

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

GABRIELA WINKLER FAVIEIRO

**ANÁLISE E DESENVOLVIMENTO DE ALGORITMO DE
RECONHECIMENTO DE PADRÃO BASEADO EM LÓGICA
PARACONSISTENTE PARA USO EM CONTEXTOS
INCOSISTENTES E SUA APLICAÇÃO NA CLASSIFICAÇÃO
DE MOVIMENTOS DO MEMBRO SUPERIOR ATRAVÉS DE
SINAIS MIOELÉTRICOS**

Porto Alegre

2019

GABRIELA WINKLER FAVIEIRO

**ANÁLISE E DESENVOLVIMENTO DE ALGORITMO DE
RECONHECIMENTO DE PADRÃO BASEADO EM LÓGICA
PARACONSISTENTE PARA USO EM CONTEXTOS
INCOSISTENTES E SUA APLICAÇÃO NA CLASSIFICAÇÃO
DE MOVIMENTOS DO MEMBRO SUPERIOR ATRAVÉS DE
SINAIS MIOELÉTRICOS**

Tese de doutorado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica, da Universidade Federal do Rio Grande do Sul, como parte dos requisitos para a obtenção do título de Doutor em Engenharia Elétrica.

Área de concentração: Controle e Automação - Instrumentação.

ORIENTADOR: Dr. Alexandre Balbinot

Porto Alegre

2019

GABRIELA WINKLER FAVIEIRO

**ANÁLISE E DESENVOLVIMENTO DE ALGORITMO DE
RECONHECIMENTO DE PADRÃO BASEADO EM LÓGICA
PARACONSISTENTE PARA USO EM CONTEXTOS
INCOSISTENTES E SUA APLICAÇÃO NA CLASSIFICAÇÃO
DE MOVIMENTOS DO MEMBRO SUPERIOR ATRAVÉS DE
SINAIS MIOELÉTRICOS**

Esta tese foi julgada adequada para a obtenção do título de Doutor em Engenharia Elétrica e aprovada em sua forma final pelo Orientador e pela Banca Examinadora.

Orientador: _____

Prof. Dr. Alexandre Balbinot, UFRGS

Doutor pela Universidade Federal do Rio Grande do Sul – Porto Alegre, Brasil

Banca Examinadora:

Prof. Dr. Alexandre Rosa Franco, NKI

Doutor pela University of New Mexico – Albuquerque, Estados Unidos da América

Prof. Dr. Altamiro Susin, PPGEE - UFRGS

Doutor pela Institut National Polytechnique de Grenoble – Grenoble, França

Prof. Dr. Fernando Santos Osório, ICMC - USP

Doutor pela Institut National Polytechnique de Grenoble – Grenoble, França

Prof. Dr. Teodiano Freire Bastos Filho, PPGEE - UFES

Doutor pela Universidad Complutense de Madrid – Madrid, Espanha

Prof. Dr. Valner Brusamarello, PPGEE - UFRGS

Doutor pela Universidade Federal de Santa Catarina – Florianópolis, Brasil

Coordenador do PPGEE: _____

Prof. Dr. João Manoel Gomes da Silva Jr.

Porto Alegre, Outubro de 2019.

“O que sabemos é uma gota;
o que ignoramos é um oceano.”

Isaac Newton

a minha filha Helena, por ceder um pouco do seu precioso tempo comigo;
a Minha Família, pelo porto seguro que sempre me proporcionam em épocas de
Tempestade.

AGRADECIMENTOS

O Doutorado é uma jornada solitária, na qual temos que provar para nós mesmos que somos capazes e, ao mesmo tempo, ter humildade suficiente em aceitar a vastidão de conhecimento, admitindo que ainda há muito para se aprender. No entanto, apesar de ser um caminho solitário, muitas pessoas fazem parte dessa travessia, servindo como corrimão, te ajudando a superar obstáculos e até em um abraço ou palavras de carinho que te motivam a continuar buscando teus sonhos.

Com isso gostaria de agradecer a meus queridos e estimados colegas e professores do IEE por todo o auxílio nas tarefas desenvolvidas durante essa empreitada e principalmente pelo apoio tanto técnico quanto psicológico nas horas de necessidade. Em especial à Karina Moura e Vinícius Cene, que sempre proporcionaram discussões saudáveis durante essa trajetória de 4 anos, que considero não somente colegas mas amigos que eu levo desse Doutorado.

A meus pais, meus irmãos e meu marido por todo apoio. Sem vocês eu não teria conseguido.

Ao professor Balbinot por sempre acreditar em meu potencial e pela paciência que teve comigo em meus momentos de dificuldade.

Ao Programa de Pós-Graduação em Engenharia Elétrica/PPGEE, pela oportunidade de realização de trabalhos em minha área de pesquisa e à CAPES pela provisão da bolsa de estudos.

RESUMO

Algoritmos de aprendizado de máquina, em especial os de reconhecimento de padrão, tem evoluído cada vez mais no campo científico. No entanto, a aplicação de tais algoritmos em sinais físicos é sempre um desafio, pois eventos indesejáveis podem ocorrer quando os sinais são adquiridos fora de um ambiente controlado. Dentre diversas aplicações, o reconhecimento de movimentos através de sinais de sEMG é especialmente complicado, pois os mesmos estão sujeitos a diversos tipos de contaminantes que podem degradar o sinal. Essas degradações alteram as características dos sinais mioelétricos, dificultando a capacidade dos algoritmos de reconhecimento de padrão em discriminar classes de movimento. Este trabalho apresenta contribuições no desenvolvimento de um algoritmo de reconhecimento de padrão que leva em consideração o tratamento desses sinais não-ideais. O método *Paraconsistent Random Forest* desenvolvido agrega as vantagens de classificadores híbridos, entre elas, a baixa suscetibilidade a ruído usando uma abordagem de *Random Forest* e a capacidade da Lógica Paraconsistente em lidar com dados inconsistentes em sua estrutura teórica sem invalidar suas conclusões. Diversas bases de dados padrão na área de aprendizagem de máquina foram utilizadas para validar o método, cujas taxas de acertos foram equiparáveis aos métodos tradicionais e se sobressaíram na classificação de bases de dados degradadas. O desempenho do método *Paraconsistent Random Forest* também foi avaliado na área de reconhecimento de movimentos do membro superior através de sinais mioelétricos. Vários procedimentos experimentais foram empregados para analisar a viabilidade e robustez do método quanto a contaminantes típicos da área de eletromiografia de superfície, como artefatos de movimento, ruído térmico e perda do contato eletrodo-pele. O método *Paraconsistent Random Forest* se mostrou promissor para o uso em contextos onde ocorrem degradação dos dados de entrada, apresentando um decréscimo inferior a 20% na predição dos movimentos em comparação aos métodos tradicionais que apresentaram, na mesma situação, decréscimos de até 90%, invalidando o modelo. Todos os experimentos foram validados estatisticamente utilizando conceito de projeto fatorial completo.

Palavras-chave: Sinais Mioelétricos. Lógica Paraconsistente. Reconhecimento de padrão. *Paraconsistent Random Forest*.

ABSTRACT

Machine learning algorithms, in special the pattern recognition ones, have been evolved significantly over the past years in the scientific community. However, the application of such algorithms in real world signals remains a challenge, since undesirable effects can occur when the acquisition is not made in a controlled environment. Among the various applications, the pattern recognition movements using sEMG signals is particularly complicated, since they are susceptible to several types of contaminants that may degrade the signal. These degradations change the myoelectric signal behavior, making it difficult for the pattern recognition algorithm to detect the movement class. In this work, a pattern recognition algorithm was developed in order to mitigate these problems. The method, *Paraconsistent Random Forest*, result of this thesis, adds the advantages of hybrid classifiers, among them, low noise susceptibility due to a Random Forest approach and the Paraconsistent Logic ability to handle contradictory signals in its theoretical structure without invalidating the conclusions. The results showed that the *Paraconsistent Random Forest* was capable to classify several benchmark databases with scores comparable to traditional methods. In this experiment, the Paraconsistent Random Forest theory also extends to the application of upper limb motion recognition through myoelectric signals. Several experimental procedures were executed in order to analyze the feasibility and robustness when compared to other algorithms and applying signal contaminants typical of the area of electromyography, such as movement artifacts, thermal noise and electrode-skin contact loss. The Paraconsistent Random Forest resulted in a promising method for use in contexts where signal degradation was presented, achieving an accuracy decrease lower than 20% when compared to traditional methods that, in the same context, presented an accuracy decrease of up to 90%, invalidating the model. All experiments were statistically validated using the full factorial design concept.

Keywords: Myoelectric Signal. Paraconsistent Logic. Pattern Recognition. Paraconsistent Random Forest.

LISTA DE FIGURAS

Figura 1 - Exemplo de árvore de decisão simbólica “ficar em casa ou ir para o parque”.....	22
Figura 2 - Diagrama de Hasse de quatro vértices.....	29
Figura 3 - (a) Reticulado finito de Hasse e (b) Reticulado finito de Hasse com anotações da Lógica Paraconsistente Anotada.....	33
Figura 4 - Quadrado unitário no plano cartesiano (QUPC).....	35
Figura 5 - L: Reticulado τ em um novo sistema de coordenadas.....	36
Figura 6 - Conversão de QUPC em pontos de L.....	36
Figura 7 - Projeções do Grau de Certeza e Grau de Contradição.....	38
Figura 8 - Projeção do Grau de Certeza Real.....	39
Figura 9 - Gráfico das características entre valores de graus de evidência resultante relacionados aos valores de intervalo de evidência resultante.....	40
Figura 10 - Diagrama de Blocos do sistema de análise paraconsistente proposto.....	61
Figura 11 - Diagrama de blocos representando a RAP para 3 características de entrada.....	61
Figura 12 - Diagrama de blocos da Unidade de Reconhecimento de Padrão (URP) proposta.....	65
Figura 13 - Diagrama de blocos do uso de várias URPs quando se utiliza mais de uma característica extraída do sinal.....	66
Figura 14 - Descrição dos experimentos realizados na avaliação dos métodos.....	79
Figura 15 - Posicionamento dos eletrodos na configuração bipolar.....	82
Figura 16 - Diagrama de blocos de aquisição de dados.....	83
Figura 17 - Diagrama de blocos do processo de pré-processamento de sinais de sEMG.....	84
Figura 18 - Diagrama de blocos das duas categorias de avaliação utilizadas.....	93
Figura 19 - (a) Curva de potência para o projeto de experimentos, (b) histograma dos resíduos e análise dos resíduos em relação ao (c) ordem de observação e (d) valor ajustado.....	98
Figura 20 - Gráfico de efeitos da taxa de classificação média para movimentos, conjunto de características e base de dados.....	100
Figura 21 - Gráfico de comparação das taxas de acerto médias obtidas para cada base de dados e para cada combinação de características do sinal.....	100
Figura 22 - Comparativo de taxas de acerto entres as bases de dados, discriminadas por movimento, utilizando as características fMED e RMS do sinal como entrada... ..	102
Figura 23 - Taxa de acerto média considerando os diferentes métodos e bases de dados.....	105
Figura 24 - Comparação da taxa de classificação média para as bases de dados selecionadas e o método PRF proposto.....	108
Figura 25 - Gráfico da taxa de classificação em comparação aos métodos PRF, LDA, SVM e KNN com adição de valores inválidos, sendo as bases de dados: CNC, WIN, GLA, IRS, DBC e PBC.....	112
Figura 26 - Gráfico da taxa de classificação em comparação aos métodos PRF, LDA, SVM e KNN com adição de ruído, sendo as bases de dados: CNC, WIN, GLA, IRS, DBC e PBC.....	116
Figura 27 - Taxa de acerto média na análise de base de dados de sinal de sEMG íntegro....	122
Figura 28 - Gráfico de interação entre médias da taxa de acerto para os métodos e bases de dados considerando os sinais de sEMG íntegros.....	123
Figura 29 - Gráfico da interação entre os fatores métodos e bases de dados.....	126
Figura 30 - Gráfico da interação entre os fatores métodos e tipo de contaminante.....	127
Figura 31 - Gráfico da interação entre os fatores métodos e canais com contaminação.....	127
Figura 32 - Gráfico da taxa de acerto média considerando métodos e configurações de contaminação dos canais de sEMG para o contaminante artefato de movimento.....	131

Figura 33 - Gráfico da % de decréscimo na taxa de acerto considerando métodos e combinações de degradação dos sinais para o contaminante ruído térmico.	134
Figura 34 - Gráfico da % de decréscimo na taxa de acerto considerando métodos e configurações de canais contaminados por perda de contato eletrodo-pele.	138
Figura 35 - Curva de Potência Estatística (a), Histograma dos resíduos (b) e análise dos resíduos pela ordem de observação (c) e pelo valor ajustado (d), para Projeto de experimentos referente ao Experimento sEMG 1, considerando os métodos de classificação... ..	157
Figura 36 - Curva de Potência Estatística (a), Histograma dos resíduos (b) e análise dos resíduos pela ordem de observação (c) e pelo valor ajustado (d), para Projeto de experimentos referente ao Experimento BDP, considerando dados com valores ausentes.....	158
Figura 37 - Curva de Potência Estatística (a), Histograma dos resíduos (b) e análise dos resíduos pela ordem de observação (c) e pelo valor ajustado (d), para Projeto de experimentos referente ao Experimento BDP, considerando dados com inserção de ruído.	159
Figura 38 - (a) Curva de potência estatística para projeto de experimentos de sinais de sEMG íntegros, (b) histograma dos resíduos e análise dos resíduos e pela (c) ordem de observação e (d) pelo valor ajustado.	160
Figura 39 - (a) Curva de potência estatística, (b) histograma dos resíduos e análise dos resíduos pela (c) ordem de observação (d) e pelo valor ajustado, para Projeto de experimentos referente aos sinais de sEMG com contaminantes.....	161
Figura 40 - (a) Curva de potência estatística, (b) histograma dos resíduos e análise dos resíduos pela (c) ordem de observação (d) e pelo valor ajustado, para Projeto de experimentos referente a base de dados de sinais de sEMG contaminado com artefatos de movimento.	162
Figura 41 - (a) Curva de potência estatística, (b) histograma dos resíduos e análise dos resíduos pela (c) ordem de observação (d) e pelo valor ajustado, para Projeto de experimentos referente a base de dados de sinais de sEMG contaminado com ruído térmico. ...	163
Figura 42 - (a) Curva de potência estatística, (b) histograma dos resíduos e análise dos resíduos pela (c) ordem de observação (d) e pelo valor ajustado, para Projeto de experimentos referente a base de dados de sinais de sEMG contaminado com perda de contato eletrodo-pele.....	164

LISTA DE TABELAS

Tabela 1 - Descrição do Algoritmo A-1 do Nó de Análise Paraconsistente.....	41
Tabela 2 - Descrição dos passos do algoritmo de ParaExtr _{ctr} (DA SILVA FILHO, 2009).	67
Tabela 3 - Descrição do Algoritmo A-2, <i>Paraconsistent Random Forest</i>	70
Tabela 4 - Descrição do Algoritmo A-3 - Árvore de Decisão Paraconsistente.	75
Tabela 5 - Descrição do Algoritmo A-4, inferência dos valores para a Árvore de Decisão Paraconsistente.	76
Tabela 6 - Descrição do Algoritmo A-5, ParaExtr _{ctr} (DA SILVA FILHO, 2009).	77
Tabela 7 - Tabela com resumo das bases de dados Ninapro.	80
Tabela 8 - Quadro-resumo das características clínicas dos voluntários com amputação.	81
Tabela 9 - Descrição dos movimentos.	82
Tabela 10 - Resumo das bases de sinais de sEMG utilizadas.	85
Tabela 11 - Características extraídas do sinal e suas definições matemáticas.	86
Tabela 12 - Características extraídas do sinal e suas definições matemáticas.	88
Tabela 13 - Modelos matemáticos dos contaminantes de sinal de sEMG propostos por (DE MOURA; BALBINOT, 2018).	89
Tabela 14 - Descrição dos canais contaminados neste estudo.	91
Tabela 15 - Características das bases de dados.	92
Tabela 16 - Parâmetros dos métodos de classificação utilizados para fins de comparação. ...	95
Tabela 17 - Análise da variância (experimento de 3 fatores).	99
Tabela 18 - Análise das médias das características utilizando método Tukey (95% de confiabilidade).	101
Tabela 19 - Análise da variância considerando como fatores métodos e bases de dados.	103
Tabela 20 - Análise das médias dos métodos utilizando o método de Tukey (95% de confiabilidade).	104
Tabela 21 - Análise do tempo de processamento de uma amostra para diferentes topologias da RAP.	106
Tabela 22 - Análise da variância considerando como fatores métodos, bases de dados e a% de dados ausentes.	108
Tabela 23 - Decréscimo na taxa de acerto média na presença de dados ausentes.	109
Tabela 24 - Comparação do decréscimo na taxa de acerto média na presença de dados ausentes com os métodos PRF, SVM, KNN e LDA.	110
Tabela 25 - Análise das médias dos métodos utilizando método Tukey (95% de confiabilidade) para o experimento BDP com dados ausentes.	111
Tabela 26 - Análise da variância considerando como fatores métodos, bases de dados e a% de ruído adicionado.	113
Tabela 27 - Decréscimo na taxa de acerto média na presença de dados com ruído.	114
Tabela 28 - Comparação do decréscimo na taxa de acerto média ao se adicionar $\alpha\%$ dados com ruído para os métodos PRF, SVM, KNN e LDA.	115
Tabela 29 - Análise das médias dos métodos utilizando método Tukey (95% de confiabilidade).	117
Tabela 30 - Comparativo do método PRF com o método FRF.	119
Tabela 31 - Projeto de experimentos referente à análise dos classificadores utilizando sinal de sEMG original.	120
Tabela 32 - Análise das médias dos métodos utilizando o método de Tukey (95% de confiabilidade).	121
Tabela 33 - Análise das médias das bases de dados utilizando o método de Tukey (95% de confiabilidade).	121

Tabela 34 - Projeto de experimentos referente à análise dos classificadores utilizando sinal de sEMG com contaminantes.	124
Tabela 35 - Análise das médias dos métodos utilizando o método de Tukey (95% de confiabilidade).	125
Tabela 36 - Análise das médias das bases de dados utilizando o método de Tukey (95% de confiabilidade).	126
Tabela 37 - Projeto de experimentos referente a análise dos métodos de classificação utilizando sinal de sEMG com contaminante artefato de movimento.	128
Tabela 38 - Análise das médias dos métodos utilizando o método de Tukey (95% de confiabilidade).	129
Tabela 39 - Decréscimo na taxa de classificador de acordo com o método e configuração de canais contaminados por artefatos de movimento.	130
Tabela 40 - Projeto de experimentos referente a análise dos métodos de classificação utilizando sinal de sEMG com contaminante ruído térmico.	132
Tabela 41 - Análise das médias dos métodos utilizando o método de Tukey (95% de confiabilidade).	133
Tabela 42 - Decréscimo na taxa de classificador de acordo com o método e canais contaminados por ruído térmico.	133
Tabela 43 - Projeto de experimentos referente à análise dos métodos de classificação utilizando sinal de sEMG com contaminante perda de contato.	135
Tabela 44 - Análise das médias dos métodos utilizando método Tukey (95% de confiabilidade).	136
Tabela 45 - Decréscimo na taxa de classificador de acordo com o método e canais contaminados por perda de contato eletrodo-pele.	136
Tabela 46 - Comparação entre resultados do sistema de Sensor Virtual e o método PRF para classificação de movimento de voluntários com membro superior intacto.	144

LISTA DE ABREVIATURAS

AF	Artefato de Movimento
ANFIS	Adaptive Neuro-Fuzzy Inference System
BDP	Base de Dados Padrão
CART	<i>Classification and Regression Trees</i>
DTA	Decréscimo na Taxa de Acerto Média
fMED	Frequência Mediana
ID3	<i>Induction Decision Tree</i>
IEE	base de dados Laboratório IEE/UFRGS
KNN	<i>k-Nearest Neighbors</i>
LDA	<i>Linear Discriminant Analysis</i>
LP	Lógica Paraconsistente
LPA2v	Lógica Paraconsistente Anotada de 2 valores
MAV	<i>Mean absolute value</i>
MFL	<i>Maximum Fractal Length</i>
MP	Movimento Padrão
NAP	Nó de Análise Paraconsistente
NPA	base de dados Ninapro Amputados
NPI	base de dados Ninapro Intactos
ParaExtr _{ctr}	Algoritmo Paraconsistente Extrator de efeitos da Contradição
PC	Perda de Contato eletrodo-pele
PRF	<i>Paraconsistent Random Forest</i>
RAP	Rede de Análise Paraconsistente
RF	<i>Random Forest</i>
RMS	<i>Root Mean Square</i>
RNA	Rede Neural Artificial
RT	Ruído Térmico
sEMG	Eletromiografia de Superfície
SKW	obliquidade
SVM	<i>Support Vector Machine</i>
TAM	Taxa de Acerto Média
TVARMA	<i>Time-Varying Autoregressive Moving Average</i>
TVK	<i>Time-Varying Kalman filter</i>
URP	Unidade de Reconhecimento de Padrão
VAR	Variância
WL	Comprimento de Onda

SUMÁRIO

1	INTRODUÇÃO	13
2	FUNDAMENTOS TEÓRICOS	21
2.1	PRINCÍPIOS SOBRE ÁRVORES DE DECISÃO	21
2.2	RANDOM FOREST	24
2.3	LÓGICA PARACONSISTENTE	27
2.3.1	Fundamentos da Lógica Paraconsistente	28
2.3.2	Lógica Paraconsistente Anotada Evidencial	32
2.3.3	Nó de Análise Paraconsistente	40
2.4	CONCEITOS BÁSICOS DE SINAIS MIOELÉTRICOS	41
3	ESTADO DA ARTE	44
3.1	APLICAÇÕES DE RECONHECIMENTO DE PADRÃO EMPREGANDO LÓGICA PARACONSISTENTE	44
3.2	TÉCNICAS DE RECONHECIMENTO DE MOVIMENTOS BASEADAS EM SINAIS MIOELÉTRICOS	46
3.3	SISTEMAS DE CLASSIFICAÇÃO ROBUSTOS A CONTAMINANTES NO SINAL DE sEMG	52
3.4	CARACTERÍSTICAS DE UM SISTEMA DE CONTROLE MIOELÉTRICO IDEAL	58
4	METODOLOGIA	60
4.1	ANÁLISE PRELIMINAR DA LÓGICA PARACONSISTENTE NA CLASSIFICAÇÃO DE SINAIS MIOELÉTRICOS	60
4.1.1	Transformação das Características em Graus de Evidência	63
4.1.2	Procedimento de Inicialização da RAP	63
4.1.3	Reconhecimento de Padrões	64
4.1.4	Grau de Evidência Resultante por Classe de Movimento Padrão	66
4.1.5	Limitações do Método	67
4.2	DESENVOLVIMENTO DO MÉTODO PARACONSISTENT RANDOM FOREST	68
4.2.1	Concepção da <i>Paraconsistent Random Forest</i>	69
4.2.2	Partição do domínio com conjuntos <i>fuzzy</i>	70
4.2.3	Concepção da Árvore de Decisão Paraconsistente	71
4.2.4	Inferência dos Valores	75
4.2.5	Estratégias para Tomada de Decisão	76
4.3	METODOLOGIA DE TESTES	78
4.3.1	Descrição das Bases de Sinais de sEMG	80
4.3.2	Experimento sEMG 1	85
4.3.3	Experimento sEMG 2	86
4.3.4	Experimento BDP	91
4.3.5	Avaliação dos métodos	93
4.4	PROJETO DE EXPERIMENTOS PROPOSTO	95
5	RESULTADOS E DISCUSSÕES	97
5.1	RESULTADOS E DISCUSSÕES DO EXPERIMENTO sEMG 1	97
5.2	RESULTADOS E DISCUSSÕES DO EXPERIMENTO BDP	107
5.3	RESULTADOS E DISCUSSÕES DO EXPERIMENTO sEMG 2	119
5.4	DISCUSSÕES ENTRE OS EXPERIMENTOS E ESTUDOS SIMILARES	140
6	CONCLUSÕES E CONSIDERAÇÕES FINAIS	145
	REFERÊNCIAS	148
	APÊNDICE A – PROJETOS DE EXPERIMENTOS	157

1 INTRODUÇÃO

Os processos de tomada de decisão estão cada vez mais dependentes de sistemas inteligentes, visto que o seu uso está altamente disseminado em comunidades científicas. O estágio atual das pesquisas ainda aponta para questões em aberto, merecendo estudos adicionais que aprimorem os métodos de reconhecimento de padrão existentes, focando na adaptabilidade do sistema aos sinais de entrada, bem como, na sua susceptibilidade a contaminantes e contradições¹ intrínsecas do sinal.

Em um sistema de tomada de decisão inteligente muitas vezes são utilizados algoritmos baseados em aprendizado supervisionado. Esse tem por objetivo criar um conjunto de regras com o intuito de determinar a classe de qualquer objeto a partir de um conjunto de atributos inicial (BONISSONE *et al.*, 2010; BREIMAN, 2001; QUINLAN, 1986; WANG *et al.*, 2015). No entanto, dependendo do tipo de aplicação, nem sempre os atributos disponíveis são capazes de fornecer informações suficientes para diferenciar, com absoluta certeza, a classe à qual o objeto pertence, demonstrando que existe uma contradição na base de dados adquirida. Por exemplo, se o conjunto de treinamento, cujas classes são conhecidas antecipadamente, contiver dois objetos com valores idênticos e mesmo assim esses objetos pertencerem a classes diferentes, será impossível distingui-los. Ou seja, apesar das soluções sofisticadas introduzidas por esses algoritmos, em aplicações do dia a dia muitas vezes os sinais de entrada não são ideais², incluindo dados inconsistentes, ou até mesmo a falta de dados, que acabam por interferir no processo de classificação.

Portanto, um dos desafios no desenvolvimento de algoritmos de aprendizagem de máquina não é somente elaborar um classificador visando elevar a predição do modelo utilizando bases de dados bem comportadas, mas essencialmente desenvolver um sistema de

¹ Contradição (ou Inconsistência) na Lógica Paraconsistente é definida de forma que para uma mesma proposição p existem evidências contrárias entre si (p e $\neg p$ são verdadeiros para uma ou mais evidências) (BLAIR; SUBRAHMANIAN, 1989; DA SILVA FILHO *et al.*, 2016; MIDDELBURG, 2011). Por exemplo, na caracterização de movimentos, para uma classe-alvo podem existir um ou mais atributos do sinal que se contradizem em relação ao modelo gerado *a priori* (baseado no conjunto de dados de treinamento). Essa contradição pode ser ocasionada por um modelo inadequado dos dados de entrada, por degradações do sinal em virtude da presença de contaminantes ou outros fatores externos à aquisição de dados (DE MOURA; BALBINOT, 2018; SAMUEL, 2019).

² Sinais não ideais são definidos, neste trabalho, como sinais nos quais podem ocorrer variações significativas nas características do mesmo para eventos semelhantes. Sinais não ideais usualmente são afetados por fatores externos a sua aquisição, como, por exemplo, degradações do sinal decorrentes de contaminações.

classificação que permita o uso de sinais não ideais, buscando ampliar a sua aplicabilidade, visto que sistemas que utilizam sinais físicos frequentemente são afetados por fatores externos à aquisição dos mesmos, como é o caso de sinais biológicos (ADEWUYI *et al.*, 2016; AHMADIZADEH *et al.*, 2017; CENE; FAVIEIRO; BALBINOT, 2015; DE MOURA; BALBINOT, 2018; GU *et al.*, 2018; MCCOOL *et al.*, 2014; MUCELI; JIANG; FARINA, 2014; PHINYOMARK *et al.*, 2013; SAMUEL, 2019; STANGO; NEGRO; FARINA, 2015; VIDOVIC *et al.*, 2016; XIAORONG ZHANG; HE HUANG; QING YANG, 2013; ZHANG; HUANG, 2015). Dessa forma, considerando que podem existir degradações nos dados fornecidos ao algoritmo de reconhecimento de padrão, o uso de classificadores com tomada de decisão binária pode não ser a melhor opção. Sendo assim, o uso de uma lógica não clássica pode trazer como vantagem um acréscimo de robustez³ ao algoritmo de decisão ao se considerar que podem ocorrer contradições no sistema.

Dentre diversas lógicas não clássicas existentes, a Lógica Paraconsistente (LP) é assim considerada pelo fato de admitir o tratamento de sinais contraditórios em sua estrutura teórica sem invalidar as conclusões (DA SILVA FILHO; ABE; TORRES, 2008). O seu uso tem ganhado visibilidade nos últimos anos ao ser aplicada com sucesso em diversos campos como análise de imagens, diagnósticos médicos e reconhecimento de voz e caracteres (ABE, 2010; ABE; LOPES; NAKAMATSU, 2014; DA SILVA FILHO, 2012; DUNIN-KĘPLICZ *et al.*, 2015; SOUZA; ABE; NAKAMATSU, 2013). Esses estudos relataram que a LP é apropriada na solução de problemas de reconhecimento de padrão em função de sua capacidade em lidar com dados inconsistentes. Abordagens similares com uso da LP, cada uma com diferentes topologias (BARBON JÚNIOR; GUIDO; VIEIRA, 2009; DA SILVA FILHO *et al.*, 2016; GUIDO *et al.*, 2013), também foram utilizadas no tratamento de situações não determinísticas. Além disso, de acordo com (DA SILVA FILHO; ABE; TORRES, 2008), a Lógica Paraconsistente apresenta vantagens pela sua simplicidade computacional que revela potencial para o uso em sistemas embarcados. Portanto, é possível afirmar que a utilização dos conceitos da Lógica Paraconsistente aliada a sistemas inteligentes de reconhecimento de padrão pode proporcionar uma solução interessante para áreas que envolvam sinais físicos sujeitos a contaminações, de forma a tentar mitigar o impacto dessas degradações do sinal no modelo de classificação.

³ Um algoritmo é considerado robusto, neste trabalho, quando o mesmo tem baixa susceptibilidade a variações nos atributos/características de entrada (dados inconsistentes), de forma a minimizar o erro adicionado por essas variações na predição das classes (FANG; LI, 2017; GENG *et al.*, 2018; PAN *et al.*, 2015).

Adicionalmente, uma das possíveis áreas que pode se beneficiar com algoritmos baseados em lógicas não clássicas são aplicações que envolvem o uso de dados biológicos, como, por exemplo, a área relacionada a processamento de sinais de eletromiografia de superfície (sEMG). Sistemas gerenciados por sinais de sEMG contêm a informação neural das tarefas motoras executadas. De fato, o sinal de sEMG contém informação similar aos dados adquiridos diretamente das terminações nervosas, uma vez que existe uma associação de um para um entre os potenciais de ação trafegando entre os axônios de cada neurônio motor e o potencial de ação composto das fibras musculares (potencial de ação da unidade motora) (FARINA et al., 2014). O uso de eletromiografia de superfície é muito extenso na área de controle inteligente de próteses, devido a possibilidade de aquisição do sinal de sEMG de forma não invasiva (ATZORI et al., 2015; BALBINOT; FAVIEIRO, 2013; CENE; FAVIEIRO; BALBINOT, 2015; FARRELL, 2011; FARRELL; WEIR, 2007; FAVIEIRO, 2012; GENG et al., 2018; HUDGINS; PARKER; SCOTT, 1993; KWON; KIM, 2011; LORRAIN; JIANG; FARINA, 2010; MONSIFROT et al., 2013; SAMUEL, 2019; ZHANG; HUANG, 2015). No entanto, realizar um controle natural e intuitivo utilizando o sinal de sEMG não é uma tarefa trivial, considerando a natureza estocástica do sinal e fatores adicionais como contaminantes e dificuldades experimentais, como exemplo, eletrodos mal posicionados ou desconectados (BETTHAUSER et al., 2018; CENE; FAVIEIRO; BALBINOT, 2015; CENE; BALBINOT, 2018; DE MOURA; BALBINOT, 2018; GENG et al., 2018; GU et al., 2018; SAMUEL, 2019; VIDOVIC et al., 2016; ZHANG; HUANG, 2015). Por esses motivos, as pesquisas que envolvem sistemas inteligentes de controle de próteses de membro superior têm progredido e apresentando crescente interesse da comunidade científica. Atualmente, existem muitos algoritmos de reconhecimento de padrão e aprendizagem de máquina que são empregados para o controle mioelétrico como *Linear Discriminant Analysis* (LDA) (GENG; ZHOU; LI, 2012; OSKOEI; HU, 2006; RIILLO et al., 2014; ZHANG; HUANG, 2015), Redes Neurais Artificiais (RNA) (BALBINOT; JÚNIOR; FAVIEIRO, 2013; OSKOEI; HU, 2008; RIILLO et al., 2014), Lógica Fuzzy e Neuro-Fuzzy (BALBINOT; FAVIEIRO, 2013; FAVIEIRO; BALBINOT, 2011; KHEZRI; JAHED, 2011) e *Support Vector Machine* (SVM) (ALKAN; GÜNAY, 2012; QIAN et al., 2010; RIILLO et al., 2014; ROSSI et al., 2015; WANG et al., 2009), entre diversas outras técnicas.

Em consequência, inúmeros trabalhos evidenciaram que diversos classificadores podem atingir taxas de acerto superiores a 80% para mais de dez movimentos distintos, embora muitas dessas pesquisas não tenham avaliado os efeitos em pessoas com amputação (FAVIEIRO; BALBINOT, 2011; GUO et al., 2015; KHUSHABA; AL-ANI; AL-JUMAILY,

2010; LI; SCHULTZ; KUIKEN, 2010; MOMEN; KRISHNAN; CHAU, 2007). No entanto, de acordo com (FARINA *et al.*, 2014; SAMUEL, 2019), essas conquistas acadêmicas ainda não são utilizadas em larga escala em aplicações comerciais, porque os sistemas propostos na literatura atual não abrangem todos os critérios necessários para serem aceitos pelos pacientes, tais como: controle natural e intuitivo, sistemas adaptativos e de tempo-real, pouco ou nenhum treinamento, número mínimo de eletrodos, robusto a degradações do sinal causadas por contaminantes, entre outros fatores. Por esses motivos, existem muitos estudos concentrados na usabilidade do sistema como um todo, objetivando viabilizar próteses com controle inteligente e intuitivo (CENE; FAVIEIRO; BALBINOT, 2015; DE MOURA; BALBINOT, 2018; FANG; LI, 2017; GENG *et al.*, 2018; GU *et al.*, 2018; LIU *et al.*, 2015; SAMUEL, 2019; XIAORONG ZHANG; HE HUANG; QING YANG, 2013; ZHANG; HUANG, 2015).

Um dos problemas relatados no uso de algoritmos de aprendizagem de máquina na classificação de movimentos, como por exemplo SVM, Redes Neurais e Árvores de Decisão, consiste em que muitas vezes os dados utilizados no treinamento supervisionado do método não representam de forma fidedigna as classes de movimentos por causa de limitação das informações fornecidas e degradações no sinal causadas por contaminantes, que podem interferir no reconhecimento de movimentos (ADEWUYI *et al.*, 2016; AHMADIZADEH *et al.*, 2017; CENE; FAVIEIRO; BALBINOT, 2015; DE MOURA; BALBINOT, 2018; GU *et al.*, 2018; MCCOOL *et al.*, 2014; MUCELI; JIANG; FARINA, 2014; PHINYOMARK *et al.*, 2013; SAMUEL, 2019; STANGO; NEGRO; FARINA, 2015; VIDOVIC *et al.*, 2016; XIAORONG ZHANG; HE HUANG; QING YANG, 2013; ZHANG; HUANG, 2015). Além disso, em aplicações com uso de sinais biológicos, como a classificação de movimentos, não é possível a escolha de um único sinal na distinção entre classes, o que torna ainda mais difícil a sua correta classificação.

Logo, considerada a natureza estocástica dos sinais de sEMG, a limitação de informação na aquisição de sinais de sEMG, e os fatores adicionais que podem contaminar o sinal, como ruído térmico, perda do contato eletrodo-pele e artefatos de movimento, é possível admitir que podem existir contradições nos sinais em relação ao modelo de classificação gerado *a priori*. Dessa forma, é importante o emprego de um algoritmo que possa levar em consideração tais elementos em sua estrutura teórica com o intuito de minimizar o prejuízo causado por esses fatores na taxa de acerto do sistema, sem necessidade de contínuas intervenções de ajustes, calibrações e contínuos treinamentos do modelo gerado. Um sistema de caracterização de movimentos que seja capaz de manter uma taxa de classificação estável pode aumentar o impacto clínico e comercial das próteses mioelétricas quando o mesmo

estiver exposto a condições não ideais, como por exemplo, na aquisição de sinais degradados por contaminantes.

Deste modo, em função das colocações anteriores, o presente trabalho tem como principal objetivo o estudo e o desenvolvimento de um algoritmo de reconhecimento de padrão baseado em Lógica Paraconsistente para o uso em contextos que envolvam dados não ideais, seja pela falta de informação ou pela degradação dos mesmos, causadas por fatores externos, com o intuito de mitigar os efeitos desses contaminantes na taxa de classificação correta dos padrões. Adicionalmente, para demonstrar a aplicação do algoritmo proposto, é realizado um estudo na área de reconhecimento de movimentos do membro superior baseada em sinais mioelétricos.

A primeira contribuição deste trabalho é a avaliação e o desenvolvimento de uma Rede de Análise Paraconsistente que foi explicitamente modelada para manipular sinais de sEMG, de uma determinada base, para classificação de 17 movimentos distintos de membro superior utilizando 12 canais de sEMG. A Rede de Análise Paraconsistente (RAP) é formada por uma rede de nós interconectados, sendo que sua arquitetura depende diretamente da aplicação correspondente. Essa rede foi proposta com o intuito de realizar uma avaliação preliminar do uso da Lógica Paraconsistente na classificação de movimentos do membro superior através de sinais mioelétricos. Diante dessa premissa, os maiores desafios e dificuldades foram em determinar uma topologia que melhor representasse o modelo de movimentos e que pudesse se beneficiar da análise paraconsistente na classificação de movimentos através de sinais mioelétricos. Após as análises da RAP, foi concluído que a Lógica Paraconsistente, no tratamento de sinais mioelétricos, obteve taxas de acerto similares a outros algoritmos largamente utilizados na área, sendo considerada adequada para sua utilização na classificação de movimentos.

Em função desta análise preliminar da LP, foi proposto um algoritmo mais versátil, que pudesse ser adaptável a diferentes configurações e aplicações. Com isso, foi estudada e desenvolvida uma técnica híbrida que agregasse as vantagens da Lógica Paraconsistente com aprendizado de máquina baseado em Árvores de Decisão e *Random Forest*.

As Árvores de Decisão são bem conhecidas na comunidade científica, especialmente pela sua compreensibilidade. Esta característica atrai muitos usuários interessados em entender o domínio, as regras simbólicas que podem ser extraídas da Árvore e até seu uso em sistemas de decisão baseados em regras. Além disso, sua taxa de acerto, especialmente quando usada com um *ensemble* de Árvores de Decisão, é comparável ou superior a outros algoritmos bem conhecidos na área de reconhecimento de padrão. Outra consideração importante é que

as Árvores de Decisão têm poucos parâmetros que precisam de ajuste, o que simplifica a definição da estrutura do modelo (JANIKOW, 1998; QUINLAN, 1986). Por essas razões, ainda há existem pesquisas na área (BONISSONE *et al.*, 2010; BREIMAN, 2001; WANG *et al.*, 2015). Outro algoritmo bem conhecido é a *Random Forest* (RF). A RF pode ser considerada uma extensão das Árvores de Decisão porque esta metodologia leva em consideração o uso de várias Árvores de Decisão não correlacionadas para a tomada de decisão, aumentando a taxa de acerto com baixa suscetibilidade ao ruído, especialmente quando se trata de sinais estocásticos (BREIMAN, 2001).

Portanto, a segunda contribuição desta tese se encontra no desenvolvimento de um algoritmo de reconhecimento de padrão no qual se utilizam conceitos de *Random Forest* e Lógica Paraconsistente, que podem ser utilizados em diversas aplicações. O propósito deste algoritmo foi combinar a Lógica Paraconsistente com as Árvores de Decisão de forma a obter-se o melhor da capacidade da LP para lidar com dados inconsistentes, a compreensibilidade gerada pela Árvore de Decisão e a baixa suscetibilidade a ruído do sinal de entrada usando uma abordagem de *Random Forest*. Na realização desta tese, o algoritmo aqui proposto recebeu a denominação de *Paraconsistent Random Forest* (PRF). O PRF é um algoritmo inteligente baseado em Lógica Paraconsistente que consegue se adaptar aos mais diversos tipos de dados e aplicações utilizando aprendizagem supervisionada. Esse algoritmo também apresenta como vantagens sua facilidade de utilização, pelo fato de necessitar poucos parâmetros, e por sua baixa susceptibilidade a sinais com contaminantes. O algoritmo PRF foi comparado a outros algoritmos de aprendizagem supervisionada altamente conceituados na área de aprendizagem de máquina se sobressaindo de maneira significativa na utilização de bases de dados padrão com redução artificial de dados de entrada e na presença de ruído. Essa análise demonstrou que o algoritmo está apto a ser utilizado em contextos no qual podem ocorrer perda de informação que, por sua vez, resultam em inconsistências com o modelo do classificador gerado *a priori*.

A contribuição final deste trabalho é o uso da *Paraconsistent Random Forest* aplicada no reconhecimento de 17 movimentos do braço através de sinais mioelétricos. O método *Paraconsistent Random Forest* é comparado aos seguintes métodos tradicionais na área de reconhecimento de padrão: *Support Vector Machines* (SVM), *Linear Discriminant Analysis* (LDA), *k-Nearest Neighbors* (KNN) e *Random Forest* (RF), em termos de taxa de acerto e susceptibilidade a contaminantes do sinal, que podem ocorrer de forma recorrente na aquisição desse tipo de sinal, como apresentam os seguintes trabalhos (CENE; FAVIEIRO; BALBINOT, 2015; DE MOURA; BALBINOT, 2018; FANG; LI, 2017; FRASER *et al.*,

2014; GENG *et al.*, 2018; MCCOOL *et al.*, 2014; SAMUEL, 2019; XIAORONG ZHANG; HE HUANG; QING YANG, 2013; ZHANG; HUANG, 2015). A aplicação da *Paraconsistent Random Forest* na área de reconhecimento de movimentos visou analisar possíveis benefícios que a Lógica Paraconsistente pudesse agregar no processamento de dados com degradações recorrentes causadas por contaminantes comumente relatados na área. Os contaminantes considerados foram os artefatos de movimento, ruído térmico e perda de contato eletrodo-pele de acordo com seu uso em outros estudos (DE MOURA; BALBINOT, 2018; FRASER *et al.*, 2014; GU *et al.*, 2018; HUANG *et al.*, 2010; MCCOOL *et al.*, 2014; PAN *et al.*, 2015; XIAORONG ZHANG; HE HUANG; QING YANG, 2013; ZHANG; HUANG, 2015).

Portanto, após essa breve introdução destacando o objetivo principal deste trabalho, assim como, as contribuições na respectiva área, este trabalho é organizado da seguinte forma: no Capítulo 2 é apresentada uma revisão de Fundamentos Teóricos que serviram como base para o desenvolvimento dos algoritmos de reconhecimento de padrão propostos nesta tese envolvendo Lógica Paraconsistente, Árvores de Decisão e *Random Forest*. O Estado da Arte, descrito no Capítulo 3, abrange métodos de reconhecimento de padrão na área de classificação de movimentos através de sEMG e também aplicações de Lógica Paraconsistente em sistemas de classificação de dados.

Em sequência, no Capítulo 4 são descritas as metodologias de desenvolvimento de dois algoritmos de reconhecimento de padrão baseados em Lógica Paraconsistente propostos nesta tese: a Rede de Análise Paraconsistente e a *Paraconsistent Random Forest*. A Rede de Análise Paraconsistente representa uma proposta de topologia preliminar utilizando unicamente a Lógica Paraconsistente para classificação de movimentos do membro superior através de sinais de sEMG. Por sua vez, o algoritmo *Paraconsistent Random Forest* proposto utiliza conceitos da Lógica Paraconsistente agregados a Árvores de Decisão e *Random Forest* para o desenvolvimento de um algoritmo híbrido de reconhecimento de padrão.

No Capítulo 4 também são apresentadas as bases de dados e a metodologia de testes que inclui o uso de bases de dados padrão utilizadas mundialmente na área de *Machine Learning*, além da sua aplicação na área de reconhecimento de movimentos do membro superior através de sinais mioelétricos, empregando sinais de diversos sujeitos, com e sem amputação transradial do antebraço, para a caracterização de 17 movimentos do membro superior, através de base de dados de livre acesso Ninapro (ATZORI *et al.*, 2012) e base de dados do laboratório IEE/UFRGS. Os procedimentos experimentais foram empregados com o objetivo de avaliar a viabilidade e robustez do método em relação a contaminações, comparando com outros algoritmos e aplicando os conceitos de sinais degradados por

contaminantes típicos da área de eletromiografia, mais especificamente artefatos de movimento, ruído térmico e perda do contato eletrodo-pele.

Os resultados e discussões dos diversos testes realizados são apresentados no Capítulo 5. Através das discussões, é apresentada a eficiência do algoritmo em lidar com contaminações nos dados de entrada, proporcionando uma redução dos erros de classificação decorridos da contaminação do sinal em comparação com outros algoritmos. Além disso, é realizada a comparação dos resultados gerados neste trabalho com os resultados obtidos por outros pesquisadores na área e de trabalhos anteriores do grupo de pesquisa do Laboratório IEE/UFRGS. Por fim, as conclusões e considerações finais, assim como, algumas propostas para trabalhos futuros, são apresentadas no Capítulo 6.

2 FUNDAMENTOS TEÓRICOS

Neste capítulo serão abordados brevemente conceitos de Árvores de Decisão, *Random Forest* e Lógica Paraconsistente, que serviram de base para os algoritmos de reconhecimento de padrão desenvolvidos nesta tese. Adicionalmente, também serão abordados, de maneira sucinta, as principais etapas do processo de contração muscular, características de eletromiografia de superfície. É importante ressaltar que muitos termos foram mantidos em inglês (no seu original) devido ao melhor entendimento do mesmo do ponto de vista técnico e científico.

2.1 PRINCÍPIOS SOBRE ÁRVORES DE DECISÃO

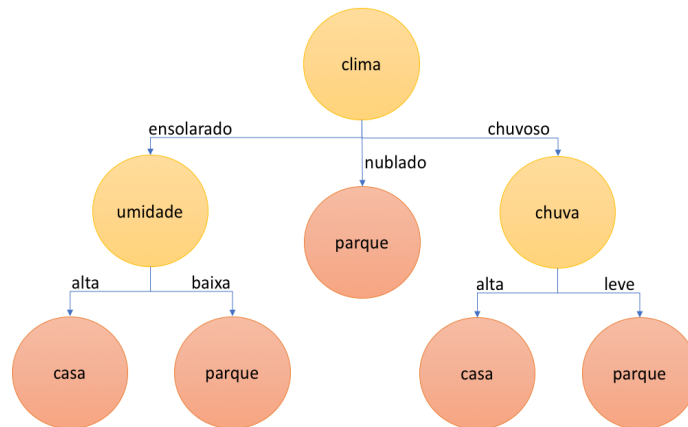
A Árvore de Decisão é um algoritmo amplamente utilizado na literatura por diversas razões. Uma das principais é a compreensibilidade do modelo gerado (JANIKOW, 1998). Outro motivo é sua capacidade de classificação correta dos dados, especialmente quando utilizada em um conjunto de Árvores de Decisão, sendo comparável ou superior a outros algoritmos tradicionais na área de reconhecimento de padrão (BREIMAN, 2001; DIETTERICH, 2000). Além disso, outra consideração importante é que as Árvores de Decisão têm poucos parâmetros que precisam de ajuste, o que simplifica a etapa de definição do modelo (BONISSONE *et al.*, 2010; BREIMAN, 2001; JANIKOW, 1998; QUINLAN, 1986; UMANOL *et al.*, 1994).

O principal objetivo de qualquer classificador é realizar a predição da classe alvo baseado em diversos atributos de entrada. Mais especificamente, cada nó da Árvore de Decisão corresponde a uma das variáveis de entrada e esse nó é ramificado em todos os possíveis valores do atributo de entrada. Em comparação, os nós-folha representam a classe alvo de acordo com os valores dos atributos de entrada, representando um caminho da raiz até a folha (BUHRMAN; WOLF, 2002; DIETTERICH, 2000; WANG *et al.*, 2015), conforme exemplo simplificado ilustrado na Figura 1.

Uma das maneiras de treinar o classificador baseado em Árvore de Decisão é dividindo-se cada atributo de entrada em subconjuntos relevantes para a classificação, utilizando limiares de divisão. Esse processo é repetido em cada subconjunto derivado de maneira recursiva, chamada de particionamento recursivo. A recursão é concluída quando o subconjunto de um nó não adiciona mais valor às previsões ou quando toda a subdivisão é pertencente à mesma classe alvo. Esse processo de indução *top-down* de Árvores de Decisão

é um exemplo de um algoritmo popularmente chamado de guloso⁴ e é uma estratégia comum para construir Árvores de Decisão baseadas em treinamento supervisionado. No entanto, essa técnica não funciona para domínios simbólicos.

Figura 1 - Exemplo de árvore de decisão simbólica “ficar em casa ou ir para o parque”.



Fonte: Autor

Uma técnica disseminada que consegue manipular domínios simbólicos é o algoritmo *Induction Decision Tree* (ID3) proposto por (QUINLAN, 1986). Adicionalmente, o método ID3 assume domínios discretos com cardinalidades pequenas. Isso pode ajudar a aumentar a compreensão do modelo gerado, se comparado, por exemplo, ao algoritmo CART (*Classification and Regression Tree*). O CART cria inúmeros limiares para particionar os dados numéricos, reduzindo a compreensão do modelo gerado devido ao extenso particionamento de dados (JANIKOW, 1998). No entanto, embora o modelo gerado pelo método ID3 seja mais compreensível, é necessário o particionamento de dados *a priori*. O ID3, assim como o CART, usa o particionamento recursivo para aprender e criar um modelo discriminativo, no qual os detalhes são representados em uma árvore.

A principal característica de árvores ID3, além do fato de precisarem de um particionamento de domínio *a priori*, é que a ID3 utiliza um método baseado em ganho de informação para realizar a escolha do atributo do nó. Esse método depende de duas suposições (QUINLAN, 1986). Assumindo que um conjunto de dados C contém p objetos da classe P e n objetos de uma classe N , as suposições são:

⁴ O algoritmo popularmente chamado de guloso é um algoritmo simples, usualmente utilizado em problemas de otimização. O algoritmo guloso escolhe em cada iteração o objeto mais relevante entre as opções fornecidas, de forma a efetuar uma escolha local ótima em cada iteração com a intenção de encontrar um ótimo global (QUINLAN, 1986).

1. Qualquer Árvore de Decisão correta para C classificará objetos na mesma proporção que sua representação em C. Um objeto arbitrário deverá pertencer à classe P com probabilidade $p / (p + n)$ e à classe N com probabilidade $n / (p + n)$;
2. Quando uma Árvore de Decisão é utilizada para classificar um objeto, a saída gerada é uma classe, ou seja, a Árvore de Decisão pode ser considerada como base para geração de classes P ou N, sendo que a informação $I(p,n)$ esperada para gerar essas classes é dada por (1):

$$I(p, n) = -\frac{p}{p+n} \log_2 \frac{p}{p+n} - \frac{n}{p+n} \log_2 \frac{n}{p+n} \quad (1)$$

Se um atributo A com valores $\{A_1, A_2 \dots A_v\}$ é utilizado como raiz da Árvore de Decisão, C é particionado em $\{C_1, C_2 \dots C_v\}$, onde C_i contém os objetos em C que tem o valor A_i de A e, assumindo que C_i contém p_i objetos da classe P e n_i da classe N, a informação esperada para a subárvore de C_i é $I(p_i, n_i)$. A informação esperada para a árvore com o atributo A como raiz é obtida através da média ponderada $E(A)$ indicada na Equação (2):

$$E(A) = \sum_{i=1}^v \frac{p_i+n_i}{p+n} I(p_i, n_i), \quad (2)$$

na qual o peso para o i-ésimo tronco da Árvore de Decisão depende da proporção de objetos em C que pertencem a C_i . Isso significa que o ganho de informação ao se utilizar o atributo A é:

$$\text{ganho}(A) = I(p, n) - E(A) \quad (3)$$

Segundo (QUINLAN, 1986) uma boa estratégia seria selecionar o atributo que será ramificado considerando maximizar o ganho de informação. Para isto, o ID3 examina todos os possíveis candidatos a atributos (A) e escolhe o atributo que maximize o ganho (A) (QUINLAN, 1986).

Um grande problema com as Árvores de Decisão é sua grande variação. Muitas vezes, uma pequena alteração nos dados pode resultar em um conjunto muito diferente de divisões, tornando a interpretação do modelo precária. A principal razão para essa instabilidade é a natureza hierárquica do processo: o efeito de um erro na divisão superior é propagado para todos os nós abaixo dele. Pode-se mitigar esse efeito até certo ponto tentando usar um critério

de divisão mais estável, mas a instabilidade inerente não é removida. É o custo associado pela estimativa de uma estrutura simples baseada em Árvore de Decisão a partir dos dados. Dessa forma, o uso de classificadores baseados em conjuntos (*ensembles*) tenta reduzir essa variância, realizando a média de diversas Árvores de Decisão (CLARKE; FOKOUE; ZHANG, 2009).

2.2 RANDOM FOREST

As Árvores de Decisão também são amplamente usadas como classificadores baseados em conjuntos. Para criar um conjunto, existem várias técnicas disponíveis (DIETTERICH, 2000). Uma das técnicas amplamente utilizadas na literatura é a *Random Forest* (BREIMAN, 2001). O método *Random Forest* (BREIMAN, 2001) é uma modificação do algoritmo *bagging*⁵ que constrói uma grande coleção de árvores não correlacionadas e, em seguida, calcula a média delas.

O método que deu origem à *Random Forest* foi proposto pela primeira vez por (HO, 1995) que afirmou que conjuntos de árvores (florestas) que são divididos com hiperplanos oblíquos podem aumentar sua taxa de classificação correta à medida que crescem sem sofrer de *overtraining*, desde que as florestas sejam aleatoriamente restritas. Neste processo, os dados de treinamento são projetados em um subespaço escolhido aleatoriamente antes de ajustar cada árvore ou cada nó. O conceito, criado por Amit e Geman, de se utilizar um subconjunto aleatório das decisões disponíveis ao dividir um nó, no contexto do crescimento de uma única árvore, também foi influente no projeto da *Random Forest*. Finalmente, a idéia de otimização aleatória dos nós, onde a decisão em cada nó é realizada por um procedimento randomizado, ao invés de uma otimização determinística, foi introduzida por Dietterich (DIETTERICH, 2000).

