

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

**MAURÍCIO CAGLIARI TOSIN**

**APLICAÇÃO DE ALGORITMOS DE SELEÇÃO  
DE CARACTERÍSTICAS NA CLASSIFICAÇÃO DE  
MOVIMENTOS DO SEGMENTO MÃO-BRAÇO ATRAVÉS  
DO PROCESSAMENTO DE SINAIS DE sEMG**

Porto Alegre

2018

**MAURÍCIO CAGLIARI TOSIN**

**APLICAÇÃO DE ALGORITMOS DE SELEÇÃO  
DE CARACTERÍSTICAS NA CLASSIFICAÇÃO DE  
MOVIMENTOS DO SEGMENTO MÃO-BRAÇO ATRAVÉS  
DO PROCESSAMENTO DE SINAIS DE sEMG**

Dissertação de mestrado 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 Mestre em Engenharia Elétrica.

Área de concentração: Engenharia de Controle e Automação - Instrumentação.

**ORIENTADOR: Prof. Dr. Alexandre Balbinot**

Porto Alegre

2018

MAURÍCIO CAGLIARI TOSIN

**APLICAÇÃO DE ALGORITMOS DE SELEÇÃO  
DE CARACTERÍSTICAS NA CLASSIFICAÇÃO DE  
MOVIMENTOS DO SEGMENTO MÃO-BRAÇO ATRAVÉS  
DO PROCESSAMENTO DE SINAIS DE sEMG**

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

Orientador: \_\_\_\_\_

Prof. Dr. Alexandre Balbinot, UFRGS

Doutor pela UFRGS – Porto Alegre, Brasil

Banca Examinadora:

Prof. Dr. Henrique Takachi Moriya, USP

Doutor pela Universidade de São Paulo – São Paulo, Brasil

Prof. Dra. Karin Becker, UFRGS

Doutora pela Facultés Universitaires Notre-Dame de la Paix – Namur, Bélgica

Prof. Dr. Altamiro Amadeu Susin, UFRGS

Doutor pela Institut National Polytechnique de Grenoble – Grenoble, França

Coordenador do PPGEE: \_\_\_\_\_

Prof. Dr. Valner João Brusamarello

Porto Alegre, março de 2018.

## **DEDICATÓRIA**

Dedico este trabalho aos meus pais Moacir e Cleide, minha namorada Juliana e meus irmãos Roberta, Márcio e Manoela.

## **AGRADECIMENTOS**

Quero agradecer primeiramente aos meus pais por sempre acreditar em mim e por me proporcionarem todas as condições para que eu conquistar meus objetivos.

A minha namorada Juliana, por estar sempre do meu lado e me dar força em todos os momentos difíceis ao longo do desenvolvimento deste trabalho.

Aos meus irmãos pelo companheirismo e apoio.

Aos colegas do Laboratório do IEE pelo auxílio na execução deste trabalho e parceria na resolução de dúvidas e sugestões.

Ao meu orientador, Prof. Alexandre Balbinot pela sua orientação e conselhos valiosos que foram fundamentais no desenvolvimento deste trabalho.

Ao Programa de Pós-Graduação em Engenharia Elétrica, PPGEE, pela oportunidade de realização de trabalhos em minha área de pesquisa.

Ao CNPq pela provisão da bolsa de mestrado.

## RESUMO

É desafiador obter taxas de acerto elevadas no que cerne a tarefa de classificar movimentos do segmento mão braço a partir de sinais de sEMG. Inúmeros trabalhos têm sido desenvolvidos nessa área explorando a filtragem do sinal, extração de características e classificador, porém muito poucos atacam o problema sob o aspecto da seleção de características. O presente trabalho visa implementar três algoritmos diferentes de seleção de características para classificação de tais sinais: *Support Vector Machine Recursive Feature Elimination* (SVM-RFE), Monte Carlo *Feature Selection*, e *Singular Value Decomposition* (SVD) *Entropy*. Para a referida análise foram utilizados sinais da Base de Dados internacional Ninapro e também foram adquiridos sinais no laboratório de Instrumentação Eletro-eletrônica (IEE). Para o estágio de classificação foi utilizado o método *Regularized Extreme Learning Machine* (RELM). Contudo, o algoritmo proposto consiste em 4 etapas fundamentais. Primeiramente, 13 características foram extraídas dos sinais de sEMG: 11 no domínio do tempo e duas no domínio da frequência. A correlação entre todas as 156 características obtidas (12 canais x 13 características) foi analisada e retirou-se do conjunto as que apresentaram correlação superior a 0,9. Na sequência, as que restaram da etapa anterior foram ranqueadas a partir dos 3 métodos de seleção de características supracitados. Por fim, o conjunto ordenado de características passou pelo estágio de classificação onde determinou-se o melhor grupo dentre elas. Como resultado, obteve-se taxa de acerto média de 85,42% para a Base 2 Ninapro, 77,46% para a Base 3 Ninapro e 86,74% para a Base do IEE.

**Palavras-chave:** *Eletromiografia, Support Vector Machine Recursive Feature Elimination, Monte Carlo Feature Selection, Singular Value Decomposition Entropy, Regularized Extreme Learning Machine.*

## ABSTRACT

It is challenging to obtain high classification rates in the task of classifying hand-arm segment movements by sEMG signals analysis. Several works have been developed in this area exploring signal filtering, feature extraction and classifier, but very few of them faces the problem under the feature selection point of view. The present work aims to implement three different feature selection algorithms to classify such signals: Support Vector Machine Recursive Feature Elimination (SVM-RFE), Monte Carlo Feature Selection, and Singular Value Decomposition (SVD) Entropy. For this analysis, signals from the International Ninapro Database were used, and additionally it was acquired signals in the Electro-Electronic Instrumentation (IEE) Laboratory. For classification stage, it was used the Regularized Extreme Learning Machine (RELM) method. However, the proposed algorithm consists of 4 fundamental steps. First, 13 characteristics were extracted from the sEMG signals: 11 in the time domain and two in the frequency domain. The correlation between all 156 features obtained (12 channels x 13 characteristics) was analyzed and those which presented a correlation value above 0,9 were removed from the set. In sequence, the remaining features were ranked by the above mentioned feature selection methods. Finally, the ordered set of features pass through the classification stage where the best group among them was determined. As a result, the mean classification rate obtained was 85,42% for Ninapro Database 2, 77,46% for Ninapro Database 3 and 86,74% for IEE Database.

**Keywords: Electromyography, Support Vector Machine Recursive Feature Elimination, Monte Carlo Feature Selection, Singular Value Decomposition Entropy, Regularized Extreme Learning Machine.**

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>14</b>
1.1	JUSTIFICATIVAS	14
1.2	OBJETIVOS	16
<b>2</b>	<b>REVISÃO DA LITERATURA</b>	<b>18</b>
2.1	CONCEITOS BÁSICOS SOBRE ELETROMIOGRAFIA	18
2.1.1	Músculo Esquelético	20
2.1.2	Neurônio Motor	22
2.1.3	Unidade Motora	24
2.1.4	Potencial de Ativação	26
2.1.5	Princípios sobre o Sinal de EMG	27
2.2	PROCESSAMENTO DE SINAIS DE ELETROMIOGRAFIA	30
2.2.1	Seleção de Características	32
2.2.2	Classificação	46
<b>3</b>	<b>METODOLOGIA EXPERIMENTAL</b>	<b>54</b>
3.1	DESCRIÇÃO DAS BASES DE DADOS	55
3.1.1	Base de Dados Ninapro	55
3.1.2	Base de Dados do IEE	58
3.2	SEGMENTAÇÃO E A CORRESPONDENTE EXTRAÇÃO DE CARACTERÍSTICAS	74
3.3	MÉTRICA DE CORRELAÇÃO	79
3.4	RANKING DAS CARACTERÍSTICAS	81
3.4.1	Algoritmo SVM-RFE	81
3.4.2	Monte Carlo <i>Feature Selection</i>	83
3.4.3	SVD <i>Entropy</i>	86
3.5	CLASSIFICAÇÃO E SELEÇÃO DE CARACTERÍSTICAS	87
3.5.1	Classificador <i>Regularized Extreme Learning Machine</i>	87
3.5.2	Algoritmo de Seleção de Características e Classificação	92
3.6	VALIDAÇÃO ESTATÍSTICA DOS RESULTADOS	94
3.6.1	Projeto de Experimentos com a Base de Dados do IEE	94
3.6.2	Projeto de Experimentos com as Bases de Dados Ninapro e a Base do IEE	95
<b>4</b>	<b>RESULTADOS E DISCUSSÕES</b>	<b>96</b>
4.1	VALIDAÇÃO DOS MODELOS DOS PROJETOS DE EXPERIMENTOS	97
4.1.1	Validação do Projeto de Experimentos com a Base de Dados do IEE	97
4.1.2	Validação do Projeto de Experimentos com as Bases de Dados Ninapro e a Base de Dados do IEE	98
4.2	RESULTADOS DO PROJETO DE EXPERIMENTOS COM A BASE DE DADOS DO IEE	99
4.2.1	Análise do Fator Controlável Movimento	101
4.2.2	Análise do Fator Controlável Voluntário	103
4.2.3	Análise do Fator Controlável Tipo de Ensaio	105
4.2.4	Análise do Fator Controlável Método de Seleção	107
4.2.5	Análise da Interação entre Voluntário e Movimento	108
4.2.6	Análise da Interação entre Método de Seleção e Movimento	110
4.2.7	Análise da Interação entre Voluntário e Método de Seleção	111

4.2.8	Análise da Interação entre Tipo de Ensaio e Movimento .....	112
4.2.9	Análise da Interação entre Tipo de Ensaio e Voluntário .....	114
4.2.10	Análise da Interação Entre Tipo de Ensaio e Método de Seleção de Características .....	115
4.2.11	Análise da Interação entre Voluntário, Método de Seleção e Tipo de Ensaio .....	116
4.2.12	Análise da Interação entre Voluntário, Movimento e Tipo de Ensaio .....	117
4.2.13	Análise da Interação entre Voluntário, Método de Seleção e Movimento .....	119
4.2.14	Análise da Interação entre Tipo de Ensaio, Método de Seleção e Movimento .....	120
4.2.15	Análise da Interação entre Voluntário, Movimento, Método de Seleção e Tipo de Ensaio.....	121
4.3	RESULTADOS DO PROJETO DE EXPERIMENTOS COM AS BASES NINAPRO E ENSAIO A DA BASE DO IEE .....	122
4.3.1	Análise do Fator Controlável Base de Dados.....	124
4.3.2	Análise do Fator Controlável Movimento .....	126
4.3.3	Análise da Interação entre Base de Dados e Movimento .....	128
4.4	ANÁLISE DA TAXA DE ACERTO MÉDIA POR VOLUNTÁRIO PARA AS BASES NINAPRO .....	130
4.5	ANÁLISE DO NÚMERO DE CARACTERÍSTICAS SELECIONADAS.....	133
4.6	COMPARAÇÃO DOS RESULTADOS COM TRABALHOS RELACIONADOS..	134
5	CONCLUSÃO.....	143
6	PROPOSTAS PARA TRABALHOS FUTUROS.....	146
	ANEXO A: TERMO LIVRE E ESCLARECIDO.....	153
	ANEXO B: ANAMNESE.....	155

## LISTA DE FIGURAS

Figura 1 Organização e identificação dos tecidos conjuntivos do músculo esquelético (adaptado de TORTORA; DERRICKSON, 2012).	21
Figura 2 Visão microscópica das fibras musculares (adaptado de TORTORA; DERRICKSON, 2012).	22
Figura 3 Estrutura de um neurônio multipolar (adaptado de TORTORA; DERRICKSON, 2012).	23
Figura 4 Partes formadoras de uma unidade motora (adaptado de POZZO; FARINA; MERLETTI, 2004).	25
Figura 5 Potencial de ação da membrana muscular (adaptado de POZZO; FARINA; MERLETTI, 2004).	26
Figura 6 Sinal de sEMG do primeiro voluntário da base do Ninapro (adaptado livremente da base de ATZORI <i>et al.</i> , 2014).	28
Figura 7 Influência do posicionamento dos eletrodos no sinal de EMG (DE LUCA, 1997).	29
Figura 8 Típico diagrama de blocos para a tarefa de classificação de sinais de EMG.	32
Figura 9 Hiperplano de separação das classes (HEARST <i>et al.</i> , 1998).	36
Figura 10 Fluxograma do método SVM-RFE.	39
Figura 11 Funcionamento do método de seleção de características Monte Carlo (adaptado livremente de DRAMINSKI <i>et al.</i> , 2008).	43
Figura 12 Estrutura da rede neural utilizada pelo ELM (adaptado de AKUSOK <i>et al.</i> , 2015).	48
Figura 13 Diagrama de blocos da metodologia proposta.	55
Figura 14 Configurações isométricas de mãos (adaptado de ATZORI <i>et al.</i> , 2014).	56
Figura 15 Movimentos básicos de pulso (adaptado de ATZORI <i>et al.</i> , 2014).	56
Figura 16 Procedimento de aquisição (adaptado de ATZORI <i>et al.</i> , 2014).	58
Figura 17 Sistema experimental proposto para a geração da Base de Dados do IEE.	60
Figura 18 Fluxograma do sistema de aquisição: (a) para os ensaios sequenciais; (b) para os ensaios aleatorizados.	62
Figura 19 Arranjo experimental utilizado nas aquisições na geração da Base de Dados do IEE: (1) corresponde ao monitor onde os movimentos eram mostrados, (2) eletromiógrafos, (3a-c) cabos do eletromiógrafo, (4) eletrodos de superfície, (5) placa de aquisição, (6) borneira, (7) computador (adaptado de CENE, 2016).	63
Figura 20 Posicionamento dos eletrodos (adaptado de CENE, 2016).	64
Figura 21 Ensaio realizado por um dos voluntários.	66
Figura 22 Diagrama de blocos do algoritmo de ajuste dos tempos de início e término de cada movimento.	67
Figura 23 Exemplo de sinal de sEMG pré-processado: (a) retificado, (b) retificado e filtrado e (c) retificado, filtrado e normalizado.	68
Figura 24 Exemplo da segmentação efetuada. Em vermelho está indicado os períodos de execução dos vídeos de movimento (nível 1) que limitam tais janelas.	69
Figura 25 Segmento do sinal dividido em 3 partes: repouso, movimento, repouso.	70
Figura 26 Parte do sinal de sEMG com a segmentação antes e após o ajuste.	72
Figura 27 Fluxograma da rotina desenvolvida para o ajuste dos rótulos das amostras.	73
Figura 28 Segmentação realizada para a extração de características (adaptado de TOSIN <i>et al.</i> , 2017).	74

Figura 29 Sistema de extração de características implementado (adaptado de ORTIZ-CATALAN; BRÅNEMARK; HÅKANSSON, 2013).....	75
Figura 30 Gráfico do comprimento da curva em função da escala ( $k$ ) (adaptado de ARJUNAN; KUMAR, 2010).....	77
Figura 31 Evolução da taxa de acerto do classificador em função do fator $C$ . ....	83
Figura 32 Evolução da distância entre rankings em função do número de subconjuntos ( $s$ ). .	85
Figura 33 Fluxograma do método SVD <i>Entropy</i> .....	86
Figura 34 Evolução da taxa de acerto do classificador em função do número de neurônios...	89
Figura 35 Evolução da taxa de acerto do classificador em função do fator $\gamma$ . ....	90
Figura 36 Fluxograma do Algoritmo RELM. ....	91
Figura 37 Fluxograma do algoritmo de classificação e seleção de características.....	93
Figura 38 Análise dos resíduos do Projeto de Experimentos com a Base de Dados do IEE. ...	98
Figura 39 Análise dos resíduos do Projeto de Experimentos com as Bases Ninapro e a do IEE. ....	99
Figura 40 Dispersão da taxa de acerto obtida para cada movimento.....	101
Figura 41 Gráfico do sinal de sEMG adquirido durante execução do ensaio A. ....	103
Figura 42 Dispersão da taxa de acerto obtida para cada voluntário. ....	105
Figura 43 Dispersão da taxa de acerto obtida para cada tipo de ensaio.....	107
Figura 44 Dispersão da taxa de acerto obtida para cada método.....	108
Figura 45 Interação entre voluntários e movimento. ....	109
Figura 46 Interação entre método e movimento. ....	111
Figura 47 Análise da interação entre voluntário e método de seleção.....	112
Figura 48 Análise da interação entre movimento e tipo de ensaio. ....	113
Figura 49 Análise da interação entre voluntário e tipo de ensaio.....	114
Figura 50 Análise da interação entre método de seleção e tipo de ensaio. ....	115
Figura 51 Interação entre voluntário, método e tipo de ensaio. ....	116
Figura 52 Interação entre voluntário, ensaio e movimento. ....	118
Figura 53 Análise da interação entre método de seleção, voluntário e movimento. ....	119
Figura 54 Análise da interação entre método de seleção, tipo de ensaio e movimento.....	120
Figura 55 Interação entre voluntário, método, movimento e ensaio. ....	121
Figura 56 Dispersão da taxa de acerto obtida para cada base de dados.....	124
Figura 57 Relação entre o fator controlável movimento com a taxa de acerto média.....	127
Figura 58 Relação entre a interação entre base de dados e movimento com a taxa de acerto média.....	129
Figura 59 Relação entre voluntário e a taxa de acerto média para as Bases Ninapro. ....	131
Figura 60 Relação entre o método de seleção de características e o número médio de características selecionadas. ....	133

## LISTA DE TABELAS

Tabela 1 Parâmetros ajustados para os ensaios da Base 2 Ninapro.....	90
Tabela 2 Parâmetros ajustados para os ensaios da Base 3 Ninapro.....	91
Tabela 3 Identificação dos movimentos e dos respectivos músculos envolvidos (adaptado de CENE, 2016). .....	96
Tabela 4 Análise de Variância do Projeto de Experimentos com a Base de Dados do IEE. .	100
Tabela 5 Análise de Variância do Projeto com as Bases Ninapro e Ensaio A da Base do IEE. ....	123
Tabela 6 Trabalhos relacionados à classificação de movimentos a partir de sinais de sEMG. ....	135

## LISTA DE ABREVIATURAS

Ach: Acetilcolina

ANOVA: *Analysis of Variance*

ATP: Trifosfato de Adenosina

CAAE: Certificado de Apresentação para Apreciação Ética

CFO: Comprimento da Forma de Onda

CFORF: Comprimento da Forma de Onda da Resposta em Frequência

CSP: *Common Spatial Pattern*

CZ: Cruzamentos por Zero

DF: Dimensão Fractal

DFH: Dimensão Fractal de Higuchi

DP: Desvio Padrão

EEG: Eletroencefalografia

ELM: *Extreme Learning Machine*

EMG: Eletromiografia

FFT: *Fast Fourier Transform*

FM: Frequência Média

GLR: *Generalized Likelihood Ratio*

ID: *Induction Decision*

IEE: Instrumentação Eletro-Eletrônica

IQR: *Inter Quartile Range*

LDA: *Linear Discriminant Analysis*

MD-DTW: *Multidimensional Dynamic Time Warping*

MDA: Média da Diferença Absoluta

MDF: Máxima Dimensão Fractal

MIS: Mudanças de Inclinação do Sinal

MMG: *Mechanomyogram*

MUAP: *Motor Unit Action Potential*

Ninapro: *Non-Invasive Adaptive Hand Prosthetics*

NMI: *Normalized Mutual Information*

OVO: *One Versus One*

PCA: *Principal Component Analysis*

RELM: *Regularized Extreme Learning Machine*

RMS: *Root Mean Square*

RST: *Rough Set Theory*

sEMG: Eletromiografia de Superfície

SNC: Sistema Nervoso Central

SVD: *Singular Value Decomposition*

SVM: *Support Vector Machine*

SVM-RFE: *Support Vector Machine Recursive Feature Elimination*

UFRGS: Universidade Federal do Rio Grande do Sul

VMA: Valor Médio Absoluto

## **1 INTRODUÇÃO**

O estudo da eletromiografia consiste na análise da atividade elétrica presente nas membranas excitáveis dos músculos. Através dela, é possível obter informações importantes a respeito do sistema neuromuscular, as quais podem ser utilizadas na detecção e prevenção de disfunções, como por exemplo, na medicina esportiva para melhorar a performance de atletas de alto rendimento, dentre outras aplicações.

Além destas áreas, a avaliação dos sinais elétricos oriundos das membranas musculares pode ser empregada na identificação de padrões de movimentos para o desenvolvimento de próteses com maior grau de liberdade. É dentro deste contexto que o presente trabalho foi desenvolvido.

### **1.1 JUSTIFICATIVAS**

Atualmente, com o avanço da tecnologia, busca-se cada vez mais qualidade de vida e conforto. Tarefas que antes eram custosas de se fazer, hoje podem ser executadas sem mesmo precisar sair de casa. É o caso do pagamento de contas, realização de compras, consulta à conta bancária, e outros tipos de serviço. Desta forma, consegue-se aproveitar melhor o tempo livre exercendo atividades que trazem prazer e até mesmo desfrutando da companhia das pessoas de quem se gosta.

Seguindo essa linha de pensamento, as pessoas portadoras de deficiência, em especial as que sofreram amputações dos membros superiores, também clamam por qualidade de vida e, principalmente, independência. Contudo, pode-se dizer que tais ensejos estão associados ao uso e às características da prótese que substitui o membro retirado. Quanto maior o número de graus de liberdade habilitados pelo braço mecânico, maior o nível de independência alcançado pelo usuário e, conseqüentemente, maior será sua qualidade de vida.

Nos dias atuais, já existem próteses comerciais com controle mioelétrico. Entretanto, muitas dessas próteses são baseadas na detecção de um valor de limiar da amplitude do sinal de eletromiografia de superfície (sEMG) e utilizam, em sua grande maioria, apenas dois eletrodos. Assim, a determinação do movimento a ser executado pela prótese é dada através da identificação de ativação do correspondente músculo. Por conseguinte, o sistema fica restrito a apenas alguns movimentos de baixa complexidade, tais quais, abrir e fechar a mão, reduzindo consideravelmente sua usabilidade e adaptação a diferentes cenários (RIILLO *et al.*, 2014), (ZHANG *et al.*, 2017b).

No entanto, para o desenvolvimento de próteses com número maior de graus de liberdade, é necessária a implementação de um sistema mais robusto de reconhecimento de padrões de modo a ser possível a identificação de diferentes tipos de movimentos. Seguindo esse objetivo, vários estudos têm sido conduzidos na análise de sinais de sEMG para o desenvolvimento de classificadores que distinguem movimentos do segmento mão-braço. Dentre eles, pode-se citar: (MOURA; FAVIEIRO; BALBINOT, 2016), (RIILLO *et al.*, 2014), (ZHAI *et al.*, 2016), (KARABULUT *et al.*, 2017), (ZHANG *et al.*, 2017b), (GENG *et al.*, 2017), (CENE; BALBINOT, 2016).

O presente trabalho está inserido nesta mesma linha de pesquisa, ou seja, está envolvido na tarefa de classificação de movimentos do segmento mão-braço a partir de sinais de sEMG. Contudo, o foco desta pesquisa está na etapa de seleção de características, a qual antecede o estágio de classificação propriamente dito. Deste modo, três algoritmos de seleção de características foram implementados e testados na referida tarefa: *Support Vector Machine Recursive Feature Elimination* (SVM-RFE), *Monte Carlo Feature Selection*, e *Singular Value Decomposition* (SVD) *Entropy*. A escolha de tais métodos se deu a partir dos bons resultados já apresentados quando aplicados na tarefa de seleção de genes para classificação de câncer, a citar: (BISMAR *et al.*, 2014), (XIE *et al.*, 2011), (YIN *et al.*, 2016b).

Este trabalho se justifica no fato de que a grande maioria das pesquisas na área de classificação de sinais de sEMG estão direcionados somente na análise da etapa de classificação e os mais diversos algoritmos existentes para tal. A etapa de seleção de características normalmente é negligenciada em tais pesquisas, sendo empregada em poucos trabalhos (ZHOU *et al.*, 2016). Logo, viu-se uma oportunidade de avaliar o desempenho de alguns algoritmos de seleção de características neste nicho acadêmico, e, conseqüentemente, visualizar o impacto da inclusão desta etapa nas taxas de acerto resultantes do processo de classificação.

De uma forma mais global, o presente trabalho é justificado pela necessidade de proporcionar melhores condições de vida às pessoas que sofreram amputação de membros superiores, isto é, pela oferta de próteses com maior número de graus de liberdade.

## 1.2 OBJETIVOS

O objetivo principal deste trabalho é o de avaliar a aplicabilidade de algoritmos de seleção de características típicos da área de seleção de genes para classificação de câncer, na tarefa de classificação de sinais de sEMG para a discriminação de movimentos do segmento mão-braço. Para tanto, três métodos de seleção de características foram utilizados. Os já citados SVM-RFE, Monte Carlo e SVD *Entropy*.

Os referidos métodos foram originalmente desenvolvidos para redução de dimensionalidade de banco de dados em problemas de classificação de genes para detecção de câncer (BISMAR *et al.*, 2014), (XIE *et al.*, 2011), (YIN *et al.*, 2016b), (ALBASHISH *et al.*, 2015). A saber, pesquisas nesta área lidam com bases de dados constituídas de milhares de atributos para poucas centenas de amostras, sendo, assim, necessária a realização de redução de dimensionalidade. Por conseguinte, o propósito desta pesquisa é o de analisar o comportamento de tais algoritmos e verificar se são aplicáveis também na tarefa de classificação de sinais de sEMG.

Para a análise de tais algoritmos, foi desenvolvida uma base de dados do Laboratório de Instrumentação Eletro-Eletrônica (IEE) da UFRGS. Para este fim, quatro voluntários com membros intactos foram recrutados e instruídos a executar uma série de ensaios contendo uma sequência específica de movimentos. Durante a realização de tais ensaios, sinais de sEMG foram adquiridos a partir de 12 pares de eletrodos de superfície estrategicamente posicionados ao longo do braço e antebraço.

Além dos sinais adquiridos experimentalmente, o presente trabalho objetivou a análise das bases 2 e 3 disponibilizadas pela equipe do Ninapro (ATZORI *et al.*, 2014). A base 2 é constituída por sinais de sEMG adquiridos de 40 voluntários intactos, ao passo que a base 3 é formada por sinais de 11 indivíduos com amputação transradial. Dez voluntários da base 3 e toda a base 2 foram selecionados para o presente estudo.

Para a etapa de classificação foi avaliado o desempenho do algoritmo baseado em redes neurais *Regularized Extreme Learning Machine* (RELM), uma das áreas foco da pesquisa de doutorado de (CENE, 2017 - "Proposta de Classificador Redundante Híbrido Para a Caracterização de Sinais Mioelétricos Baseado em Extreme Learning Machines").

Por fim, os resultados obtidos da implementação dos algoritmos de seleção e classificação nas três bases de dados supracitadas, foram avaliados estatisticamente através da realização de um Projeto de Experimentos do Tipo Fatorial Completo. Para isso, dois projetos de experimentos foram desenvolvidos considerando-se como variável de resposta a taxa de acerto média do classificador. Em um deles, avaliou-se a influência dos fatores controláveis tipo de ensaio, movimento, método de seleção de características e voluntário. Já no outro, testou-se a influência dos fatores base de dados, movimento e método de seleção de características. Adicionalmente a estes dois, um novo Projeto de Experimentos foi conduzido de modo a avaliar e comparar diretamente a atuação dos métodos de seleção implementados. Neste último considerou-se como variável de resposta o número de características selecionadas.

## 2 REVISÃO DA LITERATURA

### 2.1 CONCEITOS BÁSICOS SOBRE ELETROMIOGRAFIA

Uma das características mais importantes dos seres vivos é a sua capacidade de se locomover e de se movimentar. O tecido responsável por tal função é o músculo que age como atuador do sistema esquelético. É dificultosa a realização *in vivo* do estudo das propriedades funcionais dos músculos devido à dificuldade em se colocar sensores em série com os tendões, além do fato de que, normalmente, vários grupos musculares estarem conectados a um mesmo segmento do esqueleto (POZZO; FARINA; MERLETTI, 2004). Em face de tal impedimento, faz-se uso do estudo da atividade elétrica das fibras musculares para se determinar a ação do referido tecido durante uma contração. A essa análise dá-se o nome de Eletromiografia.

Apesar dos músculos serem alvo de pesquisas durante séculos, apenas recentemente passou-se a utilizar a sua atividade elétrica para caracterizá-lo. Em meados do século XVIII, Galvani (1737-1798) demonstrou que um estímulo elétrico pode provocar a contração de um o músculo. A primeira leitura de sinal de Eletromiografia, realizada com o uso de eletrodos superficiais, foi executada por Raymond em 1949 através de um tipo primitivo de galvanômetro (POZZO; FARINA; MERLETTI, 2004).

Atualmente, o estudo de sinais de Eletromiografia é aplicado em diversas áreas, entre elas: reabilitação motora, na ciência do esporte atuando na melhoria de desempenho de atletas de alto rendimento, medicina clínica, desenvolvimento de próteses com maior grau de liberdade (MOURA; FAVIEIRO; BALBINOT, 2016), (RIILLO *et al.*, 2014), (ZHAI *et al.*, 2016), (KARABULUT *et al.*, 2017), (ZHANG *et al.*, 2017b), (GENG *et al.*, 2017), (CENE; BALBINOT, 2016). Especificamente nessa última, muitos trabalhos com promissores resultados vêm sendo publicados, o que tem atraído a atenção de diversos pesquisadores e motivou o desenvolvimento do presente trabalho.

Define-se o mecanismo relacionado à geração da atividade elétrica presente nas membranas musculares, e consequente contração do músculo, através das seguintes etapas (GUYTON; HALL, 2011):

1. um potencial de ação é transmitido às fibras musculares por intermédio do nervo motor cujas terminações nervosas estão conectadas à membrana muscular;
2. íons de Cálcio e Sódio são então difundidos na membrana muscular, causando uma despolarização local, o que, por sua vez, abre canais de Sódio dependentes de tensão elétrica, gerando um potencial de ação;
3. o potencial gerado pela despolarização da membrana se difunde através da mesma, fazendo com que o Retículo Sarcoplasmático (rede de túbulos envolvido por membrana e preenchido com líquido, presente no Citoplasma da fibra muscular) libere íons de Cálcio para as Miofibrilas (estruturas cilíndricas formadoras da fibra muscular);
4. os íons de Cálcio liberados geram uma força de atração entre as fibras de Actina e de Miosina (filamentos formadores das Miofibrilas: a primeira consiste em filamentos delgados e a segunda espessos), resultando no deslizamento de uns sobre os outros. Tal movimento caracteriza o processo contrátil;
5. os íons de cálcio são então bombeados de volta ao Retículo Sarcoplasmático, onde são retidos até que novo potencial de ação ocorra. O processo de remoção de tais íons faz com que a contração cesse.

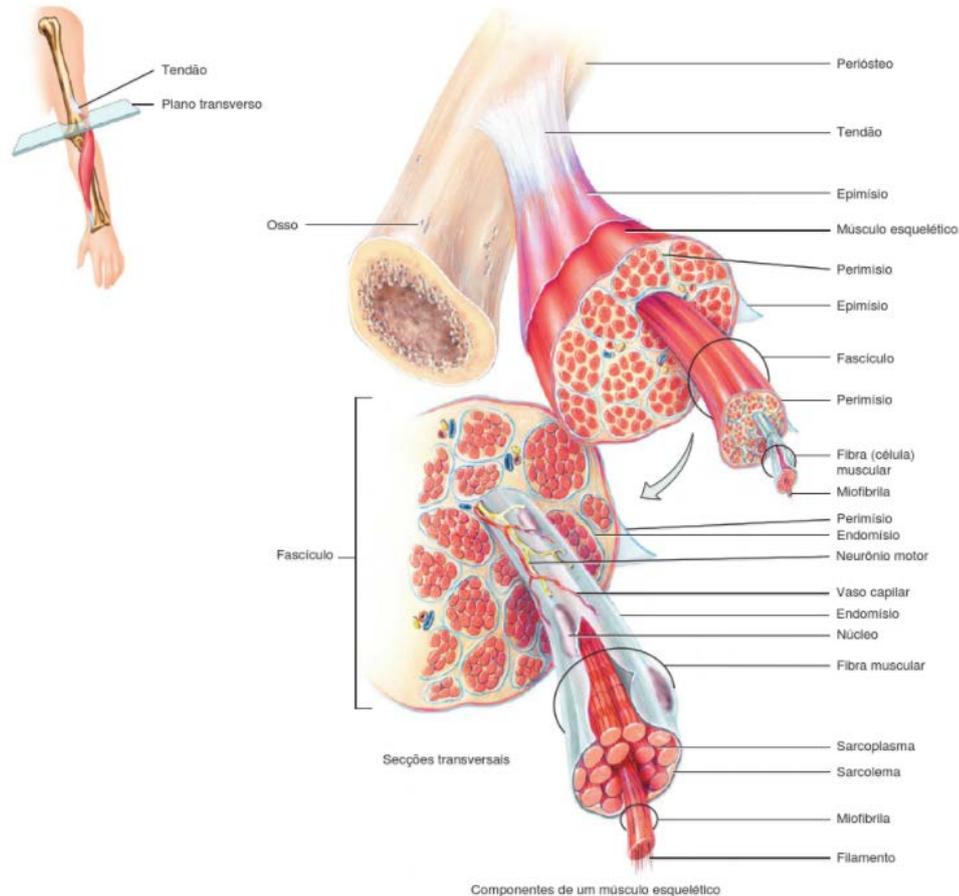
Nas seguintes subseções serão descritos mais detalhadamente cada componente associado à geração da atividade elétrica presente nas membranas musculares, bem como, a caracterização do sinal elétrico resultante.

### 2.1.1 Músculo Esquelético

Existem três tipos de tecidos muscular: o cardíaco, o liso e o esquelético. Este último será abordado nessa subseção. Como o nome já sugere, a maior parte dele está fixada nos ossos e é responsável pelo movimento do esqueleto. É estriado (possui faixas proteicas claras e escuras que podem ser vistas através de um microscópio) e, por realizar contrações e relaxar a partir de controle consciente, é de natureza voluntária (TORTORA; DERRICKSON, 2012).

