

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA E
INFORMÁTICA INDUSTRIAL**

BRUNA MACHADO MULINARI

**APLICAÇÃO DAS TRAJETÓRIAS V-I PARA MÉTODOS DE
EXTRAÇÃO DE CARACTERÍSTICAS E CLASSIFICAÇÃO DE
CARGAS ELÉTRICAS**

DISSERTAÇÃO

CURITIBA

2020

BRUNA MACHADO MULINARI

**APLICAÇÃO DAS TRAJETÓRIAS V-I PARA MÉTODOS DE EXTRAÇÃO
DE CARACTERÍSTICAS E CLASSIFICAÇÃO DE CARGAS ELÉTRICAS**

**Application of V-I trajectory in Feature Extraction Methods and Electric Load
Classification**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial, da Universidade Tecnológica Federal do Paraná (UTFPR), como requisito parcial para obtenção do título de “Mestre em Ciências” - Área de Concentração: Engenharia de Computação.
Orientador: Prof. Dr. André Eugenio Lazzaretti
Coorientador: Prof. Dr. Elder Oroski

Curitiba

2020



[4.0 Internacional](https://creativecommons.org/licenses/by/4.0/)

Esta licença permite que outros distribuam, remixem, adaptem e criem a partir do seu trabalho, mesmo para fins comerciais, desde que lhe atribuam o devido crédito pela criação original.



**Ministério da Educação
Universidade Tecnológica Federal do Paraná
Câmpus Curitiba**



BRUNA MACHADO MULINARI

APLICAÇÃO DAS TRAJETÓRIAS V-I PARA MÉTODOS DE EXTRAÇÃO DE CARACTERÍSTICAS E CLASSIFICAÇÃO DE CARGAS ELÉTRICAS

Trabalho de pesquisa de mestrado apresentado como requisito para obtenção do título de Mestre Em Ciências da Universidade Tecnológica Federal do Paraná (UTFPR). Área de concentração: Engenharia De Computação.

Data de aprovação: 10 de Dezembro de 2020

Prof Andre Eugenio Lazzaretti, Doutorado - Universidade Tecnológica Federal do Paraná

Prof Daniel Prado De Campos, - Universidade Tecnológica Federal do Paraná

Prof Elder Oroski, Doutorado - Universidade Tecnológica Federal do Paraná

Prof Mario Oleskovicz, Doutorado - Usp-Universidade de São Paulo

Prof Robson Ribeiro Linhares, Doutorado - Universidade Tecnológica Federal do Paraná

Documento gerado pelo Sistema Acadêmico da UTFPR a partir dos dados da Ata de Defesa em 10/12/2020.

AGRADECIMENTOS

Aos professores do laboratório que tive a oportunidade de trabalhar e conviver, por todo conhecimento compartilhado e amizade, em especial ao meu orientador André Eugenio Lazzaretti, coorientador Elder Oroski e demais professores que participaram do grupo de pesquisa de Assinatura de Potência.

Aos amigos e colegas Clayton Hilgemberg da Costa, Hellen Cristina Ancelmo, Daniel Prado de Campos e Lucas da Silva Nolasco que contribuíram para o desenvolvimento desse trabalho. Além dos amigos, dentro e fora do laboratório, que compartilharam momentos de felicidade e me auxiliaram em momentos de insegurança.

A todos os meus familiares, principalmente minha mãe Adiaigre Mulinari, a minha irmã Érica Fernanda e meus avós Neli e José, por todo carinho, paciência e torcida.

À Agência Nacional de Energia Elétrica (ANEEL) e à Companhia Paranaense de Energia Elétrica (COPEL) pelo suporte. Este trabalho teve financiamento via projeto de Pesquisa & Desenvolvimento dentro do programa da ANEEL, código PD2866-0464/2017.

A mente que se abre a uma nova ideia jamais
voltará ao seu tamanho original. (Albert
Einstein)

RESUMO

MULINARI, Bruna Machado. **Aplicação das trajetórias V-I para métodos de extração de características e classificação de cargas elétricas**. 2020. 126 f. Dissertação (Dissertação em Engenharia Elétrica e Informática Industrial) – Universidade Tecnológica Federal do Paraná. Curitiba, 2020.

Devido à crescente demanda de energia elétrica, programas de eficiência energética começaram a aderir a soluções de gerenciamento e conservação de energia. Para colaborar na conscientização e gerenciamento do uso da energia de consumidores, técnicas de monitoramento não-intrusivo de carga permitem que o consumo individualizado dos aparelhos de uma instalação sejam informados para o usuário. Para implementar tal abordagem existem quatro etapas fundamentais: aquisição de dados, detecção de eventos, extração de características e classificação de cargas. O desempenho da última etapa depende de como são realizadas as etapas anteriores, principalmente a etapa de extração de características que envolve a descrição quantitativa do comportamento, também denominado como assinatura de carga, de um aparelho elétrico. Nesse contexto, este trabalho se concentra no desenvolvimento de duas abordagens de extração de características baseadas nas assinaturas de cargas denominadas de trajetórias V-I. Pela abordagem tradicional de extração, novas características são propostas para utilização em estado estacionário e transitório dessa assinatura de carga. A segunda abordagem consiste no desenvolvimento de um método que descreve a própria assinatura de carga por meio dos descritores da Séries de Fourier bidimensional (2D). Com essas abordagens propostas, avaliaram-se os resultados na classificação de cargas com diferentes classificadores (k-vizinhos mais próximos, *Ensemble*, árvore de decisão, máquina de vetores suporte e análise de discriminante linear) para três bases de dados distintas e disponíveis publicamente. Além da avaliação dos recursos computacionais para os métodos propostos, como tempo e memória em um sistema embarcado, a robustez das metodologias à inserção de ruídos também foi avaliada. Os resultados obtidos para ambas as metodologias, na maioria dos casos, apresentam acurácias na etapa de classificação maiores que 90%. Adicionalmente, os resultados em termos de recursos computacionais se mostram competitivos quando comparados a outra metodologia existente na literatura, baseada na Transformada *Wavelet*.

Palavras-chave: Classificação de Cargas Elétricas. Fourier 2D. NILM. Reconhecimento de Padrões. Trajetórias V-I.

ABSTRACT

MULINARI, Bruna Machado. **Application of V-I trajectory in Feature Extraction Methods and Electric Load Classification**. 2020. 126 p. Dissertation (Master's Degree in Course Name) – Universidade Tecnológica Federal do Paraná. Curitiba, 2020.

Due to the growing demand for electricity, energy efficiency programs have started to include energy management and conservation solutions. In order to collaborate in the awareness and management of consumers' energy use, non-intrusive load monitoring techniques allow the individual consumption of devices in an installation to be informed to the user. To implement this approach, there are four main steps: data collection, event detection, feature extraction, and load identification. The performance of the last step depends on how the previous steps are performed, especially the feature extraction, which involves the quantitative description of the behavior of an electrical appliance, also defined as load signature. In this context, this work focuses on the development of two approaches based on load signatures called V-I trajectories. In the traditional extraction approach, new characteristics are proposed, for use in steady-state and transitory state of the load signature. The second approach consists in developing a method that describes the load signature itself, through the two-dimensional Fourier series (2D) descriptors. With those proposed approaches, the results in load identification are evaluated with different classifiers (k-nearest neighbors, Ensemble, decision tree, support vector machine, and linear discriminant analysis) for three distinct and publicly available databases. Besides the evaluation of computational resources of the proposed methods, such as time and memory in an embedded system, the robustness of the methods to noise insertion is evaluated. The results obtained for both methods, in most cases, present classification accuracy higher than 90%. Also, the use of computational resources provides competitive results when compared to another method in the literature (state-of-the-art), based on the Wavelet Transform.

Keywords: Classification of Electric Loads. Fourier 2D. NILM. Pattern Recognition. V-I Trajectory.

LISTA DE ILUSTRAÇÕES

Figura 1 – Exemplo de um Sistema Não-Intrusivo de Monitoramento de Carga.	21
Figura 2 – Aquisições da base de dados LIT-SYN de: (a) Tensão e (b) Corrente das cargas sendo conectadas e desconectadas na instalação. A ordem de ligação/desligamento das cargas: Micro-ondas, Carregador de celular e Secador.	22
Figura 3 – Etapas de um Sistema Não-Intrusivo de Monitoramento de Carga.	22
Figura 4 – Exemplo de mudança de estado de operação S1 para S2 de um aparelho.	24
Figura 5 – Categorias de aparelhos de acordo com os estados de operação.	25
Figura 6 – Categorias de modelos de detecção de eventos.	25
Figura 7 – Categorias de características.	26
Figura 8 – Trajetórias V-I de: (a) Furadeira em velocidade máxima (750 W) e (b) em velocidade média (750 W); (c) Lâmpada LED (1,6 W); (d) Lâmpada de halogêneo (105 W) da base de dados COOLL e (e) Exaustor de Fumaça (23 W); (f) Lâmpada LED (6 W) da base de dados LIT-SYN.	39
Figura 9 – Trajetórias V-I exibindo a característica <i>ar</i> : (a) do Exaustor de Fumaça (23 W) da base de dados LIT-SYN e (b) Pistola Elétrica de Pintura (1050 W) da base de dados COOLL.	41
Figura 10 – Diferença de Tensão e Corrente para <i>ar</i> : (a) Exaustor de Fumaça (23 W) da base de dados LIT-SYN e (b) Decapante Elétrico (1050 W) da base de dados COOLL.	42
Figura 11 – Trajetórias V-I destacando <i>lpa</i> para: (a) a Furadeira Elétrica em velocidade máxima (701 W) da base de dados COOLL e (b) Tensão e corrente da respectiva carga.	43
Figura 12 – Trajetórias V-I destacando <i>asy</i> : (a) Furadeira em velocidade média (750 W) e (b) Lâmpada halógena (105 W), ambas da base de dados COOLL.	43
Figura 13 – Tensão e Corrente: (a) Furadeira em velocidade média (750 W) e (b) Lâmpada halógena (105 W), ambas da base de dados COOLL, destacando <i>asy</i> devido a condução de corrente.	44
Figura 14 – Trajetórias V-I destacando <i>r</i> : (a) Ventilador (80 W) da base de dados LIT-SYN e Furadeira em velocidade máxima (750 W) da base de dados COOLL.	45
Figura 15 – Diferença entre as correntes: (a) Ventilador (80 W) da base de dados LIT-SYN e Furadeira em velocidade máxima (750 W) da base de dados COOLL, para a característica <i>r</i>	45
Figura 16 – Exemplo de alguns vetores utilizados na Equação 8 para o cálculo de <i>sc</i> da Furadeira em velocidade média (750 W) da base de dados COOLL.	46
Figura 17 – Trajetórias V-I destacando <i>sc</i> : (a) Furadeira em velocidade média (750 W) e (b) Lâmpada halógena (100 W), ambas da base de dados COOLL.	46
Figura 18 – Corrente: (a) Furadeira em velocidade média (750 W) e (b) Lâmpada halógena (100 W), ambas da base de dados COOLL.	47
Figura 19 – Trajetória V-I destacando os pontos de inflexão na parte A e B utilizando o Aspirador Tornado (700 W) da base de dados COOLL.	48
Figura 20 – Trajetória V-I destacando <i>mi</i> do Aspirador de pó Dilem 2406 da base de dados COOLL.	48
Figura 21 – Trajetórias V-I com <i>sh</i> distintos: (a) Monitor LED (26 W) da base de dados LIT-SYN e (b) Esmerilhadeira (2000 W) da base de dados COOLL.	49

Figura 22 – Tensão e Corrente: (a) Monitor LED (26 W) da base de dados LIT-SYN e (b) Esmerilhadeira (2000 W) da base de dados COOLL, para diferenciar a característica <i>sh</i>	50
Figura 23 – Trajetórias V-I destacando <i>alr</i> : (a) Painel de LED (13 W) da base de dados LIT-SYN e Lâmpada Fluorescente Compacta (15W) da base de dados COOLL.	50
Figura 24 – Tensão e Corrente: (a) Painel de LED (13 W) da base de dados LIT-SYN e Lâmpada Fluorescente Compacta (15W) da base de dados COOLL, com diferentes valores de <i>alr</i>	51
Figura 25 – Comparação entre a Série de Fourier tradicional e a forma Elíptica de Fourier. A representação harmônica exemplifica os métodos utilizando 2 harmônicos.	53
Figura 26 – Representação inversa da Fourier Elíptica utilizando (a) dois, (b) dois e (c) três harmônicos. Os harmônicos são representadas pelas elipses em cinza, os fasores pelas linhas pontilhadas e os contornos resultantes pela linha preta em destaque.	54
Figura 27 – Exemplo da predição do KVMP considerando $k=3$	55
Figura 28 – Exemplo do hiperplano ótimo de separação no modelo MVS.	57
Figura 29 – Exemplo de uma Árvore de decisão para distinguir quatro animais.	58
Figura 30 – Organização da metodologia.	62
Figura 31 – Bancada de aquisição de dados do subconjunto LIT-SYN.	67
Figura 32 – Exemplo de como é realizada a captura de um ciclo de rede da tensão e corrente.	68
Figura 33 – Exemplo do processo de subtração: (a) Conexão do Resistor (80 W) e posterior conexão da lâmpada LED (6 W); (b) Trajetória V-I do segundo evento, sem a subtração; (c) Trajetória V-I do segundo evento após a subtração e (d) Trajetória V-I original da carga lâmpada LED.	69
Figura 34 – Exemplo de como mensurar a característica <i>len</i> utilizando a carga: Exaustor de fumaça (23 W) da base de dados LIT-SYN	71
Figura 35 – Exemplo da característica <i>md</i> para o Ventilador (34 W) da base de dados COOLL.	72
Figura 36 – Exemplo da característica <i>ov</i> para o Aspirador de Pó Dyson (1400 W) da base de dados COOLL.	73
Figura 37 – Exemplo da característica <i>asl</i> para a Lixadeira Bosch (150 W) da base de dados COOLL.	74
Figura 38 – Exemplo das características <i>dpb</i> e <i>angp</i> do Motor de 457,1 W da base de dados LIT-SIM.	75
Figura 39 – Exemplo das características <i>dc</i> e <i>angc</i> da Lâmpada de LED (1,6 W) da base de dados COOLL.	78
Figura 40 – Reconstrução das trajetórias V-I utilizando a Fourier 2D, com variação das harmônicas de ordem h , para: (a) Furadeira em velocidade máxima (750 W), (b) Aspirador Tornado (700 W) e (c) Aspirador Dilem da base de dados COOLL.	79
Figura 41 – Visão geral do processo de classificação.	82
Figura 42 – Divisão do Conjunto Total de Características de um Método de Extração por meio dos Ids.	82
Figura 43 – SK-S7G2 Starter Kit.	83
Figura 44 – Montagem do sistema NILM com a metodologia multiagente.	84
Figura 45 – Valores medianos da característica <i>len</i> para cada classe.	87
Figura 46 – Valores medianos da característica <i>md</i> para cada classe.	88
Figura 47 – Valores medianos da característica <i>dpb</i> para cada classe.	89

Figura 48 – Valores medianos da característica <i>angp</i> para cada classe.	90
Figura 49 – Valores medianos da característica <i>vss</i> para cada classe.	91
Figura 50 – Valores medianos da característica <i>ov</i> para cada classe.	91
Figura 51 – Valores medianos da característica <i>dbang</i> para cada classe.	92
Figura 52 – Valores medianos da característica <i>dc</i> para cada classe.	92
Figura 53 – Valores medianos da característica <i>asl</i> para cada classe.	93
Figura 54 – Valores medianos da característica <i>dbd</i> para cada classe.	94
Figura 55 – Valores medianos da característica <i>dbc</i> para cada classe.	94
Figura 56 – Valores medianos da característica <i>angc</i> para cada classe.	95
Figura 57 – Correlação entre as características para a base de dados COOLL.	96
Figura 58 – Correlação entre as características para a base de dados LIT-SIM.	97
Figura 59 – Correlação entre as características para a base de dados LIT-SYN.	97
Figura 60 – Correlação entre as características da F2D para a base de dados COOLL. . .	106
Figura 61 – Correlação entre as características da F2D para a base de dados LIT-SYM 5. .	106
Figura 62 – Correlação entre as características da F2D para a base de dados LIT-SYN. .	107
Figura 63 – Avaliação do desempenho de classificação com a inserção de ruído nas amostras de entrada para os métodos analisados.	109

LISTA DE TABELAS

Tabela 1 – Palavras-chave e número de publicações encontradas.	31
Tabela 2 – Características do banco de dados final.	32
Tabela 3 – Tipos de aparelhos da base de dados COOLL.	64
Tabela 4 – Tipos de aparelhos no subconjunto LIT simulado - LIT-SIM	65
Tabela 5 – Tipos de aparelhos no subconjunto LIT sintético - LIT-SYN	66
Tabela 6 – Acurácias para o método tradicional das trajetórias V-I com a trajetória normalizada.	99
Tabela 7 – Acurácias para o método tradicional das trajetórias V-I com a trajetória sem normalização.	99
Tabela 8 – Avaliação de acurácias e tempo de extração dos conjuntos.	100
Tabela 9 – Resultados obtidos da seleção de características com a eliminação reversa.	101
Tabela 10 – Resultados obtidos da seleção de características com a seleção direta.	102
Tabela 11 – Avaliação da robustez das características ao utilizar diferentes níveis de ruído.	103
Tabela 12 – Acurácias com a variação do número de harmônicos analisadas.	104
Tabela 13 – Acurácias para o método da Fourier 2D com a trajetória V-I normalizada.	104
Tabela 14 – Acurácias para o método da Fourier 2D com a trajetória V-I sem normalização.	105
Tabela 15 – Avaliação da robustez da metodologia F2D ao utilizar diferentes níveis de ruído.	105
Tabela 16 – Comparação das metodologias de extração de características.	108
Tabela 17 – Avaliação da robustez da TW ao utilizar diferentes níveis de ruído.	109
Tabela 18 – Avaliação do desempenho de classificação ao realizar o treinamento em um contexto e o teste em outro.	110

LISTA DE ABREVIATURAS, SIGLAS E ACRÔNIMOS

SIGLAS

ACP	Análise de Componentes Principais
ADL	Análise de Discriminante Linear
AEF	Análise Elíptica de Fourier
alr	Área dos segmentos direito e esquerdo
angc	Ângulo do centroide
angp	Ângulo entre o ponto máximo e mínimo
ar	Área
asl	Valor médio de inclinação
asy	Assimetria
D	Variação da admitância instantânea
dbang	Diferença entre ângulos
dbc	Diferença entre centroides
dbd	Diferença entre distâncias
dc	Distância do centroide
dpb	Distância entre o ponto máximo e o ponto mínimo
ER	Eliminação Reversa
F2D	Fourier bidimensional
FPGA	do inglês <i>Field Programmable Gate Array</i>
I	Corrente
itc	Corrente atual
KVMP	K-Vizinhos Mais Próximos
len	Comprimento
lpa	Direção do <i>loop</i>
md	Distância Máxima
ME	Maximização da Expectativa
mi	Pico do segmento médio
MVS	Máquina de Vetor Suporte
NT	Número total de amostras
ov	Sobressinal
PM	Perceptron em Multicamadas
r	Curvatura da linha média
RF	do inglês <i>Random Forest</i>
RNA	Redes Neurais Artificiais

RNC	Rede Neural Convolucional
sc	Auto Intersecção
SD	Seleção Direta
sh	Inclinação do segmento médio
SNR	Relação Sinal-Ruído do inglês <i>Signal-to-Noise Ratio</i>
tv	Limiar utilizado para minimizar os efeitos de ruídos
TW	Transformada <i>Wavelet</i>
vss	Variações da inclinação do sinal

ACRÔNIMOS

ALM	Monitoramento de Carga do inglês <i>Appliance Load Monitoring</i>
COOLL	do inglês <i>Controlled On/Off Loads Library</i>
ILM	Monitoramento Intrusivo de Carga do inglês <i>Intrusive Load Monitoring</i>
LIT	Laboratório de Inovações
MOM	Método Oculto de Markov
NILM	Monitoramento Não-Intrusivo de Carga do inglês <i>Non-Intrusive Load Monitoring</i>
PLAID	do inglês <i>Plug Load Appliance Identification Dataset</i>
REDD	do inglês <i>Reference Energy Disaggregation Dataset</i>
SIM	Simulada
SYN	Sintética
WHITED	do inglês <i>Worldwide Household and Industry Transient Energy Data Set</i>

SUMÁRIO

1	INTRODUÇÃO	15
1.1	JUSTIFICATIVA	17
1.2	OBJETIVOS	18
1.2.1	Objetivo Geral	18
1.2.2	Objetivos Específicos	19
1.3	ESTRUTURA DO DOCUMENTO	19
2	ESTADO DA ARTE	21
2.1	NILM	21
2.2	MÉTODOS DE EXTRAÇÃO	27
2.3	MÉTODOS DE CLASSIFICAÇÃO	29
2.4	TRAJETÓRIA V-I	30
2.5	CONTRIBUIÇÕES DESTE TRABALHO	37
3	FUNDAMENTAÇÃO TEÓRICA	38
3.1	MÉTODO TRADICIONAL DA TRAJETÓRIA V-I	38
3.1.1	Corrente atual	40
3.1.2	Variação da admitância instantânea	40
3.1.3	Área	40
3.1.4	Direção do <i>loop</i>	41
3.1.5	Assimetria	42
3.1.6	Curvatura da linha média	44
3.1.7	Auto Intersecção	44
3.1.8	Pico do segmento médio	47
3.1.9	Inclinação do segmento médio	49
3.1.10	Área dos segmentos direito e esquerdo	49
3.2	MÉTODO DA FOURIER 2D	51
3.3	CLASSIFICADORES	55
3.3.1	K-Vizinhos Mais Próximos	55
3.3.2	Máquina de Vetor Suporte	56
3.3.3	Análise de Discriminantes Lineares	57
3.3.4	Árvore de Decisão	58
3.3.5	<i>Ensemble</i>	59
3.4	SELEÇÃO DAS CARACTERÍSTICAS	59
4	METODOLOGIA	62
4.1	BASE DE DADOS	63
4.1.1	COOLL	63
4.1.2	Base de dados LIT	63
4.2	PRÉ-PROCESSAMENTO	67
4.3	MÉTODO TRADICIONAL DA TRAJETÓRIA V-I	70
4.3.1	Comprimento	70
4.3.2	Máximo vetor	71
4.3.3	Sobressinal	71
4.3.4	Variação da inclinação do sinal	72

4.3.5	Valor médio da inclinação	73
4.3.6	Distância entre os pontos máximo e mínimo	74
4.3.7	Ângulo entre os pontos máximo e mínimo	74
4.3.8	Diferença entre distâncias	75
4.3.9	Diferença entre os ângulos	76
4.3.10	Distância do centroide até a origem	76
4.3.11	Ângulo do centroide	77
4.3.12	Distância entre centroides	77
4.4	FOURIER 2D	78
4.5	CLASSIFICAÇÃO	80
4.6	SISTEMA EMBARCADO	82
5	RESULTADOS E DISCUSSÃO	85
5.1	MÉTODO TRADICIONAL DAS TRAJETÓRIAS V-I	85
5.1.1	Análise das novas características	86
5.1.2	Classificação de cargas com a metodologia tradicional	98
5.2	FOURIER 2D	103
5.3	COMPARAÇÕES ENTRE METODOLOGIAS	107
6	CONCLUSÃO	112
6.1	TRABALHOS FUTUROS	113
6.2	TRABALHOS PUBLICADOS	114
	REFERÊNCIAS	117
	APÊNDICE A – RESUMO DAS CARACTERÍSTICAS EXTRAÍDAS DAS TRAJETÓRIAS V-I	125

1 INTRODUÇÃO

A energia elétrica tornou-se um fator crucial para o crescimento da sociedade, sendo base das produções industriais modernas, bem como fonte de comodidade para atividades diárias dos seres humanos (LIU, 2020). Em consequência, nas últimas décadas, o registro do consumo de energia tem aumentado em um ritmo acelerado o qual tem despertado alertas para estudiosos e governantes, pois pode levar a crises energéticas, mudanças climáticas (com a queima de combustíveis fósseis para geração de energia), e afetar a economia. Nesse contexto, para a utilização mais consciente e eficiente dos recursos energéticos, a conservação e o gerenciamento de energia têm se tornado partes integrantes de programas de desenvolvimento sustentável de diversos países (GOPINATH *et al.*, 2020).

Além do incentivo à utilização de tecnologias mais econômicas e eficientes, alguns programas investem em despertar a consciência energética dos usuários para redução do consumo de energia. Neste sentido, a disponibilização de informações sobre os aparelhos que mais consomem energia em uma instalação pode fornecer aos usuários oportunidades para a conservação e gerenciamento de energia. Entretanto, as faturas de energia geralmente apresentam o consumo total da instalação e não o consumo individual de cada aparelho, dificultando a criação de hábitos mais conscientes e econômicos da energia (WÓJCIK *et al.*, 2019).

Para fornecer informações individuais do consumo de cada aparelho é necessário que um sistema de monitoramento de carga - ALM (do inglês, *Appliance Load Monitoring*) seja implementado na instalação do consumidor. Essa implementação do sistema pode ser realizada de duas maneiras: de modo intrusivo e de modo não-intrusivo na instalação (LIU, 2020).

O monitoramento intrusivo de carga - ILM (do inglês, *Intrusive Load Monitoring*) consiste na instalação de sensores em cada uma das cargas da instalação, realizando o monitoramento das cargas em uma central. Embora esse tipo de sistema seja simples e apresente uma alta precisão de monitoramento, ele apresenta alguns problemas decorrentes da necessidade de um medidor para cada aparelho monitorado, entre eles: (i) o aumento do custo do sistema a cada nova carga a ser monitorada; (ii) cada medidor necessita estar equipado com uma interface de comunicação para envio de informações para a central; e (iii) a necessidade de uma fonte de alimentação para cada um dos medidores na instalação (WÓJCIK *et al.*, 2019).

Já o monitoramento não-intrusivo de carga - NILM (do inglês, *Non-Intrusive Load Monitoring*) se baseia na informação coletada por um único conjunto de sensores na entrada do

circuito elétrico, geralmente no medidor de energia já existente da instalação. Comparado ao sistema ILM, esse tipo de sistema é muito mais econômico pois não necessita de modificação ou expansão do *hardware* caso novas cargas sejam adquiridas na instalação. Entretanto, visto que o NILM utiliza um sinal agregado das medições de todas as cargas, nesse tipo de sistema é necessário envolver o processo de desagregação de cargas para estimar os consumos individuais de cada aparelho da instalação (GOPINATH *et al.*, 2020).

Para realizar o processo de desagregação de cargas pode-se fazer uso de técnicas de reconhecimento de padrões que utilizem as informações prévias a respeito do comportamento de um aparelho elétrico para reconhecê-lo em um sinal composto¹ desconhecido (FIGUEIREDO *et al.*, 2012). Esse comportamento único de um aparelho, capaz de distingui-lo dos demais, também pode ser denominado como uma assinatura de carga. As assinaturas utilizadas para descrever o comportamento de uma carga são obtidas a partir das coletas do sinal agregado e podem ser: formas de tensão e corrente, medidas de potência (ativa, aparente e reativa), análise harmônica, envoltórias de corrente ou potência e trajetórias V-I (ZOHA *et al.*, 2012).

No processo de desagregação é necessário extrair características da assinatura da carga que possibilitem a maior distinção entre tipos diferentes de aparelhos, para uma etapa posterior de identificação da carga (classificação). As características podem ser obtidas com métodos que utilizem o comportamento da carga no estado estacionário e/ou estado transitório. Os métodos de extração que utilizam o estado estacionário da carga tendem a ser menos susceptíveis a ruídos (WANG *et al.*, 2018). Entretanto, o período de transição de estado de uma carga pode agregar informações discriminantes na classificação das cargas (ZOHA *et al.*, 2012). Por esse motivo, recentemente alguns autores têm focado na combinação de características das assinaturas no dois estados da cargas (ANCELMO *et al.*, 2019b; MULINARI *et al.*, 2019; LAZZARETTI *et al.*, 2020).

Uma das assinaturas que apresenta uma boa taxonomia² entre aparelhos distintos de uma instalação é a trajetória V-I. Essa assinatura cria formas bidimensionais a partir da plotagem de um ciclo do componente fundamental da tensão por um ciclo da corrente medidos do sistema. Geralmente, os métodos de extração de características que se baseiam nessa trajetória atuam sobre a assinatura obtida no estado estacionário da carga, como foi inicialmente proposto em Lam *et al.* (2007). Outros trabalhos que empregaram as trajetórias V-I com o objetivo de aumentar o desempenho da identificação de cargas, desenvolveram metodologias para complementar as

¹ Refere-se à aquisição agregada de duas ou mais cargas conectadas à rede.

² Distribuição sistemática das amostras em categorias dadas as características semelhantes.

informações extraídas sobre a forma da assinatura (HASSAN *et al.*, 2014; WANG *et al.*, 2018), ou utilizaram a trajetória em metodologias diferentes da proposta inicial, como em formato de imagem para alimentar Redes Neurais Artificiais - RNA (DU *et al.*, 2016; BAETS *et al.*, 2018b; LIU *et al.*, 2019b). Embora alguns desses trabalhos tenham apresentado resultados razoáveis para a etapa de identificação das cargas, superiores a 88%, grande parte dos autores se concentrou na obtenção de novas características apenas em estado estacionário ou em metodologias que exigem maior processamento computacional.

Dada a inexistência de trabalhos que realizam a extração de características sobre as trajetórias V-I, obtidas em estado transitório e estacionário, e com o intuito de aumentar a discriminabilidade entre determinadas cargas, que, conseqüentemente, eleva a acurácia de classificação, este trabalho parte para duas abordagens de desenvolvimento baseados na assinatura da trajetória V-I. A primeira visa descrever e aplicar as novas características, em estado estacionário e transitório, propostas inicialmente em Mulinari *et al.* (2019). Já a segunda abordagem utiliza como entrada para os modelos de classificação a própria descrição matemática da forma da assinatura, por meio da Série de Fourier 2D (bidimensional). Além disso, visto que poucos trabalhos tratam da robustez da metodologia de extração e a complexidade computacional envolvida, este trabalho avalia a utilização de recursos de memória e tempo de processamento, ao aplicar as duas metodologias propostas em um sistema embarcado, além de analisar os resultados obtidos com a inclusão de ruído aos sinais de entrada utilizados.

1.1 JUSTIFICATIVA

Além dos motivos específicos acima apresentados, este trabalho também apresenta uma motivação mais geral, conforme detalhado a seguir.

A demanda de energia mundial deve aumentar cerca de 2,1% ao ano segundo as previsões da Agência Internacional de Energia (IEA, 2019). Essa estimativa provoca preocupações a governantes e pesquisadores de diversas áreas e resulta no desenvolvimento de novas soluções com foco em eficiência energética para redução do consumo de energia. Alguns estudos apontam que disponibilizar *feedbacks* do consumo de energia para o usuário podem auxiliar na economia em 9% ao ano (ARMEL *et al.*, 2013).

O Brasil, com o Ministério de Minas e Energia, apresenta um Plano Nacional de

Eficiência Energética com premissas e diretrizes para formulações de políticas energéticas³ para órgãos Federais e demais setores (comércio, indústria, transporte, educação, entre outros) para aumentar a eficiência e conservação de energia. Por exemplo, por meio de medidas que estimulem a substituição de equipamentos ineficientes e planos para realização de ações integradas com escolas e comunidades para tratar temas como o combate ao desperdício. Adicionalmente, o Protocolo Internacional de Medição e Verificação desenvolvido em 1994 estabelece conceitos e possibilidades para a economia de água e energia. Nesse documento existem recomendações de práticas para quantificar os resultados de investimentos em eficiência energética, de modo a incentivar a implementação de políticas e projetos no contexto de conservação (Efficiency Valuation Organization, 2012).

Neste cenário, técnicas que realizem o monitoramento do consumo de cargas podem viabilizar a conscientização e economia de energia em diferentes setores ao disponibilizar essas informações para o consumidor. O monitoramento não-intrusivo pode oferecer uma solução de acompanhamento e gerenciamento, eficiente e econômica, para projetos de eficiência, definidos pelo Protocolo Internacional de Medição e Verificação (Efficiency Valuation Organization, 2012). Este trabalho, visa contribuir nesse cenário, por meio do desenvolvimento de métodos de extração de características a serem aplicados em sistemas de monitoramento não-intrusivo, com intuito de fornecer as informações de identificação de aparelhos com o melhor desempenho possível.

1.2 OBJETIVOS

Dada a necessidade crescente por métodos de acompanhamento e gerenciamento de consumo, bem como as limitações existentes nos trabalhos correlatos nessa área, os objetivos para este trabalho são definidos nas próximas subseções.

1.2.1 Objetivo Geral

Propor métodos de extração de características, baseados na assinatura de carga das trajetórias V-I, para realizar a classificação de cargas elétricas.

³ Entre as políticas energéticas destaca-se o Programa de Eficiência Energética da Agência Nacional de Energia Elétrica, disponível em: <https://www.aneel.gov.br/programa-eficiencia-energetica>.

1.2.2 Objetivos Específicos

Para atingir o objetivo geral, os seguintes objetivos específicos foram determinados:

- Avaliar as metodologias de extração de características existentes na literatura que empregam as trajetórias V-I;
- Desenvolver novas características, com a trajetória em estado transitório e estacionário, para complementar as informações extraídas na metodologia proposta por Wang *et al.* (2018);
- Avaliar a utilização da metodologia proposta por Wang *et al.* (2018), junto as características propostas, na classificação de cargas elétricas, com o uso de três bases de dados públicas;
- Analisar o impacto de cada nova característica proposta pela metodologia tradicional de extração com as trajetórias V-I e a correlação existente com as demais já propostas na literatura;
- Avaliar o desempenho de conjuntos menores de características, obtidos por meio da aplicação de técnicas de seleção de características;
- Desenvolver uma nova abordagem de extração de características que empregue a metodologia das séries de Fourier bidimensional nas trajetórias V-I, em estado estacionário e transitório;
- Avaliar o desempenho na classificação da metodologia que utiliza as séries de Fourier bidimensionais;
- Comparar as metodologias propostas nesta pesquisa com outra metodologia existente na literatura, frequentemente utilizada na classificação de cargas elétricas;
- Avaliar o desempenho de execução e de utilização de memória das metodologias propostas em um sistema embarcado.

1.3 ESTRUTURA DO DOCUMENTO

Este trabalho encontra-se dividido em seis capítulos. No segundo capítulo é descrito o funcionamento das etapas da aplicação, onde as metodologias propostas se encaixam e o estado

da arte ao qual o trabalho se baseia. Já o terceiro capítulo descreve os métodos e técnicas que serão empregados. Na sequência o quarto capítulo descreve as metodologias propostas e os processos utilizados. Por fim, o quinto capítulo apresenta os resultados obtidos das análises e aplicações das metodologias, seguido do sexto capítulo em que são feitas as conclusões sobre o trabalho como um todo.

2 ESTADO DA ARTE

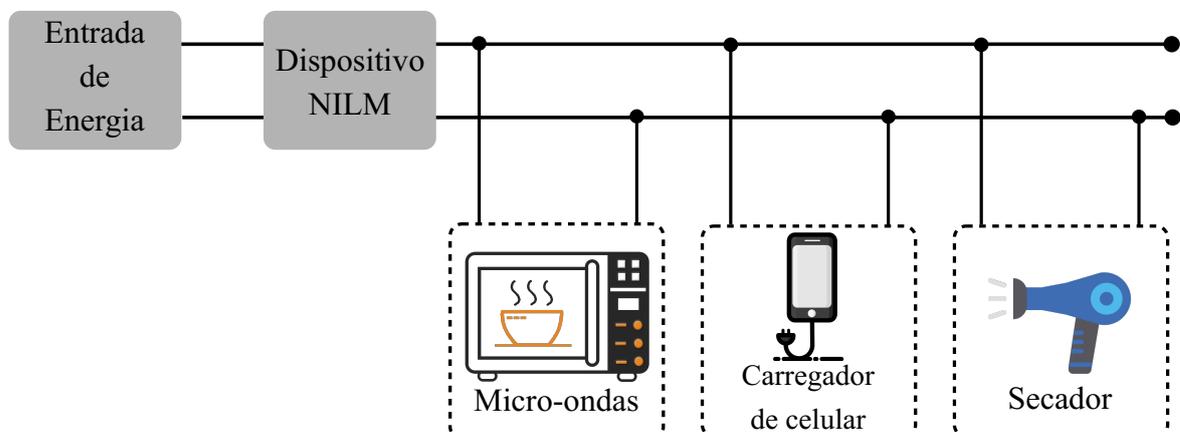
Neste capítulo é realizado o levantamento do estado da arte do método de extração de características utilizando as trajetórias V-I dentro do contexto de sistemas de monitoramento não-intrusivo de cargas. Primeiramente, são apresentadas as definições, além das etapas que podem compor esse tipo de sistema. Na sequência, são detalhadas as etapas do sistema na qual este trabalho irá se concentrar, bem como o principal método aplicado.

2.1 NILM

O primeiro a tratar de um sistema para desagregação dos consumos individuais dos aparelhos eletroeletrônicos utilizando um único ponto de medição foi Hart (1992). O sistema NILM proposto por Hart previa a detecção da condição ligado ou desligado de um aparelho e sua posterior identificação, considerando os dados de potência ativa e potência reativa obtidas por meio desse único ponto de medição.

Atualmente os sistemas NILM, ao utilizar um único ponto de medição para obter os consumos individuais, mantêm o mesmo princípio de funcionamento do sistema proposto por Hart. Porém, os sistemas atuais estão em constante desenvolvimento para possibilitarem, além da desagregação do consumo individualizado, o fornecimento do comportamento da carga, em um funcionamento normal, multi-estágio, danificado e com desgaste, entre outros (WÓJCIK *et al.*, 2019). Um exemplo da instalação de sistema NILM é apresentado na Figura 1.

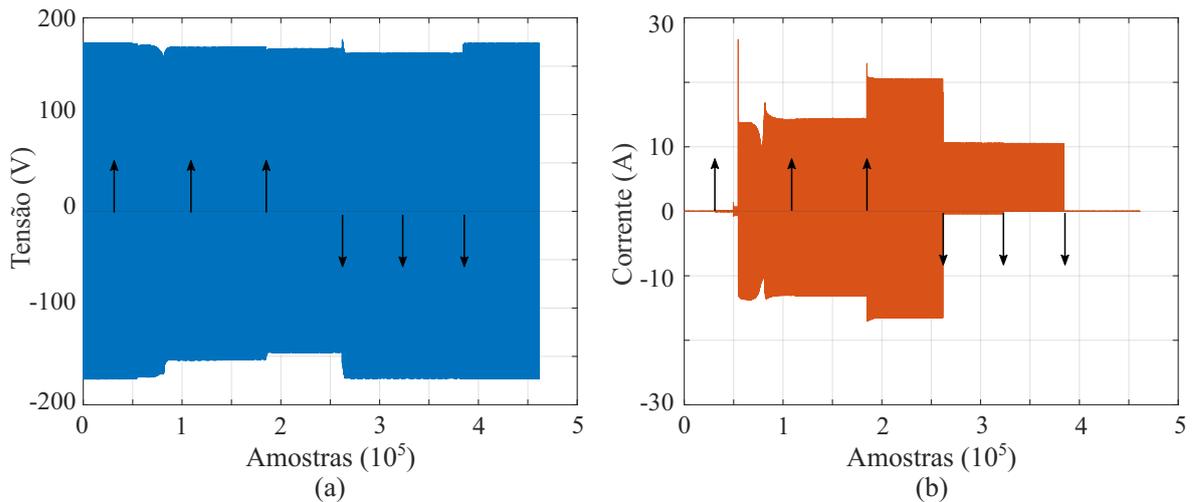
Figura 1 – Exemplo de um Sistema Não-Intrusivo de Monitoramento de Carga.



Fonte: Adaptado de Gopinath *et al.* (2020).

A Figura 2 apresenta as aquisições de tensão e corrente, com frequência de amostragem de aproximadamente 16 kHz, que são obtidas de um sistema NILM, conectado no painel principal da instalação, quando as cargas são ligadas e desligadas ao longo do tempo, indicadas pelas setas na figura.

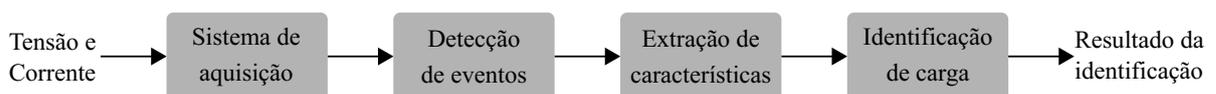
Figura 2 – Aquisições da base de dados LIT-SYN de: (a) Tensão e (b) Corrente das cargas sendo conectadas e desconectadas na instalação. A ordem de ligação/desligamento das cargas: Micro-ondas, Carregador de celular e Secador.



Fonte: Autoria própria.

O objetivo do sistema NILM é decompor dos sinais agregados, como os obtidos na Figura 2, os sinais individualizados das cargas para que seja possível distinguir quando cada aparelho encontra-se em funcionamento. A Figura 3 apresenta as etapas de um sistema NILM baseado na detecção de evento, que atua de modo diferente dos sistemas com métodos baseados em estados (GOPINATH *et al.*, 2020; LIU, 2020).

Figura 3 – Etapas de um Sistema Não-Intrusivo de Monitoramento de Carga.



Fonte: Adaptado de Ruano *et al.* (2019).

A etapa de aquisição de dados possui a função de coletar os dados de interesse, geralmente os sinais analógicos de tensão e corrente da entrada de energia, e transformá-los em sinais digitais com uma faixa de valores apropriada para o sistema digital utilizado para o processamento desses dados. Nessa etapa, um parâmetro importante é a frequência de amostragem desses sinais, pois interfere na escolha das técnicas utilizadas nas etapas futuras, bem como no

desempenho obtido pelo sistema. A frequência de amostragem de um medidor pode ser dividida em dois grupos: de baixa frequência e de alta frequência (ZOHA *et al.*, 2012).

Ainda não existe um consenso na área de NILM para estabelecer quais os valores de frequência definem um medidor de energia de baixa ou alta frequência. Atualmente a maioria dos medidores comerciais fornecem um frequência de amostragem de no máximo 1 Hz, sendo utilizado como um limite para definição de baixa frequência (RAFSANJANI *et al.*, 2018). Ainda, de acordo com Gao *et al.* (2015), as frequências de amostragem superiores a 4 kHz, consideradas no trabalho como alta frequência, devem ser utilizadas para implementar classificadores mais confiáveis de aparelhos elétricos do que os classificadores que utilizam aquisições com frequências menores, pois são capazes de descrever as aquisições de tensão e corrente, e conseqüentemente a assinatura de carga, com maiores detalhes.

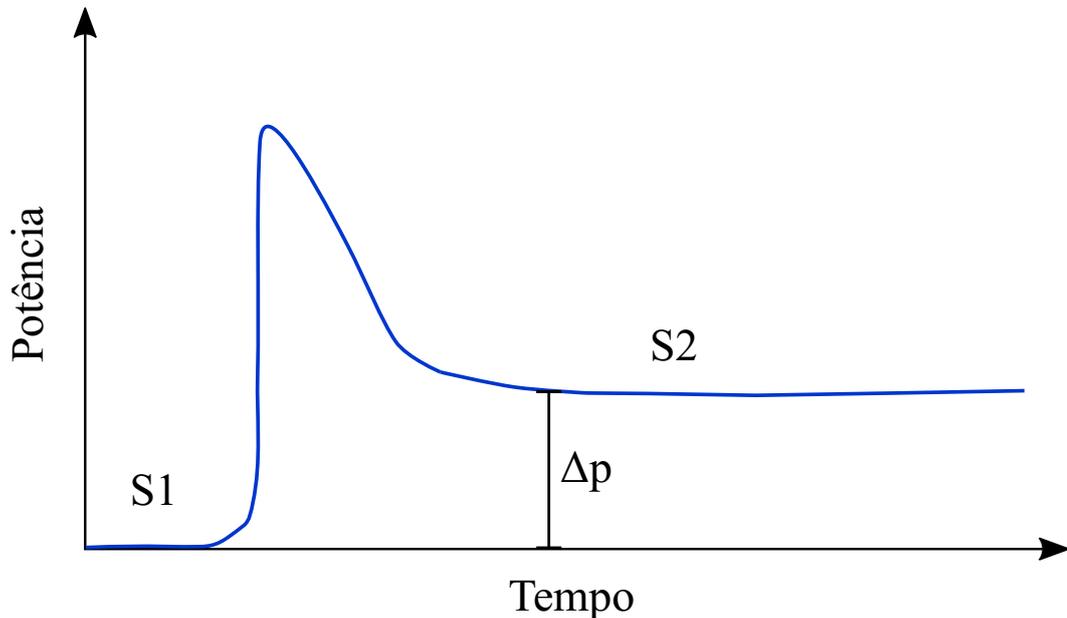
Os medidores de alta frequência apresentam um custo mais elevado quando comparado aos de baixa frequência por possuírem dispositivos com mais capacidade de processamento e componentes com mais alta resolução, como conversores Analógico-Digital (A/D) (RUANO *et al.*, 2019). Ainda assim, esse tipo de medidor permite a implementação de um sistema com uma análise mais profunda das características das aquisições, como análise de períodos transitórios e identificação de uma maior ordem de componentes harmônicos, resultando em uma desagregação de carga mais eficiente (GOPINATH *et al.*, 2020).

