Todas as etapas da metodologia são ilustradas no diagrama apresentado na Figura 4.

Figura 4 – Metodologia da abordagem proposta.



Fonte: Autor.

4.1 Aquisição de Dados

A primeira etapa metodológica, correspondente ao bloco A do diagrama apresentado na Figura 4, refere-se à aquisição dos dados. As imagens de microscopia óptica utilizadas neste estudo são provenientes de uma investigação previamente realizada com aços elétricos não orientados laminados a frio do trabalho de Freitas (Freitas, 2011). O material analisado consistiu em amostras processadas com reduções de espessura de 50% (1,2 mm) e 70% (0,6 mm). A composição química dos aços empregados é apresentada na Tabela 2.

Tabela 2 – Composição química do aço elétrico GNO utilizada.

<i>Concentração dos Elementos (% em peso)</i>									
<i>C</i>	<i>Mn</i>	<i>Si</i>	<i>P</i>	<i>S</i>	<i>Al</i>	<i>Nb</i>	<i>V</i>	<i>Ti</i>	<i>B</i>
0.050	0.290	1.280	0.025	0.014	0.036	—	—	—	—

Fonte: Freitas (Freitas, 2011)

Cada amostra de aço foi submetida à recristalização parcial em temperaturas de 575, 580, 585, 590, 595, 600, 605, 610, 615 e 620 °C. O processo foi realizado por meio de tratamento térmico em forno de laboratório, com incrementos de 5 °C entre as amostras, visando à obtenção de diferentes estágios microestruturais, com frações recristalizadas variando de 0% a 100%.

Na sequência, as amostras foram submetidas ao preparo metalográfico para a avaliação de sua microestrutura por meio de microscopia óptica. A Figura 5 apresenta o microscópio óptico (fabricante: OLYMPUS; modelo: BX51M) utilizado na aquisição das imagens, bem como uma amostra embutida em baquelite. Dessa forma, as imagens obtidas por meio do microscópio constituíram o conjunto de dados empregado na classificação automatizada do estado cristalográfico dos aços elétricos GNO.

Figura 5 – Microscópio Óptico utilizado para aquisição das imagens digitais através da fotomicrografia, destacando o material de fixação e a amostra a ela acoplada.



Fonte: Freitas (Freitas, 2011)

4.2 Definição de Classe

A etapa seguinte, correspondente ao bloco B do diagrama apresentado na Figura 4, consistiu na definição das classes das amostras de acordo com seu estado microestrutural. Ao todo, foram coletadas 116 imagens de microscopia óptica. Desse conjunto, 100 imagens correspondem ao estado parcialmente recozido (*Partial Annealed* – PA), enquanto 16 imagens correspondem aos estados laminado a frio (*Cold Rolled* – CR) e recristalização primária (*Partial Recrystallization* – PR). A Tabela 3 apresenta a quantidade de amostras de acordo com o tipo de laminação a frio, a ampliação utilizada e a classe associada, além de detalhar a subdivisão das amostras da classe parcialmente recozida em função da temperatura de recozimento.

Tabela 3 – Quantidade de amostras de imagens digitais de microscopia por temperatura e classe.

<i>Steel</i>	<i>Zoom</i>	<i>CR</i>	<i>Temperatura aplicada (°C) - PA</i>										<i>PR</i>
			575	580	585	590	595	600	605	610	615	620	
50%	200x	1	1	1	1	1	1	1	1	1	1	1	1
50%	500x	1	1	1	1	1	1	1	1	1	1	1	1
70%	200x	3	5	7	2	5	2	4	4	4	3	3	3
70%	500x	3	5	5	2	6	2	6	4	5	3	3	3

Fonte: Autor.

4.3 Aumento de Dados Sintéticos

Após a aquisição das imagens de microscopia óptica e a identificação das classes, foi observado um desbalanceamento entre a classe PA (100 imagens) e as classes minoritárias CR e PR (8 imagens cada). Para lidar com esse desbalanceamento e aumentar a representatividade das classes minoritárias, a etapa C da metodologia (Figura 4) empregou o modelo SinDiffusion, baseado em modelos de difusão (Wang *et al.*, 2025), o qual foi utilizado para gerar imagens sintéticas das classes CR e PR.

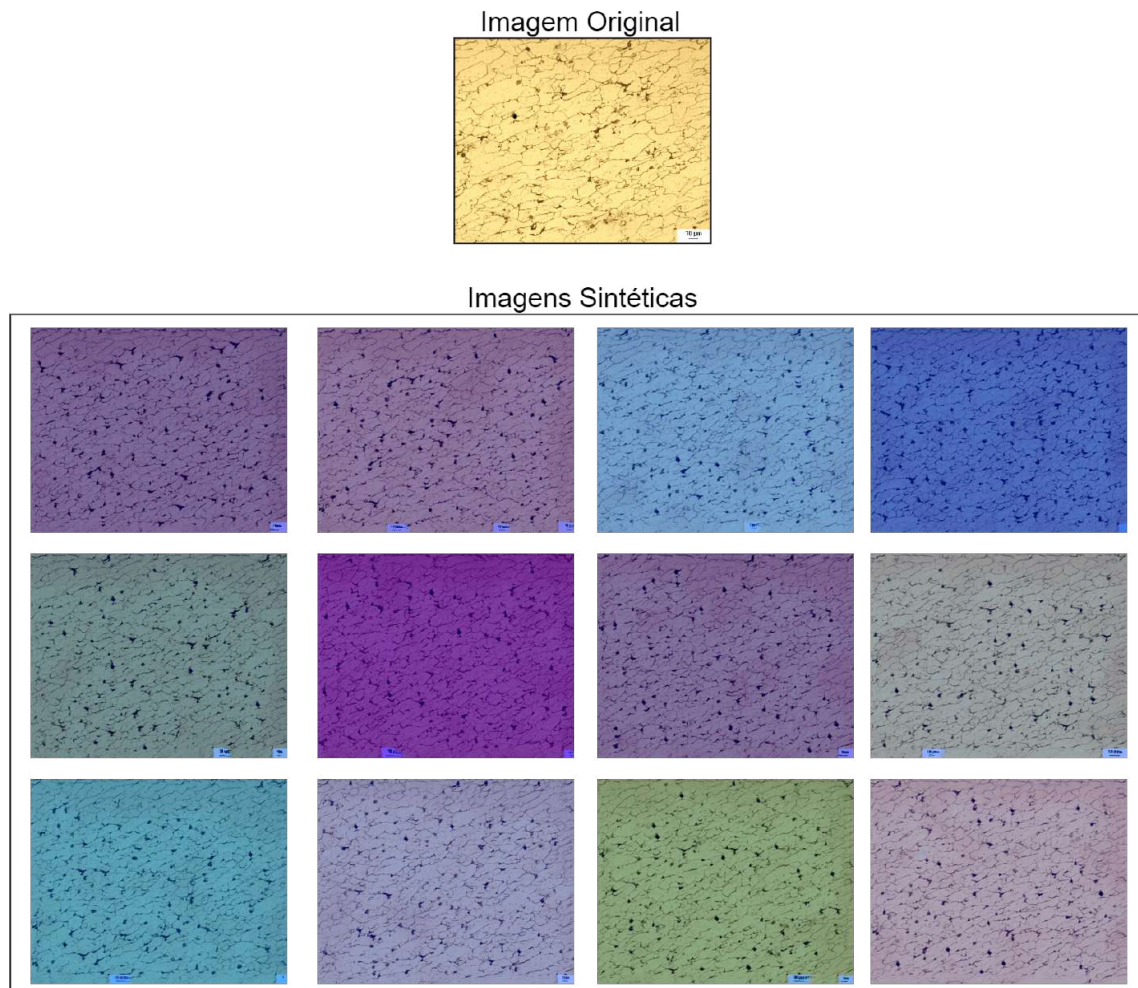
Diferentemente dos modelos de difusão tradicionais, o SinDiffusion é capaz de gerar múltiplas amostras sintéticas a partir de poucas imagens ou até mesmo de uma única imagem. Com base nessa característica, foram treinadas 16 instâncias do modelo SinDiffusion, uma para cada imagem real pertencente às classes CR e PR.

A adoção do modelo SinDiffusion como gerador de amostras sintéticas foi motivada por dois fatores principais. O primeiro refere-se ao reduzido número de amostras reais nas classes minoritárias, conforme apresentado na Tabela 3. O segundo está relacionado às características das estruturas cristalográficas analisadas, as quais inviabilizam o uso de modelos treinados com mais de uma imagem, uma vez que essa abordagem pode resultar na

geração de amostras sintéticas com características mistas. Como consequência, o conjunto de dados original poderia ser complementado com amostras distorcidas, o que levaria à redução da acurácia na classificação do estado do aço durante a fase de teste.

Para lidar com o problema de desbalanceamento amostral, foram geradas 12 imagens sintéticas para cada imagem real pertencente às classes PR e CR. Como resultado, obteve-se um total de 192 amostras sintéticas, 96 para cada classe. A Figura 6 apresenta a comparação entre uma imagem real de microscopia óptica da classe CR, com 50% de redução de espessura e ampliação de 500 $\times$, e as 12 imagens sintéticas geradas a partir dela pelo modelo SinDiffusion.

Figura 6 – Análise comparativa da estrutura cristalográfica das 12 imagens sintéticas geradas pelo *SinDiffusion* a partir do treinamento da imagem digital de microscopia óptica da amostra laminada a frio do tipo de laminação 50% com zoom de 500X.



Fonte: Autor.

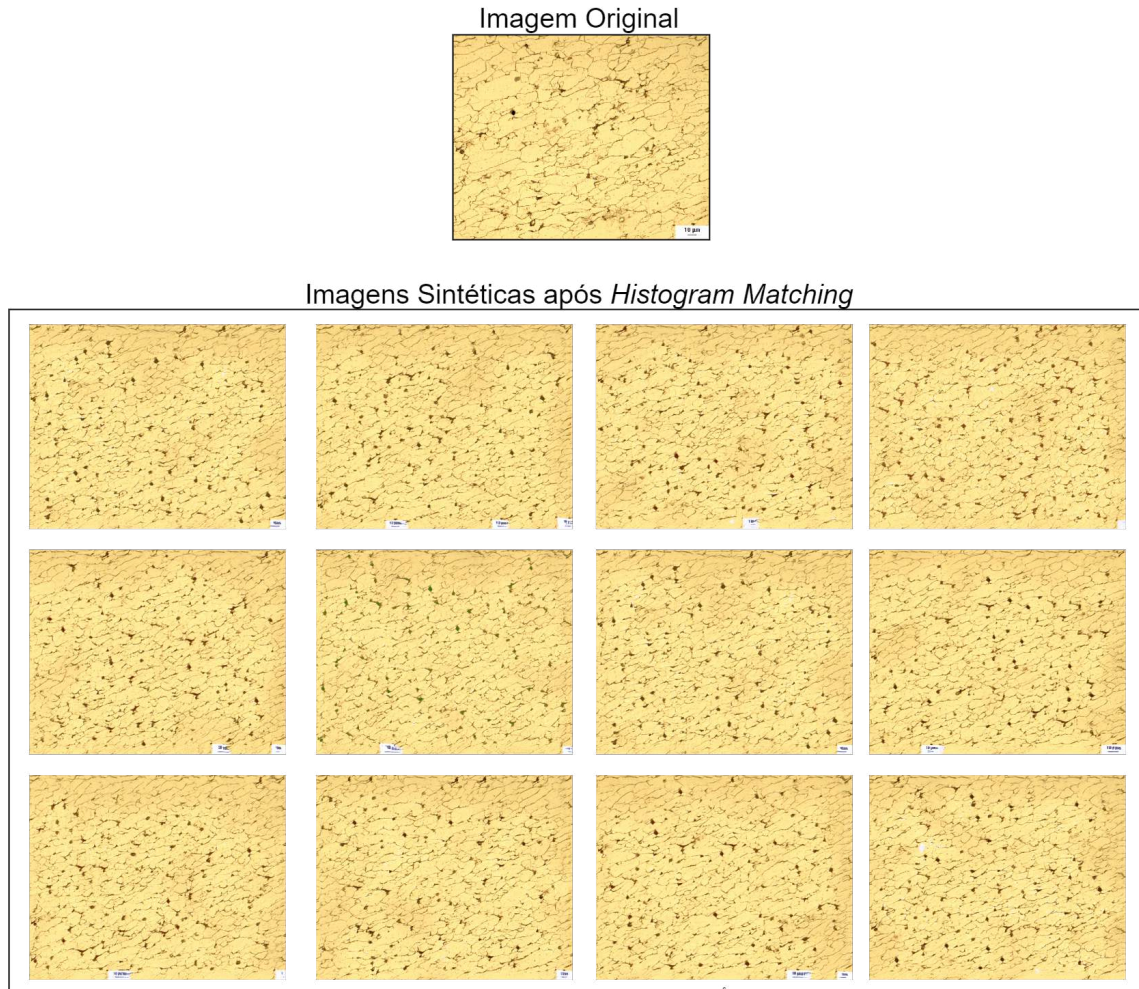
4.4 Harmonização Visual

Conforme observado na Figura 6, as imagens sintéticas geradas pelo modelo *SinDiffusion* apresentam padrões microestruturais semelhantes à da imagem original, preservando as principais características microestruturais, com exceção do indicador de ampliação localizado no canto inferior direito. Observa-se, entretanto, que a aplicação de variações estatísticas nos três canais de cor pelo modelo resultou em amostras sintéticas com tonalidades cromáticas distintas das imagens reais, sem comprometer a representação morfológica dos grãos.

Para contornar esse problema, a etapa D da metodologia aplicou uma técnica de ajuste de tonalidade conhecida como *Histogram Matching*. Trata-se de um método clássico que transforma o histograma de uma imagem de modo a coincidir com o histograma de uma imagem de referência (Gonzalez, 2009). Dessa forma, todos os dados sintéticos foram ajustados utilizando o histograma da imagem original que serviu como referência para sua geração. A Figura 7 apresenta os resultados das imagens sintéticas mostradas na Figura 6 após a aplicação da técnica de *Histogram Matching*.

Após a aplicação desta técnica, um especialista na área de ciência dos materiais realizou uma avaliação qualitativa da plausibilidade física de quatro imagens sintéticas selecionadas aleatoriamente, sem conhecimento prévio de sua origem. Duas imagens foram classificadas como reais, uma como praticamente indistinguível de uma imagem real, e uma foi corretamente identificada como sintética. Nos dois últimos casos, indícios sutis apontaram uma origem artificial: em uma das imagens, observou-se uma regularidade excessiva das regiões escuras; na outra, uma concentração localizada de grãos alongados, enquanto o alongamento induzido por deformação tipicamente ocorre de forma mais homogênea. De modo geral, apesar da presença ocasional de artefatos não físicos, as imagens sintéticas preservam, em grande medida, características metalúrgicas consistentes e apresentam elevada fidelidade visual.

Figura 7 – Análise comparativa da estrutura cristalográfica das 12 imagens sintéticas geradas pelo SinDiffusion a partir do treinamento da imagem digital de microscopia óptica da amostra da classe CR (Laminado a frio) do tipo de laminação 50% com zoom de 500X, após a aplicação da técnica de *histogram matching*.



Fonte: Autor.

4.5 Augmentation Clássico

Os resultados das imagens sintéticas após a aplicação do *Histogram Matching* demonstraram visualmente a eficácia da técnica na mitigação da variação de cor observada nas amostras artificiais. Como consequência, as imagens ajustadas passaram a apresentar elevada similaridade com a imagem original. No entanto, uma similaridade excessiva em relação aos dados reais pode reduzir a diversidade do conjunto de dados aumentado, o que, por sua vez, pode impactar negativamente a capacidade de generalização dos modelos de aprendizado de máquina (Liu *et al.*, 2025; Goncalves *et al.*, 2020).