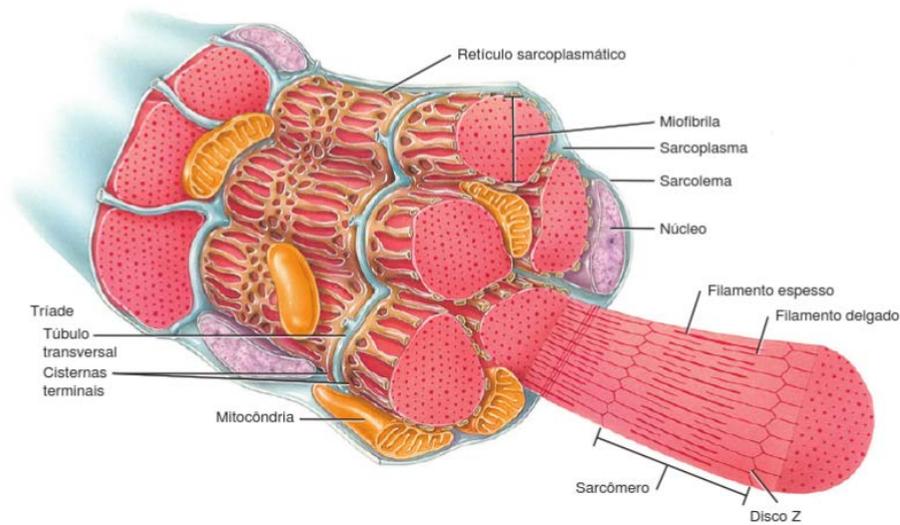
O tecido muscular é circundado por tecido conjuntivo, que tem por função proteger e possibilitar o acesso de nervos, vasos sanguíneos e linfáticos ao músculo. A *fáscia* é uma densa faixa de tecido conjuntivo que circunda e sustenta o músculo. Dela derivam-se três camadas do mesmo tecido que atuam na proteção e fortalecimento do músculo esquelético: o *Epimísio*, o *Perimísio*, e o *Endomísio*. O primeiro reveste o músculo inteiro; o segundo engloba feixes contendo de 10 a 100 ou mais fibras musculares, enquanto que o último cobre a fibra muscular. A extensão das três camadas se unificam e formam os tendões que, por sua vez, fazem a ligação do músculo com os ossos (TORTORA; DERRICKSON, 2012). A Figura 1 ilustra os componentes formadores do músculo esquelético.

O tecido muscular é formado por células cilíndricas alongadas: as fibras musculares. No citoplasma de tais fibras, denominado *Sarcoplasma*, estão situadas as Mitocôndrias. Estas são organelas celulares que possuem, dentre outras funções, a de produzir ATP (Trifosfato de Adenosina) durante uma contração muscular. Também estão presentes no *Sarcoplasma* moléculas de Mioglobina que fazem o armazenamento de Oxigênio utilizado pelas Mitocôndrias para a produção de ATP. Para a realização de uma contração, o músculo necessita de uma grande quantidade de ATP (TORTORA; DERRICKSON, 2012).



**Figura 1 Organização e identificação dos tecidos conjuntivos do músculo esquelético (adaptado de TORTORA; DERRICKSON, 2012).**

As fibras musculares são formadas por estruturas cilíndricas, as Miofibrilas. Estas por sua vez são compostas por filamentos proteicos que podem ser espessos ou delgados. Os primeiros apresentam basicamente moléculas da proteína Miosina em sua composição enquanto que os últimos são constituídos principalmente por moléculas da proteína Actina. Tais filamentos não se estendem por todo o comprimento da fibra muscular e estão sobrepostos formando compartimentos chamados Sarcômeros (TORTORA; DERRICKSON, 2012). A interação entre tais filamentos na presença de um potencial de ação caracteriza o processo contrátil do músculo. A Figura 2 apresenta uma visão microscópica da fibra muscular identificando seus componentes principais.

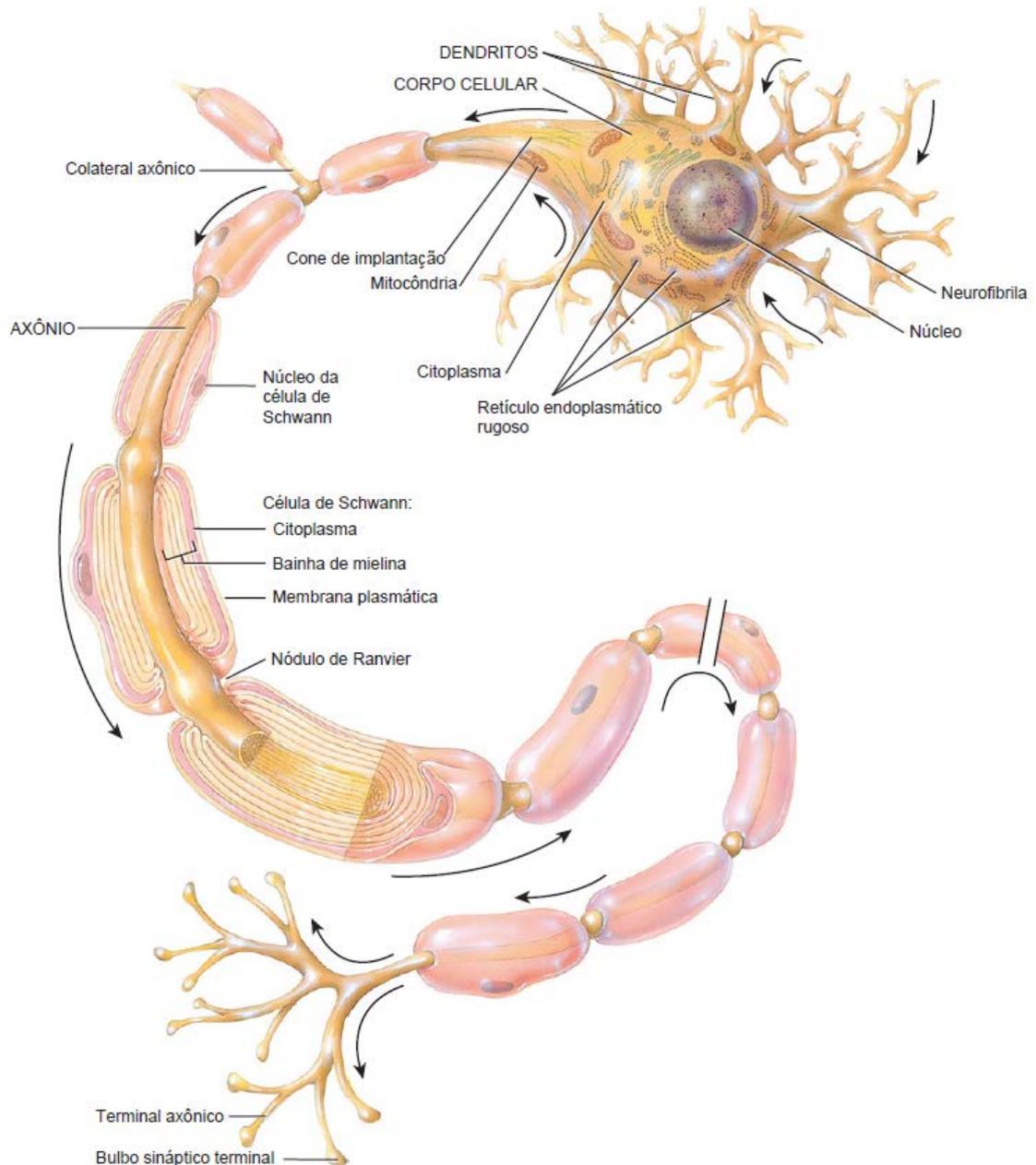


**Figura 2** Visão microscópica das fibras musculares (adaptado de TORTORA; DERRICKSON, 2012).

### 2.1.2 Neurônio Motor

Como o nome sugere, o neurônio motor está associado à função de receber estímulos do sistema nervoso central e transmitir aos músculos e glândulas do corpo. Quanto à sua estrutura, ele é classificado como multipolar, uma vez que ele possui múltiplas extensões celulares, vários dendritos e um axônio (TORTORA; DERRICKSON, 2012). Tal célula é formada basicamente por um corpo celular, dendritos e axônio. Este último se divide, em sua extremidade, em várias ramificações que se aproximam da membrana muscular para formar a junção neuromuscular. A Figura 3 mostra o desenho de um típico neurônio multipolar.

O neurônio motor recebe o estímulo elétrico através de terminais pré-sinápticos localizados na superfície dos dendritos e do corpo celular. Esses terminais são oriundos de ramificações do axônio de outros neurônios e podem ser excitatórios ou inibitórios dependendo da substância transmissora secretada por eles (HALL; GUYTON, 2011).



**Figura 3 Estrutura de um neurônio multipolar (adaptado de TORTORA; DERRICKSON, 2012).**

Na ocorrência de um potencial de ativação nos terminais pré-sinápticos, íons de Cálcio são liberados na membrana dos referidos terminais causando a despolarização da mesma. Conseqüentemente, moléculas de neurotransmissores são liberadas na fenda sináptica (espaço situado entre os terminais pré-sinápticos e o neurônio pós-sináptico) que irá excitar a membrana neuronal ou inibir dependendo dos receptores presentes na membrana do neurônio pós-

sináptico. Esse processo é gerido basicamente por duas estruturas principais, presentes nos terminais pré-sinápticos: as vesículas transmissoras, que são responsáveis pela liberação das moléculas neurotransmissoras, e a Mitocôndria, que produz ATP para o fornecimento da energia requerida na sintetização das moléculas referidas anteriormente (HALL; GUYTON, 2011).

### **2.1.3 Unidade Motora**

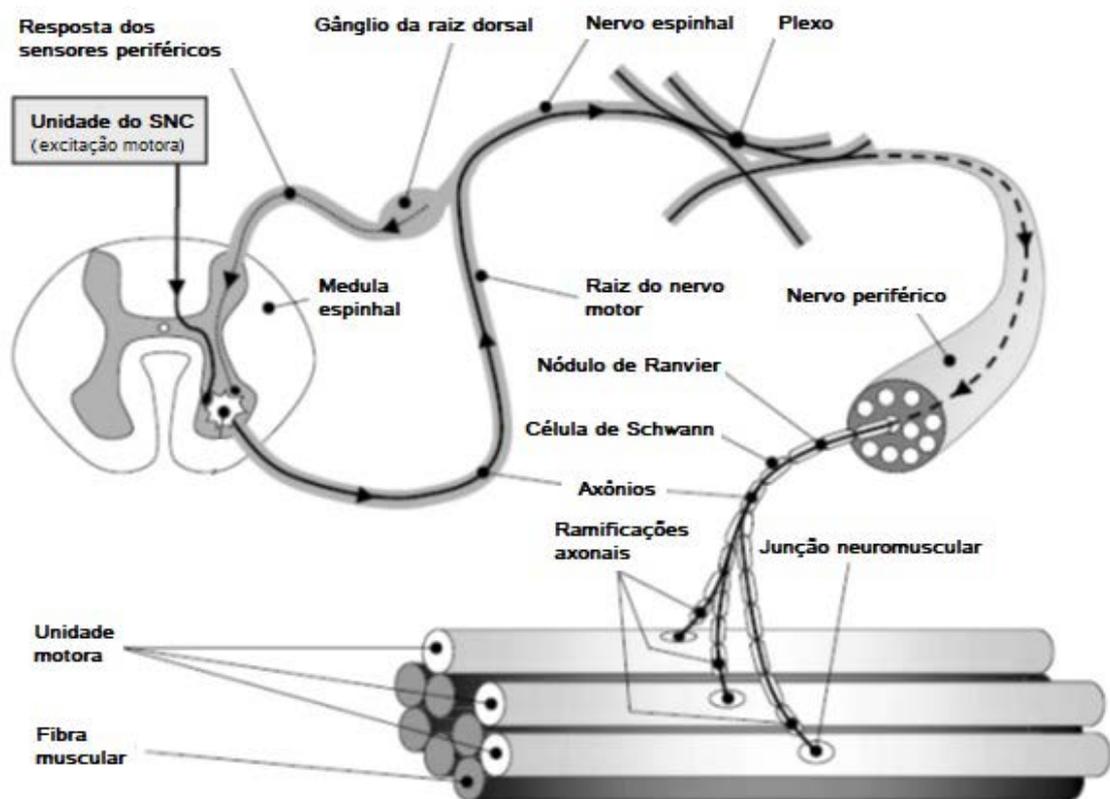
Para que a contração muscular ocorra, antes é necessário que um estímulo elétrico seja transmitido às fibras musculares através do neurônio motor. O conjunto formado pelas fibras musculares e seu respectivo neurônio motor é chamado de unidade motora. Um único neurônio motor geralmente está associado a um número elevado de fibras musculares. Em músculos que exercem força elevada como o Bíceps Braquial esse número pode chegar a algumas unidades de milhar. Já nos músculos que executam movimentos mais delicados, como o responsável pela abertura dos olhos, se encontram apenas algumas unidades de neurônios motores (TORTORA; DERRICKSON, 2012).

O neurônio motor quando adentra o músculo, se divide em ramificações chamadas terminais axônicos, os quais se aproximam do Sarcolema (camada de tecido conjuntivo que envolve a fibra muscular), porém não o tocando. As extremidades dos terminais axônicos alargam-se e dão forma aos bulbos sinápticos. A região do Sarcolema próxima aos bulbos sinápticos chama-se placa motora terminal e o espaço situado entre eles denomina-se fenda sináptica. À sinapse formada pela placa motora terminal e o bulbo sináptico dá-se o nome de junção neuromuscular. Um estímulo elétrico é transmitido através do neurônio motor ao músculo seguindo a sequência de etapas (TORTORA; DERRICKSON, 2012):

1. liberação de Acetilcolina (ACh) nos bulbos sinápticos e transmitida na fenda sináptica;

2. ativação dos receptores de ACh na placa motora terminal fazendo com que cátions, principalmente íons de Sódio, sejam liberados no Sarcolema;
3. geração do potencial de ativação: a partir da liberação de íons de Sódio no Sarcolema, origina-se um potencial de ativação na membrana, situada na região da placa motora terminal, que se propaga pela mesma;
4. degradação da ACh: o efeito da Acetilcolina é temporário e se degrada rapidamente pela ação da enzima Acetilcolinesterase, presente na fenda sináptica.

A Figura 4 ilustra os componentes da unidade motora.



**Figura 4 Partes formadoras de uma unidade motora (adaptado de POZZO; FARINA; MERLETTI, 2004).**

### 2.1.4 Potencial de Ativação

Estando o músculo em repouso, ao conectar um eletrodo em uma única célula muscular se leria um valor de potencial de repouso na faixa de 70 a 90mV negativos. Tal valor está relacionado com a proporção de íons presentes na membrana muscular. Na ocorrência de um estímulo elétrico oriundo do neurônio motor, uma região da membrana muscular é despolarizada dando origem a um potencial elétrico em tal região. Essa área em questão corresponde à zona de despolarização ou densidade de carga (POZZO; FARINA; MERLETTI, 2004). Durante uma contração, a liberação de íons de Sódio na fibra muscular acarreta no aumento do nível de potencial da zona de despolarização, atingindo valores na faixa de 50 a 75mV positivos (HALL; GUYTON, 2011). A Figura 5 apresenta um gráfico representando a resposta elétrica da membrana muscular após um estímulo proveniente do neurônio motor.

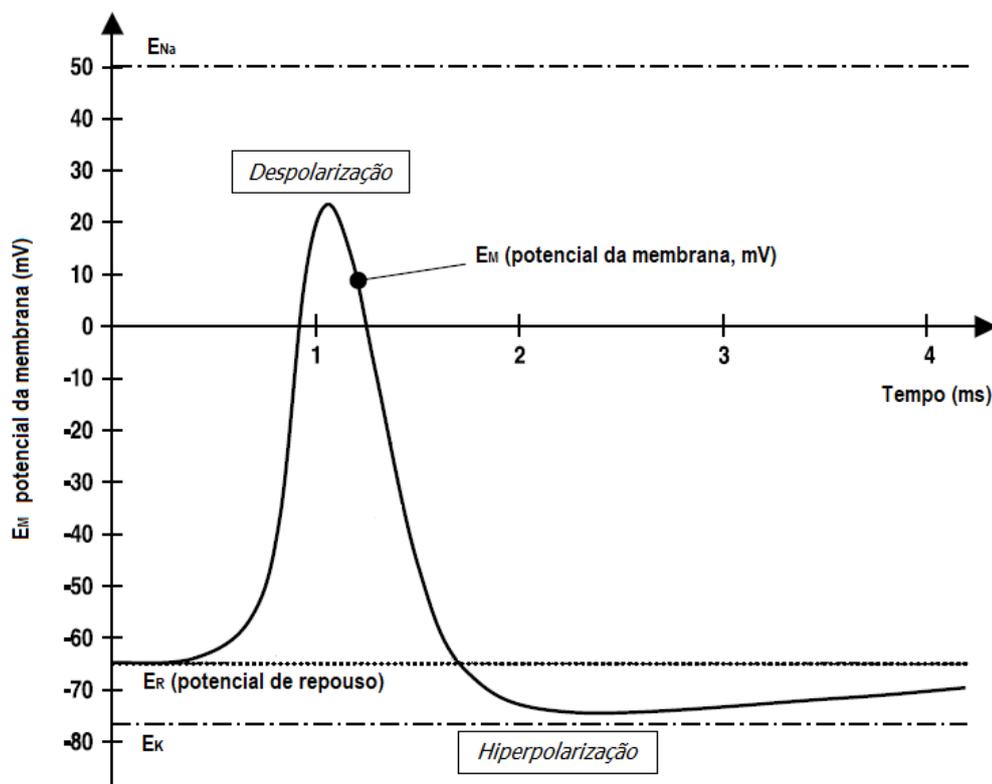


Figura 5 Potencial de ação da membrana muscular (adaptado de POZZO; FARINA; MERLETTI, 2004).

O potencial elétrico gerado na membrana muscular propaga-se sobre a mesma a uma velocidade que depende do diâmetro e tipo da fibra muscular, denominado velocidade de condução. Ele origina-se na junção neuromuscular e difunde-se até o tendão (POZZO; FARINA; MERLETTI, 2004). Segundo (ROSENFALCK, 1969), o mesmo pode ser representado analiticamente no domínio espacial, através da Equação (1):

$$V_m(z) = \begin{cases} Az^3 e^{-\lambda z} + B & ; z > 0 \\ 0 & ; z \leq 0 \end{cases} \quad (1)$$

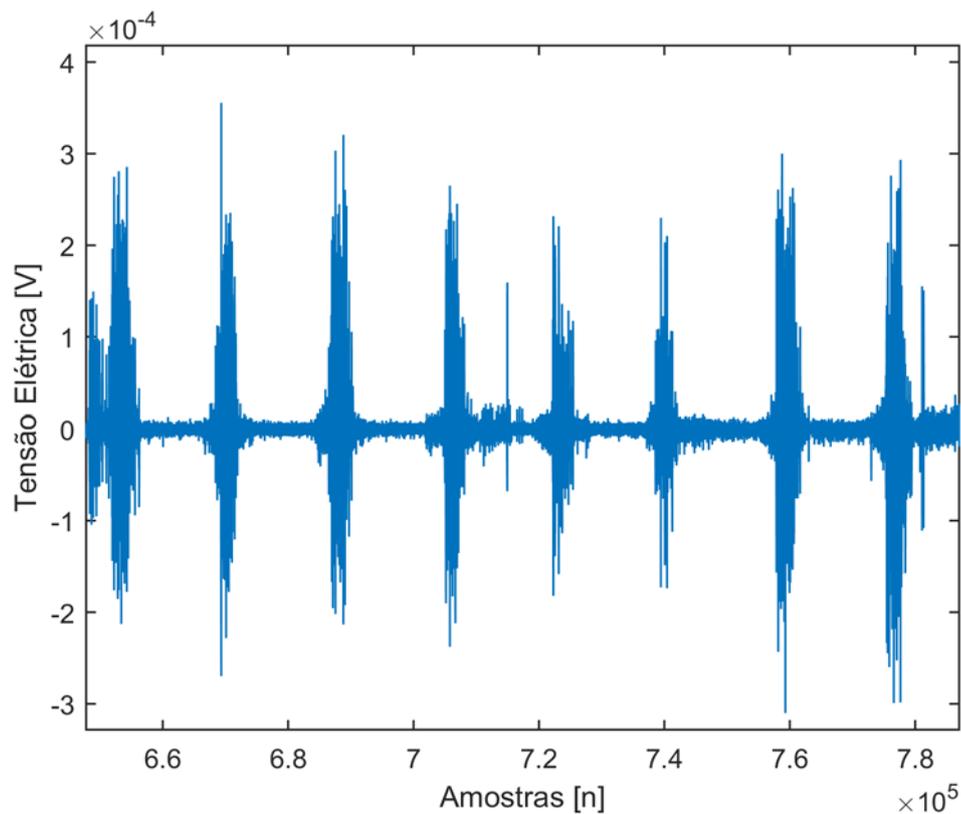
onde:  $A = 96\text{mV}\cdot\text{mm}^{-3}$ ;  $B = -90\text{mV}$ ;  $\lambda = 1\text{mm}^{-1}$ .

### 2.1.5 Princípios sobre o Sinal de EMG

O sinal de eletromiografia (EMG) pode ser adquirido de modo invasivo ou não invasivo, a partir do uso de eletrodos invasivos ou de superfície, respectivamente. A obtenção direta dos sinais elétricos das fibras musculares realizada de forma invasiva é possível a mais de 80 anos, a partir do trabalho de Adrian e Bronk, que realizaram uma análise dos impulsos elétricos gerados nas fibras musculares da pata traseira de um gato na presença de um estímulo externo, através de agulhas concêntricas (ADRIAN; BRONK, 1929). Contudo, o fato deste não ser um método muito prático e ser de natureza invasiva, pode colocar limitações em seu uso em certas aplicações (FARINA; NEGRO, 2012).

Em contrapartida, muitos trabalhos estão sendo desenvolvidos a partir do estudo do sinal de eletromiografia de superfície (sEMG) nas áreas de controle motor e também em tecnologias de reposição de unidades motoras, como por exemplo, no desenvolvimento de próteses (FARINA; NEGRO, 2012). Uma das grandes vantagens de tal método em relação ao primeiro consiste na facilidade em se fixar os eletrodos e realizar a obtenção do sinal de EMG. Além disso, os eletrodos podem ser fixados sem desconforto aos pacientes, minimizando os riscos de infecções, assim como, dependendo da área de aplicação e de interesse, sem a necessidade de supervisão médica (MERLETTI; DE LUCA, 1989).

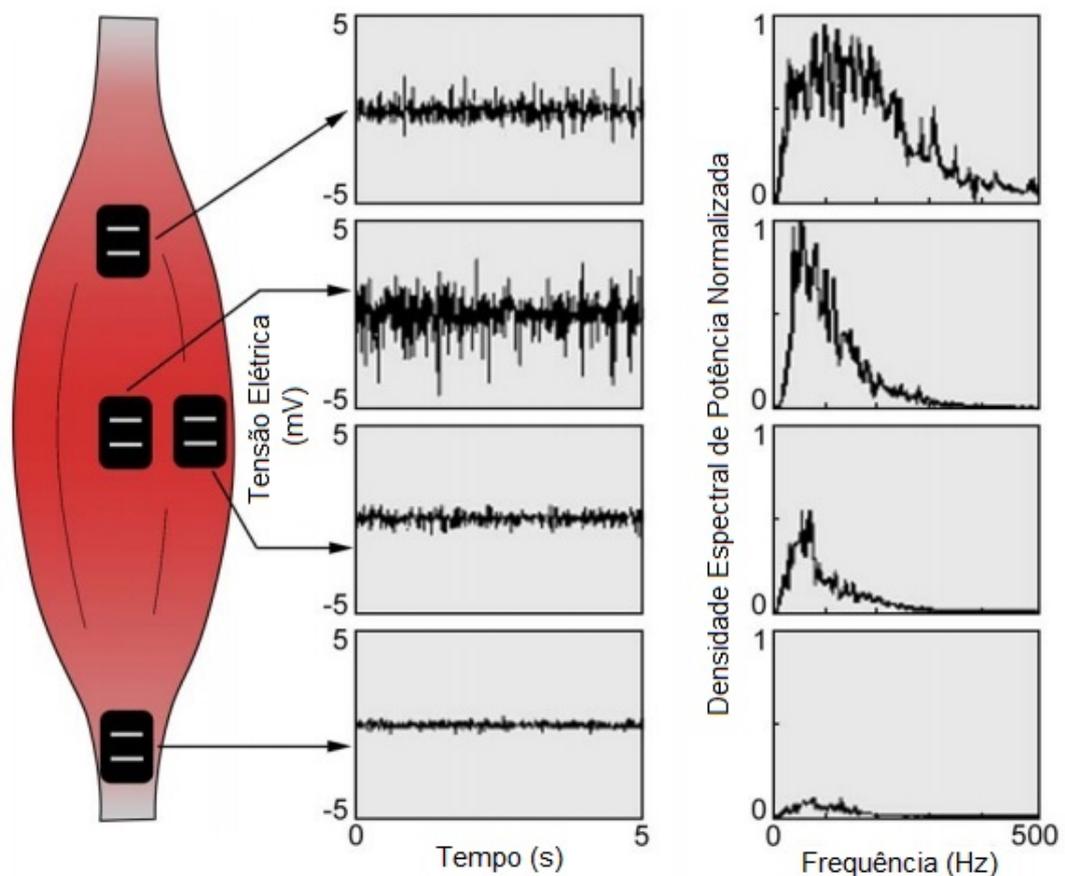
O sinal de sEMG obtido a partir de um eletrodo localizado sobre um dado músculo consiste na superposição de unidades de potencial de ação motora (MUAP – *Motor Unit Action Potential*) provenientes de unidades motoras concorrentes (MERLETTI; DE LUCA, 1989). A presença de tecidos biológicos entre as fibras musculares e os eletrodos acabam por influenciar fortemente o sinal detectado, atuando como um filtro passa-baixa (POZZO; FARINA; MERLETTI, 2004). A faixa de frequências que o sinal de sEMG está contido estende-se de 15 a 500Hz e sua amplitude varia de -5 a 5mV (ATZORI *et al.*, 2012), (KONRAD, 2005). Como exemplo, a Figura 6 ilustra parte do sinal de EMG do primeiro voluntário da base de dados 2 do Ninapro.



**Figura 6** Sinal de sEMG do primeiro voluntário da base do Ninapro (adaptado livremente da base de ATZORI *et al.*, 2014).

Como qualquer medição de uma grandeza física, a obtenção do sinal de eletromiografia realizada de modo não invasivo sofre influência de uma série de fatores que podem afetar

significativamente a qualidade do sinal adquirido. Tais circunstâncias podem ser intrínsecas ou extrínsecas em relação ao processo em si. As primeiras estão relacionadas às características fisiológicas, anatômicas e bioquímicas do músculo analisado. Já as últimas referem-se aos aspectos do sistema de aquisição utilizado, mais direcionado à estrutura dos eletrodos empregados e o posicionamento dos mesmos na superfície da pele. Por exemplo, o tamanho da área do eletrodo em contato com a epiderme está relacionado com o número de unidades motoras captadas; a distância entre eles determina a banda de frequência do sinal de EMG adquirido; a localização dos mesmos próxima às bordas do músculo analisado pode determinar o aumento de *crosstalk* (interferência de músculos vizinhos no sinal adquirido) (DE LUCA, 1997). A Figura 7 exemplifica graficamente a influência da localização dos eletrodos sobre a pele no sinal de EMG adquirido.



**Figura 7** Influência do posicionamento dos eletrodos no sinal de EMG (DE LUCA, 1997).

## 2.2 PROCESSAMENTO DE SINAIS DE ELETROMIOGRAFIA

Por se tratar de um sinal de natureza estocástica, os dados brutos de eletromiografia são pouco representativos por si só. Portanto, é necessário que tal informação seja representada de forma quantificável para que se possa compreendê-la (REAZ; HUSSAIN; MOHD-YASIN, 2006). Vários métodos de processamento têm sido utilizados para caracterizar sinais de eletromiografia. Eles baseiam-se na extração de características no domínio do tempo e/ou frequência do sinal, as quais carregam informações úteis para estimar as propriedades do músculo e quantificar o desempenho muscular (MERLETTI; DE LUCA, 1989).

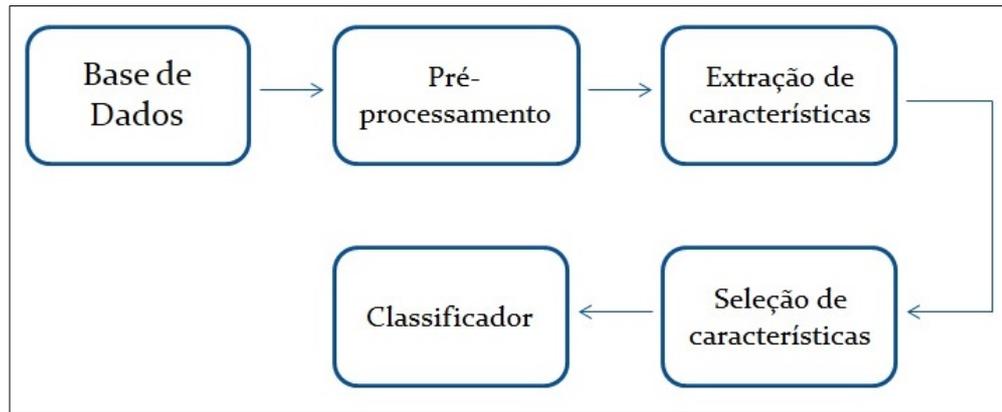
As características extraídas de sinais de EMG podem ser obtidas no domínio do tempo, frequência ou no domínio do tempo e frequência simultaneamente, sendo que cada uma delas representa um aspecto diferente do sinal. As oriundas no domínio temporal são tipicamente baseadas na amplitude do sinal e estão relacionadas ao número de unidades motoras abrangidas pelo eletrodo de superfície. Conseqüentemente, são fortemente influenciadas pelo fenômeno do cancelamento de amplitude e do ruído presente no sinal adquirido. A maior parte delas está associada à energia do sinal, a duração da contração muscular, e também indicam a relação com a força empregada (ISON; ARTEMIADIS, 2014).

Já as características relacionadas ao domínio da frequência indicam informações quanto à forma e a velocidade das MUAP's. Elas são importantes na detecção de fadiga muscular e na distinção de sinais não estacionários, porém apresentam alto custo computacional e baixa resolução temporal (ISON; ARTEMIADIS, 2014). Por fim, as características extraídas no domínio do tempo e frequência oferecem uma representação multi resolução do sinal. Normalmente tal análise é realizada a partir da transformada *Wavelet*. Esse método de processamento representa padrões estáticos de movimentos dinâmicos, tanto transientes quanto estáveis. Desse modo, características dessa natureza apresentam melhor desempenho na separação de dados oriundos de movimentos dinâmicos. Contudo, a não representação física

deste tipo de característica e o elevado custo computacional acabam por inviabilizá-la em algumas aplicações (ISON; ARTEMIADIS, 2014).

Em aplicações envolvendo a análise de sinais de EMG a escolha das características que melhor os representam é uma etapa de extrema importância. Dentre os diversos nichos de trabalhos relacionados ao estudo de eletromiografia, uma das mais nobres e que vem atraindo a atenção de vários pesquisadores é a associada com classificação de tais sinais para a caracterização de movimentos do segmento mão-braço. Essa área apresenta diversos trabalhos publicados, como por exemplo, (ABDELMASEEH; CHEN; STASHUK, 2016), (RIILLO *et al.*, 2014), (ZHAI *et al.*, 2016), (LI *et al.*, 2017), (ZHANG *et al.*, 2017b), (YAN; LIU, 2013), (CENE, 2016) que são promissores na direção do desenvolvimento de próteses com maior número de graus de liberdade. No Capítulo 4 tais trabalhos serão descritos detalhadamente onde será realizada uma análise comparativa dentre seus resultados e os obtidos no presente estudo.

Para exemplificar, a Figura 8 apresenta um diagrama de blocos típico do processo de classificação de sinais de EMG. O processo se inicia com uma base de dados contendo os sinais de EMG adquiridos, já filtrados e amplificados. Após, os sinais podem passar por um estágio de pré-processamento, o qual consiste em filtros adicionais, retificação e normalização, antes de serem extraídas as características para melhor os representar. Na sequência há a etapa de seleção de características que se trata da escolha das que melhor discriminam e separam os dados de entrada nas diferentes classes para introduzir no classificador que é normalmente a etapa final do processo.



**Figura 8** Típico diagrama de blocos para a tarefa de classificação de sinais de EMG.

De todas as etapas ilustradas no fluxograma uma das mais importantes e crucial para o desempenho do classificador é a que diz respeito a seleção de características. Para a execução de tal tarefa foram desenvolvidos uma série de métodos, com grande parte das aplicações relacionadas a classificação de genes para a detecção de câncer (ZHANG *et al.*, 2017a), (AN; WANG; WEI, 2017), (ABHISHEK; SINGH, 2017), (HAN *et al.*, 2017). No que tange a análise de sinais de EMG poucos são os trabalhos que empregam métodos de seleção de características na literatura pesquisada. Normalmente, na grande maioria dos trabalhos, é explorada somente a etapa de classificação e os mais diversos algoritmos existentes para tal (ZHOU *et al.*, 2016).

Nas próximas seções serão explorados mais detalhadamente as etapas de seleção de características e de classificação. Nelas serão expostos os algoritmos utilizados no presente trabalho.

### 2.2.1 Seleção de Características

Seleção de características é uma tarefa bastante comum nas áreas de estatística, reconhecimento de padrões e aprendizagem de máquina. Ela objetiva encontrar o conjunto de atributos mais útil para o processo em questão dentro de um determinado espaço de características disponíveis (ROKACH; CHIZI; MAIMON, 2006). Tratando-se de classificação

de dados, o intuito de se executar uma seleção de atributos está em separar apenas as características mais discriminantes para o processo de aprendizado, ou seja, somente as que conseguem discriminar as diferentes classes (TANG; ALELYANI; LIU, 2015).

Outra função importante de tal tarefa é o de redução de dimensionalidade. Em aplicações, como por exemplo, na classificação de genes para a detecção de câncer, normalmente se lida com milhares de características para apenas poucas amostras (normalmente menos de uma centena). Dentre esses milhares de genes, muitos são redundantes ou simplesmente irrelevantes para o processo de classificação e sua presença pode prejudicar a qualidade do treinamento, gerando *overfitting* e consequente perda de generalização (ZHOU; TUCK, 2007).

