

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

Análises Experimentais de Algoritmos de Aprendizagem de
Máquina na Classificação de Distúrbios Elétricos

Reneilson Yves Carvalho Santos

SÃO CRISTÓVÃO/ SE

2019

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

Reneilson Yves Carvalho Santos

Análises Experimentais de Algoritmos de Aprendizagem de
Máquina na Classificação de Distúrbios Elétricos

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação (PROCC) da Universidade Federal de Sergipe (UFS) como parte de requisito para obtenção do título de Mestre em Ciência da Computação.

SÃO CRISTÓVÃO/ SE

2019

Reneilson Yves Carvalho Santos

Análises Experimentais de Algoritmos de Aprendizagem de Máquina na Classificação de Distúrbios Elétricos

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação (PROCC) da Universidade Federal de Sergipe (UFS) como parte de requisito para obtenção do título de Mestre em Ciência da Computação.

BANCA EXAMINADORA

Prof. Dr. Edward David Moreno Ordonez, Presidente
Universidade Federal de Sergipe (UFS)

Prof. Dr. Carlos Alberto Estombelo-Montesco, Co-Orientador
Universidade Federal de Sergipe (UFS)

Prof. Dr. Danton Diego Ferreira, Membro
Universidade Federal de Lavras (UFLA)

Prof. Dr. André Britto de Carvalho, Membro
Universidade Federal de Sergipe (UFS)

Análises Experimentais de Algoritmos de Aprendizagem de Máquina na Classificação de Distúrbios Elétricos

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Sergipe, como parte dos requisitos necessários à obtenção do título de Mestre em Ciência da Computação.

São Cristóvão - SE, 21 de Fevereiro de 2019

Prof. Dr. Edward David Moreno Ordonez
Orientador

Prof. Dr. Carlos Alberto Estombelo-Montesco
Co-Orientadora

Prof. Dr. Danton Diego Ferreira
Membro

Prof. Dr. André Britto de Carvalho
Membro

Dedicatória

Aos meus pais, José Antônio e Josenúzia, meus guias e exemplos por toda a vida.

Agradecimentos

A Deus, pela oportunidade de concluir mais esta etapa da minha vida.

A CAPES/CNPQ, pela bolsa concedida durante o mestrado, essencial à realização deste trabalho.

Aos meus pais e meus irmãos, pela ajuda durante este período na medida de suas possibilidades, em especial Paula, por fazer-se presente em muitos momentos dessa caminhada.

A meu orientador Prof. Dr. Edward David Moreno Ordonez, que vem me ensinando sobre muitas coisas da vida acadêmica, dando-me oportunidades como a de desenvolver de modo autônomo, mas sempre sobre sua preciosa orientação, guiando-me por este caminho tão largo cujo um dos frutos foi este trabalho.

Ao meu coorientador prof. Dr. Carlos Alberto Estombelo-Montesco que foi meu orientador durante quase toda minha graduação e desde lá vem dando suporte na minha jornada acadêmica.

A meu amigo Joel, parceiro de muitas disciplinas e que me auxiliou muito neste trabalho. E ainda Walter e Danilo, que também fizeram parte deste meu caminhar.

A todos aqueles que não citei, mas que à sua maneira contribuíram para a realização deste trabalho.

Resumo

Nos últimos anos, a análise da qualidade da energia elétrica voltou à tona com o avanço em diversas áreas computacionais e os graves problemas relacionados à deterioração desta, podendo causar riscos à vida, redução de vida útil de equipamentos e consequentes impactos ambientais e econômicos. Neste sentido, verificou-se ainda que o interesse em pesquisas relacionadas à classificação dos distúrbios que causam a deterioração da rede elétrica voltou a crescer, principalmente a criação de sistemas passíveis de implementação em ambientes reais. Desta forma, este trabalho buscou verificar, experimentalmente, as mais diversas vertentes relacionadas ao processo de classificação de distúrbios elétricos e criação de sistemas embarcados com este fim. Foi, portanto, analisado experimentalmente a acurácia de algoritmos de classificação (obtendo como melhor resultado a acurácia do Random Forest, tanto para sinais sem ruído quanto para sinais ruidosos), eficiência em ambientes embarcados (obtendo a Árvore de Decisão como melhor eficiência nestes ambientes), e análise de classificação de distúrbios múltiplos (no qual analisou-se 22 distúrbios distintos, com e sem ruído e atingiu-se uma acurácia acima de 90% para ambientes não ruidosos com o Random Forest). Além disso, construiu-se um *benchmark*¹ de distúrbios simulados e disponibilizou-o para a comunidade (assim como seu código fonte em Python), de forma a facilitar a comparação entre diferentes propostas nesta área.

¹<https://drive.google.com/open?id=1suJiEi9KYjLKCBwZHbJMP44Gl2RhXwuC>

Abstract

In the last years, the electric power quality analysis came back to the fore with the advance in many computational areas and the serious issues related to the deterioration of it, that can cause damages to the life, lifetime reduction of the electrical devices, and, consequent environment and economic impacts. In this sense, it was verified that the interest in researches related to the classification of the disturbances that cause the deterioration of the energy quality returned to grow, mainly the creation of systems that can implement it in a real environment. Then, this work aims to verify, experimentally, many strands related to the process of electrical disturbance classification and implementation of embedded systems with the purpose of implement the classification process. Therefore, it was analyzed experimentally the accuracy of different machine learning algorithms (obtaining as a better result the accuracy of the Random forest, both in noisy and not noisy signals), efficiency of the algorithms in embedded environments (in which the Decision Tree was the best), and the analysis of the classification of multiple disturbances (in which 22 disturbances were analyzed and the Random Forest get an accuracy above 90% in not noisy environment). Besides that, it was implemented a benchmark² of simulated disturbances and make it available for the community (as well as the source code in Python), aiming to stay easier make comparisons among different proposes of different authors.

²<https://drive.google.com/open?id=1suJiEi9KYjLKCBwZHbJMP44Gl2RhXwuC>

Lista de Figuras

3.1	Uso de algoritmos de detecção ou detecção e classificação nos 56 estudos selecionados.	13
3.2	Algoritmos de Pré-Processamento	14
3.3	Algoritmos de classificação	17
3.4	Algoritmos de classificação	19
3.5	Monitoramento em Tempo Real e Offline	22
3.6	Tipos de distúrbios estudados nos estudos primários	24
3.7	Número de publicações por ano	29
3.8	Resultado da Classificação para diferentes frequências fundamentais obtidas por [Li et al. 2016]	34
3.9	Metodologia utilizada em [Gonzalez-Cordoba et al. 2015]	35
3.10	Arquitetura utilizada em [Zheng et al. 2014]	36
3.11	Digrama de blocos do sistema proposto em [Carniato et al. 2009]	39
3.12	Digrama de blocos do sistema proposto em [Zhang and Li 2010]	40
3.13	Comparativo entre trabalhos selecionados	43
3.14	Patentes depositadas a cada ano.	44
3.15	Patentes depositadas por ano de acordo com o primeiro depósito.	45
3.16	Países de depósito das patentes encontradas.	46
3.17	Classificação Internacional de Patentes das 7 patentes encontradas.	47
3.18	Classificação completa CIP presente nas patentes selecionadas.	49
5.1	Projeto do Experimento	67
5.2	Operação do Experimento	68
6.1	Pré-Processamento Distúrbios (Experimento 2)	79

6.2	Diagrama de Uso de Técnicas de Pré-Processamento (Experimento 2) . . .	80
6.3	Acurácia Geral (Sem Ruído)	85
6.4	Acurácia (Sem Ruído) - Seno Puro	86
6.5	Acurácia (Sem Ruído) - Interrupção	87
6.6	Acurácia (Sem Ruído) - Afundamento de Tensão	88
6.7	Acurácia (Sem Ruído) - Elevação de Tensão	88
6.8	Acurácia (Sem Ruído) - Oscilação	89
6.9	Acurácia (Sem Ruído) - Flutuação de Tensão	90
6.10	Acurácia (Sem Ruído) - Harmônicas	90
6.11	Acurácia (Sem Ruído) - <i>Notching</i>	91
6.12	Acurácia (Sem Ruído) - Transiente Impulsivo	92
6.13	Acurácia (Sem Ruído) - Nível DC	92
6.14	Comportamento da Acurácia dos Classificadores Sob Efeito de Ruído . . .	93
6.15	Diagrama de Uso de Técnicas de Pré-Processamento (Experimento 3) . . .	95
6.16	<i>Box-plot</i> - 30dB	100
6.17	<i>Box-plot</i> - 50dB	101
6.18	<i>Box-plot</i> - 60dB	101
6.19	<i>Box-plot</i> - 70dB	102
6.20	<i>Box-plot</i> - Sem ruído	102
6.21	Gráfico Acurácia Distúrbios	103
6.22	Matriz de Confusão 30dB	103
6.23	Formas de Onda (Seno- <i>Notching</i> -Impulso)	104
6.24	Formas de Onda (<i>Swell-Flicker</i>)	104
6.25	Diagrama de Uso de Técnicas de Pré-Processamento (Experimento 4) . . .	106
6.26	Transformada de Hilbert x Dilatação	113
6.27	Diagrama de Uso de Técnicas de Pré-Processamento (Experimento 5) . . .	114
6.28	<i>Box-plot</i> - Ambiente Ideal	118
6.29	Matriz de Confusão Ambiente Ideal (Experimento 5)	119
6.30	<i>Swell</i> + Harm. x <i>Swell</i> + Harm. + Osc.	120
6.31	<i>Box-plot</i> - SNR de 70dB	120
6.32	Sinal em Ambiente Ideal e a 70dB	121

6.33	<i>Box-plot</i> - SNR de 60dB	122
6.34	<i>Box-plot</i> - SNR de 50dB	122
6.35	<i>Box-plot</i> - SNR de 30dB	123
6.36	Matriz de Confusão - SNR de 30dB (Experimento 5)	124
A.1	Transiente Impulsivo	148
A.2	Transiente Oscilatório	149
A.3	Interrupção	150
A.4	Afundamento de Tensão	151
A.5	Elevação de Tensão	152
A.6	Harmônicas	156
A.7	<i>Notching</i>	158
A.8	Ruído	159
A.9	Flutuação de Tensão	160
A.10	Variação na Frequência	161
C.1	Modelo em Diagrama de Blocos de Aprendizagem Supervisionada (Adap- tado de [Haykin et al. 2009])	167
C.2	Representação Neurônio Biológico (Fonte: [Silva et al. 2010])	168
C.3	Topologia da Rede Perceptron Multicamadas (Adaptação de: [Haykin et al. 2009])	169
C.4	(a) Filtro Passa Alta Ideal (b) Filtro Passa Baixa Ideal (Fonte: [Theodoridis 2008])	175
C.5	Função Escalar da db4 (Fonte: [Wasilewski])	176
C.6	Função Escalar da db4 (Fonte: [Wasilewski])	176

Lista de Tabelas

2.1	Variações de Longa Duração no Valor RMS (Fonte: [Association et al. 2009])	7
2.2	Comparativo de Sistemas Embarcados	9
3.1	Algoritmos de Pré-Processamento	15
3.2	Algoritmos de Classificação	18
3.3	Plataformas de Hardware utilizadas nas referências	20
3.4	Tipo de monitoramento utilizado nas referências	22
3.5	Distúrbios Elétricos por Referência	25
3.6	Comparativo Entre Tempo no DSP e no PC (Adaptado de: [Salem et al. 2007a])	41
3.7	Classificação internacional de patentes	48
3.8	Métodos utilizados nas patentes encontradas	50
3.9	Implementação abordada nas patentes	52
3.10	Abordagem utilizada nas patentes	53
4.1	Distúrbios	59
4.2	Variáveis	59
4.3	Dados Nas Pastas Múltiplos e Isolados	61
4.4	Análise Comparativa	63
6.1	Tabela hiperparâmetros	73
6.2	Média e Desvio das Métricas Avaliadas	74
6.3	Médias e Desvio do <i>Random Forest</i> com Nove Árvores	74
6.4	Classificação do melhor algoritmo de acordo com a acurácia média	75
6.5	Classificação dos algoritmos em termo das métricas de eficiência	76
6.6	Hiper-parâmetros Algoritmos MLP	81

6.7	Hiper-parâmetros Algoritmos SVM	82
6.8	Melhores Algoritmos para Cada Distúrbio	86
6.9	Características	97
6.10	Parâmetros <i>Random Forest</i>	99
6.11	Teste Z para Acurácia Média	99
6.12	Tabela hiperparâmetros da Árvore de Decisão	107
6.13	Média e Desvio das Métricas Avaliadas	108
6.14	Teste Z para Acurácia Média	108
6.15	Tabela hiperparâmetros <i>Random Forest</i>	115
6.16	Acurácia do Classificador Final	115
6.17	Comparação Trabalhos Relacionados	125
A.1	Variações de Longa Duração no Valor RMS (Fonte: [Association et al. 2009])	153
E.1	Total de patentes selecionadas com critérios aplicados em cada base	187

Lista de Siglas

QEE - Qualidade da Energia Elétrica

RNA - Redes Neurais Artificiais

FPGA - *Field Programmable Gate Array*

ASIC - *Application Specific Integrated Circuit*

SVM - *Support Vector Machine*

MLP - *Multilayer Perceptron*

ICA - *Independent Component Analysis*

PCA - *Principal Component Analysis*

FFT - *Fast Fourier Transform*

DWT - *Discrete Wavelet Transform*

IEEE - *Institute of Electrical and Electronics Engineers*

Sumário

1	Introdução	1
1.1	Objetivo da Dissertação	4
1.2	Organização da Proposta	5
2	Fundamentação Teórica	7
2.1	Qualidade da Energia Elétrica	7
2.2	Plataformas de Hardware	9
2.3	Aprendizagem de Máquina	10
3	Trabalhos Relacionados	12
3.1	Análise dos Resultados do Mapeamento do Estado da Arte	13
3.1.1	Quais os algoritmos utilizados para detecção e/ou classificação dos distúrbios elétricos?	13
3.1.2	Quais as plataformas de hardware utilizadas na detecção e/ou classificação de distúrbios elétricos?	19
3.1.3	Os sistemas de monitoramento de distúrbios elétricos trabalham em tempo real ou offline?	21
3.1.4	Quais os tipos de distúrbios elétricos mais abordados nos sistemas propostos?	24
3.1.5	Considerações Finais	29
3.2	Resumo	30
3.2.1	Rough-set-based feature selection and classification for power quality sensing device employing correlation techniques	30

3.2.2	Real-time system for automatic classification of power quality disturbances	32
3.2.3	Detection and Classification of Power Quality Disturbances Using Double Resolution S-Transform and DAG-SVMs	33
3.2.4	Automatic detection and classification of electrical disturbances by means of empirical mode decomposition method	34
3.2.5	An embedded-based distributed private cloud: Power quality event classification	36
3.2.6	A real-time power quality disturbances classification using hybrid method based on s-transform and dynamics	37
3.2.7	Development of hardware and software for three-phase power quality disturbances detection, classification and diagnosis using Kalman Filter theory	38
3.2.8	DSP-FPGA based real-time power quality disturbances classifier . .	39
3.2.9	Power quality disturbance detection using DSP based continuous wavelet transform	40
3.2.10	Considerações Finais	41
3.3	Análise de Resultados Mapeamento Estado da Técnica	44
3.3.1	Considerações Finais	54
3.4	Considerações Finais do Capítulo	54
4	SED-Bench: <i>benchmark</i> de Distúrbios Elétricos Simulados	56
4.1	Introdução	56
4.2	Metodologia	57
4.2.1	Desenvolvimento do <i>Benchmark</i>	58
4.3	SED-Bench	60
4.3.1	Organização	60
4.3.2	Comparativo	62
4.4	Conclusão do Capítulo	62
5	Metodologia Experimental	64
5.1	Definição do Experimento	65

5.1.1	Definição de Objetivo	65
5.1.2	Planejamento	65
5.2	Operação do Experimento	67
6	Experimentos	70
6.1	Análise da Classificação de Distúrbios Elétricos em Raspberry Pi (Experimento 1)	71
6.1.1	Formulação de Hipótese	72
6.1.2	Resultados	73
6.1.3	Conclusão do Experimento	77
6.2	Algoritmos para Classificação de Distúrbios Elétricos Ruidosos sem Extração de Características (Experimento 2)	78
6.2.1	Formulação de Hipótese	80
6.2.2	Resultados	81
6.2.3	Conclusão do Experimento	94
6.3	Classificação de Distúrbios Elétricos usando Extração de Características (Experimento 3)	94
6.3.1	Formulação de Hipótese	96
6.3.2	Resultados	99
6.4	Análise de Tempo de Processamento de Sistema de Monitoramento em No-deMCU e Raspberry Pi (Experimento 4)	105
6.4.1	Formulação de Hipótese	106
6.4.2	Resultados	107
6.4.3	Conclusão do Experimento	111
6.5	Classificação de Distúrbios Isolados e Múltiplos (Experimento 5)	112
6.5.1	Formulação de Hipótese	114
6.5.2	Resultados	115
6.5.3	Conclusão do Experimento	126
7	Conclusão	127
	Referências	129

A	Qualidade da Energia Elétrica	146
A.1	Transientes	147
A.2	Transiente Impulsivo	147
A.3	Transiente Oscilatório	148
A.4	Variações de Curta Duração no Valor RMS	149
A.4.1	Interrupção	149
A.4.2	Afundamento de Tensão	151
A.4.3	Elevação de Tensão	152
A.5	Variações de Longa Duração no Valor RMS	153
A.6	Desbalanceamento	154
A.7	Distorções na Forma de Onda	155
A.7.1	Nível DC	155
A.7.2	Harmônicas	156
A.7.3	Inter-harmônicas	157
A.7.4	<i>Notching</i>	158
A.7.5	Ruído	159
A.8	Flutuações de Tensão	160
A.9	Variações de Frequência	161
B	Sistemas Embarcados	162
B.1	Sistemas Embarcados Microprocessados	163
B.2	Sistemas Embarcados Microcontrolados	164
B.3	ASICs e FPGAs	164
C	Aprendizagem de Máquina	166
C.1	Aprendizagem Supervisionada	166
C.1.1	Perceptron Multicamadas (MLP)	167
C.1.2	Máquina de Vetor de Suporte (SVM)	171
C.1.3	Árvore de Decisão (CART)	173
C.1.4	Random Forest	173
C.2	Algoritmos de Pré-Processamento de Dados	174
C.2.1	Transformada de Wavelet Discreta	174

C.2.2	Transformada de Fourier Discreta	176
C.2.3	Filtro <i>Notch</i>	177
C.2.4	Transformada de Hilbert	178
C.2.5	Filtro Gaussiano	178
C.2.6	Estatística Descritiva	179
D	Mapeamento Sistemático da Literatura	181
D.1	Questões de Pesquisa	181
D.2	Estratégia de Busca e Seleção	182
D.3	Critérios de Seleção	183
E	Mapeamento de Patentes	184
E.1	Metodologia	185
E.1.1	Questões de Pesquisa	185
E.1.2	Estratégia de Busca e Seleção	185
E.1.3	Critérios de Seleção	186

Capítulo 1

Introdução

A Qualidade da Energia Elétrica (QEE) é a habilidade de gerar ondas senoidais puras, com mínima variação ($\pm 5\%$) na frequência fundamental (50Hz ou 60Hz, a depender) [Kennedy 2000]. No entanto, muitos fatores podem afetar a QEE nas redes elétricas de distribuição ou de transmissão, como, por exemplo, a crescente poluição eletromagnética devido à difusão de sistemas semicondutores [Landi and Luiso 2008], a presença de tempestades com raios que afetem o sistema, ou ainda a geração distribuída de energia elétrica, um conceito novo surgido com o crescente aumento de fontes alternativas de energia [Gupta and Seethalekshmi 2019].

Frente a este quadro, há uma preocupação em relação à QEE por parte de setores de produção e outros que dependam da energia elétrica para um bom funcionamento, como hospitais, servidores, comércios, indústrias, redes de telecomunicação, entre outros [Edomah 2009]. Sabe-se que, nestes ambientes, qualquer distúrbio elétrico pode acarretar problemas irreversíveis, e até vitais; em hospitais, por exemplo, a interrupção elétrica (mesmo momentânea) pode acarretar em perda de dados, e, conseqüentemente, afetar diretamente tratamentos [Corner et al. 1993].

Além disso, o custo de correção nos equipamentos geralmente são mais baixos quando tomadas as devidas providências o mais rápido possível [Edomah 2009].

Entre os impactos causados pelos distúrbios a equipamentos industriais estão o encurtamento da vida útil, que leva à necessidade de um maior capital investido para compra de outros que possam suprir os deteriorados; o mal funcionamento, podendo gerar perda da qualidade no produto final; e até mesmo a perda imediata, ocasionando parada no processo

de produção [Edomah 2009]. Tudo isso leva ao aumento de custos dentro de ambientes industriais.

Nesse sentido, a pesquisa apontada em [Samotyj 2001] corrobora com o fato de estar havendo custo adicional nas indústrias devido a problemas relacionados à QEE. Segundo a pesquisa, houve, na década de 90, nos Estados Unidos, um gasto médio superior a 45,7 bilhões de dólares por ano devido a distúrbios elétricos somente no setor de indústria e comércio. Sendo que, consoante [Seymour and Horsley 2001, Garcia et al. 2009], muitos dos distúrbios são ocasionados pelos próprios equipamentos eletro-eletrônicos presentes nestes setores.

Está, portanto, havendo um aumento de interesse em pesquisas relacionadas à QEE [Association et al. 2009], inclusive no que diz respeito à classificação e detecção dos distúrbios elétricos [Bollen et al. 2010]. Bollen et al (2010) afirmam que, somente no ano 2007, mais de 800 artigos foram publicados nesta área.

Além disso, com o avanço das pesquisas a respeito de smartgrids [Farhangi 2010, Fang et al. 2012] tem crescido também o interesse a respeito de aplicações que possam auxiliar no melhoramento da qualidade da energia elétrica, e, como os próprios smartgrids já possuem sistemas que permitem tal adaptação, a produção de sistemas capazes de comunicarem-se com eles para poder melhorar toda a rede tem evoluído.

Além disso, como já mencionado anteriormente, o avanço da implementação de redes energéticas que possuam a geração de energia distribuída apesar de trazer vantagens como confiabilidade, segurança, eficiência, baixa perda de potência e baixa emissão de poluentes (devido ao acréscimo de produção de energia por meios alternativos [Guerrero et al. 2010]), elas também trazem problemas relacionados à qualidade da energia elétrica [Gupta and Sethalekshmi 2019].

Inclusive, a introdução dos smartgrids e da geração distribuída de energia elétrica tem gerado problemas que até então não eram observados ou eram pouco estudados na literatura, como a análise de supra-harmônicas e subharmônicas [Mendes et al. 2019], sendo que os trabalhos a respeito do tema são bastante recentes [Sobrinho et al. 2018, Liu et al. 2018, Yalcin et al. 2018, Kachhepati 2016].

Como consequência, muitos sistemas de tempo real foram propostos como forma de monitorar os distúrbios, como apresentado em [Ribeiro et al. 2016, Gonzalez-Cordoba et al.

2015, He et al. 2013, Dalai et al. 2013, Li et al. 2016, Carniato et al. 2009, Sobrinho et al. 2018]. Além destes existem outros trabalhos relacionados à criação ou aplicação de determinadas técnicas de computação inteligente para detectar e/ou classificar os distúrbios elétricos, na busca de maior eficiência e/ou eficácia no processo, como apresentado em [Ferreira 2010, Ferreira et al. 2011, Barbosa and Ferreira 2013, Wang and Chen 2019, Lei et al. 2018, Pinzón and Colomé 2019, Malik et al. 2019], sem necessariamente aplicá-los em um sistema de monitoramento embarcado.

Além disso, no geral, a maioria dos artigos trabalham com distúrbios simulados [Ribeiro et al. 2016, Li et al. 2016, Gonzalez-Cordoba et al. 2015, He et al. 2013, Zhang and Li 2010, Salem et al. 2007a], alguns, no entanto, conseguem trabalhar com sinais gerados através de equipamentos elétricos (como motores, geradores de distúrbios elétricos, etc) [Dalai et al. 2013, Li et al. 2016, Gonzalez-Cordoba et al. 2015, He et al. 2013, Carniato et al. 2009, Zhang and Li 2010, Salem et al. 2007a], e outros poucos utilizam sinais reais [Dalai et al. 2013].

A utilização de dados simulados no problema de detecção de distúrbios elétricos dá-se pelo fato de os autores apresentarem propostas relacionadas exclusivamente ao algoritmo proposto e não a implementação direta do sistema em um ambiente real. Existe ainda uma dificuldade em encontrar dados reais disponibilizados ou formar parcerias com empresas fornecedoras de energia elétrica para consegui-los. Além disso, o pesquisador necessitaria de uma visão de um especialista na área para classificar todos os distúrbios presentes nos dados para posterior utilização.

Adicionalmente, existem poucos trabalhos que disponibilizam os dados utilizados nos experimentos, alguns ainda trabalham com a caracterização dos distúrbios elétricos, mas, não apresentam um conjunto de dados para ser utilizados pelos projetistas de sistemas de detecção de distúrbios elétricos [Rens 2009, McGranaghan 2005, Lassila et al. 2005]. Um dos poucos conjunto de dados com acesso disponibilizado são os dados de distúrbios reais da IEEE [Power and Society 03], entretanto, além de ser uma quantidade baixa (menos de 100 dados), eles não são rotulados.

Dessa forma, em virtude da falta de dados para classificação (sejam simulados, reais ou gerados artificialmente por meio de equipamentos elétricos), do crescente interesse na área devido a utilização de smartgrids e o crescimento da geração de energia elétrica distribuída que traz mais problemas para a QEE, além dos impactos econômico, ambiental e social

causado pelos distúrbios elétricos, buscou-se formas de mitigar este problema analisando diferentes caminhos.

Portanto, enxergou-se a possibilidade de contribuir para a área fazendo análises experimentais dos seguintes quesitos:

- Análise de Algoritmos de Classificação: análise de acurácia e tempo de processamento;
- Análise de Algoritmos de Pré-Processamento: melhorar classificação em ambientes ruidosos;
- Análise de Técnicas de Extração de Características: melhorar classificação em ambientes ruidosos e reduzir o tempo de processamento;
- Aplicação em Sistemas Embarcados: buscar melhor plataforma de hardware em custo/benefício;
- Análise de Efeito de Ruídos nos Sinais: mitigar efeito de ruído na classificação;
- Análise de Classificação de Distúrbios Múltiplos: análise de acurácia;
- Disponibilização de Distúrbios Elétricos Simulados: criação e disponibilização de benchmark com dados simulados;

1.1 Objetivo da Dissertação

Avaliar experimentalmente diferentes formas de aprimorar o monitoramento de distúrbios elétricos utilizando-se de algoritmos de aprendizagem de máquina.

Objetivos Específicos

São objetivos específicos:

- Implementar e disponibilizar *benchmark* de distúrbios elétricos simulados;
- Avaliar algoritmos de classificação na tarefa de detecção de distúrbios elétricos;
- Avaliar algoritmos de pré-processamento para classificação de distúrbios elétricos;

- Avaliar sistemas embarcados com a implementação de algoritmos de classificação de distúrbios elétricos;

1.2 Organização da Proposta

A presente dissertação é subdividida em quatro partes distintas.

A primeira parte consiste na apresentação de um mapeamento da literatura e do estado da técnica sobre a Qualidade da Energia Elétrica, especialmente sobre os distúrbios elétricos e sistemas de monitoramento destes, a partir da qual obteve-se uma compreensão sobre como está o estado da arte (produção de artigos, dissertações, teses e patentes) nesta área de pesquisa.

Em seguida buscou-se uma fundamentação teórica a respeito dos temas que optou-se por abordar na dissertação, como os algoritmos de Aprendizagem de Máquina a serem utilizados, os algoritmos de pré-processamento, de extração de características e as plataformas de hardware.

Consequente ao mapeamento e à fundamentação teórica, realizou-se diversos experimentos com algoritmos de classificação, de pré-processamento e de extração de características. Sendo alguns experimentos focados na implementação dos algoritmos em plataformas de hardware para a verificação de seu comportamento em relação a tempo de processamento, potência consumida e memória utilizada.

Por fim, em paralelo foi desenvolvido um *benchmark* de distúrbios elétricos simulados com base em sua formulação matemática, de forma a proporcionar dados para os experimentos e disponibilizá-los para a comunidade acadêmica.

Em resumo, os capítulos estão organizados da seguinte forma:

- O capítulo 2 apresenta a fundamentação teórica necessária para compreensão do tema (abordando Qualidade da Energia Elétrica, Sistemas Embarcados e Aprendizagem de Máquina);
- O capítulo 3 apresenta os trabalhos correlatos para a solução do problema;
- O capítulo 4 apresenta como deu-se a criação do *benchmark* de distúrbios elétricos simulados.

- O capítulo 5 apresenta qual a metodologia utilizada nos experimentos realizados;
- O capítulo 6 apresenta todos os cinco experimentos realizados na dissertação relacionados a análises da classificação de distúrbios elétricos usando algoritmos de aprendizagem de máquina;
- Por fim, o capítulo 7 apresenta a conclusão do trabalho realizado e trás alguns possíveis trabalhos futuros que podem ser realizados como continuidade do trabalho realizado.

Capítulo 2

Fundamentação Teórica

Neste capítulo são tratados conceitos envolvidos na temática da QEE, de Sistemas Embarcados e de algoritmos e técnicas de detecção e classificação de anomalias. Deve-se ressaltar que não pretende-se apresentar um estudo extensivo sobre todos estes temas, são apresentados apenas os pontos fundamentais para o entendimento do conteúdo a ser abordado nesta dissertação, mais informações a respeito de cada um dos temas pode ser encontrado no apêndice.

2.1 Qualidade da Energia Elétrica

A análise de distúrbios elétricos tornou-se tão imprescindível que foi necessário criar uma padronização como forma de auxiliar a demanda necessária por sistemas de monitoramento, haja visto o aumento de equipamentos e dispositivos que deterioram a QEE.

No que diz respeito aos distúrbios elétricos que influenciam diretamente na QEE, a seguir é apresentada uma tabela, adaptada de [Association et al. 2009], que resume os distúrbios analisados na Seção A.

Tabela 2.1: Variações de Longa Duração no Valor RMS (Fonte: [Association et al. 2009])

Categoria	Frequência	Duração	Magnitude
Transiente Impulsivo	-	-	-

Continua na próxima página

Tabela 2.1 – continuando da página anterior

Categoria	Frequência	Duração	Magnitude
Nanosegundos	5ns	< 50ns	-
Microsegundos	1 μ s	50ns - 1ms	-
Milissegundos	0.1ms	> 1ms	-
Transiente Oscilatório	-	-	-
Baixa Frequência	< 5kHz	< 0.3 - 50ms	0 - 400%
Média Frequência	5 - 500kHz	20 μ s	0 - 800%
Alta Frequência	0.5 - 5MHz	5 μ s	0 - 400%
Afundamento de Tensão	-	-	-
Instantâneo	-	0.5 - 30 ciclos	10% - 90%
Momentâneo	-	30 ciclos - 3s	10% - 90%
Temporário	-	>3s - 1min	10% - 90%
Elevação de Tensão	-	-	-
Instantâneo	-	0.5 - 30 ciclos	110% - 180%
Momentâneo	-	30 ciclos - 3s	110% - 140%
Temporário	-	>3s - 1min	110% - 120%
Afundamento de Tensão	-	-	-
Momentâneo	-	0.5 ciclos - 3s	< 10%
Temporário	-	>3s - 1min	< 10%
Variações de Longa Duração RMS	-	-	-
Interrupção sustentada	-	> 1min	0%
Subtensão	-	> 1min	80% - 90%
Sobretensão	-	> 1min	110% - 120%
Sobrecarga de Corrente	-	> 1min	-
Desbalanceamento	-	-	-
Tensão	-	Estado Estacionário	0 - 0.1%
Corrente	-	Estado Estacionário	1 - 30%
Distorção na Forma de Onda	-	-	-
Nível DC	-	Estado Estacionário	0 - 0.1%

Continua na próxima página

Tabela 2.1 – continuando da página anterior

Categoria	Frequência	Duração	Magnitude
Harmônicas	0 - 9kHz	Estado Estacionário	0 - 20%
Interharmônicas	0 - 9kHz	Estado Estacionário	0 - 2%
<i>Notching</i>	-	Estado Estacionário	0 - 0.1%
Ruído		Estado Estacionário	0 - 1%
Flutuação de Tensão	< 25Hz	Intermitente	-
Variações na Frequência	-	< 10s	$\pm 0.10\text{Hz}$

Tendo em vista, portanto, a necessidade de sistemas que auxiliem a identificar os problemas elétricos, surge como alternativa a utilização de sistemas embarcados, que são mais baratos e facilitam a rápida realização de projetos eletrônicos, além de, como possuem um projeto dedicado a uma função específica, tende a ser mais eficiente energeticamente que um computador de propósito geral.

2.2 Plataformas de Hardware

Existem diversos tipos de plataformas de hardware que podem ser utilizadas para a construção de sistemas embarcados, dentre os quais pode-se citar os sistemas microprocessados, microcontrolados além dos ASICs e FPGAs. Na Tabela 2.2 é apresentado de forma resumida um comparativo entre tais tipos de sistemas embarcados.

Tabela 2.2: Comparativo de Sistemas Embarcados

Tipo de Sistemas	Vantagens	Desvantagens
Microprocessados	- Flexibilidade - Generalização	- Baixa eficiência energética - Necessidade de Periféricos - Alto custo

Continua na próxima página

Tabela 2.2 – continuando da página anterior

Tipo de Sistema	Vantagens	Desvantagens
Microcontrolados	- Baixo Custo - Maior Configurabilidade	- Baixo Poder Computacional - Memória Mais Limitada
ASICs	- Custo Reduz com Produção em Larga Escala	- Aplicação Específica
FPGAs	- Reconfigurabilidade - Paralelismo Inerente - Capacidade de 'Desenhar' Hardware por Software - Alta Velocidade de Processamento - Confiabilidade - Reuso	- Médio Custo

2.3 Aprendizagem de Máquina

Só a utilização de sistemas embarcados não constrói um sistema de monitoramento capaz de realizar a análise dos sinais recebidos e identificar quais são distúrbios e quais não são. É neste ponto que surge a necessidade de algoritmos capazes de aprender e classificar os distúrbios elétricos para que as devidas providências sejam tomadas.

Neste sentido, pode-se utilizar diferentes algoritmos de aprendizagem de máquina (SVM, MLP, *Random Forest*, Redes Bayesianas, Algoritmos de Detecção de Novidades, Árvore de Decisão, CNN, entre outros), baseando-se na forma como os algoritmos pertencentes a cada categoria de aprendizado consegue reconhecer padrões a partir de determinado modelo de aprendizagem.

O Apêndice C apresenta mais detalhadamente o funcionamento dos algoritmos de aprendizagem de máquina utilizados ao longo da dissertação, bem como os algoritmos e técnicas para pré-processamento dos dados e extração de características, que podem ser consideradas

fases no processo de reconhecimento de padrões.

Capítulo 3

Trabalhos Relacionados

Neste capítulo apresenta-se uma visão geral do estado da arte relacionado ao desenvolvimento de sistemas embarcados para o monitoramento da Qualidade de Energia Elétrica (em especial a análise e classificação de distúrbios elétricos).

Neste contexto, apresenta-se um estudo utilizando o método de mapeamento sistemático do estado da arte e da técnica, proposto por [Petersen et al. 2008]. A metodologia seguida é apresentada com maiores detalhes no Apêndice D.

A seguir são apresentados os resultados obtidos por meio dos mapeamentos realizados.

Neste breve estudo utilizaram-se 56 trabalhos relacionados ao tema, no intuito de responder algumas questões de pesquisa que auxiliarão no desenvolvimento da dissertação.

Em seguida, foram escolhidos 9 dos 56 trabalhos para se realizar uma análise mais detalhada. Tais trabalhos foram escolhidos com base no número de distúrbios analisados pelos sistemas propostos.

Por fim, é apresentado um mapeamento, também utilizando o método proposto em [Petersen et al. 2008], extrapolando-o para análise de patentes relacionadas ao tema de sistemas embarcados para o monitoramento de distúrbios elétricos encontradas nas bases *Derwent Innovations Index* e INPI.

3.1 Análise dos Resultados do Mapeamento do Estado da Arte

Utilizando os 56 artigos selecionados após aplicados os critérios de inclusão e exclusão, foi feita a análise detalhada de cada artigo para posterior extração de informações que possam responder às questões de pesquisa apresentadas em D.1, e conseguir, como consequência, uma visão geral a respeito do estado da arte em termos de artigos na área abordada.

A seguir são apresentados os resultados dos estudos relevantes relacionando-os com as questões de pesquisa apresentadas na seção D.1.

3.1.1 Quais os algoritmos utilizados para detecção e/ou classificação dos distúrbios elétricos?

Esta questão visa analisar a tendência da utilização das técnicas de identificação de distúrbios elétricos (detecção ou detecção e classificação). Na Figura 3.1 é apresentado um gráfico o qual ilustra que, apesar da maioria das pesquisas estarem relacionadas ao processo de classificação e detecção, ainda uma parte significativa faz apenas a detecção dos distúrbios elétricos, desta forma, deixando em aberto a possibilidade de trabalhos futuros que permitam realizar a classificação dos distúrbios detectados.

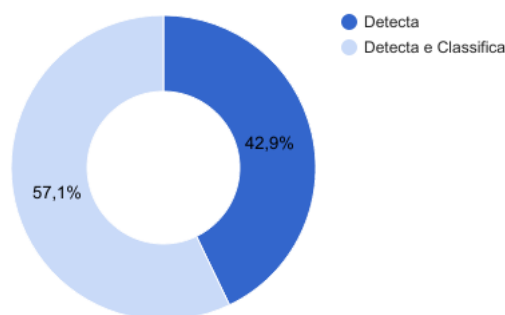


Figura 3.1: Uso de algoritmos de detecção ou detecção e classificação nos 56 estudos selecionados.

Dentro destes algoritmos de detecção e classificação, ver-se-á a seguir quais foram utilizados para a realização destas tarefas, tendo em vista, em sua grande maioria, a implementação em uma plataforma cujo poder de processamento é inferior ao de um ambiente computacional mais robusto.

Como forma de ilustrar a utilização dos algoritmos de pré-processamento, na Figura 3.2 é apresentado um gráfico de frequência trazendo-os.



Figura 3.2: Algoritmos de Pré-Processamento

Como pode-se visualizar na Figura 3.2, o algoritmo mais utilizado para pré-processamento, extração de características e/ou seleção de características dos distúrbios elétricos é a Transformada de Wavelet Contínua, isso deve-se provavelmente ao fato de sua implementação ser facilmente transferida para o ambiente embarcado, mesmo este possuindo uma menor capacidade de processamento. A Tabela 3.1 apresenta os artigos que utilizam os algoritmos apresentados na Figura 3.2.

Tabela 3.1: Algoritmos de Pré-Processamento

Algoritmo	Referência
DFT (Transformada Discreta de Fourier)	[Zhang and Li 2010, Morales-Velazquez et al. 2017, Angrisani et al. 1999, Kapisch et al. 2016]
FFT (Fast Fourier Transform)	[Zheng et al. 2014, He et al. 2013, Salem et al. 2005, Gherasim et al. 2003, Xiang-wu et al. 2006, Yongxian et al. 2011, Wang et al. 2012, Erişti et al. 2013, Yang and Wen 2006, Liang and Da 2011, Amaris et al. 2008]
Wavelet Contínua	[Xu et al. 2014, Salem et al. 2007a, Daponte et al. 2004, Duan and Guo 2012, Choong et al. 2005, Huang et al. 2002, Wang et al. 2012, Wang and Zhu 2002, Kapisch et al. 2016, Gorecki 2014, Salem et al. 2007a, Amaris et al. 2008]
Wavelet Discreta	[Back et al. 2016, Daponte et al. 2004, Arrais et al. 2014, Szweda 2009]
Morfologia Matemática	[Rodriguez et al. 2016, Radil et al. 2007, Ramos et al. 2008]
ICA	[Ruiz-Llata et al. 2011]
Valor RMS	[Zheng et al. 2014, Zhang and Li 2010, Yau et al. 2015, Radil et al. 2007, Ramos et al. 2008, Wang and Zhu 2002, Meena et al. 2009, Meena et al. 2011, Malaysia 2011, Yingkayun et al. 2009]
Chaos Synchronization	[Yau et al. 2015]
Teoria da Extensão	[Yau et al. 2015]
Estatística de Alta Ordem	[Moreno-Garcia et al. 2013, Ribeiro et al. 2016]
Transformada S	[He et al. 2013, Wang et al. 2012, Salem et al. 2007a]
STFT	[Salem et al. 2007a, Moussa et al. 2004]
Continua na próxima página	

Tabela 3.1 – continuando da página anterior

Algoritmo	Referência
Transformada Slunt	[Hsieh et al. 2010]
Transformada de Hilbert	[Granados-Lieberman et al. 2013, Valtierra-Rodriguez et al. 2016]
SVD (Decomposição em valores singulares)	[Wang et al. 2012]
DRST (Double Resolution S Transform)	[Li et al. 2016]
EMD (Empirical Mode Decomposition)	[Gonzalez-Cordoba et al. 2015]
Entropia de Shanon	[Gonzalez-Cordoba et al. 2015]
Filtro de Kalman	[Amaris et al. 2008, Carniato et al. 2009]
Correlação Cruzada	[Dalai et al. 2013]
RST (Rough set)	[Dalai et al. 2013]
Inovation Detection	[Kapisch et al. 2016]
Filtro Adaptativo	[He et al. 2013]
Regularidade	[Artioli et al. 2004]
Single-Phase Rotating Frame	[Sadigh and Smedley 2016, Malaysia 2011]
Transformada d-q	[Montero-Hernandez and Enjeti 2001]
Janela Flutuante	[Meena et al. 2009]
Histerese	[Cuma et al. 2013]
Predição Linear	[Jiasheng and Shijie 2010]
TRF (Time-Frequency Representation)	[Wang et al. 2004]

A maioria dos autores utilizam algoritmos já conhecidos na literatura para o processo de pré-processamento, extração de características e/ou seleção de características, apenas [Artioli et al. 2004] propõe um novo algoritmo baseado na regularidade do sinal de entrada, no

entanto, faz apenas a detecção de distúrbios transientes na onda senoidal de tensão.

Os artigos que apresentam a utilização do valor RMS e morfologia matemática normalmente fazem a detecção de distúrbios com um maior número de características no domínio do tempo (afundamento e elevação de tensão, por exemplo), sendo necessário utilizar outros algoritmos para verificar a existência de distúrbios com características inerentes na frequência (como *notching* e harmônicas).

Os trabalhos que apresentam detecção e classificação utilizam os algoritmos de detecção para reduzir a carga de trabalho no processo de classificação, seja reduzindo a quantidade de dados (como no caso da Wavelet) ou então alterando o domínio no qual o sinal é apresentado do tempo para a frequência (como a FFT).

Em relação aos **algoritmos de classificação**, a Figura 3.3 serve como forma de ilustrar a utilização dos algoritmos de classificação nos 56 trabalhos seleccionados, na qual pode-se observar que os algoritmos baseados em *if-else* (algoritmos que utilizam lógica condicional de acordo com os valores obtidos na fase de detecção), o algoritmo de Redes Neurais Artificiais e o algoritmo de Árvore de Decisão são bastante utilizados.

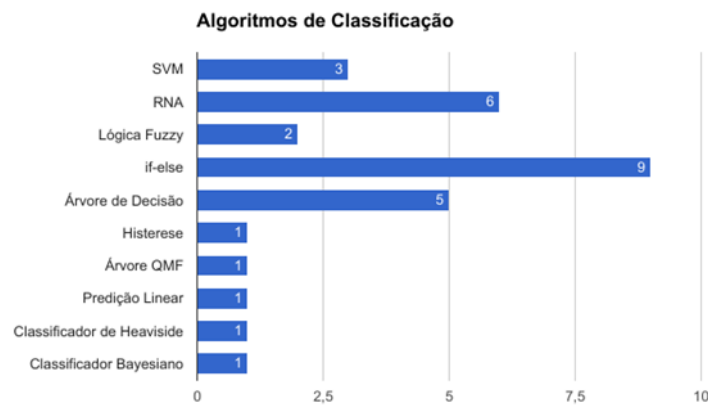


Figura 3.3: Algoritmos de classificação

Baseando-se nos resultados apresentados na Figura 3.3, na Tabela 3.2 é detalhado em qual trabalho cada algoritmo é utilizado.

Tabela 3.2: Algoritmos de Classificação

Algoritmo	Referência
SVM (<i>Support Vector Machine</i>)	[Li et al. 2016, Ruiz-Llata et al. 2011, Wang et al. 2012]
RNA (Perceptron Multicamadas)	[Ribeiro et al. 2016, Gonzalez-Cordoba et al. 2015, Granados-Lieberman et al. 2013, Choong et al. 2005, Daponte et al. 2004]
Lógica Fuzzy	[Choong et al. 2005, Valtierra-Rodriguez et al. 2016]
Lógica Condicional	[Arrais et al. 2014, Moreno-Garcia et al. 2013, Ramos et al. 2008, Radil et al. 2007, Salem et al. 2007a, Wang and Zhu 2002, Jiasheng and Shijie 2010, Huang et al. 2002, Meena et al. 2011, Moreno-Garcia et al. 2013]
Árvore de Decisão	[Ribeiro et al. 2016, Li et al. 2016, Zhang and Li 2010, Wang et al. 2004, Rodriguez et al. 2016]
Histeres	[Cuma et al. 2013]
Árvore QMF	[Daponte et al. 2004]
Predição Linear	[Jiasheng and Shijie 2010]
Classificador de Heaviside	[Wang et al. 2004]
Classificador Bayesiano	[Ribeiro et al. 2016]

Como pode-se ver pelo quadro, apesar das propostas relativas a distúrbios elétricos terem começado a partir da década de 90, somente em 2001 Montero-Hernandez e Enjeti (2001) fizeram um sistema embarcado que tinha capacidade de realizar a classificação dos distúrbios elétricos, neste caso utilizou-se uma abordagem baseada em lógica condicional sobre os valores obtidos com a Transformada d-q, sendo possível, portanto, classificar os distúrbios em afundamento de tensão, elevação de tensão e interrupção.

Desde então, outros autores propuseram diferentes abordagens para realizar a classificação como em [Ribeiro et al. 2016], no qual é utilizado a árvore de decisão em conjunto com o algoritmo de Redes Neurais Perceptron.

Neste contexto, dada a diversidade de algoritmos utilizados tanto na classificação como na detecção dos distúrbios elétricos, a seguir são apresentados em quais plataformas tais algoritmos são implementados.

3.1.2 Quais as plataformas de hardware utilizadas na detecção e/ou classificação de distúrbios elétricos?

A resposta para esta pergunta é ilustrada na Figura 3.4, que apresenta um gráfico de frequência com os sistemas embarcados utilizados nos sistemas de monitoramento da qualidade da energia elétrica. A partir desta figura, pode-se observar que os pesquisadores trabalham, em sua maioria, com DSP (*digital signal processor*) e FPGA (*field programmable gate array*) como plataforma para implementação dos algoritmos de detecção e classificação.

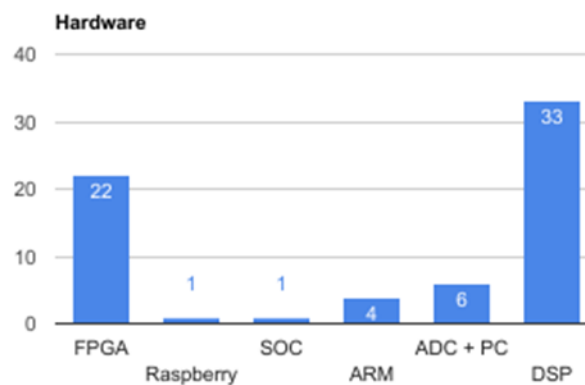


Figura 3.4: Algoritmos de classificação

É digno de nota que, em todas as propostas vistas na literatura, foi utilizado um circuito ADC (conversor analógico-digital) para fazer a captura dos sinais dos sensores conectados à rede elétrica (sejam os que fazem experimento controlado ou em ambiente real), não obstante, no gráfico é apresentado com rótulo "ADC + PC" apenas os trabalhos que utilizam somente um circuito ADC na forma de hardware "embarcado" e aplicam o algoritmo de classificação e/ou detecção em microcomputador.

Alguns trabalhos apresentam ainda a combinação de FPGA e DSP ou ARM e DSP aproveitando melhor as qualidades de cada dispositivo.

Pode-se salientar também que apenas em [Zheng et al. 2014] é utilizado Raspberry (com

o algoritmo FFT, cálculo do valor RMS e uma estrutura do tipo *if-else*). Já em [Yau et al. 2015] (que utiliza um algoritmo baseado na teoria do caos para detecção), é apresentada a possibilidade de trabalhar com o algoritmo em um SoC (*System on Chip*), apesar de detectar apenas quatro tipos de distúrbios isolados (afundamento e elevação de tensão, interrupção e harmônicas). A Tabela 3.3 apresenta as referências nas quais cada uma das plataformas apresentadas no gráfico da Figura 3.4 são utilizados.