Em produtos mais recentes no mercado, como o medidor não-intrusivo de energia Neurio, a aquisição é realizada utilizando um alicate-amperímetro e os dados obtidos são enviados para um conjunto de utilitários, por um aplicativo de celular, para serem processados e exibidos (Neurio Energy Monitor, 2020).

A etapa seguinte à de aquisição dos dados é a de detecção de eventos. De modo simplificado, um evento ocorre quando um aparelho modifica seu estado de operação, sendo possível detectá-lo quando ocorre a diferença de dois valores característicos subsequentes, ultrapassando um limite pré-determinado de detecção (WÓJCIK *et al.*, 2019). Os valores característicos dependem da metodologia de detecção que se deseja aplicar, como os métodos de envelope do sinal de corrente e a variação da amplitude da potência ativa (RUANO *et al.*, 2019) (WÓJCIK *et al.*, 2019).

Um exemplo de ocorrência de um evento é dado na Figura 4, em que a carga passa do estado S1 para S2. Por meio da envoltória da potência é possível, considerando um limiar de variação de amplitude menor do que Δp , detectar um novo evento.

Figura 4 – Exemplo de mudança de estado de operação S1 para S2 de um aparelho.

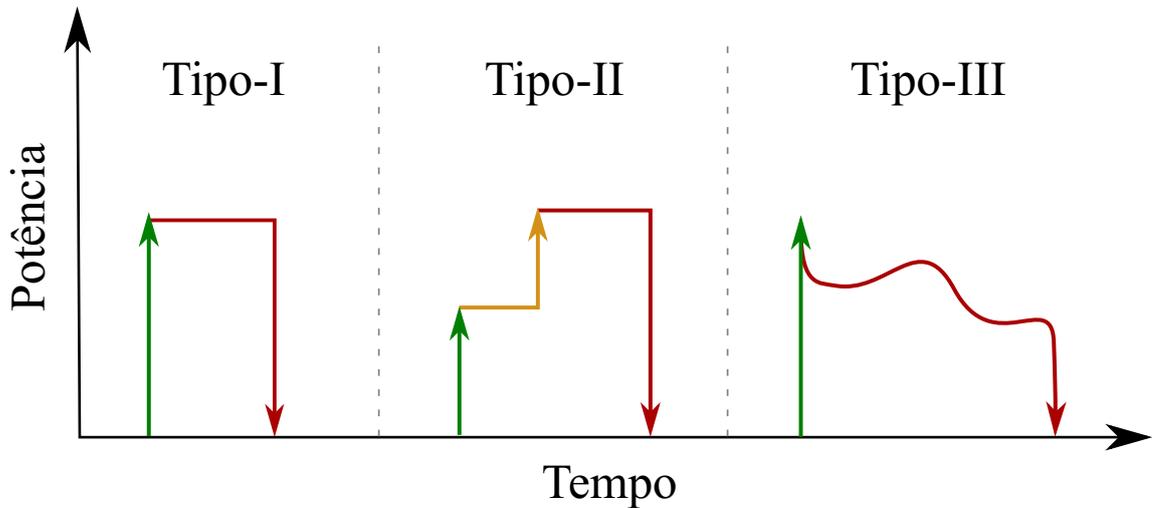


Fonte: Autoria própria.

A etapa de detecção pode se tornar complexa, visto que em uma instalação as cargas podem ser categorizadas em quatro tipos diferentes, dado o seu estilo de operação. As cargas do Tipo-I são aquelas que apresentam somente dois estados de operação (ligado-desligado), por exemplo, uma torradeira. As cargas do Tipo-II são as cargas que exibem um número finito de estados, também chamadas de multi-estados, como as máquinas de lavar roupa. Já nas cargas do Tipo-III se enquadram os aparelhos contínuos variáveis, pois não apresentam um número fixo de estados, como as lâmpadas com *dimmer*. Por fim, o Tipo-IV contém as cargas que permanecem ligadas por um longo período de tempo, consumindo constantemente energia, como detectores de fumaça (ZOHA *et al.*, 2012). A Figura 5 exibe uma representação para as três primeiras categorias visto que a última categoria opera continuamente.

De acordo com Liu (2020), os modelos de detecção de eventos podem ser divididos, como mostra a Figura 6, em quatro categorias: modelos heurísticos, modelos probabilísticos, modelo de correspondência e modelos híbridos. Nos modelos heurísticos se enquadram os modelos que definem um limiar, para considerar o evento ou não, que é configurado com base em um conhecimento prévio. Os métodos probabilísticos utilizam de informações estatísticas, como desvio padrão e variância, para detectar a mudança de operação de um sinal agregado e definir a ocorrência de um evento. Já os modelos de correspondência utilizam características obtidas do sinal agregado mais atual coletado e comparam com valores das mesmas características pré-

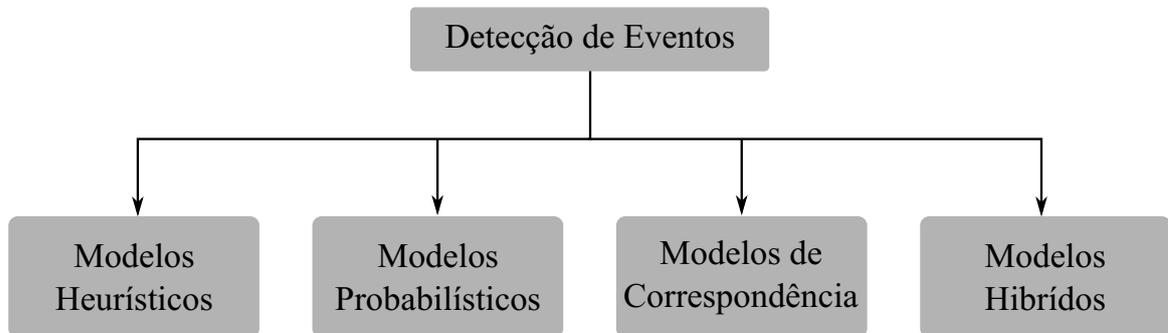
Figura 5 – Categorias de aparelhos de acordo com os estados de operação.



Fonte: Adaptado de Gopinath *et al.* (2020).

definidas utilizadas como modelo. A correspondência de características geralmente é realizada utilizando medidas de distância e métodos de similaridade (GOPINATH *et al.*, 2020). Por fim, os modelos híbridos agrupam os modelos que utilizam diferentes metodologias para composição dos seus resultados (LIU, 2020).

Figura 6 – Categorias de modelos de detecção de eventos.



Fonte: Adaptado de Liu (2020).

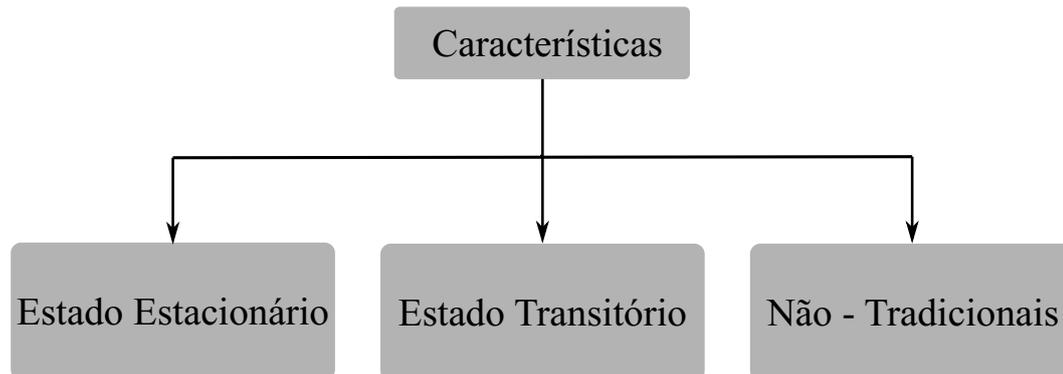
Ao detectar um novo evento, a próxima etapa de um sistema NILM é identificar qual aparelho alterou o seu estado de operação. Todo tipo de aparelho exibe um padrão único de consumo, com características associadas, denominado como assinatura de carga (HART, 1992). Esse comportamento único do dispositivo em operação possibilita a identificação e desagregação de uma dada carga de um sinal agregado (ARGATU *et al.*, 2018).

O desempenho na etapa de identificação dos sistemas NILM depende das particularidades encontradas em cada característica que descreve uma assinatura de carga. Na etapa de extração de características são coletados ou calculados esses parâmetros intrínsecos às assina-

ras, para que nas etapas seguintes seja possível que o sistema reconheça qual é o equipamento que modificou sua operação (FAUSTINE *et al.*, 2017).

As características e, por consequência os métodos utilizados para extrair esses recursos, podem ser categorizados como mostra a Figura 7, em metodologias de: estado estacionário, estado transitório e não-tradicionais (ABUBAKAR *et al.*, 2016). Os métodos de estado estacionário extraem características sobre a assinatura obtida quando as medidas de tensão e corrente da carga, após mudar seu estado de operação, se estabelecem no estado de regime permanente. Já os métodos de estado transitório extraem os recursos da assinatura obtida assim que ocorrem as transições de estado de operação. Por fim, os métodos não-tradicionais utilizam informações que não necessitam ser os dados elétricos, como horários de operação do aparelho, frequência de uso, sensores de luz e temperatura, para agregar as características obtidas nos métodos tradicionais (ZOHA *et al.*, 2012).

Figura 7 – Categorias de características.



Fonte: Adaptado de Abubakar *et al.* (2016).

Como discutido anteriormente, as características extraídas nessa etapa são altamente dependentes da taxa de amostragem do sistema, visto que para empregar metodologias que utilizam o estado transitório de uma carga é necessário uma frequência de amostragem mais alta (maior que 4 kHz) (GAO *et al.*, 2015).

Após a extração das características inerentes à assinatura de carga utilizada, a última etapa de um sistema NILM, baseado na detecção de um evento, consiste na identificação de qual foi o aparelho que realizou uma transição. Essa tarefa de identificação pode ser realizada por algoritmos que implementam modelos de aprendizagem de máquina para realizar a classificação de uma ou mais cargas (LIU, 2020).

A utilização dos algoritmos de aprendizagem podem exigir que o sistema passe por uma fase de treinamento, antes que possa ser colocado em operação e classifique os aparelhos

com base nesse aprendizado. A maneira como é realizada essa etapa de treinamento é o que categoriza os modelos de classificação em: supervisionados, não-supervisionados e híbridos (que envolve a junção dos dois modelos anteriores) (GOPINATH *et al.*, 2020).

Para o treinamento de um modelo supervisionado é necessário que um conjunto de dados rotulados seja utilizado para ensinar o modelo de classificação. Assim, em uma base de dados todas as características utilizadas para representar um determinado aparelho devem acompanhar um padrão (rótulo) que corresponda a essa categoria de carga (ZOHA *et al.*, 2012). Já na classificação não-supervisionada, os modelos dispensam a necessidade de um conjunto de dados rotulados e agrupam os exemplos de acordo com um número de classes e a quantidade de interações determinadas (RUANO *et al.*, 2019).

Geralmente, os modelos supervisionados apresentam um bom desempenho de classificação e conseguem identificar corretamente os tipos de dispositivos utilizados em seu treinamento. Entretanto, a implementação desses modelos é mais trabalhosa do que dos métodos não-supervisionados, pois necessita que o processo de atribuição de rótulo dos dados seja realizada manualmente (RUANO *et al.*, 2019). Além disso, as abordagens não-supervisionadas tem a capacidade de se adaptarem às mudanças frequentes no ambiente (GOPINATH *et al.*, 2020).

2.2 MÉTODOS DE EXTRAÇÃO

Como discutido na seção anterior, os métodos de extração de características podem ser separados em três categorias: transitório, estacionário e não-tradicional. Visto que a terceira categoria é utilizada como complementar às duas primeiras, somente elas serão tratadas nesta seção.

As características de uma assinatura de potência estão intimamente relacionadas à coleta dos dados utilizados e ao método de classificação do aparelho (RUANO *et al.*, 2019). Geralmente, os sistemas que apresentam frequências de amostragem pequenas (menor ou igual a 1 Hz) fazem uso das características extraídas de assinaturas em estado estacionário. Nessa categoria, de baixa frequência de amostragem e utilizando as assinaturas em estado estacionário, os recursos extraídos mais recorrentes na literatura são as séries temporais, como tensão e corrente, potência ativa (P), potência reativa (Q), potência aparente (S), fator de potência, entre outros (HERNÁNDEZ *et al.*, 2019).

As séries temporais de potência do aparelho podem ser utilizadas de várias maneiras. O recurso extraído mais empregado nesse tipo de metodologia é a potência aparente, sendo

exclusivamente utilizado para classificação de cargas em alguns trabalhos (LE *et al.*, 2015) (FIGUEIREDO *et al.*, 2014). Também, é possível encontrar trabalhos que utilizam mais recursos agregados para melhorar o desempenho de classificação, como os trabalhos que combinam potência ativa e reativa (CHANG *et al.*, 2014b) (KELLY; KNOTTENBELT, 2015) e trabalhos que combinam algum dos recursos de potência, com medidas de tensão ou corrente eficaz - RMS, do inglês *Root Mean Square* (MAKONIN, 2014) (WELIKALA *et al.*, 2019).

Ao utilizar os modelos de extração de características em estado estacionário com sistemas de aquisição em baixa frequência de amostragem, os recursos extraídos nesse regime de operação podem ser semelhantes em aparelhos que apresentem uma faixa de potência muito próxima (GOPINATH *et al.*, 2020). Como solução é possível elevar a frequência de amostragem para permitir que mais detalhes sobre a forma de onda de cada aparelho sejam obtidos, como harmônicos em ordem superiores, além da própria forma de onda da corrente e tensão brutas (RUANO *et al.*, 2019).

Com a frequência de amostragem mais elevada, um dos métodos que tem chamado atenção da comunidade científica é o da extração de característica utilizando como assinatura de carga as trajetórias V-I, por apresentar resultados mais eficazes de classificação quando comparado a abordagens com recursos de potência (ZOHA *et al.*, 2012).

As trajetórias V-I correspondem as formas bidimensionais obtidas com os sinais normalizados de um ciclo da tensão por um ciclo da corrente, ambos em estado estacionário (RUANO *et al.*, 2019). Lam *et al.* (2007) utilizaram essa trajetória para obter oito características derivadas da forma obtida, como área, assimetria, entre outras. Posteriormente em Wang *et al.* (2018), os autores utilizaram mais dois recursos intrínsecos a esse tipo de assinatura para obter melhores resultados na classificação dos aparelhos.

Ainda, ao considerar um sistema com frequência de amostragem elevada, é possível utilizar as informações das transições das medidas de potência, corrente e tensão das cargas, antes de se estabelecerem em estado estacionário (FAUSTINE *et al.*, 2017). As características obtidas do comportamento de transição das cargas podem distinguir os equipamentos com maior precisão comparadas as características de estado estacionário, visto que ocorre menos sobreposição de informações com cargas de potência parecidas (ZOHA *et al.*, 2012).

Entre os métodos de estado transitório, os que utilizam a transformada Wavelet tem apresentado resultados relevantes, uma vez que essa transformada decompõe um sinal em tempo e frequência, pois tornou-se muito efetiva na distinção de cargas não-lineares. Em Chang *et al.*

(2013) e Chang *et al.* (2014a) os autores calcularam os coeficientes da transformada Wavelet e reduziram seus tamanhos utilizando um espectro de potência e de energia, respectivamente, visto que utilizar diretamente todos os coeficientes obtidos como características iria requerer mais tempo e mais memória no processo.

Além dos métodos que se destacam empregando apenas um estado de funcionamento da carga, existem abordagens que associam métodos ou recursos que utilizam os dois estados de uma carga, transitório e estacionário, para aumentar a acurácia de classificação do sistema. Em Gao *et al.* (2015) são utilizados os recursos de diferentes métodos de extração, entre eles as trajetórias V-I, potência ativa e reativa, forma de onda da corrente em um período, entre outros, para melhorar o desempenho de classificação. Outro trabalho que realiza uma abordagem semelhante ao apresentado por Sadeghianpourhamami *et al.* (2017) que agrupa diferentes métodos de extração, entre eles as trajetórias V-I, Wavelet, potências e envoltórias de potência. Além disso, esse trabalho realiza uma seleção de características para encontrar um conjunto mais eficaz de características. Já em Mulinari *et al.* (2019) e Ancelmo *et al.* (2019b) os autores propõem a extração de características em ambos os estados, estacionário e transitório, para métodos anteriormente definidos para somente um dos estados, como as trajetórias V-I. Por fim, em Renaux *et al.* (2018), Lazzaretti *et al.* (2020) é proposto um sistema multiagente que associa os resultados de diferentes classificadores com diversos métodos de extração de características, entre eles as trajetórias V-I, Prony, níveis de energia da Transformada Wavelet Discreta, informações sobre o estado da vetorização e método de envelope de potência.

2.3 MÉTODOS DE CLASSIFICAÇÃO

Na última etapa do sistema NILM, na qual ocorre a identificação da carga que mudou seu estado de operação, pode-se empregar métodos de aprendizagem de máquina que utilizam a abordagem supervisionada ou a não-supervisionada. Como já tratado na Seção 2.1, o que diferencia uma abordagem da outra é a existência da etapa de treinamento ou a necessidade de dados rotulados.

Na abordagem supervisionada, é necessário utilizar um conjunto de dados, denominado de conjunto de treinamento, que conterá as características que representam uma parcela das amostras de cada classe na qual se deseja que o sistema consiga distinguir, além do conjunto de rótulos equivalentes a cada amostra do conjunto. Esse conjunto de treinamento irá direcionar o modelo para a distinção das classes. Outra parcela de amostras é utilizada para que, após

o período de treinamento de um método supervisionado, seja possível testá-lo com amostras diferentes das utilizadas na etapa de treinamento.

Entre os modelos supervisionados mais utilizados em NILM estão as RNA, com o *perceptron* em multicamadas - PM (CHANG, 2012) (CHANG *et al.*, 2014a), Redes Neurais Convolucionais - RNC (WU; WANG, 2019) (BAETS *et al.*, 2018b) e redes neurais profundas (KELLY; KNOTTENBELT, 2015) (KIM *et al.*, 2017), além dos modelos mais tradicionais de aprendizagem de máquina como a máquina de vetores suporte - MVS (ALTRABALSI *et al.*, 2014) (HASSAN *et al.*, 2014), k-vizinhos mais próximos - KVMP (FIGUEIREDO *et al.*, 2011) (BERGES *et al.*, 2011), classificadores Naive Bayes (MATHIS *et al.*, 2014), Árvores de decisão (LIAO *et al.*, 2014) (GILLIS *et al.*, 2016) e classificadores *Random Forest*- RF (LIU *et al.*, 2019a) (RUANO *et al.*, 2019).

Visto que criar um conjunto de dados rotulados das cargas para uma aplicação pode ser caro e trabalhoso, as abordagens não-supervisionadas têm chamado a atenção pois podem realizar a classificação das cargas de uma instalação sem a necessidade prévia de treinamento do modelo (FAUSTINE *et al.*, 2017). Entre os modelos mais utilizados na abordagem não-supervisionada, pode-se destacar os Modelos Ocultos de Markov - MOM (BONFIGLI; SQUARTINI, 2020) (LIU, 2020), o método por agrupamento *k-means* (YANG *et al.*, 2015) e maximização das expectativas - ME (FIGUEIREDO *et al.*, 2014).

As abordagens de classificação, supervisionada e não-supervisionada, contêm suas vantagens quando comparadas entre si. As classificações supervisionadas exibem melhores resultados na identificação de cargas, enquanto as não-supervisionadas dispensam a necessidade da criação/utilização de um banco de dados rotulados. Porém, em ambos os casos, as abordagens exigem a existência de características robustas, capazes de distinguir bem as cargas, para sistemas NILM eficazes (GOPINATH *et al.*, 2020).

2.4 TRAJETÓRIA V-I

Nesta seção foi realizado um levantamento mais aprofundado do estado da arte do método de extração de características empregando as trajetórias V-I no contexto de NILM, visto que este trabalho se concentrou nesse método para classificação de cargas elétricas em uma instalação. A metodologia utilizada para seleção das publicações que compõe esta parte do trabalho é baseada no processo estruturado descrito em Ensslin *et al.* (2013).

A primeira etapa do processo para seleção das publicações consistiu na definição dos

eixos de pesquisas, isto é, na escolha das palavras-chave que definem a aplicação e a metodologia a ser utilizada neste trabalho, para orientar as buscas e criação de um banco de dados inicial. As buscas foram realizadas no *software* ‘*Publish or Perish*’, utilizando a associação das palavras-chave definidas e delimitadas para um intervalo de tempo. O período escolhido para delimitar as buscas foi de 2000 a 2020, dado que a utilização da metodologia pesquisada é relativamente nova na área.

A Tabela 1 apresenta as palavras-chave utilizadas e o número de metadados das publicações que apresentavam as duas palavras no banco de dados do *Google Scholar*.

Tabela 1 – Palavras-chave e número de publicações encontradas.

Palavra-chave 1	Palavra-chave 2	Nº de metadados
	<i>V-I Trajectory</i>	229
<i>Non-Intrusive Load Monitoring</i>	<i>V-I Trajectories</i>	76
	<i>V-I Curves</i>	12
<i>NILM</i>	<i>V-I Trajectory</i>	198
	<i>V-I Trajectories</i>	72
	<i>V-I Curves</i>	10
<i>Appliance Load Monitoring</i>	<i>V-I Trajectory</i>	204
	<i>V-I Trajectories</i>	74
	<i>V-I Curves</i>	11

Fonte: Autoria própria.

No final da etapa de busca foram encontrados 886 metadados que apresentavam as restrições já mencionadas. Para que desse total de metadados fosse obtido um conjunto com as publicações mais relevantes, algumas filtragens foram realizadas, adaptadas de Ensslin *et al.* (2013).

A primeira filtragem consistiu em priorizar artigos de revistas e capítulos de livros, além da verificação de documentos duplicados no conjunto formado pelas combinações. Na sequência, a partir dos títulos dos documentos realizou-se uma seleção dos documentos que apresentavam maior correspondência a este trabalho. Por fim, realizou-se uma última filtragem por meio da leitura dos resumos, introduções e conclusões.

O conjunto final obtido para compor esta seção do trabalho é de 18 publicações. As métricas resultantes para esse conjunto de publicações é apresentado na Tabela 2.

Lam *et al.* (2007) foram os primeiros a realizar uma taxonomia de cargas de uma residência utilizando as características obtidas das assinaturas de cargas bidimensionais, geradas por meio da representação gráfica de um ciclo da tensão e da corrente da rede coletadas em estado estacionário, denominadas trajetórias V-I. Para realizar a categorização das cargas, os autores mensuraram as formas de onda da tensão e da corrente dos aparelhos, criaram as trajetórias V-I

Tabela 2 – Características do banco de dados final.

Característica	Resultado
Publicações	18
Citações	464
Anos	(2007-2019)
Citações/Publicação	25,78
<i>h-index</i>	9
<i>g-index</i>	18

Fonte: Autoria própria.

referentes a cada aparelho e extraíram características sobre a forma obtida, como: assimetria, área, curvatura da linha média, auto-intersecção, variação do segmento médio, área dos segmentos direito e esquerdo e o pico do segmento médio, definidas na seção 3.1. Ainda no trabalho, a característica direção do *loop* foi descrita mas não foi utilizada, pois segundo os autores estava integrada a característica da área. Embora esse trabalho realize uma comparação da taxonomia utilizando as trajetórias V-I com outros métodos, apresentando bons resultados, os autores não discutiram como realizaram as medições das características utilizadas.

O primeiro trabalho a apresentar e disponibilizar uma maneira de quantificar algumas das características propostas no trabalho de Lam *et al.* (2007) é o de Hassan *et al.* (2014). Nesse trabalho, os autores além de mensurar algumas das características sugeridas anteriormente, propõem uma nova característica que é determinada pela distância entre os pontos máximo e mínimo da corrente na trajetória. Para comparar o algoritmo criado com outros métodos de extração de características, os autores utilizaram um conjunto da base de dados pública denominada REDD (do inglês *Reference Energy Disaggregation Dataset*). Apesar de em alguns casos o algoritmo proposto apresentar um resultado melhor do que alguns métodos já existentes na literatura até aquele momento, os autores não consideraram aparelhos menores do que 50 W, pois a curva desses aparelhos quando em operação com aparelhos de potência maior apresentavam a trajetória similar a da carga de maior potência.

Um dos trabalhos que discute a similaridade da trajetória V-I de uma carga de alta potência com a de uma trajetória V-I obtida quando há uma carga de pequena potência operando junto a essa mesma carga de alta é o de Teshome *et al.* (2016). Nesse trabalho os autores sugerem a utilização da componente não-ativa da corrente, obtida por meio da decomposição da corrente empregando a Teoria de Fryze, para criação da trajetória V-I, pois essas novas assinaturas permitiriam uma distinção maior entre as cargas que apresentam potências similares quando comparadas as curvas tradicionais. Porém, os autores exibiram somente os resultados de similaridade das curvas e nenhum teste de classificação foi implementado ou utilizado para

comparação.

Já em Lin *et al.* (2016) são propostos algoritmos que empregam as características extraídas de várias metodologias, utilizando diferentes assinaturas de carga, para realizar a classificação dos aparelhos, como a forma de onda da corrente, as potências ativas e reativas, as harmônicas da corrente e duas características das trajetórias V-I (variação do segmento médio e área). Porém nesse trabalho, os autores realizaram a comparação das acurácias para seis tipos de aparelhos utilizando as características individualizadas. Entretanto, os resultados obtidos com os algoritmos que agrupam as metodologias apresentaram somente quatro dos aparelhos iniciais. Além disso, os resultados obtidos com múltiplos extratores¹ e com os métodos individuais são muito próximos para alguns casos, como o do método de potência.

Partindo para uma abordagem diferente da utilização das trajetórias V-I, em Du *et al.* (2016) é proposta a representação da curva V-I sobre uma matriz de células com valores binários², para extração de características específicas sobre essa grade. Os autores defendem que esse algoritmo apresenta uma boa acurácia e um baixo custo computacional quando comparado aos outros algoritmos que empregam a metodologia das trajetórias V-I tradicionais. No entanto não é realizada uma comparação com esses métodos e sim com os métodos que utilizam análises no domínio na frequência. Além disso, em Gao *et al.* (2015) são apresentados resultados de classificação com diferentes métodos de extração para a base de dados pública PLAID (do inglês *Plug Load Appliance Identification Dataset*), essa abordagem de grade binária é analisada, com uma matriz de 16x16. Os melhores resultados obtidos para essa abordagem são menores que 82% e quando combinada a outros métodos menores que 87%.

Também partindo por uma abordagem diferente do método de extração com as trajetórias V-I tradicionais, em Baets *et al.* (2018b) os autores as utilizaram na forma de imagens ponderadas como entrada em uma RNC. Entretanto os resultados obtidos para as bases de dados públicas utilizadas, PLAID e WHITED (do inglês *Worldwide Household and Industry Transient Energy Data Set*), foram inferiores a 80%. Já em Baets *et al.* (2019), os autores obtiveram resultados maiores que 85%, para as mesmas bases de dados, utilizando uma metodologia muito parecida ao trabalho anterior, porém a imagem era utilizada como entrada para uma rede neural siamesa³.

¹ Reúne característica de métodos de extração diferentes como entrada de um modelo de classificação.

² Se a trajetória V-I cruzar pela célula, essa célula é ocupada pelo valor 1 na matriz de representação.

³ É uma topologia especial de Rede Neural que apresenta duas redes neurais idênticas. Como entrada as redes recebem uma imagem das trajetórias V-I e um rótulo informando se ambas as imagens são da mesma classe ou não. A saída da rede são dois vetores que formam uma representação de dimensão inferior as duas imagens de entrada. Se as imagens pertencerem a mesma classe, a distância entre as representações será menor que um determinado limiar.

Ainda no contexto das metodologias que utilizam imagens das trajetórias V-I como entradas em uma RNC é Baptista *et al.* (2018). Nesse trabalho os autores implementaram o método em *hardware*, utilizando uma FPGA (do inglês *Field Programmable Gate Array*). O tempo e o consumo de potência foram evidenciados, pois eles obtiveram uma resposta de classificação em um tempo inferior (aproximadamente 5,7 ms) a extensão da janela de 333 ms considerada para o sistema e o consumo da placa ficou estimado em aproximadamente 1,9 W. O resultado de tempo, principalmente, é relevante pois com ele sendo bem inferior ao tempo de janela total considerado no sistema, isso implica que a classificação é realizada antes de terminar o próximo tamanho de janela, não ocorrendo sobreposição de informações. Entretanto, o desempenho de classificação desse trabalho permaneceu menor que 80%.

Outro trabalho que também utiliza a aprendizagem profunda é Zhibin e Hong (2019) no qual os autores propõem um algoritmo de triagem de dados, baseado em campo finito (domínio \mathbb{R}), antes da utilização da imagem na rede neural, visto que segundo os autores o ruído pode interferir nas amostras. Os resultados exibidos no trabalho mostraram que a metodologia proposta aumenta a acurácia no reconhecimento das cargas quando comparada à metodologia tradicional. Porém, os autores utilizaram somente o recurso da área para classificação das cargas na metodologia tradicional.

Outra abordagem interessante aplicando as trajetórias V-I aparece no trabalho de Baets *et al.* (2017), em que os descritores elípticos de Fourier são utilizados como características para a classificação de cargas. Para essa implementação, eles realizam a representação da trajetória V-I de uma carga em uma célula de grade binária de 16 por 16. Na sequência, utilizaram um algoritmo para identificação dos contornos da imagem, no qual cada pixel da imagem era examinado e combinado com um dos dezesseis casos possíveis nos quais eles determinaram anteriormente. Como esse algoritmo de identificação poderia encontrar mais do que um contorno, dependendo do tipo da trajetória, somente o contorno externo era considerado na aplicação. Após a obtenção do contorno era realizada a extração dos descritores de Fourier, descartando os descritores escalares e associando esses valores ao rótulo da carga correspondente. Por fim, os vetores resultantes eram utilizados pelos modelos de classificação (regressão logística, *Random Forest* e Rede Neural). Para validação da metodologia proposta, (BAETS *et al.*, 2017) utilizaram a base de dados PLAID e consideraram um intervalo de parâmetros para os classificadores utilizados, bem como para a quantidade de harmônicos para cálculo dos descritores. Os melhores resultados obtidos na proposta não ultrapassaram 80% na acurácia.

Já em Baets *et al.* (2018a), os autores realizam a comparação da metodologia criada anteriormente, dos descritores de Fourier e uma nova, na qual utilizam as imagens das trajetórias como entrada para uma RNC. As conclusões que eles chegaram são: (i) ambos os métodos são capazes de classificar os aparelhos; (ii) após uma determinada resolução adicionar informações, harmônicos, não é útil e pode levar a um desempenho pior e (iii) a metodologia utilizando as imagens com RNC obtém um desempenho ligeiramente maior do que a metodologia com os descritores de Fourier utilizando a RF. Porém, os melhores resultados obtidos não ultrapassaram 88% de acurácia de classificação.

Na linha tradicional de extração de características utilizando a trajetória V-I, em Wang *et al.* (2018) os autores propõem um algoritmo que realiza a estimativa de 10 características, sendo nove delas já discutidas em trabalhos anteriores (alcance da corrente, área, direção do *loop*, assimetria, curvatura da linha média, auto-intersecção, pico do segmento médio, variação do segmento médio e área dos segmentos direito e esquerdo) mais a estimativa da variação da admitância instantânea, definidas na seção 3.1. Os testes desse algoritmo foram realizados utilizando duas bases de dados, REDD e uma base do próprio grupo de pesquisa, em que os resultados obtidos (maiores que 88%) demonstraram que as trajetórias são capazes de distinguir aparelhos com características elétricas diferentes e consumo de energia semelhante, além de aparelhos com multi-estado. Contudo, os autores destacam duas principais limitações do algoritmo: ele não consegue manipular um evento no qual vários dispositivos mudam de estado em um intervalo de tempo muito curto⁴ e a falta de um banco de dados público adequado para realizar o treinamento do algoritmo não-supervisionado.

Uma alternativa para as limitações com o treinamento dos modelos de classificação não-supervisionados foi proposto em Liu *et al.* (2019b) ao utilizar a transferência de aprendizado no contexto NILM, empregando uma RNC pré-treinada para classificar imagens coloridas das trajetórias V-I representadas visualmente em um espaço de cores HSV (do inglês *Hue, Saturation and Value*). Os testes nesse trabalho foram realizados utilizando as bases de dados PLAID e WHITED e os resultados obtidos com o método proposto apresentaram uma melhora na precisão comparado a outros métodos que utilizaram as trajetórias como imagens binárias. Porém, o modelo utilizado para o pré-treinamento, o AlexNet, apresenta diversos parâmetros a serem estimados (maiores que 35k), o que torna o processamento computacional custoso para essa abordagem.

⁴ Intervalo de tempo entre dois eventos menor do que $T+d$, em que T é o período em segundos antes e depois de um evento e d é duração do evento anterior (WANG *et al.*, 2018)

Entre os trabalhos mais recentes que utilizam as trajetórias V-I tem-se Caiman e Dragomir (2019), em que os autores apresentam um método de assinaturas com base nos parâmetros obtidos a partir de funções de regressão não-linear e algoritmos genéticos para aproximar-se a trajetória de tensão-corrente de um aparelho. Nos resultados, os autores descrevem que o método proposto se aproximou muito bem das medidas para 4 classes de equipamentos mesmo quando conectados em paralelo, mas que esses resultados só seriam publicados em um trabalho futuro. Já em Yang e Soh C. S. and Yap (2019), os autores propõem um algoritmo de classificação multi-estágio que combina a metodologia de análise de componentes principais – ACP e a metodologia do KVMP para realizar o reconhecimento de imagens das trajetórias V-I. No entanto, o trabalho se concentra na determinação do valor de k , quantidade de vizinhos mais próximos, em cada estágio para que o algoritmo seja o mais eficaz na classificação.

No trabalho de Saha *et al.* (2019) os autores propõem uma nova taxonomia de cargas abrangendo um número maior de classes do que em Lam *et al.* (2007). Para essa abordagem eles criaram assinaturas de carga utilizando as componentes diretas e de quadratura dos sinais de corrente e tensão e desses mesmos sinais com deslocamento de 90° , além das assinaturas tradicionais de trajetórias V-I. Na sequência extraíram características (área, direção do *loop*, simetria e curvatura da linha média) de algumas dessas formas e realizaram o agrupamento das cargas dada a similaridade das características. Segundo os autores a taxonomia foi realizada utilizando um número menor de características extraídas quando comparadas a metodologia de Lam. Porém, é válido ressaltar que os autores realizaram um número menor de extrações se considerado somente a manipulação sobre uma das trajetórias obtidas e não sobre o número obtido de extrações de todas as novas assinaturas. Entretanto, os testes realizados nesse trabalho utilizaram somente um banco de dados com cargas individuais, não apresentando os possíveis impactos do desenvolvimento em cargas agregadas.

Por fim, em Mulinari *et al.* (2019) os autores propõem novas características para a metodologia das trajetórias V-I tradicionais. A diferença desse trabalho para os demais é que nele os autores utilizam as trajetórias obtidas em estado estacionário e transitório, destacando as vantagens de agregar informações da transição no desempenho de classificação. Além disso, os autores analisaram a possibilidade de selecionar, entre todas as características, os recursos que causavam maior influência na classificação, permitindo a criação de um conjunto menor de extração. Entretanto, as análises realizadas no trabalho abordam apenas uma base de dados com cargas individuais denominada COOLL (do inglês *Controlled On/Off Loads Library*), não tendo

sido testada a eficiência da metodologia proposta em cenários com aquisições multi-cargas, com mais de uma carga alterando seu estado de operação.

2.5 CONTRIBUIÇÕES DESTE TRABALHO

Ao averiguar o estado da arte do método de extração de características que utiliza as trajetórias V-I é possível observar que grande parte dos autores, para aumentar o desempenho da classificação de cargas, partiram para abordagens de desenvolvimento incluindo novas características extraídas da trajetória V-I somente em estado estacionário, ou para abordagens nas quais exigem maior processamento computacional, como as abordagens com imagens e redes neurais.

Este trabalho utiliza as trajetórias V-I em duas abordagens: uma que visa desenvolver novas características a respeito da forma obtida da trajetória, para o método denominado tradicional de extração, e outra abordagem que utiliza a própria descrição da forma da trajetória para classificação de cargas elétricas. Na linha da metodologia tradicional de extração, esse trabalho visa descrever as novas características proposta em Mulinari *et al.* (2019), no qual utiliza-se as trajetórias V-I obtidas em estado estacionário e transitório das cargas, além de expandir as conclusões obtidas no trabalho para outras bases de dados, principalmente as que apresentam aquisições de cargas múltiplas. Já na linha de desenvolver um método de extração que descreva a própria forma, esse trabalho emprega a transformada de Fourier bidimensional, denominada como Fourier 2D, sobre a forma obtida da trajetória nos dois estados da carga.

Ainda, tendo em vista a análise da complexidade computacional na etapa de extração de característica, buscou-se implementar as duas abordagens em um sistema embarcado. Essas implementações possibilitam realizar avaliações de recursos necessários para cada método, como memória e tempo computacional, recursos estes que são pouco exploradas nos trabalhos presentes no estado da arte das trajetórias V-I.

3 FUNDAMENTAÇÃO TEÓRICA

Conforme exposto no capítulo anterior, entre as abordagens de extração de características utilizadas em NILM, o método que mensura os recursos associadas à trajetória V-I destaca-se por apresentar diferentes formas bidimensionais para cargas elétricas distintas, permitindo uma categorização eficiente. Neste capítulo são detalhadas as características existentes na literatura para o método tradicional das trajetórias V-I, além de apresentar as definições sobre as Séries de Fourier bidimensional, a qual, posteriormente, é empregada como método de extração com essa mesma assinatura de carga. Também, são apresentados os métodos de classificação supervisionada utilizados para análise neste trabalho, bem como os métodos utilizados para seleção de características.

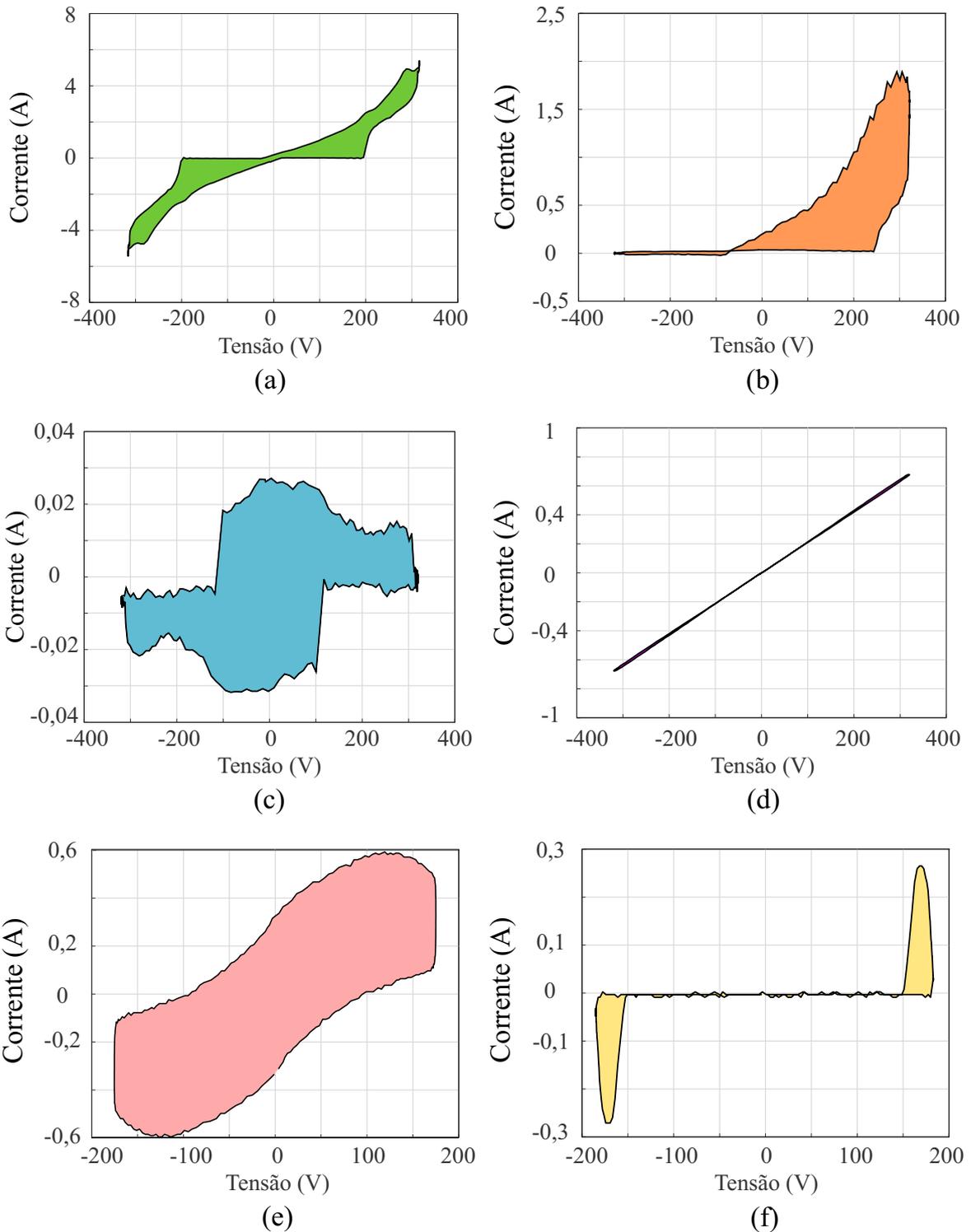
3.1 MÉTODO TRADICIONAL DA TRAJETÓRIA V-I

Lam *et al.* (2007) foram os primeiros a apresentar uma taxonomia de cargas utilizando as assinaturas obtidas de diferentes aparelhos eletrodomésticos de uma residência geradas a partir de uma metodologia bidimensional. Essa metodologia foi denominada como trajetória V-I.

A trajetória V-I consiste em obter uma assinatura de carga por meio da análise do diagrama bidimensional obtido ao plotar um ciclo da tensão (no eixo horizontal) por um ciclo da corrente (no eixo vertical), ambos os ciclos obtidos no estado estacionário de uma carga conectada a rede. Diferentes formas são obtidas ao realizar esse procedimento para aparelhos distintos ou em estados de funcionamento diferentes, como é possível observar na Figura 8.

Ao utilizar a metodologia tradicional de extração de características com as trajetórias V-I são obtidas as características sobre a forma bidimensional, tipicamente em estado estacionário. Lam *et al.* (2007) discutiram algumas dessas características e como elas se relacionavam com o tipo da carga e suas propriedades elétricas. Já Wang *et al.* (2018) propõem maneiras para quantizar essas características além de acrescentarem mais duas características às oito já propostas anteriormente. A seguir, essas características serão detalhadas.

Figura 8 – Trajetórias V-I de: (a) Furadeira em velocidade máxima (750 W) e (b) em velocidade média (750 W); (c) Lâmpada LED (1,6 W); (d) Lâmpada de halôgêneo (105 W) da base de dados COOLL e (e) Exaustor de Fumaça (23 W); (f) Lâmpada LED (6 W) da base de dados LIT-SYN.