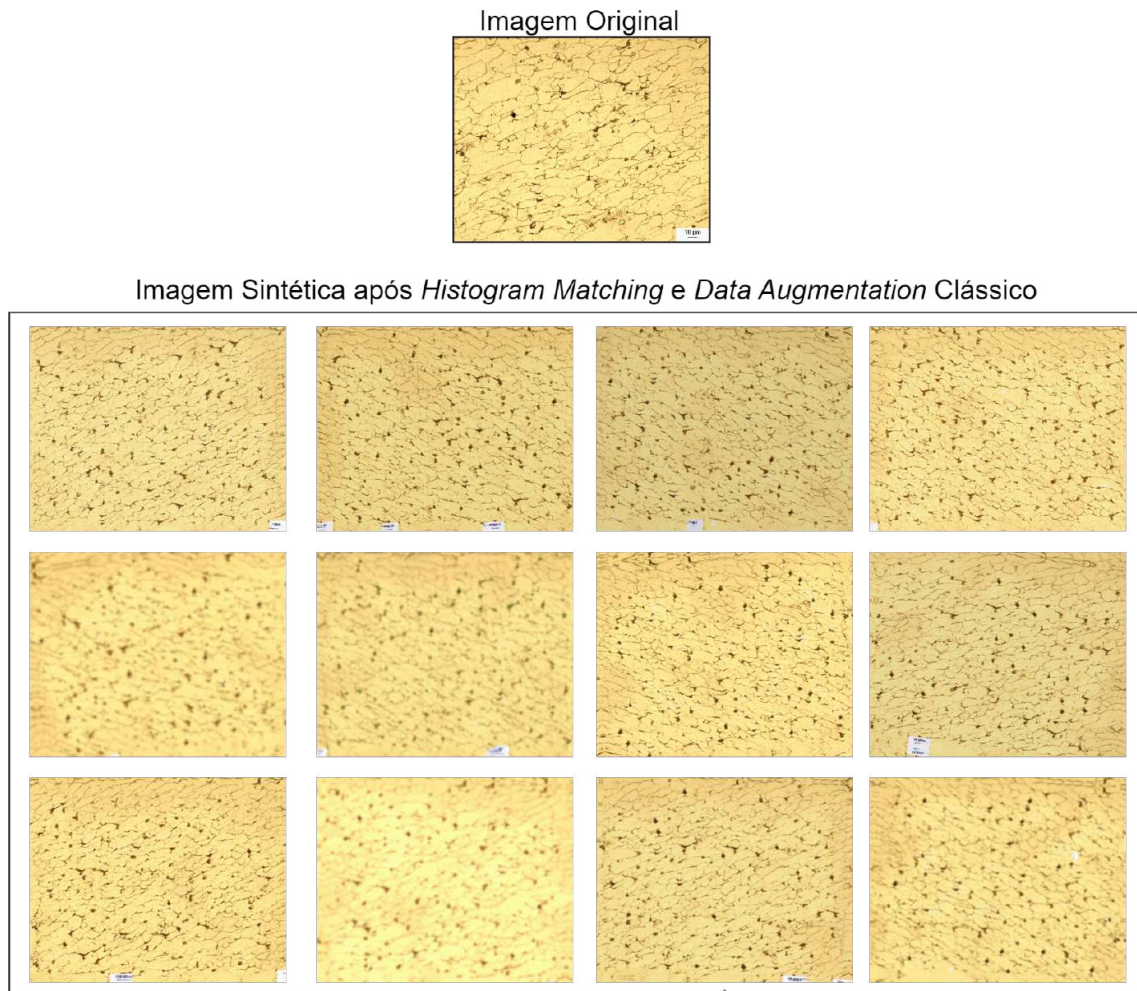
Com o objetivo de mitigar o impacto da elevada similaridade entre as amostras artificiais, a etapa seguinte da metodologia (Etapa E) consistiu na aplicação de técnicas clássicas de aumento de dados, visando ampliar a variabilidade visual das amostras. Os

métodos empregados são descritos a seguir:

- **Espelhamento Horizontal (*Horizontal Flip*)**: operação de espelhamento que inverte a imagem da esquerda para a direita. Essa técnica foi aplicada a cada imagem sintética com probabilidade de 50%.
- **Rotação (*Rotate*)**: rotação da imagem limitada a um ângulo máximo de 10° , com preenchimento das bordas realizado por meio de espelhamento, de modo a evitar regiões vazias. Essa operação foi aplicada com probabilidade de 70%.
- **Desfoque Gaussiano (*Gaussian Blur*)**: aplicação de um kernel gaussiano de dimensão 3×3 , com o objetivo de introduzir um efeito de suavização na imagem. Essa técnica foi aplicada com probabilidade de 50%.
- **Perturbação de Cor (*Color Jitter*)**: técnica que altera aleatoriamente os níveis de brilho, contraste, saturação e matiz da imagem. Neste trabalho, foram aplicadas variações aleatórias de brilho e contraste de até $\pm 10\%$ em relação aos valores originais, com probabilidade de 50%.

A seleção das técnicas de aumento de dados foi fundamentada nas características das imagens analisadas. Considerando que cada detalhe presente nas micrografias é relevante para a caracterização microestrutural, foram empregadas apenas técnicas de aumento que preservam a integridade estrutural das amostras sintéticas. A Figura 8 apresenta o resultado final da aplicação dessas técnicas após o processo de ajuste de histograma, no qual podem ser observadas variações de tonalidade e nitidez nas imagens geradas.

Figura 8 – Análise comparativa da estrutura cristalográfica das 12 imagens sintéticas geradas pelo *SinDiffusion* a partir do treinamento da imagem digital de microscopia óptica da amostra da classe CR (Laminado a frio) do tipo de laminação 50% com zoom de 500X, após a aplicação da técnica de *histogram matching* e *data augmentation* clássica.



Fonte: Autor.

4.6 Extratores de Atributos

A classificação automática dos estados microestruturais dos aços elétricos GNO requer a extração de informações relevantes capazes de representar adequadamente os padrões presentes nas imagens de microscopia óptica. Nesse contexto, na Etapa F da metodologia, foram selecionadas diferentes técnicas de extração de atributos, contemplando tanto métodos clássicos, amplamente utilizados na literatura, quanto abordagens baseadas em *deep learning*, que têm se destacado pela capacidade de capturar características mais complexas e discriminativas. Essas técnicas constituem a base para os experimentos de classificação apresentados nos capítulos subsequentes.

4.6.1 Extratores de Atributos Clássicos

Os extratores de atributos clássicos selecionados neste trabalho são amplamente empregados em diferentes domínios da extração de características em imagens, especialmente em tarefas de classificação de padrões que utilizam técnicas de aprendizado de máquina e inteligência artificial. Aplicações recentes desses métodos têm sido direcionadas a áreas como Agricultura (Devi; Amarendra, 2021; Babalola; Asad; Bais, 2023; Kilic; Yaman, 2024; Sobuj *et al.*, 2024), Saúde (Teke; Etem, 2025; Dheepak; Vaishali, 2024; Shah; Sankar, 2023), Engenharia (Daneshvari *et al.*, 2023) e Segurança (Raj; Sankar, 2022).

As técnicas de extração de atributos em imagens têm como objetivo principal capturar informações relacionadas às características intrínsecas das imagens, tais como textura, posição, curvatura e dimensão dos objetos. Os métodos clássicos aplicados neste estudo foram selecionados com o intuito de enfatizar as propriedades de textura presentes nas micrografias dos aços elétricos GNO. As subseções a seguir descrevem detalhadamente cada uma dessas técnicas.

4.6.1.1 Hu Moments

Os *Hu Moments* realizam a extração de atributos de forma invariável à orientação, à escala e à posição dos objetos presentes na imagem. Essa técnica, comumente denominada Momentos Invariantes, foi originalmente proposta por Ming-Kuei Hu (Hu, 1962). Desde então, tem sido amplamente aplicada em tarefas de reconhecimento de padrões e análise de formas.

Na análise microestrutural das imagens microscópicas dos aços elétricos GNO, esta técnica permite capturar informações globais relacionadas à distribuição espacial das regiões e formas presentes nas micrografias. A invariância a rotações, translações e variações de escala pode ser particularmente relevante, uma vez que imagens como estas podem apresentar diferentes orientações e enquadramentos decorrentes de direção de laminação, corte da amostra e zoom. Embora essa técnica não descreva diretamente características texturais finas, ela pode contribuir para a discriminação entre estados microestruturais ao representar padrões geométricos globais associados à deformação e à recristalização dos grãos.

4.6.1.2 Local Binary Pattern

Proposto por Ojala *et al.* (Ojala; Pietikäinen; Harwood, 1996), o *Local Binary Pattern* (LBP) é uma técnica que produz um descritor de textura simples, rápido e robusto, sendo amplamente utilizada para a representação de texturas em diversas aplicações de análise de imagens (Lai *et al.*, 2025). O LBP é capaz de capturar padrões locais de intensidade presentes na imagem, apresentando invariância a variações de iluminação e contraste.

Inicialmente, a técnica Local Binary Pattern (LBP) define uma matriz secundária com dimensões iguais às da imagem original, destinada a armazenar o resultado da operação para cada pixel. Para cada pixel da imagem, é considerado um conjunto de vizinhança, geralmente organizado em um arranjo circular ou quadrado ao seu redor. Cada pixel vizinho é então comparado com o pixel central: caso o valor do vizinho seja maior ou igual ao do pixel central, atribui-se o valor 1; caso contrário, atribui-se o valor 0. Após a avaliação de todos os vizinhos, os valores binários resultantes são concatenados e convertidos em um número decimal, o qual é registrado na matriz secundária na posição correspondente ao pixel central da imagem original.

4.6.1.3 Gray-Level Co-Occurrence Matrix

A técnica da Matriz de Coocorrência de Níveis de Cinza (*Gray-Level Co-occurrence Matrix* – GLCM) foi introduzida por Haralick *et al.* (Haralick; Shanmugam; Dinstein, 2007) e é amplamente utilizada para análise de textura. Essa abordagem caracteriza a distribuição espacial dos níveis de cinza em uma imagem ao quantificar a frequência com que pares de intensidades de pixels ocorrem a uma determinada distância e orientação. Ao capturar essas relações espaciais, a GLCM possibilita a extração de atributos texturais discriminativos a partir da imagem.

A técnica constrói uma matriz quadrada de dimensão $N \times N$, em que N representa o número de níveis de cinza da imagem (por exemplo, $N = 256$ para imagens de 8 bits). Cada elemento (i, j) dessa matriz armazena o número de vezes em que um pixel com intensidade i ocorre adjacente a um pixel com intensidade j , de acordo com um deslocamento previamente definido. Por exemplo, a posição $(0, 0)$ indica quantas vezes um pixel de intensidade zero é seguido por outro pixel também de intensidade zero. Para a construção da GLCM, são especificados uma distância entre pixels d e uma direção de análise ($0^\circ, 45^\circ, 90^\circ$ ou 135°), sendo que cada direção resulta em uma matriz de coocorrência distinta.

Após a obtenção das matrizes de coocorrência (GLCM), o próximo passo extrai os atributos dessas matrizes resultantes para caracterizar a textura da imagem. Esses atributos transformam a matriz em valores numéricos que sintetizam padrões de repetição, contraste e uniformidade presentes na imagem. Os principais atributos extraídos são:

- **Contraste:** mede a variação local de intensidade entre pixels vizinhos. Valores elevados indicam grandes diferenças de intensidade, estando associados a texturas mais rugosas;
- **Homogeneidade:** avalia a proximidade da distribuição da GLCM em relação à diagonal principal. Altos valores indicam que pixels vizinhos apresentam intensidades semelhantes, caracterizando texturas mais suaves;
- **Energia:** mede o grau de uniformidade da textura. Valores elevados estão

associados a padrões altamente repetitivos e bem definidos;

- **Correlação:** quantifica o grau de dependência linear entre um pixel e seus vizinhos, levando em consideração a média e o desvio padrão das linhas e colunas da matriz de coocorrência.

4.6.2 Extratores de Atributos baseados em *Deep Learning*

Nos últimos anos, arquiteturas baseadas em Redes Neurais Convolucionais e *Transformers* têm sido amplamente adotadas em diversos domínios de aplicação, que vão desde imagens médicas até sensoriamento remoto, em razão de suas robustas capacidades de extração de atributos (Zhao *et al.*, 2024). As CNNs dominaram por um longo tempo o campo da visão computacional como abordagem padrão para extração de características, devido à sua capacidade de aprender representações hierárquicas diretamente a partir dos dados (Palanisamy *et al.*, 2025; Zhao *et al.*, 2024). Entretanto, arquiteturas baseadas em *Transformers* vêm emergindo de forma crescente como extratores de atributos eficazes, principalmente em decorrência de inovações arquiteturais que reduzem o custo computacional (Palanisamy *et al.*, 2025).

Foram adotados três diferentes modelos recentemente desenvolvidos, baseados em arquiteturas de CNNs e *Transformers*. Cada um desses modelos foi empregado como extrator de atributos, e uma breve descrição de cada abordagem é apresentada a seguir:

4.6.2.1 ConvNeXt

Proposto por Zhuang Liu *et al.* (Liu *et al.*, 2022), do Meta AI, o *ConvNeXt* é uma variante moderna de modelos baseados em redes neurais convolucionais (ConvNet), projetada com o objetivo de competir com o *Vision Transformer* (ViT) em tarefas de visão computacional. Sua principal contribuição consiste na atualização do design das CNNs clássicas, incorporando diversas ideias inspiradas em arquiteturas Transformer, bem como boas práticas de treinamento desenvolvidas recentemente.

Sua arquitetura mantém a estrutura geral de uma *ResNet*. No entanto, emprega núcleos convolucionais maiores (por exemplo, 7×7) combinados com convoluções do tipo *depthwise*, inspiradas nas convoluções separáveis utilizadas em arquiteturas como *MobileNet* e *Vision Transformers*. Além disso, o modelo aumenta o número de canais e simplifica as conexões com o objetivo de melhorar a eficiência computacional. Adicionalmente, adota camadas de normalização e ativação mais modernas, como a *LayerNorm*, em substituição à *BatchNorm* em determinadas etapas da arquitetura.

O *ConvNeXt* preserva as vantagens inerentes às convoluções no tratamento de dados com forte estrutura local. Além disso, incorpora aprimoramentos que ampliam sua capacidade de modelar o contexto global das imagens (Liu *et al.*, 2022). Dessa forma,

o modelo mostra-se adequado para a extração de características associadas à estrutura cristalográfica, sendo capaz de capturar tanto padrões globais relacionados à distribuição dos grãos não orientados quanto informações locais referentes às características individuais desses grãos.

4.6.2.2 Swin Transformer

O *Swin Transformer*, desenvolvido por Ze Liu *et al.* (Liu *et al.*, 2021) na Microsoft Research Asia, é um modelo do tipo *Vision Transformer* (ViT) projetado com uma arquitetura hierárquica e eficiente, com o objetivo de atuar como uma solução padrão para tarefas de visão computacional, tais como classificação, detecção e segmentação. O termo *Swin* deriva de *Shifted Window*, técnica empregada pelo modelo para o cálculo do mecanismo de atenção do ViT em janelas locais que são deslocadas a cada camada da rede.

