



CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

JOÃO MARCUS PEREIRA LIMA E SILVA

**APLICAÇÃO DA TRANSFORMADA WAVELET PACKET E REDES NEURAIAS
ARTIFICIAIS PARA MONITORAMENTO DE CONDIÇÃO DE MOTORES DE
INDUÇÃO**

Recife
2017

João Marcus Pereira Lima e Silva

**APLICAÇÃO DA TRANSFORMADA WAVELET PACKET E REDES NEURAIS
ARTIFICIAIS PARA MONITORAMENTO DE CONDIÇÃO DE MOTORES DE
INDUÇÃO**

Tese submetida ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Pernambuco como parte dos requisitos para obtenção do título de Doutor em Engenharia Elétrica.

Área de Concentração: Processamento de Energia

Orientador: Prof. Dr. Francisco de Assis dos Santos Neves.

Coorientador: Prof. Dr. Luis Antônio Magnata da Fonte.

Recife

2017

Catálogo na fonte
Bibliotecária Valdicéa Alves, CRB-4 / 1260

- S586a Silva. João Marcus Pereira Lima e.
Aplicação da transformada wavelet packet e redes neurais artificiais para monitoramento de condição de motores de indução / João Marcus Pereira Lima e Silva - 2017.
149folhas, Il.; Tabs; Abr.e Sig.
- Orientador: Prof. Dr. Francisco de Assis dos Santos Neves.
Coorientador: Prof. Dr. Luis Antônio Magnata da Fonte.
- Tese (Doutorado) – Universidade Federal de Pernambuco. CTG.
Programa de Pós-Graduação em Engenharia Elétrica, 2017.
Include Referência e Apêndices.
1. Engenharia Elétrica. 2. Wavelets. 3. Redes neurais artificiais. 4. Análise multirresolução. 5. Assinatura de corrente. 6. Motores de indução. I. Neves, Francisco de Assis dos Santos (Orientador). II. Fonte, Luis Antônio Magnata da.(Coorientador). III. Título.
- UFPE
- 621.3 CDD (22. ed.) BCTG/2018 - 182



Universidade Federal de Pernambuco
Pós-Graduação em Engenharia Elétrica

PARECER DA COMISSÃO EXAMINADORA DE DEFESA DE
TESE DE DOUTORADO DE

JOÃO MARCUS PEREIRA LIMA E SILVA

TÍTULO

**“APLICAÇÃO DA TRANSFORMADA WAVELET PACKET E REDES NEURAIS
ARTIFICIAIS PARA MONITORAMENTO DE CONDIÇÃO
DE MOTORES DE INDUÇÃO”**

A comissão examinadora composta pelos professores: FRANCISCO DE ASSIS DOS SANTOS NEVES, DEE/UFPE; LUIZ ANTÔNIO MAGNATA DA FONTE, DEE/UFPE; JOSÉ MAURÍCIO DE BARROS BEZERRA, DEE/UFPE; RONALDO RIBEIRO BARBOSA DE AQUINO, DEE/UFPE; MÁRCIO EVARISTO DA CRUZ BRITO, DEE/UFPE e HÉLIO MAGALHÃES DE OLIVEIRA, DE/UFPE, sob a presidência do primeiro, consideram o candidato **JOÃO MARCUS PEREIRA LIMA E SILVA** **APROVADO.**

Recife, 18 de agosto de 2017.

MARCELO CABRAL CAVALCANTI
Coordenador do PPGEE

FRANCISCO DE ASSIS DOS SANTOS NEVES
Orientador e Membro Titular Interno

LUIZ ANTÔNIO MAGNATA DA FONTE
Coorientador e Membro Titular Externo

JOSÉ MAURÍCIO DE BARROS BEZERRA
Membro Titular Interno

MÁRCIO EVARISTO DA CRUZ BRITO
Membro Titular Externo

RONALDO RIBEIRO BARBOSA DE AQUINO
Membro Titular Interno

HÉLIO MAGALHÃES DE OLIVEIRA
Membro Titular Externo

Dedico esta tese a meu Eterno Professor: Elry Luiz da Silva. Ele foi a pessoa que mais amei na vida; ele era o meu pai...

AGRADECIMENTOS

Em primeiro lugar, agradeço a Deus por ter me dado toda Saúde e Energia para chegar até este momento: a concretização de um Sonho. No dia 27 de Janeiro de 2002 se iniciou uma jornada que, julgava eu, na época, ser temerosa e sem sentido. Hoje, passados quinze anos, vejo que tudo teve seu devido propósito e ocorreu na hora certa! Aqui deixo meus agradecimentos.

A minha mãe e minha família por nunca deixar de acreditar em mim e por dispensar a minha pessoa um amor incondicional, todo serviço e sacrifício... Talvez, um dia, eu consiga compreender algo tão grandioso.

Minha esposa, que sempre me apoiou e deu ânimo nas horas difíceis é testemunha de todas as vitórias e aprendizados vivenciados por mim. A ela deixo minha gratidão sem fim.

Ao Professor Luis Antônio Magnata da Fonte por, mais uma vez, acreditar e confiar no meu trabalho. Deixo aqui, não somente minha gratidão, mas um abraço fraternal e a certeza de que estará sempre na minha mais elevada estima.

Ao Professor Francisco de Assis dos Santos Neves por se mostrar sempre muito solícito e se colocar a disposição para o auxílio nas diversas partes de composição desta tese e em todo Doutorado.

Ao Professor José Maurício Barros Bezerra por toda amizade e apoio logístico, sem os quais nada disso teria se realizado.

Ao Amigo e grande mentor, Professor Alexandre Jorge Tavares de Souza por quem tenho uma enorme admiração e que viabilizou o primeiro contato com o Professor Magnata para que este meu Sonho se realizasse.

Ao Professor Geraldo Leite Torres pela ajuda no início do Mestrado. Ao amigo Herbert Sá Leitão, com quem realizei meu estágio de docência e que me ajudou na implementação em CLP dos ensaios elétricos.

A Andreia Tenório por toda dedicação a nós alunos do PPGEE. A CAPES pelo apoio financeiro.

A cada um que, de alguma forma contribuiu, ao longo destes benditos quinze anos, para a consolidação desta vitória deixo meu agradecimento e a certeza de que, em qualquer momento que precisarem, podem contar com minha amizade.

MUITO OBRIGADO !!!

João Marcus Pereira Lima e Silva

RESUMO

Esta tese apresenta e analisa resultados experimentais da aplicação de uma metodologia baseada na combinação de transformada *wavelet packet*, medidas estatísticas e redes neurais artificiais para o monitoramento de condição de motores de indução trifásicos. Tal abordagem consistiu em coleta das correntes de fase de um motor de indução, seu processamento e construção da árvore *wavelet packet*, escolha apropriada dos sinais de detalhes provenientes desta árvore que serão analisados, estabelecimento das medidas estatísticas dos sinais eleitos e formação dos vetores de entrada apresentados às redes neurais. Estas redes foram treinadas através de dois métodos numéricos: Gradientes Conjugados Escalonados e *Resilient Back-propagation*, para execução da tarefa de reconhecimento de padrões. Cinco configurações de redes neurais *perceptron* multicamada foram utilizadas para realizar a classificação das falhas mecânicas simuladas em laboratório, a saber, fissuras em 2 e 4 barras da gaiola de esquilo do rotor e danificação nas trilhas internas e externas dos rolamentos, e das falhas elétricas consistindo de faltas para terra em pontos do enrolamento de armadura localizados a 25, 50 e 75% do terminal de fase. Os vetores de entrada são constituídos com base em dez medidas estatísticas envolvendo medidas de tendência central (média e centroide), medidas de dispersão (valor RMS e desvio padrão) e uma medida de assimetria (Curtose). A configuração mais promissora foi a rede *perceptron* com quatro camadas ocultas, cada uma contendo dez neurônios. Tal rede mostrou um desempenho satisfatório tanto para as falhas elétricas, com acertos maiores que 92%, como para os defeitos mecânicos, com mais de 93% de sucesso no diagnóstico. Estes resultados superam aqueles registrados em outras investigações do gênero, recentemente divulgadas nos meios acadêmicos, como mostra o confronto apresentado ao final deste trabalho de tese. Tal constatação ratifica o acerto da metodologia escolhida para o diagnóstico de falhas em motores.

Palavras-chave: Wavelets. Redes neurais artificiais. Análise multirresolução. Assinatura de corrente. Motores de indução.

ABSTRACT

This thesis presents and analyzes experimental results of the application of a methodology based on the combination of wavelet packet transform, statistical measures and artificial neural networks for the condition monitoring of three - phase induction motors. Such an approach consisted of collecting the phase currents of an induction motor, processing and constructing the wavelet packet tree, choosing appropriate signals from the tree to be analyzed, establishing the statistical measures of the chosen signals and training of the input vectors presented to the neural networks. These networks were trained through two numerical methods: Scaled Conjugate Gradients and Resilient Back-propagation to perform the pattern recognition task. Five configurations of multilayer neural networks were used to perform the classification of simulated mechanical failures in the laboratory, namely, cracks in 2 and 4 rods of the rotor squirrel cage and damage in the internal and external tracks of the bearings, and of electrical faults consisting of ground faults at armature winding points located at 25, 50 and 75 % of the phase terminal. The input vectors are based on ten statistical measures involving measures of central tendency (mean and centroid), dispersion measures (RMS value and standard deviation) and a measure of asymmetry (Kurtosis). The most promising configuration was the textit preceptron network with four hidden layers, each containing ten neurons. Such a network showed satisfactory performance for both electrical failures, with hits greater than 92%, and for mechanical defects, with more than 93% success in diagnosis. These results surpass those registered in other researches of the genre, recently published in the academic circles, as shown by the confrontation presented at the end of this thesis. This confirms the correctness of the methodology chosen for the diagnosis of engine failures.

Keywords: Wavelet. Artificial neural networks. Multiresolution analysis. Current signature. Induction motors.

LISTA DE ILUSTRAÇÕES

Figura 1 – Plano tempo \times frequência para a STFT.	32
Figura 2 – Resolução Tempo \times frequência da STFT.	33
Figura 3 – Janela gaussiana com centro em τ	34
Figura 4 – caixa de Heisenberg.	34
Figura 5 – Subespaços expandidos através das funções escalas	40
Figura 6 – Complementares dos subespaços.	41
Figura 7 – Par de Filtros de análise.	42
Figura 8 – Banco de filtros de dois canais.	43
Figura 9 – Banco e filtros em quadratura conjugada.	43
Figura 10 – Análise multirresolução do sinal $x[n]$ até o segundo nível de resolução. . . .	44
Figura 11 – Ilustração de uma AMR até o terceiro nível.	45
Figura 12 – Ávore binária básica de uma decomposição <i>wavelet packet</i> até o segundo nível.	46
Figura 13 – Diagrama sub-bandas ou packets de frequência para o ultimo nível (terceiro nível) de decomposição, $j = 3$	47
Figura 14 – Estrutura de um neurônio artificial.	52
Figura 15 – Função degrau.	53
Figura 16 – Função degrau bipolar.	53
Figura 17 – Função rampa simétrica.	54
Figura 18 – Função logística sigmoide.	55
Figura 19 – Função tangente hiperbólica.	55
Figura 20 – Função gaussiana	56
Figura 21 – Função linear	56
Figura 22 – Exemplo de rede <i>feedforward</i> simples.	57
Figura 23 – Rede <i>Perceptron</i> simples	58
Figura 24 – Classes linearmente separáveis. As categorias triângulo e retângulo podem ser separadas através da linha pontilhada.	59
Figura 25 – Rede <i>perceptron</i> multicamada com três camadas (entrada, camada oculta e saída).	60
Figura 26 – Esquemático do <i>setup</i> montado em laboratório para realização dos ensaios com carga no motor de indução.	68
Figura 27 – Motores de indução montados para realização dos ensaios em bancada. . . .	68
Figura 28 – Resistores de 100 Ω alimentados pelo motor de indução 2 funcionando como gerador assíncrono.	69
Figura 29 – Organização dos grupos de amostras em uma matriz S.	70
Figura 30 – Diagrama representativo do processo para construção do sistema classificador de falhas.	70

Figura 31 – Montagem dos ensaios em laboratório.	71
Figura 32 – Faixas das sub-bandas de frequência.	74
Figura 33 – Árvore binária wavelet packet com os nós $W_{i,j}$ identificados também por seus números.	75
Figura 34 – Rotor 1 com suas barras danificadas através de furação	77
Figura 35 – Rotor 2, barras também danificadas através de processo de furação.	77
Figura 36 – Rolamento danificado trilha interna	78
Figura 37 – Rolamento danificado trilha externa	78
Figura 38 – Derivações do enrolamento da fase A	84
Figura 39 – Motor rebobinado com as derivações realizadas	84
Figura 40 – Montagem da Bancada para realização dos ensaios de faltas incipientes.	85
Figura 41 – Detalhe do CLP SIEMENS SIMATIC S7 - 1200.	85
Figura 42 – Esquemático do processo de obtenção e dados e classificação.	89
Figura 43 – <i>Box plots</i> das estatísticas descritivas para as configurações de redes treinadas pelo método SCG para detecção de falhas mecânicas.	95
Figura 44 – <i>Box plots</i> das estatísticas descritivas para as configurações de redes treinadas pelo método Rprop para falhas mecânicas.	96
Figura 45 – <i>Box plot</i> para as cinco configurações treinadas através do SCG para detecção de curto-circuito entre espiras a 25%.	104
Figura 46 – <i>Box plots</i> para as redes treinadas a Rprop para as falhas de 25%.	105
Figura 47 – <i>Box plots</i> para as faltas intermitentes a 50% dos enrolamentos para as redes treinadas a SCG.	106
Figura 48 – <i>Box plots</i> para falhas a 50% dos enrolamentos, redes treinadas a Rprop.	107
Figura 49 – <i>Box plot</i> para as cinco configurações treinadas através do SCG para detecção de curto-circuito entre espiras a 75%.	108
Figura 50 – <i>Box plot</i> para as cinco configurações treinadas através do Rprop para detecção de curto-circuito entre espiras a 75%.	109
Figura 51 – Histograma de erro para o treinamento com os máximos locais do quinto detalhe dos sinais de corrente de fase.	128
Figura 52 – Medida da entropia para determinação do desempenho da rede.	129
Figura 53 – Variação dos valores do gradiente da função erro ao longo do processo de treinamento da RNA.	131
Figura 54 – Matrizes de confusão para o treinamento, teste, validação e Matriz de confusão total para o primeiro caso.	132
Figura 55 – Curvas ROC de treinamento, Validação, Teste e Total para uma rede perceptron classificando sinais utilizando seus máximos locais.	134
Figura 56 – Histograma de erro para o caso com eliminação de elementos nulos do sinal de entrada.	135
Figura 57 – Curva de desempenho através da <i>Cross Entropy</i>	136

Figura 58 – Gradiente para o treino com eliminação dos elementos nulos.	137
Figura 59 – Matrizes de confusão de Treino, Teste, Validação e matriz de confusão Total para o caso da eliminação dos termos nulos dos sinais.	138
Figura 60 – Curvas ROC de treinamento, validação, teste e total para a rede neural com redução do número de nós da entrada.	139
Figura 61 – Histograma de erro para o treinamento com eliminação dos elementos nulos dos sinais para uma rede Perceptron Multicamada.	141
Figura 62 – Medida de erro médio quadrático para determinação do desempenho da rede MLP.	143
Figura 63 – Variação dos valores do gradiente da função erro ao longo do processo de treinamento da rede MLP.	144
Figura 64 – Matrizes de confusão para o treinamento, teste, validação e Matriz de confusão total para a rede MLP.	145
Figura 65 – Curvas ROC de treinamento para a rede MLP.	146

LISTA DE TABELAS

Tabela 1 – Levantamentos percentuais das ocorrências de falhas em motores de indução baseado em (EBRAHIMI et al., 2012), (MCCOY et al., 1985) e (THOMSON; FENGER, 2001).	20
Tabela 2 – Revisão de publicações sobre técnicas em uso na atualidade para monitoramento de condição de motores de indução.	25
Tabela 3 – Dados de placa dos motores utilizados nos ensaios	69
Tabela 4 – Identificação defeito - código	76
Tabela 5 – Relação dos nós com maiores valores de energia e suas respectivas entropias para cada um dos defeitos mecânicos ensaiados.	79
Tabela 6 – Relação entre valores de entropia e suas respectivas energias para cada um dos defeitos mecânicos ensaiados.	81
Tabela 7 – Combinações de medidas estatísticas.	83
Tabela 8 – Padrões adotados para distribuição dos neurônios nas camadas ocultas da rede MLP.	83
Tabela 9 – Relação dos nós com maiores valores de energia e suas respectivas entropias para faltas incipientes no estator.	86
Tabela 10 – Relação entre valores de entropia e suas respectivas energias para cada uma das faltas ensaiadas no estator.	87
Tabela 11 – Padrões adotados para distribuição dos neurônios na rede MLP e seus respectivos acertos globais para a combinação 1 e rede treinada a SCG.	90
Tabela 12 – Taxas de acertos globais (TAG) da rede MLP treinada através do <i>Resilent Backpropagation</i> e combinação de medidas estatísticas 1.	90
Tabela 13 – Comparativo de tempos de treinamento e número de iterações de uma mesma configuração da RNA para Rprop e SCG.	91
Tabela 14 – Taxas de acertos globais da rede MLP para combinação de medidas estatísticas 2.	91
Tabela 15 – Taxas de acertos globais da rede MLP para combinação de medidas estatísticas 3.	91
Tabela 16 – Taxas de acertos globais da rede MLP para combinação de medidas estatísticas 4.	92
Tabela 17 – Taxas de acertos globais da rede MLP para combinação de medidas estatísticas 5.	92
Tabela 18 – Taxas de acertos globais da rede MLP para combinação de medidas estatísticas 6.	92
Tabela 19 – Taxas de acertos globais da rede MLP para combinação de medidas estatísticas 7.	93

Tabela 20 – Taxas de acertos globais da rede MLP para combinação de medidas estatísticas 8.	93
Tabela 21 – Taxas de acertos globais da rede MLP para combinação de medidas estatísticas 9.	93
Tabela 22 – Taxas de acertos globais da rede MLP para combinação de medidas estatísticas 10.	94
Tabela 23 – Taxas de acertos globais da rede MLP para o centroide	94
Tabela 24 – Estatísticas descritivas para as taxas de acertos globais das configurações de redes treinadas a SCG.	95
Tabela 25 – Estatísticas descritivas para as cinco configurações treinadas a Rprop.	96
Tabela 26 – Percentuais de acertos para cada configuração de rede com vetores de entradas formados pelas energias dos nós 2, 4, 8 e 16 da WPT treinadas a Rprop.	97
Tabela 27 – Percentuais de acertos para o treinamento das redes neurais alimentadas com as energias dos nós 2, 4, 8 e 16 da WPT treinadas a SCG.	97
Tabela 28 – Acertos globais para as faltas incipientes de estator para combinação de medidas estatísticas 1.	98
Tabela 29 – Acertos globais para as faltas incipientes de estator para combinação de medidas estatísticas 2.	98
Tabela 30 – Acertos globais para as faltas incipientes de estator para combinação de medidas estatísticas 3.	99
Tabela 31 – Acertos globais para as faltas incipientes de estator para combinação de medidas estatísticas 4.	100
Tabela 32 – Acertos globais para as faltas incipientes de estator para combinação de medidas estatísticas 5.	100
Tabela 33 – Acertos globais para as faltas incipientes de estator para combinação de medidas estatísticas 6.	101
Tabela 34 – Acertos globais para as faltas incipientes de estator para combinação de medidas estatísticas 7.	101
Tabela 35 – Acertos globais para as faltas incipientes de estator para combinação de medidas estatísticas 8.	102
Tabela 36 – Acertos globais para as faltas incipientes de estator para combinação de medidas estatísticas 9.	102
Tabela 37 – Acertos globais para as faltas incipientes de estator para combinação de medidas estatísticas 10.	103
Tabela 38 – Acertos globais para as faltas incipientes de estator para o centroide	103
Tabela 39 – Estatísticas descritivas para as cinco configurações treinadas a SCG para falhas elétricas a 25%	104
Tabela 40 – Estatísticas descritivas para as cinco configurações treinadas a Rprop para falhas elétricas a 25%	105

Tabela 41 – Estatísticas descritivas para as cinco configurações treinadas a SCG para falhas elétricas a 50%	106
Tabela 42 – Estatísticas descritivas para as cinco configurações treinadas a Rprop para falhas elétricas a 50%	107
Tabela 43 – Estatísticas descritivas para as cinco configurações treinadas a SCG para falhas elétricas a 75%	108
Tabela 44 – Estatísticas descritivas para as cinco configurações treinadas a Rprop para falhas elétricas a 75%	109
Tabela 45 – Percentuais de acertos para cada configuração de rede com vetores de entradas formados pelas energias dos nós 2, 4, 8 e 16 da WPT das falhas elétricas treinadas a Rprop.	110
Tabela 46 – Percentuais de acertos para cada configuração de rede com vetores de entradas formados pelas energias dos nós 2, 4, 8 e 16 da WPT das falhas elétricas treinadas a SCG.	110
Tabela 47 – Estatísticas descritivas SCG <i>versus</i> Rprop.	112
Tabela 48 – Tempo médio de treinamento para a metodologia aplicada às falhas mecânicas e elétricas	112
Tabela 49 – Estatísticas descritivas SCG <i>versus</i> Rprop.	112
Tabela 50 – Tabela comparativa de métodos de monitoramento para falhas mecânicas.	113
Tabela 51 – Tabela comparativa de métodos de monitoramento para falhas em estator.	114
Tabela 52 – Redes MLP treinadas a Rprop com uma camada oculta - problemas mecânicos	147
Tabela 53 – Redes MLP treinadas a Rprop com uma camada oculta - faltas a 25%	148
Tabela 54 – Redes MLP treinadas a Rprop com uma camada oculta - faltas a 50%	148
Tabela 55 – Redes MLP treinadas a Rprop com uma camada oculta - faltas a 75%	148

LISTA DE ABREVIATURAS E SIGLAS

$\mathcal{F}\{f(t)\}$	Transformada de Fourier da função $f(t)$.
ω	Frequência de um dado sinal periódico em rad/s.
f_l	Componente de banda lateral à fundamental de corrente
f_o	Componente fundamental de corrente
j	Unidade Imaginária $\sqrt{-1}$.
$coifN$	família Coiflets de wavelets
$meyr$	família de wavelets de Meyer
$symN$	família Symlets de wavelets
CLP	Controlador Lógico Programável
CMB	Manutenção Baseada em Condição (<i>Condition-Based Maintenance</i>)
CWT	Transformada Wavelet Ccontínua (<i>Continuous Wavelet Transform</i>)
db4	Wavelet Daubechies-4
FIR	Resposta Finita a impulsos (<i>Finite Impulse Response</i>)
FPGA	<i>Field Programmable Gate Array</i>
HMM	Modelos Ocultos de Markov (<i>Hidden Markov Models</i>)
HPF	Filtro Passa-altas (<i>Highpass Filter</i>)
LPF	Filtro Passa-baixas (<i>Lowpass Filter</i>)
MCSA	Análise da Assinatura de Corrente (<i>Motor Current Signature Analysis</i>)
MIT	Motores de Indução Trifásicos
PWM	Modulação da Largura de Pulso (<i>Pulse Width Modulation</i>)
QMF	Filtros espelhados em quadratura (<i>Quadrature Mirror Filters</i>)
RBF	Função base radial (<i>Radial Basis Function</i>)
RNA	Redes Neurais Artificiais
ROC	Característica do operador receptor (<i>Receiver Operating Characteristic</i>)

Rprop	<i>Resilient back-propagation</i>
SAGNN	Redes neuronais de crescimento auto-adaptativo (<i>Self Adaptive Growing Neural Network</i>)
STFT	Transformada de Fourier de Curta Duração (<i>Short Time Fourier Transform</i>)
SVM	Máquina de Vetor Suporte (<i>Support Vector Machine</i>)
TR-LDA	Análise linear Discriminante com Razão de Rastreamento (<i>Trace Ratio Linear Discriminant Analysis</i>)
TW	Transformada Wavelet
WNN	Redes Neuronais Wavelet (<i>Wavelet Neural Networks</i>)
WPT	Transformada Wavelet Packet (<i>Wavelet Packet Transform</i>)

SUMÁRIO

1	INTRODUÇÃO	19
1.1	MOTIVAÇÃO	19
1.2	ESTADO DA ARTE	22
1.3	ROTEIRO DA TESE	29
2	ANÁLISE MULTIRRESOLUÇÃO, TRANSFORMADA WAVELET PACKET E MEDIDAS ESTATÍSTICAS DE SINAIS	31
2.1	A TRANSFORMADA WAVELET CONTÍNUA	31
2.1.1	Transformada de Fourier de curta duração	31
2.1.2	O Princípio da incerteza	35
2.1.3	Definição da transformada wavelet contínua	36
2.2	ANÁLISE MULTIRRESOLUÇÃO	37
2.2.1	Transformada wavelet discreta	37
2.2.2	Bases Ortogonais	39
2.2.3	Funções de escalonamento	39
2.3	BANCOS DE FILTROS	41
2.3.1	Filtros espelhados em quadratura	42
2.3.2	Filtros em quadratura conjugada	43
2.4	MULTIRRESOLUÇÃO	44
2.5	TRANSFORMADA WAVELET PACKET	45
2.6	MEDIDAS ESTATÍSTICAS	47
2.6.1	Média	48
2.6.2	Desvio Padrão	48
2.6.3	Curtose	49
2.6.4	Valor RMS	49
2.6.5	Centroide	49
2.7	CONCLUSÃO	50
3	REDES NEURAIS ARTIFICIAIS	51
3.1	INTRODUÇÃO	51
3.2	O NEURÔNIO ARTIFICIAL	51
3.3	ARQUITETURAS DAS RNA E PROCESSOS DE TREINAMENTO	56
3.4	PRINCIPAIS ARQUITETURAS DE RNA	57
3.4.1	Arquitetura <i>feedforward</i> simples	57
3.4.2	Arquitetura <i>feedforward</i> de camadas múltiplas	57
3.5	PROCESSOS DE TREINAMENTO	57

3.5.1	Treinamento supervisionado	57
3.5.2	Treinamento não-supervisionado	58
3.5.3	Treinamento com reforço	58
3.6	REDE PERCEPTRON	58
3.6.1	Princípio de funcionamento do <i>perceptron</i>	59
3.7	PERCEPTRON MULTICAMADAS	60
3.7.1	Princípio de funcionamento do MLP	61
3.8	ALGORITMOS DE TREINAMENTO DAS REDES	61
3.8.1	Gradiente conjugado escalonado	61
3.8.2	O método <i>resilient backpropagation</i>	63
3.9	CONCLUSÃO	66
4	METODOLOGIA	67
4.1	METODOLOGIA DE ANÁLISE	67
4.1.1	Aspectos dos dados experimentais	67
4.1.2	O experimento	67
4.1.3	Wavelet utilizada para decomposição dos sinais	71
4.1.4	Determinação do nível de decomposição	73
4.1.5	Determinação dos nós a serem escolhidos	75
4.2	O PROCESSO DE CLASSIFICAÇÃO DE FALHAS	76
4.3	FALHAS MECÂNICAS	76
4.4	FALHAS ELÉTRICAS	83
4.5	CONCLUSÃO	88
5	RESULTADOS EXPERIMENTAIS	89
5.1	RESULTADOS DE CLASSIFICAÇÃO PARA AS FALHAS MECÂNICAS	89
5.1.1	rede MLP treinada pelo método SCG com a combinação 1 de medidas estatísticas	89
5.1.2	Treinamento da rede MLP através do <i>Resilient Backpropagation</i> da combinação 1	90
5.1.3	Comparativo com treinamento de rede a partir de parâmetros provenientes da decomposição <i>wavelet packet</i> para falhas mecânicas	97
5.2	RESULTADOS EXPERIMENTAIS PARA AS FALHAS ELÉTRICAS	97
5.2.1	Comparativo com treinamento de rede a partir de parâmetros provenientes da decomposição <i>wavelet packet</i> para faltas do estator	110
5.3	CONCLUSÃO	111
6	CONCLUSÕES	113
6.1	CONCLUSÕES GERAIS	113

6.1.1	Comparativos	113
6.2	DIRETIVAS FUTURAS	114
	REFERÊNCIAS	116
	APÊNDICE A – RESULTADOS EXPERIMENTAIS - TÉCNICA INICIAL	126
	APÊNDICE B – TREINAMENTO DE NOVAS CONFIGURAÇÕES DE REDES COM OS ATRIBUTOS SELECIONADOS	147

1 INTRODUÇÃO

1.1 MOTIVAÇÃO

O motor de indução é um componente de vital importância em diversas aplicações na indústria, e geralmente está inserido em equipamentos comerciais mais complexos, como motobombas, compressores de ar, roscas transportadoras, ventiladores, entre outros. Assim, as principais preocupações de projeto do motor de indução envolvem confiabilidade, eficiência, bom desempenho e alto rendimento em variadas situações. Mesmo assim, embora o motor de indução seja robusto, de baixo custo e confiável, este muitas vezes é sujeito falhas graves (WANG; CHEN, 2009), o que pode resultar em tempo de inatividade não programado de sistemas elétricos e perdas econômicas significativas.

Com o advento da eletrônica de potência e a constante evolução dos microprocessadores, a aplicabilidade dos motores trifásicos de indução com gaiola de esquilo se tornou ainda maior no universo industrial. Contudo, motores alimentados através da técnica de modulação da largura de pulso (*Pulse Width Modulation* - PWM) - são mais sensíveis a vários tipos de falta, sem contar os problemas nos próprios equipamentos eletrônicos (LIU; TSE; WU, 2003).

Muitos esforços por parte de equipes de mantenedores têm sido direcionados para o estabelecimento de estratégias de manutenção baseada em condição - (*Condition-Based Maintenance* - CBM) paralelamente aos programas de manutenção preventiva convencionais (JARDINE; LIN; BANJEVIC, 2006) visando evitar a ocorrência falhas inesperadas, diminuir tempo de inatividade, aumentar o intervalo de tempo entre paradas programadas para manutenção padrão, reduzir os custos operacionais devidos à manutenção corretiva e evitar a operação das máquinas elétricas em uma condição insegura. Levantamentos de falha (BELL et al., 1985) (BLÖDT et al., 2008) (LI; SOPON; HE, 2012) e estudos de caso (THORSEN; DALVA, 1997) (SEERA et al., 2014) (ZHANG et al., 2011) relatam que as falhas mais apresentadas em motores de indução na indústria são faltas incipientes no estator, danos físicos aos rolamentos e ao rotor. As ocorrências destas falhas são apresentadas a seguir, na Tabela 1. As diversas falhas mecânicas que podem ocorrer em um motor são consequência de condições de operação ruins como ventilação ineficaz, acionamento inadequado, mal dimensionamento do motor para uma aplicação em questão e fatores ambientais (presença de poeira ou atmosfera agressiva com presença de substâncias químicas prejudiciais aos materiais construtivos do motor) (BENBOUZID; KLIMAN, 2003).

Tabela 1 – Levantamentos percentuais das ocorrências de falhas em motores de indução baseado em (EBRAHIMI et al., 2012), (MCCOY et al., 1985) e (THOMSON; FENGER, 2001).

Tipo de defeito	Local específico de defeito	Percentual de ocorrência
Falhas de rolamentos		41
	Luvas <i>Sleeve bearings</i>	16
	Mancais antifricção	8
	Selos	6
	Mancais esféricos	5
	Vazamentos	3
	Outros	3
Falhas de estator		37
	Isolação para terra	23
	Isolação entre espiras	4
	Amarração	3
	Cunha de isolação	1
	Carcaça	1
	Núcleo	1
	Outros	4
Problemas de rotor		10
	Gaiola	5
	Eixo	2
	Núcleo	1
	Outros	4

Em (THOMSON; RANKIN, 1987) um estudo de caso mostra os danos causados em um motor de indução de alta tensão (11kV) cujas barras da gaiola de esquilo apresentavam fissuras e, devido a alta velocidade de rotação desta máquina, se desprenderam e foram arremessadas em grande velocidade, causando sérios danos, não somente aos núcleos do estator, mas também aos enrolamentos de armadura. De fato, mesmo sendo o rompimento de barras de rotor um problema menos frequente que problemas no estator, por exemplo, tal problema mecânico já poderia ser constatado através do uso da análise de assinatura de corrente, o que representaria um avanço significativo na manutenção preditiva em um caso como este. A combinação de técnicas de análise mais elaboradas permite também que se examinem sinais em períodos transitórios, o que constitui uma vantagem adicional em se tratando do monitoramento de condição. Além disso, a combinação com ferramentas de inteligência artificial possibilita a automação do processo tornando o diagnóstico cada vez mais precoce.

Quatro estudos de caso são citados em (THOMSON; FENGER, 2001). O primeiro caso

retrata o rompimento de barras do rotor de um motor de indução de grande porte (150 HP, 440 V, 110 A, 1485 rpm). Tal motor possuía 58 ranhuras de rotor que acondicionavam as barras da gaiola (barras simples) e realizava o acionamento de um compressor. A análise da assinatura de corrente do motor foi realizada quando o mesmo já apresentava uma série de sintomas que, por sua vez, eram secundários a uma causa original ainda desconhecida. Só quando realizada a análise das correntes é que se pode constatar que o rotor apresentava barras quebradas dada a presença de componentes de banda laterais à componente fundamental, ou seja componentes cuja frequência f_l é função do escorregamento (s) do motor e da fundamental de corrente dadas por (YE; WU, 2001),

$$f_l(1 \pm 2ks)f_o, \quad k = 1, 2, 3, 4... \quad (1)$$

quando realizada a análise de Fourier do sinal.

A análise da assinatura de corrente é uma boa ferramenta para monitoramentos *online* de diversos equipamentos elétricos em virtude da sua relativa facilidade de realização (THOMSON; GILMORE, 2003). Em combinação com as wavelets, que apresentam como uma das vantagens em seu favor a capacidade de observar fenômenos transitórios (LIN et al., 2006) e decompor o sinal em várias escalas tornando possível, não somente a localização de um fenômeno no tempo, como também que se examine sinais totalmente variantes em termos de frequência, tal ferramenta dá ao usuário a possibilidade de coletar um sinal da natureza e visualizá-lo em todas as suas sutilezas.

A transformada wavelet tem mostrado uma ampla aplicabilidade na área de sistemas de potência, desde a proteção de geradores trifásicos (KHAN; OZGONENEL; RAHMAN, 2006), (PARK et al., 2009), (KHAN; OZGONENEL; RAHMAN, 2007) ou o monitoramento de outros tipos máquinas rotativas através da análise de vibração, à proteção dos elementos do sistema de potência como as linhas de transmissão (SILVA et al., 2006) e das redes de distribuição (FANUCCHI, 2014).

As redes neuronais artificiais têm sido de grande valia no que diz respeito à automação do sistema de diagnóstico, torná-lo inteligente (JAWADEKAR et al., 2014), (KOWALSKI; ORLOWSKA-KOWALSKA, 2003), proporcionando uma grande capacidade de mapeamento não-linear e autoaprendizagem, complementando de forma eficiente o reconhecimento de padrões de falha. Esta característica é particularmente interessante para o objetivo deste projeto que é apresentar uma nova metodologia em que se realiza a aplicação da Transformada *wavelet packet* aos sinais de corrente do motor de indução e desta transformada se faça o cálculo de medidas estatísticas. Tais medidas serão apresentadas a um sistema inteligente, neste caso uma rede neuronal artificial, para que se realize um reconhecimento de padrões com o propósito de se

fazer uma multiclassificação entre as várias formas de defeitos (problemas mecânicos e falhas elétricas) a que o motor possa estar sujeito. Estes problemas foram ensaiados em laboratório e todas as correntes de regime transitório e permanente de operação do motor foram registradas.

1.2 ESTADO DA ARTE

A tendência dos métodos de diagnóstico de falhas em motores de indução é, cada vez mais, a incorporação de técnicas não-invasivas que possibilitem o monitoramento contínuo das condições de funcionamento do equipamento realizando uma manutenção preditiva (SANTOS, 2013). O uso dos sistemas inteligentes adquiriu destaque nesse campo de pesquisas. Além das redes neuronais artificiais (MARTINS; PIRES; PIRES, 2007) (FILIPPETTI et al., 2000), o uso de lógica *Fuzzy*, algoritmos genéticos ou mesmo a combinação entre técnicas de Inteligência Artificial (IA) têm se mostrado promissoras. A exemplo da combinação de redes neuronais e lógica *Fuzzy*, evidenciada em (YANG et al., 2009).

Pesquisas já realizam a implementação de monitoramento em tempo real de máquinas rotativas. Abordagens através dos vetores de Park em que as correntes das fases são submetidas à transformada clássica de Park (Transformada d-q) e em seguida dispostas no plano complexo (CRUZ; CARDOSO, 2000) (CRUZ; CARDOSO, 2001), formando padrões para reconhecimento via redes neuronais tem apresentado bons resultados; (NEJJARI; BENBOUZID, 2000) registra uma efetividade de 100% e (HAJI; TOLIJAT, 2001) de 97% para falhas mecânicas com o uso dessa técnica. Para as falhas elétricas (faltas entre espiras), a técnica do vetor de Park é aprimorada pelo uso das wavelets (BARENDSE et al., 2009) possibilitando a detecção deste tipo falhas, seja em regime permanente ou transitório de operação. Esta abordagem, contudo, apresenta como desvantagem o fato de não ser multiclassificatória e realizar apenas a distinção entre “sadio” e “defeituoso” do motor. Tal desvantagem pode ser contornada, mas para isso seria necessário o uso de uma rede específica para cada possível condição defeituosa a que o motor possa estar sujeito, fazendo assim uma abordagem dita descentralizada.

Os primeiros trabalhos destinados à identificação precoce de falhas em motores elétricos de indução faziam uso de ferramentas como a transformada rápida de Fourier (*Fast Fourier Transform - FFT*) para realização de análise espectral (CRUZ; CARDOSO, 1999). A base de qualquer método de diagnóstico confiável é uma boa compreensão do comportamento da máquina em estado saudável e sob condições de falha. A assinatura de corrente é uma variável elétrica facilmente medida e muito sensível a variações de sinal provenientes de falhas (BOUZID et al., 2008). Um ponto crucial sobre a análise de assinatura de corrente de motor (MCSA) é que ela está detectando um sinal elétrico que contém componentes atuais que são um subproduto direto de componentes de fluxos rotativos específicos de falhas tais como barras da gaiola de esquilo com fissuras, excentricidade de entreferro e espiras em curto-circuito nos enrolamentos do estator (THOMSON; GILMORE, 2003). Tais métodos consistem na identificação destas componentes harmônicas específicas através de uma análise de Fourier de alta resolução da

corrente de alimentação do motor para. Deve ser apreciado o fato que, como o motor pode funcionar em condições de carga diferentes, estas condições podem impactar diretamente nas correntes e em suas componentes, o que faz a condição de carregamento fator relevante na realização do diagnóstico (THOMSON; FENGER, 2001).

A desvantagem de tal metodologia reside no fato de a transformada rápida de Fourier ser bastante eficaz para sinais periódicos mas não ter tal eficácia para analisar sinais cujas frequências variem no tempo. A Transformada wavelet vem como uma solução para superar tal obstáculo. Em combinação com o processamento estatístico de sinais e inteligência artificial, as possibilidades se expandem e podem promover um percentual de acertos cada vez mais altos.

O uso da transformada wavelet vem também se revelando um forte aliado dos sistemas inteligentes por melhorar seu desempenho, uma vez que se permite a extração de características exclusivas dos sinais que alimentam tais sistemas (ABDESH; KHAN; RAHMAN, 2008). Novas técnicas baseadas em wavelets e redes neuronais têm surgido e se tornado realidade através da implementação por processamento digital de sinais (DSP, do inglês *Digital Signal Processing*).

Algoritmos baseados em potência instantânea da transformada *wavelet packets* de sinais de tensão e corrente também já foram utilizadas por diversos autores (KHAN; OZGONENEL; RAHMAN, 2006) (KHAN; OZGONENEL; RAHMAN, 2007). Essa proposta exhibe grande simplicidade e facilidade de implementação, pois é feita em poucas etapas. Consiste, essencialmente, na aquisição das formas de onda de tensão e corrente no regime permanente da máquina, decomposição e comparação do produto dos sinais decompostos a um limiar pré-estabelecido. A desvantagem dessa abordagem reside na aquisição dos sinais, que, para os motores alimentados através de inversores de frequência, requer uma taxa de amostragem superior à frequência de chaveamento, tipicamente da ordem de 10 kHz. Tal exigência impõe um grande volume de dados a serem decompostos. Uma simplificação para essa metodologia faz uso da MCSA (*Motor Current Signature Analysis*) onde apenas sinais de corrente são necessários à tarefa de monitoramento de condição (MEHALA; DAHIYA, 2007).

Propostas combinando a wavelet contínua de Morlet às redes neuronais artificiais para a extração de características dos sinais (JAWADEKAR et al., 2014) também têm sido utilizadas por realizar uma localização *tempo × frequência* satisfatória. A dificuldade desse tratamento situa-se no esforço computacional demandado pela transformada contínua de wavelet quando comparado com a análise multirresolução. Também uma melhoria na taxa de acertos pode ser obtida pela escolha adequada da função de ativação dos neurônios da rede neuronal.

Para a realização de diagnósticos em máquinas rotativas, da metodologia para processamento dos sinais via wavelet para submissão a uma rede neuronal, uma segunda possibilidade consiste na inserção das wavelets nas redes neuronais como ativadoras dos neurônios. Essas técnicas são denominadas redes neuronais wavelet (*Wavelet Neural Networks - WNN*) (LIN et al., 2006) (WANG; GUO; DUAN, 2013) sendo utilizadas, tanto para o diagnóstico de defeitos e o monitoramento de condição em motores, quanto para a realização do controle dos

mesmos (ZHANG; WANG, 2011) (LIN; SHEN; KUNG, 2005).

No texto de (CHEN; MO, 2004), as duas possibilidades são tratadas com vistas a realizar um comparativo de desempenho, como também uma nova metodologia de treinamento das redes neurais é proposta para redução dos tempos dispendidos. Tal publicação verificou que as WNN's exibem um desempenho um pouco superior quando as redes neurais artificiais são treinadas pela metodologia proposta, porém nenhum comparativo é realizado com a WNN treinada pelos métodos de otimização já consagrados como o método de Newton, Levenberg-Marquadt, gradientes conjugados. Outra desvantagem, que também está mencionada em textos como (JINGBO; XIUQING, 2008), (SU; CHONG,) e (SHANLIN; YUZHE; HUANZHEN, 2009), diz respeito ao uso desta técnica para a análise dos sinais provenientes de vibrações da máquina.

A corrente é, de fato, uma fonte portadora de uma grande quantidade de informação para a análise e o diagnóstico das condições das máquinas rotativas. Segundo (ZHONGMING; BIN, 2000), a MCSA apresenta eficácia comprovada para a detecção de problemas nos enrolamentos do rotor e do estator, excentricidade do rotor e danos às barras da gaiola no rotor. Autores têm usado a decomposição *wavelet packet* devido a facilidade da mesma em lidar com a inerente natureza não-estacionária das correntes do estator. A partir dessa decomposição a extração dos coeficientes característicos poderá ser realizada, informando à rede neuronal sobre as sutilezas presentes no sinal analisado (YE; WU; SADEGHIAN, 2003).