A introdução da *Random Forest* propriamente dita foi feita primeiramente em um artigo de Leo Breiman (BREIMAN, 2001) que descreve um método de construção de uma floresta de árvores não correlacionadas usando um procedimento semelhante ao CART,

⁵ O método *Bagging* é baseado em uma amostragem aleatorizada com substituição dos dados de entrada para definição do conjunto de treinamento de algoritmos de aprendizagem supervisionada, especialmente comum em Árvores de Decisão. Na amostragem com substituição, algumas amostras podem ser repetidas para o conjunto de treinamento, trazendo por consequência a redução da variância e ajudando a evitar o *overfitting* (BONISSONE *et al.*, 2010; BREIMAN, 2001).

combinado com otimização aleatória dos nós e *bagging*. De forma resumida, o método *Random Forest* combina duas técnicas distintas para gerar um conjunto de Árvores de Decisão não correlacionadas. O método utiliza a técnica de *bagging*, na qual cada classificador é criado com um conjunto de dados diferente, obtido através do conjunto de dados de treinamento por uma nova amostragem com substituição. Uma adaptação da técnica de randomização proposta por (DIETTERICH, 2000) também é usada na *Random Forest*, na qual os 20 melhores atributos são selecionados em cada nó e um deles é selecionado aleatoriamente para ramificar o nó.

A *Random Forest* se comporta como muitas árvores de classificação na questão de geração do modelo e crescimento dos nós das árvores. Para classificar um novo objeto a partir de um vetor de entrada, é necessário utilizar o vetor de entrada em cada uma das árvores na floresta. Sendo assim, cada árvore gera uma classificação e, por consequência, essa classe representa o “voto” da árvore. A floresta escolhe a classificação com mais votos (considerando todas as árvores da floresta). Cada árvore em uma *Random Forest* é construída da seguinte forma:

1. Se o número de amostras do conjunto de treinamento for N , é realizada uma amostragem de N amostras aleatoriamente, mas com substituição dos dados originais. Esta amostragem com substituição é utilizada como conjunto de treinamento para o crescimento da árvore;
2. Se houver M variáveis de entrada, um número $m \ll M$ é especificado, de tal forma que em cada nó m variáveis são selecionadas aleatoriamente e a melhor divisão encontrada de acordo com as m variáveis selecionadas é utilizada para dividir o nó. O valor de m é mantido constante durante o crescimento da floresta;
3. Não é aplicado nenhum algoritmo de *pruning*⁶. Ou seja, cada árvore é gerada na maior extensão possível.

No artigo original de (BREIMAN, 2001), foi demonstrado que a taxa de erro da floresta depende de duas situações: a correlação entre qualquer árvore na floresta (quanto

⁶ O algoritmo de *pruning* é uma técnica que reduz o tamanho de Árvores de Decisão removendo galhos que fornecem pouco ganho de informação na classificação das amostras. O *pruning* reduz a complexidade da Árvore de Decisão, reduzindo o *overfit* do modelo e, por consequência, muitas vezes aumentando a taxa de acerto do classificador (JANIKOW, 1998; QUINLAN, 1986; SAFAVIAN; LANDGREBE, 1991).

maior a correlação maior a taxa de erro da floresta) e o impacto de cada árvore individual na floresta (aumentar o impacto das árvores individuais diminui a taxa de erro da floresta).

Sendo assim, é observado que a redução do número de m variáveis selecionadas reduz a correlação e o impacto, enquanto elevando o m ambos aumentam. Esse é o único parâmetro ajustável para o qual a *Random Forest* é sensível. E (BREIMAN, 2001) sugeriu a utilização de $m = \sqrt{M}$, na qual M representa o número total de variáveis de entrada.

Podem-se citar algumas vantagens do método *Random Forest*:

- Apresenta bons resultados em termos de taxa de acerto;
- Funciona eficientemente com grandes bases de dados;
- Consegue manipular milhares de variáveis de entrada sem exclusão de alguma variável;
- Fornece estimativas de quais variáveis são importantes na classificação.

Muitas aplicações de reconhecimento de padrão envolvem diversas variáveis de entrada muitas vezes na ordem das centenas ou milhares, com cada uma contendo apenas uma pequena quantidade de informação, como é o caso da área de diagnósticos médicos e de processamento de sinais biológicos. Segundo (BREIMAN, 2001), utilizar uma única Árvore de Decisão para esses tipos de classificação de dados irá prover um resultado levemente superior a uma escolha aleatória da classe de saída. No entanto, o uso de um conjunto de árvores utilizando características aleatórias na sua construção (como a *Random Forest*) pode ajudar a elevar a taxa de acerto do sistema, como foi comprovado por Breiman em um conjunto de dados simulados com 1000 variáveis de entrada, 1000 exemplos no conjunto de treinamento e um conjunto de 4000 exemplos de teste, no qual foi obtido uma taxa de acerto comparável à taxa do método de Bayes (BREIMAN, 2001).

Adicionalmente, quando se trata de aprendizagem supervisionada, um dos aspectos mais importantes é o conjunto de treinamento cujas classes são conhecidas de antemão. A idéia por trás do aprendizado supervisionado é criar um conjunto de regras que possam determinar uma classe de qualquer objeto a partir de um conjunto de atributos (QUINLAN, 1986). No entanto, os atributos disponíveis nem sempre são capazes de fornecer informações suficientes para diferenciar a classe à qual o objeto pertence sem ambiguidade. Por exemplo, se os dados de treinamento contiverem dois objetos com valores idênticos, e mesmo que esses objetos pertençam a classes diferentes, é impossível distingui-los, implicando que há uma contradição no banco de dados. Portanto, o uso de categorias binárias pode não ser a melhor opção dependendo do tipo de sinal. Nesse caso, o uso da lógica não clássica pode agregar

robustez ao algoritmo de decisão ao considerar que podem existir contradições no sistema (COBREROS, 2011). Em função disso, nesta tese foi proposto, de forma inovadora, o uso da Lógica Paraconsistente aliada ao algoritmo de *Random Forest*, de forma a obter-se o melhor da capacidade da Lógica Paraconsistente para lidar com dados inconsistentes, a compreensibilidade gerada pela Árvore de Decisão e obtendo uma baixa suscetibilidade a ruído, ao utilizar a abordagem de *Random Forest*.

2.3 LÓGICA PARACONSISTENTE

A lógica clássica ainda ocupa uma posição dominante no campo da lógica formal, incluindo tanto a lógica matemática, como a filosófica e a computacional. No entanto, no último século, o crescimento de lógicas alternativas à lógica clássica tem ganhado mais espaço. Essas lógicas são comumente chamadas de não clássicas (ABE; AKAMA; NAKAMATSU, 2015). De fato, existe mais de um tipo de lógica não clássica, considerando que as mesmas podem ser diretamente antagonistas, no sentido de negar alguns princípios básicos da lógica clássica. Por exemplo, *many-valued logic* nega o princípio binário da lógica clássica, permitindo múltiplos valores-verdade. Como exemplo, a lógica *fuzzy* pertence à família das *many-valued logic*. Dessa forma a lógica *fuzzy* consegue lidar com diversos graus de informações e também informações vagas. Assim como *fuzzy*, a Lógica Paraconsistente também é considerada uma lógica rival à lógica clássica, por lidar com contradições em sua estrutura teórica sem se tornar trivial (ABE; AKAMA; NAKAMATSU, 2015). Um dos aspectos mais interessantes do desenvolvimento de lógicas não clássicas é o fato de que elas conseguem superar algumas limitações da lógica tradicional, especialmente em aplicações que envolvem conhecimento humano. Como exemplo, a Lógica Paraconsistente tem muitas aplicações e pode servir como base para a solução de alguns problemas na engenharia, porque os mesmos, muitas vezes, envolvem informações inconsistentes (AKAMA, 2016).

A Lógica Paraconsistente foi proposta por J. Lukasiewicz e N.A. Vasilév por volta de 1910, de maneira independente. Eles visualizaram a possibilidade de uma lógica que previsse contradição em sua estrutura básica (DA SILVA FILHO; ABE; TORRES, 2008). No entanto, somente em 1954 foi desenvolvido o sistema inicial da Lógica Paraconsistente, incluindo cálculos proposicionais e todos os níveis lógicos por autoria de N.C.A. Costa (DA COSTA, 1974). Em continuidade ao desenvolvimento da técnica, em 1976 o termo "Lógica Paraconsistente" foi proposto por Francisco Miró Quesada. Ao longo desses anos foi elaborada a Lógica Paraconsistente, de maneira que é possível manipular sistemas

inconsistentes sem ser necessária a eliminação de contradições e conseguir manipular as informações sem trivialização (DA SILVA FILHO; ABE; TORRES, 2008). Nesta seção serão apresentados os conceitos básicos da Lógica Paraconsistente Anotada $P\tau$, de acordo com a proposta de (DA COSTA; ABE; SUBRAHMANIAN, 1991).

2.3.1 Fundamentos da Lógica Paraconsistente

A Paraconsistência é o estudo de sistemas lógicos nos quais a presença de contradição não implica diretamente em trivialidade (ABE; AKAMA; NAKAMATSU, 2015), isto é, as Lógicas Paraconsistentes são capazes de lidar com cenários contraditórios, evitando a trivialidade por meio da rejeição do Princípio da Explosão, no sentido de que essas teorias não se banalizam na presença de (pelo menos algumas) sentenças contraditórias. Diferente da lógica tradicional, na Lógica Paraconsistente a trivialidade não está necessariamente interligada à contradição (CARNIELLI; CONIGLIO, 2016).

Com isso, pode-se afirmar que o primeiro passo para compreender a Lógica Paraconsistente é a distinção entre trivialidade e a contradição. No entanto, primeiramente é necessário definir os termos e semânticas utilizadas na lógica para, posteriormente, realizar a definição dos termos que distinguem a Lógica Paraconsistente da Lógica Clássica.

Para introduzir a linguagem da lógica anotada proposicional $P\tau$ segundo as referências de (ABE; AKAMA; NAKAMATSU, 2015; AKAMA, 2016; BLAIR; SUBRAHMANIAN, 1989; DA COSTA; ABE; SUBRAHMANIAN, 1991), primeiramente, é necessário denotar por L a linguagem de $P\tau$. Lógicas anotadas são baseadas em uma rede finita arbitrária fixa chamada de rede de valores de verdade, denotada por $\tau = \langle |\tau|, \leq, \sim \rangle$, que é a rede completa com o ordenamento \leq e o operador $\sim: |\tau| \rightarrow |\tau|$. O símbolo \sim é denotado como negação em nível atômico de $P\tau$. Também é assumido que T é o elemento superior e \perp é o elemento inferior, respectivamente. Além disso, são utilizadas duas operações teóricas do *lattice*: \vee para o menor limite superior e \wedge para o maior limite inferior. A seguir são apresentados algumas definições e teoremas básicos para a compreensão da Lógica Paraconsistente.

Definição 3.1 (Simbologia) Os símbolos da $P\tau$ são definidos à seguir:

1. Símbolos Proposicionais: p, q, \dots
2. Constantes de Anotação: $\mu, \lambda, \dots \in |\tau|$
3. Conectivos lógicos: \wedge (conjunção), \vee (disjunção), \rightarrow (implicação), e

\neg (negação)

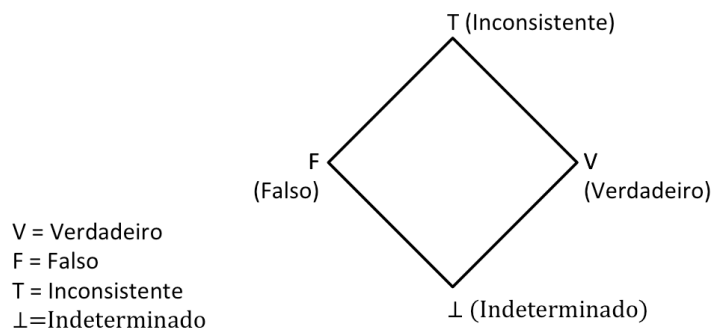
4. Parêntese: (e)

Definição 3.2 (Fórmulas) As Fórmulas são definidas como:

1. Se p é uma proposição e $\mu \in |\tau|$ é uma constante de anotação, então p_μ é uma fórmula chamada de *annotated atom*.
2. Se F é uma fórmula, então $\neg F$ é uma fórmula.
3. Se F e G são fórmulas, então $F \wedge G$, $F \vee G$, $F \rightarrow G$ são fórmulas.
4. Se p é um símbolo proposicional e $\mu \in |\tau|$ é uma constante de anotação, então uma fórmula da forma $\neg^k p_\mu$ ($k \geq 0$) é chamada de hiperliteral. Uma fórmula que não é hiperliteral é chamada de fórmula complexa.

O conjunto $\tau = \{T, V, F, \perp\}$ é o conjunto de valores de verdade da lógica de quatro valores. Os valores de verdade $\{T, V, F, \perp\}$ correspondem, respectivamente, aos valores de verdade “indefinido”, verdadeiro (no sentido da lógica de dois valores), falso (também no sentido de dois valores) e “superdefinido” (que também pode ser considerado inconsistente na lógica de dois valores). No conjunto desses objetos $\tau = \{T, V, F, \perp\}$ coloca-se uma estrutura matemática que será um reticulado com operador $\tau = \langle |\tau|, \leq, \sim \rangle$ que pode ser caracterizado pelo seguinte diagrama de Hasse (DA SILVA FILHO; ABE; TORRES, 2008), conforme ilustrado na Figura 2.

Figura 2 - Diagrama de Hasse de quatro vértices.



Fonte: adaptado de (DA SILVA FILHO; ABE; TORRES, 2008).

Definição 3.3. $(\forall x, y \in \tau), x \preceq y \Leftrightarrow x = y \vee x = \perp \vee y = T$. O ordenamento \preceq é representado pelo diagrama de Hasse da Figura 2. É utilizada a notação $x \succeq y$ sse $y \preceq x$,

onde $x, y \in \tau$. Além disso, $\langle e \rangle$ são as restrições irreflexivas de \leq e \geq , respectivamente. O conjunto τ é uma estrutura completa sob esta ordem.

Intuitivamente, o átomo de anotação $A: V$ pode ser lido como uma aproximação de “ A é conhecido por ser verdadeiro”. Da mesma forma, para o átomo de anotação $A: F$ pode-se considerar a afirmação “ A é conhecido por ser falso”.

Definição 3.4 (Propriedade Complementar) Um valor-verdade $\mu \in \tau$ tem a propriedade complementar se houver um λ tal que $\lambda \leq \mu$ e $\sim \lambda \leq \mu$. Um conjunto $\tau' \subseteq \tau$ tem a propriedade complementar se e somente se houver algum $\mu \in \tau'$ tal que μ tem a propriedade complementar.

Definição 3.5 (Intervalo) Suponha que I seja uma interpretação da linguagem L . O intervalo de I , denotado por Intervalo (I), é definido como sendo intervalo (I) = $\{\mu \mid (\exists A \in BL) I(A) = \mu\}$, onde BL denota o conjunto de todos os *ground atoms* em L .

Para $P\tau$, os *ground atoms* correspondem a variáveis proposicionais. Se o intervalo da interpretação I satisfizer a propriedade complementar, então o Teorema 3.1 pode ser estabelecido.

Teorema 3.1 Seja I uma interpretação tal que o intervalo (I) tenha a propriedade complementar. Então, há uma variável proposicional p e $\mu \in |\tau|$ tal que $v(p_\mu) = 1$ e $(\neg p_\mu) = 1$.

Proof Como o intervalo(I) tem a propriedade complementar, existe uma variável proposicional p e um $\delta \in \tau$ satisfazendo (1) $I(p) = \delta$ e (2) existe um $\gamma \in \tau$ tal que $\gamma \leq \delta$ e $\sim \gamma \leq \delta$. De acordo com (1), $I(p) \geq \delta$ é válido. Assim, temos que $v(p_\gamma) = 1$. Da mesma forma, temos que $v(\neg p_\gamma) = 1$ por (2). A partir de ambos, podemos alcançar o teorema pela Definição 3.3.

O Teorema 3.1 afirma que existe um caso em que, para alguma variável proposicional ser ao mesmo tempo verdadeira e falsa, isto é, inconsistente, o fato está intimamente ligado à noção de Paraconsistência.

Definição 3.6 (Inconsistência) Uma interpretação I é inconsistente se e somente se houver uma variável proposicional p e uma constante anotada $\mu \in |\tau|$ tal que $v(p_\mu) = v(\neg p_\mu) = 1$.

Portanto, inconsistência significa que ambos A e $\neg A$ são simultaneamente verdadeiros para algum A .

Definição 3.7 (Não trivialidade) Uma interpretação I é não trivial se e somente se houver uma variável proposicional p e uma constante anotada $\mu \in |\tau|$ tal que $v(p_\mu) = 0$.

Pela Definição 3.7, é possível afirmar que nem todo átomo é válido se uma interpretação é não trivial.

Definição 3.8 (Paraconsistência) É possível afirmar que uma interpretação I é paraconsistente se e somente se ela é tanto inconsistente quanto não trivial. $P\tau$ é chamado paraconsistente se e somente se há uma interpretação de I de $P\tau$ tal que I seja paraconsistente.

A Definição 3.8 permite o caso em que A e $\neg A$ são verdadeiros, mas alguma fórmula B é falsa em algumas interpretações paraconsistentes I . Como foi mencionado anteriormente, as Lógicas Paraconsistentes são capazes de lidar com cenários contraditórios, evitando a trivialidade por meio da rejeição do princípio da explosão.

Nesse sentido, pode-se afirmar que a Lógica Paraconsistente é robusta em relação a inconsistências (dados paraconsistentes) no sentido de não trivializar as proposições e conseguir chegar a uma conclusão válida, ainda que a mesma represente o valor lógico de inconsistência (T), definido previamente.

Definição 3.9 (Paracompleteza) Uma interpretação I é paracompleta se houver uma variável proposicional p e uma constante anotada $\lambda \in |\tau|$ tal que $v(p_\lambda) = v(\neg p_\lambda) = 0$. $P\tau$ é chamada paracompleta se e somente se houver uma interpretação I de $P\tau$ tal que I seja paracompleto.

Analisando a Definição 3.8, é possível notar que na interpretação paracompleta I , ambos A e $\neg A$ são falsos. É afirmado que $P\tau$ é não alético se e somente se for paraconsistente e paracompleto. Intuitivamente, a Lógica Paraconsistente pode lidar com informações

inconsistentes e a Lógica Paraconsistente pode lidar com informações incompletas. Isso significa que lógicas não aléticas, como lógicas anotadas, podem servir como lógicas para expressar informações inconsistentes e incompletas. Esse é um dos motivos para o estudo das lógicas anotadas aplicadas ao reconhecimento de padrão a partir de sinais biológicos.

Tais inconsistências podem facilmente ocorrer em sistemas baseados em conhecimento e não devem permitir a trivialização de um modelo. No entanto, sistemas baseados em conhecimento que se apoiam na lógica clássica não podem lidar com a situação, uma vez que o modelo é trivial. Em Blair e Subrahmanian (BLAIR; SUBRAHMANIAN, 1989) é mostrado como a lógica anotada de quatro valores pode ser usada para descrever esta situação.

O conjunto de valores de verdade com negação definida como complementação booleana forma uma lógica anotada. O intervalo unitário $[0, 1]$ dos valores de verdade com $\neg x = 1 - x$ é considerado como a base da lógica anotada para o raciocínio qualitativo ou difuso. Nesse sentido, lógicas probabilísticas e difusas poderiam ser generalizadas como lógicas anotadas. O intervalo $[0, 1] \times [0, 1]$ de valores de verdade pode ser usado para lógicas anotadas para raciocínio evidencial. Aqui, a atribuição do valor de verdade (μ_1, μ_2) à proposição p pode ser pensada como o grau de crença em p é μ_1 , enquanto o grau de descrença é μ_2 . A negação pode ser definida como $\neg (\mu_1, \mu_2) = (\mu_2, \mu_1)$. Note-se que a atribuição de $[\mu_1, \mu_2]$ a uma proposição p por uma interpretação I não necessariamente satisfaz a condição $\mu_1 + \mu_2 \leq 1$. Isso contrasta diretamente com o raciocínio probabilístico. O conhecimento sobre um determinado domínio pode ser obtido a partir de diferentes especialistas (nesse domínio) e esses especialistas podem ter visões diferentes (ABE; AKAMA; NAKAMATSU, 2015; AKAMA, 2016).

Abe foi um dos pesquisadores que se aprofundaram na Lógica Paraconsistente Anotada Evidencial, da qual são derivados diversos trabalhos utilizando a Lógica Paraconsistente, como sistemas especialistas em diversas aplicações de classificação (ABE, 1992). No próximo subcapítulo será descrito com mais detalhes a Lógica Paraconsistente Anotada.

2.3.2 Lógica Paraconsistente Anotada Evidencial

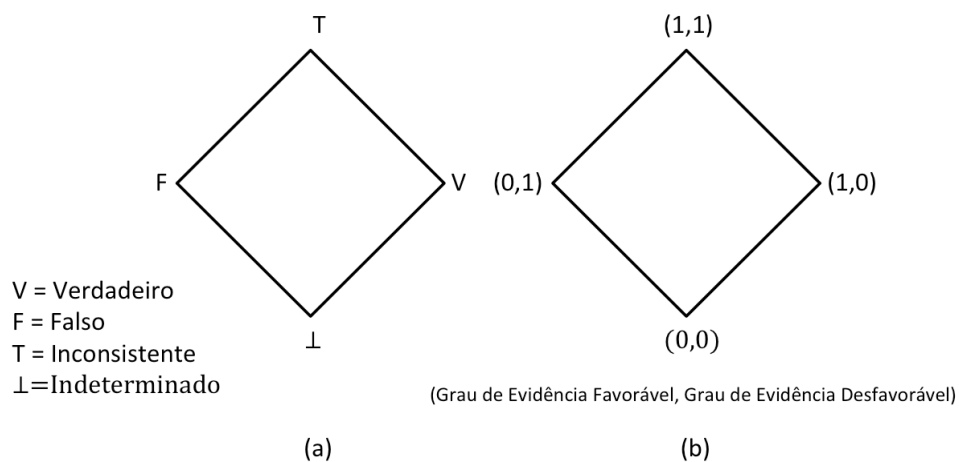
A Lógica Paraconsistente Anotada (LPA) é uma classe da Lógica Paraconsistente Evidencial que realiza o tratamento de sinais representados por anotações permitindo uma descrição e equacionamento por meio de algoritmos. Na Lógica Paraconsistente Anotada as

fórmulas proposicionais vêm acompanhadas de anotações. Cada anotação, pertence a um reticulado finito τ , que atribui valores à sua correspondente fórmula proposicional. Uma Lógica Paraconsistente Anotada (LPA) pode ter como reticulado finito, o de “quatro estados”, conforme exemplifica a Figura 3(a) (MÁRIO, 2007).

Os primeiros estudos sobre os fundamentos da Lógica Paraconsistente Anotada foram realizados por N. C. A. Da Costa em 1991 (DA COSTA; ABE; SUBRAHMANIAN, 1991). Em 1992, J. M. Abe apresentou estudos sistemáticos dessas lógicas, demonstrando teoremas de correlações e de completeza para as lógicas de primeira ordem correspondente. Estes estudos mostraram que a teoria anotada geral de conjuntos envolve, como caso particular, a Teoria dos Conjuntos Difusos, estudada por L. Zadeh em 1973 e C. V. Negoita em 1975 (ABE, 1992).

A anotação da Lógica Paraconsistente pode ter n valores (ABE, 1992), sendo que as proposições da Lógica Paraconsistente Anotada Evidencial $E\tau$ são do tipo $p(\mu, \lambda)$, onde p é uma proposição, e suas evidências $\mu, \lambda \in [0, 1]$ (intervalo real unitário fechado). Sendo assim, μ indica o Grau de Evidência favorável de p e λ indica o Grau de Evidência contrária de p (ABE, 2010). Desse modo é possível utilizar um Reticulado de Hasse (Figura 3(a)) formado pelos pares ordenados de forma que $\tau = \{(\mu, \lambda) | \mu, \lambda \in [0, 1] \subset \mathfrak{R}\}$, conforme ilustrado na Figura 3 (b).

Figura 3 - (a) Reticulado finito de Hasse e (b) Reticulado finito de Hasse com anotações da Lógica Paraconsistente Anotada.



Fonte: adaptado de (DA SILVA FILHO; ABE; TORRES, 2008).

Considerando-se o Reticulado de Hasse com a anotação de dois valores, o par (μ, λ) é denominado de constante de anotação. Esse par é um elemento de $[0,1] \times [0,1]$, ou também pode-se dizer que $(\mu, \lambda) \in [0,1]^2$, em que $[0,1]$ é o intervalo unitário real. Uma relação de ordem é definida de forma que: $(\mu_1, \lambda_1) \leq (\mu_2, \lambda_2) \Leftrightarrow \mu_1 \leq \mu_2$ e $\lambda_1 \leq \lambda_2$ (ABE; LOPES; NAKAMATSU, 2014). Analisando o Reticulado da Figura 3, é possível determinar que:

- $p(1.0, 0.0)$ representa uma proposição verdadeira (evidência favorável total e evidência contrária nula);
- $p(0.0, 1.0)$ representa uma proposição falsa (evidência favorável nula e evidência contrária total);
- $p(1.0, 1.0)$ representa uma proposição inconsistente (evidência favorável total e evidência contrária total);
- $p(0.0, 0.0)$ representa uma proposição paracompleta (evidência favorável nula e evidência contrária nula);
- $p(0.5, 0.5)$ representa uma proposição indefinida (evidência favorável igual à evidência contrária de 0.5).

Para facilitar o entendimento, considere-se o exemplo de um dado sinal de sEMG adquirido sobre o músculo bíceps de onde são extraídas duas características deste sinal mioelétrico (c_1 e c_2) Após uma transformação dessas características em Graus de Evidência μ_1 e $\mu_2 \in [0,1]^2$, e sendo $\lambda_2 = 1 - \mu_2$, tem-se a proposição $p =$ “está ocorrendo uma atividade muscular no bíceps” com dois graus de evidência $p(\mu_1, \lambda_2)$. Sendo assim, se:

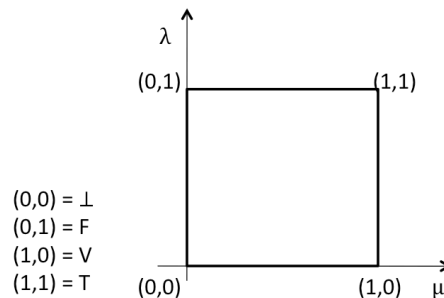
- $p(1.0, 0.0)$, em que a evidência favorável $\mu_1 = 1$ e evidência desfavorável $\lambda_2=0$, pode ser traduzida como "está ocorrendo uma atividade muscular no Bíceps" (Proposição Verdadeira);
- $p(0.0, 1.0)$, em que a evidência favorável $\mu_1 = 0$ e evidência desfavorável $\lambda_2=1$ pode ser traduzida como "não está ocorrendo uma atividade muscular no Bíceps" (Proposição Falsa);
- $p(1.0, 1.0)$ em que a evidência favorável $\mu_1 = 1$ e evidência desfavorável $\lambda_2=1$ pode ser traduzida como uma Proposição Inconsistente (as informações sobre a atividade muscular são contraditórias);
- $p(0.0, 0.0)$ em que a evidência favorável $\mu_1 = 0$ e evidência desfavorável $\lambda_2=0$ pode ser traduzida como uma Proposição Paracompleta (não existem informações suficientes sobre a atividade muscular);

- $p(0.5, 0.5)$ em que a evidência favorável $\mu_1 = 0.5$ e evidência desfavorável $\lambda_2=0.5$ pode ser traduzida como uma Proposição Indefinida.

Uma propriedade interessante que a Lógica Paraconsistente Anotada Evidencial Et apresenta é quando se analisam as negações de proposições. Seja $p(\mu, \lambda)$ uma proposição anotada, então $\neg p(\mu, \lambda) = p(\lambda, \mu)$. Pode-se perceber que a negação de p é equivalente à mesma proposição p , porém com os graus de evidência invertidos. Isso representa uma importante propriedade da lógica em que $\neg p(\mu, \lambda) \leftrightarrow p_{\neg}(\mu, \lambda)$ (ABE, 2010).

Segundo (DA SILVA FILHO; ABE; TORRES, 2008) para uma melhor representação de uma anotação na Lógica Paraconsistente Anotada de 2 valores, e também para permitir a utilização da lógica no tratamento de incertezas, podem ser realizadas algumas transformações no Quadrado Unitário no Plano Cartesiano. Para isso é adotado inicialmente um sistema de coordenadas cartesianas, e as representações das anotações (μ, λ) serão representadas por pontos nesse sistema de coordenadas, conforme esboço da Figura 4.

Figura 4 - Quadrado unitário no plano cartesiano (QUPC).



Fonte: adaptado de (DA SILVA FILHO; ABE; TORRES, 2008).

Nesse quadrado unitário, os valores do Grau de Evidência favorável μ são representados pelo eixo x e os valores de λ são representados no eixo y. Dessa forma são associados os estado lógicos: T a $(1,1)$, \perp a $(0,0)$, F a $(0,1)$ e V a $(1,0)$.

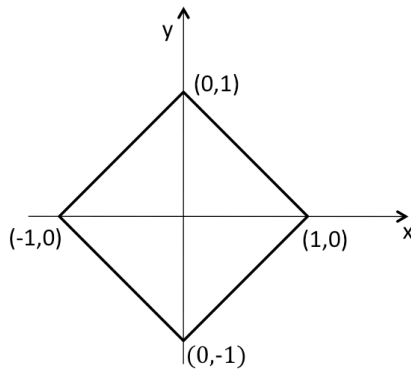
Por meio de transformações lineares em QUPC e no reticulado representativo da LPA τ , representado por \mathcal{L} na Figura 5, pode-se chegar à transformação (Equação (4)), conforme processo detalhado por (DA SILVA FILHO; ABE; TORRES, 2008):

$$T(x, y) = (x - y, x + y - 1) \quad (4)$$

Relacionando os componentes da transformação $T(x, y)$, é possível converter os pontos de QUPC em pontos de \mathcal{L} , onde $x = \mu$ (Grau de Evidência favorável) e $y = \lambda$ (Grau de Evidência desfavorável), conforme Figura 6. Assim chega-se à Equação (5):

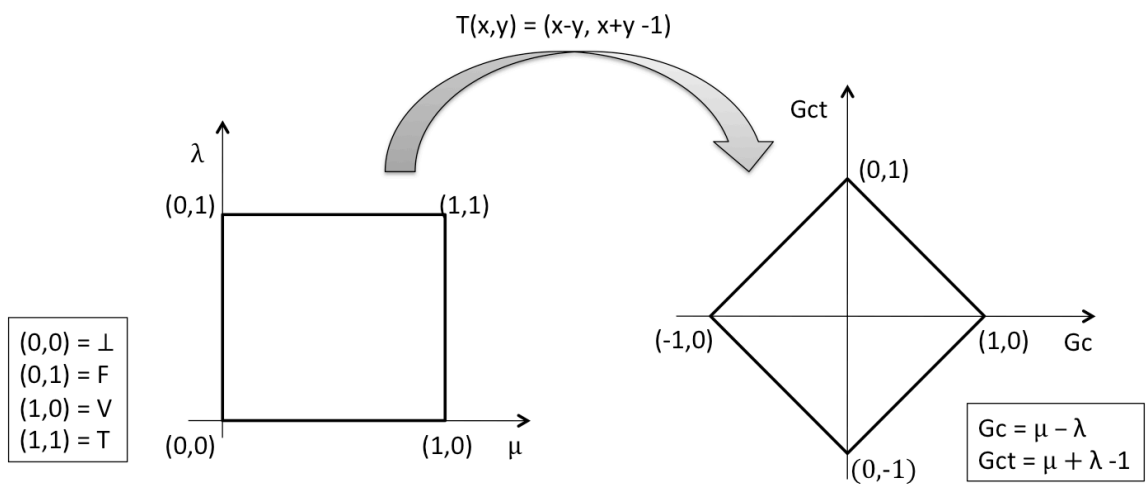
$$T(x, y) = T(\mu, \lambda) \tag{5}$$

Figura 5 - \mathcal{L} : Reticulado τ em um novo sistema de coordenadas.



Fonte: adaptado de (DA SILVA FILHO; ABE; TORRES, 2008).

Figura 6 - Conversão de QUPC em pontos de \mathcal{L} .



Fonte: adaptado de (DA SILVA FILHO; ABE; TORRES, 2008).

Do primeiro termo obtido no par ordenado da equação da transformação (Equação (4)) tem-se: $x - y = \mu - \lambda$, que é denominado Grau de certeza G_c , sendo representado pela Equação (6):

$$G_c = (\mu - \lambda) \quad (6)$$

Os valores do Grau de Certeza pertencem ao conjunto \mathfrak{R} , variam no intervalo fechado +1 e -1 e encontram-se no eixo horizontal do reticulado, o qual foi nomeado “Eixo dos Graus de Certeza”. Analisando os valores de G_c , é possível chegar à conclusão de que quando $G_c = +1$, representa o resultado lógico da análise paraconsistente Verdadeiro e, intuitivamente, quando $G_c = -1$ o resultado lógico da análise paraconsistente será Falso (DA SILVA FILHO; ABE; TORRES, 2008).

Do segundo termo obtido no par ordenado da equação da transformação (Equação (4)) obtém-se: $x + y - 1 = \mu + \lambda - 1$, que é denominado Grau de Contradição G_{ct} , de forma a ser representado pela Equação (7):

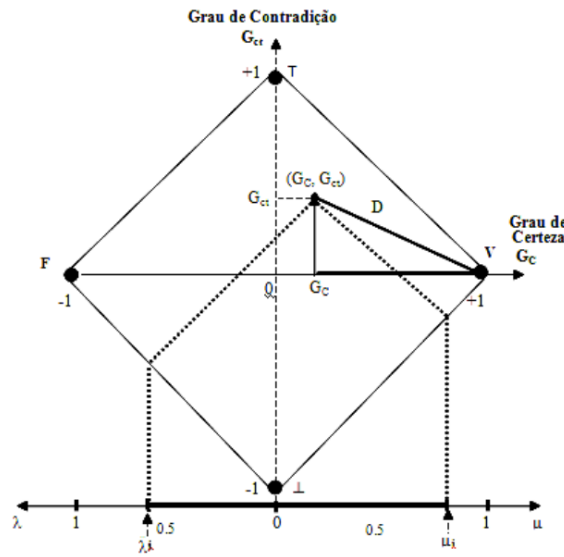
$$G_{ct} = (\mu + \lambda - 1) \quad (7)$$

Os valores do Grau de Contradição pertencem ao conjunto \mathfrak{R} , variam no intervalo fechado +1 e -1 e se encontram no eixo vertical do reticulado, o qual foi nomeado “Eixo dos Graus de Contradição”. Analisando os valores de G_{ct} , é possível chegar à conclusão de que quando $G_{ct} = +1$, representa resultado lógico da Análise Paraconsistente Inconsistente e, intuitivamente, quando $G_{ct} = -1$, o resultado lógico da análise paraconsistente será Indeterminado (DA SILVA FILHO; ABE; TORRES, 2008). Na Figura 7 é apresentado um exemplo gráfico de uma projeção dos valores de Grau de Certeza e Grau de Contradição a partir das anotações μ e λ .

Na realização da Análise Paraconsistente é calculado um valor de Grau de Certeza que se encontra no intervalo de [-1,1]. No entanto, para se utilizar o Grau de Certeza Resultante como Grau de Evidência em outra proposição, é necessário realizar a normalização de seus valores, sendo o Grau de Certeza representado pela Equação (6), e seu valor normalizado, representado pelo seu Grau de Evidência Resultante (μE), será obtido pela Equação (8):

$$\mu E = \frac{(G_c + 1)}{2} \quad (8)$$

Figura 7 - Projeções do Grau de Certeza e Grau de Contradição.



Fonte: (DA SILVA FILHO, 2009).

Existe também a possibilidade de atenuar os efeitos da contradição no Grau de Evidência da Saída, cujo valor é chamado de Grau de Certeza Real. A partir de um estudo mais detalhado, conforme mostrado em (DA SILVA FILHO; ABE; TORRES, 2008), pode-se encontrar o Grau de Certeza Real (Gcr) como um valor projetado no eixo dos Graus de Certeza do reticulado através da Equação (9) se $Gc > 0$ ou Equação (10) se $Gc < 0$:

$$Gcr = 1 - \sqrt{(1 - |Gc|)^2 + (Gct)^2}, \quad \text{para } Gc > 0 \quad (9)$$

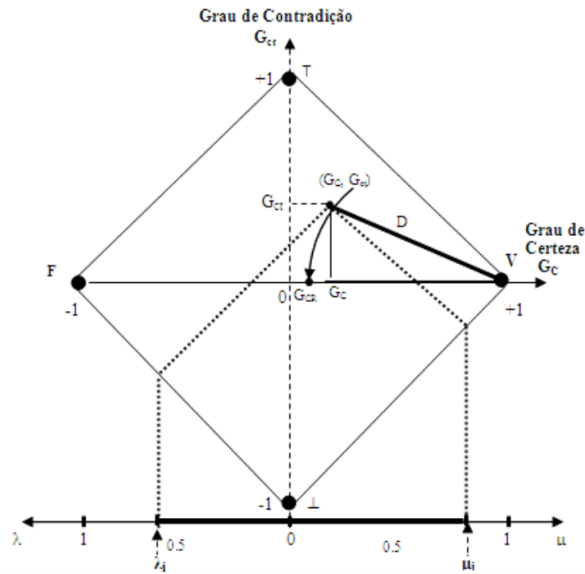
$$Gcr = \sqrt{(1 - |Gc|)^2 + (Gct)^2} - 1, \quad \text{para } Gc < 0 \quad (10)$$

A Figura 8 mostra o Grau de Certeza Real obtido pela projeção do valor da distância D sobre o eixo dos graus de certeza do reticulado para uma situação onde $Gc > 0$.

A partir de G_{cr} pode-se encontrar o seu valor normalizado que é denominado de Grau de Evidência Resultante Real μ_{ER} . Dessa forma μ_{ER} é obtido pela Equação (11):

$$\mu_{ER} = \frac{G_{cr} + 1}{2} \quad (11)$$

Figura 8 - Projeção do Grau de Certeza Real.



Fonte: (DA SILVA FILHO, 2009).

O valor do Grau de Contradição G_{ct} também pode ser normalizado, encontrando-se assim o Grau de Contradição μ_{ctr} Normalizado, representado na Equação (12):

$$\mu_{ctr} = \frac{\mu + \lambda}{2} \quad (12)$$

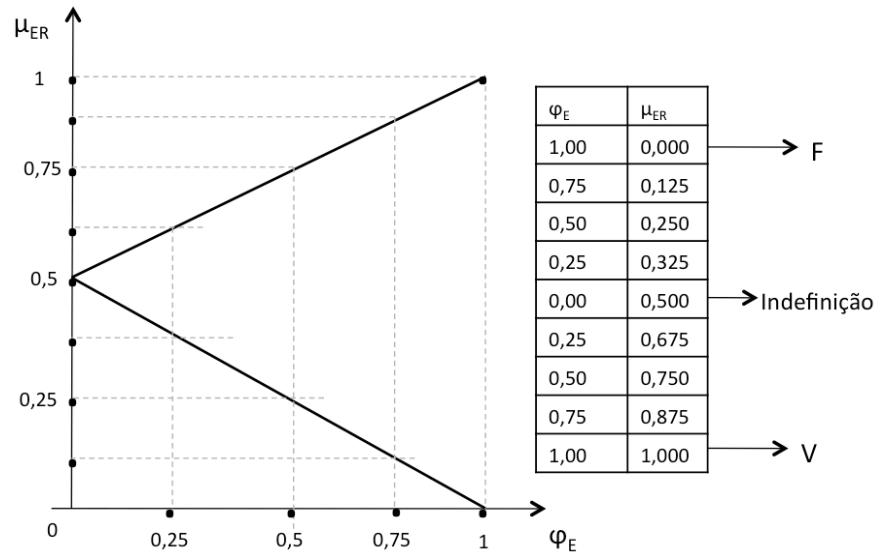
Através do Grau de Contradição Normalizado μ_{ctr} pode-se então obter um valor indicativo de quanto a contradição, na análise paraconsistente, está influenciando o valor obtido do Grau de Evidência Resultante Real. Esse indicativo é denominado de Intervalo de Evidência φ_E , e pode ser calculado pela Equação (13):

$$\varphi_E = 1 - |2\mu_{ctr} - 1| \quad (13)$$

Dessa forma, analisando o Intervalo de Evidência Resultante φ_E , percebe-se que quando o Grau de Contradição Normalizado for igual a $\mu_{ctr} = 1$, o que indica uma alta contradição com estado lógico de Inconsistente, o valor do Intervalo de Evidência será igual a zero. Da mesma forma, quando o Grau de Contradição for 0,5 (indicando que não existe influência da contradição na análise paraconsistente), o Intervalo de Evidência será igual a 1 (DA SILVA FILHO; ABE; TORRES, 2008). Pode-se também determinar um gráfico com a

relação do Intervalo de Evidência Resultante relacionado ao Grau de Evidência Resultante Real, representado na Figura 9.

Figura 9 - Gráfico das características entre valores de graus de evidência resultante relacionados aos valores de intervalo de evidência resultante.



Fonte: adaptado de (DA SILVA FILHO; ABE; TORRES, 2008).

Ao analisar o gráfico da Figura 9, é possível verificar que somente quando o Intervalo de Evidência for igual a 1 o Grau de Evidência Resultante está livre para excursionar entre o valor que representa o máximo da Falsidade e o valor que representa o máximo de Verdade.

2.3.3 Nó de Análise Paraconsistente

Os Nós de Análise Paraconsistente (NAPs) consistem em um algoritmo que consegue tratar e controlar sinais com informação imprecisa e contraditória (DA SILVA FILHO, 2009). Os NAPs têm duas saídas: o Grau de Evidência Resultante Real (μ_{ER}) e o Intervalo de Evidência Resultante (φ_E); e duas entradas: μ e λ . Os NAPs fazem tratamento de sinais conforme os fundamentos da Lógica Paraconsistente utilizando equações da metodologia LPA2v, descritas anteriormente, para obter o Grau de Evidência Resultante Real (μ_{ER}), acompanhado do seu respectivo Intervalo de Evidência Resultante (φ_E), conforme algoritmo descrito na Tabela 1. É interessante ressaltar que os NAPs podem ser interligados entre si, formando uma Rede de Análise Paraconsistente cuja arquitetura é dependente dos dados disponíveis e tipo de aplicação.

Tabela 1 - Descrição do Algoritmo A-1 do Nó de Análise Paraconsistente.

Etapas	Algoritmo
Entradas	μ_1 , sendo $0 \leq \mu_1 \leq 1$
	μ_2 , sendo $0 \leq \mu_2 \leq 1$
Equações	$\lambda_2 = 1 - \mu_2$
	$\mu_{ctr} = \frac{(\mu + \lambda)}{2}$
	$\varphi E = 1 - 2 * \mu_{crt} - 1 $
	$Gc = \mu - \lambda$
	$Gct = (\mu + \lambda) - 1$
	$D = \sqrt{(1 - Gc)^2 + (Gct)^2}$
	Se $Gc > 0$: $Gcr = 1 - D$
	Se $Gc < 0$: $Gcr = D - 1$
	Se $Gc == 0$: $Gcr = 0$
	Saídas

2.4 CONCEITOS BÁSICOS DE SINAIS MIOELÉTRICOS

Por definição, o sinal mioelétrico (sEMG) é proveniente do potencial de ação que percorre a fibra muscular, levando-a à contração. Luigi Galvani foi o primeiro a pesquisar a relação entre a contração muscular e a eletricidade em 1771. Realizando experimentos em sapos, despolarizava eletricamente as pernas dos mesmos com o toque de varas metálicas causando a contração muscular (ENDERLE; BRONZINO, 2012). O processo de contração muscular envolve diversas etapas, como a condução do impulso nervoso pelo neurônio motor, a transmissão sináptica e os potenciais de ação do músculo que, somados, representam o sinal mioelétrico, a contração das fibras musculares e, por consequência, o movimento muscular.

A célula nervosa motora, mais conhecida como neurônio motor é responsável por comandar diretamente a contração das fibras musculares (TORTORA; DERRICKSON, 2012). Cada neurônio possui um único axônio, que também é conhecido como fibra nervosa, sendo que os axônios transmitem os sinais neurais para a célula nervosa seguinte (GUYTON; HALL, 2012). Próximo aos seus terminais os axônios ramificam-se extensamente e é nessa extremidade que se encontra o botão sináptico, que contém bolsas denominadas vesículas sinápticas, onde são armazenadas substâncias químicas chamadas de neurotransmissores (TORTORA; DERRICKSON, 2012). Quando estimulado, o botão sináptico libera a

substância transmissora, fazendo com que o impulso nervoso seja transmitido (GUYTON; HALL, 2012).

O potencial de ação neural tem como objetivo propagar o impulso nervoso nos neurônios através de reações eletroquímicas. Quando um sinal é transmitido ao longo de uma fibra nervosa, o potencial da membrana passa por uma série de variações que, no seu conjunto, são chamadas de potencial de ação (GUYTON; HALL, 2012). O potencial de ação é iniciado quando ocorre uma mudança repentina no potencial de repouso negativo da membrana, tornando-se positivo, e, em seguida, ocorre seu término quando o potencial da membrana volta para seu potencial negativo (GUYTON; HALL, 2012).

Cada neurônio motor inerva diversas fibras musculares, de modo que a quantidade de fibras varia entre os tipos de músculos. A unidade motora é o conjunto de todas as fibras musculares inervadas por uma única fibra nervosa (GUYTON; HALL, 2012). De forma geral, músculos que são responsáveis por um controle mais fino são menores e compostos por poucas fibras musculares e mais fibras nervosas. Por outro lado, os músculos grandes, que não precisam de controle muito exato, podem ter unidades motoras com centenas de fibras musculares (GUYTON; HALL, 2012). Portanto, quanto menor for a taxa de inervação, maior será a capacidade de gradação fina da força total do músculo.

As fibras musculares de uma unidade motora não ficam todas agrupadas no músculo, mas, pelo contrário, ficam dispersas em outras unidades motoras, em grupos de 3 a 15 fibras. Por conseguinte, esses grupos ficam intercalados com outros grupos de diversas unidades motoras. Essa interpenetração permite que as unidades motoras distintas se contraíam em apoio umas às outras, e não de forma total como se fossem segmentos isolados (GUYTON; HALL, 2012).

De acordo com (DE LUCA, 1979), quando um neurônio envia um potencial de ação, todas as fibras musculares da sua unidade motora são estimuladas, porém não simultaneamente, ocorrendo pequenos atrasos entre as contrações. Como os potenciais de ação das fibras acabam se sobrepondo no tempo, o potencial resultante é considerado uma soma algébrica dos potenciais de ação nas diversas fibras de uma unidade motora. Esse potencial de ação resultante é chamado de MUAP (*Motor Unit Action Potential*).

Pelo fato do MUAP ter um período relativamente pequeno (1 a 12 ms), as unidades motoras devem ser ativadas repetitivamente para que se possa sustentar uma contração muscular por períodos maiores. Essa sequência resultante de MUAPs é denominada de MUAPT (*Motor Unit Action Potential Train*) (DE LUCA, 1979). É interessante ressaltar que como os músculos contêm fibras musculares de várias unidades motoras, ao se adquirir um

sinal de sEMG ele provavelmente será composto de vários MUAPTS. A captação de um MUAPT isolado só pode ser obtida se forem recrutadas apenas fibras de uma única unidade motora na vizinhança do eletrodo. De acordo com (DE LUCA, 1979) essa situação só ocorre em contrações musculares muito fracas.

Os potenciais de ação da unidade motora percorrem as fibras musculares e formam um campo eletromagnético em suas proximidades. Um eletrodo, próximo ao músculo em estudo, é capaz de detectar a resultante desses campos eletromagnéticos, sendo denominada de sinal mioelétrico (BASMAJIAN; DE LUCA, 1985).

Considerando a diferença entre os MUAPTs, a irregularidade na taxa de disparo dos neurônios motores e o fato da contração ter mais de um músculo envolvido, o sinal de sEMG pode ser descrito como sendo um processo estocástico, aproximando a função densidade espectral por uma função gaussiana, onde a amplitude instantânea do sinal é uma variável gaussiana de média zero. E características determinísticas podem ser encontradas dentro dos primeiros 200 ms do sinal de contração muscular (HUDGINS; PARKER; SCOTT, 1993).

Adicionalmente, é interessante ressaltar que a amplitude do sinal mioelétrico adquirido através de eletrodos de superfície (sEMG) está na faixa de 0 a 5 mV e a sua frequência varia na faixa de 0 a 500 Hz, sendo que a maior parte da energia se concentra entre 50 e 150 Hz (DE LUCA, 1997). Esses valores variam de acordo com o tipo de músculo analisado, o nível de contração muscular e a posição dos eletrodos.

3 ESTADO DA ARTE

Neste capítulo serão descritas pesquisas atuais abordando o uso aplicado da Lógica Paraconsistente e também trabalhos na área de reconhecimento de movimentos do membro superior através de sinais mioelétricos. Nessa pesquisa é possível observar que a Lógica Paraconsistente vem sendo utilizada com sucesso na área de reconhecimento de padrão, no entanto não foram encontrados trabalhos relatando o uso de LP aplicado na classificação de sinais de sEMG.

Na revisão do estado da arte também é observada uma tendência em pesquisas envolvendo o processamento de sinais de sEMG considerando degradações causadas por contaminantes comuns na área. De acordo com os estudos relatados, essas degradações apresentam um impacto negativo no uso clínico de próteses mioelétricas com controle natural e intuitivo, por ocasionar um decréscimo na taxa de acerto do classificador, usualmente levando longas e frequentes sessões de treinamento.

3.1 APLICAÇÕES DE RECONHECIMENTO DE PADRÃO EMPREGANDO LÓGICA PARACONSISTENTE

Neste subcapítulo serão apresentados trabalhos que empregam a Lógica Paraconsistente Anotada em aplicações que envolvem reconhecimento de padrão. Essas pesquisas abrangem áreas como EEG (ABE; LOPES; NAKAMATSU, 2014), processamento de voz (BARBON JÚNIOR; GUIDO; VIEIRA, 2009; DUNIN-KEPLICZ *et al.*, 2015; MÁRIO, 2007), reconhecimento de caracteres (SOUZA; ABE; NAKAMATSU, 2013), tratamento de incertezas (DA SILVA FILHO, 2012), entre outros (ABE, 2013).

No trabalho proposto por (SOUZA; ABE; NAKAMATSU, 2013) foi desenvolvida uma Rede Neural Paraconsistente (RNAP) para o reconhecimento de 13 caracteres MICR (*Magnetic Ink Character Recognition*), utilizada nos cheques de bancos brasileiros. As etapas de processamento incluíram desde aquisição e mapeamento da imagem, extração de característica e o reconhecimento de padrão. Foram extraídas 3 características da imagem: histograma das bordas externas, histograma da diagonal interna (direção direita para esquerda) e histograma da diagonal interna (direção esquerda para direita). Para o reconhecimento dos caracteres foi utilizada uma RNAP com a última camada contendo um algoritmo chamado de Algoritmo Paraconsistente Extrator de Contradição (DA SILVA FILHO, 2009). Também foi criada uma base de dados com 2 amostras de cada caractere, para serem utilizados como Graus de Evidência Padrão na RNAP. A base de dados de teste continha 2092 amostras para os 13

caracteres. Foi obtida uma taxa média de 97,8% de acerto que, segundo os autores, se mostra superior a outros trabalhos na área.

No estudo de (ABE; LOPES; NAKAMATSU, 2014) foi analisado o uso de RNAPs para o processamento de sinal EEG para realizar o diagnóstico de Alzheimer. Para uma análise morfológica das ondas EEG foram definidos 3 sistemas de especialistas para servirem de entrada numa RNAP de tratamento de contradição (SNAPtc). Os sistemas especialistas fazem a comparação entre as ondas de EEG adquiridas e as ondas padrão, enquanto a SNAPtc é responsável por realizar o reconhecimento do padrão. Os sistemas especialistas foram responsáveis pelas seguintes análises das ondas: número de picos, pontos similares e pontos diferentes.

Adicionalmente foi definido que para o diagnóstico de Alzheimer os sistemas especialistas eram responsáveis por analisar: decaimento da frequência, banda de concentração de altas frequências e banda de concentração de baixas frequências. Para esse estudo foi obtido 80% de sensibilidade e 73% de especificidade. Os autores concluíram que o estudo mostrou resultados promissores.

O estudo proposto por (GUIDO *et al.*, 2013) envolveu o desenvolvimento de um método baseado em Lógica Paraconsistente para reconhecimento de padrão. O método foi chamado de *Discriminative Paraconsistent Machine* (DPM), e é baseado em um modelo discriminativo de treinamento supervisionado que incorpora critérios de paraconsistência e permite o tratamento inteligente de contradições e incertezas. Segundo os autores, as DPMs podem ser aplicadas para resolver problemas em inúmeros campos da ciência. Os testes e discussões apresentados são capazes de demonstrar a eficácia do método em lidar com contradições e incertezas nos sistemas de entrada. A maior dificuldade relatada pelos autores foi na definição de um modelo adequado que conseguisse representar os conceitos de paraconsistência.

Na pesquisa de (DA SILVA FILHO *et al.*, 2016) foi utilizada a Lógica Paraconsistente Anotada de dois valores para analisar fragmentos de biopsia de tecidos de espectroscopia *Raman*. Para realizar a análise dos dados foi definida uma Rede de Análise Paraconsistente (PANnet) na qual vários Nós Paraconsistentes foram interconectados em uma PANnet. O conjunto de dados utilizados para validar o experimento compreende 146 fragmentos de biopsia de tecido por espectroscopia *Raman*, dos quais 30 representam tecido de pele normal, 96 representam tecido com carcinoma de células basais e 19 representam tecido com melanoma. Nessa base de dados, a Análise Paraconsistente foi capaz de discriminar corretamente 136 de um total de 145 fragmentos, obtendo uma taxa de diagnóstico correto de

93,793%. A aplicação de PANnet na análise de sinais de espectroscopia *Raman* foi capaz de produzir uma melhor discriminação de células do que processos estatísticos convencionais, como, por exemplo, *Principal Component Analysis (PCA)*. No entanto, apesar dos resultados parecerem promissores, o artigo carece de discussões com outros trabalhos relacionados e tampouco faz comparação com métodos inteligentes de reconhecimento de padrão, por isso não é possível analisar se os resultados obtidos são realmente superiores ao estado da arte em métodos de classificação.