Tabela 3.3: Plataformas de Hardware utilizadas nas referências

Hardware	Referências
FPGA	[Arrais et al. 2014, Choong et al. 2005, Erişti et al. 2013, Gonzalez-Cordoba et al. 2015, Granados-Lieberman et al. 2013, He et al. 2013, Hsieh et al. 2010, Huang et al. 2002, Kapisch et al. 2016, Kapisch et al. 2014, Liang and Da 2011, Meena et al. 2009, Morales-Velazquez et al. 2017, Moreno-Garcia et al. 2013, Ribeiro et al. 2016, Ribeiro et al. 2018, Rodriguez et al. 2016, Ruiz-Llata et al. 2011, Valtierra-Rodriguez et al. 2016, Xu et al. 2014, Yang and Wen 2006, Zhang and Li 2010]
Raspberry	[Zheng et al. 2014]
SoC	[Yau et al. 2015]
ARM	[Duan and Guo 2012, Lu et al. 2008, Wang et al. 2012, Yang and Wen 2006]
ADC + PC	[Back et al. 2016, Daponte et al. 2004, Kapisch et al. 2014, Moussa et al. 2004, Radil et al. 2007, Ramos et al. 2008]
Continua na próxima página	

Tabela 3.3 – continuando da página anterior

Hardware	Referências
DSP	[Amaris et al. 2008, Angrisani et al. 1999, Arrais et al. 2014, Artioli et al. 2004, Back et al. 2016, Carniato et al. 2009, Cuma et al. 2013, Dalai et al. 2013, Duan and Guo 2012, Gherasim et al. 2003, Gorecki 2014, He et al. 2013, Kapisch et al. 2016, Jiasheng and Shi-jie 2010, Li et al. 2016, Lu et al. 2008, Malaysia 2011, Meena et al. 2011, Montero-Hernandez and Enjeti 2001, Sadigh and Smedley 2016, Salem et al. 2005, Salem et al. 2006, Salem et al. 2007a, Salem et al. 2007b, Szweda 2009, Wang and Zhu 2002, Wang et al. 2004, Wang et al. 2012, Xiang-wu et al. 2006, Xu et al. 2014, Yingkayun et al. 2009, Yongxian et al. 2011, Zhang and Li 2010]

Além de saber da forma como os distúrbios são detectados e/ou classificados (qual a plataforma utilizada), é importante também entender como se dá o processo de detecção e/ou classificação, ou seja, se a tendência é se ter um trabalho cujo tempo de execução é suficiente para se ter uma resposta em tempo real ou se o foco é obter os dados para posteriormente trabalhar em cima deles. É sobre isso que se trata a terceira questão do presente estudo.

3.1.3 Os sistemas de monitoramento de distúrbios elétricos trabalham em tempo real ou offline?

A Figura 3.5 apresenta um gráfico com o percentual de trabalhos que fazem o monitoramento em tempo real e de maneira offline.

Como pode-se analisar, a grande maioria dos trabalhos são realizados para obter-se um monitoramento em tempo real, apenas cinco estudos primários ([Yang and Wen 2006, Back et al. 2016, Kapisch et al. 2016, Liang and Da 2011, Salem et al. 2006]) apresentaram uma abordagem de classificação *offline*. Na Tabela 3.4 é apresentado o resultado completo, com

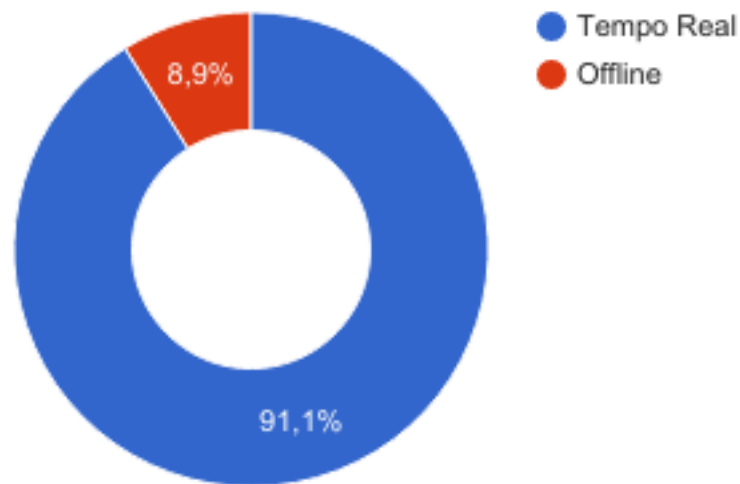


Figura 3.5: Monitoramento em Tempo Real e Offline

os artigos onde são apresentados sistemas que trabalham em tempo real e de maneira offline.

Tabela 3.4: Tipo de monitoramento utilizado nas referências

Monitoramento	Referências
Offline	[Yang and Wen 2006, Back et al. 2016, Kapisch et al. 2016, Liang and Da 2011, Salem et al. 2006]
Continua na próxima página	

Tabela 3.4 – continuando da página anterior

Monitoramento	Referências
Tempo Real	[Angrisani et al. 1999, Arrais et al. 2014, Artioli et al. 2004, Carniato et al. 2009, Choong et al. 2005, Cuma et al. 2013, Dalai et al. 2013, Daponte et al. 2004, Duan and Guo 2012, Erişti et al. 2013, Gherasim et al. 2003, Gonzalez-Cordoba et al. 2015, Gorecki 2014, Granados-Lieberman et al. 2013, He et al. 2013, Hsieh et al. 2010, Huang et al. 2002, Jiasheng and Shijie 2010, Kapisch et al. 2014, Li et al. 2016, Lu et al. 2008, Malaysia 2011, Meena et al. 2009, Meena et al. 2011, Montero-Hernandez and Enjeti 2001, Morales-Velazquez et al. 2017, Moreno-Garcia et al. 2013, Moussa et al. 2004, Radil et al. 2007, Ramos et al. 2008, Ribeiro et al. 2016, Ribeiro et al. 2018, Rodriguez et al. 2016, Ruiz-Llata et al. 2011, Sadigh and Smedley 2016, Salem et al. 2005, Salem et al. 2007a, Salem et al. 2007b, Szweda 2009, Valtierra-Rodriguez et al. 2016, Wang and Zhu 2002, Wang et al. 2004, Wang et al. 2012, Xiang-wu et al. 2006, Xu et al. 2014, Yau et al. 2015, Yingkayun et al. 2009, Yongxian et al. 2011, Zhang and Li 2010, Zheng et al. 2014]

Por fim, também é importante compreender quais são os distúrbios mais capturados pelos sistemas de monitoramento propostos nos trabalhos estudados e é sobre isso que se trata a última pergunta do mapeamento realizado.

3.1.4 Quais os tipos de distúrbios elétricos mais abordados nos sistemas propostos?

A Figura 3.6 apresenta um gráfico que traz a frequência com a qual os distúrbios são trabalhados nos 56 estudos analisados.

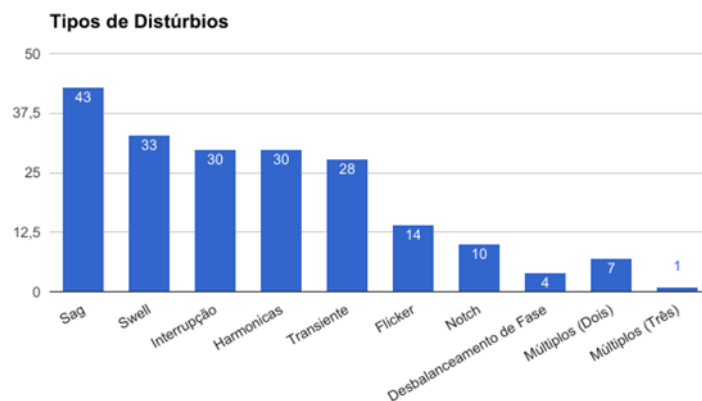


Figura 3.6: Tipos de distúrbios estudados nos estudos primários

Como pode-se perceber a partir do gráfico apresentado na Figura 3.6, o tipo de distúrbio mais comumente estudado é o *sag* (afundamento de tensão), seguido pelo *swell* (elevação de tensão). Interrupção, harmônicas e transientes são outros tipos de distúrbios bastante analisados nos sistemas estudados, inclusive, *sag* e harmônicas possuem uma grande quantidade de estudos primários que tratam unicamente deles, como em [Amaris et al. 2008, Arrais et al. 2014, Cuma et al. 2013] que detectam apenas distúrbios de afundamento de tensão e em [Erişti et al. 2013, Gherasim et al. 2003, Gorecki 2014, Wang et al. 2012] que trazem estudos focados apenas em detecção de harmônicas ou cálculo de TDH (taxa de distorção harmônica).

Poucos trabalhos apresentam a classificação de distúrbios que acontecem simultaneamente na rede elétrica (representados como "Múltiplos (Dois)" e "Múltiplos (Três)" no gráfico da Figura 3.6), são apenas sete trabalhos que apresentam soluções para tais problemas sendo que em [Ribeiro et al. 2016] é apresentada uma solução para dez distúrbios duplos (transiente oscilatório com *sag* e com *swell*, harmônicas com *sag*, *swell*, transiente oscilatório e *notching*, *notching* com *sag* e com *swell*, e transiente impulsivo com *sag* e com *swell*) e quatro triplos (harmônicas, oscilatório e *sag*, harmônicas, oscilatório e *swell*, harmôni-

cas, *notching* e *sag*, harmônicas, *notching* e *swell*). Já em [Dalai et al. 2013, Ribeiro et al. 2016, Li et al. 2016, He et al. 2013, Ruiz-Llata et al. 2011, Zhang and Li 2010, Xiang-wu et al. 2006, Yang and Wen 2006, Salem et al. 2007a] são apresentadas soluções para detecção e classificação de distúrbios duplos, sendo *sag* com harmônicas e *swell* com harmônicas os distúrbios duplos mais comuns detectados nos sistemas desenvolvidos nestes trabalhos. Na Tabela 3.5 são apresentadas todas as referências e quais distúrbios são analisados em cada uma delas.

Tabela 3.5: Distúrbios Elétricos por Referência

Distúrbio Elétrico	Referências
<i>Sag</i>	[Dalai et al. 2013, Ribeiro et al. 2016, Back et al. 2016, Li et al. 2016, Gonzalez-Cordoba et al. 2015, Zheng et al. 2014, Arrais et al. 2014, He et al. 2013, Granados-Lieberman et al. 2013, Carniato et al. 2009, Zhang and Li 2010, Salem et al. 2007a, Rodriguez et al. 2016, Wang et al. 2004, Morales-Velazquez et al. 2017, Yau et al. 2015, Moreno-Garcia et al. 2013, Duan and Guo 2012, Angrisani et al. 1999, Salem et al. 2007b, Hsieh et al. 2010, Ramos et al. 2008, Yongxian et al. 2011, Choong et al. 2005, Radil et al. 2007, Huang et al. 2002, Lu et al. 2008, Valtierra-Rodriguez et al. 2016, Wang et al. 2012, Cuma et al. 2013, Wang and Zhu 2002, Sadigh and Smedley 2016, Kapisch et al. 2016, Montero-Hernandez and Enjeti 2001, Meena et al. 2009, Malaysia 2011, Meena et al. 2011, Jiasheng and Shijie 2010, Salem et al. 2006, Moussa et al. 2004, Yingkayun et al. 2009, Amaris et al. 2008]
Continua na próxima página	

Tabela 3.5 – continuando da página anterior

Distúrbio Elétrico	Referências
<i>Swell</i>	[Dalai et al. 2013, Ribeiro et al. 2016, Back et al. 2016, Li et al. 2016, Gonzalez-Cordoba et al. 2015, Zheng et al. 2014, He et al. 2013, Granados-Lieberman et al. 2013, Carniato et al. 2009, Zhang and Li 2010, Salem et al. 2007a, Rodriguez et al. 2016, Morales-Velazquez et al. 2017, Yau et al. 2015, Moreno-Garcia et al. 2013, Duan and Guo 2012, Salem et al. 2007b, Hsieh et al. 2010, Ramos et al. 2008, Choong et al. 2005, Radil et al. 2007, Huang et al. 2002, Lu et al. 2008, Valtierra-Rodriguez et al. 2016, Wang et al. 2012, Wang and Zhu 2002, Kapisch et al. 2016, Montero-Hernandez and Enjeti 2001, Meena et al. 2009, Meena et al. 2011, Jiasheng and Shijie 2010, Moussa et al. 2004]
Interrupção	[Dalai et al. 2013, Ribeiro et al. 2016, Li et al. 2016, Gonzalez-Cordoba et al. 2015, Zheng et al. 2014, He et al. 2013, Granados-Lieberman et al. 2013, Carniato et al. 2009, Zhang and Li 2010, Salem et al. 2007a, Daponte et al. 2004, Rodriguez et al. 2016, Morales-Velazquez et al. 2017, Yau et al. 2015, Duan and Guo 2012, Angrisani et al. 1999, Salem et al. 2007b, Hsieh et al. 2010, Ramos et al. 2008, Choong et al. 2005, Radil et al. 2007, Huang et al. 2002, Lu et al. 2008, Valtierra-Rodriguez et al. 2016, Wang et al. 2012, Wang and Zhu 2002, Montero-Hernandez and Enjeti 2001, Moussa et al. 2004, Yingkayun et al. 2009]
Continua na próxima página	

Tabela 3.5 – continuando da página anterior

Distúrbio Elétrico	Referências
Harmônicas	[Dalai et al. 2013, Ribeiro et al. 2016, Back et al. 2016, Li et al. 2016, Gonzalez-Cordoba et al. 2015, Zheng et al. 2014, Xu et al. 2014, He et al. 2013, Granados-Lieberman et al. 2013, Carniato et al. 2009, Zhang and Li 2010, Salem et al. 2005, Gherasim et al. 2003, Wang et al. 2004, Morales-Velazquez et al. 2017, Yau et al. 2015, Moreno-Garcia et al. 2013, Xiang-wu et al. 2006, Hsieh et al. 2010, Ramos et al. 2008, Yongxian et al. 2011, Choong et al. 2005, Radil et al. 2007, Gorecki 2014, Erişti et al. 2013, Sadigh and Smedley 2016, Kapisch et al. 2016, Salem et al. 2006, Liang and Da 2011]
Transiente Oscilatório	[Dalai et al. 2013, Ribeiro et al. 2016, Back et al. 2016, Li et al. 2016, Gonzalez-Cordoba et al. 2015, Zheng et al. 2014, Xu et al. 2014, He et al. 2013, Carniato et al. 2009, Zhang and Li 2010, Salem et al. 2007a, Salem et al. 2005, Daponte et al. 2004, Moreno-Garcia et al. 2013, Duan and Guo 2012, Angrisani et al. 1999, Salem et al. 2007b, Ramos et al. 2008, Yongxian et al. 2011, Choong et al. 2005, Radil et al. 2007, Szweda 2009, Lu et al. 2008, Wang et al. 2012, Kapisch et al. 2014, Artioli et al. 2004, Salem et al. 2006]
Continua na próxima página	

Tabela 3.5 – continuando da página anterior

Distúrbio Elétrico	Referências
<i>Flicker</i>	[Dalai et al. 2013, Ribeiro et al. 2016, Back et al. 2016, Li et al. 2016, Gonzalez-Cordoba et al. 2015, Zheng et al. 2014, Xu et al. 2014, He et al. 2013, Carniato et al. 2009, Zhang and Li 2010, Salem et al. 2007a, Salem et al. 2005, Daponte et al. 2004, Moreno-Garcia et al. 2013, Duan and Guo 2012, Angrisani et al. 1999, Salem et al. 2007b, Ramos et al. 2008, Yongxian et al. 2011, Choong et al. 2005, Radil et al. 2007, Szweda 2009, Lu et al. 2008, Wang et al. 2012, Kapisch et al. 2014, Artioli et al. 2004, Salem et al. 2006]
<i>Notch</i>	[Gonzalez-Cordoba et al. 2015, He et al. 2013, Zhang and Li 2010, Salem et al. 2007a, Angrisani et al. 1999, Salem et al. 2007b, Szweda 2009, Kapisch et al. 2016, Salem et al. 2006]
<i>Desbalanceamento de Fase</i>	[Xiang-wu et al. 2006, Carniato et al. 2009, Xu et al. 2014, Yang and Wen 2006]
<i>Múltiplos (2x)</i>	[Dalai et al. 2013, Ribeiro et al. 2016, Li et al. 2016, He et al. 2013, Ruiz-Llata et al. 2011, Zhang and Li 2010, Xiang-wu et al. 2006, Yang and Wen 2006, Salem et al. 2007a]
<i>Múltiplos (3x)</i>	[Ribeiro et al. 2016]

É interessante observar também como deu-se o avanço no número de pesquisas nesta área nos últimos anos, com base nos artigos selecionados. O gráfico de frequência mostrado na Figura 3.7 apresenta este desenvolvimento de pesquisas relacionadas a sistemas embarcados no monitoramento de distúrbios elétricos.



Figura 3.7: Número de publicações por ano

Como pode-se visualizar, as pesquisas na área vinham crescendo e teve uma queda no ano de 2015, com apenas quatro pesquisas relacionadas, mas, voltando a crescer no ano de 2016.

3.1.5 Considerações Finais

A partir da análise dos estudos primários pode-se encontrar as tendências e pontos a serem desenvolvidos na área de classificação e detecção de distúrbios elétricos.

Descobriu-se que a maioria dos sistemas de monitoramento fazem a classificação e detecção dos distúrbios elétricos, sendo os algoritmos utilizados bastante diversificados, tendo destaque para a Transformada de Wavelet, que possui a vantagem de detectar problemas no domínio do tempo e da frequência, além de permitir a redução (compressão) dos dados (Q1).

Além disso, a maioria dos sistemas propostos trabalham em tempo real (Q3) e, quase todos apresentam a detecção do distúrbio do tipo *sag*, todavia poucos trabalhos verificam a ocorrência de distúrbios múltiplos (ou simultâneos) na rede elétrica (Q4).

Por fim, verificou-se que os sistemas embarcados são criados, em sua maioria, nas plataformas DSP (*digital signal processor*) e FPGA (*field programmable gate array*) (Q2).

3.2 Resumo

Como visto na seção anterior (D), existe na literatura uma grande quantidade de artigos relacionados à área de desenvolvimento de sistemas monitores de distúrbios elétricos (ou seja, que analisam a Qualidade da Energia Elétrica). Esta Seção busca fazer uma análise mais detalhada dos principais artigos vistos nesses estudos analisados, de forma a possuir uma base comparativa a respeito da proposta de dissertação apresentada.

Como critério de escolha dos artigos a serem abordados, escolheu-se aqueles que possuíam uma maior quantidade de distúrbios analisados. Desta forma, nove estudos foram selecionados ([Dalai et al. 2013, Ribeiro et al. 2016, Li et al. 2016, Gonzalez-Cordoba et al. 2015, Zheng et al. 2014, He et al. 2013, Carniato et al. 2009, Zhang and Li 2010, Salem et al. 2007a]), todos eles com mais de seis distúrbios analisados.

A seguir são apresentados subseções onde, em cada subseção, um artigo é apresentado, trazendo as suas principais informações, concluindo, então, com as considerações finais a respeito destes estudos e uma figura ilustrativa dos principais pontos abordados e não abordados, visualizando como tais estudos influenciam ou corroboram com a proposta apresentada.

3.2.1 Rough-set-based feature selection and classification for power quality sensing device employing correlation techniques

O artigo apresentado em [Dalai et al. 2013] apresenta um total de seis tipos de distúrbios elétricos isolados classificados pelo seu sistema proposto (afundamento de tensão, elevação de tensão, interrupção, harmônicas, transiente oscilatório, *flicker*) e dois tipos de distúrbios que acontecem simultaneamente na rede elétrica (afundamento de tensão com harmônicas e elevação de tensão com harmônicas).

Neste trabalho um microcontrolador (PIC24F) é utilizado para a construção do sistema embarcado, no qual são implementados os algoritmos de correlação cruzada, RST (*Rough Set Theory*), este último se referindo a uma técnica de encontrar o menor conjunto de características que podem ser utilizadas para a classificação dos distúrbios propostos.

Frente a isto, o que o sistema proposto faz é dividir a classificação dos distúrbios elétricos em duas etapas. A primeira refere-se ao cálculo do conjunto de características a partir do

sinal de entrada, enquanto que a segunda está relacionada à classificação dos dados utilizando as características extraídas na primeira etapa.

A escolha do cálculo de correlação para extração de características deu-se, segundo os autores, devido à sua baixa carga computacional. Dessa forma o sistema proposto consegue ser implementado em ambientes que exigem a restrição de resposta em tempo real e a um baixo custo. Uma outra vantagem apontada no artigo em relação à técnica de correlação é que ela minimiza o efeito dos ruídos presentes no sinal extraído.

Enquanto isso, a técnica RST é utilizada para a seleção ótima das características utilizadas. O RST é apropriado quando o conhecimento das características é limitado ou redundante, otimizando assim a seleção dos dados para posterior classificação, o que é inteiramente fundamental para redução de tempo de classificação e agilidade em termos de tempo de resposta do sistema proposto.

Desta forma, a maior contribuição apresentada em [Dalai et al. 2013] refere-se ao fato da aplicação da técnica RST, minimizando assim o conjunto de dados necessário para a classificação dos distúrbios elétricos, agilizando o processo e diminuindo o poder computacional necessário para realização desta tarefa.

Como resultado final, o sistema proposto consegue uma classificação com 97,1% de acurácia com dados reais, 96,72% para dados simulados com SNR de 50db, 94,44% para 40db e 92,36% para 30db, sendo que, em dados reais os distúrbios piores classificados foram a elevação com harmônicas, as oscilações e os transientes impulsivos.

A vantagem da utilização desta técnica em detrimento de técnicas mais sofisticadas para classificação dos dados está no fato de utilizar apenas regras do tipo *if-else*, reduzindo assim necessidade de treinamento e consumindo pouca memória e pouco poder de processamento, é por isso que tal técnica é possível de ser implementada em um ambiente de baixo poder computacional e consegue ter resultados semelhantes (até melhores) que técnicas mais robustas, como SVM, Lógica Fuzzy, Redes Neurais Artificiais e Transformada de Wavelet, como apresentado no próprio artigo.

3.2.2 Real-time system for automatic classification of power quality disturbances

O sistema de classificação proposto em [Ribeiro et al. 2016] é diferente dos demais no que diz respeito a utilização de uma plataforma embarcada na construção do sistema. Neste trabalho é apresentado o FPGA apenas como forma de captura dos dados reais, que, por sua vez são passados para um computador com LabVIEW ®. No entanto, apesar de não ser totalmente embarcado, este sistema possui um diferencial em relação aos demais no que diz respeito à quantidade de distúrbios detectados, são 20 distúrbios, dos quais seis são isolados (afundamento de tensão, elevação de tensão, noth, transiente impulsivo, transiente oscilatório, harmônicas) e 14 múltiplos, sendo que destes 10 são duplos (dois distúrbios simultâneos) e 4 são triplos (três distúrbios simultâneos).

Para chegar a este resultado bastante expressivo no que diz respeito à quantidade de distúrbios, as técnicas utilizadas neste trabalho foram a utilização de um filtro notch de 2ª ordem (para eliminação da frequência fundamental), cálculos para extração de características (cumulantes de segunda ordem e valor RMS) e diferentes classificadores em uma árvore de decisão (classificador bayesiano e Perceptron).

Os algoritmos perceptron são utilizados nas camadas superiores da árvore de decisão e os classificadores bayesianos são utilizados nas "folhas" da árvore para apresentar o resultado final da classificação.

A taxa de acerto obtida com a utilização desta técnica está acima dos 99%, como exposto no trabalho, sendo o distúrbio de harmônica com transiente oscilatório o que possui maior erro na classificação ($94,18 \% \pm 1,86\%$).

Apesar da alta taxa de acerto para a classificação dos distúrbios elétricos e de possuir uma resposta em tempo real, o sistema proposto não possui um detector de distúrbios elétricos, ou seja, a todo momento estará realizando os cálculos em sua busca. Além disso, não é apresentado no trabalho a classificação do sinal puro (senoide pura), logo, o sistema pode apresentar problemas de classificação neste caso, o que poderia ser evitado caso houvesse um detector.

3.2.3 Detection and Classification of Power Quality Disturbances Using Double Resolution S-Transform and DAG-SVMs

A abordagem adotada por [Li et al. 2016] consiste na utilização da Transformada S com resolução dupla (a *Double Resolution S-Transform*, ou DRST), utilizada para a extração de características em tempo real dos distúrbios elétricos; e na utilização do algoritmo de classificação SVM diferente do convencional, o DAG-SVM, que utiliza menos poder computacional. A junção destes algoritmos permite a implementação do monitor de distúrbios elétricos em ambiente embarcado com resposta em tempo real.

Nesse contexto, deve-se observar que o ambiente embarcado proposto é um processador digital de sinais (DSP), que possui, em tese, pouca memória e poder de processamento reduzido quando comparado com outras plataformas embarcadas como FPGA e plataformas microprocessadas. É por este motivo, portanto, que é realizado o treinamento do DAG-SVM anteriormente, com os valores otimizados da saída do algoritmo DRST, desta forma, os parâmetros obtidos são gravados no DSP para que este possa realizar a classificação, visto que o algoritmo já está treinado.

Além disso, deve-se ressaltar que o algoritmo DRST possui algumas características interessantes para o problema de classificação de distúrbios elétricos. Além de extrair características de frequência e tempo (permitindo uma maior variedade na detecção dos distúrbios), ele também tem a propriedade de ser menos sensível a ruídos do que, por exemplo, a Transformada S pura.

Enquanto isso, o algoritmo de classificação utilizado, o DAG-SVM, consiste de um algoritmo baseado em uma árvore na qual seus nós são algoritmos SVM, o termo *DAG* no início do nome do algoritmo significa *Direct acyclic graph*, ou seja, grafo acíclico direto (uma árvore). Logo, o treinamento deste algoritmo é semelhante ao de uma SVM pura, no entanto, possui uma maior capacidade de classificação, pois, a cada fase é feita uma nova análise, assim, os níveis mais baixos da árvore possuem poucos dados para classificação e precisam, assim, de menos recurso computacional para realizar a classificação.

Desta forma, com base nesses algoritmos, uma simulação realizada pelos autores com base em dados simulados no Matlab ®sem aplicação de ruído obteve uma média de acurácia de 97,3% quando a DAG-SVM foi treinada com 30 exemplos, 98,6% para 50 exemplos e

99,3% para 100 exemplos (lembrando que o treinamento é realizado em computador e passado para o DSP posteriormente), sendo que o distúrbio de interrupção foi o que encontrou menor acurácia na classificação, 94%, 94,5% e 96% para 30, 50 e 100 exemplos, respectivamente.

Em ambiente simulado com uma fonte trifásica capaz de simular os distúrbios, o método proposto obteve o resultado apresentado na Figura 3.8, no qual variou-se a frequência fundamental e verificou-se a adaptabilidade do método. Na figura, um eixo representa os tipos de distúrbios (classificados como F1 a F9), outro representa o valor da frequência fundamental e o último a média da acurácia geral, ou seja, o gráfico serve para mostrar a variação na qualidade do algoritmo utilizado quando varia-se a frequência fundamental do sinal.

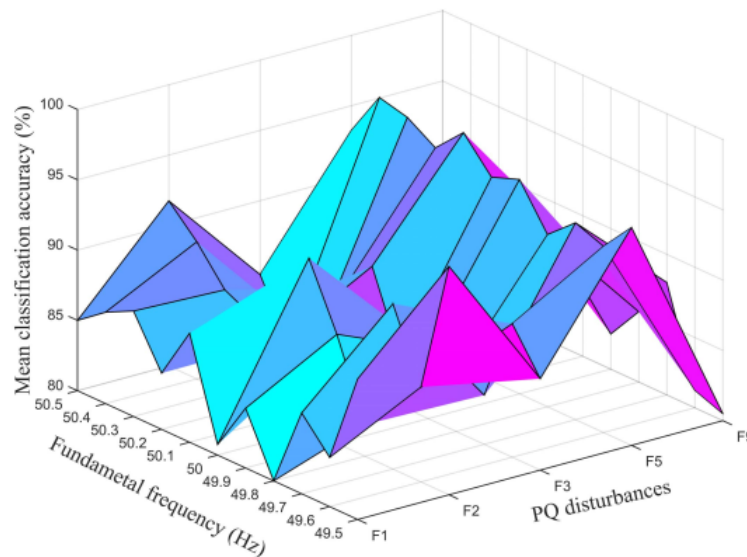


Figura 3.8: Resultado da Classificação para diferentes frequências fundamentais obtidas por [Li et al. 2016]

3.2.4 Automatic detection and classification of electrical disturbances by means of empirical mode decomposition method

Em [Gonzalez-Cordoba et al. 2015] o problema de monitoramento de distúrbios elétricos é abordado com a utilização de FPGA baseado em DAS (*data aquisition system*), sistema de aquisição dos dados, que foi utilizado para a captura dos dados e repassados para um computador realizar a classificação dos distúrbios.

Desta forma, foi implementado em computador um método para classificação dividido em três passos, o primeiro consiste na aplicação do algoritmo de EMD (*Empirical Mode Decomposition*), seguido pelo algoritmo para extração de características (a entropia de Shanon e a energia), por fim uma rede neural é aplicada para realizar a classificação. A Figura 3.9 apresenta a metodologia utilizada para o sistema de monitoramento proposto.

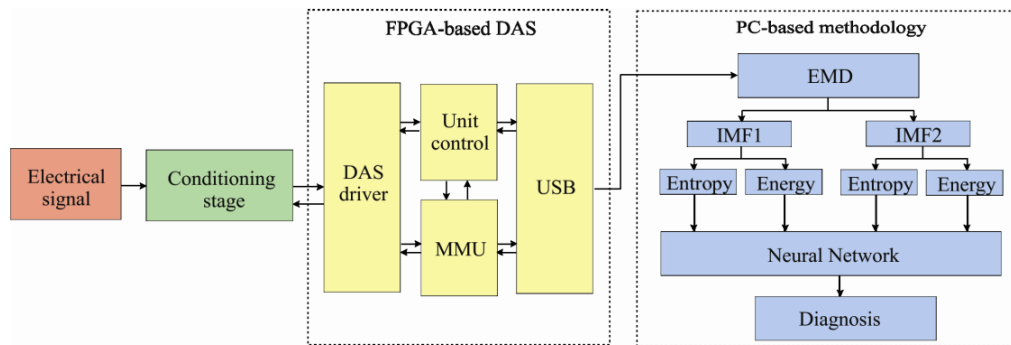


Figura 3.9: Metodologia utilizada em [Gonzalez-Cordoba et al. 2015]

Com base nesta metodologia, foi possível obter uma acurácia acima ou igual a 94% para todos os distúrbios classificados (interrupção, afundamento e elevação de tensão, *flicker*, transiente oscilatório, *notching* e harmônicas) além do sinal sem distúrbio, sendo que *flicker* foi o que possuiu menor acurácia, com exatos 94%. Estes dados foram obtidos a partir de simulação realizada com base na modelagem matemática dos distúrbios elétricos.

Em ambiente simulado com um gerador de distúrbios e um motor, o método proposto obteve uma acurácia acima de 90% para todos os distúrbios, sendo que alguns obtiveram uma acurácia de 100% na classificação (sinal puro e os distúrbios de interrupção, afundamento e elevação de tensão).

A desvantagem do método utilizado, apesar dos bons resultados obtidos, consiste na falta de aplicação do método em distúrbios múltiplos, além disso, não é apresentado o tempo de resposta do sistema proposto, não é realizada uma comparação com outros trabalhos realizados, não foi apresentada a base estatística utilizada para aquisição dos resultados experimentais (por exemplo, quantos dados foram testados para obter essa acurácia de 100%, qual o tamanho da janela utilizada), e, por fim, o sistema, sendo implementado em um computador encarece-o, além de não possuir flexibilidade na implantação em ambientes reais (visto que a comunicação com o FPGA é feita via USB).

3.2.5 An embedded-based distributed private cloud: Power quality event classification

O sistema proposto em [Zheng et al. 2014] faz a utilização de um sistema microprocessado embarcado para a realização da tarefa de monitoramento da rede elétrica na busca por distúrbios que possam influenciar no funcionamento dos equipamentos. A plataforma microprocessada escolhida foi o Raspberry Pi devido a sua ampla utilização nos dias atuais quando se trata de aplicações com sistemas embarcados, além de possuir diversos dispositivos internos que facilitam o projeto do sistema proposto.

A arquitetura para a realização da captura e análise dos distúrbios elétricos é apresentada na Figura 3.10.

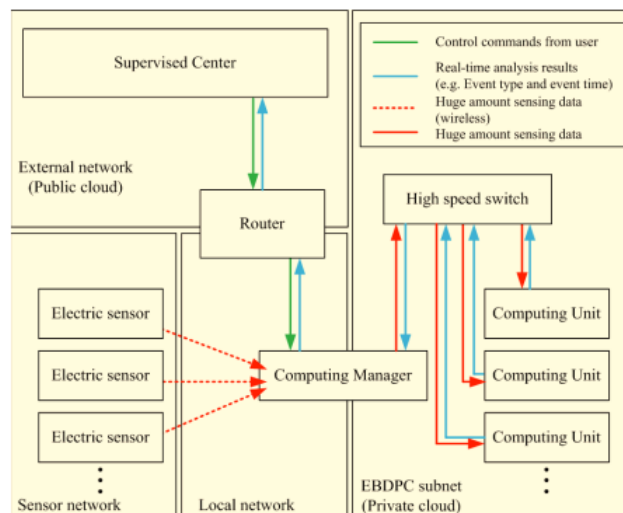


Figura 3.10: Arquitetura utilizada em [Zheng et al. 2014]

A utilização desta arquitetura tem o intuito de criar um sistema que seja altamente escalável, ou seja, que permita a adição de novas unidades computacionais (e sensores) para realizar o monitoramento de múltiplos ambientes. Nesta estrutura, os sensores são conectados ao computador de gerenciamento, que passa os dados via *switch* para as unidades de processamento, que realizam as operações de classificação dos distúrbios.

Os algoritmos utilizados foram o FFT (*Fast Fourier Transform*) e o cálculo do valor RMS, a partir destes valores, os resultados vão para uma estrutura *if-else* capaz de classificar os dados obtidos em um dos distúrbios detectados pelo sistema.

A classificação realizada por este sistema em um ambiente simulado a partir da mode-

lagem matemática, sem acréscimo de ruído e com uma amostra de 35 ciclos da senoide, obteve um resultado de 100% de acurácia (realizando a aplicação em 1000 amostras diferentes) para todos os distúrbios detectados (afundamento e elevação de tensão, interrupção, transiente impulsivo, harmônicas e *flicker*). Apesar da alta acurácia, deve-se ressaltar a falta de distúrbios múltiplos na classificação do sistema proposto e a falta de um teste estatístico para obtenção de uma conclusão com sem enviesamento.

3.2.6 A real-time power quality disturbances classification using hybrid method based on s-transform and dynamics

O trabalho realizado em [He et al. 2013] apresentou uma abordagem de sistema de monitoramento de distúrbios elétricos utilizando DSP e FPGA com a implementação de um método híbrido baseado em Transformada S e *Dynamics*.

O algoritmo *Dynamics* foi utilizado devido a sua baixa complexidade computacional, o que facilita sua implementação em ambiente embarcado e permite obter uma resposta em tempo real, o que é essencial no problema do monitoramento da Qualidade da Energia Elétrica.

A Transformada S é utilizada em conjunto com o *Dynamics*, formando um método híbrido que realiza a extração de cinco características distintas para serem analisadas e, posteriormente, classificadas de acordo com a nomenclatura proposta em [Association et al. 2009].

Neste contexto, a classificação utilizada no trabalho é realizada a partir de uma árvore de decisão, onde são utilizadas as cinco características extraídas utilizando o método híbrido com a Transformada S, o *Dynamics* e algumas outras combinações de informações (como um cálculo sobre a amplitude do sinal, obtido a partir da Transformada S).

Com a implementação deste algoritmo em ambiente computacional para obter a acurácia, foram gerados 100 distúrbios aleatoriamente (com dez ciclos cada um) e aplicados ao sistema com diferentes níveis de ruído. O resultado obtido foi de 99,27% de acurácia para o algoritmo proposto com os dados simulados com ruído a 40dB, 97,91% para distúrbios com ruído a 30dB e 94,36% com ruído a 20dB.

Frente ao fato de que o sistema possui uma capacidade de classificar distúrbios isolados

(harmônicas, *flicker*, afundamento e elevação de tensão, interrupção, transiente oscilatório, transiente impulsivo, e *notching*) e múltiplos (harmônicas com afundamento de tensão e harmônicas com elevação de tensão), ao realizar a classificação, o distúrbio de interrupção é o que possui o menor valor de acurácia, com 85% para a classificação do sinal com ruído a 20dB.

Feita a análise sobre dados simulados e mostrando uma alta eficácia na classificação frente a outros dois métodos abordados, os autores realizaram então os testes em ambiente *in vitro*, com um gerador de distúrbios conectado ao sistema proposto, no entanto, nestes testes foi analisada apenas a eficiência do algoritmo implementado na plataforma embarcada, demonstrando a possibilidade de trabalhar-se em tempo real (com tempo de resposta entre 0,03s e 0,08s).

3.2.7 Development of hardware and software for three-phase power quality disturbances detection, classification and diagnosis using Kalman Filter theory

O sistema proposto em [Carniato et al. 2009] é semelhante àqueles propostos em [Ribeiro et al. 2016] e [Gonzalez-Cordoba et al. 2015], no qual o sistema embarcado desenvolvido tem o papel de adquirir os dados e, possivelmente, realizar alguma pré-operação sobre os dados. No caso do estudo de [Carniato et al. 2009] foram propostos três sistemas embarcados, um para a aquisição dos dados, um para a confecção do filtro de Kalman (DSP) e um outro condicionador de sinal, posteriormente, os dados são passados para um computador com Matlab/Simulink ®e feita a classificação do distúrbio detectado. Na Figura 3.11 é apresentada a arquitetura do sistema proposto.

Apesar do sistema proposto ser menos flexível que os demais, por sua implementação ser em computador, além de não apresentar o resultado da eficácia e eficiência do método proposto e nem qual o algoritmo de classificação utilizado, esse trabalho possui uma vantagem sobre os demais que consiste na análise trifásica da rede elétrica, enquanto que os demais estudos até aqui apresentados trabalham somente com a análise de distúrbios acontecendo em uma única fase, tanto é que em [Carniato et al. 2009] é detectado e classificado o distúrbio do tipo desbalanceamento de fase, pouco observado na maioria dos sistemas propostos na

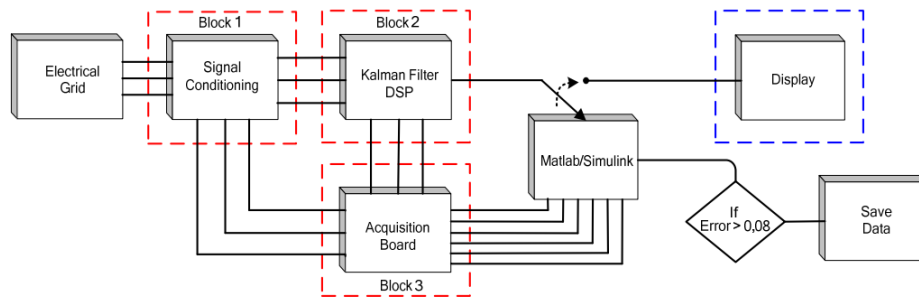


Figura 3.11: Digrama de blocos do sistema proposto em [Carniato et al. 2009]

literatura.

Além disso, uma vantagem deste sistema é a apresentação de maneiras de mitigar o problema elétrico e como se proteger deles (em termos tanto de proteção de equipamentos quanto de segurança pessoal de quem trabalha com tais equipamentos afetados pelos distúrbios).

3.2.8 DSP-FPGA based real-time power quality disturbances classifier

A proposta apresentada em [Zhang and Li 2010] traz uma abordagem com a utilização de um ambiente embarcado desenvolvido sobre as plataformas DSP e FPGA, com implementação dos algoritmos de DFT (Trasnformada Discreta de Fourier), para análise no domínio da frequência; série de valores RMS, para análise no domínio do tempo; e por fim um classificador de Árvore de Decisão Baseada em Regras (ou RBDT), um algoritmo bastante simples, o que dá ao sistema implementado a capacidade de realizar a tarefa de classificação dos distúrbios elétricos em tempo real.

Os distúrbios que podem ser detectados por este sistema são variados, sendo sete distúrbios isolados (afundamento e elevação de tensão, interrupção, harmônicas, *notching*, *flicker* e transiente oscilatório), e dois distúrbios múltiplos simultâneos (harmônicas com afundamento de tensão e harmônicas com elevação de tensão). Como pode-se notar, há a falta do distúrbio de desbalanceamento de fase, entretanto, o sistema proposto, diferente da maioria aqui abordado, apresenta uma solução para classificação de distúrbios que acontecem em várias fases, apesar de apresentar apenas os tipos de distúrbios monofásicos.

A partir das características coletadas com base nos cálculos de DFT e valor RMS, é possível construir a árvore de decisão somente com base nos valores destes cálculos. Este

tipo de algoritmo utiliza apenas uma estrutura *if-else*, que, como dito anteriormente, agiliza o processo de classificação.

A partir de simulações, o sistema proposto obteve resultados de acurácia de 99%, 97,5% e 94% com SNR (taxa de sinal-ruído) de 50dB, 40dB e 30dB respectivamente, sendo que na simulação foram utilizadas 100 amostras de cada tipo de distúrbio com 2500 amostras em 10 ciclos. O distúrbio de *notching* fora aferido como o que é mais sensível ao ruído utilizando as características apresentadas no artigo, possuindo uma acurácia abaixo de 90% quando o SNR é de 30dB.

A Figura 3.12 apresenta como deu-se a construção do sistema embarcado proposto.

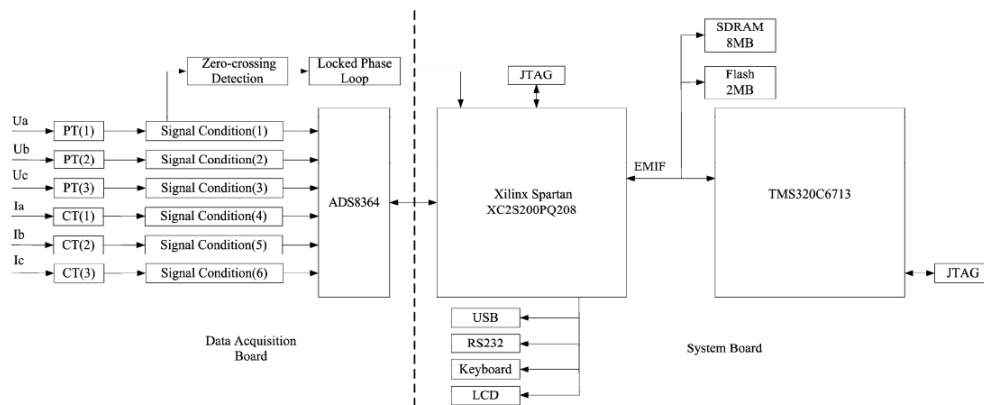


Figura 3.12: Digrama de blocos do sistema proposto em [Zhang and Li 2010]

Em relação a aplicação em ambiente real, apesar de não ter sido feita, utilizou-se uma abordagem mais próxima do real com a utilização de equipamentos geradores de distúrbios elétricos, utilizando-se o sistema proposto, os autores apresentaram uma eficiência na classificação que varia de 0,09s a 0,12s em média, um resultado bastante promissor para um sistema embarcado com resposta em tempo real (a eficácia não fora verificada neste processo).

3.2.9 Power quality disturbance detection using DSP based continous wavelet transform

A proposta de sistema de monitoramento de distúrbios elétricos em tempo real apresentada por [Salem et al. 2007a] traz uma implementação em tempo real aplicada em um sistema

embarcado de baixo custo consistindo de um DSP com a implementação do algoritmo da Transformada de Wavelet.

Este sistema é capaz de detectar a presença de distúrbios como afundamento de tensão, elevação de tensão, transiente impulsivo, *notching*, interrupção e harmônicas, no entanto, nenhum algoritmo de classificação é apresentado, ou seja, o sistema proposto busca apenas demonstrar a eficiência do método da Transformada de Wavelet para a detecção quando utiliza-o implementado em DSP, em detrimento ao mesmo algoritmo num ambiente Matlab®, como apresentam os autores nesse artigo.

Visto que o intuito do artigo consiste em fazer essa análise de tempo comparativo entre o ambiente embarcado e um computacional mais robusto (resultados mostrado na Tabela 3.6), sem implementação de uma rede neural capaz de identificar o tipo de distúrbio ocorrido na rede elétrica, esse estudo pode servir como base preliminar para a aplicação do método apresentado juntamente com um classificador satisfatório para obtenção de um sistema completo, que faça a detecção e a classificação dos distúrbios elétricos.

Tabela 3.6: Comparativo Entre Tempo no DSP e no PC (Adaptado de: [Salem et al. 2007a])

Distúrbio	Tempo DSP (s)	Tempo Matlab (s)
Afundamento de Tensão	0.4338	4.2500
Elevação de Tensão	0.4338	4.2650
Transiente Impulsivo	0.4338	4.2660
Distúrbios Múltiplos	0.4338	4.2660
<i>Notching</i>	0.4338	4.2650
Interrupção	0.4338	4.2650

3.2.10 Considerações Finais

Como visto, alguns dos trabalhos apresentados trouxeram uma plataforma embarcada apenas como forma de aquisição dos dados elétricos, como em [Ribeiro et al. 2016, Gonzalez-Cordoba et al. 2015] nos quais são apresentados FPGA como coletor dos sinais elétricos, e em [Carniato et al. 2009] no qual o sistema embarcado utilizado para a captura é o DSP. Os demais sistemas, apresentados em [Dalai et al. 2013, Li et al. 2016, Zheng et al. 2014, He et al.

2013,Zhang and Li 2010,Salem et al. 2007a], utilizam diretamente o sistema embarcado para aquisição e classificação dos dados coletados.

Pode-se perceber também que nenhum dos estudos realizados fazem a classificação de todos os distúrbios isolados, o que chega mais próximo disso é o trabalho apresentado em [Carniato et al. 2009], que, dos distúrbios isolados, apenas o *notching* não é classificado pelo sistema. No entanto, neste trabalho não foi apresentada a acurácia na classificação, portanto, não pode-se ter uma comparação com os demais sistemas que não apresentam essa alta quantidade de distúrbios isolados detectados. Além disso, este mesmo trabalho é o único dos nove selecionados que possui um sistema capaz de apresentar possíveis soluções para a mitigação do problema ocorrido na rede elétrica.

Outro fator que chama a atenção é que, apesar de alguns trabalhos apresentarem um experimento *in vitro*, utilizando geradores de distúrbios elétricos, nenhum deles traz uma implementação em ambiente real, com um estudo de caso onde possa-se comprovar a eficiência e eficácia do sistema proposto.

Como forma de resumir um comparativo entre os trabalhos, na Figura 3.13 são apresentados os dados coletados a partir dos trabalhos selecionados para obter-se uma noção mais geral de pontos fracos e fortes de cada trabalho.

Para concluir a análise de trabalhos relacionados, na Seção E apresenta-se um breve mapeamento das patentes relacionadas a sistemas de monitoramento de distúrbios elétricos nas bases INPI e *Derwent Innovations Index*.