Os algoritmos de seleção de características podem ser divididos em: baseados em filtro, *wrapper* e *embedded*. Os métodos baseados em filtro fazem a escolha das características sem a influência de algoritmos de classificação. A avaliação e seleção é realizada seguindo a análise de algum critério estatístico. Normalmente são utilizadas métricas, como por exemplo, distância, dependência, consistência, informação e correlação (TANG; ALELYANI; LIU, 2015).

Métodos do tipo *wrapper* avaliam a qualidade de uma característica a partir da taxa de acerto de um dado algoritmo de classificação. Por realizar a seleção através do treinamento e determinação de um classificador, tal método é mais custoso computacionalmente do que o baseado em filtro. Por outro lado, dispositivos *wrapper* possuem melhores taxas de acerto segundo resultados publicados (AGGARWAL; REDDY, 2013). Existem várias estratégias de busca do melhor conjunto de características, tais quais, *hill-climbing* e algoritmos genéticos, porém os mais vantajosos computacionalmente e melhores na questão do *overfitting* são os chamados *greedy*. Eles operam a partir de duas formas distintas: seleção para frente e eliminação para trás. No primeiro caso o algoritmo começa testando apenas uma característica

e vai adicionando mais uma a cada iteração mantendo-a ou não dependendo do resultado da classificação. Já o segundo começa testando todos os atributos e vai eliminando o menos relevante a cada iteração (TANG; ALELYANI; LIU, 2015).

Por fim, métodos *embedded* preenchem a lacuna entre os outros dois tipos. Eles apresentam as vantagens dos *wrapper* e dos baseados em filtro uma vez que eles também fazem uso de um modelo de classificador, e são menos custosos computacionalmente do que os primeiros. Tais algoritmos realizam a avaliação das características durante a fase de treinamento do classificador (TANG; ALELYANI; LIU, 2015). Normalmente é discriminado um valor de importância para cada característica de acordo com a sua relevância para o processo de classificação, de modo a fazer um *ranking* entre elas. A escolha do melhor conjunto de características é então realizada a partir da eliminação das que tiverem uma menor importância relativa associada até que se tenha um número adequado segundo alguma heurística pré-determinada. Um exemplo de algoritmo desse tipo é o SVM-RFE (*Support Vector Machine Recursive Feature Elimination*).

#### **2.2.1.1 Support Vector Machine Recursive Feature Elimination (SVM-RFE)**

O SVM-RFE consiste em um algoritmo de seleção de características do tipo *embedded*. Desenvolvido por (GUYON *et al.*, 2002) para a tarefa de seleção de genes para classificação de câncer em duas classes (tem ou não tem câncer), esse método faz uso de um algoritmo estado da arte em matéria de classificação que é o SVM (MOURA; FAVIEIRO; BALBINOT, 2016), (GENG *et al.*, 2017), (XIAO; MENON, 2017). A seguir serão abordados os algoritmos da SVM e do método SVM-RFE.

### ***Support Vector Machine (SVM)***

Máquina de Vetores de Suporte (SVM) teve origem no trabalho de (BOSER; GUYON; VAPNIK, 1992) para atacar tarefas de classificação binárias. No caso das SVM's lineares, o método baseia-se no ajuste dos parâmetros de um hiperplano que separe as duas classes em questão. Tais parâmetros são otimizados de modo a maximizar a distância dentre elas. De acordo com (BOSER; GUYON; VAPNIK, 1992) este hiperplano é representado pela Equação (2):

$$f(x) = \mathbf{w} \cdot \mathbf{x} + b = 0 \quad (2)$$

onde o vetor  $\mathbf{x}$  são as instâncias de entrada,  $\mathbf{w}$  é o vetor de pesos a ser ajustado e  $b$  é um *bias*. O hiperplano da Equação (2) separa o espaço dos dados em  $\mathbf{w} \cdot \mathbf{x} + b > 0$  e  $\mathbf{w} \cdot \mathbf{x} + b < 0$ . Fazendo uso de uma função sinal, têm-se todas as amostras que forem alocadas num ponto maior do que zero do hiperplano identificadas com o número 1, e as que ficarem abaixo de zero são indicadas com -1, dando origem aos hiperplanos  $H_1$  e  $H_2$  ilustrados na Figura 9.

O objetivo do treinamento de uma SVM consiste em maximizar a distância de separação das duas classes, ou seja, do valor de “d” ilustrado na Figura 9. Como tal cota é dada em função do vetor de pesos  $\mathbf{w}$ , o algoritmo trabalha no seguinte problema de otimização (LORENA; DE CARVALHO, 2007):

$$\text{Minimizar}_{\mathbf{w},b} \quad \frac{1}{2} \|\mathbf{w}\|^2 \quad (3)$$

$$\text{Com as restrições: } y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1 \geq 0, \quad \forall i = 1, \dots, n$$

Tais restrições têm por objetivo impedir que dados de treinamento ocorram por entre as margens de separação.

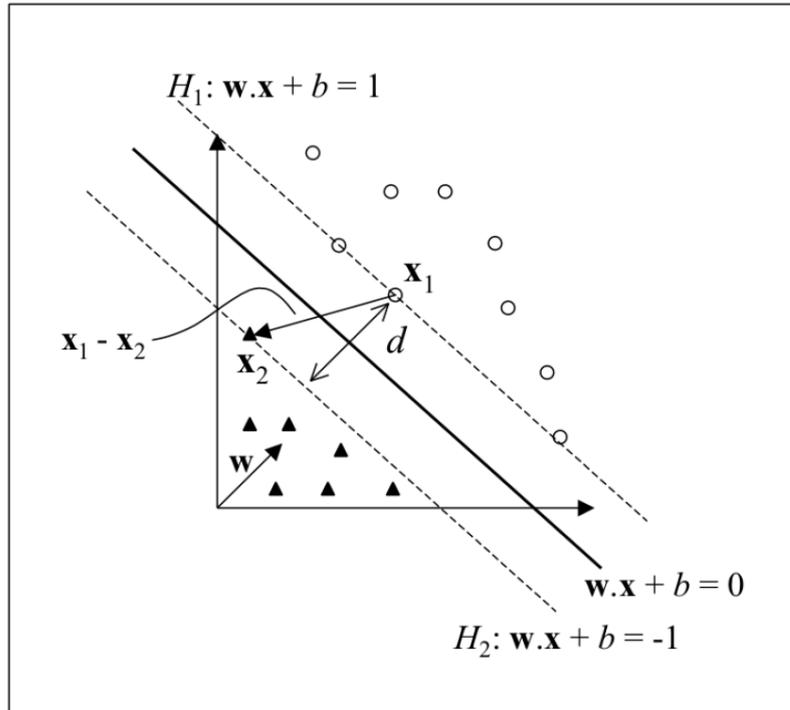


Figura 9 Hiperplano de separação das classes (HEARST *et al.*, 1998).

Como a resolução do problema de otimização exposto pela Equação (3) é de elevada complexidade matemática, introduz-se a função Lagrangeana para a simplificação da solução.

Deste modo o problema transforma-se na seguinte maximização:

$$\text{Maximizar } \alpha \sum_{i=1}^n \alpha_i - \frac{1}{2} \cdot \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j (\mathbf{x}_i \cdot \mathbf{x}_j) \quad (4)$$

$$\text{Com as restrições: } \begin{cases} \alpha_i \geq 0, \quad \forall i = 1, \dots, n \\ \sum_{i=1}^n \alpha_i y_i = 0 \end{cases}$$

Essa forma de representação do problema é chamada de dual enquanto que a primeira é conhecida como primal. A vantagem da segunda em relação à primal está na simplificação das restrições e da dependência dos dados de entrada e dos rótulos apenas, facilitando na sua resolução. Deste modo, os parâmetros  $\mathbf{w}$  e  $b$  são calculados a partir das Equações (5) e (6) respectivamente (BOSER; GUYON; VAPNIK, 1992):

$$\mathbf{w} = \sum_{i=1}^n \alpha_i y_i \mathbf{x}_i \quad (5)$$

$$b = \frac{1}{N_{sv}} \sum_{\mathbf{x}_j \in SV} \frac{1}{y_j} - \mathbf{w} \cdot \mathbf{x}_j \quad (6)$$

onde  $N_{sv}$  é o número de vetores de suporte (*support vector*) que correspondem aos dados cujos multiplicadores de Lagrange  $\alpha_i$  são diferentes de zero. Em outras palavras, são os dados que se encontram sobre os hiperplanos de separação das classes,  $H_1$  e  $H_2$ .

A proposta apresentada anteriormente é válida apenas para dados linearmente separáveis. Como em situações reais tal condição é rara, permite-se uma certa margem de relaxamento nas restrições. A partir dessa suavização nas margens de separação passam a ser aceitos dados localizados entre os hiperplanos e até mesmo algum erro de classificação. Para tal efeito no classificador, adiciona-se o termo  $\xi_i$  que atua como uma variável de folga nas restrições da tarefa original (BOSER; GUYON; VAPNIK, 1992). Assim, a forma primal do problema de otimização passa a ser:

$$\text{Minimizar}_{\mathbf{w}, b, \xi} \quad \frac{1}{2} \|\mathbf{w}\|^2 + C \left( \sum_{i=1}^n \xi_i \right) \quad (7)$$

Com as restrições:  $y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 - \xi_i$ ,  $\xi_i \geq 0$ ,  $\forall i = 1, \dots, n$

E a forma dual passa a ser:

$$\text{Maximizar}_{\alpha} \quad \sum_{i=1}^n \alpha_i - \frac{1}{2} \cdot \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j (\mathbf{x}_i \cdot \mathbf{x}_j) \quad (8)$$

Com as restrições:  $\begin{cases} 0 \leq \alpha_i \leq C, \quad \forall i = 1, \dots, n \\ \sum_{i=1}^n \alpha_i y_i = 0 \end{cases}$

O fator  $C$  representa um termo regulador do classificador. Ele fornece um peso à minimização dos erros frente a minimização da complexidade do modelo gerado. Trata-se de um parâmetro ajustável do classificador. Quanto maior for o valor ajustado para  $C$  menor será a complexidade do modelo, porém maior será o erro.

### Algoritmo da SVM-RFE

O algoritmo de seleção de características SVM-RFE faz uso de um classificador para atribuir valores de importância para as mesmas, sendo, desse modo, do tipo *embedded*. O método consiste em analisar os valores dos pesos ajustados durante o treinamento (vetor  $\mathbf{w}$  da Equação (2) do hiperplano de separação das classes) para determinar a relevância das características para o processo de classificação. Portanto, segundo (GUYON *et al.*, 2002), um valor de importância é atribuído às características de acordo com:

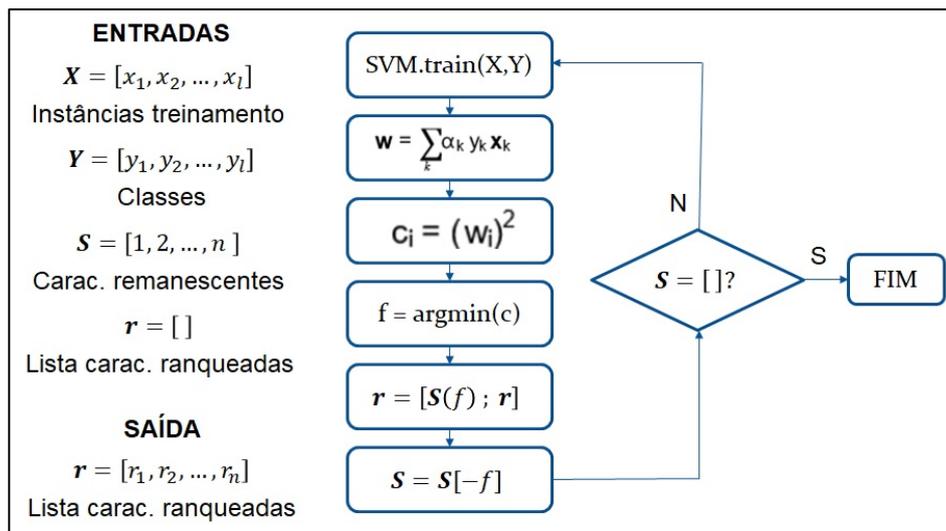
$$c_i = (w_i)^2 \quad (9)$$

onde  $c$  é o valor atribuído à característica  $i$  e  $w_i$  é o peso ajustado para tal característica durante o treinamento. O fator  $c$  é então utilizado como critério de ordenamento. Quanto maior for o valor de tal critério mais relevante será a característica para o processo de classificação.

O algoritmo é inicializado com todos os atributos, os quais são utilizados para treinar um classificador por SVM. Calcula-se então os valores de  $c$  associados a cada um deles e retira-se do conjunto aquele que tiver o menor valor. A seguir é realizado novo treinamento com as características remanescentes, e, da mesma forma, a considerada menos relevante é subtraída. O algoritmo segue iterativamente até que não permaneça mais nenhum atributo. Como resultado, se obtém uma lista com todas características ordenadas por importância. A Figura 10 apresenta o fluxograma básico deste algoritmo.

Tal método de seleção de características é amplamente utilizado em diversas áreas do conhecimento (YIN *et al.*, 2016b), (FU *et al.*, 2014), (WANG *et al.*, 2011), (KIM, 2011), (ZHOU; TUCK, 2007). Como pode-se citar, diversos trabalhos na área de diagnóstico em sistemas de classificação de câncer fazem uso desta ferramenta, justificando sua importância e utilidade em processamento de dados. Na pesquisa de (YIN *et al.*, 2016b) a SVM-RFE foi utilizado na tarefa de classificação de dados de câncer de mama. O referido método foi comparado com o uso direto da SVM em tal análise, apresentando melhores taxas de acerto.

No trabalho de (YIN *et al.*, 2016a) utilizou-se o algoritmo de seleção de características no problema de identificação de câncer de próstata. Porém, os autores constataram que este método não lidou bem com características redundantes. Ao retirar as características que apresentaram correlação maior do que 0,9 antes de aplicar a SVM-RFE observou-se um aumento na taxa de acerto do classificador. Já na pesquisa de (ALBASHISH *et al.*, 2015) o algoritmo de seleção de características foi usado no problema de diagnóstico de câncer de próstata. Nele a importância atribuída a uma característica foi resultado de uma combinação linear entre o fator determinado pela SVM-RFE e pelo fator informação mútua normalizada (NMI). Este fator atribui um valor mais baixo para características redundantes, amenizando sua influência no correspondente classificador.



**Figura 10 Fluxograma do método SVM-RFE.**

### 2.2.1.2 Monte Carlo Feature Selection

O método de seleção de características Monte Carlo foi desenvolvido no trabalho de (DRAMINIŃSKI *et al.*, 2008). Ele consiste num algoritmo do tipo *embedded* que utiliza informações do treinamento e taxas de acerto obtidas da criação de inúmeras árvores de decisão para atribuir valores de importância relativa às características.

## Árvores de Decisão

Árvores de decisão fazem parte de um tipo de método de aprendizagem de máquina que representam o conhecimento adquirido por meio de regras. Por esse motivo o processo de aprendizado fica mais compreensível e intuitivo.

Tal método baseia-se em sucessivas partições das instâncias de treinamento de acordo com testes realizados sobre suas características. Cada um dos atributos mede um aspecto importante do objeto a ser classificado e contém um conjunto de valores discretos mutuamente exclusivos. Deste modo, a cada nível da árvore uma característica é testada e os exemplos de entrada são então divididos entre os possíveis valores de tal atributo. Esse processo segue com sucessivas repartições dos dados a partir de avaliações dos atributos até que em cada um dos subconjuntos resultantes apresentem apenas instâncias de uma única classe. Na árvore resultante cada folha representa uma classe, os atributos testados constituem nós com um galho para cada saída possível (QUINLAN, 1986).

Em uma árvore de decisão o processo de classificação de um objeto normalmente se inicia pela raiz da árvore, passando por cada nó que, por sua vez, conduzirá por uma série de galhos e novos nós até que se chega à folha que representa a classe da referida instância (QUINLAN, 1986).

Existem diversas possibilidades de árvores para o mesmo problema de classificação. As árvores com menor grau de complexidade teoricamente apresentam maior conhecimento sobre a estrutura dos dados e, conseqüentemente, tendem a ter um melhor desempenho na classificação de novas instâncias (QUINLAN, 1986). Por isso a escolha do atributo o qual se fará a partição do *dataset* é de grande importância.

Pode-se comparar uma árvore de decisão a uma fonte de uma mensagem, onde a decisão de atribuir uma classe específica a um dado exemplo de entrada corresponde à mensagem. Deste modo considerando os conceitos da Teoria da Informação, pode-se calcular o ganho de

informação obtido com a divisão imposta ao conjunto de dados por um atributo, como por exemplo, a redução na entropia dos dados, provocada pela repartição (QUINLAN, 1986).

Assim, de acordo com (QUINLAN, 1986) considerando-se que a entropia das instâncias de treinamento de um dado conjunto  $S$  corresponde à Equação (10), pode-se calcular o ganho obtido pelo teste em um atributo  $A$  a partir da Equação (11):

$$E(S) = \sum_{i=1}^N p_i \log_2 p_i \quad (10)$$

$$G(A) = E(S) - \sum_{j=1}^M \frac{|S|}{|S_j|} \cdot E(S_j) \quad (11)$$

onde:  $p_i$  é a probabilidade de ocorrência da classe  $i$ ,  $N$  é o número de classes,  $S_j$  é o conjunto de dados que correspondem ao  $j$ -ésimo valor do atributo  $A$ ,  $E(S_j)$  é a entropia dos dados relacionados ao  $j$ -ésimo valor do atributo  $A$ , e  $M$  é o número de valores distintos da referida característica.

O conceito de ganho de informação associado a um atributo é utilizado pelo algoritmo ID3 (*Induction Decision 3*). Este algoritmo faz a escolha da característica que será utilizada no nó da árvore de modo a maximizar o ganho. Sendo assim é possível obter árvores de baixa complexidade e boa capacidade de generalização para instâncias diferentes das utilizadas no treinamento, fazendo dele um dos algoritmos mais utilizados para tarefa de classificação (ZOU *et al.*, 2012), (SUAREZ; SALCEDO, 2017).

#### **Algoritmo do Monte Carlo Feature Selection**

O método de seleção de características Monte Carlo faz uso intenso de árvores de decisão para atribuir um valor de importância relativa às características. Para tanto ele considera a taxa de acerto média de todas as árvores em que o atributo em questão foi selecionado, o

ganho relacionado a ele e o número de instâncias presentes no nó em que ele atua (DRAMÍŃSKI et al., 2008).

O funcionamento deste algoritmo baseia-se na construção de milhares de árvores de decisão utilizando subconjuntos de atributos selecionados aleatoriamente da base de dados. Considerando que a base tenha  $d$  características, o algoritmo gera  $s$  subconjuntos com  $m$  atributos selecionados aleatoriamente. Onde  $m$  é um parâmetro fixo e é muito menor do que  $d$ . As instâncias de treinamento de cada um dos subconjuntos (que é igual ao número total da base de dados) são divididas em aproximadamente 66% para construção da árvore e o restante para teste. Essa divisão é feita  $t$  vezes para cada subconjunto em permutações aleatorizadas da base de dados. Ao todo  $st$  árvores são criadas (DRAMÍŃSKI et al., 2008). A Figura 11 ilustra o funcionamento deste algoritmo.

Segundo (DRAMÍŃSKI et al., 2008), com base nesse funcionamento, estima-se um valor de importância relativa a um dado atributo  $gk$  através da Equação (12):

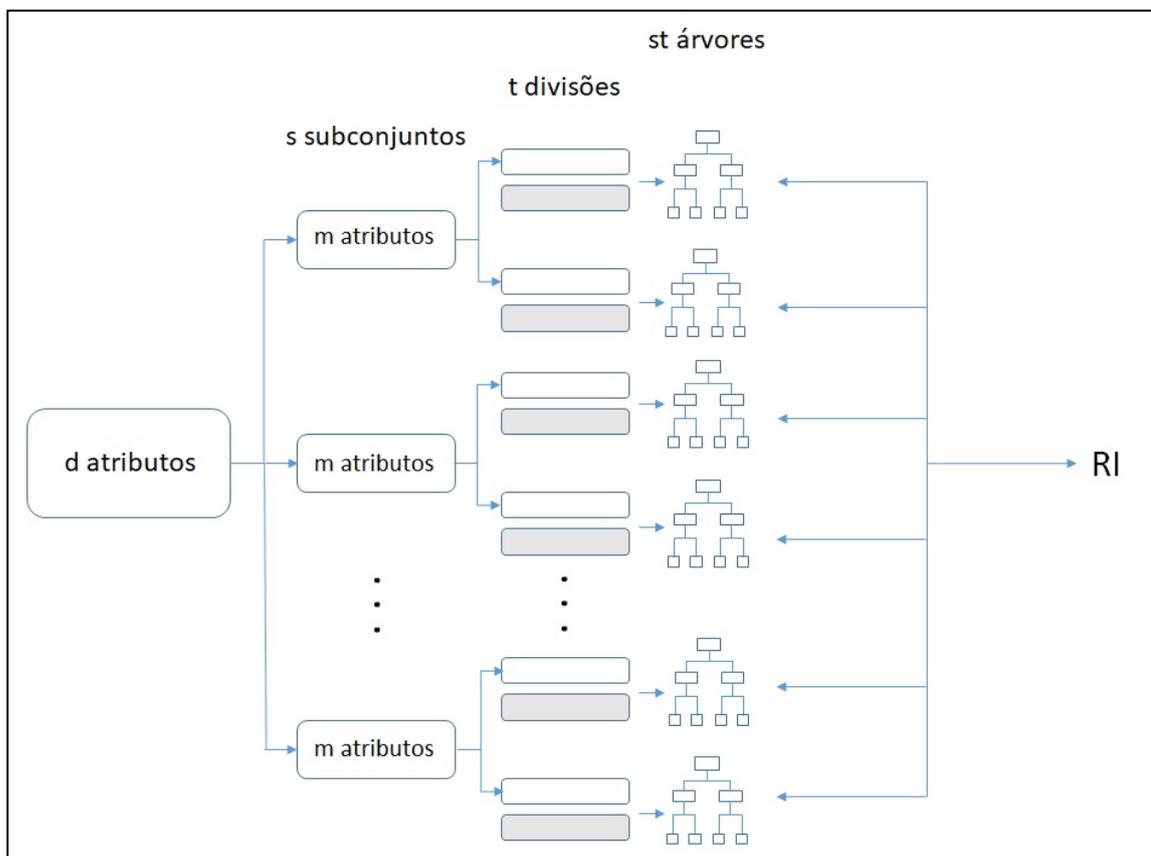
$$RI(gk) = \sum_{\tau=1}^{st} (wAcc)^u \sum_{n_{gk}(\tau)} IG(n_{gk}(\tau)) \left( \frac{\text{no. in } n_{gk}(\tau)}{\text{no. in } (\tau)} \right)^v \quad (12)$$

onde  $n_{gk}(\tau)$  corresponde a todos os nós da árvore  $\tau$  em que ocorre uma divisão no atributo  $gk$ ,  $st$  é o número total de árvores construídas,  $IG(n_{gk}(\tau))$  é o ganho de informação da divisão efetuada no nó  $n_{gk}(\tau)$  da árvore  $\tau$ ,  $(\text{no. in } n_{gk}(\tau))$  é o número de instâncias no nó  $n_{gk}(\tau)$ ,  $(\text{no. in } (\tau))$  é o número total de instâncias da árvore  $\tau$ , e  $u$  e  $v$  são valores constantes reais positivos. O fator  $wAcc$  é a taxa de acerto média por classe da árvore  $\tau$  e é dada pela Equação (13):

$$wAcc = \frac{1}{c} \sum_{n=1}^c \frac{n_{ii}}{n_{i1} + n_{i2} + \dots + n_{ic}} \quad (13)$$

onde  $c$  é o número de classes, e  $n_{ij}$  é o número de instâncias pertencentes à classe  $i$  que foi classificada como sendo classe  $j$ .

Analisando as Equações (12) e (13) é possível verificar que quanto maior o número de árvores em que o atributo é utilizado, e a taxa de acerto média de tais árvores, maior será a importância relativa do mesmo. Da mesma forma, quanto maior o ganho de informação no nó onde tal atributo foi testado, e maior for o número de instâncias presentes no referido nível da árvore, maior será sua importância relativa. Os fatores  $u$  e  $v$  são parâmetros de entrada e determinam pesos distintos para a taxa de acerto média das árvores e o ganho de informação adquirido a partir do atributo possibilitando ao usuário priorizar um em relação ao outro (DRAMÍŃSKI et al., 2008).



**Figura 11** Funcionamento do método de seleção de características Monte Carlo (adaptado livremente de DRAMINSKI *et al.*, 2008).

O método contempla ainda mais duas etapas de validação e uma de confirmação dos resultados, com o objetivo de avaliar a significância estatística dos mesmos. A primeira validação é feita repetindo-se o algoritmo várias vezes com as classes permutadas

aleatoriamente mantendo as instâncias fixas. O intuito dessa verificação é confrontar a distribuição de taxas de acerto obtidas com as permutações das classes, com a taxa de acerto média obtida nas *st* árvores criadas originalmente, de modo a comprovar que elas são significativamente diferentes. Assim se prova que as classes não são independentes dos atributos, ou, em outras palavras, que os atributos possuem informação que os relaciona com as classes (DRAMINIŃSKI et al., 2008).

O segundo ponto de validação objetiva a verificação de que é possível se obter resultados aceitáveis de classificação utilizando um pequeno conjunto de atributos selecionados dentre os mais importantes segundo o algoritmo. Isso é realizado através da classificação de milhares de árvores montadas a partir de subconjuntos de  $b$  atributos selecionados aleatoriamente dentre os primeiros  $2b$  atributos mais importantes, e milhares de árvores montadas sobre conjuntos de  $b$  atributos selecionados randomicamente dentre os demais. As instâncias de cada um destes subconjuntos são divididas em treinamento e teste em uma série de permutações aleatorizadas. Assim, obtém-se duas distribuições de taxas de acerto média: a dos subconjuntos de  $b$  atributos selecionados dentre os  $2b$  mais importantes, e a oriunda das árvores criadas a partir dos subconjuntos formados pelos demais atributos. Deste modo, através da comparação entre tais distribuições visa-se a certificação de que a primeira é significativamente superior à segunda (DRAMINIŃSKI et al., 2008).

Adicionalmente, existe uma etapa de confirmação que consiste na divisão dos dados em validação final e teste final. O primeiro compreende 75% dos dados e o segundo o restante. Com o grupo detentor da maior parcela dos dados executa-se o algoritmo de modo a obter-se o ranking dos atributos. Após, com o outro conjunto de instâncias realiza-se a segunda etapa de validação, explicada anteriormente. Assim, a confirmação da significância dos resultados da classificação dos atributos mais importantes pode ser dita incondicionalmente verdadeira (DRAMINIŃSKI et al., 2008).

### 2.2.1.3 Singular Value Decomposition (SVD) Entropy

Desenvolvido por (VARSHAVSKY *et al.*, 2006), tal método de seleção de características faz uso do conceito de entropia dos valores singulares da matriz de dados para ranquear os atributos. Deste modo, trata-se de um algoritmo não supervisionado, uma vez que não utiliza as classes para determinar a relevância das características. O algoritmo leva em consideração a contribuição de cada atributo na referida entropia para a determinação do ranking. Quanto maior a contribuição mais elevado será seu ordenamento.

Dada uma matriz de dados  $A$  com  $n$  linhas e  $m$  colunas, onde  $n$  corresponde às instâncias representadas por  $m$  atributos, sua decomposição em valores singulares é realizada a partir da representação da mesma pela multiplicação de outras 3 matrizes, conforme Equação (14) (OKUN, 2013):

$$A = USV' \quad (14)$$

onde:  $U$  é uma matriz unitária de dimensões  $(n \times n)$  formada pelos autovetores da matriz  $AA'$  também chamados de vetores singulares à esquerda de  $A$ ;  $V'$  é a transposta da matriz unitária  $V$  de dimensões  $(m \times m)$  constituída pelos autovetores da matriz  $A'A$ , também chamados de vetores singulares à direita de  $A$ ;  $S$  corresponde a uma matriz diagonal de dimensões  $(n \times m)$  cujos elementos da diagonal principal são os valores singulares da matriz  $A$  que são obtidos através da raiz quadrada dos autovalores da matriz  $A'A$ .

A entropia dos valores singulares de uma matriz de dados foi definida originalmente no trabalho de (ALTER; BROWN; BOTSTEIN, 2000) e é definida pela Equação (15):

$$E = -\frac{1}{\log m} \sum_{i=1}^m S_i \log S_i \quad (15)$$

onde:

$$S_i = \frac{\sigma_i^2}{\sum_{j=1}^m \sigma_j^2} \quad (16)$$

corresponde ao valor normalizado relativo do autovalor  $i$  da matriz  $A'A$ , e  $\sigma_i^2$  está relacionado ao autovalor  $i$  da matriz  $A'A$ .

Esse valor de entropia varia entre 0 e 1. A entropia será mínima quando a distribuição dos dados da matriz está superordenada e esta pode ser representada por um único autovetor associado ao único autovalor diferente de zero. Em contrapartida, a entropia será máxima para uma matriz de dados desordenada, representada por um conjunto de autovalores distribuídos uniformemente, ou seja, todos com mesma magnitude.

Assim, pode-se definir a contribuição de um dado atributo para a entropia dos autovalores da base de dados através da Equação (17) (VARSHAVSKY *et al.*, 2006):

$$CE_i = E(A_{[n:m]}) - E(A_{[n:(m-1)]}) \quad (17)$$

onde,  $E(A_{[n:m]})$  é a entropia dos autovalores da matriz  $A$  e  $E(A_{[n:(m-1)]})$  é a entropia da matriz  $A$  subtraindo-se o atributo analisado. Deste modo, se na remoção do atributo o valor da entropia da base de dados não sofrer alteração, diz-se que sua contribuição é nula, logo ela não terá muita relevância para a tarefa de classificação. Por outro lado, caso sua subtração influir numa redução na entropia da matriz  $A$ , sua contribuição será positiva e, conseqüentemente, ele terá certa importância.

Após a determinação da contribuição de cada atributo, os mesmos são ordenados de forma decrescente formando assim um ranking por importância.

### 2.2.2 Classificação

#### *Extreme Learning Machine (ELM)*

Redes neurais artificiais, é um método clássico de aprendizado de máquina e é amplamente utilizado em tarefas de classificação de sinais de EMG (RIILLO *et al.*, 2014), (KARABULUT *et al.*, 2017), (BASPINAR; VAROL; YILDIZ, 2012). As primeiras

informações sobre neurocomputação surgiram na publicação de (MCCULLOCH; PITTS, 1943) e desde então despertou atenção de diversos pesquisadores na área. A ideia deste método está na obtenção de aprendizado de modo similar ao que ocorre nos cérebros dos animais, por meio de uma rede de neurônios artificiais. Assim, o conhecimento adquirido é armazenado através de forças de conexão dos neurônios, representadas por pesos sinápticos. O treinamento é realizado então através do ajuste destes pesos até que o objetivo almejado seja alcançado (HAYKIN, 2001).

Existem diversos algoritmos de aprendizado para o treinamento de redes neurais. Um dos mais utilizados é o baseado na descida do gradiente, o qual realiza o aprendizado através da localização do mínimo da função de erro do sistema. O mínimo da função, por sua vez, é encontrado por meio da minimização do gradiente desta função. Assim, todos os parâmetros da rede são ajustados iterativamente na direção da minimização do gradiente. Por conseguinte, o tempo de convergência deste método acaba sendo elevado além de correr o risco de convergir em um mínimo local (HUANG; ZHU; SIEW, 2004).

Diferentemente de tais métodos de aprendizado, a *Extreme Learning Machine* (ELM) é um algoritmo de classificação e/ou regressão baseado em redes neurais que possui um aprendizado rápido em comparação com os descritos anteriormente. Desenvolvido na pesquisa de (HUANG; ZHU; SIEW, 2004), o referido método de classificação consiste na formação de uma rede neural de camada única através de aprendizado *feedforward*. A Figura 12 ilustra a típica topologia formada pelo ELM.

O ELM não faz o ajuste de todos os parâmetros da rede neural como o fazem os métodos tradicionais. Seu funcionamento consiste na determinação de valores aleatórios para os parâmetros relacionados aos neurônios da camada oculta (os pesos  $w$  e os *biases*  $b$  conforme representação na Figura 12) e, a partir de tais valores, os pesos dos neurônios da camada de saída ( $\beta$  conforme representação na Figura 12) são então obtidos analiticamente (HUANG; ZHU; SIEW, 2004).

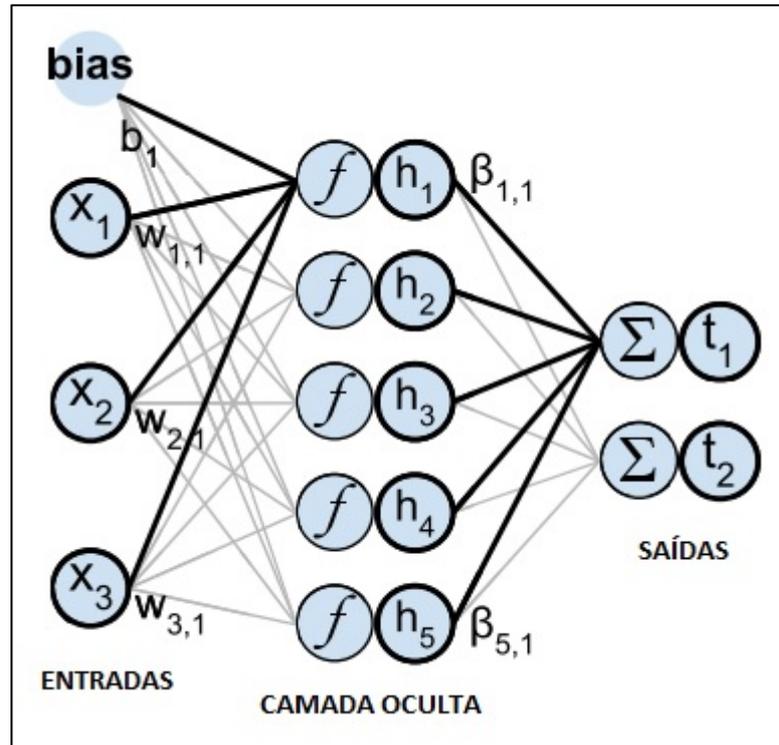


Figura 12 Estrutura da rede neural utilizada pelo ELM (adaptado de AKUSOK *et al.*, 2015).