Conforme apresentado, a Lógica Paraconsistente tem sido aplicada em diferentes áreas, nas quais, de acordo com o exposto, a mesma se apresenta apropriada na solução de problemas de reconhecimento de padrão, especialmente em contextos com dados inconsistentes e paracompletos. Nesse sentido, o uso da LP traz grandes chances de prover uma análise interessante de sinais biológicos como sinais de sEMG ao se aproveitar da habilidade de um algoritmo para lidar diretamente com sinais sujeitos a degradações, uma vez que a extração de todas as informações necessárias do sinal muscular para conseguir caracterizar o movimento correto que o membro está realizando é uma tarefa complexa e muitas vezes imprecisa, conforme será apresentado a seguir.

3.2 TÉCNICAS DE RECONHECIMENTO DE MOVIMENTOS BASEADAS EM SINAIS DE SEMG

Pelo fato de vários músculos que executam os movimentos da mão estarem localizados no antebraço, o uso do sinal mioelétrico desses músculos preservados é uma boa opção de controle para uma prótese do segmento mão-braço. É interessante salientar que, com o uso dos sinais provenientes dos músculos, o comando de um movimento da prótese pode ser realizado de maneira intuitiva, de forma que o indivíduo possa se acostumar mais facilmente com o uso da prótese. Como exemplo, diversos trabalhos tratam do projeto e construção de próteses mioelétricas para membros superiores (HUDGINS; PARKER; SCOTT, 1993).

Pode-se controlar a prótese mediante o reconhecimento dos sinais mioelétricos, relacionando esses padrões com os movimentos que a prótese deve realizar (FAVIEIRO; BALBINOT, 2011a; FAVIEIRO, 2012). O reconhecimento de padrões do sEMG é eficiente tanto para determinar o tipo de movimento que a prótese deve executar, de acordo com as características do sinal, assim como para evitar que sinais de outros músculos que não estão relacionados com a intenção de movimento possam ativá-la. Para realizar esse reconhecimento, torna-se necessário o desenvolvimento de um sistema de reconhecimento de padrão do sinal de sEMG em que podem-se implementar diversos algoritmos distintos. Alguns

exemplos de algoritmos de reconhecimento de padrão usualmente utilizados para a classificação de movimentos são: *Linear Discriminant Analysis* (LDA), sistemas *Fuzzy*, Redes Neurais, *Support Vector Machines* (SVM) (FAVIEIRO, 2012; FAVIEIRO; BALBINOT; BARRETO, 2011; GIJSBERTS *et al.*, 2014; GUO *et al.*, 2015; LORRAIN; JIANG; FARINA, 2010; MOURA; FAVIEIRO; BALBINOT, 2016), entre outros.

Um enfoque de pesquisas nos últimos anos é a preocupação com a usabilidade do dispositivo, baseado em sinais mioelétricos, em estudos clínicos, não somente enfatizando a taxa de acerto do algoritmo de classificação em um ambiente controlado de laboratório. Portanto, são as seguintes as preocupações frequentes no desenvolvimento de sistemas de reconhecimento de padrão baseados em sinais mioelétricos: adaptação automática do sistema a mudanças do sinal, atraso de sistema, movimentos mais naturais, realimentação para o usuário etc. (ANTFOLK *et al.*, 2013; CENE; FAVIEIRO; BALBINOT, 2015; FARRELL, 2011; LIU *et al.*, 2015; XIAORONG ZHANG; HE HUANG; QING YANG, 2013; ZHANG; HUANG, 2015).

Para a análise de sinais de sEMG, o uso de sistemas inteligentes é essencial para o aprendizado e adaptação em ambientes que sofrem constantes mudanças (BALBINOT; FAVIEIRO, 2013). Por essa razão, o uso de técnicas computacionais híbridas é vantajoso, o que é a essência da técnica *neuro-fuzzy*, que é composta de redes neurais que conseguem lidar com o reconhecimento de padrão, bem como se adaptar às mudanças, e a lógica *fuzzy*, que incorpora o conhecimento humano em uma solução de tomadas de decisão. Técnicas neuroadaptativas fornecem um método em que o processo de modelagem *fuzzy* pode ser realizado através de aprendizado de informações da base de dados. Por exemplo, a técnica conhecida como *Adaptive Neuro-Fuzzy Inference System* (ANFIS) constrói um sistema de inferência *fuzzy* em que seus parâmetros das funções de pertinência são ajustados utilizando um método híbrido.

A ANFIS emprega tipicamente o algoritmo de *back-propagation* para os parâmetros associados com as funções de pertinência de entradas e a estimação de mínimos quadrados para os parâmetros associados com as funções de pertinência das saídas. Por exemplo, o estudo conduzido por (FAVIEIRO; BALBINOT, 2011a) utilizou uma rede ANFIS para a classificação de 5 classes de movimentos do braço, utilizando 3 pares de eletrodos de superfície, tendo-se obtido uma taxa de classificação média de 93,3%.

Na pesquisa realizada por (CHAN *et al.*, 2000) foi desenvolvido um sistema *fuzzy*, cujas regras e funções de pertinência foram treinadas utilizando o algoritmo *back-propagation*, e a função de pertinência utilizada foi a gaussiana. O sistema visava a

classificação de 4 movimentos distintos: extensão e flexão do antebraço e extensão e flexão do punho. O sinal foi adquirido utilizando dois canais localizados nos músculos bíceps e tríceps e foram extraídas 5 características do sinal. A segmentação utilizada foi uma janela fixa de 240 ms a partir do início da contração muscular. O respectivo estudo obteve uma taxa média de acertos de 91% para 4 voluntários. Nesse estudo também foi realizada uma comparação com a técnica de Redes Neurais e se mostrou superior em termos de taxa de acerto.

Outro estudo que utilizou a técnica *fuzzy* foi desenvolvido por (AJIBOYE; WEIR, 2005) cujo sistema também almejava caracterizar 4 movimentos distintos: extensão do punho, flexão do punho, contração da mão e abdução do punho. Os músculos definidos foram: extensor dos dedos, extensor ulnar do carpo, flexor superficial dos dedos e flexor ulnar do carpo. A segmentação do sinal foi baseada no movimento completo e foi extraída uma única característica de cada janela: o valor RMS. O treinamento foi realizado *offline*, em que momentos intercalados do ensaio foram utilizadas para gerar a estrutura *fuzzy* inicial, aplicando a técnica de agrupamento de dados *fuzzy c-means* (FCM) e definindo a estrutura de regras de maneira adaptativa para o usuário em questão. No estudo foi realizada uma comparação do número de *clusters* ideal (menor taxa de erro) e chegou-se à conclusão de que a faixa ideal de *clusters* para o respectivo estudo foi na faixa de 50 a 100 *clusters*. Utilizando 50 *clusters* foi obtido um resultado de pelo menos 86% de acerto, considerando 2 voluntários.

Também foi realizado um estudo que utilizou a técnica *fuzzy c-means* para particionar o espaço de características do sinal mioelétrico (MOMEN; KRISHNAN; CHAU, 2007). Esse estudo, assim como (AJIBOYE; WEIR, 2005), extraiu somente o valor RMS do segmento do sinal, na qual foram utilizadas janelas fixas de 200 ms. Foram utilizados dois canais localizados nos músculos do extensor e flexor do antebraço. Os movimentos definidos foram selecionados pelos 7 voluntários cujos resultados apresentaram uma taxa de acertos média de $92,7\% \pm 3,2\%$ para 4 movimentos distintos. Um dos usuários realizou 9 movimentos distintos, obtendo uma Taxa de Acerto Média de 48,9%.

Uma técnica original foi proposta por (KHUSHABA; AL-ANI; AL-JUMAILY, 2010) chamada de *Orthogonal Fuzzy Neighborhood Discriminant Analysis* (OFNDA). Diferentemente de outras técnicas, o objetivo da OFNDA é minimizar a distância entre amostras pertencentes da mesma classe e maximizar a distância entre centros de classes distintas, enquanto leva em consideração a contribuição das amostras para as classes. O sistema foi testado em 7 voluntários, com 5 a 10 movimentos distintos, utilizando somente dois canais localizados no antebraço dos mesmos. A segmentação foi realizada utilizando

janelas de tamanho fixo de 187,5 ms, sendo extraídas 13 características do sinal por canal: combinação dos primeiros 6 coeficientes AR (*auto-regressive*), o valor RMS, o comprimento de onda, média do valor absoluto, a obliquidade dos dados segmentados, mudança do sinal de inclinação, número de cruzamentos por zero e a integral do valor absoluto. Os resultados variaram de 87% até 97% de Taxa de Acerto Média, de 10 até 5 classes distintas de movimentos, respectivamente.

A técnica LDA (*Linear Discriminant Analysis*) também é amplamente utilizada devido à sua simplicidade e custo computacional baixo. A técnica LDA é muito conhecida na redução de dimensionalidade em aplicações de aprendizagem de máquina. As características do LDA incluem a projeção da base de dados em um espaço dimensional reduzido que melhor representa a separabilidade das classes e, ao mesmo, tempo previne contra o *overfitting* do modelo. Adicionalmente, o processo de redução da dimensionalidade traz grandes benefícios em termos de custo computacional, uma vez que existe uma menor quantidade de dados para serem processados (LI; SCHULTZ; KUIKEN, 2010; RAURALE, 2014).

O estudo realizado por (LI; SCHULTZ; KUIKEN, 2010) utilizou um classificador LDA para analisar janelas fixas de 150 ms, em que são extraídas 4 características por canal: média do valor absoluto, número de cruzamentos por zero, comprimento de onda e número de mudanças do sinal de inclinação. Ao todo foram utilizados 12 canais de aquisição distribuídos no antebraço, pulso e mão, para a classificação de 10 movimentos distintos. Os movimentos determinados incluem a extensão e flexão do punho, a rotação do antebraço, a contração da mão e 5 movimentos diferentes de pegada. O treinamento utilizou as sessões ímpares de cada ensaio e recorreu-se às sessões pares para avaliar o sistema. A técnica proposta obteve uma taxa de acerto de $94\% \pm 3\%$. Também foi realizada uma análise utilizando 8 canais, o que gerou um decréscimo de 1% a 3% de taxa de acerto.

O estudo proposto por (RAURALE, 2014) utilizou somente dois eletrodos localizados nos músculos posteriores e anteriores do antebraço para a caracterização de seis movimentos distintos. A classificação do sinal é realizada através de LDA com uma taxa média de classificação correta dos movimentos de 80-86% no processamento *online*. As características extraídas do sinal foram o valor médio absoluto do sinal, variância, cruzamento por zero, comprimento do sinal, mudança do sinal positivo-negativo e a diferença entre o valor médio absoluto entre amostras. Nesse estudo também foi avaliada a classificação do sinal em etapas diferentes da aquisição e tratamento do sinal de sEMG, em que a taxa de classificação foi superior ao se utilizar o sinal bruto (com valor de 86,6%) sendo que a menor taxa de acerto

desse estudo de caso ocorreu ao se utilizar a média absoluta do sinal, com 80,0% de taxa de acerto.

As técnicas descritas anteriormente também podem aplicar paradigmas de adaptação supervisionados, conforme a pesquisa realizada por (SENSINGER; LOCK; KUIKEN, 2009). Nesse estudo foi proposta uma comparação utilizando a técnica LDA com ou sem treinamento supervisionado. Para o estudo de caso foi concluído que o uso de treinamento supervisionado reduziu o erro do sistema em pelo menos 26% ao longo do tempo. A adaptação supervisionada consiste em retreinar o sistema quando se tem conhecimento da saída esperada do mesmo. Sendo assim, o erro do sistema pode ser definido como a diferença entre o valor predito pelo modelo e o valor desejado (o alvo). Para realizar essa comparação foram extraídas 4 características do sinal: média do valor absoluto, número de cruzamentos por zero, comprimento de onda e número de mudanças no sinal de inclinação. O sinal foi segmentado em tamanhos fixos de 150 ms, utilizando 12 canais de aquisição. Os movimentos classificados envolviam a extensão e flexão do punho, a rotação do antebraço, a contração da mão e 5 movimentos de pegada. O erro médio do sistema foi de $30\% \pm 16\%$ para a técnica LDA sem treinamento supervisionado e de $20\% \pm 8\%$ ao utilizar o treinamento supervisionado. Pode-se perceber uma diferença na taxa de acerto do sistema ao longo do tempo na classificação dos movimentos.

A técnica *Support Vector Machines* (SVM) consiste em um algoritmo de treinamento supervisionado e tem se mostrado muito popular entre tarefas de aprendizagem de máquina que envolvem classificação e regressão (OSKOEI; HU, 2008). Dado um conjunto de exemplos de treinamento, cada um pertencendo a uma de duas categorias distintas, o algoritmo SVM constrói um modelo de classificação linear binária não probabilística. O SVM consiste basicamente em um hiperplano de separação ótimo em um espaço de características de treinamento de alta dimensionalidade, que são mapeadas usando uma função *kernel* (que pode ser linear ou não). O uso de função *kernel* não linear aumenta de forma significativa o poder de aprendizado e generalização do algoritmo. No entanto, também pode aumentar o risco de *overfitting*, o que leva a uma baixa generalização do modelo.

O estudo proposto por (GUO *et al.*, 2015) tem o objetivo de comparar dois classificadores (SVM e RNA) e 4 características dos sinais: RMS, *Peak Weight*, MM (modelo muscular) e DFA (*Detrended Fluctuation Analysis*). Para esse estudo foram selecionados 8 movimentos e para o classificador RNA foi obtida a Taxa de Acerto Média de 88,7% na etapa de treinamento, enquanto o SVM teve um desempenho melhor no experimento online, com uma Taxa de Acerto Média de 85,9%. Em termos de atraso do sistema, o algoritmo SVM foi

mais eficiente na etapa de treinamento. Entretanto, nos testes de processamento online a RNA obteve um atraso menor nos processamentos dos dados. Em relação às características, a utilização do MM teve melhor taxa de acertos para os dois classificadores: RNA com 92,01% e SVM com 94,3%, nos testes online.

Outro estudo muito interessante é de (ATZORI *et al.*, 2015) que faz uma comparação de 4 algoritmos distintos comuns na área (LDA, SVM, *Random Forests* e *k-Nearest Neighbors*) para a classificação de 50 movimentos utilizando 4 características do sinal (RMS, estatísticas no domínio do tempo (TD) descrita por (HUDGINS; PARKER; SCOTT, 1993), Histograma e Transformada *Wavelet* Discreta). A base de dados utilizada foi a Ninapro de livre acesso (ATZORI *et al.*, 2012). Para os voluntários amputados foi obtida a melhor classificação média de 46,27% de taxa de acerto utilizando o método SVM. Já para os voluntários com membro superior intacto a melhor classificação foi obtida utilizando o *Random Forest* com taxa média de acerto de 75,27%.

Na pesquisa de (MOURA; FAVIEIRO; BALBINOT, 2016) foi apresentado um estudo investigando o uso de sinais mioelétricos do antebraço para classificação de 17 movimentos específicos da mão usando 12 canais de sEMG e o classificador SVM. O estágio de classificação utilizou três tipos de funções *kernel* (linear, polinomial, base radial) para comparação dos resultados. O estudo conclui que o uso unicamente da característica RMS como entrada para os classificadores SVM já obtém mais de 80% de taxa de classificação correta em todos as funções *kernel* analisadas. A taxa de acerto média alcançada usando o classificador polinomial SVM foi de 83,7%, e para o classificador linear SVM foi de 80,8%. O classificador SVM com base radial apresentou uma taxa de acerto mais estável na classificação para diferentes voluntários e movimentos, obtendo a melhor média de taxa de acerto de 85,1% e menor desvio padrão.

O estudo proposto por (MATSUBARA; MORIMOTO, 2013) tinha por objetivo uma metodologia inovadora em que a interface mioelétrica pudesse facilmente se adaptar a múltiplos usuários, utilizando uma abordagem de separação de estilo e conteúdo. Nesse estudo eles propuseram encarar o sinal de sEMG como um modelo bi-linear, que é basicamente composto de dois fatores lineares: (1) o fator dependente do usuário (chamado de estilo) e (2) o fator dependente do movimento (chamado de conteúdo). Ao se decompor o sinal nesses dois fatores, os fatores movimento-dependentes podem ser utilizados como características independentes do usuário. Para cada novo usuário da prótese é proposto um método de adaptação que extrai as características usuário-independentes através de algumas poucas interações com os dados do usuário novo. Para os testes, foram classificados 5 movimentos

distintos e para movimento novo do usuário só é necessária uma única observação do movimento para identificar os fatores usuário-dependente. A metodologia deles inclui 4 eletrodos de superfície localizados em músculos específicos: *flexor carpi ulnaris*, *extensor carpi radialis*, *flexor digitorum profundus* e *extensor digitorum*. Os movimentos da mão foram realizados sequencialmente (flexão da mão, extensão da mão, fechamento da mão e abertura da mão) e 11 voluntários do sexo masculino participaram dos ensaios. Também é importante ressaltar que nesse estudo foi utilizada a técnica de segmentação de janelas deslizantes, com o tamanho da janela de 128 amostras e deslocamento de 25 amostras. Para cada janela foi extraída a característica RMS do sinal para cada canal. Com isso foram obtidas 80 amostras (características) para cada movimento. Utilizando a técnica SVM foi obtida uma taxa de acerto de 54%, e utilizando o método proposto resultou em uma Taxa de Acerto Média de 73% para todos os usuários e movimentos. A técnica, apesar de ser um estudo preliminar, se mostrou promissora para o uso multiusuário e, provavelmente, este tipo de abordagem será fonte de novas pesquisas nos próximos anos.

3.3 SISTEMAS DE CLASSIFICAÇÃO ROBUSTOS A CONTAMINANTES NO SINAL DE sEMG

Ao longo dos anos, diversos métodos computacionais foram propostos com o objetivo de obter um algoritmo robusto a mudanças no sinal. No entanto, definir um sistema de classificação visando à adaptabilidade do mesmo a mudanças no sinal de sEMG ainda representa um desafio porque os potenciais de ação adquiridos através de eletrodos de superfície são sensíveis ao reposicionamento dos eletrodos, mudanças na impedância eletrodo-pele, movimento relativo do músculo em relação ao eletrodo, entre diversos outros fatores. Adicionalmente, as distribuições das formas dos potenciais de ação variam de forma significativa entre sujeitos, por causa de diferenças anatômicas (DE LUCA, 1979). Esses fatores influenciam substancialmente o desempenho de algoritmos de controle mioelétricos. Por exemplo, é de conhecimento geral que mudanças no posicionamento do eletrodo têm um efeito negativo nas taxas de acerto dos métodos de controle mioelétricos (CENE; FAVIEIRO; BALBINOT, 2015; MUCELI; JIANG; FARINA, 2014).

Adicionalmente, fontes de ruído externas e artefatos também podem influenciar os sinais de sEMG, ocasionando frequentes erros de classificação dos movimentos. A maior parte dos contaminantes que degradam o sinal de sEMG consistem em ruído térmico, artefatos de movimentos, ruído de linha, perda do contato eletrodo-pele, deslocamento do eletrodo e saturação do amplificador. Existem alguns trabalhos como (FRASER *et al.*, 2014; MCCOOL

et al., 2014; ZHANG; HUANG, 2015) em que o objetivo é identificar e classificar as interferências destrutivas apresentadas nos sinais de sEMG. O processo de aquisição do sinal de sEMG é de fato uma etapa crítica que pode afetar diretamente as etapas posteriores do processamento, levando a um decréscimo na taxa de classificação correta do sistema. Nesse sentido, existem estudos com uma abordagem prévia ao processamento dos dados, visando elevar a taxa de acerto do classificador ao realizar sessões de treinamento considerando ambientes de realidade virtuais (WOODWARD; HARGROVE, 2019). O uso de ambientes virtuais, considerando o uso clínico de próteses, facilita o retreino do modelo de classificação pelos próprios usuários, além de promover uma sessão de treinamento apreciando movimentos ativos que, segundo os resultados do estudo, conseguem elevar a taxa de acerto média do sistema em ensaios de longa duração.

O trabalho proposto por (CENE; FAVIEIRO; BALBINOT, 2015) apresentou o desenvolvimento de um sistema com Redes Neurais Auto Adaptativas em que o sistema é capaz de realizar um retreinamento para acompanhar mudanças nos sinais de entrada. A proposta é realizar um constante monitoramento dos sinais de sEMG para identificar continuamente os melhores sinais de sEMG para serem utilizados como entrada da Rede Neural. O objetivo principal deste trabalho foi identificar não idealidades do sinal (eletrodos soltos ou com baixo SNR) e descartar canais que são considerados com baixa significância para o sistema. Para avaliar os resultados, o algoritmo desenvolvido foi comparado a uma Rede Neural com as mesmas características, exceto pelo fato de que ela não descarta nenhum canal. De acordo com os resultados obtidos na análise de sinais de 5 voluntários distintos, houve um acréscimo na taxa de classificação de 5,0 a 27,6% na utilização do sistema adaptativo. Um ponto que deve ser ressaltado é que mesmo com um acréscimo de 5%, o processamento de um número menor de dados pode gerar um ganho em termos de tempo de processamento do sistema na classificação dos dados.

No estudo de (ZHANG; HUANG, 2015) foi apresentado um algoritmo LDA-adaptativo capaz de mitigar a taxa de erro do classificador considerando a presença de contaminantes nos canais de sEMG. Similarmente ao trabalho de (CENE; FAVIEIRO; BALBINOT, 2015), a pesquisa de (ZHANG; HUANG, 2015) propôs uma solução que envolve o monitoramento dos canais de sEMG em busca de degradações e a atuação removendo os sinais desses canais contaminados do modelo de classificação ao realizar um novo treinamento do mesmo. O monitoramento dos sinais de sEMG é realizado através de um sensor de detecção de falhas que monitora as características de cada canal de sEMG individualmente com o objetivo de detectar contaminantes. Se nenhum contaminante é

detectado, o sensor entrega o vetor de características dos sinais de sEMG para o modelo LDA para a classificação do movimento. Caso o sensor detecte degradações nas características do sinal de sEMG, através da análise de um limiar baseado na distância de Mahalanobis, as características do canal de sEMG contaminado são removidas do vetor de características disponibilizado ao modelo LDA, realizando rápido retreino do mesmo ao reutilizar alguns parâmetros do modelo LDA gerado *a priori*. O retreino do classificador, aproveitando alguns parâmetros da LDA, ao invés de realizar um novo treinamento do zero, trouxe um ganho significativo no desempenho do sistema.

A metodologia de testes, do método LDA-adaptativo proposto por (ZHANG; HUANG, 2015), envolveu a classificação de 7 classes de movimento (repouso, extensão e flexão de punho, supinação e pronação de punho, fechamento e abertura da mão), considerando 6 canais de sEMG sob músculos específicos (flexor ulnar do carpo, flexor radial do carpo, extensor ulnar do carpo, extensor radial do carpo, extensor longo do polegar e palmar longo). A segmentação do sinal envolveu janelas deslizantes de 160 ms com incremento de 20 ms e a extração de 4 características dos sinais no domínio do tempo (valor médio absoluto, número de cruzamentos por zero, comprimento de onda, número de mudanças de sinal). Adicionalmente, os sinais de sEMG foram adquiridos de 5 sujeitos com membro superior intacto. Nos resultados foi observado que para 1 canal contaminado, o decréscimo na taxa de acerto média foi de aproximadamente 50% para o contaminante artefato de movimento, 40% para o contaminante perda de contato eletrodo-pele e 25% para o contaminante ruído de base (SNR de 20 dB) e a classificação com LDA tradicional. Utilizando a abordagem LDA-adaptativa, o decréscimo na taxa de acerto foi de aproximadamente 3% para todos os contaminantes avaliados. Considerando 3 canais contaminados, o decréscimo na taxa de acerto variou na faixa de 20% para o método adaptativo, dependendo do contaminante, contra decréscimos variando entre 30 a 90% sem o retreino do modelo. Nesse estudo foi possível observar os efeitos negativos que a presença de contaminantes acarreta na taxa de acerto do classificador e, também, que o retreino do modelo se apresenta como uma solução interessante para mitigar esses efeitos. No entanto, o constante retreino do modelo acarreta em um custo computacional e de memória que deve ser levado em consideração na aplicação dessa solução em ensaios clínicos.

O estudo proposto por (DE MOURA; BALBINOT, 2018) aplica o conceito de sensor virtual na área de eletromiografia de superfície com o objetivo de minimizar problemas na classificação de movimentos tipicamente relacionados à degradação do sinal mioelétrico. Essa degradação do sinal geralmente ocasiona uma redução na taxa de classificação dos

movimentos caracterizados por sistemas inteligentes. Esse estudo envolvendo sensores virtuais avaliou os cinco contaminantes mais comuns em sinais de sEMG: artefatos de movimento, saturação do sinal, deslocamentos de eletrodos, interferência de linha e interferência do sinal de ECG. Primeiramente, o estudo propõe um sensor tolerante a falhas, baseado em um modelo SVM, para identificar a presença dos contaminantes avaliados nesse estudo. Também é proposto um sensor virtual foi desenvolvido através de duas abordagens distintas: *Time-Varying Autoregressive Moving Average (TVARMA)* e *Time-Varying Kalman filter (TVK)* que modelam artificialmente os canais de sEMG considerados contaminados, baseado em uma matriz de correlação em relação aos outros canais de aquisição, que não apresentam contaminantes. Os resultados da classificação do movimento foram apresentados considerando o uso do classificador SVM sem nenhuma atuação na presença de contaminantes, a substituição do sinal degradado pelo sensor virtual proposto baseado em modelos TVARMA e TVK e o retreinamento do classificador considerando os modelos do sensor virtual. Dessa forma, foi possível analisar que o sistema proposto por (DE MOURA; BALBINOT, 2018), sem o uso de técnicas de retreinamento de classificadores, apresentou um acréscimo na taxa da classificação média de 4% a 38% em relação ao uso de sinais degradados por deslocamento de eletrodos, artefatos de movimento e saturação do sinal. Considerando o retreino do classificador, utilizando os sinais de sEMG modelados por TVARMA e TVK, o erro máximo apresentado foi de 15% considerando 8 canais contaminados, sendo significativamente superior a todas as outras soluções analisadas. No entanto, apesar da proposta de (DE MOURA; BALBINOT, 2018) apresentar resultados significativamente superiores em relação ao uso exclusivo do classificador SVM, a mesma exibe algumas limitações. Uma delas é a restrição que o sistema proporciona em somente atuar no modelo na presença de contaminantes conhecidos previamente. Outra limitação é observada na criação dos modelos TVARMA e TVK de sinais de sEMG que, por serem efetuadas sob demanda, apresentam um custo computacional que inviabiliza o uso do sistema em processamento *online*.

O estudo de (VIDOVIC *et al.*, 2016) também confirmou que existe um decréscimo significativo na taxa de classificação de movimentos devido aos problemas que ocorrem durante a aquisição do sinal fora do ambiente controlado de laboratório, que dificultam a classificação de sinais de sEMG para controle mioelétrico de próteses. Nesse estudo foi proposto o método de adaptação supervisionado para superar essa limitação. A abordagem foi baseada na adaptação de um classificador treinado usando apenas um pequeno conjunto de calibração, que incorpora os aspectos relevantes e requer menos de 1 minuto de gravação de

dados. Na análise de dados *offline*, a taxa de classificação permaneceu acima de 92% em cinco dias, com a abordagem proposta, enquanto diminuiu para 75% sem adaptação.

A técnica proposta por (PAN *et al.*, 2015) visa suavizar os efeitos do contaminante deslocamento de eletrodos ao considerar uma abordagem de eletromiografia de alta resolução utilizando o método *common spatial patterns* (CSP) para a extração de características do sinal de sEMG. Segundo os autores, a técnica CSP é amplamente utilizada em interfaces cérebro-maquina na classificação de sinais de eletroencefalografia (EEG). A metodologia de testes envolveu o uso de 192 eletrodos ao redor da junta rádio-humeral para a classificação de 11 classes de movimentos do segmento mão-braço, considerando 9 sujeitos com membro superior intacto. Além disso, para avaliar o aumento da robustez do método, em relação a deslocamentos de eletrodos na predição do classificador, foram considerados para os testes os deslocamentos de eletrodos na direção longitudinal e transversal em relação as fibras musculares. Adicionalmente, esse método foi comparado com os conjuntos de características amplamente utilizados na literatura: *time-domain* (TD), *time-domain autoregressive* (TDAR) e *variogram* (Variog), utilizando o classificador LDA e segmentação dos sinais de sEMG por janelas deslizantes de tamanho 200 ms com incremento de 50 ms. Analisando os resultados, foi possível observar que, considerando todas as configurações de deslocamento de eletrodo, as características CSP atingiram um acréscimo na taxa de acerto superior a 10% em relação as características TD e superior a 5% em relação as características TDAR. Em relação as características Verilog, só foi observado um acréscimo significantes considerando os deslocamentos longitudinais. A técnica CSP se mostra promissora para uso em classificadores visando mitigar os efeitos dos deslocamentos de eletrodo na taxa de acerto do método, no entanto a pesquisa carece de um estudo mais aprofundado com o objetivo de avaliar o comportamento das características extraídas, pela técnica CSP, considerando outros tipos de contaminantes que comumente afetam o sinais de sEMG (SAMUEL, 2019).

A proposta de (GENG *et al.*, 2018) também utiliza eletromiografia de alta resolução, no entanto o objetivo do estudo é apresentar um método robusto a contaminação do sinal por ruído branco gaussiano. O método de classificação utilizado é o *Sparse Representation-based Classification* (SRC). Segundo os autores, a contaminação do sinal de sEMG por ruído branco é inevitável devido a interferência eletromagnéticas, ruído interno do amplificador operacional e ruído térmico presente nos elementos eletrônicos. Além disso, a variação da interferência eletromagnética pode levar a mudanças nos níveis de ruído, dificultando a detecção dos mesmos em sistemas que utilizam sinais de sEMG no dia a dia. Por isso, um método que possibilite a redução dos efeitos desse contaminante na predição do classificador

é conveniente em aplicações baseadas em sinais de sEMG. Para validar a hipótese que o método SRC é robusto a ruído branco, a metodologia de testes considerou sinais de sEMG de 12 voluntários com lesões cerebrais traumáticas e 5 voluntários sobreviventes de derrame cerebral. A aquisição dos sinais de sEMG abrangeu o uso de 56 canais, com quantidade de classes de movimento distintas entre os voluntários, de acordo com o nível de mobilidade dos mesmos. Também foram utilizados, para fins comparativos, os classificadores LDA, KNN, RF, e SVM. Em relação aos testes envolvendo a adição de ruído branco, foram considerados os SNR de 5 dB, 10 dB, 15 dB, 20 dB, 25 dB e 30 dB. É interessante ressaltar que cada classificador utilizou conjuntos distintos de características extraídas do sinal de sEMG, visando otimizar as taxas de acerto dos diferentes classificadores. Os resultados demonstraram que o método SRC apresentou taxas de acerto média similares aos métodos SVM, KNN, LDA e RF na presença de sinais contaminados com ruído branco gaussiano considerando a variação de SNR na faixa de 0 a 15 dB. No entanto, considerando o ruído branco com SNR entre 20 a 30 dB o método SRC se mostrou significativamente superior aos outros métodos comparados. Considerando a contaminação do ruído branco somente nos conjuntos de treino ou teste, o método SRC foi superior considerando todos os níveis de SNR considerados. O uso do método SRC se mostra promissor na classificação de movimentos de membro superior, considerando a presença de diferentes níveis de ruído branco gaussiano, no entanto seria interessante avaliar o desempenho do método SRC em relação a outros contaminantes que comumente afetam os sinais de sEMG.

A pesquisa proposta por (BETTHAUSER *et al.*, 2018) afirma que sinais mioelétricos podem ser usados para prever os movimentos pretendidos de um amputado para controle de prótese. No entanto, os efeitos não treinados, como alterações na posição dos membros, influenciam as características dos sinais mioelétricos, dificultando a capacidade dos algoritmos de reconhecimento de padrão em discriminar classes de movimento. Apesar das frequentes e longas sessões de treinamento, essas influências condicionais podem resultar em baixa taxa de acerto na classificação e abandono do dispositivo. No estudo é apresentado como solução um método robusto (baixa susceptibilidade a mudanças no sinal de entrada) de classificação adaptativa baseada na dispersão, que é significativamente menos sensível a desvios de sinal resultantes de condições não treinadas. Nesse estudo foram relatadas melhorias significativas de desempenho ($p < 0,001$) em posições de membros não treinados em todos os grupos de sujeitos. Essa abordagem garantiu melhor desempenho das classes não treinadas a partir de treinamento simplificado, trazendo possíveis benefícios clínicos aos

amputados: melhor desempenho com uma menor carga de treinamento em termos de frequência e duração.

3.4 CARACTERÍSTICAS DE UM SISTEMA DE CONTROLE MIOELÉTRICO IDEAL

De acordo com (FARINA *et al.*, 2014), idealmente, um sistema de controle de próteses do membro superior deveria cumprir os seguintes critérios:

- O controle deve ser intuitivo para o usuário e para a geração de movimentos naturais, ou seja, a prótese deve gerar movimentos simultâneos e proporcionais aos graus de liberdade desejados, potencialmente controlando múltiplas articulações;
- O sistema deve ser robusto em relação à mudança na impedância eletrodo/pele, atividades diversas, posição do membro, mudanças no sinal de sEMG em diferentes dias de uso, etc;
- Deve ser um sistema adaptativo para o usuário, para que mudanças lentas no sinal, como fadiga e suor, sejam compensadas automaticamente;
- O sistema deve utilizar um número mínimo de eletrodos e não deve ser sensível a uma posição/orientação precisa dos eletrodos;
- O sistema deve utilizar um treinamento/calibração rápida e fácil (idealmente sem treinamento), sem a necessidade de retreinamentos frequentes;
- A prótese deve providenciar realimentação em funções relevantes, como por exemplo força (controle de malha fechada);
- O sistema deve ter uma complexidade computacional baixa, de forma que seja possível a implementação em uma prótese utilizando *hardware* de baixo consumo de energia;
- E, acima de tudo, deve executar os movimentos intencionados em tempo real (com atrasos de no máximo 120 ms) (FARRELL, 2011).

Nenhum dos estudos descritos neste capítulo consegue satisfazer todos os critérios para o sistema de controle mioelétrico ideal definidos por (FARINA *et al.*, 2014). Entretanto, as pesquisas mais recentes têm mostrado um empenho para cumprir alguns desses critérios, como, por exemplo, os estudos focados na qualidade do sinal ou próteses multiusuários, previamente citados. Também existem pesquisas como (ANTFOLK *et al.*, 2013), com resultados promissores em relação à realimentação para o usuário da prótese.

Adicionalmente, inúmeras pesquisas descritas neste capítulo confirmam que existe um decréscimo significativo na taxa de classificação de movimentos devido aos problemas que ocorrem durante a aquisição do sinal em condições exteriores ao ambiente controlado de laboratório, que dificultam a tradução direta de sinais de sEMG em um controle mioelétrico fino de próteses. Essas degradações do sinal influenciam as características dos sinais mioelétricos, dificultando a capacidade dos algoritmos de reconhecimento de padrão em discriminar classes de movimento. Portanto, de acordo com as linhas de pesquisas apresentadas na atualidade na área de reconhecimento de movimentos baseados em sinais mioelétricos, pode-se perceber que o objetivo das mesmas não está mais somente focado em elaborar um classificador com alta taxa de acertos em laboratório, mas, essencialmente, desenvolver um sistema de classificação que consiga se adaptar a degradações no sinal de entrada, buscando ampliar a sua aplicabilidade prática, visto que sistemas que utilizam sinais físicos muitas vezes são afetados por fatores externos à aquisição dos mesmos.

Nesse sentido, a utilização de lógicas não clássicas, como a Lógica Paraconsistente, pode agregar robustez ao algoritmo de decisão, considerando-se que o sEMG é um sinal complexo, do qual é difícil extrair todas as informações necessárias para caracterizar um único movimento. Além disso, os movimentos do membro superior possuem diversos graus de liberdade que incluem o uso de inúmeras unidades motoras para a sua efetividade. Logo, considerada a natureza estocástica do sinal de sEMG, a quantidade muitas vezes limitada de eletrodos, e os fatores adicionais, como artefatos de movimento e perda de contato eletrodo-pele, é possível considerar que podem existir contradições nos dados em relação ao modelo de classificação gerado *a priori*. Dessa forma, é interessante analisar o emprego de um algoritmo que possa levar em consideração tais elementos em sua estrutura teórica, com o intuito de minimizar o prejuízo causado por esses fatores na taxa de acerto do sistema sem a necessidade de treinamentos contínuos para evitar a degradação do modelo de classificação, uma vez que, considerando o uso clínico de próteses mioelétricas, sessões de treinamento longas e frequentes podem resultar em baixo desempenho e abandono do dispositivo.

4 METODOLOGIA

Este capítulo versa sobre uma análise preliminar da Lógica Paraconsistente através de uma Rede de Análise Paraconsistente dedicada a classificação de sinais de sEMG, bem como a concepção e o desenvolvimento de um método inovador de reconhecimento de padrão baseado em Lógica Paraconsistente e *Random Forest*. Também são abordadas a metodologia para o teste e a avaliação dos métodos aplicados na área de reconhecimento de movimentos do membro superior utilizando sinais mioelétricos, com a descrição das bases de dados utilizadas.

4.1 ANÁLISE PRELIMINAR DA LÓGICA PARACONSISTENTE NA CLASSIFICAÇÃO DE SINAIS MIOELÉTRICOS

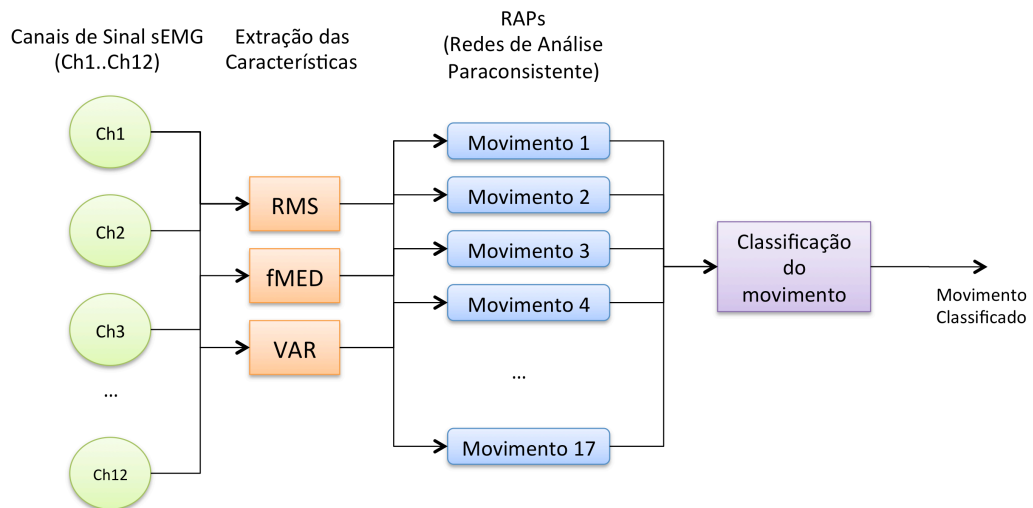
No primeiro método proposto nesta tese, os conceitos da Lógica Paraconsistente foram usados para o desenvolvimento de uma Rede de Análise Paraconsistente (RAP) específica para a classificação de movimentos do membro superior baseados em sinais mioelétricos, de tal forma que o reconhecimento de padrão foi considerado como tendo um grau intrínseco de certeza que leva em conta as contradições e incertezas do sistema. Esse método é considerado uma análise preliminar que visava avaliar se a Lógica Paraconsistente se mostrava adequada para a classificação de movimentos através de sinais mioelétricos. Adicionalmente, foi objetivado identificar vantagens e limitações da Lógica Paraconsistente, de forma a propor um método inovador que vinculasse conceitos da LP com aprendizagem de máquina para a concepção e desenvolvimento de um método de ampla aplicabilidade e com baixa susceptibilidade a degradações nos sinais de entrada.

Considerando essa premissa, para a classificação de 17 movimentos distintos foi desenvolvido um sistema de Análise Paraconsistente composto por 17 RAPs, uma para cada Movimento Padrão (MP) executado pelo segmento mão-braço definido na metodologia, conforme indicado no diagrama de blocos da Figura 10. O diagrama de blocos representa os 12 canais de aquisição do sinal de sEMG, a extração das características do sinal, as 17 RAPs, representando as 17 classes de movimento selecionadas e um bloco de tomada de decisão da classe vencedora (o movimento determinado pelo modelo).

No diagrama da Figura 11 é ilustrado com mais detalhes o processo de reconhecimento de padrão utilizando a RAP proposta. O diagrama apresenta as quatro camadas primárias do algoritmo de reconhecimento de padrão: a aquisição do sinal de sEMG, a extração e transformação das características do sinal em Graus de Evidência, o reconhecimento de

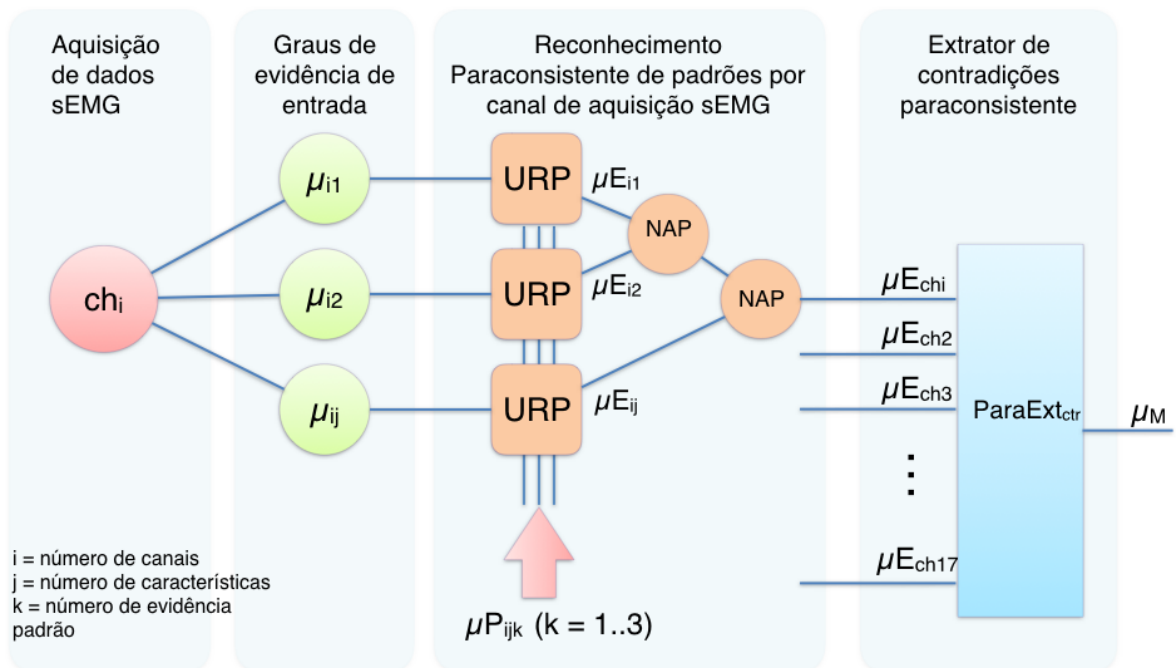
padrão por canal de aquisição, empregando Unidades de Reconhecimento de Padrão (URP), e, por último, o cálculo do Grau de Evidência Resultante por classe de Movimento Padrão, através do método Extrator de Contradições Paraconsistente. O Grau de Evidência por Movimento Padrão, resultante de cada RAP, é então utilizado para a classificação do movimento analisado.

Figura 10 - Diagrama de Blocos do sistema de análise paraconsistente proposto.



Fonte: Autor

Figura 11 - Diagrama de blocos representando a RAP para 3 características de entrada.



Fonte: Autor

A RAP é formada por vários nós conectados, denominados Nós de Análise Paraconsistente (NAP), descritos previamente no Capítulo 2, em uma topologia projetada explicitamente para a aplicação de classificação de movimentos em função dos sinais mioelétricos, sendo que para cada classe de movimento é definida uma RAP específica.

No entanto, antes de começar a análise dos dados, é necessário inicializar cada RAP com as características padrão da classe de movimento desejada (base de dados de conhecimento). As características padrão são utilizadas na terceira camada, com o objetivo de comparar as entradas com características padrão armazenados em cada RAP. A comparação fornece um vetor de Graus de Evidência que representa o quão próximos ou o quão distantes as características de entrada estão das características padrão, para cada classe de movimento.

Uma vez que todas as 17 RAPs têm a mesma arquitetura, a única diferença entre as mesmas é o valor de inicialização, que pode ser considerado similar ao "treinamento supervisionado" dos algoritmos de aprendizagem de máquinas. Cada RAP tem as seguintes características:

1. Para cada canal de análise do sinal de sEMG são necessárias n URPs (uma para cada tipo de característica extraída);
2. Para cada tipo de característica extraída são necessárias 3 amostras da característica (por movimento) para ser utilizada como padrão de um movimento x . Essas características padrão são utilizadas para inicializar a RAP, adaptando-a para o Movimento Padrão desejado;
3. Cada característica extraída do sinal é transformada de forma a poder ser utilizada como um Grau de Evidência na entrada do sistema;
4. As saídas das URPs são analisadas por um NAP (caso tenha mais de uma URP por canal) para calcular um único Grau de Evidência Resultante do canal;
5. Ao final da Análise Paraconsistente por canal é gerado um único Grau de Evidência Resultante do canal;
6. A última etapa da RAP é gerar um Grau de Evidência Resultante do movimento, na qual é utilizado um algoritmo chamado Algoritmo Paraconsistente Extrator de Efeitos da Contradição, que tem como entrada uma lista composta dos Graus de Evidência Resultante de cada canal e tem como saída o Grau de Evidência Resultante do movimento.

Após a Análise Paraconsistente do sinal de entrada em cada uma das 17 RAPs, esse sinal é classificado em um dos Movimentos Padrão.

4.1.1 Transformação das Características em Graus de Evidência

Após a etapa de extração das características é necessário adaptar os valores para que possam ser utilizados como entrada do Sistema de Análise Paraconsistente. Como mencionando anteriormente, os Graus de Evidência utilizados como entrada em Sistemas Paraconsistentes devem se encontrar no intervalo $[0;1]$. Para isso foi realizada uma modelagem preliminar das características com uma variação linear diretamente proporcional à grandeza medida. O Grau de Evidência de Entrada μ_i foi determinado pela Equação (14):

$$\mu_i(x) = \begin{cases} \frac{x-a_1}{a_2-a_1}, & x \in [a_1, a_2] \\ 0, & x > a_2 \\ 1, & x < a_1 \end{cases}, \quad (14)$$

onde x representa o valor da característica em questão e os valores de a_1 e a_2 são calculados individualmente para cada característica, para cada canal e para cada voluntário, representando, respectivamente, os valores mínimos e máximos calculados (baseado no conjunto de características).

4.1.2 Procedimento de Inicialização da RAP

Para cada RAP, foi adquirido um banco de dados de conhecimento contendo amostras de movimento de cada classe de movimento específica. A transformação das características em Graus de Evidência, através de um modelo linear, também é utilizada para condicionar as amostras de características padrão em Graus de Evidência Padrão. Os Graus de Evidência Padrão ($\mu_{P_{ijk}}$) representam, respectivamente, o canal i , o atributo j e a amostra do movimento k . O banco de dados de conhecimento é utilizado durante o primeiro estágio da Análise Paraconsistente, a fim de fornecer um grau padrão de evidência do movimento a ser analisado com o Grau de Evidência de entrada na Camada de Análise Paraconsistente.

Para este estudo foi determinado que o tamanho k do banco de dados de conhecimento é igual a 3 amostras de cada classe de movimento, uma limitação imposta pela quantidade de amostras de movimento do banco de dados de sEMG (total de 6 amostras por classe de movimento). É importante ressaltar que essas características são extraídas por canal. Então, por exemplo, a RAP da Figura 11 contém 3 características extraídas: 3 características por canal vezes 3 amostras de cada característica vezes 12 canais, gerando um total de 108 Graus

de Evidência Padrão. Deste modo, no exemplo, cada RAP será inicializada com 108 Graus de Evidência Padrão. É importante observar que para cada RAP essas Evidências Padrão são representativas de cada Movimento Padrão específico.

Outro fato importante a observar é que cada Nó Paraconsistente tem o mesmo peso na topologia da rede. Portanto, ter uma única amostra de movimento no banco de dados de conhecimento não é indicado, porque, na área de reconhecimento de movimento utilizando os sinais de sEMG, os sujeitos geralmente não executam os movimentos da mesma maneira. Por esse motivo, é essencial ter no mínimo duas amostras de cada movimento. Além disso, no presente trabalho as amostras de movimento foram escolhidas aleatoriamente no banco de dados de sinais de sEMG para compor o banco de dados de conhecimento. É importante ressaltar que esta aleatorização não apresenta garantias de que os movimentos escolhidos sejam representativos da classe de movimento desejada.

4.1.3 Reconhecimento de Padrão

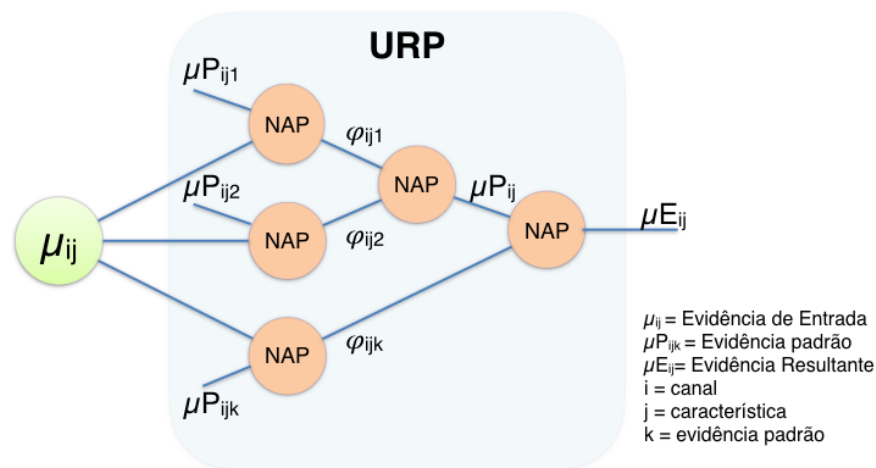
Considerando-se a área de processamento de sinais de sEMG, concorda-se que as características extraídas do sinal de sEMG têm uma influência significativa no reconhecimento das classes de movimento (GUO *et al.*, 2015; HUDGINS; PARKER; SCOTT, 1993; PHINYOMARK *et al.*, 2013; TOSIN *et al.*, 2017). Quanto mais discriminadoras forem as características do sinal entre diferentes movimentos, mais fácil será criar um modelo que possa classificá-los. Considerando essa abordagem, a solução proposta inicialmente realiza uma análise separada de cada característica escolhida por canal, na qual se pode averiguar o impacto da mesma na base de dados de conhecimento.

No processo de reconhecimento, os Graus de Evidência obtidos da extração de características são comparados aos Graus de Evidência Padrão por uma Unidade de Reconhecimento de Padrão. Uma Unidade de Reconhecimento de Padrão (URP) é definida com o propósito de calcular a distância entre três Graus de Evidência Padrão ($\mu_{P_{ij1}} \dots \mu_{P_{ij3}}$), e o Grau de Evidência de Entrada apresentado (μ_{ij}), criando um Grau de Evidência Resultante para cada característica de entrada ($\mu_{E_{ij}}$), com base no Intervalo de Evidência (ϕ_{ijk}) calculado pelo NAP, conforme apresentado na Figura 12. Esse processo é repetido para cada atributo de entrada de cada canal das 17 classes de movimentos.

Cada Unidade de Reconhecimento de Padrão (URP) tem por finalidade calcular um Grau de Evidência Resultante que representa o quanto as características do movimento analisado (entrada) estão próximas das características padrão daquela URP específica. Para

isso foi definido anteriormente que para cada URP são necessários 3 Graus de Evidência Padrão e, como entrada, a URP recebe o Grau de Evidência do movimento analisado. Conforme o diagrama de blocos da Figura 12, a URP é composta na camada inicial de 3 NAP que tem a finalidade de verificar o quão próximo o Grau de Evidência de Entrada está dos Graus de Evidência Padrão armazenados. Para isso cada NAP da primeira camada recebe como entrada o Grau de Evidência de Entrada e um Grau de Evidência Padrão. Sendo assim, a partir da Análise Paraconsistente é possível utilizar a saída Intervalo de Certeza (φ) que representará essa distância entre as entradas.

Figura 12 - Diagrama de blocos da Unidade de Reconhecimento de Padrão (URP) proposta.



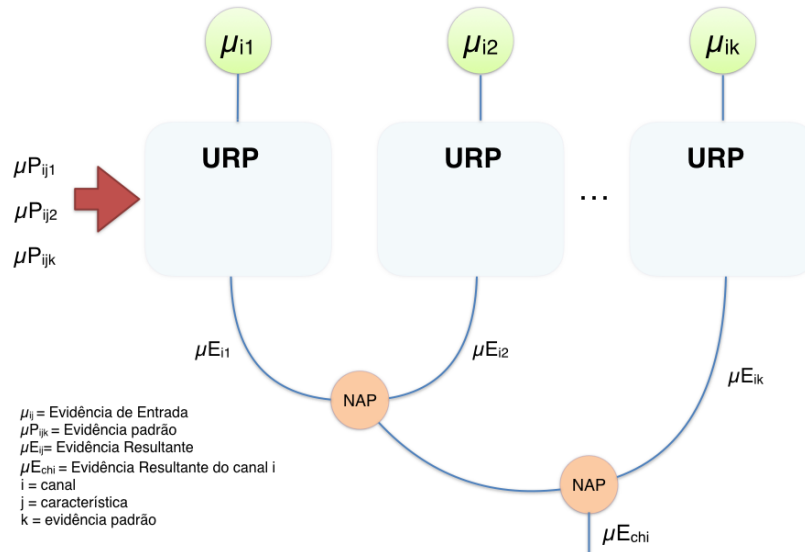
Fonte: Autor

Na segunda camada da URP é realizada uma Análise Paraconsistente para gerar um único Grau de Evidência Resultante, utilizando também duas NAP, conforme metodologia descrita previamente. Para cada tipo de característica extraída existe uma URP para realizar o processamento de reconhecimento de padrão que gera na saída um Grau de Evidência Resultante por característica dessa análise, conforme indicado na Figura 13. As saídas das URPs de cada característica extraída são analisadas por um NAP (caso tenha mais de uma URP por canal) para calcular um único Grau de Evidência Resultante do canal (baseados nos Graus de Evidência Resultante por característica). Esse valor será utilizado para calcular o Grau de Evidência Resultante por classe de Movimento Padrão.

A seguir, um único Grau de Evidência é calculado para cada um dos 12 canais de sEMG, usando o Grau de Evidência Resultante das URPs (uma URP para cada característica

extraída) como entrada do NAP e calculando o Grau de Evidência Resultante (μE_{ch}), de acordo com a Figura 13. Essa etapa gera uma lista de Graus de Evidência ($G\mu$) de valores combinados de cada canal ($\mu E_{ch1} \dots \mu E_{ch17}$) de todos os atributos para cada classe de movimento.