Referência	Distúrbio										Plataforma	Algoritmos	Acurácia Máxima	Tempo de Execução	Teste com Ruído	Tempo de Execução?	Apresenta Solução?	Análise Energética	Distúrbios Simulados	Distúrbios Gerados	Distúrbios Reais
	1	2	3	4	5	6	7	8	9	10											
[Dalai et al. 2013]	X	X	X	X		X	X				DSP	Correlação Cruzada, RST, if-else	97.10%	-	N	S	N	N	S	S	
[Ribeiro et al. 2016]	X	X	X	X	X	X					FPGA + PC	Filtro Notching, Valor RMS, Árvore de Decisão (Perceptron + Redes Bayesianas)	99.58%	-	N	N	N	S	N	N	
[Li et al. 2016]	X	X	X	X		X	X				DSP	DRST, DAG-SVM	99.30%	0,007s a 0,06s	S	S	N	S	S	N	
[Gonzalez-Cordoba et al. 2015]	X	X	X	X		X	X	X			FPGA + PC	EMD, Entropia de Shanon e Energia, RNA	94%	-	N	N	N	S	S	N	
[Zheng et al. 2014]	X	X	X	X	X		X				Raspberry Pi	FFT, Valor RMS, if-else	100%	-	N	N	N	S	N	N	
[He et al. 2013]	X	X	X	X	X	X	X	X			DSP+FPGA	Dynamics, Transformada S	99.27%	0,03s a 0,08s	S	S	N	S	S	N	
[Carniato et al. 2009]	X	X	X	X	X	X	X		X		DSP + PC	Filtros Kalman	-	-	N	N	N	S	S	N	
[Zhang and Li 2010]	X	X	X	X		X	X	X			DSP+FPGA	DFT, Valor RMS, RBDT	99%	0,09s a 0,12s	S	S	N	S	S	N	
[Salem et al. 2007]	X	X	X		X			X			DSP	Trasformada de Wavelet	-	0,4338s	N	S	N	S	S	N	
PROPOSTA	X	X	X	X	X	X	X	X	X	X	NodeMCU ou Raspberry	FFT, DWT, Filtro Notch, Filtro Gaussiano, TH, ED, SVM, MLP, RF, Dt	-	-	S	S	S	N	N		
1 Afundamento de Tensão 2 Elevação de Tensão 3 Interrupção 4 Harmônicas 5 Transiente Impulsivo 6 Transiente Oscilatório											7 Flicker 8 Notch 9 Desbalanceamento de Fase 10 Nível DC 2x Múltiplos (Dois) 3x Múltiplos (Três)										

Figura 3.13: Comparativo entre trabalhos selecionados

3.3 Análise de Resultados Mapeamento Estado da Técnica

Utilizando as 7 patentes selecionadas após aplicados os critérios de inclusão e exclusão explicitados no Apêndice E, foi feita a análise detalhada de cada uma para posterior extração de informações que possam responder às questões de pesquisa propostas e, consequentemente, alcançar o objetivo do mapeamento.

Os resultados obtidos respondendo a cada uma das questões são apresentados separadamente a seguir.

Quais os anos em que foram depositadas as patentes relacionadas a distúrbios elétricos?

Esta questão busca analisar como está o desenvolvimento de produtos patenteados relacionados à área da Qualidade de Energia Elétrica, envolvendo, especificamente, monitoramento de distúrbios elétricos, desde sua captura a possível tratamento.

Frente a isto, o histograma apresentado na Figura 3.14 mostra o desenvolvimento da área nos últimos anos (vale ressaltar que a pesquisa realizada não restringiu a data de depósito das patentes e que o gráfico apresenta todas as encontradas nas bases mencionadas, inclusive aquelas que possuem mesmo conteúdo, mas foram depositadas em diversos países, e, possivelmente, em diferentes anos).

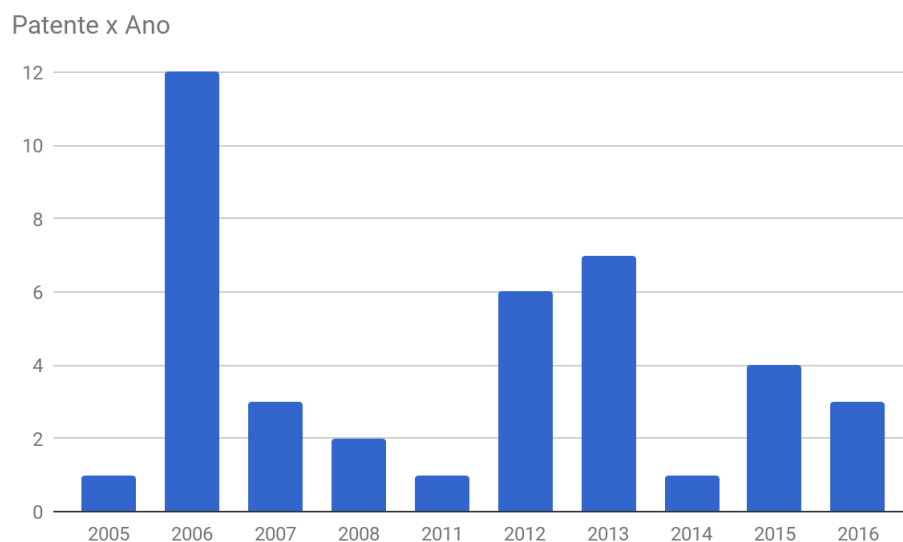


Figura 3.14: Patentes depositadas a cada ano.

Como pode ser visualizado a partir da Figura 3.14, o ano em que foi depositada a maior quantidade de patentes foi o ano de 2006, sendo que apenas a patente [Montreuil 2009] possui 12 patentes depositadas neste ano, ou seja, todas representadas no gráfico. Os anos de 2013 e 2012 são os seguintes na relação número de patentes depositadas, com 7 e 6 patentes, respectivamente.

Como forma de visualizar melhor a relação quantidade de patentes por ano, o gráfico apresentado na Figura 3.3 traz uma abordagem diferente da apresentada na Figura 3.14. Nesse, são apresentadas as patentes de acordo com o primeiro ano de depósito delas, independente do país onde tenha ocorrido.

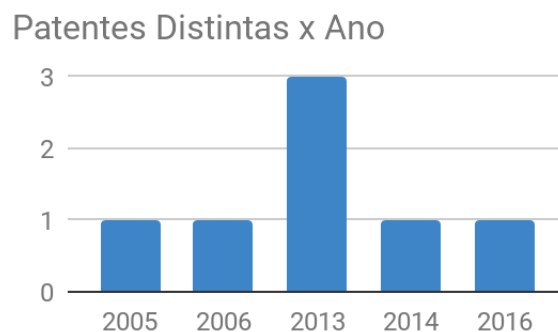


Figura 3.15: Patentes depositadas por ano de acordo com o primeiro depósito.

Como pode-se notar a partir da Figura 3.3, houve um aumento no número de patentes depositadas no ano de 2013, com 3 distintas, enquanto que nos demais anos houve apenas uma ao ano. Tendo em vista que, de acordo com [Kennedy 2000], houve um crescente aumento no número de pesquisas relacionadas ao monitoramento de distúrbios elétricos devido às despesas decorrentes deste problema nas instalações elétricas, o mesmo padrão parece não se repetir quando se trata de patentes.

Este número baixo de patentes relacionadas ao tema abre espaço para os pesquisadores da área, que estão desenvolvendo técnicas que auxiliem o monitoramento de distúrbios elétricos mais eficiente e eficaz, transformem seus trabalhos em patentes e, possivelmente produtos, de maneira a auxiliar a indústria, que vem perdendo considerável renda relacionada aos problemas de distúrbios elétricos [Kennedy 2000].

Quais os países que mais têm patentes depositadas relacionadas a distúrbios elétricos?

A resposta para esta pergunta é ilustrada na Figura 3.16, em que é apresentado um histograma exibindo os países onde foram depositadas as patentes relacionadas ao monitoramento, detecção ou classificação de distúrbios elétricos, visualizando, desta forma, oportunidades para pesquisadores desenvolverem sistemas nesta área e depositarem patentes em países onde ainda não as possui.

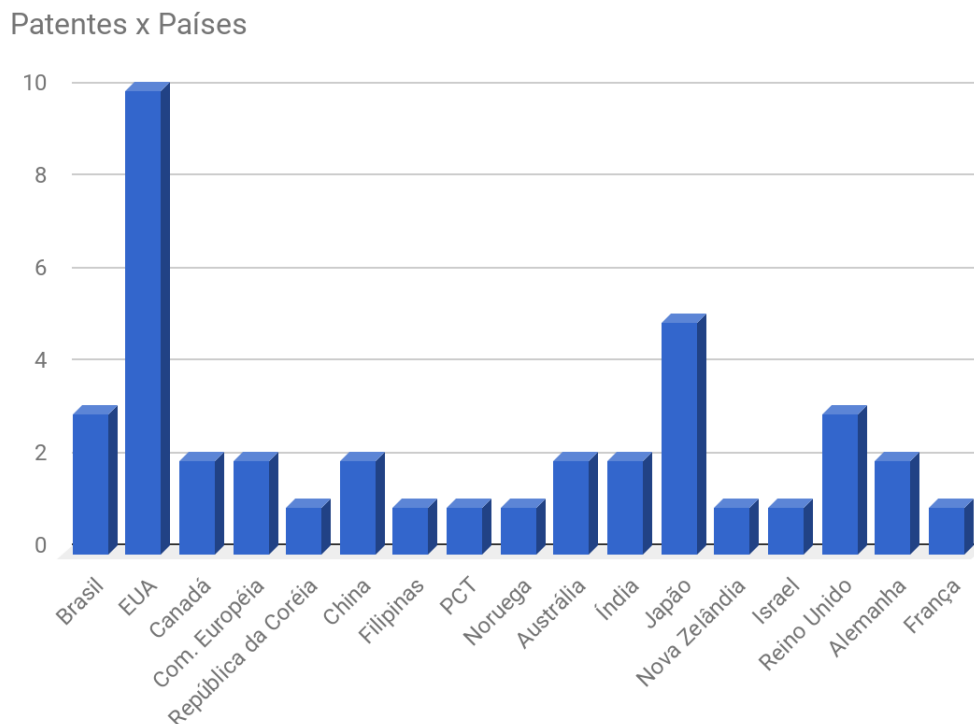


Figura 3.16: Países de depósito das patentes encontradas.

Como pode-se visualizar a partir do gráfico apresentado na Figura 3.16, os Estados Unidos é o país com maior número de patentes depositadas (10), seguido pelo Japão, com 5, Brasil e Reino Unido com 3, Canadá, Comunidade Europeia, China, Austrália, Índia e Alemanha com 2 e, por fim, com uma patente apenas tem-se a República da Coreia, Filipinas, Noruega, Nova Zelândia, Israel e França, além de uma patente publicada segundo o Tratado de Cooperação de Patente (PCT, sigla em inglês). Este tratado, segundo [Wipo 2017], facilita o trabalho dos órgãos responsáveis dos países assinantes no que diz respeito à análise da patente, além de auxiliar o próprio inventor no que diz respeito à busca por potencial de

proteção para o seu invento mediante a concessão das patentes.

Vale ressaltar que o tratado PCT não emite as patentes, apenas auxilia os membros assinantes do Tratado a tomarem as decisões relacionadas à concessão ou não da patente para determinado inventor.

Sendo assim, apesar de um baixo número na quantidade de patentes relacionadas ao tema, as existentes abrangem uma quantidade de países razoável (mais de 10), sendo a patente [Montreuil 2009] a que possui a maior abrangência, sendo ela depositada em 8 países, na comunidade Europeia e no PCT (por ter iniciado o processo pelo PCT, foi facilitado a entrada em diversos outros países).

Quais as áreas (de acordo com o CIP - Classificação Internacional de Patentes) que possui o maior número de patentes relacionadas?

Na Figura 3.17 é apresentado um histograma mostrando em quais áreas, segundo a Classificação Internacional de Patentes, as patentes foram depositadas.

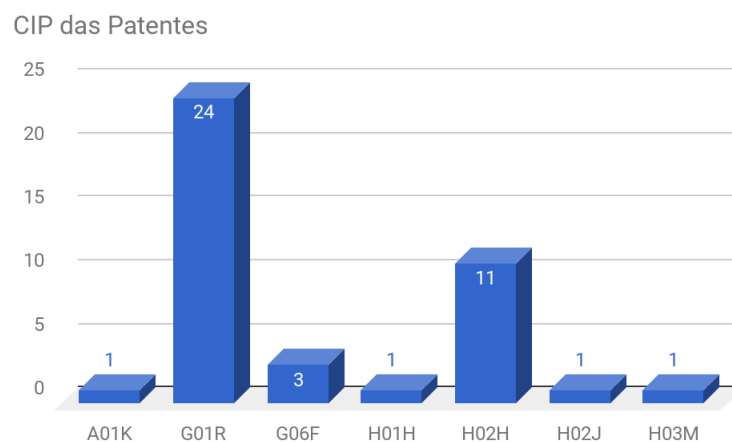


Figura 3.17: Classificação Internacional de Patentes das 7 patentes encontradas.

Antes de analisar o gráfico, na Tabela 3.7 é apresentada, de acordo com [INPI 2017], o significado de cada uma das classificações apresentadas no gráfico presente na Figura 3.17.

Tabela 3.7: Classificação internacional de patentes

CIP	Significado
A01K	Necessidades Humanas - Agricultura - Criação de Animal em Geral (especialmente pecuária)
G01R	Física - Instrumentos - Medidas de Variáveis Elétricas (ou Magnéticas)
G06F	Física - Computação; Cálculo; Contagem - Processamento Digital de Dados Elétricos
H01H	Eletricidade - Elementos Básicos de Eletricidade - Relé; Switch; Seletor; Dispositivos de Proteção Emergencial
H02H	Eletricidade - Geração, Conversão ou Distribuição de Energia Elétrica - Arranjos de Circuito de Proteção de Emergência
H02J	Eletricidade - Geração, Conversão ou Distribuição de Energia Elétrica - Arranjos de Circuitos ou Sistemas para Fornecimento ou Distribuição de Energia Elétrica; Sistemas para Armazenar Energia Elétrica
H03M	Eletricidade - Circuito Eletrônico Básico - Codificador, Decodificador ou Conversor Digital em geral

Como pode-se analisar pelas informações contidas na Tabela 3.7, a maioria das classificações estão de acordo com o que se espera de dispositivos capazes de detectar ou proteger circuitos de possíveis distúrbios elétricos, o único que possui uma característica diferente é a classificação A01K, que se refere à agricultura, isso ocorre devido à patente [Montreuil 2009] possuir uma aplicação direta em ambientes de pecuária, evitando que a presença de distúrbios elétricos afete locais como fazendas.

A classificação G01R é a que possui o maior número de patentes, o que já era esperado, visto que todas as patentes precisam apresentar uma forma de aferir determinada medida elé-

trica para poder trabalhar em cima desta e, posteriormente, verificar a presença ou ausência de distúrbio naquela medida.

No gráfico da Figura 3.17 é apresentada uma grande quantidade de patentes relacionadas, por exemplo, à classificação G01R (24 patentes), sendo que existe apenas 7 patentes. Isto acontece devido a classificação G01R ainda possuir subclassificações, como por exemplo G01R-013/02 referente a arranjos para exibir variáveis elétricas ou formas de onda em formato digital, e G01R-019/04 referente a arranjos para medir correntes ou tensões, ou para indicar a presença de sinal deste tipo medindo o valor de pico da onda ou de pulsos, além de várias outras subclassificações.

Frente a isto, o gráfico apresentado na Figura 3.18 mostra com maiores detalhes a classificação dividida em suas subcategorias.

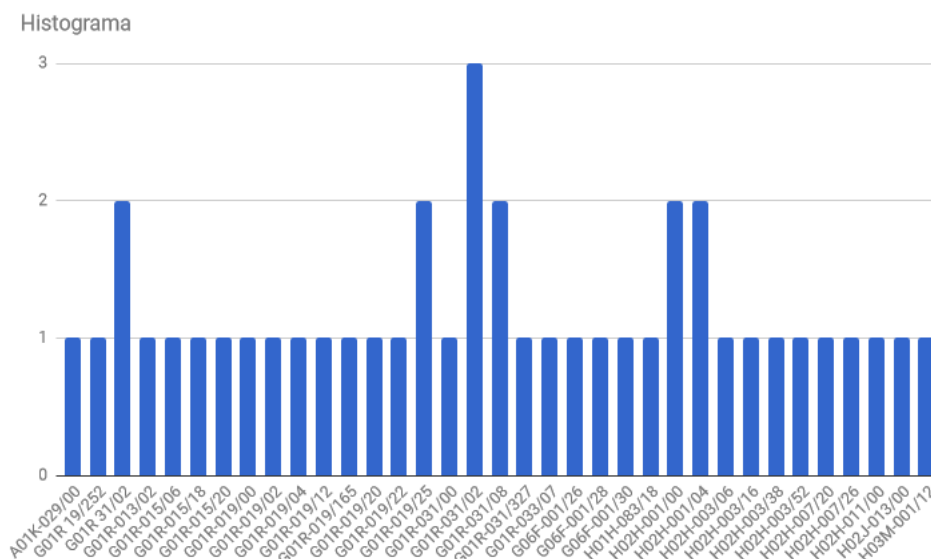


Figura 3.18: Classificação completa CIP presente nas patentes selecionadas.

Como pode-se ver, o maior número de patentes (3) está depositado com a classificação G01R-031/02, que se refere a circuitos ou sistemas para verificação (teste) da energia elétrica, verificando a presença de descontinuidade, vazamentos ou conexão incorreta na linha elétrica. O significado das demais classificações pode ser encontrado em [INPI 2017].

Tendo analisado até então detalhes relacionados às patentes, parte-se agora para uma análise mais aprofundada ao seu conteúdo, para isto, respondeu-se às três perguntas restantes, relacionadas às técnicas utilizadas nas patentes selecionadas.

Quais as técnicas utilizadas nas patentes relacionadas ao monitoramento de distúrbios elétricos?

Esta pergunta visa identificar, dentro das patentes existentes, as técnicas utilizadas para monitoramento dos distúrbios elétricos, desta forma, os pesquisadores desta área podem visualizar oportunidades para criação de novas patentes com a criação de novas técnicas diferentes daquelas apresentadas nas patentes existentes.

Na Tabela 3.8 apresentam-se as técnicas utilizadas por cada uma das sete patentes encontradas seguindo o método explicitado no presente trabalho.

Tabela 3.8: Métodos utilizados nas patentes encontradas

Método	Patentes
Análise de Fluxo Magnético	[Montreuil 2009], [Ge Aviation 2013], [Davis 2013]
Cálculo de Indutância	[Ge Aviation 2013], [Davis 2013]
Análise de Probabilidade	[Anderson et al. 2013], [Kim 2016]
Detecção de Novidade	[Ribeiro et al. 2015]
Transformada de Wavelet Discreta	[Juiz de Fora and Ltda 2014]
Comparação de Ponto de Inflexão	[Kim 2016]
Rede Neural Artificial (RNA)	[Ribeiro et al. 2015]
Erro do Sinal	[Ribeiro et al. 2015]
Subamostragem	[Ribeiro et al. 2015]

Como pode-se visualizar, a maior parte das patentes encontradas [Montreuil 2009], [Ge Aviation 2013] e [Davis 2013] trabalham a partir do fluxo magnético, sendo que as [Ge Aviation 2013] e [Davis 2013] fazem a análise de indutância do circuito. Estes métodos são, na realidade, utilizados diretamente a partir do aparelho de medida utilizado, que consegue capturar variações na indutância do sinal e, a partir daí, inferir a existência de

um determinado problema elétrico na rede, suposto como distúrbio elétrico. A depender da variação nesta indutância do circuito, um circuito de proteção é acionado, sendo que, em [Ge Aviation 2013] e [Davis 2013] é possível identificar onde exatamente aconteceu o distúrbio e, portanto, desativar apenas aquela parte do mesmo para evitar maiores danos aos equipamentos.

A patente [Anderson et al. 2013] apresenta um cálculo de probabilidade de ter acontecido um distúrbio elétrico na rede, conhecendo-se *a priori* a topologia de como os equipamentos estão conectados. Desta forma, esta técnica é capaz de detectar tanto o distúrbio quanto a localidade na qual o mesmo aconteceu.

A patente [Juiz de Fora and Ltda 2014] apresenta a utilização do algoritmo de Novidade e a Wavelet Discreta, sendo o algoritmo de Novidade utilizado para detectar o distúrbio elétrico e a Wavelet Discreta utilizada para a compactação do sinal detectado e posterior armazenamento deste.

A patente [Kim 2016], que também utiliza de técnicas de probabilidade para verificação do sinal faz, na realidade, uma comparação entre os dados capturados (a partir do ponto de inflexão) e a onda senoidal perfeita para a eletricidade (50 ou 60Hz, sem harmônicas). A partir desta comparação, verifica-se a probabilidade daquela onda capturada possuir ou não um distúrbio elétrico (normalmente harmônicas ou ruídos). Nesta patente, além de apresentar uma forma de detecção, busca-se também a eliminação do distúrbio do sinal.

Por fim, na patente [Ribeiro et al. 2015] é apresentada a utilização de técnicas mais avançadas de classificação dos sinais capturados, usando algoritmos de subamostragem, para diminuir a quantidade de dados a serem calculados (reduzindo assim a necessidade de um sistema robusto), e algoritmos de Erro do Sinal, também verificado a partir de comparações do sinal subamostrados com sinais previamente conhecidos. A partir da detecção, uma RNA é utilizada para classificação dos distúrbios capturados. Devido a subamostragem feita sobre o sinal, a RNA utilizada não ocupa muita memória e a técnica apresentada pode ser implementada em diversos dispositivos.

Em que plataforma são implementados o sistema de monitoramento das patentes?

Esta pergunta visa identificar como vem sendo implementado os sistemas propostos para o monitoramento da rede elétrica para identificação, detecção ou classificação dos distúrbios

elétricos.

Desta forma, respondendo a esta pergunta, é possível observar oportunidades para que os pesquisadores possam construir aparelhos, dispositivos, circuitos ou métodos capazes de resolver o problema e, conseqüentemente, patentear suas ideias.

A Tabela 3.9 apresenta a abordagem utilizada pelos inventores de cada uma das sete patentes selecionadas.

Tabela 3.9: Implementação abordada nas patentes

Implementação	Patentes
Sensor e circuitos de detecção	[Montreuil 2009], [Ge Aviation 2013], [Davis 2013]
Computador	[Anderson et al. 2013]
FPGA	[Juiz de Fora and Ltda 2014]
Microprocessador	[Kim 2016]
Método	[Ribeiro et al. 2015]

Como pode-se visualizar na Tabela 3.9, a utilização de sensores com circuitos de detecção foi a abordagem mais empregada. Na realidade, todas as patentes necessitam de sensores para a detecção do distúrbio elétrico e aferição da medida elétrica desejada, no entanto, as patentes [Montreuil 2009], [Ge Aviation 2013] e [Davis 2013], que apresentaram como método de detecção o cálculo de indução, utilizam sensores de indução para a detecção e, juntamente com um circuito e, possivelmente, com uma topologia conhecida, consegue identificar onde exatamente ocorreu o problema elétrico e acionar um esquema de proteção. Portanto, a utilização do termo "Sensor e circuitos de detecção" visualizado na Tabela 3.9 refere-se apenas ao fato destas patentes não utilizarem recursos mais robustos para o monitoramento dos circuitos elétricos.

Na patente [Anderson et al. 2013], como visto na resposta à Q4, é utilizada uma técnica de probabilidade, calculada com base em dados previamente conhecidos (tanto da topologia do circuito elétrico quanto de variáveis elétricas sem distúrbios) e armazenados em um computador, capaz de realizar o cálculo da probabilidade e, conseqüentemente, detectar o

distúrbio elétrico.

Nas patentes [Juiz de Fora and Ltda 2014] e [Kim 2016] são apresentadas a utilização de dispositivos embarcados para a detecção dos distúrbios, sendo que na [Juiz de Fora and Ltda 2014] é utilizado um FPGA e na [Kim 2016] é utilizado um sistema multiprocessado (também para o armazenamento de dados para o cálculo da probabilidade).

Finalmente, a patente [Ribeiro et al. 2015] traz apenas o método de como pode ser implementado um classificador de distúrbios elétricos, mas não traz sua implementação física.

Nas patentes são apresentadas técnicas capazes de detectar os distúrbios em tempo real ou de maneira offline?

Esta última pergunta do presente mapeamento busca identificar se os métodos utilizados conseguem fazer a captura dos distúrbios elétricos em tempo real ou se são realizados de maneira offline para posterior análise. Na Tabela 3.10 é apresentada a resposta para esta pergunta.

Tabela 3.10: Abordagem utilizada nas patentes

Abordagem	Patentes
Tempo real	[Montreuil 2009], [Ge Aviation 2013], [Anderson et al. 2013], [Davis 2013], [Juiz de Fora and Ltda 2014], [Kim 2016], [Ribeiro et al. 2015]
<i>Offline</i>	[Juiz de Fora and Ltda 2014]

Como pode-se visualizar a partir da Tabela 3.10, todas as patentes apresentam a solução para a detecção em tempo real, apenas na patente [Juiz de Fora and Ltda 2014] é apresentado, além da abordagem de tempo real, uma possibilidade de trabalhar de maneira offline, armazenando os dados capturados para posterior análise, utilizando a técnica de Wavelet Discreta para a compactação dos dados.

O fato de todas as patentes possuírem em seu escopo técnicas com capacidade de detecção em tempo real possivelmente está ligado ao fato de que a maioria delas ([Montreuil

2009], [Ge Aviation 2013], [Anderson et al. 2013], [Davis 2013], [Kim 2016]) visualizarem, além da detecção, a proteção do circuito que está sendo monitorado, desta forma, faz-se necessário a implementação em tempo real das técnicas de detecção.

3.3.1 Considerações Finais

A partir da análise das patentes foi possível verificar que apesar da pequena quantidade (apenas 7), existe entre elas uma variedade em relação a áreas de aplicação como na agricultura [Montreuil 2009], em ambientes com circuitos indutivos [Montreuil 2009], [Ge Aviation 2013], [Davis 2013], com circuitos de topologia conhecida [Anderson et al. 2013] e ambientes mais gerais [Juiz de Fora and Ltda 2014], [Kim 2016], [Ribeiro et al. 2015].

Foi possível, ainda, verificar a tendência no desenvolvimento de patentes relacionadas ao tema ao longo dos anos (Q1), além dos países onde estão sendo depositadas estas patentes (Q2).

Além disso, a Classificação Internacional de Patentes (ou CIP) permitiu verificar também em quais áreas há mais patentes depositadas relacionadas ao tema de detecção de distúrbios elétricos (Q3), o que pode auxiliar futuros pesquisadores a fazer seus depósitos também com estas classificações.

Partindo para a área mais técnica, verificou-se como eram trabalhados os métodos apresentados nas patentes, primeiro verificando quais são eles (Q4), seguido da análise de onde são implementados (Q5) e finalizando com a abordagem tomada pelos métodos, se estes possuem resposta em tempo real ou não (Q6).

3.4 Considerações Finais do Capítulo

Como foi visualizado ao longo deste capítulo, existem muitos trabalhos relacionados à área de detecção e classificação de distúrbios elétricos, o que já era esperado devido aos problemas e danos causados por falhas na Qualidade da Energia Elétrica em diversos setores.

Apesar de uma grande quantidade de artigos relacionados ao tema, o número de patentes ainda está reduzido, apesar da abrangência em termos de áreas e locais de publicação, existem na literatura muitos trabalhos que apresentam novas técnicas que melhoram na classificação (em termos de acurácia) ou apresentam uma nova técnica mas não são patenteados

nem nos seus países de origem, isso abre possibilidades para se criar patentes a partir de trabalhos nesta área, desde que estes apresentem novidades em termos de técnicas quando comparado ao estado da arte (artigos e patentes relacionados ao tema).

Além disso, muitos trabalhos, como os destacados na Seção 3.2, não apresentam uma abordagem prática, com um estudo de caso onde mostre a sua solução funcionando em ambiente real.

Outro ponto a ser destacado é a falta de sistemas que colaborem de maneira inteligente com seu usuário, ou seja, um sistema que seja capaz de se adaptar ao ambiente no qual está inserido, apresentando soluções baseadas nos equipamentos conectados à rede elétrica onde o sistema está instalado.

Além disso, nenhum dos sistemas abordados pelos estudos mapeados apresentam a detecção e classificação de todos os distúrbios apresentados em [Association et al. 2009] (transiente impulsivo, transiente oscilatório, afundamento de tensão, elevação de tensão, interrupção, desbalanceamento de fase, nível DC, harmônicas, interharmônicas, *notching*, ruído, flutuação de tensão e variação na frequência), deixando a solução para tal problema em aberto.

Dessa forma, pode-se perceber que a presente proposta aparece como uma forma de resolver estes problemas deixados em aberto pelos trabalhos estudados, ou seja, apresentar um sistema embarcado que seja capaz de classificar a maioria dos distúrbios monofásicos apresentados em [Association et al. 2009] (transiente impulsivo, transiente oscilatório, afundamento de tensão, elevação de tensão, interrupção, nível DC, harmônicas, *notching* e flutuação de tensão), alguns distúrbios duplos (transiente oscilatório com *sag* e com *swell*, harmônicas com *sag* e com *swell*, transiente impulsivo com *sag* e com *swell*, *notching* com *sag* e com *swell*), assim como os triplos apresentados em [Ribeiro et al. 2016] (harmônicas, oscilatório e *sag*, harmônicas, oscilatório e *swell*, harmônicas, *notching* e *sag*, harmônicas, *notching* e *swell*), sendo verificado a eficácia e a eficiência dos sistemas propostos.

Para se alcançar tal objetivo, um benchmark de distúrbios simulados foi criado e é apresentado em maiores detalhes no capítulo a seguir.

Capítulo 4

SED-Bench: *benchmark* de Distúrbios Elétricos Simulados

4.1 Introdução

Como já visto no capítulo anterior, há muitos trabalhos na literatura que visam a criação ou o aprimoramento de técnicas de detecção e classificação de distúrbios elétricos [Ribeiro et al. 2016, Gonzalez-Cordoba et al. 2015, Dalai et al. 2013, Li et al. 2016, Zhang and Li 2010, Ruiz-Llata et al. 2011, Ozgonenel et al. 2013]. Entretanto, poucos trabalhos disponibilizam os dados utilizados nos experimentos para serem utilizados por outros autores que visam a utilização dos mesmos para a realização de uma comparação justa entre o algoritmo proposto e o do artigo a ser comparado. Alguns trabalhos que se enquadram nesta categoria comparativa são [Ribeiro et al. 2016, Li et al. 2016, Gonzalez-Cordoba et al. 2015, He et al. 2013, Zhang and Li 2010, Salem et al. 2007a], que utilizam dados simulados baseados na padronização da IEEE [Association et al. 2009].

Em [Dalai et al. 2013] são utilizados dados reais, entretanto, não é especificado se tais dados são provenientes de algum fornecedor de energia parceiro e tais dados também não são disponibilizados para posterior utilização por outros pesquisadores.

Alguns artigos relacionados à produção de *benchmark* de afundamento de tensão e/ou elevação de tensão podem ser encontrados na literatura [Yang et al. 2012, Rens 2009, Mc-Granaghan 2005, Lassila et al. 2005], entretanto, seus objetivos não é a criação de um banco de dados de distúrbios elétricos, ao invés disso, pretende-se ali realizar a caracterização dos

sinais elétricos anômalos para que alguém que vá simular, ou criar um algoritmo de detecção, considere as características principais destes tipos de distúrbios ali explicitados.

Por fim, existem ainda os trabalhos que utilizam distúrbios gerados artificialmente, estando estes, portanto, no meio termo entre os distúrbios simulados e os reais [Dalai et al. 2013, Li et al. 2016, Gonzalez-Cordoba et al. 2015, He et al. 2013, Carniato et al. 2009, Zhang and Li 2010, Salem et al. 2007a]. Normalmente tais tipos de trabalhos utilizam geradores de distúrbios elétricos ou componentes elétricos, como motores e solenoides, que geram determinados tipos de distúrbios (afundamento de tensão, elevação de tensão e harmônicas, por exemplo).

Frente a tais trabalhos, nota-se a necessidade de se ter um *benchmark* de distúrbios elétricos (ainda que simulados) para que haja uma padronização que permita uma comparação mais justa entre algoritmos implementados por diferentes autores.

Assim sendo, desenvolveu-se um *benchmark* de distúrbios elétricos simulados (chamado SED-Bench), variando-se frequência, taxa de sinal-ruído (SNR) e os diferentes tipos de distúrbios isolados (interrupção, abaixamento de tensão, elevação de tensão, transiente oscilatório, transiente impulsivo, flutuação de tensão, nível DC, harmônicas e *notching*), duplos (combinação de dois isolados num mesmo ciclo) e triplos (combinação de três isolados num mesmo ciclo), além da senoide pura.

Nas próximas Seções são apresentados com maiores detalhes como se deu a implementação do *benchmark* SED-Bench bem como sua composição.

4.2 Metodologia

Esta Seção apresenta a metodologia seguida para criar o *benchmark* SED-Bench (de distúrbios elétricos simulados seguindo o padrão IEEE [Association et al. 2009]) e que contenha:

- Distúrbios com diferentes frequências fundamentais (50Hz e 60Hz);
- Distúrbios com ruídos sob diferentes valores de SNR (20dB, 30dB, 50dB, 60dB, 70dB);
- Distúrbios isolados, duplos e triplos (baseado nos distúrbios encontrados nos experimentos apresentados nos trabalhos relacionados);

- Disponibilizar o código gerador de distúrbios elétricos para a comunidade acadêmica bem como o *benchmark* proposto;

Assim sendo, pretende-se abranger a maioria dos possíveis cenários nos quais os distúrbios elétricos monofásicos podem ser apresentados e, por conseguinte, disponibilizar para a comunidade acadêmica uma forma de verificar a validade dos distúrbios presentes no *benchmark*, sendo disponibilizado o código fonte (em linguagem Python) que gerou tais distúrbios.

4.2.1 Desenvolvimento do *Benchmark*

Para o desenvolvimento do benchmark, analisou-se primeiramente a modelagem matemática dos diferentes distúrbios elétricos, seguindo o padrão determinado em [Association et al. 2009] e, tendo como base a modelagem matemática apresentada na Seção A.

Utilizou-se uma taxa de amostragem de 10KHz, dessa forma, de acordo com o Teorema de Nyquist [McClellan et al. 2003], a frequência máxima do sinal amostrado deve ser de 5KHz. Assim sendo, alguns distúrbios de alta e média frequência foram deixados de fora do *benchmark*, como é o caso dos distúrbios de transiente oscilatório, que entrou apenas nos casos de baixa frequência, e o transiente impulsivo com variações na ordem de nanosegundos.

Além disso, foram utilizados 500 amostras, o que dá um total de três ciclos para os sinais com frequência de 60Hz e aproximadamente três ciclos para os sinais com frequência de 50Hz. A quantidade de ciclos foi escolhida com intuito de utilizar a menor quantidade possível de amostras, de forma a proporcionar maior agilidade nos algoritmos; a partir de então, verificou-se que a frequência do distúrbio *notching* é a mais baixa dentre os distúrbios simulados (mas não menor do que a frequência fundamental), logo, pode-se verificar um distúrbio *notching* (e não confundi-lo com um distúrbio de transiente impulsivo) utilizando os 3 ciclos propostos.

Outra variável importante no *benchmark* é a frequência fundamental, de 50Hz ou 60Hz. Seu valor teve pequenas variações de $\pm 5\%$, dentro dos padrões permitidos nas redes de distribuição elétrica [Kennedy 2000].

Por fim, tem-se dois outros pontos importantes no *benchmark*: o valor de SNR e a quantidade de distúrbios no sinal. Houve uma variação no valor SNR de 20dB, 30dB, 50dB, 60dB, 70dB e o sinal sem ruído, para que os usuários do *benchmark* possam verificar comportamen-

tos de algoritmos com um aumento gradual do ruído nos dados, apesar que, segundo [Tomic et al. 2007] o valor de SNR varia entre 50dB e 70dB apenas.

Já em relação à quantidade de distúrbios, a Tabela 4.1 apresenta quais foram os distúrbios isolados, duplos e triplos inseridos no *benchmark*.

Tabela 4.1: Distúrbios

Isolados	Duplos	Triplos
Seno	Sag + Harm	Sag + Harm + Osc
Nível DC	Swell + Harm	Swell + Harm + Osc
Oscilação	Sag + <i>Notching</i>	Sag + Harm + <i>Notching</i>
Trans. Impulsivo	Sag + <i>Notching</i>	Swell + Harm + <i>Notching</i>
<i>Flicker</i>	Sag + Osc	
Sag	Swell + Osc	
Swell	Sag + Imp	
Harmônica	Swell + Imp	
Interrupção		
<i>Notching</i>		

Dessa forma, para resumir, a Tabela 4.2 apresenta as variáveis alteradas para criação do *benchmark* e seus respectivos valores.

Tabela 4.2: Variáveis

Variável	Valores
Taxa de Amostragem	10KHz
Frequência Fundamental	50Hz e 60Hz ($\pm 5\%$)
Tamanho do Sinal	500 Amostras
SNR	20dB, 30dB, 50dB, 60dB, 70dB e Sem Ruído
Distúrbios	10 Isolados, 8 Duplos, 4 Triplos
Número de Dados de Treino	1000/Classe
Número de Dados de Teste	50 Pastas com 200/Classe

Dessa forma, criou-se vários conjuntos de dados (tanto para teste quanto para treino):

- Distúrbios Isolados sem Ruído (60Hz)
- Distúrbios Isolados com Ruído (60Hz)(20dB, 30dB, 50dB, 60dB, 70dB)
- Distúrbios Isolados sem Ruído (50Hz)
- Distúrbios Isolados com Ruído (50Hz)(20dB, 30dB, 50dB, 60dB, 70dB)
- Distúrbios Múltiplos sem Ruído (60Hz)
- Distúrbios Múltiplos com Ruído (60Hz)(20dB, 30dB, 50dB, 60dB, 70dB)
- Distúrbios Múltiplos sem Ruído (50Hz)
- Distúrbios Múltiplos com Ruído (50Hz)(20dB, 30dB, 50dB, 60dB, 70dB)

Ficando, portanto, a critério do pesquisador que utilizar o *benchmark* quais os dados que ele utilizará.

4.3 SED-Bench

O *benchmark* SED-Bench foi implementado na Linguagem Python, armazenado em arquivo utilizando a biblioteca *numpy* [Jones et al. 01 a] e disponibilizado através do link¹. O código gerador do SED-Bench está disponível para os pesquisadores através do link².

4.3.1 Organização

Os dados estão organizados da seguinte maneira:

¹<https://drive.google.com/open?id=1suJiEi9KYjLKCBwZHbJMP44Gl2RhXwuC>

²https://drive.google.com/open?id=1s1KNQ63__MOy6CU6neGoARY3GSyTNKdr

```

SED - Benchmark
├── bench_test
│   ├── 50 Hz
│   │   ├── Multiplos
│   │   │   └── ...
│   │   └── Isolados
│   │       └── ...
│   └── 60 Hz
│       ├── Multiplos
│       │   └── ...
│       └── Isolados
│           └── ...
├── bench_train
│   ├── 50 Hz
│   │   ├── Multiplos
│   │   │   └── ...
│   │   └── Isolados
│   │       └── ...
│   └── 60 Hz
│       ├── Multiplos
│       │   └── ...
│       └── Isolados
│           └── ...
└── Source Code
    └── disturbances_generator.py

```

Sendo que os dados estão salvos com nomes como mostra a Tabela 4.3.

Local	Padrão de Arquivos
Dados de Teste	bc0_(0..49).npz; bc20_(0..49).npz; bc30_(0..49).npz; bc50_(0..49).npz; bc60_(0..49).npz
Dados de Treino	bc0_0.npz; bc20_0.npz; bc30_0.npz; bc50_0.npz; bc60_0.npz; bc70_0.npz

Tabela 4.3: Dados Nas Pastas Múltiplos e Isolados

Em que os número após o *bc* representa:

- 0 - Sem ruído
- 20 - SNR de 20dB
- 30 - SNR de 30dB
- 50 - SNR de 50dB
- 60 - SNR de 60dB
- 70 - SNR de 70dB

Logo, se um usuário do SED-Bench quiser utilizar os dados de treino com distúrbios múltiplos, com 20dB de SNR, com uma frequência de 50Hz, basta acessar os dados na pasta SED - Benchmark/bench_train/50Hz/Multiplos/bc20_0.npy. Lembrando que os distúrbios encontrados nos arquivos múltiplos são tanto os distúrbios duplos quanto os triplos, apresentados na Tabela 4.1.

4.3.2 Comparativo

Como o único trabalho que disponibiliza dados para o uso de outros autores é o apresentado em [Association et al. 2009]. Na Tabela 4.4 é feita uma análise comparativa entre estes dados disponibilizados e o *benchmark* proposto (SED-Bench).

4.4 Conclusão do Capítulo

Pesquisas recentes buscam melhorar as maneiras de detectar e classificar os distúrbios elétricos para, com isto, evitar grandes perdas tanto de equipamentos quanto de recursos. No entanto, apesar de existir modelagens matemáticas para tais distúrbios, não existe na literatura um trabalho que disponibilize os dados para a comunidade acadêmica, tornando difícil a realização de comparações entre algoritmos propostos por diferentes autores.

Desta forma, o *benchmark* proposto tem o intuito de facilitar esta comparação entre os diferentes trabalhos de pesquisa sobre distúrbios elétricos, além de disponibilizar o código fonte, de forma a permitir que diferentes autores possam sugerir modificações, como por

Tabela 4.4: Análise Comparativa

Item	IEEE Std 1159-2009 [Association et al. 2009]	Proposto
Quantidade	< 100	1000 de cada classe
Classes	Indisponível	22
Ruído	Indisponível	SNR de 20dB, 30dB, 50dB, 60dB e 70dB
Frequência	60Hz	50Hz e 60Hz
Natureza dos Dados	Reais	Simulado
Dados Disponíveis	Sob demanda	Aberto

exemplo a inclusão de uma nova faixa de SNR, a adição de uma nova frequência fundamental, o aumento na quantidade de ciclos da senoide e a variação na quantidade de pontos amostrais (consequentemente modificando a taxa de amostragem).

Capítulo 5

Metodologia Experimental

A proposta central desta dissertação consiste de diferentes experimentos relacionados à análise da classificação de distúrbios elétricos utilizando diferentes algoritmos de aprendizagem de máquina.

Desta forma, este capítulo apresenta a metodologia seguida por todos os cinco experimentos, listados a seguir, cujos resultados são apresentados no próximo capítulo.

- Experimento 1: Análise de Algoritmos de Classificação no Raspberry Pi;
- Experimento 2: Análise de Algoritmos para Classificação de Distúrbios Isolados na presença de Ruído sem Extração de Características;
- Experimento 3: Análise de Algoritmos para Classificação de Distúrbios Isolados na presença de Ruído com Extração de Características;
- Experimento 4: Comparação de Plataformas de Hardware (NodeMCU e Raspberry Pi);
- Experimento 5: Análise da Classificação de Distúrbios Múltiplos usando *Random Forest*;

Sendo assim, seguindo a metodologia experimental descrita em [Wohlin et al. 2012], as próximas seções apresentam a definição e a execução de forma generalizada dos experimentos mencionados.

5.1 Definição do Experimento

5.1.1 Definição de Objetivo

O objetivo dos experimentos consistem da avaliação, segundo determinadas métricas (acurácia, tempo de processamento, memória utilizada e/ou potência consumida, a depender do experimento), usando diferentes algoritmos para classificação de distúrbios elétricos, verificando seu comportamento em ambientes ruidosos ou não ruidosos utilizando como base a modelagem matemática dos distúrbios elétricos e o SED-Bench apresentado no Capítulo 4.

5.1.2 Planejamento

Seleção de Contexto: os experimentos *in vitro* utilizaram um conjunto de distúrbios simulados isolados (harmônicas, interrupção, afundamento e elevação de tensão, *notching*, transiente oscilatório, flutuação de tensão, transiente impulsivo e nível DC), além da senoide pura (também a ser classificada), sendo estes sinais os mais comuns nos trabalhos, como visto no Capítulo 3.

Ainda foi aplicado, em alguns experimentos, ruído gaussiano sobre o conjunto de dados com taxas SNR de 30dB, 50dB, 60dB e 70dB, o que permitiu verificar o comportamento das técnicas utilizadas em ambientes ruidosos (comum em redes elétricas) e o quanto isto influenciava na acurácia dos algoritmos de classificação.

Este conjunto de valores de SNR foi utilizado com base nos resultados apresentados em [Tomic et al. 2007], o qual sugere que em redes reais a variação de valor de SNR ocorre entre 50dB e 70dB. Adicionou-se então ainda o ruído com SNR a 30dB para verificar o comportamento dos algoritmos em ambientes cujo ruído seja mais alto que o normal.

No caso do último experimento, foi acrescentado os distúrbios múltiplos (duplos e triplos) para classificação, além dos distúrbios isolados.

Os algoritmos utilizados nos experimentos foram:

- Experimento 1: *Random Forest*, SVM, MLP, Árvore de Decisão;
- Experimento 2: SVM, MLP, FFT, DWT e Filtro Notch;

- Experimento 3: *Random Forest*, FFT, Transformada de Hilbert, Filtro Gaussiano, Estatística Descritiva;
- Experimento 4: Árvore de Decisão, FFT, Estatística Descritiva;
- Experimento 5: *Random Forest*, FFT, Transformada de Hilbert, Filtro Gaussiano, Dilatação, Estatística Descritiva;

Variáveis Independentes: Conjunto de dados com dez (ou vinte e dois) sinais com classes distintas, algoritmos de classificação, algoritmos de pré-processamento (quando aplicados) e técnicas para extração de características (quando aplicado).

Variáveis Dependentes: Hiperparâmetros dos algoritmos de classificação, acurácia média para cada distúrbio (τ , Eq. 5.1), acurácia média geral (μ , Eq. 5.2), e métricas de eficiência (a depender do experimento) tais como tempo de processamento, memória utilizada e potência consumida.

$$\tau_{alg}^i \quad (5.1)$$

, sendo i o distúrbio, e alg o algoritmo.

$$\mu_{alg} = 1/n \sum_{(i=1)}^n \tau_{alg}^i \quad (5.2)$$

, sendo n o número de distúrbios utilizados no experimento.

Formulação de Hipótese: Cada experimento tem uma hipótese distinta que são apresentadas em suas devidas seções da presente dissertação.

Seleção de Objetos: Todos os experimentos utilizaram os dados provenientes do benchmark SED-Bench apresentado no Capítulo 4. Deve-se ressaltar que as 50 bases de dados para captura da acurácia foram utilizadas de forma a aplicar testes estatísticos mais robustos sobre os resultados alcançados.

Alguns experimentos (1 e 4) utilizaram plataformas de hardware para implementação de um sistema embarcado, em tais experimentos foram utilizados os dados do benchmark para o treinamento, entretanto, na captura das métricas de eficiência utilizou-se o código fonte do benchmark para geração dos dados em tempo de execução. Neste caso, foi executado 50

vezes o código, gerando 20 dados de cada classe para a aquisição das métricas, sendo quatro com diferentes valores de SNR (30dB, 50dB, 60dB, 70dB e sem ruído).

Projeto do Experimento: o projeto dos experimentos podem ser resumidos em seis diferentes etapas (ou cinco, a depender do experimento), como mostra a Figura 5.1.

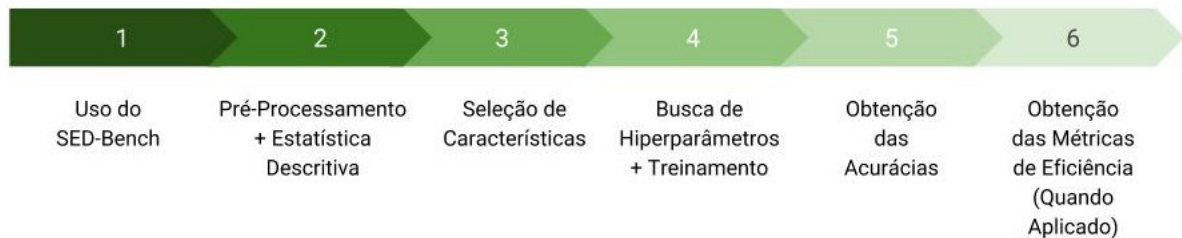


Figura 5.1: Projeto do Experimento

Instrumentação: os algoritmos utilizados foram obtidos a partir de bibliotecas provenientes da linguagem de programação Python na versão 3.5.1. Os algoritmos foram executados em um computador pessoal, com 4Gb de memória RAM, processador Intel i5-3470S (2.9GHz), sistema operacional Ubuntu 12.04.5 LTS e com o software Anaconda 3 instalado (ambiente Python no qual foram executados os experimentos).

No caso dos experimentos que utilizaram plataformas de hardware, utilizou-se como instrumento o Raspberry e/ou o NodeMCU com as características apresentadas a seguir:

- Raspberry Pi 3, com 1GB de RAM e um processador BCM2837 de quatro núcleos de 64 bits ARM Cortex-A53 a 1.2GHz;
- NodeMCU Amica, 32 bits, com 80MHz de clock, RAM de 160KB e Flash de 16Mb;

5.2 Operação do Experimento

A Figura 5.2 apresenta um resumo da operação dos experimentos realizados.

As etapas da operação marcadas com um asterisco (*) não são executadas em todos os experimentos. No caso da etapa de extração de características durante a execução, apenas o

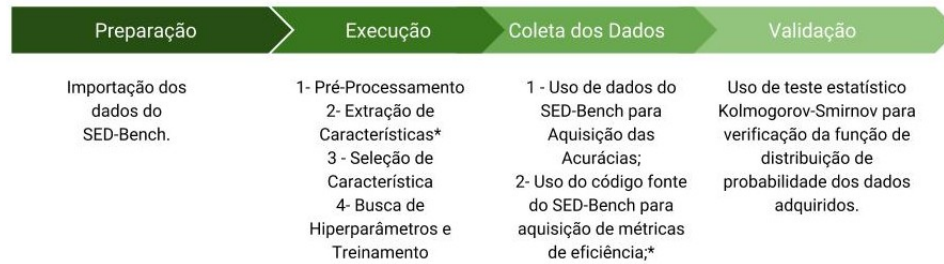


Figura 5.2: Operação do Experimento

experimento 1 e 2 não realizaram esse procedimento. Já na coleta de dados, a etapa de aquisição de métricas de eficiência foi executada apenas pelos experimentos 1 e 4 (que utilizam dispositivos de hardware).