Em (YE; WU, 2001) é apresentado um algoritmo de diagnóstico e monitoramento *online* para o motor de indução combinando *Wavelet Packet Decomposition* (WPD) e RNA. Um conjunto de coeficientes característicos, calculados através da energia do sinal proveniente de nós específicos, é construído e apresentado à rede juntamente com o escorregamento do motor. Trabalhos descrevendo propostas similares são apresentados em (GHASEMI, 2010) e (WIDODO; YANG, 2007), que utiliza máquinas de vetor suporte (*Support Vector Machine - SVM*) para realizar a distinção entre as condições de operação de motores de indução, e em (SADEGHIAN; YE; WU, 2009) para a detecção de barras com fissuras, (BACCARINI et al., 2011) usa SVM para a multiclassificação de problemas mecânicos em motores de indução a partir de sinais de vibração.

Existe também a aplicação de modelos de Markov (*Hidden Markov Models - HMM*), seja a partir da obtenção de sinais de vibração para a detecção de rolamentos defeituosos (OCAK; LOPARO, 2001), onde um HMM é treinado com características obtidas a partir dos dados de vibração coletados de um motor com rolamento normal e os dados de outros motores são apresentados ao HMM previamente treinado para ser comparado a um limiar pré-determinado e assim se realizar a classificação, ou pela extração do expoente de Lipschitz (MIAO; HUANG; FAN, 2007b) (usando a transformada wavelet para obter seus máximos locais e, conseqüentemente, o expoente) como parâmetro a ser apresentado ao HMM (MIAO; MAKIS, 2007). Além disso, há a comparação entre aplicações de SVM e HMM ambas associadas à

transformada wavelet e ao expoente de Lipschitz (MIAO; HUANG; FAN, 2007a).

Ainda sobre uso de wavelets, (CHOW; HAI, 2004) faz uso de redes neuronais recorrentes em combinação com as wavelets para detecção de faltas de origem mecânica e elétrica; (ZAREI; POSHTAN, 2007) faz uso das *wavelet packets* para detectar modulações nas correntes dos motores de indução, uma vez que estas vibrações mecânicas produzem anomalias na densidade de fluxo do entreferro, eles resultam em modulações de corrente do estator; (KANG; KIM, 2014) usa wavelets bidimensionais para analisar sinais de vibração para a detecção de problemas mecânicos em motores de indução, formando imagens em tons de cinza a partir do uso das wavelets de Shannon para serem apresentadas SVMs especialistas (uma para cada problema analisado).

A Tabela 2 apresenta resumidamente algumas outras publicações com diferentes técnicas de monitoramento de condição de motores de indução, suas vantagens e desvantagens. o objetivo é exemplificar a variedade de formas de monitoramento de condição existentes e como estas ainda podem ser melhoradas.

Tabela 2 – Revisão de publicações sobre técnicas em uso na atualidade para monitoramento de condição de motores de indução.

Publicação	Sumário da publicação	Vantagens	Desvantagens
(NIU et al., 2008)	Realiza a análise de assinaturas de correntes transitórias de motores de indução através de Análise Multirresolução (MRA) em combinação com várias formas de classificadores: <i>Support Vector Machine</i> (SVM), <i>Linear Discriminant Analysis</i> (LDA), rede de Kohonen a <i>Adaptive Resonance Theory</i> (ART-KNN) e k-vizinhos mais próximos para detecção de falhas de estator.	A combinação de quatro técnicas de inteligência artificial garante uma metodologia multiclassificatória com altas taxas de acertos (cerca de 98%).	O método realiza uma MRA do sinal até o quarto nível, e de cada um dos detalhes extrai um total de 21 atributos estatísticos que são apresentados a quatro sistemas classificadores para se tomar uma decisão através de métodos Bayesianos. Apesar de altas taxas de acertos, tal metodologia se mostra de uma complexidade elevada.
(BARAKAT et al., 2011)	Este artigo usa a variância e a curtose dos sub-sinais de uma decomposição <i>wavelet packet</i> para alimentar um classificador baseado em <i>Self Adaptive Growing Neural Network</i> (SAGNN), mostrando inclusive a relevância da combinação de medidas estatísticas e wavelets, para realizar o monitoramento de condição de motores de indução em um processo industrial.	O classificador otimiza o número de nós na SAGNN e gera sub-espacos ocultos para cada classe de saída. A combinação entre DWT e SAGNN melhora os desempenhos do multiclassificador chegando a ter taxas de acertos de 100%.	Durante o processo de treinamento, a rede tende a se adaptar criando, inclusive, nós adicionais que são associados à classe de falha em questão. Tal habilidade presente na rede demanda altos tempos de treinamento, sendo o tempo mais curto apresentado na publicação 10 minutos.

Tabela 2 Continuação da página anterior

Publicação	Sumário da publicação	Vantagens	Desvantagens
(YAZICI et al., 1997)	Relata um método adaptativo e estatístico para a detecção de falhas de rolamento. Experimentos foram conduzidos em rolamentos defeituosos com fissuras nas trilhas exteriores e danos nos elementos esféricos.	Eficácia de 100% em classificações corretas.	O procedimento de detecção requer um treinamento extensivo para extração de recursos. A detecção das falhas de rolamento usa sinais de vibração e são afetadas pela velocidade da máquina, principalmente quando o rolamento está se deteriorando.
(AYHAN; CHOW; SONG, 2005)	Utiliza a análise discriminante múltipla para detectar barras de rotor quebradas com base nas características de entrada extraídas dos sinais de corrente do estator.	As taxas de precisão reportadas de seis testes diferentes foram entre 73,12% e 99,38%.	Não realiza multiclassificação
(TAPLAK; UZMAY; YILDIRIM, 2006)	A rede neuronal foi aplicada para detectar falhas de rolamento. As características de entrada compreendidas foram amplitude, velocidade e aceleração medidas a partir de sinais de vibração. Duas avaliações baseadas em eixo livre e plena carga foram realizadas.	As taxas de precisão alcançadas foram de 76-83% e 67-74%, respectivamente.	Baixas taxas de acerto para ensaios com rotor a plena carga.
(CRUZ et al., 2017)	Realiza aquisições da corrente de armadura do estator e de vibração mecânica e realiza o processamento prévio a partir de análise multi-resolução e transformada de Fourier de curta duração como pré processamento para um sistema Fuzzy que realiza o diagnóstico a partir de experiência prévia adquirida.	Para os cenários analisados (um total de 45) o sistema conseguiu detectar a falha de desbalanceamento e sua severidade em 100% dos casos.	As fontes de aquisições de dados (vibrações e correntes), apesar de constituir uma boa combinação para alimentar o sistema classificador, tornam-se difíceis de se realizar em conjunto. Apesar de as correntes serem fáceis de adquirir, as vibrações já não desfrutaram da mesma facilidade.

Tabela 2 Continuação da página anterior

Publicação	Sumário da publicação	Vantagens	Desvantagens
(ROMERO-TRONCOSO et al., 2011)	Propõe uma nova metodologia implementável em hardware <i>Fieled Programmable Gate Array</i> (FPGA), que combina análise de entropia de informação com lógica <i>Fuzzy</i> para identificar múltiplas falhas em motores de indução.	A metodologia proposta, além de fazer implementação em <i>hardware</i> , tem eficácia de 100% na classificação de três condições faltosas (Barras com trincas, danos na trilha externa de rolamentos e desbalanceamentos de eixo).	A eficácia do sistema proposto é diretamente relacionada à granularidade (ou a quantidade de partições <i>Fuzzy</i> necessárias para separar cada classificação de elementos) nas definições da função de associação. Isso tem efeitos diretos na implementação de <i>hardware</i> , já que é necessário definir mais funções de associação para cada caso, aumentando a complexidade do sistema de detecção e a dificuldade para sua implementação.
(PÖYHÖNEN; JOVER; HYÖTY-NIEMI, 2003)	É realizada a análise de sinais de vibração através da metodologia <i>Independent Component Analysis</i> (ICA) para a formação de um conjunto de entrada para uma SVM com o objetivo de fazer a detecção de barras com fissuras em rotores de motores de indução.	Faz uso do algoritmo FastICA, pacote para <i>MatLab</i> [®] desenvolvido no Laboratório de Informática e Ciência da Informação da Universidade de Tecnologia de Helsinki, cujo desempenho computacional é de 10 a 100 superior aos métodos convencionais de gradiente para ICA. O uso de ICA mostra-se como uma vantagem na construção do conjunto de dados que é aprestado ao sistema classificador (SVM) gerando um percentual de 100% de acerto.	O sistema de monitoramento de vibrações requer o armazenamento de um grande volume de dados. A vibração é muitas vezes medida com múltiplos sensores montados em diferentes partes da máquina, pois cada máquina emite, em virtude do problema diferentes sinais de vibração. O exame de tais dados demanda tempo e é passível a erros.

Tabela 2 Continuação da página anterior

Publicação	Sumário da publicação	Vantagens	Desvantagens
(JIN et al., 2014)	Neste artigo, são considerados nove parâmetros estatísticos do domínio do tempo (Desvio padrão, pico, skewness, Curtose, Fator de crista, fator de depuração, fator de forma, fator de impacto e valor RMS). Além disso, seis parâmetros sobre a porcentagem de energia correspondente aos coeficientes wavelet também são considerados pela decomposição do sinal de vibração por uma transformada wavelet usando “db4” no nível 5. Este conjunto denso de dados é reduzido através do algoritmo <i>Trace Ratio Linear Discriminant Analysis</i> (TR-LDA) e apresentado a um classificador.	O TR-LDA atua diretamente nas distâncias euclidianas entre pontos de dados de inter e intra-classes. Fazendo com que a compacidade dentro de cada classe seja minimizada, enquanto a separabilidade de diferentes classes é maximizada criando-se, assim, conjuntos de dados linearmente separáveis e facilitando a tarefa de classificação e garantindo taxas de acertos de 100% para detecção de rolamentos defeituosos.	Tal metodologia, além de formar seus conjuntos de dados com base em vibrações, que mesmo sendo fontes de informações completas, mostram-se sensíveis a influências externas provocando erros e são difíceis de se obter exigindo um número considerável de sensores, utiliza uma quantidade muito grande de atributos estatísticos.
(KIM; PARLOS, 2002)	Neste trabalho, um sistema preditor empírico transitório, desenvolvido usando redes neuronais recorrentes dinâmicas é apresentado. Utiliza detecção de falhas baseada em modelo, por ter a capacidade de desacoplar os impactos de distúrbios e outras entradas medidas em assinaturas de correntes de motores com falhas. O preditor de corrente do motor é treinado através de uma ampla gama de motores em operação saudáveis. Os resíduos de corrente do motor resultantes (diferenças entre os valores de treinamento do preditor e valores apresentados em suas entradas) são não estacionários e uma decomposição de <i>wavelet packet</i> é realizada para separar os diferentes harmônicos e para calcular os indicadores de falha.	O sistema desenvolvido de detecção e diagnóstico de falhas possui alta eficácia de detecção, superiores a 93%. É de fácil implementação, o que permite que o sistema desenvolvido seja aplicável a motores de indução com diferentes potências, pois requer um ajuste incremental mínimo.	O uso de abordagens baseadas em modelo resulta capacidades de diagnóstico para sistemas dinâmicos mais altas, em comparação para abordagens baseadas em sinais. No entanto, existe um preço a ser pago por esse benefício por conta da necessidade de um modelo matemático preciso por mais simples que seja. Não só há muito mais esforço computacional necessário para modelar um sistema, mas também há problemas associados à sensibilidade do processo de detecção às incertezas de modelagem inevitáveis na prática.

Tabela 2 Continuação da página anterior

Publicação	Sumário da publicação	Vantagens	Desvantagens
(EBRAHIMI et al., 2012)	Este artigo apresentou um algoritmo abrangente para o diagnóstico não invasivo de falhas em máquinas de indução em que foram introduzidos dois novos índices baseados em coeficientes de wavelet da faixa de frequência inspetiva do sinal de corrente do estator. Esses dois índices permitem diagnosticar a falha e o número de barras de rotor quebradas indiferente das condições de carga	Devido à propriedade da transformada wavelet de se adequar a sinais variantes no tempo, é possível detectar a falha do motor quando este está ainda em aceleração no momento do seu acionamento. Isto é importante para o caso de motores de indução com tempo de arranque longo e também motores que são acionados com muita frequência.	Este trabalho usa como parte de seu algoritmo de diagnóstico uma modelagem do motor de indução. A modelagem é de alta complexidade para que se possa apresentar informações completas para o diagnóstico de falhas com maior sensibilidade.

1.3 ROTEIRO DA TESE

É importante, como passo inicial para realizar o objetivo desta tese, mostrar a aplicabilidade da combinação da transformada *wavelet packet* (WPT - *Wavelet Packet Transform*) com medidas estatísticas como forma de tratamento dos sinais com propósito de obtenção de informações sutis dos mesmos antes de serem aplicados para treinamento e teste das redes neuronais. Todo processamento foi realizado em ambiente *MatLab*[®] usando a *Wavelet Toolbox* e a *Neural Network Toolbox* para a implementação desde a análise dos sinais até a construção da rede neuronal através de linhas de comando. Este trabalho será apresentado em seis capítulos e dois anexos descritos a seguir.

O capítulo 2 apresenta a transformada wavelet, iniciando-se pela conceituação da transformada de tempo curto de Fourier (*Short Time Fourier Transform*), seguida da descrição da transformada de Gabor e alcançando, finalmente, a transformada de wavelet Contínua (CWT). Após tais conceitos, todo processo de discretização da transformada é apresentado juntamente com a análise multirresolução (AMR) e a transformada *wavelet packet*, usada na decomposição dos sinais das correntes de fase. Ainda neste capítulo, são definidas as medidas estatísticas extraídas dos sinais (valor RMS, curtose, média, desvio padrão e centroide) para o treinamento, testes e validação da rede.

O capítulo 3 apresenta as redes neuronais artificiais, seus elementos, arquiteturas e topologias. Neste capítulo é feita uma breve abordagem matemática dos métodos de otimização utilizados para treinamentos das redes usadas neste projeto, a saber: *resilient backpropagation* e *scaled conjugate gradient*. A topologia eleita é a perceptron multicamada (*multilayer perceptron* - MLP).

O Capítulo 4 descreve a metodologia utilizada para a análise dos sinais com a justificativa da wavelet eleita. Neste capítulo também é determinado o nível de decomposição e é descrito o

setup em laboratório para realização dos ensaios de defeitos mecânicos e falhas elétricas

No Capítulo 5 são apresentados os resultados experimentais de classificação do motor para falhas mecânicas: motor com duas e quatro barras quebradas e rolamentos danificados em suas trilhas internas e externas, e para suas falhas elétricas: falhas incipientes nos enrolamentos das três fases do estator cobrindo, 25%, 50% e 75% das bobinas dos mesmos.

O capítulo 6 apresenta as conclusões alcançadas e as propostas de expansão da metodologia para uma gama maior de defeitos que possam ser simulados. Também enumera os argumentos para a implementação em hardware da filosofia de monitoramento.

O Apêndice A descreve uma metodologia alternativa que, embora não tenha se mostrado eficaz para aplicação multiclassificatória, exibiu acertos de até 97% na detecção de barras com fissuras.

No Apêndice B a combinação de atributos estatísticos eleita preferível é aplicada a topologias de rede com apenas uma camada oculta com o propósito de avaliar seu desempenho. Os resultados mostram que mesmo em topologias de redes mais simples tal combinação é satisfatória.

2 ANÁLISE MULTIRRESOLUÇÃO, TRANSFORMADA WAVELET PACKET E MEDIDAS ESTATÍSTICAS DE SINAIS

A transformada wavelet (TW) tem sido amplamente pesquisada e discutida nos últimos anos por suas características possibilitarem a extração de aspectos locais do sinal analisado. Isso significa dizer que é possível realizar uma análise espectral de determinado ponto de um sinal não estacionário, diferentemente de outras técnicas (OLIVEIRA, 2007b). Por possuírem propriedades de localização que podem ser usadas com muita eficiência (neste sentido, eficiência significa que somente um número pequeno de coeficientes é necessário para representar um sinal com muitas informações), as wavelets têm ganhado espaço no universo da análise de sinais (PINTO, 2005).

Esta ferramenta foi introduzida em um trabalho pioneiro realizado pelo engenheiro francês Jean Morlet, onde é apresentado o conceito de “wavelet-mãe”, da qual outras funções podem ser derivadas a partir de modificações de seus parâmetros (FANUCCHI, 2014). O conceito se consolidou nas bases da teoria de wavelets e se tornou fundamental em toda a bibliografia destinada ao estudo desta ferramenta.

A proposta deste capítulo é apresentar os aspectos teóricos a respeito da TW, apresentar a AMR e sua evolução para *wavelet packet* e, por fim, mostrar como os sinais provenientes desta serão utilizados para construir o conjunto de informações a ser usado para o treinamento da Rede Neural Artificial.

2.1 A TRANSFORMADA WAVELET CONTÍNUA

2.1.1 Transformada de Fourier de curta duração

Quando se realiza a análise de um sinal através das wavelets, o mesmo está sendo observado sob uma óptica tempo-frequência. Ou seja, está sendo avaliada a forma como o conteúdo de frequências de um sinal varia com o tempo (DINIZ; SILVA; NETTO, 2014). Para mostrar inicialmente como se processa a análise tempo-frequência, será primeiro apresentada a transformada de Fourier de curta duração (do inglês *Short Time Fourier Transform* - STFT).

Quando se deseja conhecer o conteúdo de um sinal no domínio da frequência levando em consideração sua duração completa $(-\infty, \infty)$, a ferramenta mais popular é a transformada de Fourier

$$F(\omega) = \mathcal{F}\{f(t)\} = \int_{-\infty}^{+\infty} f(t)e^{-j\omega t} dt, \quad (2)$$

em que j é a unidade imaginária ($\sqrt{-1}$) e ω é a frequência de um dado sinal em rad/s.

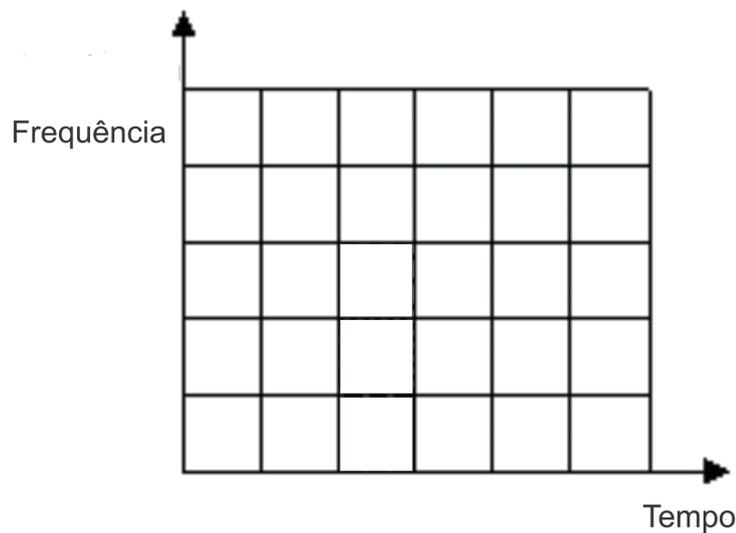
É possível se deparar com situações em que seja necessário calcular o conteúdo frequencial de um sinal apenas numa certa região temporal. Nessa situação a transformada de Fourier como apresentada em (2) não será uma ferramenta apropriada, pois levará em conta a duração completa do sinal. Num esforço para superar esta dificuldade, Dennis Gabor em 1946 fez uma adaptação à Transformada de Fourier de maneira que somente uma pequena parte do sinal no tempo fosse analisada (SANTIAGO; PEDERIVA, 2004) (MISITI et al., 1996).

A técnica em questão ficou conhecida como *Short Time Fourier Transform* e consiste em aplicar uma função janela à transformada de Fourier original:

$$STFT(\omega, \tau) = \int_{-\infty}^{+\infty} f(t)w(t - \tau)e^{-j\omega t} dt, \quad (3)$$

$w(t)$ é a função janela cuja posição é deslocada através da variável τ . Agora, a STFT transformou um sinal no domínio do tempo em uma função bidimensional. Função esta que deve ser representada em um plano tempo \times frequência exemplificado na Figura 1.

Figura 1 – Plano tempo \times frequência para a STFT.



Fonte: Próprio autor.

A STFT apresenta um comportamento dual. Seja $F(\omega)$ a transformada de Fourier de $f(t)$ e $W(\omega)$ a transformada de Fourier de $w(t)$. Pode-se, através do Teorema de Parseval (ROBERTS, 2009):

$$\int_{-\infty}^{+\infty} |f(t)|^2 dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} |F(\omega)|^2 d\omega, \quad (4)$$

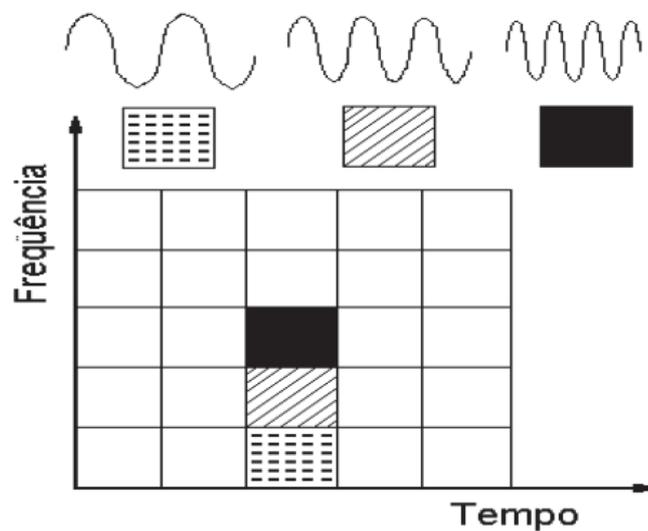
provar que (3) também pode ser escrita como

$$STFT(\omega, \tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega) W(\omega - \omega_o) e^{-j\omega\tau} d\omega. \quad (5)$$

Assim, para cada valor de ω_o , a expressão em (5) informa como evoluem no tempo as componentes de frequências de $f(t)$ em torno de $\omega = \omega_o$ (DINIZ; SILVA; NETTO, 2014).

A STFT apresenta uma limitação séria no que diz respeito à largura da janela, pois seu valor é constante para quaisquer valores de frequência que $f(t)$ possa assumir. Isso é observado na Figura 2.

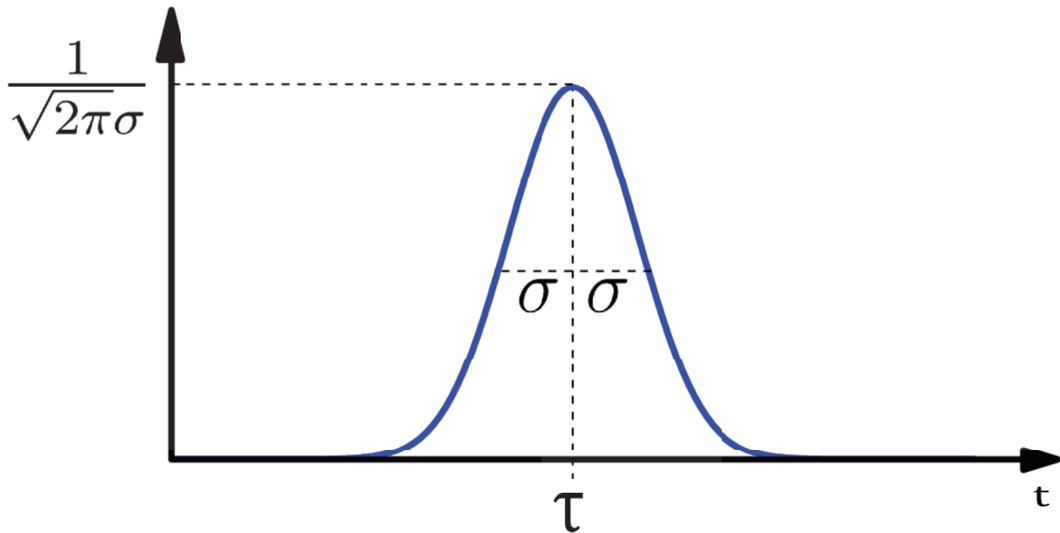
Figura 2 – Resolução Tempo \times frequência da STFT.



Fonte: (SANTIAGO; PEDERIVA, 2004)

Uma janela larga (maior número de amostras) permite uma boa resolução no domínio da frequência, mas resolução pobre no domínio do tempo. Uma faixa estreita fornecerá uma boa resolução no tempo, porém a resolução na frequência pode ser penalizada. É necessário que se estabeleça um compromisso entre as resoluções de tempo e frequência de maneira que escolha-se uma janela que possibilite a obtenção de informações em ambos os domínios. Uma escolha comum para isso é a função gaussiana, mostrada na Figura 3. Boas estimativas podem

Figura 3 – Janela gaussiana com centro em τ .

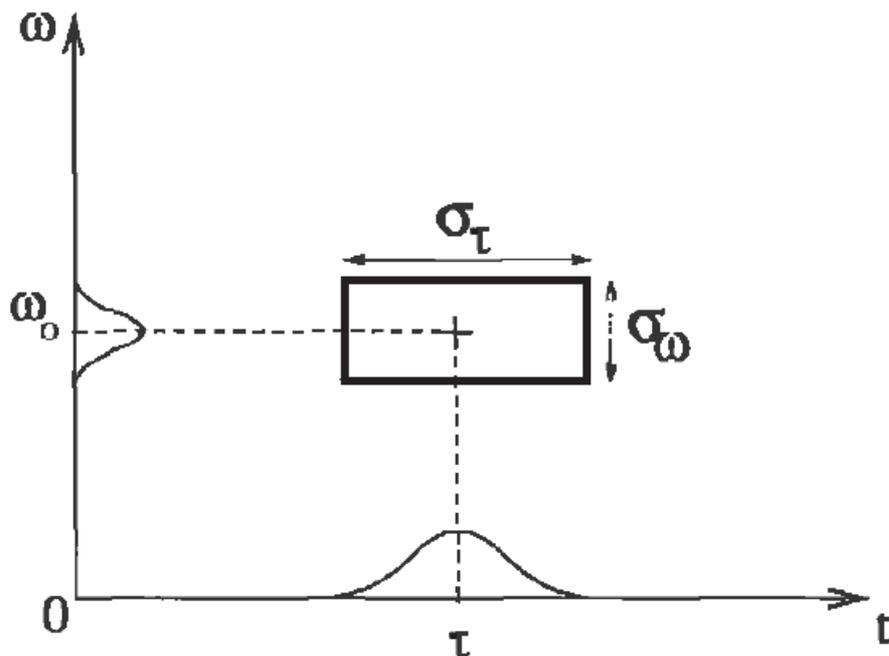


Fonte: Próprio autor

ser feitas para as larguras de $w(t)$ e $W(\omega)$ através da área de sua caixa de Heisenberg (MALLAT, 1998). Uma caixa de Heisenberg é o retângulo formado por estas duas medidas, mostrada na Figura 4.

Uma vez que $w(t)$ é par, $w_{\omega,\tau} = e^{j\omega t}w(t - \tau)$ é centrada em τ . Largura de janela em

Figura 4 – caixa de Heisenberg.



Fonte: (MALLAT, 1998)

torno de τ é independente de τ ou ω_o e é dada por

$$\sigma_\tau^2 = \frac{\int_{-\infty}^{+\infty} t^2 |w(t)|^2 dt}{\int_{-\infty}^{+\infty} |w(t)|^2 dt}. \quad (6)$$

A transformada de Fourier de $w(t)_{\omega,\tau}$ é $W_{\omega,\tau} = W(\omega - \omega_o)e^{-j\omega\tau}$ que, por sua vez é real e simétrica, pois $w(t)$ é real e simétrica (MALLAT, 1998).

Trata-se de uma translação da função com centro em ω_o ao longo do domínio da frequência cuja largura de janela é dada por

$$\sigma_\omega^2 = \frac{\int_{-\infty}^{+\infty} \omega^2 |W(\omega)|^2 d\omega}{\int_{-\infty}^{+\infty} |W(\omega)|^2 d\omega}, \quad (7)$$

σ_τ e σ_ω são os "desvios padrão" da função gaussiana no domínio do tempo e da frequência, respectivamente.

2.1.2 O Princípio da incerteza

O princípio de Gabor-Heisenberg estabelece uma cota inferior sobre o produto $\sigma_\omega\sigma_\tau$ (OLIVEIRA, 2007a)

$$\sigma_\omega\sigma_\tau \geq \frac{1}{2}. \quad (8)$$

Esse resultado implica a existência de uma resolução máxima que pode ser conjuntamente atingida na frequência e no tempo por qualquer sinal contínuo. A igualdade ocorre para sinais gaussianos, ou seja, para funções do tipo $g(t) = Ae^{-kt^2}$. Em alguns casos, as resoluções temporal e de frequências fixas da STFT se tornam uma grave desvantagem. Isso acontece normalmente com sinais altamente não-estacionários, que apresentam eventos de durações muito diferentes, exigindo graus variados de resolução. Em contrapartida, uma vez fixada a função $g(t)$, as resoluções no tempo e na frequência a ela associadas também são fixadas e somente eventos de duração comparável à de $g(t)$ podem ser convenientemente analisados. Portanto, é desejável ter uma transformada com janelas de diferentes durações, de tal forma que se possa adaptar aos eventos a serem analisados. Transformadas wavelet são uma classe de transformadas que possuem exatamente essa propriedade (DINIZ; SILVA; NETTO, 2014).

2.1.3 Definição da transformada wavelet contínua

Assim como a análise de Fourier, a análise através das wavelets inclui duas partes: a transformada wavelet e a expansão em séries. Um dos objetivos da expansão de um sinal é obter seus coeficientes de expansão através de transformadas, o que pode fornecer informações úteis a respeito do sinal analisado. A Transformada wavelet pode decompor sinais em diferentes escalas de maneira flexível.

A transformada wavelet contínua (*Continuous Wavelet Transform - CWT*) para um dado sinal $f(t)$ é definida, para $a \neq 0$, por

$$C_{a,b}\{f(t)\} = \langle f(t), \psi_{a,b}(t) \rangle = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} f(t) \psi^* \left(\frac{t-b}{a} \right) dt, \quad (9)$$

em que $\langle \rangle$ denota um produto interno entre as duas funções contínuas f e $\psi_{a,b}$, e o sobrescrito $*$ denota seu conjugado complexo. Assim como na STFT, o sinal é multiplicado por uma função de análise $\psi \left(\frac{t-b}{a} \right)$, expressa pela notação $\psi_{a,b}$, que denota uma wavelet transladada ao longo do eixo do tempo de uma medida b e escalonada de um fator a (OLIVEIRA, 2013) (LI et al., 2003).

A função $\psi_{a,b}$ não se limita apenas a funções complexas. De fato, desde que esta função atenda as condições (10) a (12), ela pode ser uma função real. Seja $\psi(t)$ uma função pertencente ao espaço das funções mensuráveis reais $\mathcal{L}^2(\mathbb{R})$, ela pode ser admitida como uma wavelet se (HUANG; HUANG, 2002):

- a) a primeira condição diz respeito à wavelet ser oscilante e ter média zero no domínio do tempo, também conhecida como condição de regularidade.

$$\int_{-\infty}^{+\infty} \psi(t) dt = 0 \quad (10)$$

- b) a segunda condição a ser atendida determina que a wavelet deve ter energia finita

$$\int_{-\infty}^{+\infty} |\psi(t)|^2 dt < \infty; \quad (11)$$

- c) a terceira condição é conhecida como condição de admissibilidade. Dado o par Transformada de Fourier $\psi(t) \longleftrightarrow \Psi(\omega)$.

$$\int_{-\infty}^{+\infty} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < \infty; \quad (12)$$

$$|\Psi(\omega)|^2 \Big|_{\omega=0} = 0 \quad (13)$$

Vale ressaltar que, uma vez atendida a condição de admissibilidade, a condição de regularidade e de energia finita também são atendidas fazendo com que a condição de admissibilidade seja condição necessária e suficiente para que uma dada função ψ seja considerada uma wavelet (GOMES, 2015) (DOERING, 2005) (COMBES; GROSSMANN; TCHAMITCHIAN, 2012).

2.2 ANÁLISE MULTIRRESOLUÇÃO

A análise multirresolução (AMR) é uma técnica criada por Yves Meyer e Stephane Mallat (MALLAT, 1989) (MALLAT,) que permite tratar de um sinal através de filtros passa-baixas, LPF - *Lowpass Filter*, e passa-altas, HPF - *Highpass Filter*, dividindo-o em sub-bandas onde cada sub-banda é analisada individualmente (OLIVEIRA, 2013). A realização da AMR permite que se obtenham detalhes contidos no sinal original. O conceito de representação de um sinal em multirresolução permite que se tire conclusões interessantes sobre a transformada wavelet, bem como uma compreensão mais profunda de sua conexão com os bancos de filtros.

Nesta seção serão apresentados, primeiramente, o processo de discretização da transformada wavelet e os bancos de filtros bem como sua aplicação na análise de sinais.

2.2.1 Transformada wavelet discreta

Dada uma família wavelet $\psi_{a,b}$ em que $a \in \mathbb{R}$ e $b \in \mathbb{R}$, com $a \neq 0$, pode-se obter sua Transformada wavelet discreta (DWT - *Discret Wavelet Transform*) fazendo a e b restritos a valores discretos. A discretização do parâmetro de dilatação se mostra natural, uma vez que se pode definir $a = a_o^m$ em que $m \in \mathbb{Z}$, e o passo de dilatação pode ser fixado $a \neq 1$ (ARRUDA, 2003).

A translação no domínio do tempo deve ser proposta de uma forma que, quando for amostrado o sinal $f(t)$, possa ser feita a reconstrução da função a partir de sua DWT com a menor redundância possível. É importante frisar também que, ao se realizar variações na escala, a largura da wavelet muda. Devido a isso, é importante que se "correlacione" as discretizações entre os fatores escala e translação.

Uma propriedade importante das wavelets é a invariância sob mudança de escala. Isso significa que ao mudar de escala, e em consequência mudar a escala do espaço de variação da função, a forma da wavelet não muda (ARRUDA, 2003). Mais precisamente

$$f_{a_o}(t) = a_o^{-\frac{1}{2}} f\left(\frac{t}{a_o}\right) \quad (14)$$

É importante preservar esta propriedade quando da discretização para que esta seja válida

também para a DWT. Uma estratégia comumente adotada para chegar a este objetivo é preservar o fator b como dependente do fator de escala a . Dessa maneira se pode escolher um fator de translação b_o e fixar um intervalo de translação $\Delta b = a_o^m b_o$. Em consequência disso o valor de b será relacionado com o valor de a por

$$b_{m,n} = na_o^m b_o, \quad (15)$$

em que $n \in Z$. Logo, a discretização em translação e em escala ocorrerá e será definida por

$$\Delta_{a_o, b_o} = \{(a_o, na_o^m b_o), m \in Z \text{ e } n \in Z\}. \quad (16)$$

Com tais escolhas dos fatores escala e translação, a função wavelet se torna

$$\psi_{m,n}(t) = a_o^{-\frac{m}{2}} \psi(a_o^{-m} t - nb_o). \quad (17)$$

Com base na expressão descrita em (17), a transformada wavelet discreta (DWT) pode ser expressa por:

$$DWT_{m,n} = C_{m,n} = |a_o^m|^{-\frac{1}{2}} \sum_{k=-\infty}^{+\infty} f(k) \psi\left\{\frac{t - nb_o a_o^m}{a_o^m}\right\}. \quad (18)$$

A CWT essencialmente mapeia um sinal unidimensional em uma representação bidimensional (tempo \times escala) que é altamente redundante. As transformadas discretas de wavelet foram introduzidas no intuito de proporcionar uma descrição mais eficiente (OLIVEIRA, 2007b). Isso ocorre devido ao fato de tanto o escalonamento da função, quanto a translação serem feitos em intervalos discretos.

Sob uma visão prática, a transformada wavelet contínua não seria uma ferramenta de aplicação viável, haja visto que por se tratar de uma transformada com variáveis contínuas iria requerer translações e escalonamentos também contínuos e isso representa um severo esforço computacional. Como já visto anteriormente, a_o e b_o são valores inteiros; o menor valor inteiro

que se pode dar à escala é $a_o = 2$. O menor passo de integração, levando em conta mesma ideia, é $b_o = 1$. Desta maneira, tem-se o que se chama de wavelet diádica; estas, por sua vez constituem bases ortogonais para $\mathcal{L}^2(R)$.

2.2.2 Bases Ortogonais

Pode-se agora construir wavelets que sejam versões transladadas e escalonadas da família

$$\psi_{m,n} = |2^m|^{-\frac{1}{2}} \psi \left\{ \frac{t - n2^m}{2^m} \right\}. \quad (19)$$

A construção destas bases pode, assim, estar relacionada com as aproximações multirresolução de sinal. Seguindo por este caminho chega-se a uma equivalência inesperada entre bases wavelet e bancos de filtros espelhados em quadratura (WEEKS, 2012). Estes bancos de filtros implementam uma transformação wavelet ortogonal rápida que exige apenas $O(n)$ operações para sinais discretos de tamanho N (MALLAT, 1998).

O efeito desta mudança de escala pode ser visualizado de uma maneira mais prática fazendo uso do conceito de resolução. Inicialmente deve-se introduzir a ideia das funções de escalonamento.

2.2.3 Funções de escalonamento

O espaço $\mathcal{L}^2(R)$ é definido como

$$v_o = \text{span}\{\phi_{m,n}(t)\}, \quad (20)$$

em que a designação *span*, por definição, é dada ao espaço formado por todas as combinações possíveis de funções presentes em $\mathcal{L}^2(R)$ (STRANG, 2016). Ou seja, subespaço v_o é fechado e seja uma função $f(t) \in v_o$, esta pode ser escrita como combinação linear das diversas translações e dilatações de $\phi_{m,n}(t)$.

$$f(t) = \sum_n C_{m,n} \phi_{m,n}, \quad (21)$$

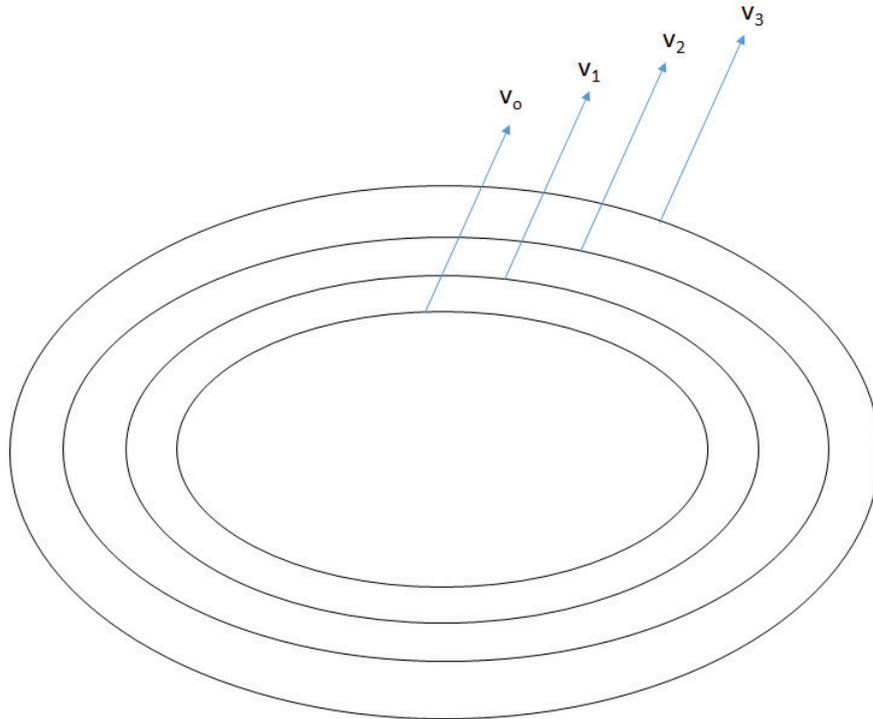
em que $C_{m,n}$ está evidenciado em (18).

Pode-se aumentar o tamanho do subespaço mudando a escala m . Ou seja, se uma dada função pertencente a $\mathcal{L}^2(R)$ não estiver no subespaço v_o , pode-se mudar para um

subespaço mais amplo v_m realizando a mudança de escala m . Seja $m \in \mathbb{Z}$ pode-se afirmar que $v_0 \subset v_1 \subset v_2 \dots \subset \mathcal{L}^2(\mathbb{R})$ ou seja, $v_m \subset v_{m+1}$.

Para um melhor entendimento da ideia de escalas e resoluções, é interessante imaginar que os elementos de um subespaço são simplesmente elementos dilatados do espaço seguinte (ARRUDA, 2003). A relação entre os espaços expandidos é ilustrada na Figura 5.

Figura 5 – Subespaços expandidos através das funções escalas



Fonte: Próprio autor

A função de escalonamento, $\phi(t)$, é determinada por meio da aplicação recursiva dos coeficientes de filtro (WEEKS, 2012). Todas as informações a respeito das funções de escalonamento e wavelet podem ser encontradas a partir de um conjunto finito de coeficientes $h[k]$, os quais são formadores do filtro passa-baixas. Uma vez que estes coeficientes tenham sido encontrados, permitindo projetar um filtro passa-baixa, os coeficientes de um filtro passa-altas ($g[k]$) podem ser determinados matematicamente a partir de $h[k]$.

A função escalonamento, então será dada por:

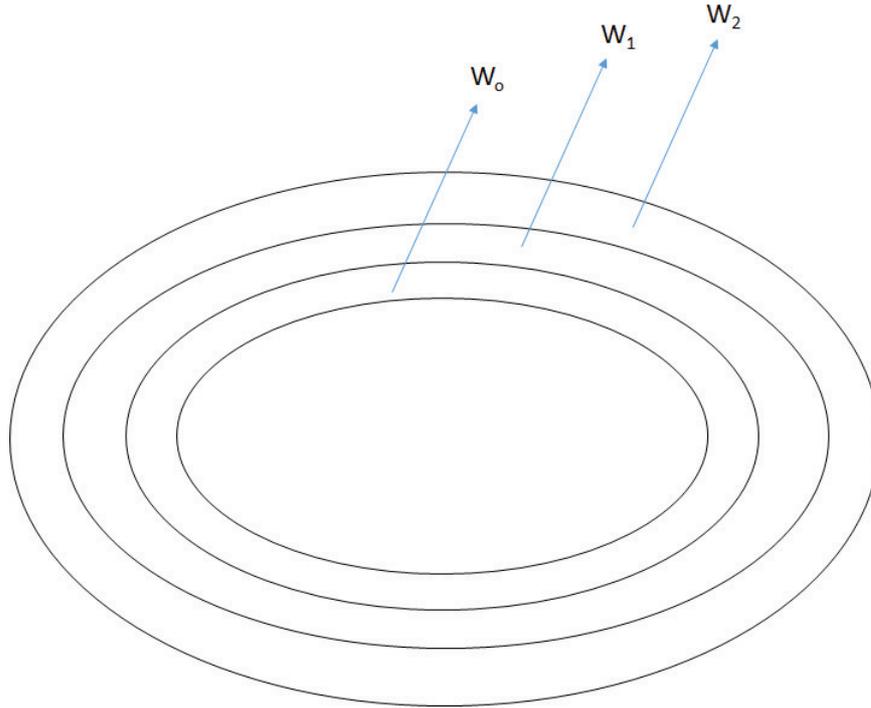
$$\phi(t) = \sqrt{2} \sum_{k \in \mathbb{Z}} h[k] \phi(2t - k) \quad \text{com} \quad \sum_{k \in \mathbb{Z}} |h[k]|^2 < \infty, \quad (22)$$

e a função wavelet será dada por

$$\psi(t) = \sqrt{2} \sum_{k \in \mathbb{Z}} g[k] \phi(2t - k). \quad (23)$$

Definindo W_m como o complemento ortogonal de V_m em V_{m+1} conforme a Figura 6:

Figura 6 – Complementares dos subespaços.