Figura 13 - Diagrama de blocos do uso de várias URPs quando se utiliza mais de uma característica extraída do sinal.



Fonte: Autor

4.1.4 Grau de Evidência Resultante por Classe de Movimento Padrão

Após a etapa de reconhecimento de padrão, para cada canal de aquisição é obtido um único Grau de Evidência que representa o quão próximo, ou longe, as características obtidas do movimento que se deseja classificar estão das características dos movimentos padrão. Ou seja, para cada classe de movimento é obtida uma lista de Graus de Evidência com 12 valores (representando 1 Grau de Evidência para cada canal de sEMG). Essa lista será a entrada de um algoritmo chamado de ParaExtr_{ctr} (Algoritmo Paraconsistente Extrator de Efeitos da Contradição) proposto por (DA SILVA FILHO, 2009). O objetivo do algoritmo é extrair gradativamente os efeitos da contradição até restar na saída um Grau de Evidência Real único. O valor resultante representará os diversos Graus de Evidência obtidos para uma preposição P , sendo, no caso em questão, uma classe de movimento.

O algoritmo ParaExtr_{ctr} utiliza os chamados Nós de Análise Paraconsistente (NAPs), descritos previamente, para extrair gradativamente os efeitos da contradição e resultar em um único Grau de Evidência Resultante Real. O algoritmo, a partir de uma lista de Graus de

Evidência de Entrada seleciona os valores mínimo (μ_{min}) e máximo (μ_{max}) dessa lista e utiliza esses valores como entrada (μ e λ) de um NAP, sendo que $\mu = \mu_{max}$ e $\lambda = 1 - \mu_{min}$. No NAP é realizada a Análise Paraconsistente das evidências μ e λ resultando em um Grau de Evidência Resultante μ_{R1} . O próximo passo é remover da lista os valores já analisados e substituí-los pelo Grau de Evidência calculado pelo NAP (μ_{R1}). Esse processo é repetido até que a lista de Graus de Evidência só contenha um único valor, que representa o Grau de Evidência Resultante Real para a classe de movimento que está sendo analisada. Os passos do algoritmo estão descritos na Tabela 2.

Sendo assim, ao final dessa etapa, cada RAP (representando cada classe de movimento) terá somente um Grau de Evidência Resultante Real μ_{ER} e, a partir desse valor, é possível classificar o sinal de entrada em um movimento.

Tabela 2 - Descrição dos passos do algoritmo de ParaExtr_{ctr} (DA SILVA FILHO, 2009).

Passos	Descrição
1	$G_{\mu} = (\mu_{ch1}, \mu_{ch2}, \mu_{ch3}, \mu_{ch4}, \mu_{ch5} \dots \mu_{ch12})$
2	$\mu_{max} = \max(G_{\mu})$
3	$\mu_{min} = \min(G_{\mu})$
4	$\mu = \mu_{max}$
5	$\lambda = 1 - \mu_{min}$
6	$\mu_{R1} = \text{NAP}(\mu, \lambda)$
7	$G_{\mu} = G_{\mu} - \mu_{max} - \mu_{min} + \mu_{R1}$
8	Retorna ao passo 2 até que G_{μ} tenha somente um único objeto
9	
10	$G_{\mu} = (\mu_{ER})$

4.1.5 Limitações do Método

O método RAP apresenta algumas limitações como o uso em processamento *online* de sinais de sEMG, no qual usualmente se faz necessário, uma segmentação em janelas deslizes. Neste caso, para cada movimento seriam extraídos diversos segmentos para análise, necessitando assim, de uma topologia mais complexa. Isso ocorre porque somente 3

amostras de cada Movimento Padrão não seriam suficientes para representar as diversas porções de um movimento, pois as características de cada segmento do sinal se alteram de acordo com a janela de tempo do movimento. Por exemplo, se é uma porção do início, meio ou fim da contração muscular.

O método também apresenta como restrição a sua baixa capacidade em se adaptar a outros contextos. Como explicado anteriormente, o método RAP foi desenvolvido para uma aplicação específica, podendo-se afirmar que o método não é adaptável a diversos tipos de sinais, características e movimentos.

De acordo com essas limitações apresentadas, foi proposto, na continuidade desta tese, um classificador híbrido, que une conceitos de Lógica Paraconsistente e do método *Random Forest*, com o objetivo de superar as restrições observadas do método RAP. Dessa forma, possibilitando a aplicação em outras áreas, por ser um algoritmo que envolve conceitos de aprendizagem de máquina.

4.2 DESENVOLVIMENTO DO MÉTODO PARACONSISTENT RANDOM FOREST

Após a avaliação preliminar da Lógica Paraconsistente através do método RAP (FAVIEIRO; MOURA; BALBINOT, 2016), foi proposto um classificador híbrido, que utiliza os conceitos de *Random Forest* agregados a Árvores de Decisão Paraconsistente, como classificador base. Dentre as diversas técnicas de *ensemble* baseadas em Árvores de Decisão, optou-se pela *Random Forest*, pois é uma das técnicas que geram melhores resultados neste contexto (BUHRMAN; WOLF, 2002). Além disso, como concluído em (BREIMAN, 2001), uma *Random Forest* também se mostra mais resistente ao ruído (quando uma porcentagem dos valores dos atributos dos dados são alterados aleatoriamente). Portanto, um sistema de múltiplos classificadores, no caso Árvores de Decisão, proporciona, como vantagem, a melhoria nos resultados se comparados aos de uma única Árvore de Decisão. Adicionalmente, o uso de Árvores de Decisão Paraconsistente agrega à *Random Forest* algumas das vantagens que foram mencionadas anteriormente para este tipo de técnica, ou seja, maior resistência a contaminantes e maior aplicabilidade a contextos incertos ou vagos. No presente estudo, a Lógica Paraconsistente foi utilizada de tal forma que é considerado que o sistema de reconhecimento de padrão possui um Grau de Certeza intrínseco que leva em conta as contradições e incertezas do sistema

A proposta da *Random Forest* de Árvores de Decisão Paraconsistente representa um conjunto de classificadores baseado em uma modificação, proposta nesta tese, na metodologia de Breiman (BREIMAN, 2001) a qual utiliza Lógica Paraconsistente para o tratamento dos dados. O método proposto é referido neste texto como *Paraconsistent Random Forest* (PRF).

Como pré-condição, é necessário dividir os dados em partições com um grau de pertinência correspondente, a fim de usar esses dados como entrada do sistema de classificação. Para calcular a inconsistência entre as amostras, é essencial que as partições divididas sejam difusas, ao invés de conjuntos *crispy* (valores nítidos). Dessa forma, considerando que os dados são *fuzzy*, uma amostra de dados pode pertencer a mais de uma partição, conseqüentemente, é possível lidar com esse tipo de inconsistência de uma maneira não trivial, baseada nos princípios da Lógica Paraconsistente. A partição de domínio *fuzzy* é considerada um estágio de pré-processamento do conjunto de dados, na qual pode ser usada qualquer metodologia de partição de domínio, desde que as partições geradas não sejam *crispy*.

Além disso, o sistema de múltiplos classificadores requer uma fase de aprendizado, onde um conjunto de dados é utilizado para construir as Árvores de Decisão, e um estágio de classificação, sendo que os dados de entrada serão classificados com um respectivo Grau de Certeza na classe correspondente.

4.2.1 Concepção da *Paraconsistent Random Forest*

Breiman (BREIMAN, 2001) propôs uma *Random Forest* na qual as Árvores de Decisão podem ser tão profundas quanto os dados permitirem, e nenhum algoritmo de poda é utilizado. No entanto, para construir cada árvore com baixa correlação de dados, a metodologia inclui dois elementos estocásticos durante a construção da mesma. O primeiro elemento estocástico se encontra na seleção do conjunto de dados de entrada, onde a técnica *bagging*⁷ é usada para selecionar os dados de cada árvore aleatoriamente. O segundo elemento estocástico é a seleção aleatória, em cada nó, de um pequeno grupo de variáveis de entrada para realizar a ramificação do mesmo. Devido a diversidade gerada pelas aleatorizações dos

⁷ O método *Bagging* é baseado em uma amostragem aleatorizada com substituição dos dados de entrada para definição do conjunto de treinamento de algoritmos de aprendizagem supervisionada, especialmente comum em Árvores de Decisão. Na amostragem com substituição, algumas amostras podem ser repetidas para o conjunto de treinamento, trazendo por consequência a redução da variância e ajudando a evitar o *overfitting* (BONISSONE *et al.*, 2010; BREIMAN, 2001).

dados na criação de Árvores de Decisão, a *Random Forest* pode apresentar uma solução melhor em termos de predição das classes, em relação à Árvores de Decisão, na presença de degradações nos dados de entrada.

Os dois parâmetros usados para projetar a *Random Forest* são o número de Árvores de Decisão e o número dos atributos do subconjunto para análise na divisão de nós. Neste caso, o tamanho do subconjunto é fixado no número de características ($\sqrt{F^N}$), proposto por Breiman (BREIMAN, 2001), sendo assim, o número de Árvores de Decisão se torna o único parâmetro no método de *Paraconsistent Random Forest* proposto, conforme descrito no Algoritmo A-2 da Tabela 3.

Tabela 3 - Descrição do Algoritmo A-2, *Paraconsistent Random Forest*.

Função **PRF** (entrada: amostras E, número de árvores T, saída: PRF)

1. Definir as partições de domínio fuzzy (DP) baseadas nas amostras de entrada E
2. Determinar o número de árvores T
3. For t = 1 to T
 - 3.1. Et = Bagging (E)
 - 3.2. ParaconsistentTrees(t) = ParaconsistentDecisionTree (Et, DP)
4. PRF = ParaconsistentTrees

fim

No Algoritmo A-2 (Tabela 3) são detalhadas as principais etapas para a seleção dos dados de treinamento que serão utilizadas para o desenvolvimento do modelo correspondente de cada Árvore de Decisão Paraconsistente, de maneira similar ao método *Random Forest* original, utilizando a técnica de *bagging*. Pode-se observar que o número de árvores de decisão é o único parâmetro de entrada para a construção da *Paraconsistent Random Forest*.

4.2.2 Partição do domínio com conjuntos *fuzzy*

O algoritmo *Fuzzy C-means* (FCM) foi utilizado para obter a partição do domínio da entrada do conjunto de dados, criando *clusters* a partir das entradas e definindo funções de pertinência gaussianas com base nos *clusters* calculados. No agrupamento *Fuzzy C-Means*, cada dado pertence a um ou mais *clusters* com um grau de pertinência especificado por classificações de amostras variando entre 0 e 1. Esse método, desenvolvido por Dunn em 1973

e aperfeiçoado por Bezdek em 1981, é comumente utilizado no reconhecimento de padrão (HOPPNER; KLAWONN, 2001) e baseia-se na minimização da seguinte função objetiva (15):

$$J_m = \sum_{i=1}^N \sum_{j=1}^C u_{ij}^m \|x_i - c_j\|^2, \quad 1 \leq m < \infty \quad (15)$$

onde m é qualquer número real maior que 1, u_{ij} é o grau de pertinência de x_i no *cluster* j , x_i é o i -ésimo de dados medidos d -dimensionais, c_j é o centro do *cluster* da dimensão d e $\|\cdot\|$ é qualquer norma que expresse a semelhança entre os dados medidos e o centro. A partição de dados ocorre através de uma iteração da função objetivo J_m , na qual ocorre a atualização do grau de pertinência u_{ij} e os centros de *clusters* c_j por (16):

$$u_{ij} = \frac{1}{\sum_{k=1}^C \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}}, \quad c_j = \frac{\sum_{i=1}^N u_{ij}^m \cdot x_i}{\sum_{i=1}^N u_{ij}^m} \quad (16)$$

O critério de parada para a iteração depende da seguinte condição (17):

$$\max_{ij} \left\{ \left| u_{ij}^{(k+1)} - u_{ij}^k \right| \right\} < \varepsilon \quad (17)$$

onde ε é um critério de parada entre 0 e 1, enquanto k representa os passos da iteração. O processo de iteração converge para um mínimo local da função objetivo J_m .

Como mencionado anteriormente, as funções de pertinência são responsáveis pelo comportamento *fuzzy* do algoritmo, no qual os dados não pertencem exclusivamente a um *cluster* bem definido. Nesse sentido, cada dado pode pertencer a vários *clusters* com valores diferentes de grau de pertinência, calculados pela sua respectiva função de pertinência.

4.2.3 Concepção da Árvore de Decisão Paraconsistente

A Árvore de Decisão Paraconsistente é baseada em uma modificação do algoritmo ID3 proposto por Quinlan (QUINLAN, 1986) e na *Random Forest* de Breiman (BREIMAN, 2001). Normalmente, no algoritmo ID3, um atributo é dividido em uma ou mais categorias, e o grau de pertinência das amostras de treinamento nesse conjunto particionado é binária, ou

seja, uma amostra pode pertencer ou não a essa partição. Entretanto, na modificação proposta, os conjuntos particionados são *fuzzy* e, portanto, elementos da lógica não clássica devem ser empregados. Desta maneira, a Lógica Paraconsistente é empregada para lidar com as inconsistências encontradas na partição do domínio *fuzzy*. À parte das adaptações citadas, os elementos básicos para a construção de uma Árvore de Decisão são respeitados. Estas são as notações usadas para descrever o algoritmo proposto nesta tese:

- C representa o conjunto de classes;
- F^N denota o conjunto de atributos que conduz ao nó N;
- O conjunto de atributos em um nó N é representado por $F^N = \{F_1, F_2, F_3, \dots, F_i\}$;
- P^N e I^N indicam, respectivamente, o Grau de Certeza total dos exemplos e o ganho de informação do nó N;
- $I^{S_{Fi}}$ representa a informação calculada para uma determinada característica i (F_i);
- $G_i^N = I^N - I^{S_{Fi}}$, indica o ganho de informação calculado para uma característica i (F_i) em um determinado nó N;
- P_k^N representa o grau de certeza propagado para uma C_k em um nó N;
- $X^N = \{X_j^N\}$ é o conjunto de graus de certeza propagados dos exemplos j que pertencem a um nó N;
- μf_k^c indica o grau de pertinência de uma partição *fuzzy* específica;

Através da técnica de *bagging*, a Árvore de Decisão Paraconsistente é inicializada com um subconjunto aleatório de atributos. Adicionalmente, o algoritmo necessita inicializar as variáveis X^{Root} e I^{Root} com o valor 1, pois o nó raiz contém todos os exemplos de treinamento e representa todo o espaço de descrição da Árvore de Decisão Paraconsistente específica, uma vez que nenhuma restrição é imposta ainda. Em cada iteração, o algoritmo determina um subconjunto aleatório dos atributos disponíveis para análise e seleciona o melhor para a divisão de nós. De acordo com (BONISSONE *et al.*, 2010) um dos mecanismos de seleção de atributos mais populares é aquele que maximiza o ganho de informação. Este mecanismo é computacionalmente simples, já que assume independência de atributos. Logo, para calcular o ganho de informação (I^N) do nó, são utilizadas as equações (18) e (19):

$$I^N = - \sum_{k=1}^{|C|} \left(\frac{P_k^N}{P^N} \cdot \log_2 \frac{P_k^N}{P^N} \right), \quad P^N = \sum_{k=1}^{|C|} P_k^N \quad (18)$$

$$P_k^N = \sum_{j=1}^{|E|} X_j^N \quad (19)$$

No algoritmo ID3 original, a função P_k^N representa a soma total de amostras pertencentes a uma classe específica k . Quando a partição é binária, uma amostra pode pertencer ou não a essa partição. Desta maneira, o valor de P_k^N indica o número total de amostras j que satisfaz as condições que levam ao nó N.

Na criação das Árvores de Decisão Paraconsistente, propostas nesta tese, é considerado o uso de conjuntos *fuzzy*, uma vez que uma amostra pode pertencer a um ou mais *clusters*, diferentemente do algoritmo ID3 que suporta somente o uso de categorias binárias. Sendo assim, foram necessárias algumas modificações no algoritmo ID3 para permitir o uso de partições *fuzzy*. A principal mudança, em relação ao algoritmo ID3, está na função X_j^N . Nesta função, o valor de influência de cada amostra é calculado pelo Grau de Certeza Paraconsistente de cada amostra pertencente ao nó anterior e pelo Grau de Pertinência da amostra pertencente à partição *fuzzy* específica μf_k^c . Portanto, o valor de X_j^N indica o Grau de Certeza de que a amostra j satisfaz às condições que levam ao nó N. A função foi definida pela Equação (20):

$$X_j^N = NAP(X_j^{N-1}, \mu f_k^c) \quad (20)$$

A Equação (18) utiliza o conceito de Nó de Análise Paraconsistente (NAP), descrito previamente no Algoritmo A-1 do Capítulo 2, para calcular o Grau de Certeza correspondente ao nó N, para cada exemplo j pertencente ao nó X_j^N .

Outro aspecto importante a considerar é que os dados de um atributo podem nem sempre estar disponíveis, como é o caso de muitas aplicações. No caso da ausência de valores para uma característica ($\mu f_k^c = NaN$), o resultado do $X_j^N = PAN(X_j^{N-1}, \mu f_k^c)$ é 0,5 representando o estado lógico de Indeterminado, de acordo com os princípios da Lógica Paraconsistente. Após o cálculo do ganho de informação do nó I^N , o atributo que possui o valor mais substancial de ganho de informação G_i^N é selecionado para realizar a divisão do nó através da função indicada pela Equação (21):

$$G_i^N = I^N - I^{SFi} \quad (21)$$

Como em cada iteração o algoritmo determina um subconjunto aleatório dos atributos disponíveis para análise e seleciona o que apresenta o maior ganho de informação G_i^N , alguns atributos (incluindo o melhor) podem não ser considerados para cada divisão. No entanto, um atributo excluído em uma divisão do nó pode ser usado por outras divisões na mesma Árvore de Decisão Paraconsistente, uma vez que o algoritmo seleciona os novos atributos para realizar a divisão do nó recursivamente, considerando apenas os atributos que ainda não foram selecionados.

Dessa forma, quando um atributo é selecionado para a divisão do nó, as partições de domínio do mesmo são utilizadas para gerar condições que levam a nós filhos. Logo, o conjunto de amostras presente no nó que está sendo dividido é particionado em subconjuntos que pertencerão aos nós filhos, de acordo com suas correspondências para essas condições. Desta maneira, pode-se concluir que na construção da Árvore de Decisão Paraconsistente nenhum nó é criado se não houver uma amostra de treinamento para suportá-lo.

Além disso, a Árvore de Decisão Paraconsistente é construída com cada nó não terminal representando o atributo selecionado no qual os dados foram divididos, bem como os nós terminais (ou nós folha) que representam o rótulo de classe do subconjunto final dessa ramificação. Os critérios de parada para o crescimento da Árvore de Decisão Paraconsistente (obtenção de um nó folha) dependem de uma das seguintes condições:

- Um nó é puro, ou seja, todas as amostras restantes pertencem à mesma classe, e o nó é transformado em uma folha e rotulado com a classe das amostras;
- Não há mais atributos disponíveis para a divisão do nó, mas as amostras ainda não pertencem à mesma classe. Em seguida, o nó é transformado em uma folha e rotulado usando votação majoritária para selecionar a classe com mais amostras associadas no subconjunto.

O Algoritmo A-3, representado pela Tabela 4, foi projetado para que as Árvores de Decisão Paraconsistente possam ser construídas sem considerar todos os atributos para dividir os nós, caso não exista necessidade de um atributo para obter um caminho específico.

Conforme a descrição dos Algoritmos A-2 e A-3, representados na Tabela 3 e Tabela 4, a *Paraconsistent Random Forest* proposta é capaz de integrar os conceitos básicos da *Random Forest* de Breiman com a Lógica Paraconsistente.

Tabela 4 - Descrição do Algoritmo A-3 - Árvore de Decisão Paraconsistente.

Função *ParaconsistentDecisionTree* (entradas: Et, DP, X)

1. Definir $X^{N=root} = 1, I^{N=Root} = 1$
2. Se nó folha (critério de parada)
 - 2.1. Classe do nó folha = C_K
 - 2.2. fim
3. Calcular o ganho de informação do nó (I^N)
4. Selecionar um subconjunto aleatório de atributos ($\sqrt{F^N}$) para possíveis divisões do nó
 - 4.1. Calcular o Ganho de Informação para cada atributo $F_i (I^{SF_i})$
5. Selecionar o atributo F_i para divisão do nó que maximize a equação: $G_i^N = I^N - I^{SF_i}$
6. Particionar o nó pelo atributo selecionado (F_i). Verificar todas as partições de domínio do atributo que contém um exemplo do conjunto E^N
7. Criar sub-nós com todos as partições de domínio validos para o atributo F_i
8. *ParaconsistentDecisionTree* (entrada: E^N, DP, X): função recursiva

fim

4.2.4 Inferência dos Valores

O procedimento de avaliação da Árvore de Decisão Paraconsistente é distinto ao do método ID3 (QUINLAN, 1986). A principal diferença é baseada no fato de que uma amostra percorre todos os caminhos da Árvore de Decisão Paraconsistente, e todas as folhas geram um Grau de Certeza associado à sua respectiva classe.

O Grau de Certeza de cada folha é calculado pelo Grau de Propagação do respectivo caminho usando a função X_j^N descrita na Equação (20). Assim, cada nó-folha possui uma classe associada com seu respectivo Grau de Certeza calculado, conforme mostrado no algoritmo A-4 da Tabela 5.

Nesse estágio, a avaliação de uma amostra gera vários Graus de Certeza e classes associadas. Para chegar à decisão final e definir a classe vencedora da floresta, é necessário definir uma ou mais abordagens de tomada de decisão, que serão apresentadas a seguir.

Tabela 5 - Descrição do Algoritmo A-4, inferência dos valores para a Árvore de Decisão Paraconsistente.

Função *ParaconsistentDecisionEvaluation* (entradas: e, DP, X, saída: leaf)

1. Definir $X^{N=root} = 1$
2. Se nó folha:
 - 2.1. leaf(N) = $[C_k, X_j^N]$
 - 2.2. fim
3. $X_j^{N+1} = PAN(X_j^N, \mu f_k^c)$
4. *ParaconsistentDecisionEvaluation* (in: e, DP, X^{N+1}): função recursiva

fim

4.2.5 Estratégias para Tomada de Decisão

Com base na avaliação da Árvore de Decisão Paraconsistente anteriormente mencionada, cada folha i de cada árvore t possui uma classe correspondente (C_k) com o respectivo Grau de Certeza (U_{ER}), que pode ser representado pelo vetor folha(i,t)= $[C_k, U_{ER}]$. Logo, para atribuir a classificação final relativa a uma amostra, é necessário definir algumas estratégias. De fato, é possível dividir as estratégias em dois grupos diferentes: estratégia (1), considerando a floresta como um classificador único; e estratégia (2), considerando a classificação individual de cada Árvore de Decisão Paraconsistente na floresta para a tomada de decisão da classe a qual a amostra pertence. A seguir cada uma das estratégias propostas será discutida.

Estratégia (1): considerando o Grau de Certeza de todas as folhas na floresta, e para determinar a classe com o grau associado mais alto, duas abordagens distintas são propostas:

- Floresta Média (AF): calcular o Grau de Certeza médio para cada classe, considerando o valor de todas as folhas na floresta e escolher a classe com *arg max*, com:

$$\text{AvgU}_{er}(c_k) = \text{média}(\text{folha}(C_k == k));$$

$$\text{ForestClass} = \text{argmax}(\text{AvgU}_{er});$$

- Floresta Paraconsistente (PF): calcular o Grau de Certeza para cada classe usando o Algoritmo Paraconsistente Extrator de Contradições (ParaExtrctr), descrito na Tabela 6, considerando o valor de todas as folhas na floresta, e escolher a classe com *arg max*, como segue:

$$\text{AvgU}_{er}(c_k) = \text{ParaExtr}(\text{folha}(C_k == k));$$

$$\text{ForestClass} = \text{argmax}(\text{AvgU}_{er});$$

O algoritmo ParaExtr_{ctr} proposto por (DA SILVA FILHO, 2009) é um algoritmo capaz de subtrair os efeitos de contradição de bancos de dados de conhecimento incertos que empregam Nós de Análise Paraconsistente (NAPs). Considerando que o Algoritmo PRF fornece uma lista de Graus de Certeza para cada classe associada, representando os valores calculados para cada caminho possível das Árvores de Decisão Paraconsistente, o Algoritmo ParaExtr_{ctr} seleciona os Graus de Evidência máximo e mínimo (μ_{\max} and μ_{\min}) para compor a entrada NAP. O objetivo principal dos NAPs é extrair o efeito da contradição, calculando o Grau de Evidência Resultante entre as entradas μ ($\mu = \mu_{\max}$) e λ ($\lambda = 1 - \mu_{\min}$). A seguir, os dois Graus de Evidência selecionados de G_μ são substituídos pelo Grau de Evidência Resultante (μ_{R1}). Esse processo é repetido até que apenas um Grau de Evidência permaneça na lista de Graus de Evidência (G_μ), conforme representado no Algoritmo-A5 da Tabela 6. No final deste processo, cada classe terá um Grau de Evidência único (μM), que é o valor representativo de todas as folhas consideradas.

Tabela 6 - Descrição do Algoritmo A-5, ParaExtr_{ctr} (DA SILVA FILHO, 2009).

Função **ParaExtr**(entrada: G_μ , saída: μM)

1. Receber uma lista de entrada com graus de certeza $G_\mu = (\mu E_1, \mu E_2, \mu E_3, \mu E_4, \mu E_5 \dots)$
2. Encontrar o valor máximo na lista $\mu_{\max} = \max(G_\mu)$
3. Encontrar o valor mínimo na lista $\mu_{\min} = \min(G_\mu)$
4. Calcular o grau de evidencia resultante usando os valores μ_{\max} e μ_{\min} como entrada do NAP

$$\mu_{R1} = \text{NAP}(\mu_{\max}, \mu_{\min})$$

5. Substituir na lista os valores μ_{\max} e μ_{\min} e adicionar o Grau de Evidencia Resultante μ_{R1} calculado:

$$G_\mu = G_\mu - \mu_{\max} - \mu_{\min} + \mu_{R1}$$

6. Ir para o passo 2 até que o tamanho da lista G_μ seja igual a 1.
7. O valor de saída é o ultimo grau de evidencia restante na lista $\mu M = G_\mu$

fim

Estratégia (2): utiliza as mesmas estratégias da Floresta. No entanto, as estratégias são aplicadas individualmente a cada Árvore de Decisão Paraconsistente. Primeiro, é necessário determinar uma classe vencedora para cada Árvore de Decisão Paraconsistente da floresta (usando a abordagem *arg max*) e, em seguida, usando a votação majoritária, decidir a classificação global da floresta. Nessa abordagem as duas possíveis opções são:

- **Árvore Média com Votação Majoritária (ATMV):** calcula o Grau Médio de Certeza para cada classe, considerando o valor de todas as folhas em cada Árvore de Decisão Paraconsistente. A escolha da classe vencedora da Árvore de Decisão Paraconsistente é feita com base no *arg max*. Finalmente, o conceito de votação majoritária é usado para decidir uma classe única da floresta, considerando a classificação individual de cada Árvore de Decisão Paraconsistente;
- **Árvore Paraconsistente com Votação Majoritária (PTMV):** consiste em calcular o Grau de Certeza para cada classe usando o algoritmo $\text{ParaExtr}_{\text{ctr}}$, levando em consideração o valor de todas as folhas em cada Árvore de Decisão Paraconsistente. Primeiramente, é necessário escolher a classe vencedora da Árvore de Decisão Paraconsistente com *arg max* e então definir uma única classificação da floresta usando o conceito de votação majoritária baseado na classificação única de cada Árvore de Decisão Paraconsistente da floresta.

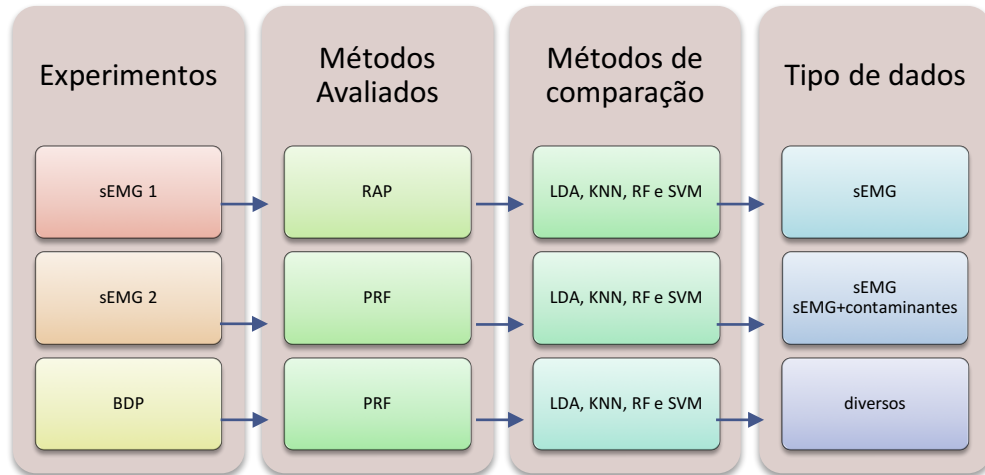
4.3 METODOLOGIA DE TESTES

A metodologia de testes para os algoritmos de reconhecimento de padrão desenvolvidos nesta tese compreende o uso de conjuntos de dados distintos, os quais são utilizados para avaliar a Taxa de Acerto Média (TAM) dos algoritmos propostos. Além disso, são apresentados experimentos com algumas modificações nos conjuntos de dados com o objetivo de avaliar o comportamento e a taxa de acerto dos algoritmos propostos perante a dados não ideais.

Ao todo foram propostos três experimentos para os testes e avaliações dos algoritmos desenvolvidos. Cada experimento utiliza um conjunto de dados único, sendo que existem dois conjuntos de dados que envolvem características extraídas de sinais mioelétricos do membro superior e um conjunto de dados contendo bases de dados padrão amplamente utilizadas na área de aprendizagem de máquina (LICHMAN, 2013), conforme diagrama da Figura 14. É importante enfatizar que o uso de bases de dados padrão é de extrema importância para validação do método de aprendizagem de máquina PRF, proposto nesta tese, pois garante

uma comparação com outros métodos reportados na literatura (BONISSONE *et al.*, 2010; BREIMAN, 2001), além de possibilitar avaliação do algoritmo com diferentes tipos de dados e aplicações.

Figura 14 - Descrição dos experimentos realizados na avaliação dos métodos.



Fonte: Autor

O experimento sEMG 1 envolve a avaliação da lógica paraconsistente através do método RAP. Considerando que o método RAP é um modelo fixo de classificação de dados de caráter preliminar, o mesmo não suporta o uso de uma base com grande quantidade de dados devido a sua simplicidade e por não possuir nenhuma etapa de treinamento de forma a otimizar seus parâmetros. Logo, os testes do método RAP utilizam uma base de dados de sinais de sEMG cujas características extraídas do sinal abrangem o segmento contendo a duração completa de cada movimento do membro superior.

O experimento sEMG 2 envolve a avaliação do método PRF no reconhecimento de movimentos através de sinais de sEMG. No entanto, como o método PRF é mais versátil em comparação ao RAP, o experimento sEMG 2 utiliza a metodologia de segmentação do sinal de sEMG através de janelas deslizantes, que não é suportada pelo algoritmo RAP. Esse tipo de segmentação é amplamente utilizado em processamento de sinais de sEMG, especialmente em contextos de processamento *online*.

O experimento BDP envolve a utilização de diversas bases de dados padrão na área de aprendizagem de máquina, e assim como o experimento sEMG 2, envolve somente a avaliação do método PRF, uma vez que esse tipo de dados não é suportado pela RAP proposta nesta tese.

Os sinais de sEMG do membro superior empregados para o desenvolvimento desses experimentos são provenientes da base de dados de livre acesso Ninapro (ATZORI *et al.*, 2012) e sinais adquiridos pelo Laboratório IEE/UFRGS (DE MOURA; BALBINOT, 2018; FAVIEIRO; MOURA; BALBINOT, 2016). Suas metodologias de aquisição serão descritas a seguir. No entanto, apesar dos experimentos denominados de sEMG1 e sEMG2 se basearem nas mesmas bases de sinais mioelétricos, o tipo de segmentação e características extraídas dos sinais são distintas, sendo assim, são considerados experimentos diferentes.

4.3.1 Descrição das Bases de Sinais de sEMG

O projeto Ninapro (ATZORI *et al.*, 2012) tem por objetivo auxiliar pesquisas de processamento de sinais mioelétricos do membro superior com a disponibilização de uma base de dados pública, em que seja possível realizar uma comparação direta entre metodologias diferentes. Foram disponibilizadas até o momento 7 bases de dados de sinais de sEMG. Entretanto, somente as bases de dados 2, 3 e 7 do projeto Ninapro foram utilizadas para os experimentos de teste, nesta tese, devido as mesmas apresentarem metodologias de aquisição similares, além da compatibilidade das mesmas com a base de dados do Laboratório IEE/UFRGS em termos de quantidade e localização dos eletrodos, bem como a sequência de movimentos realizados pelos voluntários. Na Tabela 7 é apresentado um resumo com as características das bases de dados selecionadas do projeto Ninapro para a metodologia de testes utilizadas neste trabalho. Para a base de dados 3 e 7 do projeto Ninapro também foram disponibilizadas as características clínicas dos voluntários com diferentes níveis de amputação do membro superior (consultar Tabela 8).

Tabela 7 - Tabela com resumo das bases de dados Ninapro.

	Base de Dados 2	Base de Dados 3	Base de Dados 7
Sujeitos intactos	40	0	20
Sujeitos com amputação transradial	0	11	2
Número de eletrodos de sEMG	12	12	12
Número total de movimentos	50	50	41
Repetições de cada movimento	6	6	6
Grupo de movimentos realizados	B, C, D, E	B, C, D, E	B, C, D

Legenda: O Grupo A de movimentos corresponde a movimentos básicos dos dedos (flexões e extensões); o Grupo B é relacionado a configurações isométricas e isotônicas (“posturas das mãos”); o Grupo C são os movimentos básicos do punho; o Grupo D são movimentos do tipo “garra” e funcionais; o Grupo E são movimentos de dedos para medição de padrão de força.

Fonte: Adaptado de (ATZORI *et al.*, 2014).

Tabela 8 - Quadro-resumo das características clínicas dos voluntários com amputação.

Sujeito	Amputação	Causa da amputação	Antebraço restante (%)	Anos após amputação	Sensação de membro fantasma (0-5)	Uso de prótese
1	Direita	Acidente	50	13	2	Mioelétrica
2	Esquerda	Acidente	70	6	5	Cosmética
3	Direita	Acidente	30	5	2	Mioelétrica
4	Ambas	Acidente	40	1	1	Não usa
5	Esquerda	Acidente	90	1	2	Cinemática
6	Esquerda	Acidente	40	13	4	Cinemática
7	Direita	Acidente	0	7	0	Não usa
8	Direita	Acidente	50	5	2	Mioelétrica
9	Direita	Acidente	90	14	5	Mioelétrica
10	Direita	Acidente	50	2	5	Mioelétrica
11	Direita	Câncer	90	5	4	Mioelétrica
12	Direita	Acidente	50	6	*	gancho
13	Direita	Câncer	50	18	*	gancho

* informação não disponibilizada

Fonte: Adaptado de (ATZORI *et al.*, 2014).

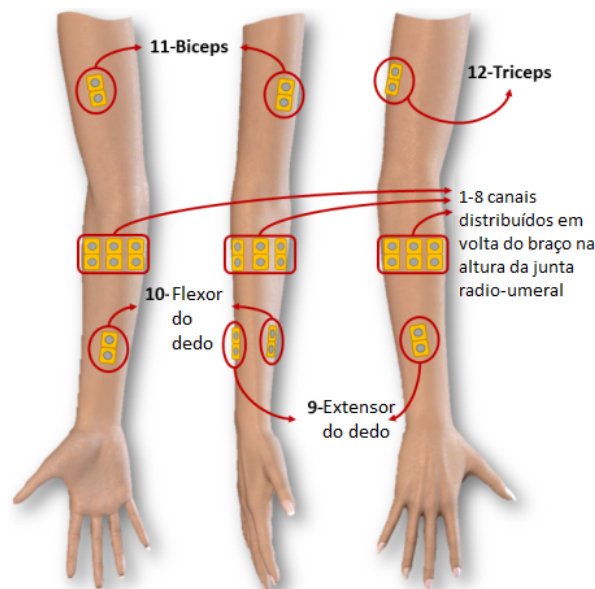
Os sinais de sEMG dos 40 sujeitos não amputados (idade entre 26 e 45 anos) da base de dados 2, 11 sujeitos amputados (idade entre 32 e 67 anos) da base de dados 3 e 2 sujeitos amputados (idade 28 e 54 anos) da base de dados 7 do Ninapro foram utilizados nos experimentos chamado de sEMG1 e sEMG2, que serão descritos na sequência, para os testes dos métodos RAP e PRF desenvolvidos. Em relação aos movimentos executados pelos sujeitos, o projeto Ninapro (ATZORI *et al.*, 2014) definiu uma sequência (não aleatória) de forma a encorajar movimentos repetitivos de maneira quase inconsciente. O grupo de movimentos selecionados para avaliação nesta tese é denominado "B" pelos autores do estudo (ATZORI *et al.*, 2015) e seus movimentos estão descritos na Tabela 9.

O posicionamento dos eletrodos é o mais semelhante possível em todas as bases de dados, conforme detalhado na Figura 15. Ao total, são 12 pares de eletrodos divididos em 8 pares de eletrodos igualmente espaçados ao redor do antebraço do voluntário (2 cm abaixo do cotovelo), na altura da junta rádio-umeral e mais 4 pares de eletrodos localizados em músculos específicos: bíceps, tríceps, flexor dos dedos e extensor dos dedos.

Tabela 9 - Descrição dos movimentos.

Movimento	Descrição
1	Polegar para cima
2	Extensão dos dedos indicador e médio
3	Flexão dos dedos anular e mínimo
4	Flexão do polegar e extensão dos outros dedos
5	Abdução dos dedos
6	Fechar a mão
7	Apontando dedo indicador
8	Adução dos dedos
9	Supinação do punho (eixo de rotação: dedo médio)
10	Pronação do punho (eixo de rotação: dedo médio)
11	Supinação do punho (eixo de rotação: dedo mínimo)
12	Pronação do punho (eixo de rotação: dedo mínimo)
13	Flexão do punho
14	Extensão do punho
15	Desvio radial do punho
16	Desvio ulnar do punho
17	Extensão do punho com mão fechada

Fonte: Adaptado de (ATZORI *et al.*, 2014).

Figura 15 - Posicionamento dos eletrodos na configuração bipolar.

Fonte: Adaptado de (MOURA; FAVIEIRO; BALBINOT, 2016).

Adicionalmente, com o objetivo de padronizar a aquisição dos sinais mioelétricos para todos os voluntários, a metodologia de ensaios, do Laboratório IEE/UFRGS, baseou-se na metodologia definida no Projeto Ninapro (ATZORI *et al.*, 2012) com relação aos movimentos, repetições, utilização de estímulo visual, número de eletrodos e posicionamento dos mesmos. É importante observar que, até o presente momento, para os sinais de sEMG adquiridos pelo

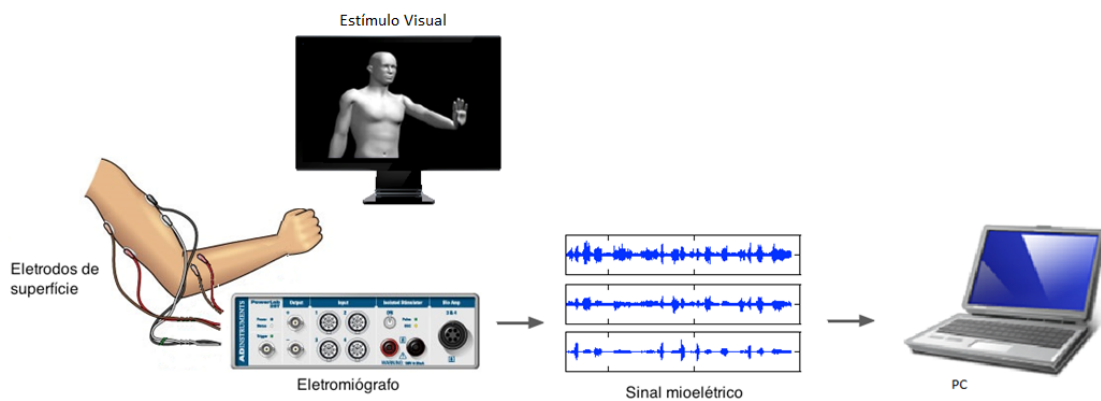
Laboratório IEE/UFRGS não é realizado monitoramento da força aplicada para os movimentos executados. Os principais aspectos considerados da metodologia de ensaios do projeto Ninapro (ATZORI *et al.*, 2012) são descritos a seguir:

- Cada ensaio é composto por 6 repetições de cada um dos 17 movimentos;
- Os movimentos são realizados sequencialmente;
- Cada movimento tem duração total de 5 segundos;
- O tempo de descanso entre movimentos é de 3 segundos.

A base de dados do Laboratório IEE/UFRGS está em constante atualização e na época de desenvolvimento deste trabalho a base continha um total de 20 ensaios de sujeitos entre 20 a 35 anos com membros superiores intactos, cujos sinais de sEMG foram utilizados nos experimentos de testes sEMG1 e sEMG2. Os ensaios estão regulamentados pelo Conselho de Ética da UFRGS com projeto de Código 11253312.8.0000.5347.

Na Figura 16, encontra-se o diagrama de blocos simplificado do fluxo da aquisição de sinais de sEMG, demonstrando visualmente os principais blocos do sistema. A aquisição da base de dados dos sinais de sEMG, tanto na base Ninapro como na desenvolvida no Laboratório IEE/UFRGS, envolve os seguintes itens: o estímulo visual, equipamentos e materiais utilizados, a digitalização dos sinais de sEMG e o correspondente armazenamento dos dados em arquivos para posterior análise. O processo de aquisição inicia com o estímulo visual, o qual permite ao usuário visualizar e repetir cada movimento exibido na tela por um modelo de braço virtual, de modo que todos os voluntários executem os movimentos da melhor maneira possível simultaneamente com o vídeo. O voluntário é instruído a replicar os movimentos utilizando força moderada, porém é importante ressaltar que não existe nenhum sistema de *feedback* para o usuário.

Figura 16 - Diagrama de blocos de aquisição de dados.



Fonte: Autor

Após o sinal de sEMG ser adquirido, o mesmo passa por algumas etapas de pré-processamento que precedem a extração das características do sinal para utilização no classificador. Essas etapas consistem em: filtro digital, ajuste de *offset*, retificação de onda completa, normalização, segmentação do sinal e extração de características, conforme demonstrado no diagrama de blocos da Figura 17.

Figura 17 - Diagrama de blocos do processo de pré-processamento de sinais de sEMG.



Fonte: Autor

Foi utilizado um filtro passa-banda FIR de ordem 20, com frequência de corte em 20-500 Hz e amostragem de 2kHz, de acordo com o sinal de sEMG adquirido. Essa frequência de corte foi selecionada porque, usualmente, sinais de sEMG adquiridos a partir de músculos do membro superior têm uma distribuição em frequência concentrada entre 50 Hz ~ 500 Hz (DE LUCA, 1997). O filtro utilizado foi baseado em (CENE, 2016) para eliminar qualquer ruído que possa ainda existir fora da faixa de interesse do sinal.

Adicionalmente, para a utilização dos segmentos de sinal em algoritmos de treinamento supervisionado, se faz necessária a definição de classes-alvo para cada segmento. Para a definição das classes-alvo de cada segmento do sinal da base de dados Ninapro (ATZORI *et al.*, 2012) foi utilizado o arquivo chamado “*stimulus*”, disponibilizado pela base de dados, que contempla informações como início e fim de cada movimento, bem como qual movimento realizado. Esse arquivo foi gerado a partir da temporização dos vídeos e de um tratamento posterior, que utiliza um algoritmo chamado de *Generalized Likelihood Ratio* para realinhar os limites de cada movimento (ATZORI *et al.*, 2014) e realizar um ajuste fino dos tempos de início e fim dos mesmos. De mesma forma, a base de dados IEE também fornece um arquivo “*stimulus*”, que é realizado de forma similar ao processo da base Ninapro, conforme descrito em (TOSIN, 2018).

Na Tabela 10 se encontra, de forma resumida, as principais características das bases de sinais de sEMG utilizadas e sua respectiva nomenclatura que será utilizada nos experimentos sEMG 1 e sEMG 2.

Tabela 10 - Resumo das bases de sinais de sEMG utilizadas.

Base de sinais de sEMG	Sigla	Sujeitos com amputação	Sujeitos com membro superior intacto
Laboratório IEE/UFRGS	IEE	0	20
Ninapro 2	NPI	0	40
Ninapro 3 e 7	NPA	13	0

Adicionalmente, é importante ressaltar que a base de dados do Laboratório IEE/UFRGS para sinais de sEMG, assim como, todos os procedimentos experimentais, são fonte de diversos trabalhos do Laboratório IEE, no decorrer dos anos, tendo ampla participação da correspondente autora desta tese (BALBINOT; FAVIEIRO, 2013; CENE; FAVIEIRO; BALBINOT, 2015; CENE, 2016; CENE; BALBINOT, 2015, 2018; FAVIEIRO; BALBINOT, 2011a; FAVIEIRO, 2012; FAVIEIRO; MOURA; BALBINOT, 2016; MOURA; FAVIEIRO; BALBINOT, 2016; TOSIN *et al.*, 2017). Mais especificamente, algumas contribuições da autora se encontram na participação nos ensaios de aquisição dos sinais, definição do protocolo experimental e desenvolvimento do software de aquisição de dados considerando o sincronismo dos modelos do braço virtual e das placas de aquisição de dados utilizadas.

4.3.2 Experimento sEMG 1

O experimento sEMG1 utiliza como entrada, para os algoritmos de reconhecimento de padrão, características extraídas dos sinais de sEMG provenientes do projeto Ninapro e adquiridos no Laboratório IEE/UFRGS, descritas previamente. O objetivo dos algoritmos de reconhecimento de padrão é classificar corretamente os 17 movimentos do segmento mão-braço descritos na Tabela 9.

Neste experimento a segmentação do sinal de sEMG foi baseada no arquivo de “*stimulus*”, o qual disponibiliza, de maneira aproximada, os marcadores de início e fim de cada movimento realizado. Portanto, cada segmento do sinal que foi analisado continha aproximadamente a duração completa de um movimento (5 segundos). A partir dos segmentos do sinal, foram extraídas características dos mesmos, para seu uso como entrada dos classificadores utilizados.

A extração de características do sinal de sEMG ainda é um desafio, devido à complexidade intrínseca de escolher as características mais representativas do sinal. Para esta análise, três características foram extraídas de cada segmento de sinal: *Root Mean Square*

(RMS), Frequência Mediana (fMED) e Variância (VAR). Essas características foram empregadas em estudos anteriores e apresentaram bons resultados quanto à taxa de acerto, quando extraídas em referência ao segmento de duração total de um movimento (CENE; FAVIEIRO; BALBINOT, 2015; CENE; BALBINOT, 2015; FAVIEIRO; MOURA; BALBINOT, 2016; MOURA; FAVIEIRO; BALBINOT, 2016). Na Tabela 11 estão descritas todas as características extraídas em termos de equações matemáticas e suas respectivas abreviações. Em todas as equações, $sEMG_i$ representa os valores amostrados do sinal mioelétrico no instante i no segmento k , n representa o número de amostras da janela, P representa o espectro de potência do sinal de sEMG na posição j .

Tabela 11 - Características extraídas do sinal e suas definições matemáticas.

Característica	Sigla	Definição matemática
Root Mean Square	RMS	$RMS_k = \sqrt{\frac{1}{n} \sum_{i=1}^n sEMG_i^2}$
Frequência Mediana	fMED	$\frac{1}{2} \sum_{j=1}^{fMED_k} P_j = \frac{1}{2} \sum_{j=fMED_k}^n P_j$
Variância	VAR	$VAR_k = \frac{1}{n} \sum_{i=1}^n (sEMG_i - \mu)^2$ $\mu = \frac{1}{n} \sum_{i=1}^n sEMG_i$

No experimento sEMG 1 a avaliação dos algoritmos envolveu o uso de todas as combinações possíveis das três características extraídas do sinal, possibilitando uma comparação quantitativa das taxas de acerto. Essa comparação possibilita, para esse estudo de caso, avaliar se alguma combinação obteve uma taxa de acerto mais elevada em relação às outras.

4.3.3 Experimento sEMG 2

Assim como no Experimento sEMG 1, o Experimento sEMG 2 também utiliza como entrada, para o método PRF, características extraídas dos sinais de sEMG provenientes do

projeto Ninapro e adquiridos no Laboratório IEE/UFRGS, com o objetivo de classificar os 17 movimentos do segmento mão-braço. No entanto, a principal diferença para o experimento sEMG 1 se encontra no tipo de segmentação do sinal utilizada para a extração das características, bem como a quantidade e tipos de características do sinal utilizadas. O experimento sEMG 2 é baseado na metodologia desenvolvida por (DE MOURA; BALBINOT, 2018) com o intuito de posterior comparação do sistema tolerante a falhas desenvolvida por De Moura, com o algoritmo PRF proposto nesta tese.

A metodologia do experimento sEMG2 em relação à segmentação do sinal de sEMG se baseia em métodos de processamentos *online*, em que usualmente é utilizada a técnica de janelas sobrepostas, com tamanhos de janela variando entre 20-300 ms (FARRELL, 2011). Já foi demonstrado que existe uma relação entre a taxa de acerto da classificação e o tamanho da janela de dados utilizada, na qual janelas maiores apresentam maior taxa de classificação correta dos movimentos (ATZORI *et al.*, 2015; ENGLEHART; HUDGINS, 2003; FARRELL, 2011). Assim a metodologia de (DE MOURA; BALBINOT, 2018) propõe o tamanho da janela deslizante em 300 ms com o incremento de 75 ms, pelo seu extenso uso na literatura.

Na metodologia proposta por (DE MOURA; BALBINOT, 2018), na qual o experimento sEMG 2 se baseia, as características extraídas do sinal que apresentaram melhores resultados de taxa de acerto de classificação média para os sujeitos analisados foram: *rRoot Mean Square* (RMS), *Mean Absolute Value* (MAV), Frequência Mediana (fMED), Comprimento de Onda (WL), *Maximum Fractal Length* (MFL) e Obliquidade (SKW). Além disso, essas características utilizadas em (DE MOURA; BALBINOT, 2018) foram selecionadas com base no resultado da classificação de outros estudos com o mesmo banco de dados (ENGLEHART; HUDGINS, 2003; MOURA; FAVIEIRO; BALBINOT, 2016; TOSIN *et al.*, 2017) e também pela avaliação de características dos sinais de sEMG realizada por (PHINYOMARK *et al.*, 2013). Na Tabela 12 estão descritas todas as características extraídas em termos de equações matemáticas e suas respectivas abreviações. Em todas as equações, $sEMG_i$ representa os valores amostrados do sinal mioelétrico no instante i no segmento k , n representa o número de amostras da janela, P representa o espectro de potência do sinal de sEMG na posição j e M é o comprimento máximo do espectro de potência.

Para avaliar o método PRF de forma mais abrangente, o experimento sEMG 2, além de utilizar características extraídas de janelas deslizantes dos sinais de sEMG originais para a avaliação da *Paraconsistent Random Forest*, também utiliza os sinais de sEMG contaminados artificialmente, segundo metodologia de (DE MOURA; BALBINOT, 2018). O uso de sinais

contaminados artificialmente tem o objetivo de ocasionar possíveis contradições nos modelos gerados *a priori* e, com isso, poder avaliar o efeito que essas degradações do sinal de sEMG tem sobre a taxa de acerto no classificador utilizado.

Como a proposta desta tese envolve o estudo, concepção e desenvolvimento de um método de reconhecimento de padrão robusto a degradações nos sinais de entrada foi optado por utilizar contaminantes adicionados artificialmente no sinal de sEMG de forma a possibilitar a validação do método com a garantia da degradação do sinal de sEMG e assim avaliar os efeitos dessas degradações na predição do classificador. Para isso, os modelos dos classificadores foram gerados a partir do sinal de sEMG íntegro e testados com características extraídas do sinal de sEMG degradado com contaminantes. Adicionalmente, a utilização de sinais de sEMG com adição de contaminantes artificiais possibilita avaliar os efeitos de cada contaminante analisado separadamente.

Tabela 12 - Características extraídas do sinal e suas definições matemáticas.

Característica	Sigla	Definição matemática
Root mean square	RMS	$RMS_k = \sqrt{\frac{1}{N} \sum_{i=1}^N sEMG_i^2}$
Frequência mediana	fMED	$\frac{1}{2} \sum_{j=1}^{fMED_k} P_j = \frac{1}{2} \sum_{j=fMED_k}^n P_j$
Mean absolute value	MAV	$MAV_k = \frac{1}{N} \sum_{i=1}^N sEMG_i $
Waveform Length	WL	$WL_k = \sum_{i=1}^{N-1} sEMG_{i+1} - sEMG_i $
Maximum Fractal Length	MFL	$MFL_k = \log_{10} \left(\sqrt{\sum_{i=1}^{N-1} (sEMG_i - sEMG_{i+1})^2} \right)$
Obliquidade	SKW	$SKW_k = \frac{\frac{1}{N} \sum_{i=1}^N (sEMG_i - \overline{sEMG})^3}{\left(\sqrt{\frac{1}{N} \sum_{i=1}^N (sEMG_i - \overline{sEMG})^2} \right)^3}$

Fonte: adaptado de (DE MOURA; BALBINOT, 2018).

Em relação a avaliação dos tipos de contaminantes na área de reconhecimento de movimentos, foram selecionados três tipos de contaminantes de sinal de sEMG: artefato de movimento, perda de contato e ruído térmico, por seu uso extenso na literatura e por degradarem o sinal de sEMG a ponto de ocasionar uma redução na predição do classificador, que por vezes é necessário o retreinamento do modelo do mesmo (DE MOURA; BALBINOT, 2018; FRASER *et al.*, 2014; MCCOOL *et al.*, 2014; XIAORONG ZHANG; HE HUANG; QING YANG, 2013; ZHANG; HUANG, 2015). É importante ressaltar que, nesta tese, foram utilizados os modelos matemáticos dos contaminantes propostos por (DE MOURA; BALBINOT, 2018) para efetuar a contaminação artificial dos sinais de sEMG, conforme descrição na Tabela 13.

Tabela 13 - Modelos matemáticos dos contaminantes de sinal de sEMG propostos por (DE MOURA; BALBINOT, 2018).