Com o objetivo de reduzir o custo computacional, o *Swin Transformer* adota janelas de atenção não sobrepostas, denominadas *window-based self-attention*. Em cada camada, o deslocamento dessas janelas permite a interação entre diferentes regiões da imagem, possibilitando ao modelo capturar dependências globais de forma eficiente. Sua estrutura hierárquica inicia-se a partir de pequenos *patches* e realiza operações de *patch merging*, gerando representações progressivamente mais compactas, de maneira análoga às operações de *downsampling* empregadas em arquiteturas baseadas em CNNs.

A arquitetura do Swin Transformer resulta em complexidade computacional linear, em contraste com a complexidade quadrática observada nos modelos ViT convencionais. Essas características tornam o *Swin Transformer* uma arquitetura mais eficiente, escalável e versátil para diferentes tarefas de visão computacional (Liu *et al.*, 2021). Dessa forma, esse modelo pode apresentar desempenhos tão bons quanto os demais modelos baseados em *deep learning* para extração de atributos, aliando tais resultados a vantagens em termos de escalabilidade e eficiência, aspectos particularmente relevantes para aplicações reais de classificação automática.

4.6.2.3 DINOv2

O DINOv2 é um modelo de visão computacional autossupervisionado proposto por Maxime Oquab *et al.* (Oquab *et al.*, 2023), do Meta AI. Ele foi desenvolvido com o objetivo de aprender representações visuais genéricas e de alta qualidade sem a necessidade do uso de imagens rotuladas.

O DINOv2 é construído sobre a arquitetura de *Vision Transformers* e emprega um método de aprendizado autossupervisionado, no qual o modelo aprende a extrair representações visuais consistentes a partir de diferentes visualizações de uma mesma imagem. Nesse processo, versões distintas da imagem de entrada são analisadas pelo próprio modelo,

que é treinado para produzir descritores semelhantes independentemente das variações visuais aplicadas. Essa estratégia permite o aprendizado de atributos discriminativos sem a necessidade de rótulos previamente definidos. Em outras palavras, a abordagem do modelo emprega uma estrutura do tipo professor–aluno, na qual o modelo é treinado para gerar representações consistentes a partir de diferentes visões de uma mesma imagem (Oquab *et al.*, 2023).

Essa arquitetura foi treinada a partir de um conjunto de dados massivo e cuidadosamente curado, composto por imagens diversificadas e de alta qualidade, previamente filtradas com o objetivo de remover ruídos. Além disso, o modelo é capaz de gerar *embeddings* visuais (atributos) que preservam tanto a semântica quanto a estrutura espacial da imagem de entrada.

4.7 Modelos para Classificação

Uma vez concluída a etapa de extração de atributos por meio das diferentes técnicas selecionadas, a Etapa F concentrou-se na classificação automática dos estados microestruturais dos aços elétricos GNO. Para isso, foram adotados diferentes algoritmos de aprendizado de máquina amplamente empregados em tarefas de reconhecimento e classificação de padrões. Esses classificadores foram utilizados para analisar os atributos extraídos das imagens de microscopia óptica pelos métodos previamente apresentados, permitindo investigar o desempenho de distintas abordagens na identificação automática das classes microestruturais dos materiais analisados.

4.7.1 Support Vector Machine

A técnica de *Support Vector Machine* (SVM) representa uma classe robusta de algoritmos de aprendizado de máquina fundamentados na teoria do aprendizado estatístico (Vapnik, 2013). O método foi inicialmente proposto como uma generalização não linear do algoritmo *Generalized Portrait*. Posteriormente, Smola e Vapnik (Smola, 1998) estenderam essa abordagem para o domínio de problemas de regressão, ampliando significativamente sua aplicabilidade em diferentes tarefas de aprendizado supervisionado.

O modelo SVM resolve um problema de otimização com restrições quadráticas, cujo objetivo é minimizar uma combinação entre o risco empírico — como empregado, por exemplo, em redes neurais convencionais — e um termo de regularização, responsável por impor suavidade à solução obtida. Essa solução é alcançada por meio da resolução de um problema convexo de programação quadrática no espaço dual e apresenta, de forma notável, as propriedades desejáveis de otimização global e esparsidade.

4.7.2 Least Squares Support Vector Machine

A técnica *Least Squares Support Vector Machine* (LSSVM) foi proposta por Suykens e Vandewalle em 1999 (Suykens; Vandewalle, 1999) como uma adaptação da abordagem de SVM de Vapnik (Subseção 4.7.1), com o objetivo de acelerar sua formulação, preservando, ao mesmo tempo, suas principais vantagens.

O algoritmo LSSVM difere do SVM padrão em dois aspectos principais. Primeiramente, o LSSVM emprega uma abordagem de otimização baseada em mínimos quadráticos para determinar o hiperplano ótimo, utilizando a Soma dos Erros Quadráticos (Sum of Squared Errors – SSE) como função objetivo. A segunda distinção está na simplificação de sua formulação original: o LSSVM substitui as restrições de desigualdade estritas por restrições de igualdade, de modo que o problema de otimização pode ser resolvido por meio de um sistema linear das condições de Karush–Kuhn–Tucker (KKT) (Fletcher, 2000).

4.7.3 Multilayer Perceptron

O Perceptron Multicamadas (*Multi-Layer Perceptron* – MLP) é uma arquitetura de redes neurais eficaz para a resolução de problemas tipicamente não lineares, sendo amplamente aplicada em tarefas de classificação de padrões e de aproximação de funções. Sua estrutura é composta por múltiplas camadas de neurônios, incluindo uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída (Principe; Euliano; Lefebvre, 1999).

A camada de entrada recebe os dados, que são associados a pesos sinápticos responsáveis por influenciar a ativação dos neurônios da primeira camada oculta. Esses neurônios podem produzir saídas lineares ou não lineares, dependendo da função de ativação adotada. As saídas das camadas ocultas são determinadas pelas ativações dos neurônios da camada anterior, moduladas pelos respectivos pesos sinápticos. Por fim, a camada de saída representa o tipo específico de tarefa que a rede busca executar, como classificação ou regressão.

O MLP foi treinado por meio do algoritmo de retropropagação do erro (back-propagation), uma extensão da regra de aprendizado de Widrow–Hoff (Widrow; Hoff, 1960). A atualização dos pesos foi realizada utilizando o otimizador Adam (Adaptive Moment Estimation) (Kingma; Ba, 2017), enquanto a função de perda de entropia cruzada (*cross-entropy loss*) (Goodfellow; Bengio; Courville, 2016) foi empregada para quantificar o erro de classificação e orientar o processo de otimização.

4.7.4 Extreme Learning Machine

A técnica de *Extreme Learning Machine* (ELM) foi introduzida por Huang *et al.* em 2004 (Huang; Zhu; Siew, 2004). Trata-se de uma rede neural com uma única camada

oculta, amplamente reconhecida por sua elevada velocidade de processamento e por sua aplicabilidade a uma ampla variedade de problemas.

Em contraste com a MLP, a ELM apresenta a característica marcante de empregar apenas uma única camada oculta. Além disso, o ajuste de seus pesos sinápticos segue uma abordagem fundamentalmente distinta daquela adotada pelas redes neurais convencionais. Os pesos sinápticos dos neurônios da camada oculta são inicializados de forma aleatória e permanecem fixos ao longo de todo o processo de treinamento.

Por outro lado, os pesos sinápticos da camada de saída são ajustados por meio da técnica conhecida como mínimos quadrados em *batch* (*batch least squares*). Nessa abordagem, todas as amostras de treinamento são apresentadas simultaneamente durante a atualização dos pesos. Essa característica distintiva da ELM a torna uma ferramenta eficiente e de rápido processamento para uma ampla variedade de tarefas de classificação, regressão e aproximação de funções.

4.7.5 k-Nearest Neighbors

O algoritmo *k-Nearest Neighbors* (KNN), introduzido por Cover e Hart em 1967 (Cover; Hart, 1967), é um método fundamental de aprendizado de máquina caracterizado por sua abordagem simples e baseada em instâncias. Ele realiza a classificação das amostras de acordo com a classe majoritária entre seus k vizinhos mais próximos, utilizando métricas de distância como as distâncias Manhattan, Euclidiana ou de Minkowski para quantificar a proximidade entre os dados.

A escolha do valor de k e da métrica de distância influencia significativamente o desempenho do algoritmo, afetando sua sensibilidade ao ruído e a eficiência computacional global. Embora o KNN seja amplamente conhecido por sua facilidade de implementação e por não assumir hipóteses sobre a distribuição subjacente dos dados, sua principal limitação reside no potencial aumento do custo computacional quando aplicado a conjuntos de dados de grande dimensão.

Técnicas de otimização, como as árvores *k-d*, foram desenvolvidas para melhorar sua eficiência, tornando o KNN uma ferramenta valiosa para análises preliminares de dados, apesar de sua simplicidade. Sua lógica direta e elevada interpretabilidade também o tornam uma opção atrativa em projetos de aprendizado de máquina, especialmente durante a fase exploratória.

4.7.6 Random Forest

A técnica de *Random Forest* (RF) ganhou popularidade principalmente a partir dos trabalhos de Leo Breiman e Adele Cutler (Cutler; Cutler; Stevens, 2012). Ela opera como um conjunto (*ensemble*) de árvores de decisão, cujas predições agregadas determinam a

classe atribuída a cada amostra. Como resultado, trata-se de uma técnica inerentemente robusta ao sobreajuste (*overfitting*).

Cada árvore realiza um procedimento de amostragem aleatória com reposição. Esse processo, comumente denominado *bootstrap sampling*, garante que cada árvore seja exposta a um subconjunto distinto dos dados de treinamento, promovendo, assim, diversidade dentro do conjunto (*ensemble*). Em seguida, um subconjunto aleatório de atributos é selecionado a cada divisão da árvore, o que aumenta ainda mais a heterogeneidade entre as árvores ao introduzir variabilidade no processo de criação dos nós.

Essas características permitem que o algoritmo RF apresente um desempenho robusto em conjuntos de dados de alta dimensionalidade e com relações não lineares. No entanto, o método pode se tornar computacionalmente oneroso quando um grande número de árvores é utilizado no treinamento, e seu desempenho pode ser degradado em cenários que envolvem dados altamente esparsos.

5 Resultados e Discussão

Este capítulo apresenta os resultados obtidos ao longo do desenvolvimento deste trabalho e está organizado de forma a permitir uma análise sistemática do desempenho das metodologias propostas. Inicialmente, são apresentados os resultados relacionados à geração de dados sintéticos e à extração de atributos, avaliando o impacto antes e após as técnicas de *data augmentation*. Em seguida, são descritos as configurações experimentais e os critérios adotados para a avaliação dos modelos. Por fim, os resultados são discutidos de maneira comparativa, destacando seus impactos e implicações para a classificação automática dos estados cristalográficos.

Assim, são apresentados os resultados das simulações computacionais referentes tanto à geração de dados sintéticos quanto à classificação dos estados cristalográficos de aços elétricos de grãos não orientados. Foram empregadas diferentes técnicas de extração de atributos em combinação com os algoritmos de aprendizado de máquina (MLP, RF, SVM, LSSVM, ELM e KNN). O desempenho dos métodos avaliados foi analisado com base em métricas de classificação, com ênfase na acurácia e na capacidade de generalização dos modelos. Todos os experimentos foram implementados em Python e executados em um sistema equipado com processador Intel i7, 16 GB de memória RAM e sistema operacional Windows 11.

5.1 Data Augmentation

Com o objetivo de validar a metodologia proposta para a geração de amostras sintéticas, foi conduzido um estudo preliminar no qual o desempenho dos classificadores foi analisado antes e após a aplicação de técnicas clássicas de aumento de dados (*data augmentation*). Essa análise inicial tem como finalidade investigar o impacto do enriquecimento do conjunto de treinamento sobre a capacidade de generalização dos modelos, servindo como base comparativa para as abordagens de geração de dados sintéticos apresentadas posteriormente.

5.1.1 Cenário de Criação de Dados Sintéticos

O conjunto de dados original é composto por 117 amostras, das quais 101 pertencem à classe *Partial Annealed* (PA). Em contraste, as classes minoritárias — *Cold Rolled* (CR) e *Primary Recrystallization* (PR) — são representadas por apenas 8 amostras cada. Esse acentuado desbalanceamento entre as classes, aliado à baixa representatividade estatística das classes minoritárias, motivou a adoção de uma estratégia de geração de dados sintéticos.

Para esse fim, foram geradas 12 amostras artificiais a partir de cada imagem real, resultando em 96 amostras sintéticas para cada uma das classes CR e PR.

Para evitar vazamento de dados (*data leakage*) durante a etapa de divisão entre conjuntos de treinamento e teste, cada amostra sintética foi alocada ao mesmo subconjunto da respectiva amostra real que lhe deu origem. Essa estratégia garante que imagens artificiais derivadas de uma mesma fonte não sejam distribuídas simultaneamente entre os conjuntos de treinamento e teste, preservando, assim, a integridade do processo de avaliação dos modelos.

Dessa forma, as amostras reais foram embaralhadas aleatoriamente e divididas em conjuntos de treinamento e teste nas proporções de 80–20 para a classe PA e de 6–2 para as classes CR e PR. Após a incorporação das amostras sintéticas, a distribuição entre as classes tornou-se mais equilibrada, resultando em divisões de 80–20 para a classe PA e de 78–20 para as classes CR e PR, respectivamente.

5.1.2 Configuração de Simulação dos Extratores de Atributos

Após a definição da metodologia de divisão das imagens em conjuntos de treinamento e teste, procedeu-se à etapa de extração de atributos. Para a técnica de Momentos de Hu, foram adotados os sete momentos invariantes como descritores. No caso do descritor *Local Binary Pattern* (LBP), utilizou-se o método uniforme com raio de $R = 3$ pixels, resultando em $P = 8 \cdot R = 24$ vizinhos ao redor do pixel central. Após a aplicação do operador LBP, foi construído um histograma com 26 bins ($P + 2$), abrangendo valores no intervalo de 0 a 25. O vetor resultante do histograma foi então utilizado como representação final das características texturais das imagens.

O descritor baseado na Matriz de Coocorrência de Níveis de Cinza (*Gray-Level Co-Occurrence Matrix* – GLCM) foi configurado considerando um deslocamento espacial de 5 pixels entre os pares de níveis de cinza e três direções angulares de análise: 0° , 45° e 90° . A partir de cada matriz GLCM, foram extraídas cinco propriedades estatísticas: contraste, dissimilaridade, homogeneidade, energia e correlação. Para cada imagem, o vetor de atributos foi composto pelos valores dessas métricas calculados para as três direções consideradas.

Para os extratores de atributos baseados em redes neurais profundas, as imagens foram processadas utilizando os três canais de cor RGB. Com o objetivo de aumentar a eficiência computacional durante a etapa de extração de atributos, foram empregadas as versões *Tiny* das arquiteturas ConvNeXt e Swin Transformer.

No caso do DINOv2, foi adotada a configuração ViT-S/14, correspondente à variante *Small* do *Vision Transformer*. Essa versão apresenta um número reduzido de parâmetros em comparação com as variantes *Base* e *Large*, proporcionando, assim, maior eficiência

computacional durante o processo de extração de atributos.

Todos os modelos foram empregados em suas versões pré-treinadas no conjunto de dados *ImageNet-1K*, aproveitando o conhecimento previamente adquirido para aprimorar a extração de atributos relevantes, método comumente chamado de *Transfer Learning*. Dessa forma, os extratores de atributos foram utilizados exclusivamente em modo de inferência, sem qualquer ajuste de pesos (*fine-tuning*), uma vez que o objetivo foi unicamente a obtenção das representações visuais.