Fonte: Adaptado de Mulinari *et al.* (2019).

3.1.1 Corrente atual

A característica denominada como corrente atual (itc) é a que está relacionada à magnitude da potência ativa de uma carga (WANG *et al.*, 2018). Essa característica é mensurada utilizando a corrente (I) do aparelho obtida no ciclo, por meio da seguinte diferença:

$$itc = \max(I) - \min(I). \quad (1)$$

Essa característica já havia sido determinada por Hassan *et al.* (2014) e voltou a ser utilizada por Wang *et al.* (2018) posteriormente.

3.1.2 Variação da admitância instantânea

A variação da admitância instantânea (D) avalia a constância dos valores de admitância de uma carga. Isso ajuda na distinção de cargas resistivas de cargas não-resistivas, pois essas cargas apresentam os valores da variação menores do que comparados aos valores de variações das cargas não-resistivas. Wang *et al.* (2018) obtiveram essa característica por meio de:

$$s_i = \frac{I_i}{V_i} \quad i \in \{1, 2, \dots, NT\} \cap |V_i| > vl, \quad (2)$$

$$D = std(s),$$

em que NT refere-se ao número de amostras no ciclo da trajetória, vl é um valor limiar da tensão para excluir valores muito próximo a zero e std a função de desvio padrão.

3.1.3 Área

A característica da área (ar) mensura o valor da área interna da trajetória V-I. Essa característica é diretamente proporcional à magnitude do deslocamento de fase entre a tensão e a corrente da carga. Desse modo, quanto maior o deslocamento entre essas medidas maior é a área interna da trajetória V-I. Por esse motivo, cargas resistivas apresentam trajetórias com áreas menores quando comparadas a outros tipos de carga, aproximando-se a uma reta.

Wang *et al.* (2018) determinaram uma etapa de divisão da trajetória em duas partes, denominadas por parte A e parte B, antes do dimensionamento da área de uma trajetória V-I. Essa divisão é realizada de acordo com os índices dos pontos máximo (i_{Vmax}) e mínimo (i_{Vmin}) da tensão, considerando o índice do ponto de máximo o inicial. Desse modo, as partes são definidas

por:

$$\begin{aligned} A &= ((V_q, I_q) | q \in \{1, 2, \dots, i_{Vmin}\}), \\ B &= ((V_q, I_q) | q \in \{i_{Vmin} + 1, \dots, NT\}). \end{aligned} \quad (3)$$

Após a realização da divisão da trajetória nas partes A e B é possível mensurar a área de uma trajetória por meio da equação:

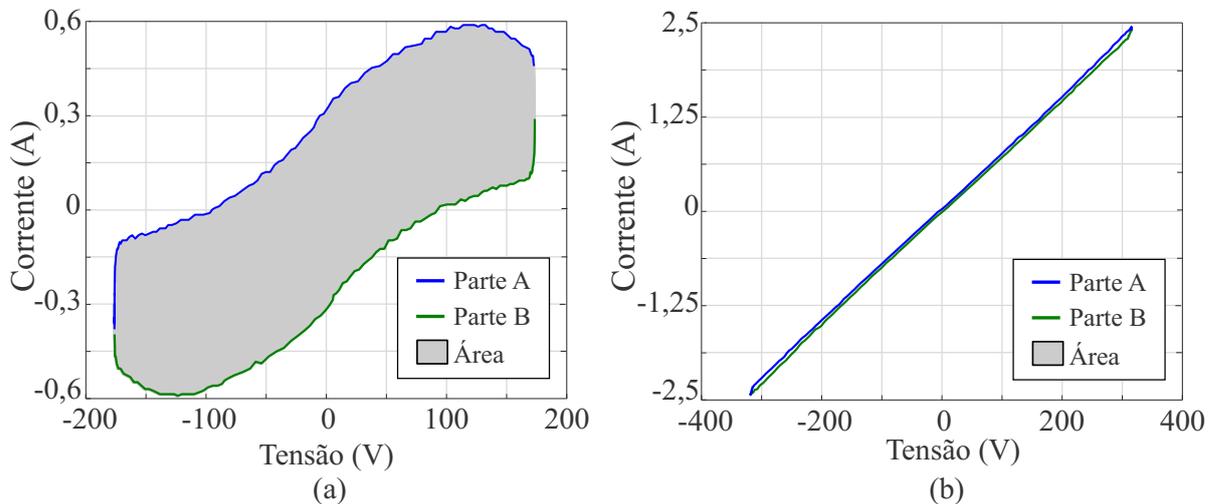
$$ar = \sum_i \frac{1}{2} |V_j - V_i| (|I_{i'} - I_i| + |I_{j'} - I_j|), \quad i \in \{1, \dots, i_{Vmin} - 1\}, \quad (4)$$

em que i' e j' representam os índices da parte B para os quais a tensão está mais próxima dos dois pontos consecutivos i e j , obtidos por meio de:

$$\begin{aligned} \operatorname{argmin}_{i'} |V_{i'} - V_i|, \quad i \in \{i_{Vmin} + 1, \dots, NT\}, \\ \operatorname{argmin}_{j'} |V_{j'} - V_j|, \quad j \in \{i_{Vmin} + 1, \dots, NT\}. \end{aligned} \quad (5)$$

A Figura 9 apresenta as trajetórias V-I de duas cargas com áreas distintas, indicando suas respectivas partes A e B. Nessa figura a carga (a) apresenta um deslocamento de fase maior que a carga apresentada em (b), como é possível verificar na Figura 10 em que são apresentadas as medidas de tensão e corrente referentes a cada uma dessas trajetórias (a) e (b).

Figura 9 – Trajetórias V-I exibindo a característica ar : (a) do Exaustor de Fumaça (23 W) da base de dados LIT-SYN e (b) Pistola Elétrica de Pintura (1050 W) da base de dados COOLL.

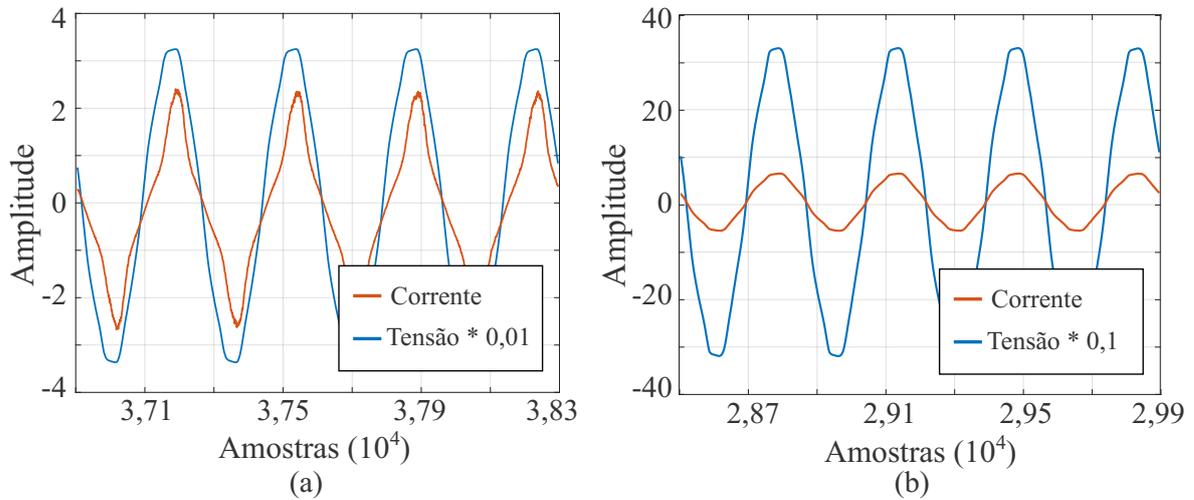


Fonte: Autoria própria.

3.1.4 Direção do *loop*

A característica denominada como direção do *loop* (*lpa*) consiste em obter o sentido do deslocamento de fase entre a tensão e a corrente da carga, relacionando-o aos tipos das

Figura 10 – Diferença de Tensão e Corrente para ar: (a) Exaustor de Fumaça (23 W) da base de dados LIT-SYN e (b) Decapante Elétrico (1050 W) da base de dados COOLL.



Fonte: Autoria própria.

cargas (resistiva, capacitiva e indutiva). Se a carga apresenta um deslocamento de fase positivo, a direção do deslocamento é horária e tem-se a corrente adiantada em relação à tensão, indicando uma carga com característica capacitiva. Caso contrário, se o sentido da direção é anti-horária, representando um deslocamento de fase negativo, tem-se a corrente atrasada em relação a tensão, caracterizando uma carga indutiva. Em casos de um deslocamento de fase nulo, tem-se que a carga é puramente resistiva. A direção de giro pode ser estimada por meio da equação:

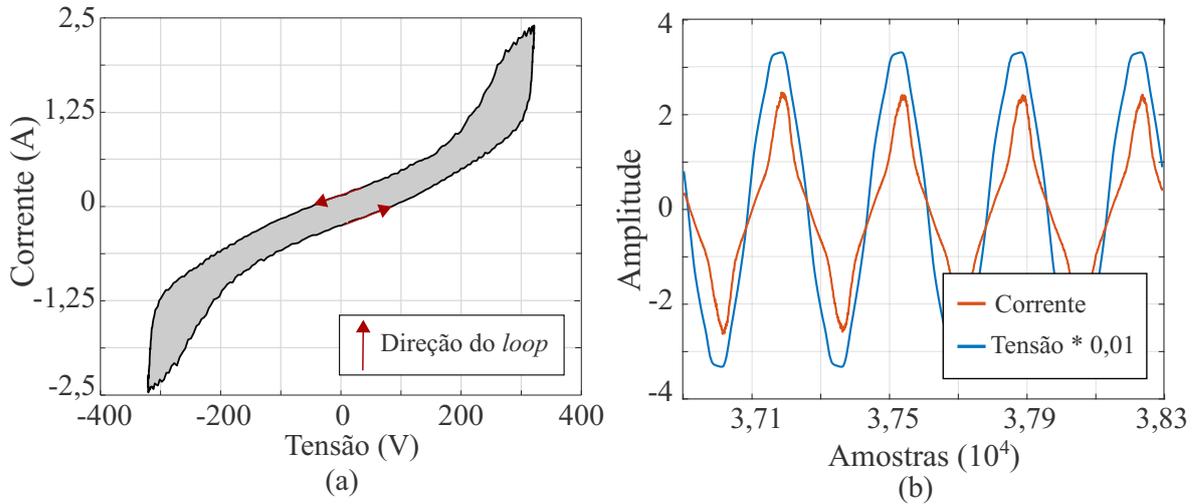
$$lpa = \sum_u \frac{1}{2} (V_{u+1} - V_u)(I_{u+1} - I_u), \quad u \in \{1, \dots, NT - 1\}. \quad (6)$$

A direção do *loop* na carga apresentada na Figura 11(a) é anti-horária, indicando uma carga com a corrente atrasada em relação a tensão, como se pode verificar nas medidas dessa carga apresentadas na Figura 11(b).

3.1.5 Assimetria

A assimetria (*asy*) é utilizada para avaliar a semelhança entre a trajetória V-I de uma carga e essa trajetória rotacionada em 180°. Essa característica se relaciona à condução da corrente nos ciclos positivos e negativos. Se as trajetórias original e rotacionada são simétricas, isso significa que a condução da corrente ocorre da mesma maneira nos dois ciclos. Caso contrário, quanto mais assimétricas as curvas, maior é a diferença na forma da corrente no ciclo positivo e do ciclo negativo. Em Wang *et al.* (2018), os autores propõem mensurar essa

Figura 11 – Trajetórias V-I destacando *lpa* para: (a) a Furadeira Elétrica em velocidade máxima (701 W) da base de dados COOLL e (b) Tensão e corrente da respectiva carga.

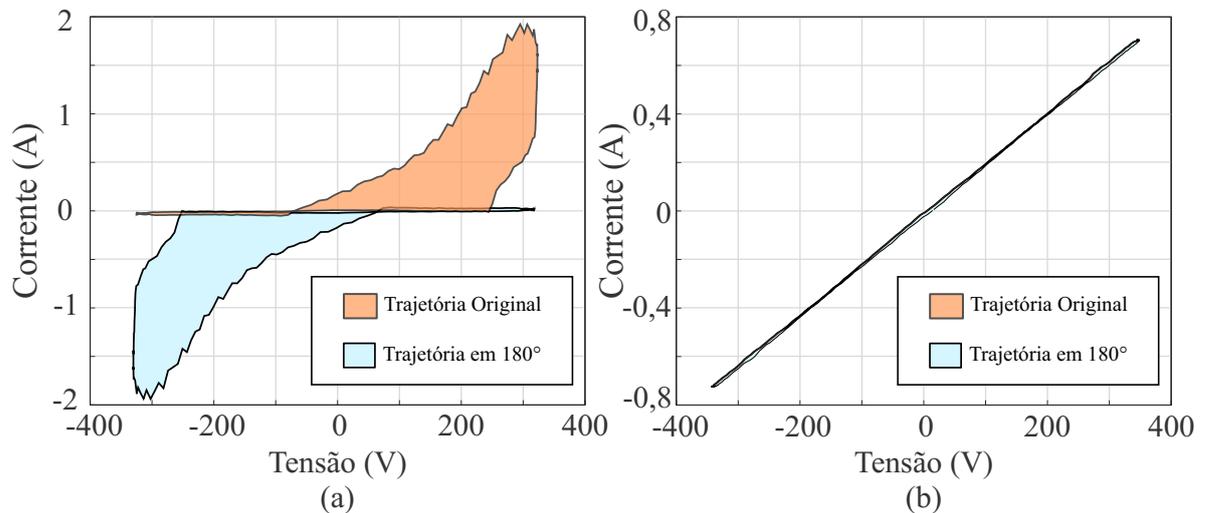


Fonte: Autoria própria.

característica utilizando a distância de Hausdorff¹ entre as trajetórias.

A Figura 12 apresenta dois exemplos de trajetórias: (a) assimétrica e (b) simétrica, por consequência da condução da corrente nos ciclos positivo e negativo, como se pode verificar na Figura 13, em que são apresentadas as medidas de tensão e corrente das cargas, respectivamente.

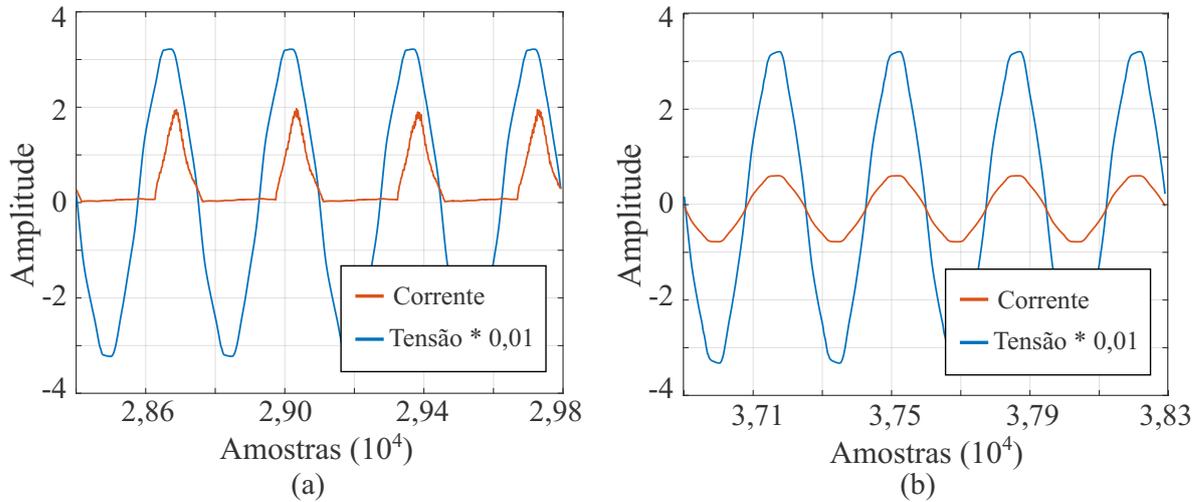
Figura 12 – Trajetórias V-I destacando *asy*: (a) Furadeira em velocidade média (750 W) e (b) Lâmpada halógena (105 W), ambas da base de dados COOLL.



Fonte: Autoria própria.

¹ A distância de Hausdorff é capaz de medir quão similares são dois conjuntos de pontos dispersos em um espaço bidimensional (DUBUISSON; JAIN, 1994).

Figura 13 – Tensão e Corrente: (a) Furadeira em velocidade média (750 W) e (b) Lâmpada halógena (105 W), ambas da base de dados COOLL, destacando *asy* devido a condução de corrente.



Fonte: Autoria própria.

3.1.6 Curvatura da linha média

A curvatura da linha média (r) determina a distorção de uma linha média na trajetória. Inicialmente essa linha média é obtida por meio da expressão:

$$M = \left\{ (x_{M,i}, y_{M,i}) \mid x_{M,i} = \frac{V_i + V_{i'}}{2}, y_{M,i} = \frac{I_i + I_{i'}}{2} \right\}, i \in \{1, 2, \dots, i_{Vmin}\}. \quad (7)$$

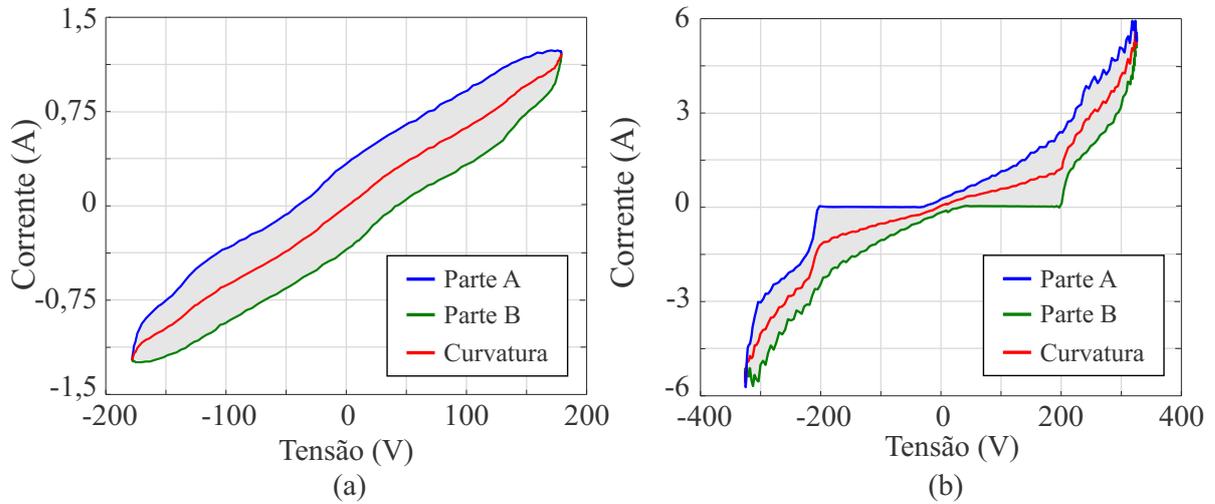
De acordo com Wang *et al.* (2018) a curvatura pode ser estimada por meio do coeficiente de correlação linear de todos os pontos que formam essa linha média.

A característica de curvatura da linha média está relacionada a não linearidade de uma carga, pois quanto mais curvada essa linha é, maior é o conteúdo harmônico da carga (LAM *et al.*, 2007). A Figura 14 exibe duas cargas com suas respectivas linhas médias, sendo (a) uma carga mais linear e (b) uma carga não-linear. O conteúdo harmônico de cada uma dessas cargas pode ser observado na Figura 15, em que (a) é a corrente da carga com características mais lineares e (b) é a corrente da carga com distorções harmônicas.

3.1.7 Auto Intersecção

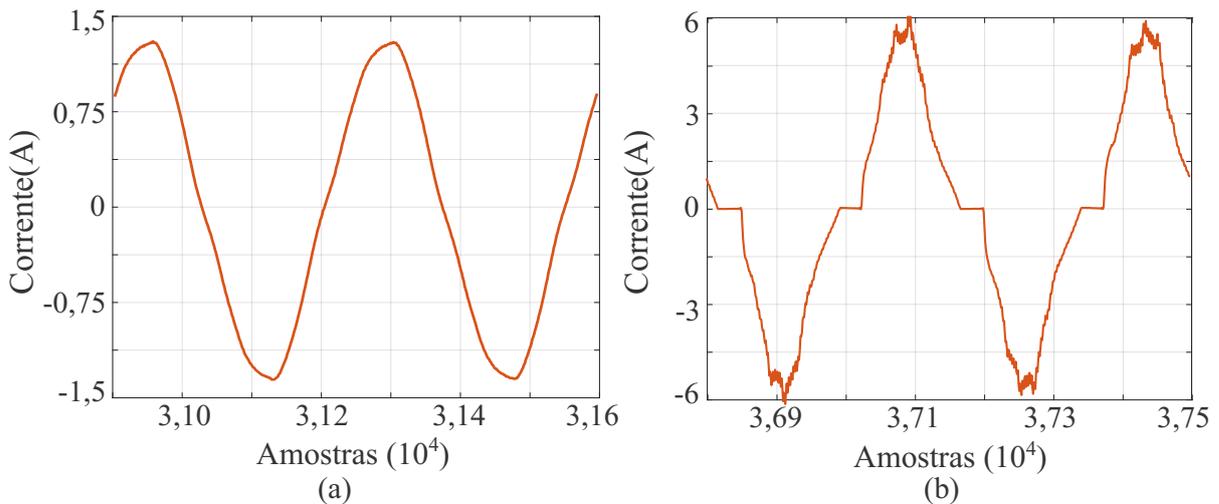
A característica da auto intersecção (sc) contabiliza a quantidade de cruzamentos a trajetória V-I contém. Alguns aparelhos não lineares podem apresentar uma curva V-I com cruzamentos, por esse motivo essa característica também está relacionada ao conteúdo harmônico de uma carga. De acordo com Lam *et al.* (2007), a quantidade dessas intersecções está

Figura 14 – Trajetórias V-I destacando r : (a) Ventilador (80 W) da base de dados LIT-SYN e Furadeira em velocidade máxima (750 W) da base de dados COOLL.



Fonte: Autoria própria.

Figura 15 – Diferença entre as correntes: (a) Ventilador (80 W) da base de dados LIT-SYN e Furadeira em velocidade máxima (750 W) da base de dados COOLL, para a característica r .



Fonte: Autoria própria.

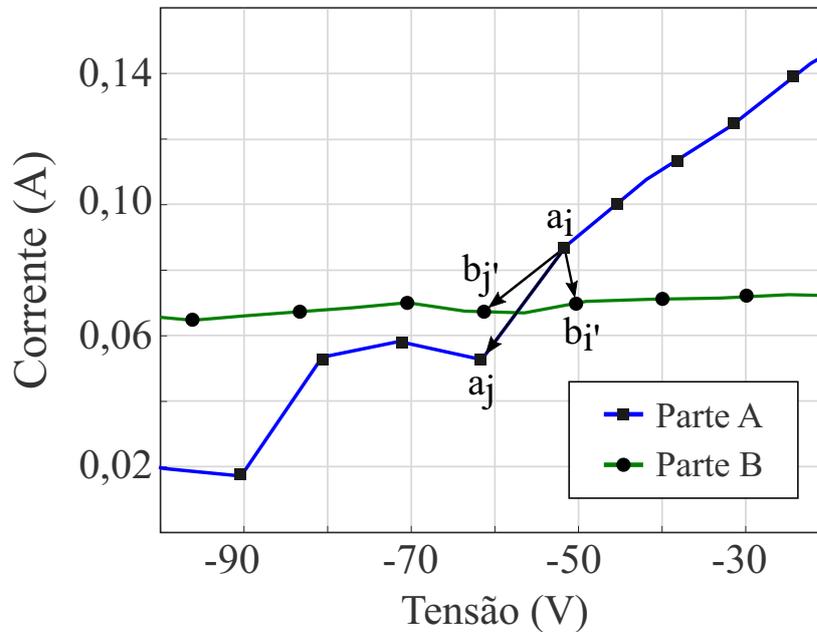
relacionada à ordem mais significativa de harmônicos que a corrente dessa carga contém. Assim, se uma trajetória tem dois cruzamentos, a corrente dessa carga apresenta o terceiro harmônico significativo. De acordo com Du *et al.* (2016), a quantidade de intersecções na trajetória pode ser obtida por meio da equação:

$$sc = \left((a_i \vec{a}_j) \times (a_i \vec{b}_{i'}) \right) \cdot (a_i \vec{a}_j) \times (a_i \vec{b}_{j'}) < 0, \quad (8)$$

em que a_i e a_j referem-se as amostras da trajetória na parte A, nos índices i e j respectivamente, definidos na Subseção 3.1.3. Da mesma maneira os pontos da parte B da trajetória são adquiridos nos índices i' e j' . Já $a_i \vec{a}_j$ significa a obtenção do vetor por meio da diferença de a_j em relação

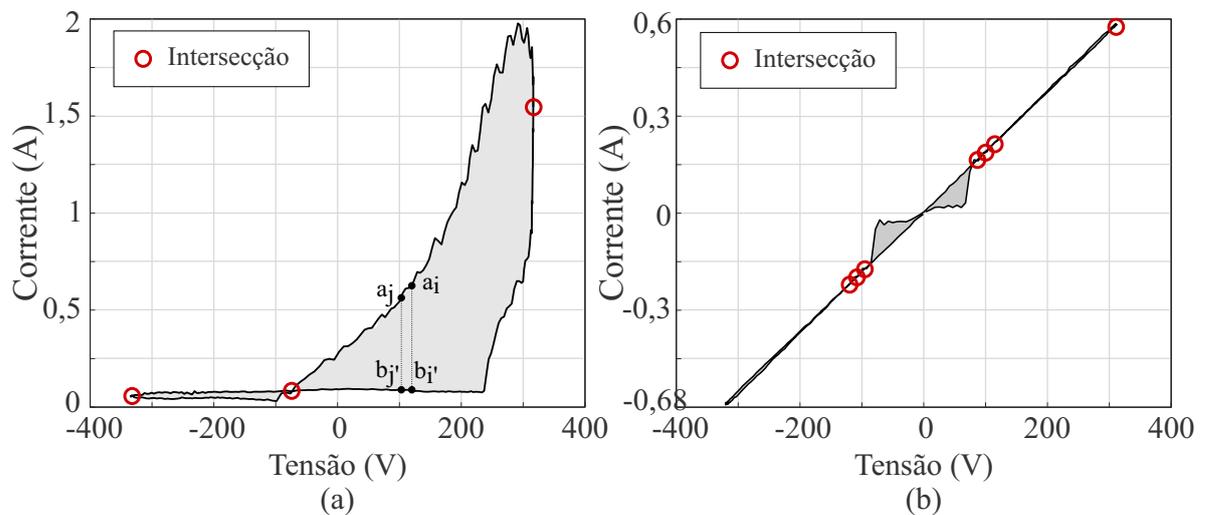
à a_i , como é possível observar na Figura 16. Já a Figura 17 exibe dois exemplos de cargas em que suas trajetórias apresentam auto-intersecções, pois ambas apresentam conteúdo harmônico. Por esse motivo, as correntes dessas cargas não são senoidais puras, como se pode visualizar na Figura 18, em que as correntes respectivas a cada trajetória são exibidas na mesma ordem.

Figura 16 – Exemplo de alguns vetores utilizados na Equação 8 para o cálculo de sc da Furadeira em velocidade média (750 W) da base de dados COOLL.



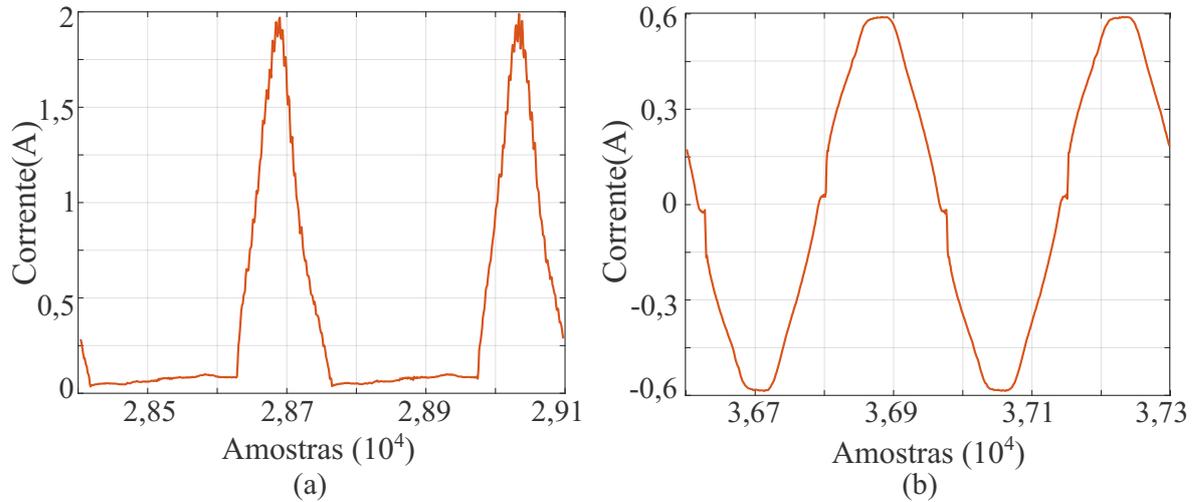
Fonte: Autoria própria.

Figura 17 – Trajetórias V-I destacando sc : (a) Furadeira em velocidade média (750 W) e (b) Lâmpada halógena (100 W), ambas da base de dados COOLL.



Fonte: Autoria própria.

Figura 18 – Corrente: (a) Furadeira em velocidade média (750 W) e (b) Lâmpada halógena (100 W), ambas da base de dados COOLL.



Fonte: Autoria própria.

3.1.8 Pico do segmento médio

Para determinar as características do pico do segmento médio (mi), inclinação do segmento médio (sh) e área dos segmentos direito e esquerdo (alr), Wang *et al.* (2018) propõem que a divisão da trajetória V-I em segmentos (esquerdo, médio e direito) seja realizada a partir dos pontos de inflexão $a1$ e $b1$, obtidos na parte A e B, respectivamente, encontrados por meio da equação:

$$bp = \max \left(\arccos \left| \frac{(\vec{m}\vec{n}) \cdot (\vec{o}\vec{p})}{|\vec{m}\vec{n}| |\vec{o}\vec{p}|} \right| \right), \quad (9)$$

em que m , n , o e p são quatro pontos consecutivos na parte da trajetória em questão. Já os pontos $a2$ e $b2$ são os pontos que ao utilizar uma linha, a partir dos pontos de inflexão $a1$ e $b2$, encontra-se a tensão mais próxima. Assim, a trajetória fica dividida por essas linhas em segmentos esquerdo, médio e direito, como é possível observar na Figura 19.

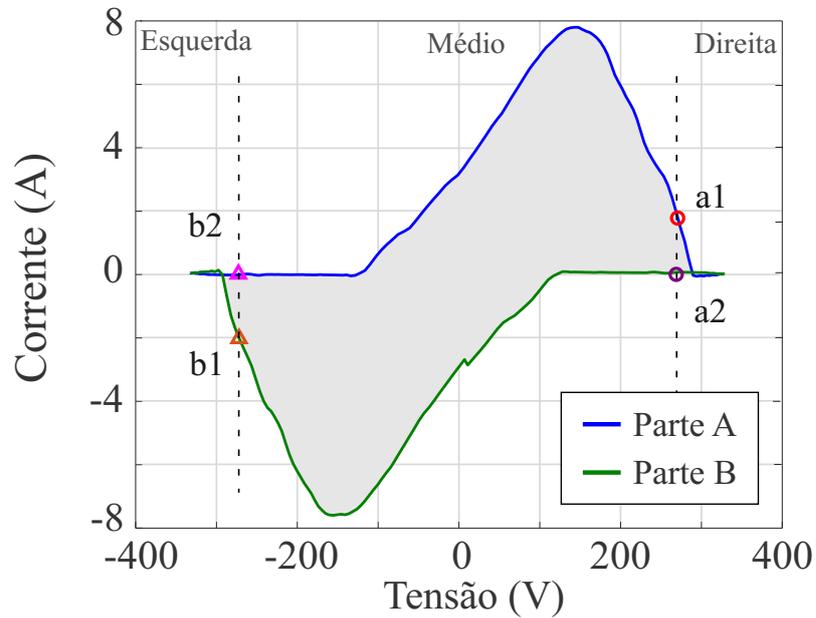
O mi é definido como a distância vertical máxima entre a trajetória e a linha reta que liga as extremidades do segmento médio. Essa característica ajuda na distinção de cargas eletrônicas com baixa potência, pois os valores para esse tipo de carga geralmente são inferiores do que outros tipos de carga. O valor do pico do segmento médio pode ser obtido por:

$$mi = \max(\max(I_r - f_a(V_r)), \max(I_t - f_b(V_t))), \quad (10)$$

$$r \in \{m_a, \dots, m_a + n_a - 1\}, \quad t \in \{m_b, \dots, m_b + n_b - 1\},$$

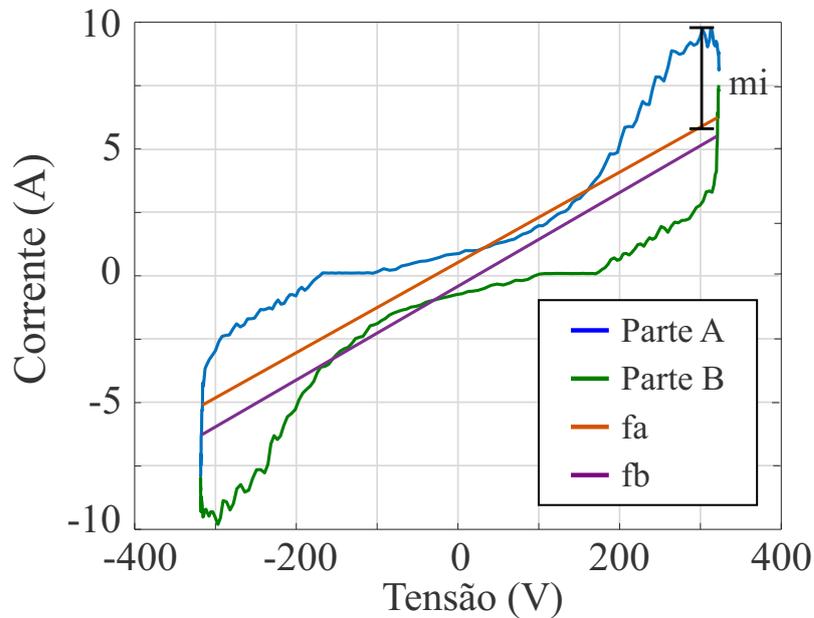
em que $f_a(\cdot)$ é a expressão que representa a linha reta determinada pelos pontos $a1$ e $b2$ e $f_b(\cdot)$ a expressão da linha reta determinada por $b1$ e $a2$. Já m_a e m_b são os primeiros pontos no segmento

Figura 19 – Trajetória V-I destacando os pontos de inflexão na parte A e B utilizando o Aspirador Tornado (700 W) da base de dados COOLL.



médio na parte A e B, além de n_a e n_b que são os números de pontos no segmento médio de cada parte respectivamente. A Figura 20 apresenta as linhas retas f_a e f_b desenhadas a partir dos pontos mencionados, além da indicação da localização do pico na carga escolhida.

Figura 20 – Trajetória V-I destacando mi do Aspirador de pó Dilem 2406 da base de dados COOLL.



Fonte: Autoria própria.

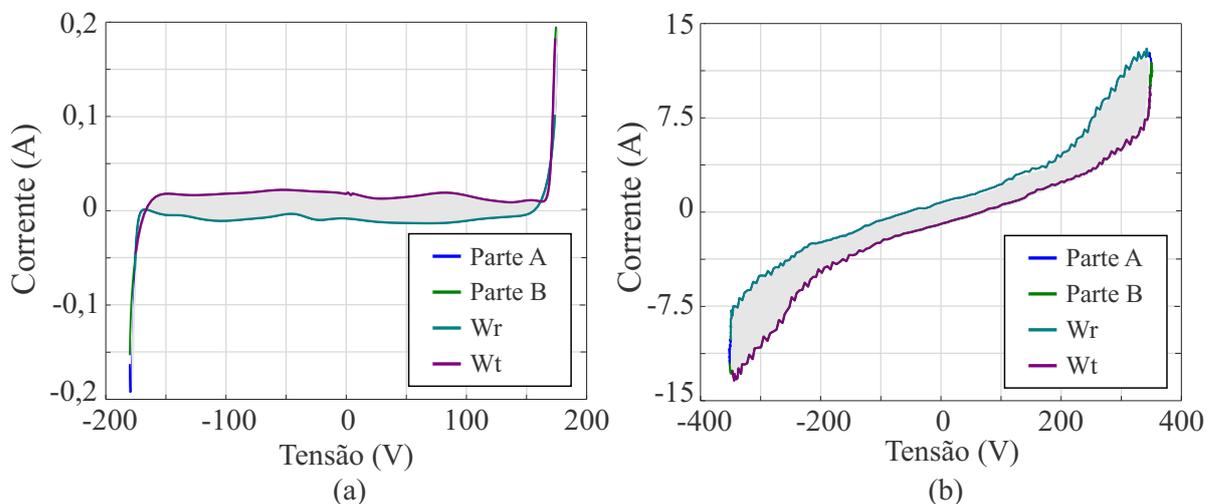
3.1.9 Inclinação do segmento médio

A inclinação do segmento médio (sh) determina o valor da inclinação do segmento médio da trajetória definido pelos pontos $a1$, $b1$, $a2$ e $b2$. A inclinação do segmento médio também auxilia na distinção das cargas de eletrônica de potência dos outros tipos de carga, pois esse segmento nesse tipo de carga é mais horizontal, com inclinação próximo a zero. Na Figura 21(a) é possível observar como o segmento médio, descrito pelos pontos Wr na parte A e Wt na parte B, apresentam uma inclinação mais horizontal quando comparada a carga apresentada na Figura 21(b). Essa característica das cargas de eletrônica, com inclinação próxima a zero, é consequência da condução da corrente nesse tipo de carga que apresenta um valor muito pequeno próximo a transição de positivo para negativo e vice-versa (LAM *et al.*, 2007), como pode-se verificar na Figura 22(a). O valor para a inclinação é obtido por:

$$sh = std(I_r) + std(I_t), \quad (11)$$

em que r e t estão definidos no mesmo intervalo dado na Equação 10.

Figura 21 – Trajetórias V-I com sh distintos: (a) Monitor LED (26 W) da base de dados LIT-SYN e (b) Esmerilhadeira (2000 W) da base de dados COOLL.

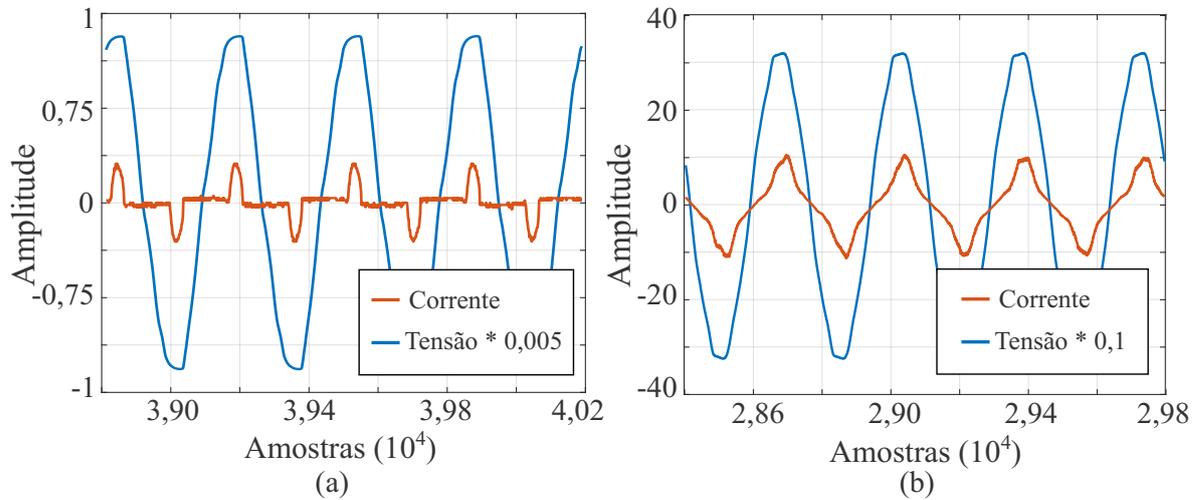


Fonte: Autoria própria.

3.1.10 Área dos segmentos direito e esquerdo

A característica denominada área dos segmentos direito e esquerdo (alr) mensura o valor das áreas dos segmentos determinados como direito e esquerdo pelos pontos de inflexão $a1$ e $b1$. Essa característica está relacionada ao deslocamento temporal entre os picos das formas de

Figura 22 – Tensão e Corrente: (a) Monitor LED (26 W) da base de dados LIT-SYN e (b) Esmerilhadeira (2000 W) da base de dados COOLL, para diferenciar a característica *sh*.



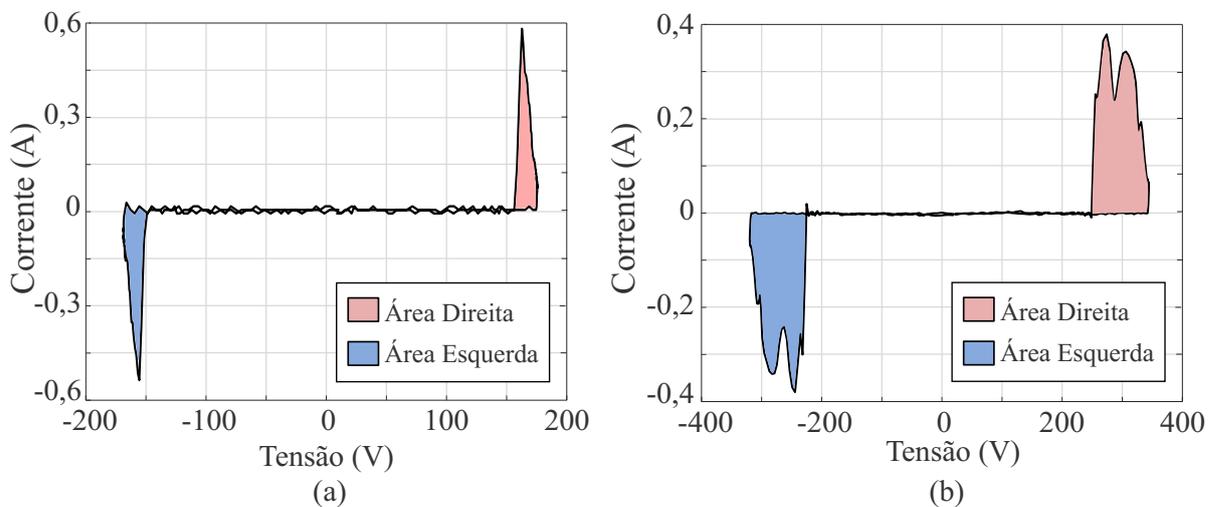
Fonte: Autoria Própria

onda de tensão e corrente, de modo que um deslocamento maior produzirá uma área maior dos segmentos. A área total desses segmentos pode ser obtida por:

$$alr = ar_{direito} + ar_{esquerdo}. \quad (12)$$

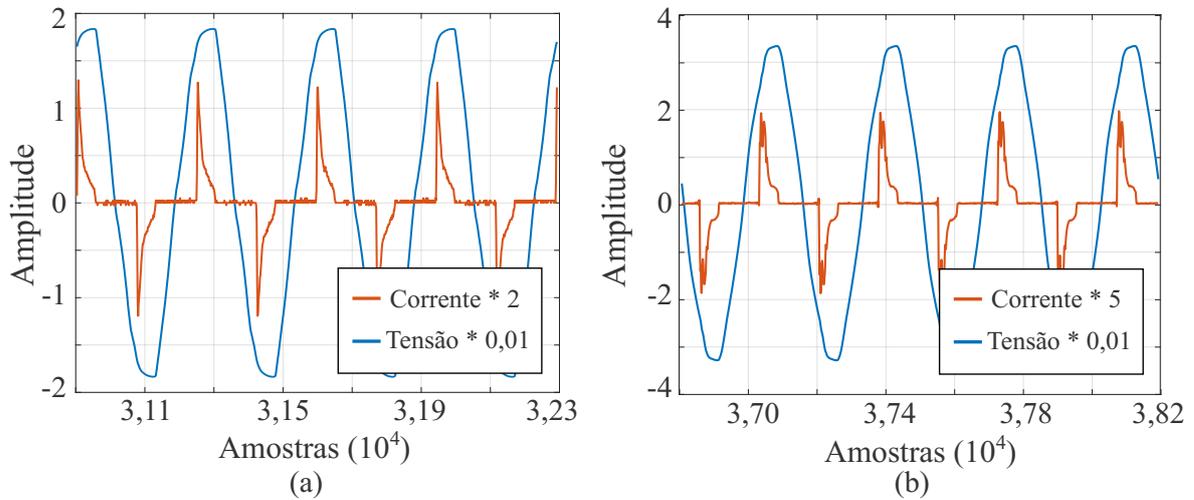
A Figura 23 exibe duas trajetórias V-I de cargas que apresentam deslocamento temporal distintos, como é possível observar na Figura 24 em que são apresentadas as respectivas medidas de tensão e corrente dessas cargas.