Fonte: Próprio autor

$$W_m \perp V_m \quad \text{e} \quad W_m \oplus V_m = V_{m+1} \quad (24)$$

em que \oplus denota a operação de soma ortogonal, que corresponde ao fechamento linear de dois espaços ortogonais (DINIZ; SILVA; NETTO, 2014). Pode-se, neste contexto, afirmar que o conjunto W_m é formado pelos detalhes que se devem acrescentar às funções de V_m para se alcançar a resolução de V_{m+1} . Daí o porquê de as funções wavelet serem ditas responsáveis por fornecer os detalhes do sinal.

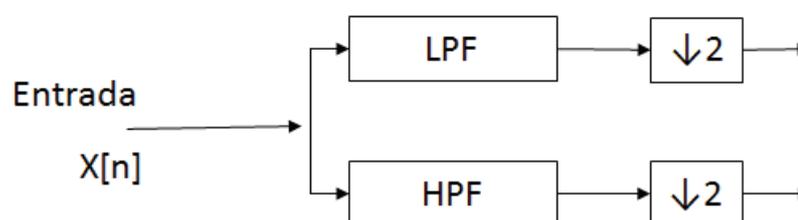
2.3 BANCOS DE FILTROS

Até o presente, foi apresentada a teoria de wavelets de maneira a se obter a transformada a partir de uma *wavelet-mãe*, porém é de interesse prático que seja possível implementar uma

wavelet a partir de um filtro. Nesta sessão é mostrado como as wavelets são utilizadas para decompor um dado sinal de entrada em sinais de aproximação e detalhes através do conceito de bancos de filtros.

Imaginando um dado sinal de entrada a ser analisado através de um par de filtros de resposta finita ao impulso (*Finite Impulse Response - FIR*) conforme apresentados na Figura 7, desde que os coeficientes deste par de filtros atenda a certos critérios, o resultado da passagem do sinal por este par de filtros será a transformada wavelet discreta deste sinal em um nível. Dá-se também a este nível de resolução o nome de *oitava*, termo emprestado livremente da terminologia musical (WEEKS, 2012).

Figura 7 – Par de Filtros de análise.



Fonte: Próprio autor.

Um dos filtros da Figura 7 é um passa-altas (HPF) que corresponde à função wavelet o outro filtro é um passa-baixa (LPF) correspondente a uma função escala.

Os HPF fornecem o conteúdo de alta frequência (detalhes) do sinal, ao passo que o filtro LPF fornecerá uma versão suavizada (aproximação) do sinal. A aproximação é uma versão semelhante do sinal, apenas em uma escala mais curta. Em ambos os casos, após os filtros, aparece um bloco com o símbolo de uma seta para baixo e o número 2 que indica a operação de subamostragem. Nesta operação o sinal é comprimido e tem seu tamanho (número de amostras) reduzido à metade e evitando as redundâncias.

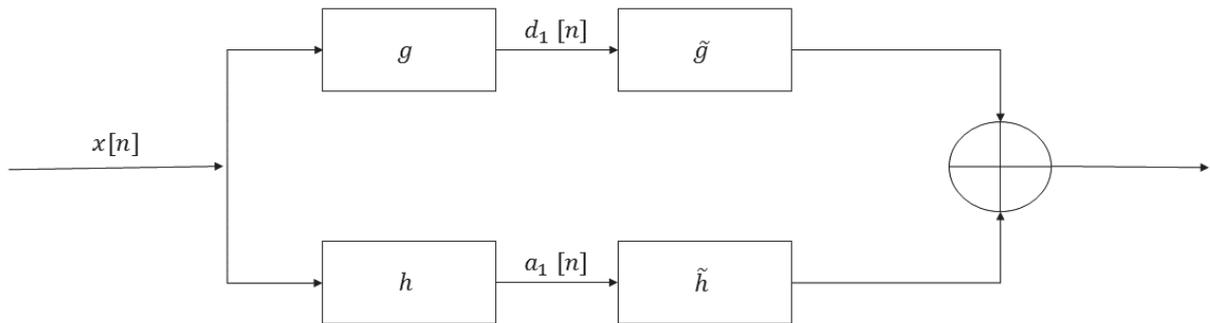
2.3.1 Filtros espelhados em quadratura

Bancos de filtros de dois canais podem, às vezes, receber nomenclaturas como *codificadores sub-bandas* muito embora codificadores de sub-banda possam ter mais de dois canais. Outra forma de serem chamados é de banco de filtros espelhados em quadratura (*Quadrature Mirror Filters - QMF*). Para se enquadrar nestas categorias, são impostas algumas condições aos coeficientes de filtros para que se evite a sobreposição espectral e se obtenha a melhor reconstrução possível.

Para um banco de filtros de dois canais conforme mostrado na Figura 8, a parte da esquerda (filtros g e h) é chamada conjunto de análise ao passo que a parte de direita (filtros \tilde{h} e \tilde{g}) é chamada conjunto de reconstrução. Na parte de análise, os filtros g e h são, respectivamente,

o filtro de detalhe e de aproximação sendo, conseqüentemente, o sinal $d_1[n]$ de detalhe resultante do filtro g e o sinal $a_1[n]$ a aproximação resultante da passagem do sinal original $x[n]$ pelo filtro $h[n]$. Em um banco de filtros espelhados em quadratura, a particularidade reside no fato de ser necessário apenas determinar os coeficientes de um dos filtros, neste caso h , e os coeficientes dos outros serem determinados a partir destes.

Figura 8 – Banco de filtros de dois canais.

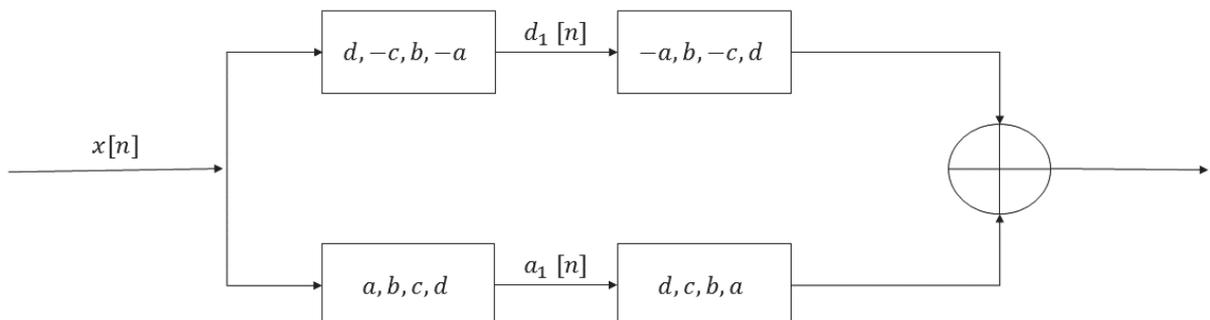


Fonte: Próprio autor

2.3.2 Filtros em quadratura conjugada

Basicamente um banco de filtros em quadratura conjugada representa um aperfeiçoamento do banco de filtros em quadratura (WEEKS, 2012). Este especifica que, dado h e seus coeficientes, g será uma versão invertida de h com sinais alternados. Ou seja, dado um banco h cujos coeficientes são $h = [a, b, c, d]$, tem-se que os coeficientes de g serão dados da forma $g = [d, -c, b, -a]$. Para reconstrução, \tilde{h} e \tilde{g} são versões invertidas de h e g conforme mostrado na Figura 9.

Figura 9 – Banco e filtros em quadratura conjugada.

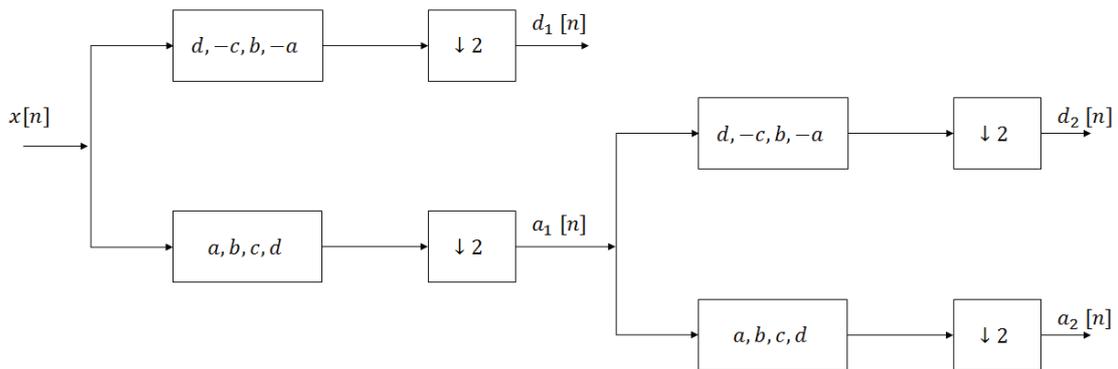


Fonte: Próprio autor.

2.4 MULTIRRESOLUÇÃO

Com base em todas as informações reunidas até o presente tópico, pode-se agora definir multirresolução. Trata-se do processo de tomar a saída de um canal de um banco de filtros e aplicá-la em um novo banco de filtros. No caso das transformadas wavelet. Para um banco de filtros de dois canais, a saída do LPF é aplicada em um novo banco de filtros de dois canais com os mesmos coeficientes do banco de filtros do qual esta saída faz parte, conforme exemplificado na Figura 10.

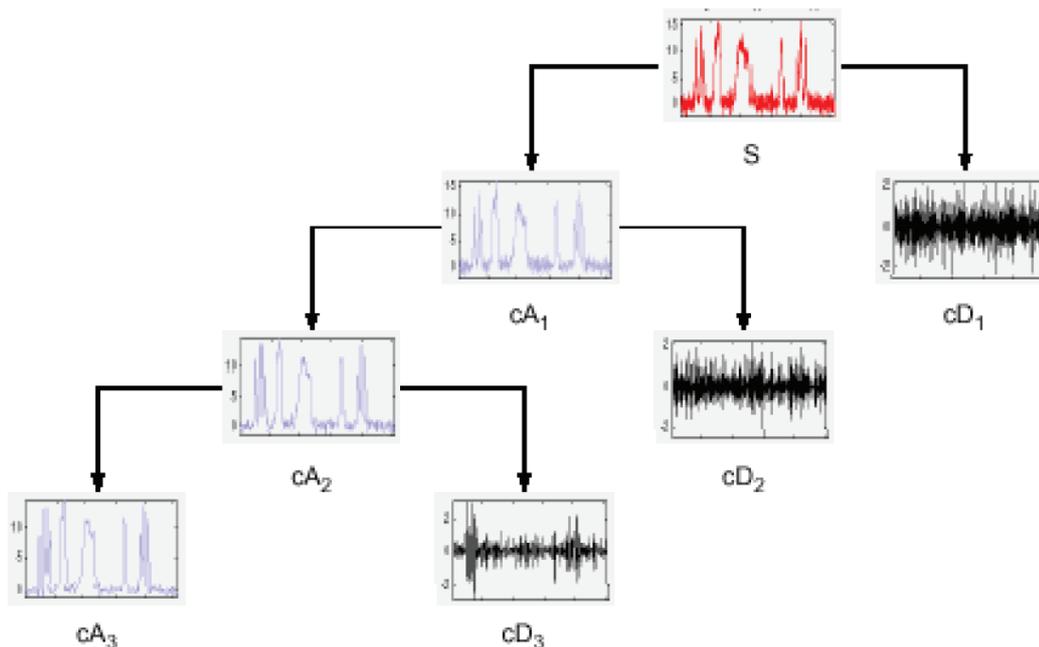
Figura 10 – Análise multirresolução do sinal $x[n]$ até o segundo nível de resolução.



Fonte: Próprio autor.

Como resultado desta operação, o que se tem na saída de cada banco de filtro subsequente é uma aproximação mais suave da função e um novo sinal de detalhes. Esta forma de decomposição pode ocorrer sucessivamente. A Figura 11 mostra o processo sendo realizado com um sinal para ser mais ilustrativo.

Figura 11 – Ilustração de uma AMR até o terceiro nível.



Fonte:(LIRA, 2004)

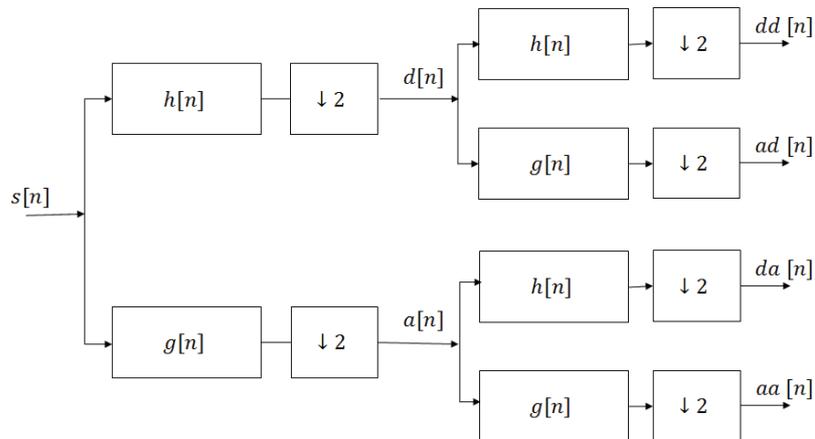
Devido ao fato de o processo ser iterativo, pode-se imaginar que seja possível decompor um sinal indefinidamente. Contudo, devido aos processos de subamostragem que ocorrem a cada nova iteração da decomposição, o número de decomposições é limitado pela quantidade de amostras que os sinais mantêm a cada nível.

2.5 TRANSFORMADA WAVELET PACKET

A transformada *wavelet packet* funciona como um “microscópio” (LIN; SONG; TIAN, 2011). Ela pode dividir o sinal em várias faixas de frequência e, dessa forma, um novo sinal “filho” dentro desta faixa pode ser recuperado.

A ideia de *wavelet packet* consiste, de uma maneira simples, em construir uma árvore binária completa de decomposição utilizando os bancos de filtros que foram apresentados na AMR (PAIVA; MACIEL; MONTENEGRO, 2005). O esboço de uma árvore binária *wavelet packet* está evidenciado na Figura 12.

Figura 12 – Árvore binária básica de uma decomposição *wavelet packet* até o segundo nível.



Fonte: Próprio autor

Ou seja, ao se construir tal árvore binária o que se faz na realidade é uma generalização do conceito de transformada wavelet discreta de maneira que, tanto os coeficientes de aproximação quanto os de detalhes, sejam decompostos desde o primeiro nível fornecendo resultados de maior resolução em tempo ou frequência (MONTEIRO, 2014).

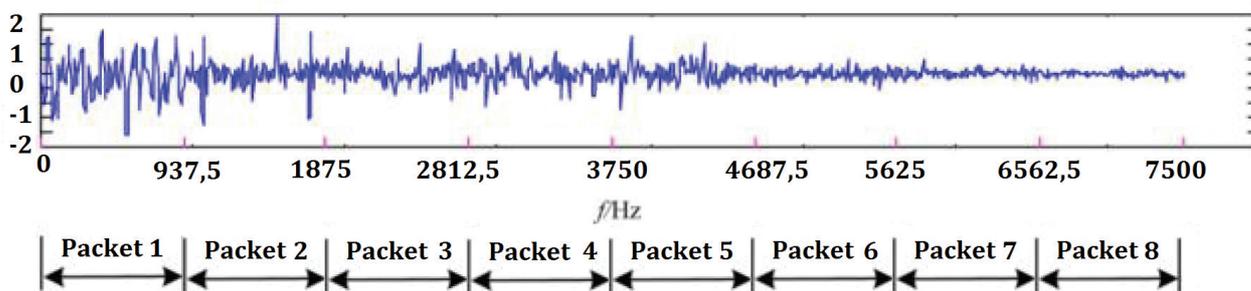
Os blocos com o símbolo “ $\downarrow 2$ ” são conhecidos como subamostradores e têm como função reduzir o número de amostras do sinal saído do filtro à metade. Essa é uma estratégia destinada a evitar a redundância das informações, fato que acontece com a transformada de wavelet contínua.

Conforme se progride nos níveis da árvore, maior é o número de “nós” que ela apresenta (entender por nó a saída de cada um dos filtros). Cada nó corresponde a uma faixa de frequência específica.

Para uma frequência de amostragem f_s , essas faixas de frequência dos nós seriam $(0 \sim \frac{f_s}{2^{n+1}})$, $(\frac{f_s}{2^{n+1}} \sim \frac{f_s}{2^n})$, $(\frac{f_s}{2^n} \sim \frac{f_s}{2^{n-1}})$...

Como resultado da decomposição de um sinal através da transformada *wavelet packet*, consegue-se uma informação muito mais apurada sobre o conteúdo de frequências do sinal (MONTEIRO, 2014). Na Figura 13 ilustra-se, a título de exemplo, as sub-bandas de frequência para 3 níveis de decomposição ($j = 3$).

Figura 13 – Diagrama sub-bandas ou packets de frequência para o ultimo nível (terceiro nível) de decomposição, $j = 3$.



Fonte: (MONTEIRO, 2014)

Nos últimos anos a transformada *wavelet packet* vem sendo utilizada nas mais diversas áreas:

- Remoção de ruídos e compressão de sinais (WALCZAK; MASSART, 1997)
- Redução de ruídos de sinais provenientes de diversas fontes como sinais de voz (CHANG et al.,) e inclusive eletrocardiogramas (TIKKANEN, 1999).
- Detecção de singularidades no sinal (MALLAT; HWANG, 1992) (VENKATAKRISHNAN; SANGEETHA; SUNDAR, 2012)

A decomposição *wavelet packet* é aplicada ao sinal de corrente para em seguida se extrair medidas estatísticas diversas (mediana, desvio padrão, curtose, valor RMS, centroide, fator de crista) dos nós da árvore binária. O intuito é gerar os valores que serão fornecidos a um sistema classificador inteligente para fazer um reconhecimento de padrões e desempenhar a tarefa de monitoramento das condições de um motor de indução.

2.6 MEDIDAS ESTATÍSTICAS

A seleção das características de um sinal é de primordial importância para um diagnóstico exato do status do motor (BARAKAT et al., 2011). Ou seja, as características de cada *packet* serão geradas e organizadas em conjuntos de dados, os quais servirão para o treinamento e teste das RNAs. A partir do conjunto de dados de treinamento, as redes neurais serão capazes de aprender a identificar se há ou não falhas num sinal de corrente, e se houver uma falha, classificá-la. Por isso, a extração de características é fundamental para um correto aprendizado das redes neurais e, conseqüentemente, um correto diagnóstico.

Na metodologia abordada neste trabalho, a decomposição *wavelet packet* é realizada e todos os seus nós têm calculada sua energia e entropia de Shannon. Os nós de detalhe com

maiores valores de energia e de entropia de Shannon, apresentadas na equação 25 (BOUZIDA et al., 2011) (ROMERO-TRONCOSO et al., 2011),

$$H(s_n) = - \sum_{k=1}^N |s_n(k)|^2 \log |s_n(k)|^2, \quad (25)$$

têm os valores de suas medidas estatísticas determinadas. As energias dos nós são calculadas pela expressão 26 (OPPENHEIM; SCHAFER, 2013), (ROBERTS, 2009)

$$E(s_n) = \sum_{k=1}^N s_n[k]^2. \quad (26)$$

Nesta sessão serão apresentadas cada uma destas medidas e como elas serão apresentadas à rede neural para a classificação da condição das máquinas de indução.

2.6.1 Média

Em estatística, média é definida como o valor que mostra para onde se concentram os dados de uma distribuição como o ponto de equilíbrio das frequências em um histograma e é definida por

$$\mu = \frac{1}{N-1} \sum_{i=1}^N s_n(i), \quad (27)$$

em que N é o número de amostras do sinal analisado.

2.6.2 Desvio Padrão

O Desvio Padrão de uma variável aleatória discreta é dado por

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (s_n(i) - \mu)^2}. \quad (28)$$

O desvio padrão pode servir como medida de incerteza. Em ciências, a precisão de medições repetidas é dada pelo desvio padrão. O desvio padrão é crucial para analisar se as medições batem com a previsão teórica (TABACNIKS et al., 2003).

2.6.3 Curtose

A Curtose tem sido frequentemente usada como um indicador quantitativo para avaliar a severidade de falhas em motores de indução (JAIN; WADHWANI; WADHWANI,) (MIAO; HUANG; FAN, 2007b). Por definição, ela é estabelecida como sendo (MATI et al., 2012),

$$K = \frac{1}{3N\sigma^4} \sum_{i=1}^N (s_n(i) - \mu)^4, \quad (29)$$

em que μ e σ são, respectivamente a média da distribuição e seu desvio padrão. O coeficiente de Curtose mede a forma como a concentração de dados se dá em relação ao pico da média de dados em uma distribuição normal (SANTOS, 2013) (CAMARGO,).

2.6.4 Valor RMS

O cálculo de parâmetros estatísticos simples, tais como o valor de raiz médio quadrado (RMS) de um sinal pode fornecer informações úteis. Por exemplo, o valor RMS da velocidade de vibração é uma medida da gravidade global da vibração para casos de diagnóstico através de vibração de motores de indução (GHATE; DUDUL, 2010). Para um dado sinal de tamanho N com média nula, seu valor RMS é dado pela expressão 30.

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N s_n(i)^2}. \quad (30)$$

2.6.5 Centroide

O centroide é uma medida que já vinha sendo usada no processamento de sinal digital, porém no domínio da frequência. Nesta situação, tal medida indica onde o “centro de massa” do espectro está. Perceptivamente, ele possui uma conexão robusta com a impressão de “brilho” quando se trata, por exemplo, do tratamento de um som (GREY; GORDON, 1978).

Em (SANTOS, 2008), o uso desta técnica permitiu compactar a informação de sinais de ultrassom no domínio do tempo para realizar a classificação do nível de poluição em isoladores em linhas aéreas de alta tensão, não importando o tamanho destes sinais, em um vetor com apenas 9 pontos composto pelas áreas e coordenadas x e y dos centroides da parte positiva do sinal (áreas do sinal que estão acima do eixo dos x), dos centroides da parte negativa do sinal

(áreas do sinal que ficam abaixo do eixo dos x) e centroide total do sinal (área calculada através soma da área da parte positiva com a área da parte negativa).

O presente trabalho usa desta filosofia e implementa o cálculo das coordenadas do centroide e da área total dos sinais oriundos das decomposições *wavelet packet* em combinação com as outras medidas estatísticas aqui apresentadas.

2.7 CONCLUSÃO

Neste capítulo são apresentadas a teoria wavelet e todo seu desdobrar desde a análise multirresolução até a transformada *wavelet packet*. Também são mostradas as medidas estatísticas e o centroide que são os valores a serem usadas para cada um dos sinais obtidos da WPT e serão apresentados à rede neural para realização da classificação de padrões.

3 REDES NEURAIS ARTIFICIAIS

3.1 INTRODUÇÃO

Uma vez tendo sido apresentada a transformada wavelet e todo o desenvolvimento que vem a resultar na transformada *wavelet packet*, tendo sido apresentadas as medidas estatísticas que serão obtidas dos sinais para serem apresentadas às RNAs, é chegado o momento de apresentar as redes neurais e os fundamentos necessários à compreensão de seu funcionamento, bem como os métodos de treinamento a serem utilizados e a topologia escolhida.

Pesquisadores das engenharias e ciências aplicadas têm trabalhado este tema desde o início da década de 1950, porém a aplicabilidade das Redes Neurais Artificiais (RNA) veio conquistar destaque somente no início da década de 1990 (SILVA; SPATTI; FLAUZINO, 2010).

No contexto deste projeto, as redes neurais vêm desempenhar um papel crucial na realização de diagnósticos de falhas em motores de indução trifásicos (MIT).

O recente sucesso das RNA em modelagem de sistemas elaborados tem sido de grande vantagem no sentido de ter potencial para realizar diagnósticos confiáveis (SU; CHONG,).

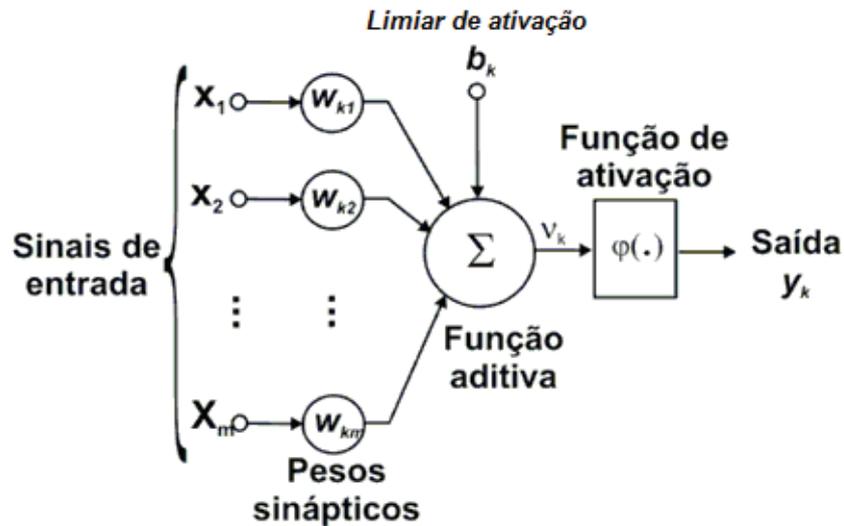
3.2 O NEURÔNIO ARTIFICIAL

A estrutura básica de uma RNA foi desenvolvida com base em modelos conhecidos de sistemas biológicos. Os Neurônios artificiais utilizados nas RNA são estruturas não-lineares, fornecem saídas tipicamente contínuas e realizam funções simples como coletar os sinais provenientes de um ambiente que são aplicados a suas entradas e agregá-los a uma função operacional de maneira a produzir uma saída. A Figura 14 mostra a representação de um neurônio artificial.

Pode-se observar na Figura 14 as seguintes partes:

1. Nós (sinais) de entrada - São a parte do neurônio que se mantém em contato com o meio externo ou recebe informações de outros neurônios. Depois de recebidos, os sinais são processados e enviados à camada de saída.
2. Pesos sinápticos - São valores destinados a ponderar as variáveis da rede. Estão sempre presentes nas conexões entre os nós de entrada e os combinadores lineares.
3. Combinador Linear - É representado pela simbolo \sum e realiza a operação de somatório dos sinais ponderados.

Figura 14 – Estrutura de um neurônio artificial.



Fonte: (HAYKIN, 2003)

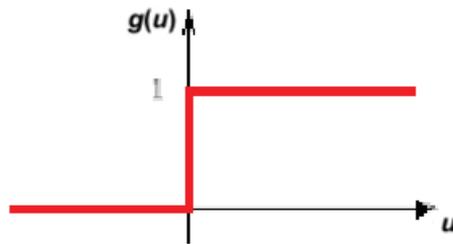
4. Limiar de ativação “Bias” - É uma variável que especifica o patamar apropriado para o combinador linear para gerar o potencial de ativação. Pode ser ajustado a zero quando necessário.
5. Potencial de Ativação (v_k) - Resultado da operação de combinação linear e é responsável por ativar ou inibir a função de ativação
6. Função de Ativação - Seu objetivo é dar o valor de saída do neurônio em valores que se estendem no intervalo entre 0 e 1.

As funções de ativação podem ser classificadas em dois tipos principais.

- Funções parcialmente diferenciáveis
- Funções totalmente diferenciáveis

As funções parcialmente diferenciáveis são do tipo que possuem pontos onde a derivada primeira não pode ser determinada. Seus três principais tipos estão apresentados nas Figuras 15, 16 e 17.

Figura 15 – Função degrau.

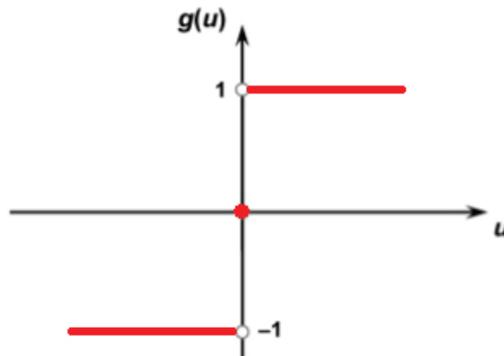


Fonte: Próprio autor

Matematicamente a função degrau (de Heaveside) é expressa por

$$g(u) = \begin{cases} 1 & \text{se } u \geq 0 \\ 0 & \text{se } u < 0 \end{cases} . \quad (31)$$

Figura 16 – Função degrau bipolar.

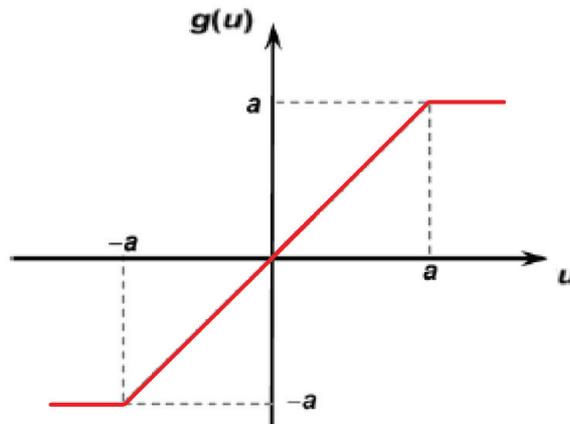


Fonte: Próprio autor

De maneira similar a função degrau bipolar (função sinal) será expressa por

$$g(u) = \begin{cases} 1 & \text{se } u \geq 0 \\ -1 & \text{se } u < 0 \end{cases} , \quad (32)$$

Figura 17 – Função rampa simétrica.



Fonte: Próprio autor

a expressão para a função rampa será

$$g(u) = \begin{cases} a & \text{se } u > a \\ u & \text{se } -a \geq u \geq -a \\ -a & \text{se } u < -a \end{cases} . \quad (33)$$

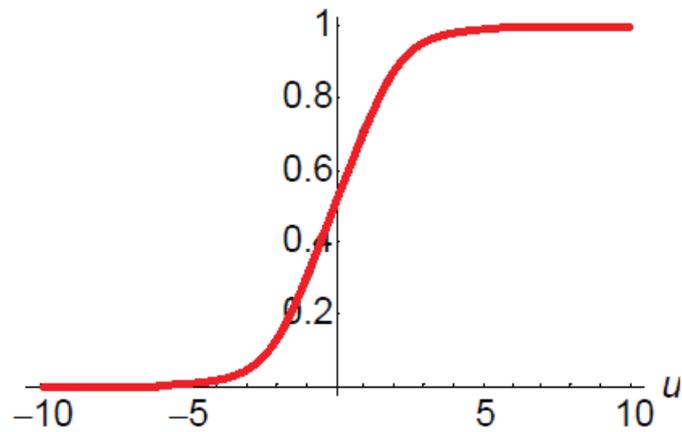
As funções de ativação totalmente diferenciáveis são contínuas em todo seu domínio e apresentam a primeira derivada para todos os pontos. São elas:

1. Logística sigmoide;
2. Tangente hiperbólica;
3. Gaussiana;

exemplificadas nas Figuras 18, 19, 20 e 21.

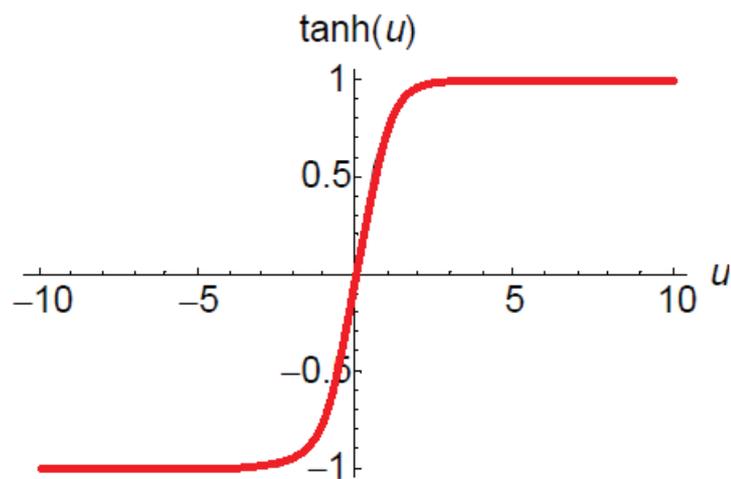
Figura 18 – Função logística sigmoide.

$$L(u) = \frac{1}{1 + e^{-\alpha u}}$$



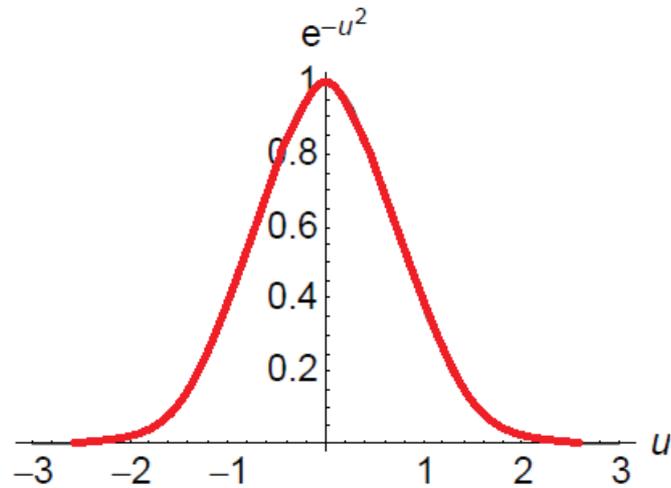
Fonte: Próprio autor.

Figura 19 – Função tangente hiperbólica.



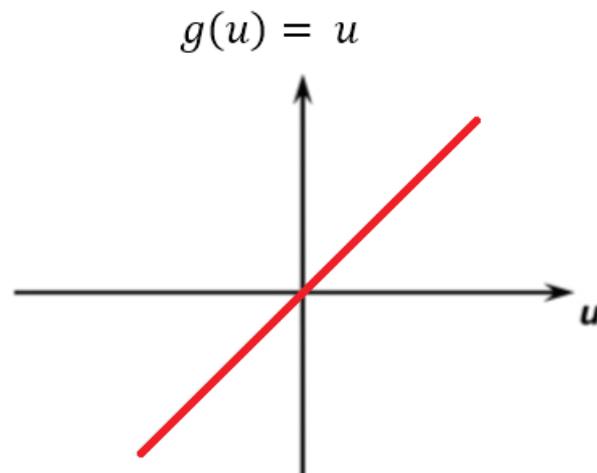
Fonte: Próprio autor.

Figura 20 – Função gaussiana



Fonte: Próprio autor.

Figura 21 – Função linear



Fonte: Próprio autor.

3.3 ARQUITETURAS DAS RNA E PROCESSOS DE TREINAMENTO

A arquitetura de uma RNA é a forma como certos neurônios são dispostos uns em relação aos outros e o direcionamento de suas conexões sinápticas. Já a topologia, dentro de uma arquitetura, trata das diversas maneiras de composição estrutural que a arquitetura pode assumir.

Por exemplo, pode-se ter duas redes de mesma arquitetura porém uma composta por 10 neurônios e outra composta por 20. Outra situação seriam duas que, mesmo tendo a mesma

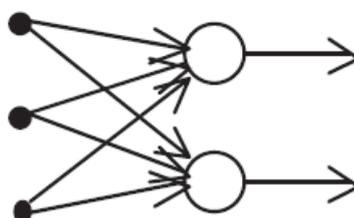
quantidade de neurônios, podem divergir em topologia apenas mudando a função de ativação dos neurônios constituintes da rede, por exemplo.

3.4 PRINCIPAIS ARQUITETURAS DE RNA

3.4.1 Arquitetura *feedforward* simples

Neste tipo de arquitetura a informação flui em um único sentido. É constituída basicamente dos nós de entrada e uma única camada de neurônios apenas, que por sua vez já vem a fornecer a saída da dita rede, como mostrado na Figura 22.

Figura 22 – Exemplo de rede *feedforward* simples.



Fonte: Próprio autor.

Dentre os principais tipos de topologias que fazem parte deste tipo de arquitetura estão o perceptron e adaline.

3.4.2 Arquitetura *feedforward* de camadas múltiplas

Nesta arquitetura, o diferencial está no fato de existir uma ou mais camadas ocultas (intermediárias) de neurônios. De fato, o número de camadas intermediárias, bem como o de neurônios, depende do nível de complexidade do problema que se deseja resolver através do uso das redes neurais (SILVA; SPATTI; FLAUZINO, 2010).

3.5 PROCESSOS DE TREINAMENTO

O treinamento de uma rede neural é o processo de encontrar valores para os pesos e para os limiares de ativação para que, por um conjunto de dados de treinamento com conhecidos valores de entrada e saída, as saídas calculadas da rede coincidam com as saídas desejadas, existentes no conjunto de dados a ela apresentados.

3.5.1 Treinamento supervisionado

Esta forma de treinamento consiste em ter, para cada amostra de entrada, as respectivas saídas desejadas. Dessa maneira, é importante que se tenha uma tabela representativa dos dados

Entrada/Saída. Durante o processo de treinamento, essa relação entre os atributos de entrada e os valores de saída atua sobre o aprendizado da rede como se fosse um professor. Daí a explicação de este processo também receber o nome popular de aprendizagem com professor (HAYKIN, 2003) (SILVA; SPATTI; FLAUZINO, 2010).

3.5.2 Treinamento não-supervisionado

De maneira diferente do Supervisionado, nesta forma de treinamento não existem as saídas desejadas. Consequentemente, a própria rede deve se auto organizar no intuito de, mediante as particularidades de um sinal de entrada, poder se obter uma classificação de saída.

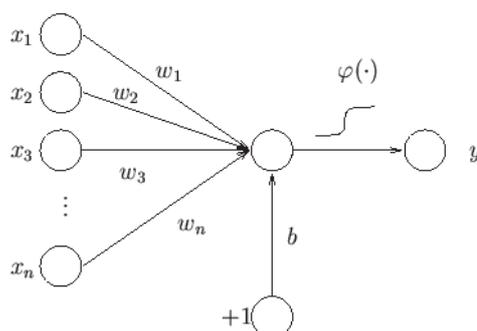
3.5.3 Treinamento com reforço

Estes métodos tem sido considerados uma variação das técnicas de aprendizado supervisionado. Os algoritmos de aprendizagem utilizados ajustam os pesos sinápticos dos neurônios com base em informações adicionais, tanto quantitativas quanto qualitativas, advindas de interações com o ambiente que está sendo mapeado, com o objetivo de melhorar o desempenho do aprendizado.

3.6 REDE PERCEPTRON

Dentre as topologias para a arquitetura *feedforward*, a *Perceptron* é a mais simples sendo, inclusive, a mais simples das configurações de RNA. A Figura 23, é um exemplo clássico de uma rede perceptron consistindo em apenas um neurônio com n entradas.

Figura 23 – Rede *Perceptron* simples



Fonte: Próprio autor

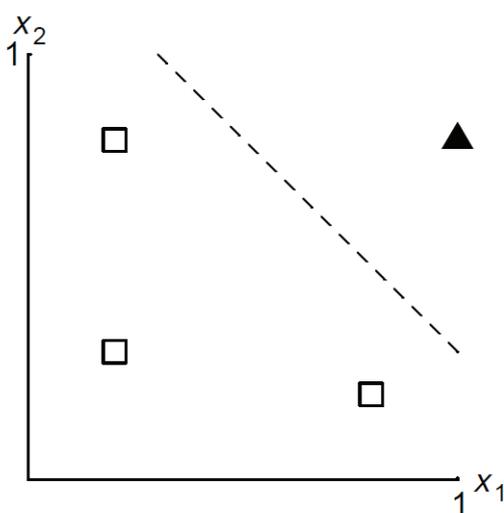
Devido à sua simplicidade, tal topologia tem ganhado preferência nas pesquisas voltadas para implementação de sistemas de diagnóstico *online* de problemas em máquinas rotativas (CASTEJÓ; LARA; GARCÍA-PRADA, 2009), (GHATE; DUDUL, 2011) (GHATE; DUDUL, 2010).

3.6.1 Princípio de funcionamento do *perceptron*

Cada um dos nós de entrada (x_i) é ponderado pelos pesos sinápticos (w_{ij}) em que i representa o número do neurônio a que se destina a sinapse e j representa o número do nó de entrada do qual parte a sinapse. Em relação às entradas, estas podem assumir qualquer valor numérico.

O Perceptron se comporta como um classificador para padrões ditos linearmente separáveis, i.e. padrões que se encontram em lados oposto de um hiperplano conforme mostrado na Figura 24 (HAYKIN, 2003). O perceptron mais simples, constituído em torno de um único neurônio, é limitado a realizar classificação em apenas duas classes. Ou seja diferenciar os dados entre duas categorias distintas, ou mesmo classificá-los como sendo pertencentes a um grupo ou não. Expandindo a quantidade de neurônios presentes nas camadas internas da rede, pode-se realizar a classificação em uma quantidade superior de classes.

Figura 24 – Classes linearmente separáveis. As categorias triângulo e retângulo podem ser separadas através da linha pontilhada.



Fonte : Próprio autor

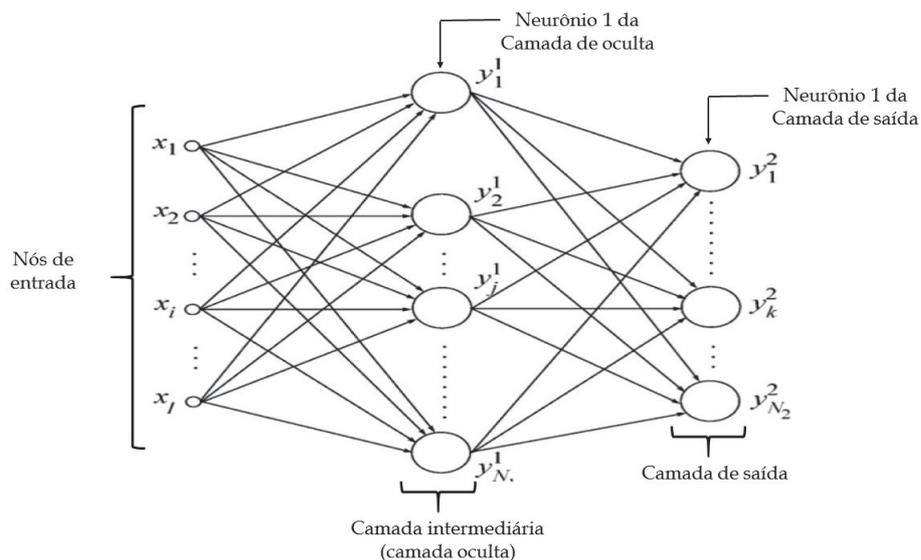
A Figura 24 mostra um plano formado pelos valores que as variáveis de entrada x_1 e x_2 podem assumir. Neste plano estão distribuídos os dados que serão apresentados à rede. nota-se que os dados de uma das categorias (quadrados) estão bem agrupados em uma região ao passo que os dados pertencentes à segunda categoria (triângulos) estão agrupados em outra região de maneira que podem ser separados por uma reta. Em regiões com mais de uma dimensão, os grupos de categorias são divididos por hiperplanos. Contudo, a grande desvantagem do perceptron reside justamente neste fato. Se as categorias não forem linearmente separáveis o perceptron não irá classificar os padrões corretamente; tornando o problema mais complicado e

sendo necessário o uso de uma topologia mais elaborada. É nesse sentido que se usa a topologia perceptron multicamada.