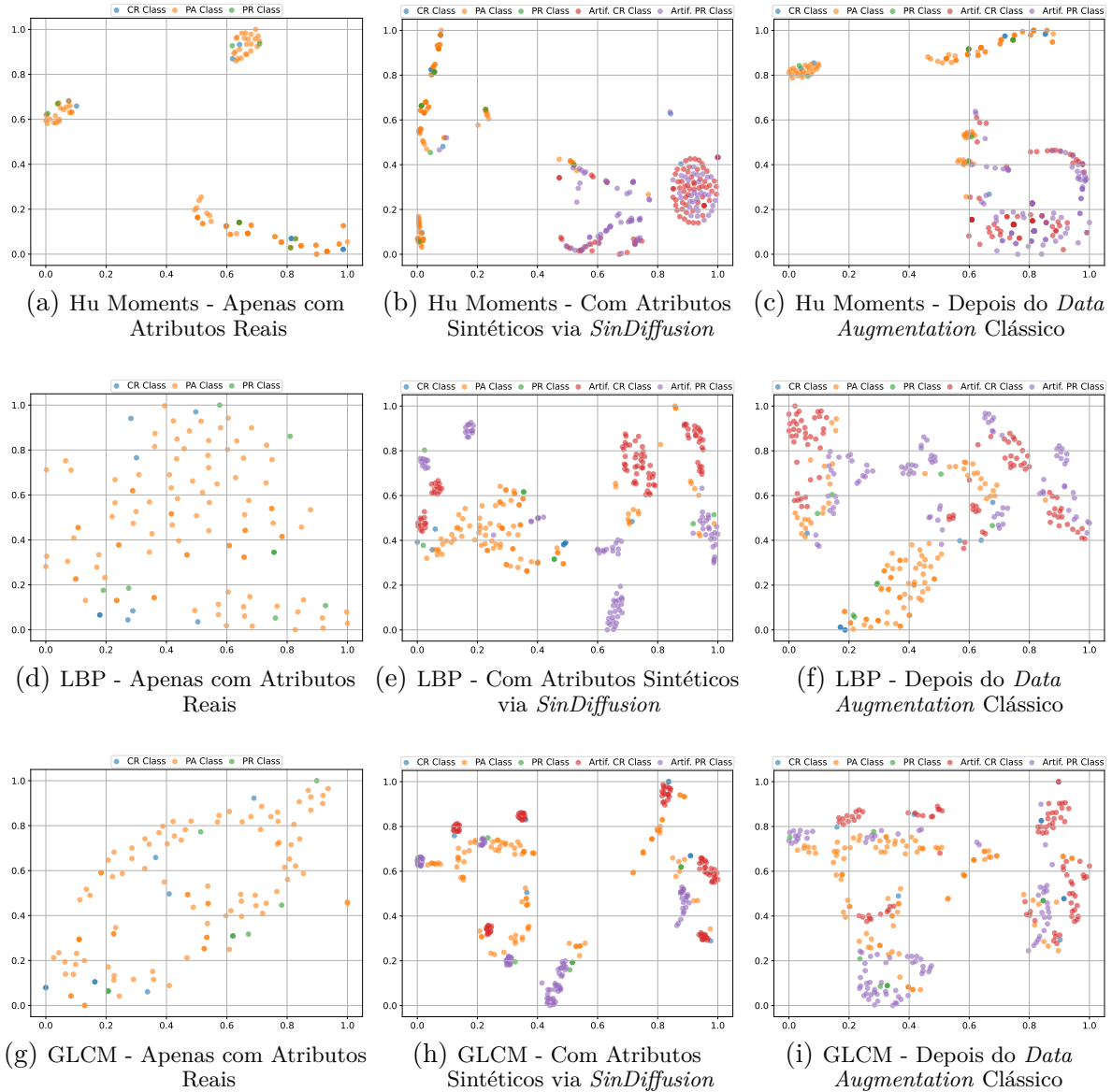
5.1.3 Representação Bidimensional dos Atributos

Para obter uma representação bidimensional da distribuição dos atributos extraídos, foi empregada a técnica *t-Distributed Stochastic Neighbor Embedding* (t-SNE) (Maaten; Hinton, 2008). Trata-se de um método não linear de redução de dimensionalidade que preserva as relações locais entre os pontos de dados, ao converter as distâncias no espaço original de alta dimensionalidade em probabilidades de vizinhança.

Essa abordagem foi aplicada para analisar o comportamento distributivo dos atributos extraídos a partir das imagens originais, das amostras sintéticas geradas pelo *SinDiffusion* e, posteriormente, à aplicação das técnicas clássicas de aumento de dados. As Figuras 9 e 10 apresentam as projeções bidimensionais obtidas pelo t-SNE em uma execução independente, tanto para os extratores de atributos tradicionais (Hu Moments, LBP e GLCM) quanto para os extratores baseados em *deep learning* (ConvNeXt, Swin Transformer e DINOv2).

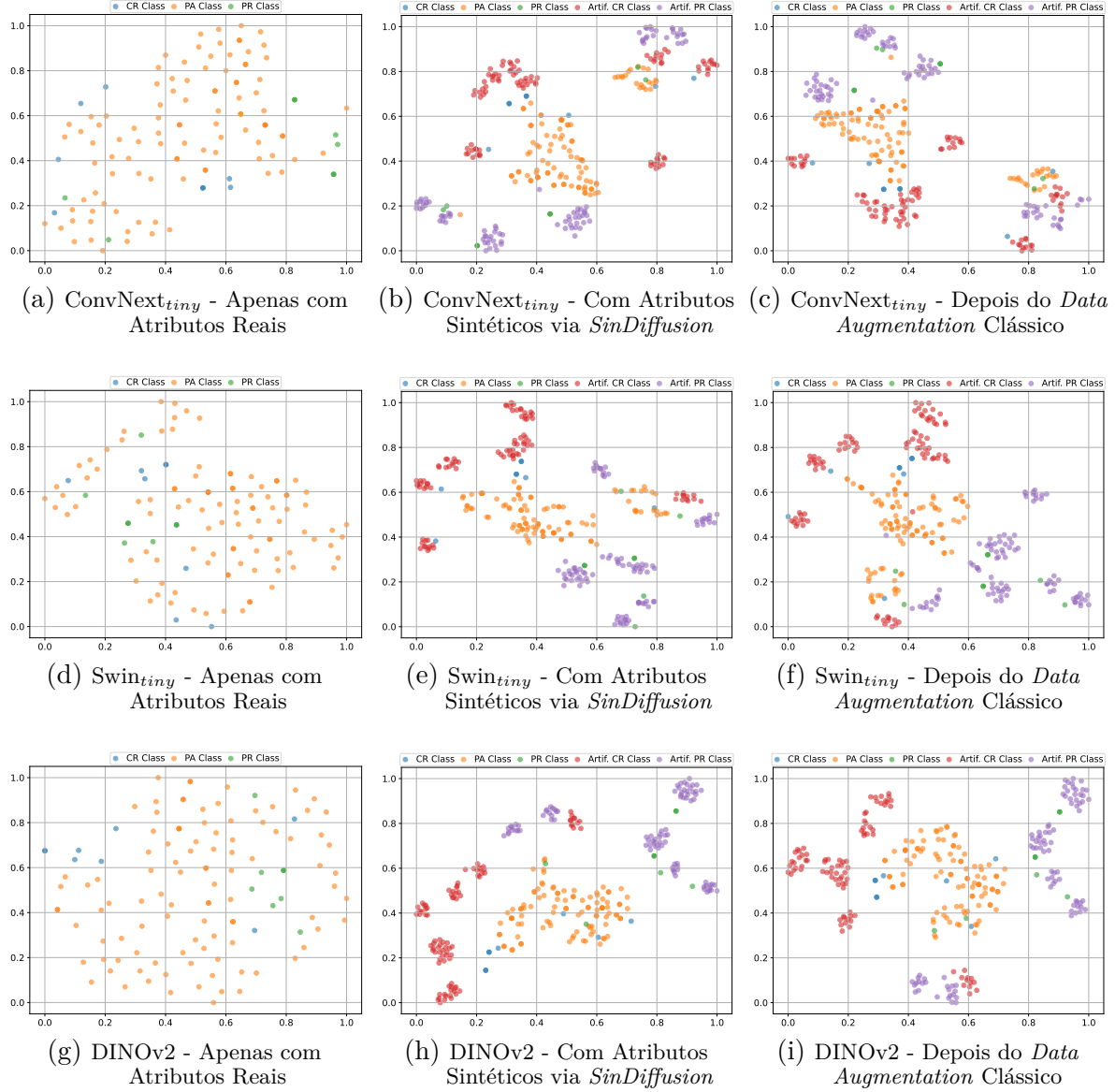
Nas Figuras, as amostras originais das classes CR, PA e PR são representadas, respectivamente, pelas cores azul, laranja e verde. As amostras sintéticas, por sua vez, são indicadas pelas cores vermelha e roxa, representando as classes CR e PR, respectivamente.

Figura 9 – Representação bidimensional por t-SNE dos atributos extraídos por técnicas clássicas para os dados reais (Figuras. 9a, 9d e 9g), após o aumento de dados sintéticos utilizando o *SinDiffusion* com ajuste de histograma (Figuras. 9b, 9e e 9h) e, por fim, após a aplicação das técnicas clássicas de aumento de dados (Figuras. 9c, 9f e 9i)..



Fonte: Autor.

Figura 10 – Representação bidimensional por t-SNE dos atributos extraídos por técnicas baseadas em *deep learning* para os dados reais (Figuras. 10a, 10d e 10g), após o aumento de dados sintéticos utilizando o *SinDiffusion* com ajuste de histograma (Figuras. 10b, 10e e 10h) e, por fim, após a aplicação das técnicas clássicas de aumento de dados (Figuras. 10c, 10f e 10i)..



Fonte: Autor.

Nas Figuras 9 e 10, observa-se uma acentuada formação de agrupamentos dos atributos sintéticos, indicando que as imagens geradas pelo *SinDiffusion* para cada amostra real das classes CR e PR apresentam características altamente semelhantes entre si. Esse comportamento torna-se particularmente evidente nos atributos extraídos pelo GLCM (Figura 9h) e em todos os extratores baseados em *deep learning* (Figuras. 10b, 10e e 10h). Tal elevado grau de homogeneidade entre as amostras sintéticas pode limitar a capacidade de generalização dos classificadores, uma vez que reduz a diversidade global do conjunto

de dados.

Por outro lado, a aplicação das técnicas clássicas de aumento de dados promoveu um aumento na variabilidade dos atributos extraídos a partir das amostras sintéticas. Esse efeito é mais evidente nas Figuras 9f, 9i e 10i, correspondentes, respectivamente, aos extratores LBP, GLCM e DINOv2. Em contraste, os extratores mais complexos ConvNeXt_{tiny} e Swin_{tiny} apresentaram alterações pouco significativas na distribuição dos atributos antes e após a aplicação do aumento de dados (Figuras 10c e 10f).

O impacto limitado das técnicas clássicas de aumento de dados sobre a variabilidade dos atributos sintéticos para os modelos ConvNeXt_{tiny} e Swin_{tiny} (Figuras 10c e 10f) sugere que esses extratores apresentam baixa sensibilidade às transformações aplicadas. Por se tratar de modelos de *deep learning* com elevada capacidade de abstração, pequenas perturbações visuais introduzidas por técnicas como *ColorJitter*, *Blur*, *Horizontal Flip* e *Rotation* tendem a modificar apenas uma fração reduzida das características extraídas. Essa invariância resulta na geração de representações sintéticas altamente semelhantes entre si, o que pode induzir efeitos de agrupamento e reduzir a diversidade efetiva das amostras utilizadas durante o treinamento dos classificadores.

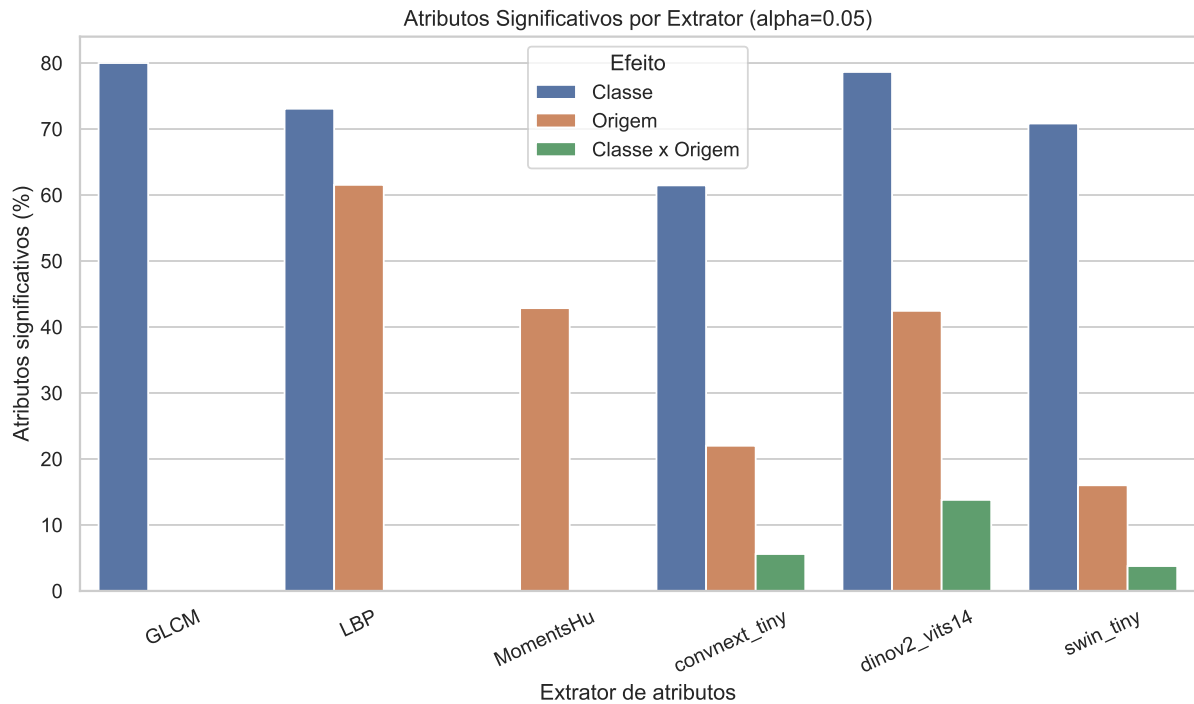
5.1.4 Análise Estatística dos Atributos

A análise bidimensional apresentou a distribuição visual dos atributos antes e após a inclusão das amostras artificiais, possibilitando observar uma redistribuição espacial das amostras. No entanto, é necessária uma análise estatística para extrair informações relevantes das amostras artificialmente criadas em relação às amostras reais.

Para isso, foi aplicada a técnica ANOVA (*Analysis of Variance*), ou Análise de Variância. Trata-se de um método estatístico utilizado para verificar a existência de diferenças significativas entre as médias de dois ou mais grupos. O objetivo principal dela é avaliar se as médias dos grupos analisados podem ser consideradas estatisticamente equivalentes ou se apresentam diferenças relevantes.