Figura 23 – Trajetórias V-I destacando *alr*: (a) Painel de LED (13 W) da base de dados LIT-SYN e Lâmpada Fluorescente Compacta (15W) da base de dados COOLL.



Fonte: Autoria própria.

Figura 24 – Tensão e Corrente: (a) Pannel de LED (13 W) da base de dados LIT-SYN e Lâmpada Fluorescente Compacta (15W) da base de dados COOLL, com diferentes valores de alr .



Fonte: Autoria própria.

3.2 MÉTODO DA FOURIER 2D

A Série de Fourier permite a representação de um sinal no domínio do tempo, contínuo ou discreto, para o domínio da frequência, por meio da representação desse sinal periódico utilizando um somatório de senos e cossenos:

$$f(n) = a_0 + \sum_{h=1}^{\infty} \left[a_h \cos \left(\frac{2\pi h}{T} n \right) + b_h \sen \left(\frac{2\pi h}{T} n \right) \right], \quad (13)$$

em que n representa as amostras obtidas do eixo n em um período de 2π , T o período fundamental do sinal, h a quantidade de harmônicos que se deseja analisar e a_h , b_h e a_0 os coeficientes do harmônico h , também chamados de descritores de Fourier (OPPENHEIM *et al.*, 1999). Esses coeficientes podem ser obtidos por meio das expressões:

$$a_0 = \frac{1}{T} \int_0^T f(n) dn, \quad (14)$$

$$a_h = \frac{2}{T} \int_0^T f(n) \cos \left(\frac{2\pi h}{T} n \right) dn, \quad (15)$$

$$b_h = \frac{2}{T} \int_0^T f(n) \sen \left(\frac{2\pi h}{T} t \right) dn. \quad (16)$$

Com a Equação 13, Fourier demonstrou que ondas complexas podem ser decompostas por meio das equações diferenciais compostas de dois termos, senos e cossenos. Assim, um sinal ou função periódicos podem ser representados e reduzidos a uma série de componentes

mais simplificadas. A amplitude dessas componentes senoidais individuais é delimitada pelos coeficientes das Equações 15 e 16 e é proporcional ao respectivo impulso espectral do harmônico no sinal original no domínio da frequência (CAPLE *et al.*, 2017). A abordagem da série de Fourier para representação de um sinal periódico também pode ser aplicada para funções com duas dimensões (x,y) . Desse modo, um contorno fechado² de um objeto pode ser descrito por:

$$f(x) = A_0 + \sum_{h=1}^{\infty} \left[A_h \cos \left(\frac{2\pi h}{T} x \right) + B_h \sin \left(\frac{2\pi h}{T} x \right) \right], \quad (17)$$

$$f(y) = C_0 + \sum_{h=1}^{\infty} \left[C_h \cos \left(\frac{2\pi h}{T} y \right) + D_h \sin \left(\frac{2\pi h}{T} y \right) \right], \quad (18)$$

em que A_h e B_h , são os descritores relativos ao harmônico h para coordenada x e C_h e D_h os descritores equivalentes para a coordenada y . Já A_0 e C_0 são os descritores escalares de Fourier, resultantes das coordenadas x e y , respectivamente, que representam o deslocamento do contorno da função em relação a origem (PERSON; FU, 1977). As Equações 14 15 e 16 podem ser aplicadas para determinar os descritores para cada coordenada, em que o método aplicado para integração nessas equações foi o Trapezoidal devido sua simplicidade de implementação e os erros de integração inerentes ao método serem aceitáveis para a implementação deste trabalho.

Entre as primeiras aplicações da abordagem bidimensional das séries de Fourier para a classificação de imagens, destaca-se o trabalho de Kuhl e Giardina (1982), em que os autores aplicaram as equações (17) e (18) para identificar e descrever quantitativamente o contorno de um avião a partir de um radar. Neste trabalho esta abordagem é denominada como Fourier 2D - F2D, porém ela também é conhecida como Análise Elíptica de Fourier - AEF³, visto que a combinação das duas equações de assinaturas para cada harmônico, em x e y , quando plotadas resulta em uma elipse. Assim, a forma de um contorno pode ser descrita por uma formação de cadeias de elipses interligadas que se movem no tempo umas sobre as outras. A Figura 25 apresenta a comparação do método tradicional utilizando a Série de Fourier com o método Elíptica de Fourier.

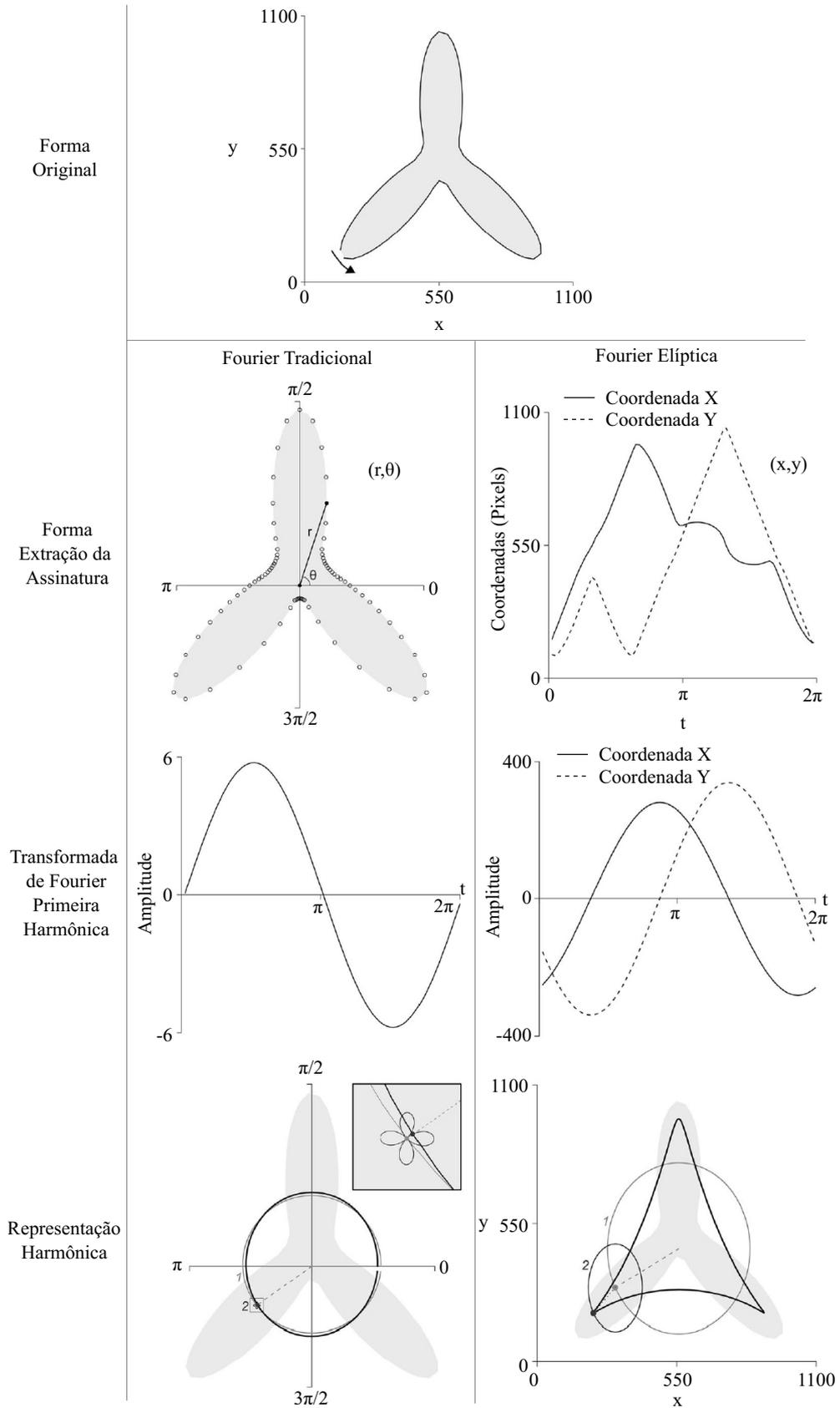
A Figura 25 demonstra a principal diferença entre o método tradicional e a forma elíptica, pois como é possível notar na etapa de extração, o método tradicional agrupa as informações em uma única função. Já o método elíptico, as coordenadas x e y são analisadas separadamente.

A posição de cada elipse, que representa um harmônico h , é centrada no ponto de tempo atual, mas na elipse do harmônico anterior $(h - 1)$. Esses pontos centrais viajam ao

² Em que o ponto inicial é igual ao ponto final.

³ Para mais detalhes, sugere-se <https://youtu.be/r6sGWTCMz2k>.

Figura 25 – Comparação entre a Série de Fourier tradicional e a forma Elíptica de Fourier. A representação harmônica exemplifica os métodos utilizando 2 harmônicos.



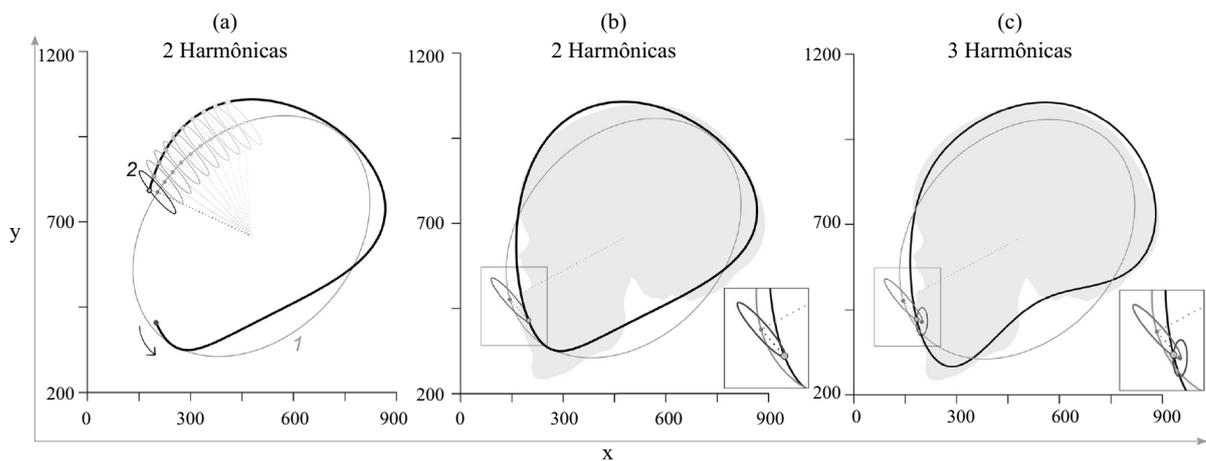
Fonte: Adaptado de Caple *et al.* (2017).

redor da própria elipse atual e se estabelece um vetor a partir do centro do harmônico anterior, denominado fasor. O contorno de uma forma utilizando a AEF é descrito devido ao movimento desse ponto central ao redor da elipse atual, bem como pelo próprio movimento dela ao redor do perímetro da elipse que a precede (CAPLE *et al.*, 2017).

A velocidade que o ponto central percorre a elipse atual é proporcional ao número do harmônico que ela representa. Isso ocorre pois o número do harmônico representa o comprimento de onda, na qual define o número de vezes que o ponto deve percorrer aquela elipse no período total de tempo. Por exemplo, um contorno descrito por dois harmônicos, o ponto viaja pela segunda elipse, referente ao segundo harmônico, duas vezes no tempo que leva para percorrer a elipse referente ao primeiro harmônico (CAPLE *et al.*, 2017).

A Figura 26 exibe a utilização de dois e três harmônicos para descrever o contorno de uma imagem de um crânio. Nessa figura, em (a) é possível observar o movimento de uma elipse sobre a outra (elipse 1 sobre a elipse 2), bem como o movimento do ponto central sobre a elipse atual. Já em (b) e (c), além da representação dos fasores, é possível notar como o contorno resultante, em destaque pelo contorno preto, apresenta mais detalhes quando a quantidade de harmônicos aumenta.

Figura 26 – Representação inversa da Fourier Elíptica utilizando (a) dois, (b) dois e (c) três harmônicos. Os harmônicos são representadas pelas elipses em cinza, os fasores pelas linhas pontilhadas e os contornos resultantes pela linha preta em destaque.



Fonte: Adaptado de Caple *et al.* (2017)

Como é possível observar nas Figuras 25 e 26 a elipse referente ao primeiro harmônico tende a descrever o tamanho geral e o comprimento do objeto. Além disso, de modo análogo a Fourier tradicional, uma quantidade maior de detalhes da forma é capturada com um número crescente de harmônicos (CAPLE *et al.*, 2017).

O resultado da AEF para um contorno qualquer são dois descritores escalares e quatro descritores vetoriais, com o tamanho proporcional à quantidade de harmônicos utilizada. É válido destacar que esses descritores não dependem de suas contrapartes e não se alteram conforme o número de harmônicos aumenta. Desse modo, os dois primeiros harmônicos permanecem os mesmos para uma forma descrita com dois ou mais harmônicos (CAPLE *et al.*, 2017).

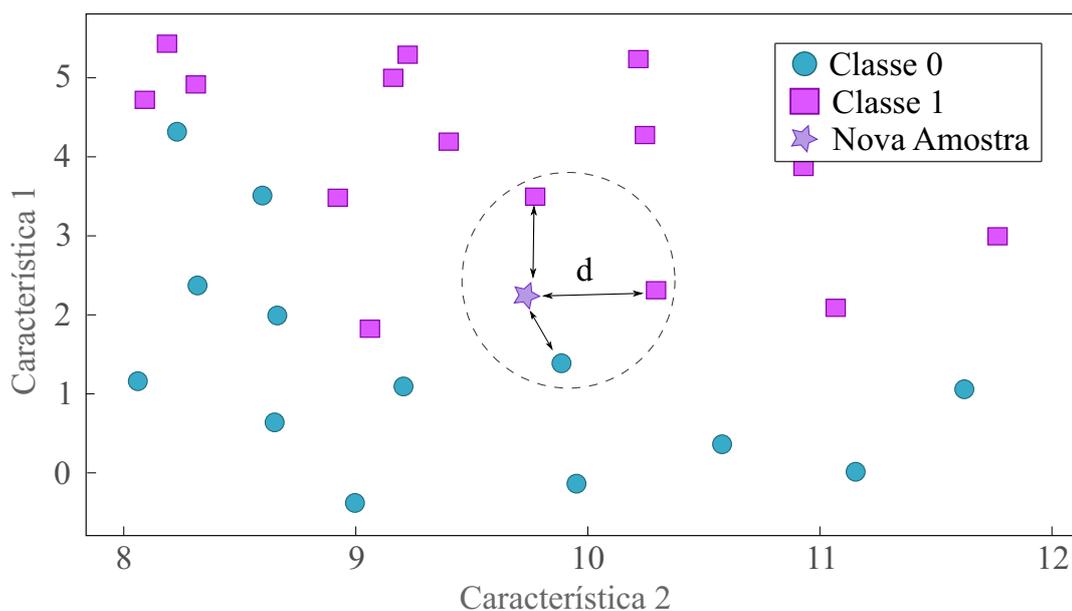
3.3 CLASSIFICADORES

Nesta seção são descritos os métodos de classificação utilizados neste trabalho. Os classificadores escolhidos são métodos de aprendizagem supervisionada e, por esse motivo, necessitam de uma base de dados rotulada para a etapa de treinamento do modelo.

3.3.1 K-Vizinhos Mais Próximos

Um dos algoritmos mais simples de aprendizagem de máquina, o método dos K-Vizinhos mais próximos (KVMP), consiste em classificar uma nova amostra, com seu respectivo valor de características, baseado nas k amostras vizinhas mais próximas, armazenadas na etapa de treinamento (CHERKASSKY; MULIER, 1998). A Figura 27 exemplifica a classificação de uma nova amostra, considerando as três amostras mais próximas, com base na distância euclidiana.

Figura 27 – Exemplo da predição do KVMP considerando $k=3$.



Fonte: Adaptado de Müller e Guido (2016).

Com a Figura 27 é possível perceber que existem dois importantes parâmetros do método: o número de vizinhos k escolhidos e a medida de distância d entre as amostras. No exemplo da figura, a nova amostra foi classificada como classe 1, pois dois dos três vizinhos pertenciam a essa classe, ou seja, os exemplos da classe 1 eram mais frequentes. Caso somente um vizinho fosse considerado, a nova amostra poderia ser considerada como classe 0, dado a distância da amostra com o exemplo dessa classe. Além disso, embora o exemplo da Figura 27 apresente somente duas classes, o método KVMP pode ser utilizado com mais classes. Nesse tipo de aplicação a lógica de funcionamento se mantém a mesma, em que se realiza a contagem de amostras vizinhas por classe e se determina qual a classe mais frequente (MÜLLER; GUIDO, 2016).

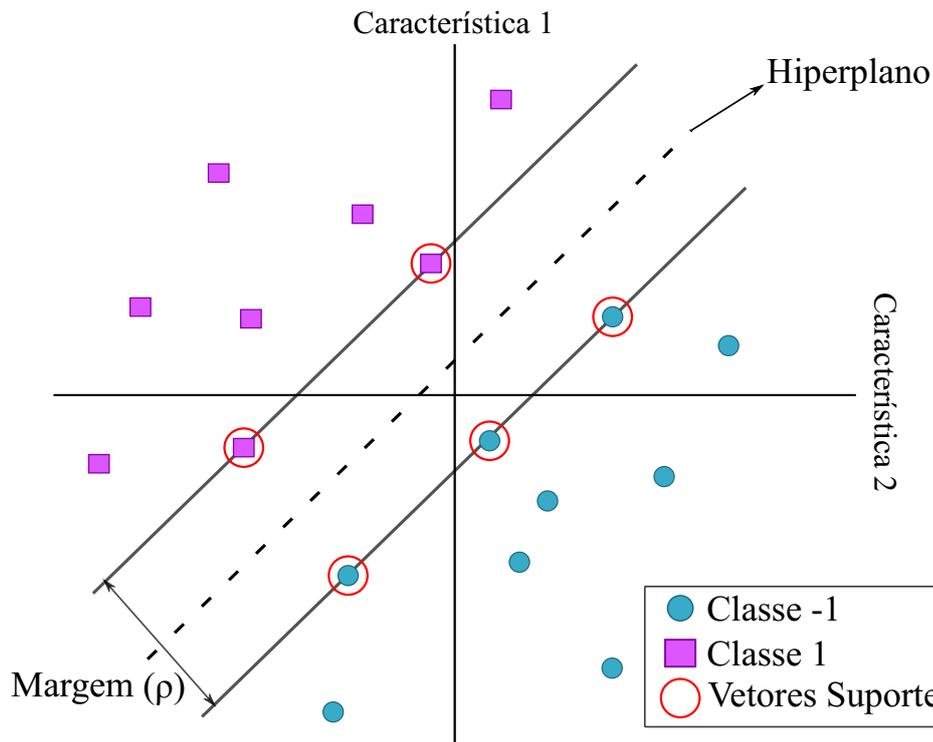
A escolha da quantidade de vizinhos k não é um valor constante para todas as aplicações, pois dependendo da base de dados utilizada a variação desse parâmetro pode gerar um desempenho melhor. Além do parâmetro k , a escolha da métrica de distância também pode influenciar a classificação e pode variar de acordo com a aplicação. Porém, a métrica mais utilizada é a distância Euclideana visto que geralmente funciona bem em diferentes configurações (MÜLLER; GUIDO, 2016).

3.3.2 Máquina de Vetor Suporte

O modelo denominado Máquina de Vetor Suporte consiste em classificar exemplos a partir de um hiperplano ótimo que separa linearmente as classes existentes na aplicação. Esse hiperplano é definido por meio da maximização da margem de separação ρ dada entre as classes do problema com o auxílio de pontos mais próximos da superfície de decisão, denominados como vetores suporte (VAPNIK, 1998). A Figura 28 mostra um exemplo do conceito de hiperplano ótimo, por meio da representação de padrões linearmente separáveis.

O hiperplano de separação, definido pela máxima margem, pode ser obtido por meio de um problema de otimização convexa, o qual pode ser solucionado por meio do método dos multiplicadores de Lagrange (VAPNIK, 1998). O conceito desse hiperplano ótimo de separação também pode ser aplicado para problemas de classificação em que os padrões não são linearmente separáveis. Nesse caso, o modelo inclui um termo para penalizar os erros de classificação para o conjunto de treinamento. Assim, o hiperplano ótimo é encontrado a partir do erro de classificação minimizado. Para equilibrar a maximização da margem e esse termo de penalização, define-se um hiperparâmetro, identificado por C , previamente escolhido pelo usuário (MÜLLER; GUIDO,

Figura 28 – Exemplo do hiperplano ótimo de separação no modelo MVS.



Fonte: Adaptado de (MÜLLER; GUIDO, 2016).

2016).

Outra abordagem para a classificação de conjuntos que não podem ser separados linearmente é a construção de um modelo não-linear que substitui o produto interno da formulação do KVMP por um *kernel* (VAPNIK, 1998). O tipo do *kernel* e seus respectivos parâmetros devem ser definidos previamente.

3.3.3 Análise de Discriminantes Lineares

O método de classificação denominado de Análise de Discriminante Linear (ADL) é uma técnica que realiza a combinação de características referentes a cada grupo (classe) para possibilitar a distinção/separação entre eles. A partir da combinação resultante, pode-se criar um classificador linear que depende apenas das médias que caracterizam cada uma dessas classes. Portanto, esse classificador parte da pressuposição de que as matrizes de covariância dessas classes são idênticas e que as funções densidade de probabilidade delas são normalmente distribuídas (Gaussianas) (BISHOP, 2006).

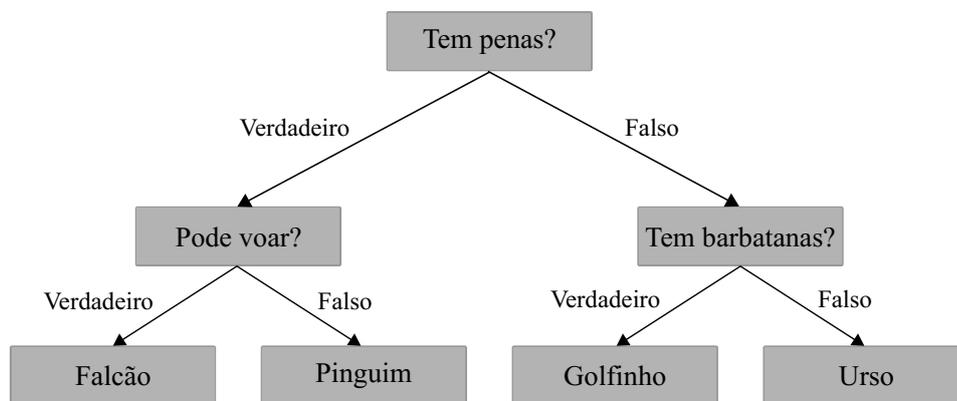
Uma variação do ADL, conhecida como discriminante linear de Fisher, pode ser utilizado com o objetivo de reduzir as características a uma dimensão inferior, projetando as amostras

de treinamento em um espaço que minimiza a distância entre as amostras de uma mesma classe e maximiza a distância de separação entre as classes distintas. Assim, a complexidade computacional e *overfitting*⁴ do modelo podem ser reduzidos, de modo que o classificador se adapta aos dados relevantes e evite se adaptar aos ruídos (CHERKASSKY; MULIER, 1998). Para essa variação do modelo é necessário definir inicialmente o número de dimensões a serem reduzidas.

3.3.4 Árvore de Decisão

O modelo de classificação de Árvore de Decisão aprende por meio de uma hierarquia de perguntas de se/então, capaz de o levar a uma decisão da classe correspondente (MÜLLER; GUIDO, 2016). Essa estrutura testa uma amostra a partir de combinações de perguntas sobre cada característica de um exemplo, nos denominados nós da árvore, até a decisão de qual classe corresponde a amostra, nas folhas (CHERKASSKY; MULIER, 1998). A Figura 29 exemplifica a utilização de uma árvore de decisão para distinguir entre quatro tipos de animais.

Figura 29 – Exemplo de uma Árvore de decisão para distinguir quatro animais.



Fonte: Adaptado de Müller e Guido (2016).

Na Figura 29, tem-se três nós que representam questões a respeito das características do conjunto. Enquanto que as quatro folhas, que contém as respostas, são os rótulos das classes existentes na aplicação.

A etapa de treinamento de um algoritmo de árvore de decisão consiste em aprender a sequência de questões que irão gerar as decisões com o menor número de impurezas⁵ a cada nível da árvore. Isto é, que as decisões sejam capazes de separar os conjuntos de modo mais

⁴ Quando um modelo se ajusta excessivamente ao conjunto de treinamento acarreta em um bom desempenho de classificação desse mesmo conjunto, mas uma alta taxa de erros ao classificar novas amostras (DUDA *et al.*, 2000).

⁵ Mais do que uma classe como resposta em cada possível decisão do nó.

homogêneo possível, com dados somente de uma mesma classe. Esse processo de redução é denominado de minimização da entropia ou ganho de informação (BREIMAN *et al.*, 1984).

Os parâmetros de uma árvore de decisão que controlam a complexidade do modelo são os parâmetros de pré-poda que estabelecem uma interrupção na construção da árvore. Geralmente os critérios tamanho máximo de profundidade da árvore, número máximo de folhas, ou número mínimo de amostras em um nó, são suficientes para prevenir o *overfitting* do modelo (MÜLLER; GUIDO, 2016).

3.3.5 *Ensemble*

O *Ensemble* é um método que reúne múltiplos modelos de aprendizagem de máquina para criar um modelo mais robusto (MÜLLER; GUIDO, 2016). Para classificar novos exemplos, esse método realiza a combinação das decisões individuais dos modelos que fazem parte do agrupamento. Geralmente essa combinação é realizada por meio da média não-ponderada das predições dos modelos, mas também pode ser realizada por média ponderada ou voto majoritário (CHERKASSKY; MULIER, 1998).

O *Ensemble* pode combinar modelos diferentes de classificadores, como uma rede neural, uma árvore de decisão e um KVMP, ou combinar modelos que apresentem estruturas semelhantes, geralmente árvores de decisões, com algoritmos como o *bagging* e o *boosting*. Uma das diferenças do *bagging* e o *boosting* é a ponderação dada a cada modelo que, por consequência, influencia em sua importância na votação final. O método *bagging* atribui um mesmo peso para todas as predições, enquanto o *boosting* atribui um peso a predição do modelo conforme seu desempenho (LANTZ, 2013). A escolha do classificador individual pode ser tratado como um parâmetro a ser definido *a priori* para o *Ensemble*. Visto que o desempenho dos modelos de classificação dependem da escolha das características utilizadas, empregar técnicas que auxiliam a seleção de um conjunto de características pode resultar em um desempenho melhor na etapa de identificação de cargas elétricas.

3.4 SELEÇÃO DAS CARACTERÍSTICAS

No processo de aprendizagem de um modelo de classificação, algumas características podem conter informações redundantes ou gerar um impacto pouco significativo na classificação. Por esse motivo, analisar uma característica, sozinha ou em um subconjunto, pode ser interessante

dado um critério de avaliação pré-definido. O critério mais utilizado é a acurácia da classificação, mas diferentes critérios podem ser utilizados e, dependente da escolha, é possível selecionar as características mais relevantes (LIU; MOTODA, 2007).

O objetivo de um processo de seleção de características não é o de reduzir o número de características a um determinado valor, mas o de encontrar um subconjunto em que o critério avaliativo seja maior ou equivalente do que quando comparado ao uso de todas as características. Por esse motivo, a seleção iterativa é finalizada quando a retirada de características não gera melhorias ou passa a prejudicar o critério avaliativo no modelo (SADEGHIANPOURHAMAMI *et al.*, 2017).

De acordo com Sadeghianpourhamami *et al.* (2017) os métodos de seleção de características podem ser divididos em três categorias: métodos de filtros, métodos *wrapper* e métodos *embedded*. Os métodos de filtros independem de um modelo de aprendizagem de máquina e a seleção de características é realizada apenas com as propriedades intrínsecas dos dados, geralmente calculando uma pontuação de relevância entre características de acordo com algum critério estatístico, por exemplo a correlação entre as características. Esses métodos costumam ser computacionalmente rápidos, porém não tão precisos por serem desacoplados de um modelo. Já nos métodos *wrapper*, subconjuntos de características são utilizados em um modelo específico de classificação ou regressão. Com base no critério de avaliação, as características são adicionadas ou removidas do subconjunto. Assim, o problema é reduzido a um problema de busca heurística e por esse motivo pode ter um custo computacional maior do que os outros modelos. Por fim, os métodos *embedded* tentam combinar as qualidades dos métodos de filtros e *wrapper*, pois esses métodos extraem a importância da característica pelo treinamento de um modelo.

Embora os métodos *wrappers* possam parecer desfavoráveis computacionalmente, visto que encontrar o melhor conjunto de características exigiria uma busca exaustiva, além de ser inviável na maioria dos casos, existem estratégias de busca eficientes que podem ser implementadas (GUYON; ELISSEEFF, 2003). Entre as estratégias para tornar as buscas mais eficientes, encontram-se os métodos de busca sequencial, em que se destacam os métodos de seleção direta e eliminação reversa (AHA; BANKERT, 1995).

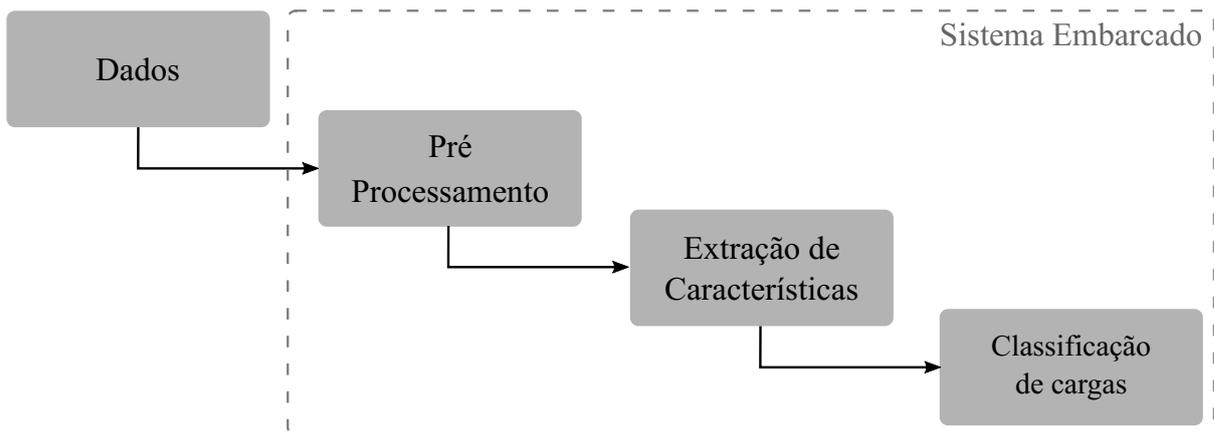
O método de seleção direta - SD começa a avaliação do modelo com um conjunto vazio e progressivamente as características vão sendo adicionadas ao subconjunto. Já o método de eliminação reversa - ER começa com todo o conjunto de característica e elimina gradativamente as características menos promissoras (GUYON; ELISSEEFF, 2003) (LIU; MOTODA, 2007),

usando a acurácia de classificação como referência.

4 METODOLOGIA

Neste capítulo estão descritas as metodologias desenvolvidas neste trabalho, conforme a sequência dada na Figura 30.

Figura 30 – Organização da metodologia.



Fonte: Autoria própria.

A primeira seção descreve as bases de dados, compostas de aquisições de cargas elétricas em uma determinada instalação, utilizadas para treinamento e teste dos modelos de classificação empregados neste trabalho. Na sequência, a seção de pré-processamento descreve as etapas realizadas para preparação e organização das medidas de tensão e corrente de cada aquisição antes da aplicação de um método de extração de características. As próximas duas seções detalham as metodologias utilizadas para extrair características de uma assinatura de carga, a trajetória V-I. A primeira metodologia descreve novas características para a forma obtida das trajetórias V-I, em estado estacionário e transitório, e a segunda metodologia descreve um novo método no contexto de NILM que extrai o próprio formato da trajetória. Já na seção de classificação de cargas é descrito o procedimento utilizado para treinar os modelos de classificação e, posteriormente, testá-los. Por fim, a última seção apresenta o sistema embarcado utilizado para a implementação das etapas anteriores e avaliação de recursos utilizados por cada método de extração desenvolvido nesse trabalho, como tempo computacional e memória.

4.1 BASE DE DADOS

Nesta seção estão descritas as três bases de dados utilizadas neste trabalho para a classificação de cargas elétricas.

4.1.1 COOLL

A base de dados COOLL reúne um conjunto de aquisições de tensão e corrente de algumas cargas em diferentes cenários de uso, com frequência fundamental de 50 Hz, capturadas em uma frequência de amostragem de 100 kHz. Esse conjunto, criado na Université d'Orleans, envolve a aquisição de cargas individuais sendo conectadas/desconectadas à rede, cujo instante de conexão e desconexão são controlados (PICON *et al.*, 2016). A COOLL é composta de 42 aparelhos, com 20 aquisições de cada um deles, conforme apresentado na Tabela 3.

O uso dos aparelhos em diferentes cenários, junto ao fato das 840 aquisições apresentarem o chaveamento controlado, permitem a análise de diferentes parâmetros dos métodos utilizados neste trabalho. Além disso, para diminuir a quantidade de amostras por aquisição da base de dados COOLL, neste trabalho as aquisições foram sub-amostradas de modo que a nova frequência de amostragem fosse de aproximadamente 17 kHz.

4.1.2 Base de dados LIT

A base de dados do Laboratório de Inovação Tecnológica em Sistemas Embarcados e Energia - LIT¹ é composta de três subconjuntos de dados (RENAUX *et al.*, 2020): Natural (LINHARES *et al.*, 2018), Simulado (ANCELMO *et al.*, 2019a) e Sintético (RENAUX *et al.*, 2018). Neste trabalho somente os subconjuntos de dados simulados e sintéticos foram utilizados.

O subconjunto de dados simulados, LIT- SIM, no *Simulink* do MATLAB, contém dados de tensão e corrente, de aparelhos simulados em uma rede de 60 Hz e frequência de amostragem de aproximadamente 15 kHz. Essa base possui sete tipos diferentes de cargas elétricas, em que cada tipo possui quatro níveis de potência distintos, simulando aparelhos diferentes, como mostra a Tabela 4.

Além da comutação de cargas no subconjunto LIT- SIM ser precisamente controlado, a nível de amostra, ele ainda disponibiliza seis variações de cenários da rede CA, em que se tem

¹ Esta base de dados foi desenvolvida no projeto PD2866-0464/2017, tendo contribuições da autora deste documento na organização de códigos para disponibilização pública dos dados.

Tabela 3 – Tipos de aparelhos da base de dados COOLL.

ID	Aparelho	Potência (W)
1	Furadeira Bosch - GSB 20-2 RE	701
2	Furadeira - MAC Allister MERH750	750
3	Furadeira - Parkside PSBM 850	850
4	Furadeira - Carrefour ID13Q-K	500
5	Furadeira - KINZO 25C070	500
6	Furadeira - ECHNI ID31	500
7	Ventilador - Calor C Serie 01123	34
8	Ventilador - STAR FS3-29SW	40
9	Moedor - Bosch GWS 20-230	2000
10	Moedor - Black&Decker KG701	710
11	Secador de Cabelo - Babyliss Expert ioniceramic	2200
12	Secador de Cabelo - Calor Airstyle	1200
13	Secador de Cabelo - Babyliss Expert 2000	1400
14	Secador de Cabelo - Ceriotti Blob	1600
15	Cortador de Grama -Metabo HS 8365 S	500
16	Cortador de Grama - Black&Decker	250
17	Cortador de Grama - Matrix Evolution 680-59	680
18	Lâmpada LED - XANLITE ALEX31-E27	16
19	Lâmpada Fluorescente - Kingfisher	15
20	Lâmpada Halógena - PHILIPS ECO	105
21	Lâmpada Halógena - Auchan M7	100
22	Removedor de Tinta - Black & Decker HG99-H1	1050
23	Platina - AEG HB 750	750
24	Roteador - Performance Power PP	1200
25	Lixadeira - Carrefour CC200ES	200
26	Lixadeira - Bosch PSS	150
27	Lixadeira - Bosch PSS 280A	250
28	Serra - Black&Decker DN 231	500
29	Serra - Brico Dépôt M1Y-YH-210L	1800
30	Serra - Bosch PST 6800	470
31	Serra - AEG SPE 65	450
32	Serra - ECHNI JS31	350
33	Serra - Bosch PST 6800 PE	470
34	Serra - Siplec-Leclerc M1Q	350
35	Serra - Bosch PKS 66	1200
36	Aspirador de Pó - Tornado 61EKJ01	700
37	Aspirador de Pó - Dilem 2406	-
38	Aspirador de Pó - Dirt Devil M2827	2300
39	Aspirador de Pó - Dyson DC 29	1400
40	Aspirador de Pó - Miele Blue Magic	2000
41	Aspirador de Pó - Miele S 511 Super Air Clean	1500
42	Aspirador de Pó - Miele S 511	1500

Fonte: Adaptado de Picon *et al.* (2016).

os cenários:

1. Ideal (LIT- SIM 1);
2. Com indutância equivalente do sistema (LIT- SIM 2);

Tabela 4 – Tipos de aparelhos no subconjunto LIT simulado - LIT-SIM

ID	Aparelho	Potência (W)
1	Motor 1	457,1
2	Motor 2	339,29
3	Motor 3	261,79
4	Motor 4	195,19
5	Ponte Completa 1	15,48
6	Ponte Completa 2	29,92
7	Ponte Completa 3	60,6
8	Ponte Completa 4	99,3
9	Circuito R 1	6775
10	Circuito R 2	3239
11	Circuito R 3	2024
12	Circuito R 4	1620
13	Circuito RL 1	1418
14	Circuito RL 2	782,2
15	Circuito RL 3	531,6
16	Circuito RL 4	209,4
17	Tiristor R 1	440
18	Tiristor R 2	367,4
19	Tiristor R 3	294,22
20	Tiristor R 4	220,72
21	Tiristor RL 1	432,8
22	Tiristor RL 2	842,6
23	Tiristor RL 3	146,8
24	Tiristor RL 4	42,36
25	Diodo1	12200
26	Diodo2	5597
27	Diodo3	2401
28	Diodo4	640,1

Fonte: Adaptado de Ancelmo *et al.* (2019a).

3. Com indutância equivalente do sistema e distorção harmônica total na tensão² de 6,8% (LIT- SIM 3);
4. Com indutância equivalente do sistema, distorção harmônica e uma adição de ruído branco Gaussiano com uma relação sinal-ruído (SNR) de 10 dB (LIT- SIM 4);
5. Com indutância equivalente do sistema, distorção harmônica e uma adição de ruído branco Gaussiano com SNR de 30 dB (LIT- SIM 5);
6. Com indutância equivalente do sistema, distorção harmônica e uma adição de ruído branco Gaussiano com SNR de 60 dB (LIT- SIM 6).

Cada cenário disponível contém 804 aquisições, referentes as cargas individuais além de combinações de cargas duplas, triplas, quádruplas e quádruplas, para as 28 cargas da Tabela 4, com possibilidade de variação de ângulo de disparo em 0°, 45° e 90°.

² Seguindo valores típicos e observados nos limites do módulo 8 do PRODIST (ANEEL, 2020)

Já o subconjunto de dados sintéticos, LIT- SYN, é denominado dessa forma pois se refere as aquisições coletadas de uma bancada onde cargas reais são conectadas, porém o instante de chaveamento das cargas na rede são controladas. Esse conjunto contém 1664 aquisições de formas de onda amostradas em 15360 Hz, com comutações de eventos (Liga/Desliga) precisamente anotados (menores que $5ms$). Essas aquisições de tensão e corrente agregadas da rede CA são monitoradas por períodos de até 40 segundos e os aparelhos utilizados nesse subconjunto são dados na Tabela 5.

Tabela 5 – Tipos de aparelhos no subconjunto LIT sintético - LIT-SYN

ID	Aparelhos	Potência(W)
1	Micro-ondas <i>Standby</i>	4,5
2	Lâmpada LED	6
3	Monitor CRT	10
4	Painel LED	13
5	Exaustor de fumaça	23
6	Monitor LED	26
7	Carregador de celular Asus	38
8	Estação de solda	40
9	Carregador de celular Motorola	-
10	Computador portátil Lenovo	70
11	Ventilador	80
12	Resistor	80
13	Computador portátil Vaio	90
14	Lâmpada Incandescente	100
15	Furadeira vel. 1	165
16	Furadeira vel. 2	350
17	Aquecedor de óleo pot. 1	520
18	Aquecedor de óleo pot. 2	750
19	Micro-ondas On	950
20	Aquecedor de Ar Nilko	1120
21	Secador de cabelo Eleganza - vel. 1	365
22	Secador de cabelo Eleganza - vel. 2	500
23	Secador de cabelo Super 4.0 - vel. 1 - Aquec. 1	660
24	Secador de cabelo Super 4.0 - vel. 1 - Aquec. 2	1120
25	Secador de cabelo Parlux - vel. 1 - Aquec. 1	660
26	Secador de cabelo Parlux - vel. 2 - Aquec. 1	885

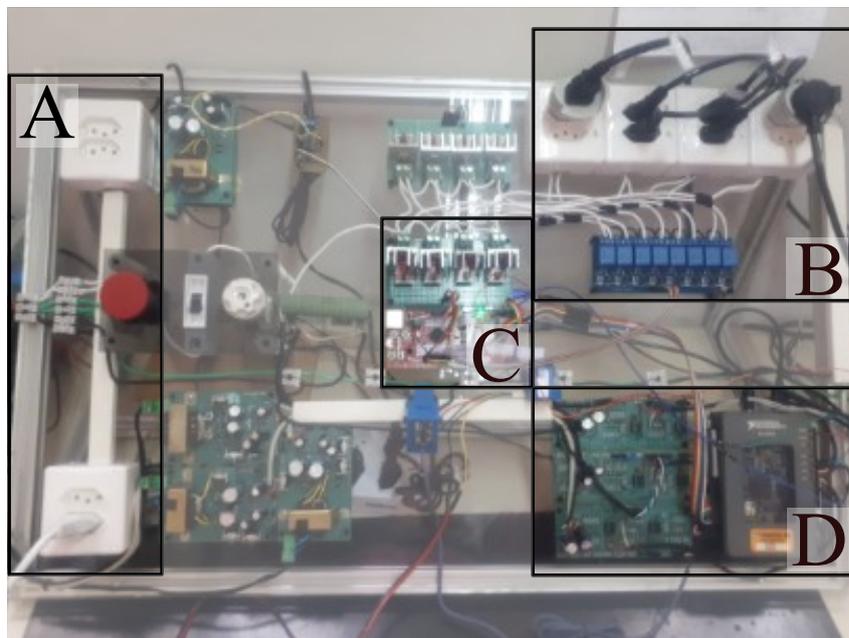
Fonte: Adaptado de Renaux *et al.* (2018).