3.7 PERCEPTRON MULTICAMADAS

Das arquiteturas de redes neuronais artificiais, os perceptrons multicamadas (do Inglês *Multilayer Perceptron* - MLP) são um dos modelos de reconhecimento de padrões (qualitativos) e de estimativa (quantitativos) mais populares (ZHANG; TIAN, 2014) (PARDO; SBERVEGLIERI, 2004). A versatilidade desta topologia as fazem ser aplicadas em problemas complexos e de natureza diversa, dando destaque nas seguintes áreas, além do reconhecimento de padrões: aproximação universal de funções, identificação e controle de processos, previsão de séries temporais e otimização de sistemas (ABREU et al., 2012). Neste projeto, a topologia perceptron multicamada, apresentada na Figura 25, é utilizada para realizar o reconhecimento de padrões tendo como fonte as correntes de fase no motor trifásico de indução. Na Figura 25, os nós de entrada são representados por x_i em que o subscrito i indica o número do nó e varia de 1 até o número de dimensões do vetor formado pelos dados apresentados à rede. Os elementos indicados por y_j^k são as saídas dos neurônios formadores das camadas ocultas. O termo k representa o número da camada e j o número do neurônio pertencente a k -ésima camada.

Figura 25 – Rede *perceptron* multicamada com três camadas (entrada, camada oculta e saída).



Fonte: Próprio autor

Além da topologia MLP, existem as funções base radial (do Inglês *Radial Basis Function* - RBF).

Em termos de arquitetura, existem ainda as redes reticuladas (aplicáveis a problemas de agrupamento e otimização) e as redes recorrentes (modelos de previsão e séries temporais).

3.7.1 Princípio de funcionamento do MLP

Um *perceptron* de múltiplas camadas pode ser visto como uma versão aprimorada do *perceptron* simples apresentando um número maior de camadas intermediárias (ocultas). Uma rede MLP típica possui três características principais:

- Os neurônios das camadas intermediárias (e, eventualmente, os da camada de saída) possuem uma função de ativação não-linear do tipo sigmoide (e.g. função logística ou tangente hiperbólica).
- A rede possui uma ou mais camadas intermediárias.
- A rede possui um alto grau de conectividade. Ou seja, todos os neurônios de uma camada são interligados com os neurônios da camada anterior e da seguinte.

3.8 ALGORITMOS DE TREINAMENTO DAS REDES

O treinamento de redes neurais com várias camadas pode ser entendido como um caso de aproximação de funções. Sendo assim, pode ser visto como um problema de otimização que tem as desvantagens de ser não-linear e não-convexo, porém tendo a vantagem de ser irrestrito. A classe de algoritmos que será intensamente estudada nesta sessão são os chamados métodos de segunda ordem, e que utilizam informações sobre a derivada segunda da função de custo (erro). Os dois algoritmos de treinamento usados para as redes neurais deste projeto são o método de gradientes escalonados conjugados e o algoritmo *resilient backpropagation*. Nesta seção, ambos os algoritmos serão explicados.

3.8.1 Gradiente conjugado escalonado

A classe de métodos de otimização chamados métodos do gradiente conjugado, tratam de problemas de grande escala (grande número de variáveis, matrizes e vetores de altas dimensionalidades) de maneira efetiva. Em comparação a este método, os gradientes conjugados escalonados apresentam a vantagem de não realizar a busca pela solução de maneira unidimensional o que torna o processo um tanto custoso do ponto de vista computacional. Em 1993 Moller (MØLLER, 1993) propôs uma variação no algoritmo de gradientes conjugados em que o passo de ajuste α é escalonado.

Seja a função erro médio dada por

$$\varepsilon_{med} = \frac{1}{N} \sum_1^N \left[d(n)_j - y(n)_j \right]^2, \quad (34)$$

em que $d(n)_j$ e $y(n)_j$ são respectivamente a saída desejada para a rede e a saída fornecida pela mesma.

Quando se trata deste método o passo de ajuste α é dado por

$$\alpha = \frac{\mathbf{d}_j^T \nabla \varepsilon(\mathbf{w})}{\mathbf{d}_j^T \nabla^2 \varepsilon(\mathbf{w}) \mathbf{d}_j}, \quad (35)$$

em que \mathbf{w} é o vetor dos pesos a serem ajustados e \mathbf{d}_j é a direção de ajuste na iteração j , dada por

$$\mathbf{d} = - \frac{\nabla \varepsilon(\mathbf{w})}{\|\nabla \varepsilon(\mathbf{w})\|}, \quad (36)$$

o termo $\nabla \varepsilon(\mathbf{w})$ é o gradiente da função erro médio. ∇ é um operador derivativo

$$\nabla = \left[\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_n} \right]; \quad (37)$$

e $\nabla^2 \varepsilon(\mathbf{w})$ é a matriz Hessiana apresentando o seguinte formato

$$\nabla^2(\varepsilon) = \begin{bmatrix} \frac{\partial^2}{\partial_1 \partial_1} & \frac{\partial^2}{\partial_1 \partial_2} & \cdots & \frac{\partial^2}{\partial_1 \partial_n} \\ \frac{\partial^2}{\partial_2 \partial_1} & \frac{\partial^2}{\partial_2 \partial_2} & \cdots & \frac{\partial^2}{\partial_2 \partial_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2}{\partial_m \partial_1} & \frac{\partial^2}{\partial_m \partial_2} & \cdots & \frac{\partial^2}{\partial_m \partial_n} \end{bmatrix}; \quad (38)$$

O cálculo da matriz hessiana é oneroso do ponto de vista computacional; a ideia de Moller é estimar um termo aproximado \mathbf{s}_j dado por

$$\mathbf{s}_j = \nabla^2 \varepsilon(\mathbf{d}_j) \approx \frac{\nabla \varepsilon(\mathbf{w}_j + \sigma_j \mathbf{d}_j) - \nabla \varepsilon(\mathbf{w}_j)}{\sigma_j} \quad 0 < \sigma_j < 1, \quad (39)$$

a aproximação tende, no limite, ao valor de $\nabla^2 \varepsilon(\mathbf{d}_j)$. De posse desta informação é possível fazer

uma combinação do método dos gradientes conjugados e do método de Levenberg-Marquadt resultam em um algoritmo diretamente aplicável a redes *feedforward*

$$\mathbf{s}_j = \nabla^2 \varepsilon(\mathbf{d}_j) \approx \frac{\nabla \varepsilon(\mathbf{w}_j + \sigma_j \mathbf{d}_j) - \nabla \varepsilon(\mathbf{w}_j)}{\sigma_j} + \delta_j \mathbf{d}_j, \quad (40)$$

O parâmetro δ é dado por

$$\delta_j = \mathbf{d}_j \mathbf{s}_j. \quad (41)$$

O elevado custo computacional demandado para o cálculo da Hessiana pode ser evitado com o uso destes artifícios.

3.8.2 O método *resilient backpropagation*

A técnica mais comum usado para treinar as redes neurais é o algoritmo retropropagação do erro. Este algoritmo requer um valor para um parâmetro chamado a taxa de aprendizagem. A eficácia da retropropagação é muito sensível ao valor da taxa de aprendizagem. O *resilient back-propagation* (Rprop) foi desenvolvido em 1993, em uma tentativa de melhorar o algoritmo de retropropagação.

O Rprop é considerado o melhor algoritmo, medido em termos de rapidez de convergência, precisão e robustez em relação a parâmetros de treino (CHEN; SU, 2010).

A ideia básica da aprendizagem por retropropagação é a aplicação repetida da regra da cadeia para calcular a influência de cada peso na rede com respeito a uma função de erro ε arbitrária.

$$\frac{\partial \varepsilon(n)}{\partial w_{ij}} = \frac{\partial \varepsilon(n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} \frac{\partial v_j(n)}{\partial w_{ij}(n)}. \quad (42)$$

A derivada parcial $\frac{\partial \varepsilon(n)}{\partial w_{ij}}$ representa um *fator de sensibilidade* para o peso sináptico w_{ij} . Uma vez que a derivada parcial para cada peso é conhecida, é alcançado o objetivo de minimizar a função de erro através da realização de uma descida de gradiente simples (RIEDMILLER;

BRAUN, 1993):

$$w_{ij}(n+1) = w_{ij}(n) - \epsilon \frac{\partial \epsilon}{\partial w_{ij}}(n), \quad (43)$$

A escolha da taxa de aprendizagem ϵ , que escalona a derivada, tem um efeito importante sobre o tempo necessário até a convergência do algoritmo. Se ela for muito pequena, são necessários muitos passos para chegar a uma solução aceitável; pelo contrário uma grande taxa de aprendizagem vai possivelmente levar à oscilação, evitando que o erro venha a cair abaixo de um certo valor.

Muitos algoritmos foram propostos para tratar de forma apropriada do problema da atualização dos pesos, fazendo algumas adaptações de parâmetros durante a aprendizagem. Eles podem aproximadamente ser separados em duas categorias: global e local. Técnicas de adaptação globais fazem uso do conhecimento de estado em totalidade da rede (por exemplo, a direção do peso da etapa anterior) para modificar os parâmetros globais. As estratégias locais utilizam apenas informações específicas dos pesos (por exemplo a derivada parcial) para se adaptar.

Ambas as categorias de algoritmos adaptativos, globais e locais, realiza uma modificação da taxa de aprendizagem de acordo com o comportamento observado da função de erro.

O Rprop é um método eficiente que realiza uma adaptação direta do vetor de pesos, com base em informações gradiente local.

A diferença crucial, em comparação às técnicas de adaptação desenvolvidas anteriormente a este método, é que o esforço de adaptação não é mais influenciado por um comportamento qualquer do gradiente.

Para conseguir isso introduz-se para cada peso um valor de atualização Δ_{ij} , que determina unicamente o tamanho da atualização para os novos pesos. Este valor de atualização adaptativo evolui durante o processo de aprendizagem com base em sua visão local da função de erro de acordo com o seguinte regra:

$$\Delta_{ij}(n) = \begin{cases} \eta^+ \times \Delta_{ij}(n-1), & \text{se } \frac{\partial \epsilon}{\partial w_{ij}}(n-1) \times \frac{\partial \epsilon}{\partial w_{ij}}(n) > 0 \\ \eta^- \times \Delta_{ij}(n-1), & \text{se } \frac{\partial \epsilon}{\partial w_{ij}}(n-1) \times \frac{\partial \epsilon}{\partial w_{ij}}(n) < 0 \end{cases} \quad (44)$$

em que $0 < \eta^- < 1 < \eta^+$.

Em termos mais simples, a regra de adaptação funciona da seguinte maneira: a cada vez que a derivada parcial do peso correspondente w_{ij} mudar de sinal, o que indica que a última

atualização era muito grande e o algoritmo “saltou” um mínimo local, o valor de atualização Δ_{ij} é diminuído de um fator η^- . Se a derivada mantém o seu sinal, o valor de atualização é ligeiramente aumentado, a fim de acelerar a convergência.

Uma vez que o valor de atualização para cada peso é adaptado, a atualização de pesos segue uma regra muito simples: se a derivada é positiva (erro a aumentar), o peso é diminuído pela sua atualização de valor, se a derivada é negativa, a atualização de valor é adicionada:

$$\Delta w_{ij}(n) = \begin{cases} -\Delta_{ij}(n) & \text{se } \frac{\partial \varepsilon}{\partial w_{ij}}(n) > 0 \\ +\Delta_{ij}(n) & \text{se } \frac{\partial \varepsilon}{\partial w_{ij}}(n) < 0 \\ \text{senão } & 0, \end{cases} \quad (45)$$

$$w_{ij}(n+1) = w_{ij}(n) + \Delta w_{ij}(n). \quad (46)$$

No entanto, existe uma exceção: Se a derivada parcial muda de sinal, isto é, o passo anterior foi muito grande e o mínimo foi perdido, a atualização anterior é revertida:

$$\Delta w_{ij}(n) = -\Delta w_{ij}(n-1), \quad \text{se } \frac{\partial \varepsilon}{\partial w_{ij}}(n-1) \times \frac{\partial \varepsilon}{\partial w_{ij}}(n) < 0. \quad (47)$$

Devido a esta “resiliência”, a derivada supostamente irá mudar de sinal mais uma vez no passo a seguir. A fim de evitar uma dupla punição da atualização de valores, não deve haver nenhuma adaptação do valor atualização na etapa seguinte. Na prática, isto pode ser feito pelo ajuste $\frac{\partial \varepsilon}{\partial w_{ij}}(n-1) = 0$ na regra de atualização de Δ_{ij} .

O *resilient back-propagation*, apesar de ser semelhante ao algoritmo de retropropagação comum, tem duas vantagens principais: o treinamento com Rprop é menos oneroso do ponto de vista computacional sendo mais eficiente que o treinamento com a retropropagação. Em segundo lugar, o Rprop não requer que se especifique quaisquer valores de parâmetros livres, ao contrário da retropropagação que precisa de valores para a taxa de aprendizagem e, geralmente, um termo impulso opcional. A principal desvantagem do Rprop é que é um algoritmo mais complexo de implementar.

3.9 CONCLUSÃO

Neste capítulo são apresentados os conceitos relacionados às Redes Neurais Artificiais necessários para compreensão da metodologia adotada neste trabalho, bem como se explica os métodos de treinamento e a topologia escolhida para desempenhar a tarefa de reconhecimento de padrões. O próximo capítulo trata dos ensaios realizados no motor trifásico de indução e dos resultados quanto à classificação de falhas elétricas e mecânicas.

4 METODOLOGIA

4.1 METODOLOGIA DE ANÁLISE

4.1.1 Aspectos dos dados experimentais

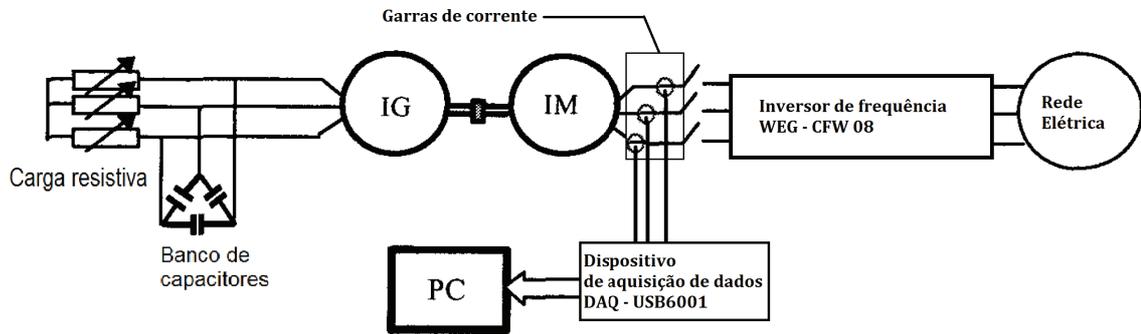
Este capítulo tem por objetivos detalhar a forma como foram conduzidos os ensaios de defeitos no motor de indução, justificar a escolha da wavelet usada para realização da decomposição dos sinais, apresentar as configurações de redes neurais adotadas e, por fim, mostrar as combinações de medidas estatísticas para realizar o monitoramento e classificação das possíveis falhas mecânicas e elétricas que ocorrem nos motores de indução. Para a elaboração deste capítulo foram ensaiados tanto defeitos mecânicos (problemas de rotor e nos rolamentos), quanto faltas incipientes entre espiras de estator.

Os dados levantados para realização de todo processo de identificação das falhas são obtidos a partir das assinaturas de corrente de fase do motor, metodologia essa também conhecida como *Motor Current Signature Analysis* (THOMSON; GILMORE, 2003). O acionamento do motor de indução de 1/3 CV e 4 polos foi através de partida direta e inversor (modelo CFW 08 de fabricação WEG). O motor recebeu diretamente uma tensão de 380V, 60 Hz de linha em seus terminais de estator para os ensaios de defeitos de estator e para os ensaios de problemas mecânicos o acionamento foi realizado através de inversor.

4.1.2 O experimento

Os sinais coletados são as correntes de linha de um motor de fabricação WEG modelo W22 de 1/3 CV, 4 polos e 1760 rpm, para a situação em vazio e com carga. A carga a ser acionada por este motor é um segundo motor de indução sendo acionado na forma de gerador assíncrono alimentando uma carga resistiva. O esquema de montagem é apresentado na Figura 26.

Figura 26 – Esquemático do *setup* montado em laboratório para realização dos ensaios com carga no motor de indução.



Fonte: Próprio autor.

O motor utilizado como gerador assíncrono foi um motor também de fabricação WEG de 1/2 CV, 4 polos e 1680 rpm. A Figura 27 mostra os motores montados para realização do ensaio.

Figura 27 – Motores de indução montados para realização dos ensaios em bancada.



Fonte: Próprio autor

O motor à esquerda é o motor 1, ensaiado, sendo acionado através do inversor e o motor à direita funcionando como gerador alimentando a carga resistiva detalhada na Figura 28 e a Tabela 3 mostra os dados de placa de ambos.

Tabela 3 – Dados de placa dos motores utilizados nos ensaios

	Motor 1 (ensaiado)	Motor 2 (carga)
Potência (CV)	1/3	1/2
Velocidade nominal (rpm)	1720	1680
Fator de potência	0,64	0,72
Tensão de alimentação (triângulo / estrela)	220/380	220/380
Corrente de operação (triângulo / estrela)	1,65/0,95	1,87/1,08

Figura 28 – Resistores de $100\ \Omega$ alimentados pelo motor de indução 2 funcionando como gerador assíncrono.

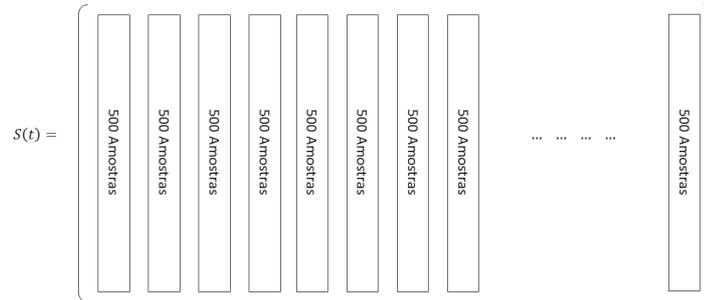


Fonte: Próprio autor

A aquisição destas correntes foi possível graças a uma placa de aquisição *National Instruments*[®] modelo *DAQ – USB6001* e de garras de corrente Minipa modelo HR30 e Agilent Technologies Modelo Agilent 1146A para, todas com resolução de 100mV/A e fundo de escala para 70 A. A faixa de corrente a ser medida é de 2,0 A para o motor 1 operando com carga.

As amostras de corrente são coletadas em conjuntos de 500 (5 ciclos). Cada um destes conjuntos de 5 ciclos de amostras são então organizadas em uma matriz, de maneira que cada conjunto seja uma coluna, conforme a Figura 29 de onde são submetidas à decomposição *wavelet packet*.

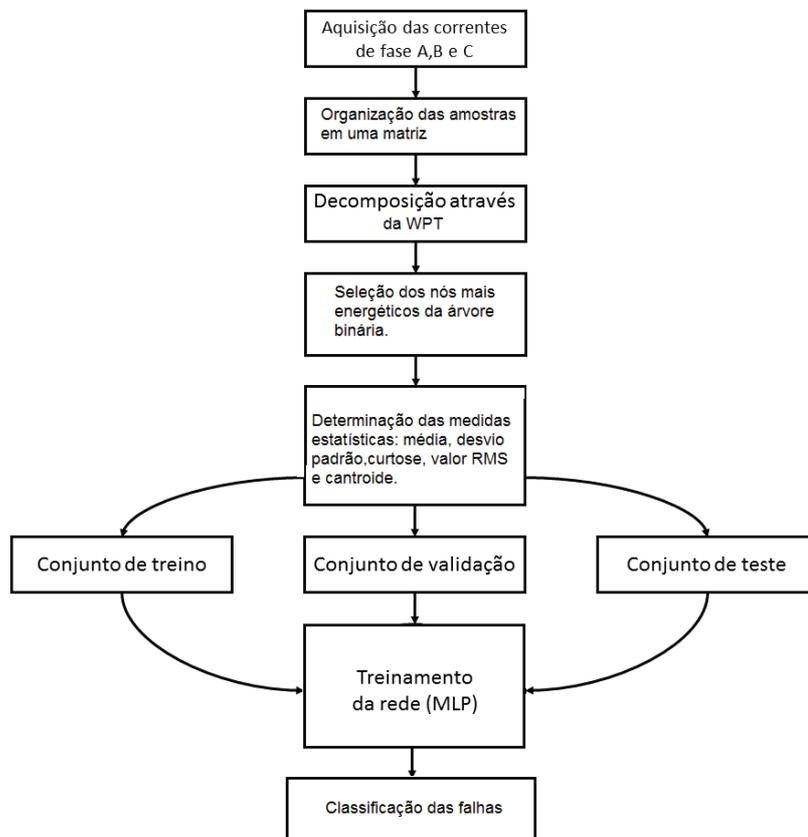
Figura 29 – Organização dos grupos de amostras em uma matriz S.



Fonte: Próprio autor

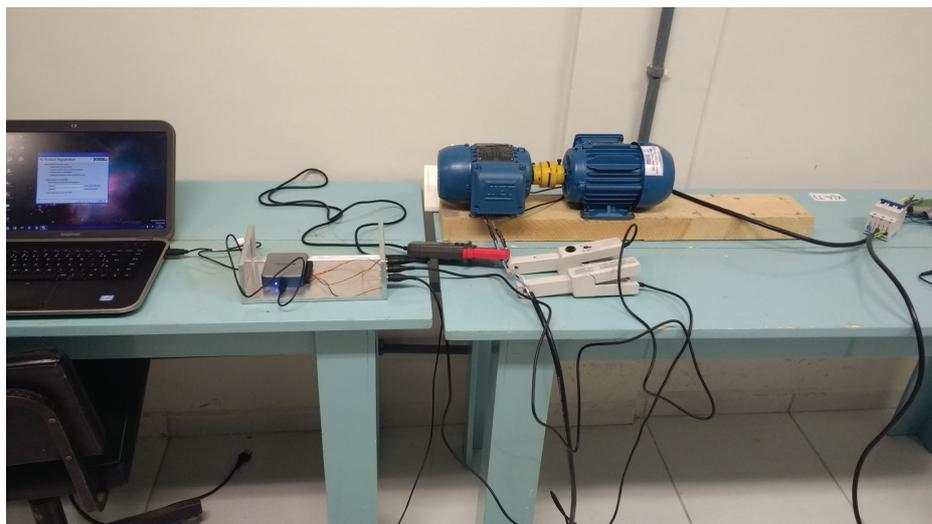
A taxa de amostragem do processo foi de $6kHz$, o que confere um total de 100 amostras por ciclo para uma corrente de regime permanente a $60Hz$. A seguir, na Figura 30, está ilustrado todo processo desde a aquisição das correntes de fase do motor até o treinamento da rede para reconhecimento de padrões.

Figura 30 – Diagrama representativo do processo para construção do sistema classificador de falhas.



Fonte: Próprio autor

Figura 31 – Montagem dos ensaios em laboratório.



Fonte: Próprio autor

Como evidenciado na Figura 30, após a obtenção dos sinais e sua separação em grupos de 500 amostras, cada um destes grupos de amostras é submetido à WPT e tem sua *wavelet packet tree* analisada com o uso da entropia de Shannon e do cálculo da energia. Uma vez que são determinados os nós que serão utilizados para a extração de informações, as medidas estatísticas são calculadas para que se formem conjuntos de treinamento, teste e validação da rede. As ferramentas computacionais utilizadas em todo este processo de análise foram a *Neural Network Toolbox* e a *Wavelet Toolbox*, ambas do *MatLab*®.

Por padrão do próprio Matlab, as amostras são separadas em proporções de 70% para conjunto de teste, 15% para conjunto de validação e 15% para o conjunto de teste. O total de amostras analisado é de 14170 para defeitos mecânicos e 11248 para as falhas elétricas.

O *setup* completo montado para realização dos experimentos em bancada está mostrado na Figura 31, o software utilizado para a aquisição dos sinais foi o *NI Signal Express 2014* também da *National Instruments*®.

4.1.3 Wavelet utilizada para decomposição dos sinais

A escolha da wavelet-mãe é parte importante no processo de análise dos sinais de assinatura de corrente, pois segundo (LIRA, 2004), wavelets com comportamentos mais suaves permitem melhor resolução no domínio da frequência, o que é importante para as falhas mecânicas, uma vez que estas produzem componentes de corrente em frequências variadas, enquanto que variações em pequenos intervalos de tempo caracterizam melhor resolução no domínio do tempo, pois permitem a localização de singularidades no sinal que podem ser provenientes das faltas incipientes no estator. Uma abordagem útil para escolha da wavelet em questão é proposta por (UZINSKI, 2013) que, através de técnicas analíticas, consegue estabelecer

critérios para que as wavelets possam ser avaliadas segundo regularidade e suavidade através de uma formulação que, basicamente é uma expansão da formulação de Sherlock & Monro (SHERLOCK; MONRO, 1998).

Estudos neste sentido podem ser úteis para que a determinação de uma função nessas condições seja mais formal e menos trabalhosa. Em outras palavras, propriedades das wavelets podem tornar sua escolha para análise de um sinal mais intuitiva, deixando de lado métodos que incluam esta escolha por exaustivos testes (UZINSKI, 2013). Propriedades como regularidade, suavidade e suporte compacto são determinantes na escolha da wavelet a ser usada (DAUBECHIES, 1992) (DAUBECHIES, 1988) (DAUBECHIES, 1990) (DAUBECHIES, 1993). Momentos nulos e suavidade estão matematicamente relacionados, pois quanto maior o número de momentos nulos de uma wavelet, mais suave ela é. Uma wavelet tem n momentos nulos quando (BŁASZCZUK; POZORSKI, 2007) (MIAO; HUANG; FAN, 2007b).

$$\int_{-\infty}^{+\infty} x^k \psi(x) dx = 0 \quad \text{para } 0 \leq k \leq n. \quad (48)$$

Um elevado número de momentos nulos permite melhor compressão das partes regulares do sinal, entretanto, o aumento do número de momentos nulos aumenta também o tamanho do suporte das wavelets. Ou seja, a escolha adequada de uma wavelet é uma questão de “negociar” entre o tamanho de seu suporte e seu número de momentos nulos (MALLAT, 1989).

Outra propriedade importante que a família de wavelets deve apresentar são as bases ortogonais. Existe uma grande quantidade de famílias que apresentam bases ortogonais, que são as famílias Daubechies (dbN), Symlets (*symN*), Coiflets (*coifN*), e Meyer (*meyr*) (NING, 2014), sendo N a ordem da função wavelet. O problema de selecionar uma wavelet pode ser visto como selecionar uma proveniente de de uma família de base ortogonal (LI; MONTI; PONCI,). Uma wavelet que venha a ser considerada adequada para a realização da MCSA deve, então atender a alguns critérios.

- A wavelet deve possuir uma quantidade considerável de momentos nulos e, conseqüentemente regularidade e suavidade.
- Ser de uma família de bases ortogonais.
- Apresentar Suporte compacto.

Existem restrições que, se atendidas, garantem que a wavelet sob análise tenha pelo menos três momentos nulos. Sejam as funções de transferência do banco de filtros (GALVÃO, 2015),

$$H^{(N)}(z) = \sum_{i=0}^{2N-1} h_i^{(N)} z^{-i} \quad , \quad G^{(N)}(z) = \sum_{i=0}^{2N-1} g_i^{(N)} z^{-i}; \quad (49)$$

se os bancos de filtros são ortogonais,

$$g_i^{(N)} = (-1)^{i+1} h_{2N-1-i}^{(N)} \quad i = 0, 1, 2, 3, \dots, 2N - 1. \quad (50)$$

As restrições para garantir três momentos nulos serão (SHERLOCK; MONRO, 1998) (GALVÃO, 2015).

$$G^{(N)}(z) \Big|_{z=1} = 0 \quad (51)$$

$$\frac{dG^{(N)}(z)}{dz} \Big|_{z=1} = 0 \quad (52)$$

$$\frac{d^2G^{(N)}(z)}{d^2z} \Big|_{z=1} = 0 \quad (53)$$

As condições 51 , 52 e 53 garantem a existência de pelo menos três momentos nulos podendo a wavelet apresentar, inclusive mais destes momentos nulos.

A wavelet usada para realizar o processamento dos sinais das correntes de fase foi da família Daubechies. São wavelets ortogonais e tem suporte compacto que apresentam uma regularidade que aumenta com seu índice N (MONTEIRO, 2014). Neste sentido, a wavelet escolhida foi a Daubechies-4 (db4).

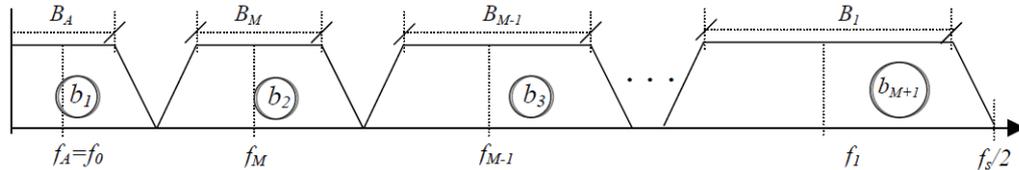
4.1.4 Determinação do nível de decomposição

O próximo passo é determinar até que nível o sinal deve ser decomposto. Um número adequado de níveis de decomposição depende na frequência de amostragem f_s do sinal a ser analisado (BOUZIDA et al., 2011). O número de sub-bandas, a ser usado na decomposição de um sinal, é escolhido de forma que a frequência fundamental do sinal, f_0 , esteja situada no centro da sub-banda das frequências mais baixas, a fim de limitar os efeitos do conteúdo espectral da fundamental nas outras sub-bandas (JAWADEKAR et al., 2014).

Para os filtros wavelet formadores da árvore wavelet packet, denotando o conjunto de sub-bandas (iniciando-se da frequência mais baixa) como $(b_1, b_2, \dots, b_M, b_{M+1})$ para um total

de M níveis, as larguras das bandas correspondentes são $(B_A, B_M, \dots, B_2, B_1)$ e os centros das frequências (equação 54) são $(f_A, f_M, \dots, f_2, f_1)$ como ilustra a Figura 32. A largura do filtro no nível k é

Figura 32 – Faixas das sub-bandas de frequência.



Fonte: (MONTEIRO, 2014)

$$B_k = \frac{f_s}{2^{k+1}} \quad k = 1, 2, 3, 4, \dots, M; \quad (54)$$

em que f_s é a frequência de amostragem do sinal, $6k Hz$.

Seja $f_0 = 60 Hz$ a frequência fundamental. Para que f_0 localize-se no centro da sub-banda b_1 , sua largura deve ser igual ao dobro da frequência fundamental, ou seja, $B_A = 2f_0$ (LIRA, 2004). Sendo assim,

$$B_A = B_M = \frac{f_s}{2^{M+1}} = 2f_0; \quad (55)$$

logo,

$$\frac{f_s}{2^{M+1}} = 2f_0 \Rightarrow \frac{f_s}{4f_0} = 2^M; \quad (56)$$

realizando os devidos algebrismos chega-se ao valor de M .

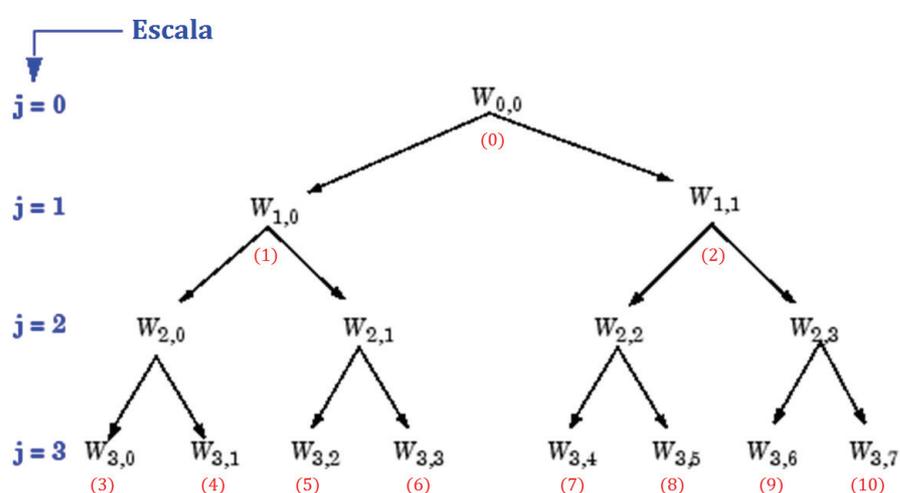
$$M = \log_2\left(\frac{f_s}{4f_0}\right) - 2; \quad (57)$$

Para os valores de f_s e f_0 utilizados, obtém-se $M = 4,64$ o que significa que o ideal é a decomposição até o quinto nível da árvore *wavelet packet*.

4.1.5 Determinação dos nós a serem escolhidos

Após a escolha da wavelet e a realização da decomposição *wavelet packet*, o que se obtém é a árvore binária *Wavelet Packet Tree* com uma estrutura já mostrada no capítulo 2. É interessante, antes de prosseguir com a metodologia, explanar sobre como os nós na árvore binária podem ser identificados. Os nós nesta estrutura seriam as saídas dos bancos de filtros. A Figura 33 mostra a estrutura da árvore com seus nós identificados por seus índices $W_{i,j}$ em que i representa o número do nó em um nível de decomposição e j representa o nível de decomposição do sinal. Os números entre parêntesis representam o número do nó.

Figura 33 – Árvore binária wavelet packet com os nós $W_{i,j}$ identificados também por seus números.



Fonte: Próprio autor

A energia do sinal é contida principalmente na parte de aproximação e uma pequena parte de detalhes. Por exemplo, a aproximação no primeiro nível contém muito mais energia do que os outros coeficientes ao mesmo nível da árvore de decomposição. Contudo, os sinais resultantes de uma operação com defeito possuem componentes de alta frequência (EKICI; YILDIRIM; POYRAZ, 2008). Por esta razão é mais distinto usar os nós de detalhe para prover informações às redes neurais.

Através da entropia de Shannon (JAIN; WADHWANI; WADHWANI,) (PIMENTA; PAVLIK, 2015) (SEERA et al., 2014), é possível determinar o quanto de informação cada um dos nós s_n da árvore binária *wavelet packet* carrega. Na metodologia aqui utilizada, os nós que apresentam os maiores valores na relação entre a entropia de Shannon e a energia é que serão utilizados.

4.2 O PROCESSO DE CLASSIFICAÇÃO DE FALHAS

Nesta seção será apresentado o classificador baseado em RNA e WPT com o uso de medidas estatísticas de sinal. No anexo A será mostrada uma segunda metodologia que foi adotada anteriormente à explanada nesta seção. Tal metodologia se mostrou eficaz no caso de identificação das condições “com defeito” e “sem defeito”. Contudo, no proceder dos ensaios e na investigação da distinção entre um número maior defeitos a eficácia da metodologia caiu consideravelmente levando à procura da abordagem apresentada aqui nesta seção.

4.3 FALHAS MECÂNICAS

A tabela 4 mostra as seis condições a que o motor está sujeito (condição sem defeito e os 5 defeitos mecânicos inseridos nos ensaios). Todas as situações foram realizadas para o motor em vazio e com carga.

Tabela 4 – Identificação defeito - código

Código da condição	Descrição
Sad	Sadio
D1	Duas barras quebradas
D2	Quatro Barras quebradas
D3	Um Rolamento danificado trilha externa
D4	Dois Rolamentos danificados Trilha externa
D5	Dois rolamentos danificados trilha interna

Nas Figuras 34 e 35 estão mostrados os dois rotores a serem usados nos ensaios já com as barras danificadas. São rotores para um motor WEG de tamanho de carcaça 63 (motor1) e com 30 barras cada. Da mesma forma na Figuras 36 e 37 exibem os rolamentos danificados usados nos ensaios.

Figura 34 – Rotor 1 com suas barras danificadas através de furação



Fonte: Próprio autor

O Rotor da Figura 34 teve quatro barras danificadas. As barras são radialmente opostas. O mesmo se repete para o eixo 2, na Figura 35, que teve duas barras radialmente opostas danificadas.

Figura 35 – Rotor 2, barras também danificadas através de processo de furação.



Fonte: Próprio autor

Figura 36 – Rolamento danificado trilha interna



Fonte: Próprio autor

Figura 37 – Rolamento danificado trilha externa



Fonte: Próprio autor

Na tabela 5 estão mostrados os nós de Detalhe da WPT que apresentaram os maiores

valores da entropia de Shannon e de energia calculada por (26).

Tabela 5 – Relação dos nós com maiores valores de energia e suas respectivas entropias para cada um dos defeitos mecânicos ensaiados.

Defeito	Nó da <i>wavelet packet tree</i>	Entropia de Shannon	Energia
D1	2	0,0341	0,1155
	4	0,0581	0,1077
	5	0,0073	0,1237
	6	0,0144	0,1126
	8	0,0161	0,0337
	12	0,0052	0,1234
	14	0,0093	0,1250
	16	0,0379	0,0354
	25	0,0041	0,1032
	30	0,0041	0,0965
D2	2	0,0291	0,1254
	4	0,0644	0,0843
	5	0,0244	0,1265
	6	0,0278	0,1543
	8	0,0477	0,0521
	10	0,0433	0,0846
	12	0,0010	0,0819
	13	0,0118	0,1174
	14	0,0116	0,0921
	16	0,0819	0,0405
	23	0,0116	0,0706
	27	0,0043	0,0726
D3	2	0,0599	0,3904
	4	0,0646	0,1989
	5	0,0080	0,2042
	6	0,0146	0,1750
	8	0,0262	0,1232
	9	0,0196	0,1077
	10	0,0428	0,0916
	11	0,0041	0,1388
	16	0,0342	0,0321
D4	2	0,2684	0,3142
	4	0,1860	0,2428
	5	0,0205	0,2825

Tabela 5 Continuação da página anterior

Defeito	Nó da <i>wavelet packet tree</i>	Entropia de Shannon	Energia
	6	0,0199	0,1547
	8	0,094	0,0405
	10	0,0418	0,1075
	12	0,0077	0,1221
	16	0,091	0,0407
	23	0,0069	0,1072
	D5	2	0,0345
4		0,0494	0,0884
5		0,0189	0,1307
6		0,0328	0,1845
8		0,0401	0,0301
9		0,0174	0,0688
10		0,032	0,0738
11		0,0085	0,1017
14		0,0154	0,1421
16		0,0592	0,0413
22		0,0256	0,0865
26		0,0077	0,0916

A tabela 5 mostra os nós 2 e 4 estando entre os nós de maiores valores de entropia e de energia, o que sugere que os detalhes de tais nós realmente sejam portadores de uma quantidade significativa de informações referentes às condições de operação do motor. Isso também pode se dever ao fato de serem nós dos primeiros níveis de decomposição e, por esta razão, levarem sempre uma quantidade significativa de informações. Os nós 8 e 10 (exceto para D1) e 16 também presentes entre os maiores valores para entropia de Shannon da árvore binária *wavelet packet*.

Os nós 2, 4, 6, 8 e 16 são os únicos que apresentam valores relevantes tanto de energia, quanto de entropia (acima de 10^{-2}) em todos os defeitos. O nó 5 se apresenta dentre os mais energéticos em todos os defeitos, porém seus valores de entropia de Shannon são sempre inferiores ao limiar de 10^{-2} . Levando, assim, a não ser considerado como nó viável, a princípio. O nó 6, apesar de superar o limiar, mostra ser dos nós que apresentam menores valores para as entropias.

A combinação da energia e entropia de Shannon dos sinais da *wavelet packet tree* pode ser usada como recurso para medir a quantidade de informações portadas pelos nós da árvore (ZHENG-YOU; XIAOQING; GUOMING, 2006). Nesse sentido, um critério viável para seleção dos nós da decomposição é a relação entre a energia e a entropia Shannon do sinal proveniente

do nó (BAFROUI; OHADI, 2014) (KANKAR; SHARMA; HARSHA, 2011), de maneira que, os nós selecionados são aqueles que apresentem os maiores valores para esta relação. A Tabela 6 exhibe os nós e as relações entre suas entropias e energias.

Tabela 6 – Relação entre valores de entropia e suas respectivas energias para cada um dos defeitos mecânicos ensaiados.

Defeito	Nó da <i>wavelet packet tree</i>	Relação <i>Entropia/Energia</i>
D1	2	0,2954
	4	0,5394
	5	0,0591
	6	0,1279
	8	0,4778
	12	0,0042
	14	0,0744
	16	1,0706
	25	0,0397
	30	0,0430
D2	2	0,2320
	4	0,7591
	5	0,1930
	6	0,1801
	8	0,9173
	10	0,5083
	12	0,0122
	13	0,1005
	14	0,1259
	16	2,0222
	23	0,1643
27	0,0592	
D3	2	0,1534
	4	0,3248
	5	0,0391
	6	0,0834
	8	0,2126
	9	0,1820
	10	0,4672
	11	0,0295
	16	1,0654
	2	0,8542

Tabela 6 Continuação da página anterior

Defeito	Nó da <i>wavelet packet tree</i>	Relação <i>Entropia/Energia</i>
D4	4	0,7660
	5	0,0725
	6	0,1286
	8	2,3210
	10	0,3888
	12	0,0631
	16	2,2359
	23	0,0643
D5	2	0,1296
	4	0,5588
	5	0,1446
	6	0,1778
	8	1,3322
	9	0,2529
	10	0,4336
	11	0,0835
	14	0,1083
	16	1,4334
	22	0,2959
	26	0,0841

A observação da Tabela 6 revela que, para D1, os nós que apresentam os maiores valores da relação *Energia/Entropia* são 2, 4, 6, 8 e 16. Para o defeito D2, os nós foram 2, 4, 5, 6, 8, 10 e 16.

O defeito D3 exibe os nós 4, 8, 9, 10 e 16, a presença do nó 9 sugere que este possa ser um nó característico do defeito D3 por este ser o único que apresenta este nó entre os de melhores relações *Energia/Entropia*. Para D4 os nós foram 2, 4, 8 e 16 e no defeito D5 os nós foram 2, 4, 8, 9, 10 e 16. O nó 16 merece destaque por apresentar em todos defeitos valores superiores a 1.

A presença dos nós 4, 8 e 16 detre os que apresentam as maiores relações *Energia/Entropia* sugere que estes, assim como o nó 2 que sempre esteve entre os de maiores valores de entropia e energia, sejam relacionados diretamente com a o tipo de máquina ensaiada (motor de indução) e à wavelet utilizada para decompor o sinal.

São dez combinações possíveis, i.e. $\binom{5}{3}$ de medidas estatísticas. Estas são mostradas na Tabela 7.

Além destas dez combinações das medidas estatísticas, será analisado o desempenho

Tabela 7 – Combinações de medidas estatísticas.

Número da combinação	Medidas estatísticas combinadas
1	RMS ; curtose ; Desvio padrão
2	RMS ; curtose ; Média
3	RMS ; curtose ; centroide
4	RMS ; Média ; Desvio padrão
5	RMS ; Média ; centroide
6	RMS ; Desvio padrão ; centroide
7	desvio padrão ; centroide ; Média
8	desvio padrão ; centroide ; curtose
9	desvio padrão ; curtose ; Média
10	centroide ; curtose ; Média

do centroide sozinho com o propósito de analisar o potencial deste atributo. No Capítulo 5, todas estas combinações serão aplicadas a treinamento e serão avaliadas com o propósito de se determinar qual a tem um melhor desempenho geral.