A ANOVA foi aplicada com o objetivo de avaliar três aspectos principais dos atributos extraídos por cada um dos extratores considerados, com foco nas amostras pertencentes às classes minoritárias CR e PR que geraram amostras artificiais. O primeiro aspecto corresponde ao fator classe, que avalia a existência de diferenças estatisticamente significativas entre as distribuições das classes CR e PR. O segundo aspecto refere-se ao fator origem, que investiga possíveis discrepâncias entre amostras reais e artificiais. Por fim, o terceiro aspecto considera o termo de interação entre classe e origem, permitindo verificar se o efeito da origem dos dados se manifesta de maneira uniforme entre as classes ou se há variações específicas por classe. A Figura 11 apresenta o histograma ANOVA com indicação de porcentagem de significância dos atributos para cada extrator utilizado.

Figura 11 – Análise ANOVA em histograma dos atributos de cada extrator para distribuição estatística entre classes minoritárias (CR e PR), entre origem (real e artificial) e, entre classe e origem.



Fonte: Autor.

Dentre os extratores de atributos clássicos, o GLCM apresentou o resultado mais favorável. Pela ANOVA, 80% dos atributos tiveram efeito significativo de classes, indicando que o extrator consegue diferenciar estatisticamente CR e PR. Ao passo que 0% dos atributos tiveram efeito significativo de origem e, consequentemente, interação. Esse comportamento sugere que o GLCM reconhece diferenças microestruturais entre as classes minoritárias, mas não percebe diferenças sistemáticas entre a imagem real e a sintética.

O LBP também é discriminativo entre as classes, apresentando 73,1% de efeito dos atributos entre classes. Em contrapartida, 61,5% dos atributos apresentaram efeito significativo de origem, indicando forte diferença estatística entre dados reais e sintéticos. Além disso, o efeito de interação foi de 0%, sugerindo que essa diferença real contra sintético é um comportamento global e não depende da classe. Com base nisso, o LBP separa bem CR e PR, mas também acaba capturando assinaturas estatísticas dos dados sintéticos. Isso é um indicativo que os sintéticos podem ter diferenças texturais detectáveis por LBP.

Dentre os extratores clássicos, o *Moments Hu* obteve o pior desempenho, pois não teve capacidade de separar as classes CR e PR, ao mesmo tempo que 42,9% dos atributos apresentaram efeito significativo de origem. Ou seja, este extrator não parece útil para diferenciar CR e PR, mas detecta diferenças entre sintética e real, sendo um sinal desfavorável, pois ele percebe mais a origem dos dados do que a classe microestrutural.

A análise de variância aplicada aos extratores de atributos baseados em *deep learning* evidenciou um comportamento consistente no que diz respeito à discriminação entre as classes minoritárias CR e PR. Todos os modelos apresentaram elevada proporção de atributos com efeito significativo para o fator classe, com destaque para o DINOv2 (78,6%), seguido pelo Swin-Tiny (70,8%) e pelo ConvNeXt-Tiny (61,5%), indicando forte capacidade de separação entre as classes no espaço latente. Em contrapartida, observou-se variação no impacto do fator origem, sendo relativamente baixo para o Swin-Tiny (16,0%) e moderado para o ConvNeXt-Tiny (22,0%), mas significativamente mais elevado para o DINOv2 (42,4%), sugerindo maior sensibilidade deste último às diferenças entre amostras reais e artificiais.

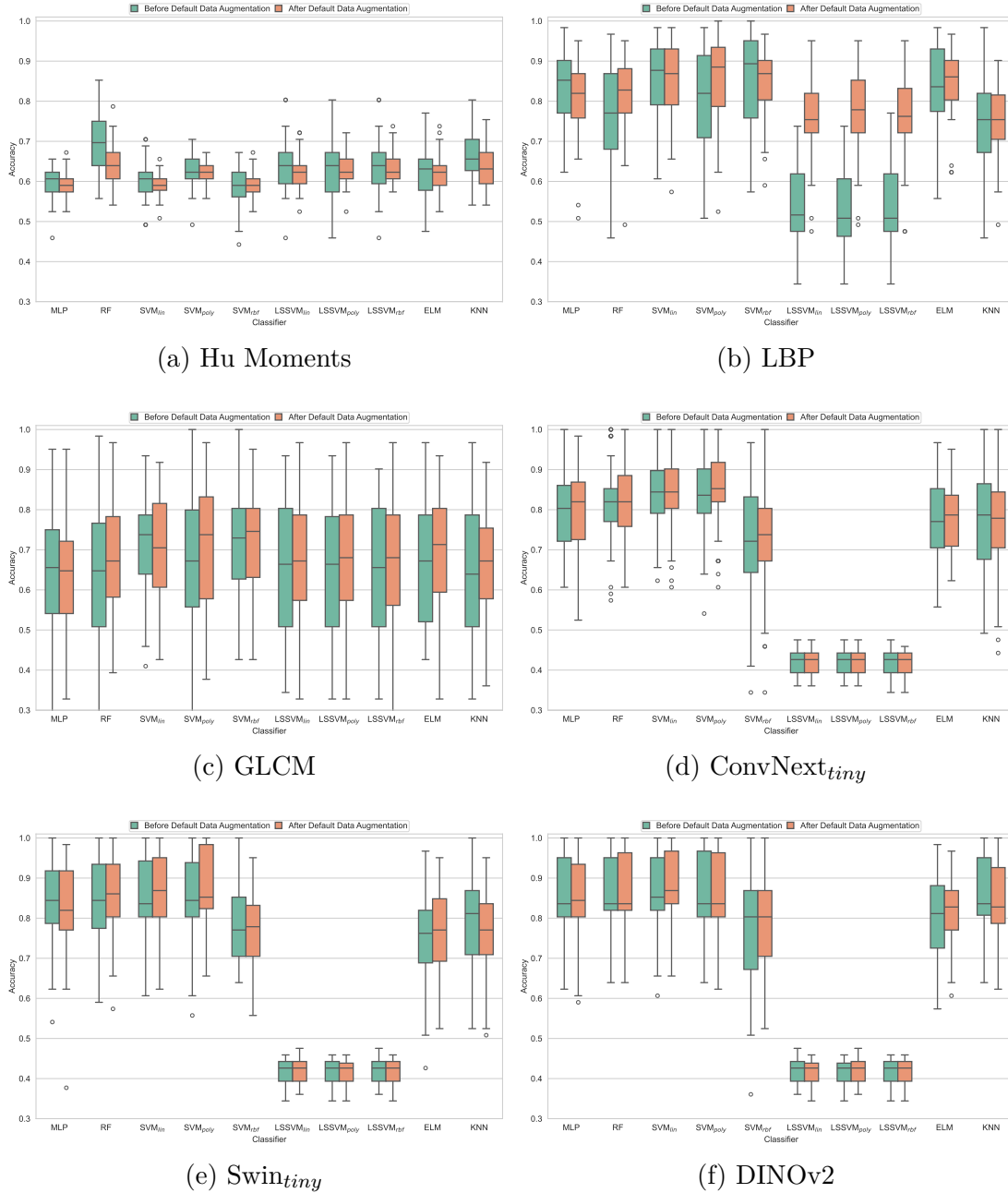
Em relação ao termo de interação entre classe e origem, todos os modelos apresentaram baixa incidência de atributos significativas, com valores de 3,8% para o Swin-Tiny, 5,6% para o ConvNeXt-Tiny e 13,8% para o DINOv2, indicando que, de modo geral, o efeito da origem tende a se manifestar de forma relativamente uniforme entre as classes. Ainda assim, o DINOv2 apresenta uma interação mais pronunciada, sugerindo maior variabilidade na influência da origem dependendo da classe considerada. De forma consolidada, os resultados indicam que, embora todos os extratores profundos sejam eficazes na discriminação entre CR e PR, o Swin-Tiny apresenta o melhor equilíbrio entre capacidade discriminativa e baixa sensibilidade à origem dos dados, enquanto o ConvNeXt-Tiny ocupa uma posição intermediária e o DINOv2, apesar de altamente discriminativo, demonstra maior propensão a capturar características associadas à origem artificial dos dados.

5.1.5 Resultados de Simulação

Para avaliar se o aumento da variabilidade introduzido pelo processo de aumento de dados sintéticos resultaria em uma melhora no desempenho dos classificadores, foram realizadas simulações comparando os cenários antes e após a aplicação das técnicas clássicas de *data augmentation*. Os resultados das simulações foram obtidos a partir de 50 execuções independentes e sintetizados por meio de diagramas de caixa (*boxplots*).

A Figura 12 apresenta os diagramas de caixa correspondentes aos extratores clássicos e aos extratores baseados em *deep learning*. Cada Figura é composta por subgráficos individuais para cada extrator, nos quais são apresentados os resultados de acurácia dos classificadores antes (diagramas de caixa em verde) e após (diagramas de caixa em laranja) a aplicação das técnicas de aumento de dados.

Figura 12 – Comparação, por meio de boxplots, da acurácia dos classificadores na fase de teste antes (boxplots verdes) e após (boxplots laranja) a aplicação de técnicas padrão de aumento de dados, ao longo de 50 execuções independentes para cada um dos extratores de características.



Fonte: Autor.

Os resultados apresentados na Figura 12 indicam que as características extraídas pelos extratores clássicos (Figuras 12a, 13a e 12c) apresentaram uma maior discrepância de desempenho antes e após a aplicação das técnicas de aumento de dados. Observa-se, ainda, que, na maioria dos casos, o desvio padrão dos resultados após o aumento de dados foi reduzido, sugerindo uma melhora consistente na estabilidade das simulações.

Os resultados dos boxplots para os extratores baseados em *deep learning* (Figuras 12d, 12e e 12f) indicam uma menor variação de desempenho antes e após a aplicação do aumento de dados. Esse comportamento corrobora a hipótese de que tais extratores apresentam baixa sensibilidade às transformações aplicadas. Como consequência, a alta dimensionalidade dos vetores de características gerados por esses modelos foi minimamente afetada pelas variações introduzidas nas imagens sintéticas por meio das técnicas selecionadas.

Os resultados das simulações apresentados nos boxplots indicam uma melhoria geral no desempenho após a aplicação das técnicas de aumento de dados, especialmente no caso dos extratores clássicos. Esse comportamento sustenta a metodologia adotada, a qual foi validada e estabelecida como base para as simulações de classificação do aço elétrico GNO apresentadas na seção seguinte.

5.2 Simulações dos Classificadores

Foi aplicada a técnica de *grid search* para a otimização dos hiperparâmetros de cada classificador, empregando validação cruzada com cinco partições (*5-fold cross-validation*, *5-fold CV*) ao longo de todo o processo. Em cada execução independente, a validação cruzada foi realizada com o objetivo de garantir uma estimativa confiável dos hiperparâmetros.

Para os classificadores SVM e LSSVM, foram avaliados os *kernels* lineares, polinomiais e RBF (radial basis function). No caso do LSSVM, o parâmetro de regularização γ foi ajustado por meio de validação cruzada *5-fold*, bem como o grau do polinômio (d) e a largura do *kernel* gaussiano (σ). Para o classificador SVM, o parâmetro C , responsável por controlar o compromisso entre a maximização da margem e o erro de classificação, foi igualmente ajustado por validação cruzada.

O número de neurônios na camada oculta (q) foi considerado o principal hiperparâmetro ajustado para a MLP com uma única camada oculta. Os demais parâmetros, tais como a taxa de aprendizado, o número de épocas e o uso de dropout, foram definidos empiricamente com o objetivo de limitar a complexidade do modelo e reduzir o custo computacional. Em contrapartida, para a rede ELM, foram otimizados tanto o número de neurônios na camada oculta (q) quanto a função de ativação empregada, considerando as funções ReLU, tangente hiperbólica e sigmoide logística.¹

Para o classificador KNN, foi adotada empiricamente a distância euclidiana, sendo o número de vizinhos mais próximos (k) ajustado como o principal hiperparâmetro. No caso do classificador RF, foram otimizados os hiperparâmetros relacionados ao número de

¹ *Dropout* é uma técnica de regularização em redes neurais na qual uma fração dos neurônios é desativada aleatoriamente durante o treinamento, reduzindo o sobreajuste e promovendo representações mais robustas.

estimadores (árvores), bem como ao número de nós e de folhas por árvore, explorando a eficiência computacional do modelo.

Para validar os resultados das simulações computacionais, foram utilizados a média e o desvio padrão da acurácia, bem como os valores médios de PPV, sensibilidade e F-score.

Matematicamente, acurácia (Acc) (Bishop; Nasrabadi, 2006) ou taxa de classificação (CR - *Classification Rate*) é calculada como

$$Acc = CR = \frac{TP + TN}{TP + TN + FP + FN}, \quad (5.1)$$

Onde TP, TN, FP, e FN representam os valores de verdadeiros positivos (*True Positive*), verdadeiros negativos (*True Negative*), falsos positivos (*False Positive*), e falsos negativos (*False Negative*), respectivamente. Além disso o Valor Positivo Verdadeiro (PPV - *Positive Predict Value*), Sensibilidade (Se - *Sensitivity*), e F-score (Bishop; Nasrabadi, 2006) são descritas por

$$PPV = \frac{TP}{TP + FP}, \quad (5.2)$$

$$Se = \frac{TP}{TP + FN}, \quad (5.3)$$

$$F-Score = 2 \cdot \frac{PPV \cdot Se}{PPV + Se}. \quad (5.4)$$

A métrica PPV (Eq. 5.2), também conhecida como Precisão, indica a proporção de cada amostra classificada como pertencente à classe que verdadeiramente a pertence. A métrica sensibilidade (Eq. 5.3), também conhecida como *recall*, indica a proporção de cada amostra classificada como pertencente à classe em razão do número total de amostras que verdadeiramente pertencem àquela classe. Por fim, o F-score (Eq. 5.4) pode ser interpretado como a média harmônica da PPV e sensibilidade. Por tanto, quanto mais próximo de 1 as métricas estejam, melhor é a performance do modelo.

Ao longo das simulações também foram calculados a média e o desvio padrão da acurácia na fase de treinamento, com o objetivo de investigar possíveis indícios de sobreajuste (*overfitting*). A análise dessas métricas durante o treinamento é particularmente relevante, considerando o número limitado de amostras reais nas classes minoritárias CR e PR, bem como a elevada cardinalidade dos atributos extraídos.

Adicionalmente, matrizes de confusão agregadas foram construídas considerando todas as execuções independentes para os classificadores com melhor desempenho na fase de teste. Essa análise tem como objetivo identificar possíveis padrões de erro de

classificação entre as classes, uma vez que a observação das métricas de média de PPV, sensibilidade e F-score não mostra o desempenho por classe.

As Tabelas 4 e 5 apresentam a média e o desvio padrão da acurácia nas fases de treinamento e teste, bem como os valores médios de precisão (PPV), sensibilidade e F-score na fase de teste, calculados ao longo de 50 execuções independentes.

Tabela 4 – Média e desvio padrão das fases de treinamento e teste para cada estimador, bem como os valores de PPV, sensibilidade (Se) e F-score referentes às amostras de teste, calculados ao longo de 50 execuções independentes para os extratores de atributos clássicos.