Do mesmo modo que o subconjunto simulado, o conjunto sintético é composto por aquisições de cargas individuais e de cargas combinadas, dos aparelhos apresentados na Tabela 5, em que podem ser utilizadas todas as aquisições (LIT- SYN) ou com agrupamentos de cargas específicos, com uma carga (LIT-SYN 1), duas (LIT-SYN 2), três (LIT-SYN 3) ou oito (LIT-SYN 8) cargas por aquisição. Além disso, para cada carga ou combinação de cargas são realizadas 16 variações de ângulos de disparo, para conexão das cargas na instalação.

A Figura 31 apresenta a estrutura desenvolvida para realizar as aquisições do subcon-

junto sintético. Na parte A encontram-se a conexão de energia, fonte de alimentação auxiliar e os soquetes para as fontes de alimentação do equipamento. Logo a direita, na parte inferior desse bloco, encontra-se uma placa de proteção do sistema. Já na parte B, tem-se oito módulos de tomadas que permitem o monitoramento de até oito cargas, além de logo abaixo estarem os relés, que realizam a conexão/desconexão dos módulos, controlados pela placa presente no bloco C. Além da placa de controle, o bloco C apresenta os sensores de tensão e corrente. Por fim, o bloco D exibe as placas de condicionamento e módulo incorporado MyRIO que fazem a aquisição das formas obtidas com o monitoramento das cargas conectadas (RENAUX *et al.*, 2020).

Figura 31 – Bancada de aquisição de dados do subconjunto LIT-SYN.



Fonte: Adaptado de Renaux *et al.* (2020).

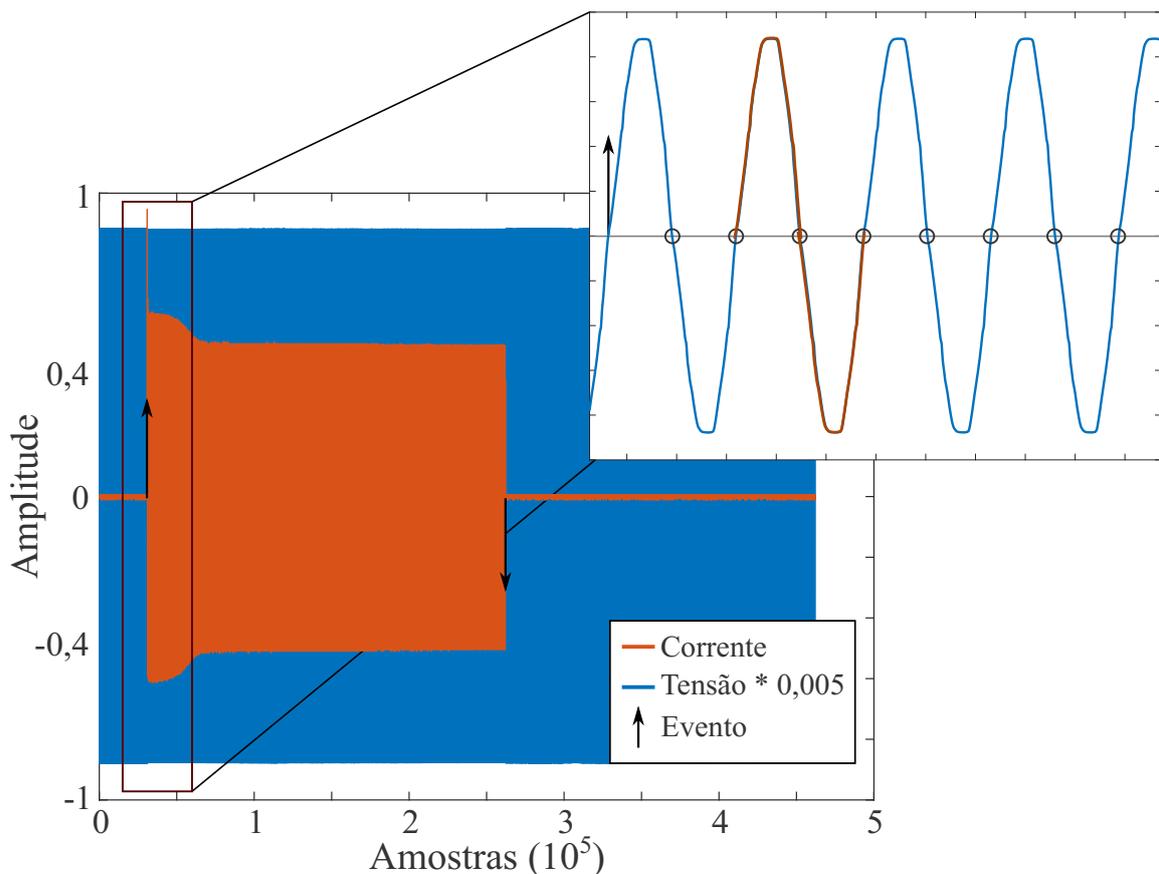
4.2 PRÉ-PROCESSAMENTO

Como todas as bases de dados utilizadas neste trabalho apresentam marcações detalhadas dos instantes nos quais ocorrem os eventos de conexão e desconexão dos aparelhos analisados, as etapas que antecedem a extração de características são baseadas nas amostras identificadas com eventos de acionamento das cargas, e não em uma amostra resultante de um método de detecção, visto que esse trabalho não abrange essa implementação (LAZZARETTI *et al.*, 2020).

O pré-processamento começa com o algoritmo para captura de um ciclo da rede, da tensão e da corrente da carga, no qual primeiramente obtém-se as amostras que apresentam uma

marcação de evento do tipo *ON*, na estrutura fornecida de uma aquisição. A partir dessa amostra de evento, uma janela de amostras equivalente a três ciclos da rede é analisada para determinar os índices em que ocorrem as passagens por zero, tendo a medida de tensão como referência. O primeiro índice identificado como início de um ciclo da medida de tensão, partindo do semiciclo positivo para o semiciclo negativo, é adotado como a primeira amostra do conjunto utilizado para a trajetória V-I. A Figura 32 apresenta um exemplo de como é capturado um ciclo na aquisição.

Figura 32 – Exemplo de como é realizada a captura de um ciclo de rede da tensão e corrente.



Fonte: Autoria própria.

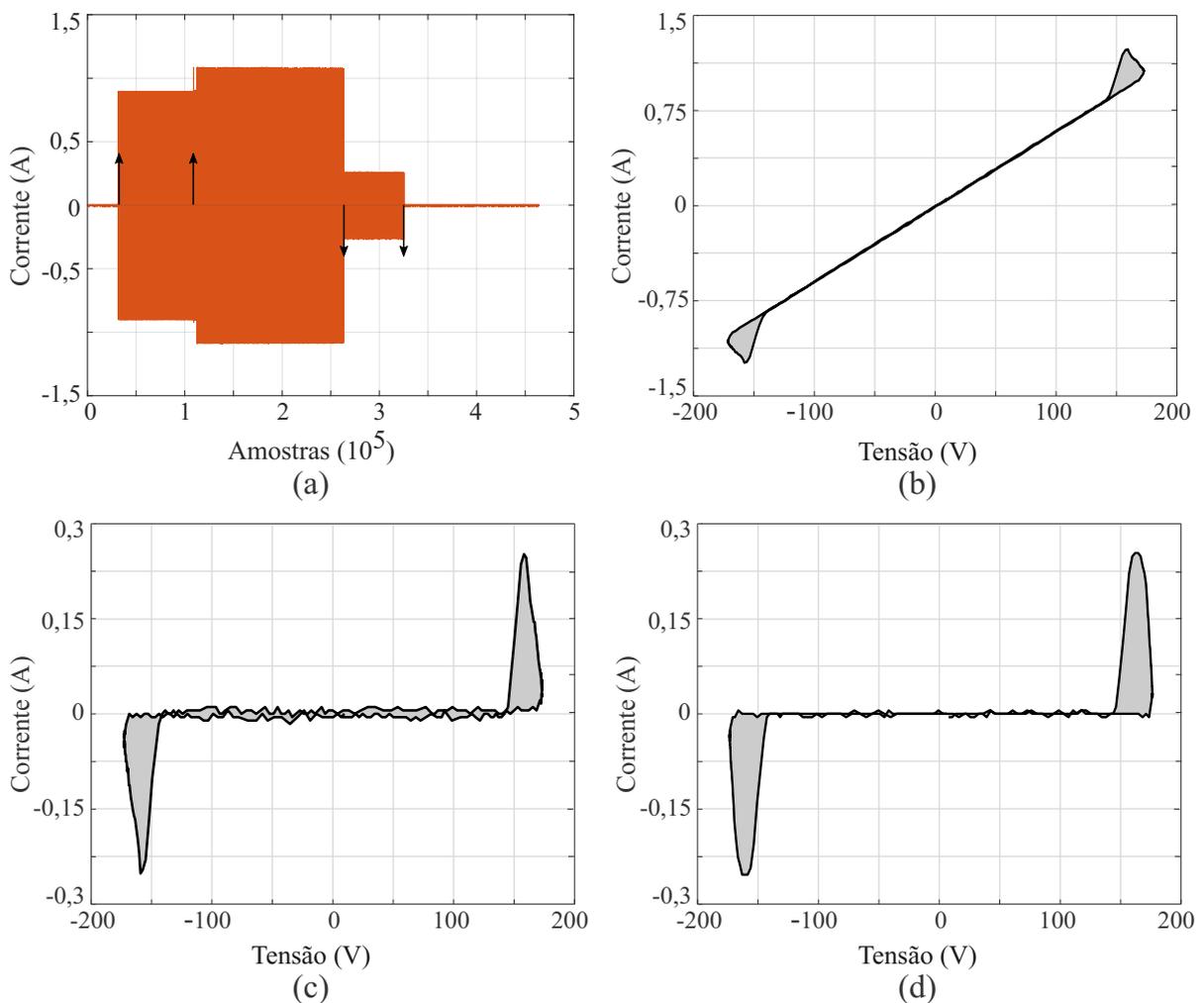
Visto que o conjunto de índices da amostra, de tensão e corrente, são encontrados logo após a conexão de uma carga, a assinatura de carga é obtida no estado transitório. Para obter a assinatura no estado estacionário, o algoritmo equivalente é aplicado, porém a janela de três ciclos começa a ser analisada após um deslocamento de tempo de aproximadamente 60 ciclos³ da frequência fundamental após a amostra identificada como um *ON*.

Após obter as amostras da assinatura, em estado transitório e estacionário, a etapa de desacoplamento é realizada no pré-processamento. Essa etapa ocorre pois uma carga pode ser

³ O valor do deslocamento foi determinado após a análise da base de dados COOLL, em que notou-se que maioria das cargas tornavam-se constantes após essa quantidade de ciclos.

conectada na instalação enquanto outras cargas já estão em funcionamento. Nesse contexto, a trajetória V-I resultante, sem qualquer desacoplamento, tenderia ao formato da carga com a maior potência conectada na instalação. Um exemplo disso é dado na Figura 33(a), em que a lâmpada LED (6 W) é conectada após a carga Resistor (80 W) já encontrar-se ligada. A trajetória V-I obtida para esse segundo evento, sem realizar qualquer desacoplamento, é apresentado em Figura 33(b). Essa trajetória resultante apresenta o formato muito próximo ao formato obtido da carga Resistor e não da carga mais atual conectada, a lâmpada LED.

Figura 33 – Exemplo do processo de subtração: (a) Conexão do Resistor (80 W) e posterior conexão da lâmpada LED (6 W); (b) Trajetória V-I do segundo evento, sem a subtração; (c) Trajetória V-I do segundo evento após a subtração e (d) Trajetória V-I original da carga lâmpada LED.



Fonte: Autoria própria.

O método de desacoplamento é utilizado para obter a forma da trajetória V-I da carga atual, no exemplo a lâmpada LED, mais próxima a forma dessa carga analisada individualmente na instalação, Figura 33(d). Um método simples de desacoplamento é proposto em trabalhos como Baets *et al.* (2018a) e em Wang *et al.* (2018) e consiste em realizar a diferença entre os

sinais de corrente da trajetória antes e depois a ocorrência de um evento. Enquanto que para a medida de tensão, eles propõem realizar a média dos valores de antes e depois a ocorrência do evento. Da forma:

$$i_{traj} = i_{depois} - i_{antes}, \quad (19)$$

$$v_{traj} = \frac{v_{depois} + v_{antes}}{2}. \quad (20)$$

Apesar das trajetórias V-I resultantes do desacoplamento em um contexto multi-cargas, apresentarem pequenas diferenças com as formas das trajetórias em um contexto de aquisições de cargas individuais, devido as propriedades não-lineares, ainda é possível notar características específicas de cada tipo de carga, visto que a parcela linear parece ser a maior, como é possível observar na Figura 33(c).

4.3 MÉTODO TRADICIONAL DA TRAJETÓRIA V-I

Diferente dos trabalhos anteriores na literatura, a metodologia proposta em Mulinari *et al.* (2019) utiliza, além da trajetória V-I obtida em estado estacionário, a trajetória V-I no estado transitório da carga para extrair algumas características. Os autores utilizaram essa metodologia para aumentar o desempenho na etapa de classificação de cargas de uma instalação, visto que o período transitório de uma carga pode conter características discriminantes. Esta seção visa descrever com mais detalhes as novas características propostas em Mulinari *et al.* (2019).

4.3.1 Comprimento

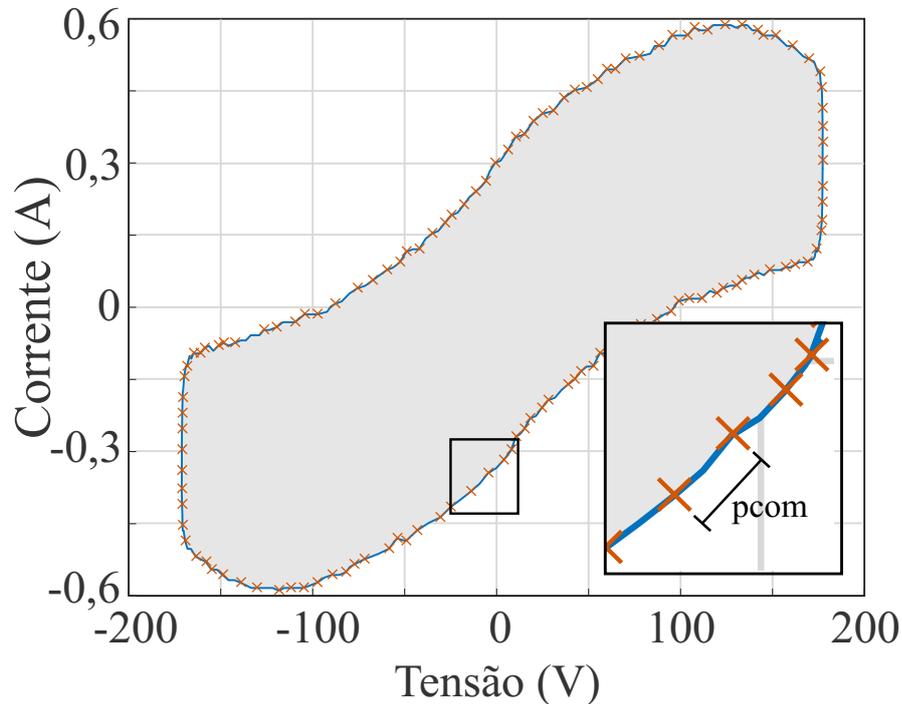
A característica denominada comprimento (*len*) corresponde ao somatório das distâncias euclidianas entre os pontos consecutivos da trajetória V-I em estado estacionário. Expressa por:

$$pcom[i] = \|([V[i] - V[i - 1], I[i] - I[i - 1]])\|, \quad (21)$$

$$len = \sum_{i=2}^{NT-1} pcom[i].$$

Essa característica está relacionada às características de amplitude e frequência do sinal, conforme discutido brevemente em Tkach *et al.* (2010). A Figura 34 exemplifica como é obtido *pcom* da Equação 21 e por consequência como é mensurada a característica *len*.

Figura 34 – Exemplo de como mensurar a característica *len* utilizando a carga: Exaustor de fumaça (23 W) da base de dados LIT-SYN



Fonte: Autoria própria.

4.3.2 Máximo vetor

A característica do máximo vetor (*md*) refere-se a obter o valor do vetor com maior amplitude entre todos os formados por cada ponto da trajetória V-I, em estado estacionário, e a origem (0,0), isto é:

$$md = \max \left(\sqrt{V_i^2 + I_i^2} \right), \quad i \in \{1, 2, \dots, NT\}. \quad (22)$$

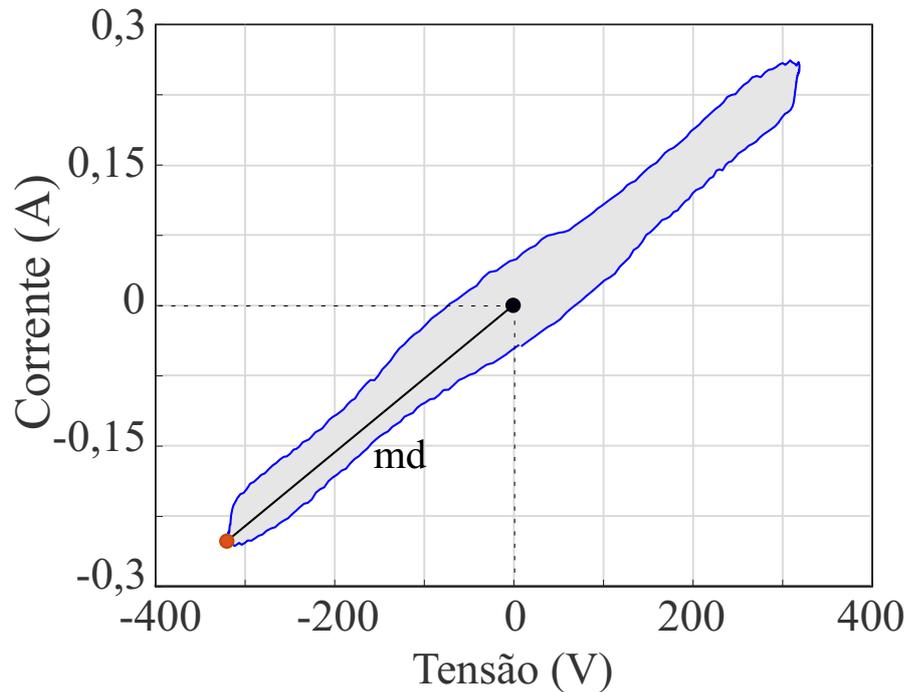
Essa característica está relacionada a maior potência instantânea de uma carga no seu estado estacionário. A Figura 35 apresenta a característica *md* obtida de uma carga da base de dados CO-OLL.

4.3.3 Sobressinal

Como o nome sugere, a característica do sobressinal (*ov*) quantifica o valor aproximado do sobressinal de uma carga, por meio da diferença dos valores máximos das amostras de corrente da trajetória V-I obtida durante o transitório e a trajetória em estado estacionário:

$$ov = \max(I_{tran}) - \max(I_{estac}), \quad (23)$$

Figura 35 – Exemplo da característica md para o Ventilador (34 W) da base de dados COOLL.



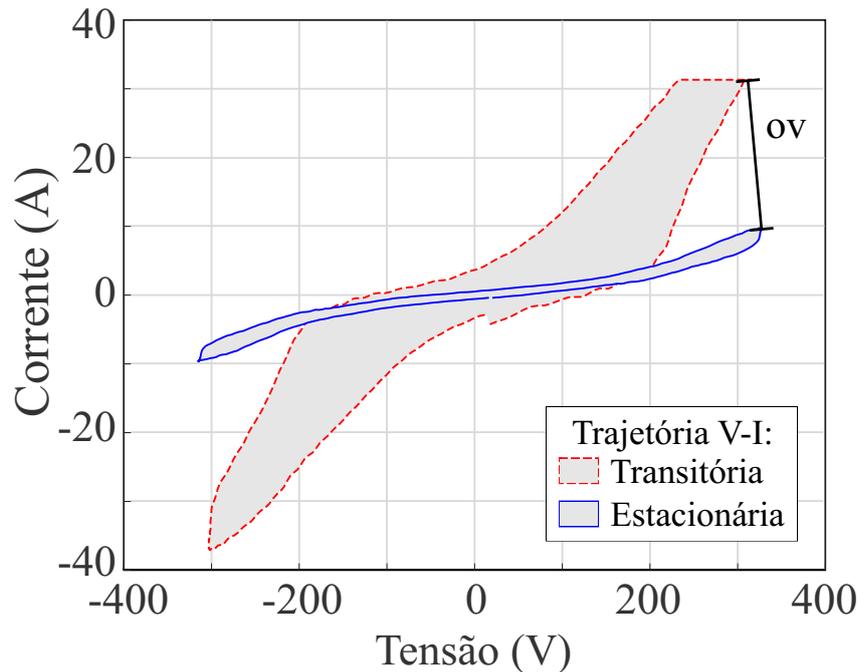
Fonte: Autoria própria.

em que I_{tran} é a medida de corrente da trajetória no período transitório da carga e I_{estac} no estado estacionário. Essa característica procura destacar cargas que apresentam uma resposta transitória menor ou maior que seu valor em estado permanente, como as cargas com corrente de energização elevada. A Figura 36 apresenta como essa característica é obtida.

4.3.4 Variação da inclinação do sinal

A característica da variação da inclinação do sinal (vss) contabiliza a quantidade de alternâncias que ocorrem entre inclinações positivas e negativas, na trajetória V-I estacionária. A inclinação entre dois pontos consecutivos é proporcional à diferença de amplitude entre essas amostras, sendo que o sinal dessa diferença é capaz de indicar se a inclinação é positiva ou negativa. Por esse motivo, realiza-se o produto de duas inclinações consecutivas de modo que o sinal resultante só é negativo se houver variação da inclinação da trajetória. Phinyomark *et al.*

Figura 36 – Exemplo da característica *ov* para o Aspirador de Pó Dyson (1400 W) da base de dados COOLL.



Fonte: Autoria própria.

(2012) definem essa característica por meio de:

$$\begin{aligned}
 amostra[i] &= \sqrt{(V[i] - V[1])^2 + (I[i] - I[2])^2}, \quad i \in \{1, 2, \dots, NT\}, \\
 f[i] &= (amostra[i] - amostra[i - 1])(amostra[i] - amostra[i + 1]), \\
 vss &= \sum_{i=1}^{NT} (f[i] > tv),
 \end{aligned} \tag{24}$$

em que tv é um limiar utilizado para minimizar os efeitos de ruídos. Essa característica pode ser interpretada com uma outra forma de representar informações de frequência dos sinais que compõem a trajetória (HUDGINS *et al.*, 1993).

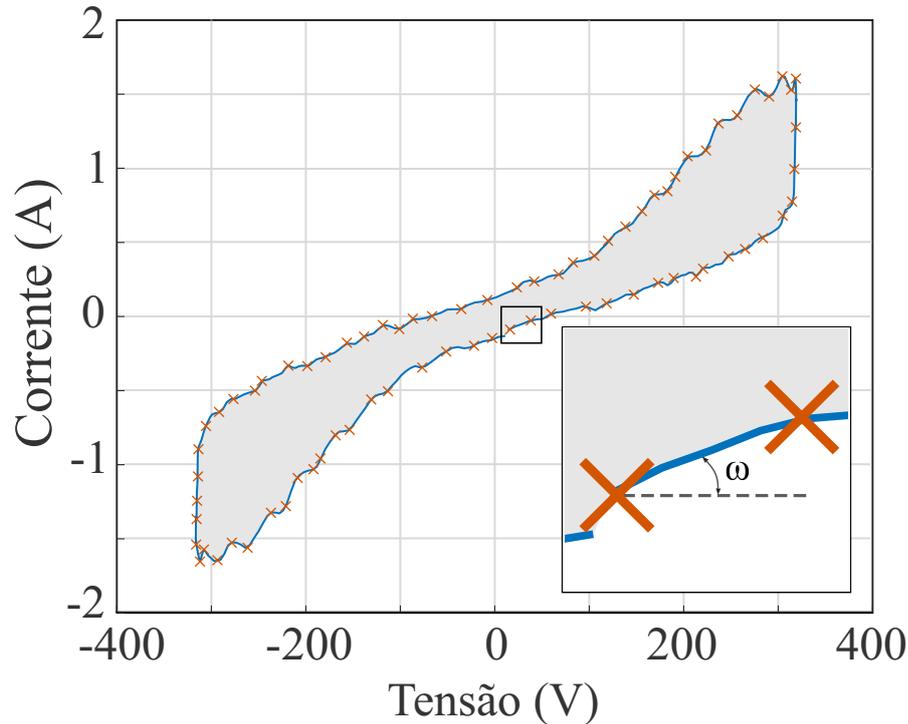
4.3.5 Valor médio da inclinação

O valor médio da inclinação (*asl*) refere-se a característica que mensura a variação média das inclinações entre todos os pontos da trajetória V-I em estado estacionário e pode ser obtida por meio de:

$$\begin{aligned}
 \omega[i] &= \frac{I[i] - I[i - 1]}{V[i] - V[i - 1]}, \quad i \in \{2, \dots, NT\}, \\
 asl &= \frac{\sum_{i=2}^{NT} \omega}{NT}.
 \end{aligned} \tag{25}$$

A Figura 37 exemplifica como ω é obtido entre os pontos da trajetória.

Figura 37 – Exemplo da característica *asl* para a Lixadeira Bosch (150 W) da base de dados COOLL.



Fonte: Autoria própria.

4.3.6 Distância entre os pontos máximo e mínimo

A característica denominada como distância entre os pontos máximo e mínimo (*dpb*) quantifica a distância euclidiana entre o ponto máximo e o ponto mínimo na trajetória V-I, obtida no estado estacionário, da forma:

$$dpb = \sqrt{(V_{imax} - V_{imin})^2 + (I_{imax} - I_{imin}^2)}, \quad (26)$$

em que V_{imax} e I_{imax} referem-se aos valores de tensão e corrente no ponto máximo da trajetória e V_{imin} e I_{imin} referem-se a essas medidas no ponto mínimo. Essa característica, exibida na Figura 38, pode ajudar a diferenciar cargas que apresentem maiores valores de potência com propriedades mais resistivas.

4.3.7 Ângulo entre os pontos máximo e mínimo

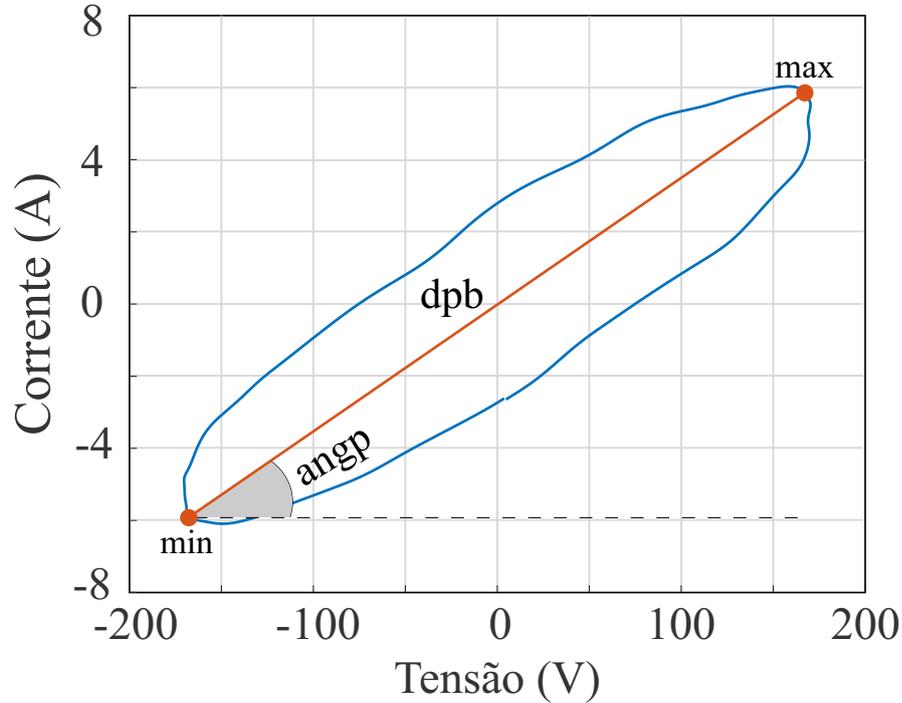
O ângulo entre o ponto máximo e mínimo (*angp*) refere-se a característica que mensura o ângulo formado pelo ponto máximo e o ponto mínimo da trajetória V-I no estado estacionário,

por meio da expressão:

$$angp = \arctan \left(\frac{I_{imax} - I_{imin}}{V_{imax} - V_{imin}} \right). \quad (27)$$

Essa característica é representada na Figura 38.

Figura 38 – Exemplo das características *dpb* e *angp* do Motor de 457,1 W da base de dados LIT-SIM.



Fonte: Autoria própria.

4.3.8 Diferença entre distâncias

Por meio das características *dpb* e *angp*, extraída das trajetórias obtidas em estado estacionário e transitório, é possível determinar as duas características: a diferença entre distâncias e a diferença entre ângulos.

A diferença entre distâncias (*dbd*) mensura a diferença entre as medidas de distância *dpb* dos dois estados das cargas (transitório e estacionário), da forma:

$$dbd = dpb_{tran} - dpb_{estac}. \quad (28)$$

4.3.9 Diferença entre os ângulos

A diferença entre os ângulos ($dbang$) refere-se a característica que quantifica a diferença entre os ângulos $angp$ dos dois estados das cargas (transitório e estacionário), por meio de:

$$dbang = angp_{tran} - angp_{estac}. \quad (29)$$

A intenção em obter dpb e $angp$ é a de encontrar diferentes métricas para distinguir cargas que apresentem variações de potência do período transitório para o estacionário.

4.3.10 Distância do centroide até a origem

As próximas três características (distância do centroide até a origem, ângulo do centroide e diferença entre centroides) dependem dos centroides das trajetórias V-I para mensurar seus valores, como visto na Figura 39, em estado estacionário e transitório. Portanto, antes de expressar essas características é necessário definir o centroide da trajetória. Um polígono regular, definido por um número NT de pontos, pode ser decomposto em um número finito de pequenas partes (GERE; GOODNO, 2009). Desse modo, o centroide desse polígono, localizado em C_x no eixo x e em C_y no eixo y , pode ser calculado utilizando os centroides C_i e as áreas A_i referentes a cada uma dessas partes, da forma:

$$C_x = \frac{\sum_{i=1}^{NT-1} C_{ix} A_i}{\sum_{i=1}^{NT-1} A_i}, \quad C_y = \frac{\sum_{i=1}^{NT-1} C_{iy} A_i}{\sum_{i=1}^{NT-1} A_i}. \quad (30)$$

Dado que as pequenas partes que formam o polígono original sejam triângulos formados por dois pontos consecutivos e a origem, tem-se que seus centroides são dados pelo baricentro, da forma:

$$C_{ix} = \frac{1}{3}(x_i + x_{i+1}), \quad C_{iy} = \frac{1}{3}(y_i + y_{i+1}). \quad (31)$$

A área referente a cada parte pode ser dada por:

$$A_i = \frac{1}{2}(x_i y_{i+1} - x_{i+1} y_i). \quad (32)$$

Dado que A refere-se ao somatório das áreas das pequenas partes, pode-se utilizar as definições dadas nas Equações 31 e 32 nas formas das trajetórias V-I, para obter os centroides,

de modo que:

$$\begin{aligned} C_x &= \frac{1}{6A} \sum_{i=1}^{NT-1} (V[i] - V[i+1])(V[i]I[i+1] - V[i+1]I[i]), \\ C_y &= \frac{1}{6A} \sum_{i=1}^{NT-1} (I[i] - I[i+1])(V[i]I[i+1] - V[i+1]I[i]). \end{aligned} \quad (33)$$

É válido destacar que as Equações 33 e 32 são utilizadas para polígonos sem intersecção (GERE; GOODNO, 2009). Entretanto, de acordo com alguns testes realizados neste trabalho que realizavam a comparação entre os valores de centroide obtido com as equações descritas e os valores obtidos com a função *centroid* do MATLAB, as equações oferecem boas aproximações quando aplicadas a polígonos desse tipo.

A característica denominada de distância do centroide à origem (dc) mensura a distância euclidiana entre o ponto do centroide (C_x, C_y), obtido por meio da Equação 33, para a trajetória em estado estacionário, até o centro do plano (0,0), da forma:

$$dc = \sqrt{C_{xestac}^2 + C_{yestac}^2}. \quad (34)$$

Essa característica pode distinguir cargas que apresentem condução de corrente diferente nos ciclos positivo e negativo.

4.3.11 Ângulo do centroide

O ângulo do centroide ($angc$) obtém o ângulo formado pelo ponto do centroide, da trajetória estacionária, e o ponto de origem do plano, por meio de:

$$angc = \arctan\left(\frac{C_{yestac}^2}{C_{xestac}^2}\right). \quad (35)$$

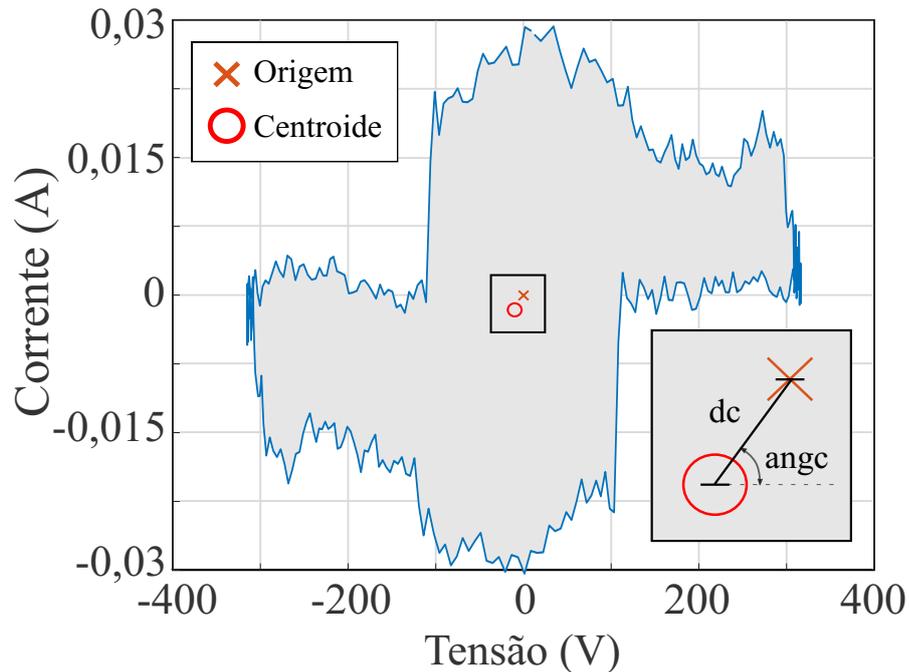
A Figura 39 exemplifica como são mensuradas as características dc e $angc$.

4.3.12 Distância entre centroides

A distância entre centroides (dbc) refere-se a característica que quantifica a distância euclidiana entre o centroide da trajetória V-I obtida em estado transitório da trajetória obtida no estado estacionário, da forma:

$$dbc = \sqrt{(C_{xtran} - C_{xestac})^2 + (C_{ytran} - C_{yestac})^2}. \quad (36)$$

Figura 39 – Exemplo das características *dc* e *angc* da Lâmpada de LED (1,6 W) da base de dados COOLL.



Fonte: Autoria própria.

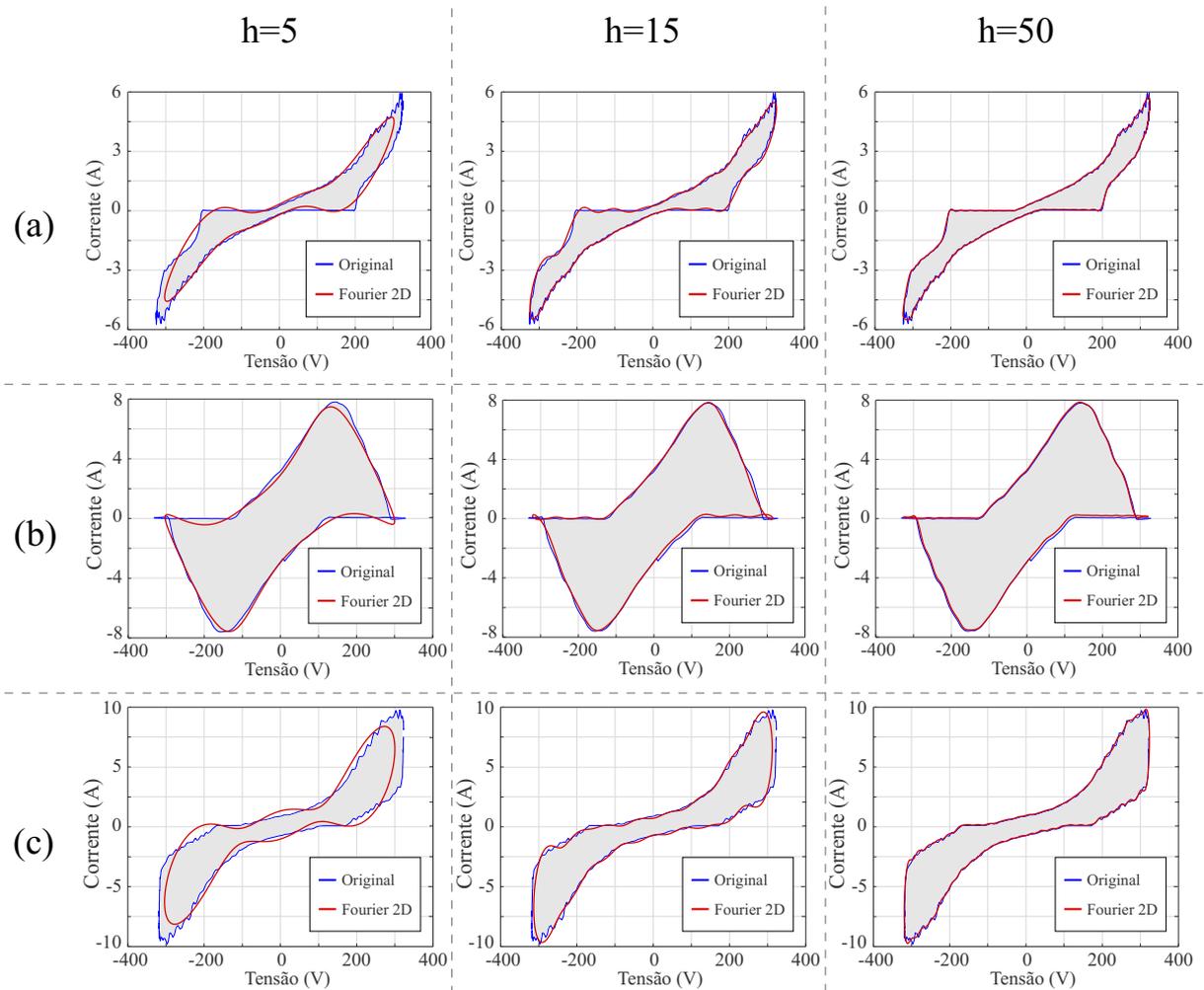
4.4 FOURIER 2D

Neste trabalho a Fourier bidimensional - F2D, descrita na Subseção 3.2, foi aplicada nas trajetórias V-I como um novo método de extração de características no contexto de NILM. Assim, de modo distinto ao método tradicional da trajetória V-I que consiste em encontrar atributos a respeito da forma obtida, a Fourier 2D, por meio de seus descritores, pode descrever a própria forma. Além disso, diferente da aplicação dada em Baets *et al.* (2017) que utilizava a Fourier 2D sobre um contorno obtido de uma imagem da trajetória, este trabalho emprega essa abordagem sobre o contorno descrito pelas próprias medidas de tensão e corrente que compõem a curva V-I, minimizando etapas anteriores a extração de características, como a criação de imagens da curva e algoritmos de detecção de bordas.

A Figura 40 apresenta algumas reconstruções das trajetórias V-I utilizando os descritores de Fourier obtidos considerando a medida de tensão como eixo x, na Equação 17, e medida de corrente como eixo y, na Equação 18.

A partir da Figura 40 é possível notar algumas particularidades discutidas em Caple *et al.* (2017), como a influência da quantidade de harmônicos na reconstrução de um contorno, como também o fato de contornos mais arredondados serem praticamente descritos utilizando poucas harmônicas (menores que 10).

Figura 40 – Reconstrução das trajetórias V-I utilizando a Fourier 2D, com variação das harmônicas de ordem h , para: (a) Furadeira em velocidade máxima (750 W), (b) Aspirador Tornado (700 W) e (c) Aspirador Dilem da base de dados COOLL.



Fonte: Autoria própria.

A ferramenta da F2D nas trajetórias V-I consegue descrever formatos distintos, mesmo com poucos harmônicos, além de dispor de uma base de representação com funções ortonormais de cosseno e seno, resultando em componentes que não apresentam projeção umas sobre as outras, isto é, não apresentam correlação entre si. Essa particularidade é conveniente para os modelos de classificação distinguirem melhor as classes existentes na aplicação (BISHOP, 2006).

Este trabalho também avaliou as influências da utilização da F2D para extrair as características da trajetória V-I em estado transitório, além das características de estado estacionário. Embora o desempenho de classificação aumentasse com a utilização dessa abordagem, o número de características finais era duplicado, o que, por consequência, aumentava o tempo de extração.

Por esse motivo, outras análises ⁴ foram realizadas de modo a manter o menor número possível de características, preservando um bom desempenho (maiores que 90%) de classificação e que permitisse a utilização das trajetórias dos dois estados das cargas.

A metodologia entre as analisadas, que apresentou um desempenho considerável e manteve uma quantidade de características próximas a extração realizada somente sobre a trajetória em estado estacionário, consistiu na criação de uma única trajetória V-I a partir dos valores médios das trajetórias em estado transitório e estacionário, adicionalmente com as informações de potência ativa de cada estado. Por fim, o número de características resultantes ao aplicar essa abordagem é:

$$n_{car} = 4h + 4, \quad (37)$$

em que h representa a quantidade de harmônicos analisados e o valor constante é resultado dos dois descritores escalares obtidos da F2D mais os dois valores de potência ativa dos estados transitório e estacionário.

4.5 CLASSIFICAÇÃO

O processo de classificação utilizando modelos supervisionados de aprendizagem necessitam passar por duas etapas: de treinamento e de teste. Por esse motivo, após extrair todas as características de um determinado banco de dados é necessário dividir esse conjunto em dois subconjuntos, um para o treinamento do modelo e outro para o teste. Neste trabalho, os bancos de dados foram divididos em 80% para treinamento e os outros 20% dos dados para teste.

Antes de serem utilizados por todos os modelos de classificadores os dois subconjuntos, de treino e teste, são normalizados no intervalo de $[-1, 1]$. Vale ressaltar que esta normalização ocorre apenas no domínio das características e não tem relação com as medidas de tensão e corrente utilizadas para criação das assinaturas de carga. Além disso, ela é realizada pois alguns classificadores tem um desempenho melhor quando as características se encontram em um intervalo menor de valores, como o caso do MVS (BISHOP, 2006).

Neste trabalho, durante a etapa de treinamento, o subconjunto de treinamento foi utilizado em uma otimização Bayesiana (SNOEK *et al.*, 2012) para selecionar os parâmetros

⁴ Que avaliavam o desempenho de classificação ao utilizar a trajetória V-I somente em um dos estados, a trajetória obtida pela média dos dois estados, a trajetória obtida pela mediana, uma trajetória resultante (média, mediana) dos dois estados mais características de potência, entre outras.

referentes a cada modelo de classificação, como: (i) o número de vizinhos k no KVMP; (ii) a regularização⁵ e os parâmetros do núcleo do MVS; (iii) os modelos de aprendizagem para o *Ensemble*⁶; (iv) o tamanho máximo da árvore de decisão; (v) e o número de dimensões para o discriminante linear. Ao fim dessa etapa, os modelos são criados com base no conjunto de dados utilizado no treinamento.

Neste trabalho, utilizou-se a otimização Bayesiana disponível na *toolbox* de aprendizagem de máquina e estatística do *software* MATLAB, que utiliza uma abordagem similar a de validação cruzada padrão Cherkassky e Mulier (1998), além de ser uma abordagem menos custosa computacionalmente conforme apresentado em Snoek *et al.* (2012).