Durante a etapa de execução, obteve-se a melhor configuração dos algoritmos de classificação utilizando a técnica *GridSearch* (disponível na biblioteca *scikit-learn* [Pedregosa et al. 2011] em Python). Para tanto, variam-se os hiper-parâmetros em busca da melhor acurácia.

Além disso, na etapa de treinamento (ainda na execução) foi utilizada a técnica K-fold, com K=5, de forma a conseguir um treinamento melhor dos algoritmos, evitando superajuste.

A técnica *GridSearch* permite a busca dos melhores hiperparâmetros ao mesmo tempo e que o algoritmo é treinado. Tal busca é feita usando *brute force* (força-bruta), onde vários hiper-parâmetros são passados e todas as combinações entre eles são feitas, com o intuito de encontrar aquela na qual o algoritmo encontre um valor máximo da métrica que busca-se otimizar (neste caso a acurácia).

Na etapa de coleta de dados, utilizou-se como auxílio as bibliotecas *scikit-learn* [Pedregosa et al. 2011], *numpy* [Jones et al. 01 a] e *joblib* [Developpers 09], a partir das quais foi possível realizar a coleta dos resultados e salvá-los em arquivo para futuras análises estatísticas.

No caso da coleta de dados dos experimentos que utilizaram as plataformas de hardware, a coleta de dados foi feita por meio de inserção de códigos para captura do tempo de processamento da classificação. Por outro lado, para captura da potência consumida usou-se dois multímetros, um na função de amperímetro e outro na função de voltagem; e para aquisição da memória consumida usou-se a interface *htop* do Linux para aferição da memória usada pelo processo de classificação.

Por fim, em relação à validação dos dados, a partir do resultado do teste Kolmogorov-

Smirnov, normalmente um outro é utilizado para análise das hipóteses de cada experimento, podendo ser um teste paramétrico ou não-paramétrico, a depender da função de distribuição de probabilidade dos dados.

Capítulo 6

Experimentos

Este capítulo apresenta os experimentos realizados para análise de formas de melhorar a classificação dos distúrbios elétricos utilizando algoritmos de aprendizagem de máquina.

O primeiro experimento decorre sobre diferentes classificadores implementados sem aplicação de técnicas de pré-processamento ou extração de características nos dados de entrada, além de fazer uma análise em ambiente embarcado.

O segundo experimento acrescenta a utilização de técnicas de pré-processamento mas ainda sem a utilização de extração de características.

O terceiro experimento por fim acrescenta extratores de características utilizando estatística descritiva e de alta ordem.

O quarto experimento apresenta uma comparação da criação de sistemas de classificação em plataformas de hardware (NodeMCU e Raspberry Pi), fazendo a comparação em termos de tempo de processamento.

Por fim, o quinto experimento faz uma análise da classificação de distúrbios múltiplos utilizando o algoritmo *Random Forest* e diferentes formas de pré-processamento (FFT, filtro gaussiano, transformada de Hilbert e dilatação), além de estatística descritiva e cálculos matemáticos para a extração das características.

6.1 Análise da Classificação de Distúrbios Elétricos em Raspberry Pi (Experimento 1)

Como consequência do crescente problema com a QEE [Association et al. 2009, Samotyj 2001, Targosz and Manson 2007], muitos sistemas de monitoramento vêm sendo propostos como forma de prevenir, mitigar ou agilizar o processo de correção na rede elétrica, evitando assim que danos maiores possam ocorrer nos equipamentos [Ribeiro et al. 2016, Gonzalez-Cordoba et al. 2015, Dalai et al. 2013, Carniato et al. 2009, Back 2013] devido aos distúrbios elétricos.

Para a construção de sistemas que possam realizar tal procedimento, faz-se necessário a implementação de algoritmos capazes de detectar e/ou classificar os distúrbios na rede elétrica. Além disso, sendo requerido respostas em tempo real, a utilização de sistemas embarcados para este fim tem crescido, como visto no mapeamento presente na Seção D.

Neste contexto, o presente trabalho tem como objetivo apresentar uma comparação entre algoritmos distintos para classificação dos distúrbios elétricos no ambiente embarcado Raspberry Pi 3, sendo eles:

- MLP (Rede Neural Perceptron Multicamadas), devido sua ampla utilização em trabalhos relacionados a criação de sistemas de monitoramento de distúrbios elétricos [Ribeiro et al. 2016, Li et al. 2016, Zhang and Li 2010, Wang et al. 2004, Rodriguez et al. 2016];
- SVM, um algoritmo mais sofisticado e pouco utilizado na classificação de distúrbios elétricos até 2006, mas considerado uma boa alternativa devido à melhor generalização na classificação [Bollen and Gu 2006] sendo, até então, utilizado em alguns trabalhos em conjunto com outros algoritmos de pré-processamento como em [Li et al. 2016, Ruiz-Llata et al. 2011, Wang et al. 2012, Cerqueira et al. 2008]
- Árvore de Decisão, utilizado em trabalhos como [Zhang and Li 2010, Gonzalez-Cordoba et al. 2015, Li et al. 2016, Ribeiro et al. 2016, Meena et al. 2018], normalmente aplicado quando se tem um conjunto de características extraídas por algum outro algoritmo de pré-processamento.

- Florestas Aleatórias (chamado a partir deste ponto como *Random Forest*), que tem tido resultados promissores em termos de acurácia quando da classificação de distúrbios elétricos [Borges et al. 2015, Borges et al. 2016, Markovska and Taskovski 2017, Feng et al. 2018]

Como forma de avaliar como cada algoritmo se comporta no Raspberry Pi e compará-los, seguiu-se a metodologia experimental descrita no Capítulo 5. Abaixo são apresentados as hipóteses, resultados e conclusões do experimento.

Deve-se ressaltar que este experimento não trabalhou com algoritmos de pré-processamento, extração de características, nem dados ruidosos, ou seja, o dado bruto foi utilizado na classificação, aplicando-se apenas técnicas de seleção de características (PCA, no caso do MLP e do SVM, e os demais algoritmos, *Random Forest* e Árvore de Decisão fazem isso automaticamente pela forma como são implementados).

6.1.1 Formulação de Hipótese

As questões de pesquisa para este experimento são: algum dos algoritmos (SVM, MLP, Árvore de Decisão ou *Random Forest*) pode ser considerado melhor, em termos de eficácia, na tarefa de classificação dos distúrbios elétricos? Qual o melhor algoritmo para implementação em ambiente embarcado quando considerado a eficiência (tempo de processamento, uso de memória e potência consumida)?

Para a primeira questão pode-se considerar a acurácia média como métrica para avaliar e comparar os resultados, enquanto que para a segunda questão pode-se considerar as métricas capturadas pela implementação no Raspberry Pi 3 dos algoritmos treinados.

Dentro destas questões, as seguintes hipóteses podem ser verificadas (a saber, DT = Árvore de Decisão e RF = *Random Forest*):

Hipótese 1

H0: Não há diferença entre as acurácias médias dos algoritmos $\mu_{SVM} = \mu_{MLP} = \mu_{DT} = \mu_{RF}$.

H1: Algum algoritmo possui acurácia diferente em relação aos demais. ($\exists \mu_{alg_x} \neq \mu_{alg_y}$).

Hipótese 2

H0: Não há diferença entre as métricas de eficiência dos algoritmos $\mu_e^{SVM} = \mu_e^{MLP} =$

$$\mu_e^{DT} = \mu_e^{RF}.$$

H1: Algum dos algoritmos possui eficiência diferente em relação aos demais. ($\exists \mu_e^{algx} \neq \mu_e^{algy}$).

Esta segunda hipótese se desdobra em 3 hipóteses diferentes, uma para cada métrica de eficiência analisada (tempo de processamento, uso de memória e potência consumida).

6.1.2 Resultados

Primeiramente, como forma de ter uma visualização geral dos resultados obtidos, as Tabelas 6.1 e 6.2 apresentam os melhores hiperparâmetros selecionados para cada algoritmo e as métricas obtidas para estes, respectivamente.

Tabela 6.1: Tabela hiperparâmetros	
Algoritmos	hiperparâmetros
MLP	Número de Iterações: 1000
	Função de Ativação: Tangente Hiperbólica
	Função de Atualização dos Pesos: Método Quasi-Newton (<i>lbfgs</i>)
	Camadas: (5, 15, 10)
	Fator de Regularização L2 (α) : 1e-3
	Números Componentes PCA : 5
<i>Random Forest</i>	Número de Árvores: 500
	Critério de Seleção : Entropia
	Número Máximo de Características: $\sqrt{Características}$
Árvore de Decisão (CART)	Critério de Seleção: Entropia
	Pre Ordenação : Sim
	Número Máximo de Características : \log_2
	<i>Splitter</i> : <i>best</i>
SVM	Kernel: RBF
	Fator de Regularização (C) : 1e+6
	Abordagem de Classificação : Um contra Um (<i>ovo</i>)
	Números Componentes PCA : 8

Tabela 6.2: Média e Desvio das Métricas Avaliadas

Algoritmo	Acurácia	Tempo (μ s)	Memória	Potência (W)
MLP	85.24% \pm 2.9%	249.48 \pm 5.39	4.6%	1.605
RF	99.44% \pm 0.11%	33660 \pm 978	6.0%	1.652
SVM	92.32% \pm 0.32%	870 \pm 6	4.5%	1.601
DT	95.74% \pm 0.33%	140 \pm 6	4.4%	1.642

Como pode-se observar, o tempo de processamento do *Random Forest* é muito superior ao dos demais algoritmos, em relação à Árvore de Decisão, por exemplo, o *Random Forest* possui um tempo mais de 240 vezes maior, desta forma, tendo em vista que o período de captura do sinal para classificação é de 0.05s (três ciclos de 60Hz), e o de classificação com o *Random Forest* é 0.033s, muitos dados seriam perdidos enquanto o algoritmo faz a predição, mais de metade dos três ciclos.

Diante deste problema, os hiperparâmetros do *Random Forest* foram revistos, diminuindo o número de árvores e verificando o quanto era perdido em termos de acurácia. Assim sendo, verificou-se que, com os mesmos hiperparâmetros representados na Tabela 6.1, modificando apenas o número de árvores para nove, obteve-se o resultado de acurácia, tempo de processamento, uso de memória e potência consumida apresentados na Tabela 6.3.

Tabela 6.3: Médias e Desvio do *Random Forest* com Nove Árvores

Algoritmo	Acurácia	Tempo (μ s)	Memória	Potência (W)
RF (9 Árvores)	99.00% \pm 0.33%	855 \pm 44	4.6%	1.645

Como pode-se verificar pela Tabela 6.3, houve uma redução de, aproximadamente, 46 vezes no tempo de processamento em comparação com o *Random Forest* com 500 árvores, sendo que a acurácia média teve uma redução de apenas 0.44%. Assim sendo, optou-se por utilizar o *Random Forest* com 9 árvores. Portanto, quando se tratar do algoritmo de *Random Forest* deste ponto em diante desta Seção, estar-se-á referindo ao *Random Forest* com 9 árvores de decisão.

Nesse contexto, como forma de avaliação das métricas apresentadas e sua interpretação, a seguir é apresentado mais detalhadamente esses resultados com as devidas avaliações estatísticas, tendo-se em vista as hipóteses levantadas.

Análise e Interpretação

Os resultados encontram-se divididos em duas etapas distintas:

1. Análise da acurácia média dos diferentes algoritmos (SVM, MLP, Árvore de Decisão e *Random Forest*) (Hipótese 1);
2. Comparação da métricas de eficiência em ambiente embarcado (tempo de processamento, uso de memória e potência consumida) (Hipótese 2).

Análise de Acurácia Média O teste estatístico de Friedman foi selecionado e aplicado para o conjunto de dados de acurácia visto que o resultado do teste Kolmogorov-Smirnov (KS) evidenciara que os dados não possuem distribuição normal (p -value próximo de zero, podendo portanto rejeitar a hipótese nula de que os dados seguem uma distribuição normal), portanto, um teste não-paramétrico deve ser utilizado, e, devido à forma como o problema é composto, de vários tratamentos para um mesmo conjunto de dados, o teste de Friedman foi escolhido e aplicado sobre os resultados de acurácia média.

Como resultado da aplicação do Teste de Friedman, obteve-se um p -value próximo de zero, evidenciando assim, a existência de diferença estatística entre os resultados dos quatro algoritmos. A partir do teste de Friedman (implementado em Python por [Rodríguez-Fdez et al. 2015]) foi possível construir um ranqueamento, o qual indica o algoritmo que obteve melhor desempenho quando levado em consideração a acurácia, ver Tabela 6.4.

Tabela 6.4: Classificação do melhor algoritmo de acordo com a acurácia média

Algoritmo	Classificação
<i>Random Forest</i>	1º
Árvore de Decisão (CART)	2º
SVM	3º
MLP	4º

Estima-se que o *Random Forest* obteve o melhor resultado devido sua capacidade de fazer "seleção natural" das características que melhor representam o problema. Foram aplicados métodos de seleção de características (PCA) para SVM e MLP, porém, os resultados levam

a crer que a seleção com PCA não conseguiu se sobressair à seleção natural realizada no *Random Forest* para o domínio do problema em questão.

Análise das Métricas de Eficiência em Ambiente Embarcado Assim como a acurácia média, o tempo de processamento também não possuiu uma distribuição normal segundo o resultado do Teste KS aplicado (*p-value* próximo de zero), desta forma, foi aplicado o teste de Friedman, para o qual o *p-value* retornado foi de zero (ou muito próximo disso), indicando assim a existência de diferença estatística entre os conjuntos de dados. A Tabela 6.5 apresenta a classificação construída a partir do teste de Friedman para o tempo de processamento e o resultado geral do uso da memória e potência consumida (que, como não houve variação, não foi necessário aplicar teste estatístico, $\sigma = 0$).

Tabela 6.5: Classificação dos algoritmos em termo das métricas de eficiência

Algoritmo	Tempo	Memória	Potência Dissipada
DT	1°	1°	3°
MLP	2°	2°	2°
<i>Random Forest</i>	3°	2°	4°
SVM	3°	1°	1°

Os resultados disponíveis na Tabela 6.5 ilustram que o algoritmo de Árvore de Decisão possui o melhor desempenho em termos de tempo de processamento e uso de memória (nesta métrica empatando com o SVM). Este resultado era esperado devido a forma como é realizada a classificação utilizando Árvores de Decisão (a partir de uma estrutura de lógica condicional), enquanto que os demais algoritmos apresentam diversas multiplicações vetoriais (SVM e MLP) ou múltiplas árvores de decisão (*Random Forest*).

Além disso, os algoritmos de MLP e SVM, que apresentaram *overfitting*, tiveram que implementar etapas de pré-processamento e seleção de características com algoritmos externos como o PCA, a regularização L2 e a normalização dos dados, enquanto que os algoritmos de *Random Forest* e Árvore de Decisão utilizaram apenas a normalização, reduzindo ainda mais o tempo de processamento deste último em relação aos demais.

No entanto, verifica-se que em termos de potência consumida o algoritmo de SVM e sobressaiu em relação aos demais, apesar de ficar em terceiro em tempo de processamento.

Desta forma, verificando-se na Tabela 6.2 que a diferença em termos de potência consumida é pouco (menor que 0.1W), a escolha do algoritmo determinada por essa métrica não é recomendada.

Sendo assim, para um desenvolvedor de sistema de monitoramento em tempo real, com estes resultados, seria mais interessante verificar apenas as métricas de tempo de processamento e acurácia, visto que as duas outras possuem pouca variação entre os algoritmos.

6.1.3 Conclusão do Experimento

O experimento apresentado nesta Seção mostrou uma abordagem experimental relacionada à Qualidade da Energia Elétrica, avaliando diferentes combinações de algoritmos de classificação para utilização em sistema de monitoramento (abordando o Raspberry Pi 3 como ambiente embarcado).

Foi possível verificar que a classificação utilizando *Random Forest* obteve um resultado superior em termos de acurácia (99%), mas ficou abaixo em termos de tempo de processamento (855ms), uso de memória e potência consumida. Em contrapartida, o algoritmo de Árvore de Decisão ficou com acurácia (95.74%) acima dos demais (SVM e MLP) e obteve o melhor resultado em tempo de processamento (140ms).

Portanto, fica a critério do desenvolvedor qual aspecto deve ser priorizado na construção de um modelo de predição para sistemas de monitoramento de distúrbios elétricos. Em um ambiente médico, por exemplo, onde falhas elétricas pode ocasionar na perda de vidas, o tempo de resposta é muito mais valioso que a acurácia do algoritmo, e, portanto, Árvore de Decisão poderia ser escolhido para este caso; já em caso de indústrias de pequeno porte, onde a falha de equipamento não ocasione em perdas desastrosas, ou mesmo em ambientes nos quais a classificação dos distúrbios é realizada de maneira offline, pode-se priorizar a acurácia, para evitar a chamada de técnicos externos devido a uma classificação incorreta de determinado distúrbio ocorrido na rede.

6.2 Algoritmos para Classificação de Distúrbios Elétricos Ruidosos sem Extração de Características (Experimento 2)

Diversos sistemas têm sido propostos para o processo de detecção de distúrbios elétricos ruidosos [Zhong et al. 2018, do Amaral Monteiro et al. 2018, Li et al. 2016, He et al. 2013, Zhang and Li 2010], trabalhando com diferentes classificadores e valores de taxa sinal-ruído que variam entre 20dB e 50dB.

Nesse contexto, buscou-se apresentar uma comparação entre dois algoritmos de classificação: o MLP, devido sua ampla utilização em trabalhos relacionados à criação de sistemas de monitoramento de distúrbios elétricos [Ribeiro et al. 2016, Li et al. 2016, Zhang and Li 2010, Wang et al. 2004, Rodriguez et al. 2016] e o SVM, um algoritmo mais sofisticado e que tem a propriedade de ser mais genérico em sua classificação, sendo um algoritmo pouco utilizado na classificação de distúrbios elétricos até 2006, mas considerado uma boa alternativa devido a esta generalização na classificação [Bollen and Gu 2006] sendo, até então, utilizado em alguns trabalhos em conjunto com outros algoritmos de pré-processamento como em [Li et al. 2016, Ruiz-Llata et al. 2011, Wang et al. 2012, Cerqueira et al. 2008].

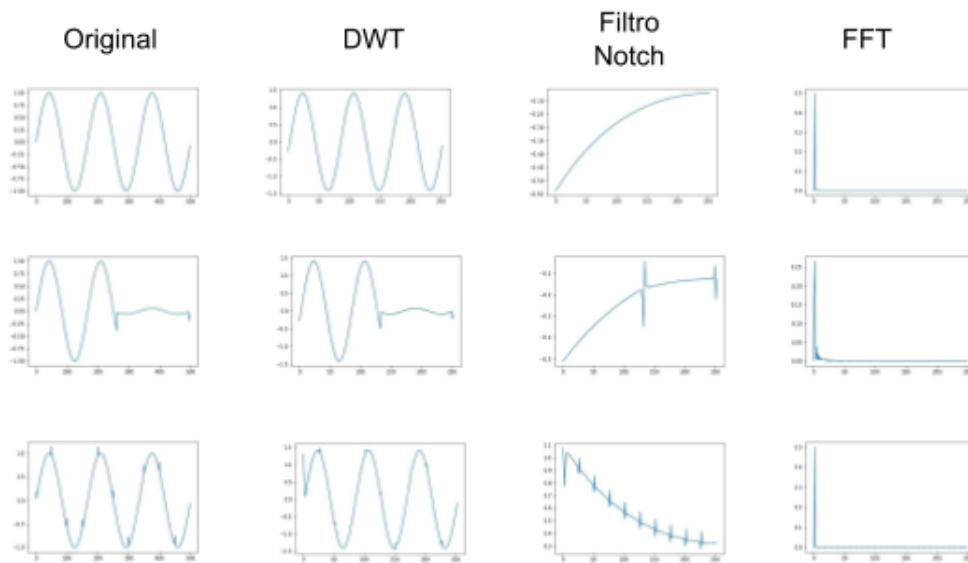
Além disso, propôs-se trabalhar com frequências de 30dB a 70dB, que são frequências mais comuns nas redes elétricas [Tomic et al. 2007].

No intuito de melhorar a classificação dos dados ruidosos e não aumentar significativamente o tempo de processamento, acrescentou-se técnicas de pré-processamento dos dados para serem também analisadas, sendo elas facilmente implementadas em dispositivos embarcados (FFT, DWT e filtro notch), além da seleção de características utilizando o algoritmo de *Random Forest* para tal (devido ao resultado encontrado no experimento 1). Os resultados destas técnicas em alguns dos distúrbios classificados é apresentado na Figura 6.1.

Como pode-se visualizar pela imagem, o DWT foi utilizado para reduzir pela metade a quantidade de amostras do sinal, o Filtro Notch reduz a amplitude na frequência fundamental do sinal (60Hz), e a FFT transforma os dados para o domínio da frequência.

Ainda em relação às técnicas de pré-processamento empregadas, para o cálculo da DWT utilizou-se a Wavelet mãe *db-4* e obteve-se os coeficientes de aproximação do sinal; da FFT

Figura 6.1: Pré-Processamento Distúrbios (Experimento 2)



utilizou-se a magnitude do sinal transformado; e, para o filtro notch utilizou-se um filtro IIR implementado com auxílio da biblioteca Scipy [Jones et al. 01 b] em Python, usando como parâmetros a frequência fundamental de 60Hz, a frequência amostral de 10KHz e o fator de qualidade igual a 0.1 para atenuar ainda mais a amplitude na frequência especificada.

Os algoritmos foram utilizados em separado para verificação de qual seria o melhor utilizando-os separadamente, como mostra a Figura 6.2, entretanto, poder-se-ia aplicar as técnicas em conjunto, como por exemplo aplicar a FFT após a aplicação da DWT, reduzindo assim tempo de processamento da FFT (apesar da perda de alguns detalhes).

Um experimento controlado foi implementado para análise comparativa das técnicas de classificação, seguindo a metodologia descrita no Capítulo 5.

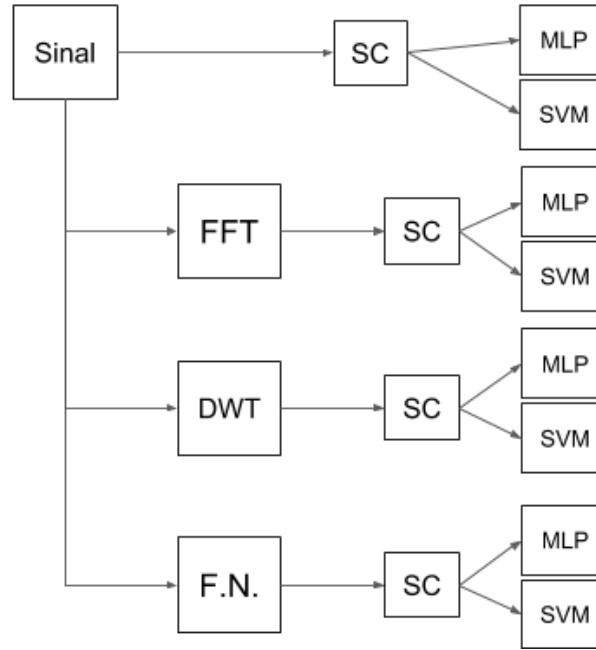


Figura 6.2: Diagrama de Uso de Técnicas de Pré-Processamento (Experimento 2)

6.2.1 Formulação de Hipótese

As questões de pesquisa para este experimento são: os algoritmos SVM ou MLP puros ou alguma combinação destes com FFT, DWT ou filtro notch pode ser considerado melhor para classificação geral dos distúrbios elétricos? Qual a melhor técnica, caso exista, para classificar cada distúrbio específico? Qual a influência do ruído na acurácia dos algoritmos?

Para a primeira e terceira questão pode-se considerar a acurácia média geral (μ) para avaliar e comparar os resultados, e, para a segunda questão, a acurácia média individual (τ). Neste contexto, as seguintes hipóteses podem ser verificadas (a saber, S=SVM, M=MLP, F=FFT, W=DWT, N=Filtro Notch):

Hipótese 1 (Para cada um dos distúrbios)

H0: Não há diferença estatística entre as acurácias médias individuais, ou seja, $(\tau_S^i = \tau_{SW}^i = \tau_{SN}^i = \tau_{SF}^i = \tau_M^i = \tau_{MW}^i = \tau_{MN}^i = \tau_{MF}^i)$.

H1: Existe diferença entre as acurácias médias individuais. $(\exists \tau_{algx}^i \neq \tau_{algy}^i)$

Hipótese 2 (Para cada algoritmo)

H0: Não há diferença entre a acurácia média geral das técnicas de classificação $(\mu_S = \mu_{SW} = \mu_{SN} = \mu_{SF} = \mu_{MW} = \mu_M = \mu_{MN} = \mu_{MF})$.

H1: Existe diferença entre a acurácia média geral das técnicas de classificação. ($\exists \mu_{algx} \neq \mu_{algy}$).

Hipótese 3 (Comportamento das técnicas sob ruído)

H0: Não há diferença entre a acurácia média geral das técnicas de classificação em diferentes ambientes ($\mu_{algx}^{30dB} = \mu_{algx}^{50dB} = \mu_{algx}^{60dB} = \mu_{algx}^{70dB} = \mu_{algx}^{Ideal}$).

H1: Existe diferença entre a acurácia média geral das técnicas de classificação. ($\exists \mu_{algx}^{ruido_y} \neq \mu_{algx}^{ruido_z}$).

6.2.2 Resultados

Antes de apresentar os resultados relacionados às hipóteses levantadas, as Tabelas 6.6 e 6.7 apresentam as melhores configurações obtidas para cada algoritmo utilizado no presente trabalho (MLP e SVM), a saber, *lbfgs* refere-se ao Método Broyden–Fletcher–Goldfarb–Shanno com memória limitada e *OVO* refere-se ao processo de fazer o treinamento do algoritmo SVM utilizando a técnica um-contra-um.

Tabela 6.6: Hiper-parâmetros Algoritmos MLP

Hiper-parâmetro	MLP	MLP e FFT	MLP e DWT	MLP e Filtro Notch
Função de ativação	<i>tanh</i>	Logística	Logística	Logística
Coefficiente de Regularização (α)	0.1	0.1	0.1	0.1
Número de Neurônios na Camada Escondida	100	100	100	100
Função de Resolução	<i>lbfgs</i>	<i>lbfgs</i>	<i>lbfgs</i>	<i>lbfgs</i>

Como pode-se visualizar, na Tabela 6.6, a função de resolução, o coeficiente de regularização e a quantidade de neurônios na camada escondida são iguais, independente do

Tabela 6.7: Hiper-parâmetros Algoritmos SVM

Hiper-parâmetro	SVM	SVM e FFT	SVM e DWT	SVM e Filtro Notch
Kernel	RBF	RBF	Polinomial	RBF
Técnica de Decisão	OVO	OVO	OVO	OVO
Coeficiente de Regularização (C)	10^5	10^4	10^6	10^5

pré-processamento utilizado, e, em relação a função de ativação, apenas o MLP, com o próprio dado como entrada do algoritmo, possui a função tangente hiperbólica como melhor, os demais utilizam a função logística. Deve-se ressaltar que a técnica de parada utilizada é o número de épocas ou a partir do momento em que não há melhoria na métrica de acurácia no processo de treinamento.

A função de resolução *lbfgs* (Broyden–Fletcher–Goldfarb–Shanno com memória limitada), busca uma otimização numérica para um dado problema, neste caso, a minimização do erro. O coeficiente de regularização (α) faz com que o dado seja regularizado, de forma a evitar problemas na classificação como o superajuste aos dados de entrada (*overfitting*), um valor de α baixo indica uma penalidade maior no conjunto de entrada.

O número de neurônios na camada escondida foi variado entre 2, 5, 10, 50 e 100, e, como pode-se visualizar na Tabela 6.6, todos os algoritmos utilizaram 100 neurônios, talvez um número maior de neurônios melhorasse a acurácia de treinamento, mas, poderia deixar o classificador com superajuste, portanto, as vezes, mesmo que a acurácia seja mais alta, é interessante utilizar um com acurácia mais baixa mas que esteja corretamente ajustado, nem com *overfitting* nem com *underfitting*. Além disso, quanto maior o número de neurônios, maior o número de pesos e, conseqüentemente, maior o tempo de processamento e a memória utilizada pelo algoritmo.

Por fim, em relação à função de ativação no MLP, a diferença entre a tangente hiperbólica e a função logística, para este problema, era baixa em todos os algoritmos, caso fosse visualizar uma aplicação futura em ambiente embarcado, a função logística é preferível, pois é mais fácil de ser implementada em hardware.

Já em relação aos hiper-parâmetros do algoritmo de classificação SVM, percebe-se que houve uma regularidade na técnica de decisão, sendo que a técnica *OVO* (um-contra-um) su-

perou a técnica um-contra-todos (OVR). Deve-se ressaltar que a técnica OVR é mais rápida, então, pode-se verificar em um trabalho futuro uma análise de diferentes configurações de SVM, com análise de tempo, além da acurácia.

Em relação ao Kernel utilizado, o RBF foi melhor em três resultados e o Polinomial apenas no caso em que o DWT foi utilizado. Deve-se lembrar que a técnica de kernelização permite que os dados não linearmente separáveis possam ser classificados utilizando SVM. Ambos os kernels possuem complexidades próximas, dependendo do valor do expoente polinomial, desta forma, a utilização baseada na melhora da acurácia vale a pena.

Por fim, o fator de regularização C variou entre 10.000 e 1.000.0000, sendo que, quanto maior o valor de C , maior a penalização imposta sobre os dados, isso significa que, em determinados algoritmos não foi necessário uma penalização grande sobre os dados, enquanto outros sim, sendo que foi utilizado uma variação de 1.000 a 1.000.000 no *GridSearch*.

É importante acrescentar que o valor dos fatores de regularização para ambos os algoritmos podem ter sido modificados, apesar do resultado do *GridSearch* pois, embora tenha-se obtido uma boa acurácia no conjunto de treino, no de teste a acurácia pode vir a ser baixa, indicando superajuste dos classificadores. Neste momento, faz-se necessário a modificação manual do parâmetro para obter os classificadores com um ajuste mais próximo do ideal.

Sendo assim, de posse dos algoritmos de classificação com seus melhores hiperparâmetros para classificação dos distúrbios elétricos, pode-se fazer a análise dos resultados da acurácia visando verificar as hipóteses levantadas.

O teste estatístico Kolmogorov-Smirnov foi utilizado para verificação da normalidade dos dados, para tanto, um nível de confiança de 95% foi aplicado. A partir do resultados identificou-se que, tanto para as acurácias gerais (de cada algoritmo e em cada ambiente, ruidoso ou não) quanto para as acurácias individuais, a distribuição de probabilidade não é normal, pois os p -values retornados (medida que indica a probabilidade do conjunto avaliado seguir a distribuição normal), foram próximos a zero (menor que 10^{-10}).

Diante desse resultado e com intuito de avaliar as hipóteses apresentadas, um teste não-paramétrico foi aplicado sobre os dados. Escolheu-se o teste de Friedman para avaliação das hipóteses 1 e 2, devido à forma como o problema é apresentado (vários tratamentos para um mesmo conjunto de dados), enquanto que para a hipótese 3 escolheu-se o teste Kruskal Wallis, visto que, diferentemente das demais hipóteses, esta trabalha com diferentes

conjuntos de dados inseridos num mesmo tratamento. Dessa forma, sabendo-se os testes a serem utilizados para avaliação das hipóteses, a próxima subseção apresenta os resultados e suas respectivas análises e interpretações.

Análise e Interpretação

Como mencionado, o teste de Friedman (implementado em Python por [Rodríguez-Fdez et al. 2015]) foi utilizado para análise estatística dos resultados devido à configuração do problema nas hipóteses 1 e 2. Uma vantagem deste teste é que, quando aplicado para mais de duas variáveis, realiza uma avaliação pareada entre todas as variáveis e calcula o valor do ranque, ou seja, a diferença estatística entre elas.

Sendo assim, o teste de Friedman torna-se uma escolha natural para avaliação de duas das três hipóteses apresentadas: hipótese 1 (comparação entre as acurácias das técnicas implementadas para cada distúrbio) e hipótese 2 (comparação entre as acurácias das técnicas implementadas). Já para a análise da terceira hipótese levantada, como mencionado anteriormente, utilizou-se o teste Kruskal Wallis.

De forma a organizar os resultados alcançados, a seguir é apresentado separadamente a análise de cada uma das hipóteses.

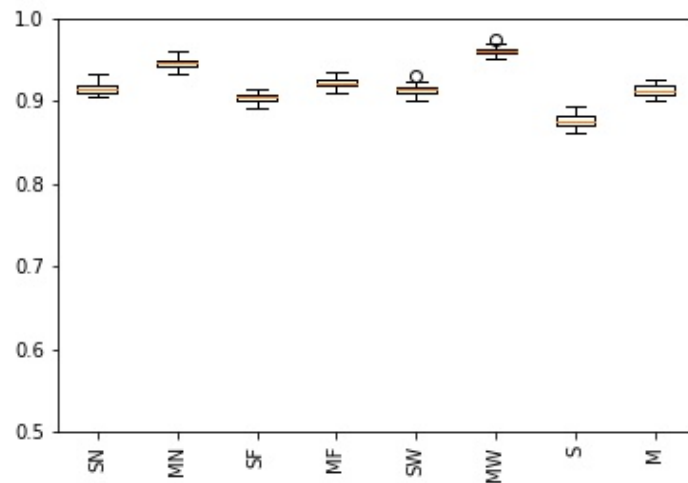
Acurácia Geral O teste de Friedman para as acurácias gerais retornou um *p-value* de zero, isso indica que há diferença estatística nos dados de entrada (neste caso as acurácias). E, pelo próprio resultado do ranqueamento do teste, o algoritmo com maior ranque foi o MLP utilizando FFT como pré-processamento dos dados.

Como forma de verificar graficamente os valores de acurácia, a Figura 6.3 apresenta um gráfico *box-plot*, que representa os resultados da acurácia obtido nas 50 bases de dados utilizadas (com os distúrbios em ambientes ideais, ou seja, sem ruído).

Como pode-se verificar graficamente, o melhor algoritmo de fato é o MLP com DWT, com uma média de acurácia de 96.05%, enquanto que o segundo (MLP com Filtro Notch) possui uma acurácia média de 94.52%.

Esse resultado demonstra que, em ambiente sem ruído, a técnica de DWT é uma boa escolha pois reduz a quantidade de detalhes do sinal de entrada, o que acaba reduzindo também quando aplicada em ambientes ruidosos (que foram utilizados no processo de treinamento

Figura 6.3: Acurácia Geral (Sem Ruído)



do algoritmo).

Apesar do MLP com DWT ser melhor no geral, é interessante verificar se outras técnicas podem ser melhores na classificação para um tipo específico de distúrbio. Tal resultado é analisado a seguir, referente à avaliação da Hipótese 1 deste experimento.

Acurácia para Cada Distúrbio Segundo o resultado do teste de Friedman, o distúrbio do tipo Flutuação de Tensão e a senóide pura possuem p -value igual a 1, indicando que qualquer algoritmo pode ser utilizado para classificação destes sinais sem perda relevante na acurácia. Todos os demais distúrbios possuíram um p -value de zero ou próximo disso, indicando que há diferença estatística entre os algoritmos.

A Tabela 6.8 apresenta os melhores algoritmos para cada sinal de entrada segundo o ranqueamento realizado pelo teste de Friedman.

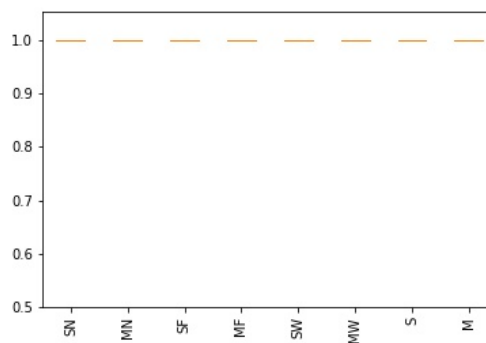
Tabela 6.8: Melhores Algoritmos para Cada Distúrbio

Distúrbio	Melhor(es) Algoritmo(s)
Seno Puro	Qualquer um
Interrupção	MLP com FFT
Afundamento de Tensão	MLP com Filtro Notch
Elevação de Tensão	SVM com DWT
Harmônica	SVM com DWT
Flutuação de Tensão	Qualquer um
Oscilação	MLP com DWT
<i>Notching</i>	SVM, SVM com FFT, MLP, MLP com FFT
Impulsivo	MLP com DWT
Nível DC	SVM com DWT

A seguir são apresentados os gráficos do resultado da acurácia para cada um dos sinais analisados.

- **Seno Puro**

Figura 6.4: Acurácia (Sem Ruído) - Seno Puro



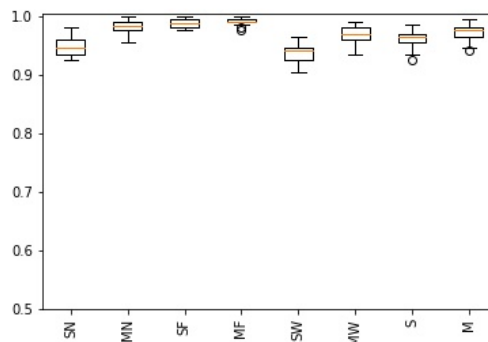
Como pode-se visualizar na Figura 6.4, todos os algoritmos conseguiram classificar o seno puro (sem incidência de ruído) de forma satisfatória, sem variação na acurácia independente do algoritmo. Logo, dentre os algoritmos utilizados, pode-se concluir que o melhor para ser utilizado em um ambiente que necessite de classificação em tempo real (se levado em consideração apenas a senoide pura), os melhores algoritmos

seriam o MLP e o SVM puro, pois evitariam o pré-processamento do sinal, reduzindo assim o tempo total de classificação.

• Interrupção

A Figura 6.5 apresenta os resultados da acurácia dos algoritmos na classificação do distúrbio de interrupção.

Figura 6.5: Acurácia (Sem Ruído) - Interrupção



O algoritmo o MLP com FFT (99.16%) e SVM com FFT (98.74%) foram os que obtiveram melhores acurácias médias, indicando que o algoritmo FFT é um bom algoritmo de pré-processamento para o caso deste distúrbio.

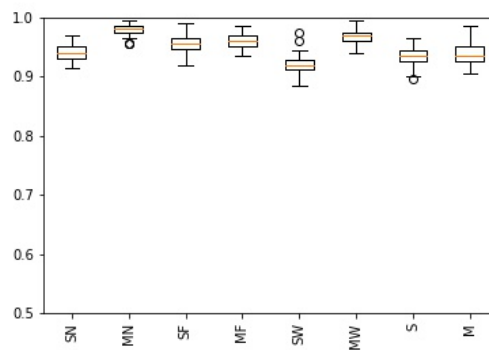
Além disso, pode-se perceber que DWT e o distúrbio puro (sem pré-processamento) foram os que possuíram os piores resultados em média. Isso provavelmente ocorre devido à semelhança (no domínio do tempo) do distúrbio de interrupção com o distúrbio de afundamento de tensão. A diferença fica mais evidenciada, portanto, no domínio da frequência e por isso o filtro notch (que retira a frequência fundamental) e a FFT (que apresenta a amplitude do sinal no domínio da frequência) possuem os melhores resultados.

• Afundamento de Tensão

A Figura 6.6 apresenta os resultados para o distúrbio de Afundamento de Tensão.

O algoritmo MLP com filtro notch obteve a melhor acurácia média (97.94%), seguido pelo MLP com DWT (96.96%). Este resultado é interessante devido ao fato de que o DWT ter conseguido um alto valor de acurácia com o MLP mas o menor valor médio

Figura 6.6: Acurácia (Sem Ruído) - Afundamento de Tensão

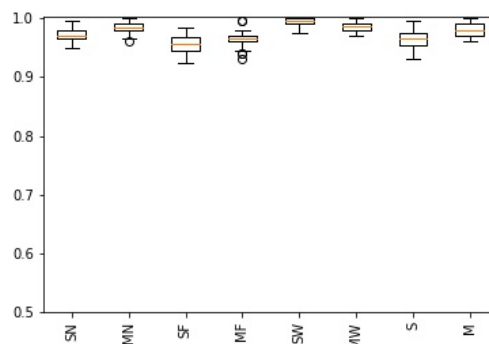


com o SVM, indicando que, neste caso específico, o SVM não consegue construir um vetor de suporte adequado.

• Elevação de Tensão

A Figura 6.7 apresenta os resultados da acurácia para o distúrbio de Elevação de Tensão.

Figura 6.7: Acurácia (Sem Ruído) - Elevação de Tensão

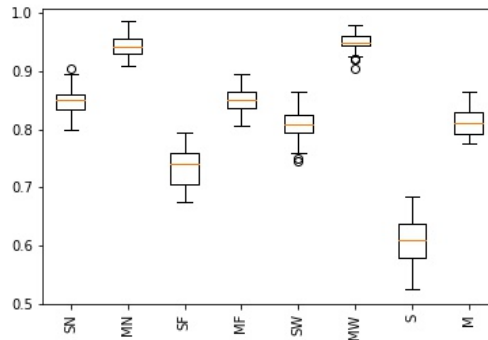


Os resultados para este tipo de distúrbios foram bons para a maioria dos algoritmos, destoando apenas naqueles que utilizam a FFT como pré-processamento que possuíram acurácias abaixo dos 97%). Enquanto isso, o algoritmo SVM com DWT obteve o melhor resultado. Com o algoritmo DWT como pré-processamento, o classificador obteve melhores resultados que o algoritmo puro (sem aplicação de pré-processamento) devido à redução de ruído nos casos em que houve incidência destes, e, sendo estes dados utilizados durante o treinamento do classificador, estes dados influenciam na estrutura final do algoritmo.

- **Transiente Oscilatório (Oscilação)**

A Figura 6.8 apresenta os resultados para o distúrbio de Oscilação.

Figura 6.8: Acurácia (Sem Ruído) - Oscilação



A acurácia dos distúrbios começam a abaixar quando começa a classificação de distúrbios elétricos que podem ser confundidos com o ruído gaussiano, como é o caso do transiente de oscilação. Neste caso, os melhores algoritmos ficam sendo o MLP com filtro notch (94.43%) e o MLP com DWT (95.02%), que são os únicos com média de acurácia acima dos 90%. Os demais possuem valores baixos, inclusive chegando próximo a 50% no caso do SVM puro (60.84%).

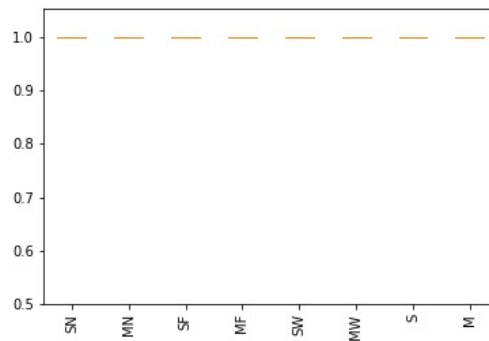
Caso o treinamento dos algoritmos utilizasse somente distúrbios sem ruído, os valores de acurácia seriam melhores, como mostrado no experimento 1, apresentado na seção anterior.

Para obter melhores valores de acurácia em distúrbios neste estilo pode-se acrescentar o uso de algoritmos de redução ou mitigação de ruídos ou ainda utilizar técnicas de extração de características que consigam capturar a essência dos distúrbios mesmo em ambientes ruidosos.

- **Flutuação de Tensão**

A Figura 6.9 apresenta os resultados para o distúrbio de Flutuação de Tensão (*flicker*).

Figura 6.9: Acurácia (Sem Ruído) - Flutuação de Tensão



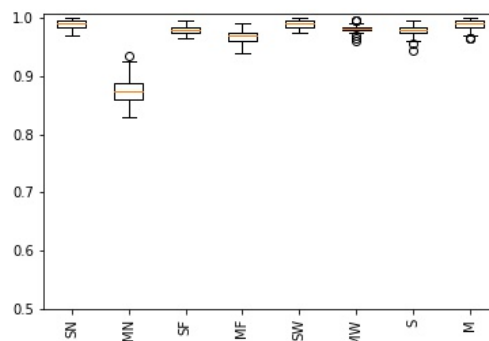
Pode-se visualizar que para esse distúrbio, semelhante ao que ocorrera com a senoide pura, pode ser classificado de maneira satisfatória com todos os algoritmos utilizados no experimento.

Esperava-se um bom resultado principalmente com a utilização da FFT e filtro notch como algoritmo de pré-processamento, devido à variação na frequência do sinal que seriam facilmente identificados com tais algoritmos. Os bons resultados com DWT e sem utilização de pré-processamento deve ter-se dado por alguma característica encontrada através dos algoritmos MLP e SVM em seu processo de treinamento mas que passa invisível a olhos humanos.

• Harmônicas

A Figura 6.10 apresenta os resultados para o distúrbio do tipo Harmônicas.

Figura 6.10: Acurácia (Sem Ruído) - Harmônicas



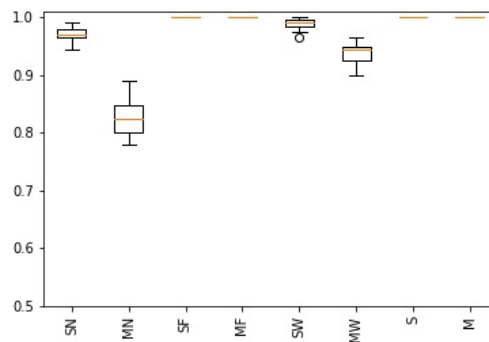
Como pode-se verificar, apenas o algoritmo MLP com filtro Notch está com um valor de acurácia abaixo dos 90%, destoando de todos os demais com bons valores de

acurácia, chegando a 99.03% no caso do SVM com DWT.

- **Notching**

A Figura 6.11 apresenta os resultados de acurácia para o distúrbio Notching, percebe-se que a acurácia média foi de 100% e sem variação no caso de vários algoritmos como SVM com FFT, MLP com FFT, SVM puro e MLP puro. Em contrapartida, houve um baixo valor de acurácia média para o MLP com filtro notch (82.59%) e MLP com DWT (94.00%).

Figura 6.11: Acurácia (Sem Ruído) - *Notching*



Novamente são problemas no processo de escolha das características pelo MLP que não conseguiu capturar a essência do distúrbio utilizando-se tais técnicas de pré-processamento e não uma deficiência da técnica, como mostra os bons resultados obtidos pelo SVM com tais técnicas utilizadas.

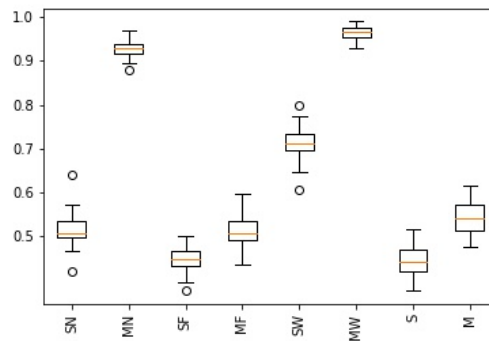
- **Transiente Impulsivo**

A Figura 6.12 apresenta os resultados de acurácia para o distúrbio de transiente impulsivo (*spike*).

Novamente percebe-se uma baixa na acurácia dos algoritmos para este tipo de distúrbio, o que já era esperado devido à forma como o distúrbio é composto (variação de pico na onda senoidal com uma duração de até $1.2\mu s$). Desta forma, este distúrbio é dificilmente detectável quando o sinal está também com ruído (a depender da amplitude e duração do pico).

Apenas os algoritmos MLP com filtro notch e MLP com DWT conseguiram acurácias médias acima dos 90%.

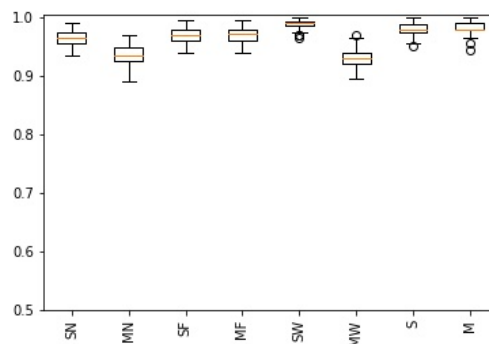
Figura 6.12: Acurácia (Sem Ruído) - Transiente Impulsivo



• Nível DC

Por fim, a Figura 6.13 apresenta os resultados de acurácia para o distúrbio de elevação no nível DC da tensão.