Contaminante	Sigla	Definição matemática
Artefatos de movimento	AF	$y_{MA}(t) = 1500r \cdot t^2 \cdot e^{-140t} \sin(60.2\pi(t - 20))$ <p>onde, o modelo se repete com um número aleatório r de 1 a 10 para cada amostra de tempo t [1,2000].</p>
Ruído térmico	RT	$y_{DE}(t) = \sqrt{var} \cdot randn(t)$ $var = 10^{-SNR/10}$ <p>onde, a variância var é definida através do SNR, no caso de 15 dB, e $randn$ são números aleatórios normalmente distribuídos.</p>
Perda de contato eletrodo-pele	PC	$y_S(t) = 12,6 \cdot \left(\frac{\sin(203.2\pi t)}{100} + \frac{3 \cdot \sin(210.2\pi t)}{200} + \frac{\sin(213.2\pi t)}{200} + \frac{4 \cdot \sin(220.2\pi t)}{5} + \frac{9 \cdot \sin(226.2\pi t)}{1000} + \frac{3 \cdot \sin(230.2\pi t)}{200} + \frac{4 \cdot \sin(236.2\pi t)}{500} \right).$

Fonte: adaptado de (DE MOURA; BALBINOT, 2018).

O contaminante artefato de movimento (também chamado de artefato de contato) é caracterizado por um ruído de baixa frequência causado por movimentos nos cabos conectados aos sensores do sEMG e também por movimentos do sensor em relação a pele (HUANG *et al.*, 2010; ZHANG; HUANG, 2015). Esse contaminante é comum em aquisições de sinal de sEMG e pode levar a desvios da magnitude do sinal de sEMG (HUANG *et al.*,

2010). No estudo de (FRASER *et al.*, 2014), o contaminante artefato de movimento foi adquirido a partir de sinais de sEMG provenientes de eletrodos localizados em partes inativas do corpo de voluntários em condições ambulatoriais. Esses dados foram multiplicados por um fator de ganho e adicionados aos sinais de sEMG para simular o contaminante artefato de movimento. Na proposta de (DE MOURA; BALBINOT, 2018) o contaminante artefato de movimento foi adquirido a partir de ensaios experimentais, na qual foi o artefato de movimento foi simulado através de uma pequena batida no eletrodo de superfície. O sinal contendo o artefato de movimento foi analisado em tempo e frequência para a definição de um modelo matemático que representasse o contaminante.

O contaminante perda de contato eletrodo-pele ou perda de contato pode causar variações na impedância entre o eletrodo e a pele de forma que a informação do sinal de sEMG é degradada. A simulação da perda de contato eletrodo-pele seguiu diretrizes de (DE MOURA; BALBINOT, 2018; ZHANG; HUANG, 2015) no qual uma amostra experimental desse contaminante foi adquirida ao puxar gentilmente o eletrodo bipolar durante um momento de repouso do usuário, cuidando para que o eletrodo de superfície não fosse totalmente removido. A partir dessa amostra (DE MOURA; BALBINOT, 2018) realizou uma análise no tempo e frequência, propondo um modelo matemático que simulasse artificialmente o contaminante “perda de contato”, que no estudo de (DE MOURA; BALBINOT, 2018) é chamado de saturação do sinal.

O contaminante ruído térmico, na forma de ruído branco gaussiano, é comumente observado em ensaios de sEMG de longa duração, devido a interferências magnéticas, ruído interno do amplificador e ruído térmico dos elementos eletrônicos (GENG *et al.*, 2018; ZHANG; HUANG, 2015). Em diversos estudos o ruído térmico, ou ruído de base, é simulado através de adição de ruído gaussiano branco, com valores de SNR variando entre 0 dB a 30 dB (GENG *et al.*, 2018; ZHANG; HUANG, 2015). Na proposta de (DE MOURA; BALBINOT, 2018) foi realizada a adição de ruído branco gaussiano com SNR de 15 dB para simular artificialmente o contaminante “deslocamento de eletrodo”, de acordo com ensaios experimentais. Nesta tese também foi utilizado a adição de ruído branco gaussiano com SNR de 15 dB, mas a nomenclatura utilizada é ruído térmico por ser mais fidedigna ao encontrado na literatura.

Analisando a pesquisa de (DE MOURA; BALBINOT, 2018) foi possível observar que as configurações de canais afetados que têm maior influência na taxa de acerto do modelo SVM foram os eletrodos posicionados em músculos específicos. Por essas razões, foram selecionadas, da metodologia proposta por (DE MOURA; BALBINOT, 2018), 8

configurações de contaminação dos canais, descritas na Tabela 14, que envolvem os músculos específicos, nas quais foi avaliada a necessidade do uso de um método mais robusto aos contaminantes do sinal de sEMG.

Tabela 14 - Descrição dos canais contaminados neste estudo.

Caso de Estudo	Canais Contaminados
9	Eletrodo sobre o músculo <i>flexor digitorum</i>
10	Eletrodo sobre o músculo <i>extensor digitorum</i>
11	Eletrodo sobre o músculo bíceps
12	Eletrodo sobre o músculo tríceps
13	Eletrodos sobre os músculos <i>flexor digitorum</i> e <i>extensor digitorum</i>
14	Eletrodos sobre os músculos bíceps e tríceps
15	Eletrodos sobre os músculos <i>flexor digitorum</i> , <i>extensor digitorum</i> , bíceps e tríceps
16	Eletrodos uniformemente espaçados abaixo da junta radio-umeral (8 pares).

Fonte: adaptado de (DE MOURA; BALBINOT, 2018)

Adicionalmente, para a tomada de decisão do método PRF, optou-se pela estratégia ATMV, similar ao método *Random Forest* descrita anteriormente. A ATMV determina uma classe vencedora para cada Árvore de Decisão Paraconsistente com *arg max* e aplica o conceito de votação majoritária para decidir uma classe única da floresta, considerando a classificação individual de cada Árvore. Essa estratégia foi selecionada pela facilidade dos cálculos e tempo de processamento, uma vez que as Árvores de Decisão geradas para o processamento de sinais de sEMG são extensas, devido ao número de partições de domínio e quantidade de atributos de entrada. Importante ressaltar que não foi notada nenhuma diferença significativa na taxa de classificação, com base em ensaios preliminares, entre os métodos de tomada de decisão avaliados.

4.3.4 Experimento Base de Dados Padrão (BDP)

O experimento BDP tem o objetivo de avaliar o algoritmo PRF com bases de dados genéricas, sem uma aplicação-alvo, como o estudo de caso dos experimentos sEMG 1 e sEMG. As bases de dados utilizadas para essa metodologia de teste foram selecionadas a partir do Repositório UCI (LICHMAN, 2013), que contém bases de dados padrão para uso na avaliação de algoritmos de classificação. Esse repositório é amplamente utilizado para a avaliação e *benchmarks* de algoritmos de reconhecimento de padrão (BONISSONE *et al.*, 2010; DIETTERICH, 2000; FRAIMAN; JUSTEL; SVARC, 2010; LICHMAN, 2013;

WANG *et al.*, 2015). As bases de dados selecionadas e suas características são especificadas na Tabela 15.

Tabela 15 - Características das bases de dados.

Base de dados	Sigla	# amostras	# atributos	# classes
<i>Iris Plant</i>	IRS	150	4	3
<i>Breast Cancer Wisconsin (Prognostic)</i>	PBC	198	34	2
<i>Breast Cancer Wisconsin (Diagnostic)</i>	DBC	569	32	2
<i>Glass Identification</i>	GLA	214	10	6
<i>Wine</i>	WIN	178	13	3
<i>Cancer</i>	CNC	216	4000	2

Fonte: adaptado de UCI (LICHMAN, 2013)

Além da análise de classificação do algoritmo com as bases de dados originais, também foram realizados dois procedimentos experimentais aplicando o conceito de dados não ideais nas bases de dados padrão selecionadas, para avaliar a viabilidade e robustez do método PRF a variantes dos atributos de entrada, baseado na metodologia de (BONISSONE *et al.*, 2010).

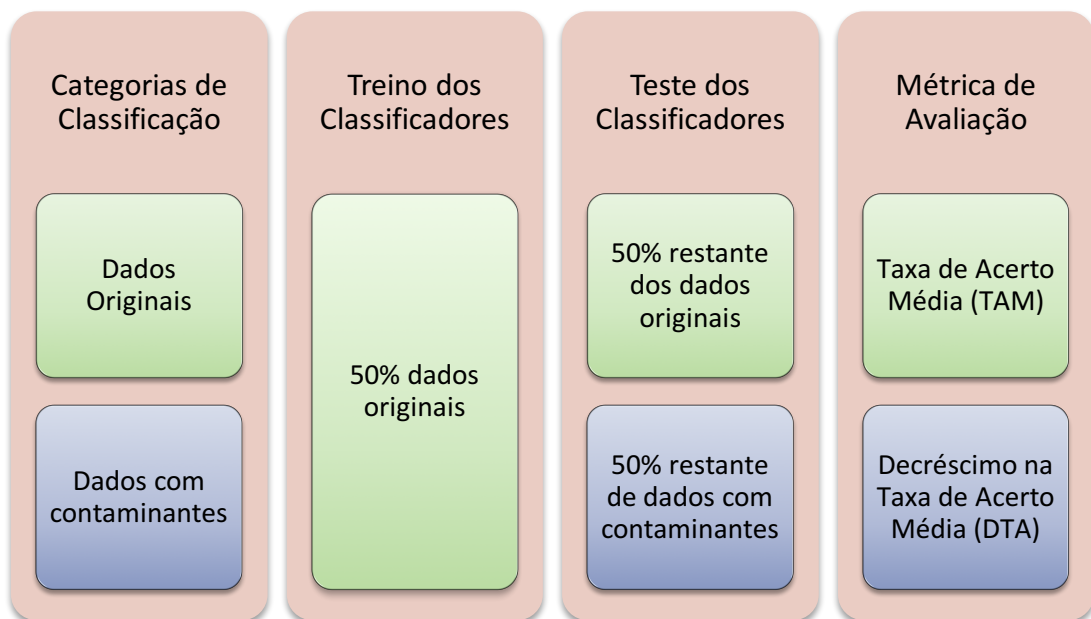
O primeiro procedimento envolveu uma redução artificial dos dados de entrada, chamado de experimento com dados ausentes, ao efetuar uma substituição de atributos de entrada por valores inválidos NaN (*Not-a-Number*). Esse procedimento experimental avalia o método na ausência de dados. De maneira similar, no segundo experimento foi realizada uma análise dos dados na presença de ruído, que foi simulado ao realizar uma substituição de atributos de entrada por um valor aleatório uniformemente distribuído.

Nos dois experimentos a quantidade de dados substituídos foi calculada de acordo com $a\% (E \cdot M)$, em que $a\%$ representa a porcentagem de valores a ser substituída, E representa o número total de amostras da base de dados e M é o número de atributos que cada amostra contém. Os atributos a serem substituídos foram selecionados aleatoriamente entre todas as amostras do conjunto de teste do classificador. É importante ressaltar que o treino dos classificadores foi realizado com uma parcela do conjunto de dados original. Adicionalmente, a porcentagem de dados inválidos foi definida empiricamente com valores variando de 5 a 90%.

4.3.5 Avaliação dos métodos

O procedimento de avaliação dos métodos é similar em todos os experimentos e foi aplicado em 2 configurações diferentes. O primeiro caso corresponde a classificação dos dados íntegros, na qual não foi adicionado nenhum contaminante. Neste caso, 50% dos dados do sinal íntegro (sem adição de contaminantes) foram usados para treinamento e os outros 50% para o teste. No segundo caso, foram utilizados os mesmos modelos dos classificadores gerados a partir de dados do sinal íntegro, no entanto, os testes do modelo foram efetivados considerando atributos dos sinais com a presença de contaminantes, conforme representado pelo diagrama da Figura 18.

Figura 18 - Diagrama de blocos das duas categorias de avaliação utilizadas.



Fonte: Autor

Adicionalmente, no experimento sEMG1 foi utilizada uma metodologia de validação cruzada dos resultados chamada de *k-fold* (WEBB; COPSEY, 2011) pelo fato de existir uma limitação no número de amostras do experimento. Esse método é utilizado quando a quantidade de dados da amostra é pequena e a utilização de uma validação convencional (em que os dados são separados em 2 categorias de treino e teste) pode não representar de maneira realista o desempenho do modelo. O método *k-fold* se baseia em dividir os dados em 2 grupos de tamanhos pré-definidos, porém realiza todas as combinações possíveis de divisões das

amostras nesses grupos. Assim como nos outros experimentos, é utilizado 50% dos dados para categoria de treino e 50% para categoria de teste, representados por 3 amostras. Portanto, utilizando o método *k-fold* foram possíveis 20 combinações distintas das duas categorias. Para o resultado final da validação cruzada, é calculada a média dos resultados de todas as possíveis combinações.

Com relação à métrica de avaliação dos métodos, foi considerada a Taxa de Acerto Média (TAM) calculada segundo a Equação (22):

$$\%TAM = \sum_{j=1}^c \frac{TM_j}{C}, \quad TM_j = \frac{CC_j}{N_j}, \quad (22)$$

onde C representa o número total de classes, j representa uma classe específica, N indica o número total de amostras para uma classe j e CC é o número total de amostras classificadas corretamente para uma classe j .

Para os testes que envolveram os sinais com contaminantes também foi calculado, além da Taxa de Acerto Média, o Decréscimo na Taxa de Acerto (DTA) em relação à classificação TAM dos sinais íntegros. O DTA foi calculado segundo a Equação (23):

$$\%DTA = \frac{TAM - TAMD}{TAM}, \quad (23)$$

em que TAM é a Taxa de Acerto Média, considerando os dados íntegros, e $TAMD$ é a Taxa de Acerto Média para os dados com degradação.

Além dos testes realizados nos algoritmos desenvolvidos nesta tese, as bases de dados também foram classificadas utilizando métodos clássicos na área de *machine learning* para fins de comparação dos modelos. Similarmente, foram utilizados exatamente os mesmos conjuntos de treino e teste para a classificação dos dados usando os métodos clássicos na área de reconhecimento de padrão. Os métodos escolhidos para comparação são o SVM, LDA, KNN e *Random Forest* pela sua utilização em larga escala, especialmente na área de reconhecimento de movimentos do membro superior através de sinais mioelétricos (ALKAN; GÜNAY, 2012; ATZORI *et al.*, 2015; DE MOURA; BALBINOT, 2018; GUO *et al.*, 2015; OSKOEI; HU, 2008; PHINYOMARK *et al.*, 2013; RAURALE, 2014; ROSSI *et al.*, 2015; TOSIN *et al.*, 2017). Na Tabela 16 são detalhados os métodos utilizados e seus respectivos

parâmetros principais, cujas funções computacionais foram fornecidas pela ferramenta Matlab 2018.

Tabela 16 - Parâmetros dos métodos de classificação utilizados para fins de comparação.

Classificadores	Sigla	Parâmetros
<i>Random Forest</i>	RF	Número de árvores de decisão = 100
<i>Linear Discriminant Analysis</i>	LDA	-
<i>Support Vector Machines</i>	SVM	Função <i>kernel</i> = Gaussiana Método multiclasse = <i>One-versus-one</i>
<i>k- Nearest Neighborhood</i>	KNN	Número de vizinhos = 10 Métrica de distância = Euclidiana

4.4 PROJETO DE EXPERIMENTOS PROPOSTO

Com o objetivo de realizar afirmações sobre os resultados obtidos em forma de Taxa de Acerto do classificador é realizada uma análise estatística. Para validação estatística da metodologia de testes foi utilizado o “Projeto e Análise de Experimentos de Três Fatores - Modelo de Efeitos Fixos de Três Fatores” (MONTGOMERY, 2012). Dessa forma, as diferenças das médias obtidas entre os diferentes métodos e parâmetros podem ser analisadas com uma confiabilidade de 95%, com base na metodologia de projeto de experimentos do tipo fatorial completo.

Para cada experimento analisado foi calculada a potência estatística do modelo que garante o número mínimo de amostras para garantir uma confiabilidade de 95% na análise dos resultados. No entanto, somente o cálculo da potência estatística não assegura que as análises entre os fatores possam ser consideradas significativas. Para isso é necessária a validação do modelo gerado, o que é imprescindível para garantir algumas condições: as observações devem estar adequadamente descritas pelo modelo, os resíduos devem ser distribuídos normalmente e de forma independente e a variância deve ser constante (MONTGOMERY, 2012).

Diante dessas premissas, somente após garantir que a potência estatística do modelo está adequada, com número de amostras suficiente e que o modelo do projeto de experimentos está validado em uma distribuição normal é possível avaliar se a diferença das médias dos fatores analisados é significativa. A ferramenta computacional utilizada para a análise estatística foi Minitab V. 17. Entretanto, a mesma apresenta limitações quanto ao número de

interações possíveis (2000 interações) quando é utilizado mais de um fator. Por esse motivo se fez necessário o cálculo de mais de um Projeto Fatorial Completo para que seja possível analisar de forma mais abrangente os fatores controláveis e suas interações. Em todos os projetos de experimentos foi utilizada como variável de resposta a Taxa de Acerto Média de classificação de movimentos. O projeto de experimento fatorial completo de três fatores proposto possui o seguinte modelo de efeitos (24):

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_k + (\tau\beta)_{ij} + (\tau\gamma)_{ik} + (\beta\gamma)_{jk} + (\tau\beta\gamma) + \epsilon_{ijkl} \quad (24)$$

onde, y_{ijkl} corresponde a resposta de nível ijk na repetição l , μ corresponde ao efeito médio geral, τ_i é o efeito do i -ésimo nível do fator A, β_j é o efeito do j -ésimo nível do fator B, γ_k é o efeito do k -ésimo nível do fator C, $(\tau\beta)_{ij}$ é o efeito da interação entre A e B, $(\tau\gamma)_{ik}$ é o efeito da interação entre A e C, $(\beta\gamma)_{jk}$ é o efeito da interação entre B e C, e ϵ_{ijkl} é um componente de erro aleatório tendo uma distribuição normal com média zero e variância σ^2 . Para maiores detalhes sobre os procedimentos matemáticos utilizados nos cálculos de Projetos de Experimentos recomenda-se a consulta a (MONTGOMERY, 2012).

5 RESULTADOS E DISCUSSÕES

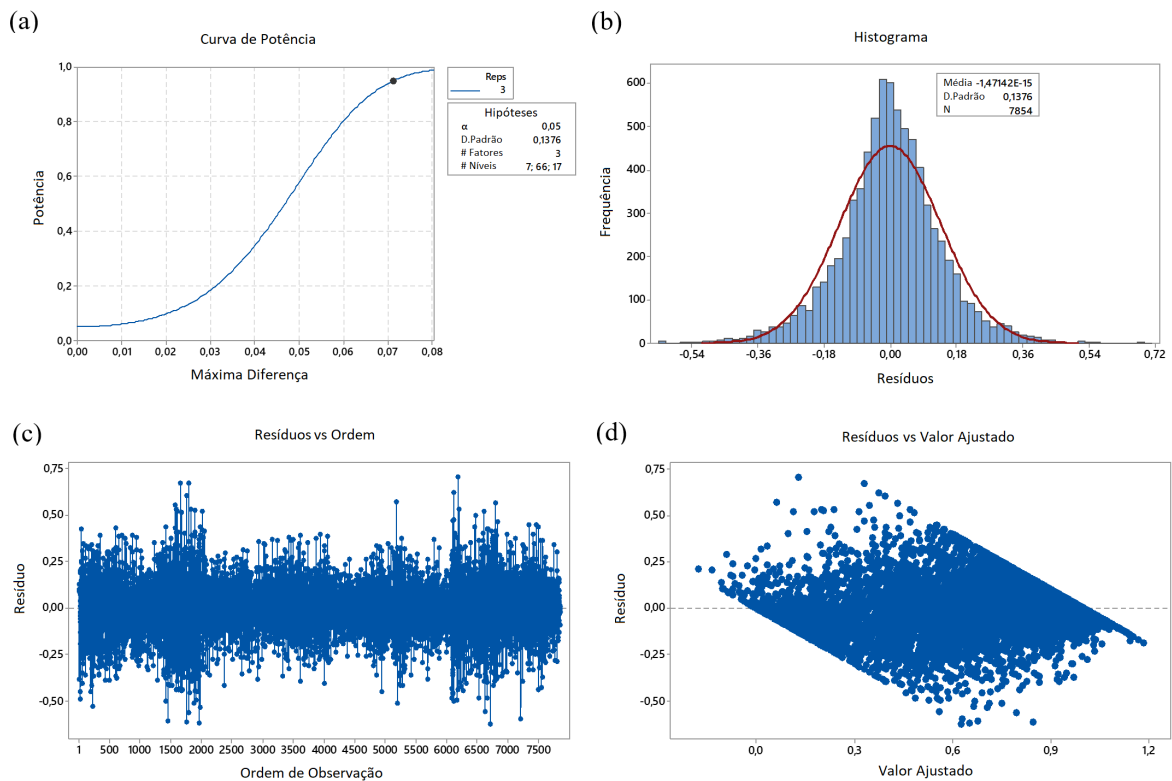
Neste capítulo são apresentados os resultados e discussões dos algoritmos de reconhecimento de padrão desenvolvidos nesta tese. Primeiramente, é feita uma avaliação independente das contribuições de cada um dos métodos propostos de acordo com a metodologia de teste descrita anteriormente. Nessa avaliação individual é efetuada uma análise estatística e os resultados são discutidos. Em um segundo momento é realizada uma discussão envolvendo os dois métodos propostos e estudos similares conduzidos por outros pesquisadores, de forma a examinar a contribuição do uso da Lógica Paraconsistente no contexto de classificação de movimentos do membro superior, através do uso de sinais mioelétricos, considerando o impacto de contaminantes no sinal de sEMG na eficiência do classificador.

5.1 RESULTADOS E DISCUSSÕES DO EXPERIMENTO sEMG 1

O experimento sEMG 1 compreende a análise preliminar do uso da Lógica Paraconsistente na classificação de movimentos do membro superior baseada em sinais de sEMG. Em uma primeira análise do experimento, foi analisada a potência estatística do mesmo de acordo com a metodologia de projeto “Modelo de Efeitos Fixos de Três Fatores” (MONTGOMERY, 2012). O projeto de experimento elaborado utiliza como variável de resposta a taxa de acerto média e como fatores controláveis os movimentos realizados, sujeitos e características de entrada do método. Sendo assim, avaliando a potência estatística do projeto de experimentos realizado, é possível afirmar que, considerando uma potência de 95%, a máxima diferença das médias calculada é de 7,12%, conforme pode ser visualizado na Figura 19(a).

Conforme esclarecido anteriormente, no capítulo de metodologia, somente o cálculo da potência estatística não garante que as análises entre os fatores possam ser consideradas significativas. Para isso é necessário a validação do modelo gerado realizando uma análise dos resíduos. Efetuando uma análise do histograma de resíduos, representado na Figura 19 (b), é correto concluir que os resíduos possuem distribuição normal, com média aproximadamente em zero, sem a presença de *outliers*. Portanto, pode-se considerar que os resultados não são afetados por desvios da normalidade e que os dados são satisfatórios para validação do modelo do projeto de experimentos.

Figura 19 - (a) Curva de potência para o projeto de experimentos, (b) histograma dos resíduos e análise dos resíduos em relação ao (c) ordem de observação e (d) valor ajustado.



Fonte: Autor

Outra condição importante que necessita ser avaliada é se os resíduos estão distribuídos aleatoriamente e de forma independente. Dessa forma, observando o gráfico de resíduos apresentado na Figura 19 (c) não é possível notar nenhum padrão de comportamento, sendo os resíduos uniformemente distribuídos de maneira aleatorizada e com amplitude constante. Na Figura 19 (d) é apresentado o gráfico dos resíduos com o valor da resposta ajustado. Observa-se nesse gráfico que a variância não é homogênea, um comportamento esperado, uma vez que a variável de resposta é limitada entre os valores de 0% a 100%. No entanto, a violação da suposição de variância homogênea em casos em que a ANOVA é balanceada não invalida o modelo, uma vez que os resultados do teste F são apenas levemente afetados (MONTGOMERY, 2012). Portanto, como o projeto de experimentos gerado é balanceado (todos os fatores contêm a mesma quantidade de observações) pode-se considerar o modelo válido para a análise estatística do projeto de experimentos.

Dessa forma, após a garantia de que a potência estatística está adequada, com número de amostras suficiente, e que o modelo do projeto de experimentos está validado em uma

distribuição normal, é possível avaliar se a diferença das médias dos fatores analisados é significativa, conforme os resultados da ANOVA apresentada na Tabela 17. A análise de variância para o experimento de três fatores com confiabilidade de 95% mostrou que os principais efeitos devido à influência dos sujeitos, características e movimento de entrada são significativos, como representado na Tabela 17, na qual o valor do teste P é inferior a 0,05. Portanto, é possível afirmar que as taxas de classificação baseadas no conjunto de atributos de entrada para cada um dos 66 sujeitos e 17 movimentos são significativamente distintas uma da outra e, por consequência, os sinais mioelétricos também são distintos e podem ser tratados como canais distintos pela Rede de Análise Paraconsistente desenvolvida. Além disso, os resultados desse modelo mostraram que as interações entre os fatores também são significativas.

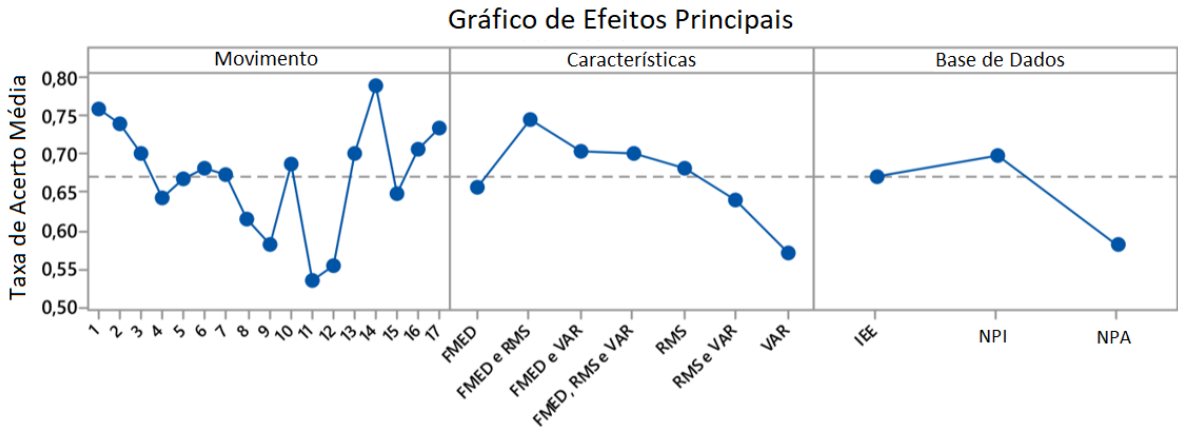
Tabela 17 - Análise da variância (experimento de 3 fatores).

Fatores [†]	Graus de liberdade	Soma dos quadrados	Média Quadrada	F	P
A	6	21,232	3,53865	148,43	0,000
B	65	95,020	1,46185	61,32	0,000
C	16	36,477	2,27981	95,63	0,000
A*B	390	13,698	0,03512	1,47	0,000
A*C	96	4,424	0,04608	1,93	0,000
B*C	1040	382,065	0,36737	15,41	0,000
Erro	6240	148,763	0,02384		
Total	7853	701,679			

[†] Fator A: características de entrada, Fator B: Sujeitos e Fator C: Movimentos realizados

Analisando o gráfico de efeitos principais, representado na Figura 20, pode-se observar que os atributos de entrada que obtiveram a maior taxa de acerto média, superior a 70%, são compostos pela combinação das características Frequência Mediana (fMED) e RMS do sinal de sEMG. Além disso, a base de dados NPA, que compreende os sinais de sEMG adquiridos dos sujeitos com amputação parcial do membro superior pelo projeto Ninapro, apresenta menor taxa de acerto média, inferior a 60%, em comparação com as bases de dados com voluntários com membro superior intacto (NPI e IEE). Em relação às classes de movimento, os movimentos de supinação e pronação do punho (classes 11 e 12) representaram a menor taxa de acerto média (<55%), considerando todas as bases de dados e características.

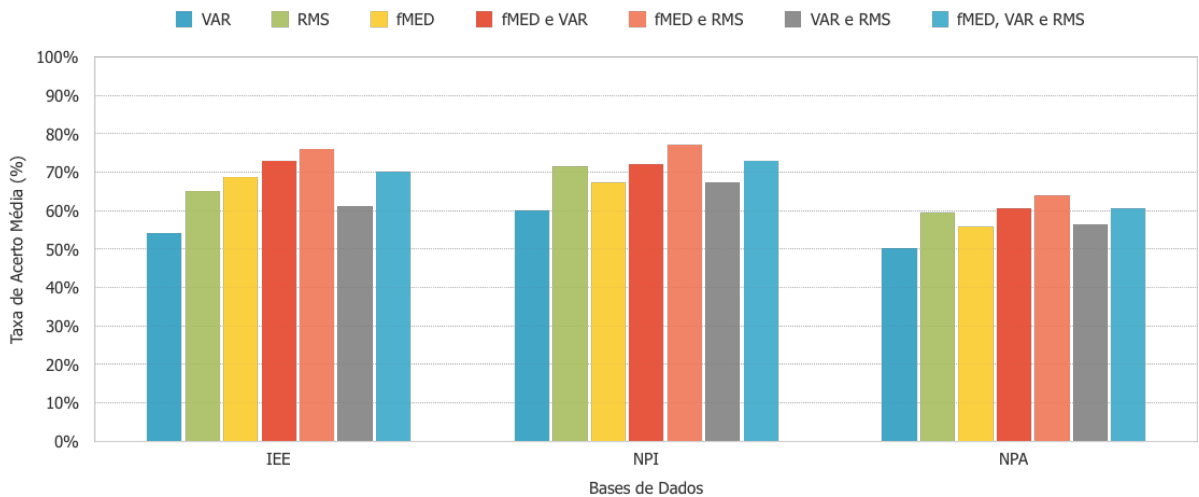
Figura 20 - Gráfico de efeitos da taxa de classificação média para movimentos, conjunto de características e base de dados.



Fonte: Autor

Para uma melhor visualização dos resultados é apresentado na Figura 21 um gráfico comparativo da taxa de acerto média para cada base de dados utilizada neste experimento, contemplando as diferentes combinações de atributos de entrada.

Figura 21 - Gráfico de comparação das taxas de acerto médias obtidas para cada base de dados e para cada combinação de características do sinal.



Fonte: Autor

Analisando a Figura 21 pode-se perceber que os resultados envolvendo as características fMED e RMS apresentaram as melhores taxas de acerto média em todas as bases de dados utilizadas, sendo que as bases de dados com voluntários com membros intactos (NPI e IEE) obtiveram taxas de classificação médias de aproximadamente 76,0% e 77,0%,

respectivamente. A base de dados NPA, com sujeitos com amputação do membro superior, para essa mesma condição de entrada, obteve uma taxa de acerto média de 63,9%. Por outro lado, a característica variância apresentou os piores resultados em termos de taxa de acerto média para todas as bases de dados, com taxa de acerto média de 54,2% e 60% para as bases de dados com sujeitos com membro superior intacto IEE e NPI, respectivamente, e uma taxa de acerto média de 50,2% para a base de dados NPA.

Analisando-se a diferença das taxas de acerto médias através do método de Tukey, apresentado na Tabela 18, é possível afirmar, com um nível de confiabilidade de 95%, que as características de entrada fMED e RMS apresentaram resultados médios significativamente superiores que todas as outras características de entrada utilizadas nesse experimento, exceto pela combinação fMED e VAR que, de acordo com essa análise estatística, não apresentou uma diferença de média significativa. Portanto, pode-se concluir que as características de entrada que obtiveram as melhores taxas de acerto médias foram ‘fMED e RMS’ e ‘fMED e VAR’. No entanto, para um estudo mais aprofundado do método, foi escolhido o conjunto de entrada ‘fMED e RMS’ pelo fato de que a característica RMS do sinal de sEMG é amplamente utilizada na área de classificação de movimentos do membro superior como pode-se verificar em alguns dos seguintes trabalhos (BALBINOT; FAVIEIRO, 2013; CENE; FAVIEIRO; BALBINOT, 2015; CENE; BALBINOT, 2018; FAVIEIRO; BALBINOT, 2011a; GUO *et al.*, 2015; MOMEN; KRISHNAN; CHAU, 2007; MOURA; FAVIEIRO; BALBINOT, 2016; RIILLO *et al.*, 2014; TOSIN *et al.*, 2017).

Tabela 18 - Análise das médias das características utilizando método Tukey (95% de confiabilidade).

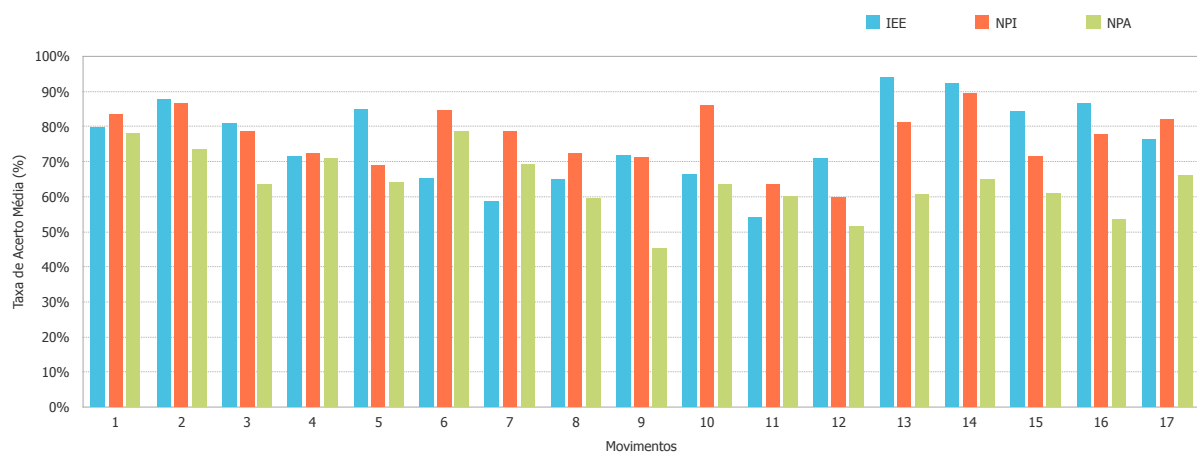
Características	Amostras	Médias	Grupo [†]
fMED e RMS	1122	72,3%	A
fMED e VAR	1122	68,5%	A B
fMED, RMS e VAR	1122	67,9%	B C
RMS	1122	65,4%	B C D
fMED	1122	63,9%	C D
VAR e RMS	1122	61,6%	D
VAR	1122	54,8%	E

[†] Médias que não compartilham uma letra são significativamente diferentes.

Em uma segunda análise do método RAP foram estudados os resultados em termos de taxa de acerto média por movimento, para cada uma das bases em relação ao uso das características de fMED e RMS como entrada da rede, conforme representado na Figura 22.

Observando o gráfico, verifica-se que existe uma diferença entre as taxas de acerto dos movimentos para as diferentes bases de dados. Conforme comentado anteriormente, as médias das taxas de acerto para as características fMED e RMS para as bases de dados com voluntários com membro superior intacto (IEE e NPI) obtiveram resultados próximos (76,0% e 77,0% respectivamente). No entanto, nota-se claramente na Figura 22 que os movimentos com as melhores e piores taxas de acerto médias são referentes a movimentos distintos. Analisando a base de dados IEE o movimento com a maior taxa de acerto média é o movimento 13 (flexão do punho) com 94,4%, e o movimento com a menor taxa de acerto é o movimento 11 (supinação do punho) com 54,4%. Adicionalmente, analisando a base de dados NPI, os movimentos com maior e menor taxa de acerto média são os movimentos 14 (extensão do punho) e 12 (pronação do punho), com taxas de acerto médias de 89,6% e 71,5%, respectivamente. Logo, é possível constatar que os movimentos de extensão e flexão do punho são fáceis de executar e apresentam uma ativação muscular bem característica, facilitando sua classificação correta. Por sua vez, os movimentos de supinação (movimento 11) e pronação (movimento 12) do punho são movimentos com sinais similares em relação aos movimentos 9 e 10 e, portanto, com baixa discriminação, considerando-se as características empregadas no modelo proposto, o que pode ocasionar um erro de classificação entres os mesmos.

Figura 22 - Comparativo de taxas de acerto entres as bases de dados, discriminadas por movimento, utilizando as características fMED e RMS do sinal como entrada.



Fonte: Autor

É interessante ressaltar que, na aquisição dos sinais de sEMG da base IEE, foi constatada uma dificuldade na execução dos movimentos 9, 10, 11 e 12 que envolvem a pronação e supinação do punho, uma vez que a diferença entre os mesmos é o eixo de rotação. Essa dificuldade é claramente observada nos resultados em termos de taxa de acerto média,

pois, ao realizar a análise da matriz de confusão, é possível avaliar que os erros de classificação ocorrem, em sua maioria, entre esse grupo de movimentos, suportando a percepção observada nos ensaios, tanto da dificuldade de execução correta dos mesmos por parte dos voluntários, como a dificuldade do classificador em diferenciá-los.

Utilizando as características de entrada fMED e RMS também foi realizado um experimento que inclui a comparação da RAP com outros métodos de reconhecimento de padrão. Os métodos escolhidos foram LDA, RF, KNN e SVM por seu extenso uso na área de classificação de movimentos através de sinais mioelétricos (ALKAN; GÜNAY, 2012; ATZORI *et al.*, 2015; FRAIMAN; JUSTEL; SVARC, 2010; MOURA; FAVIEIRO; BALBINOT, 2016; NAIK; AL-TIMEMY; NGUYEN, 2016; TOSIN *et al.*, 2017; ZHANG; HUANG, 2015). O experimento considerou como fatores controláveis os métodos e as bases de dados. A variável de resposta foi a taxa de acerto média. O experimento foi validado, em uma distribuição normal, de acordo com a análise dos resíduos, e considerando uma potência de 95% a máxima diferença das médias calculada foi de 2,5%, conforme gráficos apresentados na Figura 35 do APÊNDICE A. Na Tabela 19 encontra-se a análise da variância do experimento, na qual se observa que os fatores métodos de classificação e bases de dados são considerados significativos de maneira independente (valor de $P < 0,05$). No entanto, ao se considerar a interação dos fatores, as taxas de acerto médias não apresentam diferenças significativas.

Tabela 19 - Análise da variância considerando como fatores métodos e bases de dados.

Fatores [†]	Graus de liberdade	Soma dos quadrados	Média Quadrada	F	P
A	4	68,005	17,0012	325,78	0,000
B	2	22,666	11,3329	217,16	0,000
A*B	8	0,738	0,0922	1,77	0,079
Erro	6343	331,021	0,0522		
Total	6357	436,020			

[†] Fator A: Métodos de classificação e Fator B: Base de dados

Com o intuito de avaliar a diferença das taxas de acerto médias dos métodos de classificação selecionados neste experimento, foi realizada uma análise das médias utilizando o método Tukey com 95% de confiabilidade, conforme resultados apresentados na Tabela 20. A partir dessa análise confirma-se que o método LDA obteve a melhor taxa de acerto média de 84,4%, com diferença significativa para os outros métodos testados. O método proposto

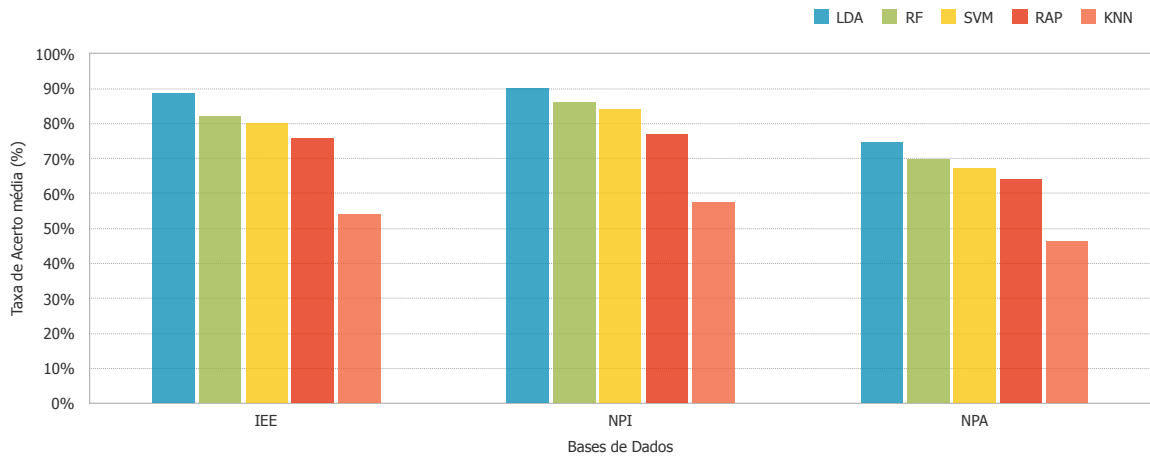
RAP obteve uma média de 72,3%, que é considerada inferior aos métodos RF e SVM, que obtiveram uma taxa de acerto média de 79,3% e 77,1%, respectivamente. Adicionalmente, a análise das médias revelou que os métodos RF e SVM obtiveram taxas de acerto médias não significativas entre si, ou seja, não é possível afirmar que o método RF é superior ao SVM, em termos de taxa de acerto média, neste experimento (considerando um nível de confiança de 95%). O método KNN obteve a menor taxa de acerto média desse experimento atingindo o valor de 52,5%. A taxa de acerto média do método KNN pode ser explicada pela baixa amostragem de dados para treinamento do modelo, uma vez que o método KNN utilizado considerou como métrica de similaridade a distância Euclidiana, que é calculada entre os dados de treinamento e teste para realizar a tomada de decisão e, devido à baixa quantidade de amostras, o número k de vizinhos considerados no método é limitado.

Tabela 20 - Análise das médias dos métodos utilizando o método de Tukey (95% de confiabilidade).

Métodos	Amostras	Médias	Grupo [†]
LDA	1309	84,4%	A
RF	1309	79,3%	B
SVM	1309	77,1%	B
RPA	1309	72,3%	C
KNN	1309	52,5%	D

[†] Médias que não compartilham uma letra são significativamente diferentes

Na Figura 23 são apresentadas as taxas de acerto médias considerando os diferentes métodos de classificação e as bases de dados. Nesse gráfico é possível verificar que as taxas de acerto médias dos métodos seguem um mesmo padrão para as diferentes bases de dados, sendo que a base de dados NPA, que abrange os sinais de sEMG de sujeitos com amputação parcial do antebraço, é a base de dados que apresenta a menor taxa de acerto média. Também observa-se no gráfico que o método LDA obteve a maior taxa de acerto média para todas as bases de dados, com taxas de 90,1%, 88,5% e 74,70% para as bases NPI, IEE e NPA respectivamente. O método RAP proposto obteve taxas de acerto médias aproximadas de 76% para a base de dados IEE, 77% para a base de dados NPI e 63% para bases de dados NPA. Também é possível observar que as taxas de acerto médias do método RAP para as bases de dados IEE e NPA se aproximam do método SVM, que atingiu valores médios de 80% e 67%, respectivamente. Sendo assim, a diferença da taxa de acerto média para os métodos SVM e RAP, para essas bases, foi de aproximadamente de 4%.

Figura 23 - Taxa de acerto média considerando os diferentes métodos e bases de dados.

Fonte: Autor

É importante ressaltar que, apesar do método RAP apresentar resultados inferiores aos métodos LDA, RF e SVM, o mesmo é um modelo mais simplificado em comparação aos outros, uma vez que a RAP não possui nenhuma etapa de treinamento de forma a otimizar seus parâmetros, sendo considerado um modelo fixo de classificação de dados. Adicionalmente, considerando que a RAP é um modelo preliminar de análise da Lógica Paraconsistente na aplicação de reconhecimento de movimentos do membro superior através de sinais de sEMG, pode-se assegurar que os resultados obtidos em termos de taxa de acerto média são promissores para continuar a investigação do uso da Lógica Paraconsistente nessa área.

Além da taxa de acerto, a preocupação com o desempenho e latência do sistema, ao abordar o tema de controle de próteses mioelétricas, está cada vez mais em evidência, uma vez que a usabilidade da prótese é de extrema importância, conforme estudos realizados por (FARRELL, 2011). Dessa forma, foi realizada uma avaliação do tempo de processamento do algoritmo RAP para cada amostra de movimento, em um conjunto de 20.400 amostras. Os resultados obtidos em ms e seus respectivos desvios padrão estão representados na Tabela 21. De acordo com (FARRELL, 2011), a latência do sistema deve ser de no máximo 100 ms, para não afetar o desempenho da prótese pelos usuários. Na Tabela 21 também é possível observar que existe uma proporção, aproximadamente linear, entre o tempo de processamento e o número de características de entrada, na qual pode-se concluir que a RAP tem uma complexidade computacional $O(n)$.

Tabela 21 - Análise do tempo de processamento de uma amostra para diferentes topologias da RAP.

Topologia da RAP	Tempo de processamento (ms)
3 características	20,6 ± 0,3
2 características	14,5 ± 0,2
1 característica	7,9 ± 0,3

Considerando uma abordagem de processamento paralelo para a topologia de 2 características de entrada, o tempo de processamento poderia ser reduzido para aproximadamente 1 ms por amostra, uma vez que existem 17 redes de análise independentes entre si, em termos de cálculos matemáticos, uma para cada classe de movimento.

Em relação à RAP proposta, em paralelo a outros estudos reportados na literatura, é importante ressaltar que a comparação de métodos de classificação na área da sEMG se torna complicada pelo fato de cada estudo usar diferentes classes de movimento, conjunto de características, quantidade e posicionamento de eletrodos, entre outros. Por exemplo, considerando-se o banco de dados Ninapro, a pesquisa proposta por (GIJSBERTS *et al.*, 2014) avaliou o banco de dados utilizando 40 sujeitos saudáveis e 40 classes de movimento e obteve uma taxa de acerto de classificação de 72,2% a 75,41% (dependendo do atributo de entrada), valores próximos às taxas de acerto calculadas para a RAP. Além disso, em (GIJSBERTS *et al.*, 2014) foi calculado o atraso total do sistema, que atingiu entre 300 e 1200 ms (dependendo de um parâmetro de suavização k , representando o tamanho da janela deslizante). Neste estudo, verificou-se que a taxa de erro aumentou significativamente, acima de 20%, para $k = 1$ e, neste caso, os níveis de atraso foram os maiores observados, acima de 300 ms.

No estudo proposto por (LI; SCHULTZ; KUIKEN, 2010), cinco sujeitos com amputação transradial realizaram dez movimentos do punho e da mão com uma taxa de acerto de classificação de 84,4% e um tempo médio de classificação do movimento de aproximadamente 190 ms. O classificador utilizado foi o LDA e o tempo computacional para cada janela de análise foi de aproximadamente 3 ms. Outro classificador bem conhecido na área de classificação de movimentos baseado em sinais de sEMG é o método SVM. A investigação proposta por (MOURA; FAVIEIRO; BALBINOT, 2016) utilizou SVM para classificar 17 classes de movimento do punho e mão com uma taxa de acerto média de 80,8 a 85,1% para sujeitos saudáveis, dependendo da função *kernel*. A metodologia experimental apresentada em (MOURA; FAVIEIRO; BALBINOT, 2016) também foi baseada no projeto Ninapro, e a característica de entrada foi o RMS. Como observado, em relação à taxa de acerto

da classificação, o RAP obteve resultados comparáveis às taxas de acerto encontradas em outras pesquisas utilizando sinais de sEMG para classificação de movimento de membros superiores.

No que se refere às limitações do algoritmo proposto, o trabalho enfatiza que a proposta da RAP não visa resolver os problemas em aberto na área de classificação de movimentos através de sinais de sEMG, como o controle natural e intuitivo, mas sim realizar uma investigação preliminar da Lógica Paraconsistente aplicada na área de reconhecimento de movimentos. Considerando essa premissa, os resultados obtidos em termos de taxa de acerto e tempo de processamento se mostraram promissores para investir em uma abordagem que utilize conceitos da Lógica Paraconsistente e aprendizagem de máquina.

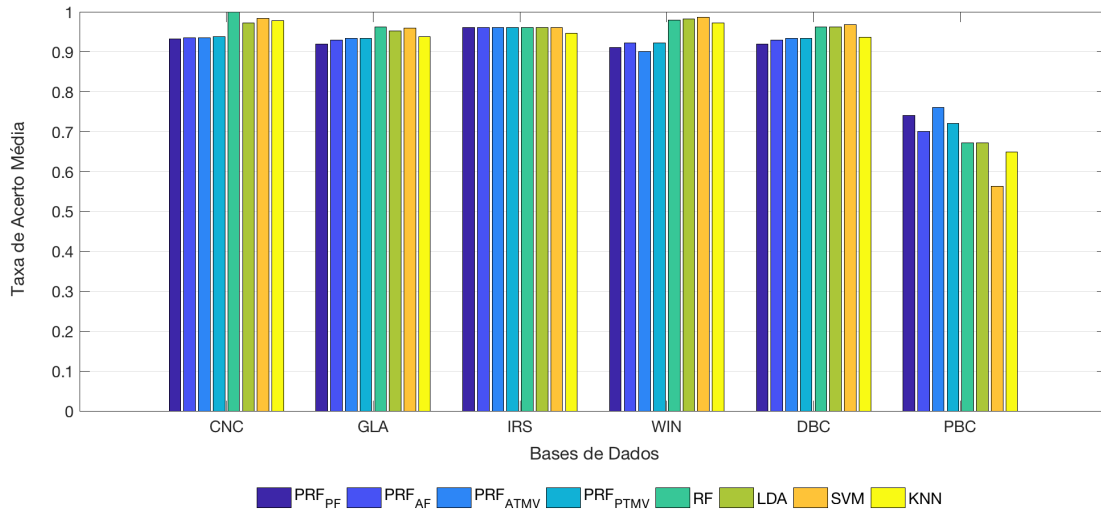
5.2 RESULTADOS E DISCUSSÕES DO EXPERIMENTO BDP

O experimento BDP compreende a análise do algoritmo *Paraconsistent Random Forest* (PRF) desenvolvido utilizando bases de dados padrão na área de aprendizagem de máquina. O objetivo desse experimento é avaliar a viabilidade do algoritmo na classificação de dados padrão ouro e comparar resultados com outros métodos de reconhecimento de padrão bem fundamentados na literatura. A avaliação de algoritmos empregando bases de dados padrão é de extrema importância, pois permite a comparação de resultados com outras pesquisas da comunidade científica.

Em uma primeira análise, a taxa de acerto média (TAM) foi calculada para todos os algoritmos de tomada de decisão propostos para a *Paraconsistent Random Forest*. A TAM é demonstrada na Figura 24 em comparação com outros algoritmos selecionados (LDA, KNN, SVM e RF) e seis bancos de dados padrão selecionados do repositório UCI (LICHMAN, 2013), conforme descrito previamente na metodologia de testes. Ao avaliar os resultados do método PRF com aqueles obtidos por um conjunto de classificadores, foi possível observar que o método PRF obteve taxas de acerto médias próximas aos classificadores SVM, LDA, KNN e RF, conforme apresentado na Figura 24. A maior diferença nas taxas de acerto foi observada para base de dados WIN na qual o método PRF apresentou uma taxa de acerto média de aproximadamente 93% contra uma taxa de 98% para os métodos SVM e LDA. Sendo assim, considerando todas as bases, as diferenças máximas nas taxas de acerto, comparando o método PRF com os outros classificadores, foi de no máximo 5%. Também é interessante observar que, para a base de dados PBC, o método PRF_{ATMV} apresentou os

melhores resultados, na comparação com os outros métodos, com uma taxa de acerto média de aproximadamente 76%.

Figura 24 - Comparação da taxa de classificação média para as bases de dados selecionadas e o método PRF proposto.



Fonte: Autor

Tabela 22 - Análise da variância considerando como fatores métodos, bases de dados e a% de dados ausentes.

Fatores [†]	Graus de liberdade	Soma dos quadrados	Média Quadrada	F	P
A	7	2,5729	0,36756	99,00	0,000
B	5	5,3889	1,07778	290,30	0,000
C	6	1,4001	0,23334	62,85	0,000
A*B	35	1,5070	0,04591	12,37	0,000
A*C	42	0,5170	0,01231	3,32	0,000
B*C	30	0,7165	0,02388	6,43	0,000
Erro	210	0,7797	0,00371		
Total	335	12,9820			

[†] Fator A: Métodos de classificação, Fator B: Base de dados e Fator C: a% de dados ausentes

Com o intuito de verificar a habilidade do algoritmo em lidar com dados não ideais, foi realizado um experimento no qual uma porcentagem (a%) de valores ausentes foi introduzida na base de dados. É importante ressaltar que os modelos de classificação foram criados utilizando toda a informação disponível do conjunto de treino e os valores ausentes só foram introduzidos no conjunto de teste. Para visualizar o impacto na predição do modelo, dependendo da porcentagem (a%) de dados ausentes definida, foi calculado o decréscimo na taxa de acerto média. Em relação à análise estatística, o projeto de experimentos fatorial

completo considerou como fatores controláveis os métodos, as bases de dados e a $\alpha\%$ de dados ausentes. A variável de resposta foi a taxa de acerto média. O experimento foi validado, em uma distribuição normal, de acordo com a análise dos resíduos, e considerando uma potência de 95% a máxima diferença das médias calculada foi de 10,8%, conforme gráficos apresentados na Figura 36 do APÊNDICE A. Na Tabela 22 encontra-se a análise da variância do experimento, na qual percebe-se que os fatores métodos de classificação, bases de dados e $\alpha\%$ de dados ausentes são considerados significativos de maneira independente (valor de $P < 0,05$). As interações entre os fatores também foram consideradas significativas.

A Tabela 23 representa o decréscimo na taxa de acerto média (DTA), ao se introduzir uma porcentagem de dados ausentes, para o algoritmo PRF e as suas quatro variações dos algoritmos de tomada de decisão. Observa-se que, no pior caso, com 90% dos dados ausentes, o valor máximo de decréscimo DTA foi somente de 19,45% para a base de dados IRS considerando a abordagem ATMV. Observando o caso de 5% de valores ausentes, o algoritmo PRF foi capaz de manter a taxa de acerto média, sem decréscimo significativo na classificação dos dados.

Tabela 23 - Decréscimo na taxa de acerto média na presença de dados ausentes.