Extrator	Classificador	<i>Fase de Treino</i>		<i>Fase de Teste</i>		
		<i>Acc (%)</i>	<i>Acc (%)</i>	<i>PPV</i>	<i>Se</i>	<i>f-score</i>
Hu Mom.	<i>MLP</i>	64,65 ± 5,36	59,28 ± 3,52	0,586	0,588	0,587
	<i>RF</i>	75,58 ± 3,59	63,93 ± 5,04	0,619	0,634	0,607
	<i>SVM_{lin.}</i>	63,92 ± 1,90	59,41 ± 2,98	0,582	0,589	0,585
	<i>SVM_{rbf}</i>	66,65 ± 1,63	62,03 ± 2,44	0,594	0,615	0,590
	<i>SVM_{poly.}</i>	63,09 ± 1,53	59,28 ± 3,01	0,583	0,588	0,585
	<i>LSSVM_{lin.}</i>	68,29 ± 2,58	62,36 ± 4,07	0,595	0,618	0,590
	<i>LSSVM_{rbf}</i>	68,53 ± 2,31	63,28 ± 3,65	0,605	0,627	0,594
	<i>LSSVM_{poly.}</i>	68,63 ± 2,34	63,41 ± 3,41	0,606	0,628	0,598
	<i>ELM</i>	68,89 ± 3,07	61,90 ± 4,72	0,594	0,614	0,580
	<i>KNN</i>	69,03 ± 3,41	63,44 ± 5,28	0,610	0,629	0,615
LBP	<i>MLP</i>	95,55 ± 3,31	80,16 ± 8,68	0,801	0,801	0,801
	<i>RF</i>	99,69 ± 0,83	81,41 ± 9,08	0,816	0,813	0,813
	<i>SVM_{lin.}</i>	97,67 ± 1,87	84,95 ± 9,03	0,855	0,848	0,849
	<i>SVM_{rbf}</i>	99,75 ± 0,59	85,67 ± 10,21	0,861	0,855	0,855
	<i>SVM_{poly.}</i>	96,68 ± 1,22	84,89 ± 8,82	0,856	0,847	0,848
	<i>LSSVM_{lin.}</i>	99,97 ± 0,13	75,74 ± 9,96	0,773	0,754	0,752
	<i>LSSVM_{rbf}</i>	99,99 ± 0,06	76,66 ± 10,01	0,781	0,764	0,762
	<i>LSSVM_{poly.}</i>	100,0 ± 0,00	76,03 ± 10,10	0,775	0,757	0,755
	<i>ELM</i>	98,40 ± 1,09	84,43 ± 7,70	0,848	0,843	0,843
	<i>KNN</i>	95,35 ± 1,62	75,08 ± 9,08	0,759	0,748	0,743
GLCM	<i>MLP</i>	90,90 ± 8,48	63,15 ± 13,70	0,641	0,631	0,633
	<i>RF</i>	99,60 ± 0,61	68,36 ± 13,87	0,744	0,679	0,679
	<i>SVM_{lin.}</i>	92,89 ± 2,15	70,10 ± 12,01	0,713	0,699	0,701
	<i>SVM_{rbf}</i>	99,06 ± 1,07	71,15 ± 15,03	0,765	0,708	0,710
	<i>SVM_{poly.}</i>	97,75 ± 0,89	72,62 ± 12,70	0,784	0,723	0,726
	<i>LSSVM_{lin.}</i>	99,97 ± 0,14	68,36 ± 14,07	0,769	0,679	0,678
	<i>LSSVM_{rbf}</i>	99,97 ± 0,14	68,62 ± 13,94	0,768	0,682	0,681
	<i>LSSVM_{poly.}</i>	99,98 ± 0,08	68,46 ± 14,21	0,772	0,680	0,680
	<i>ELM</i>	98,95 ± 0,94	69,31 ± 14,39	0,744	0,690	0,691
	<i>KNN</i>	97,08 ± 1,08	66,98 ± 11,66	0,741	0,666	0,666

Observando a Tabela 4, o extrator de atributos LBP foi o único, entre os extratores clássicos, a alcançar uma acurácia superior a 80,00%. Os melhores resultados para LBP foram obtidos com os classificadores ELM e SVM, utilizando *kernel* linear, RBF e polinomial. Entre eles, a SVM com *kernel* RBF apresentou o melhor desempenho, atingindo 85,67% de acurácia na fase de teste. O ELM destacou-se por apresentar um dos menores desvios padrão entre os métodos de melhor desempenho, com $Acc = 84,43\% \pm 7,70\%$.

Tabela 5 – Média e desvio padrão das fases de treinamento e teste para cada estimador, bem como os valores de PPV, sensibilidade (Se) e F-score referentes às amostras de teste, calculados ao longo de 50 execuções independentes para os extratores de atributos baseados em *deep learning*.

Extractor	Classificador	<i>Fase de Treino</i>		<i>Fase de Teste</i>		
		<i>Acc (%)</i>	<i>Acc (%)</i>	<i>PPV</i>	<i>Se</i>	<i>f-score</i>
ConvNeXt _{Tiny}	<i>MLP</i>	100,0 ± 0,00	79,97 ± 11,52	0,802	0,797	0,795
	<i>RF</i>	99,88 ± 0,21	82,56 ± 10,18	0,825	0,823	0,821
	<i>SVM_{lin.}</i>	99,25 ± 0,60	84,46 ± 9,90	0,845	0,842	0,840
	<i>SVM_{rbf}</i>	99,47 ± 0,41	85,31 ± 9,58	0,853	0,851	0,849
	<i>SVM_{poly.}</i>	99,95 ± 0,22	73,05 ± 13,11	0,772	0,726	0,719
	<i>LSSVM_{lin.}</i>	100,0 ± 0,00	41,74 ± 2,71	0,489	0,420	0,412
	<i>LSSVM_{rbf}</i>	100,0 ± 0,00	41,70 ± 2,75	0,536	0,420	0,404
	<i>LSSVM_{poly.}</i>	100,0 ± 0,00	41,54 ± 2,76	0,505	0,418	0,408
	<i>ELM</i>	99,07 ± 0,85	78,46 ± 8,16	0,784	0,782	0,778
	<i>KNN</i>	97,89 ± 0,75	76,69 ± 11,71	0,785	0,763	0,760
Swin _{Tiny}	<i>MLP</i>	100,0 ± 0,00	82,59 ± 11,01	0,835	0,824	0,824
	<i>RF</i>	99,99 ± 0,06	85,38 ± 10,09	0,857	0,852	0,852
	<i>SVM_{lin.}</i>	99,89 ± 0,33	86,52 ± 9,91	0,873	0,863	0,863
	<i>SVM_{rbf}</i>	99,98 ± 0,08	86,85 ± 10,32	0,873	0,866	0,865
	<i>SVM_{poly.}</i>	99,92 ± 0,20	76,98 ± 9,55	0,827	0,766	0,768
	<i>LSSVM_{lin.}</i>	100,0 ± 0,00	41,87 ± 2,65	0,579	0,422	0,405
	<i>LSSVM_{rbf}</i>	100,00 ± 0,00	41,44 ± 2,72	0,505	0,417	0,405
	<i>LSSVM_{poly.}</i>	100,0 ± 0,00	41,31 ± 3,02	0,456	0,415	0,407
	<i>ELM</i>	99,53 ± 0,84	77,15 ± 10,18	0,780	0,769	0,766
	<i>KNN</i>	98,67 ± 0,83	76,49 ± 10,10	0,819	0,761	0,764
DINOv2	<i>MLP</i>	99,96 ± 0,29	85,25 ± 9,75	0,852	0,851	0,850
	<i>RF</i>	99,67 ± 0,31	86,33 ± 9,52	0,864	0,861	0,860
	<i>SVM_{lin.}</i>	99,47 ± 0,73	87,25 ± 9,95	0,873	0,870	0,869
	<i>SVM_{rbf}</i>	99,57 ± 0,47	85,93 ± 9,85	0,861	0,857	0,855
	<i>SVM_{poly.}</i>	99,81 ± 0,32	78,75 ± 11,61	0,829	0,784	0,780
	<i>LSSVM_{lin.}</i>	100,0 ± 0,00	41,28 ± 2,87	0,482	0,415	0,411
	<i>LSSVM_{rbf}</i>	100,0 ± 0,00	41,80 ± 2,61	0,577	0,421	0,413
	<i>LSSVM_{poly.}</i>	100,0 ± 0,00	41,57 ± 2,70	0,534	0,419	0,409
	<i>ELM</i>	99,14 ± 0,64	81,08 ± 8,95	0,813	0,808	0,805
	<i>KNN</i>	98,19 ± 0,58	84,59 ± 9,37	0,848	0,844	0,842

Por outro lado, os classificadores KNN e LSSVM com *kernel* linear, RBF e polinomial apresentaram desempenho inferior na fase de teste, apesar de alcançarem alta acurácia durante o treinamento, com valores atingindo até 100% no treino e caindo drasticamente para números inferiores a 77% no teste. Essa discrepância indica a presença de *overfitting*, especialmente na LSSVM, em que a diferença de acurácia entre treinamento e teste chegou a aproximadamente 25 pontos percentuais. Um padrão semelhante também foi observado para o extrator de atributos GLCM.

A principal razão para esse comportamento provavelmente decorre da limitada disponibilidade de amostras reais para as classes minoritárias, CR e PR, o que restringe sua representatividade no espaço de atributos. Consequentemente, mesmo com a aplicação do aumento de dados utilizando o *SinDiffusion* e técnicas clássicas, a introdução de

variabilidade por meio de dados sintéticos mostrou-se insuficiente para determinados extratores de atributos e classificadores.

Ao avaliar os resultados dos classificadores para os extratores de atributos baseados em *deep learning*, apresentados na Tabela 5, destacam-se os classificadores que alcançaram acurácia superior a 80% em todos os extratores. Tanto o ConvNeXt_{tiny} quanto o Swin_{tiny} apresentaram desempenho semelhante, com os SVMs com *kernels* lineares e RBF se destacando como os melhores classificadores. O RF também apresentou resultados relevantes para o Swin_{tiny}. Adicionalmente, alguns classificadores apresentaram o mesmo comportamento de *overfitting* observado nas simulações com extratores clássicos.

Entretanto, a análise de variância (ANOVA) aplicada às características extraídas permite uma interpretação mais criteriosa desses resultados. Observou-se que, embora todos os extratores apresentem boa capacidade discriminativa entre as classes CR e PR, há diferenças relevantes quanto à sensibilidade à origem dos dados. Nesse contexto, o Swin_{tiny} apresentou o melhor equilíbrio entre alta significância para o fator *class* e baixa significância para o fator *origin*, indicando que suas representações são capazes de separar adequadamente as classes sem incorporar de forma significativa a assinatura estatística das amostras artificiais. Esse comportamento sugere maior robustez e melhor capacidade de generalização em cenários com dados sintéticos.

Em contraste, o extrator DINOv2 apresentou o maior número de classificadores com acurácia média de teste superior a 80%, alcançando o melhor resultado global, com $Acc = 87,25\%$, utilizando uma SVM linear. Esse desempenho indica que o modelo ViT-S/14 do DINOv2 é altamente eficaz na captura de características discriminativas relevantes para a classificação microestrutural de aços elétricos GNO.

Contudo, a análise ANOVA revelou que o DINOv2 também apresenta a maior proporção de características com efeito significativo para o fator de origem, evidenciando maior sensibilidade às diferenças entre amostras reais e artificiais. Esse resultado sugere que parte de seu elevado desempenho pode estar associada à captura de padrões relacionados à origem dos dados, e não exclusivamente às características intrínsecas das classes. Dessa forma, embora o DINOv2 apresente elevada capacidade discriminativa, o Swin_{tiny} se destaca como a alternativa mais equilibrada, por preservar a separabilidade entre classes ao mesmo tempo em que minimiza a influência da assinatura artificial, resultando em representações mais consistentes e confiáveis para o problema proposto.

Destaca-se que o desempenho do Swin_{tiny} foi obtido mesmo considerando o distanciamento entre o domínio das imagens de microscopia óptica e os dados utilizados em seu pré-treinamento, o que evidencia sua capacidade de extrair representações robustas e transferíveis. Além disso, o elevado desempenho aliado ao uso de classificadores lineares simples indica que as características extraídas apresentam alto grau de separabilidade entre as classes, sem depender de estruturas complexas de decisão. Em conjunto com

os resultados da análise ANOVA, esse comportamento reforça que o Swin_{tiny} é capaz de capturar predominantemente padrões intrínsecos às classes, minimizando a influência da origem dos dados e contribuindo para uma melhor capacidade de generalização.

Além disto, o comportamento observado para o estimador LSSVM evidenciou elevado *overfitting*, principalmente para os extratores baseados em *deep learning*, atingindo acurácia próxima de 100% na fase de treino e apresentando queda significativa na fase de teste. Tal fenômeno pode ser explicado pela própria formulação do LSSVM, que, diferentemente do SVM convencional, utiliza todos os vetores de treinamento na construção da solução ótima, tornando-se mais sensível à redundância presente nos dados.

Considerando que o conjunto de treinamento foi expandido com amostras sintéticas em uma proporção de 1:12 por meio do *Sindiffusion* aplicadas às imagens reais de microscopia óptica, houve um aumento significativo de redundância estatística entre as classes. Essa redundância impacta diretamente a matriz de *kernel*, tornando-a altamente correlacionada e favorecendo soluções com baixa capacidade de generalização. A análise da ANOVA, esse efeito é ainda mais crítico para extratores que apresentam maior sensibilidade à origem dos dados, como o DINOv2, potencializando o *overfitting* observado. Em contraste, o SVM convencional apresentou menor grau de *overfitting*, devido à sua propriedade esparsa, na qual apenas os vetores de suporte contribuem para a definição da fronteira de decisão, reduzindo a influência de redundâncias introduzidas pelas amostras sintéticas.

Outra métrica importante a se destacar é o tempo de processamento e a escalabilidade da metodologia proposta, pois elas impactam diretamente a viabilidade prática da metodologia para a classificação de aços elétricos GNO. A Tabela 6 apresenta a complexidade operacional aproximada dos quatro melhores extratores de atributos.

Tabela 6 – Escalabilidade computacional dos extratores de atributos avaliados.

Extrator	Complexidade Aproximada	Escalabilidade
LBP	$O(HWP) \approx O(7,16 \times 10^6)$	Escala linearmente com o número de pixels da imagem ($O(HW)$).
ConvNeXt-Tiny	$O(\sum_s D_s H_s W_s (7^2 C_s + 8 C_s^2)) \approx O(4,27 \times 10^9)$	Escala aproximadamente de forma linear com a resolução espacial, com forte dependência do número de canais.
DINOv2 ViT-S/14	$O(L(N^2 d + N d^2)) \approx O(7,59 \times 10^8)$	Escala quadraticamente com o número de tokens ($O(N^2)$), sendo sensível ao aumento da resolução.
Swin-Tiny	$O(\sum_s D_s N_s (M^2 d_s + d_s^2)) \approx O(4,17 \times 10^8)$	Escala aproximadamente de forma linear com o número de tokens, devido ao uso de janelas locais.

Os termos utilizados nas expressões de complexidade representam as principais variáveis estruturais de cada modelo. As variáveis H e W correspondem às dimensões espaciais da imagem de entrada, enquanto P representa o número de vizinhos considerados no descritor LBP. Para os modelos baseados em redes profundas, C_s e d representam a dimensionalidade dos canais ou embeddings, D_s e L indicam o número de camadas ou blocos, e N corresponde ao número de tokens ou patches gerados a partir da imagem. No caso do Swin, M define o tamanho da janela de atenção local. Dessa forma, a análise de complexidade permite compreender como o custo computacional cresce em função tanto da resolução da imagem quanto da estrutura interna de cada modelo.

A análise da Tabela 6 evidencia diferenças relevantes na escalabilidade computacional entre os extratores de atributos avaliados. O descritor LBP apresenta a menor complexidade, com crescimento linear em função do número de pixels, o que o torna altamente eficiente do ponto de vista computacional. Entretanto, conforme observado na análise ANOVA, ele apresenta elevada proporção de atributos com efeito significativo para o fator origem, indicando forte sensibilidade à origem dos dados, o que limita sua capacidade de generalização em cenários mais realistas.