Figura 6.13: Acurácia (Sem Ruído) - Nível DC



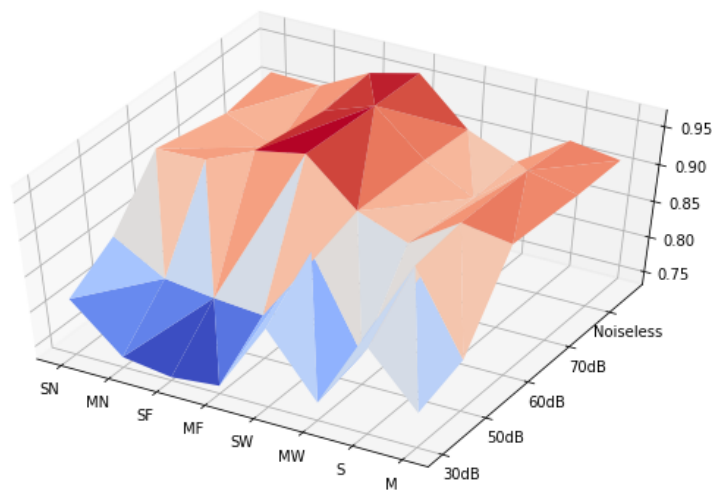
Este foi um distúrbio facilmente classificável pelos algoritmos propostos, todos com valores de acurácia acima dos 90%, sendo o SVM com DWT o que possui a maior média (98.78%).

Novamente não pode-se concluir nada a respeito das técnicas de pré-processamento aplicadas para este tipo de distúrbio específico, pode-se concluir apenas que qualquer uma delas pode ser aplicada com bons resultados.

Desta forma, pode-se notar que alguns distúrbios específicos possuíram um resultado bastante satisfatório com diversas técnicas (100% para todos os 50 conjuntos de dados utilizados), como é o caso do distúrbio de *Notching* e Flutuação de Tensão, além da senoide pura.

Comportamento das Técnicas em Ambientes Ruidosos O teste Kruskal Wallis para a terceira hipótese levantada mostrou que, para todas as técnicas de classificação utilizadas, há diferença estatística quando os dados estão sob efeito de ruído gaussiano (todos os oito p -values, um para cada algoritmo, deram resultados da ordem de 10^{-6} ou menos), sendo o ambiente sem ruído onde eles possuem o melhor comportamento (maior acurácia média). Isso pode ser visto na Figura 6.14, que apresenta a acurácia dos algoritmos propostos com os dados sob efeito de ruído ou não.

Figura 6.14: Comportamento da Acurácia dos Classificadores Sob Efeito de Ruído



Como pode-se verificar, há uma queda na acurácia a medida em que a taxa de sinal-ruído vai diminuindo. Em outras palavras, aumentando-se o ruído, diminui-se a acurácia, mesmo utilizando os algoritmos de FFT, DWT e Filtro Notch.

Assim posto, percebe-se empiricamente que é necessário uma técnica de eliminação ou redução de ruído sobre os sinais de entrada ou ainda um algoritmo capaz de extrair as características dos distúrbios mesmo com os sinais ruidosos, apesar de tais técnicas aumentarem o tempo de processamento.

Esta hipótese foi testada pois, mesmo sob o efeito de ruído, as técnicas poderiam se comportar de forma satisfatória, o que não provou-se verdade, entretanto, uma análise experimental era fundamental, mesmo que tal conclusão pareça, a primeira vista, evidente.

6.2.3 Conclusão do Experimento

Por fim, verificou-se que, sem a presença de ruído, todos os algoritmos conseguem alcançar uma alta taxa de acurácia. Comparando-a aos trabalhos encontrados na literatura, os algoritmos aplicados ficam abaixo apenas do trabalho [Ozgonenel et al. 2013], que alcançou 100% de acurácia utilizando a técnica de SVM juntamente com várias outras técnicas de pré-processamento.

Foi possível, portanto, verificar que o melhor algoritmo para classificação de distúrbios elétricos em ambiente ideal (sem ruído) foi o MLP, utilizando DWT como técnica de pré-processamento dos dados, o que é corroborado pelos resultados individuais de acurácia média, nos quais o DWT aparece acompanhando o melhor classificador para 5 dos 10 distúrbios classificados.

Além disso, alguns distúrbios em ambientes ideais tiveram uma taxa de acerto de 100%, como é o caso do distúrbio de Flutuação de Tensão (para qualquer algoritmo utilizado), o *Notching* (com os algoritmo SVM com FFT, SVM puro, MLP com FFT e MLP puro) e a senoide pura para qualquer algoritmo utilizado).

Desta forma, a depender do distúrbio que se deseja avaliar, ou seja, aquele que tem mais incidência na rede, pode-se escolher determinado algoritmo em detrimento de um outro.

Entretanto, para alcançar um valor mais alto de acurácia (principalmente em ambientes ruidosos) buscou-se novas formas de trabalhar com os tais dados, para isso utilizou-se alternativas providas de diferentes técnicas já bem consolidadas na literatura como os filtros gaussianos e a transformada de Hilbert, de forma a não utilizar o próprio sinal como entrada do algoritmo de classificação, mas um conjunto de características calculadas utilizando estatística descritiva. A próxima Seção apresenta a construção e o resultado do experimento *in vitro*.

6.3 Classificação de Distúrbios Elétricos usando Extração de Características (Experimento 3)

Para alcançar um resultado melhor em ambientes ruidosos (com SNR de 70dB, 60dB, 50dB e 30dB), um experimento *in vitro* foi criado, seguindo a metodologia descrita no Capítulo 5.

O experimento reproduzido é semelhante ao apresentado na seção anterior, modificando em alguns pontos fundamentais, dentre os quais:

1. Utilização do algoritmo Random Forest para classificação (melhor resultado no experimento 1);
2. Utilização de diferentes técnicas de pré-processamento como Transformada de Hilbert, para envelopamento dos dados, Filtro Gaussiano, para retirada do ruído branco, e a FFT para utilização dos dados no domínio da frequência.
3. Utilização de conjunto de características, extraídas a partir de estatística descritivas e análise dos sinais, ao invés dos dados originais como entrada do algoritmo de classificação (experimento 1) ou dos dados somente pré-processados (experimento 2);

A preferência do *Random Forest* em detrimento das técnicas da seção anterior (SVM e MLP) deu-se devido a sua capacidade de apresentar o percentual de importância de determinada característica dentro do processo de classificação, o que auxiliou na busca por características que aumentassem a eficácia do algoritmo. Uma análise realizada com as mesmas características utilizando SVM, MLP e ainda Árvore de Decisão (CART) alcançou resultados inferiores.

A escolha e aplicação dos algoritmos de pré-processamento nos experimentos 3 e 5 deu-se conforme o diagrama apresentado na Figura 6.15.

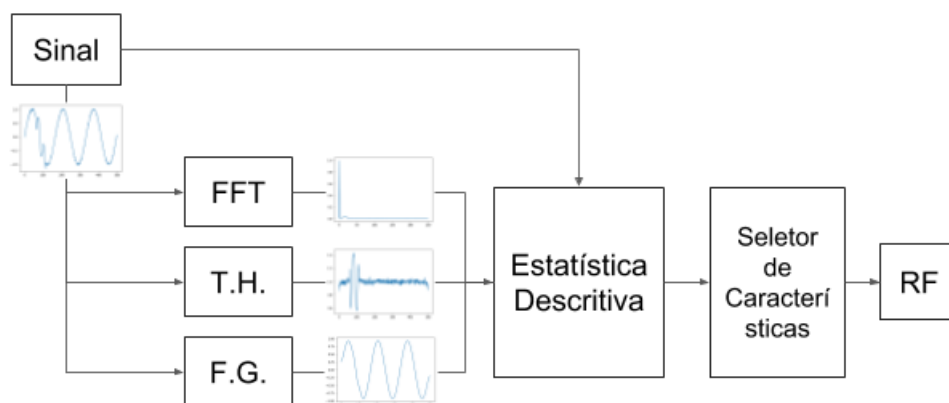


Figura 6.15: Diagrama de Uso de Técnicas de Pré-Processamento (Experimento 3)

Depois de aplicado a FFT, foi utilizado o espectro da magnitude do sinal; em relação à Transformada de Hilbert, utilizou-se também a magnitude do sinal; por fim, em relação ao pré-processamento usando o filtro Gaussiano unidimensional, utilizou-se um valor de $\sigma = 10$, um valor alto para retirar totalmente o ruído do sinal de entrada mas não perder a forma da onda.

Além disso, as características apresentadas na Tabela 6.9 foram utilizadas para a classificação dos sinais de entrada.

Estas características foram escolhidas com base na teoria por trás das métricas utilizadas juntamente com uma base empírica, analisada com base em testes no Python. As métricas energia do sinal, sinais filtrados I, valor máximo sinais filtrados, kurtose sinal filtrado, obliquidade sinal filtrado, máximo/mínimo/média t. Hilbert, soma sinal filtrado e desvio padrão sinal filtrado são utilizados para capturar, principalmente, distúrbios que tenham uma maior influência sobre a amplitude do sinal (afundamento de tensão, elevação de tensão, interrupção, nível-DC e seno puro). Enquanto isso, as demais métricas que utilizam a FFT ou medidas estatísticas descritivas sobre o sinal de entrada puro ou sobre o ruído (sinal de entrada subtraído do sinal filtrado), são, no geral, para capturar os sinais que são melhor caracterizados no domínio da frequência (transiente oscilatório, *notching*, transiente impulsivo, harmônicas e flutuação de tensão).

É importante ainda ressaltar que, em certos momentos, foi utilizado os valores elevado ao quadrado ou à quarta potência, principalmente da FFT ou do sinal ruidoso. Isso foi realizado para obter melhor diferenciação entre o pico (ponto máximo) do sinal e os demais valores.

Por fim, na etapa de seleção das características, o próprio algoritmo *Random Forest* realiza internamente a escolha das melhores características para compor cada uma das árvores presentes no seu sistema de classificação.

6.3.1 Formulação de Hipótese

As questões de pesquisa para este experimento são: o algoritmo *Random Forest* com as técnicas de pré-processamento e extração de características empregadas possui acurácia média geral acima de 90% em ambiente com 30dB de SNR?, a acurácia permanece constante nos demais casos (SNR de 70dB, 60dB, 50dB, 30dB e distúrbios sem ruídos)?

Tabela 6.9: Características

Característica	Equação
Energia dos Sinais	$\sum x^2$
Sinais Filtrados 1	$\sum x_F * (max(x_F) - mean(x_F))$
Valor Máximo Sinais Filtrados	$max(abs(x_F))$
Variância FFT Sinal Filtrado	$var(fft(x_F))$
Máximo - Média FFT Sinal	$max(fft(x)) - mean(fft(x))$
Soma FFT Sinal	$100 * \sum fft(x)^4$
Máximo FFT * Soma FFT Sinal Filtrado	$max(fft(x_F)) * \sum fft(x_F)$
Máximo FFT Sinal	$max(fft(x))$
Máximo FFT Sinal Filtrado	$max(fft(x_F))$
Soma Ruído	$\sum (x - x_f)^2$
Kurtose Sinal Filtrado	$kurt(x_F)$
Obliquidade Sinal Filtrado	$skew(x_F)$
4º Momento Sinal	$momento(x, 4)$
2º Momento Sinal	$momento(x, 2)$
Obliquidade/Desvio Padrão/ Soma FFT Sinal	$skew(fft(x)) + 1000 * std(fft(x)) + 10 * \sum fft(x)$
Máx Sinal Filtrado * Soma Sinal	$max(x_F) * 100 * \sum fft(x)^4$
Primeiro Valor Ruído	$(x - x_F)[0]$
Soma Primeiros Valores FFT Sinal e Sinal Filtrado	$fft(x)[0] + fft(x_F)[0]$
Máximo - Média FFT Ruído	$max(fft(x - x_F)) - mean(fft(x - x_F))$
Kurtose T. Hilbert	$kurt(th(x))$
2º e 4º Momento T. Hilbert	$momento(th(x), 2) * momento(th(x), 4)$
2º Momento FFT Sinal Filtrado	$momento(fft(x_F), 2)$
Máx/Mín/Média T. Hilbert	$(max(th(x)) + min(th(x))) * mean(th(x))$
Soma Sinal Filtrado	$\sum x_F$
Desvio Padrão Sinal Filtrado	$std(x_F)$

Com base nestas perguntas podem-se levantar as seguintes hipóteses:

Hipótese 1

H0:

A acurácia é inferior a 90% em sinais com SNR de 30dB.

H1:

A acurácia é superior a 90% em sinais com SNR de 30dB.

Hipótese 2

H0:

A acurácia do algoritmo para sinais com SNR de 50dB é inferior ou igual a acurácia para sinais com SNR de 30dB.

H1:

A acurácia do algoritmo para sinais com SNR de 50dB é superior à acurácia para sinais com SNR de 30dB.

Hipótese 3

H0:

A acurácia do algoritmo para sinais com SNR de 60dB é inferior ou igual a acurácia para sinais com SNR de 50dB. H1:

A acurácia do algoritmo para sinais com SNR de 60dB é superior à acurácia para sinais com SNR de 50dB.

Hipótese 4

H0:

A acurácia do algoritmo para sinais com SNR de 70dB é inferior ou igual a acurácia para sinais com SNR de 60dB. H1:

A acurácia do algoritmo para sinais com SNR de 70dB é superior à acurácia para sinais com SNR de 60dB.

Hipótese 5

H0:

A acurácia do algoritmo para sinais sem ruído é inferior ou igual à acurácia para sinais com SNR de 70dB.

H1:

A acurácia do algoritmo para sinais sem ruído é superior à acurácia para sinais com SNR de 70dB.

6.3.2 Resultados

Antes de apresentar os resultados relacionados às hipóteses levantadas, a Tabela 6.10 apresenta os melhores hiper-parâmetros obtidos para o algoritmo de classificação utilizado no presente trabalho (*Random Forest*).

Tabela 6.10: Parâmetros *Random Forest*

Parâmetro	Valor
Critério	Entropia
Número de Árvores	50
Máximo de Características	Automático

De posse do *Random Forest* com seus melhores hiper-parâmetros, pode-se fazer a análise dos resultados da acurácia visando verificar as hipóteses levantadas.

A Tabela 6.11 apresenta os resultados do teste realizado com as acurácias médias do algoritmo nos diferentes conjuntos de dados disponíveis no benchmark com distúrbios simples e frequência de 60Hz.

Tabela 6.11: Teste Z para Acurácia Média

Número	Hipótese	<i>p-value</i>	Acurácia Média
1	$\mu_{30dB} > 90\%$	$8.609e^{-253}$	$(92.59 \pm 0.22)\%$
2	$\mu_{30dB} > \mu_{50dB}$	0.0	$(98.82 \pm 0.53)\%$
3	$\mu_{50dB} > \mu_{60dB}$	0.071	$(98.89 \pm 0.25)\%$
4	$\mu_{60dB} > \mu_{70dB}$	0.773	$(98.86 \pm 0.23)\%$
5	$\mu_{70dB} > \mu_{SemRudo}$	0.604	$(98.84 \pm 0.25)\%$

Os valores de *p-value* próximo de zero apresentados na Tabela 6.11, indicam que as hipóteses nulas podem ser rejeitadas, e, como utilizou-se um teste unilateral a direita, pode-se concluir que o algoritmo garante um valor de acurácia superior a 90% para ambientes cujo

valor de SNR chega a 30dB (Hipótese 1) e acurácia maior para sinais com SNR de 50dB que com SNR de 30dB (Hipótese 2). A Hipótese 3, de que os valores de acurácia são iguais para sinais com 50dB e 60dB deu um p -value de 0.07, logo não pode-se rejeitar fortemente a hipótese nula, o mesmo acontece com as hipóteses 4 e 5, que tem p -values mais próximos de 1, logo, não podemos concluir se as acurácias são iguais ou diferentes para estes casos, o que pode-se notar pela média de acurácia próxima nestes casos.

Análise e Interpretação

Iniciando pelo resultado da acurácia do algoritmo para cada distúrbio individualmente em diferentes cenários com e sem ruído, as Figuras 6.16, 6.17, 6.18, 6.19 e 6.20 apresentam gráficos *box-plot* onde pode-se identificar a média de acurácia do algoritmo para cada distúrbio bem como a variação.

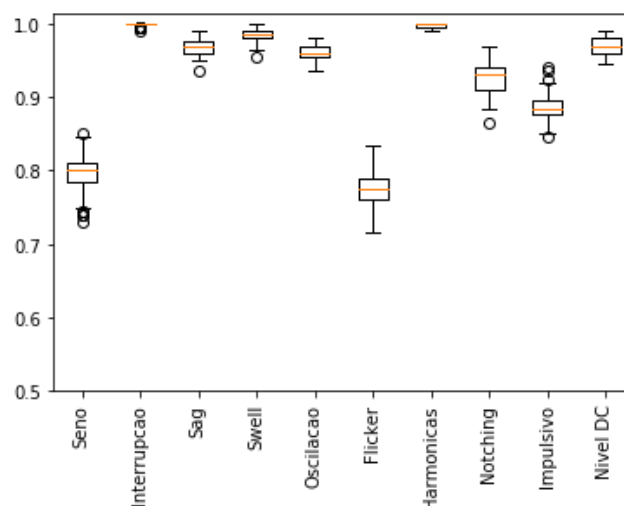


Figura 6.16: *Box-plot* - 30dB

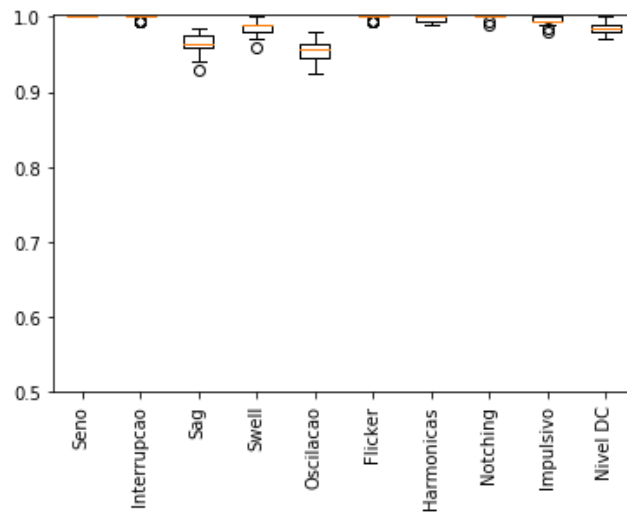


Figura 6.17: *Box-plot* - 50dB

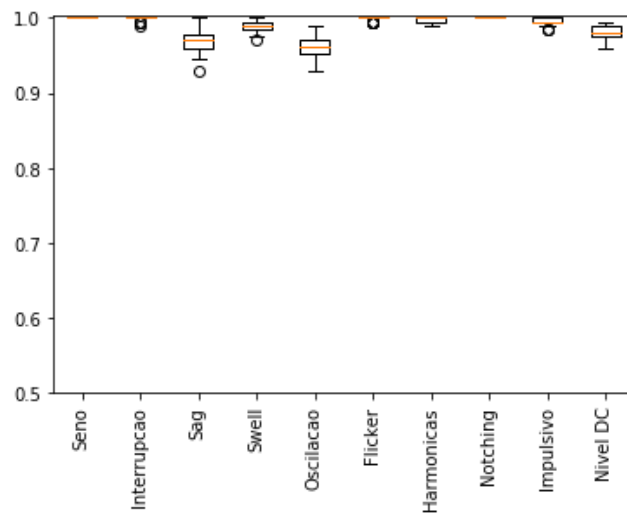


Figura 6.18: *Box-plot* - 60dB

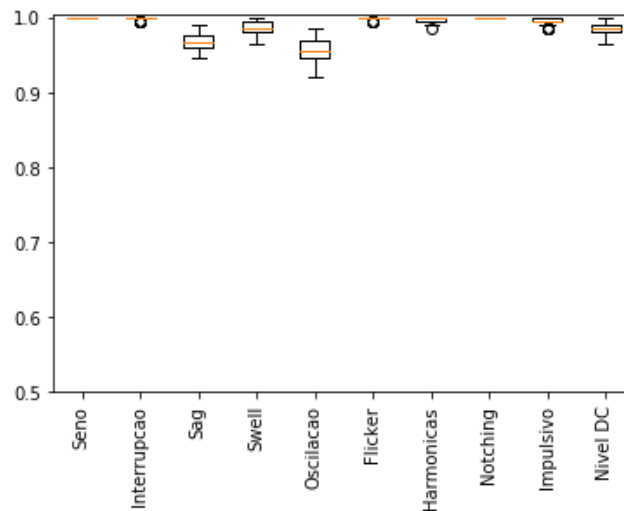


Figura 6.19: *Box-plot* - 70dB

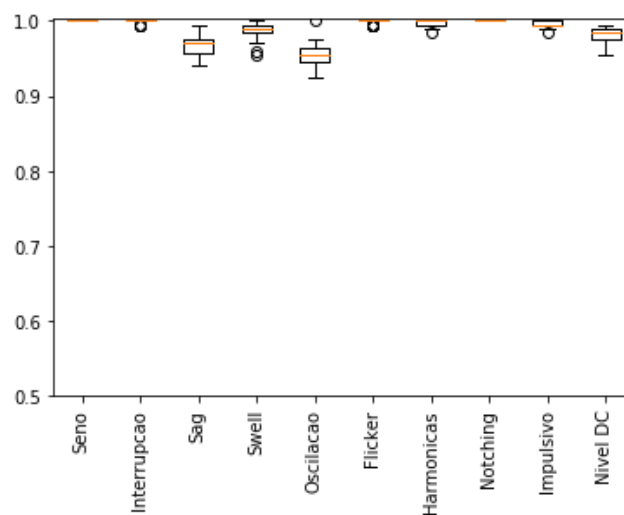


Figura 6.20: *Box-plot* - Sem ruído

Como pode-se verificar, há uma queda na acurácia em determinados tipos de distúrbio com o aumento do valor do ruído (redução da SNR). A Figura 6.21 apresenta melhor esta curva de tendência de cada distúrbio com a variação no valor de SNR.

No gráfico da Figura 6.21 pode-se perceber que há quatro sinais que possuem valores baixo de acurácia média para o ambiente de 30dB (Flutuação de Tensão, Seno Puro, Transiente Impulsivo e *Notching*). A Figura 6.22 apresenta a matriz de confusão para o ambiente de 30dB de SNR.

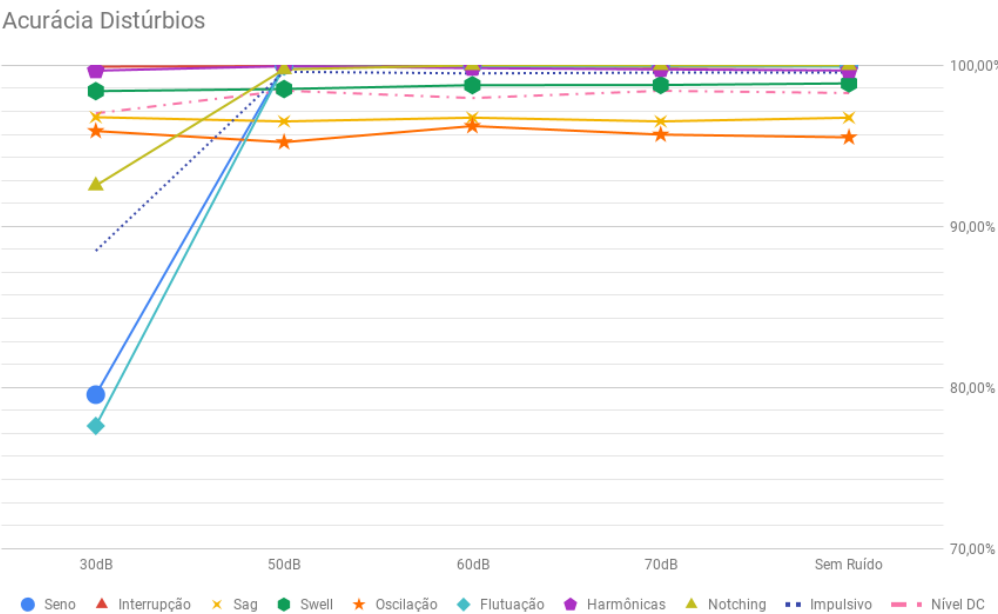


Figura 6.21: Gráfico Acurácia Distúrbios

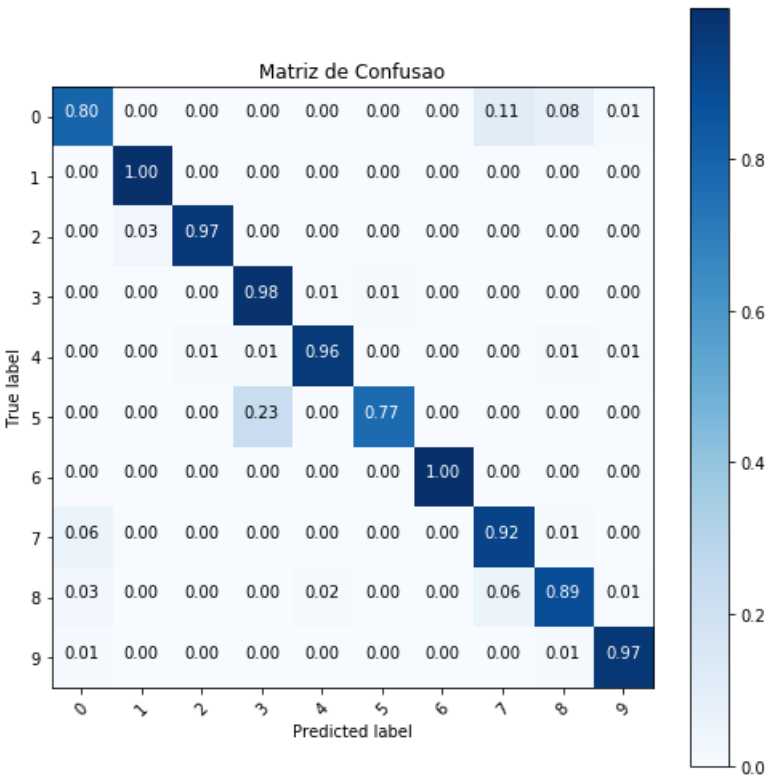


Figura 6.22: Matriz de Confusão 30dB

Na Figura 6.22, os valores numéricos representam: 0 - seno, 1 - interrupção, 2 - *sag*, 3 - *swell*, 4 - oscilação, 5 - *flicker*, 6 - harmônicas, 7 - *notching*, 8 - transiente impulsivo, 9 - elevação no nível DC. A partir desta figura, pode-se notar que os distúrbios de *notching*, transiente impulsivo e o seno estão sendo confundidos durante a classificação, pois, o alto valor do ruído faz com que os distúrbios sejam difíceis de serem identificados, como mostra a Figura 6.23, além disso o distúrbio *flicker* fora classificado como distúrbio de elevação de tensão (*swell*), como pode-se verificar pela Figura 6.24 tais distúrbios podem ser facilmente confundidos.

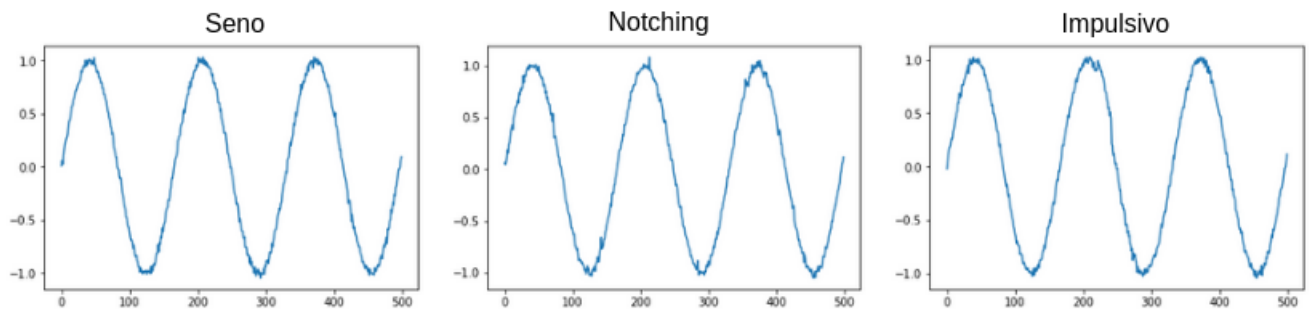


Figura 6.23: Formas de Onda (Seno-*Notching*-Impulso)

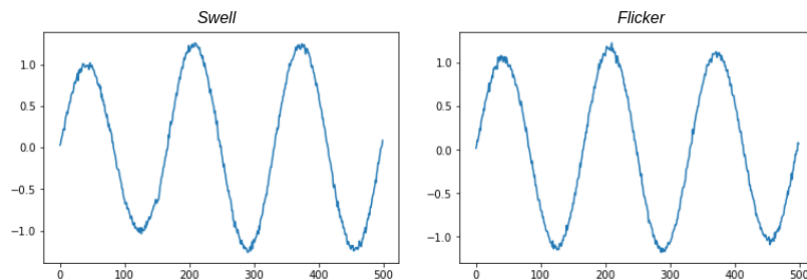


Figura 6.24: Formas de Onda (*Swell*-*Flicker*)

Ameaças à Validade

Os problemas relacionados à validade de conclusão foram mitigados ao utilizar um benchmark disponibilizado online para que outros pesquisadores possam utilizar e comparar com os resultados aqui apresentados, no entanto, há uma limitação quanto a quantidade de distúrbios utilizados, poder-se-ia aplicar o algoritmo em toda a base criada, com os distúrbios

duplos e triplos, no entanto, experimentos com distúrbios múltiplos foram feitos em paralelo e serão apresentados em capítulos seguintes da presente dissertação.

Conclusão do Experimento

O resultado alcançado mostrou que é possível melhorar os resultados da classificação em ambientes ruidosos utilizando técnicas de pré-processamento dos dados (FFT, Filtragem gaussiana e Transformada de Hilbert), extração de características (estatística descritiva e cálculos matemáticos) e seleção de características (feita pelo próprio algoritmo *Random Forest*).

Outras técnicas como o ICA e o cálculo da FFT utilizando janela deslizante foram utilizadas, entretanto, tais técnicas não conseguiram reduzir o ruído gaussiano e melhorar acurácia na classificação. O ICA falhava principalmente em distúrbios do tipo transiente oscilatório, transiente impulsivo e *notching*, não conseguindo separar os sinais do ruído. Já a técnica de janela deslizante não mostrou-se melhor que a aplicação da FFT pura sobre os dados. Desta forma, optou-se por não utilizar tais técnicas que aumentariam o tempo de processamento da classificação, não tendo resultados significativos na acurácia.

6.4 Análise de Tempo de Processamento de Sistema de Monitoramento em NodeMCU e Raspberry Pi (Experimento 4)

Um experimento *in vitro*, seguindo a metodologia apresentada no Capítulo 5, foi realizado para comparação da implementação de um sistema de monitoramento de distúrbios elétricos utilizando o NodeMCU¹, um sistema microcontrolado, e o Raspberry Pi 3, um sistema microprocessado.

O experimento 1, apresentado no presente capítulo, foi voltado especificamente para o Raspberry Pi, em contrapartida, neste buscou-se baratear o custo do sistema final, portanto, focou-se principalmente em sistemas microcontrolados de baixo custo, como o Arduino. No entanto, devido a baixa memória do Arduino UNO, o algoritmo aqui proposto não pode ser implementado nessa plataforma, sendo o NodeMCU portanto uma alternativa viável,

¹http://www.nodemcu.com/index_en.html

visto o seu baixo custo. O Raspberry entra no experimento, portanto, apenas para efeito de comparação, pois, imagina-se que este conseguirá realizar uma classificação muito mais rápida devido ao seu alto poder de processamento, logo, pode ser interessante ao projetista do sistema de monitoramento de distúrbios elétricos ter uma base com a qual comparar os resultados obtidos através do NodeMCU.

Neste experimento utilizou-se apenas a FFT como forma de pré-processamento pois adicionar novas técnicas iriam aumentar o tempo de processamento no ambiente embarcado, e, além disso, observou-se que a FFT, junto com técnicas de extração de características, consegue pegar os detalhes dos ruídos no domínio da frequência. A Figura 6.25 apresenta um diagrama de como deu-se o processo de classificação dos dados.

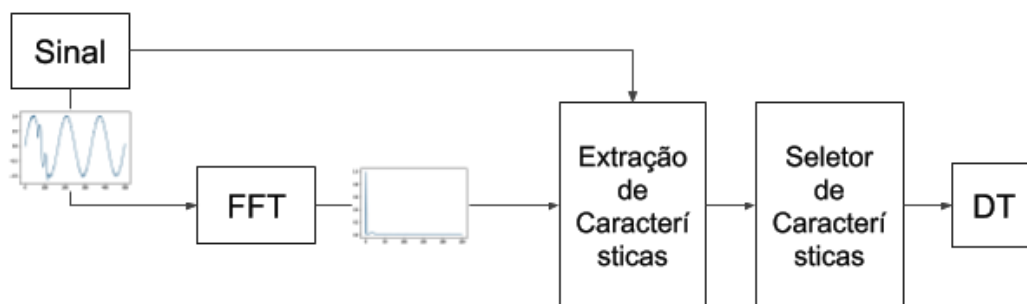


Figura 6.25: Diagrama de Uso de Técnicas de Pré-Processamento (Experimento 4)

Além disso, como forma de verificar como se comporta o sistema proposto em ambientes embarcados, utilizou-se dados ruidosos com valores de SNR a 30dB, 50dB, 60dB e 70dB, como mencionado na metodologia utilizada (Capítulo 5).

6.4.1 Formulação de Hipótese

As questões de pesquisa para este experimento são: é possível implementar um algoritmo de classificação em ambiente embarcado microcontrolado com taxa de acerto superior a 90% para distúrbios sem ruído? Existe alguma diferença na acurácia ao modificar o valor de SNR no sinal de entrada? Existe alguma diferença de tempo de processamento entre as diferentes plataformas de hardware onde o algoritmo é implementado (NodeMCU e Raspberry Pi 3)?

Para a primeira e segunda questão pode-se considerar a acurácia média como métrica para avaliar e comparar os resultados, enquanto que para a terceira questão pode-se considerar o

tempo de processamento em cada uma das plataformas.

Hipótese 1 (Acurácia sem ruído)

H0: A acurácia média do algoritmo de Árvore de Decisão é inferior a 90% para os sinais sem ruído.

H1: A acurácia média do algoritmo de Árvore de Decisão é superior a 90% para os sinais sem ruído.

Hipótese 2 (Comportamento sob efeito de ruído)

H0: Não há diferença entre a acurácia média geral das técnicas de classificação em diferentes ambientes ($\mu_{dt}^{30dB} = \mu_{dt}^{50dB} = \mu_{dt}^{60dB} = \mu_{dt}^{70dB} = \mu_{dt}^{SemRuído}$).

H1: Existe diferença entre a acurácia média geral das técnicas de classificação. ($\exists \mu_{dt}^{ruído_y} \neq \mu_{dt}^{ruído_z}$).

Hipótese 3 (Tempo de processamento)

H0: Não há diferença entre o tempo de processamento no NodeMCU e no Raspberry Pi $\mu_t^{node} = \mu_t^{rasp}$.

H1: Os tempos de processamento são diferentes nos dispositivos ($\mu_t^{node} \neq \mu_t^{rasp}$).

6.4.2 Resultados

Primeiramente, como forma de ter uma visualização geral dos resultados obtidos, as Tabelas 6.12 e 6.13 apresentam, respectivamente, os melhores hiperparâmetros selecionados para a Árvore de Decisão e as métricas obtidas.

Tabela 6.12: Tabela hiperparâmetros da Árvore de Decisão

Hiperparâmetro	Valor
Critério de Seleção	Entropia
Altura Máxima	15
<i>Splitter</i>	<i>best</i>
Nº Máximo de Características	$\sqrt{Características}$

A seguir é apresentado uma análise destes resultados, além das devidas verificações estatísticas tendo em vista as três hipóteses levantadas.

Tabela 6.13: Média e Desvio das Métricas Avaliadas

Métrica	Valor	Ambiente
Acurácia (Sem Ruído)	94.75% \pm 0.04%	desktop
Acurácia (70dB)	94.63% \pm 0.05%	desktop
Acurácia (60dB)	94.65% \pm 0.05%	desktop
Acurácia (50dB)	93.74% \pm 0.06%	desktop
Acurácia (30dB)	78.48% \pm 0.08%	desktop
Tempo de Processamento	(48.77 \pm 0.21)ms	Raspberry Pi 3
Tempo de Processamento	(279.12 \pm 12.42)ms	NodeMCU

Análise e Interpretação

Os resultados encontram-se divididos em duas etapas distintas:

1. Análise da acurácia média do algoritmo com diferentes taxas de sinal ruído (Hipóteses 1 e 2);
2. Comparação dos tempos de processamento nos dois sistemas implementados (Hipótese 3).

Análise de Acurácia Média Como forma de averiguar estatisticamente a hipótese 1, utilizou-se o teste estatístico Z para amostra única visto que possui-se um número de amostras superior a 30 [Manly 2008], apresentado na Tabela 6.14. Foi utilizado o Teste Z neste caso devido a forma como a hipótese fora escrita, de forma a comparar apenas dois valores distintos (a média de acurácia com um valor específico), podendo-se, portanto, utilizar um teste unilateral, neste caso, à direita (verificar se um dado valor, em média, é maior que outro).

Tabela 6.14: Teste Z para Acurácia Média

Hipótese	<i>p-value</i>	Acurácia Média
$\mu_{Ideal} > 90\%$	0.0	94.75%

O valor de *p-value* igual a zero apresentado na Tabela 6.14, indica que a hipótese nula

pode ser fortemente rejeitada, e, como utilizou-se um teste unilateral à direita, pode-se concluir que o algoritmo garante um valor de acurácia superior a 90% para ambientes ideais, ou seja, sem presença de ruído nos sinais de entrada (o que já era esperado pelo alto valor médio de acurácia obtido, 94.75%).

Em relação à hipótese 2, verificou-se, inicialmente, que o resultado do teste KS para os 50 valores de acurácia média retornou um valor de *p-value* igual a zero para todos os cinco conjuntos de dados (com ruído a SNR de 30dB, 50dB, 60dB e 70dB, e sem ruído), o que indica que pode-se rejeitar a hipótese nula (de que os dados seguem uma distribuição normal).

Sabendo-se então que, estatisticamente, os dados não possuem distribuição normal, aplicou-se um teste não-paramétrico para verificação da segunda hipótese levantada, o Teste Kruskal Wallis (disponibilizado em Python na biblioteca *scipy* [Jones et al. 01 b]) devido a forma como esta hipótese foi levantada (vários conjuntos de dados independentes com um mesmo tratamento).

O resultado do teste retornou um *p-value* de $1.00 * 10^{-34}$, o que indica que a hipótese nula pode ser rejeitada, ou seja, o algoritmo não se comporta de maneira igualitária em todos os ambientes, o que já era esperado, visto que o uso simples da técnica de FFT e alguns extratores de características baseados em estatísticas descritivas de baixa ordem (como média, variância e desvio padrão), são simples demais para alcançar um resultado eficaz quando se trata da classificação de dados com altos valores de ruído.

As Seções 6.2 e 6.3 apresentam, de maneira mais incisiva, como se pode fazer melhorias na acurácia do algoritmo quando se trabalha com sinais ruidosos, principalmente a Seção 6.3, a qual uma série de medidas estatísticas de alta ordem, além das técnicas de pré-processamento, são utilizadas para alcançar-se um resultado de acurácia superior a 90% mesmo em ambiente com uma alta taxa de ruído (SNR de 30dB).

Como o objetivo deste experimento foi fazer uma análise mais aprofundada da implementação em ambiente embarcado, a análise da acurácia na presença de ruído serve apenas de apontamento para os pesquisadores analisarem como esta técnica se comportaria em um ambiente ruidoso, caso viesse a escolher o NodeMCU como plataforma para o sistema embarcado de monitoramento.

Caso a escolha do projetista venha a ser a utilização do Raspberry Pi 3, pode-se então

implementar os algoritmos apresentados na Seção 6.3, visto que a presença do ambiente Miniconda permite tal implementação de maneira simples, sendo que, em comparação ao trabalho aqui apresentado, haveria um aumento no tempo de processamento relacionado à escolha do algoritmo de classificação (*Random Forest*) e do aumento de técnicas de pré-processamento e extração de características dos sinais de entrada.

Comparativo de Plataformas Embarcadas Assim como a acurácia média, o tempo de processamento também não possuiu uma distribuição normal segundo o resultado do Teste KS aplicado (*p-value* de 0.0 para a média de tempo no Raspberry e no NodeMCU), desta forma, foi aplicado o teste de Friedman, devido a forma com o problema é composto (um mesmo conjunto de dados aplicado a diferentes tratamentos, representado pelos dispositivos nos quais os algoritmos foram implementados).

O resultado do teste de Friedman retornou um *p-value* de 0.0, indicando assim a existência de diferença estatística entre os conjuntos de dados. A Tabela 6.13, apresentada anteriormente, mostra os valores de tempo de processamento em cada um dos ambientes e, como pode-se ver, o valor de tempo de processamento do Raspberry é muito inferior ao do tempo no NodeMCU.

Estes resultados ilustram o que era inteiramente esperado, devido ao poder de processamento do Raspberry e as otimizações das bibliotecas Python em relação à forma de programação realizada no NodeMCU (utilizando C-Arduino).

Entretanto, o resultado ainda mostra que, apesar de uma superioridade de mais de 5 vezes em relação ao tempo de processamento, o NodeMCU ainda consegue realizar este cálculo de maneira relativamente rápida. No entanto, caso apenas um NodeMCU seja utilizado em um sistema de monitoramento, basicamente 12 ciclos do sinal serão perdidos, ou seja, se o programa utiliza 3 ciclos para classificar, significa que apenas 1/5 do sinal é classificado.

Uma alternativa para sanar este problema é utilizar múltiplos NodeMCUs em processo de classificação paralela, sendo necessário para tanto 5 NodeMCUs para realizar a classificação de maneira adequada.

Enquanto isso, o tempo de processamento no Raspberry passa 2,5 ciclos do sinal para ser processado, nesse caso, 45% do sinal é classificado, enquanto 55% do sinal é perdido enquanto o dispositivo está sendo utilizado para classificação dos dados.

Entretanto, no caso do Raspberry Pi 3, devido a seus múltiplos núcleos, é possível utilizar um dos núcleos para a aquisição do sinal e armazenamento interno para que possa ser analisado todo o sinal da rede. Entretanto, isso pode gerar um atraso e estouro de espaço com o tempo, logo, não é uma alternativa muito viável.

Outra alternativa para o Raspberry Pi seria a utilização de um algoritmo mais ágil de detecção, fazendo com que o sistema só analise o sinal a partir do momento em que alguma anomalia seja verificada, tal algoritmo poderia ser, por exemplo, um cálculo de entropia, um classificador Bayesiano simples ou mesmo um algoritmo de detecção de novidade. Dessa forma, um núcleo ficaria responsável pela detecção e os demais pela classificação (já que exige maior poder de processamento). A mesma solução poderia ser implementada pelo NodeMCU, só que com este dispositivo seriam necessário utilizar dois microcontroladores, um no papel de detector e outro no papel de classificador. Essa solução diminuiria a perda de dados, mas não a eliminaria por completo.

A seguir, apresenta-se o que pode-se concluir dos resultados adquiridos a partir do experimento realizado.

6.4.3 Conclusão do Experimento

É visível, pelos diversos trabalhos na área, que a quantidade de sistemas de monitoramento de distúrbios elétricos tem crescido, principalmente devido a difusão dos equipamentos eletrônicos de carga não-linear. Este experimento surge, portanto, para dar luz às diferentes plataformas com as quais um desenhista de sistemas deste tipo pode se deparar.

Dentre os diversas formas de implementação deste tipo de sistema (microcontroladores, DSPs, sistemas microprocessados, FPGAs, SoCs, entre outros), este experimento resolveu abordar dois deles, um sistema microcontrolado usando a plataforma NodeMCU Amica, e um sistema microprocessado usando o Raspberry Pi 3, que também pode ser visto como um minicomputador.

Devido a alta performance de sistemas microprocessados, já era de se esperar um resultado superior em termos de tempo de processamento do sistema microcontrolado. No entanto, o resultado obtido a partir do experimento foi interessante ao mostrar que a diferença nesse quesito foi de menos de seis vezes em média.

Entretanto, apesar da diferença não ser muito alta, a quantidade média de tempo utili-

zada para se classificar um distúrbio em um sistema microcontrolado (279.12ms) ocasiona perda de dados (aproximadamente 12 ciclos do seno são perdidos) enquanto o sistema está processando 3 ciclos.

Dessa forma, o desenvolvedor de um sistema embarcado para o monitoramento de distúrbios elétricos pode levar em consideração ainda a inserção de sistemas paralelos, ou seja, diversos NodeMCU (ou outro microcontrolador ou DSP) trabalhando em paralelo para que não haja perda do sinal. Outra alternativa seria a inserção de um sistema detector de distúrbio que por sua vez acionaria o classificador, tal sistema mitigaria o problema, apesar de ainda assim poder haver perda de dados quando o detector conseguir detectar alguma anomalia e no mesmo momento o microcontrolador ainda estiver trabalhando na classificação de um distúrbio anterior.

O mesmo problema se dá com o sistema microprocessado, apesar que em menor escala, visto seu menor tempo de processamento (48.77ms).

Desta forma, fica como trabalhos futuros a realização de experimentos que adicionem um sistema detector de distúrbios elétricos e a paralelização dos algoritmos nas diferentes plataformas para obter resultados mais rápidos na classificação dos distúrbios.

6.5 Classificação de Distúrbios Isolados e Múltiplos (Experimento 5)

Para a verificação do comportamento de classificação de distúrbios elétricos múltiplos e isolados realizou-se um experimento *in vitro*, seguindo a metodologia apresentada no Capítulo 5, no qual escolheu-se um algoritmo de aprendizagem de máquina supervisionado (*Random Forest*) com base nos resultados conhecidos dos experimentos anteriores.

Optou-se por, além de verificar a qualidade da classificação dos distúrbios múltiplos em ambientes ideais (simulados sem a presença de ruído), avaliar também o comportamento do classificador proposto na presença de ruído nos sinais de entrada, de forma a poder avaliar qual o impacto do ruído sobre o sistema.

Deve-se lembrar ainda que o próprio ruído nos dados é considerado um distúrbio elétrico pelo padrão da IEEE [Association et al. 2009], entretanto, no presente experimento (assim como nos demais que trabalharam com sinais ruidosos), considerou-se o ruído não como

distúrbio a ser classificado, mas como uma falha presente que deveria ser mitigada pelo algoritmo de classificação. Entretanto, em ambientes sensíveis a ruídos, pode ser importante a presença de um sistema de monitoramento deste tipo específico de distúrbio.

Em relação aos algoritmos de pré-processamento dos dados, neste experimento adicionou-se o algoritmo de morfologia matemática além da FFT, Transformada de Hilbert e filtragem gaussiana. Fora adicionado a dilatação do sinal com uma janela de tamanho 100, de forma a capturar melhor o envelopamento do sinal sem a captura de detalhes como acontece com a transformada de Hilbert. A Figura 6.26 apresenta a diferença entre o envelopamento usando a transformada de Hilbert e a dilatação da morfologia matemática.

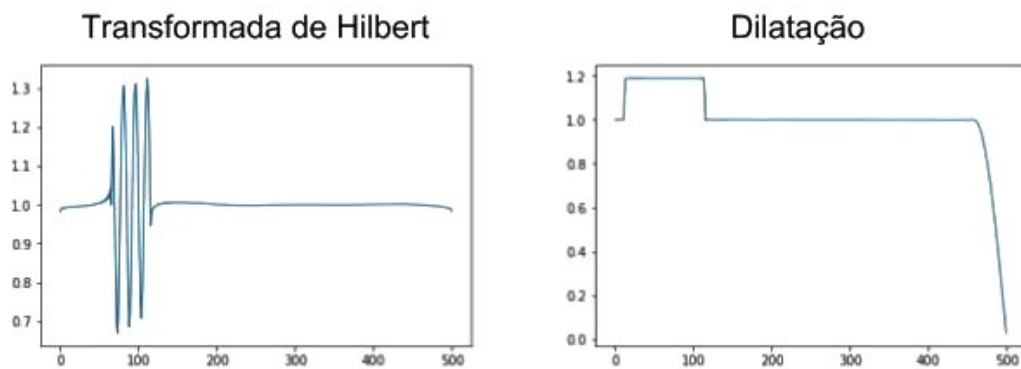


Figura 6.26: Transformada de Hilbert x Dilatação

Desta forma, o classificador final segue os passos mostrados no diagrama da Figura 6.27.

Como técnica de seleção de características utilizou-se o próprio algoritmo *Random Forest* com sua seleção intrínseca para criação das diferentes árvores de decisão que compõem a "floresta". Deve-se ainda ressaltar que utilizou o classificador para realizar a identificação dos diferentes tipos de distúrbios, sejam eles isolados ou múltiplos, ou seja, não separou-se os distúrbios do sinal para classificar um a um, mas supôs o distúrbio múltiplo como uma classe à parte.