Assim, definindo uma rede neural de camada única com  $N$  entradas e  $\tilde{N}$  neurônios, e função de ativação dada por  $g(x)$ , a expressão para a saída da rede pode ser representada pela Equação (18):

$$\sum_{i=1}^{\tilde{N}} \beta_i g(\mathbf{w}_i \cdot \mathbf{x}_j + b_i) = \mathbf{t}_j, \text{ para } j = 1, \dots, N \quad (18)$$

onde  $\beta$  é o vetor de pesos que conecta os neurônios da camada oculta aos da saída,  $\mathbf{w}$  é o vetor de pesos que liga os neurônios da camada oculta aos da entrada,  $b$  é o valor de limiar associado a cada neurônio, e  $\mathbf{t}$  é o vetor com as classes a serem preditas.

As  $N$  equações definidas na Equação (18) podem ser escritas de forma compacta a partir da transformada linear ilustrada através da Equação (19):

$$\mathbf{H}\beta = \mathbf{T} \quad (19)$$

onde  $\mathbf{H}$  é a matriz de saídas da camada oculta e é dada por:

$$H = \begin{bmatrix} g(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) & \cdots & g(\mathbf{w}_{\tilde{N}} \cdot \mathbf{x}_1 + b_{\tilde{N}}) \\ \vdots & \ddots & \vdots \\ g(\mathbf{w}_1 \cdot \mathbf{x}_N + b_1) & \cdots & g(\mathbf{w}_{\tilde{N}} \cdot \mathbf{x}_N + b_{\tilde{N}}) \end{bmatrix}_{N \times \tilde{N}} \quad (20)$$

Deste modo, uma vez conhecida a matriz  $H$  o problema torna-se a resolução de um sistema de equação lineares. Como na grande maioria dos casos reais o número de instâncias de treinamento é muito maior do que o número de neurônios, tal sistema encontra solução na determinação da matriz inversa generalizada de  $H$  (HUANG; ZHU; SIEW, 2004).

Segundo (BARTLETT, 1998) para uma grande rede neural com ínfimo valor de pesos e que possui pequeno erro de classificação nas instâncias de treinamento, a capacidade de generalização da rede está associada com a minimização dos pesos ao invés do aumento do número deles. Com base nisso, o algoritmo ELM busca a solução da Equação (19) que minimiza o erro quadrático e que também é o de menor norma (HUANG; ZHU; SIEW, 2004).

Por conseguinte, o algoritmo resolve o sistema da Equação (19) através da determinação da matriz inversa generalizada (ou pseudoinversa) de Moore-Penrose associada a  $H$ . Isso porque, de acordo com o Teorema 2.1 de (SERRE, 2002), a solução de um sistema de equações obtida a partir da matriz inversa generalizada de Moore-Penrose é a de norma mínima que minimiza o erro médio quadrático.

Segundo a definição da matriz inversa generalizada de Moore-Penrose, para uma matriz  $A$  de dimensões  $(m:n)$ , será sua pseudoinversa a matriz  $G$  de dimensões  $(n:m)$  caso sejam satisfeitas as seguintes condições:

$$\begin{aligned} AGA &= A \\ GAG &= G \\ AG^T &= AG \\ GA^T &= GA \end{aligned} \quad (21)$$

Definindo um sistema de equações lineares do tipo  $Ax = y$ ,  $x_0$  será a solução de menor norma que minimiza o erro médio quadrático se:

$$\|x_0\| \leq \|x\| \quad \forall x \in \{x: \|Ax - y\| \leq \|Az - y\| \quad \forall z \in R^n\} \quad (22)$$

Em outras palavras, uma solução para um dado sistema de equações será o de mínimo erro médio quadrático de norma mínima se ele for o que apresentar a menor norma dentre todas as possíveis soluções de mínimo quadrático.

Voltando ao algoritmo ELM, o vetor de pesos dos neurônios da camada de saída ( $\beta$ ) é então calculado a partir da Equação (23):

$$\beta = H^\dagger T \quad (23)$$

onde  $H^\dagger$  é a pseudoinversa da matriz  $H$  e o vetor  $\beta$  (pesos dos neurônios da camada de saída) corresponde à única solução de mínima norma por mínimos quadrados do sistema dado pela Equação (19).

Deste modo, o algoritmo do classificador ELM pode ser descrito resumidamente pelos seguintes passos:

- atribui valores aleatoriamente ao vetor de pesos dos neurônios da camada oculta ( $w$ ) e aos seus respectivos limiares ( $b$ );
- a partir dos valores designados, calcula a matriz de saídas da camada oculta  $H$ ;
- calcula os pesos de saída  $\beta$  através da Equação (23).

### ***Regularized Extreme Learning Machine (RELM)***

Apesar de todas as vantagens advindas do método ELM, o mesmo possui certas limitações. Uma delas diz respeito ao fato do algoritmo ser baseado na minimização do risco empírico (erro de classificação das instâncias de treinamento) para formar o modelo da rede, o que acaba gerando modelos sobre ajustados (*overfit*). Além disso, este algoritmo não lida bem com *outliers* (DENG; ZHENG; CHEN, 2009).

Com o objetivo de atuar nestas limitações do referido método, foi desenvolvido o algoritmo *Regularized Extreme Learning Machine* (RELM) por (DENG; ZHENG; CHEN, 2009). Diferentemente do algoritmo original, o RELM possui aprendizado baseado na ponderação entre a minimização do risco estrutural e do risco empírico.

Segundo a teoria de aprendizado estatístico, um modelo apresenta um bom nível de generalização quando seu treinamento apresenta uma ponderação entre a minimização do risco empírico e a do risco estrutural (DENG; ZHENG; CHEN, 2009). O primeiro está relacionado com o erro de classificação e o segundo com a maximização da margem de separação entre as classes.

Assim, a versão regularizada do método ELM introduz um fator de ponderação na minimização dos dois riscos supracitados, de modo a melhorar a capacidade de generalização do modelo gerado. Com isso, o parâmetro  $\gamma$  é inserido para ponderar o risco empírico (representado pela soma do erro quadrático) frente ao risco estrutural (representado pelo termo  $\|\beta\|^2$ ), e o algoritmo do RELM pode ser descrito a partir do seguinte problema de otimização:

$$\min \frac{1}{2} \|\beta\|^2 + \frac{1}{2} \gamma \|D\varepsilon\|^2 \quad (24)$$

Com as restrições:  $\sum_{i=1}^{\tilde{N}} \beta_i g(\mathbf{w}_i \cdot \mathbf{x}_j + b_i) - \mathbf{t}_j = \varepsilon_j, j = 1, \dots, N$

Adicionalmente, um fator extra de ponderação (representado por  $D$  na Equação (24)) pode ser inserido no vetor de erros de modo a minimizar a influência da presença de *outliers* no processo de aprendizado.

Na Equação (24)  $D$  é uma matriz diagonal cujos valores  $[v_1, v_2, \dots, v_N]$  são os pesos relacionados ao erro de cada instância de treinamento e  $\varepsilon = [\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N]$  são os erros associados a cada amostra de treinamento.

Aplicando a função Lagrangeana ao problema descrito na Equação (24) se obtém:

$$\min \frac{1}{2} \gamma \|D\varepsilon\|^2 + \frac{1}{2} \|\beta\|^2 - \sum_{j=1}^N \alpha_j \left( \sum_{i=1}^{\tilde{N}} \beta_i g(\mathbf{w}_i \cdot \mathbf{x}_j + b_i) - \mathbf{t}_j - \varepsilon_j \right) \quad (25)$$

Ou de forma compacta:

$$\min \frac{1}{2} \gamma \|D\varepsilon\|^2 + \frac{1}{2} \|\beta\|^2 - \alpha (\mathbf{H}\beta - \mathbf{T} - \varepsilon) \quad (26)$$

Onde  $\alpha_j$  são os multiplicadores de Lagrange.

Assim, tomando as derivadas parciais em relação a  $\alpha$ ,  $\beta$  e  $\varepsilon$  e igualando-as a zero chega-se à seguinte expressão para o vetor de pesos da camada de saída:

$$\beta = \left( \frac{I}{\gamma} + \mathbf{H}^T D^2 \mathbf{H} \right)^\dagger \mathbf{H}^T D^2 \mathbf{T} \quad (27)$$

Se a matriz de pesos do erro  $D$  for igual à identidade ( $I$ ), se obtém a versão desbalanceada do algoritmo e o vetor  $\beta$  passa a ser calculado a partir da Equação (28):

$$\beta = \left( \frac{I}{\gamma} + \mathbf{H}^T \mathbf{H} \right)^\dagger \mathbf{H}^T \mathbf{T} \quad (28)$$

Analisando a Equação (28) observa-se que o algoritmo original do ELM é um caso particular do RELM desbalanceado com  $\gamma$  tendendo a infinito.

Os pesos  $[v_1, v_2, \dots, v_N]$  relacionados aos erros e que compõem a diagonal da matriz  $D$ , são obtidos a partir da Equação (29):

$$v_j = \begin{cases} 1, & \left| \frac{\varepsilon_j}{\hat{s}} \right| \leq c_1 \\ \frac{c_2 - \left| \frac{\varepsilon_j}{\hat{s}} \right|}{c_2 - c_1}, & c_1 \left| \frac{\varepsilon_j}{\hat{s}} \right| \leq c_2 \\ 10^{-4}, & \text{caso contrário} \end{cases} \quad (29)$$

onde  $\hat{s}$  corresponde a uma estimativa robusta do desvio padrão dos erros da versão desbalanceada do RELM, e é obtido através da Equação (30):

$$\hat{s} = \frac{IQR}{2 \cdot 0,6745} \quad (30)$$

A variação interquartil (IQR - *Inter Quartile Range*) é a diferença entre o 75º e o 25º quartis do erro (DAVID, 1998). Uma vez que em uma distribuição Gaussiana existem muito poucas amostras acima de  $2,5\hat{s}$ , as constantes  $c_1$  e  $c_2$  são tipicamente ajustadas em 2,5 e 3 respectivamente (DENG; ZHENG; CHEN, 2009).

Assim, o algoritmo do RELM pode ser descrito através da seguinte sequência de passos:

1. atribuir valores aleatoriamente ao vetor de pesos dos neurônios da camada oculta ( $w$ ) e aos seus respectivos limiares ( $b$ );

2. a partir dos valores designados, calcular a matriz de saídas da camada oculta  $H$ ;
3. calcular os pesos de saída ( $\beta$ ) da versão desbalanceada do método através da Equação (28);
4. calcular o valor dos multiplicadores de Lagrange ( $\alpha$ ) e os erros ( $\varepsilon$ ) através das Equações (31) e (32) respectivamente:

$$\alpha = -\gamma(\mathbf{H}\beta - \mathbf{T})^T \quad (31)$$

$$\varepsilon_i = \frac{\alpha_i}{\gamma} \quad \text{para } i = 1, 2, \dots, N \quad (32)$$

5. calcular  $\hat{s}$  e  $v_j$  através das Equações (30) e (29);
6. determinar o valor dos pesos de saída ( $\beta$ ) por meio da Equação (27).

### 3 METODOLOGIA EXPERIMENTAL

O presente trabalho visa a aplicação de três métodos diferentes de seleção de características na tarefa de classificação de movimentos do segmento mão-braço a partir de sinais de eletromiografia. O algoritmo proposto divide-se em 4 etapas fundamentais: extração de características, análise da correlação entre elas, ranking das características e, por fim, uma etapa onde é feita a seleção final das características a partir dos resultados de classificação.

Dessa forma, o algoritmo proposto consiste na análise de dados oriundos de duas bases distintas: a do Ninapro (ATZORI *et al.*, 2014) e a obtida no Laboratório de Instrumentação Eletroeletrônica (IEE) da UFRGS. Na etapa seguinte, 13 características foram extraídas de modo a melhor representar os sinais de sEMG: 11 no domínio do tempo e duas no domínio da frequência. A correlação entre todas as 156 possibilidades obtidas (12 canais x 13 características) foi analisada e retirou-se do conjunto as que apresentaram correlação superior a 0,9. Na sequência, as que restaram da etapa anterior foram ranqueadas a partir de 3 métodos distintos: SVM-RFE, Monte Carlo e SVD *Entropy*. Por fim, o conjunto ordenado de características passou pelo estágio de classificação onde determinou-se o melhor grupo dentre elas. O algoritmo de classificação utilizado nesta etapa foi o RELM. O diagrama de blocos da Figura 13 mostra o passo a passo da metodologia utilizada.

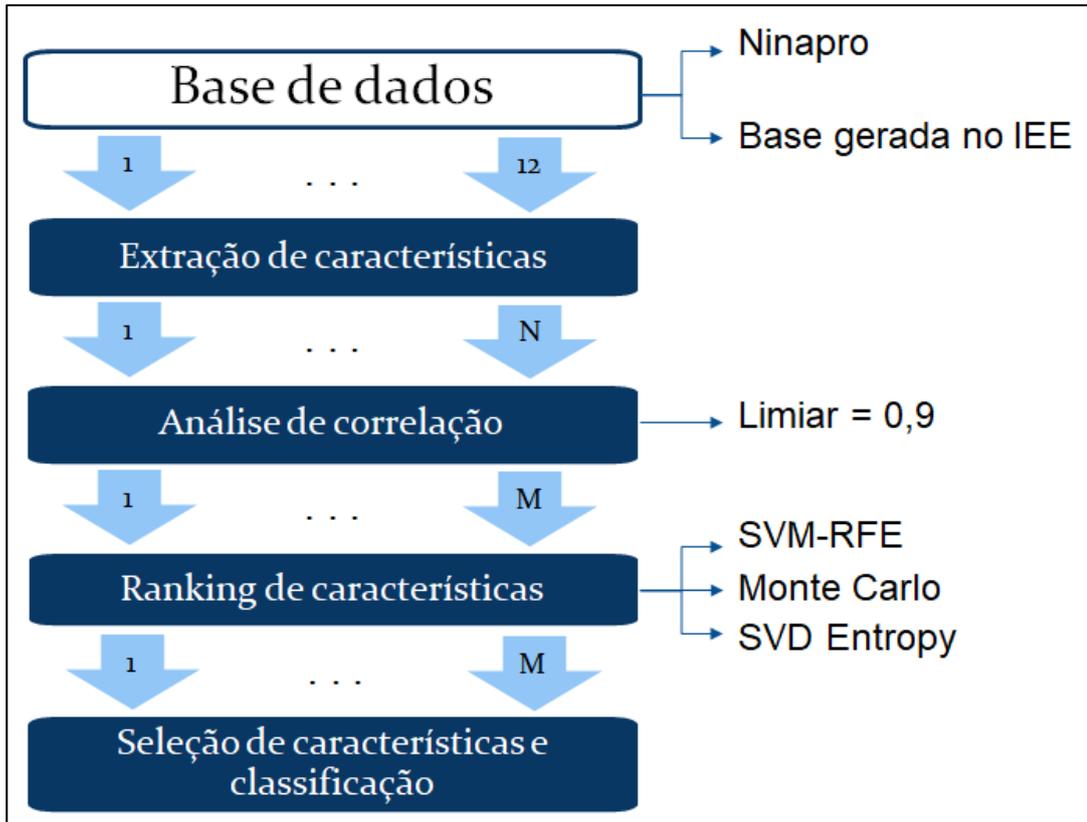


Figura 13 Diagrama de blocos da metodologia proposta.

### 3.1 DESCRIÇÃO DAS BASES DE DADOS

#### 3.1.1 Base de Dados Ninapro

O banco de dados Ninapro (*Non-Invasive Adaptive Hand Prosthetics*) é uma base de dados internacional disponibilizada. Essa base é resultado do trabalho de (ATZORI *et al.*, 2014) e a sua concessão tem por principal objetivo o de possibilitar que os resultados de pesquisas na área de classificação de sinais de eletromiografia pudessem ser comparados, visto que são raros os bancos de dados disponibilizados.

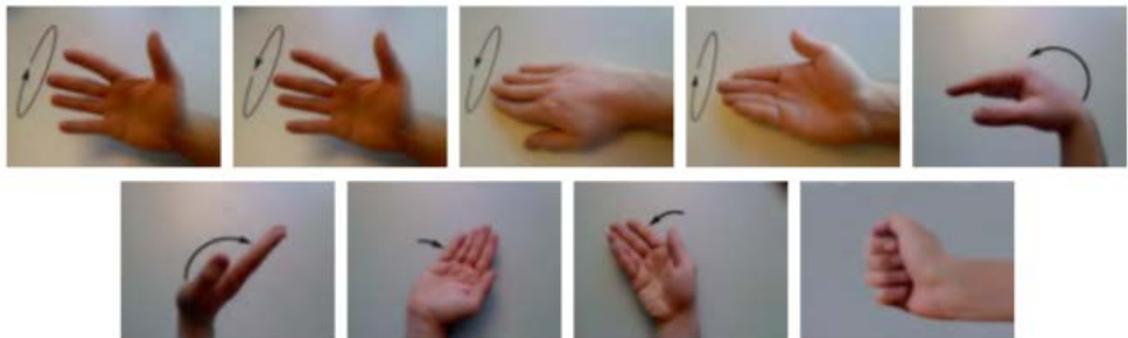
Atualmente ela é composta por 7 bases distintas contendo sinais de sEMG adquiridos em 6 exercícios diferentes de voluntários amputados e não amputados. Neste trabalho foram utilizadas as bases 2 e 3, da classe de exercícios B. Na primeira constam dados de ensaios realizados em 40 voluntários com membros intactos enquanto que na outra constam dados de

ensaios de 11 adultos com amputação transradial. A análise foi realizada nos 40 sujeitos da base 2 e em 10 sujeitos da base 3. Um indivíduo da base 3 foi excluído do experimento devido a qualidade do sinal de sEMG. Para tal sujeito conseguiu-se taxa de acerto média de 20,2%, significativamente inferior à média dos demais voluntários desta base (77,5%).

A classe de exercícios B é constituída por 17 movimentos: 8 configurações isométricas de mão (Figura 14) e 9 movimentos básicos de pulso (Figura 15). O voluntário executa cada um deles 6 vezes de modo sequencial. A sequência de movimentos é reproduzida em um *laptop* e o voluntário é instruído a fazer a replicação. A execução de cada movimento possui um tempo de duração de 5 segundos e são alternados com um período de repouso de 3 segundos.



**Figura 14** Configurações isométricas de mãos (adaptado de ATZORI *et al.*, 2014).



**Figura 15** Movimentos básicos de pulso (adaptado de ATZORI *et al.*, 2014).

O sinal de EMG foi adquirido através de 12 eletrodos diferenciais de superfície. Foram utilizados eletrodos *Trigno Wireless* do fabricante Delsys. O sinal foi adquirido a uma taxa de 2kHz e conta com um nível de ruído máximo de 750nVrms. Os eletrodos foram fixados no braço e antebraço do voluntário. Dois deles foram alocados no braço, na região de máxima

atividade do bíceps e do tríceps. Os 10 eletrodos restantes foram colocados no antebraço da seguinte forma: dois na região de máxima atividade do músculo *flexor digitorum* e do músculo *extensor digitorum*, e oito igualmente espaçados em torno do antebraço (ATZORI *et al.*, 2014).

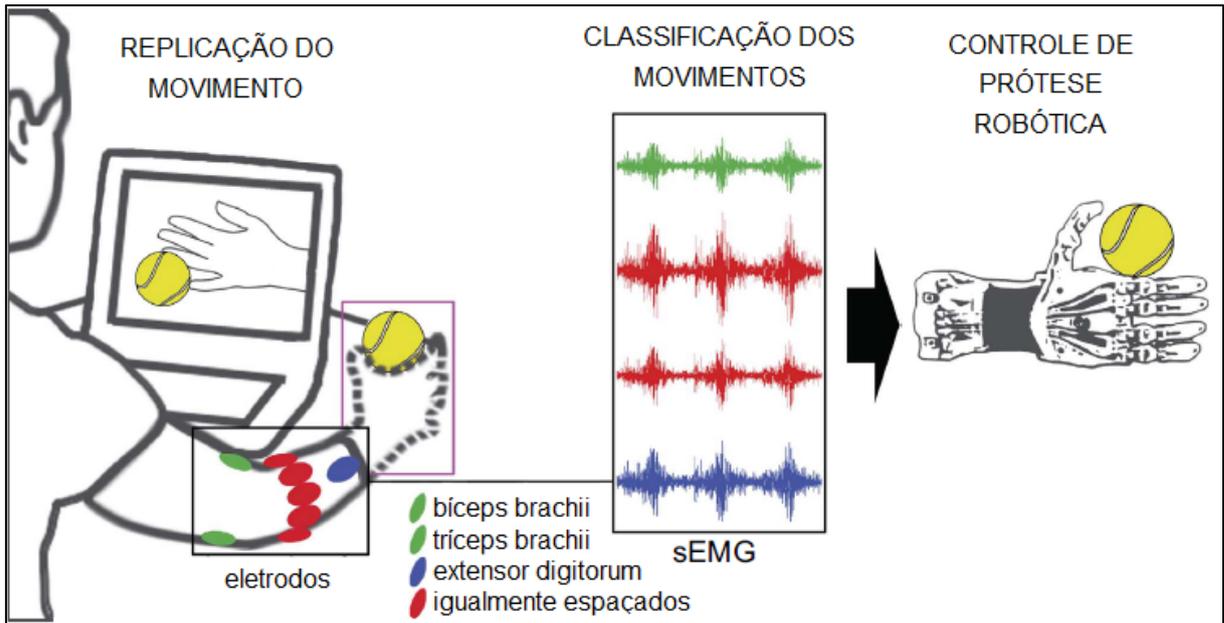
O sinal de sEMG disponibilizado nesta base conta com a identificação dos tempos de início e de término de cada movimento baseado nos tempos de execução dos vídeos. Porém, devido ao tempo de reação do voluntário que gera atrasos no início e término da execução do movimento, essa identificação acaba contendo muitas amostras nomeadas incorretamente, principalmente nas regiões próximas às extremidades do sinal do movimento. Dessa forma, para corrigir tal desvio, foi realizado, pela equipe do Ninapro, o reajuste dos tempos de início e término de cada movimento baseado na máxima verossimilhança entre a sequência de dados correspondente a repouso-movimento-repouso.

Resumidamente o procedimento de aquisição dos sinais de sEMG da Base Ninapro é dado pela seguinte sequência de passos: os movimentos são reproduzidos na tela de um computador; o voluntário, sentado confortavelmente em uma cadeira, executa tais movimentos; eletrodos de superfície fixados no braço e antebraço obtêm o sinal de EMG; o sinal é transmitido a um *laptop* onde é feito processamento e o treinamento de um classificador; o classificador pode ser testado através do estímulo de um braço robótico. A Figura 16 ilustra o passo a passo do sistema proposto pela equipe Ninapro.

Para mais informações a respeito da Base Ninapro, consultar, como sugestão, as seguintes referências:

- ATZORI, M. *et al.* *Building the Ninapro database: A resource for the biorobotics community. Proceedings of the IEEE RAS and EMBS International Conference on Biomedical Robotics and Biomechatronics*, p. 1258–1265, 2012.

- ATZORI, M. *et al.* *A Benchmark Database for Myoelectric Movement Classification. Transactions on Neural Systems and Rehabilitation Engineering*, 2013.
- ATZORI, M. *et al.* *Electromyography data for non-invasive naturally-controlled robotic hand prostheses. Scientific Data*, 2014.



**Figura 16** Procedimento de aquisição (adaptado de ATZORI *et al.*, 2014).

### 3.1.2 Base de Dados do IEE

Com o intuito de criar uma base de dados alternativa a do Ninapro, aquisições de sinais de eletromiografia foram realizadas no laboratório do IEE. Assim, se obtém total controle do protocolo de aquisição, bem como conhecimento dos equipamentos utilizados, perfis dos voluntários, etc. Além de tal motivação, a nova base tem por objetivo o de obter dados de ensaios distintos ao do Ninapro, porém com os mesmos movimentos do exercício B, descritos na seção 3.1.1. Assim, 4 tipos diferentes de ensaios foram propostos aos voluntários recrutados para a formação da base:

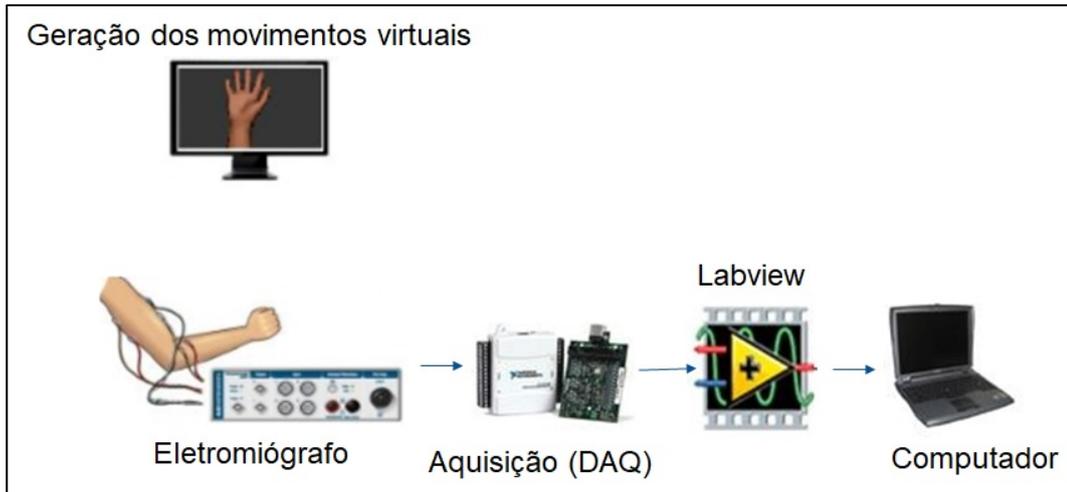
- Ensaio A: execução dos 17 movimentos de modo sequencial e com 6 repetições de cada um deles (mesmo procedimento da Base Ninapro com a única diferença estando nos tempos de repouso e execução de cada movimento);
- Ensaio B: execução dos 17 movimentos de modo sequencial e com 10 repetições de cada um deles;
- Ensaio C: execução dos 17 movimentos de modo aleatorizado e com 6 repetições de cada um deles;
- Ensaio D: execução dos 17 movimentos de modo aleatorizado e com 10 repetições de cada um deles.

A proposição do Ensaio A tem por objetivo a comparação direta das bases e a verificação da influência da aquisição nos resultados de classificação. Atualmente a base do IEE conta com dados de sinais adquiridos de 4 voluntários, os quais foram analisados no presente trabalho.

O procedimento de aquisição do sinal de sEMG ocorreu a partir da sequência de passos:

1. o movimento a ser executado pelo voluntário é reproduzido através de um vídeo em uma tela de computador;
2. o voluntário, sentado confortavelmente em uma cadeira com apoio para os braços, executa o movimento mostrado durante o período de execução do vídeo;
3. eletrodos fixados no braço e antebraço do sujeito fazem a obtenção do sinal de sEMG;
4. os sinais obtidos pelos eletrodos são transmitidos via cabo até dois eletromiógrafos, onde o sinal é filtrado e amplificado;
5. na próxima etapa o sinal é digitalizado e transmitido até um computador;
6. no computador, o sinal é lido e gravado através do software LabView 2009.

A Figura 17 ilustra o sistema descrito anteriormente.



**Figura 17 Sistema experimental proposto para a geração da Base de Dados do IEE.**

### 3.1.2.1 Estímulo Visual dos Movimentos

Para a apresentação do movimento ao voluntário, foi realizada a mesma abordagem utilizada nos trabalhos de (FAVIEIRO, 2012) e (CENE, 2016). Um modelo virtual foi desenvolvido para a representação de cada movimento. Para isso, foram utilizados os softwares MakeHuman 1.0.2 e o Blender 2.75a.

Através do primeiro é possível gerar modelos humanos (humanoides) a partir de parâmetros de entrada tais quais, altura, sexo e peso. Já o segundo consiste em um software de criação de aplicações 3D, utilizado para o desenvolvimento de animações, jogos, e tarefas do gênero.

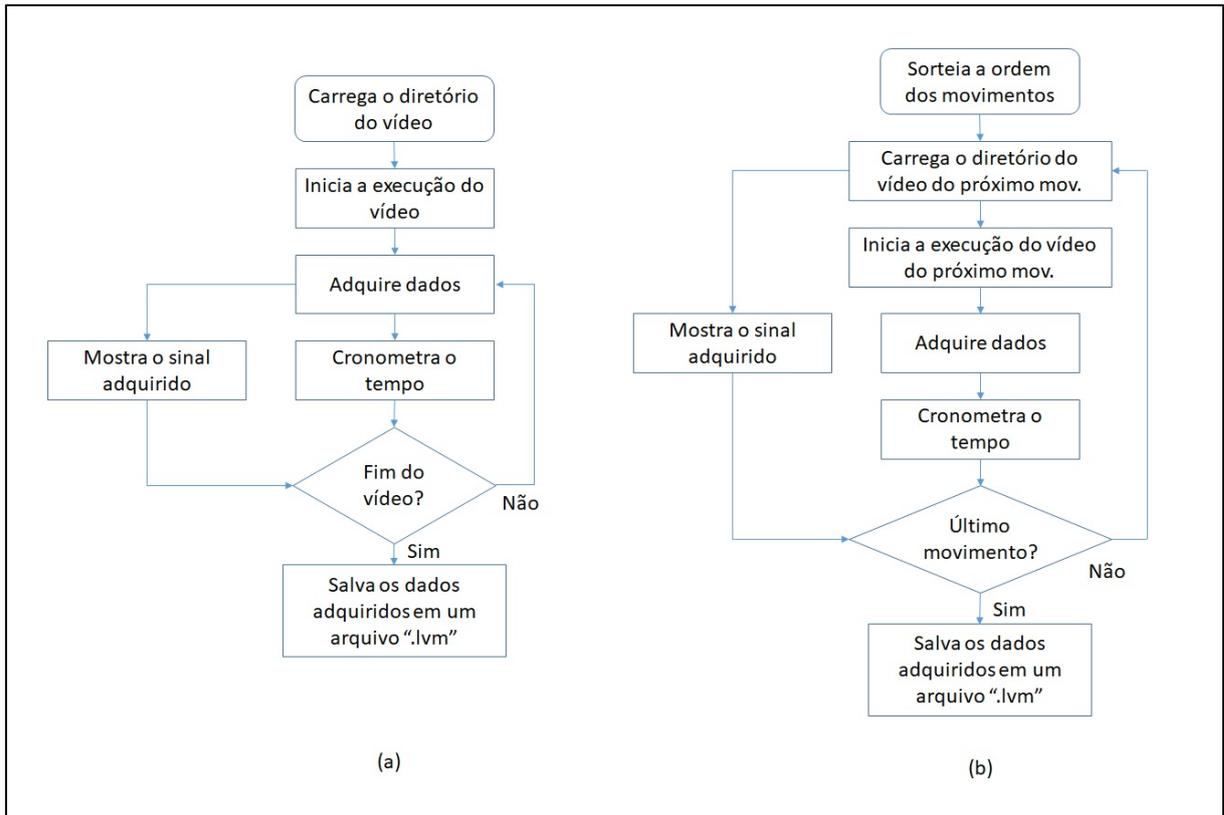
Por conseguinte, o *software* MakeHuman foi utilizado para a criação do modelo da mão humana, o qual foi exportado para um arquivo .dae que pode ser reconhecido pelo Blender. Um esqueleto é, então, gerado a partir deste modelo, cujas articulações manipuláveis são utilizadas pelo Blender para a determinação das posições que a mão deve assumir. Estas, por sua vez, são utilizadas na criação da animação do movimento. Para cada movimento foi gerado um arquivo .avi o qual pode ser lido pelo Windows Media Player.

O desenvolvimento do modelo da mão e das animações de cada um dos movimentos foi realizado como parte do trabalho de doutorado da aluna (MOURA, 2016 - "Proposta de um Modelo Adaptativo Baseado em Sensores Virtuais e TVARMA na Caracterização de Sinais Mioelétricos do Segmento Mão-Braço”).

Para a seleção do vídeo a ser executado bem como a instrução da execução, foi utilizado o *software* LabView 2009. Deste modo, duas rotinas foram desenvolvidas de acordo com o tipo de ensaio a ser realizado: uma para ensaio sequencial e outra para ensaio com os movimentos aleatorizados. Além da execução dos vídeos, as rotinas do LabView também realizam a leitura dos sinais de sEMG adquiridos e a gravação dos mesmos em arquivos .lvm. Os sinais eram mostrados na tela do computador ao voluntário e ao operador do ensaio, de modo que fosse realizado o acompanhamento da qualidade do sinal adquirido e verificação de ocorrências tais como deslocamento do eletrodo.

Durante o ensaio, dois vídeos eram mostrados ao voluntário. O principal, ocupando a maior parte da tela do monitor, indica o movimento a ser replicado pelo voluntário. Um vídeo adicional é mostrado em uma tela secundária, no canto superior esquerdo do monitor. Ele tem por objetivo o de mostrar para o voluntário, durante o período de repouso que precede o movimento, a próxima ação que será executada na tela principal. A inclusão do vídeo secundário foi necessária em virtude dos ensaios aleatorizados, de modo a dar a possibilidade da pessoa se preparar adequadamente para a próxima realização de movimento.

Assim, a única diferença entre as duas rotinas do LabView criadas está na parte de seleção e execução dos vídeos dos movimentos, já que, para os ensaios sequenciais, o exercício inteiro está contido em único vídeo enquanto que nos aleatorizados o vídeo de cada movimento é executado independentemente. A Figura 18 mostra o fluxograma de cada uma das rotinas desenvolvidas.