Inicialmente, a topologia MLP aqui utilizada será disposta com quatro camadas ocultas. Estas serão organizadas em cinco combinações apresentadas na Tabela 8. Uma vez determinada, nos resultados experimentais, qual combinação de atributos estatísticos tem melhor desempenho (ou seja, tenha taxas de acertos altas para todas as situações defeituosas), esta combinação será apresentada a redes mais simples (com apenas uma camada oculta) afim de mostrar a eficiência da metodologia ante outras configurações de rede MLP.

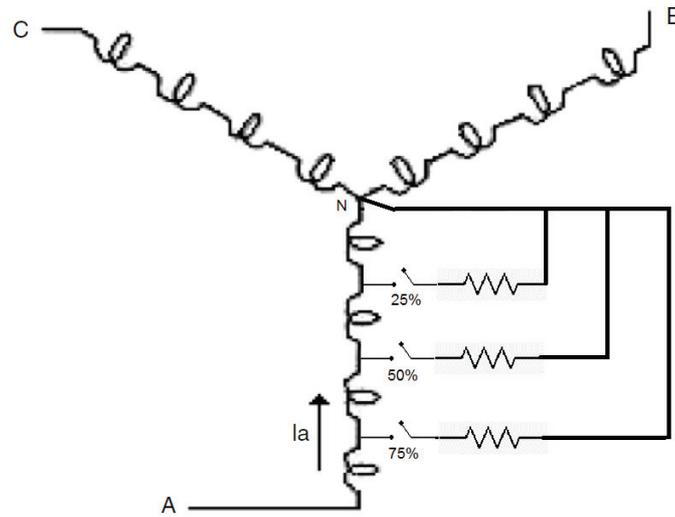
Tabela 8 – Padrões adotados para distribuição dos neurônios nas camadas ocultas da rede MLP.

Configuração	Disposição dos neurônios
A	[9, 9, 9, 9]
B	[9, 10, 10, 9]
C	[10, 10, 10, 9]
D	[10, 10, 10, 10]
E	[8, 8, 8, 8]

4.4 FALHAS ELÉTRICAS

Para realizar os ensaios elétricos de faltas incipientes nas bobinas de estator foram necessárias adaptações no motor testado. O motor foi rebobinado de maneira que apresentasse terminais disponíveis na sua parte externa referentes a derivações de 25% , 50% e 75% dos enrolamentos das fases A, B e C. A Figura 38 mostra como estas derivações foram realizadas exemplificando para a fase A.

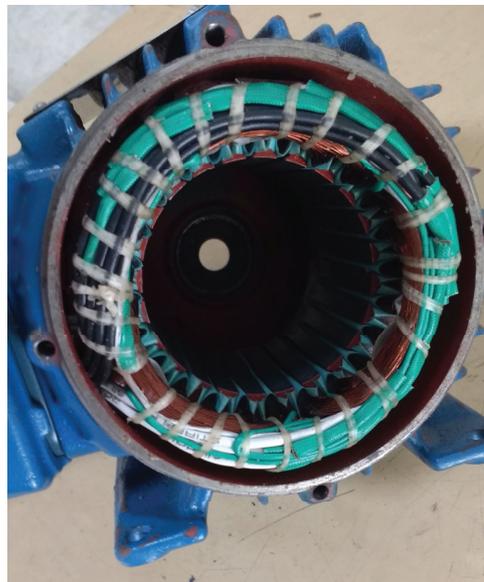
Figura 38 – Derivações do enrolamento da fase A



Fonte: Próprio autor

A Figura 39 mostra o motor já rebobinado e com as derivações já realizadas

Figura 39 – Motor rebobinado com as derivações realizadas

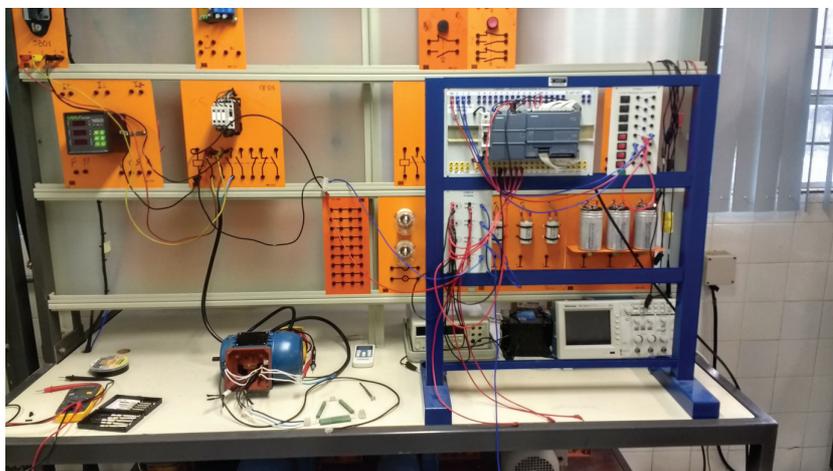


Fonte: Próprio autor

As chaves representadas na Figura 38 foram controladas via controlador lógico programável (CLP). Sendo cada uma das derivações atuantes por um intervalo de tempo de 5 minutos. Dentro deste intervalo de tempo a chave trabalha a intervalos de 5 segundos (2 segundos de chave fechada e 3 de chave aberta). A Figura 40 mostra a bancada montada com o motor

acionado via partida direta e o CLP que realizou o processo de chaveamentos das derivações das fases. Na Figura 41 está mostrado o detalhe do CLP utilizado.

Figura 40 – Montagem da Bancada para realização dos ensaios de faltas incipientes.



Fonte: Próprio autor

Figura 41 – Detalhe do CLP SIEMENS SIMATIC S7 - 1200.



Fonte: Próprio autor

A metodologia de tratamento do sinal das correntes de fase e treinamento da rede neural é a mesma utilizada para a detecção de falhas mecânicas.

A Tabela 9 mostra os valores de energia e entropia para os nós da WPT da decomposição das correntes de faltas incipientes no estator.

Tabela 9 – Relação dos nós com maiores valores de energia e suas respectivas entropias para faltas incipientes no estator.

Defeito	Nó da <i>wavelet packet tree</i>	Entropia de Shannon	Energia
25%	2	0,0012	0,1262
	4	0,0009	0,1578
	5	0,00057	0,1442
	8	0,0028	0,0864
	9	0,0017	0,1305
	10	0,0045	0,1355
	11	0,00034	0,1301
	14	0,00031	0,0883
	16	0,0320	0,0983
	40	0,00086	0,0895
50%	2	0,0010	0,2318
	4	0,0039	0,0685
	5	0,00053	0,0955
	6	0,00054	0,1171
	8	0,0334	0,0230
	13	0,0003	0,0758
	14	0,00027	0,0942
	16	0,0885	0,0342
	25	0,00015	0,1012
	33	0,045	0,0800
	52	$6,79 \cdot 10^{-5}$	0,0728
75%	2	0,0012	0,1514
	4	0,0049	0,1024
	5	0,00053	0,1299
	6	0,00054	0,1018
	8	0,0358	0,1057
	11	0,00026	0,1090
	12	0,00029	0,0937
	14	0,00025	0,1292
	16	0,0969	0,1016
	29	0,00018	0,1010

Da mesma forma que para as falhas mecânicas, será aqui feita a observação das entropias e das energias dos nós para cada defeito. No caso das faltas a 25% do enrolamento, os 2, 4, 8, 9, 10 e 16 apresentaram os maiores valores de entropia e energia. Vale destaque ao nó 5 que,

mesmo apresentando o maior valor de energia, tem uma entropia muito pequena em comparação aos nós anteriormente citados.

Observando para as falhas a 50%, os nós que apresentam bons valores de energia e entropia são 2, 8, 16 e 33. O nó 4 também apresenta valores acima de um limiar de 10^{-3} para entropia e 10^{-2} para energia.

Nas faltas a 75%, apenas os nós 2, 4, 8 e 16 tiveram bons valores de entropia e energia. O que se nota é, mais uma vez a presença dos nós 2, 4, 8 e 16 dentre os mais energéticos. A Tabela 10.

Tabela 10 – Relação entre valores de entropia e suas respectivas energias para cada uma das faltas ensaiadas no estator.

Defeito	Nó da <i>wavelet packet tree</i>	Relação <i>Entropia/Energia</i>
25%	2	0,0095
	4	0,0057
	5	0,0039
	8	0,0324
	9	0,0131
	10	0,0332
	11	0,0026
	14	0,0035
	16	0,3255
	40	0,0096
50%	2	0,0043
	4	0,0569
	5	0,0055
	6	0,0043
	8	1,4522
	13	0,0395
	14	0,0028
	16	2,5877
	25	0,0015
	33	0,5625
52	0,00093	
75%	2	0,0079
	4	0,7485
	5	0,0041
	6	0,0053
	8	0,3387
	11	0,0024

Tabela 10 Continuação da página anterior

Defeito	Nó da <i>wavelet packet tree</i>	Relação <i>Entropia/Energia</i>
	12	0,0031
	14	0,0019
	16	0,9537
	29	0,0018

As maiores relações entre energia e entropia para faltas a 25% estão são as dos nós 2, 8, 9, 10 e 16. Para 50%, são 4, 8, 13, 16 e 33. No caso das faltas a 75% a relação é maior em 2, 4, 8 e 16 apenas, que evidencia os nós 2, 4, 8 e 16 realmente como os principais portadores de informação para treinamento da rede.

No intuito de averiguar se tal fato está relacionado com wavelet selecionada, foi feita a decomposição usando a wavelet de Haar para o defeito mecânico D1. Os nós com maiores valores de energia e entropia foram 6, 10, 12, 17, 18, 23, 31, 32, 33, 34, 39 e 40 exibindo todos uma entropia acima do limiar estabelecido na metodologia aqui apresentada (10^{-2} para falhas mecânicas), sugerindo que existe uma relação com o tipo de wavelet escolhida para que se extraia informações que ainda possam existir em outros nós.

4.5 CONCLUSÃO

O presente capítulo apresentou a metodologia abordada nesta tese para realização dos ensaios e coleta dos dados. Foi apresentada a maneira como se escolheu a wavelet para analisar o sinal (db4) e como se estabeleceu até qual nível se realizou a decomposição dos sinais de corrente para detecção dos problemas elétricos e mecânicos. A topologia de rede adotada foi MLP com quatro camadas ocultas, cinco configurações de rede com as ditas quatro camadas ocultas foram utilizadas, além das dez combinações possíveis de atributos estatísticos para serem formadores dos vetores de entrada das redes. O próximo capítulo trata dos resultados experimentais da aplicação da metodologia aqui proposta.

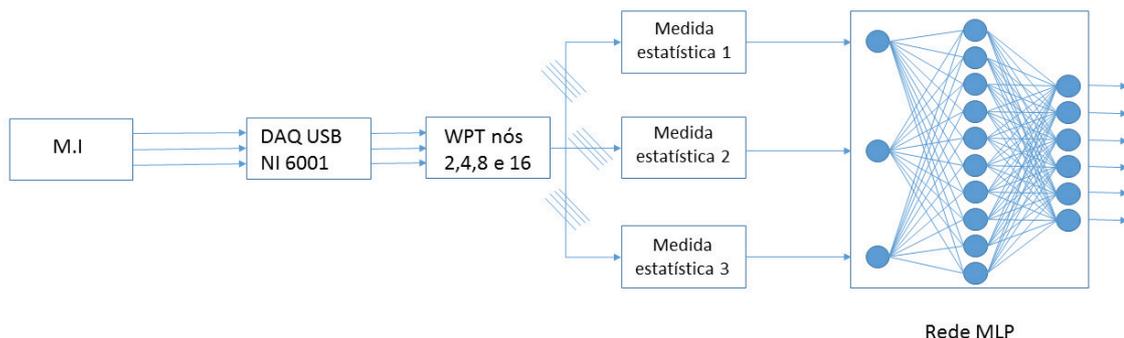
5 RESULTADOS EXPERIMENTAIS

Neste capítulo são exibidos os resultados dos ensaios, os acertos globais para todas as combinações de atributos e configurações de rede e serão desenvolvidas observações quanto ao desempenho dos métodos SCG e Rprop, seus tempos de treinamento e comparativos entre eles.

5.1 RESULTADOS DE CLASSIFICAÇÃO PARA AS FALHAS MECÂNICAS

Nesta sessão serão mostradas as taxas de acertos globais para cada uma das combinações de atributos em cada uma das configurações de rede propostas. Também é realizado um comparativo entre os dois métodos usados para treinamento das redes: *Resilient Backpropagation* e gradientes conjugados escalonados. A Figura 42 mostra o diagrama do processo geral de análise e separação dos dados e classificação dos mesmos.

Figura 42 – Esquemático do processo de obtenção de dados e classificação.



Fonte: Próprio autor

5.1.1 rede MLP treinada pelo método SCG com a combinação 1 de medidas estatísticas

Na primeira análise de resultados, será posta uma tabela comparativa para cinco configurações possíveis de redes MLP. Todas apresentam quatro camadas ocultas e seis saídas (Sadio, D1, D2, D3, D4, e D5). O valor “1” presente em uma dada saída representa o defeito em questão ao passo que as demais saídas apresentam valor 0. As medidas estatísticas utilizadas neste momento foram curtose, valor RMS e desvio padrão.

A Tabela 11 apresenta um comparativo entre os acertos globais das cinco configurações com melhores desempenhos.

Tabela 11 – Padrões adotados para distribuição dos neurônios na rede MLP e seus respectivos acertos globais para a combinação 1 e rede treinada a SCG.

Configuração	Taxa de acerto global (SCG)
A	96,2%
B	96,8%
C	94,5%
D	96,4%
E	93,9%

A configuração B apresentou um desempenho ligeiramente superior. Contudo, todas as configurações apresentaram boas taxas de acerto global.

5.1.2 Treinamento da rede MLP através do *Resilient Backpropagation* da combinação 1

A seguir, na tabela 12, são mostrados os comparativos das taxas de acertos para as cinco configurações de rede agora treinadas com a combinação de atributos 1 através do método *Resilient Backpropagation*.

Tabela 12 – Taxas de acertos globais (TAG) da rede MLP treinada através do *Resilient Backpropagation* e combinação de medidas estatísticas 1.

Configuração	TAG (Rprop)
A	93,7%
B	94,5%
C	94,0%
D	94,5%
E	93,3%

O que se pode perceber é que as taxas de acertos globais para o método *resilient backpropagation* são um pouco menores. Contudo, este método se mostra interessante devido ao fato de suas iterações serem muito menos onerosas computacionalmente. Ao se realizar uma comparação do tempo médio de treinamento de uma rede neural treinada aos dois métodos o tempo é menor quando a rede é treinada a *resilient backpropagation* (Rprop). Ainda sobre a Tabela 12, as configurações B e D apresentaram a mesma taxa de acertos e a configuração C fica em segundo lugar. O que se observa diante das duas tabelas é que as configurações B e D para ambos os métodos de treinamento apresentaram os melhores desempenhos quando do uso da combinação de atributos 1.

A tabela 13 faz um comparativo entre o número médio de iterações e o tempo médio de treinamento para uma mesma configuração (neste caso a configuração 5 em particular).

Tabela 13 – Comparativo de tempos de treinamento e número de iterações de uma mesma configuração da RNA para Rprop e SCG.

Método de treino	Tempo de treinamento	Número de Iterações
Rprop	9 segundos	636
SCG	18 segundos	711

As Tabelas 14 a 23 mostram as taxas de acertos globais para as demais configurações.

Tabela 14 – Taxas de acertos globais da rede MLP para combinação de medidas estatísticas 2.

Configuração	TAG (SCG)	TAG (Rprop)
A	94,4%	93,0%
B	93,1%	88,5%
C	94,3%	92,2%
D	87,6%	88,2%
E	92,7%	90,8%

Da Tabela 14 para a combinação de atributos estatísticos 2, observa-se que a configuração A é superior em ambos os métodos de treinamento. Em segundo lugar, vem a configuração C que, para o método SCG, apresenta um desempenho um pouco inferior à configuração A em ambos os métodos. A configuração B não apresenta o terceiro melhor desempenho para o caso do treinamento a Rprop, contudo as configurações que apresentam desempenho superior à B no treinamento a Rprop são todas de desempenho inferior a tal configuração quando se trata do treinamento a SCG.

Tabela 15 – Taxas de acertos globais da rede MLP para combinação de medidas estatísticas 3.

Configuração	TAG (SCG)	TAG (Rprop)
A	94,4%	90,7%
B	95,6%	90,7%
C	95,4%	89,9%
D	92,1%	92,2%
E	92,7%	92,9%

Para a combinação de medidas estatísticas 3 na Tabela 15, o melhor desempenho geral é da configuração B, pois apresenta o melhor desempenho a SCG e um desempenho de 90,7% a Rprop. A configuração C apresenta um desempenho um pouco menor que a B a SCG, contudo a Rprop esta configuração apresenta o pior desempenho entre todas. A configuração E apresenta o melhor desempenho a Rprop e igualmente bom a SCG, ambos próximos de 93%.

Tabela 16 – Taxas de acertos globais da rede MLP para combinação de medidas estatísticas 4.

Configuração	TAG (SCG)	TAG (Rprop)
A	91,1%	92,0%
B	92,6%	88,6%
C	91,7%	90,4%
D	92,1%	91,7%
E	80,7%	89,0%

Para o método SCG, as configurações B,D e C da Tabela 16, que evidencia a combinação 4, apresentam os melhores desempenhos necessariamente nessa ordem. Já para o método Rprop os três melhores desempenhos são, das configurações A, D e C. O que evidencia as configurações D e C com bons desempenhos em ambos os métodos e sendo a configuração D a de melhor desempenho.

Tabela 17 – Taxas de acertos globais da rede MLP para combinação de medidas estatísticas 5.

Configuração	TAG (SCG)	TAG (Rprop)
A	95,2%	83,6%
B	91,4%	89,3%
C	93,6%	87,0%
D	94,6%	87,9%
E	93,4%	83,3%

Para a combinação de medidas 5, exibida na Tabela 17, o desempenho de todas as configurações para o Rprop foi inferior a 90%, o que sugere que tal combinação não seja de fato tão eficaz, dado o fato de a diferença entre os resultados ser, em apenas uma das configurações, inferior a 2% e em todas as outras superior a 6%. A SCG, o melhor desempenho ficou por conta da configuração A.

Tabela 18 – Taxas de acertos globais da rede MLP para combinação de medidas estatísticas 6.

Configuração	TAG (SCG)	TAG (Rprop)
A	93,3%	95,4%
B	94,2%	92,7%
C	95,3%	92,7%
D	94,8%	91,8%
E	94,7%	92,0%

A Tabela 18 apresenta as taxas de acertos globais para a combinação de atributos 6. As configurações C,D e E apresentam os melhores desempenhos para o método SCG. As

configurações A, B e C mostram os melhores desempenhos para o Rprop. O que se conclui, aqui, é que a configuração C, para esta combinação de atributos dá o melhor desempenho de forma geral, por ser a melhor taxa de acertos do SCG e a segunda melhor do Rprop.

Tabela 19 – Taxas de acertos globais da rede MLP para combinação de medidas estatísticas 7.

Configuração	TAG (SCG)	TAG (Rprop)
A	92,0%	91,7%
B	93,5%	90,8%
C	94,5%	90,5%
D	94,6%	92,0%
E	94,4%	89,5%

A combinação de medidas estatísticas 7, Tabela 19, tem como melhores configurações D,C e E para SCG. A configuração D é, também, a de melhor desempenho para Rprop.

Tabela 20 – Taxas de acertos globais da rede MLP para combinação de medidas estatísticas 8.

Configuração	TAG (SCG)	TAG (Rprop)
A	85,9%	94,4%
B	96,5%	93,9%
C	95,4%	94,1%
D	94,8%	94,0%
E	95,2%	92,7%

Com exceção da combinação A para SCG, a combinação de medidas estatísticas 8, apresentada na Tabela 20, foi a que apresentou os melhores resultados para ambos os métodos de treinamento. Para o método de treinamento SCG, as melhores configurações são B,C e E. Já para o Rprop, as melhores são A, C e D. O que se observa é que a configuração C é mais interessante por ser a segunda melhor em ambos os métodos de treinamento.

Tabela 21 – Taxas de acertos globais da rede MLP para combinação de medidas estatísticas 9.

Configuração	TAG (SCG)	TAG (Rprop)
A	88,4%	87,4%
B	88,1%	84,5%
C	47,0%	86,4%
D	87,6%	87,7%
E	89,4%	86,8%

É interessante frisar, a partir da observação da Tabela 21, que a combinação de medidas estatísticas 9, além de apresentar um desempenho inferior ao apresentado pelas combinações anteriores, também apresenta em sua configuração C para o SCG o pior desempenho de todas as combinações estatísticas abordadas neste trabalho. O mais provável é que tal valor seja apenas um *outlier* (ponto que está muito distante das demais observações em uma série estatística) e não seja realmente condizente com o desempenho real da configuração, haja visto o fato de que a mesma teve um desempenho de 86,4% quando treinada a outro método.

Tabela 22 – Taxas de acertos globais da rede MLP para combinação de medidas estatísticas 10.

Configuração	TAG (SCG)	TAG (Rprop)
A	93,9%	94,2%
B	94,2%	93,5%
C	95,2%	94,1%
D	94,9%	91,0%
E	95,3%	92,1%

Para a Tabela 22, as configurações E, C e D para o SCG apresentam melhores resultados, para o Rprop são as configurações A, C e B que são as melhor desempenho. Mais uma vez, a configuração C mostra ter bom desempenho em todos os métodos.

Tabela 23 – Taxas de acertos globais da rede MLP para o centroide

Configuração	TAG (SCG)	TAG (Rprop)
A	84,4%	80,0%
B	82,9%	81,6%
C	82,6%	76,6%
D	84,9%	77,4%
E	83,1%	77,4%

A Tabela 23 vem com o propósito de mostrar o desempenho do centroide como atributo isolado para classificação de sinais. Este atributo já vem sendo utilizado para detecção de problemas em equipamentos presentes em linhas de transmissão (SANTOS, 2008). Muito embora exista uma diferença considerável entre os desempenhos para os dois métodos, o centroide mostra-se bom classificador para SCG com taxas de acertos superiores a 82%.

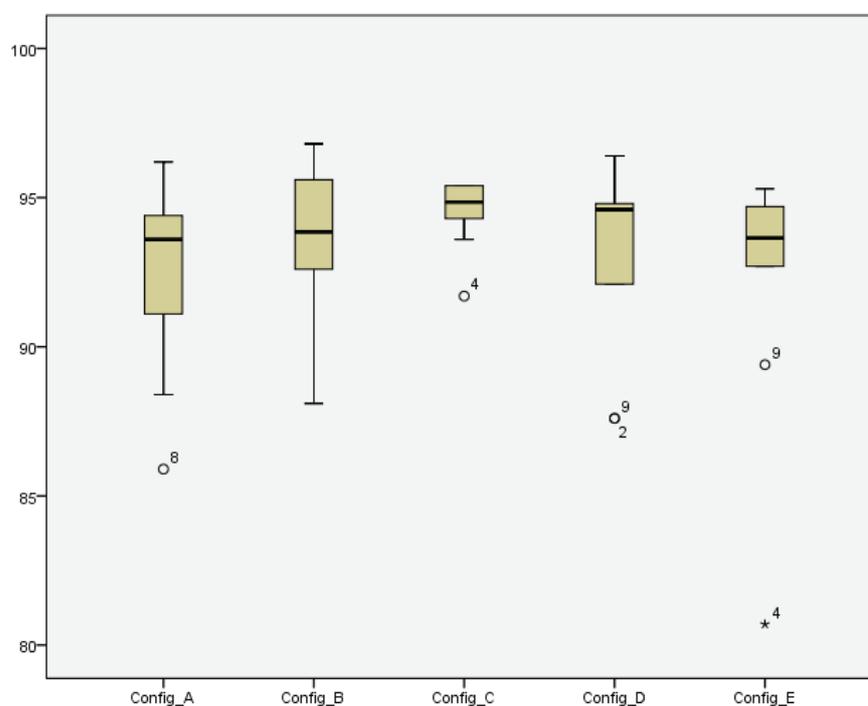
Para uma melhor avaliação de quais seriam as melhores configurações para a classificação de padrões, a Tabela 24 e os *box plots* da Figura 43 foram construídos. Vale lembrar da existência de um ponto *outstanding* na configuração C e combinação de atributos 9. Este ponto teve uma taxa de acertos muito baixa (47%).

Tabela 24 – Estatísticas descritivas para as taxas de acertos globais das configurações de redes treinadas a SCG.

	Configuração A	Configuração B	Configuração C	Configuração D	Configuração E
Média	92,48	93,60	89,69	92,95	92,24
Mediana	93,60	93,85	94,50	94,60	93,65
Desvio Padrão	3,22	2,57	15,04	3,10	4,40
Mínimo	85,90	88,10	47,00	87,60	80,70
Máximo	96,20	96,80	95,40	96,40	95,30
Primeiro Quartil	90,42	92,30	93,12	90,97	91,87
Terceiro Quartil	94,60	95,82	95,32	94,82	94,82

Pela observação da Tabela 24, as três configurações que tiveram maiores médias para o SCG foram B, D e A. Vale ressaltar que a configuração C teve sua média diminuída em virtude do *outstanding* na combinação de atributos 9. se tal ponto for desprezado, a média desta configuração sobe para 94,43% fazendo da configuração C a de melhor desempenho junto com B e D.

Figura 43 – *Box plots* das estatísticas descritivas para as configurações de redes treinadas pelo método SCG para detecção de falhas mecânicas.

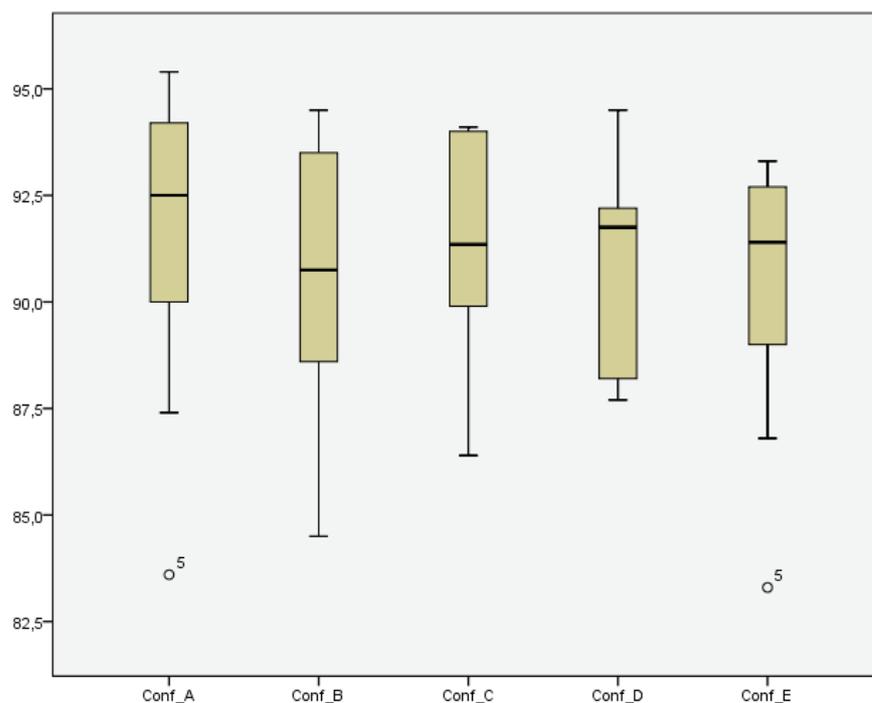


Fonte: Próprio autor

A seguir serão indicadas também as estatísticas descritivas, Tabela 25 e os *box plots* da Figura 44 para as cinco configurações para as redes treinadas a Rprop.

Tabela 25 – Estatísticas descritivas para as cinco configurações treinadas a Rprop.

	Configuração A	Configuração B	Configuração C	Configuração D	Configuração E
Média	91,54	90,70	91,13	91,10	90,24
Mediana	92,50	90,75	91,35	91,75	91,40
Desvio Padrão	83,60	84,50	86,40	87,70	83,30
Mínimo	83,60	84,50	86,40	87,70	83,30
Máximo	95,40	94,50	94,10	94,50	93,30
Primeiro Quartil	89,35	88,57	89,17	88,12	88,45
Terceiro Quartil	94,25	93,60	94,02	92,65	92,75

Figura 44 – *Box plots* das estatísticas descritivas para as configurações de redes treinadas pelo método Rprop para falhas mecânicas.

Fonte: Próprio autor

Para o método Rprop, o que se observa da Tabela 25, é que as configurações A, C e D tiveram as maiores médias. É interessante mostrar o fato de que estas três configurações tiveram as melhores médias em ambos os métodos de treinamento. Levando em consideração que a mais simples destas é a configuração A, apresentando 9 neurônios em cada uma das suas camadas ocultas, esta é provavelmente preferível diante das outras quatro configurações. Cabendo a observação do desempenho desta configuração ante as falhas elétricas para se confirmar.

5.1.3 Comparativo com treinamento de rede a partir de parâmetros provenientes da decomposição *wavelet packet* para falhas mecânicas

No intuito de mostrar a viabilidade da metodologia aqui empregada, será apresentado o resultado e o treinamento da rede neural com dados diretamente da decomposição *wavelet packet*. Na Tabela 26, são exibidos os resultados do treinamento a Rprop das cinco configurações de rede tendo como dados formadores do vetor de entrada os valores das energias dos nós 2, 4, 8 e 16.

Tabela 26 – Percentuais de acertos para cada configuração de rede com vetores de entradas formados pelas energias dos nós 2, 4, 8 e 16 da WPT treinadas a Rprop.

	A	B	C	D	E
Percentagem Classificação correta	51,88%	55,26%	54,42%	54,65%	51,97%
Percentagem Classificação Incorreta	48,11%	44,73%	45,57%	45,34%	48,02%
Maior erro do teste	10,16%	10,04%	10,03%	9,88%	10,42%
Maior erro de validação	9,82%	9,97%	10,00%	10,11%	10,40%
Maior erro Global	9,93%	10,05%	9,74%	9,84%	10,14%

Todos os resultados foram inferiores a 56%. A combinação de mais de um atributo estatístico fornece um apanhado de informações mais completo e por essa razão permitem à rede maior desempenho. Apenas a energia como atributo não foi suficiente para criar um conjunto de dados que propicie um bom percentual de acertos.

O resultado do treinamento a SCG é apresentado na Tabela 27. Igualmente ao treinamento com Rprop, o treinamento a SCG não teve taxas de acertos superiores a 56%.

Tabela 27 – Percentuais de acertos para o treinamento das redes neurais alimentadas com as energias dos nós 2, 4, 8 e 16 da WPT treinadas a SCG.

	A	B	C	D	E
Percentagem Classificação correta	57,61%	55,64%	54,56%	55,26%	54,65%
Percentagem Classificação Incorreta	42,38%	44,35%	45,43%	44,73%	45,34%
Maior erro do teste	9,47%	9,81%	9,80%	9,80%	9,97%
Maior erro de validação	9,66%	10,26%	9,85%	9,75%	9,79%
Maior erro Global	9,45%	9,67%	9,80%	9,46%	9,78%

5.2 RESULTADOS EXPERIMENTAIS PARA AS FALHAS ELÉTRICAS

Da mesma forma que foi apresentado para as falhas mecânicas, serão apresentados os percentuais de acertos globais nas Tabelas 28 a 38 para as dez combinações possíveis de medidas estatísticas nas cinco configurações de rede abordadas.

Tabela 28 – Acertos globais para as faltas incipientes de estator para combinação de medidas estatísticas 1.

	Falha a 25%		Falha a 50%		Falha a 75%	
	SCG	Rprop	SCG	Rprop	SCG	Rprop
configuração A	99,00%	98,80%	88,20%	91,16%	93,83%	94,13%
configuração B	99,00%	98,90%	90,04%	92,05%	94,36%	92,70%
configuração C	99,40%	99,10%	99,05%	90,81%	95,02%	93,18%
configuração D	99,20%	98,80%	94,13%	92,59%	93,47%	93,89%
configuração E	99,00%	98,70%	95,02%	87,90%	93,89%	93,77%

A Tabela 28 exhibe os resultados de treinamento para a combinação de parâmetros estatísticos 1 (RMS, curtose e desvio padrão). Dois destes parâmetros são referentes à dispersão dos dados (RMS e desvio padrão), enquanto o terceiro parâmetro (curtose) refere-se a assimetria da distribuição dos dados. Serão comentadas aqui as estruturas com melhor desempenho. Nesse sentido, analisando cada um dos percentuais de falha presentes na Tabela 28 isoladamente, tem-se as configurações C,D e B com o melhor desempenho para o caso de falhas a 25% do enrolamento. A configuração C apresenta o melhor desempenho para ambos os métodos de treinamento.

Para o caso de 50%, também se tem as configurações C, D e B com os desempenhos mais relevantes. Embora a configuração E apresente um desempenho de 95% para o SCG, seu desempenho para o Rprop é o de menor valor para este caso.

Para o caso de 75%, as configurações C, D e B ainda apresentam bons desempenhos, entretanto a configuração A se mostra ligeiramente superior à D. Contudo, vale ressaltar que a configuração D se mostrou superior à A para os dois casos anteriores o que leva a crer que, mesmo tendo um desempenho inferior à A, D seja preferível por realizar bons resultados conjuntamente com C e B para todos os três casos.

Tabela 29 – Acertos globais para as faltas incipientes de estator para combinação de medidas estatísticas 2.

	Falha a 25%		Falha a 50%		Falha a 75%	
	SCG	Rprop	SCG	Rprop	SCG	Rprop
configuração A	98,60%	98,40%	91,52%	82,92%	93,53%	90,81%
configuração B	99,00%	99,10%	96,50%	81,09%	93,89%	92,17%
configuração C	98,80%	98,90%	94,19%	89,44%	93,77%	93,71%
configuração D	98,60%	98,90%	98,28%	95,90%	93,89%	92,23%
configuração E	98,70%	99,00%	91,52%	87,67%	93,06%	92,29%

Os resultados apresentados na Tabela 29 mostram que, para falhas a 25%, as

configurações B,C e E apresentam os melhores desempenhos para ambos os métodos. É importante frisar que a configuração E é superior à D apenas ligeiramente (0,1% para ambos os métodos).

Para 50%, a configuração D tem o melhor desempenho seguida por B e C para o SCG. Para o Rprop, a configuração D também é a de melhor resultado seguida por C, porém B tem o pior resultado para treinamento via Rprop.

Nos resultados a 75%, constata-se que, também D, B e C são as de melhor desempenho. Levando em consideração que a configuração E é superior à D com muito pouca diferença no primeiro primeiro caso (25%), e que nos outros dois casos as configurações D, B e C foram as melhores, mais uma vez estas três configurações se mostram preferíveis para a combinação de medidas 2. Esta combinação da Tabela 29 usa de uma medida de dispersão (RMS), uma de assimetria (curtose) e uma medida de tendência central (Média).

Tabela 30 – Acertos globais para as faltas incipientes de estator para combinação de medidas estatísticas 3.

	Falha a 25%		Falha a 50%		Falha a 75%	
	SCG	Rprop	SCG	Rprop	SCG	Rprop
configuração A	98,20%	98,10%	99,46%	97,80%	93,24%	92,17%
configuração B	98,10%	98,90%	98,63%	97,09%	92,23%	94,30%
configuração C	99,00%	98,50%	97,68%	97,92%	94,13%	93,12%
configuração D	98,90%	98,90%	98,04%	96,68%	94,01%	92,23%
configuração E	99,20%	98,90%	98,34%	97,56%	93,59%	93,06%

O uso do centroide como medida de tendencia central, em lugar da média, fez com que a combinação 3, Tabela 30, tenha desempenho melhorado em relação à 2. A combinação 3 apresentou taxas de acertos superiores a 92% em todos os casos para todas as configurações. Destaque para a configuração E que se mostrou entre as de desempenho superior em todos os casos. As configurações C e D se mostraram entre os melhores desempenhos para 25% e 75%.

Para o caso de faltas a 50%, todas as configurações, em ambos os métodos de treinamento, apresentaram desempenhos superiores a 96%.

Tabela 31 – Acertos globais para as faltas incipientes de estator para combinação de medidas estatísticas 4.

	Falha a 25%		Falha a 50%		Falha a 75%	
	SCG	Rprop	SCG	Rprop	SCG	Rprop
configuração A	97,80%	96,70%	98,87%	88,73%	88,67%	88,26%
configuração B	97,80%	96,90%	98,45%	94,01%	90,21%	89,27%
configuração C	98,00%	96,50%	98,63%	87,67%	88,79%	87,67%
configuração D	97,70%	97,10%	98,81%	90,04%	89,50%	87,49%
configuração E	97,60%	97,60%	97,86%	79,96%	87,78%	88,44%

A Tabela 31 mostra uma combinação com duas medidas de dispersão (RMS e Desvio Padrão) e uma medida de tendência central (média). O destaque fica para o baixo desempenho da configuração E no caso de falhas a 50% treinado a Rprop.

Os melhores resultados no caso de faltas a 25% são das configurações C,B,A e D que é inferior à A no treinamento a SCG, porém tem desempenho superior no treinamento a Rprop. A 50%, a configuração A apresenta o melhor desempenho a SCG e o pior a Rprop. Diante disso, a configuração B se faz preferível à A pois apresenta bons resultados em ambos os métodos sendo, inclusive o melhor desempenho a Rprop. Com base nisso as três melhores configurações são, também para esta combinação, B, D e C que também se mostram as melhores no caso de 75%.

Na Combinação de mediadas estatísticas 5 (RMS, média e centroide), cujos resultados estão presentes na Tabela 32, nenhum dos resultados para detecção a 75% apresenta desempenho superior a 90%. Para as falhas elétricas a 25%, os melhores resultados estão para as redes A, E e B. Para 50%, as configurações A e E apresentam os dois melhores desempenhos. Se for levado em consideração apenas o método SCG, a configuração C tem o terceiro melhor desempenho, contudo esta é a configuração que apresenta o pior desempenho a Rprop sendo mais interessante a configuração D que tem um desempenho a Rprop superior à C.

Tabela 32 – Acertos globais para as faltas incipientes de estator para combinação de medidas estatísticas 5.

	Falha a 25%		Falha a 50%		Falha a 75%	
	SCG	Rprop	SCG	Rprop	SCG	Rprop
configuração A	97,80%	96,80%	97,45%	93,06%	88,08%	89,62%
configuração B	97,70%	97,60%	93,36%	90,81%	88,32%	87,49%
configuração C	97,30%	96,40%	95,31%	88,44%	89,27%	89,09%
configuração D	97,50%	96,80%	94,54%	91,46%	89,98%	89,62%
configuração E	97,80%	96,60%	96,50%	91,34%	88,55%	88,08%

Para a combinação 6 (RMS, desvio padrão e centroide) apresentada na Tabela 33, observa-

se a substituição da média como medida de tendência central pelo centroide em comparação à configuração 4. O que se observa é uma ligeira melhoria no desempenho para falhas a 75%. As configurações de destaque são C,B e E para 25%, para 50% D,C e B e para 75% E,D e C.

Tabela 33 – Acertos globais para as faltas incipientes de estator para combinação de medidas estatísticas 6.

	Falha a 25%		Falha a 50%		Falha a 75%	
	SCG	Rprop	SCG	Rprop	SCG	Rprop
configuração A	97,40%	97,10%	95,79%	92,82%	90,27%	89,56%
configuração B	98,00%	97,10%	97,33%	92,59%	90,10%	88,85%
configuração C	98,20%	97,00%	97,62%	94,01%	90,51%	90,33%
configuração D	97,10%	97,40%	97,92%	93,36%	90,93%	88,20%
configuração E	97,70%	97,00%	95,73%	89,38%	92,11%	88,61%

Tabela 34 – Acertos globais para as faltas incipientes de estator para combinação de medidas estatísticas 7.

	Falha a 25%		Falha a 50%		Falha a 75%	
	SCG	Rprop	SCG	Rprop	SCG	Rprop
configuração A	98,10%	97,20%	96,85%	91,34%	89,86%	89,33%
configuração B	97,60%	96,00%	95,25%	90,93%	88,61%	88,44%
configuração C	97,20%	96,60%	97,21%	85,35%	90,51%	88,79%
configuração D	97,80%	98,00%	96,44%	88,14%	89,50%	88,02%
configuração E	97,70%	97,10%	97,39%	87,72%	89,09%	85,83%

A Tabela 34 apresenta uma mudança na medida de dispersão em comparação à combinação 5. A combinação 7 aborda desvio padrão, centroide e média. De maneira geral, o desempenho não apresentou melhoras significativas quando comparado à combinação 5. Apenas uma configuração teve seu desempenho superior a 90% de acertos para as faltas em 75% dos enrolamentos de estator. Da mesma forma que na combinação 5, as melhores configurações para 50 e 75% foram E,C e A e C,D e E respectivamente. Destaque novamente para a configuração E que esteve em todos os casos com o melhor desempenho a SCG. Contudo, a Rprop esta configuração apresentou o pior desempenho no caso de 75%.

Tabela 35 – Acertos globais para as faltas incipientes de estator para combinação de medidas estatísticas 8.

	Falha a 25%		Falha a 50%		Falha a 75%	
	SCG	Rprop	SCG	Rprop	SCG	Rprop
configuração A	99,10%	98,20%	98,93%	96,80%	94,25%	93,71%
configuração B	98,20%	98,80%	98,75%	95,91%	92,88%	91,16%
configuração C	98,00%	98,40%	98,63%	96,00%	93,77%	91,28%
configuração D	99,10%	97,90%	98,63%	97,30%	94,13%	92,88%
configuração E	99,20%	98,70%	98,39%	96,80%	94,42%	91,22%

A combinação 8 da Tabela 35 apresenta uma medida de tendência central (centroide), uma medida de dispersão (desvio padrão) e a curtose. Em comparação à combinação de parâmetros 3 que difere na medida de tendência central, o desempenho não apresenta grandes diferenças, apenas uma redução nas taxas de acertos do treinamento a Rprop para falhas a 75%. Para falhas a 25%, os melhores resultados são das combinações E,C e A. Para 50%, são B,D e C e para 75% E, A e D.

Tabela 36 – Acertos globais para as faltas incipientes de estator para combinação de medidas estatísticas 9.

	Falha a 25%		Falha a 50%		Falha a 75%	
	SCG	Rprop	SCG	Rprop	SCG	Rprop
configuração A	94,70%	96,10%	95,61%	94,42%	90,63%	89,86%
configuração B	94,70%	94,80%	95,79%	94,25%	90,98%	90,98%
configuração C	94,30%	95,70%	97,45%	95,85%	91,22%	90,45%
configuração D	95,10%	95,30%	96,50%	94,72%	90,21%	87,78%
configuração E	95,50%	95,50%	96,91%	93,89%	91,34%	90,45%

A combinação da Tabela 36 também exibe uma medida de dispersão, uma medida de tendência central e a curtose. em comparação aos resultados Tabela 35 e à 30 esta tem um desempenho um pouco inferior, porém é importante frisar que as combinações de medidas estatísticas que envolvem medidas de dispersão, assimetria e tendência central (2, 3, 8 e 9) apresentam os melhores desempenhos gerais sendo a 8 inclusive a que apresenta os melhores desempenhos para as falhas mecânicas.

Tabela 37 – Acertos globais para as faltas incipientes de estator para combinação de medidas estatísticas 10.

	Falha a 25%		Falha a 50%		Falha a 75%	
	SCG	Rprop	SCG	Rprop	SCG	Rprop
configuração A	99,00%	98,40%	97,09%	91,81%	92,29%	92,76%
configuração B	98,90%	98,50%	97,51%	87,90%	94,13%	93,95%
configuração C	98,80%	98,90%	92,35%	89,56%	93,59%	91,64%
configuração D	98,60%	98,90%	90,57%	87,67%	92,59%	93,36%
configuração E	98,80%	98,30%	93,47%	80,97%	94,19%	93,65%

Da Tabela 37, as melhores configurações para 25% são A, B e C. Para 50% são B, A e D e para 75% E, B e C.