Por outro lado, os extratores baseados em *deep learning* apresentam maior custo computacional, porém oferecem maior robustez na representação dos dados. Nesse contexto, o DINOv2, embora altamente discriminativo, apresenta menor escalabilidade devido ao termo quadrático associado ao número de tokens, além de maior sensibilidade à assinatura artificial. Em contraste, o Swin_{tiny} combina complexidade aproximadamente linear com baixa sensibilidade à origem dos dados destacando-se como a alternativa mais equilibrada. Dessa forma, o Swin_{tiny} apresenta um compromisso adequado entre custo computacional, escalabilidade e robustez estatística, tornando-se a escolha mais adequada para a aplicação proposta.

5.2.1 Matrizes de Confusão

Os valores médios de PPV, sensibilidade e F-score também refletiram as tendências observadas na acurácia em todos os classificadores. Para fornecer uma análise mais detalhada dos padrões de predição de cada classe, a Figura 13 apresenta as matrizes de confusão dos classificadores de melhor desempenho obtidas a partir dos extratores LBP, ConvNeXt, Swin e DINOv2. As matrizes são exibidas como *heatmaps*, com os valores correspondentes à agregação dos resultados de 50 execuções independentes.

De acordo com as matrizes de confusão na Figura 13), a classe PA apresentou o melhor desempenho em todos os modelos avaliados, o que pode ser atribuído ao seu conjunto mais amplo e diversificado de amostras reais. Em contrapartida, os classificadores demonstraram maior dificuldade na correta identificação das amostras das classes CR e PR. Essa confusão entre estados cristalográficos adjacentes evidencia uma limitação

crítica da metodologia, uma vez que tais erros podem comprometer sua aplicabilidade em cenários práticos de inspeção.

A classificação incorreta de um aço elétrico GNO no estado de recristalização primária como laminação a frio pode impactar diretamente o processo de recozimento na produção. Consequentemente, o material já recristalizado pode ser submetido novamente ao recozimento, evoluindo para um novo estado microestrutural conhecido como recristalização secundária. Nesse estado, as propriedades magnéticas são alteradas, desviando-se de forma significativa daquelas desejadas na recristalização primária.

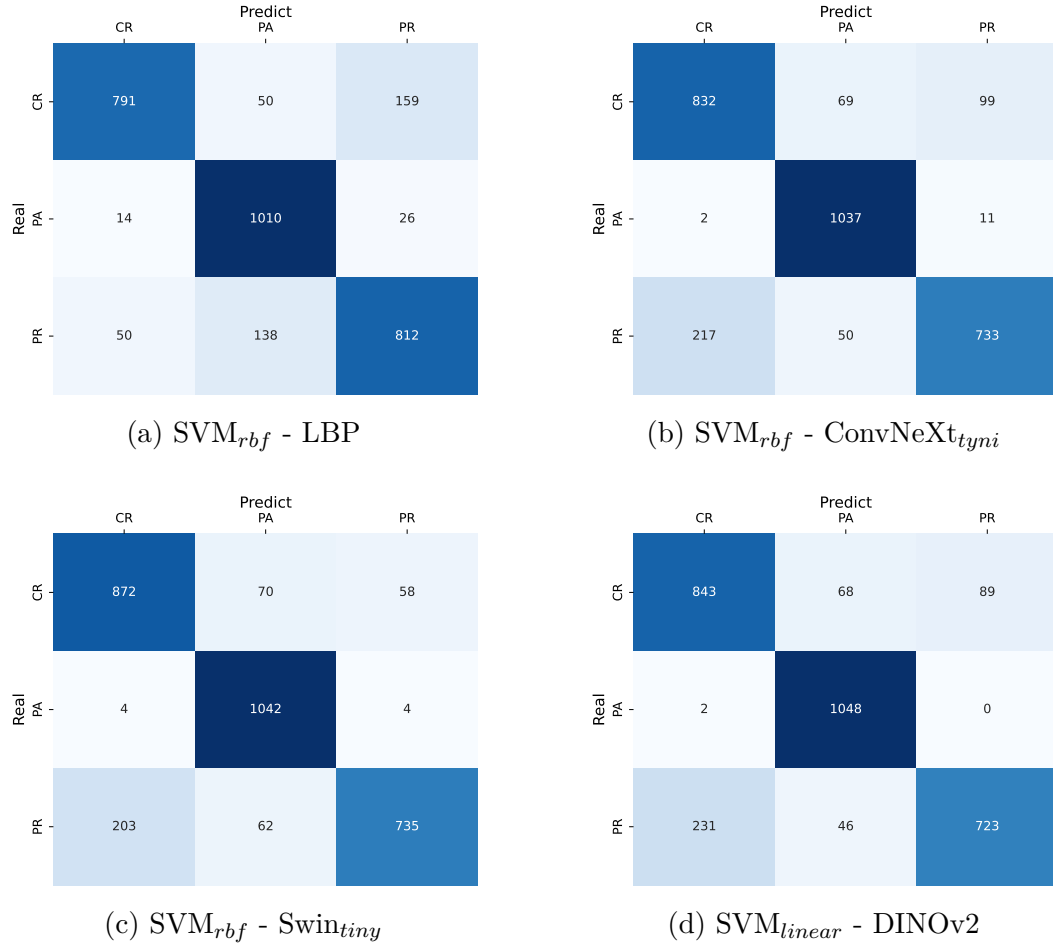
Esse comportamento está associado à baixa representatividade das classes CR e PR. A limitada variabilidade de amostras reais provavelmente favoreceu a formação de agrupamentos de características, prejudicando a capacidade de generalização e resultando em elevada acurácia durante o treinamento, seguida de uma queda perceptível no desempenho em teste, conforme apresentado na Tabela 4.

Além disso, embora o aumento de dados sintéticos contribua para mitigar a escassez de dados, ele pode mascarar parcialmente a capacidade de generalização dos classificadores em dados reais, ao reforçar padrões aprendidos a partir de amostras sintéticas. A análise de variância (ANOVA) corrobora essa interpretação, ao evidenciar que parte significativa das características extraídas apresenta sensibilidade ao fator de origem, indicando que os modelos incorporam padrões estatísticos associados à origem dos dados. Esse comportamento é refletido nas matrizes de confusão dos extratores baseados em *deep learning*, nas quais uma parcela significativa das amostras PR é classificada incorretamente como CR. Dessa forma, a confusão observada não está apenas associada à escassez de dados reais, mas também à influência da assinatura artificial introduzida pelas amostras sintéticas, reforçando a necessidade de ampliar o conjunto de dados reais para melhorar a generalização dos modelos.

Adicionalmente, a limitada diversidade do conjunto de dados decorre de sua finalidade original. O estudo do qual as amostras foram derivadas (Freitas, 2011) teve como foco a investigação dos fenômenos de recristalização durante a transição do estado laminado a frio para o estado primariamente recristalizado. Consequentemente, a maior parte das amostras pertence à classe parcialmente recristalizada, resultando em um conjunto de dados criticamente desbalanceado e com baixa variabilidade nas classes minoritárias.

O desempenho de classificação das classes minoritárias CR e PR é impactado por sua baixa representatividade, uma vez que cada classe contém apenas oito amostras reais. Essa limitada disponibilidade de dados pode ter levado à formação de agrupamentos após a geração das amostras sintéticas correspondentes, dificultando a capacidade de generalização do modelo na fase de teste. Por outro lado, uma estratégia eficaz para melhorar o desempenho dessas classes consiste em aumentar o número de amostras reais disponíveis.

Figura 13 – Matrizes de confusão exibidas como *heatmaps* para as amostras de teste, agregando os resultados de 50 execuções independentes dos classificadores de melhor desempenho, utilizando os extratores de atributos LBP, ConvNeXt, Swin e DINOv2.



De modo geral, os resultados das simulações indicam que a metodologia proposta apresenta desempenho satisfatório para aplicações práticas, com confiabilidade crescente à medida que aumentam a quantidade e a variabilidade das amostras de treinamento. Consequentemente, essa técnica pode auxiliar profissionais de metalografia na classificação dos estados microestruturais de aços elétricos GNO a partir de imagens de microscopia óptica, reduzindo, assim, a subjetividade inerente à análise humana.

Além disso, a metodologia pode ser implementada e validada de forma rápida, uma vez que a combinação do *SinDiffusion* com técnicas de aumento de dados permite sua aplicação em cenários com amostras limitadas, gerando dados sintéticos que favorecem o treinamento adequado dos classificadores.

6 Conclusão

Este estudo desenvolveu uma metodologia que integra ferramentas de aprendizado de máquina e aprendizado profundo para a classificação dos estados microestruturais de aços elétricos não orientados (GNO) a partir de imagens de microscopia óptica. Adicionalmente, técnicas de geração de imagens sintéticas por meio do *SinDiffusion*, associadas a métodos clássicos de aumento de dados, foram empregadas com o objetivo de enriquecer o conjunto de dados e viabilizar o treinamento de classificadores em cenários com amostras limitadas e desbalanceadas.

A metodologia proposta foi avaliada por meio de diferentes técnicas de extração de atributos e algoritmos de aprendizado de máquina. Entre os extratores clássicos, o descritor LBP foi o único a apresentar desempenho de classificação superior a 80%, destacando-se o SVM com *kernel* RBF, que atingiu acurácia de 85,67%. No entanto, a análise estatística baseada em ANOVA indicou elevada sensibilidade desse descritor à origem dos dados, evidenciando a presença de uma assinatura estatística associada às amostras sintéticas, o que limita sua capacidade de generalização em cenários mais realistas.

Os extratores baseados em *deep learning* apresentaram desempenho superior aos métodos clássicos, com todos os modelos avaliados gerando classificadores com acurácia acima de 80%. O SVM com *kernel* RBF destacou-se como o melhor classificador para os extratores ConvNeXt_{tiny} e Swin_{tiny}, com acurácias de 85,31% e 86,85%, respectivamente. O extrator DINOv2 apresentou o maior número de classificadores com desempenho elevado, alcançando a maior acurácia global com um SVM linear (87,33%). Entretanto, a análise de variância revelou que o DINOv2 também apresenta maior sensibilidade à origem dos dados, indicando que parte de seu desempenho pode estar associada à captura de padrões decorrentes da geração de amostras sintéticas. Em contrapartida, o Swin_{tiny} destacou-se como o extrator mais equilibrado, ao combinar elevada capacidade discriminativa com baixa sensibilidade à origem dos dados, indicando maior robustez e confiabilidade para a aplicação proposta.

As matrizes de confusão agregadas possibilitaram uma avaliação do desempenho dos melhores classificadores e extratores de atributos em nível de classe. Todos os modelos apresentaram excelente desempenho na classificação das amostras da classe Recristalização Parcial (*Partial Annealed* - PA). Em contraste, as matrizes revelaram que várias amostras da classe Laminado a Frio (*Cold Rolled* - CR) foram incorretamente classificadas como Recristalização Primária (*Primary Recrystallization* - PR). Esse comportamento evidencia que, mesmo com a aplicação de técnicas de aumento de dados, a limitada diversidade das amostras reais das classes CR e PR ainda afeta de forma significativa o desempenho das

simulações, além das assinaturas artificiais corroborarem diretamente em um desempenho superior em alguns extratores.

Os resultados indicam que o modelo proposto pode atuar como uma ferramenta de suporte à decisão para o controle de qualidade de aços elétricos GNO, permitindo a identificação automática do estágio microestrutural durante as observações em microscopia óptica e reduzindo o tempo de análise laboratorial. Além disso, graças à utilização de técnicas de aumento de dados, a metodologia apresenta potencial para aplicações em cenários com poucas amostras, possibilitando adaptação rápida a diferentes tipos de aço e contextos industriais.

Para trabalhos futuros, pretende-se estender a metodologia proposta a outros estados microestruturais de aços elétricos GNO, como o estado intermediário entre a recristalização primária e secundária. Adicionalmente, serão realizados testes com maior variabilidade de amostras reais das classes CR e PR, a fim de analisar a quantidade mínima de amostras de microscopia óptica necessária para melhorar o desempenho da classificação. Dessa forma, essa abordagem poderia ser expandida para apoiar uma gama mais ampla de aplicações práticas, abrangendo diferentes cenários industriais e de pesquisa.

6.1 Limitações do Trabalho

Apesar dos resultados promissores obtidos, o presente trabalho apresenta algumas limitações que devem ser consideradas na interpretação dos resultados e na generalização da metodologia proposta.

A primeira limitação refere-se ao tamanho e à diversidade do conjunto de dados reais utilizado. Embora técnicas de geração de dados sintéticos tenham sido empregadas para mitigar a escassez de amostras, a limitada variabilidade intrínseca das imagens reais pode ter influenciado a capacidade de generalização dos modelos, especialmente nas classes minoritárias CR e PR.

Adicionalmente, a utilização de dados sintéticos, apesar de contribuir para o balanceamento das classes, introduz uma assinatura estatística associada à origem dos dados, conforme evidenciado pela análise de variância (ANOVA). Essa característica pode levar os modelos a capturarem padrões não intrínsecos às microestruturas, impactando a robustez das representações aprendidas, além de distorcer os resultados de simulação, como foi o caso do extrator LBP.

Do ponto de vista computacional, embora modelos como o Swin_{tiny} apresentem bom equilíbrio entre desempenho e escalabilidade, sua aplicação pode ser limitada em cenários que envolvem dispositivos com recursos computacionais restritos, como sistemas embarcados baseados em microcontroladores. A necessidade de processamento de redes

profundas, aliada ao uso de operações de alta complexidade, pode inviabilizar a execução desses modelos em tempo real ou com baixo consumo energético. Dessa forma, a aplicação da metodologia proposta em ambientes industriais com restrições severas de hardware pode demandar estratégias adicionais, como otimização de modelos, compressão ou uso de arquiteturas mais leves.

Por fim, a metodologia proposta foi avaliada em um conjunto específico de condições experimentais, incluindo tipos de aço, parâmetros de processamento e condições de aquisição de imagem. Dessa forma, a generalização dos resultados para outros cenários industriais deve ser realizada com cautela, sendo necessária a validação adicional em bases de dados mais amplas e diversificadas.