Base de dados	Método	% decréscimo na taxa de acerto					
		a% de dados ausentes adicionados					
		5%	10%	30%	50%	70%	90%
IRS	PRF _{PF}	0,00%	2,82%	7,05%	12,68%	22,54%	21,13%
	PRF _{AF}	1,39%	5,55%	9,72%	11,11%	20,83%	20,83%
	PRF _{ATMV}	0,00%	2,78%	5,55%	11,11%	20,83%	19,45%
	PRF _{PTMV}	1,39%	4,17%	4,17%	8,33%	19,45%	22,22%
PBC	PRF _{PF}	0,00%	1,35%	4,05%	0,00%	0,00%	8,11%
	PRF _{AF}	1,43%	5,71%	1,43%	10,00%	10,00%	8,57%
	PRF _{ATMV}	1,32%	2,63%	7,89%	11,84%	14,47%	11,84%
	PRF _{PTMV}	0,00%	0,00%	1,39%	12,50%	11,11%	5,56%
DBC	PRF _{PF}	0,76%	0,76%	1,52%	2,68%	8,02%	6,49%
	PRF _{AF}	2,26%	1,51%	1,88%	5,66%	12,08%	9,81%
	PRF _{ATMV}	0,38%	0,00%	0,00%	-0,76%	4,51%	3,38%
	PRF _{PTMV}	0,00%	-0,38%	-0,38%	0,38%	6,39%	4,89%
GLA	PRF _{PF}	2,02%	-2,05%	10,19%	10,19%	26,52%	22,45%
	PRF _{AF}	-2,13%	2,13%	6,38%	10,64%	31,91%	17,02%
	PRF _{ATMV}	0,00%	-1,91%	-5,76%	1,93%	7,69%	13,48%
	PRF _{PTMV}	1,86%	-1,86%	3,69%	0,00%	29,63%	22,22%
WIN	PRF _{PF}	0,00%	0,00%	3,33%	3,33%	11,11%	23,33%
	PRF _{AF}	0,00%	4,87%	8,54%	14,63%	19,51%	26,82%
	PRF _{ATMV}	0,00%	2,41%	8,44%	13,25%	21,69%	25,30%
	PRF _{PTMV}	2,47%	3,70%	6,18%	11,11%	11,11%	24,69%
CNC	PRF _{PF}	0,00%	4,87%	8,54%	14,63%	19,51%	26,82%
	PRF _{AF}	0,00%	2,41%	8,44%	13,25%	21,69%	25,30%
	PRF _{ATMV}	2,47%	3,70%	6,18%	11,11%	11,11%	24,69%
	PRF _{PTMV}	0,00%	1,20%	8,44%	7,22%	12,05%	16,86%

É importante destacar que os valores negativos de decréscimo na taxa de acerto, que aparecem na Tabela 23 representam ganhos na taxa de acerto média em comparação com a taxa de acerto considerando os dados originais. Esses pequenos valores de ganho possivelmente ocorrem em casos em que a substituição aleatória dos dados não abrangeu atributos significativos para o modelo, dessa forma, não ocorrendo decréscimos nas taxas de acerto.

De maneira similar, o decréscimo na taxa de acerto média, na presença de dados ausentes, também foi calculado para outros métodos populares na área de aprendizagem de máquina (SVM, LDA, KNN e RF) e seus resultados estão apresentados na Tabela 24, em comparação com os melhores resultados obtidos do método PRF.

Tabela 24 - Comparação do decréscimo na taxa de acerto média na presença de dados ausentes com os métodos PRF, SVM, KNN e LDA.

Base de dados	Método	% decréscimo na taxa de acerto					
		a% de dados ausentes adicionados					
		5%	10%	30%	50%	70%	90%
IRS	PRF	0,00%	2,78%	4,17%	8,33%	19,45%	19,45%
	RF	0,00%	1,42%	2,82%	7,05%	12,68%	18,32%
	LDA	15,28%	29,17%	45,83%	59,72%	61,11%	63,89%
	SVM	15,28%	30,55%	45,83%	59,72%	61,11%	63,89%
	KNN	14,09%	29,58%	45,07%	59,15%	60,57%	63,38%
PBC	PRF	0,00%	0,00%	1,39%	0,00%	0,00%	5,56%
	RF	-3,92%	8,50%	12,91%	21,24%	23,37%	25,49%
	LDA	22,22%	25,49%	25,49%	25,49%	25,49%	25,49%
	SVM	11,11%	11,11%	11,11%	11,11%	11,11%	11,11%
	KNN	11,15%	23,99%	22,97%	22,97%	22,97%	22,97%
DBC	PRF	0,00%	-0,38%	-0,38%	-0,76%	4,51%	3,38%
	RF	0,36%	0,36%	0,73%	1,81%	5,45%	8,36%
	LDA	27,74%	32,85%	34,67%	34,67%	34,67%	34,67%
	SVM	28,26%	32,97%	35,14%	35,14%	35,14%	35,14%
	KNN	27,71%	32,21%	32,95%	32,95%	32,95%	32,95%
GLA	PRF	0,00%	-1,91%	-5,76%	1,93%	7,69%	13,48%
	RF	0,00%	-12,96%	-5,55%	-11,12%	0,00%	11,10%
	LDA	13,64%	33,33%	40,91%	43,95%	42,43%	42,43%
	SVM	15,88%	27,34%	46,04%	44,45%	44,45%	44,45%
	KNN	17,46%	26,99%	38,10%	39,69%	39,69%	39,69%
WIN	PRF	0,00%	0,00%	3,33%	3,33%	11,11%	23,33%
	RF	0,00%	1,20%	8,44%	7,22%	12,05%	16,86%
	LDA	30,68%	39,77%	54,55%	57,96%	59,09%	59,09%
	SVM	46,19%	59,18%	71,24%	72,17%	72,17%	72,17%
	KNN	31,03%	39,08%	54,03%	57,47%	58,62%	58,62%
CNC	PRF	0,00%	1,20%	6,18%	7,22%	11,11%	16,86%
	RF	0,00%	0,00%	3,33%	3,33%	11,11%	23,33%
	LDA	30,68%	39,77%	54,55%	57,96%	59,09%	59,09%
	SVM	46,19%	59,18%	71,24%	72,17%	72,17%	72,17%
	KNN	31,03%	39,08%	54,03%	57,47%	58,62%	58,62%

Analisando a Tabela 24, é possível afirmar que um valor de 5% de dados ausentes representa uma queda significativa na predição para os métodos SVM, KNN e LDA, com decréscimos na taxa de acerto variando entre 14% a 46%, dependendo da base de dados utilizada. Também se observa que o método RF obteve decréscimos na taxa de acerto mais próximos aos resultados obtidos com o método PRF. Considerando 90% de dados ausentes, o algoritmo PRF só apresenta um decréscimo maior na taxa de acerto média, em comparação com algoritmo RF, para as bases IRS e GLA.

Realizando uma análise das médias pelo método Tukey com 95% de confiabilidade, cujos resultados são apresentados na Tabela 25, é possível afirmar que, considerando a classificação de bases de dados com valores ausentes, tanto o método *Random Forest* quanto o *Paraconsistent Random Forest* apresentaram taxas de acerto médias significativamente superiores em relação aos métodos LDA, KNN e SVM testados neste experimento. Além disso, não foram observadas diferenças significativas nas médias entre os métodos RF e PRF, com taxas de acerto médias de 85,79% e 85,16%, respectivamente. Também pode-se comprovar que os métodos de tomada de decisão propostos para o algoritmo PRF não apresentaram diferença significativa, entre os mesmos, na taxa de acerto média.

Tabela 25 - Análise das médias dos métodos utilizando método Tukey (95% de confiabilidade) para o experimento BDP com dados ausentes.

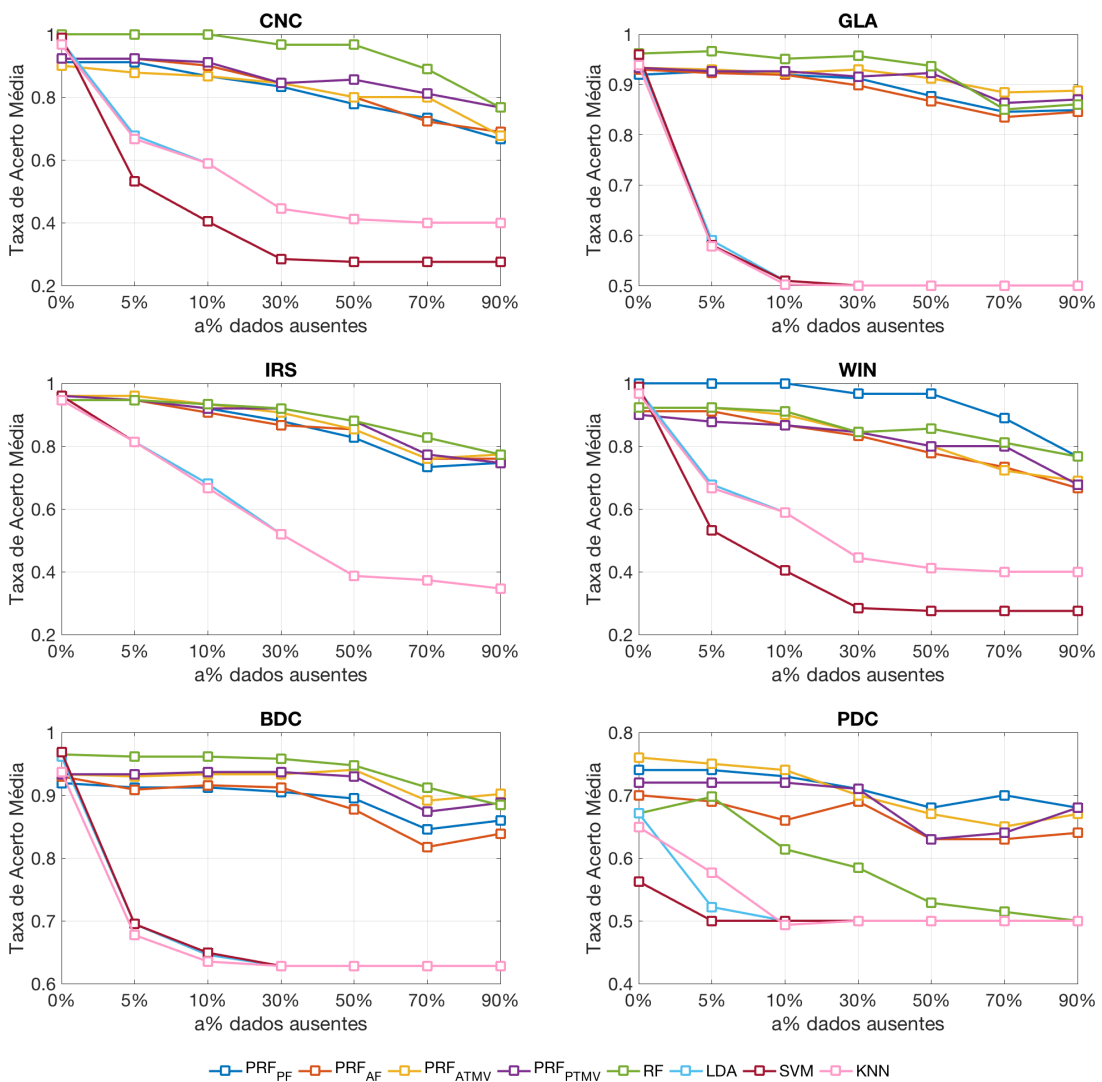
Métodos	Amostras	Médias	Grupo [†]
RF	42	85,79%	A
PRF _{ATMV}	42	85,16%	A
PRF _{PTMV}	42	84,65%	A
PRF _{PF}	42	84,56%	A
PRF _{AF}	42	82,40%	A
LDA	42	60,31%	B
KNN	42	59,73%	B
SVM	42	58,11%	B

[†] Médias que não compartilham uma letra são significativamente diferentes

Outro aspecto importante a considerar é que, conforme detalhado previamente no capítulo de metodologia, na situação de ausência de valores para uma característica ($\mu f_k^c = NaN$), o Grau de Evidência, em relação à característica ausente, é definido como 0,5. Esse valor representa, na Lógica Paraconsistente, o valor lógico de Indefinido. Portanto, para todas as partições de domínio do atributo cujo dado é inexistente, o grau de pertinência associado será o mesmo. No entanto, na abordagem da *Random Forest*, ao se deparar com dados

ausentes, a função de tomada de decisão retorna a classe com o maior número de amostras de treinamento que atingiram o nó que está sendo avaliado. De acordo com os resultados obtidos e análise estatística empregada, as duas abordagens, considerando dados ausentes, tanto do método RF como do método desenvolvido PRF, obtiveram taxas médias sem diferença significativa na taxa de acerto. Considerando essas premissas, pode-se concluir que, para dados ausentes, os dois métodos RF e PRF apresentam uma solução satisfatória, com taxas de acerto médias no mínimo 20% superiores em comparação aos outros métodos testados (SVM, LDA e KNN).

Figura 25 - Gráfico da taxa de classificação em comparação aos métodos PRF, LDA, SVM e KNN com adição de valores inválidos, sendo as bases de dados: CNC, WIN, GLA, IRS, DBC e PBC.



Fonte: Autor

A vantagem na classificação de dados com valores ausentes dos métodos RF e PRF pode ser claramente observada na Figura 25, na qual, para todas as bases de dados (exceto PDC), os dois métodos apresentaram uma curva de predição similar. Também é visível que para os métodos SVM, KNN e LDA uma ausência de 5% dos dados já causa um impacto considerável na taxa de acerto, com uma redução na taxa variando entre 9% a 45%.

Também foi realizado um experimento para avaliar o comportamento do algoritmo diante de dados com ruído. O ruído nesse experimento é caracterizado pela substituição de uma porcentagem ($\alpha\%$) das características de entrada por um valor aleatório de distribuição uniforme. É importante ressaltar que os modelos de classificação foram criados utilizando toda a informação disponível do conjunto de treino, sendo assim, o ruído só foi introduzido no conjunto de teste. Para visualizar o impacto na predição do modelo, dependendo da porcentagem ($\alpha\%$) de dados com ruído definida, foi calculado o decréscimo na taxa de acerto média. Quanto à análise estatística, o projeto de experimentos fatorial completo considerou como fatores controláveis os métodos e α ($\alpha\%$) de dados ausentes. A variável de resposta foi a taxa de acerto média. O experimento foi validado, em uma distribuição normal, de acordo com a análise dos resíduos, e considerando uma potência de 95% a máxima diferença das médias calculada foi de 7,7%, conforme gráficos apresentados na Figura 37 do APÊNDICE A. Na Tabela 26 encontra-se a análise da variância do experimento, na qual se observa que os fatores métodos de classificação e α ($\alpha\%$) de dados com ruído são considerados significativos de maneira independente (valor de $P < 0,05$). As interações entre os fatores também foram consideradas significativas.

Tabela 26 - Análise da variância considerando como fatores métodos, bases de dados e $\alpha\%$ de ruído adicionado.

Fatores [†]	Graus de liberdade	Soma dos quadrados	Média Quadrada	F	P
A	7	6,698	0,956902	123,79	0,000
B	7	3,299	0,471332	60,98	0,000
A*B	49	1,565	0,031935	4,13	0,000
Erro	320	2,474	0,007730		
Total	383	14,036			

[†] Fator A: Métodos e Fator B: $\alpha\%$ de ruído adicionado

Na Tabela 27 é representado o decréscimo na taxa de acerto média (DTA), ao se introduzir uma porcentagem de dados com ruído, para o algoritmo PRF e as suas quatro variações dos algoritmos de tomada de decisão. É possível observar que no pior caso, com

90% contaminados com ruído, o valor máximo de decréscimo DTA foi somente de 30,56% para a base de dados IRS, considerando a abordagem ATMV. Observando o caso de 5% de dados com ruído, o algoritmo PRF foi capaz de manter a taxa de acerto média, com o decréscimo máximo encontrado na base de dados IRS com valor de 2,78%. Também é interessante ressaltar que para até 30% de dados com ruído, a taxa máxima de decréscimo foi de apenas 5,56% (considerando a base de dado IRS).

Tabela 27 - Decréscimo na taxa de acerto média na presença de dados com ruído.

Base de dados	Método	% decréscimo na taxa de acerto						
		a% de ruído adicionado						
		5%	10%	15%	30%	50%	70%	90%
CNC	PRF _{PF}	-0,61%	0,61%	1,23%	2,45%	3,07%	6,44%	12,88%
	PRF _{AF}	0,31%	0,00%	1,53%	2,45%	2,45%	5,81%	10,40%
	PRF _{ATMV}	0,31%	0,61%	0,92%	2,45%	1,83%	4,89%	9,79%
	PRF _{PTMV}	0,00%	0,00%	1,22%	1,22%	2,74%	7,01%	13,72%
GLA	PRF _{PF}	0,38%	1,53%	3,05%	4,96%	13,36%	16,79%	18,32%
	PRF _{AF}	0,75%	2,26%	3,02%	5,66%	13,58%	16,60%	18,87%
	PRF _{ATMV}	0,38%	2,26%	2,63%	4,14%	10,90%	14,66%	18,42%
	PRF _{PTMV}	0,00%	1,88%	0,75%	3,76%	9,77%	12,78%	16,54%
IRS	PRF _{PF}	2,78%	5,56%	-1,39%	5,56%	6,94%	18,06%	30,56%
	PRF _{AF}	2,78%	5,56%	-1,39%	5,56%	6,94%	18,06%	31,94%
	PRF _{ATMV}	2,78%	6,94%	1,39%	9,72%	15,28%	20,83%	30,56%
	PRF _{PTMV}	2,78%	5,56%	1,39%	12,50%	15,28%	26,39%	37,50%
WIN	PRF _{PF}	0,00%	1,22%	3,66%	3,66%	13,41%	20,73%	17,07%
	PRF _{AF}	0,00%	6,02%	4,82%	7,23%	18,07%	24,10%	15,66%
	PRF _{ATMV}	-1,23%	2,47%	0,00%	1,23%	9,88%	18,52%	16,05%
	PRF _{PTMV}	1,20%	1,20%	0,00%	2,41%	7,23%	14,46%	12,05%
DBC	PRF _{PF}	0,76%	3,82%	2,29%	7,25%	9,16%	18,32%	18,70%
	PRF _{AF}	1,13%	3,77%	2,64%	7,55%	10,94%	19,25%	19,25%
	PRF _{ATMV}	1,13%	2,26%	2,63%	6,02%	9,40%	15,41%	15,79%
	PRF _{PTMV}	0,00%	2,26%	2,26%	4,14%	7,14%	15,04%	14,66%
PBC	PRF _{PF}	1,35%	-4,05%	-1,35%	0,00%	-2,70%	-2,70%	-2,70%
	PRF _{AF}	-2,86%	-1,43%	-5,71%	-1,43%	-10,00%	-8,57%	-8,57%
	PRF _{ATMV}	-1,32%	0,00%	2,63%	3,95%	0,00%	0,00%	0,00%
	PRF _{PTMV}	-4,17%	2,78%	0,00%	0,00%	-5,56%	-4,17%	-5,56%

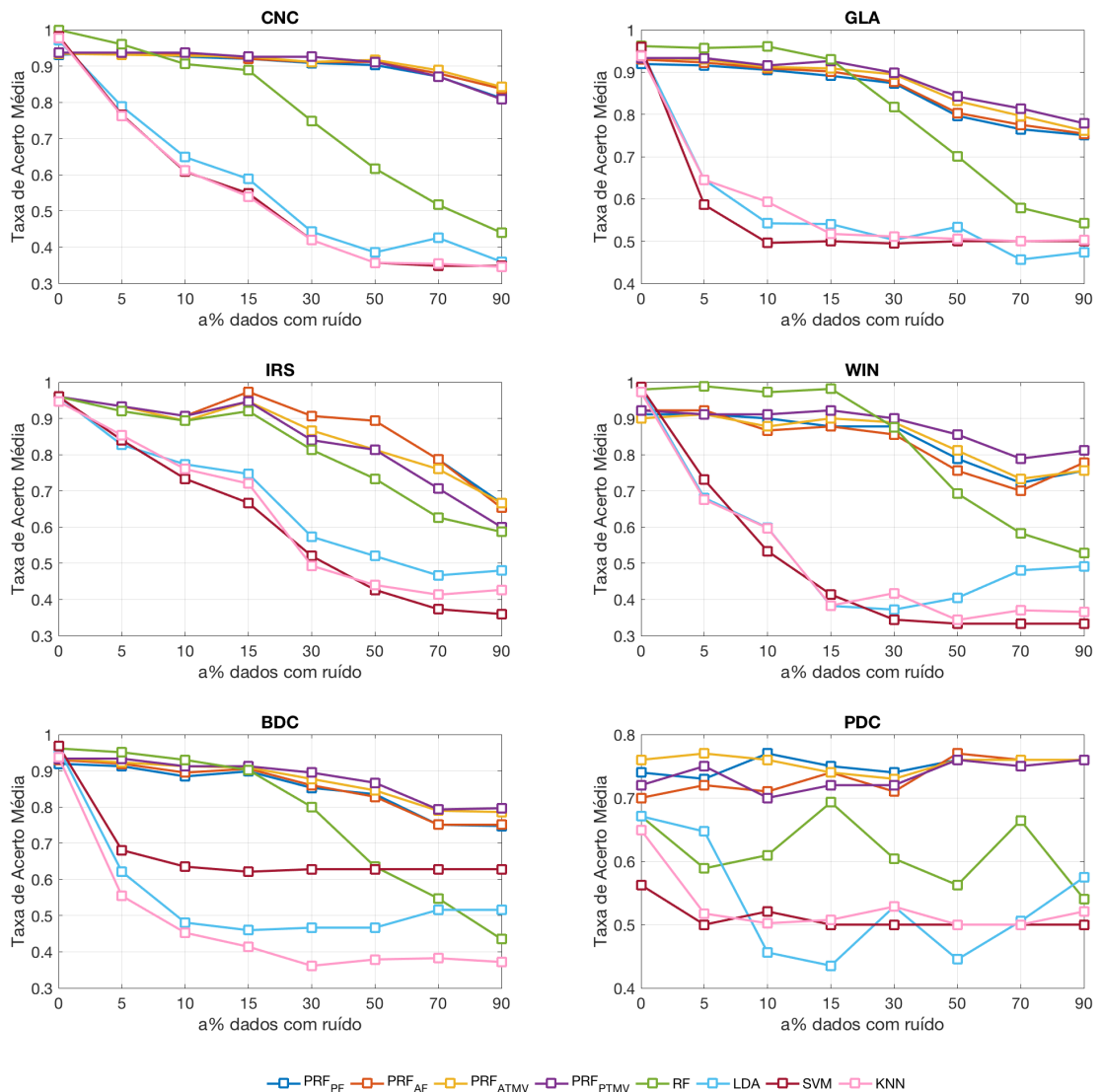
Similarmente, o decréscimo na taxa de acerto média, na presença de ruído na base de dados, também foi calculado para os métodos SVM, LDA, KNN e RF e seus resultados estão apresentados na Tabela 28, em comparação com os melhores resultados obtidos do método PRF. Analisando a Tabela 28, pode-se afirmar que, na grande maioria dos casos, um valor de 5% de dados ausentes representa uma queda significativa na predição para os métodos SVM, KNN e LDA, com decréscimos na taxa de acerto variando entre 3% e 40%, dependendo da base de dados utilizada.

Tabela 28 - Comparação do decréscimo na taxa de acerto média ao se adicionar $\alpha\%$ dados com ruído para os métodos PRF, SVM, KNN e LDA.

Base de dados	Método	% decréscimo na taxa de acerto						
		a% de ruído adicionado						
		5%	10%	15%	30%	50%	70%	90%
CNC	PRF	0,31%	0,61%	0,92%	2,45%	1,83%	4,89%	9,79%
	RF	4,00%	9,43%	11,14%	25,14%	38,29%	48,29%	56,00%
	LDA	18,82%	33,24%	39,41%	54,41%	60,29%	56,18%	62,94%
	SVM	22,09%	38,08%	44,19%	57,27%	63,66%	64,53%	64,53%
	KNN	21,93%	37,43%	44,74%	57,02%	63,45%	63,74%	64,62%
GLA	PRF	0,38%	2,26%	2,63%	4,14%	10,90%	14,66%	18,42%
	RF	0,47%	0,07%	3,29%	15,00%	27,09%	39,78%	43,55%
	LDA	32,23%	43,03%	43,22%	47,18%	43,91%	52,04%	50,21%
	SVM	38,93%	48,32%	47,89%	48,47%	47,89%	47,89%	47,89%
	KNN	31,23%	36,81%	44,87%	45,53%	46,11%	46,71%	46,41%
IRS	PRF	2,78%	6,94%	1,39%	9,72%	15,28%	20,83%	30,56%
	RF	2,78%	5,56%	1,39%	12,50%	15,28%	26,39%	37,50%
	LDA	13,89%	19,44%	22,22%	40,28%	45,83%	51,39%	50,00%
	SVM	12,50%	23,61%	30,56%	45,83%	55,56%	61,11%	62,50%
	KNN	9,86%	19,72%	23,94%	47,89%	53,52%	56,34%	54,93%
WIN	PRF	-1,23%	2,47%	0,00%	1,23%	9,88%	18,52%	16,05%
	RF	-0,95%	0,76%	-0,19%	10,68%	29,30%	40,45%	46,12%
	LDA	30,66%	39,06%	61,04%	62,08%	58,77%	51,04%	49,91%
	SVM	25,92%	45,92%	58,03%	65,07%	66,20%	66,20%	66,20%
	KNN	30,48%	38,57%	60,57%	57,05%	64,67%	61,90%	62,38%
DBC	PRF	1,13%	2,26%	2,63%	6,02%	9,40%	15,41%	15,79%
	RF	1,09%	3,28%	6,20%	16,79%	33,94%	43,07%	54,74%
	LDA	35,40%	50,00%	52,19%	51,46%	51,46%	46,35%	46,35%
	SVM	29,71%	34,42%	35,87%	35,14%	35,14%	35,14%	35,14%
	KNN	40,82%	51,69%	55,81%	61,42%	59,55%	59,18%	60,30%
PBC	PRF	-1,32%	0,00%	2,63%	3,95%	0,00%	0,00%	0,00%
	RF	12,25%	9,15%	-3,27%	9,97%	16,18%	0,98%	19,44%
	LDA	3,59%	32,03%	35,13%	21,24%	33,66%	24,67%	14,38%
	SVM	11,11%	7,41%	11,11%	11,11%	11,11%	11,11%	11,11%
	KNN	20,27%	22,64%	21,79%	18,58%	22,97%	22,97%	19,76%

Para melhor visualização dos resultados os mesmos também são apresentados de forma gráfica na Figura 26. Nesse gráfico é possível observar a superioridade do método PRF, e suas variações de tomada de decisão, em relação à taxa de acerto média, na comparação com outros métodos. Também verifica-se que, no que concerne ao método RF, o método PRF obteve taxas de acerto médias superiores em todos os casos quando considerada uma porcentagem de adição de ruído superior a 30%. Analisando as taxas de acerto médias dos métodos LDA, KNN e SVM percebe-se que para somente 5% de dados com ruído já existe uma redução significativa na taxa de acerto média, com reduções na taxa de acerto de até 40%.

Figura 26 - Gráfico da taxa de classificação em comparação aos métodos PRF, LDA, SVM e KNN com adição de ruído, sendo as bases de dados: CNC, WIN, GLA, IRS, DBC e PBC.



Fonte: Autor

Adicionalmente, realizando uma análise das médias pelo método Tukey, com 95% de confiabilidade, com resultados apresentados na Tabela 29, é possível afirmar que, considerando a classificação de bases de dados com adição de ruído, o método *Paraconsistent Random Forest* apresenta taxas de acerto médias superiores aos outros métodos testados neste experimento. Também pode-se afirmar que os métodos de tomada de decisão propostos para o algoritmo PRF não apresentaram diferença significativa, entre os mesmos, na taxa de acerto média, assim como observado no experimento com dados ausentes. Em relação aos outros métodos considerados no experimento, o RF apresentou uma taxa de acerto média superior

aos métodos LDA, KNN e SVM, cujas diferenças das médias não foram consideradas significativas para 95% de confiabilidade.

Tabela 29 - Análise das médias dos métodos utilizando método Tukey (95% de confiabilidade).

Métodos	Amostras	Médias	Grupo [†]
PRF _{ATMV}	48	85,30%	A
PRF _{PTMV}	48	85,17%	A
PRF _{PF}	48	84,59%	A
PRF _{AF}	48	84,44%	A
RF	48	76,41%	B
LDA	48	57,80%	C
KNN	48	56,88%	C
SVM	48	54,87%	C

[†] Médias que não compartilham uma letra são significativamente diferentes

Considerando os resultados e análise estatística dos dois experimentos realizados na avaliação do método PRF e suas variações, na presença de dados ausentes e com adição de ruído, é possível concluir que o algoritmo de tomada de decisão que obteve a melhor taxa de acerto média foi o PRF_{ATMV}. O algoritmo ATMV é uma modificação de algoritmo votação majoritária, amplamente utilizado em classificadores baseados em conjuntos (BONISSONE *et al.*, 2010; BREIMAN, 2001). A única diferença entre o método de votação majoritária e o método ATMV é que o mesmo necessita primeiro calcular uma média dos resultados considerando cada nó folha para a classificação individual de cada Árvore de Decisão, uma vez que, nessa proposta, uma amostra não percorre somente um único caminho na árvore, conforme explicado em mais detalhes no capítulo de metodologia.

Outro motivo pelo qual o algoritmo ATMV se destaca é pela sua eficiência computacional, necessitando de menos iterações quando comparado aos outros algoritmos propostos (PF e PTMV) que utilizam a Lógica Paraconsistente na tomada de decisão. Por exemplo, o algoritmo ParaExtr_{ctr}, depende diretamente do tamanho da árvore e a da quantidade de nós folha, ou seja, o tempo de execução pode se tornar considerável sem necessariamente refletir em um aumento na taxa de acerto, conforme confirmado pela análise estatística dos experimentos. Outro aspecto importante de ser ressaltado é que a estrutura da Árvore Paraconsistente é altamente dependente da partição de domínio *fuzzy*, o que sugere que a definição dos conjuntos *fuzzy* de entrada pode afetar diretamente a taxa de acerto da árvore criada.

Nesse experimento BDP foi possível avaliar as vantagens do método PRF em lidar com a presença de dados ausentes e de ruído que, por sua vez, podem gerar inconsistências no modelo de classificação definido *a priori*. Tais inconsistências podem facilmente ocorrer em sistemas baseados em aprendizagem supervisionada. No entanto, sistemas baseados em conhecimento que se apoiam na lógica clássica não conseguem lidar com essa situação, uma vez que o modelo gerado é trivial.

Conforme observado neste experimento, a Lógica Paraconsistente foi capaz de lidar com cenários contraditórios, representados pela ausência de dados e contaminação da base com ruído, evitando a trivialidade por meio da rejeição do princípio da explosão, no sentido de que essas teorias não se banalizam na presença de, pelo menos algumas, sentenças contraditórias. Essa característica do método PRF é obtida através de um modelo de Árvore de Decisão Paraconsistente, em que as decisões dos nós da árvore não são binárias e sim utilizam a teoria da LP para avaliar individualmente cada atributo e determinar um Grau de Evidência associado a cada classe, levando em consideração toda a informação disponível. Os Nós de Análise Paraconsistente têm um papel fundamental ao extrair gradualmente contradições no modelo e ao calcular o Grau de Evidência Real para todos os caminhos da árvore. Ou seja, a decisão final da Árvore Paraconsistente só é tomada avaliando os resultados calculados para todos os nós folha; diferentemente de uma Árvore de Decisão binária que percorre somente um único caminho. Nesse sentido, somente um atributo contraditório no conjunto de dados de entrada pode fazer com que uma Árvore de Decisão binária não seja capaz de classificar corretamente o dado apresentado.

Na literatura, outros métodos vêm sendo propostos com o objetivo de reduzir a susceptibilidade ao ruído na classificação dos dados, como é o caso da *Fuzzy Random Forest* (FRF) (BONISSONE *et al.*, 2010). O FRF utiliza uma abordagem de Árvore de Decisão *fuzzy* e também obteve resultados superiores no tratamento de dados com valores ausentes em comparação com outros métodos clássicos na área de reconhecimento de padrão. Comparando os resultados obtidos com o método PRF proposto, é possível sustentar que, para as bases de dados IRS e WIN, o método proposto PRF superou o FRF (BONISSONE *et al.*, 2010) na classificação dos dados com valores ausentes, conforme apresentado na Tabela 30. A maior diferença pode ser observada com 30% de dados ausentes, nos quais o método PRF obteve um decréscimo na taxa de acerto média em torno de 5 a 7% em comparação com o método FRF, cujo decréscimo foi de aproximadamente 14 a 16%. No entanto, é importante notar que o método FRF obteve resultados levemente superiores na classificação dos dados originais. Essa comparação suporta a prerrogativa de que a *Paraconsistent Random Forest* é um

algoritmo adequado em contextos que envolvem o tratamento de dados incompletos. Além disso, outra diferença entre o FRF e o PRF é que a FRF foi treinada e testada com o conjunto de dados modificado. Essa distinção pode levar a uma menor taxa de erro média, uma vez que o método já foi apresentado a esse tipo de dado durante a criação do modelo.

Tabela 30 - Comparativo do método PRF com o método FRF.

Base de dados	Algoritmo	Taxa de acerto média original	% decréscimo na taxa de acerto	
			$\alpha\%$ de valores ausentes	
			5%	30%
IRS	PRF	96,00%	0,00%	5,33%
	FRF (BONISSONE <i>et al.</i> , 2010)	97,33%	1,23%	16,71%
WIN	PRF	92,22%	0,00%	7,78%
	FRF (BONISSONE <i>et al.</i> , 2010)	97,88%	4,41%	14,21%

5.3 RESULTADOS E DISCUSSÕES DO EXPERIMENTO SEMG 2

O experimento sEMG2 compreende o uso e a análise do algoritmo *Paraconsistent Random Forest* (PRF) desenvolvido utilizando como entrada sinais de sEMG. O objetivo desse experimento é avaliar o método PRF na aplicação de classificação de movimentos do membro superior e comparar resultados com outros métodos de reconhecimento de padrão bem fundamentos na literatura como SVM, KNN, LDA e RF. O experimento compreende a avaliação de dados de sEMG íntegros e também sinais de sEMG com adição de contaminantes artificiais com o intuito de analisar a robustez do método em relação a degradações no sinal de entrada, conforme as bases de dados e a metodologia descritas previamente no Capítulo 4.

Em uma primeira análise, foi realizado um projeto de experimentos que envolve a análise dos resultados na classificação de 17 movimentos do membro superior, para as 3 bases de dados com sinal de sEMG, utilizando o método proposto PRF em comparação com os classificadores SVM, KNN, LDA e RF. As taxas de acerto médias dos métodos foram calculadas utilizando características extraídas do sinal de sEMG original, ou seja, sem adição de contaminantes. No Apêndice A – Projeto de Experimentos encontram-se os gráficos de resíduos e potência estatística deste experimento.

Avaliando a potência estatística do projeto de experimentos realizado é possível afirmar que, para uma potência de 95%, a máxima diferença das médias é de 1,99%; ou seja, o percentual máximo das diferenças de taxa de acerto considerado significativo com um nível

de confiabilidade de 95%, conforme pode ser visualizado na Figura 38 (a) do Apêndice - A. Através da análise de resíduos, também apresentados na Figura 38 do Apêndice - A, verifica-se que os mesmos possuem uma distribuição normal e os dados satisfazem a validação do projeto de experimentos. Dessa forma, com garantia de que a potência estatística está adequada, com número de amostras suficiente e que o modelo do projeto de experimentos está validado em uma distribuição normal, pode-se avaliar se a diferença das médias dos fatores analisados é significativa, conforme apresentado na Tabela 31. O projeto de experimentos foi elaborado considerando como fatores controláveis os métodos de classificação e as bases de dados. Como variável de resposta é considerada a taxa de acerto média.

Tabela 31 - Projeto de experimentos referente à análise dos classificadores utilizando sinal de sEMG original.

Fatores*	Graus de liberdade	Soma dos quadrados	Média Quadrada	F	P
A	4	1,706	0,42	10,41	0,000
B	2	18,422	9,21	224,83	0,000
A*B	8	2,092	0,26	6,38	0,000
Erro	5816	238,265	0,04		
Total	5830	262,251			

* Fator A: Método de classificação e Fator B: base de dados

O cálculo do projeto para o experimento de dois fatores controláveis, apresentado na Tabela 31, demonstrou que existem diferenças significativas, levando-se em consideração 95% de confiabilidade, nas taxas de acerto médias entre os métodos de classificação e também nas bases de dados, conforme valor de P inferior a 0,05. Também é possível alegar que a interação entre os fatores métodos e bases de dados é significativa.

Aprofundando a análise de quais métodos são significativos, foi realizada uma análise das médias baseada no método Tukey, com 95% de confiabilidade, com resultados apresentados na Tabela 32. Nessa análise é possível afirmar que o método PRF obteve taxas de acerto médias sem diferença significativa em comparação com os métodos LDA e KNN. Também é interessante avaliar que, considerando a taxa de acerto média, os métodos RF e SVM também não são considerados significativamente diferentes para uma confiabilidade de 95%. No entanto, os mesmos apresentam uma taxa de acerto média significativamente superior aos métodos LDA, PRF e KNN. É importante ressaltar que, apesar do método RF se mostrar superior em termos de taxa de acerto média em comparação ao método desenvolvido PRF, a diferença média entre os dois foi apenas de aproximadamente 4%.

Tabela 32 - Análise das médias dos métodos utilizando o método de Tukey (95% de confiabilidade).

Métodos	Amostras	Médias	Grupo [†]
RF	1173	62,49%	A
SVM	1173	62,05%	A
LDA	1173	58,97%	B
PRF	1173	58,46%	B
KNN	1173	58,01%	B

[†] Médias que não compartilham uma letra são significativamente diferentes

Na Tabela 33 é apresentado o resultado da análise das médias considerando as bases de dados através do método Tukey, com 95% de confiabilidade. Nessa análise pode-se assegurar, com 95% de confiabilidade, que as bases de dados possuem taxas de acerto médias distintas entre si. A base de dados NPA, que é composta por voluntários com amputação parcial do antebraço, possui a menor taxa de acerto média com 50,50% em comparação com as outras bases de dados que só contêm sinais de sEMG de sujeitos intactos, conforme também relatado em outros estudos (ATZORI *et al.*, 2015; DE MOURA; BALBINOT, 2018; FAVIEIRO; MOURA; BALBINOT, 2016; GIJSBERTS *et al.*, 2014; MOURA; FAVIEIRO; BALBINOT, 2016). No tocante às bases de dados com voluntários com membro superior intacto, IEE e NPI, a base de dados IEE apresenta resultados significativamente superiores, com uma média de 67,54% considerando todos os voluntários e movimentos. A superioridade da base de dados IEE pode ser explicada pelo fato da mesma ter sido adquirida em um ambiente controlado de laboratório, bem como por muitos dos voluntários participantes serem integrantes do grupo de pesquisa e, por sua vez, participaram de diversas coletas de dados. Essa participação frequente ocasiona uma maior familiarização com o sistema de aquisição e modelo virtual, podendo, possivelmente, elevar a taxa de acerto média.

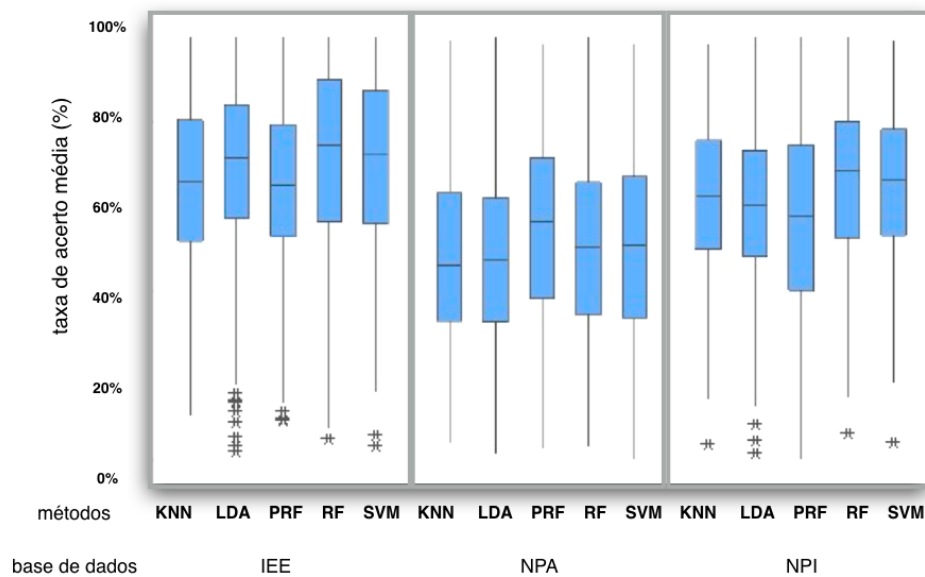
Tabela 33 - Análise das médias das bases de dados utilizando o método de Tukey (95% de confiabilidade).

Bases de dados	Amostras	Médias	Grupo [†]
IEE	1513	67,54%	A
NPI	3230	61,83%	B
NPA	1088	50,50%	C

[†] Médias que não compartilham uma letra são significativamente diferentes

Uma análise comparativa foi realizada entre os diferentes métodos de classificação para cada uma das bases de dados avaliadas, conforme representado na Figura 27. Nesse gráfico é possível visualizar que o método RF apresenta as melhores taxas de acerto médias para as bases de dados com voluntários com membro superior intacto, IEE e NPI. No entanto, para a base de dados NPA (que é composta por sinais adquiridos de voluntários com amputação transradial do antebraço), o método PRF apresentou taxas de acerto superiores aos outros. Esse fato corrobora a hipótese de que o método PRF é adequado para lidar com sinais inconsistentes e/ou com informação limitada. Sinais de sEMG adquiridos de voluntários com amputação possuem uma limitação física no posicionamento dos eletrodos, dificuldades para executar os movimentos, especialmente se os mesmos não estão acostumados ou não utilizam próteses mioelétricas, e possíveis atrofias musculares. Esses sinais também são propensos a um baixo SNR. Todos esses fatores têm influência quando da utilização de sinais de sEMG para um reconhecimento natural e intuitivo do movimento e dificultam o reconhecimento de movimentos pelos classificadores, como é observado na Figura 27.

Figura 27 - Taxa de acerto média na análise de base de dados de sinal de sEMG íntegro.

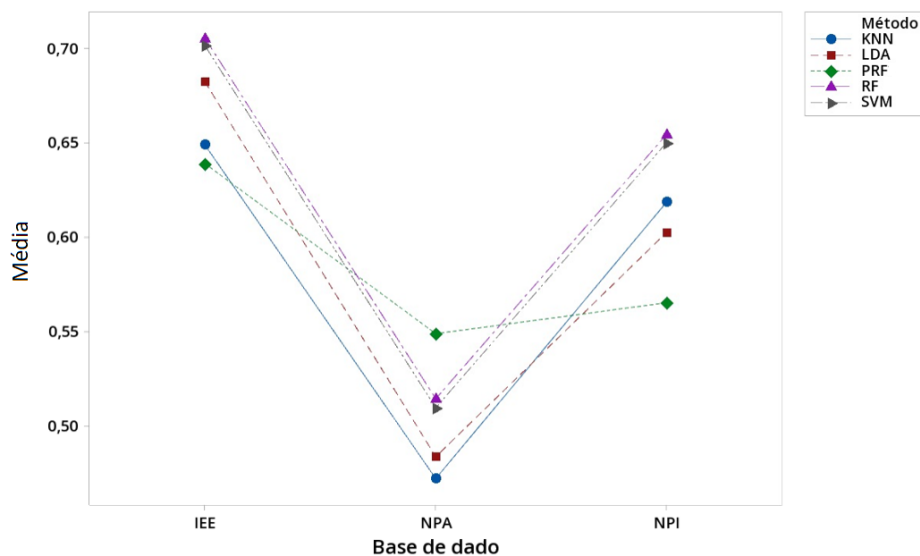


Fonte: Autor

Para analisar em maiores detalhes as taxas de acerto médias, é interessante avaliar o gráfico de interação entre os fatores, apresentado na Figura 28. Nesse gráfico existe o cruzamento entre fatores, o que pode indicar que existem diferenças significativas na taxa de acerto média considerando os métodos e bases de dados. Portanto, entre os métodos avaliados pode não existir um único que possua uma taxa de acerto média superior para todas as bases

de dados consideradas. Essa hipótese também foi evidenciada na ANOVA do experimento, Tabela 31, na qual a interação entre os fatores foi considerada significativa. Sendo assim, após uma análise das diferenças das médias pelo método de Tukey, considerando a interação dos fatores bases de dados e métodos de classificação, com 95% de confiabilidade, é possível afirmar que para a base de dados NPA o método PRF possui uma taxa de acerto significativamente superior aos métodos LDA, KNN e SVM. Nessa análise, não foi observada diferença significativa em relação ao método RF.

Figura 28 - Gráfico de interação entre médias da taxa de acerto para os métodos e bases de dados considerando os sinais de sEMG íntegros.



Fonte: Autor

Nesse projeto de experimento, considerando o sinal de sEMG original, sem nenhuma contaminação artificial adicionada, confirma-se que, de modo geral, os métodos RF e SVM apresentaram as melhores taxas de acerto médias. Logo, em uma primeira análise do método PRF e sua aplicabilidade na área de reconhecimento de movimentos, seria possível afirmar que esse método não é o mais adequado, pois os métodos RF e SVM obtiveram taxas de acerto significativamente superiores, considerando as bases de dados NPI e IEE. No entanto, apesar das diferenças serem significativas, o método PRF apresentou uma taxa de acerto próxima em relação ao método RF, com uma diferença média de 4%. Além disso, para a base de dados de voluntários com amputação NPA o método PRF apresenta uma taxa de acerto média similar ao método RF e significativamente superior aos métodos LDA, KNN e SVM.

Sendo assim, uma segunda avaliação do método foi realizada, utilizando sinais de sEMG com contaminações artificiais, para determinar se o método PRF é adequado para lidar com possíveis contradições geradas por degradações no sinal de sEMG. Para isso os modelos dos classificadores foram gerados a partir de um conjunto de dados proveniente do sinal de sEMG original e testados com dados do sinal de sEMG com adição de contaminantes em diferentes combinações de canais de aquisição, conforme detalhado previamente na metodologia de testes.

O projeto de experimentos desta segunda avaliação foi elaborado considerando como fatores controláveis os métodos de classificação, bases de dados, tipos de contaminantes e canais contaminados. A variável de resposta considerada foi a taxa de acerto média. Avaliando a potência estatística do projeto de experimentos realizado é legítimo afirmar que para uma potência de 95%, a máxima diferença das médias é de 0,71%, conforme pode ser visualizado na Figura 39 (a) do Apêndice A – Projeto de Experimentos. Através da análise de resíduos, também apresentados na Figura 39, verifica-se que os mesmos possuem uma distribuição normal e os dados satisfazem o modelo de validação do projeto de experimentos. Dessa forma, com garantia de que a potência estatística está adequada, com número de amostras suficiente e que o modelo do projeto de experimentos está validado em uma distribuição normal, é possível avaliar se a diferença das médias dos fatores analisados é significativa, conforme apresentado na Tabela 34.

Tabela 34 - Projeto de experimentos referente à análise dos classificadores utilizando sinal de sEMG com contaminantes.

Fatores [†]	Graus de liberdade	Soma dos quadrados	Média Quadrada	F	P
A	2	511	255,52	3818,16	0
B	4	2954,2	738,544	11035,84	0
C	2	53,8	26,909	402,1	0
D	7	404,1	57,735	862,72	0
A*B	8	303,2	37,903	566,38	0
A*C	4	17,8	4,461	66,66	0
A*D	14	59,5	4,25	63,5	0
B*C	8	25,1	3,134	46,84	0
B*D	28	278,9	9,959	148,82	0
C*D	14	14,2	1,016	15,18	0
A*B*C	16	11,3	0,709	10,59	0
A*B*D	56	61,6	1,099	16,43	0
A*C*D	28	5,9	0,21	3,13	0
B*C*D	56	15,4	0,275	4,11	0
A*B*C*D	112	8,9	0,079	1,19	0,086
Erro	139584	9341,3	0,067		
Total	139943	15212,0			

[†] Fator A: Tipo de contaminante, Fator B: Métodos, Fator C: Base de dados e Fator D: Canais com contaminação

O cálculo do projeto para o experimento de quatro fatores apresentado na Tabela 34 demonstrou que existem diferenças significativas, considerando 95% de confiabilidade, nas taxas de acerto médias entre todos os fatores individualmente. As interações entre os fatores também foram consideradas significativas, exceto para a interação que contempla todos os fatores, que teve um valor P de 0,086, superior a 0,05.

Com o objetivo de avaliar a taxa de acerto média dos métodos de classificação, foi realizada uma análise das médias de acordo com o método Tukey, considerando um nível de confiabilidade de 95%, conforme resultados apresentados na Tabela 35. A partir dessa análise é possível afirmar que todos os métodos de classificação analisados apresentam taxas de acerto médias significativamente distintas entre si. Adicionalmente o método PRF se mostrou superior aos métodos RF, KNN, LDA e SVM em relação à taxa de acerto média, considerando a média envolvendo todos os contaminantes, combinações de canais de sEMG contaminados e bases de dados, com uma média de 55,09% de acerto. O segundo melhor método é o RF, com uma taxa de acerto média de 35,61%. Essa análise confirma a hipótese de que a união de conceitos de Lógica Paraconsistente em uma extensão do método de *Random Forest* é capaz de mitigar os efeitos de contaminantes de sinal de sEMG na taxa de classificação correta dos padrões, provendo um método ainda mais robusto a mudanças no sinal de entrada em comparação ao método *Random Forest*.

Tabela 35 - Análise das médias dos métodos utilizando o método de Tukey (95% de confiabilidade).

Métodos	Amostras	Médias	Grupo [†]
PRF	28152	55,09%	A
RF	28152	35,61%	B
KNN	28152	18,06%	C
LDA	28152	13,90%	D
SVM	28152	12,56%	E

[†] Médias que não compartilham uma letra são significativamente diferentes

Considerando as bases de dados, a Tabela 36 apresenta a análise das médias segundo metodologia de Tukey, considerando uma confiabilidade de 95%. Como esperado, as taxas de acerto médias para as diferentes bases de dados seguiram o mesmo padrão do projeto de experimentos levando em consideração os sinais de sEMG originais (sem adição de contaminantes), na qual a base de dados IEE obteve as melhores taxas de acerto com 30,15% e a base de dados NPA obteve a menor taxa de acerto com 24,62%.

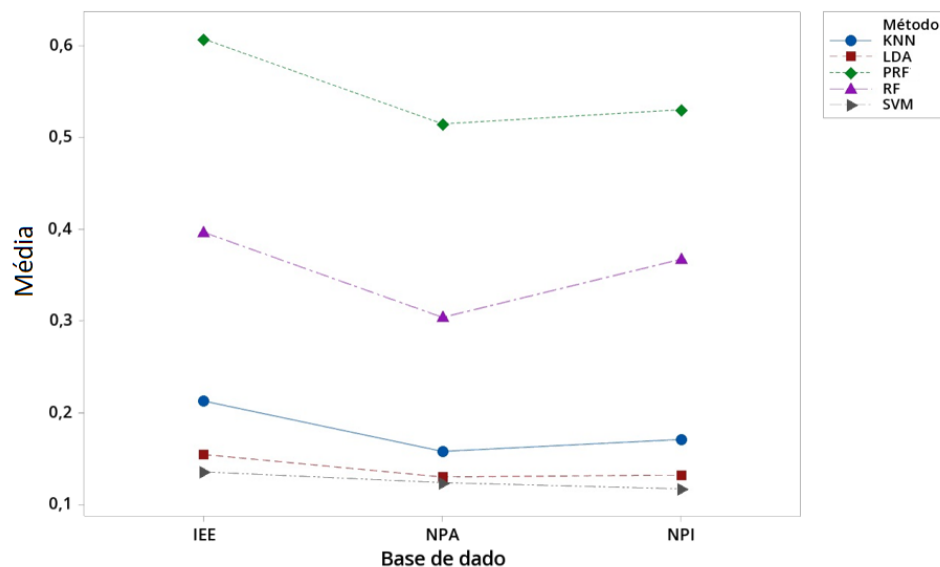
Além das análises das médias, também foram realizadas análises das interações entre os fatores. Na Figura 29 é apresentado o gráfico da interação entre os métodos de classificação e bases de dados. Nesse gráfico pode-se observar que o método PRF possui taxas de acerto superiores para todas as bases de dados.

Tabela 36 - Análise das médias das bases de dados utilizando o método de Tukey (95% de confiabilidade).

Base de dados	Amostras	Médias	Grupo [†]
IEE	36312	30,15%	A
NPI	77520	26,37%	B
NPA	26112	24,62%	C

[†]Médias que não compartilham uma letra são significativamente diferentes

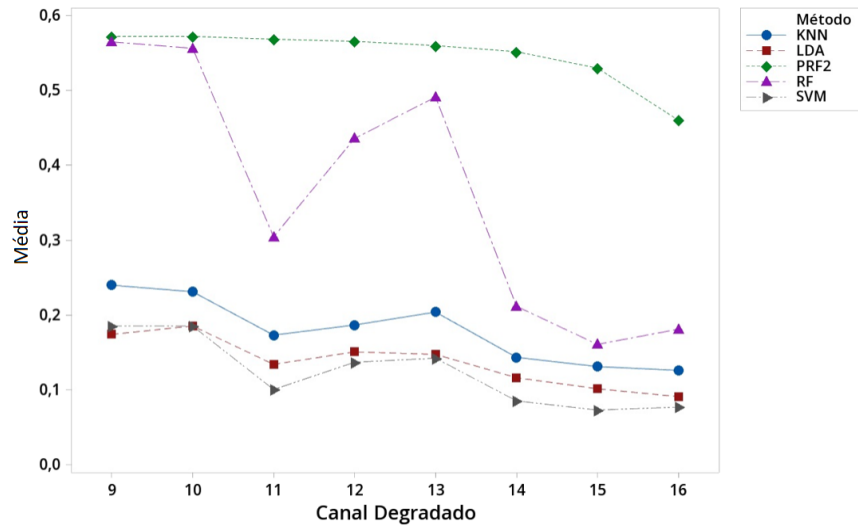
Figura 29 - Gráfico da interação entre os fatores métodos e bases de dados.



Fonte: Autor

No gráfico de interação entre os fatores métodos e tipo de contaminante (Figura 30) é possível observar que tanto o método PRF e RF aparentemente são pouco afetados, quando comparado aos demais, pelos diferentes tipos de contaminantes. No entanto, os métodos KNN, LDA e SVM apresentam taxas de acerto médias distintas entre os tipos de contaminantes. As linhas projetadas do gráfico não são paralelas, indicando que a interação entre métodos e tipo de contaminante é significativa o que ficou comprovado pelos resultados apresentados na Tabela 34 (interação A e B).