A Tabela 38 apresenta os desempenhos para o centroide “atuando” sozinho por ser uma medida que apresenta três parâmetros e conseqüentemente uma diversidade maior de informação, seu desempenho é razoável quando comparado ao desempenho de outros parâmetros apresentados isolados à rede.

Tabela 38 – Acertos globais para as faltas incipientes de estator para o centroide

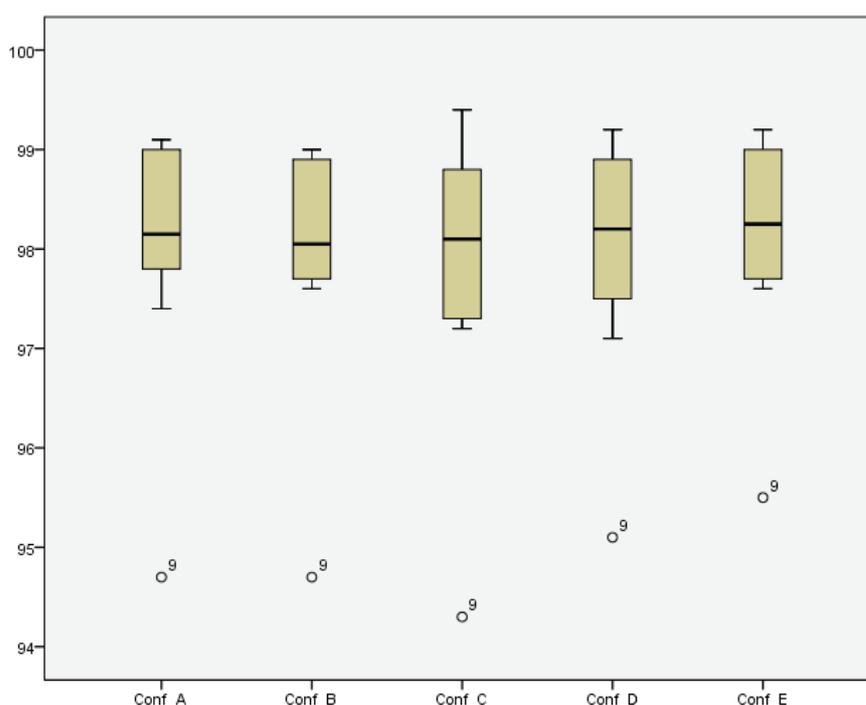
	Falha a 25%		Falha a 50%		Falha a 75%	
	SCG	Rprop	SCG	Rprop	SCG	Rprop
configuração A	91,90%	92,30%	92,76%	92,05%	83,63%	82,69%
configuração B	91,90%	92,10%	93,77%	91,93%	85,24%	84,05%
configuração C	92,50%	92,80%	94,72%	91,93%	85,24%	84,17%
configuração D	91,90%	92,40%	91,52%	90,69%	86,72%	84,64%
configuração E	91,90%	92,00%	91,52%	89,86%	83,75%	86,18%

Na Figura 45 estão os *box plots* para os desempenhos das cinco configurações de redes adotadas juntamente com a Tabela 39 que apresenta as estatísticas descritivas para as cinco configurações treinadas a SCG com faltas a 25%. A observação das médias mostra que as configurações E, A e D são as que apresentam melhor desempenho.

Tabela 39 – Estatísticas descritivas para as cinco configurações treinadas a SCG para falhas elétricas a 25%

	Configuração A	Configuração B	Configuração C	Configuração D	Configuração E
Média	97,97	97,90	97,90	97,96	98,12
Mediana	98,15	98,05	98,10	98,20	98,25
Desvio Padrão	1,28	1,24	1,45	1,23	1,13
Mínimo	94,70	94,70	94,30	95,10	95,50
Máximo	99,10	99,00	99,40	99,20	99,20
Primeiro Quartil	97,70	97,67	97,27	97,40	97,67
Terceiro Quartil	99,00	98,92	98,85	98,95	99,05

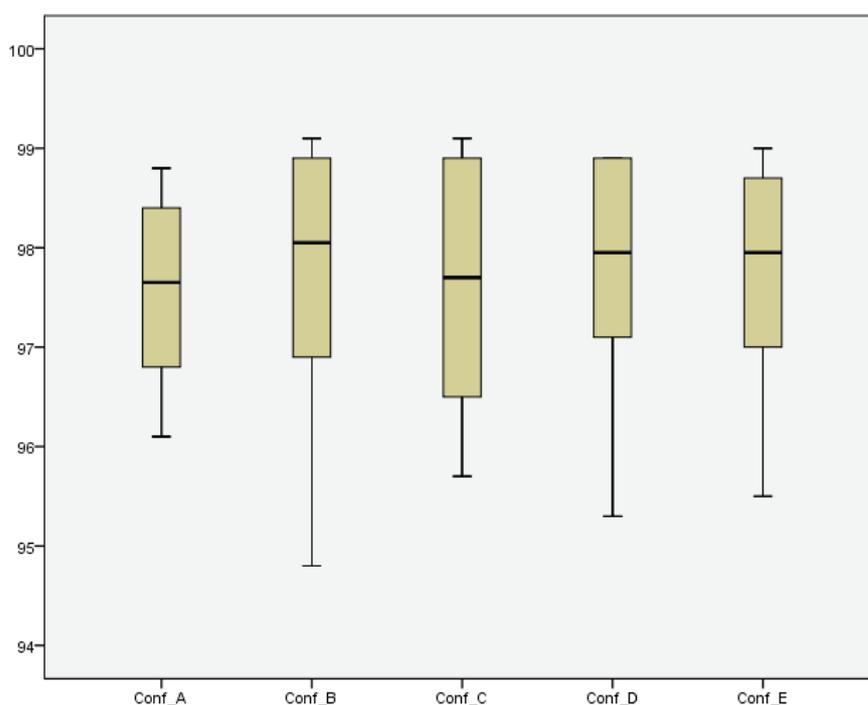
Figura 45 – *Box plot* para as cinco configurações treinadas através do SCG para detecção de curto-circuito entre espiras a 25%.



Fonte: Próprio autor

Vale a observação da configuração A, pois esta havia se mostrado provável configuração preferível entre as demais de média mais altas para a classificação de falhas mecânicas para ambos os métodos de treinamento. Contudo, esta mostrou não ser a maior média para detecção de faltas intermitentes de 25% do enrolamento com a rede treinada a Rprop, Tabela 40 e Figura 46, para este caso as melhores configurações são D, E e B.

Figura 46 – *Box plots* para as redes treinadas a Rprop para as falhas de 25%.



Fonte: Próprio autor

Tabela 40 – Estatísticas descritivas para as cinco configurações treinadas a Rprop para falhas elétricas a 25%

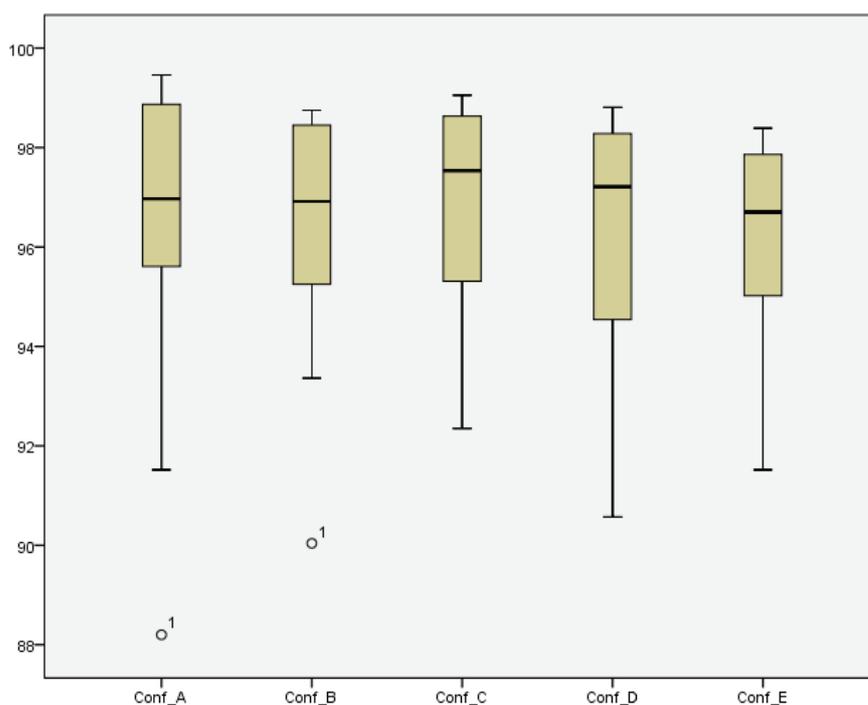
	Configuração A	Configuração B	Configuração C	Configuração D	Configuração E
Média	97,58	97,66	97,60	97,80	97,74
Mediana	97,65	98,05	97,70	97,95	97,95
Desvio Padrão	0,90	1,45	1,27	1,18	1,17
Mínimo	96,10	94,80	95,70	95,30	95,50
Máximo	98,80	99,10	99,10	98,90	99,00
Primeiro Quartil	96,77	96,67	96,47	97,02	96,90
Terceiro Quartil	98,40	98,90	98,90	98,90	98,75

Para o caso de enrolamentos apresentando falhas a 50%, a Tabela 41 exhibe as configurações C,D e B com as maiores médias para redes treinadas a SCG. A Figura 47 exhibe graficamente os desempenhos de todas as configurações. As configurações A e B apresentaram *outliers*.

Tabela 41 – Estatísticas descritivas para as cinco configurações treinadas a SCG para falhas elétricas a 50%

	Configuração A	Configuração B	Configuração C	Configuração D	Configuração E
Média	95,97	96,16	96,81	96,38	96,11
Mediana	96,97	96,91	97,53	97,21	96,70
Desvio Padrão	3,55	2,741	2,17	2,62	2,23
Mínimo	88,20	90,04	92,35	90,57	91,52
Máximo	99,46	98,75	99,05	98,81	98,39
Primeiro Quartil	94,58	94,77	95,03	94,43	94,63
Terceiro Quartil	98,88	98,49	98,63	98,36	97,98

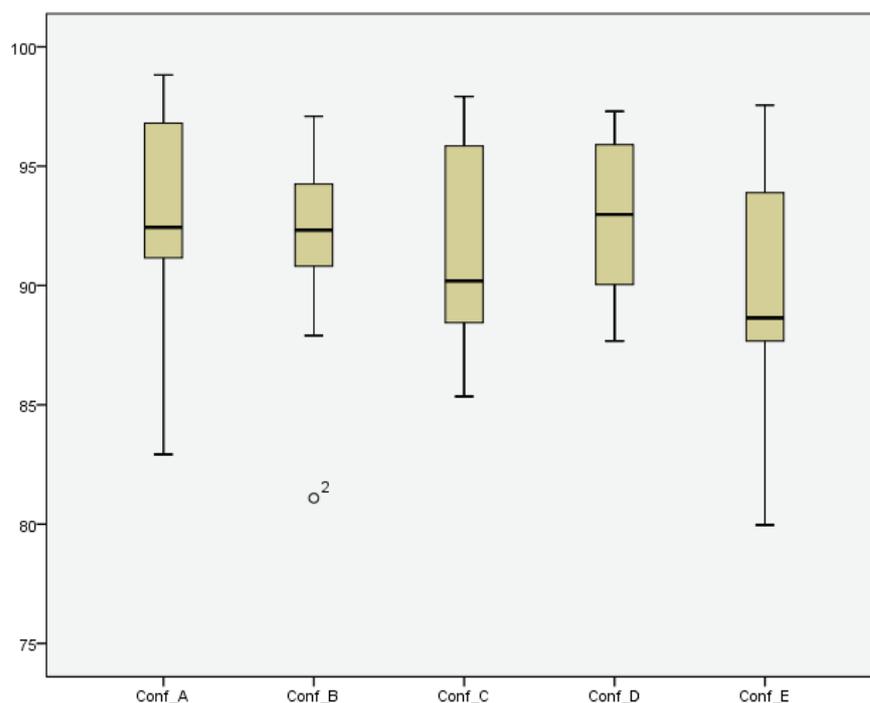
Figura 47 – *Box plots* para as faltas intermitentes a 50% dos enrolamentos para as redes treinadas a SCG.



Fonte: Próprio autor

As configurações A, D e B são as que apresentam melhores médias para o treinamento a Rprop, Tabela 42. Os *box plots* para as faltas a 50% e redes treinadas a Rprop estão na Figura 48. Os melhores desempenhos ficaram por conta das sonfigurações D, A e B.

Figura 48 – Box plots para falhas a 50% dos enrolamentos, redes treinadas a Rprop.



Fonte: Próprio autor

Tabela 42 – Estatísticas descritivas para as cinco configurações treinadas a Rprop para falhas elétricas a 50%

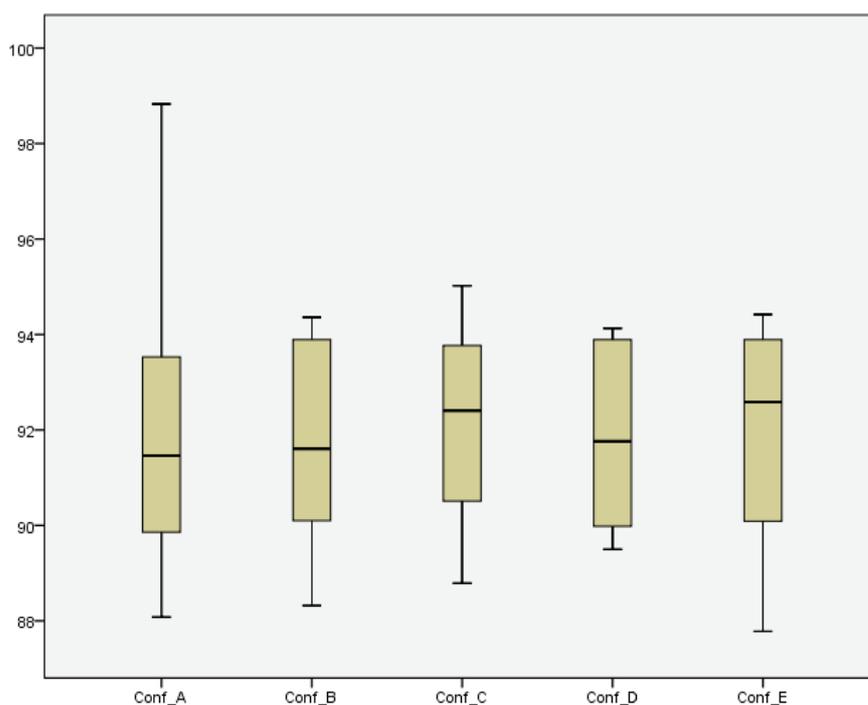
	Configuração A	Configuração B	Configuração C	Configuração D	Configuração E
Média	92,68	91,66	91,50	92,78	89,31
Mediana	92,43	92,32	90,18	92,97	88,64
Desvio Padrão	4,71	4,56	4,17	3,43	5,90
Mínimo	82,92	81,09	85,35	87,67	79,96
Máximo	98,82	97,09	97,92	97,30	97,56
Primeiro Quartil	90,55	90,08	88,24	89,56	85,99
Terceiro Quartil	97,05	94,66	95,88	96,09	94,61

A Tabela 43 apresenta as estatísticas para as faltas a 75% de enrolamentos danificados com redes treinadas a SCG, a Figura 49 apresenta os *box plots* referentes às configurações treinadas. Para as falhas a 75%, as redes que apresentaram melhores desempenhos foram C, E e D para treino a SCG e A, B e C para treino a Rprop indicados na Tabela 44 com gráficos na Figura 50.

Tabela 43 – Estatísticas descritivas para as cinco configurações treinadas a SCG para falhas elétricas a 75%

	Configuração A	Configuração B	Configuração C	Configuração D	Configuração E
Média	91,96	91,57	92,05	91,82	91,90
Mediana	91,46	91,60	92,40	91,76	92,58
Desvio Padrão	3,18	2,24	2,24	1,97	2,39
Mínimo	88,08	88,32	88,79	89,50	87,78
Máximo	98,83	94,36	95,02	94,13	94,42
Primeiro Quartil	89,56	89,72	90,20	89,86	89,70
Terceiro Quartil	93,71	93,95	93,86	93,92	93,96

Figura 49 – *Box plot* para as cinco configurações treinadas através do SCG para detecção de curto-circuito entre espiras a 75%.

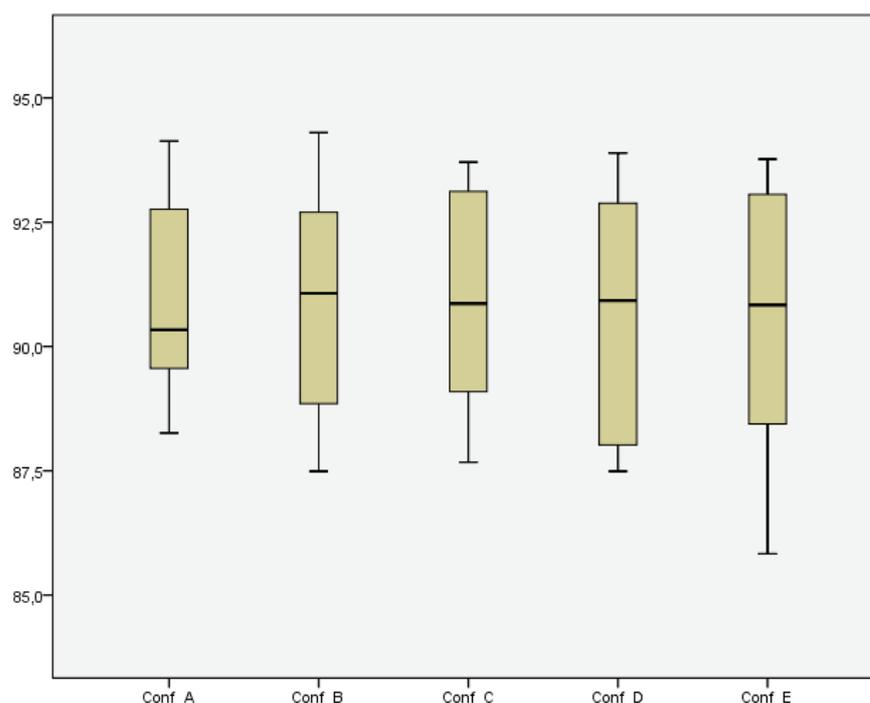


Fonte: Próprio autor

Tabela 44 – Estatísticas descritivas para as cinco configurações treinadas a Rprop para falhas elétricas a 75%

	Configuração A	Configuração B	Configuração C	Configuração D	Configuração E
Média	91,0210	90,9310	90,9260	90,5700	90,5400
Mediana	90,3350	91,0700	90,8650	90,9250	90,8350
Desvio Padrão	2,03342	2,36453	2,03843	2,58083	2,71537
Mínimo	88,26	87,49	87,67	87,49	85,83
Máximo	94,13	94,30	93,71	93,89	93,77
Primeiro Quartil	89,5025	88,7475	89,0150	87,9600	88,3500
Terceiro Quartil	92,9975	93,0125	93,1350	93,0000	93,2075

Figura 50 – *Box plot* para as cinco configurações treinadas através do Rprop para detecção de curto-circuito entre espiras a 75%.



Fonte: Próprio autor

Uma vez observadas estatísticas descritivas para todas as configurações em todos os casos das falhas elétricas, é importante definir-se uma configuração preferível. Para as falhas mecânicas a configuração A se mostrou provavelmente a melhor em virtude de sua simplicidade em comparação às configurações C e D que, juntamente com a A, tiveram as melhores médias para falhas mecânicas.

Quando se observa as estatísticas descritivas para as falhas elétricas (da Tabela 39 a 44), a configuração A e C só se mostraram entre as melhores em três das seis tabelas, ao passo

que a configuração D se mostrou entre as melhores em cinco das seis tabelas de estatísticas. Esta configuração também está entre as três candidatas das falhas mecânicas. Dado o fato de se apresentar entre os melhores desempenhos para ambos os métodos para as falhas mecânicas e em quase todos os casos das falhas elétricas, a configuração D é a preferível para o monitoramento para ambos os tipos de faltas.

5.2.1 Comparativo com treinamento de rede a partir de parâmetros provenientes da decomposição *wavelet packet* para faltas do estator

Da mesma forma que foi realizado para os problemas mecânicos, nesta seção as Tabelas 45 e 46 apresentam o desempenho das cinco configurações de rede para o caso dos vetores de entrada serem formados unicamente pelos valores das energias dos nós 2, 4, 8 e 16.

Tabela 45 – Percentuais de acertos para cada configuração de rede com vetores de entradas formados pelas energias dos nós 2, 4, 8 e 16 da WPT das falhas elétricas treinadas a Rprop.

	A	B	C	D	E
Percentagem Classificação correta	79,13%	77,88%	76,76%	81,26%	72,08%
Percentagem Classificação Incorreta	20,86%	22,11%	23,23%	18,73%	27,91%
Maior erro do teste	7,25%	7,70%	7,93%	6,94%	9,17%
Maior erro de validação	7,88%	7,77%	7,79%	7,45%	8,54%
Maior erro Global	7,42%	7,38%	7,28%	6,93%	8,70%

Tabela 46 – Percentuais de acertos para cada configuração de rede com vetores de entradas formados pelas energias dos nós 2, 4, 8 e 16 da WPT das falhas elétricas treinadas a SCG.

	A	B	C	D	E
Percentagem Classificação correta	82,98%	83,75%	77,23%	84,05%	80,49%
Percentagem Classificação Incorreta	17,01%	16,24%	22,76%	15,94%	19,50%
Maior erro do teste	5,98%	5,93%	7,97%	5,94%	7,00%
Maior erro de validação	6,19%	6,40%	7,81%	6,39%	7,20%
Maior erro Global	5,76%	5,95%	6,89%	5,92%	6,77%

O que se observa das duas tabelas é que o desempenho para as faltas no estator é superior ao desempenho para as falhas mecânicas. Tendo, inclusive, taxas de acertos apreciáveis, porém estas são, de maneira geral inferiores às obtidas quando se realiza a combinação de medidas estatísticas.

5.3 CONCLUSÃO

Neste capítulo foram apresentados os resultados obtidos mediante uso da metodologia proposta nesta tese. Na aplicação para as falhas mecânicas o método apresentou desempenho considerável tendo taxas de acerto, inclusive próximas dos 97% na configuração de melhor resultado. Para as falhas elétricas o desempenho foi igualmente satisfatório apresentando taxas de acertos, em sua maioria e para ambos os métodos de treinamento, com valores acima dos 92%.

É importante frisar que as configurações de rede apresentadas aqui foram as mais simples (menor número de camadas e menor quantidade de neurônios por camada). Com isso, nada impede que um melhor resultado seja obtido com configurações de redes mais complexas. Realizando análises estatísticas dos desempenhos das configurações de rede

O propósito neste trabalho é mostrar a combinação do uso das wavelets com redes neurais artificiais com o para realização de classificação e reconhecimento de padrões. Ainda nesse sentido, as medidas estatísticas foram usadas três a três para os nós de detalhes mais energéticos da WPT. Essa maneira garantiu eficiência computacional, tendo na maioria dos ensaios um tempo de treinamento da rede neural inferior a 8 segundos para as falhas elétricas com boas taxas de acerto, o que leva à conclusão de que a metodologia é eficaz e promissora uma vez que pode receber incrementos para melhora de desempenho.

Identificar uma configuração que fosse preferível entre as cinco também é uma conclusão importante. Comparações foram feitas entre os desempenhos e entre as médias de acertos, chegando-se à conclusão que a configuração D é preferível. Dentre as combinações de medidas estatísticas, a combinação 8 se apresenta com o melhor desempenho entre as classificações de problemas mecânicos e entre as melhores para as faltas incipientes de estator, sendo então a escolhida.

O método de treinamento Rprop se mostra também mais interessante, pois o esforço computacional para este método é muito menor que para o SCG, conforme evidenciado anteriormente na Tabela 13 para ps defeitos mecânicos, seu tempo de treinamento para a RNA é de 9 segundos e 636 iterações contra 18 segundos e 711 iterações do SCG. Na Tabela 47, estão ilustrados as estatísticas descritivas para que se possa fazer uma comparação de desempenho entre ambos os métodos. Mesmo tendo um desempenho um tanto abaixo do SCG, o *resilient backpropagation* ainda é preferível. A Tabela 48 mostra os tempos médios de treinamentos para o caso de detecção de faltas intermitentes.

Tabela 47 – Estatísticas descritivas SCG *versus* Rprop.

	SCG	Rrpop
Média	92,19	90,94
Mediana	94,20	91,75
Desvio Padrão	7,20	2,96
Mínimo	47,00	83,30
Máximo	96,80	95,40

Tabela 48 – Tempo médio de treinamento para a metodologia aplicada às falhas mecânicas e elétricas

	SCG	RBprop
Falhas Elétricas	7,81 segundos	7,75 segundos

Ainda para as falhas elétricas, a Tabela 49 mostra o comparativo entre os dois métodos para detecção de falhas a 25% e para este caso a diferença é de 0,3% entre as médias. O que mostra, mais uma vez a eficácia do Rprop.

Tabela 49 – Estatísticas descritivas SCG *versus* Rprop.

	SCG	Rrpop
Média	97,97	97,67
Mediana	98,10	97,95
Desvio Padrão	1,22	1,16
Mínimo	94,30	94,80
Máximo	99,40	99,10

6 CONCLUSÕES

6.1 CONCLUSÕES GERAIS

Esta trabalho apresentou os resultados experimentais de uma metodologia de análise da assinatura de corrente para um motor de indução, a qual utilizou a transformada *wavelet packet* com as redes neurais artificiais com o intuito de identificar a presença de falhas elétricas e mecânicas neste tipo de equipamento. O procedimento constou, basicamente na extração, a partir de processamento estatístico, dos atributos de nós de detalhe específicos da *wavelet packet tree*, exatamente os que apresentam a maior entropia de Shannon, para formação dos vetores de entrada de uma rede perceptron multicamada para fazer o reconhecimento de padrões.

As elevadas taxas de acertos globais obtidas com a aplicação da metodologia descrita credenciam a mesma para o fim de diagnóstico de falhas em motores de indução. Este capítulo pretende, justamente, certificar essa posição, procedendo a um confronto com outras pesquisas registradas na literatura especializada.

6.1.1 Comparativos

Os resultados deste trabalho foram confrontados com os dados constantes nas publicações mais recentes acerca do tema, não importando se o método descrito possuía ou não o caráter multiclassificatório, ou seja, presta-se-ia tanto para o diagnóstico de falhas elétricas como mecânicas. As Tabelas 50 e 51, a primeira relativa aos defeitos de natureza mecânica e a segunda para as ocorrências elétricas, descrevem os tópicos desta comparação, assinalando a metodologia em uso e a efetividade dos acertos no diagnóstico das falhas para cada caso.

No trabalho de (PALÁCIOS et al., 2016) tanto as redes perceptron multicamada e as redes *radial basis functions* foram utilizadas para verificação das falhas em estator. Para o

Tabela 50 – Tabela comparativa de métodos de monitoramento para falhas mecânicas.

Referências	(SHI; CHEN; VAGAPOV, 2013)	(MATIĆ et al., 2012)	(SU; CHONG,)	(HAJI; TOLIYAT, 2001)	(LEE et al., 2006)	Este trabalho
Motor	5 HP	1.5 kW	597 kW	3HP	0.385 kW	1/3 HP
Metodologia	DWT	H	STFT	DQTrans	CWT e DFT	WPT
Fonte de dados	MCSA	MCSA	Vib	MCSA	MCSA	MCSA
Método de classificação	PNN	SVM	MLP - NN	BMEC	Fuzzy	MLP - NN
Defeitos em Rotor	Sim	Sim	Sim	Sim	Sim	Sim
Defeitos em Rolamentos	Não	Não	não	Não	Sim	Sim
Efetividade do método (%)	85	100	97	97.9	90	< 93
Multiclassificatório	Não	não	não	Não	Sim	Sim

DWT = Discret Wavelet Transform ; PNN = Probabilistic Neural Network ; H = Transformada de Hilbert ; STFT = Short Time fourier Transform ; DQTrans = transformada DQ ; SVM = Support Vector Machine ; BMEC = Bayes Minimal Error Classifier

Tabela 51 – Tabela comparativa de métodos de monitoramento para falhas em estator.

Referências	(PALÁCIOS et al., 2016)	(BAZAN et al., 2017)	Este trabalho
Motor	0,75 kW	0,74 / 1,48 kW	1/3 HP
Metodologia	D + PCA	MI	WPT + ME
Fonte de dados	MCSA	MCSA	MCSA
Método de classificação	MLP / RBF	MLP / RBF	MLP
Deteção de descargas parciais	Sim	Sim	Sim
Determinação da severidade da falta	não	Sim	Sim
Efetividade do método (%)	< 90	< 93	< 92

PCA = Principal Component Analysis ; STFT = Short Time fourier Transform ; DQTrans = transformada DQ ; SVM = Support Vector Machine ; D = Discretização do sinal ; ME = Medidas estatísticas ; MI = Mutual Information

caso de (BAZAN et al., 2017), dois motores foram testados para realizar os ensaios um com 0,74kW e outro com 1,48kW. Além disso, a técnica escolhida para analisar os sinais foi a *mutual information*, medida estatística da quantidade de informação que uma variável aleatória contém acerca da outra. Ou seja, a informação mútua é definida como a quantidade de incerteza que é reduzida em uma dada variável, a partir do conhecimento provido pela outra. Assim é realizada a medida da informação mútua entre as correntes de fase.

Está patente nas Tabelas 46 e 47 que a taxa de acertos do método proposto, inferior a 93% para as falhas mecânicas e 92% para os defeitos elétricos, somente é superada pelos procedimentos não multiclassificatórios, de âmbito, portanto mais restrito. Essa constatação constitui-se um handicap em favor do modelo desenvolvido neste trabalho. É importante ressaltar que, ao contrário das demais metodologias analisadas neste tópico, também a severidade da falha presente numa fase do motor é passiva de monitoração pelo procedimento proposto nesta tese, o que concede um status de originalidade ao mesmo.

6.2 DIRETIVAS FUTURAS

Conforme mencionado anteriormente, outras configurações de rede e novas topologias podem ser exploradas com o intuito de aprimorar o desempenho do método. O uso de outras medidas estatísticas também seria uma alternativa promissora como, por exemplo, a Informação Mútua. Esta opção poderia ser explorada e utilizada em combinação com as outras medidas aqui exploradas.

Outra proposta que se mostra interessante é o estabelecimento de diretrizes para implementação em hardware desta metodologia. A combinação de DSP e FPGA (*Field Programmable Gate Array*) já tem sido, inclusive abordada deteção de descargas parciais no sistema de potência (MOTA, 2011) (JAIN; SINGH,), mostrando a viabilidade do desenvolvimento, ou mesmo da possibilidade de elaboração de projetos para equipamentos de monitoramento em tempo real.

O uso de ferramentas de redução de dimensionalidade com o propósito de compactar os conjuntos de dados e torná-los mais separáveis assim como *Trace Ratio Linear Discriminant*

Analysis (TR-LDA).

Pelo fato de fundamentar-se na classificação de padrões mediante a observação de medidas estatísticas extraídas de nós de detalhe fornecidos pela análise *wavelet packet*, a metodologia aqui proposta também se mostra promissora para realizar análise de outras máquinas elétricas (rotativas ou não) inclusive em regime transitório. Acredita-se que o tipo de máquina deverá influenciar nos nós mais energéticos da WPT.

REFERÊNCIAS

- ABDESH, M.; KHAN, S.; RAHMAN, M. A. A new wavelet based diagnosis and protection of faults in induction motor drives. In: IEEE. *Power Electronics Specialists Conference, 2008. PESC 2008. IEEE*. [S.l.], 2008. p. 1536–1541. Citado na página 23.
- ABREU, T. et al. Metodologia híbrida utilizando os modelos arima e redes neurais artificiais para previsão de cargas elétricas. In: *Congresso Brasileiro de Automática. Campina Grande*. [S.l.: s.n.], 2012. Citado na página 60.
- ARRUDA, E. F. de. *Análise de distúrbios relacionados com a qualidade da energia elétrica utilizando a transformada Wavelet*. Tese (Doutorado) — Universidade de São Paulo, 2003. Citado 2 vezes nas páginas 37 e 40.
- AYHAN, B.; CHOW, M.-Y.; SONG, M.-H. Multiple signature processing-based fault detection schemes for broken rotor bar in induction motors. *IEEE Transactions on Energy Conversion*, IEEE, v. 20, n. 2, p. 336–343, 2005. Citado na página 26.
- BACCARINI, L. M. R. et al. Svm practical industrial application for mechanical faults diagnostic. *Expert Systems with Applications*, Elsevier, v. 38, n. 6, p. 6980–6984, 2011. Citado na página 24.
- BAFROUI, H. H.; OHADI, A. Application of wavelet energy and shannon entropy for feature extraction in gearbox fault detection under varying speed conditions. *Neurocomputing*, Elsevier, v. 133, p. 437–445, 2014. Citado na página 81.
- BARAKAT, M. et al. Self adaptive growing neural network classifier for faults detection and diagnosis. *Neurocomputing*, v. 74, n. 18, p. 3865–3876, 2011. ISSN 09252312. Citado 2 vezes nas páginas 25 e 47.
- BARENDSE, P. S. et al. The application of wavelets for the detection of inter-turn faults in induction machines. In: *2009 IEEE Int. Electr. Mach. Drives Conf. IEMDC '09*. [S.l.: s.n.], 2009. p. 1401–1407. ISBN 9781424442522. Citado na página 22.
- BAZAN, G. H. et al. Stator fault analysis of three-phase induction motors using information measures and artificial neural networks. *Electric Power Systems Research*, Elsevier, v. 143, p. 347–356, 2017. Citado na página 114.
- BELL, R. et al. Report of large motor reliability survey of industrial and commercial installations. i. *IEEE Transactions on Industry applications*, Institute of Electrical and Electronics Engineers, v. 21, n. 4, p. 853–864, 1985. Citado na página 19.
- BENBOUZID, M. E. H.; KLIMAN, G. B. What stator current processing-based technique to use for induction motor rotor faults diagnosis? *IEEE Trans. Energy Convers.*, v. 18, n. 2, p. 238–244, 2003. ISSN 08858969. Citado na página 19.
- BŁASZCZUK, J.; POZORSKI, Z. Application of the lipschitz exponent and the wavelet transform to function discontinuity estimation. *Scientific Research of the Institute of Mathematics and Computer Science*, v. 6, n. 1, p. 23–30, 2007. Citado na página 72.

BLÖDT, M. et al. Models for bearing damage detection in induction motors using stator current monitoring. *Industrial Electronics, IEEE Transactions on*, IEEE, v. 55, n. 4, p. 1813–1822, 2008. Citado na página 19.

BOUZID, M. B. K. et al. An effective neural approach for the automatic location of stator interturn faults in induction motor. *IEEE Trans. Ind. Electron.*, v. 55, n. 12, p. 4277–4289, 2008. ISSN 02780046. Citado na página 22.

BOUZIDA, A. et al. Fault diagnosis in industrial induction machines through discrete wavelet transform. *IEEE Trans. Ind. Electron.*, v. 58, n. 9, p. 4385–4395, 2011. ISSN 0278-0046. Disponível em: <http://ieeexplore.ieee.org/xpl/freeabs{_}all.jsp?arnumber=5645>. Citado 2 vezes nas páginas 48 e 73.

BRAGA, A. Curvas roc: aspectos funcionais e aplicações. 2000. Citado na página 133.

CAMARGO, T. A. B. d. S. *Transformada Wavelet Aplicada a Análise de Falhas em Rolamentos*. Tese (Doutorado). Citado na página 49.

CASTEJÓ, C.; LARA, O.; GARCÍA-PRADA, J. C. Automated diagnosis of rolling bearings using mra and neural networks. *Mech. Syst. Signal Process.*, v. 24, p. 289–299, 2009. Citado na página 58.

CHANG, S. et al. Speech enhancement for non-stationary noise environment by adaptive wavelet packet. n. I, p. 561–564. Citado na página 47.

CHEN, C.; MO, C. A method for intelligent fault diagnosis of rotating machinery. *Digit. Signal Process. A Rev. J.*, v. 14, n. 3, p. 203–217, 2004. ISSN 10512004. Citado na página 24.

CHEN, C.-S.; SU, S.-L. Resilient back-propagation neural network for approximation 2-d gdop. In: CITESEER. *Proceedings of the International MultiConference of Engineers and Computer Scientists*. [S.l.], 2010. v. 2, p. 900–904. Citado na página 63.

CHOW, T. W.; HAI, S. Induction machine fault diagnostic analysis with wavelet technique. *IEEE Transactions on Industrial Electronics*, IEEE, v. 51, n. 3, p. 558–565, 2004. Citado na página 25.

COMBES, J.-M.; GROSSMANN, A.; TCHAMITCHIAN, P. *Wavelets: Time-Frequency Methods and Phase Space Proceedings of the International Conference, Marseille, France, December 14–18, 1987*. [S.l.]: Springer Science & Business Media, 2012. Citado na página 37.

CRUZ, A. G. de A. et al. A hybrid system based on fuzzy logic to failure diagnosis in induction motors. *IEEE Latin America Transactions*, IEEE, v. 15, n. 8, p. 1480–1489, 2017. Citado na página 26.

CRUZ, S.; CARDOSO, A. Rotor cage fault diagnosis in three-phase induction motors by the total instantaneous power spectral analysis. *Conf. Rec. 1999 IEEE Ind. Appl. Conf. Thirty-Forth IAS Annu. Meet. (Cat. No.99CH36370)*, v. 3, p. 1929–1934, 1999. ISSN 0197-2618. Citado na página 22.

CRUZ, S.; CARDOSO, A. Rotor cage fault diagnosis in three-phase induction motors by extended park's vector approach. *Electric Machines & Power Systems*, Taylor & Francis, v. 28, n. 4, p. 289–299, 2000. Citado na página 22.

CRUZ, S. M.; CARDOSO, A. M. Stator winding fault diagnosis in three-phase synchronous and asynchronous motors, by the extended park's vector approach. *IEEE Transactions on industry applications*, IEEE, v. 37, n. 5, p. 1227–1233, 2001. ISSN 00939994. Citado na página 22.

DAUBECHIES, I. Orthonormal bases of compactly supported wavelets. *Communications on pure and applied mathematics*, Wiley Online Library, v. 41, n. 7, p. 909–996, 1988. Citado na página 72.

DAUBECHIES, I. The wavelet transform, time-frequency localization and signal analysis. *IEEE transactions on information theory*, IEEE, v. 36, n. 5, p. 961–1005, 1990. Citado na página 72.

DAUBECHIES, I. *Ten lectures on wavelets*. [S.l.]: SIAM, 1992. Citado na página 72.

DAUBECHIES, I. Orthonormal bases of compactly supported wavelets ii. variations on a theme. *SIAM Journal on Mathematical Analysis*, SIAM, v. 24, n. 2, p. 499–519, 1993. Citado na página 72.

DINIZ, P. S.; SILVA, E. A. da; NETTO, S. L. *Processamento Digital de Sinais-: Projeto e Análise de Sistemas*. [S.l.]: Bookman Editora, 2014. Citado 4 vezes nas páginas 31, 33, 35 e 41.

DOERING, D. *O uso da transformada Wavelet bi-dimensional no conceito do espaço de escala*. Tese (Doutorado), 2005. Citado na página 37.

EBRAHIMI, B. M. et al. Novel indices for broken rotor bars fault diagnosis in induction motors using wavelet transform. *Mech. Syst. Signal Process.*, Elsevier, v. 30, p. 131–145, 2012. ISSN 08883270. Citado 3 vezes nas páginas 11, 20 e 29.

EKICI, S.; YILDIRIM, S.; POYRAZ, M. Energy and entropy-based feature extraction for locating fault on transmission lines by using neural network and wavelet packet decomposition. *Expert Systems with Applications*, Elsevier, v. 34, n. 4, p. 2937–2944, 2008. Citado na página 75.

FANUCCHI, R. Z. *A Detecção e a Localização de Falhas de Alta Impedância em Sistemas de Distribuição utilizando Transformadas Wavelets e Redes Neurais Artificiais*. 176 p. Tese (Doutorado) — Universidade de São Carlos, 2014. Citado 2 vezes nas páginas 21 e 31.

FILIPPETTI, F. et al. Recent developments of induction motor drives fault diagnosis using ai techniques. *IEEE transactions on industrial electronics*, v. 47, n. 5, p. 994–1004, 2000. ISSN 02780046. Citado na página 22.

GALVÃO, J. U. . H. P. . F. V. . M. D. . R. Additional constraints to ensure three vanishing moments for orthonormal wavelet filter banks and transient detection. In: COMPUTACIONAL, S. B. de Matemática aplicada e (Ed.). [S.l.: s.n.], 2015. p. 17 – 29. Citado 2 vezes nas páginas 72 e 73.

GHASEMI, S. *An Application Of A Hybrid Wavelet-SVM Based Approach For Induction Motor Fault Detection*. Tese (Doutorado), 2010. Citado na página 24.

GHATE, V. N.; DUDUL, S. V. Optimal mlp neural network classifier for fault detection of three phase induction motor. *Expert Systems with Applications*, Elsevier, v. 37, n. 4, p. 3468–3481, 2010. Citado 2 vezes nas páginas 49 e 58.

GHATE, V. N.; DUDUL, S. V. Cascade neural-network-based fault classifier for three-phase induction motor. *IEEE Trans. Ind. Electron.*, v. 58, n. 5, 2011. Citado na página 58.

- GOMES, L. V. J. *From Fourier Analysis to Wavelets*. 1st ed. 2015. ed. [S.l.]: Springer International Publishing, 2015. (IMPA Monographs 3). ISBN 3319220748,978-3-319-22074-1,978-3-319-22075-8. Citado na página 37.
- GREY, J. M.; GORDON, J. W. Perceptual effects of spectral modifications on musical timbres. *The Journal of the Acoustical Society of America*, ASA, v. 63, n. 5, p. 1493–1500, 1978. Citado na página 49.
- HAJI, M.; TOLIYAT, H. A. Pattern recognition-a technique for induction machines rotor broken bar detection. *IEEE Transactions on Energy Conversion*, IEEE, v. 16, n. 4, p. 312–317, 2001. Citado 2 vezes nas páginas 22 e 113.
- HAYKIN, S. *Redes Neurais - Princípios e Prática*. [S.l.]: Bookman, 2003. ISBN 8573077182. Citado 3 vezes nas páginas 52, 58 e 59.
- HUANG, Y.-C.; HUANG, C.-M. Evolving wavelet networks for power transformer condition monitoring. *Power Delivery, IEEE Transactions on*, IEEE, v. 17, n. 2, p. 412–416, 2002. Citado na página 36.
- JAIN, S.; WADHWANI, A. K.; WADHWANI, S. Wavelet transform and back propagation neural network: A tool for induction motor fault detection and classification. Citado 2 vezes nas páginas 49 e 75.
- JAIN, S. K.; SINGH, S. N. Low-order dominant harmonic estimation using adaptive wavelet neural network. *IEEE Trans. Ind. Electron.* ISSN 02780046. Citado na página 114.
- JARDINE, A. K.; LIN, D.; BANJEVIC, D. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical systems and signal processing*, Elsevier, v. 20, n. 7, p. 1483–1510, 2006. Citado na página 19.
- JAWADEKAR, A. et al. Artificial neural network-based induction motor fault classifier using continuous wavelet transform. *Syst. Sci. Control Eng. An Open Access J.*, v. 2, p. 684–690, 2014. Disponível em: <<http://dx.doi.org/10.1080/21642583.2014.956266>>. Citado 3 vezes nas páginas 21, 23 e 73.
- JIN, X. et al. Motor bearing fault diagnosis using trace ratio linear discriminant analysis. v. 61, n. 5, p. 2441–2451, 2014. ISSN 0278-0046. Disponível em: <<http://ieeexplore.ieee.org.http://ieeexplore.ieee.org/document/6583974/>>. Citado na página 28.
- JINGBO, L.; XIUQING, W. Vibration pattern recognition and classification of electric generator in power system using wavelet analysis. *2008 27th Chinese Control Conf.*, n. 1, p. 34–37, 2008. Citado na página 24.
- KANG, M.; KIM, J. M. Reliable fault diagnosis of multiple induction motor defects using a 2-d representation of shannon wavelets. *IEEE Trans. Magn.*, v. 50, n. 10, 2014. ISSN 00189464. Citado na página 25.
- KANKAR, P. K.; SHARMA, S. C.; HARSHA, S. P. Fault diagnosis of ball bearings using continuous wavelet transform. *Applied Soft Computing*, Elsevier, v. 11, n. 2, p. 2300–2312, 2011. Citado na página 81.