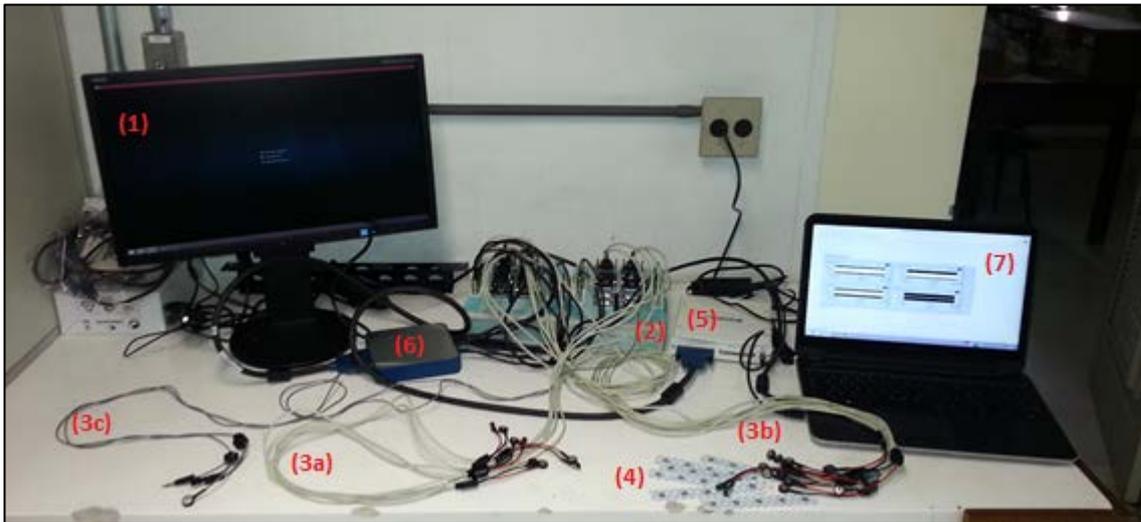
**Figura 18 Fluxograma do sistema de aquisição: (a) para os ensaios sequenciais; (b) para os ensaios aleatorizados.**

### 3.1.2.2 Equipamentos e Materiais Utilizados

Para a obtenção dos sinais de sEMG, foram utilizados 12 eletrodos de superfície bipolares passivos do fabricante SOLIDOR, modelo MSGST-06. Os eletrodos foram conectados a dois eletromiógrafos de 8 canais ligados em paralelo. Assim, 8 eletrodos foram conectados a um deles e 4 ligados ao outro. Foram escolhidos eletromiógrafos do fabricante EMG System do Brasil, modelo EMG 830 C para as aquisições.

O eletromiógrafo realiza filtragem e amplificação do sinal de sEMG adquirido, uma vez que a banda de interesse de um sinal de sEMG está contida na faixa de 15 a 500Hz e sua amplitude varia de -5 a 5mV. Assim, um filtro passa faixa de quarta ordem na banda de interesse e uma amplificação de aproximadamente 2000 vezes são aplicados ao sinal, o que possibilitou a aquisição de sinais com relação sinal ruído de 37 dB.

Para a digitalização e transferência dos dados ao computador foi utilizado uma placa da National Instruments NI USB-6289 (NATIONAL INSTRUMENTS, 2014). Ela contém 32 entradas analógicas e permite uma conversão A/D com 18 bits de resolução a um limite de 625 kS/s. No presente trabalho realizou-se a aquisição a uma taxa de amostragem de 2kHz, similar à utilizada na base do Ninapro. A Figura 19 apresenta uma foto do arranjo experimental utilizado para a geração da Base de Dados do IEE.



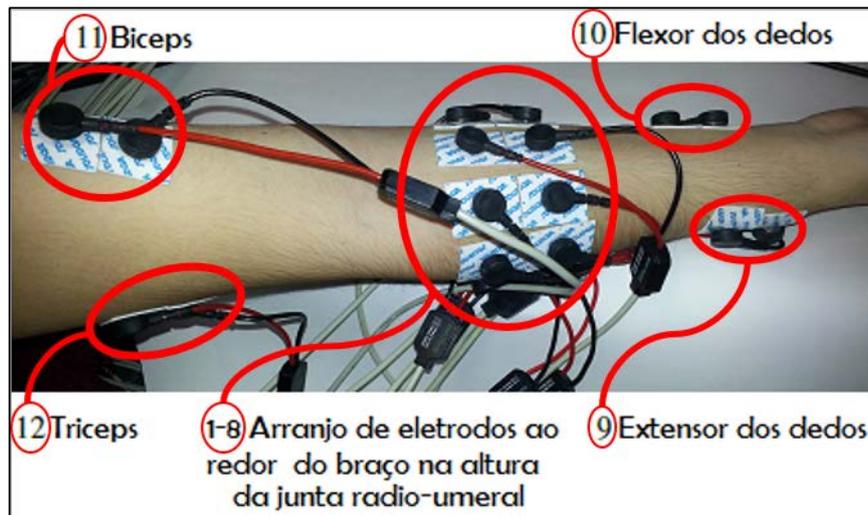
**Figura 19** Arranjo experimental utilizado nas aquisições na geração da Base de Dados do IEE: (1) corresponde ao monitor onde os movimentos eram mostrados, (2) eletromiógrafos, (3a-c) cabos do eletromiógrafo, (4) eletrodos de superfície, (5) placa de aquisição, (6) borneira, (7) computador (adaptado de CENE, 2016).

### 3.1.2.3 Procedimentos do Processo de Aquisição

Como em qualquer experimento envolvendo seres humanos, a primeira etapa do estudo consiste na apresentação e esclarecimento do Termo de Consentimento Livre e Esclarecido aos voluntários participantes. Deve constar no termo todos os detalhes do experimento a ser realizado, bem como a finalidade e os riscos aos quais o voluntário está submetido, em caso de consentimento à participação do estudo. Junto ao termo é anexado também um breve formulário onde constam dados do voluntário e características do mesmo que são relevantes ao estudo foco do ensaio. Ambos, o termo de consentimento e o formulário, estão expostos nos anexos A e B

respectivamente. O presente trabalho faz parte do projeto aprovado pelo Comitê de Ética CAAE: 11253312.8.0000.5347.

O posicionamento dos eletrodos no braço e antebraço dos voluntários seguiu o padrão adotado no trabalho da Base de Dados Ninapro<sup>1</sup>. Quatro deles foram fixados na região de máxima atividade dos músculos: *flexor digitorum*, *extensor digitorum*, bíceps e tríceps. Os demais foram alocados igualmente espaçados em torno do antebraço. Para a identificação da região de máxima atividade dos referidos músculos, foi utilizado o método palmar. Adicionalmente, um par de eletrodos foram fixados na testa do voluntário para atuarem como referência para os eletromiógrafos. A Figura 20 apresenta uma foto deste posicionamento.



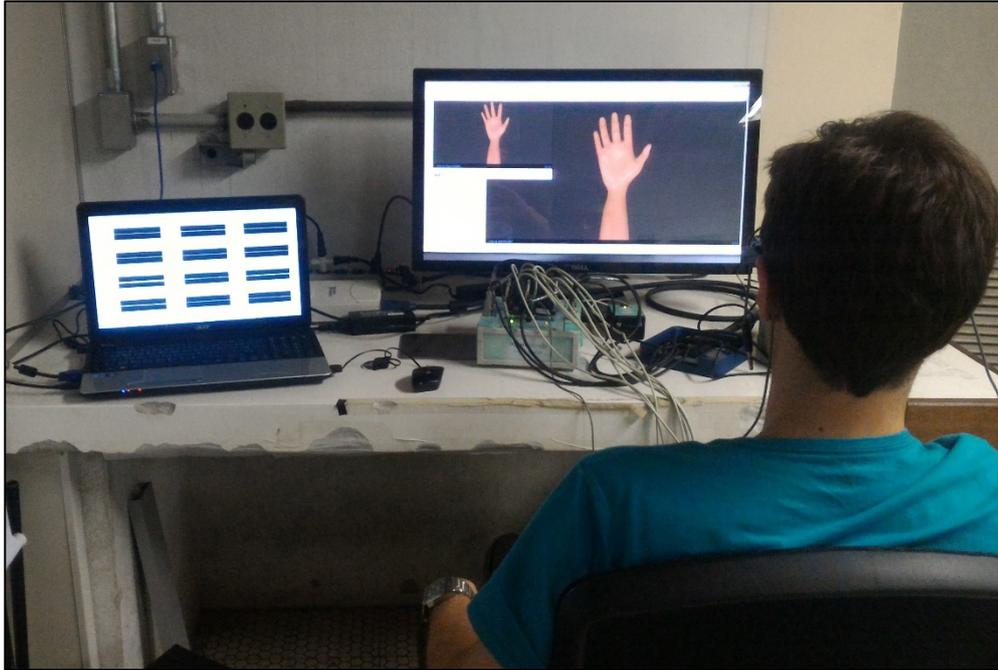
**Figura 20** Posicionamento dos eletrodos (adaptado de CENE, 2016).

<sup>1</sup> É importante ressaltar que o procedimento de posicionamento dos eletrodos não é a melhor forma para obtenção de sinais de sEMG em amputados devido a possibilidade da não existência de partes de regiões do segmento mão-braço (dependendo do grau de amputação). Neste tipo de situação seria importante a seleção de outro grupo muscular, por exemplo, no trapézio para a obtenção de determinados canais de sEMG o que minimizaria os problemas em determinados canais na aquisição dos sinais de sEMG no grupo de amputados da base NINAPRO e por consequência a melhora na correspondente classificação. Na Base de Dados do IEE com amputados pretende-se realizar o posicionamento de duas formas, ou seja, mantendo o procedimento de posicionamento da Base de Dados Ninapro e um novo procedimento proposto na geração dessa base. Este tipo de discussão não faz parte da proposta desta dissertação, mas é importante a mesma devido ao uso de dados de sEMG da referida Base Ninapro.

De modo a minimizar a influência de ruídos relacionados à fixação dos eletrodos na pele e obter um sinal com a melhor qualidade possível, o procedimento para a fixação dos sensores seguiu as instruções descritas no trabalho de (FAVIEIRO, 2012):

- raspagem dos pelos na região de colagem dos eletrodos;
- limpeza da pele;
- fixação dos 4 eletrodos responsáveis pela obtenção dos sinais de sEMG dos músculos *flexor digitorum*, *extensor digitorum*, bíceps e tríceps na região de máxima atividade dos referidos músculos;
- colocação dos eletrodos de modo a não interferir no movimento executado pelo voluntário;
- fixação dos cabos para que os mesmos não se movimentem durante o ensaio, evitando que os eletrodos se desloquem.

Antes da execução do ensaio, os movimentos são mostrados ao voluntário e o mesmo é instruído a repetir cada um deles por 3 vezes de modo a se familiarizar com o exercício. Cada movimento possui 5 segundos de duração alternados com períodos de repouso de 9 segundos, aproximadamente. O repouso é dividido em 3 partes: os primeiros 2 segundos o voluntário descansa; nos 5 subsequentes é mostrado o próximo movimento a ser executado (na tela secundária); nos últimos 2 o voluntário se prepara para a realização do próximo movimento. Durante o repouso, uma contagem regressiva é mostrada na tela principal indicando o tempo que falta até a execução do próximo movimento. Isso é feito com o objetivo de facilitar a preparação do voluntário. A Figura 21 ilustra um dos ensaios realizados.



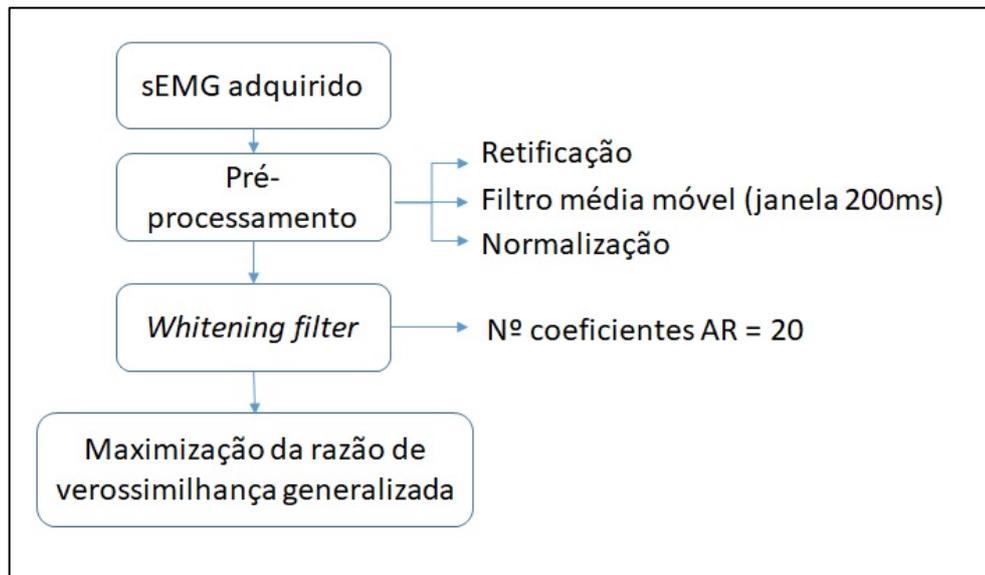
**Figura 21** Ensaio realizado por um dos voluntários.

Ao todo os voluntários executaram cada um dos 4 tipos de ensaios (A, B, C e D) 3 vezes, totalizando 12 ensaios por indivíduo. Cada voluntário executou no máximo 3 ensaios por dia de modo a evitar cansaço e possível fadiga muscular. Portanto, a Base do IEE possui 48 sinais de sEMG oriundos de 4 tipos de ensaios distintos realizados por 4 voluntários com 3 repetições.

#### **3.1.2.4 Ajuste dos Tempos de Início e Término de Cada Movimento**

Logo após a realização da aquisição, a identificação dos segmentos de movimento e de repouso ao longo de cada ensaio é dada pelos tempos de início e término da execução do vídeo de cada movimento. Porém, devido ao tempo de reação do voluntário, essa identificação acaba contendo muitas amostras nomeadas incorretamente, principalmente nas regiões próximas às transições entre sinais, ou seja, as transições entre um determinado movimento e repouso e entre repouso e início de um dado movimento.

Para realizar a correção nos tempos de início e término de cada movimento automaticamente, foi desenvolvido um algoritmo baseado no utilizado na Base de Dados Ninapro. A implementação foi realizada no software Matlab R2015a. O diagrama de blocos com as principais etapas do algoritmo desenvolvido está resumido na Figura 22.



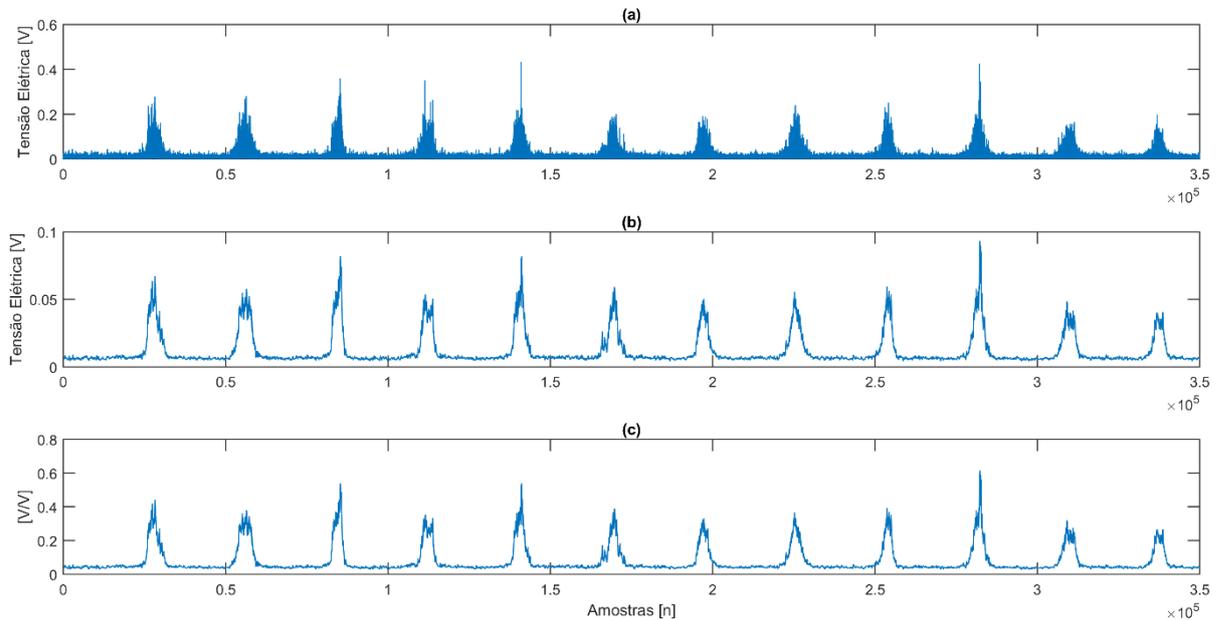
**Figura 22 Diagrama de blocos do algoritmo de ajuste dos tempos de início e término de cada movimento.**

Primeiramente o sinal passa por uma etapa de pré-processamento constituída por:

- retificação: consiste na determinação do valor absoluto do sinal de modo a torná-lo positivo. A Figura 23 (a) mostra parte do sinal de um dos ensaios após retificação (operação similar a realizada pelo retificador de onda completa implementado por componentes eletrônicos);
- filtro média móvel: filtro aplicado em janelas de 100 ms (200 amostras) com o objetivo de suavizar a forma de onda do sinal. A Figura 23 (b) mostra o mesmo segmento de sinal da Figura 23 (a) após a filtragem;
- normalização do sinal: essa etapa foi realizada a partir da Equação (33), de modo a representar o sinal dentro de uma faixa compreendida entre 0 e 1. A Figura 23 (c) mostra o mesmo segmento de sinal da Figura 23 (b) após a normalização.

$$x_k = \frac{x_k}{\max(\mathbf{x})} \quad (33)$$

onde  $x_k$  é a amostra  $k$  do sinal de sEMG  $\mathbf{x}$  de um dado canal.



**Figura 23** Exemplo de sinal de sEMG pré-processado: (a) retificado, (b) retificado e filtrado e (c) retificado, filtrado e normalizado.

Na etapa seguinte, o sinal foi dividido em janelas limitadas pelos tempos de início e término da execução dos vídeos. As análises subsequentes foram executadas somente nestes segmentos do sinal. A Figura 24 ilustra uma parte do sinal com duas janelas representadas.

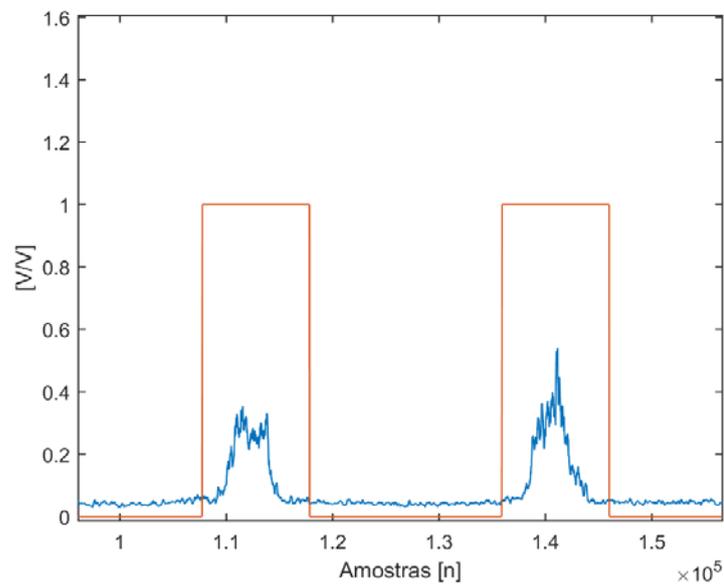
Seguindo as instruções de (STAUDE; WOLF, 1999), aplicou-se um *whitening filter* em cada janela do sinal para remover componentes autorregressivas que são irrelevantes para detecção de variações nos parâmetros estatísticos no referido sinal. De acordo com (STAUDE; WOLF, 1999), o sinal de sEMG pode ser representado em função de seus componentes autorregressivos por meio da Equação (34).

$$x_k = - \sum_{i=1}^p a_i x_{k-i} + w_k \quad (34)$$

onde  $a_i$  são os coeficientes autorregressivos que representam o sinal de sEMG  $x$ , e  $w$  é uma variável aleatória dada pela Equação (35):

$$w_k = u_k + n_k \quad (35)$$

onde  $u$  corresponde à contribuição normal das unidades motoras englobadas pelo eletrodo durante uma contração muscular, e  $n$  representa as ativações espontâneas das unidades motoras, mesmo na ausência de uma contração. Este último componente representa o ruído intrínseco do sistema (STAUDE; WOLF, 1999).

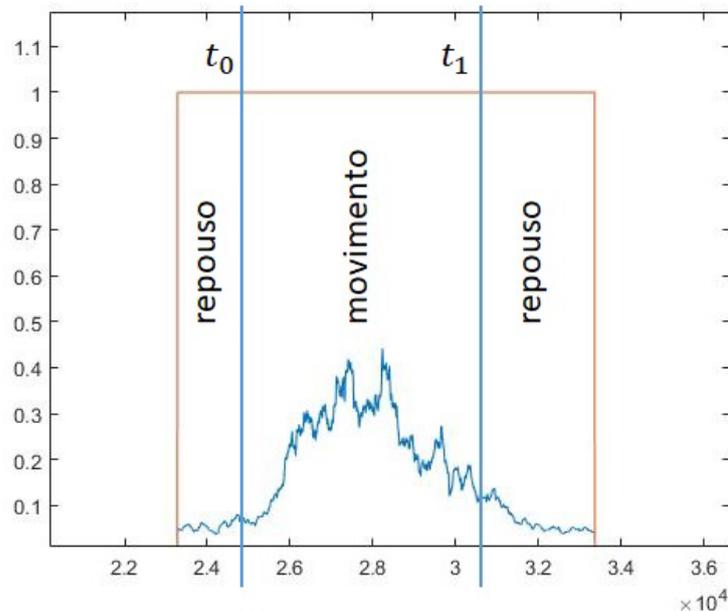


**Figura 24 Exemplo da segmentação efetuada. Em vermelho está indicado os períodos de execução dos vídeos de movimento (nível 1) que limitam tais janelas.**

Ainda segundo (STAUDE; WOLF, 1999), a informação presente no sinal de sEMG que varia de forma mais significativa na presença de uma alteração do sinal (presença de um evento, como por exemplo, mudança de estado de repouso para contração do músculo) está contida na variância do ruído ( $w$ ). As componentes autorregressivas ( $a_i$ ) não são, portanto, relevantes para a detecção de mudanças no sinal de sEMG.

Assim, no presente trabalho utilizou-se uma representação a partir de apenas 20 componentes autorregressivas do sinal, a qual mostrou ser suficiente. Tal modelo é igual ao utilizado pela equipe Ninapro (ATZORI *et al.*, 2013).

Por fim, a determinação dos tempos de início e término de cada movimento foi realizada a partir da maximização da razão da verossimilhança generalizada (GLR – *Generalized Likelihood Ratio*) entre o modelo estatístico do sinal de repouso e o do sinal de movimento. Para tanto, as janelas referidas anteriormente foram divididas em 3 partes (repouso, movimento, repouso) conforme ilustrado na Figura 25. Portanto, a tarefa consiste na busca pelos tempos  $t_0$  e  $t_1$  que fazem o ajuste mais adequado desta divisão, ou seja, a partir da maximização da soma das probabilidades de cada amostra fazer parte do modelo estatístico que melhor representa o segmento de sinal de repouso e de movimento.



**Figura 25 Segmento do sinal dividido em 3 partes: repouso, movimento, repouso.**

Para a obtenção do modelo estatístico que melhor representa os segmentos de sinal correspondentes a repouso e movimento, utilizou-se o conceito de máxima verossimilhança. Tal conceito tem por objetivo o de avaliar quão bem um dado conjunto de amostras é

representado por uma distribuição de probabilidades com específicos parâmetros estatísticos. Em outras palavras, este conceito é utilizado para a estimação das características estatísticas da distribuição de probabilidades de uma população a partir de um conjunto limitado de amostras.

Segundo (MYUNG, 2003), para um dado vetor de parâmetros estatísticos  $w$  e um conjunto de amostras  $y$ , a função de verossimilhança é dada pela Equação (36):

$$\mathcal{L}(w|y) = f(y|w) \quad (36)$$

onde  $f(y|w)$  é a função densidade de probabilidade de  $y$  cujos parâmetros são dados pelo vetor  $w$ , e  $\mathcal{L}$  é o valor da verossimilhança calculada sobre o conjunto de amostras  $y$  em função do vetor de parâmetros  $w$ .

Considerando que o vetor  $y$  é formado por  $N$  amostras estatisticamente independentes, a Equação (36) pode ser reescrita a partir da Equação (37):

$$\mathcal{L}(w|y_1, \dots, y_N) = f(y_1, \dots, y_N|w) = \prod_{i=1}^N f(y_i|w) \quad (37)$$

Contudo, a estimação da máxima verossimilhança resume-se na busca pelo vetor de parâmetros  $w$  que maximize a Equação (37).

Voltando ao problema do ajuste dos tempos  $t_0$  e  $t_1$ , de acordo com (ATZORI *et al.*, 2013) a maximização da razão da verossimilhança generalizada entre as partes correspondentes a repouso, movimento e repouso, consiste na maximização da expressão dada pela Equação (38):

$$\left[ \sum_{i=1}^{t_0-1} \ln(f(y_i|w_0)) + \sum_{j=t_0}^{t_1-1} \ln(f(y_j|w_1)) + \sum_{k=t_1}^T \ln(f(y_k|w_0)) \right] \quad (38)$$

onde  $f(y_i|w)$  é a probabilidade da amostra  $y_i$  fazer parte de uma distribuição de probabilidades de parâmetros estatísticos  $w$ , sendo  $w$  obtido por máxima verossimilhança, considerando que os três segmentos da janela fazem parte de uma distribuição Gaussiana. Os índices 0 e 1

representam segmento de repouso e movimento respectivamente. O símbolo  $T$  é o número de amostras da janela.

Assim, o processo de ajuste dos tempos  $t_0$  e  $t_1$  pela maximização do GLR se deu a partir de busca exaustiva por meio da combinação de todos os valores possíveis de  $t_0$  e  $t_1$  dentro da janela. Essa busca ocorreu respeitando as condições dadas pelas Equações (39) a (41):

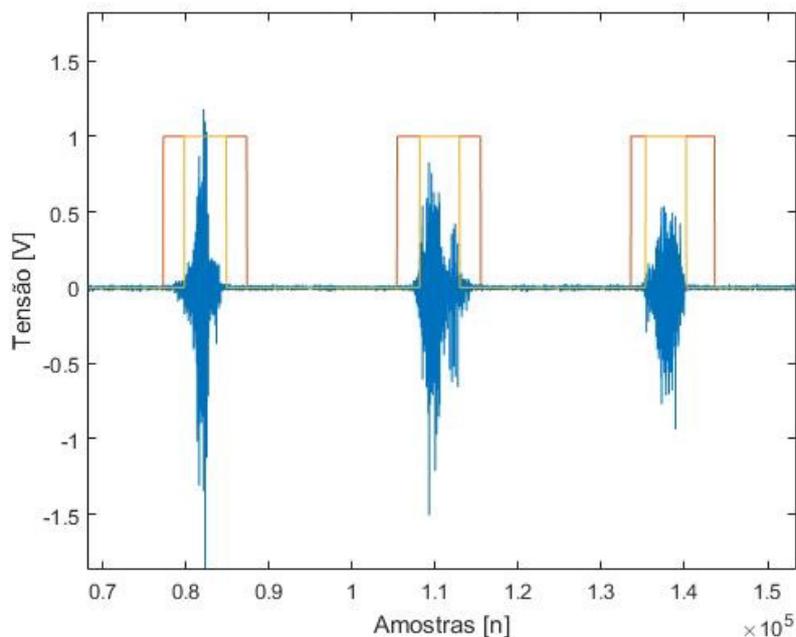
$$t_1 - t_0 \geq 0,3T \quad (39)$$

$$t_0 \geq 0,1f_a = 100 \text{ ms} \quad (40)$$

$$S_1 \geq S_0 \quad (41)$$

onde  $S_1$  é a variância das amostras do segmento associado ao movimento,  $S_0$  é a variância do repouso e  $f_a$  é a frequência de amostragem.

A Figura 26, como exemplo, mostra um segmento do sinal de sEMG adquirido (azul) com a segmentação baseada nos tempos de execução dos vídeos (vermelho) e a obtida após o ajuste (laranja). A Figura 27 apresenta o fluxograma da rotina desenvolvida no Matlab R2015a.



**Figura 26** Parte do sinal de sEMG com a segmentação antes e após o ajuste.

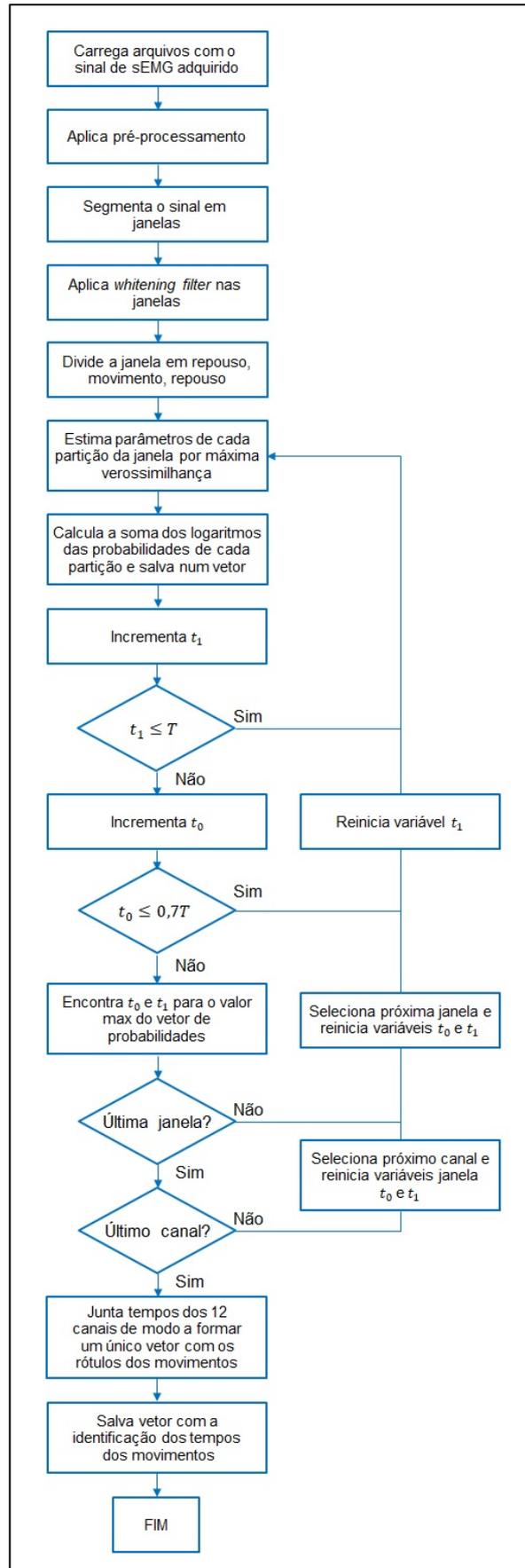
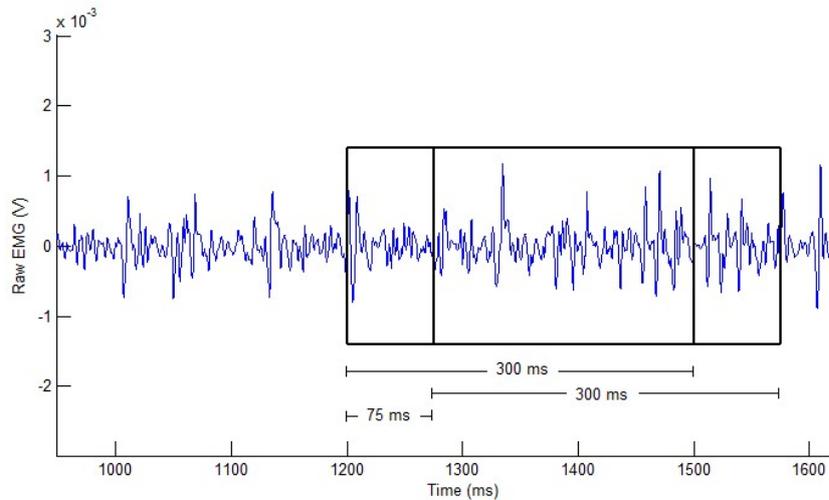


Figura 27 Fluxograma da rotina desenvolvida para o ajuste dos rótulos das amostras.

### 3.2 SEGMENTAÇÃO E A CORRESPONDENTE EXTRAÇÃO DE CARACTERÍSTICAS

As características foram extraídas em janelas de 300 ms com deslocamentos de 75 ms, seguindo as instruções de (RIILLO *et al.*, 2014). A Figura 28 ilustra a segmentação realizada para esta etapa do trabalho.

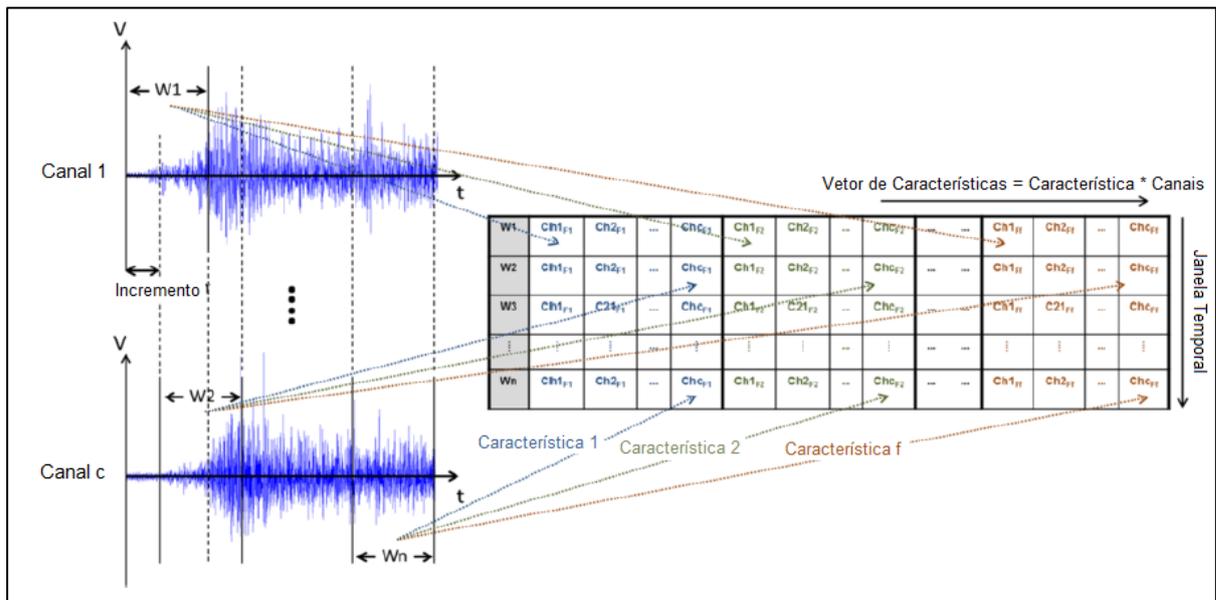


**Figura 28 Segmentação realizada para a extração de características (adaptado de TOSIN *et al.*, 2017).**

Em cada uma das janelas supracitadas foram obtidas 13 características nos domínios do tempo e frequência. Para tanto foram utilizados os códigos abertos disponíveis na plataforma BioPatRec. A referida plataforma foi desenvolvida no trabalho de (ORTIZ-CATALAN; BRÅNEMARK; HÅKANSSON, 2013) e tem por objetivo o de disponibilizar uma plataforma de pesquisa para o desenvolvimento e avaliação de algoritmos para controle de próteses. Ela foi desenvolvida no Matlab e é constituída por módulos independentes. Tais módulos contêm algoritmos nas seguintes áreas: processamento de sinais de sEMG, extração e seleção de características, classificação, e aplicações em tempo real (ORTIZ-CATALAN; BRÅNEMARK; HÅKANSSON, 2013).