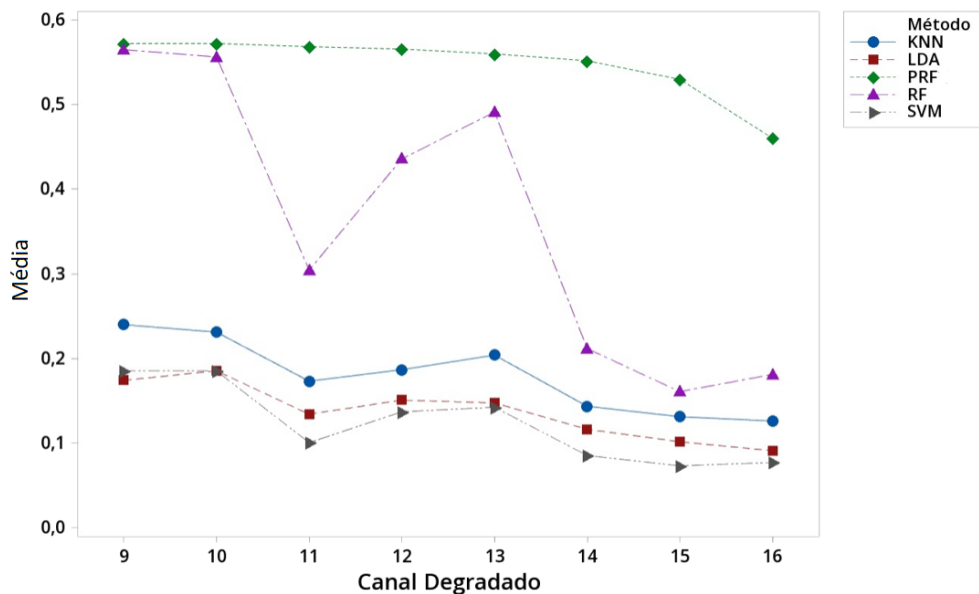
Figura 30 - Gráfico da interação entre os fatores métodos e tipo de contaminante.



Fonte: Autor

Adicionalmente, analisando graficamente a interação entre os canais contaminados e os métodos de classificação, apresentada na Figura 31, é possível verificar que as linhas não são paralelas, indicando que a interação entre os fatores é significativa. Também se observa que o método RF sofre uma maior influência na taxa de acerto média de acordo com a configuração de canais de sEMG no qual o contaminante foi adicionado.

Figura 31 - Gráfico da interação entre os fatores métodos e canais com contaminação.



Fonte: Autor

Também foram realizadas as análises individuais para cada tipo de contaminante. O primeiro experimento de avaliação da classificação do sinal de sEMG com adição de contaminantes compreende a contaminação por artefatos de movimento. No Apêndice A – Projeto de Experimentos encontram-se os gráficos de resíduos e da potência estatística deste experimento. Os fatores controláveis A, B e C correspondem, respectivamente, aos métodos de classificação, bases de dados e combinação de canais de sEMG contaminados. O projeto de experimentos foi avaliado quanto a sua potência estatística (Figura 40), concluindo-se que, para uma potência de 95%, a máxima diferença das médias calculada foi de 1,54%. Além disso, o modelo foi considerado adequado para a análise da diferença entre médias, de acordo com a análise de resíduos apresentada na Figura 40. O cálculo do projeto para o experimento de três fatores é exibido na Tabela 37 e o mesmo demonstrou que as principais variações dos fatores referentes às bases de dados e classificadores, configuração de canais com contaminante e tipo de contaminante adicionado ao sinal são significativas, ou seja, as variações dos três principais efeitos são diferentes. As interações entre os fatores controláveis também são significativas, devido ao valor P ser inferior a 0,05.

Tabela 37 - Projeto de experimentos referente a análise dos métodos de classificação utilizando sinal de sEMG com contaminante artefato de movimento.

Fatores [†]	Graus de liberdade	Soma dos quadrados	Média Quadrada	F	P
A	4	1361,14	340,286	5357,86	0,000
B	2	8,59	4,296	67,65	0,000
C	7	88,27	12,609	198,54	0,000
A*B	8	11,98	1,498	23,58	0,000
A*C	28	127,30	4,547	71,59	0,000
B*C	14	4,28	0,306	4,82	0,001
A*B*C	56	6,10	0,109	1,72	
Erro	46528	2955,06	0,064		
Total	46647	4957,93			

[†] Fator A: Métodos de classificação, Fator B: Base de dados e Fator C: Canais com contaminação

A análise de diferenças entre médias com 95% de confiabilidade em relação aos classificadores está representada na Tabela 38. Nessa análise é possível afirmar que o método PRF obteve uma taxa de acerto média de 54,84%, sendo significativamente superior em comparação com todos os outros métodos propostos, considerando somente os sinais de sEMG contaminados por artefatos de movimento e todas as combinações de canais contaminados. Também pode-se concluir que todos os métodos RF e KNN obtiveram uma taxa de acerto média sem diferenças significativas de aproximadamente 38%.

Tabela 38 - Análise das médias dos métodos utilizando o método de Tukey (95% de confiabilidade).

Métodos	Amostras	Médias	Grupo [†]
PRF	9384	54,84%	A
RF	9384	38,12%	B
KNN	9384	37,51%	B
LDA	9384	27,27%	C
SVM	9384	24,59%	D

[†] Médias que não compartilham uma letra são significativamente diferentes

Na sequência, realizou-se uma análise mais detalhada no decréscimo da taxa de acerto quanto ao contaminante, configuração de canais contaminados, bases de dados e métodos de classificação avaliados, com o objetivo de avaliar se o método PRF apresenta taxas de acerto médias superiores aos outros métodos em todas as configurações de adição de contaminantes testadas. Na Tabela 39 encontra-se o decréscimo calculado para a taxa de acerto levando-se em consideração o método, a base de dados e as configurações de contaminação do sinal para o contaminante artefato de movimentos. Em uma primeira análise, o método PRF apresentou taxas inferiores de decréscimo na classificação em comparação com os outros métodos e combinações de canais contaminados, com um erro adicional de somente 21,65% para uma configuração com 8 canais de sEMG contaminados por artefatos de movimento. Considerando um único canal contaminado (configurações 9 a 12), o método PRF apresentou um erro adicional de, no máximo, 4,5%. O gráfico das taxas de acerto médias, resultante dos dados da tabela, também é apresentado para uma melhor visualização dos resultados na Figura 32.

Observando o gráfico de taxas de acerto médias considerando os métodos de classificação e as configurações de sinais de sEMG com contaminante artefato de movimento, representado na Figura 32, o método PRF apresenta uma taxa de acerto média mais estável em relação aos canais contaminados quando comparado com os outros métodos avaliados nesta tese. Isso suporta a teoria de que a Lógica Paraconsistente é menos suscetível a degradações no sinal, pelo fato de que a mesma consegue lidar com algumas evidências inconsistentes, no caso atributos do sinal de sEMG com adição de contaminantes, sem invalidar as conclusões e de forma a produzir um Grau de Certeza Resultante levando em consideração todos os atributos apresentados ao modelo. Comparando com os métodos LDA, KNN e SVM, que dependem muito da posição espacial dos atributos no plano, esses métodos

acabam sendo mais passíveis a alterações significativas nas evidências de entrada (atributos do sinal de sEMG). Por esse motivo, a peculiaridade da *Paraconsistent Random Forest* em avaliar todos os possíveis caminhos da árvore é essencial para uma maior robustez quanto aos contaminantes do sinal: dessa forma é possível avaliar o conjunto completo dos atributos e identificar para qual classe as evidências de entrada possuem o maior Grau de Certeza associado.

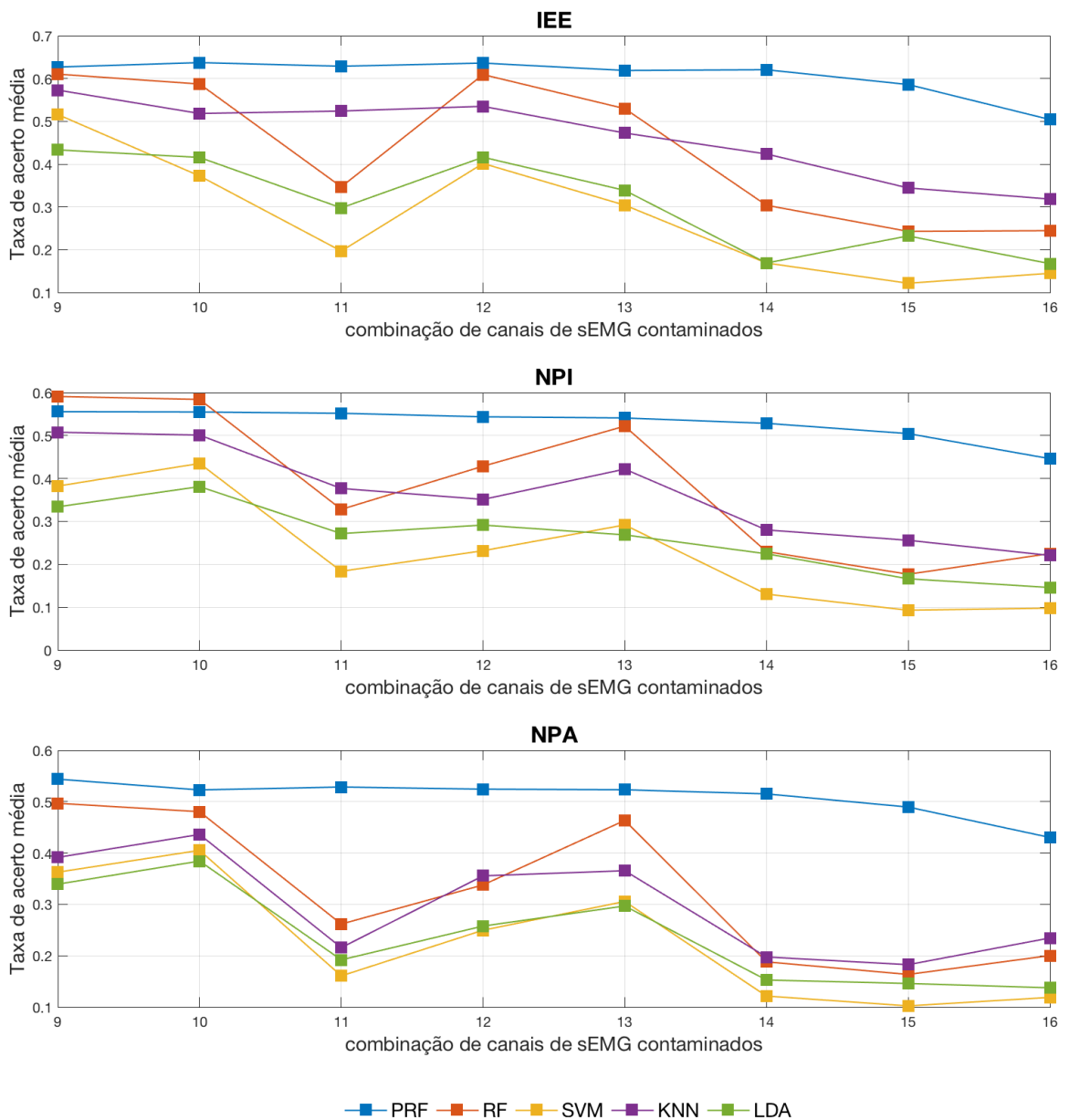
Tabela 39 - Decréscimo na taxa de classificador de acordo com o método e configuração de canais contaminados por artefatos de movimento.

Base de dados	Método	% decréscimo na taxa de acerto							
		configurações de canais contaminados							
		9	10	11	12	13	14	15	16
IEE	PRF	1,92%	0,30%	1,63%	0,47%	3,18%	2,91%	8,31%	21,11%
	RF	13,51%	16,77%	50,83%	13,65%	24,92%	56,94%	62,01%	65,34%
	SVM	26,47%	46,93%	71,95%	42,77%	56,65%	75,92%	82,67%	79,36%
	KNN	11,73%	20,14%	19,26%	17,57%	27,18%	34,71%	46,95%	50,96%
	LDA	36,52%	39,10%	56,37%	39,02%	50,38%	75,25%	65,99%	75,47%
NPI	PRF	2,42%	2,55%	3,09%	4,53%	5,00%	7,20%	11,40%	21,65%
	RF	9,70%	10,75%	49,89%	34,52%	20,27%	64,93%	68,96%	65,63%
	SVM	41,20%	33,10%	71,75%	64,36%	55,14%	79,93%	85,70%	84,94%
	KNN	17,94%	19,01%	39,11%	43,24%	31,82%	54,73%	58,62%	64,34%
	LDA	44,58%	36,78%	54,94%	51,65%	55,43%	62,75%	72,41%	75,79%
NPA	PRF	0,97%	4,79%	3,75%	4,52%	4,74%	6,19%	10,85%	21,64%
	RF	3,48%	6,66%	49,17%	34,37%	9,97%	63,47%	70,22%	61,10%
	SVM	28,90%	20,55%	68,49%	51,06%	40,08%	76,20%	79,94%	76,71%
	KNN	17,14%	7,70%	54,35%	24,76%	22,65%	58,20%	61,38%	50,37%
	LDA	29,90%	20,60%	60,35%	46,76%	38,68%	68,45%	69,86%	71,59%

Também pode-se atestar, analisando-se a Figura 32, que a configuração de canais de sEMG contaminado 11, representado o músculo bíceps, é relevante para a taxa de classificação correta para os métodos RF, SVM, LDA e KNN. É possível supor que esse canal seja muito significativo para os 17 movimentos selecionados e que qualquer degradação no mesmo afete de maneira significativa o modelo de reconhecimento de padrão gerados, causando um erro adicional na taxa de acerto superior a 80% em alguns casos. O mesmo padrão é observado para as configurações de contaminação 14 e 15, que também envolvem o canal de aquisição de sinais de sEMG na região do bíceps. Essa tendência não ocorre para o método PRF, que se apresenta mais tendencioso ao decréscimo da taxa de acerto em relação à quantidade de canais degradados. Uma das explicações dessa diferença de comportamento é o fato de que o método PRF não depende de maneira significativa de uma única entrada, pois o método avalia todos os possíveis caminhos da árvore gerada no lugar de apenas realizar tomadas de decisão binárias; diferentemente do método RF que percorre um caminho único

em cada Árvore de Decisão gerada. Adicionalmente, o método SVM, que obteve os maiores decréscimos na taxa de acerto, entre 20% e 80%, conforme apresentado na Tabela 39, é baseado em hiperplanos, ou seja, qualquer alteração em uma das dimensões/atributos pode fazer com que o dado seja classificado em um hiperplano diferente do esperado.

Figura 32 - Gráfico da taxa de acerto média considerando métodos e configurações de contaminação dos canais de sEMG para o contaminante artefato de movimento.



Fonte: Autor

Na sequência realizou-se a análise da classificação do sinal de sEMG com adição do contaminante ruído térmico. No Apêndice A – Projeto de Experimentos encontram-se os gráficos de resíduos e potência estatística deste experimento. Os fatores controláveis A, B e C correspondem, respectivamente, aos métodos de classificação, bases de dados e combinação de canais de sEMG contaminados. O experimento foi avaliado considerando uma potência de 95% e a máxima diferença das médias calculada foi de 1,25%, conforme gráficos apresentados na Figura 41.

Além disso, o modelo foi considerado adequado para a análise da diferença entre médias, de acordo com a análise de resíduos apresentadas na Figura 41. O cálculo do projeto para o experimento de três fatores é apresentado na Tabela 40 e o mesmo demonstrou que as principais variações dos fatores referentes às bases de dados e classificadores, configuração de canais com contaminante e tipo de contaminante adicionado ao sinal são significativas, ou seja, as variações dos três principais efeitos são diferentes. As interações entre os fatores controláveis também são significativas, devido ao valor P ser inferior a 0,05.

Tabela 40 - Projeto de experimentos referente a análise dos métodos de classificação utilizando sinal de sEMG com contaminante ruído térmico.

Fatores [†]	Graus de liberdade	Soma dos quadrados	Média Quadrada	F	P
A	4	1361,14	340,286	5357,86	0,000
B	2	8,59	4,296	67,65	0,000
C	7	88,27	12,609	198,54	0,000
A*B	8	11,98	1,498	23,58	0,000
A*C	28	127,30	4,547	71,59	0,000
B*C	14	4,28	0,306	4,82	0,000
A*B*C	56	6,10	0,109	1,72	0,001
Erro	46528	2955,06	0,064		
Total	46647	4957,93			

[†] Fator A: Métodos de classificação, Fator B: Base de dados e Fator C: Canais com contaminação

Também foi realizada uma análise de diferenças entre médias em relação aos classificadores com o método de Tukey (95% de confiabilidade) com resultados apresentados na Tabela 41. Nessa análise é cabível afirmar que o método PRF obteve uma taxa de acerto média significativamente superior em comparação com todos os outros métodos propostos, considerando somente os sinais de sEMG contaminados por ruído térmico e todas as combinações de canais contaminados. Também pode-se afirmar que todos os métodos de classificação analisados obtiveram taxas de acerto médias com diferenças significativas entre si.

Tabela 41 - Análise das médias dos métodos utilizando o método de Tukey (95% de confiabilidade).

Métodos	Amostras	Médias	Grupo [†]
PRF	9384	55,51%	A
RF	9384	34,40%	B
KNN	9384	10,40%	C
LDA	9384	8,18%	D
SVM	9384	6,81%	E

[†] Médias que não compartilham uma letra são significativamente diferentes

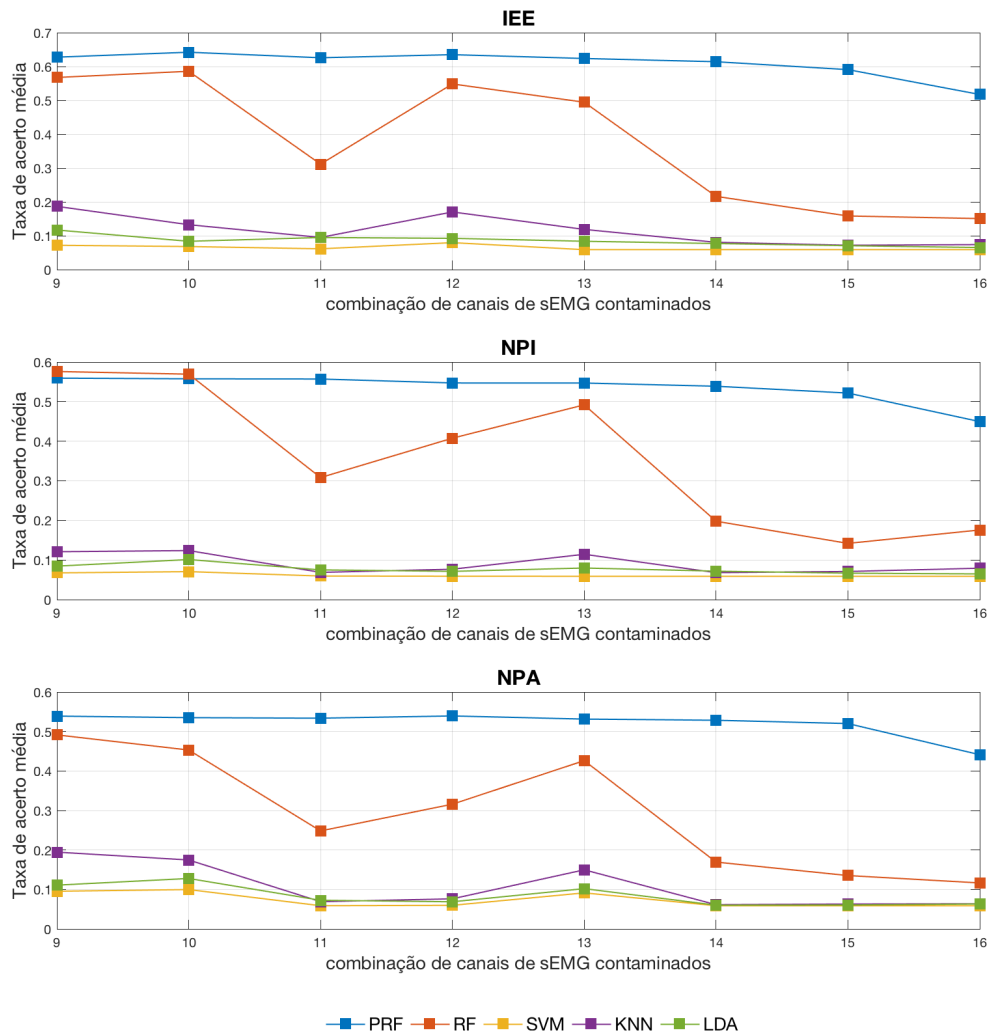
Tabela 42 - Decréscimo na taxa de classificador de acordo com o método e canais contaminados por ruído térmico.

Base de dados	Método	% decréscimo na taxa de acerto							
		configurações de canais contaminados							
		9	10	11	12	13	14	15	16
IEE	PRF	1,80%	-0,49%	2,08%	0,64%	2,41%	3,91%	7,56%	19,04%
	RF	19,55%	16,95%	55,85%	22,25%	29,91%	69,34%	75,22%	78,64%
	SVM	89,76%	90,29%	91,25%	88,65%	91,62%	91,62%	91,62%	91,62%
	KNN	71,22%	79,57%	85,36%	73,76%	81,71%	87,58%	88,92%	88,63%
	LDA	82,87%	87,75%	86,11%	86,47%	87,73%	88,72%	89,65%	90,46%
NPI	PRF	1,79%	2,11%	2,18%	3,93%	3,97%	5,39%	8,41%	21,07%
	RF	12,01%	13,03%	52,92%	37,73%	24,87%	69,77%	75,08%	73,14%
	SVM	89,63%	89,14%	90,85%	90,93%	90,96%	90,96%	90,96%	90,96%
	KNN	80,48%	79,99%	88,90%	87,60%	81,51%	89,01%	88,52%	87,20%
	LDA	85,98%	83,23%	87,56%	88,17%	86,78%	88,09%	88,93%	89,25%
NPA	PRF	1,84%	2,57%	2,80%	1,77%	3,22%	3,75%	5,28%	19,65%
	RF	4,53%	11,97%	51,78%	38,62%	17,19%	67,11%	75,38%	77,37%
	SVM	81,29%	80,45%	88,47%	88,27%	82,10%	88,47%	88,47%	88,47%
	KNN	58,84%	63,05%	85,40%	83,85%	68,32%	87,01%	86,67%	86,50%
	LDA	77,09%	73,60%	85,02%	85,79%	78,93%	87,56%	87,48%	86,94%

Na Tabela 42 encontra-se o decréscimo calculado para a taxa de acerto levando em consideração o método, a base de dados e as configurações de contaminação do sinal para o contaminante ruído térmico. Em uma primeira análise, nota-se o efeito negativo do contaminante ruído térmico para os métodos LDA, KNN e SVM, que apresentam um valor elevado de decréscimo na taxa de acerto entre 58% a 91%, considerando somente 1 canal contaminado (configurações 9 a 12). Em comparação, o método PRF apresentou as menores taxas de decréscimo na classificação com taxas de decréscimo entre 0,46% para 1 canal contaminando até 48,48%, considerando 8 canais de sEMG contaminados (configuração 16). Um gráfico das taxas de acerto médias por base de dados, configuração de canais

contaminados e métodos também é apresentado para uma melhor visualização dos resultados na Figura 33.

Figura 33 - Gráfico da % de decréscimo na taxa de acerto considerando métodos e combinações de degradação dos sinais para o contaminante ruído térmico.



Fonte: Autor

Observando a Figura 33 é interessante notar que a taxa de decréscimo é mais elevada para o contaminante ruído térmico do que para o artefato de movimentos, indicando que o sinal de sEMG contaminado por ruído térmico altera de forma mais significativa as características calculadas sobre o mesmo. Avaliando os canais degradados, para a combinação 16 ocorre a maior taxa de decréscimo na classificação para todos os métodos. Isso está associado ao fato de que essa degradação do sinal ocorre em 2/3 dos canais de aquisição, totalizando 8 canais com contaminação. Também se observa que para esse contaminante, os métodos SVM, LDA e KNN são afetados de uma forma quase constante para todas as combinações, mostrando que esses métodos apresentam uma alta sensibilidade a alterações

nos atributos de entrada e que os mesmos foram afetados de forma mais significativa pelo contaminante ruído térmico, do que os métodos PRF e RF, de forma a inviabilizar o modelo gerado (taxas de acerto médias de aproximadamente 6 a 8%). Quanto ao método PRF, ele prossegue com a mesma tendência do contaminante artefato de movimento, descrito previamente, apresentando um acréscimo no erro proporcional à maior quantidade de canais contaminados.

Por último, realizou-se a análise da classificação do sinal de sEMG com adição do contaminante perda de contato. No Apêndice A – Projeto de Experimentos encontram-se os gráficos de resíduos e da potência estatística deste experimento. Os fatores controláveis A, B e C correspondem, respectivamente, aos métodos de classificação, bases de dados e combinação de canais de sEMG contaminados. O projeto de experimentos foi avaliado quanto a sua potência estatística (Figura 42), concluindo-se que para uma potência de 95% é a diferença máxima das médias é de 1,37%. Além disso, o modelo foi considerado adequado para a análise da diferença entre médias, de acordo com a análise de resíduos apresentadas na Figura 42. O cálculo do projeto para o experimento de três fatores é apresentado na Tabela 43 e o mesmo demonstrou que as principais variações dos fatores referente às bases de dados e classificadores, configuração de canais com contaminantes e tipo de contaminante adicionado ao sinal são significativas, ou seja, as variações dos três principais efeitos são diferentes. As interações entre os fatores controláveis também são significativas, devido ao valor P ser inferior a 0,05.

Tabela 43 - Projeto de experimentos referente à análise dos métodos de classificação utilizando sinal de sEMG com contaminante perda de contato.

Fatores [†]	Graus de liberdade	Soma dos quadrados	Média Quadrada	F	P
A	4	1474,26	368,564	6080,92	0,000
B	2	4,62	2,308	38,09	0,000
C	7	58,99	8,427	139,04	0,000
A*B	8	15,27	1,909	31,49	0,000
A*C	28	143,51	5,126	84,57	0,000
B*C	14	1,38	0,098	1,62	0,065
A*B*C	56	4,75	0,085	1,40	0,026
Erro	46528	2820,06	0,061		
Total	46647	4937,91			

[†] Fator A: Métodos de classificação, Fator B: Base de dados e Fator C: Canais com contaminação

A análise de diferenças entre médias com 95% de confiabilidade em relação aos classificadores também foi realizada e está representada na Tabela 44. Nessa análise é possível afirmar que o método PRF obteve uma taxa de acerto média de 54,95%, sendo significativamente superior em comparação com todos os outros métodos propostos, considerando somente os sinais de sEMG contaminados por artefatos de movimento e todas as combinações de canais contaminados. O segundo melhor método em termos de taxa de acerto média foi o RF com uma taxa média de 34,3%. Também pode-se afirmar que todos os métodos SVM, LDA e KNN obtiveram uma taxa de acerto média sem diferenças significativas, com taxas aproximadas de 6%.

Tabela 44 - Análise das médias dos métodos utilizando método Tukey (95% de confiabilidade).

Métodos	Amostras	Médias	Grupo [†]
PRF	9384	54,93%	A
RF	9384	34,30%	B
KNN	9384	6,28%	C
LDA	9384	6,28%	C
SVM	9384	6,24%	C

[†] Médias que não compartilham uma letra são significativamente diferentes

Tabela 45 - Decréscimo na taxa de classificador de acordo com o método e canais contaminados por perda de contato eletrodo-pele.

Base de dados	Método	% decréscimo na taxa de acerto							
		configuração de canais contaminados							
		9	10	11	12	13	14	15	16
IEE	PRF	2,21%	1,00%	2,22%	0,67%	3,54%	3,77%	8,26%	20,43%
	RF	19,46%	17,39%	56,41%	24,65%	30,11%	69,34%	74,43%	78,41%
	SVM	91,62%	91,62%	91,62%	91,62%	91,62%	91,62%	91,62%	91,62%
	KNN	90,94%	90,94%	90,94%	90,94%	90,69%	90,94%	90,94%	90,94%
	LDA	91,38%	91,38%	91,38%	91,38%	91,32%	91,38%	91,38%	91,38%
NPI	PRF	2,51%	2,28%	3,16%	4,53%	4,76%	7,18%	10,41%	22,95%
	RF	11,98%	13,08%	53,39%	38,08%	24,78%	69,84%	74,93%	72,84%
	SVM	90,96%	90,96%	90,96%	90,96%	90,96%	90,96%	90,96%	90,96%
	KNN	90,50%	90,50%	90,35%	90,33%	90,40%	90,50%	90,45%	90,50%
	LDA	90,19%	90,09%	90,01%	90,24%	89,76%	90,03%	90,24%	90,24%
NPA	PRF	0,97%	3,97%	2,99%	3,82%	4,77%	5,26%	9,42%	19,80%
	RF	5,23%	11,85%	51,84%	38,57%	17,23%	67,26%	75,87%	76,96%
	SVM	82,10%	82,10%	88,47%	88,47%	82,10%	88,47%	88,47%	88,47%
	KNN	80,59%	81,33%	87,56%	87,56%	80,89%	87,56%	87,56%	88,02%
	LDA	82,33%	82,33%	87,85%	87,85%	82,33%	87,85%	87,85%	87,85%

A Tabela 45 apresenta o decréscimo calculado para a taxa de acerto levando-se em consideração o método, as bases de dados e as configurações de contaminação do sinal para o contaminante perda de contato. Em uma primeira análise, nota-se o efeito negativo do contaminante de perda de contato para os métodos LDA, KNN e SVM, nos quais observa-se

um valor elevado de decréscimo na taxa de acerto entre 80,59 e 91,62%, um decréscimo ainda mais acentuado que no caso do contaminante ruído térmico. Em comparação, assim como nos outros contaminante analisados, o método PRF apresentou taxas inferiores de decréscimo na classificação quando contrastado com os outros métodos e combinações de canais degradados, apresentando taxas de decréscimo entre 0,28% a 22,95%.

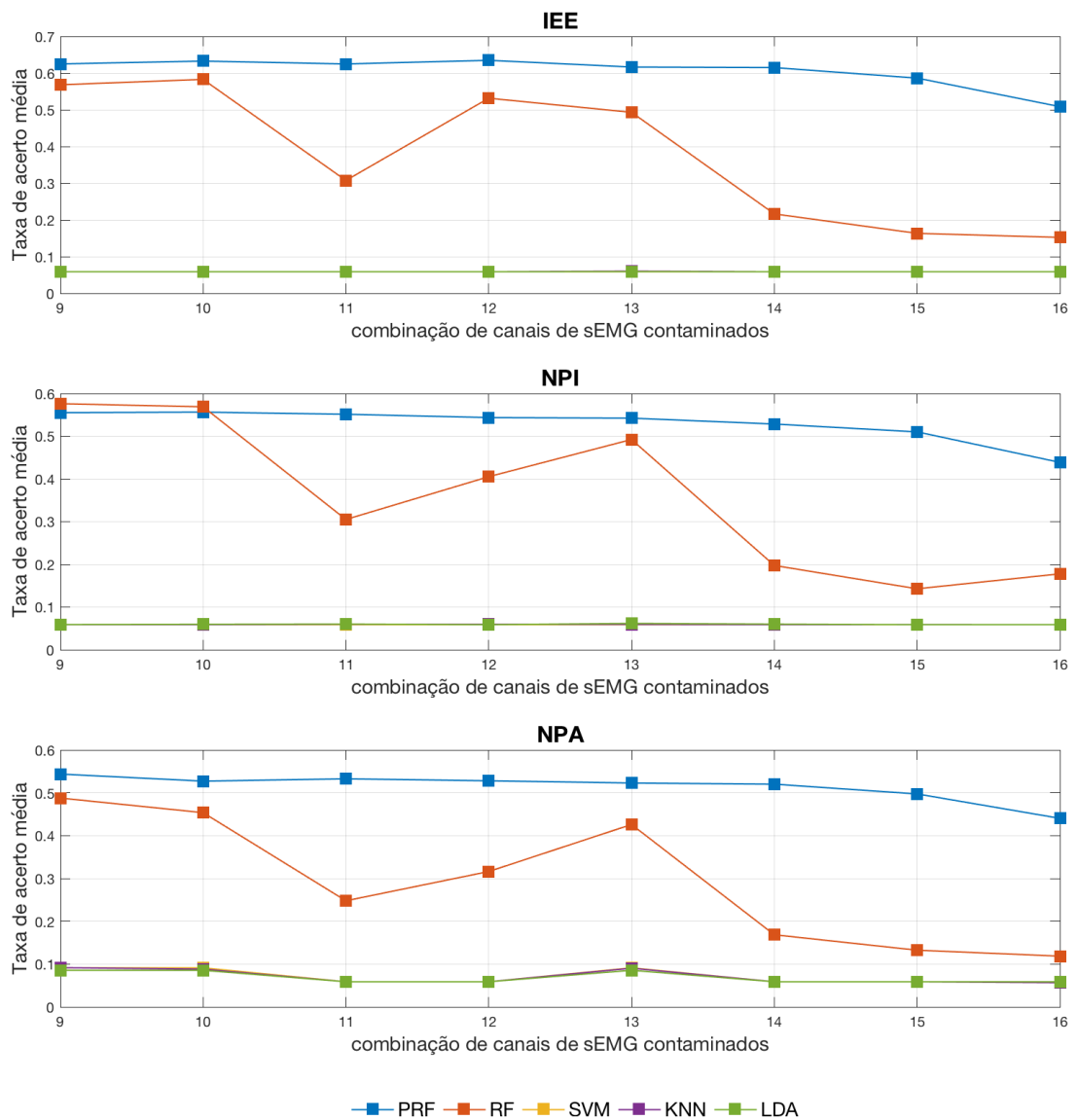
É interessante perceber que os contaminantes perda de contato e ruído térmico apresentaram uma tendência similar no decréscimo da taxa de acerto para todos os métodos avaliados. O contaminante artefato de movimento foi o que menos afetou o sinal e, como visto anteriormente, o mesmo apresenta uma taxa de acerto média significativamente superior aos outros contaminantes avaliados. Adicionalmente, o gráfico de tendências resultante dos dados da tabela também é apresentado para uma melhor visualização dos resultados na Figura 34.

Ao analisar a Figura 34, assim como observado para o contaminante ruído térmico, para o contaminante perda de contato os métodos de SVM, LDA e KNN apresentam uma taxa de acerto aproximadamente constante e são afetados de tal forma que o modelo é considerado inadequado para a classificação dos dados, com taxas de acerto médias aproximadas de 6%. Quanto ao método PRF, ele segue a mesma tendência do contaminante artefato de movimento e ruído térmico, descrita previamente, se mostrando mais tendencioso ao decréscimo da taxa de acerto em relação à quantidade de canais degradados, com taxas de acerto similares.

Considerando todos os resultados apresentados, é importante ressaltar que métodos de aprendizagem de máquina baseados em treinamento supervisionado dependem da qualidade dos dados de entrada para criar um modelo adequado do sistema. No entanto, nem sempre é possível prever todas as possíveis combinações de dados de entrada para gerar um modelo que consiga atuar sem decréscimos significativos na taxa de acerto em uma situação na qual ocorrem contaminações não previstas nos dados analisados. Por esses motivos é necessário evitar o *overfit* do modelo gerado, para que exista uma certa tolerância a essas situações. O *overfit* é uma condição na qual o modelo gerado tem uma correspondência muito forte com o conjunto de treinamento fornecido e, dessa forma, não consegue classificar amostras futuras. No entanto, como foi observado para os métodos KNN, LDA e SVM, no caso do modelo gerado se deparar com situações em que os atributos de entradas foram contaminados de uma forma a modificar de maneira significativa o sinal, o modelo se torna obsoleto, no sentido de que o mesmo não consegue mais realizar o reconhecimento dos padrão do sinal. Nesse caso é necessário realizar um novo treinamento do modelo, contabilizando os sinais degradados ou removendo-os do sistema, conforme análises de diversos estudos na área de processamento

de sinais de sEMG (CENE; FAVIEIRO; BALBINOT, 2015; DE MOURA; BALBINOT, 2018; ZHANG; HUANG, 2015).

Figura 34 - Gráfico da % de decréscimo na taxa de acerto considerando métodos e configurações de canais contaminados por perda de contato eletrodo-pele.



Fonte: Autor

No caso do método *Random Forest*, um conjunto de árvores não correlacionadas é gerado, utilizando a técnica de *bagging*, de forma a obter uma menor susceptibilidade a contaminantes no sinal de entrada, garantindo a não ocorrência de *overfit* nas Árvores de Decisão criadas. No entanto, como foi observado, se o atributo degradado for significativo para o modelo RF empregado, as taxas de acerto podem decair na ordem de 30 a 40%. Isso

ocorre porque quanto mais significativo o atributo é para o modelo, mais próximo do nó raiz ele se encontra, podendo gerar um erro que será propagado até o nó folha da árvore. O método RF tenta minimizar esse problema ao escolher um subconjunto aleatório de atributos para cada divisão do nó. No entanto, apesar do uso dessas técnicas para diminuir a correlação entre árvores, ainda assim observa-se um padrão do método RF, para todos os contaminantes, no decréscimo na taxa de acerto no que concerne a contaminação do canal sobre o músculo bíceps, que se mostrou ser significativo para a classificação do movimento.

A proposta de *Paraconsistent Random Forest* desenvolvida nesta tese, além de utilizar as técnicas de *bagging* e o subconjunto aleatório de atributos em cada divisão de nó, de forma similar à técnica RF, também acrescenta o conceito de que as divisões de cada nó não precisam ser binárias, no sentido de que um conjunto de atributos de entrada pode percorrer mais de um caminho na árvore. Além disso, ao se realizar uma Análise Paraconsistente dos dados é calculado um Grau de Certeza que é propagado entre os todos os nós até o nó folha, de forma a analisar todos os possíveis caminhos e atributos, conseguindo extrair o máximo de informação, mesmo com as possíveis inconsistências nos dados do modelo. Dessa forma é possível afirmar que a PRF apresenta uma tendência maior ao decréscimo da taxa de acerto média em relação à quantidade de atributos de entrada contaminados e menor em relação a qual atributo em si, como é o caso da RF.

Outro ponto importante, é que os resultados deste experimento confirmaram a hipótese de que o método PRF é um método robusto a contaminações nos sinais de sEMG, conseguindo mitigar os efeitos desses contaminantes na predição das classes de movimento, com decréscimos na taxa de acerto variando aproximadamente entre 1% e 20%, dependendo da quantidade de canais de sEMG contaminados. Outra vantagem apresentada, é que o método PRF não apresentou diferenças nas taxas de acerto médias entre os tipos de contaminantes, sendo possível concluir que o método é robusto, quando comparado a métodos amplamente usado, a degradações nos sinais de entrada de forma generalizada, não apresentando comportamento específico para os diferentes tipos de contaminantes.

No que diz respeito à complexidade computacional, o método PRF apresenta uma complexidade similar ao método *Random Forest*, utilizando o algoritmo ID3 como classificador base, durante a construção da estrutura da árvore. No entanto, no seu uso como classificador, o custo computacional para a classificação de uma amostra é proporcional ao tamanho da árvore gerada. Diferentemente de uma árvore ID3, na Árvore Paraconsistente desenvolvida, uma amostra percorre todos os caminhos possíveis gerando um valor por folha.

O método proposto apresenta como limitação a sua necessidade de um particionamento dos dados *a priori*, o que não ocorre no caso de uso de algoritmos como KNN, LDA e SVM. No caso do método RF, é realizado seu próprio particionamento dos dados utilizando o algoritmo CART. É importante ressaltar que durante os testes observou-se que o método PRF com sinais de sEMG, assim como já tinha sido constatado no experimento com dados padrão, é extremamente dependente dos métodos e parâmetros utilizados no particionamento de domínio *fuzzy*. Além disso, verificou-se uma dificuldade em encontrar um particionamento *a priori* dos sinais de sEMG em categorias *fuzzy* que conseguisse fornecer uma taxa de classificação satisfatória, especialmente devido à característica estocástica do sinal de sEMG. O sinal de sEMG apresenta uma grande variação das características dependendo da porção do movimento, como início, meio e fim, conforme já avaliado em outros estudos (CENE; BALBINOT, 2018). Essa dificuldade encontrada no particionamento dos dados afeta diretamente a taxa de classificação correta dos mesmos. Sendo assim, pode-se concluir que o uso de técnicas mais sofisticadas, que se adaptem melhor a esse tipo de sinal, pode resultar em um acréscimo na taxa de acerto do método proposto nesta tese.

5.4 DISCUSSÕES ENTRE OS EXPERIMENTOS E ESTUDOS SIMILARES

Primeiramente, é importante sublinhar que existe uma dificuldade em realizar comparações entre estudos de caracterização de movimentos através de sinais de sEMG pela falta de padronização em relação à quantidade de eletrodos, posicionamento dos mesmos, características extraídas do sinal, tamanho dos segmentos de análise, entre outros. Sabendo disso, nos estudos que envolvem contaminantes no sinal de sEMG escolheu-se empregar como métrica de comparação, quando possível, o decréscimo da taxa de acerto média por ser uma avaliação mais imparcial entre os métodos.

Considerando essas premissas, com o objetivo de comparar os resultados dos experimentos sEMG1 e sEMG2, é importante destacar as diferenças entre os métodos RAP e PRF aplicados na área de reconhecimento de movimentos através de sinais de sEMG. O método RAP proposto tem uma topologia fixa para a caracterização de movimentos utilizando 12 canais de aquisição de sEMG e até 3 características extraídas por canal. O método RAP também apresenta um valor fixo de amostras de treinamento (somente 3 amostras) e por causa dessa limitação, o método não é adequado para uso em sinais com segmentação de janelas deslizantes. Em comparação, a metodologia híbrida PRF, que une os conceitos de Lógica Paraconsistente e *Random Forest*, é um método versátil que pode ser utilizado em diversas

aplicações, conforme demonstrado no experimento BDP. Além disso, o método PRF não apresenta limitação quanto ao número de atributos de entrada, nem quanto à quantidade de amostras de treinamento.

Em relação à comparação dos métodos PRF e RAP com métodos tradicionais da área, a avaliação preliminar da Lógica Paraconsistente, através da Rede de Análise Paraconsistente, atingiu resultados próximos aos encontrados na literatura para a aplicação de reconhecimento de padrão baseadas em sinal de sEMG. No entanto, nos testes comparativos com os métodos RF, SVM, KNN e LDA, apesar de resultados semelhantes, o método RAP foi o que apresentou a menor taxa de acerto média. Na avaliação do método PRF, os resultados obtidos, considerando os sinais de sEMG sem contaminações, foram significativamente similares aos métodos KNN e LDA e com uma taxa de acerto média somente 4% inferior em comparação ao método RF. Além disso, considerando a base de dados de voluntários com amputação parcial do antebraço, o método PRF obteve taxas de acerto médias consideravelmente superiores aos métodos RF, SVM, LDA e KNN. Dessa forma, pode-se afirmar que o uso da Lógica Paraconsistente agregada a conceitos de aprendizagem de máquina, representada pelo método PRF, apresenta vantagens que tornam o método superior em relação à RAP na classificação de movimentos de membro superior. Em particular, os resultados obtidos pelo método PRF com bases de dados com contaminantes são promissores. A Lógica Paraconsistente cumpriu com seus princípios ao lidar eficientemente com os dados com contaminações, de forma que o decréscimo na taxa de acerto foi menor em comparação com outros algoritmos.

Na área de reconhecimento de movimentos, um ponto que merece destaque é a motivação por trás da escolha do classificador. Diversos estudos afirmam que a escolha do método para classificação de movimentos através de sinais de sEMG é feita de acordo com a facilidade de implementação do método (PHINYOMARK *et al.*, 2013; SAMUEL, 2019; XIAORONG ZHANG; HE HUANG; QING YANG, 2013; ZHANG; HUANG, 2015). Nesse sentido, o método LDA se destaca como um dos métodos mais utilizados na área por sua baixa parametrização (SAMUEL, 2019). Considerando essa premissa, é interessante ressaltar que o algoritmo proposto PRF só depende do parâmetro número de árvores para sua implementação e, em comparação com o método LDA, o método PRF obteve resultados significativamente similares considerando o sinal de sEMG original e resultados significativamente superiores em todos os casos de canais de sEMG contaminados. Sendo assim, é possível concluir que o método PRF se mostra como uma alternativa ao método LDA, considerando os aspectos de taxa de acerto média e baixa parametrização.

Além da escolha de um método simplificado para a classificação de movimentos, um dos focos das pesquisas atuais na área de classificação de movimentos do membro superior baseadas em sinais de sEMG é prover um sistema que mitigue os efeitos de contaminantes do sinal de sEMG na predição dos movimentos. Para atingir esse objetivo, foram propostas diversas soluções que envolvem desde sistemas que realizam um retreinamento do modelo de forma automática. O trabalho proposto por (CENE; FAVIEIRO; BALBINOT, 2015) apresenta o desenvolvimento de um sistema com Redes Neurais Auto Adaptativas em que o sistema é capaz de realizar um retreinamento para acompanhar contaminantes nos sinais, como deslocamento/remoção de eletrodos. De acordo com os resultados obtidos na análise de sinais de 5 voluntários distintos houve um acréscimo na taxa de classificação de 5,0-27,6% na utilização do sistema adaptativo, com a remoção de no máximo dois canais contaminados de sEMG. Na avaliação do método PRF, considerando a contaminação de perda de contato eletrodo-pele em até dois canais de sEMG, os resultados apresentados têm uma taxa de acerto média superior, variando entre 10% e 50% em comparação com outros métodos tradicionais considerados. Importante ressaltar que os sinais de sEMG, características extraídas e quantidade de canais de sEMG são diferentes entre os estudos, mas, realizando uma comparação indireta, o método PRF aparenta apresentar resultados melhores em termos de acréscimo na taxa de acerto do sistema. Outro ponto é que a solução proposta por (CENE; FAVIEIRO; BALBINOT, 2015) necessita realizar um constante monitoramento dos sinais de sEMG para identificar os melhores sinais de sEMG para serem utilizados como entrada da rede neural, além de ser necessário o armazenamento de um conjunto de dados de treinamento para um retreinamento contínuo do modelo. Essas demandas consomem tempo de processamento e memória, sendo outro ponto negativo em comparação com o algoritmo PRF, que aparenta ser mais robusto quanto a contaminantes no sinal de sEMG.

Em um estudo similar de (ZHANG; HUANG, 2015) foi apresentado um algoritmo LDA-adaptativo capaz de mitigar canais com contaminantes e realizar um rápido retreinamento do algoritmo de classificação ao reutilizar alguns parâmetros da LDA ao invés de realizar um novo treinamento do zero, o que trouxe um ganho significativo no desempenho do sistema. Nos testes realizados com contaminantes no sinal de sEMG, o uso do algoritmo LDA-adaptativo obteve um erro adicional de 2,8% (1 canal de sEMG contaminado), 6% (2 canais contaminados) e 15% (3 canais contaminados) em comparação com a classificação dos sinais de sEMG sem a presença de contaminantes. Conforme foi avaliado no experimento sEMG 2, o método PRF também apresenta uma queda menos significativa na taxa de acerto do modelo, comparado com outros classificadores tradicionais, com erro adicional

aproximado de no máximo 20% considerando 8 canais de sEMG com contaminações. Para somente 1 canal de sEMG com contaminação, o erro adicional do algoritmo PRF variou entre aproximadamente 1% e 4%. No entanto, apesar de o algoritmo LDA-adaptativo proposto por (ZHANG; HUANG, 2015) apresentar resultados similares ao método PRF em termos de decréscimo na taxa de acerto na presença de contaminantes, o mesmo ainda depende do monitoramento do sinal de sEMG para a presença de contaminantes para então realizar uma atuação direta e controlada sobre o modelo do classificador. Em comparação, o método PRF apresenta a clara vantagem de não necessitar atuar no modelo de forma direta, evitando um retreinamento desnecessário, na presença de sinais contaminados. Ainda assim, é interessante ressaltar que a solução proposta por (ZHANG; HUANG, 2015) tem vantagens em relação a outros estudos que exigem o monitoramento e atuação no modelo como o caso de (CENE; FAVIEIRO; BALBINOT, 2015; DE MOURA; BALBINOT, 2018), não somente pelo rápido retreino do modelo, mas por atuar em qualquer tipo de contaminante que degrade o sinal de sEMG, assim como o método PRF.

O estudo proposto por Moura (DE MOURA; BALBINOT, 2018) também visa mitigar o erro de classificação na presença de contaminantes no sinal de sEMG ao aplicar o conceito de sensor virtual. No caso, após a identificação de contaminantes no sinal de sEMG, o canal contaminado é substituído por um sinal modelado através dos métodos *Time-Varying Autoregressive Moving Average* (TVARMA) e *Time-Varying Kalman filter* (TVK). Esse sistema foi desenvolvido para a identificação de somente 5 contaminantes do sinal de sEMG: artefatos de movimento, saturação do sinal, deslocamentos de eletrodos, interferência na linha de força e interferência do ECG. Como a metodologia de testes desta tese é baseada em uma adaptação dos testes propostos por (DE MOURA; BALBINOT, 2018), criou-se uma tabela para uma comparação direta entre os resultados alcançados (Tabela 46) na classificação de movimentos através de sinais de sEMG obtido com voluntários com membro superior intacto. É importante ressaltar que nesta tese foi utilizada uma terminologia distinta dos termos definidos por (DE MOURA; BALBINOT, 2018) objetivando representar de forma mais fidedigna os modelos dos contaminantes em relação a outros estudos na literatura. Sendo assim, foi definido o uso do termo “perda de contato eletrodo-pele” em substituição ao termo “saturação do sinal” e “ruído térmico” em relação ao termo “deslocamento de eletrodo”.

Na tabela comparativa é possível concluir que, considerando os contaminantes artefato de movimento, ruído térmico (deslocamento de eletrodo) e perda do contato eletrodo-pele (saturação do sinal) e todas as configurações de canais contaminados, o método PRF obteve um erro adicional inferior ao se comparar com a proposta de substituição dos canais

contaminados por modelos TVARMA e TVK. Também foi analisada a proposta com retreinamento do modelo SVM com sinais modelados pelos métodos TVK e TVARMA e os resultados encontrados foram próximos aos obtidos pelo método PRF. No entanto, o uso de retreinamento do algoritmo SVM associado com a criação de um modelo do sinal de sEMG utilizando os métodos TVARMA e TVK tem um custo computacional elevado, inviabilizando essa solução para uso clínico. Nesse sentido, o uso do método PRF se mostra mais promissor pois não necessita de monitoração do sinal, não possui o custo computacional extra para realizar a substituição dos canais de sEMG por modelos matemáticos e não necessita de retreinamento constante. O método PRF também é robusto a qualquer tipo de contaminante que degrade o sinal de sEMG, não necessitando de um sistema à parte para identificar alguns contaminantes-alvo.

Tabela 46 - Comparação entre resultados do sistema de Sensor Virtual e o método PRF para classificação de movimento de voluntários com membro superior intacto.

Método	Artefato de Movimento	Decréscimo na taxa de acerto média (%)	
		Ruído Térmico (Deslocamento de Eletrodo)	Perda de Contato Eletrodo-Pele (Saturação do Sinal)
PRF	4,97%	4,61%	5,25%
SVM+TVARMA	42,19%	41,92%	41,79%
SVM+TVK	40,7%	40,84%	40,56%
Retreino com TVARMA	5,69%	5,69%	5,97%
Retreino com TVK	5,97%	5,83%	5,83%

Comparando o método PRF com outros estudos, cujo objetivo era mitigar os efeitos de contaminantes no sinal de sEMG na taxa de predição do modelo, conclui-se que o uso do método PRF apresenta uma alternativa interessante na área por conseguir atenuar o erro adicionado por degradações do sinal de sEMG com resultados similares aos reportados na literatura. Adicionalmente o método PRF ainda apresenta como vantagens não necessitar de um retreinamento do modelo e um sistema de avaliação contínua do sinal de sEMG. Além disso, o método PRF é um algoritmo de baixa parametrização e possui como característica ser robusto a contaminações no sinal de maneira não específica.

6 CONCLUSÕES E CONSIDERAÇÕES FINAIS

Neste trabalho, foram apresentadas algumas contribuições, as quais visam o desenvolvimento de um classificador que seja robusto a degradações no sinal de entrada causadas por contaminantes. Essas degradações do sinal, por sua vez, modificam as características dos sinais, dificultando a capacidade dos algoritmos de reconhecimento de padrão em discriminar classes, impactando diretamente na predição dos mesmos. O uso de Lógica Paraconsistente associado a algoritmos de aprendizagem de máquina teve por objetivo lidar diretamente com essas degradações nos sinais causadas por contaminantes externos, buscando mitigar os seus efeitos na taxa de acerto. Dentre diversas aplicações, uma das áreas que pode se beneficiar dessa característica está entre aquelas que envolvem sinais de sEMG, como por exemplo a classificação de movimentos do membro superior, cujos sinais são comumente afetados por contaminantes como artefato de movimento, perda de contato eletrodo-pele e ruído térmico. Tais efeitos impactam diretamente no uso clínico de próteses mioelétricas com controle natural e intuitivo.

Logo, a primeira contribuição desta tese foi o estudo e desenvolvimento de uma Rede de Análise Paraconsistente especificamente concebida para classificação de 17 movimentos distintos de membro superior utilizando 12 canais de sEMG. Essa rede foi proposta com o objetivo de efetuar uma análise preliminar da utilização da Lógica Paraconsistente como método de classificação. Na avaliação do método, o mesmo também foi comparado com outros classificadores conceituados na área de processamento de sinais mioelétricos, sendo eles, KNN, LDA, SVM e RF. De acordo com os resultados obtidos e análise estatística, a Lógica Paraconsistente foi bem-sucedida, como método preliminar, em tratar sinais mioelétricos, com uma taxa de acerto média próxima ao método SVM. No entanto, foi observada a necessidade de utilizar um algoritmo mais adaptável a diferentes atributos de entrada e aplicações. Com isso, foi proposta uma técnica híbrida que agregasse as vantagens da Lógica Paraconsistente com aprendizado de máquina baseado em *Random Forest*.

Assim, a segunda contribuição dessa tese foi o estudo, concepção e desenvolvimento do algoritmo *Paraconsistent Random Forest* que une os conceitos de *Random Forest* e Lógica Paraconsistente. A *Paraconsistent Random Forest* se mostrou um algoritmo robusto a degradações nos dados de entrada, adaptável a diferentes aplicações e de baixa parametrização. Para validar o método, o mesmo foi testado com diversas bases de dados padrão na área de aprendizagem de máquina, obtendo resultados similares a métodos como SVM, KNN, LDA e RF. Adicionalmente, vários procedimentos experimentais foram

empregados para analisar a viabilidade e robustez do método quanto a degradação dos atributos de entrada, aplicando os conceitos de redução artificial de informação, através de dados ausentes e com adição de ruído aleatório uniformemente distribuído. Através da análise estatística dos resultados foi observado que, ao aplicar redução de informação através de dados ausentes, o método PRF foi significativamente superior aos métodos LDA, KNN e SVM, com decréscimos máximos na taxa de acerto de 24% quando 90% dos dados estão ausentes. Quando comparado com o método RF, os resultados foram similares. No entanto, na análise com dados ruidosos, o método PRF se sobressaiu a todos os outros métodos comparados, apresentando decréscimos de apenas 31% na taxa de acerto quando adicionado ruído em 90% dos dados de entrada. Logo, é possível afirmar que o método PRF desenvolvido apresentou vantagens que permitem lidar de forma mais eficiente com dados ruidosos e dados ausentes, típicos de aplicações que envolvem sinais físicos, resultando em menores decréscimos na taxa de acerto em relação a outros métodos.