KHAN, M.; OZGONENEL, O.; RAHMAN, M. A. Real time implementation of a wavelet power based algorithm for three-phase synchronous generator protection. In: IEEE. *Electrical and Computer Engineering, 2006. CCECE'06. Canadian Conference on*. [S.l.], 2006. p. 182–185. Citado 2 vezes nas páginas 21 e 23.

KHAN, M.; OZGONENEL, O.; RAHMAN, M. A. Diagnosis and protection of stator faults in synchronous generators using wavelet transform. In: IEEE. *Electric Machines & Drives Conference, 2007. IEMDC'07. IEEE International*. [S.l.], 2007. v. 1, p. 184–189. Citado 2 vezes nas páginas 21 e 23.

KIM, K.; PARLOS, A. G. Induction motor fault diagnosis based on neuropredictors and wavelet signal processing. *IEEE/ASME Transactions on mechatronics*, v. 7, n. 2, p. 201–219, 2002. ISSN 10834435. Citado na página 28.

KOWALSKI, C. T.; ORLOWSKA-KOWALSKA, T. Neural networks application for induction motor faults diagnosis. *Mathematics and Computers in Simulation*, Elsevier, v. 63, n. 3, p. 435–448, 2003. Citado na página 21.

LEE, S.-H. et al. Fourier and wavelet transformations for the fault detection of induction motor with stator current. In: SPRINGER. *International Conference on Intelligent Computing*. [S.l.], 2006. p. 557–569. Citado na página 113.

LI, R.; SOPON, P.; HE, D. Fault features extraction for bearing prognostics. *Journal of Intelligent Manufacturing*, Springer, v. 23, n. 2, p. 313–321, 2012. Citado na página 19.

LI, W.; MONTI, A.; PONCI, F. Fault detection and classification in medium voltage dc shipboard power systems with wavelets and artificial neural networks. *IEEE Trans. Instrum. Meas.*, n. 11. Citado na página 72.

LI, X. et al. Application of wavelet packet analysis for fault detection in electro-mechanical systems based on torsional vibration measurement. *Mechanical Systems and Signal Processing*, Elsevier, v. 17, n. 6, p. 1219–1235, 2003. Citado na página 36.

LIN, F.-J.; SHEN, P.-H.; KUNG, Y.-S. Adaptive wavelet neural network control for linear synchronous motor servo drive. *IEEE Trans. Magn.*, v. 41, n. 12, p. 4401–4412, 2005. ISSN 0018-9464. Disponível em: <<http://ieeexplore.ieee.org/document/1556712/>>. Citado na página 24.

LIN, L.; SONG, J.; TIAN, M. Characteristic analysis of asynchronous motor fault signal based on wavelet packet decomposition. In: IEEE. *Electrical Machines and Systems (ICEMS), 2011 International Conference on*. [S.l.], 2011. p. 1–3. Citado na página 45.

LIN, L. et al. Application of wavelet neural network on turbo-generator set fault diagnosis system. In: IEEE. *Control Conference, 2006. CCC 2006. Chinese*. [S.l.], 2006. p. 1311–1314. Citado 2 vezes nas páginas 21 e 23.

LIRA, M. M. d. S. *Wavelets na compactação e processamento de sinais de distúrbios em sistemas de potência classificados via redes neurais artificiais*. 117 p. Tese (Doutorado) — UFPE, 2004. Citado 4 vezes nas páginas 45, 71, 74 e 126.

LIU, M.; TSE, C. K.; WU, J. A wavelet approach to fast approximation of steady-state waveforms of power electronics circuits. *Int. J. Circuit Theory Appl.*, v. 31, n. 6, p. 591–610, 2003. ISSN 00989886. Citado na página 19.

- MALLAT, S. *A wavelet tour of signal processing*. [S.l.: s.n.], 1998. ISBN 012466606X. Citado 3 vezes nas páginas 34, 35 e 39.
- MALLAT, S.; HWANG, W. L. Singularity detection and processing with wavelets. *IEEE Trans. Inf. Theory*, v. 38, n. 2, p. 617–643, 1992. ISSN 15579654. Citado na página 47.
- MALLAT, S. G. Multiresolution approximations and wavelet orthonormal bases of $L^2(\mathbb{R})$. *Trans. Am. Math. Soc.*, n. 1, p. 69–69. ISSN 0002-9947. Citado na página 37.
- MALLAT, S. G. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 11, n. 7, p. 674–693, 1989. Citado 2 vezes nas páginas 37 e 72.
- MARTINEZ, E. Z.; LOUZADA-NETO, F.; PEREIRA, d. B. B. A curva roc para testes diagnósticos. *Cad Saúde Coletiva*, v. 11, n. 1, p. 7–31, 2003. Citado na página 133.
- MARTINS, J. F.; PIRES, V. F.; PIRES, A. Unsupervised neural-network-based algorithm for an on-line diagnosis of three-phase induction motor stator fault. *IEEE Transactions on Industrial Electronics*, v. 54, n. 1, p. 259–264, 2007. ISSN 02780046. Citado na página 22.
- MATI, D. et al. Support vector machine classifier for diagnosis in electrical machines: Application to broken bar. *Expert Syst. Appl.*, v. 39, p. 8681–8689, 2012. Citado na página 49.
- MATIĆ, D. et al. Support vector machine classifier for diagnosis in electrical machines: Application to broken bar. *Expert Systems with Applications*, Elsevier, v. 39, n. 10, p. 8681–8689, 2012. Citado na página 113.
- MCCOY, R. et al. *Improved motors for utility applications. Volume 1. Industry assessment study: update and analysis. Final report*. [S.l.], 1985. Citado 2 vezes nas páginas 11 e 20.
- MEHALA, N.; DAHIYA, R. Motor current signature analysis and its applications in induction. *Int. J.*, v. 2, n. 1, p. 29–35, 2007. Citado na página 23.
- MIAO, Q.; HUANG, H.-Z.; FAN, X. A comparison study of support vector machines and hidden markov models in machinery condition monitoring. *Journal of Mechanical Science and Technology*, Springer, v. 21, n. 4, p. 607–615, 2007. Citado na página 25.
- MIAO, Q.; HUANG, H. Z.; FAN, X. Singularity detection in machinery health monitoring using lipschitz exponent function. *J. Mech. Sci. Technol.*, 2007. ISSN 1738494X. Citado 3 vezes nas páginas 24, 49 e 72.
- MIAO, Q.; MAKIS, V. Condition monitoring and classification of rotating machinery using wavelets and hidden markov models. *Mechanical systems and signal processing*, Elsevier, v. 21, n. 2, p. 840–855, 2007. Citado na página 24.
- MISITI, M. et al. Wavelet toolbox. *The MathWorks Inc., Natick, MA*, 1996. Citado na página 32.
- MØLLER, M. F. A scaled conjugate gradient algorithm for fast supervised learning. *Neural networks*, Elsevier, v. 6, n. 4, p. 525–533, 1993. Citado na página 61.
- MONTEIRO, M. V. V. *Detecção De Falhas Em Motores Elétricos Utilizando A Transformada Wavelet Packet E Métodos De Redução De Dimensionalidade*. Tese (Doutorado) — UNICAMP, 2014. Citado 4 vezes nas páginas 46, 47, 73 e 74.

MOTA, H. de O. *Processamento de sinais de descargas parciais em tempo real com base em wavelets e seleção de coeficientes adaptativa espacialmente*. Tese (Doutorado), 2011. Citado na página 114.

NEJJARI, H.; BENBOUZID, M. E. H. Monitoring and diagnosis of induction motors electrical faults using a current park's vector pattern learning approach. *IEEE Trans. Ind. Appl.*, v. 36, n. 3, p. 730–735, 2000. ISSN 00939994. Citado na página 22.

NING, W. G. . J. Wavelet-based disturbance analysis for power system wide-area monitoring. v. 63, n. 11, November 2014. Citado na página 72.

NIU, G. et al. Decision-level fusion based on wavelet decomposition for induction motor fault diagnosis using transient current signal. *Expert Systems with Applications*, Elsevier, v. 35, p. 918–928, 2008. Citado na página 25.

OCAK, H.; LOPARO, K. A new bearing fault detection and diagnosis scheme based on hidden markov modeling of vibration signals. *IEEE Int. Conf. Acoust. Speech, Signal Process. Proc. (Cat. No.01CH37221)*, v. 5, p. 3141–3144, 2001. ISSN 0-7803-7041-4. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=940324>>. Citado na página 24.

OLIVEIRA, H. de. *Análise de Fourier e Wavelets*. [S.l.]: Editora Universitária da UFPE, 2007. Citado na página 35.

OLIVEIRA, H. M. de. *Análise de sinais para engenheiros*. [S.l.]: Brasport, 2007. Citado 2 vezes nas páginas 31 e 38.

OLIVEIRA, M. O. Proteção diferencial adaptativa de transformadores de potência baseada na análise de componentes wavelets. 2013. Citado 2 vezes nas páginas 36 e 37.

OPPENHEIM, A.; SCHAFER, R. *Processamento em tempo discreto de sinais*. [S.l.]: Pearson, 2013. Citado na página 48.

PAIVA, H. M.; MACIEL, H. S.; MONTENEGRO, C. Wavelet-packet identification of dynamic systems in frequency subbands. 2005. Citado na página 45.

PALÁCIOS, R. H. C. et al. Fault identification in the stator winding of induction motors using pca with artificial neural networks. *Journal of Control, Automation and Electrical Systems*, Springer, v. 27, n. 4, p. 406–418, 2016. Citado 2 vezes nas páginas 113 e 114.

PARDO, M.; SBERVEGLIERI, G. Remarks on the use of multilayer perceptrons for the analysis of chemical sensor array data. *IEEE Sensors Journal*, IEEE, v. 4, n. 3, p. 355–363, 2004. Citado na página 60.

PARK, C.-W. et al. Generator fault detection technique using detailed coefficients ratio by daubechies wavelet transform. In: IEEE. *Power & Energy Society General Meeting, 2009. PES'09. IEEE*. [S.l.], 2009. p. 1–7. Citado na página 21.

PIMENTA, M. F.; PAVLIK, S. A. Conversor a/d baseado em compressão por transformadas wavelet. 2015. Citado na página 75.

PINTO, S. C. D. *Análise de Formas 3D usando Wavelets 1D, 2D e 3D*. Tese (Doutorado) — Universidade de Sao Paulo, 2005. Citado na página 31.

PÖYHÖNEN, S.; JOVER, P.; HYÖTYNIEMI, H. Independent component analysis of vibrations for fault diagnosis of an induction motor. In: *Proceedings of IASTED International Conference Circuits, Signals, and Systems*. [S.l.: s.n.], 2003. p. 19–21. Citado na página 27.

RIEDMILLER, M.; BRAUN, H. A direct adaptive method for faster backpropagation learning: The rprop algorithm. In: IEEE. *Neural Networks, 1993., IEEE International Conference on*. [S.l.], 1993. p. 586–591. Citado na página 64.

ROBERTS, M. J. *Fundamentos de sinais e sistemas*. [S.l.]: McGraw Hill Brasil, 2009. Citado 2 vezes nas páginas 32 e 48.

ROMERO-TRONCOSO, R. J. et al. Fpga-based online detection of multiple combined faults in induction motors through information entropy and fuzzy inference. *IEEE Transactions on Industrial Electronics*, IEEE, v. 58, n. 11, p. 5263–5270, 2011. Citado 2 vezes nas páginas 27 e 48.

SADEGHIAN, A.; YE, Z.; WU, B. Online detection of broken rotor bars in induction motors by wavelet packet decomposition and artificial neural networks. In: *IEEE Trans. Instrum. Meas.* [S.l.: s.n.], 2009. ISBN 0780370678. ISSN 00189456. Citado na página 24.

SANTIAGO, D.; PEDERIVA, R. Diagnóstico de falhas em máquinas rotativas utilizando transformada de wavelet e redes neurais artificiais. *Faculdade de Engenharia Mecânica, Universidade Estadual de Campinas, Tese de Doutorado*, 2004. Citado 2 vezes nas páginas 32 e 33.

SANTOS, F. M. d. C. *Identificação de falhas em motores de indução trifásicos usando sistemas inteligentes*. Tese (Doutorado) — Universidade de São Paulo, 2013. Citado 2 vezes nas páginas 22 e 49.

SANTOS, G. S. de Melo e. *Aplicação de Redes Neurais para Diagnóstico de Poluição em Isoladores de Alta Tensão*. 2008. Citado 2 vezes nas páginas 49 e 94.

SEERA, M. et al. Condition monitoring of induction motors: a review and an application of an ensemble of hybrid intelligent models. *Expert Systems with Applications*, Elsevier, v. 41, n. 10, p. 4891–4903, 2014. Citado 2 vezes nas páginas 19 e 75.

SHANLIN, K. S. K.; YUZHE, K. Y. K.; HUANZHEN, Z. H. Z. Application of wavelet transform for fault pattern recognition and analysis of power system generator. *2009 Chinese Control Decis. Conf.*, p. 3908–3911, 2009. Citado na página 24.

SHERLOCK, B. G.; MONRO, D. M. On the space of orthonormal wavelets. *IEEE Trans. SIGNAL Process.*, v. 46, n. 6, 1998. Citado 3 vezes nas páginas 72, 73 e 126.

SHI, P.; CHEN, Z.; VAGAPOV, Y. Fault detection and diagnosis of induction machines based on wavelet and probabilistic neural network. *International Journal of Computer Applications*, Foundation of Computer Science, v. 69, n. 14, 2013. Citado na página 113.

SILVA, I. N. da; SPATTI, D. H.; FLAUZINO, R. A. Redes neurais artificiais para engenharia e ciências aplicadas curso prático. *Artliber*, SciELO Brasil, 2010. Citado 4 vezes nas páginas 51, 57, 58 e 127.

SILVA, K. et al. Haar wavelet-based method for fast fault classification in transmission lines. In: IEEE. *Transmission & Distribution Conference and Exposition: Latin America, 2006. TDC'06. IEEE/PES*. [S.l.], 2006. p. 1–5. Citado na página 21.

STRANG, G. *Introduction to Linear Algebra*. 5. ed. [S.l.]: Wellesley-Cambridge Press, 2016. ISBN 0980232775,9780980232776. Citado na página 39.

SU, H.; CHONG, K. T. Induction machine condition monitoring using neural network modeling. *IEEE Trans. Ind. Electron.*, n. 1, p. 241–249. ISSN 02780046. Citado 3 vezes nas páginas 24, 51 e 113.

TABACNIKS, M. H. et al. Conceitos básicos da teoria de erros. *São Paulo, USP*, 2003. Citado na página 48.

TAPLAK, H.; UZMAY, I.; YILDIRIM, Ş. An artificial neural network application to fault detection of a rotor bearing system. *Industrial Lubrication and Tribology*, Emerald Group Publishing Limited, v. 58, n. 1, p. 32–44, 2006. Citado na página 26.

THOMSON, W.; RANKIN, D. Case histories of rotor winding fault diagnosis in induction motors. p. 798–819, 1987. Citado na página 20.

THOMSON, W. T.; FENGER, M. Current signature analysis to detect induction motor faults. *IEEE Ind. Appl. Mag.*, Ontario, v. 7, n. 4, p. 26–34, 2001. ISSN 10772618. Citado 3 vezes nas páginas 11, 20 e 23.

THOMSON, W. T.; GILMORE, R. J. Motor current signature analysis to detect faults in induction motor drives—fundamentals, data interpretation, and industrial case histories. p. 145–156, 2003. Citado 3 vezes nas páginas 21, 22 e 67.

THORSEN, O.; DALVA, M. Condition monitoring methods, failure identification and analysis for high voltage motors in petrochemical industry. p. 109–113, 1997. Citado na página 19.

TIKKANEN, P. E. Nonlinear wavelet and wavelet packet denoising of electrocardiogram signal. *Biol. Cybern.*, v. 80, n. 4, p. 259–267, 1999. ISSN 0340-1200. Citado na página 47.

UZINSKI, J. C. *Momentos nulos e regularidade wavelet na detecção de falhas em sinais*. Tese (Doutorado), 2013. Citado 3 vezes nas páginas 71, 72 e 126.

VENKATAKRISHNAN, P.; SANGEETHA, S.; SUNDAR, M. Measurement of lipschitz exponent (le) using wavelet transform modulus maxima (wtmm). *International Journal of Scientific & Engineering Research*, v. 3, n. 6, 2012. Citado na página 47.

WALCZAK, B.; MASSART, D. Noise suppression and signal compression using the wavelet packet transform. *Chemom. Intell. Lab. Syst.*, v. 36, n. 2, p. 81–94, 1997. ISSN 01697439. Citado na página 47.

WANG, G.; GUO, L.; DUAN, H. Wavelet neural network using multiple wavelet functions in target threat assessment. *The Scientific World Journal*, Hindawi Publishing Corporation, v. 2013, 2013. Citado na página 23.

WANG, H.; CHEN, P. Fault diagnosis method based on kurtosis wave and information divergence for rolling element bearings. *WSEAS Transactions on System*, v. 8, n. 10, p. 1155–1165, 2009. Citado na página 19.

WEEKS, M. *Parocessamento digital de sinais usando MatLab e wavelets*. [S.l.]: LTC, 2012. ISBN 9788521621416. Citado 4 vezes nas páginas 39, 40, 42 e 43.

WIDODO, A.; YANG, B.-S. Application of nonlinear feature extraction and support vector machines for fault diagnosis of induction motors. *Expert Systems with Applications*, v. 33, n. 1, p. 241–250, 2007. Citado na página 24.

YANG, B.-S. et al. Fault diagnosis of induction motor based on decision trees and adaptive neuro-fuzzy inference. *Expert Systems with Applications*, Elsevier, v. 36, n. 2, p. 1840–1849, 2009. Citado na página 22.

YAZICI, B. et al. An adaptive, on-line, statistical method for bearing fault detection using stator current. In: IEEE. *Industry Applications Conference, 1997. Thirty-Second IAS Annual Meeting, IAS'97., Conference Record of the 1997 IEEE*. [S.l.], 1997. v. 1, p. 213–220. Citado na página 26.

YE, Z.; WU, B. Online rotor bar breakage detection of three phase induction motors by wavelet packet decomposition and artificial neural network. In: IEEE. *Power Electronics Specialists Conference, 2001. PESC. 2001 IEEE 32nd Annual*. [S.l.], 2001. v. 4, p. 2209–2216. Citado 2 vezes nas páginas 21 e 24.

YE, Z.; WU, B.; SADEGHIAN, A. Current signature analysis of induction motor mechanical faults by wavelet packet decomposition. *IEEE Trans. Ind. Electron.*, v. 50, n. 6, 2003. Citado na página 24.

ZAREI, J.; POSHTAN, J. Bearing fault detection using wavelet packet transform of induction motor stator current. *Tribol. Int.*, v. 40, n. 5, p. 763–769, 2007. ISSN 0301679X. Citado na página 25.

ZHANG, L.; TIAN, F. Performance study of multilayer perceptrons in a low-cost electronic nose. *IEEE Transactions on Instrumentation and Measurement*, IEEE, v. 63, n. 7, p. 1670–1679, 2014. Citado na página 60.

ZHANG, P. et al. A survey of condition monitoring and protection methods for medium-voltage induction motors. *Industry Applications, IEEE Transactions on*, IEEE, v. 47, n. 1, p. 34–46, 2011. Citado na página 19.

ZHANG, X.; WANG, Y. Wavelet-neural-network based robust sliding-mode control for induction motor. *J. Inf. Comput. Sci.*, v. 8, n. 7, p. 1209–1216, 2011. Disponível em: <<http://www.joics.com>>. Citado na página 24.

ZHENG-YOU, H.; XIAOQING, C.; GUOMING, L. Wavelet entropy measure definition and its application for transmission line fault detection and identification;(part i: Definition and methodology). In: IEEE. *Power System Technology, 2006. PowerCon 2006. International Conference on*. [S.l.], 2006. p. 1–6. Citado na página 80.

ZHONGMING, Y. E.; BIN, W. U. A review on induction motor online fault diagnosis. In: *3rd Power Electron. Motion Control Conf*. [S.l.: s.n.], 2000. p. 1353–1358. ISBN 7-80003-464-X. Citado na página 24.

APÊNDICE A – RESULTADOS EXPERIMENTAIS - TÉCNICA INICIAL

METODOLOGIA DE ANÁLISE

Aspectos dos dados experimentais

Os dados levantados para realização de todo processo de identificação de falhas são obtidos a partir das assinaturas de corrente de fase do motor, metodologia essa também conhecida como (MCSA). O acionamento do motor de indução de 1/3 CV e 4 polos foi através de partida direta. O motor recebeu diretamente uma tensão de 380V, 60 Hz de linha em seus terminais de estator.

As aquisições foram realizadas com o motor acionado em vazio através do dispositivo *NI – DAQ – USB6001* de fabricação *National Instruments*[®].

Wavelet utilizada para decomposição dos sinais

A escolha da wavelet-mãe é parte importante no processo de análise dos sinais de assinatura de corrente, pois segundo (LIRA, 2004), wavelets com comportamentos mais suaves permitem melhor resolução no domínio da frequência, enquanto que variações em pequenos intervalos de tempo caracterizam melhor resolução no domínio do tempo. Uma abordagem útil para escolha da wavelet em questão é proposta por (UZINSKI, 2013) que, através de técnicas analíticas, consegue estabelecer critérios para que as wavelets tenham regularidade e suavidade através de uma formulação que, basicamente é uma expansão da formulação de Sherlock & Monro (SHERLOCK; MONRO, 1998).

Neste sentido a wavelet escolhida foi a Daubechies-4.

Classificador de falhas

Nesta seção será apresentado o classificador baseado em RNA e DWT. Basicamente, foram realizados três treinamentos com a rede neural. O primeiro consiste em, diretamente, aplicar o sinal do quinto detalhe obtido através da AMR, o que se tornou inviável devido à quantidade de nós de entrada necessária para receber os cinco ciclos de cada uma das três fases, além do esforço computacional necessário para treinamento da rede.

O passo seguinte foi a redução da quantidade de informação presente nesses sinais abrangendo somente seus máximos locais. Do ponto de vista computacional, o esforço para realização desta tarefa foi reduzido. O percentual de acertos também foi significativo (cerca de 96% das amostras de treinamento, teste e validação), porém a quantidade de nós de entrada da rede permaneceu a mesma.

A nova alternativa adotada foi a remoção dos elementos nulos dos sinais do segundo caso, que utiliza apenas os máximos locais (redução de dimensionalidade). Essa alternativa se mostrou interessante, pois além de uma redução de esforço computacional em comparação às duas tentativas anteriores, o número de nós de entrada fica consideravelmente reduzido a quase metade da quantidade dos nós nos casos anteriores.

As arquiteturas de rede usadas foram uma rede alimentada adiante (*feedforward*) com camada única camada oculta constando de dez neurônios nesta. A segunda arquitetura de rede é uma *feedforward* de camadas múltiplas do tipo *multilayer perceptron* com especificamente cinco camadas: a camada com os nós de entrada, três camadas ocultas (10, 26 e 2 neurônios respectivamente) e a camada de saída com dois neurônios também.

A rede de camada oculta única foi treinada através do método *Scaled conjugate gradient backpropagation*, ou apenas.

A segunda topologia de rede foi treinada fazendo uso da técnica de *Resilient backpropagation* devido ao fato de redes multicamadas normalmente usarem funções de transferência sigmoide nas camadas ocultas. Estas funções são conhecidas por efeito de saturar as saídas dos neurônios para os valores limites deste tipo de função (0 e 1 no caso da sigmoide). As derivadas parciais destas funções assumiriam valores muito pequenos, próximos de zero, o que implicaria em lentidão no processo de treinamento. Nesta condição, variações muito pequenas no gradiente da função erro combinado com faixas de saturação das funções de ativação, fazem com que o processo de convergência seja muito lento (SILVA; SPATTI; FLAUZINO, 2010).

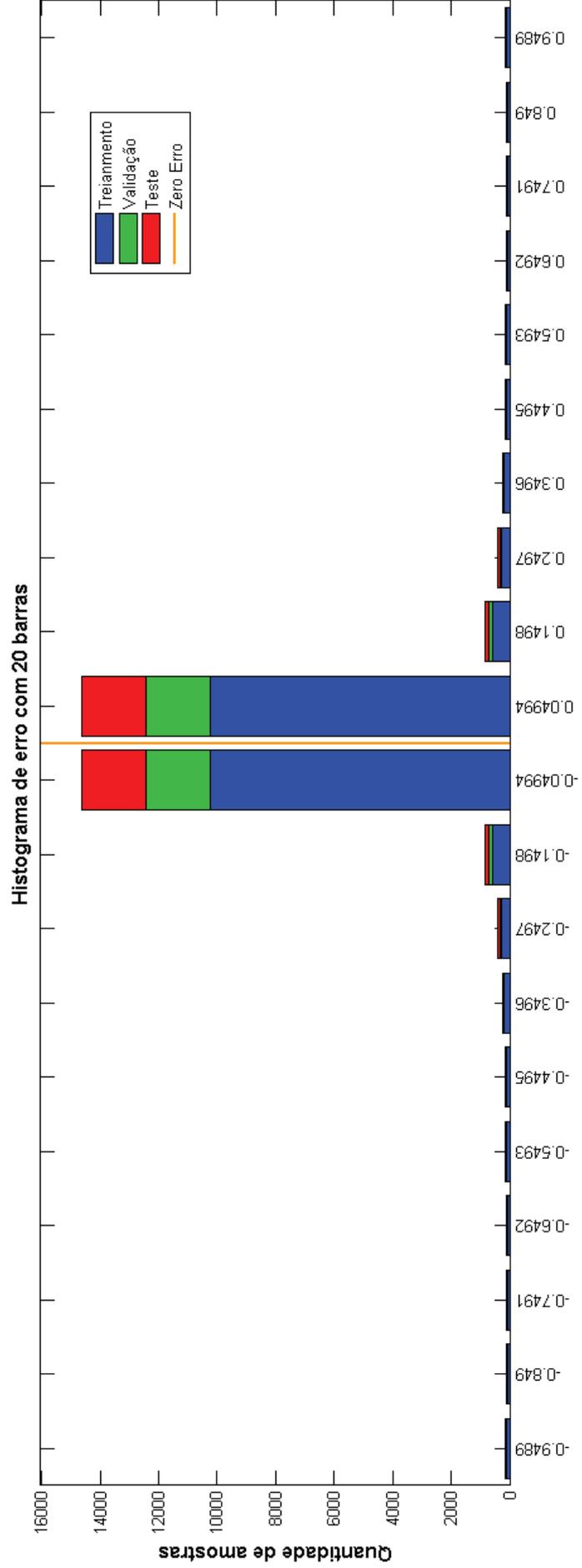
RESULTADOS PARA REDE PERCEPTRON COM APENAS UMA CAMADA

Treinamento utilizando os máximos locais

Os primeiros resultados aqui apresentados serão para o caso da rede treinada com os máximos locais do sinal do quinto detalhe da decomposição wavelet dos sinais de corrente de fase do motor. Inicialmente serão apresentados os resultados para a rede *perceptron* com apenas uma camada oculta.

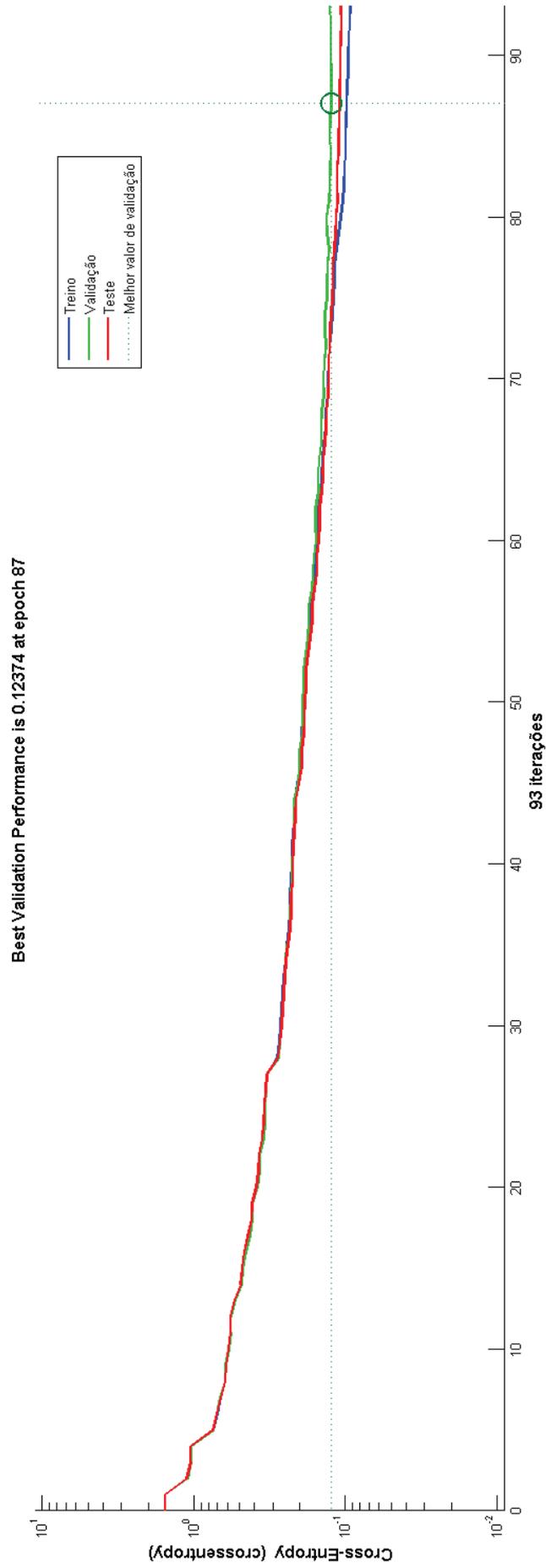
A Figura 51 Ilustra um histograma de 20 barras. Neste histograma estão indicadas as quantidades de amostras em relação a um dado percentual de erro. A linha central deste tipo de gráfico é também conhecida como linha de erro zero. A grandeza de referência para a medição do desempenho desta rede foi a *Cross Entropy*. Quanto mais próximo de zero estiver esta grandeza, menor será o erro cometido pela rede a variação desta grandeza pode ser observada na Figura 52.

Figura 51 – Histograma de erro para o treinamento com os máximos locais do quinto detalhe dos sinais de corrente de fase.



Erros = saídas desejadas - saídas fornecidas
Fonte: Próprio autor através da *Neural Networks Toolbox* do *MatLab*®

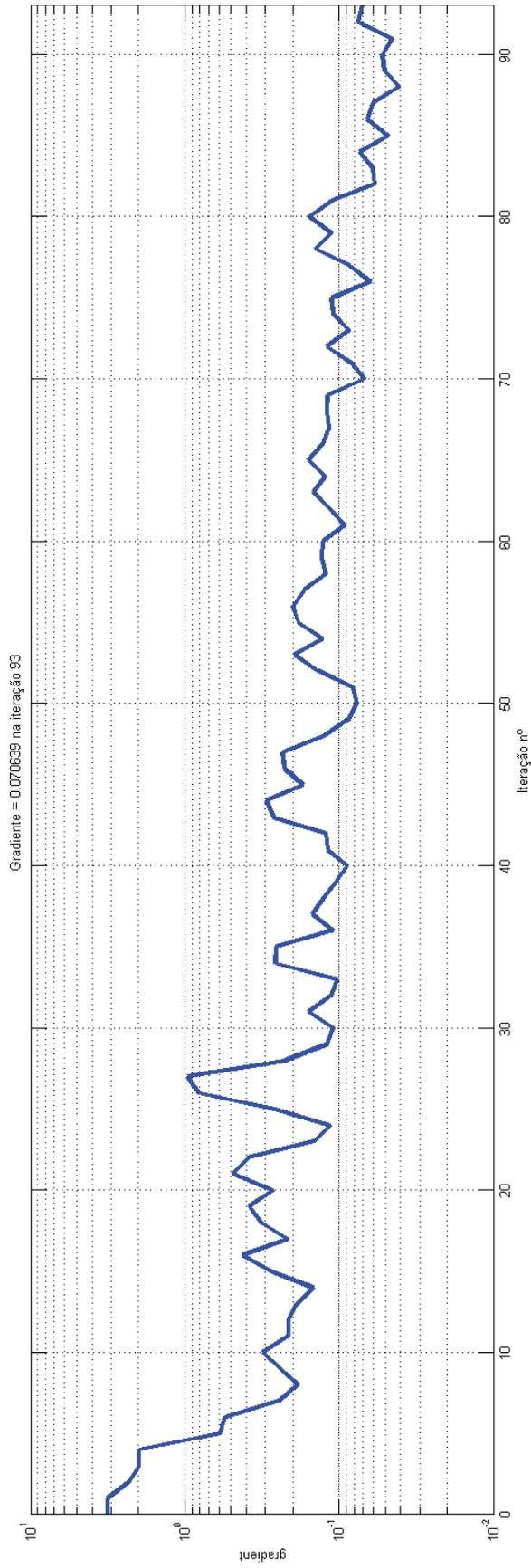
Figura 52 – Medida da entropia para determinação do desempenho da rede.



Fonte: Próprio autor através da *Neural Networks Toolbox* do *MatLab*®

Outro fator que pode dizer muito sobre o processo de treinamento da rede é o gradiente da função erro. Observar sua variação ao longo do processo pode ser bastante informativo no que diz respeito à convergência do processo de treinamento da rede. A Figura 53 ilustra bem este processo.

Figura 53 – Variação dos valores do gradiente da função erro ao longo do processo de treinamento da RNA.

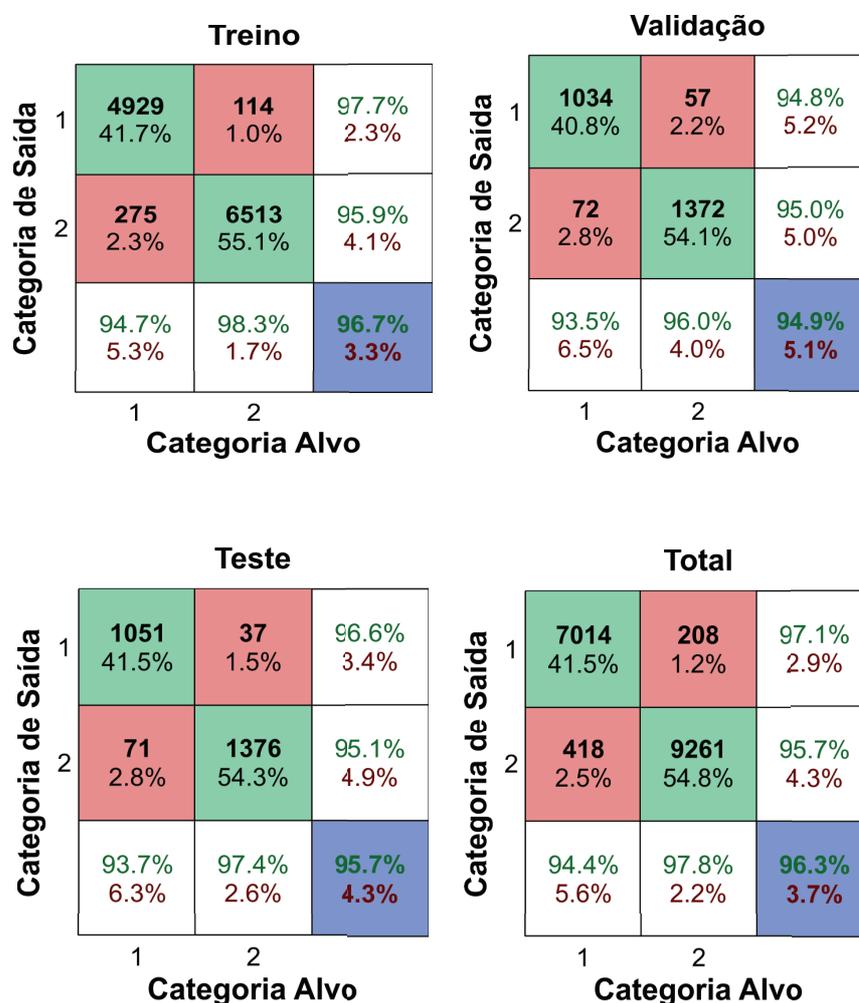


Fonte: Próprio autor através da *Neural Networks Toolbox* do *MatLab*®

Matrizes de Confusão

Outra medida que também é efetiva na verificação do desempenho de um modelo de classificação é a sua matriz de confusão. Esta ferramenta apresenta número de classificações corretas e as classificações previstas para categoria em um dado conjunto de exemplos. Os valores presentes na diagonal principal indicam o total de acertos. Ou seja, a quantidade de exemplos que a RNA aponta como sendo de uma categoria e estes de fato o são. A Figura 54 mostra as quatro matrizes de confusão para a primeira rede neural treinada (perceptron camada única) utilizando os máximos locais do quinto detalhe do sinal de corrente.

Figura 54 – Matrizes de confusão para o treinamento, teste, validação e Matriz de confusão total para o primeiro caso.



Fonte: Próprio autor

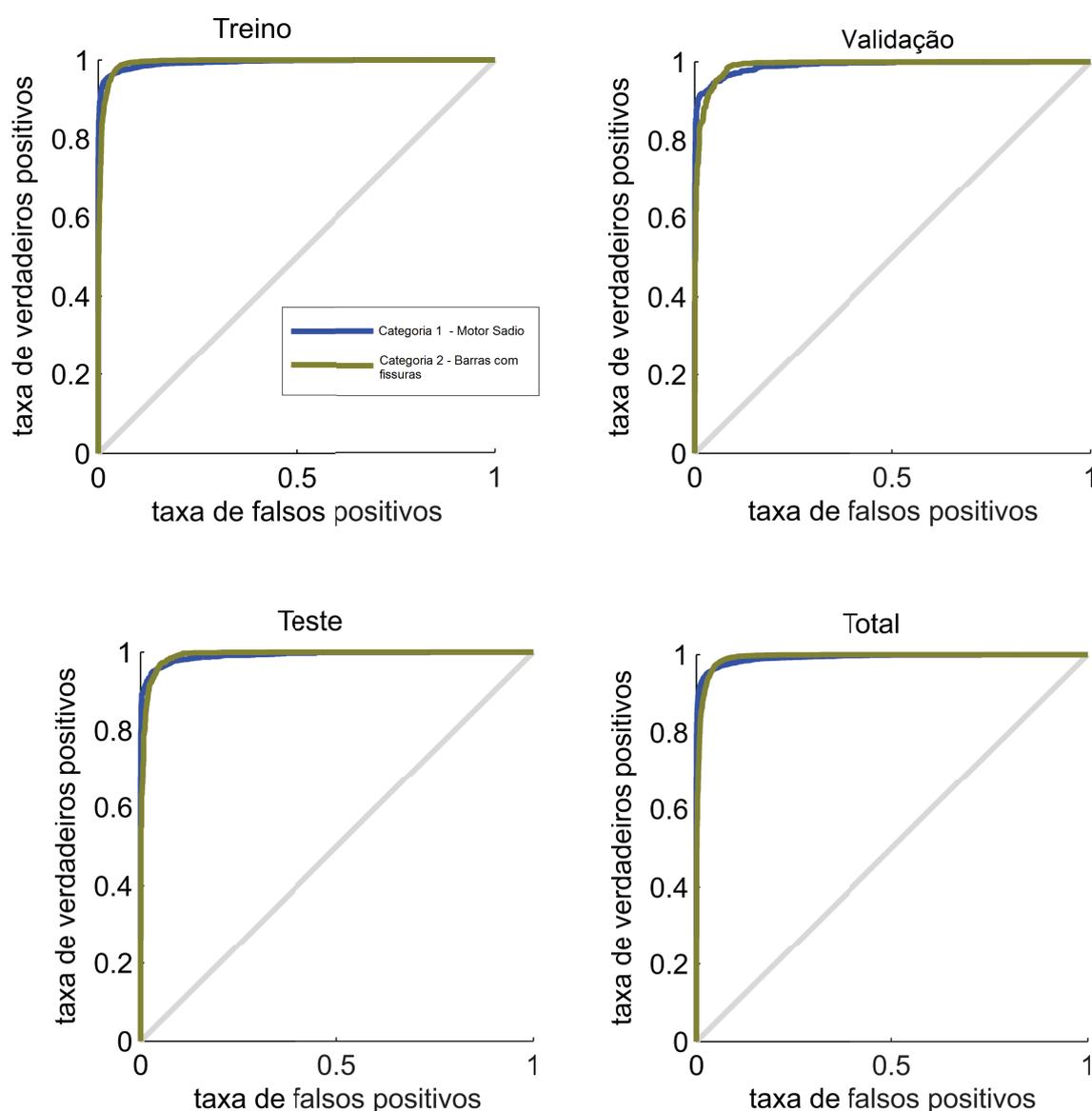
Curvas ROC

O próximo indicador de performance da rede são as curvas ROC (*Receiver Operating Characteristic*).

A análise ROC teve origem na teoria de decisão estatística e foi pela primeira vez usada durante a Segunda Guerra Mundial. O propósito desta ferramenta era verificar a eficiência dos operadores de radar (*receiver operators*) em distinguir um sinal de um ruído. A potencial utilidade da análise ROC em avaliar diagnósticos médicos foi desde então utilizada por vários autores e, subsequentemente, foi aplicada com sucesso a uma grande variedade de testes de diagnóstico e em particular no diagnóstico de imagem médica (BRAGA, 2000).

A Curva ROC tem sido bastante usada pela comunidade de Aprendizagem de Máquina pois, em geral, avaliar apenas a taxa de acerto de um classificador é uma métrica muito simples. A ideia central ao se utilizar estas curvas é ver em um gráfico bidimensional a quantidade relativa de resultados positivos verdadeiros e de falsos positivos (é o caso deste formato de classificação abordado neste trabalho) ou de negativos e falsos negativos. A taxa de verdadeiros positivos é colocada no eixo das ordenadas, enquanto a taxa de falsos positivos é disposta no eixo das abcissas (MARTINEZ; LOUZADA-NETO; PEREIRA, 2003). A Figuras 55 foi gerada a partir da função “plotroc” da *Neural Networks Toolbox* do *MatLab*[®] e evidenciam as curvas ROC para a RNA *perceptron* treinada com os máximos locais.