Referências

- ABBASS, K. *et al.* A review of the global climate change impacts, adaptation, and sustainable mitigation measures. *Environmental science and pollution research*, Springer, v. 29, n. 28, p. 42539–42559, 2022.
- AHN, Y.-K. *et al.* Texture evolution of non-oriented electrical steel analyzed by ebsd and in-situ xrd during the phase transformation from γ to α . *Materials Today Communications*, Elsevier, v. 25, p. 101307, 2020.
- AYAD, A. *et al.* Quantifying primary recrystallization from ebsd maps of partially recrystallized states of an if steel. *Materials Characterization*, Elsevier, v. 171, p. 110773, 2021.
- AZIMI, S. M. *et al.* Advanced steel microstructural classification by deep learning methods. *Scientific reports*, Nature Publishing Group UK London, v. 8, n. 1, p. 2128, 2018.
- BABALOLA, E.-O.; ASAD, M. H.; BAIS, A. Soil surface texture classification using rgb images acquired under uncontrolled field conditions. *IEEE Access*, IEEE, v. 11, p. 67140–67155, 2023.
- BISHOP, C. M.; NASRABADI, N. M. *Pattern recognition and machine learning*. [S.l.]: Springer, 2006. v. 4.
- BRANDON, D.; KAPLAN, W. D. *Microstructural characterization of materials*. [S.l.]: John Wiley & Sons, 2013.
- BULGAREVICH, D. S. *et al.* Pattern recognition with machine learning on optical microscopy images of typical metallurgical microstructures. *Scientific reports*, Nature Publishing Group UK London, v. 8, n. 1, p. 2078, 2018.
- CHEN, B. *et al.* Recent progress of digital reconstruction in polycrystalline materials. *Archives of Computational Methods in Engineering*, Springer, p. 1–52, 2025.
- CHU, S.; ZHOU, B.; MAO, B. State of the art and prospects on the metallurgical design and manufacturing process of grain-oriented electric steels. *Journal of Magnetism and Magnetic Materials*, Elsevier, v. 614, p. 172739, 2025.
- COOPER, C. *et al.* Surface roughness prediction in machining using two-stage domain-incremental learning with input dimensionality expansion. *Journal of Manufacturing Systems*, Elsevier, v. 80, p. 503–510, 2025.
- COVER, T.; HART, P. Nearest neighbor pattern classification. *IEEE transactions on information theory*, IEEE, v. 13, n. 1, p. 21–27, 1967.
- CROITORU, F.-A. *et al.* Diffusion models in vision: A survey. *IEEE transactions on pattern analysis and machine intelligence*, Ieee, v. 45, n. 9, p. 10850–10869, 2023.
- CUI, R. *et al.* A modified residual stress dependent jile-atherton hysteresis model. *Journal of Magnetism and Magnetic Materials*, Elsevier, v. 465, p. 578–584, 2018.

- CUI, S.-g. *et al.* Variations in microstructure and mechanical properties along thickness direction in a heavy high strength low alloy steel plate. *Journal of Materials Research and Technology*, Elsevier, v. 26, p. 9190–9202, 2023.
- CUTLER, A.; CUTLER, D. R.; STEVENS, J. R. Random forests. In: *Ensemble machine learning*. [S.l.]: Springer, 2012. p. 157–175.
- DANESHVARI, M. H. *et al.* Efficient lbp-glm texture analysis for asphalt pavement raveling detection using extreme gradient boost. *Construction and Building Materials*, Elsevier, v. 401, p. 132731, 2023.
- DENG, Y.; ZOU, Y.; GUANG, F. Unveiling the energy efficiency paradox: Industrial automation and energy rebound in china. *Energy*, Elsevier, p. 136482, 2025.
- DEVI, M. B.; AMARENDRA, K. Machine learning-based application to detect pepper leaf diseases using histgradientboosting classifier with fused hog and lbp features. In: *Smart Technologies in Data Science and Communication: Proceedings of SMART-DSC 2021*. [S.l.]: Springer, 2021. p. 359–369.
- DHEEPAK, G.; VAISHALI, D. Optimizing brain tumor classification through innovative texture-based mri analysis: A fusion of novel glm and lbp with svm ensemble for enhanced accuracy. In: IEEE. *2024 Global Conference on Communications and Information Technologies (GCCIT)*. [S.l.], 2024. p. 1–6.
- DIMIAN, G. C. *et al.* The influence of digitalization on greenhouse gas emissions in european union. the analysis of mediating effect of renewable energy consumption. *Journal of the Knowledge Economy*, Springer, p. 1–41, 2025.
- DUARTE, L. M. *et al.* A novel approach based on pattern recognition techniques to evaluate magnetic properties of a non-grain oriented electrical steel in the secondary recrystallization process. *Measurement*, Elsevier, v. 167, p. 108135, 2021.
- DURMAZ, A. R. *et al.* A deep learning approach for complex microstructure inference. *Nature communications*, Nature Publishing Group UK London, v. 12, n. 1, p. 6272, 2021.
- FLETCHER, R. *Practical methods of optimization*. [S.l.]: John Wiley & Sons, 2000.
- FREITAS, F. N. C. *Influência da recristalização no comportamento da textura cristalográfica e nas perdas magnéticas por histerese de um aço elétrico de grão não orientado (GNO)*. Tese (Tese de Doutorado) — Universidade Federal do Ceará, Fortaleza, Brasil, 2011.
- GAO, H. *et al.* A high-level feature channel attention unet network for cholangiocarcinoma segmentation from microscopy hyperspectral images. *Machine Vision and Applications*, Springer, v. 34, n. 5, p. 72, 2023.
- GHASEMIAN, S. *et al.* An overview of global energy scenarios by 2040: identifying the driving forces using cross-impact analysis method. *International Journal of Environmental Science and Technology*, Springer, v. 21, n. 11, p. 7749–7772, 2024.
- GOLA, J. *et al.* Advanced microstructure classification by data mining methods. *Computational Materials Science*, Elsevier, v. 148, p. 324–335, 2018.

- GONCALVES, A. *et al.* Generation and evaluation of synthetic patient data. *BMC medical research methodology*, Springer, v. 20, n. 1, p. 108, 2020.
- GONZALEZ, R. C. *Digital image processing*. [S.l.]: Pearson education india, 2009.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. [S.l.]: MIT Press, 2016. <<http://www.deeplearningbook.org>>.
- GUAN, Y. *et al.* Burden of the global energy price crisis on households. *Nature Energy*, Nature Publishing Group UK London, v. 8, n. 3, p. 304–316, 2023.
- GUI, J. *et al.* A review on generative adversarial networks: Algorithms, theory, and applications. *IEEE transactions on knowledge and data engineering*, IEEE, v. 35, n. 4, p. 3313–3332, 2021.
- GUO, S. *et al.* Machine learning for metal additive manufacturing: Towards a physics-informed data-driven paradigm. *Journal of Manufacturing Systems*, Elsevier, v. 62, p. 145–163, 2022.
- HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. H. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, IEEE, n. 6, p. 610–621, 2007.
- HAYAKAWA, Y. Electrical steels. In: CABALLERO, F. G. (Ed.). *Encyclopedia of Materials: Metals and Alloys*. Oxford: Elsevier, 2022. p. 208–213. ISBN 978-0-12-819733-2. Disponível em: <<https://www.sciencedirect.com/science/article/pii/B9780128197264000041>>.
- HIEMATH, P. *et al.* Hybrid machine learning and regression framework for automated phase classification and quantification in sem images of commercial steels. *Discover Materials*, Springer, v. 5, n. 1, p. 156, 2025.
- HU, M.-K. Visual pattern recognition by moment invariants. *IRE transactions on information theory*, IEEE, v. 8, n. 2, p. 179–187, 1962.
- HUANG, G.-B.; ZHU, Q.-Y.; SIEW, C.-K. Extreme learning machine: a new learning scheme of feedforward neural networks. In: IEEE. *2004 IEEE international joint conference on neural networks (IEEE Cat. No. 04CH37541)*. [S.l.], 2004. v. 2, p. 985–990.
- IVO, R. F. *et al.* Non-grain oriented electrical steel photomicrograph classification using transfer learning. *Journal of Materials Research and Technology*, Elsevier, v. 9, n. 4, p. 8580–8591, 2020.
- IVO, R. F. *et al.* Study and classification of the crystallographic orientation distribution function of a non-grain oriented electrical steel using computer vision system. *Journal of Materials Research and Technology*, Elsevier, v. 8, n. 1, p. 1070–1083, 2019.
- JAIN, V. *et al.* Origin of goss texture in grain oriented electrical steel: Role of shear bands. *Materialia*, Elsevier, v. 22, p. 101398, 2022.
- KARAMI, R.; BUTLER, D.; TAMIMI, S. Manufacturing of non-grain-oriented electrical steels. *The International Journal of Advanced Manufacturing Technology*, Springer, v. 133, n. 3, p. 1083–1109, 2024.

- KILIC, I.; YAMAN, O. Classification method of plants with support vector machine (svm) using local binary gaussian model (lbgp) for smart agriculture. In: *2024 32nd Signal Processing and Communications Applications Conference (SIU)*. [S.l.: s.n.], 2024. p. 1–4.
- KINGMA, D. P.; BA, J. *Adam: A Method for Stochastic Optimization*. 2017. Disponível em: <<https://arxiv.org/abs/1412.6980>>.
- KUMAR, T. *et al.* Image data augmentation approaches: A comprehensive survey and future directions. *IEEE Access*, IEEE, 2024.
- LAI, C.-L. *et al.* Integrating local binary patterns and convolutional neural networks with a balanced gray encoded pooling mechanism. *Multimedia Tools and Applications*, Springer, p. 1–27, 2025.
- LI, Z.-H.; WANG, G.-D.; LIU, H.-T. Effects of processing routes on recrystallization texture development and magnetic properties of 0.10 mm ultrathin non-oriented electrical steel sheets for high-frequency application. *Journal of Alloys and Compounds*, Elsevier, v. 935, p. 167984, 2023.
- LI, Z.-H. *et al.* Dependence of recrystallization behavior and magnetic properties on grain size prior to cold rolling in high silicon non-oriented electrical steel. *Journal of Alloys and Compounds*, Elsevier, v. 888, p. 161576, 2021.
- LIU, D. *et al.* Synthetic data generation for augmenting small samples. *arXiv preprint arXiv:2501.18741*, 2025.
- LIU, Z. *et al.* Swin transformer: Hierarchical vision transformer using shifted windows. In: *Proceedings of the IEEE/CVF international conference on computer vision*. [S.l.: s.n.], 2021. p. 10012–10022.
- LIU, Z. *et al.* A convnet for the 2020s. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. [S.l.: s.n.], 2022. p. 11976–11986.
- MAATEN, L. v. d.; HINTON, G. Visualizing data using t-sne. *Journal of machine learning research*, v. 9, n. Nov, p. 2579–2605, 2008.
- MAO, Y. *et al.* An ai framework for time series microstructure prediction from processing parameters. *Scientific Reports*, Nature Publishing Group UK London, v. 15, n. 1, p. 24074, 2025.
- MOSES, A. *et al.* *Electrical Steels Volume 1: Fundamentals and Basic Concepts*. [S.l.]: The Institution of Engineering and Technology, 2019.
- MUMUNI, A.; MUMUNI, F. Data augmentation: A comprehensive survey of modern approaches. *Array*, Elsevier, v. 16, p. 100258, 2022.
- NEUNDLINGER, L. *et al.* Microstructural and textural evolution of double stage cold rolled non-grain oriented electrical steel. *Journal of Magnetism and Magnetic Materials*, Elsevier, v. 597, p. 172032, 2024.
- OJALA, T.; PIETIKÄINEN, M.; HARWOOD, D. A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, Elsevier, v. 29, n. 1, p. 51–59, 1996.

- OPOKU, R. *et al.* Optimization of industrial energy consumption for sustainability using time-series regression and gradient descent algorithm based on historical electricity consumption data. *Sustainability Analytics and Modeling*, Elsevier, v. 2, p. 100004, 2022.
- OQUAB, M. *et al.* Dinov2: Learning robust visual features without supervision. *arXiv preprint arXiv:2304.07193*, 2023.
- PALANISAMY, B. *et al.* Transformers for vision: A survey on innovative methods for computer vision. *IEEE Access*, v. 13, p. 95496–95523, 2025.
- PRINCIPE, J. C.; EULIANO, N. R.; LEFEBVRE, W. C. *Neural and adaptive systems: fundamentals through simulations with CD-ROM*. [S.l.]: John Wiley & Sons, Inc., 1999.
- RAJ, A.; SANKAR, D. Now-lbp glcm feature as the texture descriptor of the camera fingerprint in image forensics. In: IEEE. *2022 IEEE 19th India Council International Conference (INDICON)*. [S.l.], 2022. p. 1–6.
- Rebouças Filho, P. P. *et al.* New approach to evaluate a non-grain oriented electrical steel electromagnetic performance using photomicrographic analysis via digital image processing. *Journal of Materials Research and Technology*, Elsevier, v. 8, n. 1, p. 112–126, 2019.
- RUSANOVSKY, M.; BEERI, O.; OREN, G. An end-to-end computer vision methodology for quantitative metallography. *Scientific Reports*, Nature Publishing Group UK London, v. 12, n. 1, p. 4776, 2022.
- SCOTT, D. A. *Metallography in archaeology and art*. [S.l.]: Springer, 2019.
- SHAH, P. J.; SANKAR, D. Viral pneumonia detection from chest x-rays using glcm and lbp features. In: IEEE. *2023 Global Conference on Information Technologies and Communications (GCITC)*. [S.l.], 2023. p. 1–6.
- SHORTEN, C.; KHOSHGOFTAAR, T. M. A survey on image data augmentation for deep learning. *Journal of big data*, Springer, v. 6, n. 1, p. 1–48, 2019.
- SIMÕES, G. M. F.; LEDER, S. M. Energy poverty: The paradox between low income and increasing household energy consumption in brazil. *Energy and Buildings*, Elsevier, v. 268, p. 112234, 2022.
- SISTANINIA, M. *et al.* Designing deformation texture in non-oriented electrical steel for enhanced magnetic properties. *International Journal of Mechanical Sciences*, Elsevier, v. 290, p. 110090, 2025.
- SMOLA, A. Learning with kernels phd thesis. *Technische Universitat Berlin*, 1998.
- SOBUJ, M. S. I. *et al.* Leveraging pre-trained cnns for efficient feature extraction in rice leaf disease classification. In: IEEE. *2024 International Conference on Advances in Computing, Communication, Electrical, and Smart Systems (iCACCESS)*. [S.l.], 2024. p. 01–06.
- SUN, L. *et al.* Nanostructural metallic materials: Structures and mechanical properties. *Materials Today*, Elsevier, v. 38, p. 114–135, 2020.

- SUYKENS, J. A.; VANDEWALLE, J. Least squares support vector machine classifiers. *Neural processing letters*, Springer, v. 9, n. 3, p. 293–300, 1999.
- TAMIMI, S. *et al.* Mechanical properties and crystallographic texture of non-oriented electrical steel processed by repetitive bending under tension. *Materials Science and Engineering: A*, Elsevier, v. 835, p. 142665, 2022.
- TEKE, M.; ETEM, T. Cascading glcm and t-sne for detecting tumor on kidney ct images with lightweight machine learning design. *The European Physical Journal Special Topics*, Springer, p. 1–16, 2025.
- VAPNIK, V. *The nature of statistical learning theory*. [S.l.]: Springer science & business media, 2013.
- WANG, W. *et al.* Sindiffusion: Learning a diffusion model from a single natural image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE, 2025.
- WIDROW, B.; HOFF, M. E. *Adaptive Switching Circuits*. [S.l.], 1960.
- WINTER, K. *et al.* Effect of sub-micron deformations at opposing strain rates on the micromagnetic behaviour of non-oriented electrical steel. *Nature Communications*, Nature Publishing Group UK London, v. 15, n. 1, p. 9010, 2024.
- YANG, X.; ZHOU, H.; GAO, J. Enhancing renewable energy productivity and energy efficiency of energy projects: How does cost of capital influence? *Energy Strategy Reviews*, Elsevier, v. 57, p. 101608, 2025.
- YANG, Z. *et al.* A survey of automated data augmentation algorithms for deep learning-based image classification tasks. *Knowledge and Information Systems*, Springer, v. 65, n. 7, p. 2805–2861, 2023.
- YAZDANI, S. *et al.* Generative ai in depth: A survey of recent advances, model variants, and real-world applications. *Journal of Big Data*, Springer, v. 12, n. 1, p. 1–43, 2025.
- ZHAO, S.; SONG, J.; ERMON, S. Towards deeper understanding of variational autoencoding models. *arXiv preprint arXiv:1702.08658*, 2017.
- ZHAO, X. *et al.* A review of convolutional neural networks in computer vision. *Artificial Intelligence Review*, Springer, v. 57, n. 4, p. 99, 2024.