No presente trabalho utilizou-se apenas o módulo que trata de extração de características. As características extraídas foram armazenadas em matrizes onde as colunas

representam as características e as linhas são as janelas temporais, sendo a primeira coluna preenchida com o vetor contendo o rótulo de identificação de cada janela (número do movimento ou repouso). A Figura 29 ilustra o processo de extração realizado. As matrizes contendo as características extraídas foram salvas em arquivos .xlsx.



**Figura 29 Sistema de extração de características implementado (adaptado de ORTIZ-CATALAN; BRÅNEMARK; HÅKANSSON, 2013).**

A seguir estão descritas todas as características extraídas no domínio do tempo. Em todas as equações,  $w$  representa o número de amostras da janela e  $x_i$  é a  $i$ -ésima amostra do sinal de sEMG  $x$ .

**Root Mean Square (Raiz Média Quadrática - RMS):** apresenta informação em relação à amplitude da contração executada. É calculada a partir da Equação (42):

$$RMS = \sqrt{\frac{1}{w} \sum_{i=1}^w x_i^2} \quad (42)$$

**Média da Diferença Absoluta (MDA):** corresponde ao valor esperado da diferença absoluta entre duas variáveis aleatórias. Considerando uma amostra aleatória constituída por  $w$

instâncias, tal característica pode ser calculada a partir da média aritmética das diferenças entre todas as combinações das instâncias formadoras da amostra (ver Equação (43)):

$$MDA = \frac{1}{w^2} \sum_{i=1}^w \sum_{j=1}^w |x_i - x_j| \quad (43)$$

**Dimensão Fractal (DF):** representa o nível de irregularidade de um fractal (formas que não podem ser explicadas pela geometria Euclidiana). Diferentemente da dimensão Euclidiana, a fractal pode ser descrita por números fracionários. Isso se deve ao fato dela descrever o nível de ocupação pela forma e não pelo espaço onde o objeto está inserido (BACKES; BRUNO, 2005). No contexto do presente trabalho, leva-se em conta a forma de onda do sinal de sEMG. A Equação (44) mostra como ele foi computado.

$$DF = \frac{\ln w}{\left(\ln w + \frac{d}{L}\right)} \quad (44)$$

onde  $d$  é a máxima diferença absoluta entre duas amostras subsequentes formadoras da janela, e  $L$  é a soma de tais diferenças.

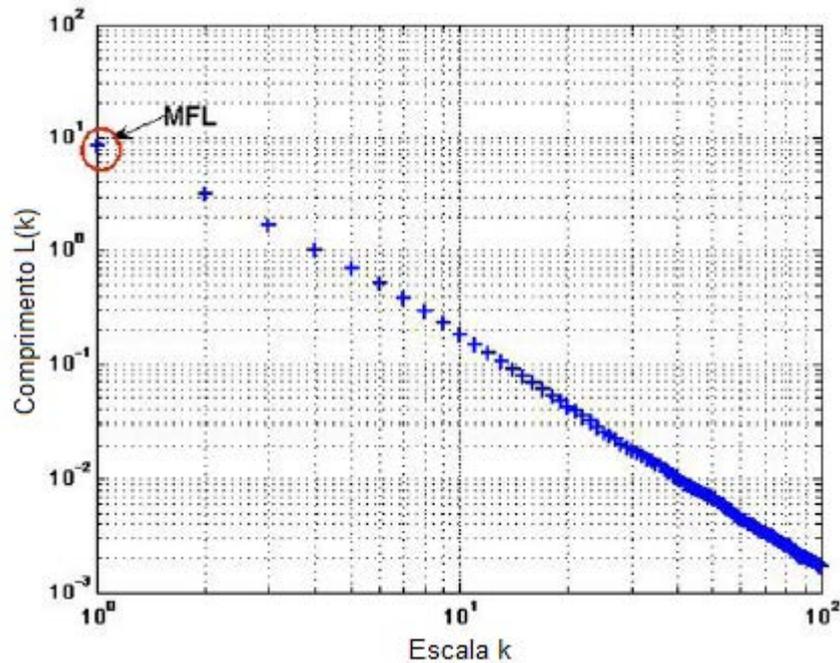
**Dimensão Fractal de Higuchi (DFH):** representa o cálculo da dimensão fractal pelo método de Higuchi. Segundo (ARJUNAN; KUMAR, 2010), para a realização do cálculo primeiramente é realizada a determinação do comprimento da curva para uma dada escala  $k$  conforme Equação (45):

$$L_m(k) = \frac{\left\{ \left( \sum_{i=1}^{\lfloor \frac{w-m}{k} \rfloor} |x(m+ik) - x(m+(i-1)k)| \right) \frac{w-1}{\lfloor \frac{w-m}{k} \rfloor k} \right\}}{k} \quad (45)$$

onde  $m$  é o instante inicial da janela,  $k$  é o intervalo de tempo (escala) para o qual se está calculando o comprimento da curva, e o termo  $\frac{w-1}{\lfloor \frac{w-m}{k} \rfloor k}$  representa um termo normalizador.

O comprimento da curva para um dado intervalo  $k$  é então o valor médio sobre as  $k$  amostras do intervalo. A curva será considerada fractal de dimensão  $D$  se  $\langle L(k) \propto k^{-D} \rangle$

(ARJUNAN; KUMAR, 2010). A Figura 30 ilustra um exemplo de gráfico do  $\log L(k)$  em função de  $\log k$ .



**Figura 30 Gráfico do comprimento da curva em função da escala ( $k$ ) (adaptado de ARJUNAN; KUMAR, 2010).**

Deste modo, a dimensão fractal de Higuchi corresponde à inclinação da reta do gráfico da Figura 30. Em outras palavras, é dado a partir da Equação (46):

$$DFH = \frac{d \log L(k)}{d \log(k)} \quad (46)$$

**Máxima Dimensão Fractal (MDF):** de acordo com a definição da dimensão fractal de Higuchi, determina-se a Máxima Dimensão Fractal como sendo o valor médio do comprimento da curva na menor escala ( $k$ ). Tal valor está indicado na Figura 30 e pode ser calculado a partir da Equação (47):

$$MDF = L(1) \quad (47)$$

**Valor Médio Absoluto (VMA):** representa o valor médio absoluto da amplitude do sinal de sEMG na janela analisada (Equação (48)):

$$VMA = \frac{1}{w} \sum_{i=1}^w |x_i| \quad (48)$$

**Potência:** indica a potência do sinal de sEMG na janela analisada (Equação (49)):

$$P = \frac{\sum_{i=1}^w x_i^2}{w} \quad (49)$$

**Mudanças de Inclinação do Sinal (MIS):** indica o número de vezes que o sinal muda sua inclinação. É uma forma simplificada de se representar a frequência do sinal. É dado pela Equação (50):

$$MIS = \frac{1}{w} \sum_{i=2}^{w-1} f[(x_i - x_{i-1}) \cdot (x_i - x_{i+1})] \quad (50)$$

onde  $f$  é uma função representada pela Equação (51):

$$f(x) = \begin{cases} 1, & x \geq \text{limiar} \\ 0, & \text{caso contrário} \end{cases} \quad (51)$$

No presente trabalho utilizou-se a média absoluta do sinal como limiar.

**Desvio Padrão (DP):** representa a dispersão do sinal em torno da média e é determinado pela da Equação (52):

$$DP = \sqrt{\frac{1}{w-1} \sum_{i=1}^w (x_i - \bar{x})^2} \quad (52)$$

onde  $\bar{x}$  é a média do sinal de sEMG na janela analisada.

**Comprimento da Forma de Onda (CFO):** caracteriza o comprimento cumulativo da forma de onda do sinal determinado pela Equação (53):

$$CFO = \sum_{i=1}^{w-1} |x_{i+1} - x_i| \quad (53)$$

**Cruzamentos por Zero (CZ):** corresponde ao número de vezes que o sinal muda de sinal. Tal característica também está relacionado à frequência do sinal. Ele pode ser calculado pela Equação (54):

$$CZ = \sum_{i=1}^{w-1} ((x_i > 0) \wedge (x_{i+1} < 0)) \vee ((x_i < 0) \wedge (x_{i+1} > 0)) \quad (54)$$

Respeitando a condição imposta pela Equação (55):

$$|x_i - x_{i+1}| \geq \text{limiar} \quad (55)$$

No presente trabalho utilizou-se a média absoluta do sinal como limiar.

No domínio da frequência foram extraídas as características.

**Frequência Média (FM):** indica a frequência média da banda em que o sinal está inserido. É calculada através do somatório da multiplicação da frequência pela potência espectral, dividido pela soma total da potência espectral da janela analisada (ver Equação (56)):

$$FM = \frac{\sum_{i=1}^M f_i P_i}{\sum_{i=1}^M P_i} \quad (56)$$

onde  $f_i$  é a frequência correspondente ao  $i$ -ésimo termo do espectro de frequências da janela do sinal de sEMG,  $P_i$  é a potência relacionada à frequência  $f_i$ , e  $M$  é o tamanho do vetor da resposta em frequência da janela.

**Comprimento da Forma de Onda da Resposta em Frequência (CFORF):** caracteriza o comprimento cumulativo da forma de onda da resposta em frequência do sinal. Esta métrica foi obtida pela Equação (57):

$$CFORF = \sum_{i=1}^{M-1} |F_{i+1} - F_i| \quad (57)$$

onde  $F_i$  indica a amplitude da resposta em frequência do sinal associada à frequência  $f_i$ .

### 3.3 MÉTRICA DE CORRELAÇÃO

A correlação entre características é utilizada como um meio de se reduzir a dimensionalidade de um banco de dados. Tal método de seleção é do tipo baseado em filtro e é não supervisionado uma vez que ele avalia apenas a interação entre as características e, portanto, não faz uso de um classificador para executar tal tarefa.

A correlação entre dois vetores pode ser calculada através da Equação (58):

$$r_{ij} = \frac{\sum_{k=1}^n (c_{ik} - \bar{c}_i)(c_{jk} - \bar{c}_j)}{\sqrt{\sum_{k=1}^n (c_{ik} - \bar{c}_i)^2} \cdot \sqrt{\sum_{k=1}^n (c_{jk} - \bar{c}_j)^2}} \quad (58)$$

onde,  $c_{ik}$  corresponde ao  $k$ -ésimo termo da característica  $c_i$ ,  $c_{jk}$  corresponde ao  $k$ -ésimo termo da característica  $c_j$ ,  $\bar{c}$  é a média do vetor de características  $c$ .

Tal indicador estatístico pode resultar em valores compreendidos no intervalo [-1, 1]. Os extremos deste intervalo indicam correlação máxima, ou seja, máxima similaridade entre dois vetores. Valores negativos representam relação inversa entre dois vetores de características, ao passo que valores positivos indicam uma relação direta.

Seguindo as instruções de (YIN *et al.*, 2016a), introduziu-se no presente trabalho uma etapa a qual consiste na análise da correlação entre as características da base de dados. Tal estágio foi aplicado antes da execução do ranking das características de modo a filtrar aquelas redundantes e irrelevantes para o processo de classificação. Segundo (YIN *et al.*, 2016a), algoritmos como SVM-RFE não lidam bem com a presença de características redundantes sendo, contudo, recomendável a eliminação destas antes da aplicação de tais algoritmos.

Baseando-se no trabalho de (YIN *et al.*, 2016a), estipulou-se um nível de correlação igual a 0,9 como limiar para a determinação de redundância entre duas características. Deste modo, aquelas que apresentaram correlação maior do que 0,9 foram eliminadas da base de dados e não seguiram nas etapas subsequentes.

Esta etapa do trabalho foi implementada através do software R-3.3.2. A análise da correlação foi realizada por meio da biblioteca Caret do referido software, a qual disponibiliza funções para o cálculo da correlação e remoção das características que apresentarem valor acima de determinado limiar.

### 3.4 RANKING DAS CARACTERÍSTICAS

#### 3.4.1 Algoritmo SVM-RFE

O algoritmo SVM-RFE foi implementado através do software R-3.3.2. Para o treinamento do classificador SVM utilizou-se o código aberto desenvolvido em C++ da biblioteca LIBSVM (CHANG; LIN, 2011), presente no pacote e1071 do referido software.

Originalmente, o método SVM foi desenvolvido para a classificação de problemas binários, porém o mesmo pode ser estendido para tarefas de múltiplas classes através da expansão da tarefa em múltiplos classificadores binários. Um dos métodos mais utilizados para se executar tal decomposição é o chamado *one versus one* (OVO). Seu funcionamento é baseado no treinamento de um classificador para cada par de classes. Assim, para uma determinada instância da base de dados, será atribuída a classe à qual tiver o maior número de votos dentre todos os classificadores. Por conseguinte, para um problema de  $k$  classes, ele realiza o treinamento de  $k(k - 1)/2$  classificadores (HUANG *et al.*, 2016).

O supracitado método de decomposição é o utilizado pela biblioteca LIBSVM para o treinamento de SVM's em problemas de múltiplas classes. A determinação do valor de importância associado a cada característica foi realizada da mesma forma do que em (TOSIN *et al.*, 2017), semelhante à executada em problemas binários, dada pela Equação (9). A única diferença está no fato do peso considerado referente a uma dada característica ser a média dos pesos ajustados a ela em cada um dos classificadores. Assim, o valor de importância é dado através da Equação (59):

$$c_i = \frac{1}{k(k - 1)/2} \sum_{j=1}^{k(k-1)/2} w_{ij} \quad (59)$$

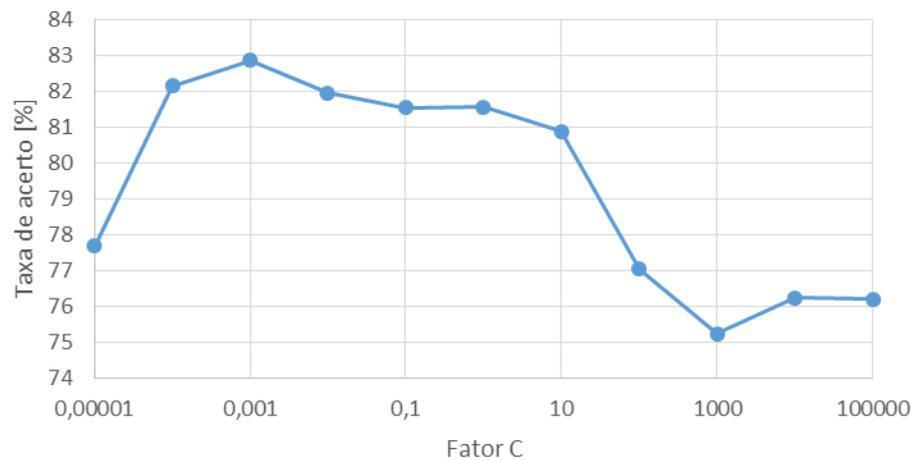
O algoritmo do SVM-RFE foi implementado em sua configuração original conforme (GUYON *et al.*, 2002), sendo, desde modo, escolhido o kernel linear para o treinamento das

SVM's. A estrutura do algoritmo implementado segue o fluxograma representado na Figura 10 apenas com a devida adaptação no que tange a determinação do valor de importância atribuído a cada característica (Equação (59)).

Para o ajuste do termo regulador ( $C$ ) do classificador foi realizado o treinamento de uma série de SVM's com o valor de tal fator variando de  $10^{-5}$  a  $10^5$  seguindo uma progressão geométrica de razão igual a 10. A base de dados foi dividida inicialmente em 70% para o treinamento das SVM's e o restante foi separado para o teste do classificador. Durante a fase de treinamento, implementou-se validação cruzada em cada um dos experimentos executados. Para tanto as instâncias foram divididas em 5 partes iguais sendo 4 delas utilizadas no treinamento do classificador e 1 separada para teste. Esse procedimento foi repetido 5 vezes de modo a todas as partições serem utilizadas para testar o classificador. Por fim, o classificador obtido após a etapa de validação foi submetido às instâncias de teste (30% separados inicialmente), e a taxa de acerto obtida foi utilizada como critério para seleção do melhor valor do fator  $C$ .

Devido ao elevado tempo de processamento despendido na análise supracitada e o número de ensaios analisados no presente trabalho (98 no total), optou-se por realizar o ajuste em apenas um ensaio de cada base. Assim, três ensaios foram escolhidos aleatoriamente: um dentre os 40 voluntários não amputados da Base 2 Ninapro, um da Base 3 Ninapro, e um da Base de Dados do IEE. O valor ajustado para o parâmetro  $C$  foi, então, generalizado para os demais ensaios de cada uma das referidas bases.

Deste modo, selecionou-se os seguintes ensaios: voluntário 3 da Base 2 Ninapro, voluntário 4 da Base 3 Ninapro, e voluntário 3, exercício C, repetição 2 da Base de Dados do IEE. A Figura 31 ilustra a variação da taxa de acerto do classificador em função do parâmetro  $C$  (escala logarítmica) para o voluntário selecionado da Base 2 Ninapro.



**Figura 31** Evolução da taxa de acerto do classificador em função do fator  $C$ .

Como conclui-se da Figura 31, o classificador obteve melhor resultado para fator  $C$  igual a 0,001. Para os demais experimentos realizados observou-se o mesmo comportamento. Portanto, utilizou-se em todos os ensaios fator  $C$  igual a 0,001. Ao término do processamento desta etapa, a lista com o ranking das características de cada ensaio foi salva em um arquivo .xlsx.

### 3.4.2 Monte Carlo *Feature Selection*

O método de seleção de características Monte Carlo foi implementado através do software R-3.3.2. Utilizou-se a biblioteca `rmcfs` presente no referido software a qual disponibiliza uma função que executa o método a partir de uma série de parâmetros de entrada fornecidos pelo usuário.

De acordo com a seção 2.2.1.2, tais parâmetros consistem em: tamanho dos subconjuntos de características ( $m$ ), número de subconjuntos formados ( $s$ ), número de divisões em treinamento e teste realizados em cada subconjunto ( $t$ ), e os fatores  $u$  e  $v$ . Estes últimos são fixados em 1 na biblioteca utilizada, uma vez que, de acordo com os experimentos realizados

em (DRAMÍŃSKI *et al.*, 2008), a variação de tais parâmetros não altera significativamente os rankings obtidos.

Para o ajuste dos demais parâmetros ( $m$ ,  $s$  e  $t$ ), procedeu-se conforme instruções de (DRAMÍŃSKI *et al.*, 2008). Para tanto, fixou-se valores para  $m$  e  $t$  e variou-se  $s$  de 1 a 100 em intervalos de 5. Porém, na iteração seguinte ao processamento com  $s$  igual a 1, utilizou-se  $s$  igual a 5 e a partir de então passou-se a incrementá-lo em 5 em cada iteração. Segundo (DRAMÍŃSKI *et al.*, 2008), a escolha do valor de  $m$  deve ser realizada levando-se em consideração o equilíbrio entre evitar que características relevantes sejam severamente mascaradas pelas que apresentarem maior valor de importância relativa e o fato de que o valor de  $s$  não seja demasiadamente elevado. Quanto maior for o valor de  $m$ , maior a chance de uma característica informativa ser mascarada pelas mais importantes de acordo com o critério de importância relativa. Em contrapartida, quanto menor for o valor de tal parâmetro, maior valor de  $s$  será exigido, uma vez que todas as características devem ter probabilidade elevada de serem selecionadas em múltiplos subconjuntos de características (DRAMÍŃSKI *et al.*, 2008). Deste modo, através de testes preliminares ajustou-se  $m$  em 16 e  $t$  igual a 5.

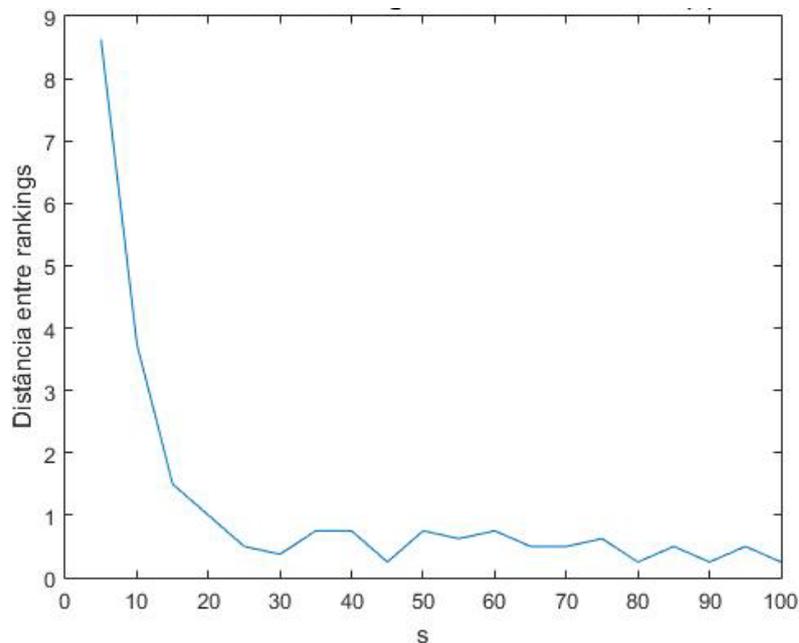
Na verificação da alteração dos rankings obtidos com a evolução do parâmetro  $s$ , foi utilizado o conceito de distância entre rankings subsequentes. Segundo (DRAMÍŃSKI *et al.*, 2008), tal distância pode ser calculada através da Equação (60):

$$Dist(s, s - 5) = \frac{1}{d_p} \sum_{g_k} |rank(g_k, s) - rank(g_k, s - 5)| \quad (60)$$

onde, o cálculo é executado somente em  $p\%$  das características as quais foram mais bem-ordenadas utilizando  $s - 5$  subconjuntos,  $rank(g_k, r)$  é o ordenamento da característica  $g_k$  para um número  $r$  de subconjuntos e  $d_p$  é o número de características que estão sendo consideradas no cálculo. Segundo (DRAMÍŃSKI *et al.*, 2008) é aconselhável que a escolha de

$p$  seja feita entre 5 e 20. No presente trabalho decidiu-se por usar  $p$  igual a 10 com base em ensaios preliminares executados com as bases de dados.

Pelas mesmas razões explicitadas na seção 3.4.1, optou-se por realizar o ajuste em apenas um ensaio de cada base. Deste modo, selecionou-se os seguintes ensaios: voluntário 1 da Base 2 Ninapro, voluntário 2 da Base 3 Ninapro, e voluntário 4, exercício B, repetição 3 da Base de Dados do IEE. A Figura 32 ilustra a variação da distância entre rankings em função do parâmetro  $s$  para o voluntário selecionado da Base 2 Ninapro.

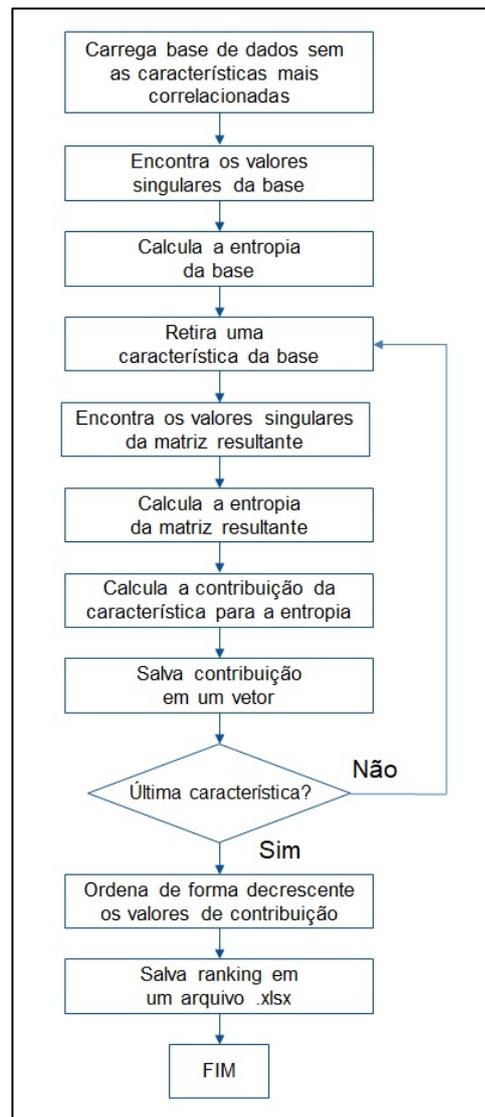


**Figura 32** Evolução da distância entre rankings em função do número de subconjuntos ( $s$ ).

Como conclui-se da Figura 32, a distância entre rankings diminui rapidamente com a evolução de  $s$  e fica praticamente estável a partir de  $s$  igual a 25. Para os demais experimentos realizados observou-se comportamento semelhante. Por conseguinte, utilizou-se em todos os ensaios parâmetro  $s$  igual a 30. Ao término do processamento desta etapa, a lista com o ranking das características de cada ensaio foi salva em um arquivo .xlsx.

### 3.4.3 SVD *Entropy*

O algoritmo do método de seleção de características *SVD Entropy* foi desenvolvido no Matlab R2015a. No presente trabalho, tal método foi utilizado para a realização de um ranking das características de acordo com a contribuição de cada uma delas com a entropia dos valores singulares da matriz de dados. Quanto maior a contribuição maior a relevância para o processo de classificação. O fluxograma do código desenvolvido está representado na Figura 33. Ao término do processamento desta etapa, a lista com o ranking das características de cada ensaio foi salva em um arquivo .xlsx.



**Figura 33 Fluxograma do método *SVD Entropy*.**

## 3.5 CLASSIFICAÇÃO E SELEÇÃO DE CARACTERÍSTICAS

### 3.5.1 Classificador *Regularized Extreme Learning Machine*

Para a etapa de classificação utilizou-se o método RELM em sua versão desbalanceada, pelo seu baixo tempo de processamento e boa capacidade de generalização, além de já ter apresentado bons resultados em outros trabalhos envolvendo classificação de movimentos, como, por exemplo, no trabalho de (CENE; FAVIEIRO; BALBINOT, 2017). A implementação do algoritmo foi realizada por meio de código aberto em linguagem do Matlab disponibilizada na *toolbox* desenvolvida em (AKUSOK *et al.*, 2015). A realização da referida *toolbox* baseou-se no trabalho de (HUANG *et al.*, 2012), o qual introduz o conceito de *kernels* à aplicação em ELM e generaliza as determinações de (DENG; ZHENG; CHEN, 2009) para qualquer função de ativação dos neurônios da camada oculta (a versão regularizada de (DENG; ZHENG; CHEN, 2009) considera apenas a função sigmoide). Todavia, no presente trabalho utilizou-se a versão regularizada do método ELM segundo (DENG; ZHENG; CHEN, 2009), ou seja, foi escolhida a função sigmoide para a ativação dos neurônios da camada oculta.

O código utilizado de (AKUSOK *et al.*, 2015) fornece uma função que executa o método RELM a partir de uma série de entradas fornecidas pelo usuário. Tais parâmetros consistem em: dados de treinamento e teste (separados em duas matrizes), tipo da função de ativação dos neurônios da camada oculta, número de neurônios da camada oculta, indicador da tarefa de aprendizagem de máquina a ser executada (classificação ou regressão) e fator de regularização ( $\gamma$ ).

Antes de executar a etapa de classificação, as características extraídas dos sinais de sEMG, são normalizadas, uma vez que cada uma delas possui faixas distintas, seguindo a orientação de (AKUSOK *et al.*, 2015). A normalização foi realizada em cada uma das

características dividindo-se os valores da coluna da respectiva característica pelo seu maior valor (Equação (61)):

$$\mathbf{x}_{k_{norm}} = \frac{\mathbf{x}_k}{\max(\mathbf{x}_k)} \quad (61)$$

onde  $\mathbf{x}_k$  é o vetor de características da  $k$ -ésima coluna da matriz de dados.

O código utilizado para a implementação do classificador oferece uma série de funções de ativação para os neurônios da camada oculta. Como já mencionado, no presente trabalho optou-se por utilizar a função sigmoide, dada pela Equação (62):

$$f(y) = \frac{1}{1 + e^{-\lambda y}} \quad (62)$$

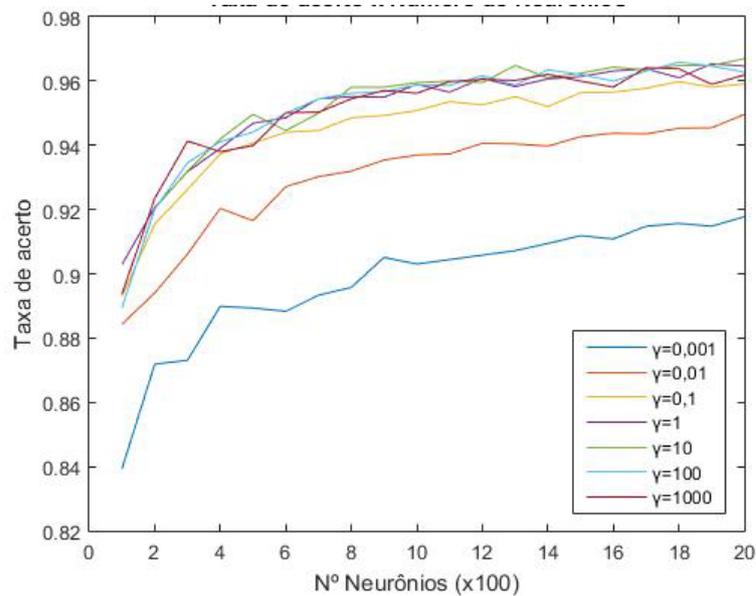
Deste modo, para um valor de  $\lambda$  suficientemente elevado, garante-se que a saída dos neurônios da camada oculta seja binário (0 ou 1).

Por conseguinte, a camada de saída da rede neural foi formada por um neurônio para cada classe a ser predita. Assim, se uma dada instância de treinamento é pertencente a uma classe  $i$ , o  $i$ -ésimo neurônio da camada de saída terá valor 1 e os demais 0.

O ajuste do número de neurônios ( $\tilde{N}$ ) e do fator regularizador  $\gamma$  foi realizado a partir de uma série de classificações variando-se ambos os termos. Num primeiro momento, testou-se o classificador com  $\tilde{N}$  variando de 100 a 2000 em uma progressão aritmética de razão 100, e  $\gamma$  variando de  $10^{-3}$  até  $10^3$  em uma progressão geométrica de razão igual a 10. O experimento foi realizado dividindo-se o conjunto de dados em 70% para treinamento e o restante para teste. A escolha dos melhores valores para tais parâmetros se deu através da avaliação da taxa de acerto observada ao submeter a rede neural obtida às instâncias do conjunto de teste. O experimento foi realizado em todos os ensaios das três bases utilizadas.

A Figura 34 ilustra o gráfico da variação da taxa de acerto em função do número de neurônios da camada oculta ( $\tilde{N}$ ) em um dos ensaios da Base de Dados do IEE. Nele estão sobrepostas as curvas da referida relação para todos os valores do fator  $\gamma$  ensaiados. Observa-

se que o valor da taxa de acerto estabiliza a partir de um número de neurônios igual a 1200. Comportamento parecido foi observado em todos os ensaios analisados das três bases. Deste modo, fixou-se tal parâmetro em 1200 para todas as bases.



**Figura 34** Evolução da taxa de acerto do classificador em função do número de neurônios.

Através dos resultados da primeira etapa verificou-se que as maiores taxas de acerto foram obtidas para valores de  $\gamma$  superiores a 1. Assim, decidiu-se por realizar mais uma etapa de testes do classificador de modo a fazer um ajuste fino de tal parâmetro. O mesmo experimento da primeira etapa foi executado, só que desta vez com  $\tilde{N}$  fixado em 1200 e  $\gamma$  variando de  $2^0$  a  $2^9$  em uma progressão geométrica de razão 2. A Figura 35 apresenta o gráfico da variação da taxa de acerto em função do número do fator ( $\gamma$ ) no ensaio de um voluntário amputado da base de dados do Ninapro.

Para os dados da base adquirida experimentalmente não foi observada variação da taxa de acerto frente a evolução do fator  $\gamma$ . Nesse caso optou-se por utilizar valor igual a 2 para tal parâmetro em todos os ensaios da base do IEE. As tabelas 1 e 2 resumem os valores de  $\tilde{N}$  e  $\gamma$  ajustados para as demais bases. A Figura 36 representa, resumidamente, o fluxograma do referido algoritmo.