Figura 55 – Curvas ROC de treinamento, Validação, Teste e Total para uma rede perceptron classificando sinais utilizando seus máximos locais.

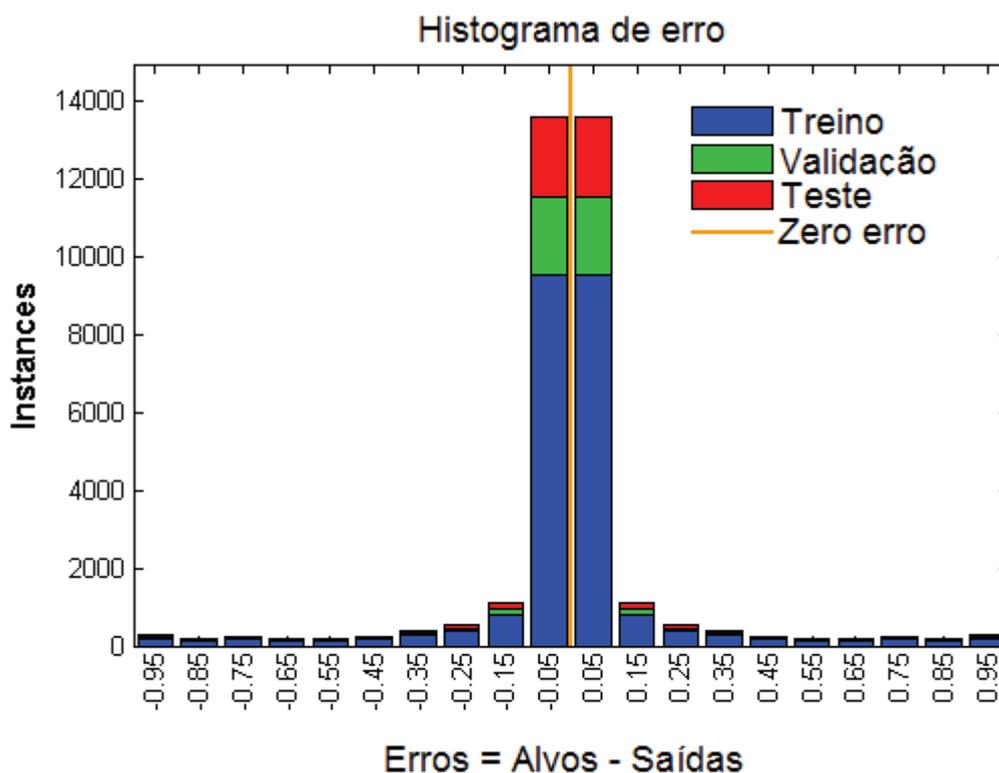


Fonte: próprio autor

Resultados do treinamento utilizando as amostras de treino sem elementos nulos

Devido ao grande número de nós de entrada que foram necessários para que a rede pudesse ser treinada com os sinais de detalhe retirados das correntes de fase, uma solução para redução da dimensionalidade destes se fez necessária. Tendo em vista que se trata de uma matriz contendo apenas os máximos locais dos sinais originais e que os elementos que não gozam dessa propriedade são substituídos por zero, uma solução viável do ponto de vista computacional é a remoção destes zeros. Do ponto de vista do treinamento da rede, foi necessário um número

Figura 56 – Histograma de erro para o caso com eliminação de elementos nulos do sinal de entrada.

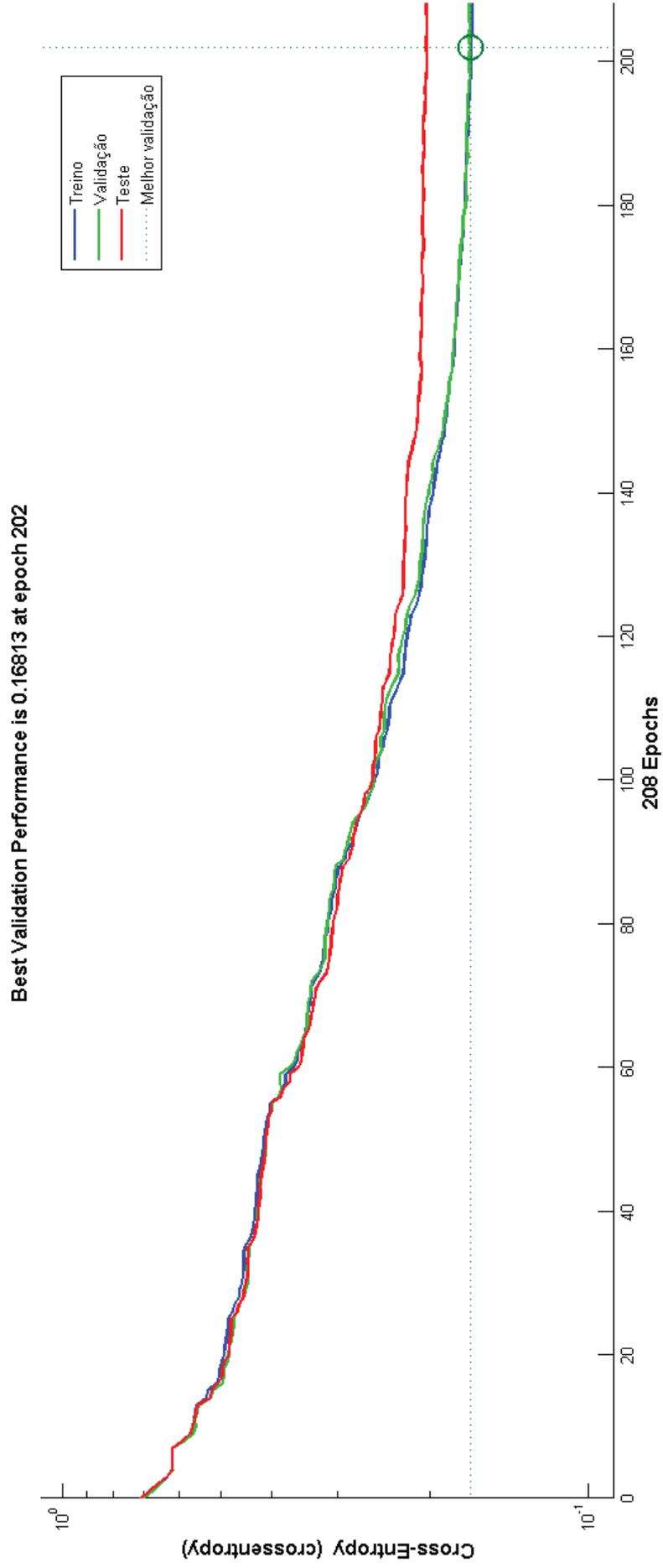


Fonte: Próprio autor

maior de iterações; contudo, a rede teve o número de nós de entrada consideravelmente reduzido (39 nós, em contrapartida a 66 nós no caso anterior). As classificações continuaram com um bom percentual de acertos (93%).

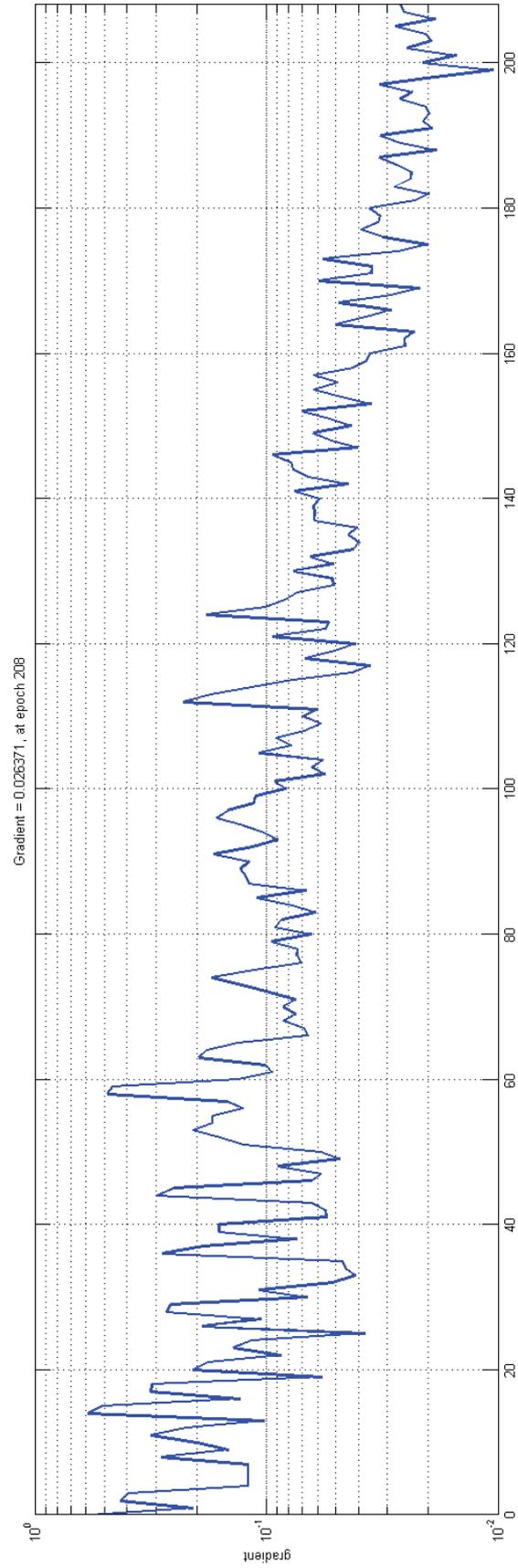
A Figura 56 mostra um novo histograma com os valores de erros e a quantidade de amostras correspondentes a estes dados valores. As figuras 57 e A exibem os valores de *Cross Entropy* e de gradientes para esta nova configuração da rede neural.

Figura 57 – Curva de desempenho através da *Cross Entropy*.



Fonte: Próprio autor

Figura 58 – Gradiente para o treino com eliminação dos elementos nulos.

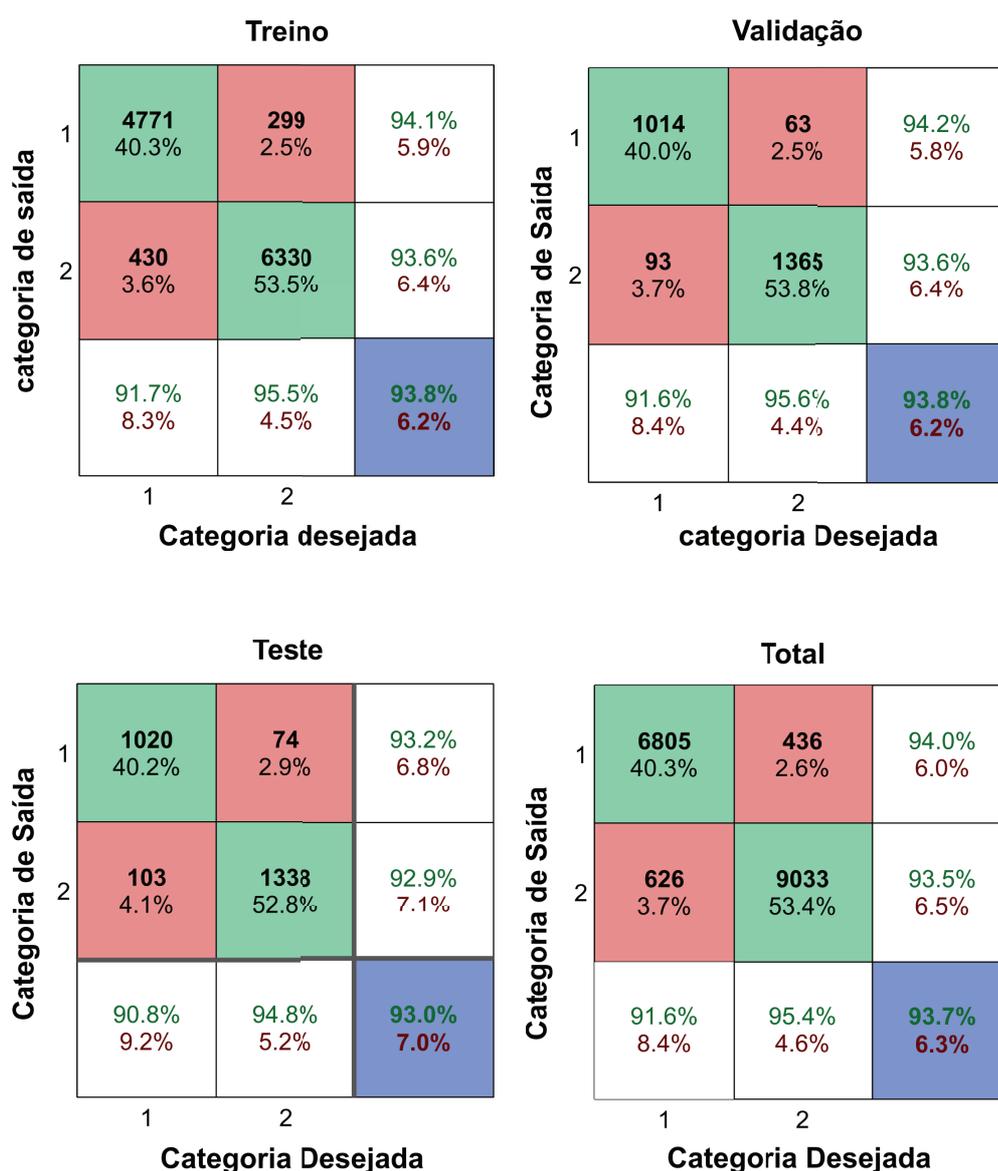


Fonte: Próprio autor

Matrizes de Confusão

Nas matrizes de confusão da Figura 59, é mostrada a diferença em valores percentuais. Na abordagem com as matrizes de máximos locais, o percentual total de acertos foi pouco mais de 96%, ao passo que com a remoção dos elementos nulos esse percentual cai para pouco mais de 93% o que não deixa de ser uma boa taxa de acertos.

Figura 59 – Matrizes de confusão de Treino, Teste, Validação e matriz de confusão Total para o caso da eliminação dos termos nulos dos sinais.



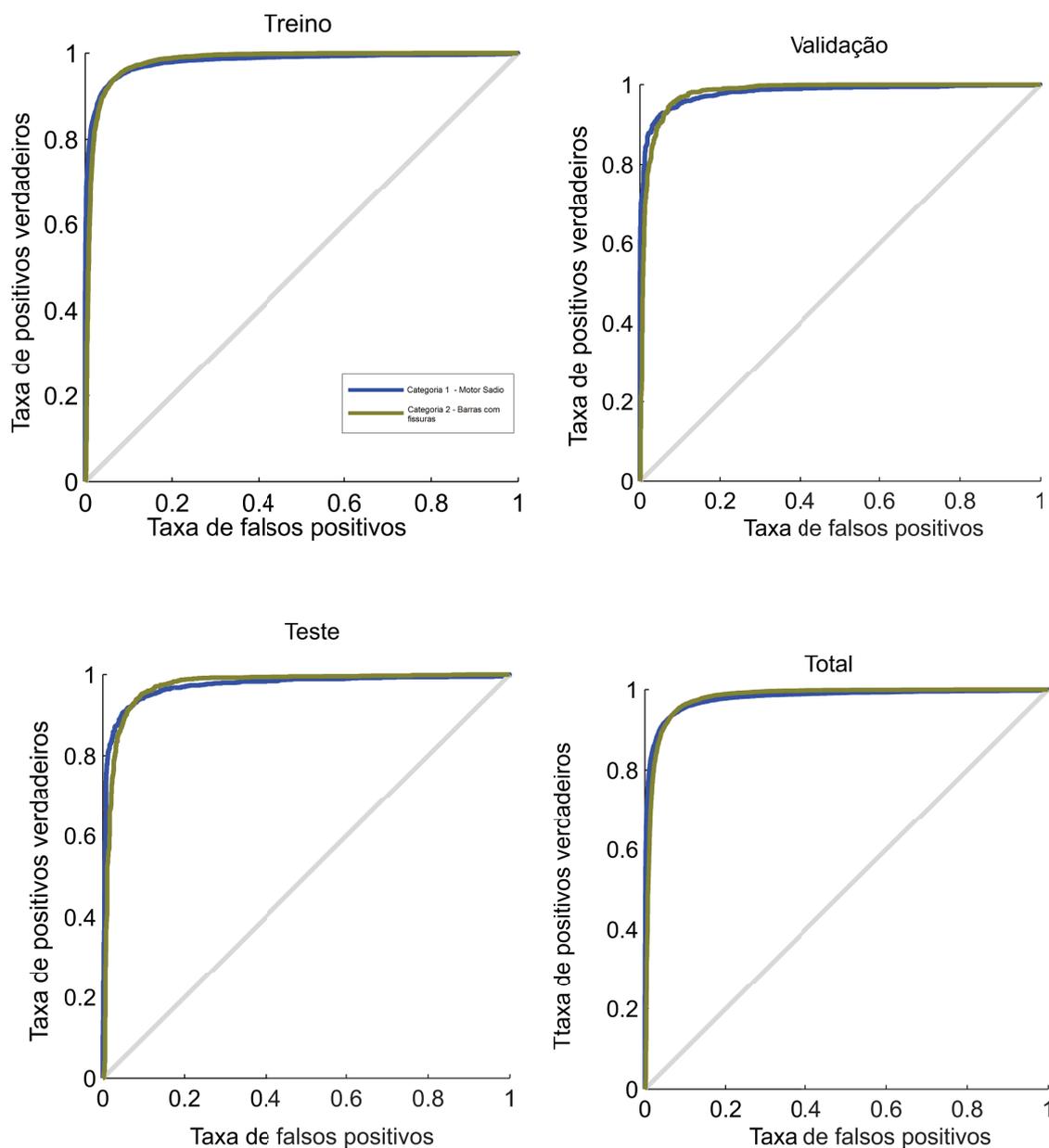
Fonte: Próprio autor

Curvas ROC

A curvas ROC na Figura 60 apresentam como particularidade o fato de serem menos acentuadas que no caso anterior. Isso se deve ao fato de a quantidade de acertos ser menor,

conforme já fora constatado anteriormente nas matrizes de confusão, isso gera uma redução nas taxas de verdadeiros positivos e, conseqüentemente uma diminuição das áreas sob as curvas além da menor acentuação.

Figura 60 – Curvas ROC de treinamento, validação, teste e total para a rede neural com redução do número de nós da entrada.



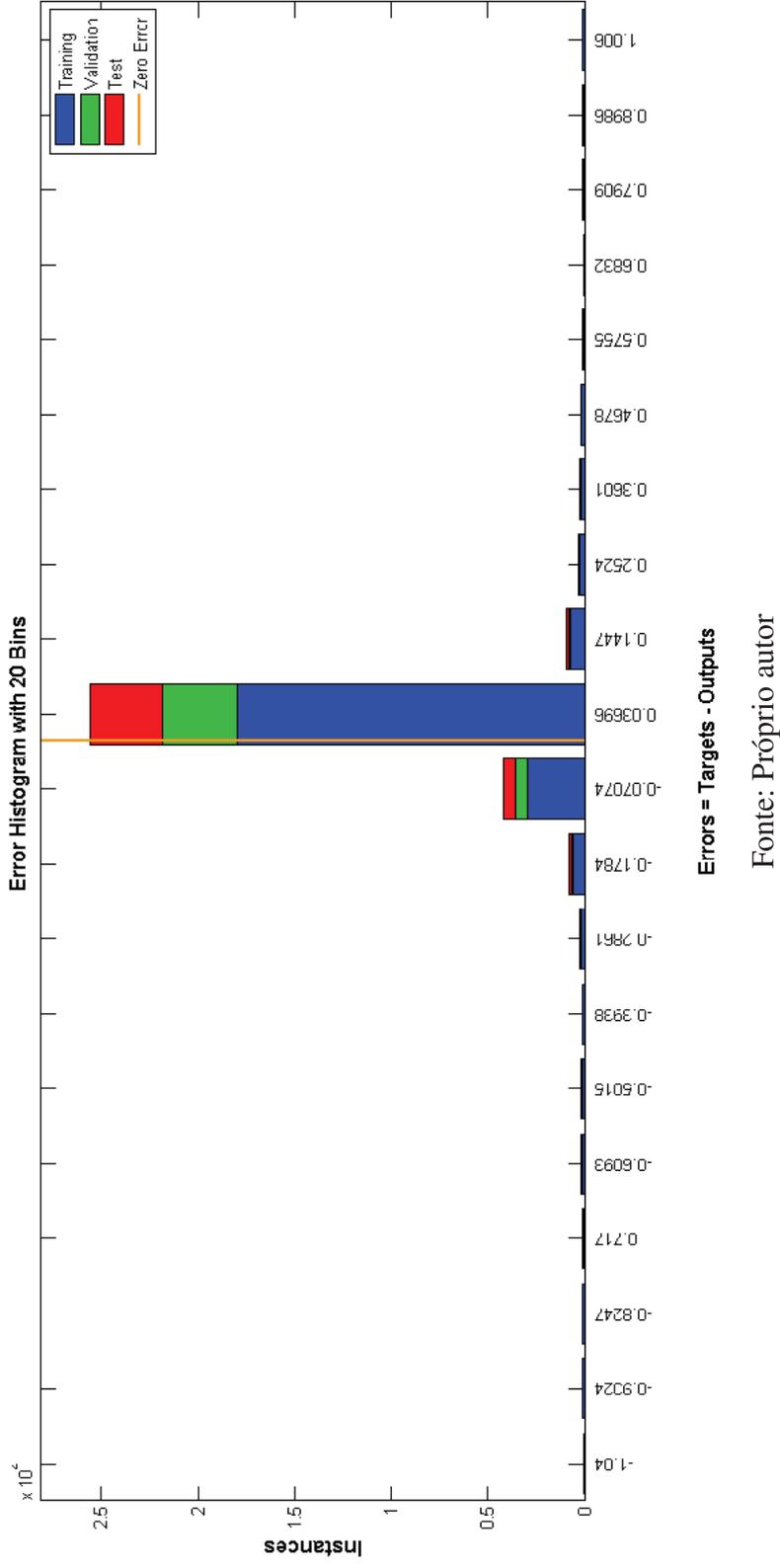
Fonte: Próprio autor

TREINAMENTO DA REDE NEURAL MLP

O resultado do treinamento com a rede de com apenas uma camada oculta para o caso da redução das dimensões dos sinais de entrada apresentou uma performance um tanto inferior para o caso de uma entrada de rede com os máximos locais e os elementos nulos. No intuito

de verificar se uma melhor performance poderia ser atingida com outra topologia de rede, uma rede *multilayer perceptron* (MLP) foi utilizada. Esta nova rede apresenta três camadas ocultas com 10, 26 e 2 neurônios, uma camada de saída também com dois neurônios e a camada de entrada com o número reduzido de neurônios de entrada (39 no caso deste conjunto de dados). O Histograma da Figura 61 mostra uma quantidade significativa de amostras que tiveram um erro bem próximo à linha de erro zero.

Figura 61 — Histograma de erro para o treinamento com eliminação dos elementos nulos dos sinais para uma rede Perceptron Multicamada.

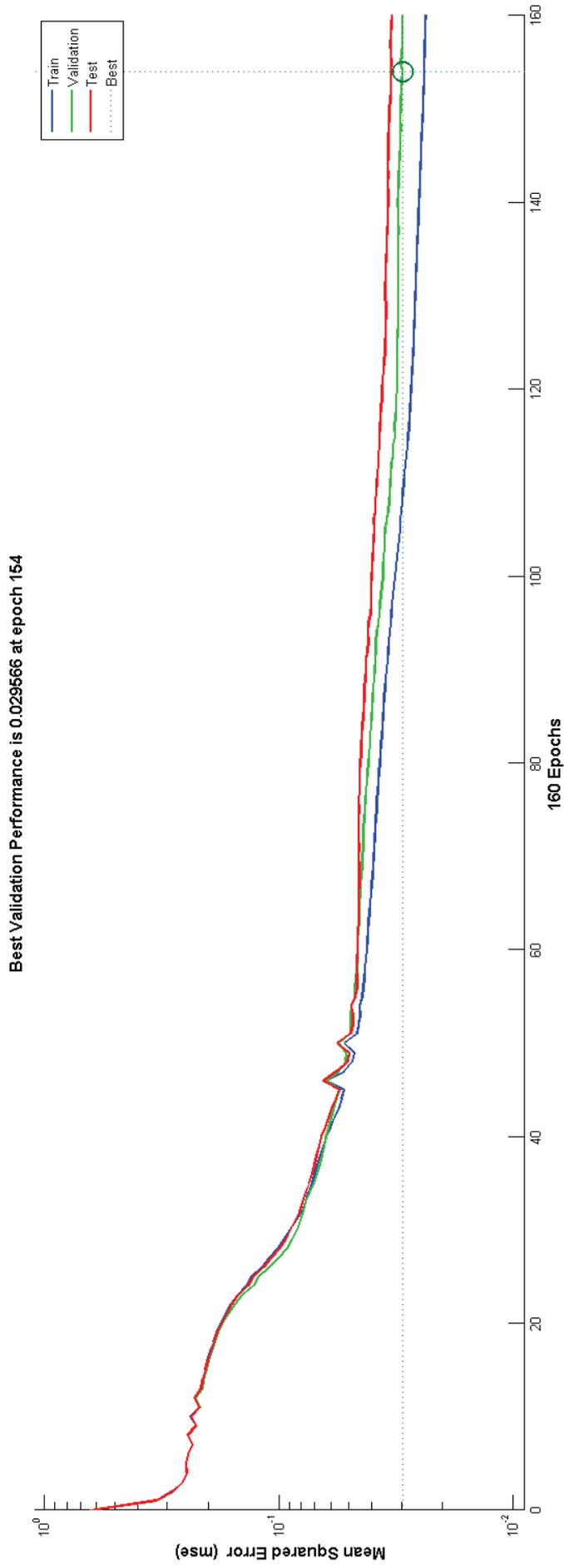


Errors = Targets - Outputs

Fonte: Próprio autor

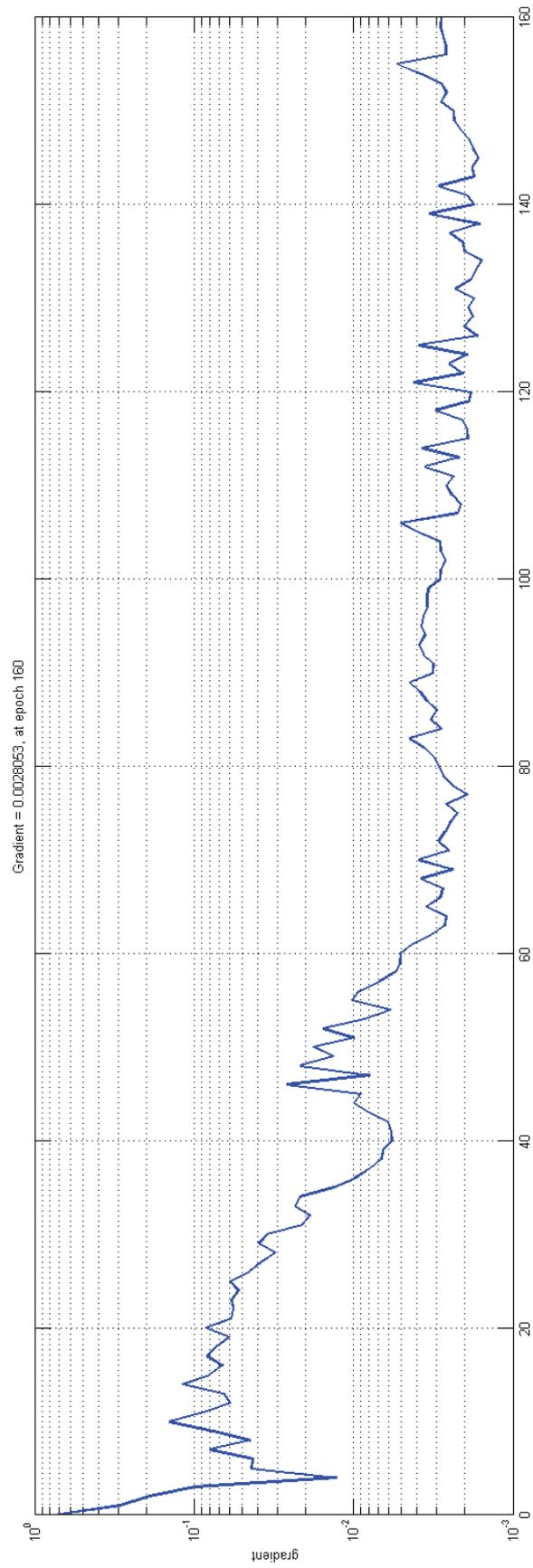
O número de iterações nesta nova topologia diminui de 208 para 160. O tempo de treinamento também diminui e os resultados são melhores que no caso com apenas uma camada oculta. A Figura 62 mostra a tendência do erro médio quadrático ao longo das 160 iterações realizadas ao longo do processo e a Figura 63 mostra o comportamento do gradiente para a mesma rede. Vale lembrar que o método utilizado aqui foi o *resilient backpropagation*.

Figura 62 – Medida de erro médio quadrático para determinação do desempenho da rede MLP.



Fonte: Próprio autor

Figura 63 – Variação dos valores do gradiente da função erro ao longo do processo de treinamento da rede MLP.

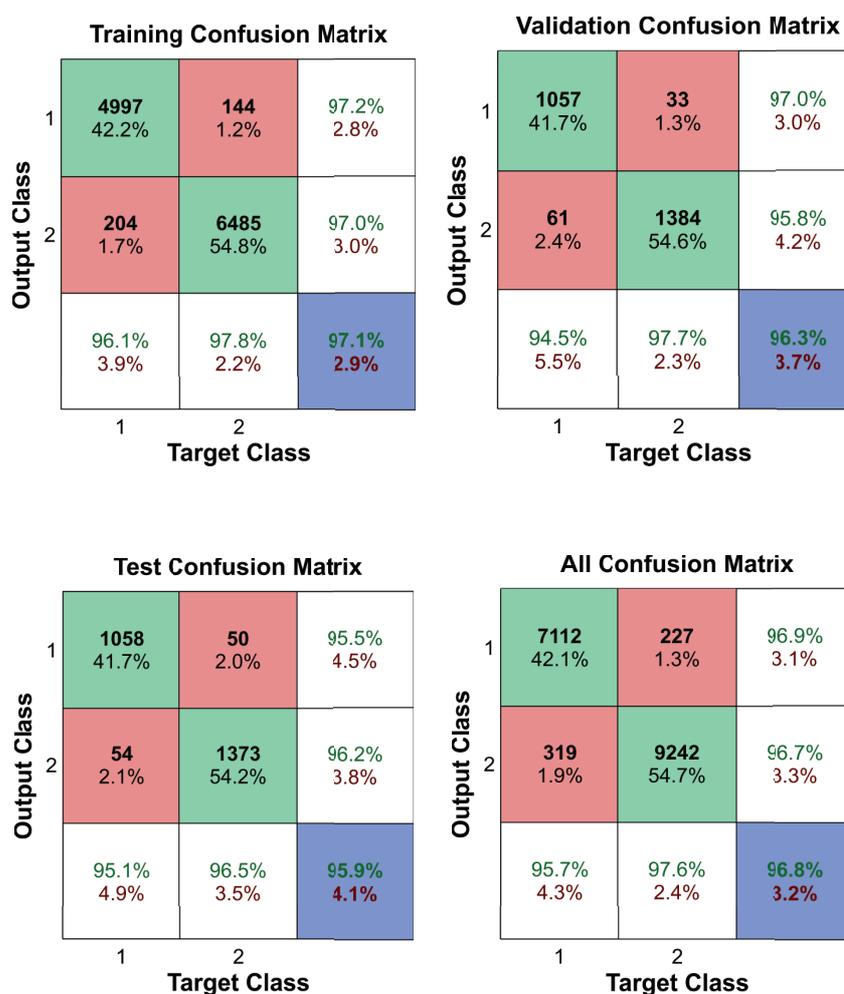


Fonte: Próprio autor

Matrizes de Confusão

As matrizes de Confusão presentes na Figura 64 mostram, para a rede MLP, um resultado superior ao caso de apenas uma camada oculta.

Figura 64 – Matrizes de confusão para o treinamento, teste, validação e Matriz de confusão total para a rede MLP.

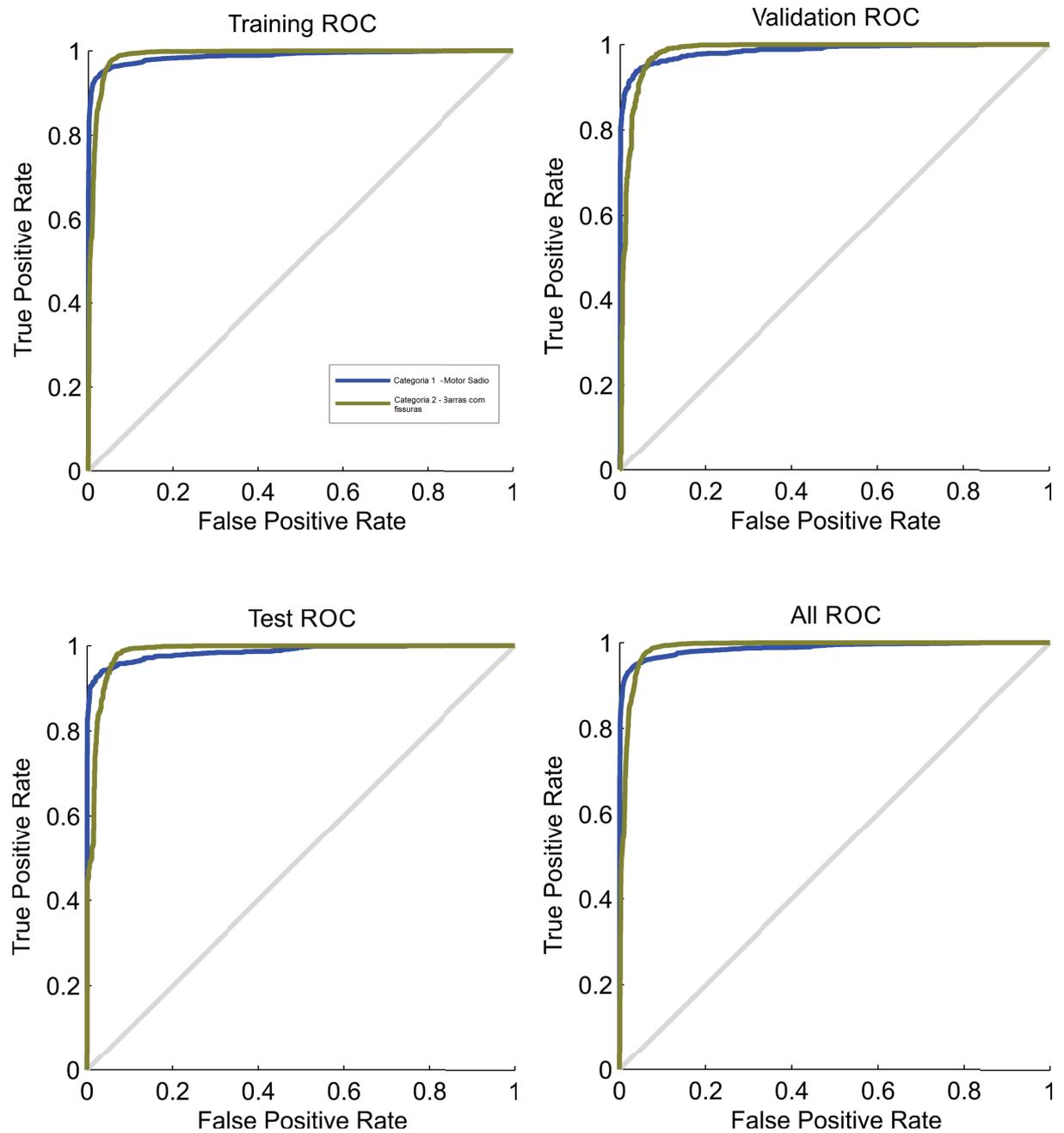


Fonte:Próprio autor

Curvas ROC

As curvas ROC para a topologia MLP apresentadas na Figura 65 mostram um comportamento mais acentuado que no caso de apenas uma camada oculta; o que significa uma maior taxa de verdadeiros positivos.

Figura 65 – Curvas ROC de treinamento para a rede MLP.



Fonte: Próprio autor.

APÊNDICE B – TREINAMENTO DE NOVAS CONFIGURAÇÕES DE REDES COM OS ATRIBUTOS SELECIONADOS

INTRODUÇÃO

No capítulo referente aos resultados experimentais, a Configuração D das redes MLP utilizadas foi eleita como preferível entre as demais. A saber, tal configuração consta com quatro camadas ocultas e dez neurônios em cada camada. A combinação de atributos que mostrou melhor desempenho em todas as camadas foi a 8 que é formada pelo centroide, desvio padrão e curtose. O objetivo deste adendo ao texto da tese é mostrar como a combinação 8 se comporta quando aplicada a topologias mais simples para a rede MLP. Todas as dez configurações aqui apresentadas tem apenas uma camada oculta e o número de neurônios nestas variam de 1 a 10.

RESULTADOS DOS TREINAMENTOS

Nas tabelas aqui aprestandas, são mostrados o número de neurônios na camada oculta, os percentuais de classificação correto e de classificações incorretas, o maiores valores de erros de teste, validação e global (ou seja, das classificações incorretas, quanto foi a maior discrepância entre o resultado esperado e o desenvolvido pela rede), número de iterações e tempo para treinamento.

As Tabelas 52 a 55 mostram todos o resultados do treinamento desde o percentual de acertos até o tepo de treinamento de cada caso.

Tabela 52 – Redes MLP treinadas a Rprop com uma camada oculta - problemas mecânicos

Neurônios na camada oculta	1	2	3	4	5	6	7	8	9	10
Classificações corretas	39,04%	72,11%	85,18%	82,22%	86,45%	90,87%	91,20%	91,11%	92,43%	83,82%
Classificações incorretas	60,96%	27,89%	14,82%	17,78%	13,55%	9,13%	8,80%	8,89%	7,57%	16,18%
Maior erro de teste	76,33%	70,57%	65,28%	66,39%	64,96%	63,28%	63,02%	62,50%	62,25%	64,59%
Maior erro validação	76,19%	69,68%	65,29%	65,83%	65,04%	62,49%	62,16%	62,64%	61,95%	64,60%
Maior erro global	76,25%	70,13%	65,57%	65,69%	64,24%	62,38%	62,08%	61,82%	61,40%	63,90%
Iterações	1894	4969	3497	3622	955	5096	3597	1900	2207	3079
tempo de treino (segundos)	21,62	61,15	47,57	51,59	14,75	82,90	62,85	28,23	35,56	51,16

Na Tabela 52, o destaque vai para as redes com 6, 7, 8 e 9 neurônios na camada oculta. Seus desempenhos foram todos superiores a 90%.

Na Tabela 53, os desempenhos e tempos de treinamento foram muito bons. A exceção da rede com apenas um neurônio, todas as outras configurações tiveram percentuais de acerto superior a 98%.

Tabela 53 – Redes MLP treinadas a Rprop com uma camada oculta - faltas a 25%

Neurônios na camada oculta	1	2	3	4	5	6	7	8	9	10
Classificações corretas	48,37%	98,40%	98,58%	99,17%	99,11%	99,17%	98,34%	98,93%	98,70%	98,10%
Classificações incorretas	51,63%	1,60%	1,42%	0,83%	0,89%	0,83%	1,66%	1,07%	1,30%	1,90%
Maior erro de teste	73,52%	53,49%	53,16%	52,78%	52,87%	52,80%	53,13%	52,76%	53,14%	53,26%
Maior erro validação	72,91%	53,11%	53,20%	53,05%	53,07%	53,17%	52,94%	53,07%	53,18%	53,06%
Maior erro global	72,89%	53,08%	52,82%	52,83%	52,54%	52,67%	52,70%	52,54%	52,74%	52,57%
Iterações	1197	315	169	239	262	180	135	169	167	136
tempo de treino (segundos)	8,64	2,42	1,46	2,10	2,50	1,84	1,49	1,63	1,75	1,47

Tabela 54 – Redes MLP treinadas a Rprop com uma camada oculta - faltas a 50%

Neurônios na camada oculta	1	2	3	4	5	6	7	8	9	10
Classificações corretas	50,92%	85,48%	89,45%	87,85%	86,43%	90,40%	92,12%	90,75%	91,88%	92,95%
Classificações incorretas	49,08%	14,52%	10,55%	12,15%	13,57%	9,60%	7,88%	9,25%	8,12%	7,05%
Maior erro de teste	75,29%	63,78%	61,59%	61,81%	62,18%	59,65%	58,47%	60,11%	59,05%	58,37%
Maior erro validação	75,49%	64,05%	62,85%	62,14%	63,05%	59,64%	58,49%	60,38%	58,99%	58,44%
Maior erro global	75,46%	63,82%	62,15%	62,34%	62,06%	59,29%	57,94%	59,56%	58,65%	57,86%
Iterações	4889	4257	4667	2819	1682	2624	3511	2029	2160	2861
tempo de treino (segundos)	34,84	32,55	38,12	24,25	16,01	26,42	37,69	19,25	22,08	29,92

Os percentuais de classificação correta da Tabela 54 foram superiores a 85% em quase todas as configurações (exceto a 1). vale destacar que os tempos de treinamento foram muito mais altos comparados à Tabela 53.

Tabela 55 – Redes MLP treinadas a Rprop com uma camada oculta - faltas a 75%

Neurônios na camada oculta	1	2	3	4	5	6	7	8	9	10
Classificações corretas	50,68%	92,18%	93,54%	92,77%	93,95%	92,65%	94,96%	93,84%	94,25%	94,49%
Classificações incorretas	49,32%	7,82%	6,46%	7,23%	6,05%	7,35%	5,04%	6,16%	5,75%	5,51%
Maior erro de teste	72,60%	58,57%	57,71%	57,33%	57,35%	58,05%	56,35%	57,14%	56,38%	57,05%
Maior erro validação	73,26%	60,11%	58,61%	57,94%	57,44%	57,83%	57,86%	57,52%	56,88%	57,17%
Maior erro global	72,63%	59,15%	57,94%	56,98%	56,89%	57,00%	56,37%	57,10%	56,29%	56,22%
Iterações	1204	229	695	967	816	413	611	300	744	438
tempo de treino (segundos)	8,47	1,76	5,78	8,46	7,64	4,11	6,55	2,90	7,60	4,65

A Tabela 55 apresenta resultados similares à Tabela 53, com percentuais de acertos um pouco inferiores. Contudo, os tempos de treinamento não são tão elevados quanto da Tabela 54 que também trata de falhas elétricas.

CONCLUSÃO

Para as configurações adotadas nos experimentos que resultaram nesta tese, a combinação de atributos estatísticos 8 (Curtose, Desvio Padrão e Centroide) mostrou-se ser a de melhor desempenho entre as dez analisadas. Com o propósito de avaliar se esta combinação de atributos era realmente satisfatória, este anexo foi criado. Topologias mais simples, com apenas uma camada oculta, receberam os vetores de entrada formados pela combinação e 8 e tiveram seus desempenhos compilados nas tabelas acima apresentadas. Em quase todos os casos (somente nas redes com apenas um neurônio na camada oculta é que isso não se consolidou) os desempenhos foram satisfatórios. Isso mostra que, de fato a combinação de atributos estatísticos 8 retirada dos nós 2, 4, 8 e 16 da WPT foram um conjunto de dados favorável à tarefa de classificação de padrões para monitoramento de condições do motores de indução.