Conhecido a metodologia utilizada pelo algoritmo, a subseção a seguir apresenta as hipóteses levantadas para o experimento.

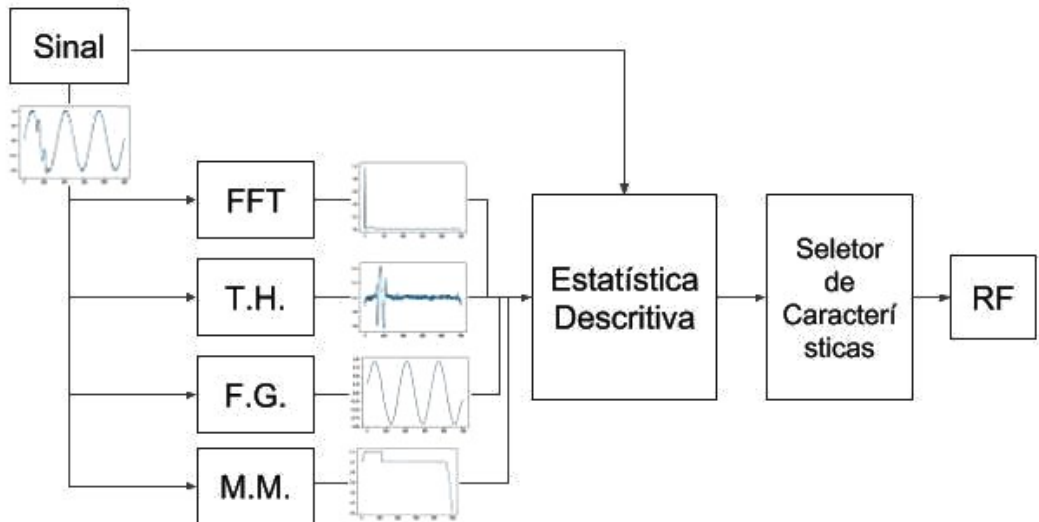


Figura 6.27: Diagrama de Uso de Técnicas de Pré-Processamento (Experimento 5)

6.5.1 Formulação de Hipótese

Com base no objetivo traçado para este experimento e considerando-se os resultados obtidos nos demais experimentos presentes nesta dissertação, pode-se então construir as seguintes questões de pesquisa: é possível alcançar uma acurácia acima dos 90% para a classificação dos 22 distúrbios propostos (isolados, duplos e triplos)?, a acurácia do classificador decai com a diminuição da taxa de sinal ruído (SNR) nos distúrbios?

Desta forma, com base nestas perguntas, pode-se então construir as seguintes hipóteses para serem testadas estatisticamente:

Hipótese 1 (Acurácia sem ruído)

H0: A acurácia média do algoritmo *Random Forest* é inferior a 90% para os sinais sem ruído.

H1: A acurácia média do algoritmo *Random Forest* é superior a 90% para os sinais sem ruído.

Hipótese 2 (Comportamento sob efeito de ruído)

H0: Não há diferença entre a acurácia média geral das técnicas de classificação em diferentes ambientes ($\mu_{rf}^{30dB} = \mu_{rf}^{50dB} = \mu_{rf}^{60dB} = \mu_{rf}^{70dB} = \mu_{rf}^{Ideal}$).

H1: Existe diferença entre a acurácia média geral das técnicas de classificação. ($\exists \mu_{rf}^{ruído_y} \neq \mu_{rf}^{ruído_z}$).

6.5.2 Resultados

Primeiramente, como forma de ter uma visualização geral dos resultados obtidos, as Tabelas 6.15 e 6.16 apresentam, respectivamente, os melhores hiperparâmetros selecionados para o *Random Forest* e as acurácias obtidas para cada um dos ambientes de simulação.

Tabela 6.15: Tabela hiperparâmetros *Random Forest*

Hiperparâmetro	Valor
Critério de Seleção	Entropia
Altura Máxima	15
Número de Árvores	100
Número de Características	Automático
<i>Splitter</i>	<i>best</i>

Tabela 6.16: Acurácia do Classificador Final

Métrica	Valor
Acurácia (Sem Ruído)	93.34% \pm 0.33%
Acurácia (70dB)	93.41% \pm 0.34%
Acurácia (60dB)	93.21% \pm 0.41%
Acurácia (50dB)	91.84% \pm 0.33%
Acurácia (30dB)	83.70% \pm 0.60%

A seguir é apresentado uma análise destes resultados, além das devidas verificações estatísticas tendo em vista as duas hipóteses levantadas.

Análise e Interpretação

Os resultados encontram-se divididos em duas etapas distintas:

1. Análise da acurácia média do classificador com distúrbios simulados em ambiente ideal (sem ruído);
2. Comparativo da acurácia média do classificador com distúrbios simulados em diversos ambientes (com e sem ruído);

Análise de Acurácia Média (Sem Ruído) Como forma de averiguar estatisticamente a hipótese 1, utilizou-se o teste estatístico Z para amostra única visto que possui-se um número de amostras superior a 30 [Manly 2008].

O resultado do teste retornou um valor de *p-value* igual a zero, o que indica que a hipótese nula pode ser rejeitada, e, como utilizou-se um teste unilateral à direita, pode-se concluir que o algoritmo tem uma acurácia superior a 90% para ambientes ideais, ou seja, sem presença de ruído nos sinais de entrada.

Análise de Classificador em Diferentes Ambientes Simulados (Hipótese 2) Em relação à hipótese 2, verificou-se, inicialmente, o resultado do teste KS para as médias de acurácia em cada um dos ambientes, o que retornou um valor de *p-value* igual a zero para todos eles (ambiente simulado sem ruído, com SNR de 70dB, 60dB, 50dB e 30dB). Como já visto em experimentos anteriores, isto indica que pode-se rejeitar a hipótese nula de que os dados seguem uma distribuição normal.

De posse desse resultado, aplicou-se o Teste Kruskal Wallis em Python (disponibilizado na biblioteca *scipy* [Jones et al. 01 b]) devido a forma como a hipótese 2 foi levantada (vários conjuntos de dados independentes com um mesmo tratamento).

O resultado do teste retornou um *p-value* de 1.18×10^{-40} , o que indica que a hipótese nula pode ser rejeitada, ou seja, o algoritmo não se comporta de maneira igualitária em todos os ambientes, o que já era esperado, visto o resultado semelhante retornado pelo experimento apresentado na Seção 6.3.

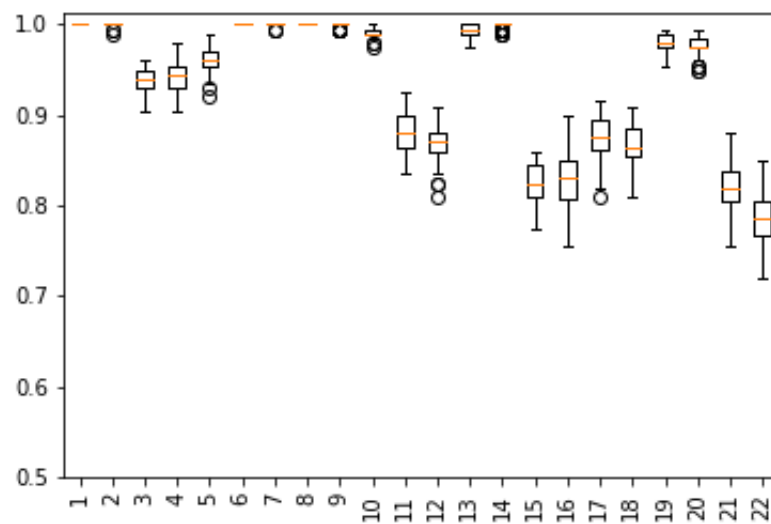
Como forma de visualizar como comportou-se o algoritmo em todos os quatro ambientes simulados para cada distúrbio particularmente, as Figuras 6.28, 6.31, 6.33, 6.34 e 6.35 apresentam gráficos *box-plot* nos quais é possível visualizar a acurácia de cada um dos 22 distúrbios analisados neste experimento. A saber, os números indicados nas imagens referem-se aos distúrbios:

1. Seno
2. Interrupção
3. Afundamento de Tensão (*sag*)
4. Elevação de Tensão (*swell*)

5. Oscilação
6. Flutuação de Tensão (*flicker*)
7. Harmônicas
8. *Notching*
9. Transiente Impulsivo (*spike*)
10. Nível DC
11. *Swell* + Harmônicas
12. *Sag* + Harmônicas
13. *Swell* + *Notching*
14. *Sag* + *Notching*
15. *Swell* + Oscilação
16. *Sag* + Oscilação
17. *Sag* + *Spike*
18. *Swell* + *Spike*
19. *Sag* + *Notching* + Harmônicas
20. *Swell* + *Notching* + Harmônicas
21. *Sag* + Oscilação + Harmônicas
22. *Swell* + Oscilação + Harmônicas

- **Ambiente Sem Ruído**

Como pode-se verificar na Figura 6.28, alguns distúrbios conseguem altos valores de acurácia (100%) sem nenhuma variação, como é o caso dos distúrbios *flicker*, *notching* e a senoide pura (sem distúrbio). Em contrapartida, a maioria dos distúrbios múltiplos

Figura 6.28: *Box-plot* - Ambiente Ideal

possuem seus valores abaixo dos 90% devido à confusão na classificação com os distúrbios isolados (lembrando que todos os distúrbios múltiplos são compostos por um destes dois distúrbios acompanhado de um ou dois outros distúrbios distintos).

Como forma de visualizar melhor para este ambiente como deu-se a confusão entre a classificação dos diferentes distúrbios, a Figura 6.29 apresenta a matriz de confusão.

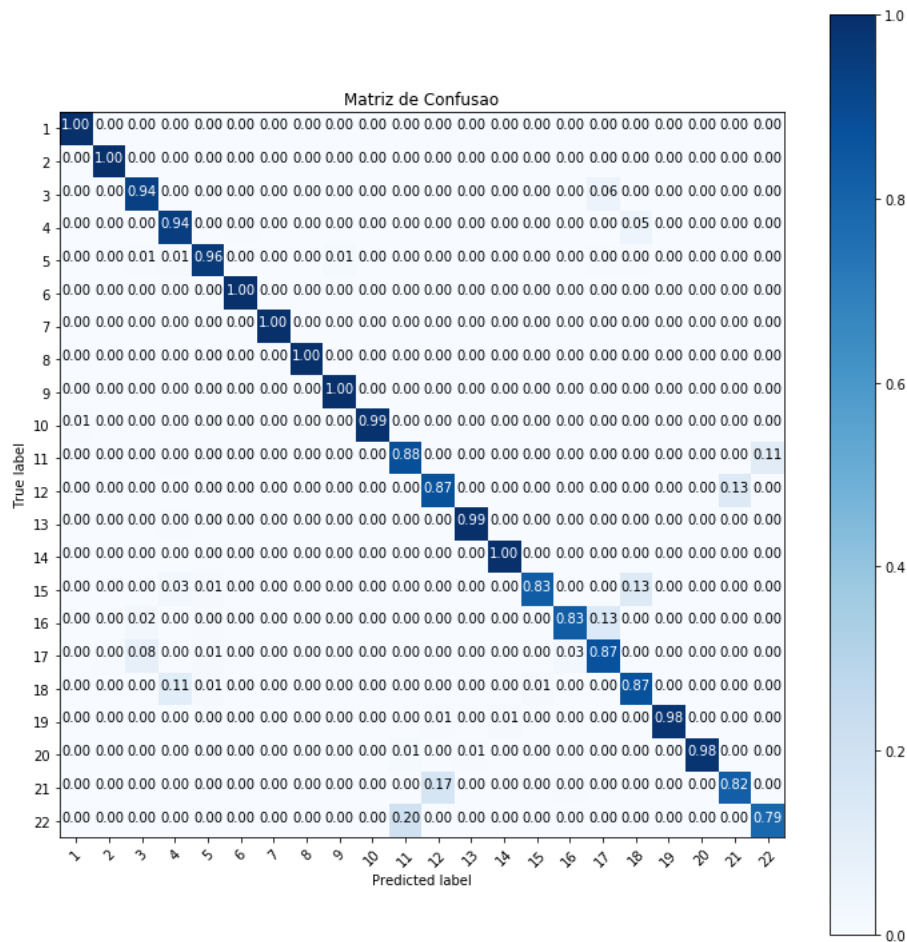


Figura 6.29: Matriz de Confusão Ambiente Ideal (Experimento 5)

Verifica-se, portanto, a partir desta matriz que, por exemplo, o distúrbio que possuiu menor acurácia (*Swell* + Oscilação + Harmônicas) fora confundido em grande parte com o distúrbio 11 (*Swell* + Harmônicas). O mesmo aconteceu com o distúrbio triplo 20 e o distúrbio 12, onde há apenas o acréscimo do distúrbio de oscilação de diferença entre um e outro. Nestes casos, quando o distúrbio de oscilação possuir uma amplitude baixa, fica difícil para o classificador identificar esta diferença, como pode-se visualizar a partir do exemplo apresentado na Figura 6.30 que mostra os distúrbios 11 e 22.

O mesmo problema acontece na classificação contrária, ou seja, o distúrbio 11 as vezes é classificado como o 22 e o distúrbio 12 é classificado como o 21. Para melhorar tal classificação para estes distúrbios múltiplos poder-se-ia tentar a implementação de um

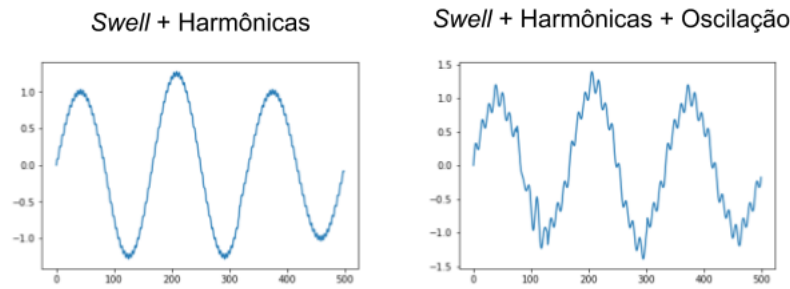


Figura 6.30: *Swell* + Harm. x *Swell* + Harm. + Osc.

algoritmo para separação dos diferentes distúrbios e somente *a posteriori* realizar a classificação, o que poderia ser feito com o algoritmo de ICA por exemplo. A aplicação desta técnica é, certamente, um possível trabalho futuro.

- **Ambiente 70dB**

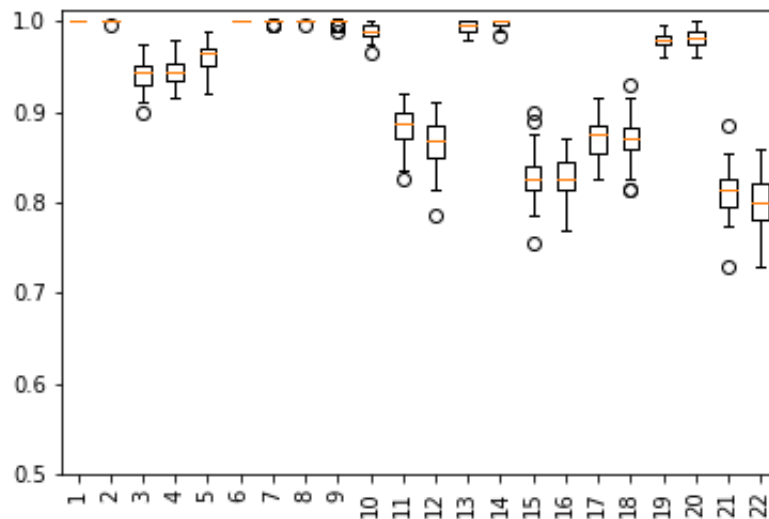


Figura 6.31: *Box-plot* - SNR de 70dB

O gráfico da Figura 6.31 é bastante semelhante ao da Figura 6.28, com pequenas variações em alguns distúrbios, isso porque o SNR de 70dB é relativamente alto, e, portanto, o ruído não influencia fortemente na classificação dos distúrbios. Inclusive,

aplicando-se o teste estatístico de Kruskal Wallis apenas nos distúrbios nos ambientes de 70dB e sem ruído o p -value é de 0.35, ou seja, não pode-se rejeitar a hipótese nula de que a média de acurácia são diferentes. A Figura 6.32 apresenta um exemplo do seno puro para visualizar como o valor do ruído é quase imperceptível no domínio do tempo, apenas no domínio da frequência pode-se verificar a presença do ruído como mostra a FFT.

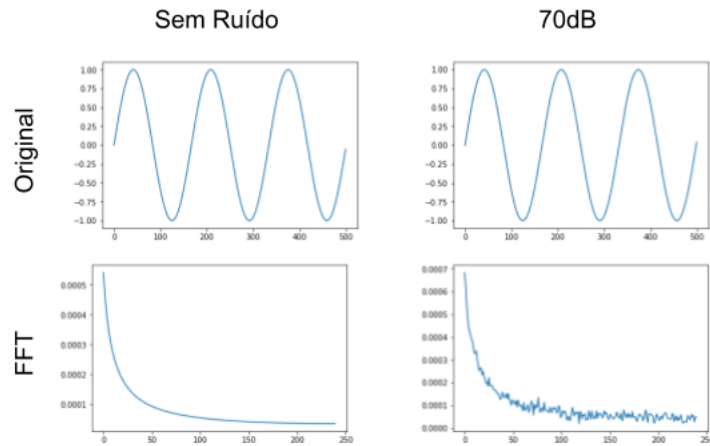
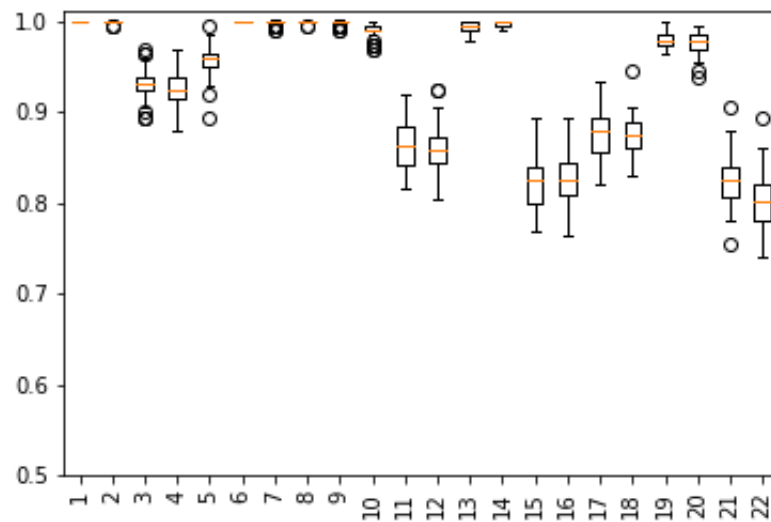


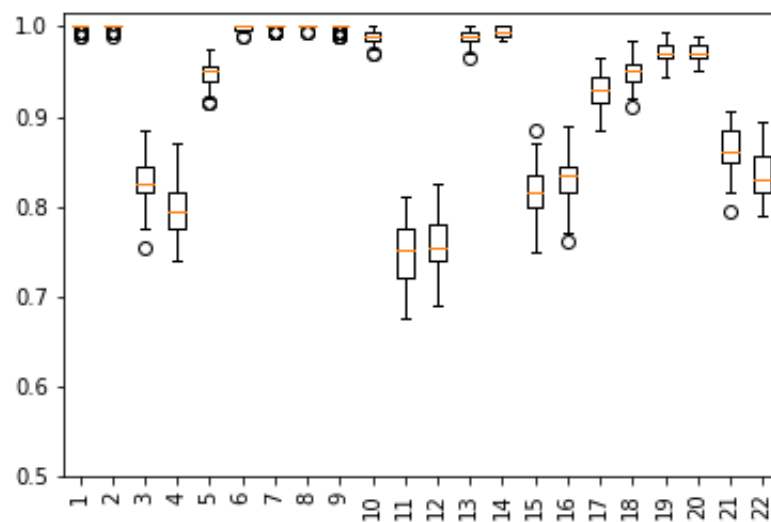
Figura 6.32: Sinal em Ambiente Ideal e a 70dB

● Ambiente 60dB

Novamente a classificação dos distúrbios com SNR a 60dB possui valores semelhantes aos do ambiente a 70dB e sem ruído, inclusive com valores de acurácia média próximos dos demais. Novamente, isso acontece devido à SNR ser alta. Os mesmos problemas ocorridos portanto com os sinais sem ruído aconteceram também neste ambiente. Inclusive, novamente aplicando-se o Kruskal Wallis entre o ambiente de 60dB e o ambiente sem ruído o p -value é de 0.10, ou seja, apesar do seu menor valor, ainda não pode-se rejeitar a hipótese nula de que as acurácias médias são iguais.

Figura 6.33: *Box-plot* - SNR de 60dB

- Ambiente 50dB

Figura 6.34: *Box-plot* - SNR de 50dB

No ambiente de 50dB, como mostra a Figura 6.34, já há uma maior variação na acurácia dos distúrbios, inclusive com valores abaixo de 90% para distúrbios isolados (*sag*

e *swell*). Apesar de possuir uma acurácia média acima dos 90%, ao ter uma acurácia baixa para dois dos distúrbios mais observados na literatura, verifica-se a necessidade de melhorar a escolha das características de forma a melhorar estes distúrbios. Novamente, a aplicação do ICA para separação e classificação separada dos sinais pode ser um bom recurso para realizar tal melhoria.

- **Ambiente 30dB**

Mesmo sabendo que ruídos com taxas de 30dB são incomuns em redes elétricas [Tomim et al. 2007], é interessante observar o comportamento do algoritmo proposto em ambientes extremos.

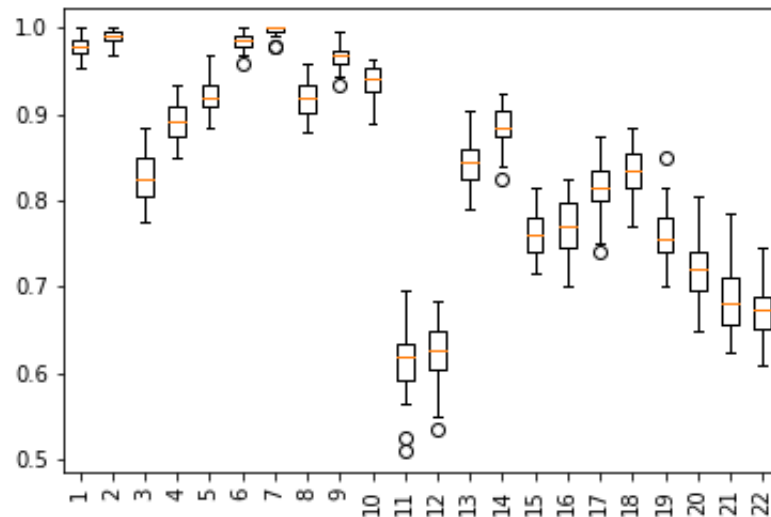


Figura 6.35: *Box-plot* - SNR de 30dB

A classificação dos distúrbios com ruídos mais acentuados, com um valor de SNR mais baixo (30dB) fez com que o valor de acurácia dos distúrbios múltiplos tivesse uma queda significativa, com acurácias abaixo dos 70% como pode ser observado pelo gráfico *box-plot* apresentado na Figura 6.35. Para compreender como deu-se a classificação dos distúrbios, a matriz de confusão é apresentada na Figura 6.36.

Novamente, pode-se verificar que a maior confusão de classificação ocorreu entre os distúrbios 11 e 22, e 12 e 21. Apesar disso, os distúrbios isolados conseguem ser

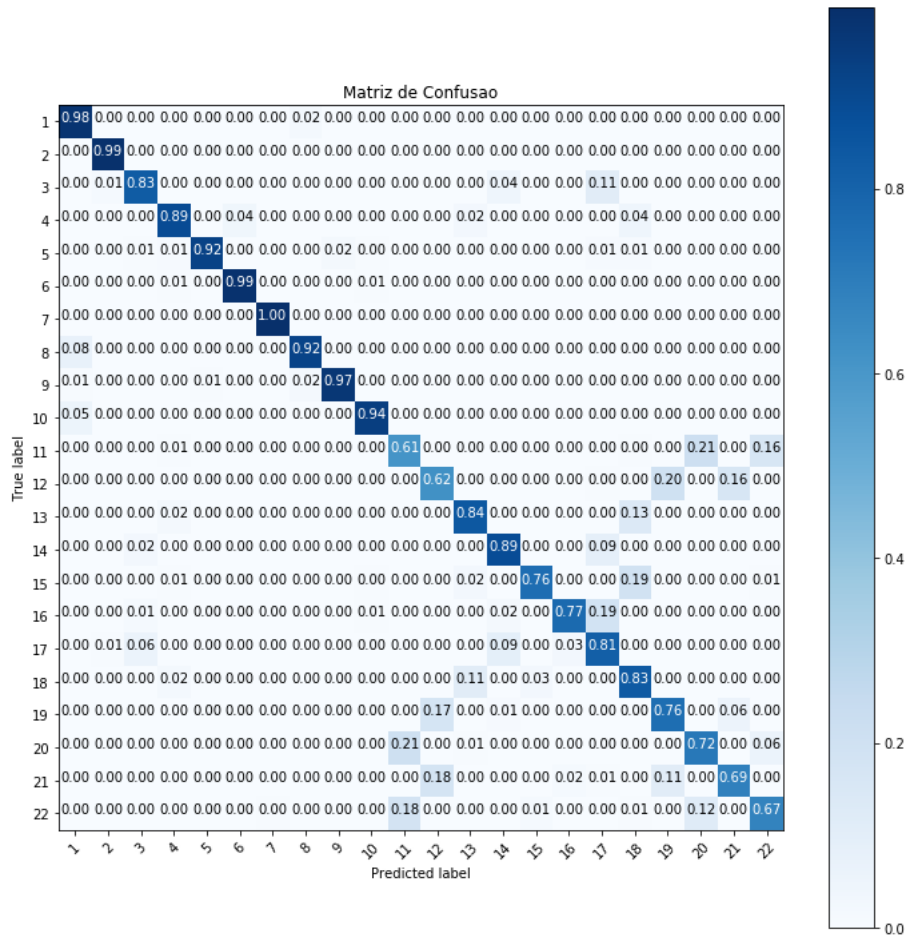


Figura 6.36: Matriz de Confusão - SNR de 30dB (Experimento 5)

classificados com valores de acurácia relativamente altos (acima de 90% para todos), se levado em consideração a baixa SNR aplicada.

• Comparativo

Como forma de comparar os resultados obtidos com um trabalho na literatura que implementou um classificador também para distúrbios múltiplos (duplos e triplos) [Ribeiro et al. 2016], a seguir é apresentada o comparativo na Tabela 6.17. Deve-se ressaltar que tal comparação serve apenas a título de ilustração visto que a implementação dos sinais são realizadas de forma diferentes e, além disso, diferentes distúrbios são analisados pelos autores, além das diferentes taxas de SNR.

Como o trabalho [Ribeiro et al. 2016] não utilizou distúrbios ruidosos, o seu classificador fica em vantagem na classificação de alguns distúrbios. Para compreender

Tabela 6.17: Comparação Trabalhos Relacionados

	Distúrbio	[Ribeiro et al. 2016]	Proposto
1	Sag	99.61±0.58	93.74±1.37
2	Swell	99.66±0.54	94.45±1.79
3	Notch	99.90±0.28	100.00±0.00
4	Transiente Impulsivo	99.90±0.28	99.92±0.18
5	Oscilação	99.83±0.29	96.01±1.43
6	Harmônicas	97.99±1.24	99.97±0.12
7	Oscilação + Sag	99.63±0.42	82.8±2.86
8	Oscilação + Swell	99.50±0.47	82.54±1.98
9	Harmônicas + Oscilação	94.18±1.86	-
10	Harmônicas + Notch	99.60±0.37	-
11	Harmônicas + Sag	99.99±0.09	86.78±1.84
12	Harmônicas + Swell	99.95±0.30	88.18±2.24
13	Notch + Sag	99.96±0.15	99.86±0.28
14	Notch + Swell	99.57±0.39	99.4±0.50
15	Spike + Sag	94.21±2.02	87.47±2.32
16	Spike + Swell	97.51±0.99	86.76±2.34
17	Harmônicas + Oscilação + Sag	96.70±1.31	81.81±2.71
18	Harmônicas + Oscilação + Swell	95.71±1.25	78.81±3.11
19	Harmônicas + Notch + Sag	97.67±1.24	98.16±0.96
20	Harmônicas + Notch + Swell	98.30±1.28	97.71±1.05
21	Seno Puro	-	100.00±0.00
22	Interrupção	-	99.96±0.17
23	Flutuação de Tensão	-	100.00±0.00
24	Nível DC	-	99.16±0.58

com deu-se o desenvolvimento do algoritmo e as técnicas utilizadas em [Ribeiro et al. 2016], um resumo do trabalho é apresentado na Subseção 3.2.2.

Com base nestes resultados e no desenvolvimento de todo o experimento apresentado,

a seguir é apresentado uma conclusão e propõe-se novos trabalhos para prosseguir com o experimento apresentado.

6.5.3 Conclusão do Experimento

O experimento proposto teve em vista a análise de viabilidade da implementação de um classificador de distúrbios elétricos capaz de classificar distúrbios isolados e múltiplos.

Os distúrbios múltiplos, apesar de menos frequentes nas redes elétricas, podem ocasionar perdas maiores, pois, como o próprio nome sugere, são múltiplos distúrbios acontecendo simultaneamente em uma mesma rede elétrica, de forma que os problemas causados por cada um deles são somados.

Logo, pode ser interessante para eventuais consumidores de energia elétrica que precisem de um ambiente mais seguro em relação às possíveis anomalias que possam ocorrer na Qualidade da Energia elétrica, preocuparem-se também com tais distúrbios.

Desta forma, o experimento apresentado obteve um resultado de 93.34% de acurácia na classificação geral dos 22 diferentes tipos de distúrbios, isto no ambiente ideal onde não há presença de ruído. Quando o ruído começa a aparecer, a acurácia do algoritmo permanece constante até decair quando chega-se a 50dB de SNR.

Desta forma, apesar de alcançar um bom resultado para ambientes ideais, é preciso trabalhar melhor a classificação de distúrbios múltiplos. Portanto, em trabalhos futuros pode-se dar continuidade a este trabalho buscando implementar algumas outras técnicas para classificação, como já mencionado, pode-se usar a técnica de ICA para tentar separar os diferentes distúrbios do sinal e conseguir classificá-los de maneira adequada mesmo quando na presença de ruído.

Capítulo 7

Conclusão

A partir do mapeamento do estado da arte, cobrindo desde artigos a patentes acerca dos sistemas de monitoramento de anomalias na Qualidade da Energia elétrica, ou mais precisamente, de distúrbios elétricos, pode-se ter um norte para seguir ao longo de todo o trabalho de pesquisa realizado para a dissertação.

Ao perceber-se, portanto, a falta de um conjunto de dados disponibilizados a todos os autores para que estes pudessem trabalhar de maneira igualitária com as suas propostas, de forma a ter uma base comparativa justa, acreditou-se que a criação de um benchmark¹ poderia ser útil neste quesito, mesmo que este seja de distúrbios elétricos simulados e não ainda de distúrbios reais ou gerados por equipamentos elétricos. O benchmark já serve como norte para futuros pesquisadores e projetistas de sistemas de monitoramento dos distúrbios elétricos, podendo utilizar, como em [Ferreira 2010], os distúrbios reais disponibilizados por [Power and Society 03] para averiguação do sistema proposto (precisando para tanto de um especialista que possa definir a qual classe pertence os distúrbios contidos naquela base).

Por conseguinte, com o objetivo de fazer análises nos setores de classificação de distúrbios em ambientes ruidosos, criação de sistemas de monitoramento de distúrbios embarcados e classificação de distúrbios múltiplos, foram criados cinco experimentos, todos seguindo a metodologia [Wohlin et al. 2012], de forma a ter resultados com maior embasamento estatístico e uma apresentação sucinta de como deu-se o experimento, disponibilizando, quando necessário, os dados e códigos utilizados no experimento.

Dessa forma, com base em todos os experimentos realizados pode-se concluir que a

¹<https://drive.google.com/open?id=1suJiEi9KYjLKCBwZHbJMP44Gl2RhXwuC>

classificação de dados com distúrbios torna-se mais viável à medida em que algoritmos mais sofisticados de pré-processamento dos dados e extração de características são utilizados. Mesmo a aplicação pura de um algoritmo de pré-processamento não garante uma alta taxa de acertos na classificação dos distúrbios, sendo sempre recomendado a utilização de extratores de características.

Além disso, pode-se ainda verificar que um dispositivo microcontrolado, mesmo com pouca memória e baixo poder de processamento, foi capaz de realizar a classificação com uma alta taxa de acurácia. Entretanto, como visto, o tempo de processamento ainda está alta, podendo deixar passar distúrbios enquanto está em processo de classificação. Sendo assim, um dispositivo com maior poder de processamento, sistema microprocessado como o Raspberry, por exemplo, pode ser utilizado com a garantia de redução no tempo de resposta e, portanto, menor perda de dados.

Por fim, verificou-se que quando se trata da classificação de distúrbios que acontecem simultaneamente na rede elétrica, já utilizado em alguns trabalhos [Dalai et al. 2013, Ribeiro et al. 2016, Li et al. 2016, He et al. 2013, Zhang and Li 2010, Salem et al. 2007b], é interessante, assim como quando há a presença de ruído, utilizar algoritmos de pré-processamento e extratores de características, visto que a troca de domínio dos distúrbios (do tempo para a frequência, por exemplo), pode facilitar a identificação dos distúrbios múltiplos.

Portanto, pode-se concluir que o presente trabalho contribuiu para o ambiente acadêmico ao acrescentar a este um benchmark de distúrbios elétricos com diversas variações nas suas variáveis, como visto no capítulo 4; apresentou soluções práticas que melhoraram a classificação de distúrbios sob efeito de ruído, além de mostrar a importância da utilização de técnicas de pré-processamento para estes casos (os distúrbios com SNR de 30dB, por exemplo, ficaram entre 70% e 80% no experimento da Seção 6.2 e ultrapassaram os 90% no experimento da Seção 6.3, ambos utilizando apenas distúrbios isolados); verificou-se ainda a implementação em diferentes plataformas de hardware e comparou-as em termos de tempo de processamento, utilizando uma técnica de classificação simples mas obtendo, ainda assim, bons resultados (acurácia acima de 90%); e por fim, implementou-se uma técnica capaz de identificar distúrbios múltiplos mesmo sob o efeito de ruído gaussiano, alcançando acurácias acima de 90% para todos os ambientes com taxa de sinal-ruído entre 50dB e 70dB (além de sinais sem ruídos), enquanto que no ambiente com 30dB a acurácia ultrapassou os 80%.

Desta forma, pode-se então dar continuidade ao trabalho realizado propondo-se, principalmente, técnicas de extração de características mais robustas, de forma a melhorar ainda mais a classificação dos distúrbios múltiplos e distúrbios em ambientes ruidosos.

Adicionalmente, pode-se propor ainda a paralelização do algoritmo em Raspberry, de forma a utilizar de maneira mais eficaz seus recursos, e, conseqüentemente, reduzir a quantidade de dados perdidos devido ao tempo de processamento. Sugere-se ainda a criação de algoritmos mais simples (menor tempo de processamento) capazes de detectar os distúrbios elétricos em tempo real. Assim, poder-se-ia utilizar um classificador mais robusto apenas quando detectado a presença de algum distúrbio (isso baratearia o projeto, podendo-se utilizar dois NodeMCU para tal processo, por exemplo).

Para a melhora da escolha das características dos distúrbios pode-se utilizar ainda algoritmos de otimização, buscando dentre as características a que melhor caracteriza cada um dos distúrbios. Inclusive, a classificação dos distúrbios múltiplos pode ainda ser feita utilizando algoritmos que sejam capazes de retirar cada distúrbio do sinal e classificá-los separadamente, como, por exemplo, o ICA.

Por fim, mas não menos importante, pretende-se construir um benchmark de distúrbios reais, criado a partir de alguma parceria com empresas distribuidoras de energia e um especialista em anomalias elétricas para a rotulação dos dados, finalizando então com uma implementação em ambiente real (estudo de caso) de um sistema embarcado classificador de distúrbios elétricos.

REFERÊNCIAS

- [Amaris et al. 2008] Amaris, H., Alvarez, C., Alonso, M., Florez, D., Lobos, T., Janik, P., Rezmer, J., and Wacławek, Z. (2008). Application of advanced signal processing methods for accurate detection of voltage dips. In *Harmonics and Quality of Power, 2008. ICHQP 2008. 13th International Conference on*, pages 1–6. IEEE.
- [Anderson et al. 2013] Anderson, J., Boss, G., Jones, A., McConnell, K., and Moore, J. (2013). Protecting electronic devices from electrical supply disturbances. US Patent App. 13/276,393.
- [Angrisani et al. 1999] Angrisani, L., Daponte, P., D’Apuzzo, M., and Pietrosanto, A. (1999). A vxi power quality analyser implementing a wavelet transform-based measurement procedure. *Measurement*, 26(2):91–102.
- [Arrais et al. 2014] Arrais, E., Roda, V. O., Neto, C. M., Ribeiro, R. L., and Costa, F. B. (2014). Fpga versus dsp for wavelet transform based voltage sags detection. In *Instrumentation and Measurement Technology Conference (I2MTC) Proceedings, 2014 IEEE International*, pages 643–647. IEEE.
- [Artioli et al. 2004] Artioli, M., Pasini, G., Peretto, L., Sasdelli, R., and Filippetti, F. (2004). Low-cost dsp-based equipment for the real-time detection of transients in power systems. *IEEE Transactions on Instrumentation and Measurement*, 53(4):933–939.
- [Association et al. 2009] Association, I. S. et al. (2009). 1159-2009. iee recommended practice for monitoring electric power quality industrial and commercial applications.
- [Back 2013] Back, J. A. (2013). Desenvolvimento de um sistema embarcado voltado ao gerenciamento da energia elétrica em processos industriais.

- [Back et al. 2016] Back, J. A., Tedesco, L. P., Molz, R. F., and Nara, E. O. B. (2016). An embedded system approach for energy monitoring and analysis in industrial processes. *Energy*, 115:811–819.
- [Barbosa and Ferreira 2013] Barbosa, B. H. G. and Ferreira, D. D. (2013). Classification of multiple and single power quality disturbances using a decision tree-based approach. *Journal of Control, Automation and Electrical Systems*, 5(24):638–648.
- [Bollen and Gu 2006] Bollen, M. H. and Gu, I. (2006). *Signal processing of power quality disturbances*, volume 30. John Wiley & Sons.
- [Bollen et al. 2010] Bollen, M. H., Ribeiro, P., Gu, I. Y., and Duque, C. A. (2010). Trends, challenges and opportunities in power quality research. *European Transactions on Electrical Power*, 20(1):3–18.
- [Borges et al. 2016] Borges, F., Silva, I., Fernandes, R., and Moraes, L. (2016). *Classification of Power Quality Disturbances Using Forest Algorithm*, pages 247–252. Springer International Publishing, Cham.
- [Borges et al. 2015] Borges, F. A., Fernandes, R. A., Lucas, A., and Silva, I. N. (2015). Comparison between random forest algorithm and j48 decision trees applied to the classification of power quality disturbances. In *Proceedings of the International Conference on Data Mining (DMIN)*, page 146. The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).
- [Breiman 2001] Breiman, L. (2001). Random forests. *Machine learning*, 45(1):5–32.
- [Breiman et al. 1984] Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. A. (1984). *Classification and regression trees*. CRC press.
- [B.V.] B.V., E. Scopus. <https://www.scopus.com/sources>. Accessed: 2017-04-21.
- [Camposano and Wilberg 1996] Camposano, R. and Wilberg, J. (1996). Embedded system design. *Design Automation for Embedded Systems*, 1(1):5–50.

- [Carniato et al. 2009] Carniato, A. A., Godoy, R. B., and Pinto, J. O. P. (2009). Development of hardware and software for three-phase power quality disturbances detection, classification and diagnosis using kalman filter theory. In *Power Electronics Conference, 2009. COBEP'09. Brazilian*, pages 451–457. IEEE.
- [Carro and Wagner 2003] Carro, L. and Wagner, F. R. (2003). Sistemas computacionais embarcados. *Jornadas de atualização em informática. Campinas: UNICAMP*.
- [Cerqueira et al. 2008] Cerqueira, A. S., Ferreira, D. D., Ribeiro, M. V., and Duque, C. A. (2008). Power quality events recognition using a svm-based method. *Electric Power Systems Research*, 78(9):1546–1552.
- [Choong et al. 2005] Choong, F., Reaz, M. B. I., and Mohd-Yasin, F. (2005). Power quality disturbance detection using artificial intelligence: A hardware approach. In *Parallel and Distributed Processing Symposium, 2005. Proceedings. 19th IEEE International*, pages 7–pp. IEEE.
- [Corner et al. 1993] Corner, G., Evans, A., Railton, R., and Tytler, S. (1993). Electrical mains disturbances introduced by emergency generators and their possible effects on medical equipment. In *EMC and Medicine, IEE Colloquium on*, pages 7–1. IET.
- [Cuma et al. 2013] Cuma, M. U., Teke, A., Meral, M. E., Bayindir, K. C., and Tumay, M. (2013). Implementation of a non-linear adaptive filter based sag detection method for dynamic voltage restorers under unbalanced fault conditions. *Journal of Power Electronics*, 13(2):304–312.
- [Dalai et al. 2013] Dalai, S., Chatterjee, B., Dey, D., Chakravorti, S., and Bhattacharya, K. (2013). Rough-set-based feature selection and classification for power quality sensing device employing correlation techniques. *IEEE Sensors Journal*, 13(2):563–573.
- [Daponte et al. 2004] Daponte, P., Di Penta, M., and Mercurio, G. (2004). Transientmeter: A distributed measurement system for power quality monitoring. *IEEE Transactions on Power Delivery*, 19(2):456–463.
- [Davis 2013] Davis, A. (2013). Determination of the location of an electrical disturbance. US Patent App. 13/745,139.

- [Developpers 09] Developpers, J. (2009–). Joblib: running python functions as pipeline jobs. [Online; acesso <14/01/2019>].
- [do Amaral Monteiro et al. 2018] do Amaral Monteiro, D., Zvietcovich, W. G., and Braga, M. F. (2018). Detection and classification of power quality disturbances with wavelet transform, decision tree algorithm and support vector machines. In *2018 Simposio Brasileiro de Sistemas Eletricos (SBSE)*, pages 1–6. IEEE.
- [Duan and Guo 2012] Duan, Y. and Guo, X. (2012). Research and design of power quality monitoring equipment. In *Measurement, Information and Control (MIC), 2012 International Conference on*, volume 2, pages 740–744. IEEE.
- [Dubey 2008] Dubey, R. (2008). *Introduction to embedded system design using field programmable gate arrays*. Springer Science & Business Media.
- [DWPISM and DWPIM 2017] DWPISM and DWPIM (2017 (Acessado em 21 de junho de 2017)). *Derwent World Patents Index® (DWPISM) and Derwent Markush Resource (DWPIM)*.
- [Edomah 2009] Edomah, N. (2009). Effects of voltage sags, swell and other disturbances on electrical equipment and their economic implications. In *IEEE Proc. of 20th International Conference on Electricity Distribution*. IET.
- [Erişti et al. 2013] Erişti, B., Yıldırım, Ö., Erişti, H., and Demir, Y. (2013). An fpga-based system for real-time monitoring of voltage harmonics. IMEKO.
- [Fang et al. 2012] Fang, X., Misra, S., Xue, G., and Yang, D. (2012). Smart grid—the new and improved power grid: A survey. *IEEE communications surveys & tutorials*, 14(4):944–980.
- [Farhangi 2010] Farhangi, H. (2010). The path of the smart grid. *IEEE power and energy magazine*, 8(1):18–28.
- [Feng et al. 2018] Feng, D., Deng, Z., Wang, T., Liu, Y., and Xu, L. (2018). Identification of disturbance sources based on random forest model. In *2018 International Conference on Power System Technology (POWERCON)*, pages 3370–3375. IEEE.

- [Ferreira 2010] Ferreira, D. D. (2010). *Análise de Distúrbios Elétricos em Sistemas de Potência*. PhD thesis, Universidade Federal do Rio de Janeiro.
- [Ferreira et al. 2011] Ferreira, D. D., Duque, C. A., Cerqueira, A. S., Ribeiro, M. V., and Seixas, J. M. d. (2011). Sistema automático de classificação de distúrbios elétricos múltiplos e isolados. *Sba: Controle & Automação Sociedade Brasileira de Automatica*, 22(1):39–48.
- [Garcia et al. 2009] Garcia, V., Kagan, E., Amasifen, J., Suematsu, A., Hayashi, R., and Chaves, L. (2009). Sistema de consultoria de qualidade de energia elétrica. In *VIII Conferência Brasileira sobre Qualidade da Energia Elétrica*.
- [Ge Aviation 2013] Ge Aviation, S. L. (2013). Aparelho, circuito e método para determinar a localização de um distúrbio elétrico em um circuito.
- [Gherasim et al. 2003] Gherasim, C., Van den Keybus, J., Driesen, J., and Belmans, R. (2003). Power measurements based on the iec 61000-4 standard 1459. In *Instrumentation and Measurement Technology Conference, 2003. IMTC'03. Proceedings of the 20th IEEE*, volume 2, pages 1042–1047. IEEE.
- [Gonzalez and Woods 2010] Gonzalez, R. C. and Woods, R. (2010). *Processamento Digital de Imagens. Tradução de Cristina Yamagami e Leonardo Piamonte*. São Paulo: Pearson Prentice Hall.
- [Gonzalez-Cordoba et al. 2015] Gonzalez-Cordoba, J. L., Mejia-Barron, A., and Valtierra-Rodriguez, M. (2015). Automatic detection and classification of electrical disturbances by means of empirical mode decomposition method. In *Power, Electronics and Computing (ROPEC), 2015 IEEE International Autumn Meeting on*, pages 1–6. IEEE.
- [Gorecki 2014] Gorecki, K. (2014). Wpt in power quality measurement system based on digital signal processor. In *Harmonics and Quality of Power (ICHQP), 2014 IEEE 16th International Conference on*, pages 783–787. IEEE.
- [Granados-Lieberman et al. 2013] Granados-Lieberman, D., Valtierra-Rodriguez, M., Morales-Hernandez, L. A., Romero-Troncoso, R. J., and Osornio-Rios, R. A. (2013).

- A hilbert transform-based smart sensor for detection, classification, and quantification of power quality disturbances. *Sensors*, 13(5):5507–5527.
- [Guerrero et al. 2010] Guerrero, J. M., Blaabjerg, F., Zhelev, T., Hemmes, K., Monmasson, E., Jemei, S., Comech, M. P., Granadino, R., and Frau, J. I. (2010). Distributed generation: Toward a new energy paradigm. *IEEE Industrial Electronics Magazine*, 4(1):52–64.
- [Gupta and Seethalekshmi 2019] Gupta, N. and Seethalekshmi, K. (2019). Review of analyzing techniques in technical challenges related to distributed generation. *International Journal of Applied Engineering*, 14(2).
- [Hahn 1996] Hahn, S. L. (1996). *Hilbert transforms in signal processing*. Artech House.
- [Haykin 2001] Haykin, S. (2001). *Redes neurais: princípios e prática*. Bookman Editora.
- [Haykin et al. 2009] Haykin, S. S., Haykin, S. S., Haykin, S. S., and Haykin, S. S. (2009). *Neural networks and learning machines*, volume 3. Pearson Upper Saddle River, NJ, USA:.
- [He et al. 2013] He, S., Li, K., and Zhang, M. (2013). A real-time power quality disturbances classification using hybrid method based on s-transform and dynamics. *IEEE Transactions on Instrumentation and Measurement*, 62(9):2465–2475.
- [Hsieh et al. 2010] Hsieh, C.-T., Lin, J.-M., and Huang, S.-J. (2010). Slant transform applied to electric power quality detection with field programmable gate array design enhanced. *International Journal of Electrical Power & Energy Systems*, 32(5):428–432.
- [Huang et al. 2002] Huang, S.-J., Yang, T.-M., and Huang, J.-T. (2002). Fpga realization of wavelet transform for detection of electric power system disturbances. *IEEE Transactions on Power Delivery*, 17(2):388–394.
- [INPI 2017] INPI (2017 (Acessado em 27 de junho de 2017)). *IPCPUB V5.3*.
- [Jiasheng and Shijie 2010] Jiasheng, L. and Shijie, C. (2010). Design of power quality disturbances monitoring device based on linear prediction mechanism. In *Challenges in Environmental Science and Computer Engineering (CESCE), 2010 International Conference on*, volume 2, pages 358–360. IEEE.