Na etapa de teste, os modelos construídos são avaliados a partir do seu desempenho em classificar o subconjunto de teste. A avaliação consiste em obter a acurácia da classificação para cada classe (ACC_l), por meio de:

$$ACC_l = \frac{NECC_l}{NTE_l}, \quad (38)$$

em que $NECC_l$ se refere ao número de classificações corretas para a classe l e NTE_l o número total de exemplos existentes no conjunto de teste dessa classe. Além de avaliar a acurácia média correspondente para todas as classes (ACC), por meio da expressão:

$$ACC = \left(\frac{1}{N_{cargas}} \right) \sum_{i=1}^{N_{cargas}} ACC_l, \quad (39)$$

em que N_{cargas} é o número de aparelhos existentes em uma base de dados, que, por consequência, é o número de classes existentes na aplicação.

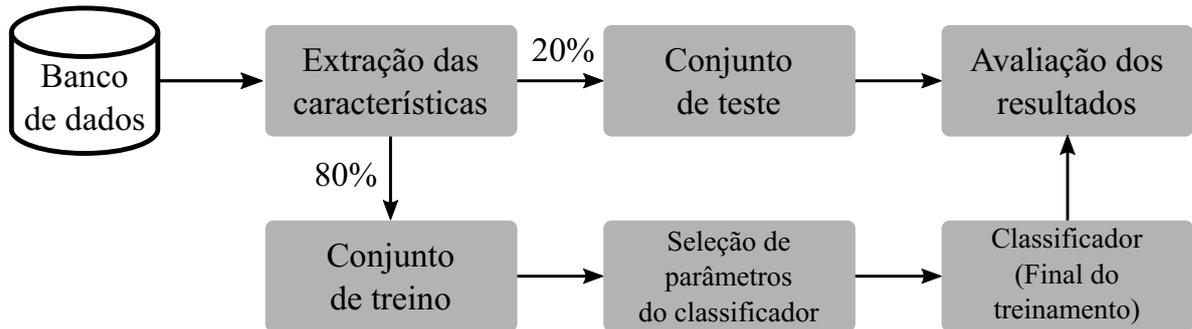
A Figura 41 resume as principais etapas do processo de classificação.

Para todos os testes de classificação realizados neste trabalho, foram utilizados 10 conjuntos diferentes de teste e de treino, para cada método de extração de características com cada uma das bases de dados utilizadas. Esse procedimento foi realizado para que os resultados apresentados sejam dados por uma média e um desvio padrão. Também, de modo a realizar uma comparação mais íntegra dos métodos de extração, a mesma ordem das cargas foram adotadas (mesmos Ids), para a composição desses 10 conjuntos de treino e teste, para cada um dos métodos analisados, como exemplifica a Figura 42.

⁵ Ajuste do valor do parâmetro C, discutido na Subseção 3.3.2

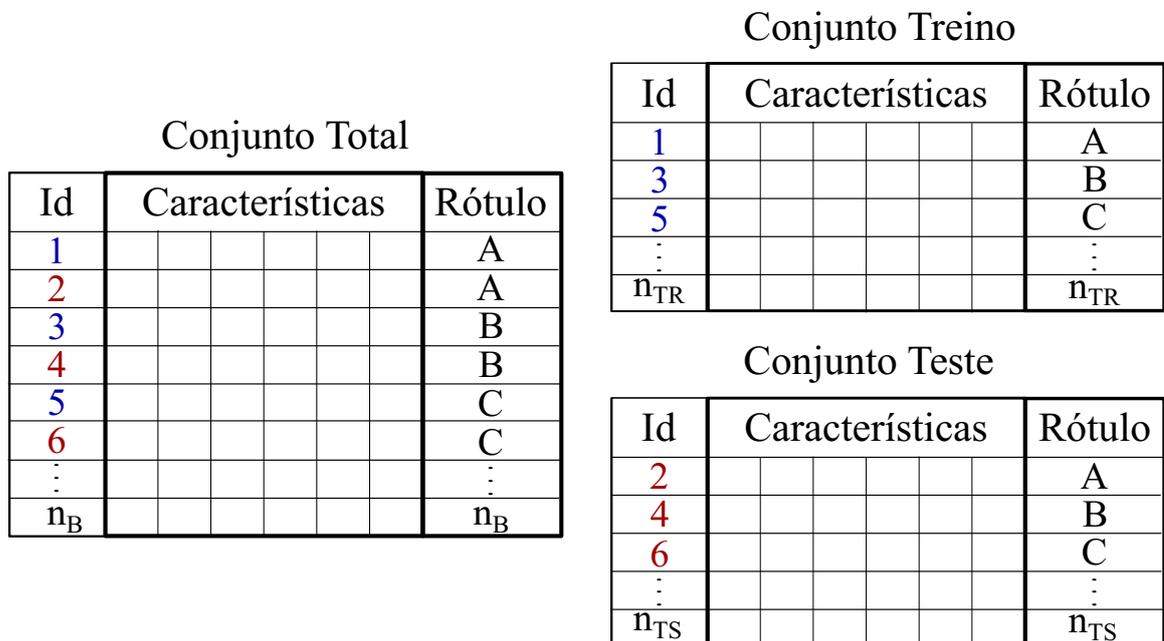
⁶ Entre os modelos de aprendizagem do Ensemble tem-se *AdaBoostM1*, *AdaBoostM2*, *Bag (Random Forest)*, *GentleBoost*, *LogitBoost*, *LPBoost*, *LSBoost*, *RobustBoost*, *RUSBoost*, *Subspace* e *TotalBoost* (SNOEK *et al.*, 2012; BISHOP, 2006).

Figura 41 – Visão geral do processo de classificação.



Fonte: Autoria própria.

Figura 42 – Divisão do Conjunto Total de Características de um Método de Extração por meio dos Ids.



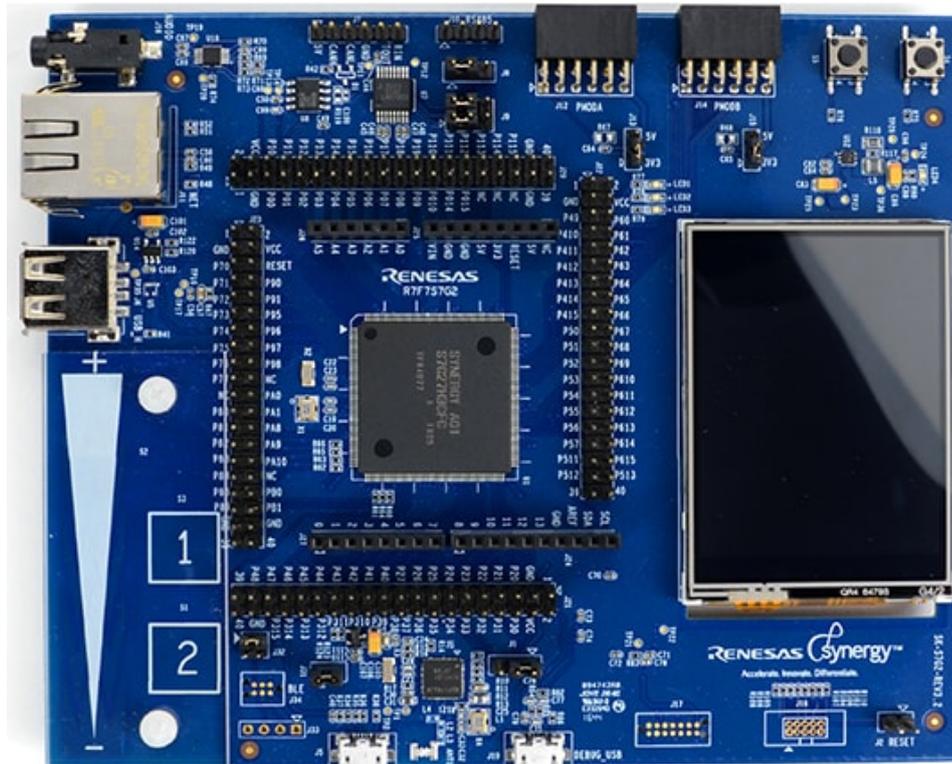
Fonte: Autoria própria.

4.6 SISTEMA EMBARCADO

Neste trabalho o sistema embarcado utilizado, para medições de tempo e memória, é de um dos *kits* de desenvolvimento da Renesas, denominado SK-S7G2 *Starter Kit*, exibido na Figura 43.

O *Starter Kit* apresenta um microcontrolador da família S7G2, um processador ARM Cortex-M4 com 4 MB de *flash* interna, SRAM de 640 KB e frequência de operação de 240 MHz. Os processadores Cortex-M4 contam com unidades de ponto flutuante e instruções de processamento de sinal digital, além de apresentar benefícios como baixo consumo de energia e baixo custo (Renesas Electronics, 2015)

Figura 43 – SK-S7G2 Starter Kit.



Fonte: Renesas Electronics (2015).

Para realizar os testes de tempo computacional no sistema embarcado para cada uma das metodologias propostas, foram realizadas dez execuções consecutivas mensurando o tempo que a metodologia levava para criar sua assinatura de carga de referência e extrair as características, por meio de um temporizador. Para criar as assinaturas os métodos que utilizam as trajetórias V-I, armazenaram e manipularam as amostras equivalente a um ciclo da frequência fundamental para a curva em estado transitório e equivalente a um ciclo para estado estacionário. Enquanto que para a transformada Wavelet, o método armazenou e manipulou as amostras equivalente a quatro ciclos em cada estado.

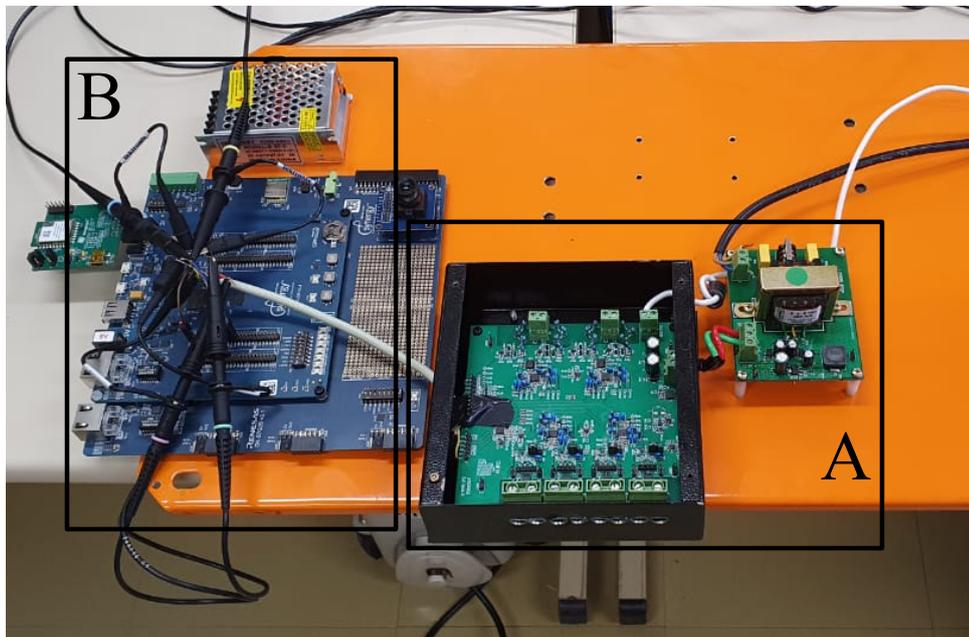
As amostras de tensão e corrente utilizadas por cada método foram salvas no próprio sistema embarcado e são decorrentes das cinco mil primeiras amostras de uma aquisição da base de dados LIT-SYN. Desse modo, foi possível validar se o sistema embarcado obtinha os mesmos resultados que as implementações realizadas no *software* MATLAB.

Já para realizar as medidas de memória utilizada por cada metodologia, utilizou-se o estimador do ambiente de programação e2 Studio que compõe a plataforma *Synergy*, da própria fabricante Renesas. Esse estimador mensura a memória aproximada necessária para armazenar todas as estruturas de uma aplicação. Por esse motivo, é válido ressaltar que em todos

os resultados obtidos de memória neste trabalho, além da alocação de memória do método de extração, considera-se a alocação dos vetores de tensão e corrente (com cinco mil amostras cada) para três canais (referente a cada fase em um sistema trifásico).

A escolha do *kit* SK-S7G2 também se deu ao fato dele pertencer a mesma família do sistema utilizado no projeto PD2866-0464/2017, em que os métodos de extração aqui propostos são parte da solução embarcada que compõe um sistema multiagentes para classificação de cargas elétricas, como exposto em Lazzaretti *et al.* (2020). A Figura 44 exibe o sistema final, que já se encontra em funcionamento. Na figura, o bloco A exibe os conversores A/D e circuitos de condicionamento para as medidas de tensão e corrente, já o bloco B apresenta o *kit* da Renesas utilizado no projeto.

Figura 44 – Montagem do sistema NILM com a metodologia multiagente.



Fonte: Autoria própria.

5 RESULTADOS E DISCUSSÃO

Neste capítulo são apresentadas as análises exploratórias sobre as novas características propostas e os resultados de classificação, empregando as metodologias descritas neste trabalho. De modo geral, os experimentos realizados bem como seus objetivos são resumidos a seguir:

- Análise das novas características propostas na metodologia tradicional das trajetórias V-I (Subseção 5.1.1): tem como objetivo discutir quais tipos de cargas cada característica pode distinguir, além de avaliar a existência de uma correlação linear entre as características;
- Classificação de cargas com a metodologia tradicional (Subseção 5.1.2): apresenta a análise da acurácia de classificação para conjuntos diferentes de características, além de avaliar a utilização dos recursos de tempo computacional e memória no sistema embarcado para a extração de cada um desses conjuntos;
- Classificação de cargas e análises do método da Fourier 2D (Subseção 5.2): tem como objetivo avaliar os resultados obtidos na acurácia de classificação com a metodologia da F2D, além dos parâmetros que podem influenciar nesse desempenho. Também, apresenta os resultados de utilização dos recursos no sistema embarcado;
- Comparativo das metodologias propostas com um método popular na literatura (Subseção 5.3): tem como objetivo avaliar o desempenho das metodologias propostas com outra metodologia por meio da avaliação das acurácias de classificação, utilização de recursos do sistema embarcado e robustez a ruído.

5.1 MÉTODO TRADICIONAL DAS TRAJETÓRIAS V-I

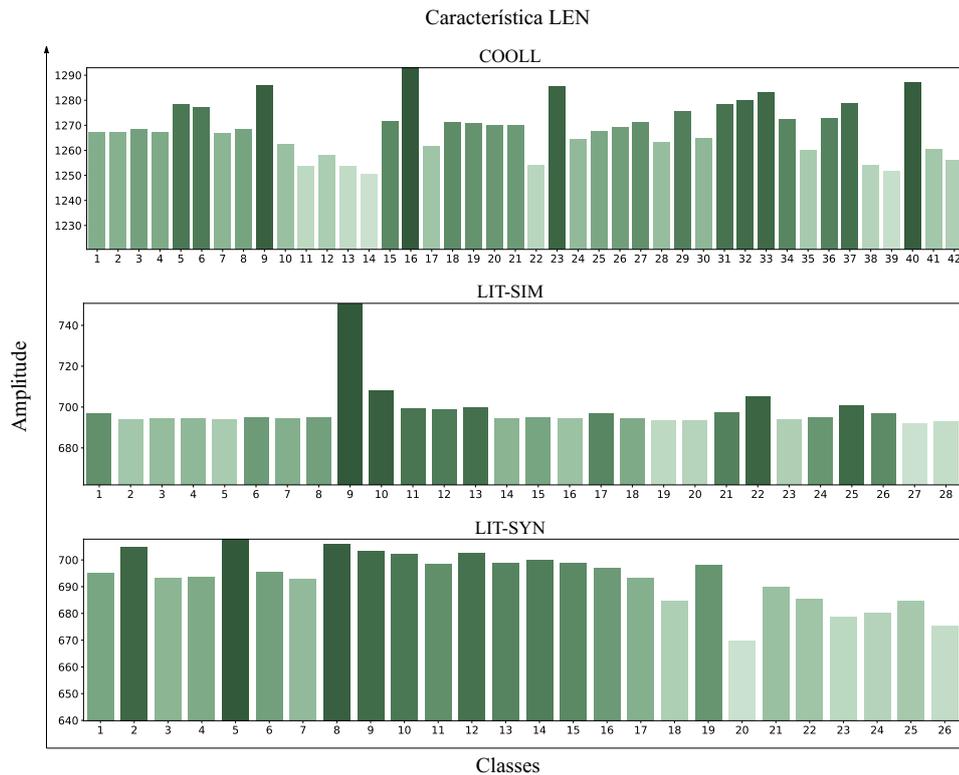
Esta seção do trabalho aborda a metodologia tradicional de extração de características sobre a trajetória V-I e encontra-se dividida em duas subseções. A primeira parte mostra as análises realizadas sobre as novas características propostas em Mulinari *et al.* (2019) e uma análise exploratória por meio da correlação linear das características propostas com as características já existentes na metodologia tradicional. Já a segunda parte exhibe os resultados obtidos empregando essas características, juntamente com as existentes no estado da arte da metodologia, na classificação de cargas elétricas.

5.1.1 Análise das novas características

Uma análise foi realizada com cada característica, descritas na seção 4.3, para observar os comportamentos nas bases de dados utilizadas e encontrar relações com os tipos de cargas que cada uma pode destacar, relacionando os números das classes conforme as cargas dadas nas Tabelas 3, 4 e 5. Para possibilitar a visualização da amplitude das medianas de cada característica os gráficos de barras foram utilizados. Além disso aplicou-se a diferenciação entre as tonalidades das barras para auxiliar a distinguir os valores da característica analisada, em que tonalidades mais escuras representam valores mais elevados, enquanto os valores menores são representados pelas barras mais claras.

- *len* (comprimento): A Figura 45 mostra os valores obtidos para cada classe nas três bases utilizadas. Para a base de dados COOLL os menores valores da característica são obtidos para as classes 14 e 39, e os maiores valores para a classe 16. Para a base de dados LIT-SIM, a classe que parece mais se distinguir das demais é a carga 9. Já para a base de dados LIT-SYN as classes mais distintas das demais ao avaliar essa característica são a 5, com o maior valor, e a 20, com o menor valor.
- *md* (máximo vetor): A Figura 46 exhibe os valores para cada classe dessa característica, nas três bases de dados utilizadas. Essa característica parece diferenciar as cargas 14 e 41 na base de dados COOLL por seus valores menores e a carga 16 por seu valor superior das demais. Na base de dados LIT-SIM as cargas 7 e 27 apresentam os menores valores da característica, enquanto as cargas 9, 10 e 22 os valores maiores. Já na base de dados LIT-SYN a carga 20 aparece com o menor valor dentre as classes para a característica e as cargas 2, 5 e 8 com os maiores valores.
- *dpb* (distância entre os pontos máximo e mínimo): A Figura 47 exhibe as amplitudes encontradas para essa característica nas três bases de dados utilizadas. Para esse atributo é possível notar o mesmo padrão que *md*, Figura 46, embora as amplitudes sejam distintas. Analisando as cargas que apresentaram valores em destaque, parece que essa característica também pode diferenciar cargas com maiores potências e com propriedades mais resistivas das demais. Na base de dados COOLL e LIT-SYN as cargas com as potências maiores tendem a exibir os valores mais baixos da característica, como as classes 14 e 41 na

Figura 45 – Valores medianos da característica *len* para cada classe.



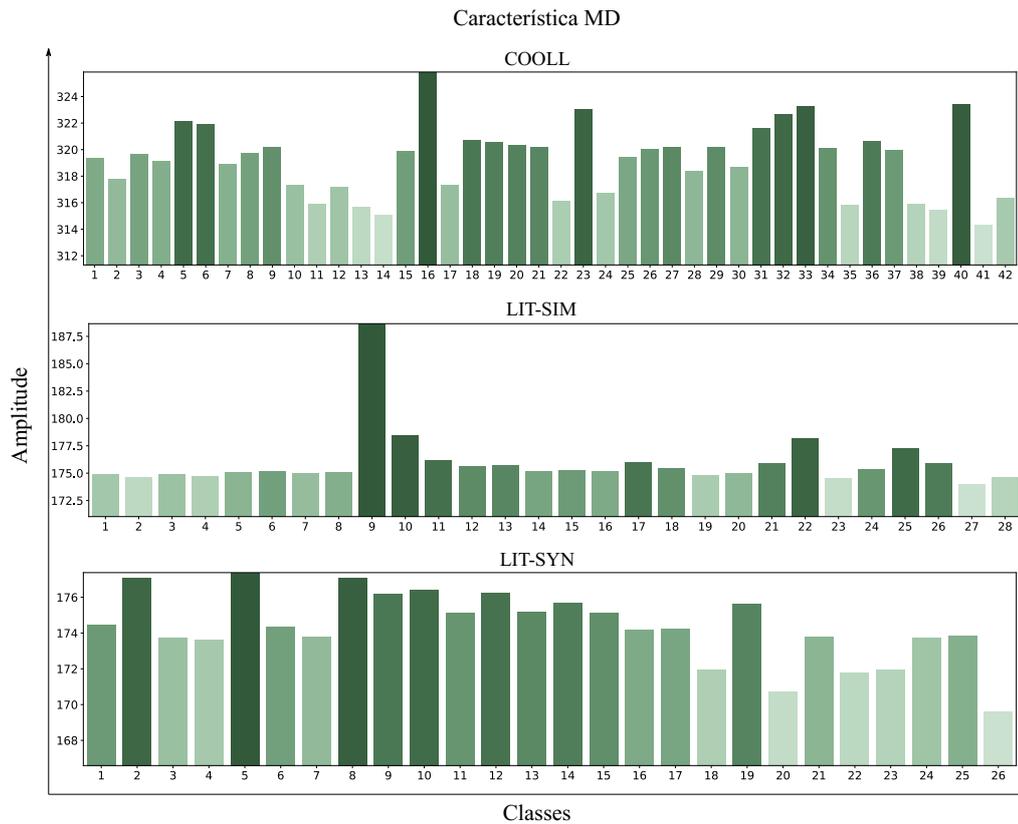
Fonte: Autoria própria.

COOLL e a classe 20 na LIT- SYN. Isso pode ocorrer pois essas bases apresentam uma impedância de entrada maior e quando essas cargas são conectadas a rede causam um 'afundamento' de tensão¹. Já para a base de dados LIT- SIM 5 o comportamento é contrário, pois essa base apresenta um impedância menor do que as outras duas bases e não ocorre o 'afundamento' de tensão.

- *angp* (ângulo entre o ponto máximo e mínimo): A Figura 48 exibe as amplitudes dos valores obtidos para essa característica para as três bases de dados. De acordo com a figura, a característica apresenta para a base de dados COOLL os menores valores para as classes 7, 8 e 18 e os maiores valores para as classes 29 e 40. Para a base de dados LIT-SIM, as classes com menores valores são a 5, 6 e 24 e com maiores valores 9 e 10. Já para a base de dados LIT-SYN os menores valores aparecem para as classes 1, 7, 13 e 19 e os maiores valores para a 20 e 26. De acordo com as classes indicadas pelas três bases, parece que a

¹ De acordo com o módulo 8 do Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional - PRODIST, o termo afundamento de tensão (do inglês *Voltage sag*) trata-se de um dos fenômenos que causa variação de tensão de curta duração (ANEEL, 2020)

Figura 46 – Valores medianos da característica *md* para cada classe.

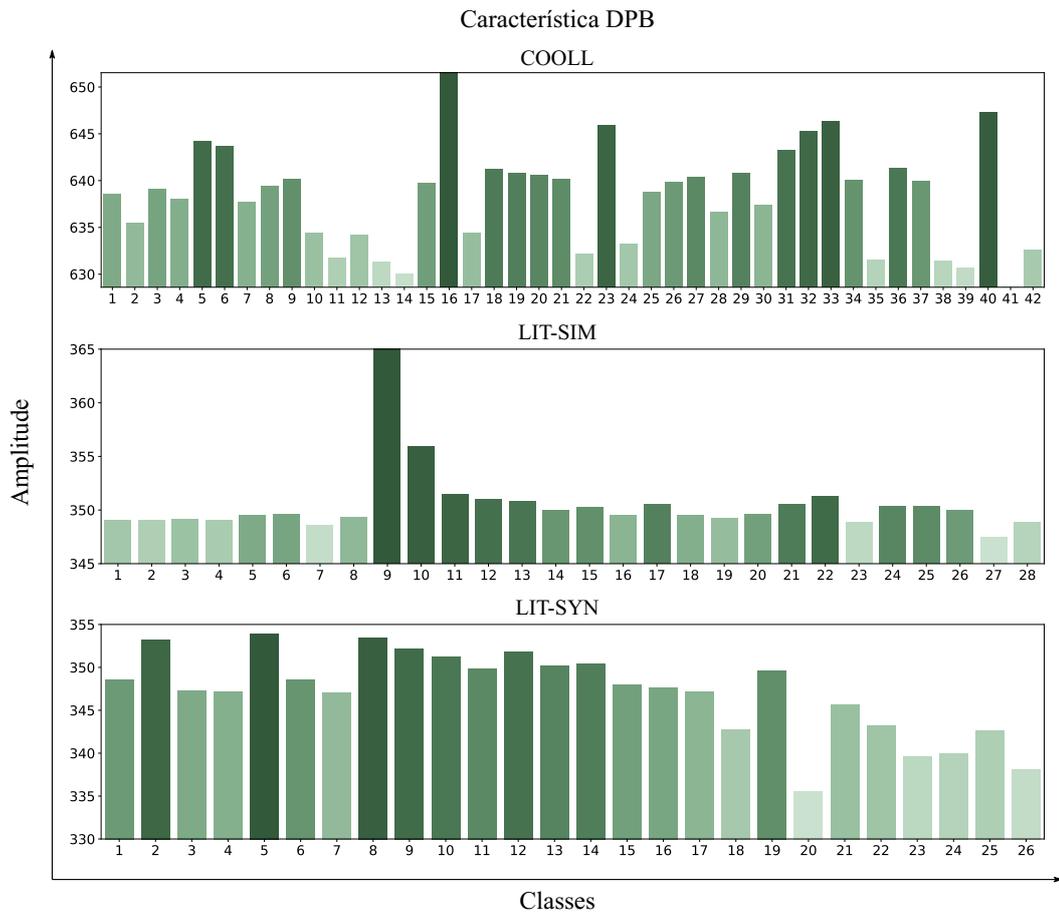


Fonte: Autoria própria.

característica apresenta valores maiores para cargas com propriedades mais resistentes e com grande potência e valores menores para cargas eletrônicas.

- *vss* (variação da inclinação do sinal): A Figura 49 mostra os valores para cada classe dessa característica para as três bases. Para base de dados COOLL a classe que apresenta maior destaque entre as outras é a 9, com um valor elevado para a característica. Já na base LIT-SIM a classe em destaque é a 26, também por seu maior valor comparado as demais classes. Na base de dados LIT-SYN a classe que se diferencia das outras é a classe 15, por seu valor elevado e as classes 6, 21, 22 e 24 por valores inferiores.
- *ov* (sobressinal): A Figura 50 apresenta os valores para cada classe dessa característica. Na base de dados COOLL, a classe que apresenta o maior valor da característica é a 35 e a classe 40 com menor valor. Para a base LIT-SIM é possível notar que as quatro primeiras classes são diferenciadas das demais pelos seus valores maiores. Já na base LIT-SYN, as classes 15 e 16 apresentam os maiores valores para característica, enquanto a classe 10

Figura 47 – Valores medianos da característica *dpb* para cada classe.

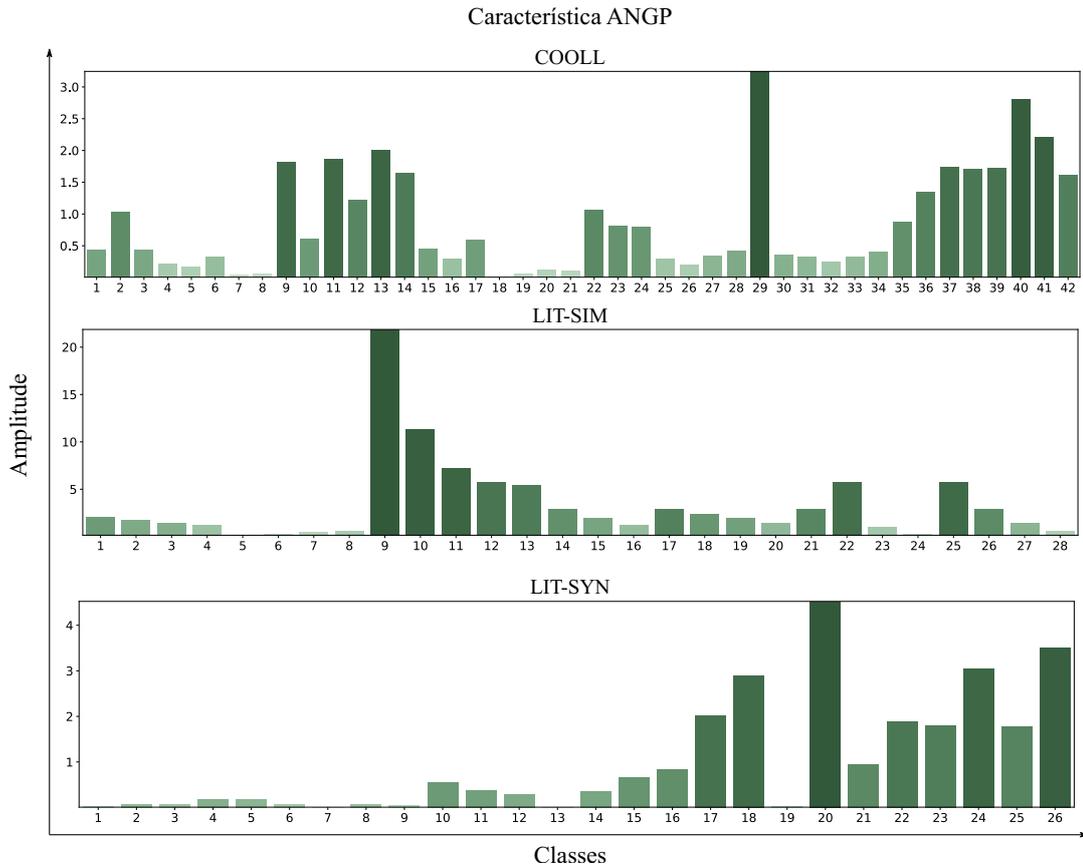


Fonte: Autoria própria.

o menor valor. Ainda por meio dessa figura, pode-se notar que essa característica pode apresentar valores positivos e negativos, pois algumas cargas podem apresentar no período transitório valores mais elevados ou valores inferiores as medidas em regime estacionário.

- *dbang* (diferença entre ângulos): A Figura 51 exibe os valores dessa característica para cada classe das bases analisadas. A partir da figura é possível notar o mesmo padrão dessa característica com a característica do sobressinal (*ov*), Figura 50.
- *dc* (distância do centroide): A Figura 52 mostra os valores dessa característica para cada classe para as três bases analisadas. De acordo com a figura, para a base de dados COOLL a classe que mais se destaca entre as demais é a 5, por apresentar um alto valor da característica. Essa classe apresenta um valor elevado da característica, pois é a única carga da base a apresentar uma forma de onda de corrente totalmente assimétrica do ciclo

Figura 48 – Valores medianos da característica *angp* para cada classe.

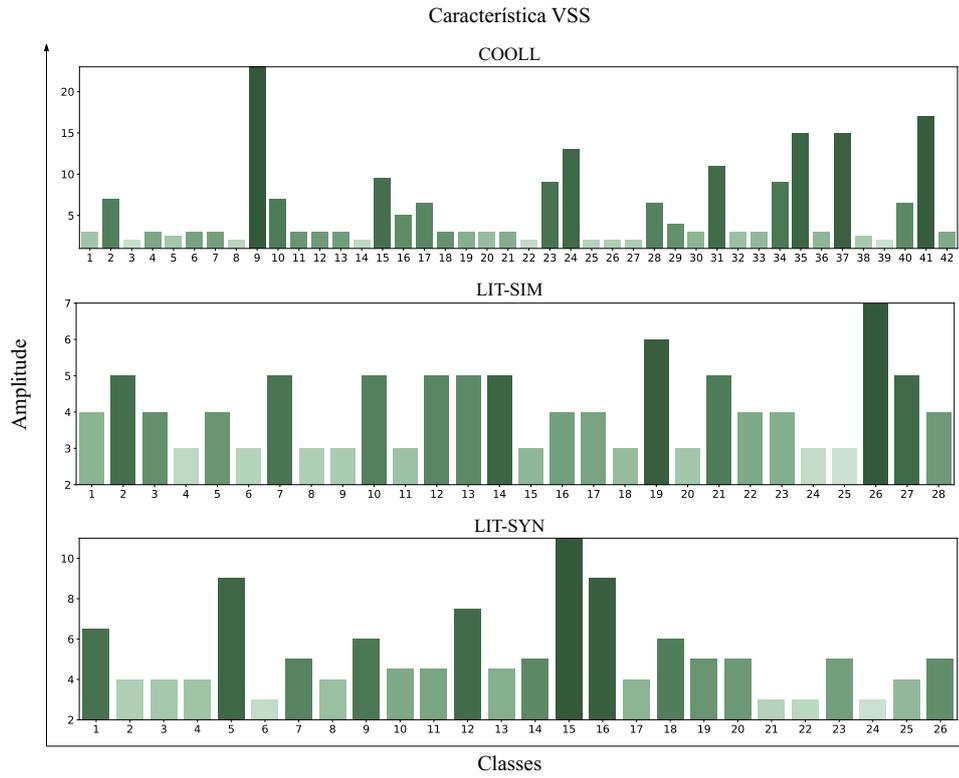


Fonte: Autoria própria.

positivo para o ciclo negativo, pois essa carga é uma furadeira na velocidade 1 (retificador meia onda) como mostra a Tabela 3. Para a base de dados LIT-SIM a classe que mais se difere é a 24 pelo seu valor elevado. Já na base de dados LIT-SYN as classes 15 e 23 parecem apresentar os maiores valores.

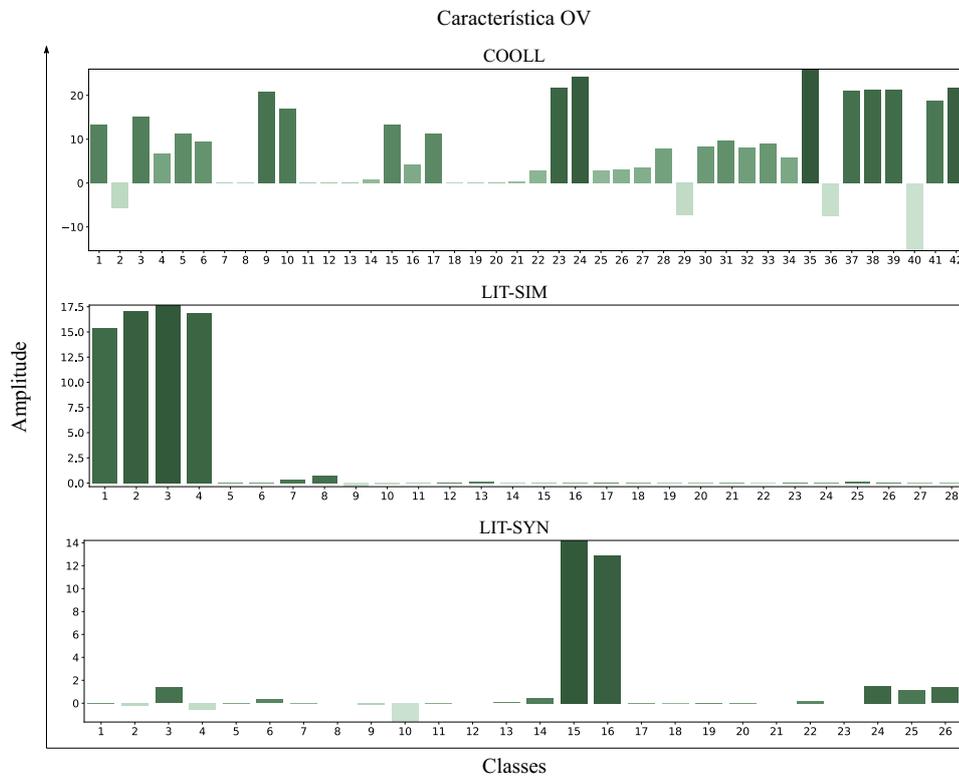
- *asl* (valor médio da inclinação): A Figura 53 exibe os valores medianos obtidos para essa característica em cada classe das bases de dados. Na base de dados COOLL, essa característica parece destacar as classes 14 e 40 por seus valores inferiores e as classes 29 e 37 por seus valores superiores. Para a base LIT-SIM a característica diferencia as classes 9, 10 e 25, com seus valores inferiores, e as classes 12, 17 e 19 por seus valores superiores. Já para a LIT-SYN a característica destaca principalmente a classe 24, pelo valor inferior, e a classe 22 pelo valor mais elevado.
- *dbd* (diferença entre distâncias): A Figura 54 apresenta os valores medianos dessa característica para cada classe das três bases de dados analisadas. Na base COOLL os maiores

Figura 49 – Valores medianos da característica *vss* para cada classe.



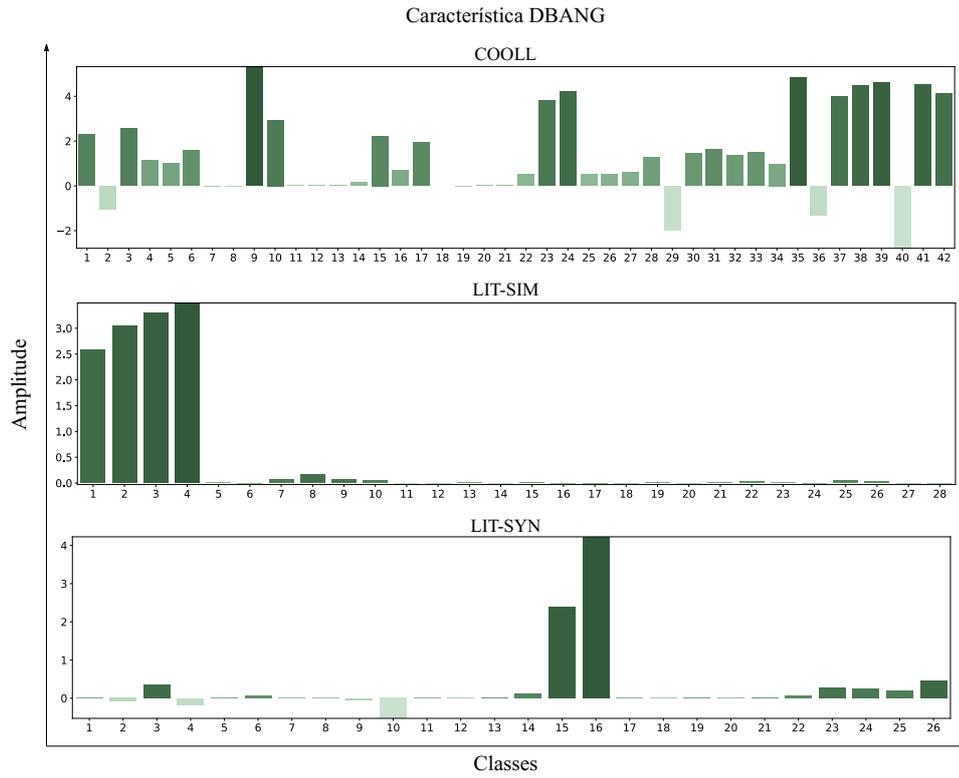
Fonte: Autoria própria.

Figura 50 – Valores medianos da característica *ov* para cada classe.



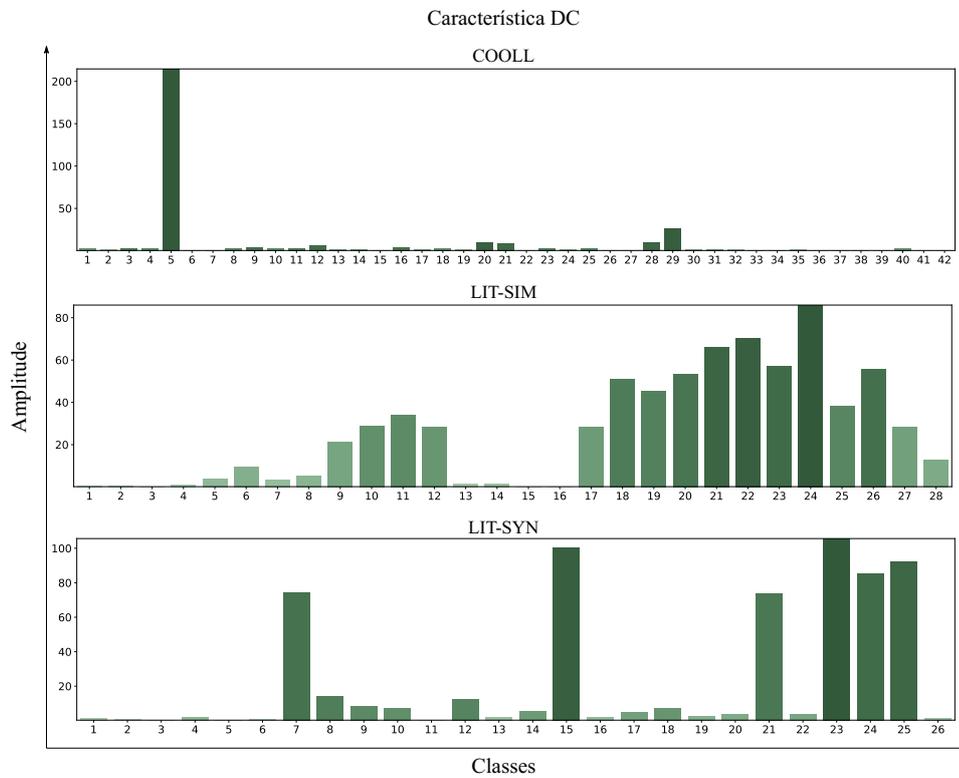
Fonte: Autoria própria.

Figura 51 – Valores medianos da característica *dbang* para cada classe.



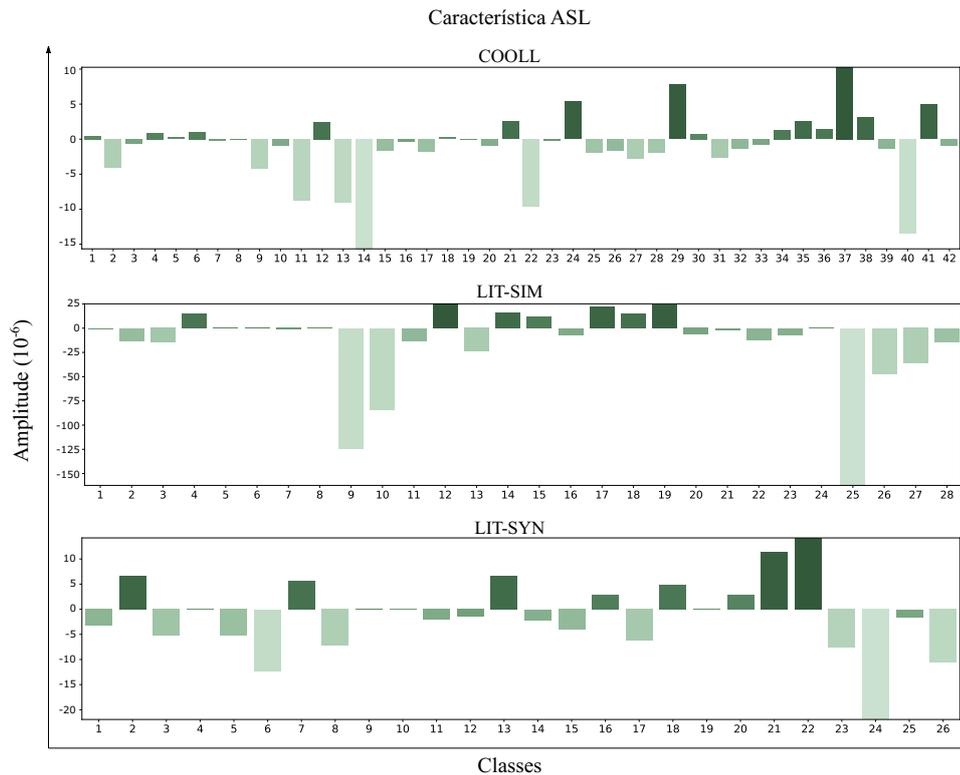
Fonte: Autoria própria.