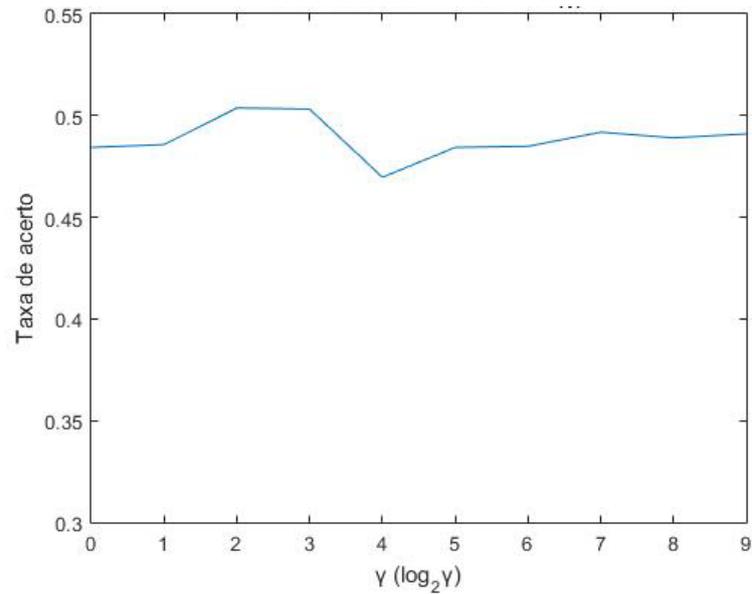


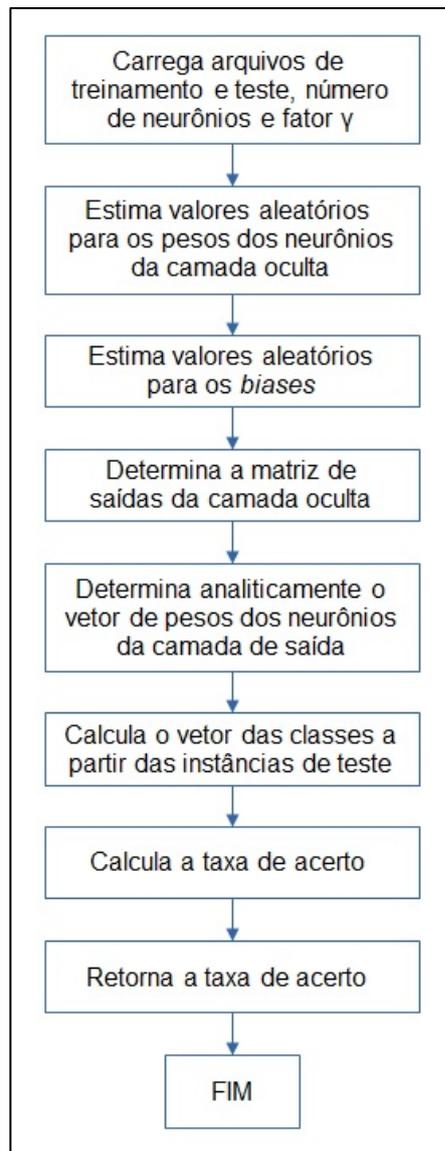
Figura 35 Evolução da taxa de acerto do classificador em função do fator  $\gamma$ .

Tabela 1 Parâmetros ajustados para os ensaios da Base 2 Ninapro.

Sujeito	Neurônios ( $\tilde{N}$ )	Fator ( $\gamma$ )	Sujeito	Neurônios ( $\tilde{N}$ )	Fator ( $\gamma$ )
1	1200	2	21	1200	8
2	1200	2	22	1200	8
3	1200	2	23	1200	4
4	1200	4	24	1200	1
5	1200	2	25	1200	2
6	1200	4	26	1200	2
7	1200	32	27	1200	1
8	1200	2	28	1200	32
9	1200	2	29	1200	8
10	1200	2	30	1200	16
11	1200	2	31	1200	4
12	1200	8	32	1200	32
13	1200	4	33	1200	4
14	1200	2	34	1200	4
15	1200	16	35	1200	64
16	1200	8	36	1200	8
17	1200	8	37	1200	8
18	1200	4	38	1200	4
19	1200	16	39	1200	4
20	1200	4	40	1200	8

**Tabela 2 Parâmetros ajustados para os ensaios da Base 3 Ninapro.**

Sujeito	Neurônios ( $\tilde{N}$ )	Fator ( $\gamma$ )	Sujeito	Neurônios ( $\tilde{N}$ )	Fator ( $\gamma$ )
1	1200	4	6	1200	4
2	1200	4	7	1200	4
3	1200	2	8	1200	4
4	1200	2	9	1200	4
5	1200	4	10	1200	4

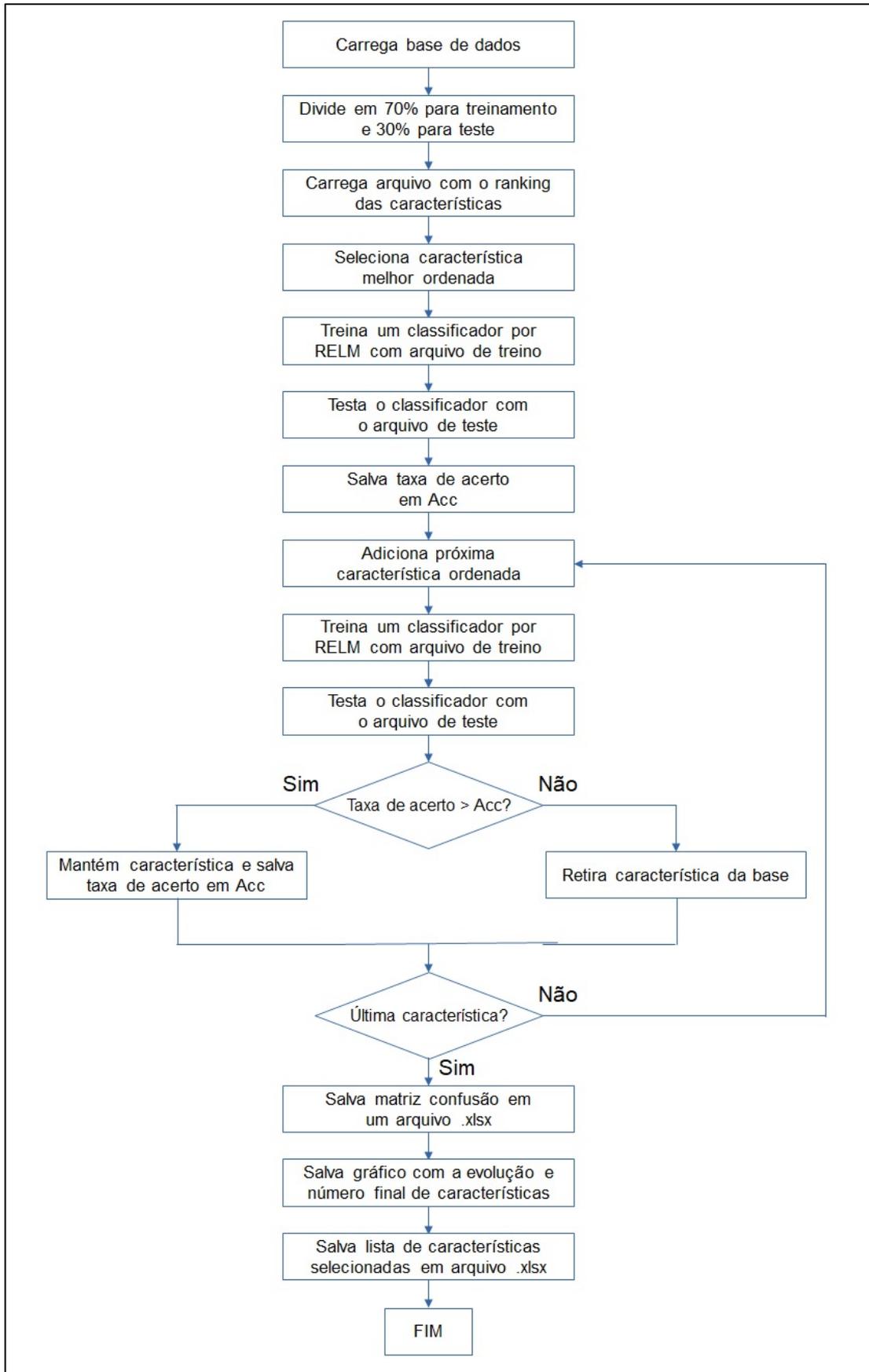


**Figura 36 Fluxograma do Algoritmo RELM.**

### 3.5.2 Algoritmo de Seleção de Características e Classificação

A última etapa do experimento proposto consiste na classificação e seleção final das características. Esta etapa tem por função executar a busca do melhor conjunto de características através da implementação de método de seleção similar aos do tipo *wrapper* sequencial para frente.

O algoritmo realiza a seleção através da classificação dos sinais com as características ordenadas por importância. A classificação é realizada iterativamente começando com a característica mais importante. Quando uma classificação termina, a próxima característica por ordem de importância é incluída no conjunto de treinamento, iniciando uma nova iteração. A taxa de acerto é analisada após cada iteração para decidir se a característica adicionada permanece ou é removida do conjunto. Se a taxa de acerto obtida for menor do que a anterior, a característica que causou essa redução é removida, caso contrário, ela permanece. Ao final obtém-se o conjunto final dos atributos selecionados e a taxa de acerto obtida. A Figura 37 representa o fluxograma de tal algoritmo.



**Figura 37 Fluxograma do algoritmo de classificação e seleção de características.**

### 3.6 VALIDAÇÃO ESTATÍSTICA DOS RESULTADOS

De modo a executar a validação estatística dos resultados obtidos, dois projetos de experimentos do tipo fatorial completo foram desenvolvidos. Em um deles considerou-se somente a Base de Dados do IEE, sendo avaliada a influência dos fatores controláveis tipo de ensaio, voluntário, movimento e método de seleção de características, na taxa de acerto do classificador (variável de resposta). Já no outro avaliou-se os resultados obtidos com as Bases Ninapro e os associados ao Ensaio A da Base do IEE. Neste Projeto de Experimentos analisou-se a influência dos fatores controláveis base de dados, movimentos, e método de seleção de características na taxa de acerto do classificador. Para a verificação da significância da influência de cada um dos fatores controláveis e da interação entre eles na variável de resposta, foi utilizado o método de análise das variâncias (ANOVA). Todos os cálculos relacionados à aplicação de tal método foram realizados através do software Minitab 17.

#### 3.6.1 Projeto de Experimentos com a Base de Dados do IEE

Neste Projeto de Experimentos foram avaliadas a influência dos fatores controláveis tipo de ensaio, voluntário, movimento, método de seleção de características e da interação de todas as combinações possíveis entre eles na taxa de acerto média do classificador. Deste modo, considerou-se somente os resultados dos experimentos realizados na Base de Dados do IEE, uma vez que as demais bases não possuem os ensaios B, C e D. Foi assim caracterizado um Projeto de Experimentos do tipo fatorial completo com 3 repetições.

O modelo dos efeitos de tal Projeto de Experimentos está descrito na Equação (63):

$$\begin{aligned}
 y_{ijklm} = & \mu + \alpha_i + \beta_j + \gamma_k + \tau_l + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\alpha\tau)_{il} + \dots \\
 & \dots + (\beta\gamma)_{jk} + (\beta\tau)_{jl} + (\gamma\tau)_{kl} + (\alpha\beta\gamma)_{ijk} + (\alpha\beta\tau)_{ijl} + \dots \\
 & \dots + (\alpha\gamma\tau)_{ikl} + (\beta\gamma\tau)_{jkl} + (\alpha\beta\gamma\tau)_{ijkl} + \epsilon_{ijklm}
 \end{aligned} \tag{63}$$

onde,  $y_{ijklm}$  corresponde a  $m$ -ésima observação dos níveis  $i, j, k, l$  dos 4 fatores controláveis analisados;  $\alpha, \beta, \gamma, \tau$  são os efeitos dos fatores controláveis;  $\mu$  é a média global da variável de resposta, e  $\epsilon$  é o erro aleatório, o qual engloba todas as fontes de variabilidade do experimento.

O número de dados analisados no Projeto A é dado pelas Equação (64):

$$\begin{aligned} \text{Número dados}(A) &= \text{voluntários} * \text{métodos seleção} * \dots \\ &\dots * \text{tipos ensaios} * \text{movimentos} * \text{repetições} \quad (64) \\ \text{Número dados}(A) &= 4 * 3 * 4 * 17 * 3 = 2448 \end{aligned}$$

### 3.6.2 Projeto de Experimentos com as Bases de Dados Ninapro e a Base do IEE

Aqui foram considerados os sinais das bases do Ninapro mais os do Ensaio A da Base de Dados do IEE. Uma vez que cada uma das bases possui sinais adquiridos de voluntários distintos, optou-se por não avaliar o fator controlável voluntário neste projeto. Assim, uma vez que as três bases de dados analisadas possuem número distintos de ensaios, neste caso obteve-se um Projeto de Experimentos desbalanceado.

É importante ressaltar que a análise baseada em um Projeto de Experimentos do tipo fatorial completo é uma aproximação devido à falta de informação completa de como foi desenvolvida e obtida a Base Ninapro. O modelo dos efeitos está descrito na Equação (65):

$$y_{ijkm} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijkm} \quad (65)$$

onde,  $y_{ijkm}$  corresponde a  $m$ -ésima observação dos níveis  $i, j, k$ , dos 3 fatores controláveis analisados;  $\alpha, \beta, \gamma$  são os efeitos dos fatores controláveis;  $\mu$  é a média global da variável de resposta, e  $\epsilon$  é o erro aleatório, o qual engloba todas as fontes de variabilidade do experimento.

O número de dados analisados no Projeto B é dado pelas Equação (66):

$$\begin{aligned} \text{Número dados}(B) &= \text{métodos seleção} * \text{movimentos} * n^{\circ} \text{ total de ensaios} \\ \text{Número dados}(B) &= 3 * 17 * 62 = 3162 \quad (66) \end{aligned}$$

## 4 RESULTADOS E DISCUSSÕES

Para melhor compreensão dos resultados expostos neste capítulo, a Tabela 3 apresenta a relação de cada um dos movimentos analisados com a correspondente identificação utilizada na descrição dos resultados. No presente estudo optou-se por desconsiderar o repouso nas análises, pois o mesmo representa mais de 50% da base de dados. Portanto, apenas 17 movimentos foram analisados. Além da identificação supracitada, a Tabela 3 ressalta o grupo muscular associado a cada um dos movimentos realizados durante os ensaios. Assim, através desta análise é possível relacionar as taxas de acerto em função dos movimentos e grupo muscular afetado.

**Tabela 3 Identificação dos movimentos e dos respectivos músculos envolvidos (adaptado de CENE, 2016).**

Identificação	Movimento Realizado	Principais Músculos Envolvidos
m01	Estender o dedo 1	Flexor profundo dos dedos, Flexor superficial dos dedos, Adutor longo do polegar.
m02	Estender os dedos 2 e 3, flexionando os demais	Flexor profundo dos dedos, Flexor superficial dos dedos, Extensor dos dedos
m03	Estender os dedos 4 e 5, flexionando os demais	Flexor profundo dos dedos, Flexor superficial dos dedos, Extensor dos dedos.
m04	Flexionar o dedo 1, estendendo os demais	Extensor dos dedos, Flexor do polegar.
m05	Estender todos os dedos	Extensor dos dedos.
m06	Flexionar todos os dedos	Flexor profundo dos dedos, Flexor superficial dos dedos.
m07	Movimento de apontar	Flexor profundo dos dedos, Flexor superficial dos dedos, Extensor dos dedos.
m08	Adução de dedos estendidos	Flexor profundo dos dedos, Flexor superficial dos dedos.
m09	Supinar o pulso (dedo 3 como eixo)	Bíceps braquial, Tríceps braquial, Supinador.
m10	Pronar o pulso (dedo 3 como eixo)	Bíceps braquial, Tríceps braquial, Pronador redondo.
m11	Supinar o pulso (dedo 5 como eixo)	Bíceps braquial, Tríceps braquial, Supinador.
m12	Pronar o pulso (dedo 5 como eixo)	Bíceps braquial, Tríceps braquial, Pronador redondo.

Continuação da Tabela 3.

m13	Flexionar o pulso	Flexor radial do carpo, Flexor ulnar do carpo, Extensor radial longo do carpo.
m14	Estender o pulso	Flexor radial do carpo, Flexor ulnar do carpo.
m15	Desvio radial do pulso	Extensor longo radial do carpo, Extensor dos dedos.
m16	Desvio ulnar do pulso	Extensor ulnar do carpo, Extensor dos dedos.
m17	Estender o pulso com a mão fechada	Flexor radial do carpo, Flexor ulnar do carpo, Palmar longo, Flexor profundo dos dedos, Flexor superficial dos dedos

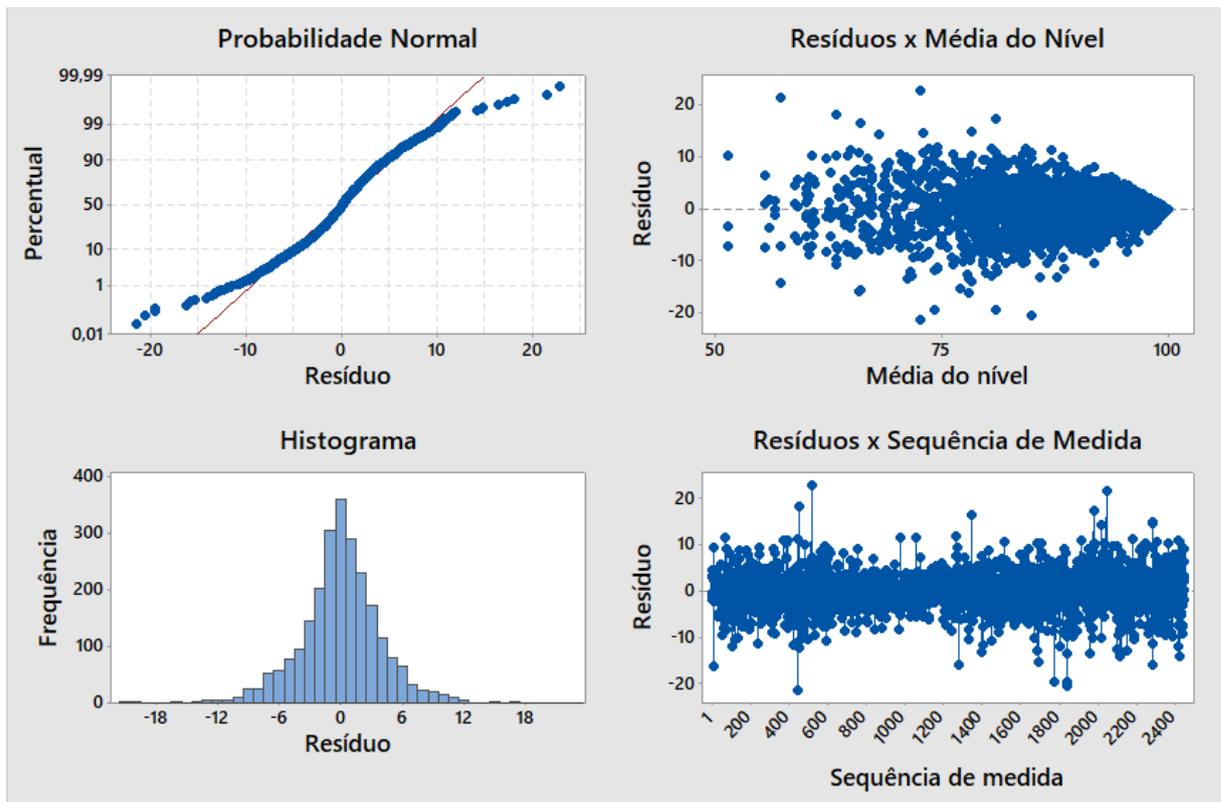
#### 4.1 VALIDAÇÃO DOS MODELOS DOS PROJETOS DE EXPERIMENTOS

De modo a validar a premissa de que os experimentos foram completamente aleatorizados, quatro verificações foram conduzidas nos dados analisados em cada um dos projetos de experimentos. A primeira baseou-se na análise do histograma dos resíduos. Para validar a premissa supracitada, o histograma dos resíduos deve seguir uma distribuição normal de média zero. A segunda consiste na verificação da relação dos resíduos em função do tempo (sequência das medidas), de modo a examinar a existência de correlação entre os resíduos. Aqui, a validação do modelo é caracterizada pela inobservância de correlação, ou seja, a curva da relação citada anteriormente não deve apresentar nenhuma estrutura aparente. A terceira está associada com a análise da dispersão dos resíduos em torno da média dos níveis dos fatores controláveis. Também nesta verificação não deve ser observado nenhuma estrutura aparente. Por fim, a quarta verificação constitui-se na avaliação do gráfico de Probabilidade Normal. Tal gráfico deve apresentar tendência linear para o erro estar normalmente distribuído.

##### 4.1.1 Validação do Projeto de Experimentos com a Base de Dados do IEE

A Figura 38 mostra a análise dos resíduos através das quatro verificações expostas na seção 4.1. Examinando o histograma dos resíduos observou-se que o mesmo segue uma

distribuição normal. Já as relações entre resíduos e a média dos níveis e entre resíduos e sequência de medidas não apresentaram estrutura aparente, o que também valida o modelo deste Projeto de Experimentos. Por fim, o gráfico de Probabilidade Normal indica tendência linear o que corrobora o fato de que o erro segue uma distribuição normal. Contudo, por meio da análise dos resíduos o modelo do Projeto de Experimentos com a Base de Dados do IEE foi validado.

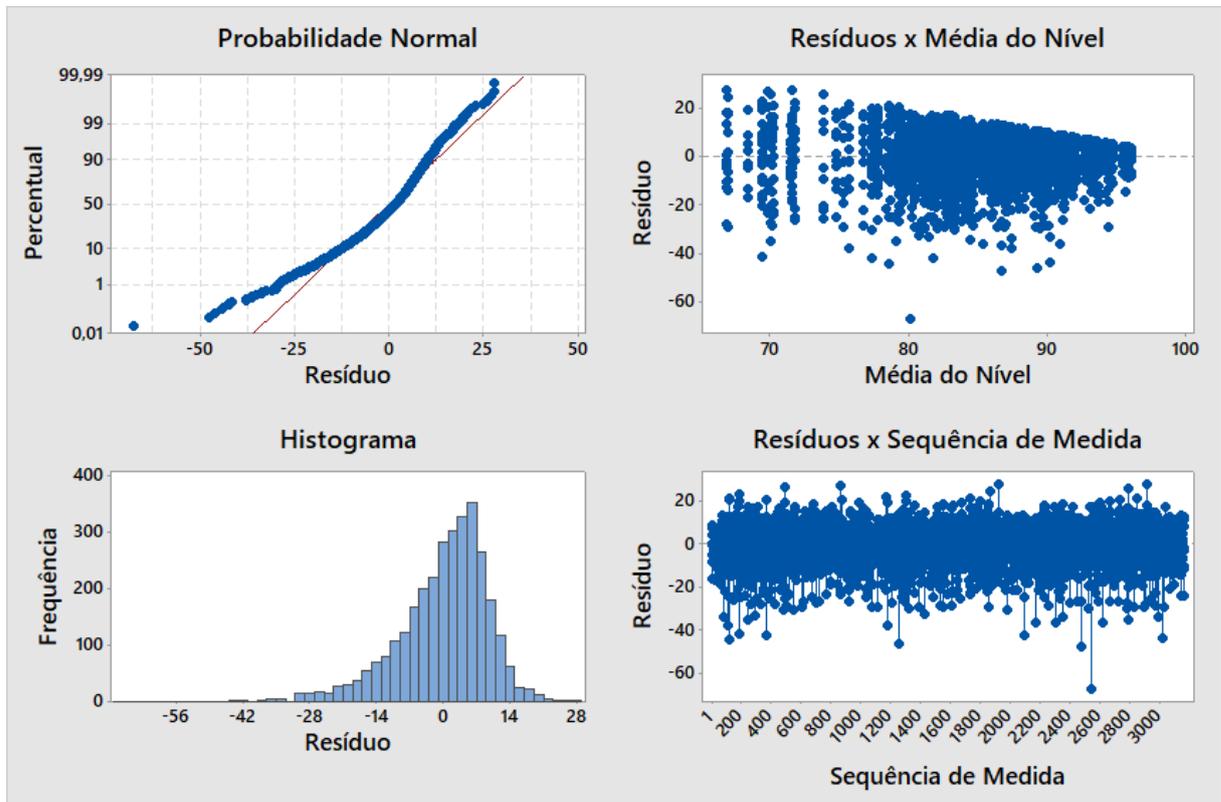


**Figura 38** Análise dos resíduos do Projeto de Experimentos com a Base de Dados do IEE.

#### **4.1.2 Validação do Projeto de Experimentos com as Bases de Dados Ninapro e a Base de Dados do IEE**

A Figura 39 ilustra a análise dos resíduos do Projeto de Experimentos com as Bases de Dados do Ninapro e a Base do IEE. O histograma dos resíduos e o gráfico de Probabilidade Normal indicam que o erro segue uma distribuição normal. Em contrapartida, as relações entre

resíduos e a média dos níveis e entre resíduos e sequência de medidas não apresentaram estrutura aparente. Tais fatos validam o modelo deste Projeto de Experimentos.



**Figura 39** Análise dos resíduos do Projeto de Experimentos com as Bases Ninapro e a do IEE.

## 4.2 RESULTADOS DO PROJETO DE EXPERIMENTOS COM A BASE DE DADOS DO IEE

Em função da grande quantidade de interações entre fatores controláveis optou-se por realizar todas as análises com base na taxa de acerto média obtida para cada fator controlável. No decorrer desta seção, serão abordados os resultados da influência de cada um dos fatores controláveis (voluntário, movimento, método de seleção e tipo de ensaio) e também da interação de todas as combinações entre eles, na taxa de acerto média obtida. A Tabela 4 apresenta os resultados obtidos do correspondente Projeto de Experimentos com a Base do IEE com um intervalo de confiança de 95%.

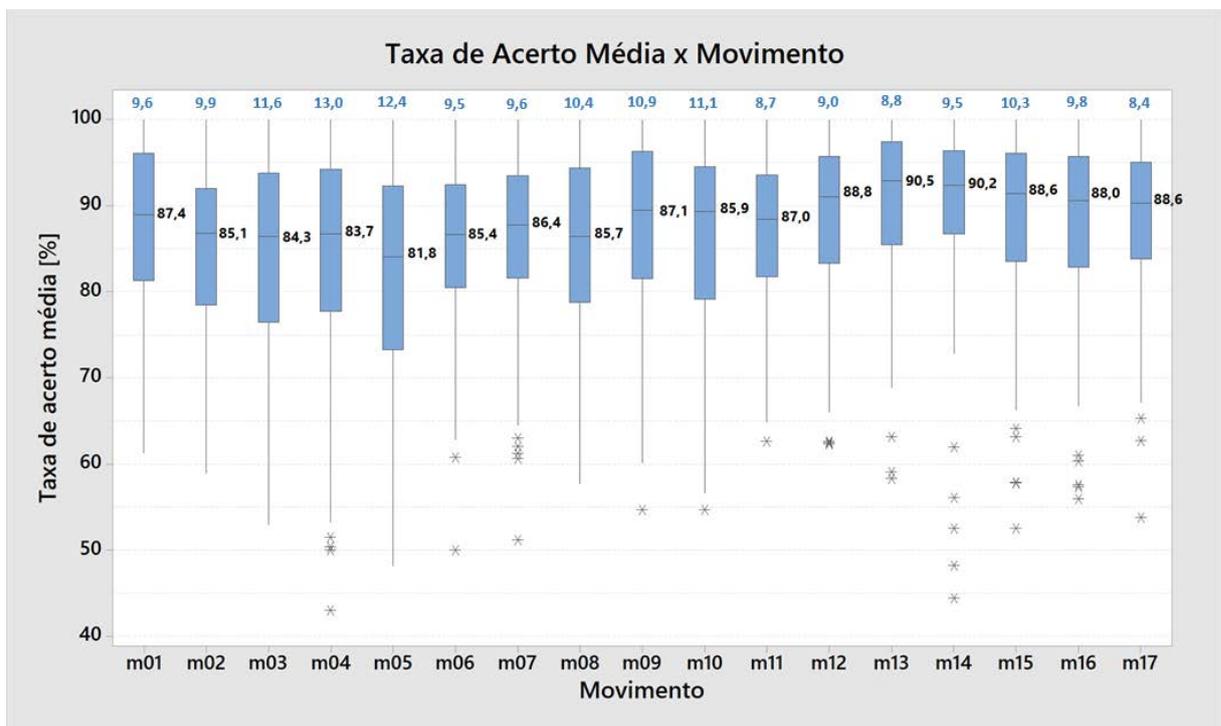
**Tabela 4 Análise de Variância do Projeto de Experimentos com a Base de Dados do IEE.**

Fonte de Variação	Soma dos Quadrados	Graus de Liberdade	Médias Quadradas	Fator F	Fator P	Resultado
Voluntário (Fator A)	40072	3	13357,3	542,98	0,00	Significativo
Método de Seleção (Fator B)	1639	2	819,4	33,31	0,00	Significativo
Ensaio (Fator C)	56870	3	18956,6	770,59	0,00	Significativo
Movimento (Fator D)	12259	16	766,2	31,15	0,00	Significativo
Interação AB	1541	6	256,8	10,44	0,00	Significativo
Interação AC	18006	9	2000,7	81,33	0,00	Significativo
Interação AD	13389	48	278,9	11,34	0,00	Significativo
Interação BC	3347	6	557,9	22,68	0,00	Significativo
Interação BD	3697	32	115,5	4,70	0,00	Significativo
Interação CD	6143	48	128	5,20	0,00	Significativo
Interação ABC	15709	18	872,7	35,48	0,00	Significativo
Interação ABD	6831	96	71,2	2,89	0,00	Significativo
Interação ACD	15600	144	108,3	4,40	0,00	Significativo
Interação BCD	6471	96	67,4	2,74	0,00	Significativo
Interação ABCD	23120	288	80,3	3,26	0,00	Significativo
Erro	40087	1632	24,6			
Total	264780	2447				

Da Tabela 4 conclui-se que, todos os fatores controláveis analisados as interações entre os mesmos apresentaram influência significativa na taxa de acerto média. Em todos eles o fator  $p$  resultou valor 0,0. Nas seções a seguir, os resultados obtidos da análise de cada um dos fatores controláveis e das combinações entre eles serão avaliados em minúcias.

#### 4.2.1 Análise do Fator Controlável Movimento

As médias (em preto) e o desvio padrão (em azul) das taxas de acerto obtidas para cada um dos movimentos considerados estão descritas no boxplot da Figura 40. Analisando os referidos dados, observa-se que o movimento m05, que é o que apresenta a menor taxa de acerto média (81,8%), possui o segundo maior desvio padrão (12,4%). A maior dispersão em torno da média está associada ao movimento m04, cujo desvio padrão resultou em 13%. Em contrapartida, o menor desvio padrão obtido corresponde ao movimento m17 com 8,4%, que está entre as maiores taxas de acerto média (88,6%). Tais resultados indicam que os movimentos m04 e m05 possuíram maior variabilidade em sua caracterização entre os ensaios, ao passo que o m17, por sua vez, apresentou a menor variação.



**Figura 40** Dispersão da taxa de acerto obtida para cada movimento.

Ainda de acordo com a referida figura, observa-se taxas de acerto inferiores para os movimentos m02, m03, m04 e m05. Tais movimentos estão relacionados a, praticamente, os mesmos grupos musculares, e, conseqüentemente, possuem padrões muito parecidos de

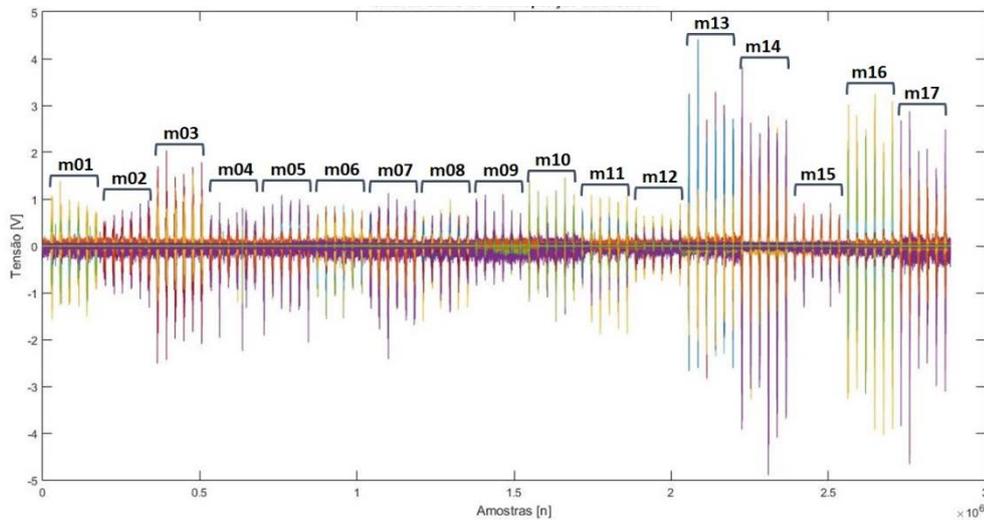
ativação muscular, especialmente o m02 e m03. Além disso, são movimentos finos, ou seja, não exigem níveis de ativação elevados para sua execução, resultando em sinais adquiridos de pequena amplitude e mais suscetíveis a ruído.

O movimento que apresentou menor taxa de acerto (81,8%) foi o m05, o qual é resultado da ativação de um único músculo principal, o Extensor dos dedos. Dentre todos movimentos analisados no presente trabalho, ele é o único que possui essa característica. Tal fato, adicionalmente aos já mencionados, prejudicam consideravelmente a qualidade da caracterização do padrão do movimento m05, visto que, a ocorrência de só haver ativação de um único músculo principal diminui a probabilidade dos eletrodos registrarem ativação muscular. Porém, deve-se considerar a questão do posicionamento dos eletrodos, uma vez que, um eletrodo bem posicionado pode caracterizar melhor um movimento específico do que uma série de eletrodos posicionados fora da região de máxima ativação do grupo muscular envolvido.

Em contrapartida, os movimentos m12, m13, m14, m15, m16 e m17 foram os que apresentaram as melhores taxas de acerto. Todos acima de 88%. Estes movimentos, em especial o m13, m14, m16 e o m17, parecem ter sido favorecidos pela disposição dos eletrodos. Durante os ensaios, observou-se que a amplitude do sinal de sEMG era maior para tais movimentos, e também abrangia um maior número de sensores. O gráfico da Figura 41 ilustra o sinal de sEMG adquirido durante a execução de uma repetição do ensaio A (exercício sequencial com 6 repetições, tal qual o realizado nas Bases Ninapro). Nele, os sinais dos 12 eletrodos estão superpostos.