- [Jones et al. 01 a] Jones, E., Oliphant, T., Peterson, P., et al. (2001–a). SciPy: Open source scientific tools for Python. [Online; acesso <05/12/2017>].
- [Jones et al. 01 b] Jones, E., Oliphant, T., Peterson, P., et al. (2001–b). SciPy: Open source scientific tools for Python. [Online; accessed <today>].
- [Juiz de Fora and Ltda 2014] Juiz de Fora, U. F. d. and Ltda, K. I. E. (2014). Registrador eficiente de distúrbios elétricos (rede).
- [Kachhepati 2016] Kachhepati, B. (2016). *Application of Short Time Fourier Transform (STFT) in Power Quality Monitoring and Event Classification*. PhD thesis, New Mexico State University.
- [Kapisch et al. 2016] Kapisch, E., Silva, L., Martins, C., Barbosa, A., Duque, C., Tavi, A., de Souza, L., et al. (2016). An implementation of a power system smart waveform recorder using fpga and arm cores. *Measurement*, 90:372–381.
- [Kapisch et al. 2014] Kapisch, E. B., Silva, L. R. M., Martins, C. H. N., Barbosa, A. S., Duque, C. A., de Andrade Filho, L. M., and Cerqueira, A. S. (2014). An electrical signal disturbance detector and compressor based on fpga platform. In *Harmonics and Quality of Power (ICHQP), 2014 IEEE 16th International Conference on*, pages 278–282. IEEE.
- [Kennedy 2000] Kennedy, B. (2000). *Power quality primer*. McGraw Hill Professional.
- [Kim 2016] Kim, J. (2016). Digital protective relay. US Patent App. 15/003,728.
- [Landi and Luiso 2008] Landi, C. and Luiso, M. (2008). Performances assessment of electrical motors in presence of disturbances on power supply. In *Power Electronics, Electrical Drives, Automation and Motion, 2008. SPEEDAM 2008. International Symposium on*, pages 167–172. IEEE.
- [Lassila et al. 2005] Lassila, J., Honkapuro, S., Viljainen, S., Tahvanainen, K., Partanen, J., Kivikko, K., Antila, S., Mäkinen, A., and Järventausta, P. (2005). Power quality factors in efficiency benchmarking. In *Electricity Distribution, 2005. CIRED 2005. 18th International Conference and Exhibition on*, pages 1–5. IET.

- [Lei et al. 2018] Lei, Z., Ji, T., Xie, C., Li, M., and Wu, Q. (2018). Power quality disturbance identification using improved particle swarm optimizer and support vector machine. In *2018 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia)*, pages 681–686. IEEE.
- [Li et al. 2016] Li, J., Teng, Z., Tang, Q., and Song, J. (2016). Detection and classification of power quality disturbances using double resolution s-transform and dag-svms. *IEEE Transactions on Instrumentation and Measurement*, 65(10):2302–2312.
- [Liang and Da 2011] Liang, W. and Da, L. (2011). Research on key technologies of fpga based power quality monitoring system. In *Computer Science and Automation Engineering (CSAE), 2011 IEEE International Conference on*, volume 2, pages 623–626. IEEE.
- [Liu et al. 2018] Liu, Q., Zhang, J., Lin, F., Xu, F., and Zhang, W. (2018). A method for supraharmonic source determination based on complex ica. In *2018 China International Conference on Electricity Distribution (CICED)*, pages 711–715. IEEE.
- [Lu et al. 2008] Lu, P., Su, S., Liu, G., Rong, H., and Zheng, L. (2008). A new power quality detection device based on embedded technique. In *Electric Utility Deregulation and Restructuring and Power Technologies, 2008. DRPT 2008. Third International Conference on*, pages 1635–1640. IEEE.
- [Malaysia 2011] Malaysia, L. (2011). New detection of voltage sag based on phase angle analysis. *Australian Journal of Basic and Applied Sciences*, 5(7):405–419.
- [Malik et al. 2019] Malik, H., Kaushal, P., and Srivastava, S. (2019). A hybrid intelligent model for power quality disturbance classification. In *Applications of Artificial Intelligence Techniques in Engineering*, pages 55–63. Springer.
- [Manly 2008] Manly, B. J. F. (2008). *Métodos Estatísticos Multivariados: Uma Introdução*. Bookman.
- [Markovska and Taskovski 2017] Markovska, M. and Taskovski, D. (2017). Optimal wavelet based feature extraction and classification of power quality disturbances using random forest. In *Smart Technologies, IEEE EUROCON 2017-17th International Conference on*, pages 855–859. IEEE.

- [Marwedel 2006] Marwedel, P. (2006). Embedded system design.
- [Marwedel 2011] Marwedel, P. (2011). Embedded system design : Embedded systems foundations of cyber-physical systems.
- [McClellan et al. 2003] McClellan, J., Schafer, R., and Yoder, M. (2003). *Signal Processing First*. Number v. 1 in Signal Processing First. Pearson/Prentice Hall.
- [McGranaghan 2005] McGranaghan, M. (2005). Overview of international voltage quality benchmarking efforts. In *Transmission and Distribution Conf. and Exhibition*, volume 2006, pages 100–102.
- [Meena et al. 2018] Meena, M., Mahela, O. P., Kumar, M., and Kumar, N. (2018). Detection and classification of complex power quality disturbances using stockwell transform and rule based decision tree. In *2018 International Conference on Smart Electric Drives and Power System (ICSEDPS)*, pages 220–226. IEEE.
- [Meena et al. 2009] Meena, P., Rao, K. U., and Ravishankar, D. (2009). Real-time detection and analysis of pq disturbances with dsp using matlab embedded link to code composer studio. In *Power Systems, 2009. ICPS'09. International Conference on*, pages 1–5. IEEE.
- [Meena et al. 2011] Meena, P., Uma Rao, K., and Deekshit, R. (2011). A simple method for real-time detection of voltage sags and swells in practical loads. *EPE Journal*, 21(3):33–40.
- [Mendes et al. 2019] Mendes, T. M., Duque, C. A., Silva, L. R., Ferreira, D. D., and Meyer, J. (2019). Supraharmonic analysis by filter bank and compressive sensing. *Electric Power Systems Research*, 169:105–114.
- [Mingoti 2005] Mingoti, S. A. (2005). *Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada*. Editora UFMG.
- [Monard and Baranauskas 2003] Monard, M. C. and Baranauskas, J. A. (2003). Conceitos sobre aprendizado de máquina. *Sistemas Inteligentes-Fundamentos e Aplicações*, 1(1).
- [Montero-Hernandez and Enjeti 2001] Montero-Hernandez, O. C. and Enjeti, P. N. (2001). A fast detection algorithm suitable for mitigation of numerous power quality disturbances.

- In Industry Applications Conference, 2001. Thirty-Sixth IAS Annual Meeting. Conference Record of the 2001 IEEE*, volume 4, pages 2661–2666. IEEE.
- [Montreuil 2009] Montreuil, M. (2009). Current sensor for measuring an electrical current in a circuit having a plurality of electrical conductors. US Patent 7,477,058.
- [Morales-Velazquez et al. 2017] Morales-Velazquez, L., de Jesus Romero-Troncoso, R., Herrera-Ruiz, G., Morinigo-Sotelo, D., and Osornio-Rios, R. A. (2017). Smart sensor network for power quality monitoring in electrical installations. *Measurement*, 103:133–142.
- [Moreno-Garcia et al. 2013] Moreno-Garcia, I., Moreno-Munoz, A., Domingo-Perez, F., Pallares-Lopez, V., Real-Calvo, R., and Santiago-Chiquero, I. (2013). Implementation of a smart grid inverter through embedded systems. *Elektronika ir Elektrotechnika*, 19(3):3–6.
- [Moussa et al. 2004] Moussa, A.-E., EI-Gammal, M., Abdallah, E., and EI-Seoud, A. (2004). Hardware-software structure for on-line power quality assessment: part i. In *Rail Conference, 2004. Proceedings of the 2004 ASME/IEEE Joint*, pages 147–152. IEEE.
- [Oppenheim 2010] Oppenheim, A. V. (2010). *Sinais e sistemas*. Prentice-Hall.
- [Ozgonenel et al. 2013] Ozgonenel, O., Yalcin, T., Guney, I., and Kurt, U. (2013). A new classification for power quality events in distribution systems. *Electric Power Systems Research*, 95:192–199.
- [Pedregosa et al. 2011] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.
- [Petersen et al. 2008] Petersen, K., Feldt, R., Mujtaba, S., and Mattsson, M. (2008). Systematic mapping studies in software engineering. In *EASE*, volume 8, pages 68–77.
- [Pinzón and Colomé 2019] Pinzón, J. D. and Colomé, D. G. (2019). Real-time multi-state classification of short-term voltage stability based on multivariate time series machine learning. *International Journal of Electrical Power & Energy Systems*, 108:402–414.

- [Power and Society 03] Power, I. and Society, E. (2003–). IEEE P1159.3: On-line documents. [Online; acesso <25/10/2018>].
- [Radil et al. 2007] Radil, T., Matz, V., Janeiro, F. M., Ramos, P., and Serra, A. C. (2007). On-line detection and classification of power quality disturbances in a single-phase power system. In *Power Engineering, Energy and Electrical Drives, 2007. POWERENG 2007. International Conference on*, pages 713–718. IEEE.
- [Ramírez-Castro and Montejo 2011] Ramírez-Castro, R. I. and Montejo, L. A. (2011). Transformada de hilbert, descomposición modal empírica y sus aplicaciones en el análisis de vibraciones libres. *Hilbert transform, empirical mode decomposition and its applications to free vibration analysis.(English)*, 11 (2), pages 121–132.
- [Ramos et al. 2008] Ramos, P. M., Janeiro, F. M., Serra, A. C., et al. (2008). Pq monitoring system for real-time detection and classification of disturbances in a single-phase power system. *IEEE Transactions on Instrumentation and Measurement*, 57(8):1725–1733.
- [Rens 2009] Rens, J. (2009). On the development of a power quality benchmarking model. In *Electrical Power Quality and Utilisation, 2009. EPQU 2009. 10th International Conference on*, pages 1–5. IEEE.
- [Ribeiro et al. 2016] Ribeiro, E., Dias, G., Barbosa, B., and Ferreira, D. (2016). Real-time system for automatic classification of power quality disturbances. In *Harmonics and Quality of Power (ICHQP), 2016 17th International Conference on*, pages 908–913. IEEE.
- [Ribeiro et al. 2018] Ribeiro, E. G., Mendes, T. M., Dias, G. L., Faria, E. R., Viana, F. M., Barbosa, B. H., and Ferreira, D. D. (2018). Real-time system for automatic detection and classification of single and multiple power quality disturbances. *Measurement*, 128:276–283.
- [Ribeiro et al. 2015] Ribeiro, M. V., Duque, C. A., and Cerqueira, A. S. (2015). Classificador de distúrbios de qualidade de energia elétrica de baixa complexidade computacional - low complexity event classification (lcec).

- [Rodriguez et al. 2016] Rodriguez, M. V., Hernandez, L. A. M., Rangel, J. P. B., and Gonzalez, A. D. (2016). Real-time monitoring of voltage variations using mathematical morphology. *IEEE Latin America Transactions*, 14(5):2138–2145.
- [Rodríguez-Fdez et al. 2015] Rodríguez-Fdez, I., Canosa, A., Mucientes, M., and Bugarín, A. (2015). STAC: a web platform for the comparison of algorithms using statistical tests. In *Proceedings of the 2015 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*.
- [Ruiz-Llata et al. 2011] Ruiz-Llata, M., Guarnizo, G., and Boya, C. (2011). Embedded power quality monitoring system based on independent component analysis and svms. In *Neural Networks (IJCNN), The 2011 International Joint Conference on*, pages 2229–2234. IEEE.
- [Sadigh and Smedley 2016] Sadigh, A. K. and Smedley, K. (2016). Fast and precise voltage sag detection method for dynamic voltage restorer (dvr) application. *Electric Power Systems Research*, 130:192–207.
- [Salem et al. 2007a] Salem, M. E., Mohamed, A., and Samad, S. A. (2007a). Power quality disturbance detection using dsp based continuous wavelet transform. *Journal of Applied Sciences*, 7(6):893–902.
- [Salem et al. 2006] Salem, M. E., Mohamed, A., Samad, S. A., and Jones, O. (2006). Dsp implementation of power quality disturbance analysis using continuous wavelet transform. In *Power and Energy Conference, 2006. PECon'06. IEEE International*, pages 94–98. IEEE.
- [Salem et al. 2005] Salem, M. E., Mohamed, A., Samad, S. A., and Mohamed, R. (2005). Development of a dsp-based power quality monitoring instrument for real-time detection of power disturbances. In *Power Electronics and Drives Systems, 2005. PEDS 2005. International Conference on*, volume 1, pages 304–307. IEEE.
- [Salem et al. 2007b] Salem, M. E., Mohamed, A., Samad, S. A., and Yahya, I. (2007b). Software tool for real time power quality disturbance analysis and classification. In *Re-*

- search and Development, 2007. SCORED 2007. 5th Student Conference on*, pages 1–5. IEEE.
- [Samotyj 2001] Samotyj, M. (2001). The cost of power disturbance to industrial and digital economy companies. *Consortium for electrical infrastructure to support a digital society, an initiative by EPRI and the Electrical Innovation Institute*.
- [Seymour and Horsley 2001] Seymour, J. and Horsley, T. (2001). Sete tipos de problemas no fornecimento.
- [Silva et al. 2010] Silva, I. N. d., Spatti, D. H., and Flauzino, R. A. (2010). *Redes Neurais Artificiais, Curso prático*. ArtLiber Ed.
- [Sobrinho et al. 2018] Sobrinho, A. S. F., Flauzino, R. A., Liboni, L. H. B., and Costa, E. C. M. (2018). Proposal of a fuzzy-based pmu for detection and classification of disturbances in power distribution networks. *International Journal of Electrical Power & Energy Systems*, 94:27–40.
- [Szweda 2009] Szweda, M. (2009). Dwt analysis of selected transient and notching disturbances. *signal (comparison with the prevent steady-state waveform)*, 3:15.
- [Targosz and Manson 2007] Targosz, R. and Manson, J. (2007). Pan-european power quality survey. In *Electrical Power Quality and Utilisation, 2007. EPQU 2007. 9th International Conference on*, pages 1–6. IEEE.
- [Theodoridis 2008] Theodoridis, S. (2008). Pattern recognition.
- [Theodoridis and Koutroumbas 2008] Theodoridis, S. and Koutroumbas, K. (2008). *Pattern Recognition*. Elsevier Science.
- [Tomic et al. 2007] Tomic, J. J., Kusljevic, M. D., and Vujicic, V. V. (2007). A new power system digital harmonic analyzer. *IEEE transactions on power delivery*, 22(2):772–780.
- [Ukil and Živanović 2006] Ukil, A. and Živanović, R. (2006). Abrupt change detection in power system fault analysis using adaptive whitening filter and wavelet transform. *Electric Power Systems Research*, 76(9):815–823.

- [Valtierra-Rodriguez et al. 2016] Valtierra-Rodriguez, M., Granados-Lieberman, D., Torres-Fernandez, J. E., Rodríguez-Rodríguez, J. R., and Gómez-Aguilar, J. F. (2016). A new methodology for tracking and instantaneous characterization of voltage variations. *IEEE Transactions on Instrumentation and Measurement*, 65(7):1596–1604.
- [Wang et al. 2012] Wang, A., Pan, F., Li, Y., and Tao, R. (2012). The design of power quality detecting system based on omap-1138. In *Control and Modeling for Power Electronics (COMPEL), 2012 IEEE 13th Workshop on*, pages 1–4. IEEE.
- [Wang et al. 2004] Wang, M., Rowe, G. I., and Mamishev, A. V. (2004). Classification of power quality events using optimal time-frequency representations-part 2: application. *IEEE Transactions on Power Delivery*, 19(3):1496–1503.
- [Wang and Chen 2019] Wang, S. and Chen, H. (2019). A novel deep learning method for the classification of power quality disturbances using deep convolutional neural network. *Applied energy*, 235:1126–1140.
- [Wang and Zhu 2002] Wang, Z. and Zhu, S. (2002). Comparative study on power quality disturbance magnitude characterization. In *Power System Technology, 2002. Proceedings. PowerCon 2002. International Conference on*, volume 1, pages 106–111. IEEE.
- [Wasilewski] Wasilewski, F. Daubeches 4 wavelet db4. <http://wavelets.pybytes.com/wavelet/db4/>. Accessed: 2019-01-11.
- [Wipo 2017] Wipo (2017 (Acessado em 27 de junho de 2017)). *Perguntas e Respostas Sobre o PCT*.
- [Wohlin et al. 2012] Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., and Wesslén, A. (2012). *Experimentation in software engineering*. Springer Science & Business Media.
- [Wolf 2008] Wolf, W. (2008). *Computers as components: principles of embedded computing system design*. Elsevier.
- [Xiang-wu et al. 2006] Xiang-wu, Y., Wei, K., He-ming, L., Li-xia, Z., Zhi-jun, H., and Rui, K. (2006). The application of dsp and virtual instrument to online power quality detection.

- In *Power Electronics, Electrical Drives, Automation and Motion, 2006. SPEEDAM 2006. International Symposium on*, pages 761–764. IEEE.
- [Xu et al. 2014] Xu, W., Xu, G., and Yuan, H. (2014). High performance distributed power quality monitoring ied used in smart grid. In *Electricity Distribution (CICED), 2014 China International Conference on*, pages 706–710. IEEE.
- [Yalcin et al. 2018] Yalcin, T., Ozdemir, M., Kostyla, P., and Leonowicz, Z. (2018). Discrete wavelet analysis of supra-harmonic emissions in smart grid (pv inverter implementation). In *2018 IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)*, pages 1–5. IEEE.
- [Yang and Wen 2006] Yang, G. and Wen, B. (2006). A device for power quality monitoring based on arm and dsp. In *Industrial Electronics and Applications, 2006 IST IEEE Conference on*, pages 1–5. IEEE.
- [Yang et al. 2012] Yang, Y., Blaabjerg, F., and Zou, Z. (2012). Benchmarking of voltage sag generators. In *IECON 2012-38th Annual Conference on IEEE Industrial Electronics Society*, pages 943–948. IEEE.
- [Yau et al. 2015] Yau, H.-T., Wang, M.-H., Wang, T.-Y., and Chen, G. (2015). Signal clustering of power disturbance by using chaos synchronization. *International Journal of Electrical Power & Energy Systems*, 64:112–120.
- [Yingkayun et al. 2009] Yingkayun, K., Premrudeepreechacharn, S., and Oranpiroj, K. (2009). A power quality monitoring system for real-time fault detection. In *Industrial Electronics, 2009. ISIE 2009. IEEE International Symposium on*, pages 1846–1851. IEEE.
- [Yongxian et al. 2011] Yongxian, S., Yuan, F., Chenlong, G., and Naibao, H. (2011). Design of electric power parameter monitoring system based on dsp and cpld. In *Intelligent Computation Technology and Automation (ICICTA), 2011 International Conference on*, volume 1, pages 638–641. IEEE.

- [Zhang and Li 2010] Zhang, M. and Li, K. (2010). Dsp-fpga based real-time power quality disturbances classifier. In *Transmission and Distribution Conference and Exposition, 2010 IEEE PES*, pages 1–6. IEEE.
- [Zheng et al. 2014] Zheng, X. Y., Chen, C. P., and Jiang, J. A. (2014). An embedded-based distributed private cloud: Power quality event classification. In *High Performance Computing and Communications, 2014 IEEE 6th Intl Symp on Cyberspace Safety and Security, 2014 IEEE 11th Intl Conf on Embedded Software and Syst (HPCC, CSS, ICESS), 2014 IEEE Intl Conf on*, pages 1237–1242. IEEE.
- [Zhong et al. 2018] Zhong, T., Zhang, S., Cai, G., and Huang, N. (2018). Power-quality disturbance recognition based on time-frequency analysis and decision tree. *IET Generation, Transmission & Distribution*, 12(18):4153–4162.

Apêndice A

Qualidade da Energia Elétrica

A qualidade da energia elétrica é definida como a entrega das medidas elétricas (tensão, corrente e/ou potência) sem grande variação. Em outras palavras, as ondas senoidais que representam estas medidas devem possuir no máximo uma variação de 5% em relação à onda perfeitamente senoidal [Kennedy 2000]. Em [Association et al. 2009] é acrescentado ainda que a QEE se refere à variedade de efeitos eletromagnéticos que podem ocorrer em determinada medida elétrica (seja da corrente, tensão ou potência elétrica).

Não obstante os desenvolvimentos na área de distribuição da energia elétrica, efeitos eletromagnéticos continuam a afetá-la, mesmo nos melhores sistemas [Edomah 2009]. Isso ocorre devido ao crescente número de equipamentos eletro-eletrônicos que geram variações elétricas [Association et al. 2009, Landi and Luiso 2008], logo, a QEE, sendo de interesse direto de consumidores e centros de distribuição tornaram a área propícia para pesquisas acadêmicas. Tanto é que a busca por maneiras de medir e monitorar a QEE tornou-se amplamente difundida (principalmente no cenário acadêmico, sendo, só no ano de 2007, produzidos mais de 800 artigos relacionados ao tema) [Bollen et al. 2010].

Frente a este cenário com grande propensão a estudos relacionados a área de QEE, o IEEE (Instituto de Engenharia Eletrotécnica e Eletrônica) criou uma padronização para o tratamento das variações não esperadas (distúrbios) nas medidas de variáveis elétricas. Sendo assim, o IEEE almeja tornar possível a análise da QEE em um consenso mundial, apresentando uma unificação da nomenclatura dada aos distúrbios elétricos (de acordo com algumas variáveis como tempo de duração, frequência e amplitude). Além disso, a padronização facilita a normalização das classificações realizadas por sistemas de monitoramento, emprega

padrões nos equipamentos de medição, definindo, por exemplo, valores como taxa de amostragem do conversor analógico digital a ser utilizado [Association et al. 2009].

Dentre os distúrbios identificados e normalizados pelo IEEE estão transientes, variações de curta e longa duração no valor RMS (na amplitude do sinal), desbalanceamento, distorções na forma de onda, flutuações na tensão, e variações na frequência da potência elétrica. Dentro de cada um dessas sete categorias existem tipos de distúrbios comumente identificados nos sistemas de monitoramento, como por exemplo o afundamento e elevação de tensão, que estão dentro da categoria de variações de curta duração no valor RMS. De modo a ter um maior conhecimento a respeito de cada tipo de distúrbio, nas seguintes subseções se apresentam maiores detalhes de cada um dos conjuntos citados (de acordo com [Association et al. 2009]).

A.1 Transientes

Também conhecido como transitórios, de acordo com [Seymour and Horsley 2001], são os tipos de perturbação energética com maior grau de prejuízo aos equipamentos conectados à rede elétrica onde eles ocorrem. Os transientes podem ser divididos em dois tipos diferentes de distúrbios, os impulsivos e os oscilatórios.

A.2 Transiente Impulsivo

Dentro dos transientes impulsivos ainda há uma subdivisão de acordo com o tempo de duração destes distúrbios, podendo ir da escala de nanosegundos a milissegundos. Consoante [Kennedy 2000], tal tempo é definido como o momento em que houve um pico de energia até o momento de estabilização da medida elétrica na qual ocorreu tal pico. Por isso, informalmente este tipo de distúrbios elétrico também é conhecido como picos de energia. Na Figura A.1 é apresentada uma ilustração deste tipo de distúrbio.

A causa mais comum de transientes impulsivos está relacionada a queda de raios em ambientes de transmissão ou distribuição de energia elétrica, em casos assim o pico pode chegar a 2000V em $1.2\mu s$ [Kennedy 2000]. Normalmente, a utilização de componentes resistivos conseguem reduzir o efeito deste tipo de distúrbio.

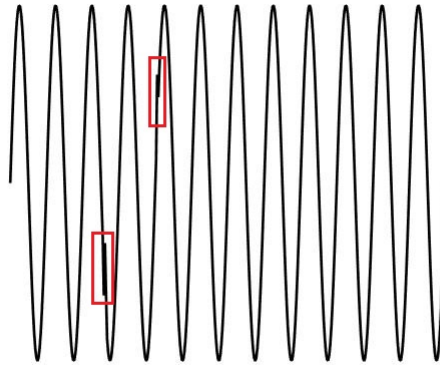


Figura A.1: Transiente Impulsivo

A modelagem matemática para o transiente impulsivo é dada por [Granados-Lieberman et al. 2013] e é apresentada na Equação A.1.

$$v(t) = A \sin(2\pi ft) + \alpha [|\sin(2\pi \beta t)|] \quad (\text{A.1})$$

em que A é a amplitude, $0.1 \leq \alpha \leq 0.5$ e $1 < \beta \leq 10$.

A.3 Transiente Oscilatório

Enquanto isso, os distúrbios do tipo transiente oscilatório são subdivididos de acordo com a frequência da oscilação, podendo ser de baixa frequência, média ou alta (<5kHz, entre 5kHz e 500kHz, entre 500kHz e 5MHz, respectivamente) [Association et al. 2009]. Na Figura A.2 é apresentado uma ilustração do distúrbio de transiente oscilatório.

A duração dos transientes oscilatórios vão da ordem de ms (para oscilações de baixa frequência), em torno de $20\mu s$ (para oscilações de média frequência) e em torno de $5\mu s$ (para oscilações de alta frequência). Normalmente, segundo [Kennedy 2000], os transientes oscilatório são causados pela ressonância devido a queda de algum raio na rede de transmissão ou distribuição de energia elétrica.

A modelagem matemática, segundo [Li et al. 2016] para o transiente oscilatório pode ser dada pela Equação A.2.

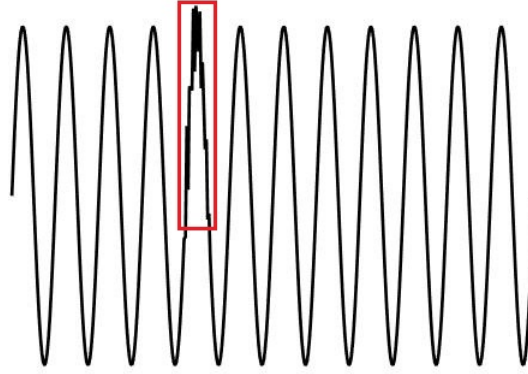


Figura A.2: Transiente Oscilatório

$$v(t) = \sin(2\pi ft) + \alpha e^{-(t-t_1)/\tau_{osc}} (u(t-t_1) - u(t-t_2)) \sin(2\pi if(t-t_1)) \quad (\text{A.2})$$

em que $0.3 \leq \alpha \leq 0.5$, $3ms \leq \tau_{osc} \leq 50ms$, $0.5T \leq t_2 - t_1 \leq 3T$, $6 \leq i \leq 18$, T o período do seno puro e $u(t)$ a função degrau (0 para $t < 0$ e 1 para $t > 0$).

A.4 Variações de Curta Duração no Valor RMS

Consoante [Association et al. 2009], a causa principal dos efeitos de variações curtas em determinada medida elétrica é causada por algum problema na energização de determinada carga conectada à rede elétrica como, por exemplo, a conexão de uma carga que requer uma alta quantidade de corrente.

A.4.1 Interrupção

Interrupção refere-se a uma perda total, ou quase total (menos de 10% do valor nominal), do fornecimento de energia em uma ou mais fases [Kennedy 2000]. O IEEE, em [Association et al. 2009], divide a interrupção em dois diferentes subconjuntos, as interrupções momentâneas e as interrupções temporárias, definidas de acordo com o tempo de duração de cada

uma. Interrupções momentânea tem em média uma duração que varia de 0.5 ciclos a 3s, enquanto que as temporárias possuem uma média de 3s a 1min. Na Figura A.3 é apresentado uma ilustração do distúrbio de interrupção momentânea.

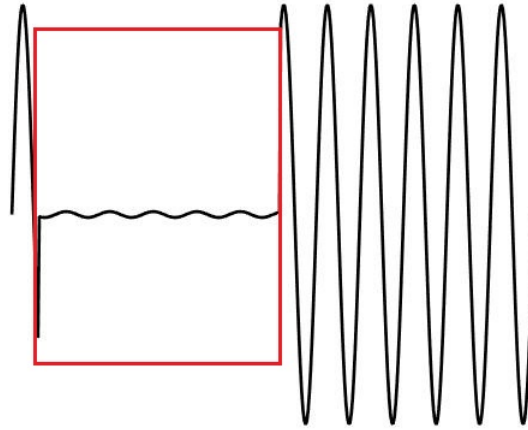


Figura A.3: Interrupção

Normalmente a causa de interrupção é resultado de alguma falha no fornecimento da energia elétrica, podendo ser causada por fontes naturais, como a queda de raios, quanto por problemas nas linhas de transmissão [Seymour and Horsley 2001]. Ainda é possível que os próprios consumidores causem um curto circuito que ocasione uma interrupção.

Frente a isto, deve-se ressaltar que a interrupção no fornecimento de energia elétrica para determinada localidade é inteiramente prejudicial, visto que pode gerar perdas irreparáveis em diferentes cenários. Em indústrias pode causar a interrupção na produção (aumentando custos), em hospitais a parada de tratamentos, em servidores a suspensão na entrega de serviços etc.

A modelagem matemática para a interrupção é apresentada na Equação A.3 [Granados-Lieberman et al. 2013].

$$v(t) = -\alpha A(u(t - t_1) - u(t - t_2)) \sin(2\pi ft) \quad (\text{A.3})$$

Em que $u(t)$ representa a função degrau, $0.9 \leq \alpha \leq 1$ e $t_1 < t_2$.

A.4.2 Afundamento de Tensão

O afundamento de tensão, como o próprio nome sugere, refere-se à redução na amplitude da tensão elétrica em um determinado período de tempo (0.5 ciclos a 1 minuto) [Seymour and Horsley 2001]. A Figura A.4 apresenta distúrbio do tipo afundamento de tensão como forma de ilustração.

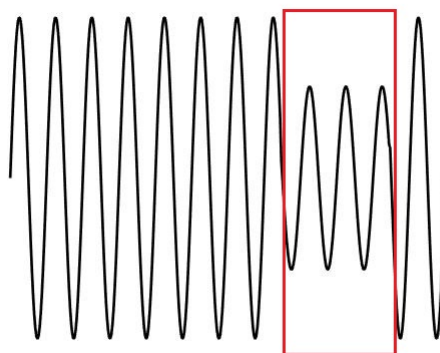


Figura A.4: Afundamento de Tensão

Este tipo de distúrbio elétrico pode causar, por exemplo, danos em motores, levando-os a estancar ou superaquecer, em computadores ocasiona falhas e em *drivers* de ajuste de velocidade desligando-os [Kennedy 2000]. Frente a este cenário, percebe-se que este é um distúrbio com alto impacto em ambientes industriais pois, a parada de um motor pode, por exemplo, interromper toda uma linha de produção.

Ainda de acordo com [Kennedy 2000], o distúrbio de afundamento pode ser causado por falhas no sistema de distribuição, transmissão ou a partir de cargas conectadas pelos consumidores da energia elétrica. Sendo assim, como forma de mitigar este problema pode-se utilizar, por exemplo, transformadores ferromagnéticos, equipamentos de armazenamento de energia (como geradores e no-breaks), UPS (*uninterruptible power supply*) ou restauradores de tensão dinâmicos (DVRs).

O afundamento de tensão pode ser subdividido em três subcategorias (de acordo com a duração): instantâneo, momentâneo e temporário. Os instantâneos possuem um período de 0.5 a 30 ciclos, os momentâneos vão de 30 ciclos a 3s e os temporários de 3s a 1min [Association et al. 2009]. Em relação a redução na amplitude da tensão, segundo o padrão 1159 do IEEE, deve ser de 10% a 90% em relação a amplitude original. Frente a isso, o

afundamento de tensão pode ser modelado matematicamente como mostra a Equação A.4 [Granados-Lieberman et al. 2013].

$$v(t) = (1 - \alpha(u(t - t_1) - u(t - t_2)))A \sin(2\pi ft) \quad (\text{A.4})$$

Em que $u(t)$ é a função degrau, $0.1 \leq \alpha \leq 0.9$, $t_1 < t_2$, A a amplitude do sinal e f a frequência.

A.4.3 Elevação de Tensão

A elevação de tensão é um tipo de distúrbio elétrico que causa o aumento no valor RMS da tensão elétrica durante um período de tempo que varia de 0.5 ciclos a 1 minuto [Seymour and Horsley 2001]. Na Figura A.5 é apresentado uma ilustração deste tipo de distúrbio.

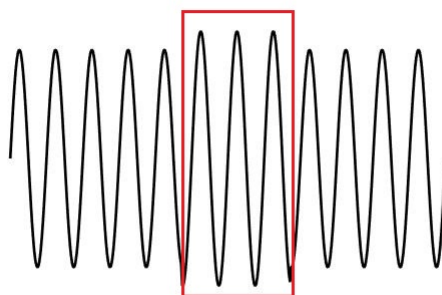


Figura A.5: Elevação de Tensão

As elevações de tensão são causadas normalmente por conexões na rede elétrica de equipamentos de alta impedância, redução abrupta de carga, falha monofásica em um sistema trifásico [Seymour and Horsley 2001], um banco de capacitor de valor elevado [Association et al. 2009] ou ainda o crescimento de um árvore que atrapalhe no sistema de distribuição [Kennedy 2000].

Apesar da grande quantidade de fatores que podem causar o distúrbio de elevação de tensão, ele ainda é menos comum que o seu inverso (afundamento de tensão) [Association et al. 2009], todavia, assim como este, a elevação de tensão também é subdividida em três subcategorias de acordo com a duração do distúrbio e o aumento relativo no valor da tensão

(instantânea, de 0.5 ciclos a 30 ciclos com aumento de 10% a 80%; momentânea, de 30 ciclos a 3s com aumento de 10% a 40%; temporária, de 3s a 1min, com aumento de 10% a 20%). Formulando matematicamente, a elevação de tensão é apresentada como na Equação A.5 [Granados-Lieberman et al. 2013].

$$v(t) = (1 + \alpha(u(t - t_1) - u(t - t_2)))A \sin(2\pi ft) \quad (\text{A.5})$$

Em que $u(t)$ é a função degrau, α depende da duração do distúrbio (podendo variar de 0.1 a 0.8), $t_1 < t_2$, A a amplitude do sinal e f a frequência.

A.5 Variações de Longa Duração no Valor RMS

As variações de longa duração no valor RMS de determinada medida elétrica possui características semelhantes das variações de curta duração, modificando apenas o tempo de duração dos distúrbios, e, de maneira a não confundir, devido a diferentes causas e efeitos na rede elétrica, o padrão 1159 do IEEE utiliza nomenclatura diferentes, interrupção passa a ser interrupção sustentada, afundamento de tensão torna-se subtensão e elevação de tensão torna-se sobretensão [Association et al. 2009]. Todos os distúrbios nesta categoria possuem duração superior a 1 minuto.

Deve-se ressaltar apenas o fato de que, no caso da interrupção sustentada, o valor da tensão durante o tempo de interrupção deve ser 0, enquanto que no caso da interrupção de curta duração o valor nominal da tensão poderia ser de até 10% o valor normal. Já a subtensão deve possuir um valor de 80% a 90% o valor RMS e a sobretensão deve ser de 10% a 20% a mais do valor RMS. A Tabela A.1 apresenta um resumo em relação a estes valores.

Tabela A.1: Variações de Longa Duração no Valor RMS (Fonte: [Association et al. 2009])

Subcategoria	Valor RMS	Tempo
Interrupção Sustentada	0	> 1min
Subtensão	-20% a -10%	> 1min
Sobretensão	+10% a +20%	> 1min

Ainda relacionado a variações de longa duração no valor RMS, existe uma quarta sub-

categoria relacionada à corrente elétrica e não à tensão, como os distúrbios até aqui apresentados (apesar que as variações ocorridas na tensão elétrica é refletida na corrente). Esta quarta subcategoria é a **sobrecarga de corrente**, que, assim como os demais distúrbios apresentados nesta subseção, ocorre com uma duração acima de 1 minuto [Association et al. 2009].

A.6 Desbalanceamento

Enquanto os demais distúrbios apresentados até aqui referem-se a problemas que podem ser monitorados monofasicamente, o distúrbio relacionado ao desbalanceamento elétrico só pode ser visualizado em ambientes nos quais o número de fases é superior a um. Por exemplo, em sistemas trifásicos, o desbalanceamento é definido como a razão entre a magnitude da sequência de componentes negativa com a magnitude da sequência de componentes positiva e é expresso em termos de porcentagem [Association et al. 2009] como apresentado na Equação A.6.

$$\%desbalanceamento = \frac{|V_{neg}|}{|V_{pos}|} \times 100\% \quad (A.6)$$

No entanto, realizar a medida de V_{neg} e V_{pos} não é simples, pois, além do valor RMS é necessário a medição do ângulo de fase. Desta forma, segundo [Association et al. 2009] é possível realizar o cálculo do desbalanceamento utilizando a fórmula apresentada na Equação A.7.

$$\%desbalanceamento = \sqrt{\frac{1 - \sqrt{3 - 6\beta}}{1 + \sqrt{3 - 6\beta}}} \times 100\% \quad (A.7)$$

Em que, β é dado por:

$$\beta = \frac{|V_{AB}|^4 + |V_{BC}|^4 + |V_{CA}|^4}{(|V_{AB}|^2 + |V_{BC}|^2 + |V_{CA}|^2)^2} \quad (A.8)$$

desta forma, é possível medir facilmente os valores de V_{AB} , V_{BC} e V_{CA} utilizando sensores disponíveis no mercado.

Assim como grande parte dos distúrbios apresentados, o desbalanceamento de fase pode ser visualizado tanto com a medida de corrente quanto de tensão elétrica [Association et al.

2009]. Além disso, esse tipo de distúrbio é normalmente evitado no momento da instalação elétrica, fazendo o balanceamento correto das cargas conectadas às fases.

Segundo [Kennedy 2000], o desbalanceamento de fase pode tolerar, especialmente em motores (mais afetados por este tipo de distúrbio), **um percentual de 2% de desbalanceamento**, valores maiores que isso podem ocasionar superaquecimento do motor. Apesar de poder ser evitado no momento da instalação, a falha de equipamentos ou o seu mau funcionamento pode fazer com que haja problemas de balanceamento e, por este motivo, é importante ter sistemas capazes de realizar o monitoramento constante do balanceamento em ambientes trifásicos.

A.7 Distorções na Forma de Onda

Os distúrbios denominados como distorções na forma de onda são definidos como aqueles que afetam diretamente a senoide pura, com desvios em relação ao seu estado estacionário [Association et al. 2009], normalmente as variações desta categoria podem ser melhor verificadas a partir do espectro de frequência do sinal.

A seguir são apresentados os cinco tipos primários de distorções na forma de onda.

A.7.1 Nível DC

Em média, a tensão fornecida por um sistema elétrico é de 0V, quando há a presença de uma tensão ou corrente contínua (um nível DC) esta média é modificada, sendo, normalmente elevada, o que significa que, graficamente, haverá um deslocamento da senoide para cima (em relação ao eixo horizontal).

O nível DC é facilmente detectado através do espectro de frequência, o valor para a frequência zero deve ser também nulo, caso haja algum valor para tal frequência significa que há um nível DC indesejado.

O aumento no nível DC da tensão ou corrente elétrica é ocasionado geralmente por falha de retificadores dentro dos equipamentos que possuem conversores de corrente alternada para contínua [Seymour and Horsley 2001].

Como consequência do aumento no nível DC problemas como superaquecimento ou saturação de transformadores podem ocorrer. Desta forma, ficam eles incapazes de entregar

a potência desejada. Além disso este tipo de distúrbio pode ocasionar instabilidade nos equipamentos de carga eletrônicas [Seymour and Horsley 2001].

A.7.2 Harmônicas

A frequência fundamental da senoide fornecida pela rede elétrica é previamente conhecida (normalmente 60Hz ou 50Hz). Quando, além da frequência fundamental, a senoide contém frequências harmônicas (que são múltiplas da fundamental) diz-se que o sinal está sob o efeito do distúrbio de harmônicas [Edomah 2009], que também pode ser encontrado na literatura como distorção harmônica [Kennedy 2000]. A Figura A.6 apresenta uma ilustração para o efeito de harmônicas em uma senoide.

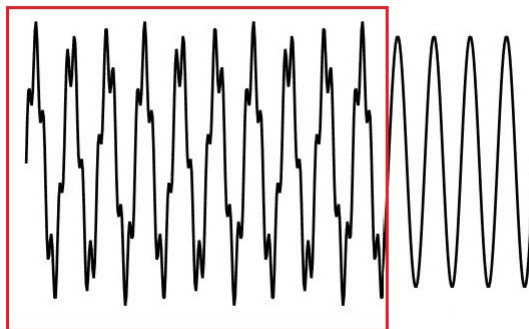


Figura A.6: Harmônicas

As harmônicas são normalmente causadas devido a cargas não-lineares conectadas à rede elétrica [Kennedy 2000], o que tem aumentado com o avanço tecnológico, visto que hoje existe um número maior de dispositivos semicondutores conectados à rede elétrica [Landi and Luiso 2008], além de sistemas de ressonância.

Os problemas relacionados à presença de harmônicas na rede elétrica estão relacionados ao superaquecimento de transformadores e motores, queima de fusíveis, má operação de dispositivos de medição [Kennedy 2000].

Frente a este problema, a indústria de informática, que construía os principais equipamentos geradores de harmônicas nas redes elétricas, com intuito de mitigar o problema, criou um design novo para as fontes comutadas (de acordo com um padrão internacional),

utilizando, basicamente, um superdimensionamento dos condutores neutros e a instalação de transformadores de fator k além de filtro de harmônicas [Association et al. 2009].

As harmônicas possuem frequência que variam entre um mínimo de 0KHz e máximo de 9kHz e uma variação de 0% a 20% em relação a amplitude normal da senoide [Association et al. 2009], sendo que, as frequências das harmônicas podem ser tanto múltiplos pares quanto ímpares da frequência fundamental [Landi and Luiso 2008]. Matematicamente, a distorção harmônica pode ser modelada como apresentado na Equação A.9 [Granados-Lieberman et al. 2013].

$$v(t) = A \sin(2\pi ft) + \sum_{m=1}^M A_m \sin(2\pi h_m ft) \quad (\text{A.9})$$

Em que A representa a amplitude da senoide pura, A_m a amplitude da harmônica m (pode ter de 0 a 20% do valor de A), M é o número total de harmônicas e h_m é o múltiplo da frequência fundamental, que, segundo [Granados-Lieberman et al. 2013], pode variar entre 2 e 40.

A.7.3 Inter-harmônicas

Muito semelhante às harmônicas, o que diferencia as inter-harmônicas em termos matemáticos é o fato de suas frequências não serem múltiplos inteiros da frequência fundamental (como acontece com aquelas) [Association et al. 2009].

Este distúrbio resulta, normalmente, de "equipamentos elétricos como conversores de frequência estática, motores a indução e dispositivos de geração de arco" [Seymour and Horsley 2001]. Apesar de possuir uma causa conhecida, o mesmo não pode-se dizer a respeito do efeito [Association et al. 2009], visto que não são conhecidos todos os possíveis efeitos deste distúrbio nos equipamentos. O que se conhece até então é que as inter-harmônicas causam efeitos visuais indesejados em monitores e lâmpadas.

Além disso, pode-se citar ainda as supra-harmônicas e subharmônicas, que são distúrbios que eram até então pouco visualizados na literatura, que, inclusive, não possuem classificação dentro do padrão da IEEE mas que vem ganhando espaço com o aumento da utilização de smartgrids, devido a necessidade de comunicação entre diversos pontos da rede de distribuição de energia [Morales-Velazquez et al. 2017]. As supra-harmônicas, por exemplo,

pode causar problemas na comunicação entre os pontos, e, por isso vem ganhando espaço em pesquisas recentes na área de análise da QEE [Mendes et al. 2019].

A.7.4 *Notching*

O distúrbio do tipo *notching* é, na realidade, o mesmo distúrbio de transiente impulsivo, porém, com uma frequência de repetição, normalmente de meio ciclo (por isso está categorizado como distorção na forma de onda) [Seymour and Horsley 2001]. A Figura A.7 ilustra este tipo de variação elétrica.

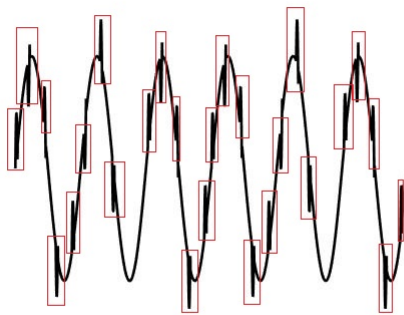


Figura A.7: *Notching*

Os *notchings* são causados geralmente por equipamentos eletrônicos como mecanismos de velocidade variável, atenuadores de luz e soldadores a arco (em seu funcionamento normal) [Seymour and Horsley 2001]. O distúrbio também pode ser causado quando a corrente é de uma fase para outra [Association et al. 2009].

Os problemas relacionados a este tipo de distúrbio dependerá, segundo [Association et al. 2009], da indutância da fonte e do isolamento entre o conversor, a magnitude da corrente e o ponto sendo monitorado. Tais problemas vão desde parada do sistema, perda de dados e problemas na transmissão de dados [Seymour and Horsley 2001].

Como forma de modelar este distúrbio, a Equação A.10 apresenta uma forma como fazê-lo [Granados-Lieberman et al. 2013].

$$v(t) = A \sin(2\pi ft) + \alpha [|\sin(2\pi \beta t)|] \quad (\text{A.10})$$

Em que A é a amplitude da tensão (ou corrente elétrica), f a frequência fundamental, β a frequência com que ocorre o *notching*, normalmente variando de 60Hz a 240Hz, e α a amplitude do distúrbio, que varia de 0 a 0.2.

A.7.5 Ruído

O ruído é um tipo de sinal elétrico que normalmente ocorre com uma frequência inferior a 200kHz e que não pode ser classificado como nenhum dos tipos de distorção da onda (harmônica, inter-harmônica, *notching*) nem mesmo como transientes [Association et al. 2009]. A Figura A.8 apresenta um sinal elétrico com ruído.

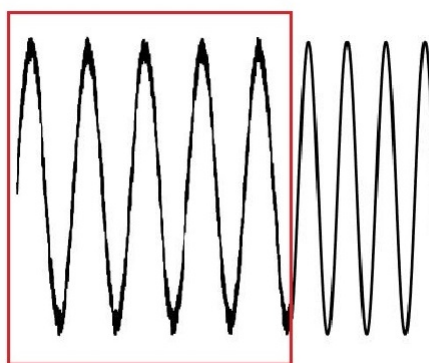


Figura A.8: Ruído

Segundo [Seymour and Horsley 2001], o ruído pode ser gerado por dispositivos eletrônicos alimentados eletricamente, ou ainda circuitos de controle, soldadoras a arco (assim como os *notching*), fonte de alimentação, transmissores de rádio, entre outros.

Entre os problemas causados pela presença de ruído na rede elétrica está o erro nos dados lidos por determinado aparelho (como sensores), distorções nos monitores de vídeo, mau funcionamento de determinados equipamentos etc [Seymour and Horsley 2001]. Além disso, um setor altamente afetado pelos ruídos elétricos é o de telecomunicações, pois seus equipamentos podem transferir o ruído por meio do canal de comunicação e perder na qualidade deste sinal [Kennedy 2000].

Existem maneiras de proteger os sinais elétricos dos ruídos, entre os quais, os mais comuns são: a instalação de filtros, blindagem de cabos, isolamentos das cargas elétricas, ins-

talação de transformadores aterrados e com blindagem, e realocação de cargas [Seymour and Horsley 2001].

A.8 Flutuações de Tensão

As flutuações de tensão, também conhecidas como *flickers* (em inglês), são um envelopamento da forma de onda fundamental em uma senoide com uma frequência inferior (não ultrapassando, normalmente, 25Hz) [Association et al. 2009]. Os valores da tensão também podem variar aleatoriamente entre os valores de 95% a 105%. É ilustrado na Figura A.9 o distúrbio de flutuação de tensão.

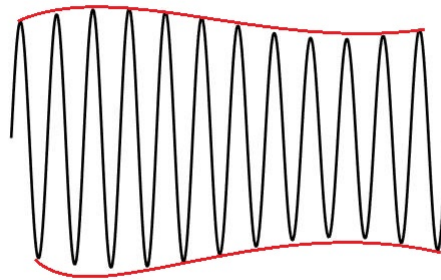


Figura A.9: Flutuação de Tensão

Este tipo de distúrbio geralmente causa maiores transtornos na produção humana que em problemas direto nos aparelhos eletro-eletrônicos. Isto deve-se ao fato de as flutuações de tensão façam com que a luminosidade fique oscilando, ocasionando problemas como dores de cabeça, estresse e irritabilidade [Kennedy 2000]. Dessa forma, como maneira de mitigar este problema pode-se utilizar SVCs (controlador estático de reativos), que, apesar de caros são efetivos na resolução do problema [Seymour and Horsley 2001].