Figura 52 – Valores medianos da característica *dc* para cada classe.



Fonte: Autoria própria.

Figura 53 – Valores medianos da característica *asl* para cada classe.

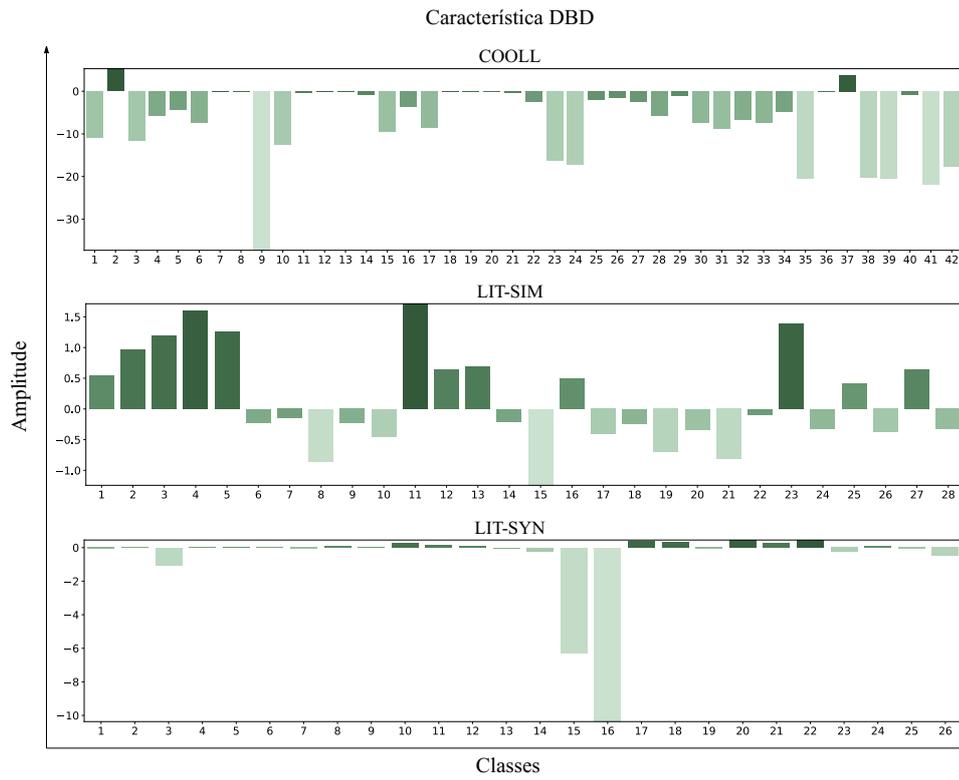


Fonte: Autoria própria.

valores para características são visualizados para as classes 2 e 37 e o menor valor para a classe 9. Para a base de dados LIT-SIM, a classe 11 apresenta o maior valor da característica e a classe 15 o menor valor. Já para a LIT-SYN as classes que mais se diferem das demais são a 15 e 16, devido aos seus valores inferiores comparados as outras.

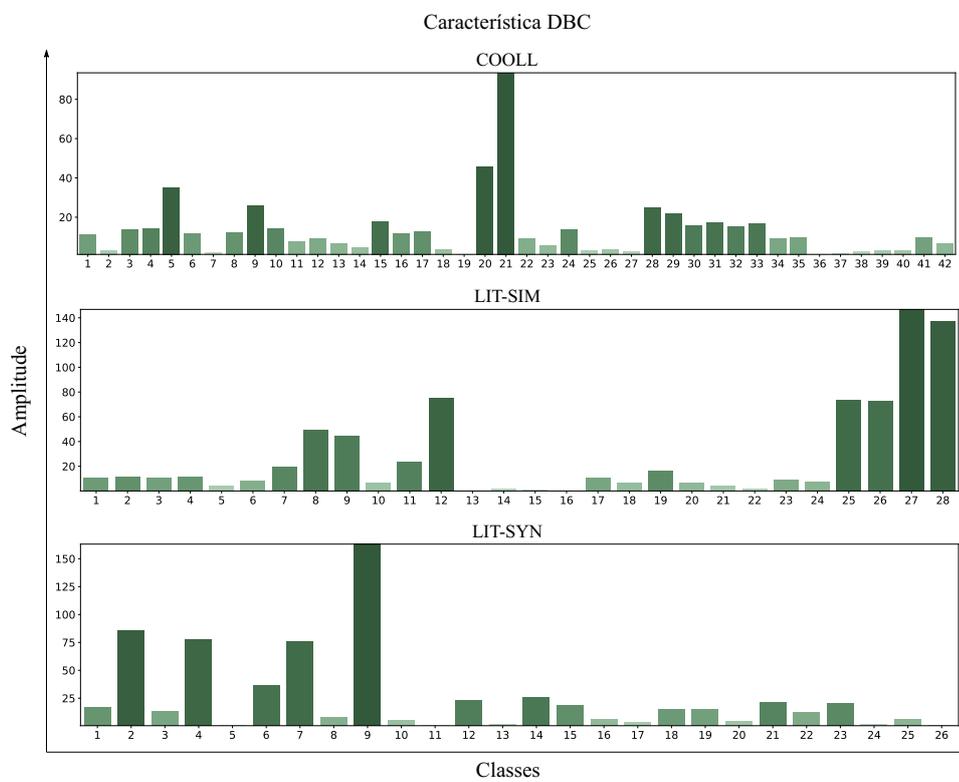
- *dbc* (diferença entre os centroides): A Figura 55 exibe os valores dessa característica para cada classe das bases de dados avaliadas. Para base de dados COOLL, as classes 20 e 21 se destacam por seus valores superiores. Para a base LIT-SIM as classes 27 e 28 recebem destaque por seus valores maiores. Já para a base de dados LIT-SYN a classe 9 é evidenciada entre as demais também por seu valor superior da característica.
- *angc* (ângulo do centroide): A Figura 56 apresenta os valores dessa característica para as classes das três bases de dados analisadas. Com a figura pode-se notar que essa característica parece destacar algumas classes principalmente com valores superiores que as demais. Para a base COOLL, destacam-se as classes 29 e 40 pelo valores superiores e a classe 41 pelo valor mais inferior que as demais. Para a base LIT-SIM a classe 9 parece se

Figura 54 – Valores medianos da característica *dbd* para cada classe.



Fonte: Autoria própria.

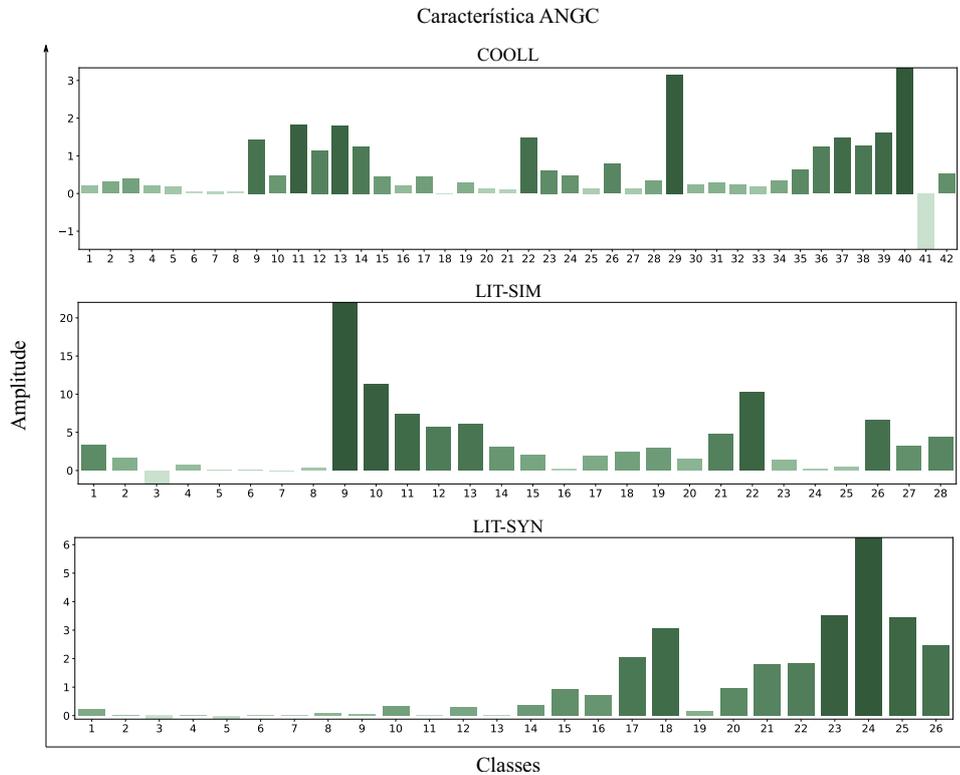
Figura 55 – Valores medianos da característica *dbc* para cada classe.



Fonte: Autoria própria.

destacar das demais pelo valor superior da característica e a classe 3 pelo valor inferior. Já para a LIT-SYN, a classe 24 é diferenciada das outras por seu valor superior de *angc*.

Figura 56 – Valores medianos da característica *angc* para cada classe.



Fonte: Autoria própria.

De modo geral, pode-se observar que determinadas características são mais influentes para alguns tipos de carga, por exemplo: (i) Cargas com propriedades resistivas: características *len*, *md*, *dpb*, *angp*, *asl*, *dbd* e *angc*; (ii) Cargas com base em eletrônica de potência: características *vss*, *dc*, *asl* e *dbc* (iii) Cargas com altas correntes de energização: característica *ov*, *dbang* e *dbd*.

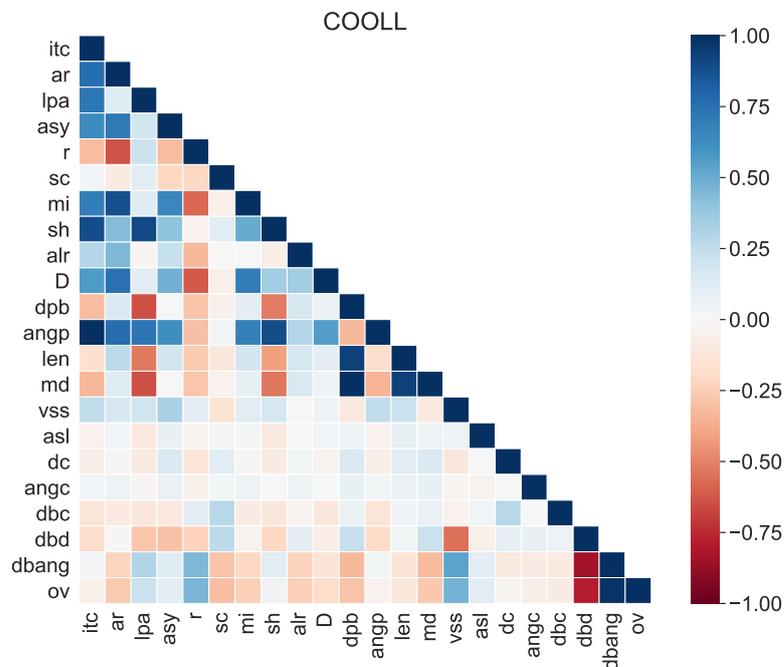
Além da análise realizada com cada uma das características, descritas na seção 4.3, esta seção também realiza uma análise exploratória preliminar da correlação linear existente entre as características para as bases de dados utilizadas. Essa análise pode informar a forma como duas variáveis se relacionam de modo linear. Neste trabalho essa medida foi obtida por meio do coeficiente de correlação de Pearson r , definido como:

$$r = \frac{\sum_{i=1}^{NT} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left[\sum_{i=1}^{NT} (x_i - \bar{x})^2 \right] \left[\sum_{i=1}^{NT} (y_i - \bar{y})^2 \right]}}, \quad (40)$$

em que x e y representam duas características distintas para os NT exemplos de um conjunto (MUKAKA, 2012).

Se o coeficiente r gera resultados mais próximos a 1, há uma correlação linear positiva entre as variáveis. Se o resultado é mais próximo a -1, existe uma correlação negativa entre as variáveis. Em geral, as características que possuem alta correlação (positiva ou negativa) entre si tendem a não aumentar o desempenho de classificação quando incluídas no conjunto (BISHOP, 2006). Além de quanto mais próximo o coeficiente for a zero, menor é a relação linear de uma variável com a outra. As Figuras 57, 58 e 59 exibem as correlações encontradas para as bases COOLL, LIT- SIM 5 e LIT- SYN respectivamente.

Figura 57 – Correlação entre as características para a base de dados COOLL.



Fonte: Autoria própria.

A partir das Figuras 57 e 59 é possível notar que as características propostas por Wang *et al.* (2018) (*itc*, *ar*, *lpa*, *asy*, *r*, *sc*, *mi*, *sh*, *alr* e *D*), em maioria, apresentam correlações iguais ou superiores a 0,5 para as bases de dados COOLL e LIT- SYN. Para a base de dados LIT- SIM, Figura 58, essas características passam a ter uma correlação com medidas menores. Já ao avaliar a correlação entre essas características e as propostas por Mulinari *et al.* (2019) (*dpb*, *angp*, *len*, *md*, *vss*, *asl*, *dc*, *angc*, *dbc*, *dbd*, *dbang* e *ov*) é possível notar que as maiores correlações exibidas, para as três bases utilizadas, são das características *itc* (corrente atual) e *lpa* (direção do *loop*) com as características *dpb* (distância entre os pontos máximo e mínimo), *angp*

(ângulo entre os pontos máximo e mínimo), *len* (comprimento) e *md* (máxima distância). Essa alta correlação é consequência dessas características estarem fortemente relacionadas a potência de uma carga.

Outra conclusão que se pode observar nas Figuras 57, 58 e 59 é como as características *dbd* (diferença entre distâncias), *dbang* (diferença entre ângulos) e *ov* (sobressinal) também apresentam uma alta correlação entre elas, pois as três estão relacionadas à diferença entre a trajetória V-I transitória para a estacionária. Por fim, também se observa como as características *vss*, *asl*, *dc*, *angc*, *dbc*, *dbd*, *dbang* e *ov* apresentam baixas correlações com as características de Wang *et al.* (2018). Isso pode ser explicado pelo fato de parte das novas características utilizarem informações sobre a transição da carga, o que não era considerado nas características anteriores e reforça a relevância da inclusão dessas características. Além disso, as outras características propostas, ainda que utilizem somente a trajetória estacionária, estão relacionadas às informações de frequência e simetria das cargas, de modo distinto às características anteriores, o que também enfatiza a importância dessa inclusão.

É válido destacar que os resultados obtidos das análises de correlação linear das características exibidos nesse trabalho não são definitivos para definir quais variáveis apresentam semelhanças entre si, visto que podem existir correlações não-lineares entre as características analisadas. Ainda assim, essa análise exploratória preliminar já mostra algumas conclusões que serão refinadas ao incluir os resultados de classificação com essas características.

5.1.2 Classificação de cargas com a metodologia tradicional

Neste trabalho os testes realizados com a metodologia tradicional das trajetórias V-I visam complementar as conclusões obtidas em Mulinari *et al.* (2019). Inicialmente, avaliou-se o impacto nas acurácias obtidas com diferentes classificadores em três bases de dados ao realizar a normalização das medidas de tensão e corrente (no intervalo de -1 a 1) que formam as trajetórias V-I, antes da extração das 22 características. A Tabela 6 apresenta os resultados obtidos com a normalização. Já a Tabela 7 os resultados sem utilizar a normalização.

Como é possível analisar pelas Tabelas 6 e 7, na maioria dos casos, as classificações realizadas com as características extraídas das trajetórias V-I não normalizadas apresentaram acurácias maiores ou equivalentes as classificações realizadas com as trajetórias normalizadas. No caso da base de dados LIT- SYN o aumento ocorreu para maior parte dos classificadores, isso pode ter ocorrido pois ao utilizar as características de uma trajetória sem normalização, as

Tabela 6 – Acurácias para o método tradicional das trajetórias V-I com a trajetória normalizada.

Classificador	Base de Dados		
	COOLL	LIT-SIM 5	LIT-SYN
<i>Ensemble</i>	99,52% ± 0,41%	93,63% ± 1,06%	95,46% ± 0,66%
KVMP	87,30% ± 3,17%	65,37% ± 2,19%	89,37% ± 1,49%
MVS	95,87% ± 2,96%	94,11% ± 2,13%	94,08% ± 1,96%
ADL	98,41% ± 1,24%	88,17% ± 0,89%	71,58% ± 2,44%
Árvore de decisão	99,05% ± 1,11%	92,13% ± 1,12%	92,06% ± 1,11%

Fonte: Autoria própria.

Tabela 7 – Acurácias para o método tradicional das trajetórias V-I com a trajetória sem normalização.

Classificador	Base de Dados		
	COOLL	LIT-SIM 5	LIT-SYN
<i>Ensemble</i>	99,21% ± 0,75%	92,61% ± 1,12%	96,54% ± 0,76%
KVMP	96,27% ± 1,80%	76,88% ± 5,53%	95,32% ± 2,06%
MVS	97,46% ± 2,01%	91,19% ± 2,13%	92,51% ± 7,08%
ADL	97,46% ± 1,34%	89,20% ± 1,18%	82,30% ± 1,18%
Árvore de decisão	98,41% ± 1,30%	91,95% ± 1,59%	93,66% ± 1,41%

Fonte: Autoria própria.

informações de potência sobre as cargas ficam mais evidentes e podem complementar a distinção de um conjunto de dados que apresenta aparelhos com potências diversificadas.

Visto que realizar a normalização da trajetória V-I não gera benefícios para a maioria dos classificadores e pode ser considerada mais uma etapa a ser considerada no tempo de extração das características, as próximas avaliações foram realizadas com as trajetórias V-I sem esse pré-processamento. Além disso, a partir das Tabelas 6 e 7 é possível notar que o melhor classificador, para todas as bases de dados consideradas, foi o *Ensemble*. Assim, os resultados seguintes dessa subseção foram realizados com a utilização somente desse modelo de classificação.

Com o objetivo de reduzir a dimensionalidade do conjunto de atributos do método da trajetória V-I tradicional, alguns conjuntos menores de características foram analisados. Para comparar esses conjuntos avaliou-se o tempo médio de extração, decorrente das dez execuções no sistema embarcado, como descrito na Subseção 4.6, e a acurácia obtida na classificação das bases de dados utilizadas. O conjunto um analisado considerou somente as dez características propostas por Wang *et al.* (2018), o dois somente as 12 características propostas por Mulinari *et al.* (2019) e o três considerou as 12 características propostas por Mulinari *et al.* (2019), mais as características de corrente atual (itc) e de variação da admitância (D) proposta por Wang *et al.* (2018). Somente as duas características foram escolhidas para agregar o terceiro conjunto pois são características que já foram empregadas em outros trabalhos além de Wang *et al.* (2018). Os resultados obtidos com esses conjuntos iniciais são mostrados na Tabela 8. Para comparação desses resultados, o conjunto que utiliza todas as características (22 no total) também foi apresentado.

Tabela 8 – Avaliação de acurácias e tempo de extração dos conjuntos.

Conjunto	Num. Caract.	Base de Dados			Tempo de extração (ms)
		COOLL	LIT-SIM 5	LIT-SYN	
Todas	22	99,21% ± 0,75%	92,61% ± 1,12%	96,54% ± 0,76%	424,432 ± 0,0033
1	10	99,05% ± 1,53%	92,92% ± 1,62%	93,71% ± 2,24%	422,543 ± 0,0035
2	12	96,11% ± 1,52%	81,28% ± 1,78%	90,83% ± 2,91%	12,133 ± 0,0004
3	14	97,86% ± 1,06%	85,89% ± 1,32%	94,53% ± 1,14%	13,095 ± 0,0005

Fonte: Autoria própria.

Como é possível notar pela Tabela 8, as melhores acurácias geralmente são obtidas utilizando todas as características. O único caso em que esse conjunto não gera a maior acurácia é para a base de dados LIT- SIM 5, em que o conjunto 1 obtém um resultado praticamente equivalente ao de todas as características. Já ao avaliar o critério de tempo de extração, é possível notar que, como esperado, o conjunto que consome mais tempo é o que contém todas as características. Entretanto, é possível notar que os características que influenciam nesse valor são os que compõem o conjunto 1, isso pode ocorrer devido ao fato delas serem calculadas por meio das partes divididas da trajetória, demandando mais tempo na criação e processamento dos conjuntos dessa divisão. Além disso, o conjunto que apresentou o menor tempo de extração é o conjunto 2, porém os resultados das classificações para esse conjunto foram os piores. Por fim, o terceiro conjunto apresentou uma redução das acurácias comparado aos valores obtidos com todas as características, mas também houve uma grande redução do tempo de extração. Portanto, avaliando somente o desempenho de classificação, o conjunto de todas as características é o mais satisfatório. Já quando se considera o desempenho e tempo de extração, o conjunto 3 é o mais adequado.

Os conjuntos apresentados na Tabela 8 foram divididos conforme as propostas de Wang *et al.* (2018) e Mulinari *et al.* (2019). Entretanto, nenhum critério ou processo foi utilizado para determinar as características que iriam compor cada um deles. Por esse motivo, com o auxílio da caixa de ferramenta do MATLAB, de estatística e aprendizagem de máquina, foi realizado o processo de seleção de características empregando os métodos de seleção direta e de eliminação reversa, descritos na Subseção 3.4, com a função *sequentialfs*. Nessa função, definiu-se o classificador *Ensemble* e a acurácia como critério de avaliação para seleção dos conjuntos. Além disso, a função realiza validações cruzadas para cada conjunto candidato de característica. Neste trabalho, para criação dos conjuntos para a validação, tanto para o método de seleção direta quanto o de eliminação reversa, utilizou-se a função *cvpartition* com tamanho igual a cinco. Desse modo cinco conjuntos aleatórios foram criados.

Para realizar a eliminação reversa, utilizaram-se três conjuntos diferentes como entrada, referentes às três bases de dados utilizadas, com as 22 características extraídas. O conjunto 4 utilizou o conjunto de características obtidas da base de dados COOLL e obteve como resultado da eliminação reversa que as características *sh* e *len* deveriam ser desconsideradas. Já o conjunto cinco, considerou os dados iniciais da LIT- SIM 5 e, como resultado da eliminação, removeu a característica *sh*. Por fim, o conjunto seis, que considerou os dados da LIT- SYN, eliminou as características *itc*, *r* e *angc*. A Tabela 9 apresenta os resultados obtidos de classificação e de tempo de extração ao utilizar os resultados da eliminação dos três conjuntos de referência para as demais bases de dados. Os resultados obtidos com todas as características também são apresentados para comparação.

Tabela 9 – Resultados obtidos da seleção de características com a eliminação reversa.

Conj.	Caract. elim.	Num. de caract.	Base de dados aplicada			Tempo de extração (ms)
			COOLL	LIT-SIM 5	LIT-SYN	
Todas	-	22	99,21% ± 0,75%	92,61% ± 1,12%	96,54% ± 0,76%	424,432 ± 0,0033
4	<i>sh</i> , <i>len</i>	20	99,68% ± 0,55%	92,27% ± 1,59%	96,14% ± 0,88%	423,104 ± 0,0035
5	<i>sh</i>	21	99,37% ± 0,82%	93,54% ± 1,33%	96,51% ± 0,82%	424,042 ± 0,0033
6	<i>itc</i> , <i>r</i> , <i>angc</i>	19	95,48% ± 1,87%	92,22% ± 1,74%	96,04% ± 1,80%	425,551 ± 0,0031

Fonte: Autoria própria.

De acordo com a Tabela 9, as melhores acurácias, considerando todas as bases e comparando aos resultados com os de todos os recursos, são obtidas ao utilizar o conjunto cinco, que remove a característica *sh* (inclinação do segmento médio). Entretanto, no quesito tempo, nenhum dos conjuntos da eliminação reversa apresenta uma redução considerável quando comparado ao conjunto com todas as características.

No processo da seleção direta, também se utilizaram os três conjuntos iniciais, referentes às três bases de dados utilizadas neste trabalho, como entrada de avaliação desse processo. O conjunto 7 utilizou as características extraídas da base de dados COOLL e resultou na seleção de cinco características: *ar*, *r*, *angp*, *md* e *dbang*. Já o conjunto 8 utilizou as características da base LIT- SIM 5 e resultou na seleção de seis variáveis de entrada: *ar*, *lpa*, *asy*, *angp*, *dbang* e *ov*. Por fim, o conjunto 9, que empregou as características da LIT- SYN, selecionou seis características: *itc*, *ar*, *r*, *dpb*, *md* e *dbang*. A Tabela 10 apresenta os resultados de classificação e tempo de extração ao utilizar as características selecionadas em cada base de referência nas demais bases de dados. Os resultados da utilização de todas as características também são apresentados na tabela para comparação.

Tabela 10 – Resultados obtidos da seleção de características com a seleção direta.

Conj.	Caract. selec.	Num. de car.	Base de dados aplicada			Tempo de extração (ms)
			COOLL	LIT-SIM 5	LIT-SYN	
Todas	-	22	99,21% ± 0,75%	92,61% ± 1,12%	96,54% ± 0,76%	424,432 ± 0,0033
7	ar, angp, r, md, dbang	5	99,60% ± 0,77%	90,84% ± 1,2%	95,41% ± 0,75%	10,022 ± 0,0004
8	ar, lpa, asy, angp, dbang, ov	6	98,17% ± 1,19%	93,78% ± 1,22%	93,84% ± 0,71%	418,538 ± 0,0032
9	itc, ar, r, dpb, md, dbang	6	98,97% ± 1,06%	90,68% ± 1,14%	96,40% ± 0,77%	10,037 ± 0,0000

Fonte: Autoria própria.

Ao analisar a Tabela 10 é possível notar que algumas características parecem se repetir nos três conjuntos resultantes, como *ar* (área), *angp* (ângulo entre os pontos máximo e mínimo) e *dbang* (diferença entre as distâncias). Essas características podem ser recorrentes pois apresentam valores bem distintos para cargas de tipos diferentes. Além disso, a partir da tabela se pode observar que o conjunto que apresenta o desempenho mais adequado, dado o tempo de extração mais reduzido, é o conjunto 7. O conjunto 8 obteve bom desempenho para as três bases de dados consideradas. Entretanto, o tempo de extração é praticamente equivalente ao tempo de extração de todos os recursos. Já o conjunto 9, embora apresente um tempo de extração reduzido, exibe um desempenho menor para duas das três bases de dados comparado ao conjunto sete. Por esse motivo, o conjunto 7 foi utilizado para comparação com o conjunto de todas as características para avaliação da robustez da metodologia a diferentes níveis de ruído, detalhada na sequência.

Para realizar a avaliação da robustez das características a inserção de ruído, a base de dados LIT- SIM 3 foi utilizada e diferentes níveis de ruído branco Gaussiano foram adicionados às medidas de tensão e corrente das aquisições. Embora a faixa de valores de ruído considerada ultrapasse os níveis reais encontrados em um sistema elétrico de potência, essa análise foi realizada para avaliar o impacto dessas alterações das amostras no desempenho de classificação. A Tabela 11 mostra os resultados obtidos desse processo.

Ao analisar a Tabela 11 é possível notar como a remoção das características do grupo 7 influenciam na sensibilidade da metodologia com a adição de ruído. Além disso, observa-se como o ruído influencia o método tradicional, mesmo empregando todas as características, pois causa a deformação das trajetórias V-I.

Outro atributo avaliado para todos os conjuntos analisados foi a memória utilizada no

Tabela 11 – Avaliação da robustez das características ao utilizar diferentes níveis de ruído.

SNR (dB)	Características	
	Todas	Conjunto 7
Sem Ruído	99,74% ± 0,30%	99,82% ± 0,20%
60	99,00% ± 0,61%	97,36% ± 0,99%
50	98,27% ± 0,79%	96,34% ± 1,08%
40	95,21% ± 0,99%	93,40% ± 1,15%
30	92,61% ± 1,12%	85,60% ± 1,59%
20	74,71% ± 2,05%	69,26% ± 1,23%
10	51,38% ± 2,44%	40,56% ± 2,55%
5	37,42% ± 2,20%	30,21% ± 1,83%

Fonte: Autoria própria.

sistema embarcado. Entretanto, houve pouca variação do espaço utilizado comparado a quando se manipula todas as características (106 kB de memória RAM e 50 kB de memória Flash).

Também, para avaliar a otimização dos tempos de extração obtidos com a utilização de todas as características (424,432 ms) e com a utilização do conjunto 7 (10,022 ms), utilizou-se a *flag O2*, que o compilador GCC (do inglês GNU *Compiler Collection*) disponibiliza. Essa *flag* realiza a otimização das execuções ao alterar as ordens em que as instruções são executadas na aplicação (FOUNDATION, 1988-2020). Com a utilização da *flag* o tempo de execução com todas as características passou para 123,336 ms e com o conjunto sete, para 8,108 ms.

5.2 FOURIER 2D

A partir da metodologia em que se utiliza a F2D na trajetória V-I média, juntamente com as características de potência ativa dos dois estados de uma carga, como descrita na Seção 4.4, os primeiros testes realizados observaram a influência da variação dos harmônicos analisados e, por consequência, a variação da quantidade de características extraídas das trajetórias V-I nas classificações das cargas com as bases de dados escolhidas. Somente o classificador *Ensemble* foi considerado nessa avaliação, pois apresenta um tempo pequeno de treinamento e acurácia relativamente superior. As acurácias obtidas nas classificações, bem como o tempo médio de extração no sistema embarcado, referente a dez execuções dessas característica para uma única forma de onda, são exibidos na Tabela 12.

Como é possível avaliar na Tabela 12, o aumento da quantidade de harmônicos analisados não gera um aumento na acurácia obtida. Esse comportamento já foi discutido em trabalhos como Caple *et al.* (2017). Considerando o tempo médio de extração e a quantidade total de características extraídas, o melhor conjunto é obtido ao utilizar cinco harmônicos. Porém, ao

Tabela 12 – Acurácias com a variação do número de harmônicos analisadas.

Harmônicos	Quant. de caract.	Base de dados			Tempo de Extração (ms)
		COOLL	LIT-SIM 5	LIT-SYN	
5	24	99,29% ± 0,79%	96,64% ± 1,26%	94,43% ± 2,34%	252,877
6	28	99,13% ± 1,37%	97,16% ± 0,78%	96,75% ± 0,91%	352,848
7	32	98,89% ± 1,76%	96,38% ± 1,11%	94,99% ± 1,22%	469,866
8	36	99,13% ± 1,27%	97,10% ± 0,78%	95,23% ± 1,97%	603,945
9	40	99,13% ± 0,95%	96,85% ± 1,16%	96,59% ± 1,46%	755,129
10	44	97,94% ± 2,22%	96,88% ± 1,18%	94,94% ± 1,08%	923,433
11	48	98,02% ± 1,55%	97,14% ± 0,78%	94,12% ± 2,83%	1108,864
12	52	98,89% ± 1,07%	96,63% ± 1,13%	95,33% ± 0,70%	1311,466
13	56	98,89% ± 1,19%	97,11% ± 1,02%	95,08% ± 1,75%	1531,251
14	60	98,65% ± 0,92%	97,16% ± 1,03%	94,89% ± 1,09%	1768,205
15	64	97,94% ± 1,55%	97,03% ± 1,03%	94,82% ± 1,33%	2022,361

Fonte: Autoria própria.

avaliar o desempenho de classificação o conjunto que apresentou resultados melhores para maioria das bases de dados é a análise com seis harmônicos. Por esse motivo, os próximos resultados foram obtidos com essa quantidade de harmônicos.

Com o intuito de diminuir o tempo de extração obtido na Tabela 12, utilizaram-se os recursos de otimização de código da plataforma Renesas, *flag O2*, para obtenção de um tempo médio de extração de 180,428 ms para o conjunto com os seis harmônicos. Além disso, outra tentativa de redução do tempo de extração utilizou a trajetória V-I sub-amostrada, na qual a F2D é calculada em um intervalo a cada duas amostras de uma trajetória V-I original. Esse procedimento não gerou alterações no desempenho de classificação mas reduziu o tempo de extração para 96,632 ms.

Na sequência, avaliou-se o impacto da normalização das medidas de tensão e corrente, no intervalo de -1 a 1, nas acurácias das classificações com diferentes modelos. Os resultados obtidos com a trajetória normalizada são apresentados na Tabela 13 e os resultados sem normalização na Tabela 14.

Tabela 13 – Acurácias para o método da Fourier 2D com a trajetória V-I normalizada.

Classificador	Base de Dados		
	COOLL	LIT-SIM 5	LIT-SYN
<i>Ensemble</i>	98,33% ± 0,87%	95,78% ± 0,93%	96,14% ± 1,18%
KVMP	96,90% ± 1,81%	85,69% ± 1,44%	90,51% ± 2,30%
MVS	98,97% ± 0,99%	93,54% ± 1,46%	91,77% ± 5,39%
ADL	99,05% ± 0,73%	90,18% ± 1,17%	75,96% ± 2,17%
Árvore de decisão	97,62% ± 0,84%	94,42% ± 0,74%	93,63% ± 1,83%

Fonte: Autoria própria.

Como no caso da extração de características utilizando a metodologia tradicional das trajetórias V-I, na subseção 5.1.2, a normalização da trajetória V-I não gera resultados melhores

Tabela 14 – Acurácias para o método da Fourier 2D com a trajetória V-I sem normalização.

Classificador	Base de Dados		
	COOLL	LIT-SIM 5	LIT-SYN
<i>Ensemble</i>	99,13% ± 1,37%	97,16% ± 0,78%	96,75% ± 0,91%
KVMP	80,48% ± 3,02%	77,18% ± 5,04%	88,07% ± 4,14%
MVS	98,33% ± 1,21%	94,55% ± 1,14%	94,87% ± 1,08%
ADL	99,13% ± 0,79%	93,65% ± 1,42%	78,49% ± 2,57%
Árvore de decisão	97,62% ± 1,30%	95,36% ± 1,03%	94,36% ± 0,99%

Fonte: Autoria própria.

do que utilizar a trajetória sem normalização. Em alguns casos, realizar a normalização gera resultados inferiores, devido ao fato de encobrir informações que distinguem cargas que, embora possam exibir um mesmo formato de trajetória, apresentem potências distintas. Por esse motivo, essa etapa de pré-processamento não é considerada nos resultados posteriores. Além disso, dadas as Tabelas 13 e 14, o classificador que apresentou os melhores resultados para duas das três bases de dados consideradas foi o *Ensemble*. Desse modo, somente esse classificador é utilizado no próximo teste.

O próximo teste avaliou a robustez da metodologia F2D com a adição de ruído branco Gaussiano aos dados de tensão e corrente da base LIT- SIM 3. A Tabela 11 mostra os resultados obtidos desse processo.

Tabela 15 – Avaliação da robustez da metodologia F2D ao utilizar diferentes níveis de ruído.

SNR (dB)	F2D
Sem Ruído	99,57% ± 0,32%
60	99,56% ± 0,36%
50	99,44% ± 0,42%
40	98,86% ± 0,73%
30	97,16% ± 0,78%
20	84,18% ± 0,71%
10	67,33% ± 2,32%
5	54,33% ± 1,80%

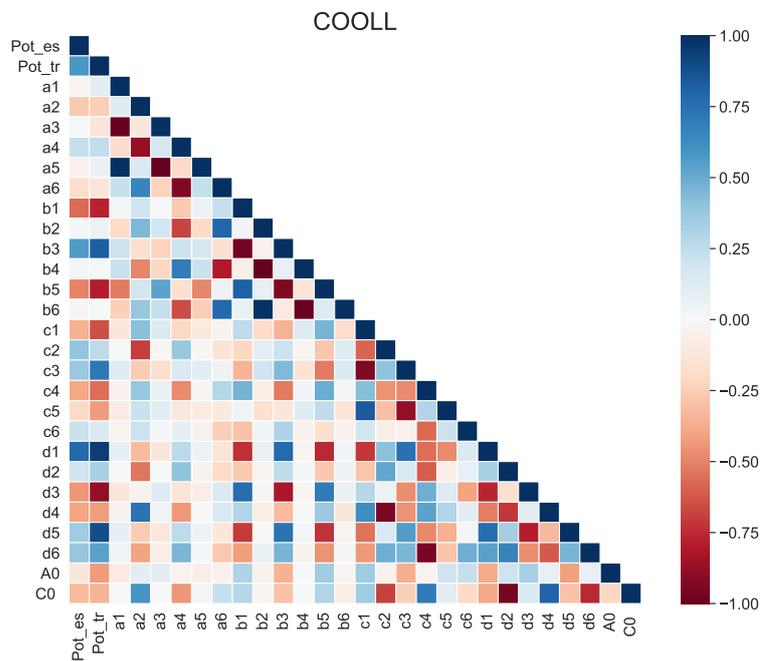
Fonte: Autoria própria.

Com os resultados obtidos da Tabela 15 é possível notar que a metodologia sofre um impacto maior quando o SNR decresce. Esse comportamento era esperado, visto que as trajetórias V-I ficam deformadas com a adição de ruído.

Outra análise exploratória realizada com a metodologia F2D, também avaliou a correlação linear entre as 28 características obtidas para os três conjuntos de dados considerados. As Figuras 60, 61 e 62 apresentam os resultados obtidos para as bases de dados COOLL, LIT- SIM 5 e LIT- SYN respectivamente.

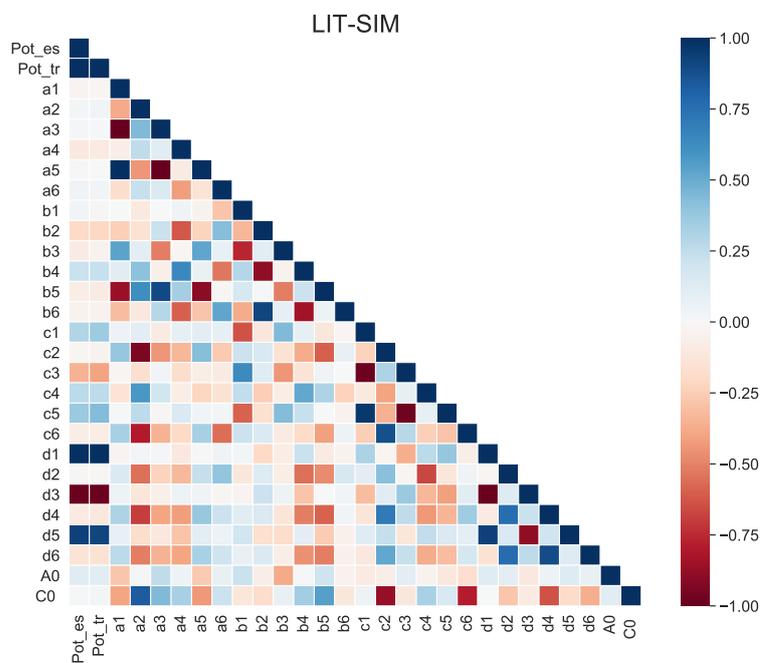
Ao analisar as Figuras 60, 61 e 62 é possível notar que as correlações positivas e

Figura 60 – Correlação entre as características da F2D para a base de dados COOLL.



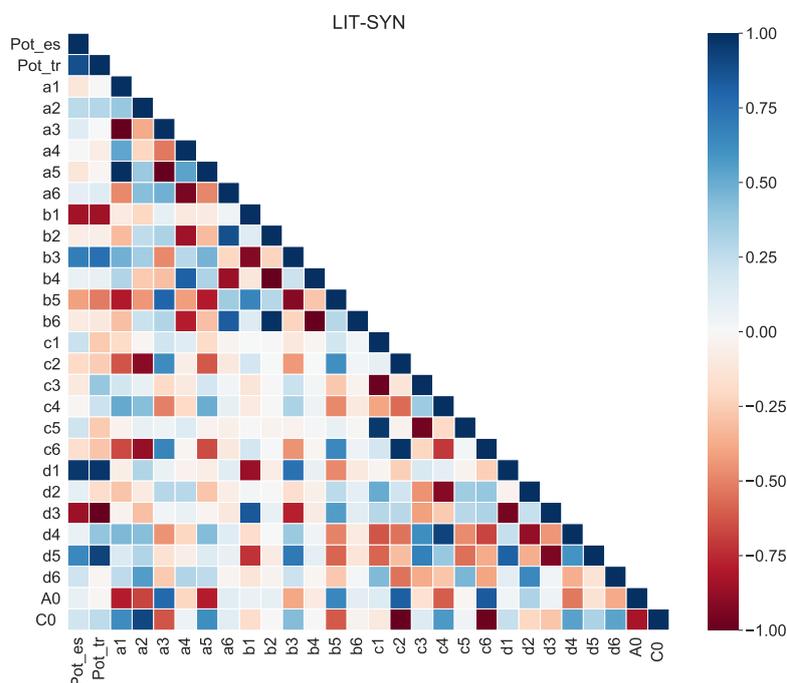
Fonte: Autoria própria.

Figura 61 – Correlação entre as características da F2D para a base de dados LIT-SYM 5.



Fonte: Autoria própria.

Figura 62 – Correlação entre as características da F2D para a base de dados LIT-SYN.



Fonte: Autoria própria.

negativas maiores podem ser percebidas entre as potências, *Pot_es* e *Pot_ts*, e entre essas características e os descritores. Isso ocorre pois os valores obtidos dos descritores, quando não se utiliza a trajetória normalizada, são influenciados pelo tamanho da curva obtida e, por consequência, carregam informações de potência da carga. Além disso, ao observar a Figura 62, comparada às outras bases de dados, as características apresentam maiores correlações entre si. Isso pode ter ocorrido pois essa base é a que apresenta a menor diversidade de aparelhos disponível.

Por fim, para essa metodologia F2D também se avaliou a quantidade de memória utilizada no sistema embarcado, considerando a análise de seis harmônicos e o resultado obtido foi de 66 kB de memória RAM e 49 kB de memória Flash.

5.3 COMPARAÇÕES ENTRE METODOLOGIAS

Esta última seção dos resultados apresenta uma comparação entre as metodologias: V-I tradicional (MULINARI *et al.*, 2019), com todas as características e com o conjunto 7, Fourier Bidimensional F2D, que é uma das propostas deste trabalho, e a metodologia que

utiliza a energia da Transformada *Wavelet* (TW) sobre o sinal de corrente² (CHANG *et al.*, 2013; LAZZARETTI *et al.*, 2020). Visto que os melhores resultados obtidos na classificação para ambas as metodologias utilizaram o classificador *Ensemble*, somente esse modelo será considerado nas comparações iniciais.

A Tabela 16 apresenta os resultados obtidos na classificação das bases de dados CO-OLL, LIT-SIM 5 e LIT-SYN com os quatro métodos de extração. Os resultados apresentados nessa tabela consideram os métodos com as otimizações discutidas nas seções 5.1.2 para V-I tradicional e 5.2 para F2D. Para a transformada *Wavelet*, também se empregou a otimização de código sugerida pela plataforma da Renesas.

Tabela 16 – Comparação das metodologias de extração de características.

Ext.	Base de dados			Tempo de extração (ms)	Memória RAM (kB)
	COOLL	LIT-SIM 5	LIT-SYN		
Trad. Todas	99,21% ± 0,75%	92,61% ± 1,12%	96,54% ± 0,76%	123,336 ± 0,0033	106
Trad. Conj. 7	99,60% ± 0,77%	90,84% ± 1,20%	95,41% ± 0,75%	8,108 ± 0,0003	103
F2D	99,13% ± 1,37%	97,16% ± 0,78%	96,75% ± 0,91%	96,632 ± 0,0572	66
TW	99,68% ± 0,55%	97,62% ± 0,89%	97,97% ± 0,49%	30,956 ± 0,0003	85

Fonte: Autoria própria.

A partir da Tabela 16 é possível notar que os desempenhos mais altos são obtidos com a utilização do método de extração com a TW. Entretanto, é possível notar que o método de extração F2D apresenta um desempenho na classificação muito próximo a essa metodologia. Além disso, quanto mais propriedades são avaliadas, como tempo médio de extração e memória utilizada, os métodos de extração propostos neste trabalho apresentam valores competitivos.

Outra análise realizada com os quatro métodos de extração foi a avaliação de robustez das metodologias quanto a adição de ruído nas aquisições. Os resultados dessa avaliação utilizando a metodologia tradicional é apresentada na Seção 5.1.2 e com para F2D na Seção 5.2. Os resultados obtidos para a TW são dados na Tabela 17. Para comparação da robustez das metodologias a Figura 63 exhibe os resultados obtidos para os quatro métodos.