Por último, a contribuição final deste trabalho está no uso da *Paraconsistent Random Forest* aplicada ao reconhecimento de 17 movimentos do braço através de sinais mioelétricos. A aplicação da *Paraconsistent Random Forest* na área de reconhecimento de movimentos visou analisar possíveis benefícios que a Lógica Paraconsistente pudesse agregar no processamento de dados com degradações recorrentes, levando a uma redução significativa na taxa de acerto que, conforme demonstrado nos resultados, em alguns casos chega a ser próxima a 90% para os métodos SVM e LDA. Conforme resultados e análises estatísticas, o método PRF se mostrou promissor para o uso em aplicações que dependem de sinais de sEMG por conseguir tratar as inconsistências inerentes à presença de contaminantes, resultando em melhores soluções para a classificação em comparação com os métodos RF, LDA, KNN e SVM. O método PRF conseguiu mitigar os efeitos desses contaminantes na predição das classes de movimento, com decréscimos na taxa de acerto variando aproximadamente entre 1% e 20%, dependendo da quantidade de canais de sEMG contaminados. Também foi constatado nos resultados que a redução na taxa de acerto tem uma maior correlação com a quantidade de canais contaminados ao invés de canais específicos, como é o caso do método *Random Forest*. Outra vantagem observada nos testes, é que o método PRF não apresentou diferenças nas taxas de acerto médias entre os contaminantes avaliados, sendo possível concluir que o método é robusto a degradações nos sinais de entrada de forma generalizada, não apresentando comportamento específico para diferentes tipos de contaminantes.

Logo, o método desenvolvido PRF contribuiu como uma solução alternativa a métodos que necessitem de retreinamento contínuo devido a degradações do modelo de classificação

causadas por contaminantes de sinal de sEMG na área de reconhecimento de movimentos do membro superior. Além disso, avaliação do método PRF em outras áreas é recomendada, especialmente as que envolvem aquisição de sinais físicos sujeitos a degradações causadas por fatores externos como, por exemplo, aplicações que compreendem o uso de sinais de EEG.

No entanto, existem aspectos que ainda necessitam de melhoria, principalmente a necessidade de um particionamento de domínio dos dados *a priori*. A dificuldade em definir as partições de domínio, especialmente para sinais estocásticos, pode acabar prejudicando o uso do método, apesar da sua vantagem comprovada em lidar com sinais contaminados. Portanto, a continuidade desta pesquisa envolve o estudo de novos métodos para particionamento de domínio *fuzzy* para dados estocásticos com o objetivo de elevar a taxa de acerto do método. Também é interessante avaliar o uso de algoritmos de *pruning* e outros algoritmos de tomada de decisão e seu impacto na predição do classificador.

Adicionalmente, considerando a área de reconhecimento de movimentos de membro superior, seria interessante realizar uma avaliação de outras técnicas que se utilizadas em conjunto com o PRF pudessem impactar na predição do modelo, aumentando as taxas de acerto na presença de contaminantes, como o método de extração de características proposto por (PAN *et al.*, 2015), ou técnicas voltadas pro treinamento supervisionado (HARGROVE; ENGLEHART; HUDGINS, 2008; WOODWARD; HARGROVE, 2019). Outra tendência é o uso de eletromiografia de alta resolução (GENG *et al.*, 2018; PAN *et al.*, 2015; STANGO; NEGRO; FARINA, 2015).

Também é importante ressaltar que para o uso do método *Paraconsistent Random Forest* em ensaios clínicos, é de extrema importância avaliar o desempenho do método em processamento *online*, preferencialmente em ensaios de longa duração. Nessa linha de pesquisa também se mostra interessante analisar o desempenho do método em relação a degradações no sinal de sEMG causadas por fadiga muscular, pois já foi amplamente reportado na literatura que a fadiga muscular altera algumas características do sinal de sEMG (ARJUNAN *et al.*, 2011; MERLETTI *et al.*, 1999; SONG *et al.*, 2009).

REFERÊNCIAS

- ABE, Jair Minoro. Paraconsistent logics and applications. In: INTERNATIONAL WORKSHOP ON SOFT COMPUTING APPLICATIONS 2010, 4., Arad. **Proceedings...** New York: IEEE, 2010. p. 11-18.
- ABE, Jair Minoro. **Aspectos de computação inteligente paraconsistente**. 1. ed. São Paulo: Instituto de Estudos Avançados da USP, 2013.
- ABE, Jair Minoro; AKAMA, Seiki; NAKAMATSU, Kazumi. **Introduction to Annotated Logics**. Cham: Springer International Publishing, 2015.
- ABE, Jair Minoro; LOPES, Helder F. S.; NAKAMATSU, Kazumi. Paraconsistent neurocomputing and brain signal analysis. **Vietnam Journal of Computer Science**, Vietnam, v. 1, n. 4, p. 219–230, 2014.
- ADEWUYI, Adenike A. *et al.* An Analysis of Intrinsic and Extrinsic Hand Muscle EMG for Improved Pattern Recognition Control. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, New York, v. 24, n. 4, p. 485–494, 2016.
- AHMADIZADEH, Chakaveh *et al.* Toward Intuitive Prosthetic Control: Solving Common Issues Using Force Myography, Surface Electromyography, and Pattern Recognition in a Pilot Case Study. **IEEE Robotics and Automation Magazine**, New York, v. 24, n. 4, p. 102–111, 2017.
- AJIBOYE, A. B.; WEIR, R. F. H. A Heuristic Fuzzy Logic Approach to EMG Pattern Recognition for Multifunctional Prosthesis Control. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, New York, v. 13, n. 3, p. 280–291, 2005.
- AKAMA, Seiki. **Towards Paraconsistent Engineering**. Kawasaki: Springer International Publishing, 2016.
- ALKAN, Ahmet; GÜNAY, Mücahid. Identification of EMG signals using discriminant analysis and SVM classifier. **Expert Systems with Applications**, New York, v. 39, n. 1, p. 44–47, 2012.
- ANTFOLK, Christian *et al.* Artificial Redirection of Sensation From Prosthetic Fingers to the Phantom Hand Map on Transradial Amputees: Vibrotactile Versus Mechanotactile Sensory Feedback. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, New York, v. 21, n. 1, p. 112–120, 2013.
- ATZORI, Manfredo *et al.* Building the Ninapro database: A resource for the biorobotics community. In: IEEE RAS & EMBS INTERNATIONAL CONFERENCE ON BIOMEDICAL ROBOTICS AND BIOMECHATRONICS (BIOROB), 4. , 2012, Roma. **Proceedings...** New York: IEEE, 2012. p. 1258-1265.
- ATZORI, Manfredo *et al.* Electromyography data for non-invasive naturally-controlled robotic hand prostheses. **Scientific Data**, New York, v. 1, Não paginado, 2014.
- ATZORI, Manfredo *et al.* Characterization of a Benchmark Database for Myoelectric

Movement Classification. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, New York, v. 23, n. 1, p. 73–83, 2015.

BALBINOT, Alexandre; FAVIEIRO, Gabriela. A Neuro-Fuzzy System for Characterization of Arm Movements. **Sensors**, Basel, Switzerland, v. 13, n. 2, p. 2613–2630, 2013.

BALBINOT, Alexandre; JÚNIOR, Adalberto Schuck; FAVIEIRO, Gabriela Winkler. Decoding Arm Movements by Myoelectric Signal and Artificial Neural Networks. **Intelligent Control and Automation**, [s. l.], n. 4, p. 87–93, 2013.

BARBON JÚNIOR, Sylvio; GUIDO, Rodrigo Capobianco; VIEIRA, Lucimar Sasso. A Neural-Network Approach for Speech Features Classification Based on Paraconsistent Logic. In: **IEEE INTERNATIONAL SYMPOSIUM ON MULTIMEDIA**, 11., 2009, San Diego. **Proceedings...** New York: IEEE, 2009. p. 567-570

BASMAJIAN, J. V.; DE LUCA, C. J. **Muscles alive**. 5th. ed. Baltimore: Williams and Wilkins, 1985.

BETTHAUSER, Joseph L. *et al.* Limb Position Tolerant Pattern Recognition for Myoelectric Prosthesis Control with Adaptive Sparse Representations From Extreme Learning. **IEEE Transactions on Biomedical Engineering**, New York, v. 65, n. 4, p. 770–778, 2018.

BLAIR, Howard A.; SUBRAHMANIAN, V. S. Paraconsistent logic programming. **Theoretical Computer Science**, Amsterdam, v. 68, n. 2, p. 135–154, 1989.

BONISSONE, Piero *et al.* A fuzzy random forest. **International Journal of Approximate Reasoning**, New York, v. 51, n. 7, p. 729–747, 2010.

BREIMAN, Leo. Random forests. **Machine Learning**, Ney York, v. 45, n. 1, p. 5–32, 2001.

BUHRMAN, Harry; WOLF, Ronald De. Complexity Measures and Decision Tree Complexity: A Survey. **Theoretical Computer Science**, Amsterdam, v. 288, p. 21–43, 2002.

CARNIELLI, Walter; CONIGLIO, Marcelo Esteban. **Paraconsistent Logic: Consistency, Contradiction and Negation**. New York: Springer International Publishing, 2016

CENE, V. H.; FAVIEIRO, G.; BALBINOT, A. Upper-limb movement classification based on sEMG signal validation with continuous channel selection. In: **2015 37TH ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY (EMBC)**, 37., 2015, Milão. **Proceedings...** New York: IEEE, 2015. p. 486-489.

CENE, VINÍCIUS HORN. **Desenvolvimento de um projeto de experimentos para a caracterização de sinais mioelétricos através do uso de regressão logística**. 2016. Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2016.

CENE, Vinicius Horn; BALBINOT, Alexandre. Upper-limb movement classification through logistic regression sEMG signal processing. In: **LATIN AMERICA CONGRESS ON COMPUTATIONAL INTELLIGENCE (LA-CCI)**, 2015, Curitiba. **Preceedings...** New

York: IEEE, 2015. p. 1-4.

CENE, Vinicius Horn; BALBINOT, Alexandre. Using the sEMG signal representativity improvement towards upper-limb movement classification reliability. **Biomedical Signal Processing and Control**, Amsterdam, v. 46, n. 1, p. 182–191, 2018.

CHAN, Francis H. Y. *et al.* Fuzzy EMG classification for prosthesis control. **IEEE Transactions on Rehabilitation Engineering**, New York, v. 8, n. 3, p. 305–311, 2000.

CLARKE, Bertrand; FOKOUE, Ernest; ZHANG, Hao Helen. **Principles and Theory for Data Mining and Machine Learning**. New York: Springer, 2015.

COBREROS, Pablo. Paraconsistent vagueness: a positive argument. **Synthese**, New York, v. 183, n. 2, p. 211–227, 2011.

DA COSTA, Newton C. A. On the theory of inconsistent formal systems. **Notre Dame Journal of Formal Logic**, Durham, v. 15, n. 4, p. 497–510, 1974.

DA SILVA FILHO, João Inácio *et al.* Paraconsistent analysis network applied in the treatment of Raman spectroscopy data to support medical diagnosis of skin cancer. **Medical & Biological Engineering & Computing**, New York, v. 54, n. 10, p. 1453–1467, 2016.

DA SILVA FILHO, João Inácio; ABE, Jair Minoru; TORRES, Germano Lambert. **Inteligência Artificial com as Redes de Análises Paraconsistentes**. 1. ed. Rio de Janeiro: LTC, 2008.

DA SILVA FILHO, João Inácio. Algoritmo Paraconsistente Extrator de efeitos da Contradição - ParaExtrctr. **Seleção documental**, São Paulo, v. 15, n. 4, p. 21–25, 2009.

DA SILVA FILHO, João Inácio. Treatment of Uncertainties with Algorithms of the Paraconsistent Annotated Logic. **Journal of Intelligent Learning Systems and Applications**, Wuhan, China, v. 4, n. 2, p. 144–153, 2012.

DE LUCA, C. J. Physiology and mathematics of myoelectric signals. **IEEE transactions on bio-medical engineering**, New York, v. 26, n. 6, p. 313–325, 1979.

DE LUCA, Carlo J. The use of surface electromyography in biomechanics. **Journal of Applied Biomechanics**, Champaign, v. 13, n. 2, p. 135–163, 1997.

DE MOURA, Karina; BALBINOT, Alexandre. Virtual Sensor of Surface Electromyography in a New Extensive Fault-Tolerant Classification System. **Sensors**, Basel, v. 18, n. 5, Não paginado, 2018.

DIETTERICH, Thomas G. An Experimental Comparison of Three Methods for Constructing Ensembles of Decision Trees. **Machine Learning**, New York, v. 40, p. 139–157, 2000.

DUNIN-KEPLICZ, Barbara *et al.* Paraconsistent semantics of speech acts. **Neurocomputing**, Amsterdam, v. 151, p. 943–952, 2015.

ENDERLE, John; BRONZINO, Joseph. **Introduction to Biomedical Engineering**. 3. ed. Burlington: Elsevier, 2012.

- ENGLEHART, Kevin; HUDGINS, Bernard. A robust, real-time control scheme for multifunction myoelectric control. **IEEE Transactions on Biomedical Engineering**, New York, v. 50, n. 7, p. 848–854, 2003.
- FARINA, Dario *et al.* The Extraction of Neural Information from the Surface EMG for the Control of Upper-Limb Prostheses: Emerging Avenues and Challenges. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, New York, v. 22, n. 4, p. 797–809, 2014.
- FARRELL, Todd R. Determining delay created by multifunctional prosthesis controllers. **The Journal of Rehabilitation Research and Development**, Washington D.C., v. 48, n. 6, 2011.
- FARRELL, Todd R.; WEIR, Richard F. The optimal controller delay for myoelectric prostheses. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, New York, v. 15, n. 1, p. 111–118, 2007.
- FAVIEIRO, G. W.; BALBINOT, Alexandre. Adaptive neuro-fuzzy logic analysis based on myoelectric signals for multifunction prosthesis control. In: 2011 ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY, 33., 2011, Boston. **Proceedings...** New York: IEEE, 2011. p. 7888-7891.
- FAVIEIRO, Gabriela. **Desenvolvimento De Um Sistema Neuro-Fuzzy Para Análise De Sinais Mioelétricos Do Segmento Mão-Braço**. 2012. Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2012.
- FAVIEIRO, Gabriela W.; MOURA, Karina O. A.; BALBINOT, Alexandre. Novel method to characterize upper-limb movements based on paraconsistent logic and myoelectric signals. In: 2016 ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY (EMBC), 38., 2016, Orlando. **Proceedings...** New York: IEEE, 2016. p. 395-398.
- FAVIEIRO, Gabriela Winkler; BALBINOT, Alexandre; BARRETO, Mara M. G. Decoding arm movements by myoelectric signals and artificial neural networks. In: ISSNIP BIOSIGNALS AND BIROBOTICS CONFERENCE, 4., 2011, Milão. **Proceedings...** New York: IEEE, 2011. p 1-4.
- FRASER, Graham D. *et al.* Automated Biosignal Quality Analysis for Electromyography Using a One-Class Support Vector Machine. **IEEE Transactions on Instrumentation and Measurement**, New York, v. 63, n. 12, p. 2919–2930, 2014.
- FURUKAWA, Jun-ichiro *et al.* Human Movement Modeling to Detect Biosignal Sensor Failures for Myoelectric Assistive Robot Control. **IEEE Transactions on Robotics**, New York, v. 33, n. 4, p. 846–857, 2017.
- GENG, Yanjuan *et al.* A Robust Sparse Representation Based Pattern Recognition Approach for Myoelectric Control. **IEEE Access**, New York, v. 6, p. 38326–38335, 2018.
- GENG, Yanjuan; ZHOU, Ping; LI, Guanglin. Toward attenuating the impact of arm positions on electromyography pattern-recognition based motion classification in transradial

- amputees. **Journal of NeuroEngineering and Rehabilitation**, New York, v. 9, n. 1, p. 74, 2012.
- GIJSBERTS, Arjan *et al.* Movement Error Rate for Evaluation of Machine Learning Methods for sEMG-Based Hand Movement Classification. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, New York, v. 22, n. 4, p. 735–744, 2014.
- GU, Yikun *et al.* Robust EMG pattern recognition in the presence of confounding factors: features, classifiers and adaptive learning. **Expert Systems with Applications**, Amsterdam, v. 96, p. 208–217, 2018.
- GUIDO, Rodrigo Capobianco *et al.* Introducing the Discriminative Paraconsistent Machine (DPM). **Information Sciences**, Amsterdam, v. 221, p. 389–402, 2013.
- GUO, Shuxiang *et al.* Comparison of sEMG-Based Feature Extraction and Motion Classification Methods for Upper-Limb Movement. **Sensors**, Basel, v. 15, n. 4, p. 9022–9038, 2015.
- GUYTON, Arthur C.; HALL, John E. **Tratado de Fisiología Médica**. 12. ed. Rio de Janeiro: Elsevier, 2012.
- HARGROVE, Levi J.; ENGLEHART, Kevin; HUDGINS, Bernard. A comparison of surface and intramuscular myoelectric signal classification. **IEEE transactions on biomedical engineering**, New York, v. 54, n. 5, p. 847–53, 2007.
- HOPNER, F.; KLAWONN, F. A New Approach to Fuzzy Partitioning. In: IFSA WORLD CONGRESS AND NAFIPS INTERNATIONAL CONFERENCE, 20., 2001, Vancouver, **Proceedings...** New York: IEEE, 2011. p. 1419–1424.
- HUANG, He *et al.* Design of a robust EMG sensing interface for pattern classification. **Journal of Neural Engineering**, [s. l.], v. 7, p. 1–10, 2010.
- HUDGINS, B.; PARKER, P.; SCOTT, R. N. A new strategy for multifunction myoelectric control. **IEEE transactions on biomedical engineering**, New York, v. 40, n. 1, p. 82–94, 1993.
- JANIKOW, C. Z. Fuzzy decision trees: issues and methods. **IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)**, New York, v. 28, n. 1, Não paginado, 1998.
- KHEZRI, Mahdi; JAHED, Mehran. A Neuro–Fuzzy Inference System for sEMG-Based Identification of Hand Motion Commands. **IEEE Transactions on Industrial Electronics**, New York, v. 58, n. 5, p. 1952–1960, 2011.
- KHUSHABA, Rami N.; AL-ANI, Ahmed; AL-JUMAILY, Adel. Orthogonal fuzzy neighborhood discriminant analysis for multifunction myoelectric hand control. **IEEE Transactions on Biomedical Engineering**, New York, v. 57, n. 6, p. 1410–1419, 2010.
- KWON, Suncheol; KIM, Jung. Real-Time Upper Limb Motion Estimation From Surface Electromyography and Joint Angular Velocities Using an Artificial Neural Network for Human Machine Cooperation. **IEEE Transactions on Information Technology in Biomedicine**, New York, v. 15, n. 4, p. 522–530, 2011.

- LI, Guanglin; SCHULTZ, Aimee E.; KUIKEN, Todd A. Quantifying pattern recognition-based myoelectric control of multifunctional transradial prostheses. **IEEE transactions on neural systems and rehabilitation engineering**, New York, v. 18, n. 2, p. 185–92, 2010.
- LICHMAN, M. **UCI Machine Learning Repository**. Irvine: University of California, School of Information and Computer Science, 2018.
- LIU, Jianwei *et al.* Towards Zero Re-training for Myoelectric Control Based on Common Model Component Analysis. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, New York, v. 11, n. 4, Não paginado, 2015.
- LORRAIN, T.; JIANG, N.; FARINA, D. Surface EMG classification during dynamic contractions for multifunction transradial prostheses. In: ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY, 3., 2010, Buenos Aires. **Proceedings...** New York: IEEE, 2010. p. 2766-2769.
- MÁRIO, Maurício C. Sistema Classificador de Sinais Projetado com Redes Neurais Artificiais Paraconsistentes. **Revista Seleção Documental**, São Paulo, v. 7, n. 2, p. 19– 24, 2007.
- MASTINU, Enzo *et al.* An Alternative Myoelectric Pattern Recognition Approach for the Control of Hand Prostheses: A Case Study of Use in Daily Life by a Dysmelia Subject. **IEEE Journal of Translational Engineering in Health and Medicine**, New York, v. 6, Não paginado, 2018.
- MATSUBARA, Takamitsu; MORIMOTO, Jun. Bilinear modeling of EMG signals to extract user-independent features for multiuser myoelectric interface. **IEEE Transactions on Biomedical Engineering**, New York, v. 60, n. 8, p. 2205–2213, 2013.
- MCCOOL, Paul *et al.* Identification of contaminant type in surface electromyography (EMG) signals. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, New York, v. 22, n. 4, p. 774–783, 2014.
- MOMEN, Kaveh; KRISHNAN, Sridhar; CHAU, Tom. Real-time classification of forearm electromyographic signals corresponding to user-selected intentional movements for multifunction prosthesis control. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, New York, v. 15, n. 4, p. 535–542, 2007.
- MONSIFROT, J. *et al.* Online estimation of EMG signals model based on a renewal process. In: IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING, 1. ,2013, Vancouver. **Proceedings...** New York: IEEE, 2013. p. 944-948.
- MONTGOMERY, Douglas C. **Design and Analysis of Experiments**. 8th. ed. New Jersey: John Wiley & Sons, 2012.
- MOURA, Karina O. A.; FAVIEIRO, Gabriela W.; BALBINOT, Alexandre. Support vectors machine classification of surface electromyography for non-invasive naturally controlled hand prostheses. In: ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY (EMBC), 38., 2016, Orlando. **Proceedings...** New York: IEEE, 2016. p. 788-791.

MUCELI, Silvia; JIANG, Ning; FARINA, Dario. Extracting signals robust to electrode number and shift for online simultaneous and proportional myoelectric control by factorization algorithms. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, New York, v. 22, n. 3, p. 623–633, 2014.

NAIK, Ganesh R.; AL-TIMEMY, Ali H.; NGUYEN, Hung T. Transradial Amputee Gesture Classification Using an Optimal Number of sEMG Sensors: An Approach Using ICA Clustering. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, New York, v. 24, n. 8, p. 837–846, 2016.

ORTIZ-CATALAN, Max; BRÅNEMARK, Rickard; HÅKANSSON, Bo. BioPatRec: A modular research platform for the control of artificial limbs based on pattern recognition algorithms. **Source Code for Biology and Medicine**, New York, v. 8, Não paginado, 2013.

OSKOEI, M. a.; HU, Huosheng Hu Huosheng. Support Vector Machine-Based Classification Scheme for Myoelectric Control Applied to Upper Limb. **IEEE Transactions on Biomedical Engineering**, New York, v. 55, n. 8, p. 1956–1965, 2008.

OSKOEI, Mohammadreza Asghari; HU, Huosheng. GA-based Feature Subset Selection for Myoelectric Classification. In: INTERNATIONAL CONFERENCE ON ROBOTICS AND BIOMIMETICS, Kunming, 2006, **Proceedings...** New York: IEEE, 2006. p. 1465-1470.

PAN, Lizhi et al. Improving robustness against electrode shift of high density EMG for myoelectric control through common spatial patterns. **Journal of NeuroEngineering and Rehabilitation**, New York, v.12, n. 110, p. 1–16, 2015.

PANCHOLI, Sidharth; JOSHI, Amit M. Portable EMG Data Acquisition Module for Upper Limb Prosthesis Application. **IEEE Sensors Journal**, New York, v. 18, n. 8, Não paginado, 2018.

PHINYOMARK, Angkoon *et al.* EMG feature evaluation for improving myoelectric pattern recognition robustness. **Expert Systems with Applications**, Amsterdam, v. 40, n. 12, p. 4832–4840, 2013.

QIAN, Huimin *et al.* Recognition of human activities using SVM multi-class classifier. **Pattern Recognition Letters**, Amsterdam, v. 31, n. 2, p. 100–111, 2010.

QUINLAN, J. R. induction fo decion trees. **Machine Learning**, New York, v. 1, p. 81– 106, 1986.

RAURALE, Sumit A. Acquisition and processing real-time EMG signals for prosthesis active hand movements. In: INTERNATIONAL CONFERENCE ON GREEN COMPUTING COMMUNICATION AND ELECTRICAL ENGINEERING (ICGCCEE), 2014, Coimbatore. **Proceedings...** New York: IEEE, 2014. p 1-6.

RIILLO, F. *et al.* Optimization of EMG-based hand gesture recognition: Supervised vs. unsupervised data preprocessing on healthy subjects and transradial amputees. **Biomedical Signal Processing and Control**, Amsterdam, v. 14, n. 1, p. 117–125, 2014.

ROSSI, Matteo *et al.* Hybrid EMG classifier based on HMM and SVM for hand gesture recognition in prosthetics. In: INTERNATIONAL CONFERENCE ON INDUSTRIAL TECHNOLOGY (ICIT), Seville, 2015, **Proceedings...** New York: IEEE, 2015. p. 1700-

1705

SAFAVIAN, S. Rasoul; LANDGREBE, David. A Survey of Decision Tree Classifier Methodology. **IEEE Transactions on Systems, Man and Cybernetics**, New York, v. 21, n. 3, p. 660–674, 1991.

SAMUEL, Oluwarotimi Williams. Intelligent EMG Pattern Recognition Control Method for Upper-Limb Multifunctional Prostheses : Advances , Current Challenges , and Future Prospects. **IEEE Access**, New York, v. 7, p. 10150–10165, 2019.

SENSINGER, J. W.; LOCK, B. A.; KUIKEN, T. A. Adaptive Pattern Recognition of Myoelectric Signals: Exploration of Conceptual Framework and Practical Algorithms. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, New York, v. 17, n. 3, p. 270–278, 2009.

SHENOY, Pradeep *et al.* Online Electromyographic Control of a Robotic Prosthesis. **IEEE Transactions on Biomedical Engineering**, New York, v. 55, n. 3, p. 1128– 1135, 2008.

SOUZA, Sheila; ABE, Jair Minoro; NAKAMATSU, Kazumi. MICR Automated Recognition based on Paraconsistent Artificial Neural Networks. **Procedia Computer Science**, [s. l.], v. 22, p. 1083–1091, 2013.

STANGO, Antonietta; NEGRO, Francesco; FARINA, Dario. Spatial Correlation of High Density EMG Signals Provides Features Robust to Electrode Number and Shift in Pattern Recognition for Myocontrol. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, New York, v. 23, n. 2, p. 189–198, 2015.

TORTORA, G. J. **Corpo Humano: Fundamentos de Anatomia e Fisiologia**. 4. ed. Porto Alegre: Artmed, 2012.

TOSIN, M. C. *et al.* SEMG feature selection and classification using SVM-RFE. In: ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY, 39., Jeju, 2017, **Proceedings...**New York: IEEE, 2017. p. 390-393.

TOSIN, Mauricio C. C. **Aplicação de Algoritmos de Seleção de Características na Classificação de Movimentos do Segmento Mão-Braço Através do Processamento de Sinais de sEMG**. 2018. Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2018.

UMANOL, M. *et al.* Fuzzy decision trees by fuzzy ID3 algorithm and its application to diagnosis systems. In: IEEE INTERNATIONAL FUZZY SYSTEMS CONFERENCE, 3rd, 1994, Orlando. **Proceedings...** New York: IEEE, 1994. p. 2113-2118.

VIDOVIC, Marina M. C. *et al.* Improving the robustness of myoelectric pattern recognition for upper limb prostheses by covariate shift adaptation. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, New York, v. 24, n. 9, p. 961– 970, 2016.

WANG, Anna *et al.* A novel pattern recognition algorithm: Combining ART network with SVM to reconstruct a multi-class classifier. **Computers and Mathematics with Applications**, [s. l.], v. 57, n. 11–12, p. 1908–1914, 2009.

WANG, Xianchang *et al.* Fuzzy rule based decision trees. **Pattern Recognition**, [s. l.], v. 48, n. 1, p. 50–59, 2015.

WEBB, Andrew R.; COPSEY, Keith D. **Statistical Pattern Recognition**. 3. ed. Chichester: Wiley, 2011.

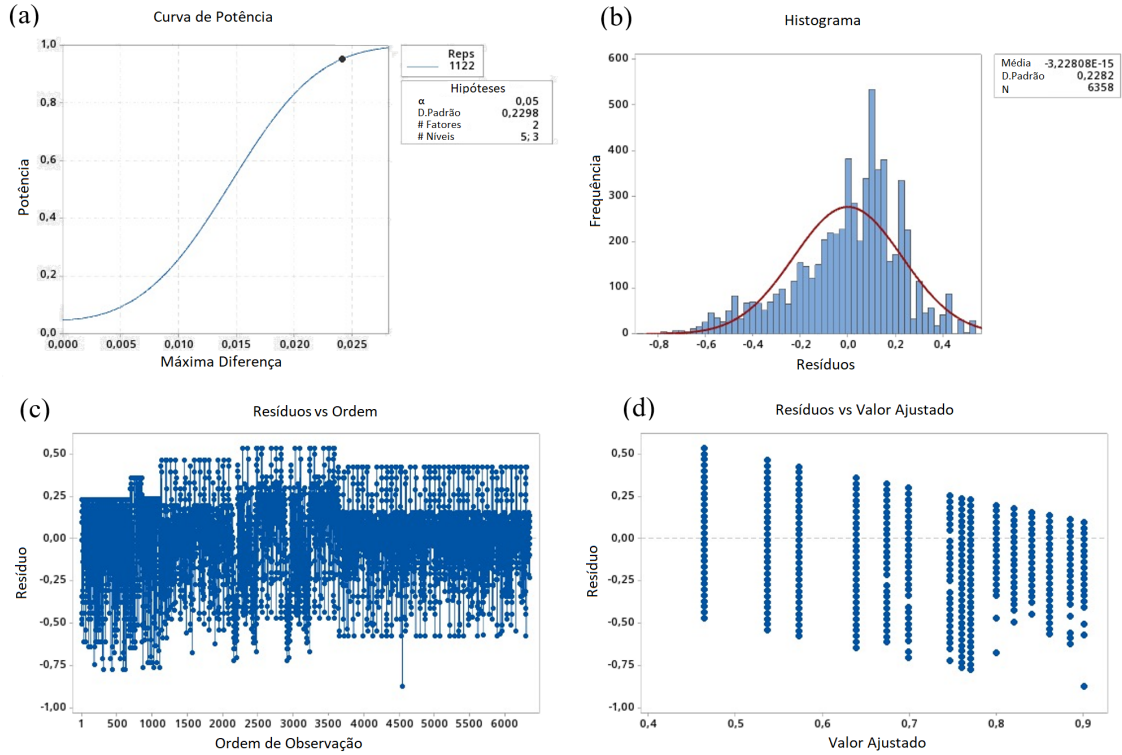
XIAORONG ZHANG; HE HUANG; QING YANG. Real-time implementation of a self-recovery EMG pattern recognition interface for artificial arms. In: ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY (EMBC), 35., 2013, Osaka. **Proceedings...** New York: IEEE, 2013. p. 5926-5929

XING, Kexin *et al.* A real-time EMG pattern recognition method for virtual myoelectric hand control. **Neurocomputing**, Amsterdam, v. 136, p. 345–355, 2014.

ZHANG, Xiaorong; HUANG, He. A real-time, practical sensor fault-tolerant module for robust EMG pattern recognition. **Journal of NeuroEngineering and Rehabilitation**, New York, v. 12, n. 1, Não paginado, 2015.

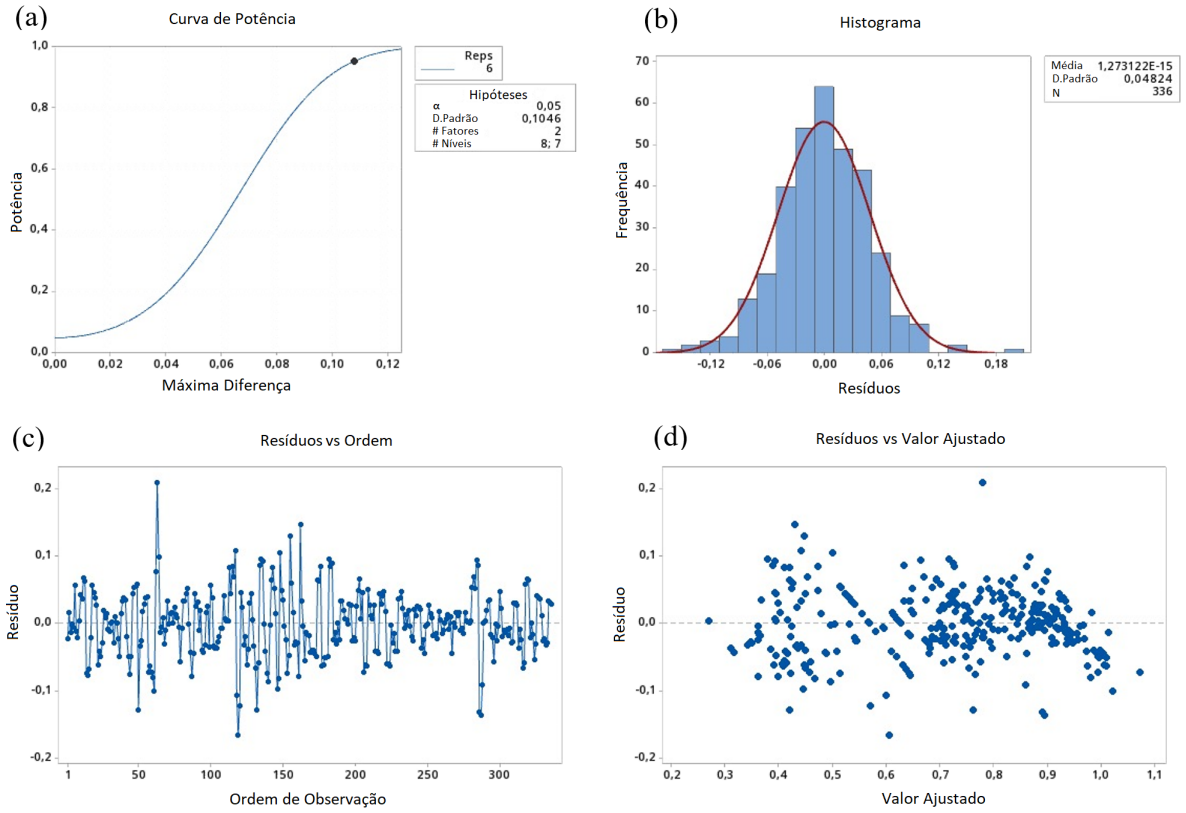
APÊNDICE A – PROJETOS DE EXPERIMENTOS

Figura 35 - Curva de Potência Estatística (a), Histograma dos resíduos (b) e análise dos resíduos pela ordem de observação (c) e pelo valor ajustado (d), para Projeto de experimentos referente ao Experimento sEMG 1, considerando os métodos de classificação.



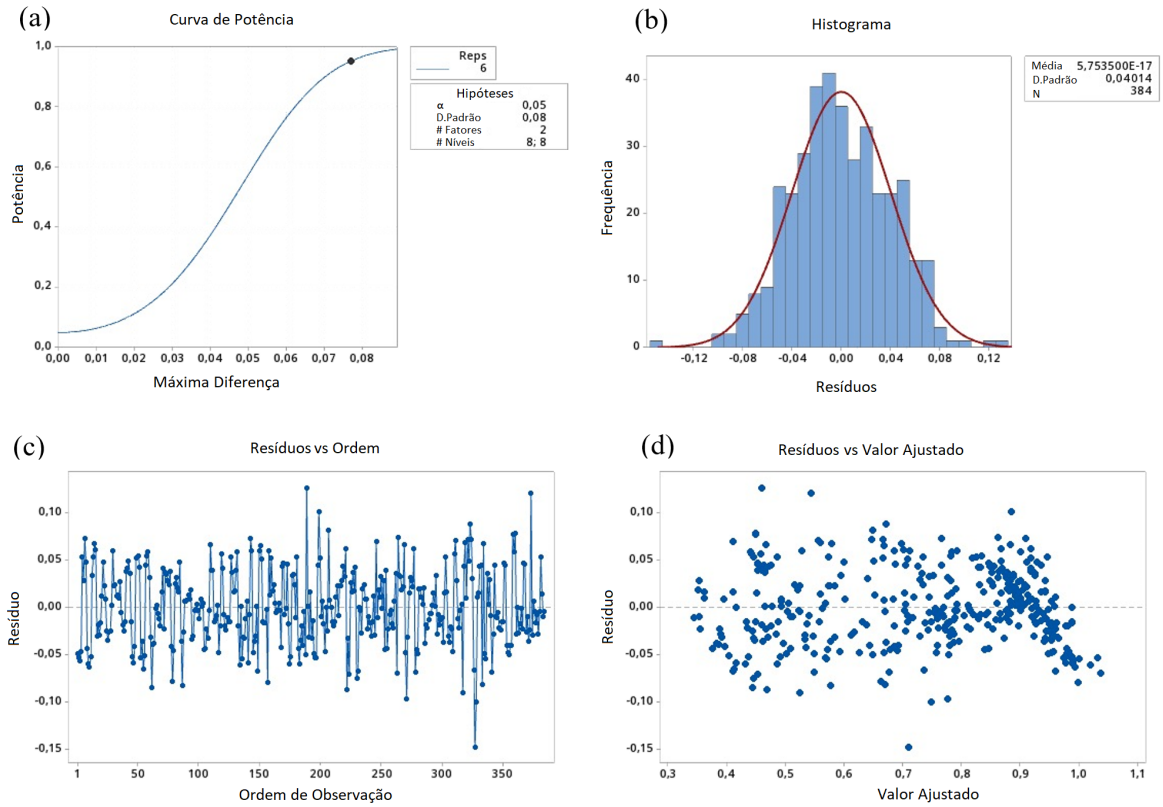
Fonte: Autor

Figura 36 - Curva de Potência Estatística (a), Histograma dos resíduos (b) e análise dos resíduos pela ordem de observação (c) e pelo valor ajustado (d), para Projeto de experimentos referente ao Experimento BDP, considerando dados com valores ausentes.



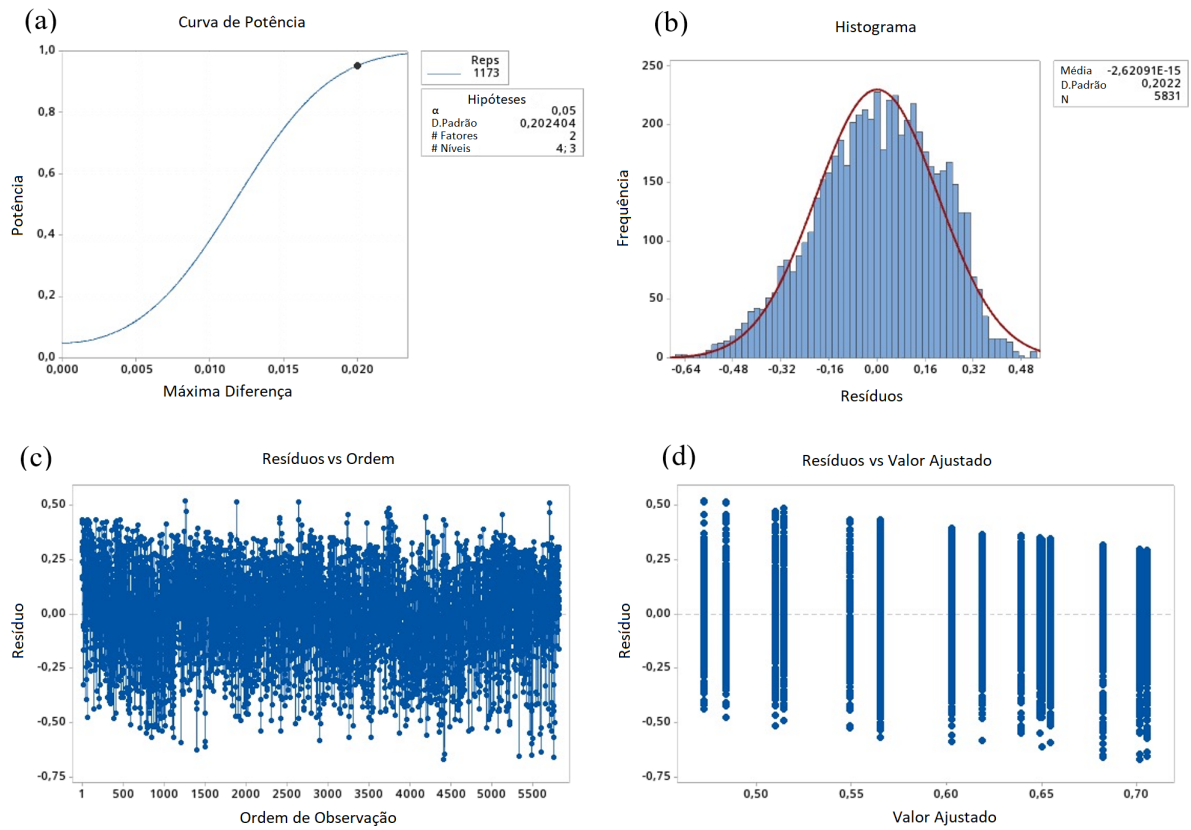
Fonte: Autor

Figura 37 - Curva de Potência Estatística (a), Histograma dos resíduos (b) e análise dos resíduos pela ordem de observação (c) e pelo valor ajustado (d), para Projeto de experimentos referente ao Experimento BDP, considerando dados com inserção de ruído.



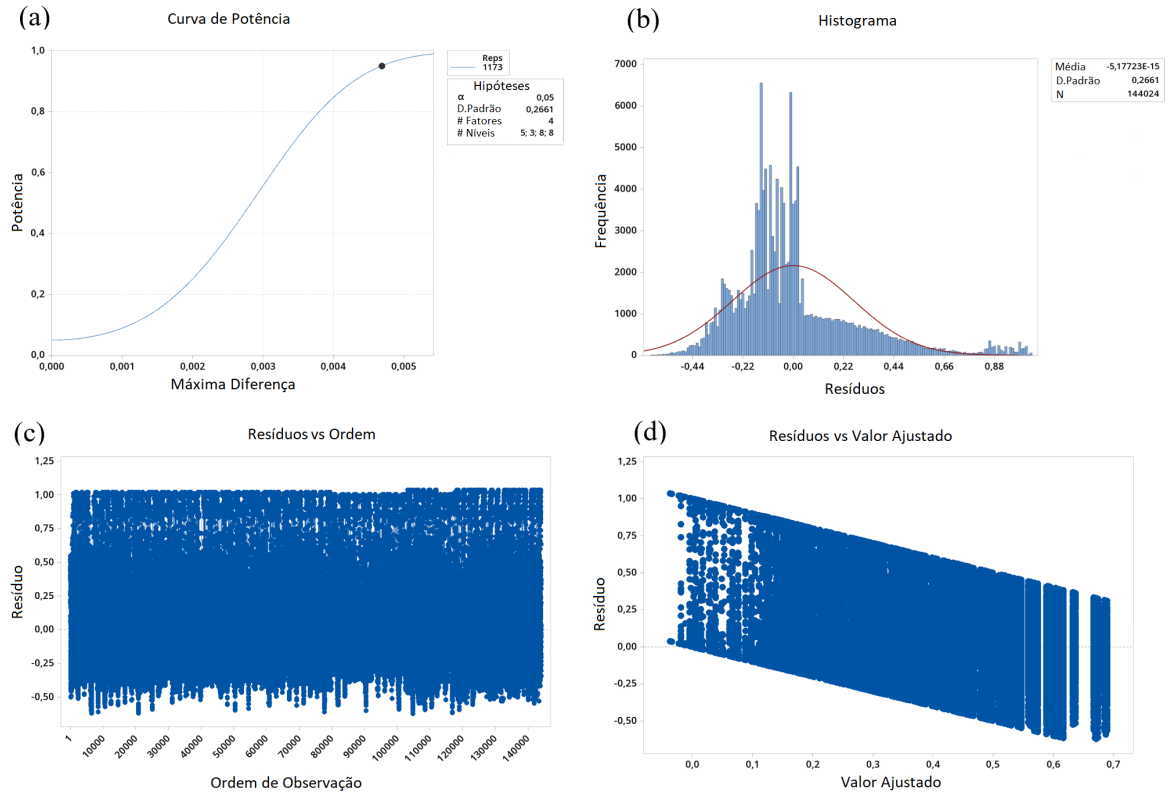
Fonte: Autor

Figura 38 - (a) Curva de potência estatística para projeto de experimentos de sinais de sEMG íntegros, (b) histograma dos resíduos e análise dos resíduos e pela (c) ordem de observação e (d) pelo valor ajustado.



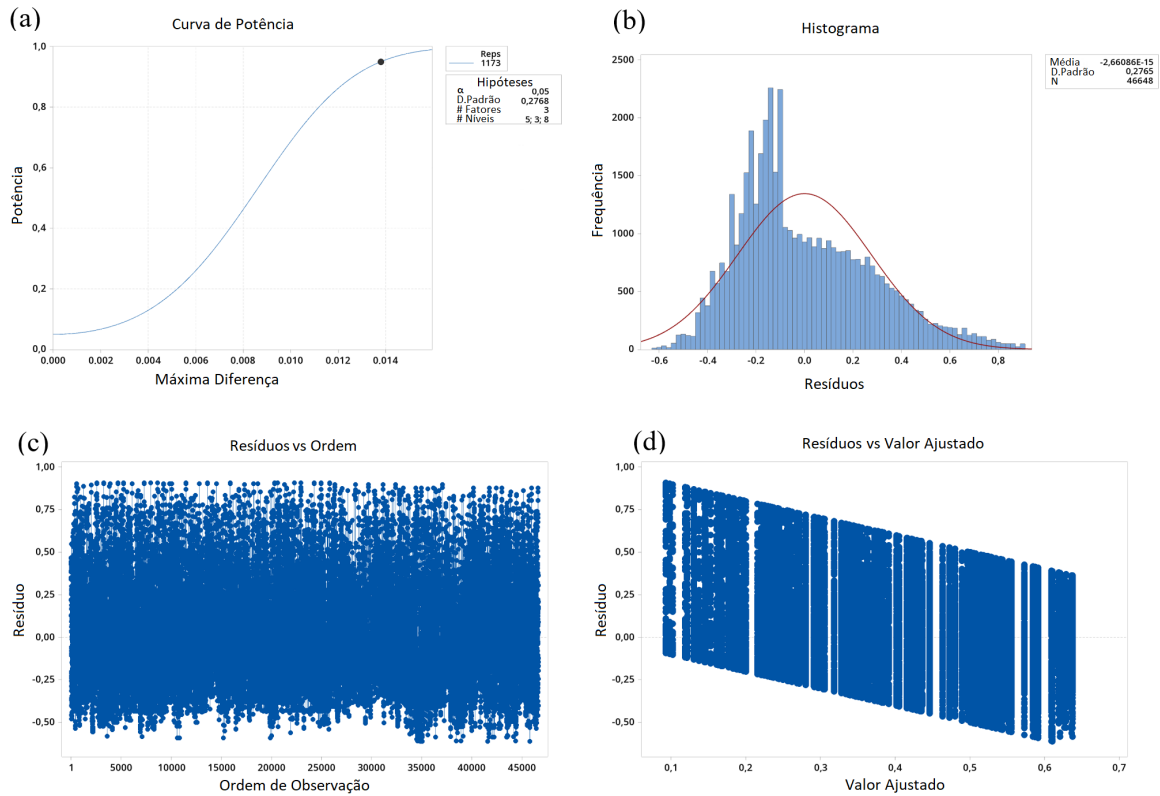
Fonte: Autor

Figura 39 - (a) Curva de potência estatística, (b) histograma dos resíduos e análise dos resíduos pela (c) ordem de observação (d) e pelo valor ajustado, para Projeto de experimentos referente aos sinais de sEMG com contaminantes.



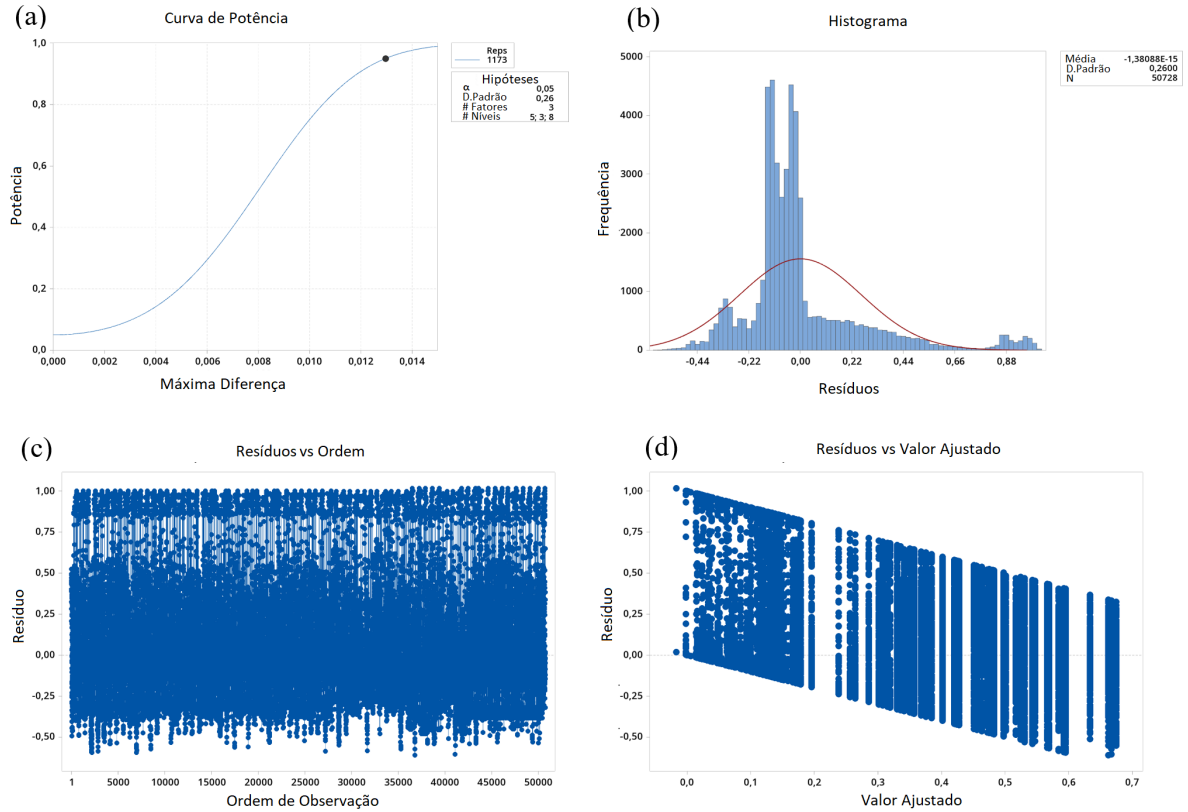
Fonte: Autor

Figura 40 - (a) Curva de potência estatística, (b) histograma dos resíduos e análise dos resíduos pela (c) ordem de observação (d) e pelo valor ajustado, para Projeto de experimentos referente a base de dados de sinais de sEMG contaminado com artefatos de movimento.



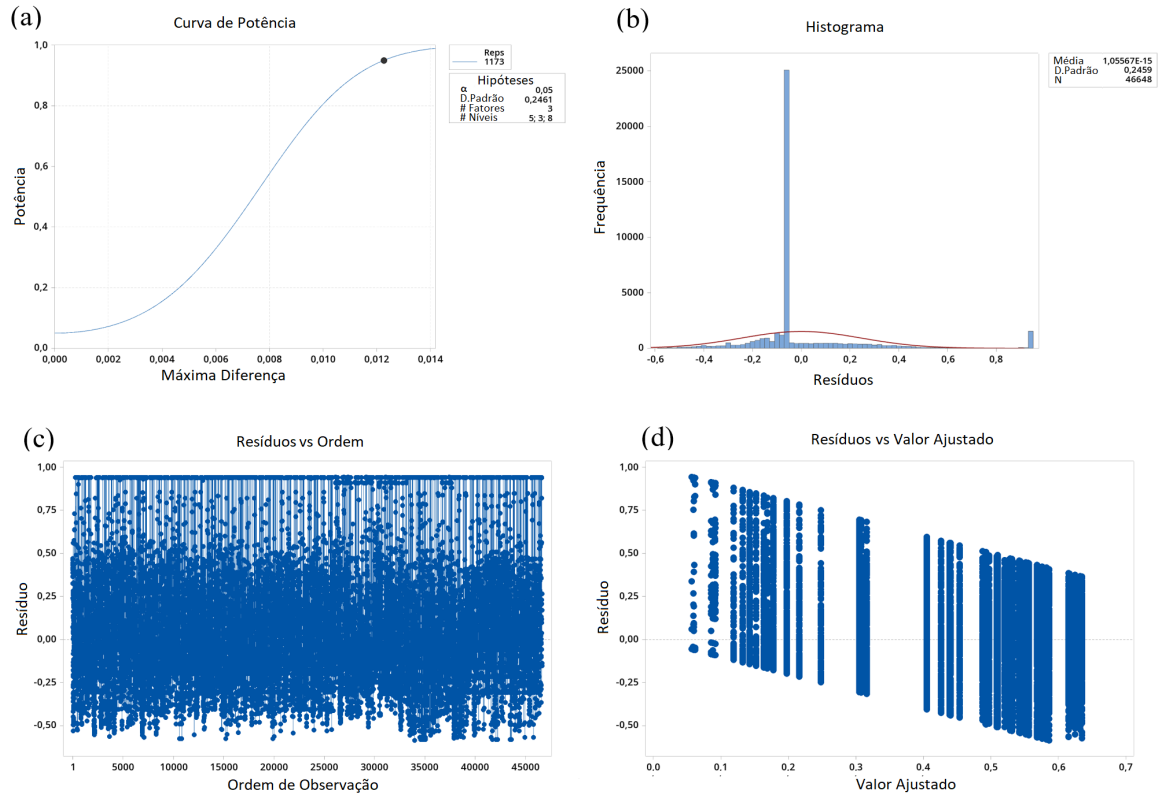
Fonte: Autor

Figura 41 - (a) Curva de potência estatística, (b) histograma dos resíduos e análise dos resíduos pela (c) ordem de observação (d) e pelo valor ajustado, para Projeto de experimentos referente a base de dados de sinais de sEMG contaminado com ruído térmico.



Fonte: Autor

Figura 42 - (a) Curva de potência estatística, (b) histograma dos resíduos e análise dos resíduos pela (c) ordem de observação (d) e pelo valor ajustado, para Projeto de experimentos referente a base de dados de sinais de sEMG contaminado com perda de contato eletrodo-pele.



Fonte: Autor