Matematicamente, o distúrbio de flutuação de tensão pode ser modelado como apresentado na Equação A.11 [Granados-Lieberman et al. 2013, Li et al. 2016].

$$v(t) = (1 + \alpha \sin(2\pi f_r t)) A \sin(2\pi f t) \quad (\text{A.11})$$

Em que $-0.05 < \alpha \leq 0.05$ [Association et al. 2009], $1 \leq f_r \leq 25$ [Seymour and Horsley 2001], A é a amplitude da senoide e f é a sua frequência fundamental.

A.9 Variações de Frequência

As variações na frequência é um dos distúrbios menos comuns nas redes elétricas estáveis (redes conectadas diretamente ao sistema elétrico) [Seymour and Horsley 2001]. O problema, portanto, é mais comum em locais onde encontra-se geradores de reserva ou uma infra-estrutura de rede pobre. Em [Association et al. 2009] é adicionado ainda que a conexão de novos dispositivos na rede pode ocasionar também numa variação da frequência elétrica momentaneamente. Na Figura A.10 é ilustrado um exemplo deste tipo de distúrbio.

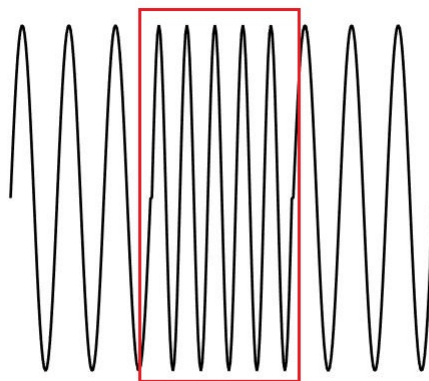


Figura A.10: Variação na Frequência

As falhas na frequência geralmente não geram problemas em equipamentos da área de tecnologia da informação, mas, causam problemas em equipamentos síncronos [Seymour and Horsley 2001] ou ainda em maquinário rotativo [Association et al. 2009]. Este problema pode ser mitigado verificando-se frequentemente como está o funcionamento de geradores locais (já que estes são os maiores causadores deste tipo de distúrbio).

Apêndice B

Sistemas Embarcados

Sistemas embarcados são dispositivos programáveis, assim como computadores, entretanto, ao contrário destes, possuem um propósito específico [Wolf 2008]. Além disso, o sistema embarcado é capaz de processar informações dentro de um outro produto maior, como por exemplo automóveis, e, normalmente não estão diretamente visíveis para o usuário [Marwedel 2006].

O avanço tecnológico, principalmente no que diz respeito à construção de uma arquitetura estruturada em chips com milhões de transistores fez com que o desenvolvimento de microprocessadores crescesse substancialmente ao ponto de pequenos aparelhos hoje serem capazes de fazer o processamento que antes seria necessário um computador do tamanho de uma sala inteira [Camposano and Wilberg 1996].

Consequentemente, este avanço fez com que a busca por soluções antes inimagináveis comesçassem a surgir. Nesta esteira, podem-se notar os sistemas embarcados presentes em diversas áreas como eletrônica automotiva, aviação, telecomunicação, equipamentos médicos, segurança, fabricação, prédios, casas e cidades inteligentes, logística, robótica, aplicações militares, entre outros [Marwedel 2011].

Apesar desta grande variedade de problemas solucionados com este tipo de sistemas, eles, geralmente, possuem algumas características em comum [Marwedel 2011]:

1. A presença de atuadores ou sensores (para realizar alguma tarefa no ambiente onde está sendo implantado, seja de monitoramento ou atuação);
2. Dependência (que engloba confiabilidade, manutenção, disponibilidade, preservação

e segurança, ou seja, verificar qual a probabilidade do sistema não falhar, o tempo necessário para reparar falhas, a probabilidade dele está disponível, garantir que não ocorra danos caso o sistema danifique e qual a segurança garantida em relação ao acesso a este sistema);

3. Eficiência (energética, tamanho do código, tempo-real, peso e custo);
4. Dedicados a um problema específico;
5. Interface de usuário dedicada;
6. Restrições de tempo-real;

Apesar do crescente cenário no que diz respeito à construção de sistemas embarcados, o projeto deste tipo de sistema computacional é muito mais complexo do que projetos realizados para computadores de propósito geral, pois, diversas restrições devem ser levadas em consideração tais como limite no consumo de potência, portabilidade, baixa disponibilidade de memória, necessidade de segurança e confiabilidade e tempo de resposta restrito [Carro and Wagner 2003]. Além disso, neste sistema o software não pode ser desacoplado do hardware, eles estão intrinsecamente conectados e a solução dada deve ser visualizada observando ambos os níveis [Marwedel 2011].

Com vistas a alcançar diferentes soluções para diferentes problemas, foi surgindo com o tempo uma grande quantidade de dispositivos capazes de implementar sistemas embarcados, os dispositivos microprocessados, microcontrolados, ASICs e FPGAs são alguns exemplos. Desta forma, é papel do desenvolvedor entender cada uma das arquiteturas e escolher aquela que melhor se adéque ao problema a ser solucionado. As subseções seguintes apresentam uma visão geral de cada uma dessas quatro arquiteturas.

B.1 Sistemas Embarcados Microprocessados

A grande vantagem em utilizar sistemas microprocessados está em sua flexibilidade, visto que seus comportamentos podem ser alterados apenas modificando o software que está sendo executado neste sistema (semelhantes a computadores de propósito geral, que também possui como base um sistema microprocessado) [Marwedel 2011]. Em contrapartida, por ter

um ambiente que pode implementar qualquer programa (desde que siga as restrições de memória), os sistemas embarcados microprocessados utilizam mais potência que as demais arquiteturas [Carro and Wagner 2003].

Exemplos de plataformas microprocessadas são: Raspberry Pi ®, Asus Tinker ®, Orange Pi ® e HiKey 960 ®. Ao contrário de microcontroladores, os microprocessadores não podem ser utilizados sozinhos na implementação de sistemas embarcados, visto que eles possuem um propósito diferente (apenas executar a programação), por isso, faz-se necessário todo um sistema com memória e periféricos externos para a construção de sistemas embarcados com esta arquitetura.

B.2 Sistemas Embarcados Microcontrolados

Um microcontrolador, base dos sistemas embarcados microcontrolados, é um dispositivo eletrônico programável que foi desenvolvido para, como o próprio nome sugere, controlar periféricos conectados ao mesmo, sejam estes sensores ou atuadores. Uma grande vantagem dos sistemas microcontrolados em relação às demais arquiteturas é o seu baixo custo. No entanto, este tipo de sistema pode ser utilizado apenas para sistemas que não necessitem de alto poder computacional (o que é suficiente para a maioria dos sistemas embarcados) [Wolf 2008].

Exemplos de plataformas que possuem microcontroladores são: o Arduino® UNO, uma plataforma eletrônica com um microcontrolador de 8 bits da Atmega®; e o Galileo®, com microcontrolador de 32 bits da Intel®. Além destas plataformas, que visam facilitar a implementação de sistemas embarcados, os próprios microcontroladores puros podem ser utilizados para construção deles, reduzindo a quantidade de energia utilizada pelo sistema (evita a utilização de módulos desnecessários encontrados nas plataformas já prontas).

B.3 ASICs e FPGAs

Para a construção de aplicações muito específicas, na qual sabe-se que a possibilidade de alteração é mínima, a utilização de ASICs (*Application Specific Integrated Circuits*, ou Circuito Integrado de Aplicação Específica) é recomendada [Carro and Wagner 2003], pois este

tipo de dispositivo permite uma criação em larga escala mais fácil e o custo de produção se reduz quando comparado às demais alternativas, além de que, sendo de propósito específico, a redução na energia utilizada (restrição de muitos sistemas embarcados) pode ser reduzida no desenvolvimento do projeto.

Enquanto isso, o FPGA é uma alternativa interessante para os desenvolvedores de sistemas embarcados digitais devido sua portabilidade de código através de vários tipos de FPGA dos mais diferentes vendedores (Xilinx®, Altera®, etc), bibliotecas de códigos reutilizáveis e disponibilidade de ferramentas de programação de baixo custo [Dubey 2008].

As vantagens do FPGA em relação às demais plataformas para criação de sistemas embarcados está relacionada a sua reconfigurabilidade, paralelismo inerente e a capacidade de 'desenhar' hardware através de software (com linguagem de descrição de hardware, como Verilog e VHDL), alta velocidade, confiabilidade, proteção sobre propriedade intelectual e reúso [Dubey 2008]. Além disso, os mais diversos fabricantes de FPGA disponibilizam placas de desenvolvimento que aceleram o processo de criação e/ou prototipação de sistemas embarcados.

Apêndice C

Aprendizagem de Máquina

A aprendizagem de máquina é uma subárea da Inteligência Artificial (IA), cujo principal objetivo é fazer técnicas computacionais que sejam capazes de aprender, ou seja, adquirir conhecimento de forma automática [Monard and Baranauskas 2003].

Grande parte da IA é baseada em técnicas adquiridas a partir de observações do comportamento da realidade, como por exemplo algoritmos genéticos, bioinspirados e redes neurais artificiais. Com a área de Aprendizagem de Máquina não é diferente, sua teoria é baseada no comportamento humano, por exemplo, a forma como as máquinas aprendem é semelhante à forma como os humanos fazem esta tarefa, ou seja, por associação ou por exemplo. E é a partir daí que surgem os termos de **Aprendizagem Supervisionada** e **Aprendizagem Não Supervisionada** [Haykin et al. 2009].

C.1 Aprendizagem Supervisionada

Quando se assume que é conhecido previamente alguma informação a respeito do dado, ou seja, se tem alguma informação *a priori* a respeito de qual classe o dado pertence (ou qual padrão ele possui), pode-se utilizar algoritmos de **Aprendizagem supervisionada** [Theodoridis and Koutroumbas 2008]. Isso significa que, como já é conhecido a informação a respeito do dado, então, pode-se *treinar* o algoritmo com base nesta informação para que ele aprenda a partir de *exemplos* e consiga, portanto, generalizar para poder classificar outros dados *semelhantes*.

Vale ressaltar que, geralmente, os dados recebidos por um algoritmo de aprendizagem

supervisionada já possui associado um *vetor de características* (ou atributos), além do *rótulo* da classe associada (que seria a informação conhecida e utilizada para o treinamento do algoritmo) [Monard and Baranauskas 2003]. Redes neurais artificiais (RNA), máquina de valor de suporte (ou SVM, *Support Vector Machine*) e redes Bayesianas são alguns exemplos deste tipo de algoritmo.

Na Figura C.1 é apresentado um modelo em diagrama de blocos (adaptado de [Haykin et al. 2009]) de aprendizagem supervisionada. Nessa figura o vetor de descrição é retirado do ambiente e passado para o sistema de aprendizagem e o "professor"(aquele que sabe o resultado que deve ser obtido com aquela entrada vinda do ambiente), o valor adquirido do sistema de aprendizagem é "comparado" com a resposta desejada e o sinal de erro, obtido da comparação, é então utilizado para ajustar o sistema de aprendizagem.

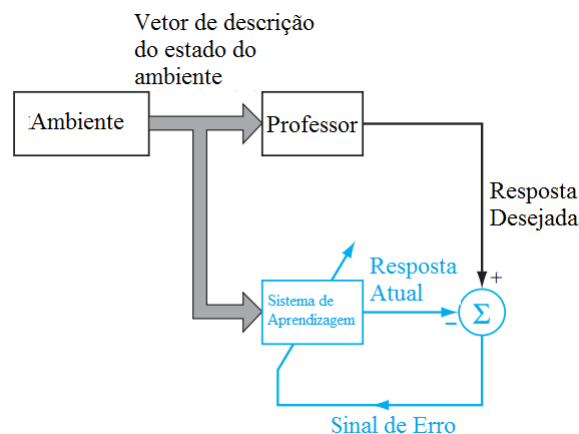


Figura C.1: Modelo em Diagrama de Blocos de Aprendizagem Supervisionada (Adaptado de [Haykin et al. 2009])

Como forma de compreender melhor os algoritmos a serem utilizados na dissertação, nas subseções a seguir são apresentados os algoritmos Perceptron Multicamadas (MLP), SVM, Árvore de Decisão e Random Forest.

C.1.1 Perceptron Multicamadas (MLP)

Segundo [Haykin et al. 2009], o trabalho em redes neurais artificiais (ou, como é comumente referido, redes neurais), tem sido motivado pelo fato de que o cérebro humano

consegue ter uma complexidade alta, com um processamento extremamente paralelo conseguindo aprender coisas novas a todo momento, logo, os cientistas da computação tentam assemelhar essas características nos computadores digitais.

A capacidade de aprendizagem do cérebro humano, segundo [Haykin et al. 2009] se dá pelo fato das células cerebrais responsáveis pelo aprendizado (os neurônios) possuírem duas características fundamentais: plasticidade e adaptabilidade. Desta forma conseguem moldar-se, adquirindo novas informações e armazenando-as em suas estruturas.

Os neurônios são células presentes em todo o corpo humano, formado por uma estrutura como a apresentada na Figura C.2. Como explicado em [Silva et al. 2010], o corpo do neurônio realiza o processamento das informações providas dos dendritos. Processada essa informação, o resultado passará por uma espécie de função que indicará se o neurônio disparará ou não um potencial de ação através do seu axônio.

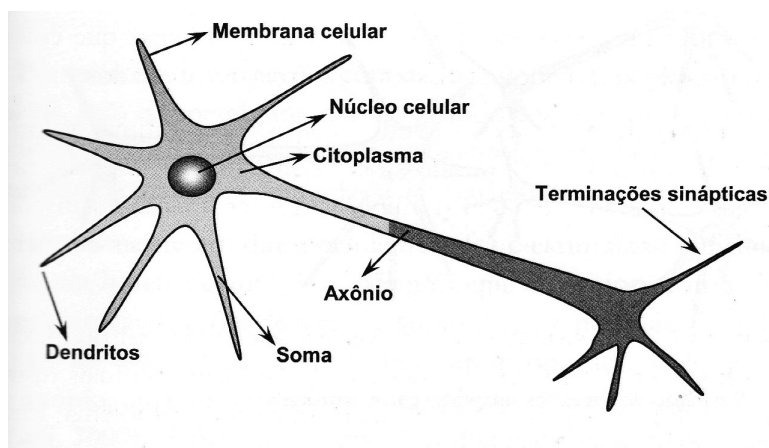


Figura C.2: Representação Neurônio Biológico (Fonte: [Silva et al. 2010])

Em [Haykin et al. 2009] resume-se a ideia acerca da rede neural artificial e sua semelhança com o modo de trabalho do cérebro humano da seguinte maneira:

Uma rede neural é um processador distribuído massivamente paralelo feito a partir de unidades de processamentos simples que têm uma propensão natural para armazenar conhecimento experimental e fazê-lo disponível para uso. Ela se assemelha ao cérebro em dois aspectos: o conhecimento é adquirido através do processo de aprendizagem e as interconexões entre neurônios fortes, conhecidas como pesos sinápticos, são usadas para armazenar o conhecimento adquirido.

Segundo [Silva et al. 2010], Perceptron é a configuração mais simples que pode ter uma rede neural, o que pode ser uma vantagem e uma desvantagem simultaneamente. A desvantagem está em sua limitação a realização de trabalhos com dados que podem ser separáveis linearmente, devido a esta limitação, em 1969, com o lançamento do livro Perceptron por Minsky e Papert, a área de redes neurais começou a decair.

Somente no final da década de 80 houve uma retomada significativa nos estudos e pesquisas de redes neurais, principalmente a partir da publicação do livro de Rumelhart, Hinton e William (*Parallel distributed processing*), que trouxe a tona o algoritmo *backpropagation*, possibilitando, a partir de então, adicionar um número maior de camadas e dando, portanto, uma solução para o problema encontrado em 1969 por Minsky e Papert [Silva et al. 2010].

Desta forma, o avanço nas tecnologias e com o algoritmo de atualização de pesos sinápticos das redes neurais *backpropagation* foi construída a rede perceptron multicamadas, apresentada topologicamente na Figura C.3.

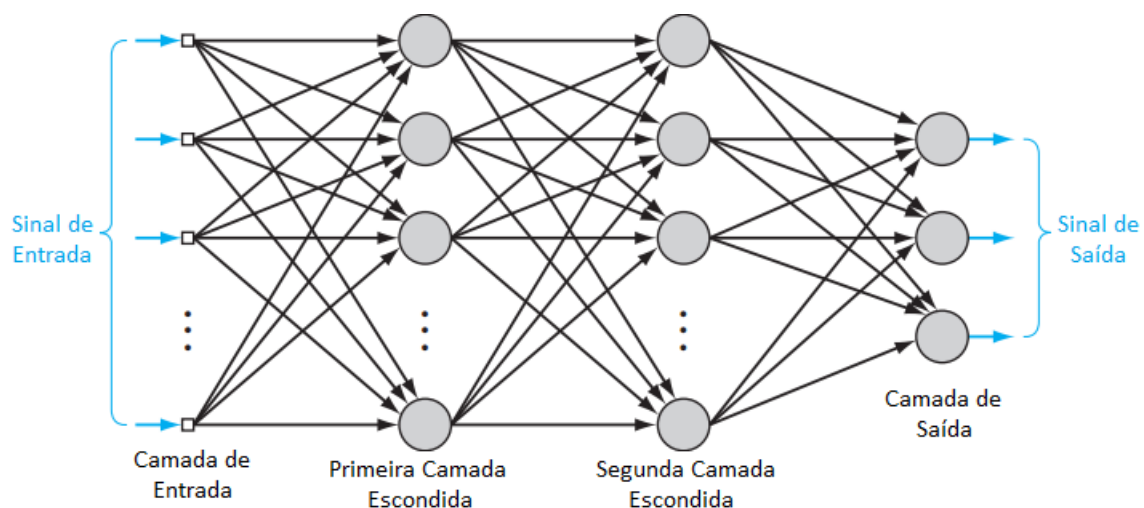


Figura C.3: Topologia da Rede Perceptron Multicamadas (Adaptação de: [Haykin et al. 2009])

Como pode ser visto na Figura C.3, a rede perceptron multicamadas possui vários estágios (camadas) de aprendizagem. No entanto, de [Silva et al. 2010], sabe-se que para atualizar os pesos da rede utiliza-se um conjunto de treinamento com o valor de saída da rede desejado já conhecido, a partir de tais valores, o peso é atualizado para se conformar

aos dados de entrada. O problema acontece que nas camadas internas da rede não existe valor desejado para suas saídas, é neste momento que utiliza-se o algoritmo de *backpropagation*, propagando assim o erro encontrado na camada de saída (que possui valor desejado conhecido) para as camadas mais internas da rede.

Segundo [Silva et al. 2010], o algoritmo de treinamento de uma rede perceptron multicamadas pode ser resumido com os passos de *forward* e *backward*, em que nos passos *forward* calcula-se os valores das saídas dos neurônios de acordo com o dado de entrada, os pesos sinápticos da rede e sua função de ativação; e, no passo *backward* é calculado inicialmente o valor de atualização do peso para a camada de saída a partir do valor desejado, como na rede perceptron original, utilizando a regra delta [Haykin et al. 2009]; a partir do valor de erro na camada de saída, este valor é passado para a camada interna anterior e calcula-se, utilizando a Equação C.1.

$$\delta_j^{(n)} = \left(\sum_{k=1}^{N_{(n+1)}} \delta_k^{(n+1)} \cdot W_{kj}^{(n+1)} \cdot g'(I_j^{(n)}) \right) \quad (\text{C.1})$$

Em que o $\delta_j^{(n)}$ representa o erro no neurônio j da camada n , $g'(\cdot)$ é a derivada da função de ativação utilizada, $I_j^{(n)}$ é o valor de saída do neurônio j da camada n após passar pela função de ativação e $W_{kj}^{(n+1)}$ é a matriz de pesos sinápticos que saem do neurônio j e vão para os neurônios da camada posterior $n + 1$. Esta regra pode ser utilizada em qualquer uma das camadas internas na rede perceptron multicamadas, no entanto, deve-se fazer uma ressalva para o fato de que na primeira camada escondida em vez de se utilizar os valores de $I_j^{(1)}$ utiliza-se o valor do sinal de entrada.

Os valores de $I_j^{(i)}$, em que (i) representa a camada e j o neurônio, são calculados a partir da fórmula apresentada na Equação C.2.

$$I_j^{(i)} = \sum_{l=0}^{n_{i-1}} W_{jl}^{(i)} \cdot Y_l^{(i-1)} \quad (\text{C.2})$$

E os valores de $Y_j^{(i)}$ são calculados a partir da Equação C.3.

$$Y_j^{(i)} = g(I_j^{(i)}) \quad (\text{C.3})$$

Na qual a função $g(\cdot)$ representa a função de ativação que deve ser contínua e diferenciável em todo o seu domínio [Silva et al. 2010]. Para o valor de $Y_j^{(0)}$ utiliza-se o valor de

entrada da rede neural.

C.1.2 Máquina de Vetor de Suporte (SVM)

A Máquina de Vetor de Suporte (ou SVM, sigla em inglês), é um algoritmo de aprendizagem supervisionada proposto na década de 90 por Vapnik [Haykin 2001], que, assim como o Perceptron, pode ser utilizada para solucionar problemas de classificação de padrões e regressão linear.

No problema de classificação de padrões, a separação entre as classes de um dado problema pode ser realizada de diversas maneiras, como por exemplo, no caso de dados bidimensionais separáveis linearmente, pode existir diversas retas que fazem a divisão entre as classes (com diferentes coeficientes angulares e lineares).

Nas redes Perceptron, o limiar entre as classes é encontrado aleatoriamente, o primeiro a ser encontrado que satisfaça algum critério (como por exemplo quantidade de dados classificados corretamente) é utilizado como solução [Silva et al. 2010]. No caso do algoritmo de SVM, as classes são separadas de maneira a maximizar a distância entre determinados vetores (chamados de vetores de suporte) e o limiar de classificação, ou seja, no caso de dados separáveis linearmente em duas dimensões, seria a distância entre determinados pontos no espaço e a reta que os separa [Haykin 2001].

No contexto de SVM, a reta que separa as classes é chamada de hiperplano, pois, tal denominação pode ser utilizada em diversos casos, não apenas nos bidimensionais. Sendo assim, o hiperplano pode ser generalizado como representado na Equação C.4 [Haykin 2001].

$$w^T x + b = 0 \quad (\text{C.4})$$

Desta forma, a função discriminante que determina a distância de um ponto \mathbf{x} no espaço até o hiperplano é dada pela Equação C.5.

$$g(x) = w_o^T x + b_o \quad (\text{C.5})$$

Na qual o subscrito o é utilizado para identificar o hiperplano ótimo, ou seja, aquele que maximiza a distância entre o vetor de suporte e o hiperplano.

Sendo assim, o problema de encontrar o hiperplano ótimo torna-se um problema de otimização com restrição, que pode ser resolvido utilizando a técnica matemática de Multiplicadores de Lagrange [Theodoridis and Koutroumbas 2008].

Ao aplicar-se tal técnica, o problema pode ser resolvido como mostra a Equação C.6.

$$\mathbf{w} = \sum_{i=1}^N \alpha_i d_i \mathbf{x}_i \quad (\text{C.6})$$

Sob a restrição:

$$\sum_{i=1}^N \alpha_i d_i = 0 \quad (\text{C.7})$$

Em que o parâmetro α representa os multiplicadores de Lagrange. Desta forma, não pode-se calcular o valor de \mathbf{w} visto que também não se conhece os valores dos multiplicadores que maximizem a função, ficando assim com duas incógnitas na equação. Sendo assim, aplica-se o teorema da dualidade, de forma a encontrar uma função que encontre a solução ótima para o valor de α sem que seja necessário o valor de \mathbf{w} [Theodoridis and Koutroumbas 2008]. Com a aplicação de tal teorema, proposto por Bertsekas em 1995, o problema torna-se [Haykin 2001]:

Dada a entrada, encontrar os multiplicadores de Lagrange que maximizem:

$$Q(\alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j d_i d_j x_i^T x_j \quad (\text{C.8})$$

Sob as restrições:

1. $\sum_{i=1}^N \alpha_i d_i = 0$
2. $\alpha_i \geq 0$ para $i = 1, 2, \dots, N$

Sendo que, ao encontrar-se os valores de α que satisfaçam o problema, pode-se calcular o valor de w_o a partir da Equação C.9.

$$w_o = \sum_{i=1}^N \alpha_{o,i} d_i x_i \quad (\text{C.9})$$

Esta solução é válida para dados que são linearmente separáveis, quando os dados não o são, utiliza-se a técnica de *kernelização* para a realização da operação de SVM [Theodoridis and Koutroumbas 2008]. Além disso, a solução apresentada é utilizada para separação

de duas classes, quando pretende-se utilizar o SVM para separação de múltiplas classes, segundo [Theodoridis and Koutroumbas 2008], o que faz-se é uma abordagem "um contra todos", ou seja, escolhe-se uma classe para encontrar o hiperplano ótimo e os demais dados como outra classe e encontra assim, utilizando o mesmo resultado aqui apresentado para realizar a separação, e segue o mesmo processo para todas as classes contidas no problema.

C.1.3 Árvore de Decisão (CART)

O algoritmo de Árvore de Decisão consiste de um classificador que utiliza verificações lógicas para realizar a classificação de um conjunto de dados. Segundo [Theodoridis 2008], as questões básicas feitas em uma árvore de decisão simples (como é o caso do algoritmo CART [Breiman et al. 1984] aqui descrito) é se determinada característica do dado de entrada é menor ou igual a dado valor (sendo x_k a característica e α o valor, uma árvore de decisão verifica se $x_k \leq \alpha$).

Desta forma, o algoritmo de impureza, baseado na Teoria da Informação de Shanon, pode ser utilizado como forma de encontrar quais são os valores de α para que a árvore construída seja otimizada (o objetivo é, portanto, encontrar os valores de α que minimizem o valor de impureza).

Sendo as árvores de decisão do tipo CART baseada em perguntas do tipo $x_k \leq \alpha$, a possibilidade é de duas respostas, uma positiva e outra negativa, portanto, a impureza dos dois nós filhos (positivo e negativo), deve minimizar o máximo possível o valor da impureza do nó pai.

Como critério de parada do algoritmo de construção da árvore (na etapa de treinamento do algoritmo), pode-se utilizar diferentes técnicas, como um limiar de impureza, escolhendo assim como classificação final a classe que possuir o número máximo de elementos. Outra técnica de parada é a definição do tamanho da árvore, podendo utilizar uma poda posterior ou parar o treinamento quando a árvore chegar a determinada "altura" pré-definida pelo usuário.

C.1.4 Random Forest

Random Forest é uma técnica de aprendizagem de máquina criada por Breiman em 2001 [Breiman 2001], que consiste de um classificador com diversas árvores de decisão

$h(X, v_k), k = 1, 2, \dots$ sendo v_k vetores aleatórios independentes distribuídos igualmente em todas as árvores.

Neste algoritmo, cada árvore de decisão computa um "voto" em uma classe como "vencedora" e ao final do processo a classificação final é construída, sendo portanto o dado classificado como a classe que possuiu o maior número de "votos" entre as árvores de decisão.

O conjunto de dados de entrada (características) usado no algoritmo é selecionado de forma aleatória (comumente usando a técnica *bootstrap*). Após a seleção dos dados os passos seguintes são a seleção de características dos preditores, a expansão das árvores de decisão até chegar a um critério de parada determinado e, por fim, a estimativa do erro de crescimento da árvore.

C.2 Algoritmos de Pré-Processamento de Dados

C.2.1 Transformada de Wavelet Discreta

Segundo [Theodoridis 2008], a Transformada de Wavelet Discreta (ou DWT) pode ser vista como um banco de filtros (arranjo de filtros passa-faixa e passa-alta que decompõem o sinal de entrada em vários sinais de saída, cada um com uma frequência específica). Desta forma, no algoritmo Wavelet, as saídas são representadas pelos coeficientes, sendo a entrada o sinal que deseja-se decompor. Desta forma, os filtros passa-baixa retornaram o que se chama de coeficientes de aproximação, enquanto os filtros passa-alta retornam o que se denominam detalhes do sinal de entrada.

A DWT apresenta as seguintes características que podem ser úteis na análise de distúrbios elétricos:

1. Filtragem dos dados de entrada;
2. Compactação dos dados;

Dessa forma, pode-se utilizar a wavelet, portanto, para capturar tanto as características de baixa frequência do sinal quanto as características de alta frequência. Além disso, o que faz a DWT ser utilizada para compactação dos dados é que, a depender da quantidade de coeficientes utilizados, a composição do dado final utilizado pode ser reduzida.

Os coeficientes de aproximação e os coeficientes de detalhes podem ser calculados a partir de uma convolução dos sinais com uma função linear para cada um dos casos. Tais funções podem ser representadas no domínio da frequência como os filtros ideais passa-alta e passa-baixa apresentados na Figura C.4.

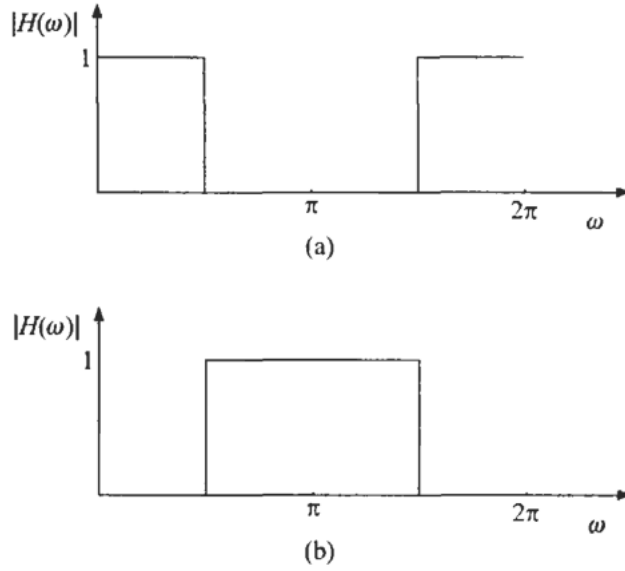


Figura C.4: (a) Filtro Passa Alta Ideal (b) Filtro Passa Baixa Ideal (Fonte: [Theodoridis 2008])

A função dos filtros passa baixa e passa alta podem ser definidas como mostra a Equação C.10 [Theodoridis 2008].

$$y_0(k) = \sum_l x(l)h_0(n-l)|_{n=2k} \quad y_1(k) = \sum_l x(l)h_1(n-l)|_{n=2k} \quad (C.10)$$

Em que h_0 representa o filtro passa baixa e h_1 o filtro passa alta, sendo portanto as funções y_0 e y_1 as funções do filtro, chamadas de filtros de análise, ou ainda os coeficientes do filtro.

Para retornar ao valor original do dado a partir dos coeficientes, ou seja, reconstruir o sinal original a partir da saída dos filtros, utilizam-se os filtros de síntese g_i como mostra a Equação C.11 [Theodoridis 2008].

$$x(n) = \sum_{i=0}^l \sum_k y_i(k)g_i(n-2k) \quad (C.11)$$

Para construir estes filtros, utiliza-se uma função que denominou-se de Wavelet-Mãe. Existe na literatura estudos experimentais que trabalham somente com a escolha da Wavelet-Mãe para determinado tipo de dado, como é o caso do trabalho em [Ukil and Živanović 2006], que adquiriu como resultado experimental a informação de que a db4 (Daubeches 4) é a melhor Wavelet-Mãe para se trabalhar com os dados de distúrbios elétricos. Como visto na Equação C.10, tem-se os filtros de análise que são determinados a partir da função escalar $\phi(h_i(n-l))$ e na Equação C.11 tem-se os filtros de síntese, calculados a partir da função wavelet (que é a Wavelet-Mãe) $\psi(g_i(n-2k))$, assim sendo, a Figura C.5 apresenta uma ilustração da função escalar da db4 e a Figura C.6 ilustra a função wavelet-mãe db4.

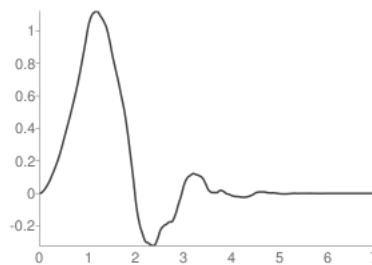


Figura C.5: Função Escalar da db4 (Fonte: [Wasilewski])

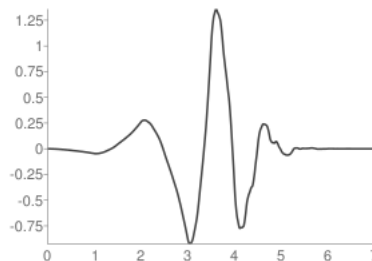


Figura C.6: Função Escalar da db4 (Fonte: [Wasilewski])

C.2.2 Transformada de Fourier Discreta

A Transformada de Fourier, discreta ou contínua, tem o intuito de capturar informações no espectro da frequência do sinal de entrada, representando, por exemplo, a amplitude do sinal em cada frequência dentro do seu espectro ou a fase do sinal [Oppenheim 2010].

A DFT (transformada de Fourier Discreta) é calculada a partir da fórmula apresentada na Equação C.12 [Theodoridis 2008].

$$y(k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x(n) \exp(-i \frac{2\pi}{N} kn), k = 0, 1, \dots, N \quad (C.12)$$

Em que N é o tamanho do sinal de entrada. Esta equação pode ser reduzida ao cálculo de uma multiplicação matricial, fazendo como apresentado na Equação C.13 [Theodoridis 2008].

$$W_N = \exp(-i \frac{2\pi}{N}) \quad (C.13)$$

Sendo a matriz W_N unitária e simétrica, pode-se reduzir o número de cálculos necessários para realização da transformada de Fourier discreta [Theodoridis 2008], que é o que se chama de FFT (*Fast Fourier Transform*), que reduz a quantidade de cálculos de um sinal com N amostras de $2N^2$ (da transformada de Fourier discreta padrão) para $2N \log_2 N$ [Gonzalez and Woods 2010].

C.2.3 Filtro *Notch*

A filtragem, no âmbito do processamento digital de sinais, consiste de um sistema que é utilizado para remover algum componente (amplitude de alguma frequência específica), ou modificar alguma característica, do sinal de entrada [McClellan et al. 2003].

Para construir sistemas de filtragem, entender o espectro de frequência é essencial, pois, os filtros normalmente trabalham na atenuação de determinada faixa de frequência. Por exemplo, filtros do tipo passa-baixa, como o próprio nome sugere, deixa passar o sinal de baixa frequência e atenua os sinais de alta frequência [Oppenheim 2010].

O filtro do tipo *notch* é um filtro seletivo em frequência, segundo a definição dada por [Oppenheim 2010], visto que é um tipo de filtragem no qual uma determinada frequência será atenuada ou, idealmente, eliminada. Como é difícil a construção de filtros que consigam eliminar totalmente determinada frequência, utilizam-se aproximações.

Como forma de implementar digitalmente o filtro *notch* é possível utilizar filtros digitais do tipo IIR de primeira e segunda ordem [McClellan et al. 2003] (forma como será utilizada neste trabalho).

C.2.4 Transformada de Hilbert

A transformada de Hilbert é uma função matemática que transforma um dado sinal no domínio do tempo para um outro ainda no domínio do tempo, mas que, busca enfatizar algumas informações a respeito de localidade no sinal de entrada [Hahn 1996]. Seu cálculo pode ser feito através da Equação C.14.

$$v(t) = h(t) * u(t) \quad (\text{C.14})$$

Em que $u(t)$ é o sinal de entrada, $v(t)$ é o sinal transformado, $*$ representa a convolução entre as funções e $h(t)$ é dado por:

$$h(t) = -\frac{1}{\pi t} \quad (\text{C.15})$$

Dessa forma, pode ser mostrado matematicamente que o resultado da transformada de Hilbert sobre um dado sinal de entrada é dado pelo envelopamento do sinal de entrada [Ramírez-Castro and Montejo 2011].

C.2.5 Filtro Gaussiano

No processamento digital de sinais, os filtros servem para retirar componentes de frequência (alta, baixa ou de valor específico) de um dado sinal de entrada. No caso dos filtros gaussianos (altamente importantes no âmbito da filtragem [Gonzalez and Woods 2010]), utiliza-se como base a função gaussiana.

A fórmula matemática do filtro gaussiano no domínio do tempo é apresentado na Equação C.16.

$$h(x) = \sqrt{2\pi}\sigma A e^{-2\pi^2\sigma^2 x^2} \quad (\text{C.16})$$

Dependendo dos valores de A e de σ o filtro pode ser passa-baixa ou passa-alta. Quanto maior o valor de σ , menor a frequência que o filtro permite passar, ou seja, em um ambiente ruidoso, menos ruído é passado pelo filtro, porém, em contrapartida, muitos detalhes de alta frequência do sinal pode ser perdido.

C.2.6 Estatística Descritiva

Normalmente as medidas de estatística descritiva mais utilizadas são a média e a variância, que são na realidade a esperança de primeira ordem (mostrada na Equação C.17) e segunda ordem (mostrada na Equação C.18).

$$\mu_x = E\{x\} = - \int_{-\infty}^{\infty} x f(x) dx. \quad (C.17)$$

$$\sigma_x = E\{(x - \mu_x)^2\} \quad (C.18)$$

A Esperança é uma medida estatística que pode ser definida discreta (quando os dados são amostrados), ou contínua (quando se têm a função densidade de probabilidade dos dados) [Manly 2008] e é definida matematicamente a partir da função densidade de probabilidade $f(x)$, como pode-se visualizar na definição de média presente na Equação C.17.

Dentro desse contexto, o momento, uma outra medida da estatística descritiva, pode ser definido como o valor esperado de x^j , em que j representa a ordem do momento que deseja-se calcular. A Equação C.19 apresenta como se calcula o momento de j -ésima ordem utilizando esperança estatística.

$$\alpha_j = E\{x^j\} \quad (C.19)$$

Segundo [Mingoti 2005], conhecer os valores dos momentos dos dados equivale, na prática, conhecer a densidade probabilística do sinal.

Existem alguns momentos que são amplamente utilizados na literatura, tanto que receberam nomes especiais, como é o caso da curtose e da obliquidade (*skewness*).

A curtose, que é uma normalização do quarto momento e pode ser definida matematicamente como apresentado na Equação C.20.

$$\alpha_4 = kurt(x) = E\{x^4\} - 3[E\{x^2\}]^2 \quad (C.20)$$

Uma propriedade interessante da curtose é que ela é a medida mais simples para não-gaussianidade de um sinal.

Enquanto isso, o terceiro momento dividido pelo desvio padrão ao cubo (como mostra a Equação C.21), chamado de obliquidade, mede a assimetria da distribuição de frequência do sinal.

$$\gamma = E\left\{\left(\frac{X - \mu}{\sigma}\right)^3\right\} \quad (\text{C.21})$$

Apêndice D

Mapeamento Sistemático da Literatura

Com o objetivo de mapear os estudos acerca dos sistemas embarcados utilizados para monitoramento de distúrbios elétricos com base em questões de pesquisa estabelecidas, foi adotado o método de mapeamento sistemático (SLM, *Systematic Literature Mapping*) apresentado em [Petersen et al. 2008], um estudo secundário no qual, para se alcançar o objetivo é seguido o processo de: definir as questões de pesquisa, fazer a busca por artigos com estudos primários, classificar e extrair informações deles e, por fim, mapear os seus resultados.

Seguindo este método, apresentam-se a seguir as questões de pesquisa que nortearam os estudos, seguido pela estratégia de busca e pela metodologia (critérios) utilizada para a seleção dos artigos de estudos primários.

D.1 Questões de Pesquisa

Com o objetivo de mapear os artigos que apresentam a utilização de sistemas embarcados na solução do problema de detecção e/ou classificação de distúrbios elétricos, verificando os algoritmos que são utilizados e se esta implementação é de tempo real ou não, as seguintes questões de pesquisa foram elaboradas:

- (Q1) Quais os algoritmos utilizados para detecção e/ou classificação dos distúrbios elétricos?
- (Q2) Quais as plataformas de hardware utilizadas na detecção e/ou classificação de distúrbios elétricos?

- (Q3) Os sistemas de monitoramento de distúrbios elétricos trabalham em tempo real ou de maneira offline?
- (Q4) Quais os tipos de distúrbios elétricos mais abordados nos sistemas propostos?

D.2 Estratégia de Busca e Seleção

Para a realização da busca por estudos relevantes foram selecionadas as bases Scopus e *Web of Science*. O portal de periódicos da Capes (<https://www.periodicos.capes.gov.br>) foi utilizado para obter acesso aos artigos disponíveis nessas bases. A escolha da base Scopus deu-se devido ao fato dela possuir em seu acervo os artigos presentes em bases importantes da área de computação e engenharia elétrica (IEEE, ACM, Elsevier, entre outras) [B.V.], já o *Web of Science* foi escolhido devido a falta de alguns artigos desta base na Scopus.

As palavras chaves definidas para busca foram: *electrical disturbance*, *embedded system*, *classification*, *detection*, *monitoring*. Como forma de abranger sinônimos (e com base em alguns artigos encontrados previamente) foram adicionadas as palavras: *FPGA*, *DSP*, *Digital Signal Processor*, *SoC* e *System on Chip* como formas de sinônimos para *embedded system*; e *power quality*, *power fault*, *power disturbance* e *voltage disturbance*, como forma de sinônimos para *electrical disturbance*. Desta forma, as strings de busca foram:

Scopus: TITLE-ABS-KEY (("EMBEDDED SYSTEM"OR "FPGA"OR "DSP"OR "SOC"OR "SYSTEM ON CHIP"OR "DIGITAL SIGNAL PROCESSOR") AND ("ELECTRICAL DISTURBANCE"OR "POWER FAULT"OR "POWER DISTURBANCE"OR "POWER QUALITY"OR "VOLTAGE DISTURBANCE") AND (monitoring OR detection OR classification))

Web of Science: TS=(("EMBEDDED SYSTEM"OR "FPGA"OR "DSP"OR "SOC"OR "SYSTEM ON CHIP"OR "DIGITAL SIGNAL PROCESSOR") AND ("ELECTRICAL DISTURBANCE"OR "POWER FAULT"OR "POWER DISTURBANCE"OR "POWER QUALITY"OR "VOLTAGE DISTURBANCE") AND (monitoring OR detection OR classification))

As pesquisas com estas strings de busca foram realizadas em abril de 2017, retornando 391 artigos na base Scopus e 168 artigos na base Web of Science, totalizando 559 trabalhos. A partir destes artigos foi iniciada a fase de seleção de acordo com os critérios a seguir

apresentados.

D.3 Critérios de Seleção

Com o intuito de filtrar os artigos encontrados, buscando-se apenas aqueles relevantes para o objetivo proposto, foram definidos os seguintes critérios de inclusão:

- Foram incluídos artigos que resolvem o problema (classificação e detecção de distúrbios elétricos) utilizando sistema(s) embarcado(s).
- Foram incluídos artigos cujo conteúdo seja relacionado ao problema da qualidade elétrica relacionada especificamente a distúrbios elétricos.
- Foram incluídos artigos cujas fontes fossem conferências ou Journals.

Os critérios de inclusão foram utilizados baseando-se no resumo (*abstract*) dos artigos encontrados.

Em paralelo aos critérios de inclusão, foram definidos critérios de exclusão, descritos a seguir:

- Foram excluídos artigos duplicados.
- Foram excluídos artigos que não disponibilizavam o texto completo.
- Foram excluídos artigos que não apresentassem uma abordagem voltada para a detecção, classificação ou monitoramento dos distúrbios elétricos.
- Foram excluídos artigos que não apresentassem os algoritmos de detecção/classificação dos distúrbios elétricos.

Com base nos critérios de inclusão e exclusão adotados, dos 559 artigos encontrados inicialmente, foram selecionados 56 para responder às questões de pesquisa apresentadas.

Apêndice E

Mapeamento de Patentes

Com vistas a alcançar o objetivo proposto, mapear as patentes depositadas sobre sistemas de monitoramento da rede elétrica com intuito de verificar a presença de distúrbios elétricos, foi adotado o mesmo método de mapeamento sistemático da Literatura (SLM, *Systematic Literature Mapping*) utilizado na Seção D, o qual define este tipo de estudo como secundário no qual para se alcançar o objetivo (mapear) é seguido um processo pré-determinado [Petersen et al. 2008]. Tal processo consiste de definir as questões de pesquisas; buscar o conteúdo que comporá a base teórica do mapeamento, que em [Petersen et al. 2008] são artigos mas nesta seção foi-se além, buscando patentes; classificar; extrair informações necessárias para responder as questões de pesquisa propostas no primeiro passo; mapear os resultados a partir das informações extraídas.

Frente ao exposto, de maneira a seguir o método, esta Seção traz os passos que conduzem o estudo, ou seja, inicia-se apresentando as questões de pesquisa, seguido da estratégia de busca nas bases de patentes, finalizando com a apresentação dos critérios de inclusão e exclusão das patentes encontradas. Após a apresentação da metodologia adotada, os resultados obtidos (com base nas questões de pesquisa) são listados, de acordo com as patentes incluídas no processo de mapeamento, finalizando, então, com uma conclusão a respeito do estudo realizado.

E.1 Metodologia

E.1.1 Questões de Pesquisa

Objetivando, portanto, a construção de um mapeamento observando as patentes depositadas aquelas que estão relacionadas a sistemas de monitoramento de distúrbios elétricos, as seguintes questões de pesquisa foram elaboradas:

- (Q1) Quais os anos em que foram depositadas as patentes relacionadas a distúrbios elétricos?
- (Q2) Quais os países que mais têm patentes depositadas relacionadas a distúrbios elétricos?
- (Q3) Quais as áreas (de acordo com o CIP - Classificação Internacional de Patentes) que possui o maior número de patentes relacionadas?
- (Q4) Quais as técnicas utilizadas nas patentes relacionadas ao monitoramento de distúrbios elétricos?
- (Q5) Em que plataforma são implementados o sistema de monitoramento das patentes?
- (Q6) Nas patentes são apresentadas técnicas capazes de detectar os distúrbios em tempo real ou de maneira offline?

E.1.2 Estratégia de Busca e Seleção

Para a realização da busca por patentes nacionais e internacionais utilizaram-se duas bases de patentes, a saber INPI (base brasileira) e *Derwent Innovations Index* (base internacional). O portal de periódicos da Capes (<https://www.periodicos.capes.gov.br>) foi utilizado para obter acesso às patentes disponíveis nessas bases.

A escolha da base *Derwent* deu-se devido a esta possuir em seu acervo patentes de diversos países (são disponibilizados patentes registradas em 48 autoridades emissoras de patentes em todo o mundo) [DWPISM and DWPIM 2017], já o INPI foi escolhido para a busca por patentes depositadas no território nacional.

As palavras chaves definidas para as buscas foram: *electrical disturbance*, *classification*, *detection*, *monitoring* para a *Derwent* e apenas "distúrbio elétrico" para o INPI. Como forma de abranger sinônimos (e com base em algumas patentes encontradas previamente nestas bases) foi adicionado na pesquisa pelas patentes no INPI o termo no plural (distúrbios elétricos), já que a pesquisa avançada naquela base só encontra as patentes que possuam a palavra completa em seu título ou resumo. Desta forma, as strings de busca foram:

- Derwent Innovations Index: Tópico: ("electrical disturbance") AND Tópico: ((monitoring OR detection OR classification)).
- INPI: distúrbio elétrico OR distúrbios elétricos.

As pesquisas com estas strings foram realizadas em junho de 2017, retornando 23 patentes na base *Derwent Innovations Index* e 7 na base do INPI, totalizando 30 patentes relacionadas ao tema. A partir destas patentes foi iniciada a fase de seleção de acordo com os critérios a seguir apresentados.

E.1.3 Critérios de Seleção

Com o intuito de filtrar as patentes encontradas, buscando-se apenas aquelas relevantes para o objetivo proposto, foram definidos os seguintes critérios de inclusão:

- Foram incluídas patentes que resolvem o problema (classificação e/ou detecção de distúrbios elétricos).
- Foram incluídas patentes cujo conteúdo seja relacionado ao problema da qualidade elétrica relacionada especificamente a distúrbios elétricos.

Os critérios de inclusão foram utilizados baseando-se no resumo (*abstract*) das patentes encontradas.

Em paralelo aos critérios de inclusão, foram definidos critérios de exclusão, descritos a seguir:

- Foram excluídas patentes que não apresentassem uma abordagem voltada para a detecção, classificação ou monitoramento dos distúrbios elétricos.

Com base nos critérios de inclusão e exclusão adotados, das 30 patentes encontradas inicialmente, foram selecionadas 7 para responder às questões de pesquisa do presente mapeamento sistemático. A Tabela E.1 apresenta o resultado do critério de seleção para cada base de patentes utilizada.

Tabela E.1: Total de patentes selecionadas com critérios aplicados em cada base

Base	Resultado da pesquisa	Selecionados pelo Critério de Seleção
<i>Derwent Innovations Index</i>	23	4
INPI	7	3