Ao avaliar a Figura 63 nota-se que a metodologia que sofre maior influência com a adição do ruído é a tradicional que utiliza um conjunto de características selecionadas, isso porque, como tratado nas subseções anteriores, as características que compõe o conjunto total da metodologia devem agregar informações em cenários de distorção da forma da trajetória. O método da Fourier 2D também é influenciado pelas distorções, porém em menor amplitude

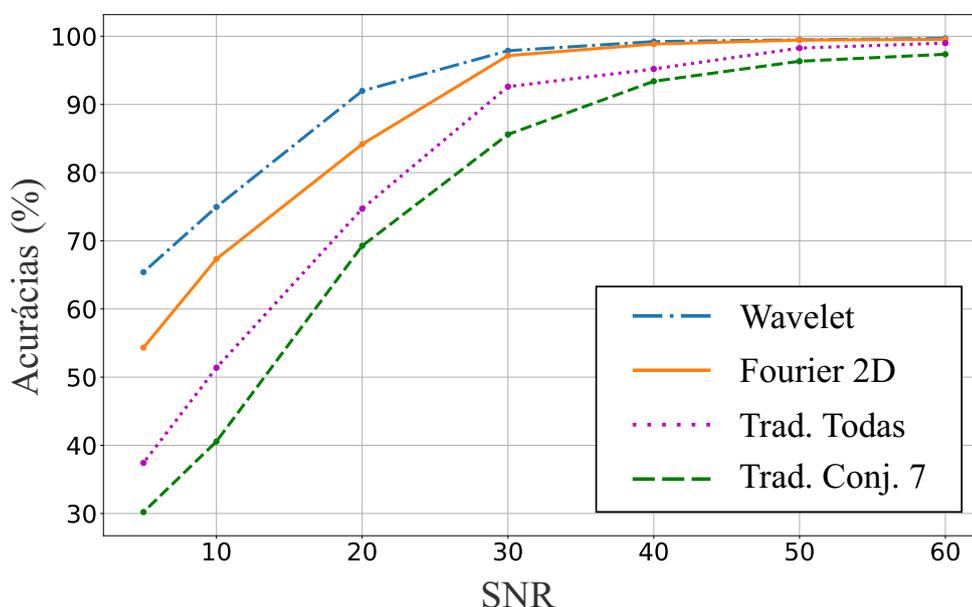
² Este é o melhor método reportado na literatura atual para estas bases de dados.

Tabela 17 – Avaliação da robustez da TW ao utilizar diferentes níveis de ruído.

SNR (dB)	Wavelet
Sem ruído	99,63% ± 0,66%
60	99,70% ± 0,17%
50	99,48% ± 0,61%
40	99,22% ± 0,50%
30	97,88% ± 1,11%
20	91,99% ± 0,81%
10	74,96% ± 1,30%
5	65,41% ± 1,71%

Fonte: Autoria própria.

Figura 63 – Avaliação do desempenho de classificação com a inserção de ruído nas amostras de entrada para os métodos analisados.



Fonte: Autoria própria.

que a metodologia tradicional, pois essa nova metodologia descreve uma forma aproximada da trajetória. Já o método de extração que utiliza a TW é o menos influenciado, isso pode ser decorrente desta metodologia utilizar somente uma das medidas, neste caso a corrente da carga.

Por fim, este trabalho também avaliou o desempenho de classificação quando o treinamento utiliza as características extraídas em um contexto e o teste com as características extraídas em outro, de modo a avaliar a generalização dos modelos de classificação utilizados (LAZZARETTI *et al.*, 2020). Para realizar esse processo utilizou-se a base de dados LIT-SYN 1 e LIT-SYN 3 para treinamento dos modelos de classificação. Enquanto os testes foram realizados com as bases de dados LIT-SYN 2, LIT-SYN 3 e LIT-SYN 8. Assim, por exemplo, um método extrai as características da base LIT-SYN 1, em que as aquisições se referem a das cargas

individuais nas quais são utilizadas para treinamento de um modelo de classificação. Na etapa de teste, as características obtidas em outro contexto, por exemplo da base LIT-SYN 2, em que as aquisições apresentam cargas duplas agregadas, são aplicadas no modelo de classificação. Os resultados obtidos nessa avaliação são apresentados na Tabela 18.

Tabela 18 – Avaliação do desempenho de classificação ao realizar o treinamento em um contexto e o teste em outro.

Ext.	Base Treino	Base Teste	Ensemble (%)	KVMP (%)	MVS (%)	ADL (%)	Árvore de decisão (%)
Trad. Todas	LIT-SYN 1	LIT-SYN 2	90,10	45,46	43,53	74,48	83,26
	LIT-SYN 1	LIT-SYN 3	69,57	43,56	45,77	63,57	64,46
	LIT-SYN 1	LIT-SYN 8	56,48	25,35	24,65	49,88	40,62
	LIT-SYN 3	LIT-SYN 8	76,39	68,29	74,42	77,89	74,65
Trad. Conj. 7	LIT-SYN 1	LIT-SYN 2	77,08	56,77	59,89	67,11	79,61
	LIT-SYN 1	LIT-SYN 3	53,72	46,21	41,35	59,28	50,32
	LIT-SYN 1	LIT-SYN 8	35,99	28,36	24,77	41,09	33,22
	LIT-SYN 3	LIT-SYN 8	78,93	68,87	76,50	73,15	77,08
F2D	LIT-SYN 1	LIT-SYN 2	92,19	31,55	33,26	65,99	93,53
	LIT-SYN 1	LIT-SYN 3	57,51	29,36	32,26	45,01	63,64
	LIT-SYN 1	LIT-SYN 8	43,63	21,99	15,97	27,55	55,67
	LIT-SYN 3	LIT-SYN 8	80,90	42,82	77,31	63,66	78,24
TW	LIT-SYN 1	LIT-SYN 2	91,67	90,99	71,13	84,15	80,21
	LIT-SYN 1	LIT-SYN 3	72,09	75,95	54,73	62,06	50,82
	LIT-SYN 1	LIT-SYN 8	51,04	51,62	33,79	30,21	33,56
	LIT-SYN 3	LIT-SYN 8	85,53	65,74	72,92	50,58	79,74

Fonte: Autoria própria.

A partir da Tabela 18 é possível avaliar a capacidade dos modelos de classificação, para diferentes extratores, em generalizar as informações obtidas na etapa de treinamento, com as aquisições de cargas individuais ou com poucas cargas agregadas, para um cenário de teste com um número maior de cargas múltiplas. Para todas as metodologias observa-se a diminuição do desempenho de classificação ao utilizar a base LIT-SYN 1 no treinamento e durante o teste aumentar a quantidade de cargas múltiplas. Isso ocorre pois todos os métodos sofrem atenuações ou deformações em suas assinaturas de cargas proporcionais ao aumento das cargas múltiplas, o que as difere das formas individuais utilizadas no treinamento. Nesse contexto, as metodologias propostas neste trabalho obtiveram desempenho maior ou equivalente a TW, principalmente no pior conjunto avaliado que treina com a LIT-SYN 1 e testa com a LIT-SYN 8. Também a partir da tabela observa-se que o aumento do desempenho da classificação para a base LIT-SYN 8, com todos os métodos de extração, ocorre quando na etapa de treinamento são adicionadas cargas múltiplas (LIT-SYN 3).

Com os resultados apresentados nesta seção de comparação é possível considerar as metodologias propostas, tradicionais ou com a F2D, como alternativas para a classificação de

cargas elétricas, principalmente quando na aplicação há restrições como tempo de execução e memória utilizada pelo método de extração. Além disso, essas metodologias apresentaram acurácias muito próximas as obtidas com o método de extração da TW. Também, essas metodologias propostas podem ser utilizadas para agregar abordagens que realizam a combinação de métodos de extração, como a apresentada em Renaux *et al.* (2018) e Lazzaretti *et al.* (2020)

Por fim, é possível realizar a comparação das duas metodologias expostas neste trabalho e, a partir dos resultados exibidos nesta seção de resultados, concluir que a metodologia que aplica a F2D para extração de características apresenta desempenho de classificação, robustez ao ruído e quantidade de memória utilizada com valores melhores ao valores obtidos para metodologia tradicional que compõe o estado da arte atual.

6 CONCLUSÃO

Entre os principais desafios de NILM está o desenvolvimento de métodos capazes de distinguir qual é o aparelho que foi conectado ou desconectado em um determinado momento na instalação. Entre as etapas de um sistema NILM, a de extração de característica visa utilizar métodos para extrair informações a respeito de cada tipo de aparelho para que posteriormente, combinados a um modelo de classificação, permitam identificar qual aparelho modificou sua condição de operação. Nesse contexto, este trabalho concentrou-se no desenvolvimento de duas metodologias de extração de características que tivessem como base a assinatura de carga denominada trajetória V-I e nas quais os dois estados, de transição e permanente, fossem utilizados.

Dado os objetivos definidos na Seção 1.2 para o trabalho, todos eles foram atendidos a partir da criação de novas métricas sobre a trajetória V-I, análises e seleção de características, classificações de cargas, avaliação da utilização de recursos de memória e tempo computacional, além da comparação das metodologias criadas com o método de extração com a TW.

As análises realizadas com a metodologia tradicional procuraram estender as conclusões obtidas e descrever com maiores informações as características propostas em Mulinari *et al.* (2019). Neste trabalho, as características foram descritas e analisadas para três bases de dados COOLL, LIT-SIM 5 e LIT-SYN, na Subseção 5.1.1, o que possibilitou avaliar para cada característica quais tipos de cargas são evidenciadas. Além disso, as análises de correlações das características para cada conjunto de dados utilizado evidenciou como as características propostas, no geral, apresentam uma baixa correlação entre si e com as características já existentes na literatura. Os resultados, na Subseção 5.1.2, exibiram que o conjunto total de características com a metodologia tradicional apresentou desempenho de classificação, combinada ao classificador *Ensemble*, superior à 99% para COOLL, 92% para LIT-SIM 5 e 96% para LIT-SYN, com tempo médio de extração no sistema embarcado inferior a 124 ms e utilização de memória RAM igual a 106 kB.

Ainda com a metodologia tradicional, avaliou-se a criação de conjuntos menores de características a partir de métodos de seleção de atributos. Em todos os conjuntos selecionados é possível observar como as características propostas são escolhidas em maior número ou não são eliminadas dos conjuntos. Com o melhor conjunto obtido por meio da seleção de características, no qual selecionaram-se somente 5 características, em que três delas são propostas

em Mulinari *et al.* (2019), também se observaram resultados relevantes, sendo o desempenho de classificação superior a 90% em todas as bases de dados, tempo médio de extração no sistema embarcado inferior a 9 ms e utilização de memória RAM igual a 103 kB, resultados presentes na Subseção 5.1.2. Todos os algoritmos referentes a cada característica na metodologia tradicional encontram-se disponíveis no *GitHub* da autora¹.

Além da metodologia tradicional, este trabalho apresentou uma nova aplicação a F2D, ao empregá-la como um método de extração sobre as trajetórias V-I. Essa metodologia exibiu algumas vantagens quando comparada ao estado da arte das trajetórias V-I, dado que apresentou um desempenho melhor de classificação, sendo superior a 99% para COOLL, 97% para LIT-SIM 5 e 96% para LIT-SYN. Além disso, se mostrou menos suscetível a adição de ruídos aos sinais de entrada do método. Finalmente, essa metodologia apresentou um tempo médio de extração inferior a 97 ms e utilização de memória RAM igual a 66 kB, resultados apresentados na Subseção 5.2.

Por fim, na Subseção 5.3 realizou-se a comparação das metodologias propostas, que utilizam a trajetória V-I, com outra metodologia existente na literatura que utiliza a TW. Nos resultados presentes nessa seção, observa-se a viabilidade de utilizar ambas as metodologias propostas para extração de características, bem como as vantagens que cada uma oferece comparadas a metodologia *Wavelet*, como menor tempo de execução com a V-I tradicional e menor utilização de memória com a F2D.

6.1 TRABALHOS FUTUROS

Para a continuação do trabalho descrito, propõem-se:

- Utilizar outros métodos para a seleção de características na metodologia tradicional de extração de características das trajetórias V-I, como a aplicação dos métodos de seleção com filtros e *embedded* desse conjunto de características;
- Avaliar a junção do método de extração com a Fourier 2D com a TW, para descrever melhor formas menos arredondadas, como descrito em (CAPLE *et al.*, 2017);
- Avaliar novas metodologias para a extração das características com a Fourier 2D, pois neste trabalho avaliaram-se metodologias que realizavam a utilização das trajetórias nos

¹ https://github.com/brunamulinari/V-I_trajectory.

dois estados, além das características de potência ativa nos dois estados. Porém, é possível avaliar a aplicação de outras características, diferentes das potências, que resultem em uma projeção mínima com os descritores;

- Utilizar os métodos simplificados e otimizados neste trabalho, como o conjunto selecionado no método tradicional e a Fourier 2D com a trajetória sub-amostrada, para avaliação no impacto da acurácia de classificação, além do tempo e memória, no sistema multiagente descrito em Lazzaretti *et al.* (2020);
- Avaliar o desenvolvimento de um método de pré-classificação ou de complemento a detecção, empregando as trajetórias V-I tradicionais selecionadas no conjunto sete, obtido na Subseção 5.1.2, por apresentarem um menor tempo de execução;
- Analisar as possibilidades de outras otimizações de recursos de tempo e memória para os dois métodos de extração de características apresentados, como reavaliar os vetores e interações criadas;
- Implementar técnicas de transferência de aprendizagem, para treinamento em um contexto e utilização em outro, ou aplicação de um modelo de aprendizagem contínuo e adaptativo com o surgimento de novas classes.

6.2 TRABALHOS PUBLICADOS

Durante o desenvolvimento deste trabalho publicou-se um artigo em congresso internacional diretamente relacionado ao presente trabalho:

MULINARI, B. M.; CAMPOS, D. P.; COSTA, C. H.; ANCELMO, H. C.; LAZZARETTI, A. E.; OROSKI, E.; LIMA, C. R. E.; RENAUX, D. P. B.; POTTKER, F.; LINHARES, R. P. A New Set of Steady-State and Transient Features for Power Signature Analysis Based on V-I Trajectory. In: Innovative Smart Grid Technologies (ISGT), 2019, Gramado.

Além disso, foi realizada a publicação de dois artigos para a revista *Energies*-MDPI (fator de impacto 2.702):

LAZZARETTI, A. E.; RENAUX, D. P. B.; LIMA, C. R. E.; **MULINARI, B. M.**; ANCELMO, H. C.; OROSKI, E.; POTTKER, F.; LINHARES, R. P.; NOLASCO, L. S.; LIMA, L. T.; OMORI, J. S.; SANTOS, R. R. A Multi-Agent NILM Architecture for Event Detection and Load Classification. *Energies MDPI* 2020, 13(17).

RENAUX, D. P. B.; POTTKER, F.; ANCELMO, H. C.; LAZZARETTI, A. E.; LIMA, C. R. E.; LINHARES, R. P.; OROSKI, E.; NOLASCO, L. S.; LIMA, L. T.; **MULINARI, B. M.**; SILVA, J. R. L.; OMORI, J. S.; SANTOS, R. R. A Dataset for Non-Intrusive Load Monitoring: Design and Implementation. *Energies MDPI*, 2020, 13(20).

Embora não apresentem resultados do presente trabalho, também foi publicado 1 artigo em congresso nacional e 6 artigos em congressos internacionais nas mesmas áreas de conhecimento, como processamento de sinais e aprendizado de máquina aplicados à sistemas elétricos. Os artigo são:

ANCELMO, H. C.; **MULINARI, B. M.**; POTTKER, F.; LAZZARETTI, A. E.; BAZZO, T. P. M.; OROSKI, E.; RENAUX, D. P. B.; LIMA, C. R. E.; LINHARES, R. P.; GAMBA, A. R. A. A New Simulated Database for Classification Comparison in Power Signature Analysis. In: 20th International Conference on Intelligent Systems Applications to Power Systems (ISAP), 2019, Nova Deli.

ANCELMO, H. C.; GRANDO, F. L.; **MULINARI, B. M.**; COSTA, C. H.; LAZZARETTI, A. E.; OROSKI, E.; RENAUX, D. P. B.; POTTKER, F.; LIMA, C. R. E.; LINHARES, R. P. A Transient and Steady-State Power Signature Feature Extraction Using Different Prony's Methods. In: 20th International Conference on Intelligent Systems Applications to Power Systems (ISAP), 2019, Nova Deli.

ANCELMO, H. C.; GRANDO, F. L.; COSTA, C. H.; **MULINARI, B. M.**; OROSKI, E.; LAZZARETTI, A. E.; POTTKER, F.; RENAUX, D. P. B. Automatic Power Signature Analysis using Prony's Method and Machine Learning-Based Classifiers. In: 2nd European Conference on Electrical Engineering & Computer Science. 2nd European Conference on Electrical Engineering & Computer Science, 2018.

ANCELMO, H. C.; **MULINARI, B. M.**; OROSKI, E.; LAZZARETTI, A. E.; LINHARES, R. P.; LIMA, C. R. E.; RENAUX, D. P. B.; POTTKER, F. Método Não Intrusivo De Classificação de Cargas Baseadas em Power Signature In: Smart Energy, 2018, Curitiba. Smart Energy. , 2018.

COSTA, C. H.; MORITZ, G. L.; LAZZARETTI, A. E.; **MULINARI, B. M.**; ANCELMO, H. C.; RODRIGUES, M. P.; OROSKI, E.; GOES, R. E. A Comparison of Machine Learning-Based Methods for Fault Classification in Photovoltaic Systems. In: Innovative Smart Grid Technologies (ISGT), 2019, Gramado. IEEE Innovative Smart Grid Technologies (ISGT), 2019.

LINHARES, R. P.; LIMA, C. R. E.; RENAUX, D. P. B.; POTTKER, F.; OROSKI, E.; LAZZARETTI, A. E.; **MULINARI, B. M.**; ANCELMO, H. C.; GAMBA, A. R.; BERNARDI, L. A.; LIMA, L. T.; One-millisecond low-cost synchronization of wireless sensor network. In: IX Brazilian Symposium on Computing Systems Engineering, 2019, Salvador.

POTTKER, F.; LAZZARETTI, A. E.; OROSKI, E.; RENAUX, D. P. B.; LINHARES, R. P.; LIMA, C. R. E.; ANCELMO, H. C.; **MULINARI, B. M.** Non-Intrusive Load Monitoring: A Multi-Agent Architecture and Results. In: 2nd European Conference on Electrical Engineering & Computer Science, 2018., 2018.

REFERÊNCIAS

ABUBAKAR, I.; KHALID, S. N. H.; MUSTAFA, M. W.; SHAREEF, H.; MUSTAPHA, M. Recent approaches and applications of non-intrusive load monitoring. **ARPN journal of engineering and applied sciences**, v. 11, p. 4609–4618, 2016.

AHA, D. W.; BANKERT, R. L. A comparative evaluation of sequential feature selection algorithms. *In: In Proceedings of the Fifth International Workshop on Artificial Intelligence and Statistics*. [S.l.]: Springer-Verlag, 1995. p. 1–7.

ALTRABALSI, H.; LIAO, J.; STANKOVIC, L.; STANKOVIC, V. A low-complexity energy disaggregation method: Performance and robustness. *In: 2014 IEEE Symposium on Computational Intelligence Applications in Smart Grid (CIASG)*. [S.l.: s.n.], 2014. p. 1–8.

ANCELMO, H.; MULINARI, B. M.; POTTKER, F.; LAZZARETTI, A.; BAZZO, T. P. M. de; OROSKI, E.; RENAUX, D.; LIMA, C. E.; LINHARES, R. A new simulated database for classification comparison in power signature analysis. *In: 20th International Conference on Intelligent Systems Applications to Power Systems*. [S.l.: s.n.], 2019. p. 1 – 7.

ANCELMO, H. C.; GRANDO, F. L.; MULINARI, B. M.; COSTA, C. H. da; LAZZARETTI, A. E.; OROSKI, E.; RENAUX, D. P. B.; POTTKER, F.; LIMA, C. R. E.; LINHARES, R. R. A transient and steady-state power signature feature extraction using different prony's methods. *In: 20th International Conference on Intelligent System Application to Power Systems (ISAP)*. [S.l.: s.n.], 2019. p. 1–6.

ANEEL. **Módulo 8 – Qualidade da Energia Elétrica**: Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional – PRODIST. [S.l.], 2020. 89 p.

ARGATU, F.C.; ARGATU, V.; ENACHE, B.A.; CEPISCA, C.; SERITAN, G.C.; ANDREI, P.C. Load monitoring solutions for the smart buildings – an overview. **The Scientific Bulletin of Electrical Engineering Faculty**, Sciendo, Berlin, v. 18, n. 2, p. 1 – 6, 2018.

ARMEL, K. C.; GUPTA, A.; SHRIMALI, G.; ALBERT, A. Is disaggregation the holy grail of energy efficiency? the case of electricity. **Energy Policy**, v. 52, p. 213 – 234, 2013. Special Section: Transition Pathways to a Low Carbon Economy.

BAETS, L. De; DEVELDER, C.; DHAENE, T.; DESCHRIJVER, D. Automated classification of appliances using elliptical fourier descriptors. **IEEE International Conference on Smart Grid Communications**, n. October, p. 153–158, 2017.

BAETS, L. De; DEVELDER, C.; DHAENE, T.; DESCHRIJVER, D. Detection of unidentified appliances in non-intrusive load monitoring using siamese neural networks. **International Journal of Electrical Power and Energy Systems**, Elsevier, v. 104, p. 645–653, 2019.

BAETS, L. De; DHAENE, T.; DESCHRIJVER, D.; BERGES, M.; DEVELDER, C. Vi-based appliance classification using aggregated power consumption data. **Proceedings - 2018 IEEE International Conference on Smart Computing, SMARTCOMP 2018**, p. 179–186, 2018.

BAETS, L. De; RUYSSINCK, J.; DEVELDER, C.; DHAENE, T.; DESCHRIJVER, D. Appliance classification using vi trajectories and convolutional neural networks. **Energy and Buildings**, Elsevier B.V., v. 158, p. 32–36, 2018.

BAPTISTA, D.; MOSTAFA, S. S.; PEREIRA, L.; SOUSA, L.; MORGADO-DIAS, F. Implementation strategy of convolution neural networks on field programmable gate arrays for appliance classification using the voltage and current (v-i) trajectory. **Energies**, v. 11, n. 9, p. 1–18, 2018.

BERGES, M.; GOLDMAN, E.; MATTHEWS, H.; SOIBELMAN, L.; ANDERSON, K. User-centered nonintrusive electricity load monitoring for residential buildings. **Journal of Computing in Civil Engineering**, v. 25, p. 471–480, 11 2011.

BISHOP, C. M. **Pattern Recognition and Machine Learning**. Berlin, Heidelberg: Springer-Verlag, 2006. ISBN 0387310738.

BONFIGLI, R.; SQUARTINI, S. Hmm based approach. *In: ____*. **Machine Learning Approaches to Non-Intrusive Load Monitoring**. Cham: Springer International Publishing, 2020. p. 31–90.

BREIMAN, L.; FRIEDMAN, J.; OLSHEN, R.; STONE, C. J. **Classification and Regression Trees**. Boca Raton: Taylor & Francis, 1984.

CAIMAN, D. V.; DRAGOMIR, T. L. Empirical voltage-current signatures for individual household consumers obtained by non-linear regression. **Studies in Informatics and Control**, ICI Bucharest, v. 28, n. 2, p. 201–212, 2019.

CAPLE, J.; BYRD, J.; STEPHAN, C. Elliptical fourier analysis: fundamentals, applications, and value for forensic anthropology. **International Journal of Legal Medicine**, v. 131, 02 2017.

CHANG, H.-H. Non-intrusive demand monitoring and load identification for energy management systems based on transient feature analyses. **Energies**, MDPI AG, v. 5, n. 11, p. 4569–4589, Nov 2012.

CHANG, H.-H.; LIAN, K.-L.; SU, Y.-C.; LEE, W.-J. Energy spectrum-based wavelet transform for non-intrusive demand monitoring and load identification. *In: 2013 IEEE Industry Applications Society Annual Meeting. [S.l.: s.n.]*, 2013. p. 1–9.

CHANG, H.-H.; LIAN, K.-L.; SU, Y.-C.; LEE, W.-J. Power-spectrum-based wavelet transform for nonintrusive demand monitoring and load identification. **IEEE Transactions on Industry Applications**, v. 50, n. 3, p. 2081–2089, 2014.

CHANG, H.-H.; WIRATHA, P. W.; N.CHEN. A non-intrusive load monitoring system using an embedded system for applications to unbalanced residential distribution systems. **Energy Procedia**, v. 61, p. 146 – 150, 2014. International Conference on Applied Energy, ICAE2014.

CHERKASSKY, V. S.; MULIER, F. **Learning from Data: Concepts, Theory, and Methods**. 1st. ed. New York, NY, USA: John Wiley & Sons, Inc., 1998. ISBN 0471154938.

DU, L.; HE, D.; HARLEY, R. G.; HABETLER, T. G. Electric load classification by binary voltage-current trajectory mapping. **IEEE Transactions on Smart Grid**, ieeexplore.ieee.org, v. 7, n. 1, p. 358–365, 2016.

DUBUISSON, M. .; JAIN, A. K. A modified hausdorff distance for object matching. *In: Proceedings of 12th International Conference on Pattern Recognition. [S.l.: s.n.]*, 1994. v. 1, p. 566–568 vol.1.

DUDA, R.O.; HART, P.E.; STORK, D.G. **Pattern Classification**. 2. ed. New York: Wiley-Interscience, 2000.

Efficiency Valuation Organization. **International Performance Measurement and Verification Protocol: Concepts and Options for Determining Energy and Water Savings**. 2012. Disponível em: <https://evo-world.org>.

ENSSLIN, L.; ENSSLIN, S. R.; PINTO, H. de M. Processo de investigação e análise bibliométrica: avaliação da qualidade dos serviços bancários. **Revista de Administração Contemporânea**, v. 17, n. 3, p. 325–349, 2013.

FAUSTINE, A.; MVUNGI, N. H.; KAIJAGE, S.; KISANGIRI, M. A survey on non-intrusive load monitoring methodies and techniques for energy disaggregation problem. **CoRR**, [abs/1703.00785](http://arxiv.org/abs/1703.00785), 2017. Disponível em: <http://arxiv.org/abs/1703.00785>.

FIGUEIREDO, M.; ALMEIDA, A.; RIBEIRO, B. Home electrical signal disaggregation for non-intrusive load monitoring systems. **Neurocomputing**, v. 96, p. 66–73, nov 2012. ISSN 0925-2312.

FIGUEIREDO, M.; RIBEIRO, B.; ALMEIDA, A. de. Electrical signal source separation via nonnegative tensor factorization using on site measurements in a smart home. **IEEE Transactions on Instrumentation and Measurement**, v. 63, n. 2, p. 364–373, 2014.

FIGUEIREDO, M. B.; ALMEIDA, A. de; RIBEIRO, B. An experimental study on electrical signature identification of non-intrusive load monitoring (nilm) systems. *In: Adaptive and Natural Computing Algorithms*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011. p. 31–40.

FOUNDATION, Free Software. **Using the GNU Compiler Collection (GCC)**: Options that control optimization. *[S.l.]*, 1988–2020. Disponível em: <https://gcc.gnu.org/onlinedocs/gcc/>.

GAO, J.; KARA, E. C.; GIRI, S.; BERGÉS, M. A feasibility study of automated plug-load identification from high-frequency measurements. *In: 2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. *[S.l.: s.n.]*, 2015. p. 220–224.

GERE, J.M.; GOODNO, B.J. **Mechanics of Materials**. Toronto: Cengage Learning, 2009.

GILLIS, J. M.; ALSHAREEF, S. M.; MORSI, W. G. Nonintrusive load monitoring using wavelet design and machine learning. **IEEE Transactions on Smart Grid**, v. 7, n. 1, p. 320–328, 2016.

GOPINATH, R.; KUMAR, M.; JOSHUA, C. P. C.; SRINIVAS, K. Energy management using non-intrusive load monitoring techniques – state-of-the-art and future research directions. **Sustainable Cities and Society**, v. 62, p. 102411, 2020.

GUYON, I.; ELISSEEFF, A. An introduction of variable and feature selection. **J. Machine Learning Research Special Issue on Variable and Feature Selection**, v. 3, p. 1157 – 1182, 01 2003.

HART, G. W. Nonintrusive appliance load monitoring. **Proceedings of the IEEE**, v. 80, n. 12, p. 1870–1891, 1992.

HASSAN, T.; JAVED, F.; ARSHAD, N. An empirical investigation of v-i trajectory based load signatures for non-intrusive load monitoring. **IEEE Transactions on Smart Grid**, IEEE, v. 5, n. 2, p. 870–878, 2014.

HERNÁNDEZ, Á.; RUANO, A.; UREÑA, J.; RUANO, M.G.; GARCIA", J.J. Applications of nilm techniques to energy management and assisted living. **IFAC-PapersOnLine**, v. 52, n. 11, p. 164 – 171, 2019. 5th IFAC Conference on Intelligent Control and Automation Sciences ICONS 2019.

HUDGINS, B.; PARKER, P.; SCOTT, R. N. A new strategy for multifunction myoelectric control. **IEEE Transactions on Biomedical Engineering**, v. 40, n. 1, p. 82–94, Jan 1993.

IEA. **World Energy Outlook 2019**. [S.l.], 2019. <https://www.oecd-ilibrary.org/content/publication/caf32f3b-en>.

KELLY, J.; KNOTTENBELT, W. Neural nilm: Deep neural networks applied to energy disaggregation. **Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments - BuildSys '15**, ACM Press, 2015.

KIM, J.; LE, T.-T.-Huong; KIM, H. Nonintrusive load monitoring based on advanced deep learning and novel signature. **Computational Intelligence and Neuroscience**, v. 2017, p. 1–22, 10 2017.

KUHL, F.; GIARDINA, C. Elliptic fourier features of a closed contour. **Comput. Graph. Image Process.**, v. 18, p. 236–258, 1982.

LAM, H. Y.; FUNG, G. S.K.; LEE, W. K. A novel method to construct taxonomy electrical appliances based on load signatures. **IEEE Transactions on Consumer Electronics**, v. 53, n. 2, p. 653–660, 2007.

LANTZ, B. **Machine Learning with R**. [S.l.]: Packt, 2013. 396 p.

LAZZARETTI, A.E.; RENAUX, D.P.B.; LIMA, C.R.E.; MULINARI, B.M.; ANCELMO, H.C.; OROSKI, E.; PÖTTKER, F.; LINHARES, R.R.; NOLASCO, L.S.; LIMA, L.T.; OMORI, J.S.; SANTOS, R.B. A multi-agent nilm architecture for event detection and load classification. **Energies**, v. 13, n. 17, p. 1–37, 2020.

LE, X.-C.; VRIGNEAU, B.; SENTIEYS, O. 11-norm minimization based algorithm for non-intrusive load monitoring. *In: 2015 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*. [S.l.: s.n.], 2015. p. 299–304.

LIAO, J.; ELAFOUDI, G.; STANKOVIĆ, L.; STANKOVIC, V. Power disaggregation for low-sampling rate data. *In: NILM Workshop 2014*. [S.l.: s.n.], 2014.

LIN, S.; ZHAO, L.; LI, F.; LIU, Q.; LI, D.; FU, Y. A nonintrusive load identification method for residential applications based on quadratic programming. **Electric Power Systems Research**, Elsevier B.V., v. 133, p. 241–248, 2016.

LINHARES, R. R.; LIMA, C. R. E.; RENAUX, D. P. B.; PÖTTKER, F.; OROSKI, E.; LAZZARETTI, A. E.; M.MULINARI, B.; ANCELMO, H. C.; GAMBA, A.; BERNARDI, L. A.; LIMA, L. T. One-millisecond low-cost synchronization of wireless sensor network. *In: VIII Brazilian Symposium on Computing Systems Engineering*. [S.l.: s.n.], 2018.

LIU, H. **Non-intrusive Load Monitoring, Theory, Technologies and Applications**. [S.l.]: Science Press, 2020. ISBN 978-981-15-1859-1.

LIU, H.; MOTODA, H. **Computational Methods of Feature Selection (Chapman & Hall/Crc Data Mining and Knowledge Discovery Series)**. [S.l.]: Chapman & Hall/CRC, 2007. ISBN 1584888784.

LIU, H.; WU, H.; YU, C. A hybrid model for appliance classification based on time series features. **Energy and Buildings**, v. 196, p. 112 – 123, 2019.

LIU, Y.; WANG, X.; YOU, W. Non-intrusive load monitoring by voltage–current trajectory enabled transfer learning. **IEEE Transactions on Smart Grid**, v. 10, n. 5, p. 5609–5619, 2019.

MAKONIN, S. **Real-Time Embedded Low-Frequency Load Disaggregation**. 08 2014. Tese (Doutorado) — Simon Fraser University, Burnaby, Canadá, 08 2014.

MATHIS, M.; RUMSCH, A.; KISTLER, R.; ANDRUSHEVICH, A.; KLAPPROTH, A. Improving the recognition performance of nialm algorithms through technical labeling. *In: 2014 12th IEEE International Conference on Embedded and Ubiquitous Computing*. [S.l.: s.n.], 2014. p. 227–233.

MUKAKA, M. M. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. **Malawi Medical Journal**, p. 69–71, 2012.

MULINARI, B. M.; CAMPOS, D. P. de; COSTA, C. H. da; ANCELMO, H. C.; LAZZARETTI, A. E.; OROSKI, E.; LIMA, C. R. E.; RENAUX, D. P. B.; POTTKER, F.; LINHARES, R. R. A new set of steady-state and transient features for power signature analysis based on v-i trajectory. *In: 2019 IEEE PES Innovative Smart Grid Technologies Conference - Latin America (ISGT Latin America)*. [S.l.: s.n.], 2019. p. 1–6.

MÜLLER, A.C.; GUIDO, S. **Introduction to Machine Learning with Python: A Guide for Data Scientists**. [S.l.]: O'Reilly Media, 2016.

Neurio Energy Monitor. **Neurio Energy Monitor**. 2020. A Generac Company. Disponível em: <https://www.neur.io/energy-monitor/>.

OPPENHEIM, A.V.; SCHAFER, R.W.; BUCK, J.R.; LEE, L. **Discrete-time Signal Processing**. [S.l.]: Prentice Hall, 1999. (Prentice Hall international editions). ISBN 9780137549207.

PERSOON, E.; FU, K.-S. Shape discrimination using fourier descriptors. **IEEE Transactions on Systems, Man, and Cybernetics**, v. 7, n. 3, p. 170–179, 1977.

PHINYOMARK, A.; PHUKPATTARANONT, P.; LIMSAKUL, C. Feature reduction and selection for emg signal classification. **Expert Systems with Applications**, Elsevier Ltd, v. 39, n. 8, p. 7420–7431, 2012.

PICON, T.; MEZIANE, M. N.; RAVIER, P.; LAMARQUE, G.; NOVELLO, C.; BUNETEL, J. C. Le; RAINGEAUD, Y. COOLL: Controlled on/off loads library, a public dataset of high-sampled electrical signals for appliance identification. **arXiv:1611.05803**, 2016.

RAFSANJANI, H. N.; AHN, C. R.; CHEN, J. Linking building energy consumption with occupants' energy-consuming behaviors in commercial buildings: Non-intrusive occupant load monitoring (niolm). **Energy and Buildings**, v. 172, p. 317 – 327, 2018.

RENAUX, D. P. B.; LINHARES, R. R.; POTTKER, F.; LAZZARETTI, A. E.; LIMA, C. E. E. de; COELHO-NETO, A.; CAMPANER, M. H. Designing a novel dataset for non-intrusive load monitoring. *In: VIII Brazilian Symposium on Computing Systems Engineering. [S.l.: s.n.]*, 2018.

RENAUX, D. P. B.; POTTKER, F.; ANCELMO, H. C.; LAZZARETTI, C. R. E. Lima A. E.; LINHARES, R. R.; OROSKI, E.; NOLASCO, L. S.; LIMA, L. T.; MULINARI, B. M.; SILVA, J. R. L. da; OMORI, J. S.; SANTOS, R. B. dos. A dataset for non-intrusive load monitoring: Design and implementation. **Energies**, v. 13, n. 20, p. 1–35, 2020.

Renesas Electronics. **Renesas Synergy: Starter Kit SK-S7G2. [S.l.]**, 2015.

RUANO, A.; HERNANDEZ, A.; UREÑA, J.; RUANO, M.; GARCIA, J. Nilm techniques for intelligent home energy management and ambient assisted living: A review. **Energies**, MDPI AG, v. 12, n. 11, p. 2203, Jun 2019.

SADEGHIANPOURHAMAMI, N.; RUYSSINCK, J.; DESCHRIJVER, D.; DHAENE, T.; DEVELDER, C. Comprehensive feature selection for appliance classification in nilm. **Energy and Buildings**, v. 151, p. 98–106, 2017.

SAHA, R.; BERA, J. N.; SARKAR, G. An improved method for load taxonomy using sample shifting technique and signature analysis. **Electric Power Components and Systems**, Taylor & Francis, v. 47, n. 1-2, p. 113–127, 2019.

SNOEK, J.; LAROCHELLE, H.; ADAMS, R. P. Practical bayesian optimization of machine learning algorithms. *In: Proceedings of the 25th International Conference on Neural Information Processing Systems. [S.l.: s.n.]*, 2012. (NIPS'12), p. 2951–2959.

TESHOME, D.; HUANG, T. D.; LIAN, Kuo-Lung. A distinctive load feature extraction based on fryze's time-domain power theory. **IEEE Power and Energy Technology Systems Journal**, ieeexplore.ieee.org, p. 1–1, 2016.

TKACH, D.; HUANG, H.; KUIKEN, T. A. Study of stability of time-domain features for electromyographic pattern recognition. **Journal of NeuroEngineering and Rehabilitation**, v. 7, n. 1, p. 21, 2010.

VAPNIK, V. N. **Statistical Learning Theory**. [S.l.]: Wiley-Interscience, 1998.

WANG, A. L.; CHEN, B. X.; WANG, C. G.; HUA, D. D. Non-intrusive load monitoring algorithm based on features of v-i trajectory. **Electric Power Systems Research**, Elsevier B.V., v. 157, p. 134–144, 2018.

WELIKALA, S.; THELASINGHA, N.; SHAFIE, A.; GODALIYADDA, G M R.; EKANAYAKE, M. P.; EKANAYAKE, J. Implementation of a robust real-time non-intrusive load monitoring solution. **Applied Energy**, v. 238, p. 1519–1529, 01 2019.

WÓJCIK, A.; LUKASZEWSKI, R.; KOWALIK, R.; WINIECKI, W. Nonintrusive appliance load monitoring: An overview, laboratory test results and research directions. **Sensors**, MDPI AG, v. 19, n. 16, p. 3621, Aug 2019. ISSN 1424-8220.

WU, Q.; WANG, F. Concatenate convolutional neural networks for non-intrusive load monitoring across complex background. **Energies**, MDPI AG, v. 12, n. 8, p. 1572, Apr 2019. ISSN 1996-1073. Disponível em: <http://dx.doi.org/10.3390/en12081572>.

YANG, C.; SOH, C.; YAP, V. A systematic approach to on-off event detection and clustering analysis of non-intrusive appliance load monitoring. **Frontiers in Energy**, v. 9, 05 2015.

YANG, C. C.; SOH C. S. AND YAP, V. V. A systematic approach in load disaggregation utilizing a multi-stage classification algorithm for consumer electrical appliances classification. **Frontiers in Energy**, Springer, v. 13, n. 2, p. 386–398, 2019.

ZHIBIN, Y.; HONG, C. Research on household appliances recognition method based on data screening of deep learning. **IFAC-PapersOnLine**, Elsevier Ltd, v. 52, n. 24, p. 140–144, 2019.

ZOHA, A.; GLUHAK, A.; IMRAN, M.; RAJASEGARAR, S. Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey. **Sensors**, MDPI AG, v. 12, n. 12, p. 16838–16866, Dec 2012.

**APÊNDICE A – RESUMO DAS CARACTERÍSTICAS EXTRAÍDAS DAS
TRAJETÓRIAS V-I**

Nome	Estado da carga	Fórmula
Área (ar)	Estacionário	$ar = \sum_i \frac{1}{2} V_j - V_i (I_{i'} - I_i + I_{j'} - I_j),$ $i \in \{1, \dots, i_{Vmin} - 1\}$
Direção do loop (lpa)	Estacionário	$lpa = \sum_u \frac{1}{2} (V_{u+1} - V_u)(I_{u+1} - I_u),$ $u \in \{1, \dots, NT - 1\}$
Assimetria (asy)	Estacionário	$asy(A,B) = \max_{a \in A} \left\{ \min_{b \in B} \{d(a,b)\} \right\}$
Curvatura da linha média (r)	Estacionário	$r = corr(x_{M,i}, y_{M,i})$
Auto intersecção (sc)	Estacionário	$sc = \left((a_i \vec{a}_j) \times (a_i \vec{b}_{i'}) \right) \cdot (a_i \vec{a}_j) \times (a_i \vec{b}_{j'}) < 0,$
Pico do segmento médio (mi)	Estacionário	$mi = \max(\max(I_r - f_a(V_r)), \max(I_t - f_b(V_t))),$ $r \in \{m_a, \dots, m_a + n_a - 1\},$ $t \in \{m_b, \dots, m_b + n_b - 1\}$
Inclinação do segmento médio (sh)	Estacionário	$sh = std(I_r) + std(I_t),$
Área dos segmentos direito e esquerdo (alr)	Estacionário	$alr = ar_{direito} + ar_{esquerdo}.$
Corrente atual (itc)	Estacionário	$itc = \max(I) - \min(I)$
Varição da admitância instantânea (D)	Estacionário	$s_i = \frac{I_i}{V_i} \quad i \in \{1, 2, \dots, NT\} \cap V_i > vl,$ $D = std(s)$
Comprimento (len)	Estacionário	$pcom[i] = \ ([V[i] - V[i - 1], I[i] - I[i - 1]])\ ,$ $len = \sum_{i=2}^{NT-1} pcom[i].$
Máximo vetor (md)	Estacionário	$md = \max \left(\sqrt{V_i^2 + I_i^2} \right),$ $i \in \{1, 2, \dots, NT\}$

Sobressinal (<i>ov</i>)	Estacionário/ Transitório	$ov = \max(I_{tran}) - \max(I_{estac}),$
Varição da inclinação do sinal (<i>vss</i>)	Estacionário	$am[i] = \sqrt{(V[i] - V[1])^2 + (I[i] - I[2])^2},$ $i \in \{1, 2, \dots, NT\},$ $f[i] = (am[i] - am[i - 1])(am[i] - am[i + 1]),$ $vss = \sum_{i=1}^{NT} (f[i] > tv)$
Valor médio da inclinação (<i>asl</i>)	Estacionário	$\omega[i] = \frac{I[i] - I[i - 1]}{V[i] - V[i - 1]}, \quad i \in \{2, \dots, NT\},$ $asl = \frac{\sum_{i=2}^{NT} \omega}{NT}$
Distância entre os pontos máximo e mínimo (<i>dbp</i>)	Estacionário	$dbp = \sqrt{(V_{imax} - V_{imin})^2 + (I_{imax} - I_{imin}^2)}$
Ângulo entre os pontos máximo e mínimo (<i>angp</i>)	Estacionário	$angp = \arctan\left(\frac{I_{imax} - I_{imin}}{V_{imax} - V_{imin}}\right)$
Diferença entre as distâncias (<i>dbd</i>)	Estacionário/ Transitório	$dbd = dpb_{tran} - dpb_{estac}$
Diferença entre os ângulos (<i>dbang</i>)	Estacionário/ Transitório	$dbang = angp_{tran} - angp_{estac}$
Distância do centroide até a origem (<i>dc</i>)	Estacionário	$dc = \sqrt{C_{xestac}^2 + C_{yestac}^2}$
Ângulo do centroide (<i>angc</i>)	Estacionário	$angc = \arctan\left(\frac{C_{yestac}}{C_{xestac}}\right)$
Distância entre centroides (<i>dbc</i>)	Estacionário/ Transitório	$dbc = \sqrt{(C_{xtran} - C_{xestac})^2 + (C_{ytran} - C_{yestac})^2}$