Como se observa na Figura 41, a amplitude do sinal adquirido para os movimentos m13, m14, m16 e o m17 é consideravelmente superior à dos demais. Conseqüentemente, a relação sinal ruído nestes movimentos é superior em comparação a dos outros, possibilitando, assim, uma melhor caracterização, facilitando, por conseguinte, a tarefa de classificação. Também vale

ressaltar que tais movimentos exigem maior esforço ao voluntário o que explica essa diferença na amplitude do sinal adquirido.



**Figura 41** Gráfico do sinal de sEMG adquirido durante execução do ensaio A.

#### 4.2.2 Análise do Fator Controlável Voluntário

Na Figura 42 está exposta a média (em preto) e o desvio padrão (em azul) da taxa de acerto obtida para cada um dos voluntários. Vale destacar que o voluntário 2 apresentou a maior taxa de acerto média com 93,7% e também foi o que obteve a menor variabilidade deste resultado de classificação entre ensaios, apresentando desvio padrão da taxa de acerto igual a 6,1%. Por outro lado, os voluntários 1 e 4 foram os que apresentaram maior dispersão da taxa de acerto entre ensaios, resultando em desvio padrão igual a 11,3% e 11,1% respectivamente. Esta constatação indica que o voluntário 2 manteve uma constância maior na forma de execução dos exercícios do que os demais voluntários.

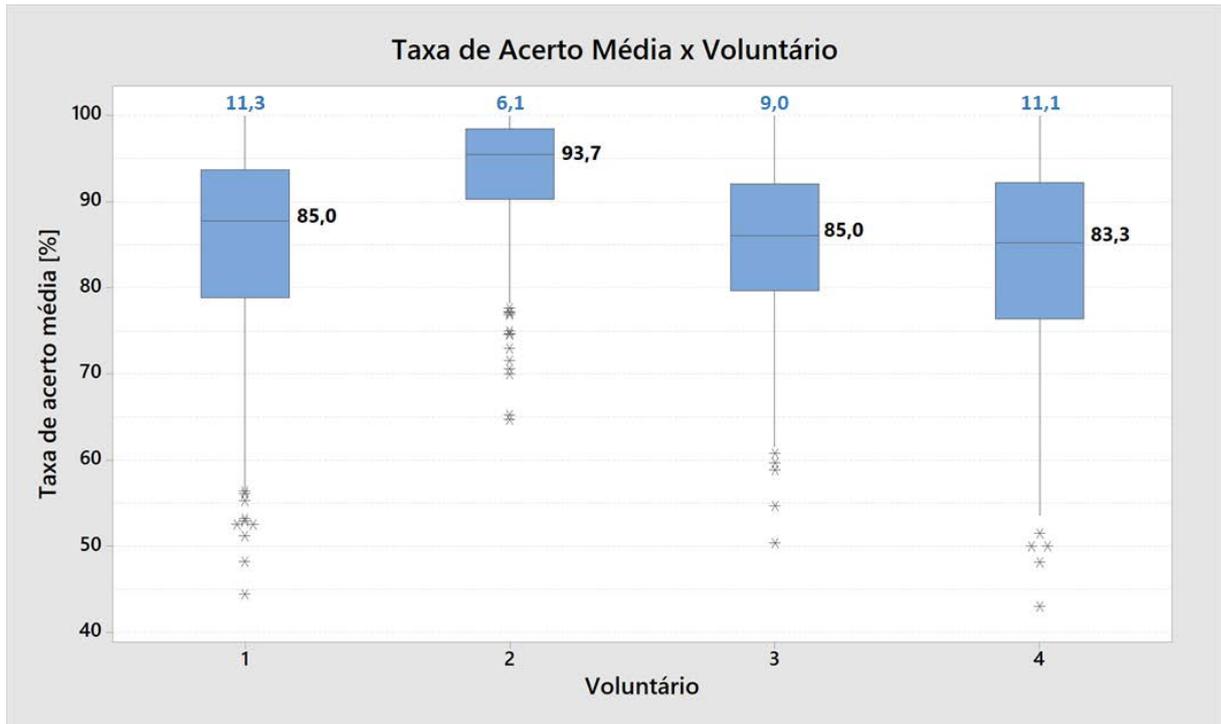
Como era esperado, o efeito do fator controlável voluntário na taxa de acerto média foi verificado como sendo significativo através da análise do correspondente Projeto de Experimentos, ou seja, é possível afirmar com grau de confiança de 95% para a base de dados avaliada que a taxa de acerto média é dependente do sujeito participante do ensaio. Sendo assim,

a aquisição de sinais de sEMG pode ser facilitada em alguns sujeitos e dificultada em outros, devido à presença considerável de tecido adiposo sob a pele, ocorrência de disfunções neuromusculares, dentre outras. Porém, o mais importante é que o comportamento do correspondente classificador é influenciado pelas características fisiológicas do sujeito.

Outra fonte de variabilidade entre voluntários reside na forma como cada um entende e executa cada um dos movimentos. O jeito que um indivíduo faz um movimento específico pode beneficiar ou prejudicar a qualidade do sinal adquirido. Se um sujeito executa o movimento realizando menor esforço muscular do que os demais indivíduos, menores serão os níveis de ativação observados e o sinal de sEMG obtido apresentará, conseqüentemente, reduzida relação sinal ruído.

Por outro lado, também deve ser levada em consideração a questão do posicionamento dos eletrodos. Por mais que se tenha padronizada a localização dos mesmos, é natural que haja variabilidade entre voluntários e até mesmo para o mesmo sujeito em ensaios realizados em dias diferentes. A caracterização dos padrões de movimentos está relacionada com o posicionamento dos eletrodos, uma vez que, como já mencionado na seção 4.2.1, um eletrodo bem posicionado, pode caracterizar melhor um movimento específico do que um conjunto de eletrodos posicionados distantes da região de máxima atividade do grupo muscular envolvido.

De acordo com a Figura 42, o voluntário 2 apresentou taxa de acerto média consideravelmente superior aos demais (93,6% enquanto que os outros três apresentaram resultados variando entre 83,3 e 85%). Uma das prováveis razões para essa diferença está no fato do voluntário 2 ser o autor do presente trabalho e um dos autores da base de dados adquirida. Certamente, a motivação do voluntário 2 em executar os movimentos da melhor forma possível era superior à dos demais e pode explicar tamanha diferença nos resultados.



**Figura 42** Dispersão da taxa de acerto obtida para cada voluntário.

#### 4.2.3 Análise do Fator Controlável Tipo de Ensaio

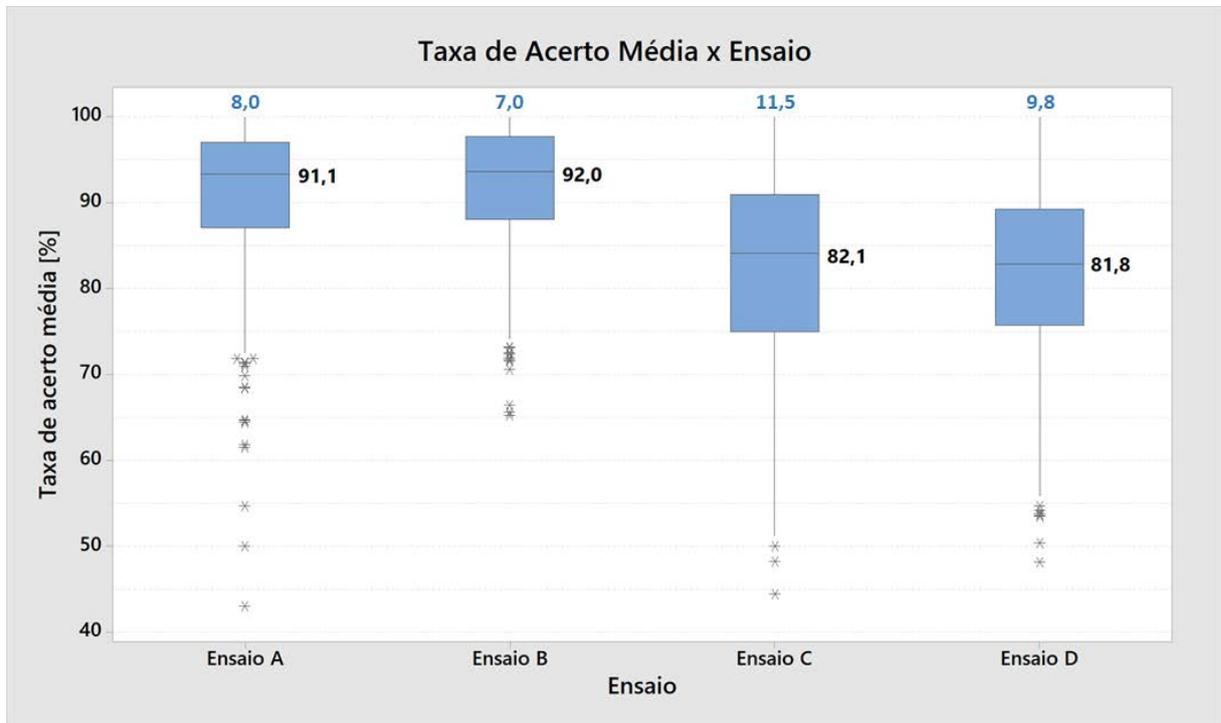
A informação da média (em preto) e desvio padrão (em azul) da taxa de acerto resultante para cada tipo de ensaio está descrita na Figura 43. Analisando os dados da referida figura concernente a questão da dispersão da taxa de acerto em torno da média, percebe-se menor variabilidade para os ensaios sequenciais em comparação com os ensaios aleatorizados. O maior desvio padrão está associado ao Ensaio B com 7,0% e o maior corresponde ao Ensaio C com 11,5%. A questão da aleatorização dos movimentos pode estar associada a maior variabilidade observada nos ensaios C e D, uma vez que, nestes tipos de ensaio, o voluntário tem pouco tempo para assimilar qual será o próximo movimento. Deste modo, em ensaios onde movimentos parecidos são apresentados na sequência, a probabilidade de o voluntário fazer confusão entre os movimentos e realizar a execução incorretamente é maior, impactando de forma negativa na caracterização dos movimentos e posterior taxa de acerto do classificador.

Analisando ainda a Figura 43 observa-se claramente que os ensaios sequenciais (A e B) resultaram em melhores taxas de acerto médio do que os que apresentava os movimentos de modo aleatorizado (C e D). Convém destacar o fato de que, durante os ensaios sequenciais, o voluntário já tem a informação de qual será o próximo movimento com bastante antecedência e a chance de ocorrer algum equívoco, como por exemplo a execução de movimento incorreto, é baixa. Já no decorrer dos ensaios aleatorizados, o movimento subsequente é uma surpresa para o sujeito e, mesmo com o período de repouso de aproximadamente 10 segundos onde é indicada a próxima ação, são maiores as possibilidades de o indivíduo executar o movimento errado, ou iniciar o movimento de forma incorreta e depois corrigir.

Além disso, deve-se levar em conta que os ensaios aleatorizados são mais longos do que os sequenciais. Isso se deve ao fato deles possuírem um tempo maior de repouso para dar mais tempo ao voluntário se preparar para o próximo movimento (10 segundos enquanto que os sequenciais possuem apenas 9). Adicionalmente, o tempo de execução dos movimentos nos ensaios aleatorizados foi maior do que os sequenciais, pois, como em tais ensaios o vídeo de cada movimento era executado independentemente em cada iteração, o software utilizado (LabView 2009) inseriu um pequeno atraso em cada uma das execuções, que era o tempo do diretório do referido vídeo ser carregado. Assim, a duração dos ensaios A e B ficou em aproximadamente 24 minutos e 40 minutos respectivamente, enquanto que os ensaios C e D demoraram cerca de 42 minutos e 1 hora respectivamente.

Por conseguinte, os exercícios aleatorizados foram mais exaustivos do que os sequenciais, além de também serem os que exigiam um maior nível de concentração para os voluntários. Isso pode explicar, também, o porquê da diferença observada nas taxas de acerto dos ensaios A e B em relação aos C e D. Porém de forma geral, essa diferença merece ser melhor investigada no decorrer do aumento da correspondente base (aumento do número de voluntários), assim como, a relação do tempo de uso do correspondente sistema por cada

voluntário. Provavelmente existe uma forte influência da usabilidade do sistema e, portanto, do tempo de exposição ao mesmo para cada voluntário.



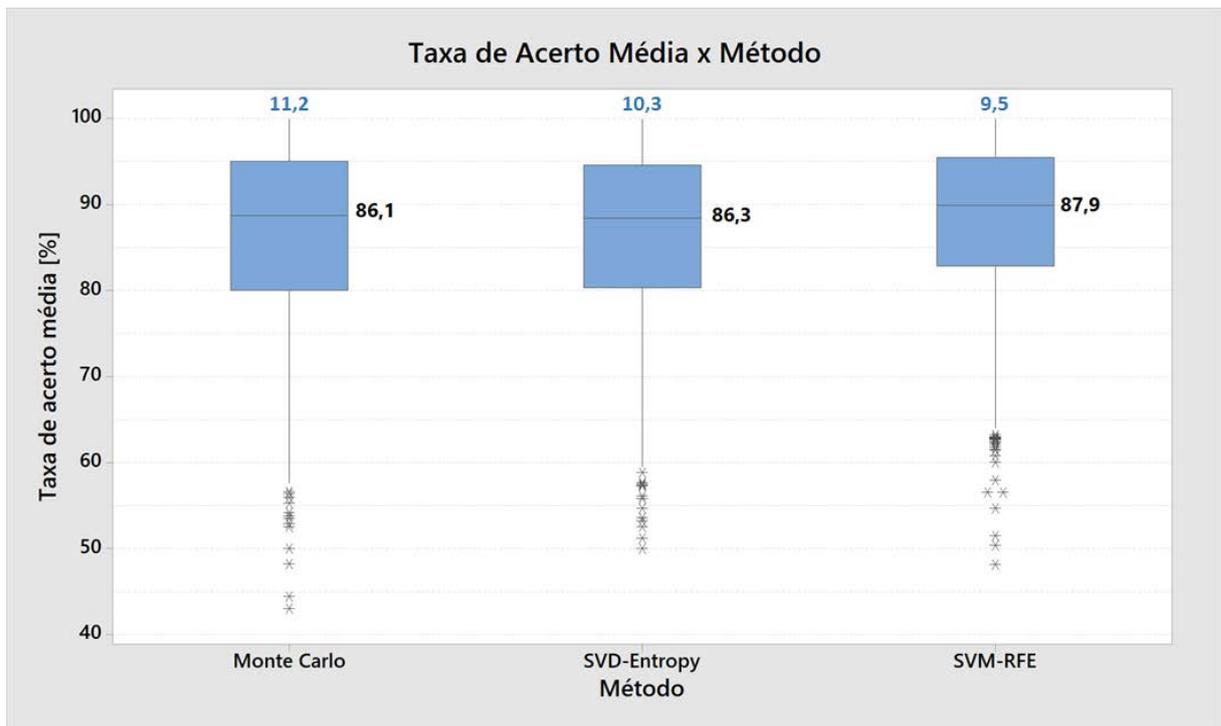
**Figura 43** Dispersão da taxa de acerto obtida para cada tipo de ensaio.

#### 4.2.4 Análise do Fator Controlável Método de Seleção

Na Figura 44 estão descritas a média (em preto) e o desvio padrão (em azul) da taxa de acerto resultante da aplicação de cada um dos três métodos de seleção de características abordados neste trabalho. Da análise destes valores observa-se que o algoritmo que apresentou maior variabilidade na taxa e acerto resultante foi o Monte Carlo, com desvio padrão igual a 11,2%. Por outro lado, a menor variação foi constatada no método SVM-RFE, o qual obteve desvio padrão de 9,5%.

De acordo com os resultados da análise das variâncias expostos na Tabela 3, verificou-se significância na influência do método de seleção de características na taxa de acerto média, uma vez que o valor de  $p$  resultou em 0,0. Analisando a Figura 44, percebe-se que o método

SVM-RFE foi o que apresentou melhor resultado (87,9%) frente aos demais. Monte Carlo e SVD *Entropy* resultaram em taxas de acerto muito próximas, com diferença de apenas 0,2%. Tal resultado indica superioridade do algoritmo SVM-RFE em relação aos outros dois métodos no concernente à taxa de acerto do classificador para os sinais da Base de Dados do IEE.



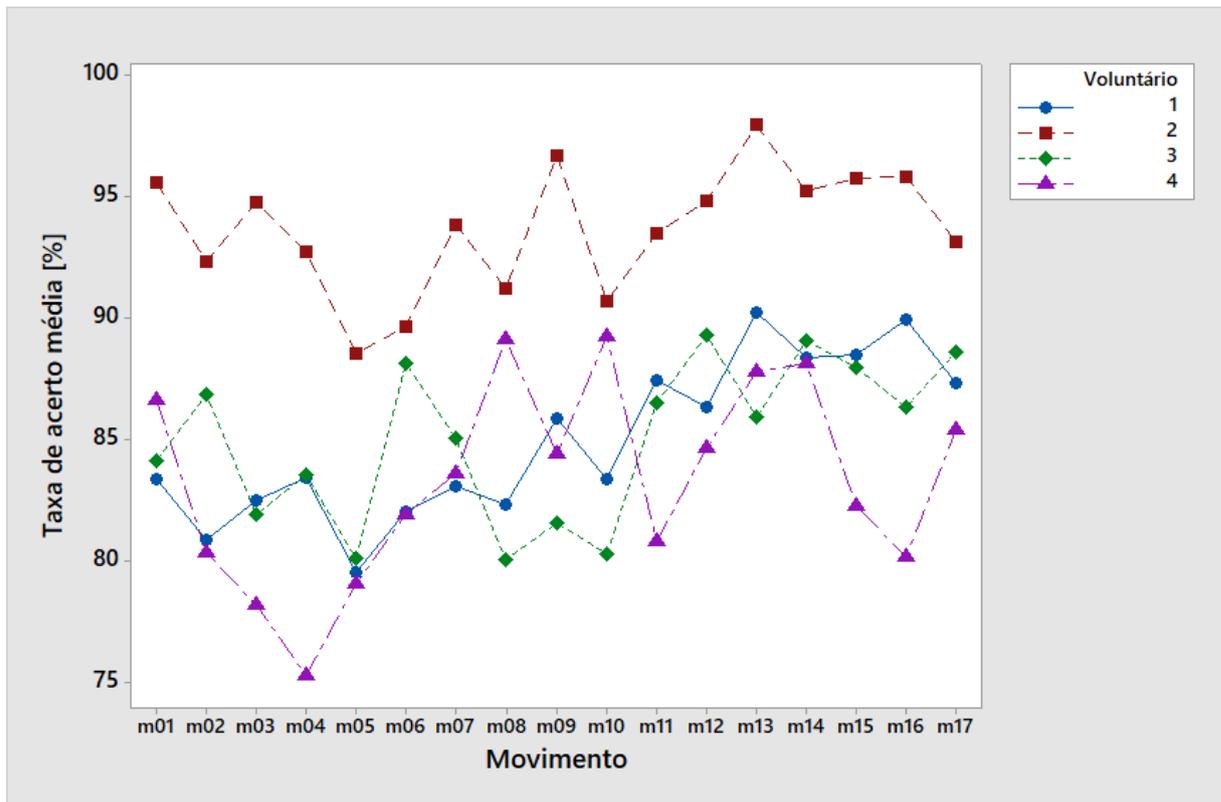
**Figura 44** Dispersão da taxa de acerto obtida para cada método.

#### 4.2.5 Análise da Interação entre Voluntário e Movimento

A Figura 45 apresenta a relação entre a taxa de acerto média e cada um dos 17 movimentos para os 4 voluntários. A partir do referido gráfico, observa-se clara influência da interação entre voluntário e movimento na taxa de acerto média. Os movimentos com maior e menor taxa de acerto média foram diferentes para cada voluntário exceto para o 1 e o 2. Para estes últimos a maior taxa de acerto foi observada no movimento m13 (90,2% para o voluntário 1 e 97,9% para o 2) e a menor para o m05 (79,5% para o voluntário 1 e 88,5% para o 2). Já para

o voluntário 3 a maior está associada a m12 (89,3%) e a menor a m05 (80,1%), enquanto que o sujeito 4 obteve maior taxa de acerto no movimento m10 (89,2%) e menor no m04 (75,3%).

Esse resultado é compreensível, visto que, cada ser humano possui características fisiológicas particulares. Logo, é aceitável que existam movimentos que sejam mais naturais para alguns voluntários e menos para outros. Porém, vale destacar que os movimentos que apresentaram as maiores taxas de acerto média (m13, m14, m15, m16, e 17) também foram os que resultaram nas maiores taxas de acerto média para cada um dos voluntários, a exceção dos movimentos m15 e m16 para o voluntário 4. Isso reforça a conclusão de que tais movimentos foram bem caracterizados através dos sinais adquiridos nas posições indicadas na Figura 20.



**Figura 45** Interação entre voluntários e movimento.

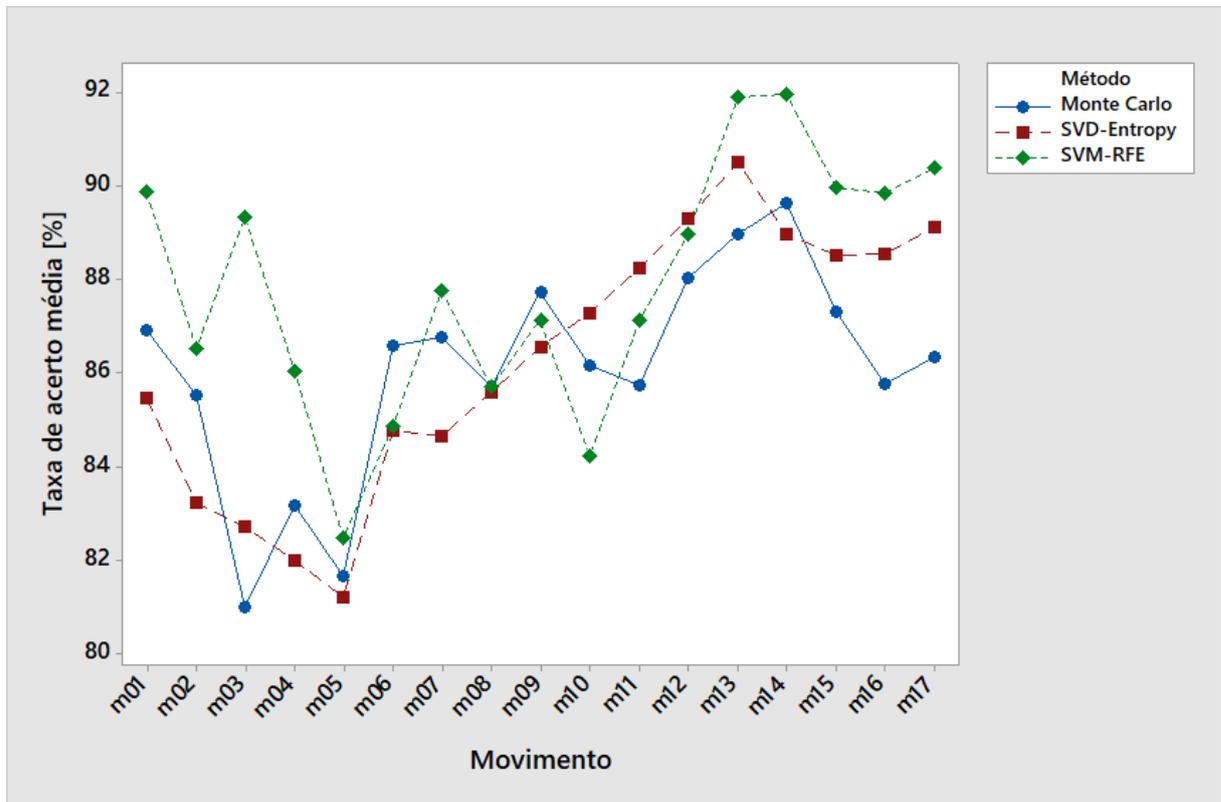
#### 4.2.6 Análise da Interação entre Método de Seleção e Movimento

A Figura 46 ilustra a relação entre a taxa de acerto média de cada um dos 17 movimentos para os 3 métodos de seleção de características implementados. Analisando o referido gráfico, observa-se que, para alguns movimentos, a taxa de acerto média varia consideravelmente entre os três métodos implementados no presente trabalho. Esse comportamento é observado principalmente no movimento m03, onde o método SVM-RFE mostrou resultado muito superior frente aos outros dois. Tal fato indica que as características selecionadas a partir do ranking realizado pelo algoritmo da SVM-RFE, representam melhor o movimento m03 do que o conjunto selecionado através da implementação dos métodos Monte Carlo e SVD *Entropy*. Por conseguinte, fica evidenciado que existe uma relação entre características e movimentos. Aqui, é importante ressaltar que, quando se fala em relação entre características e movimentos, leva-se em conta não só as características em si, mas também o número delas selecionadas para cada treinamento realizado individualmente. Isso porque, em cada execução do algoritmo proposto no presente trabalho, o número final de características selecionadas varia de ensaio para ensaio.

Em contrapartida, essa conclusão vai ao encontro dos resultados expostos em (CENE, 2016). Em seu trabalho, (CENE, 2016) avalia a influência de 3 características na representação de sinais de sEMG para classificação de movimentos do segmento mão-braço. São elas: variância, RMS e frequência mediana. Para tanto, foram testadas todas as combinações possíveis entre elas e também entre combinações dos 12 canais utilizados nas bases 2 e 3 do Ninapro (as mesmas utilizadas no presente trabalho). Como resultado da análise estatística realizada sobre os resultados obtidos dessa variação, não foi identificada influência significativa da interação entre características e movimentos na taxa de acerto média do classificador.

Porém, vale destacar que a análise realizada no presente trabalho é mais abrangente quanto ao número de características consideradas (13 contra apenas 3 de (CENE, 2016)). Logo,

a variabilidade de características entre os ensaios realizados foi maior neste trabalho do que o executado por (CENE, 2016).



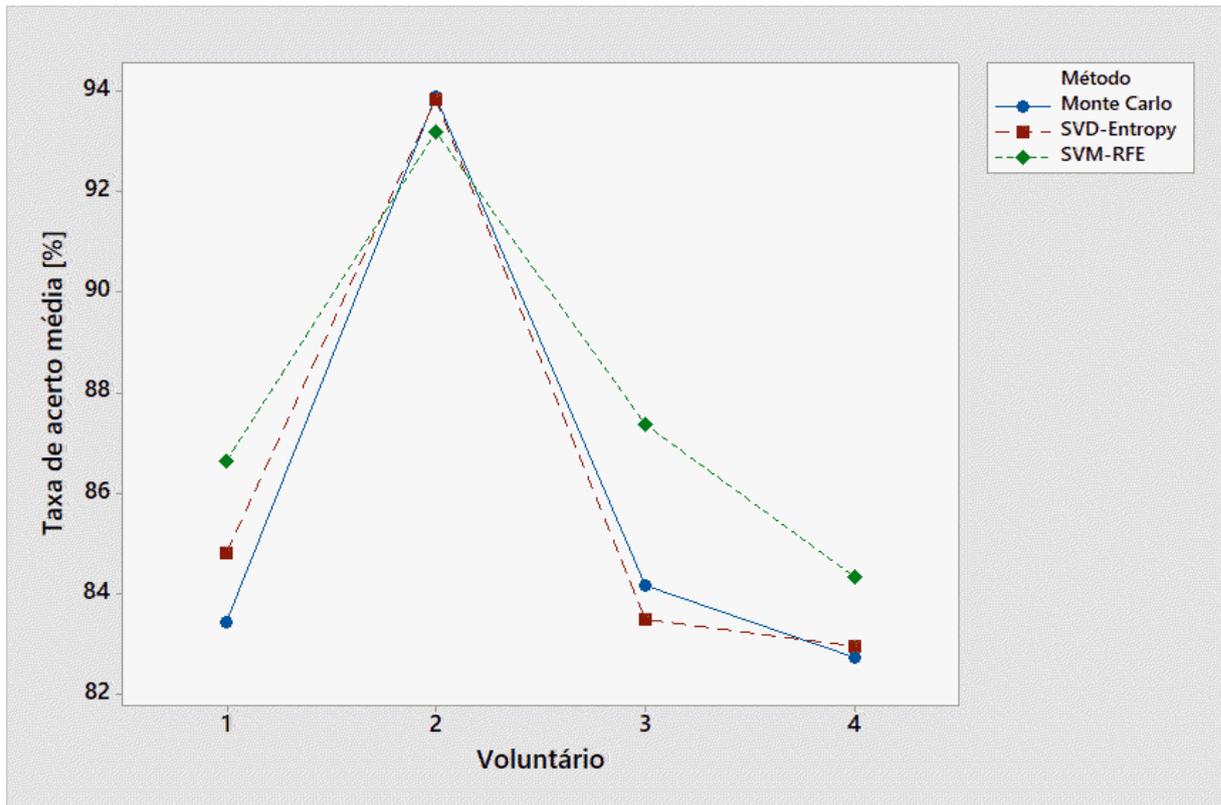
**Figura 46** Interação entre método e movimento.

#### 4.2.7 Análise da Interação entre Voluntário e Método de Seleção

A Figura 47 apresenta a relação entre o fator controlável voluntário e a taxa de acerto média para os três métodos de seleção de características considerados. Segundo resultados expostos na Tabela 4, conclui-se que a interação entre tais fatores controláveis influencia significativamente a taxa de acerto média com um nível de confiança de 95%, visto que o valor obtido para o fator  $p$  foi de 0,0.

A influência da combinação de sujeitos e métodos de seleção de características na taxa de acerto média fica evidente ao se observar que as 3 curvas representadas na Figura 47 se cruzam em uma série de pontos. Por exemplo, quando se examina a curva associada ao

algoritmo SVM-RFE (tracejado verde), nota-se que tal método é responsável pelas maiores taxas de acerto média para os voluntários 1, 3 e 4. Por outro lado, para o sujeito 2 ele corresponde a menor taxa de acerto obtida. Tal resultado indica que a escolha do método de seleção de características deve ser feita individualmente para cada sujeito.

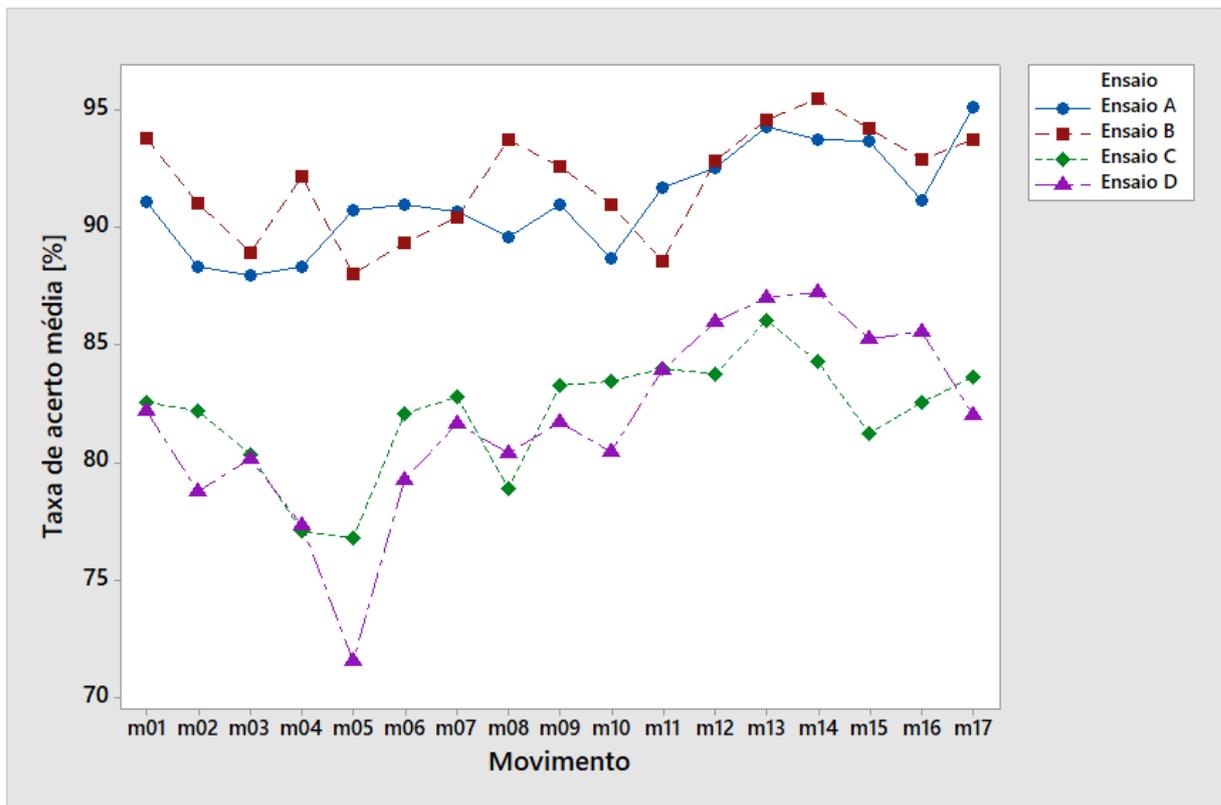


**Figura 47** Análise da interação entre voluntário e método de seleção.

#### 4.2.8 Análise da Interação entre Tipo de Ensaio e Movimento

Os resultados apontados na Tabela 4 indicam ser significativa a influência da interação dos fatores controláveis tipo de ensaio e movimento na taxa de acerto média do classificador, uma vez que o fator  $p$  encontrado ser igual a 0,0. A Figura 48 mostra a relação entre o fator controlável movimento e a taxa de acerto média para os quatro tipos de ensaio realizados. Desta figura observa-se que as curvas dos ensaios A e B estão separadas das dos ensaios C e D. Em nenhum dos 17 movimentos a taxa de acerto média dos ensaios C e D foi superior à dos outros

dois. Entretanto, quando se analisa os pares de curvas separadamente, nota-se que tanto as curvas dos ensaios A e B quanto as curvas referentes aos ensaios C e D se cruzam em uma série de pontos, o que justifica a hipótese de que a taxa de acerto média varia significativamente a medida em que se alteram o tipo de ensaio e de movimento simultaneamente.

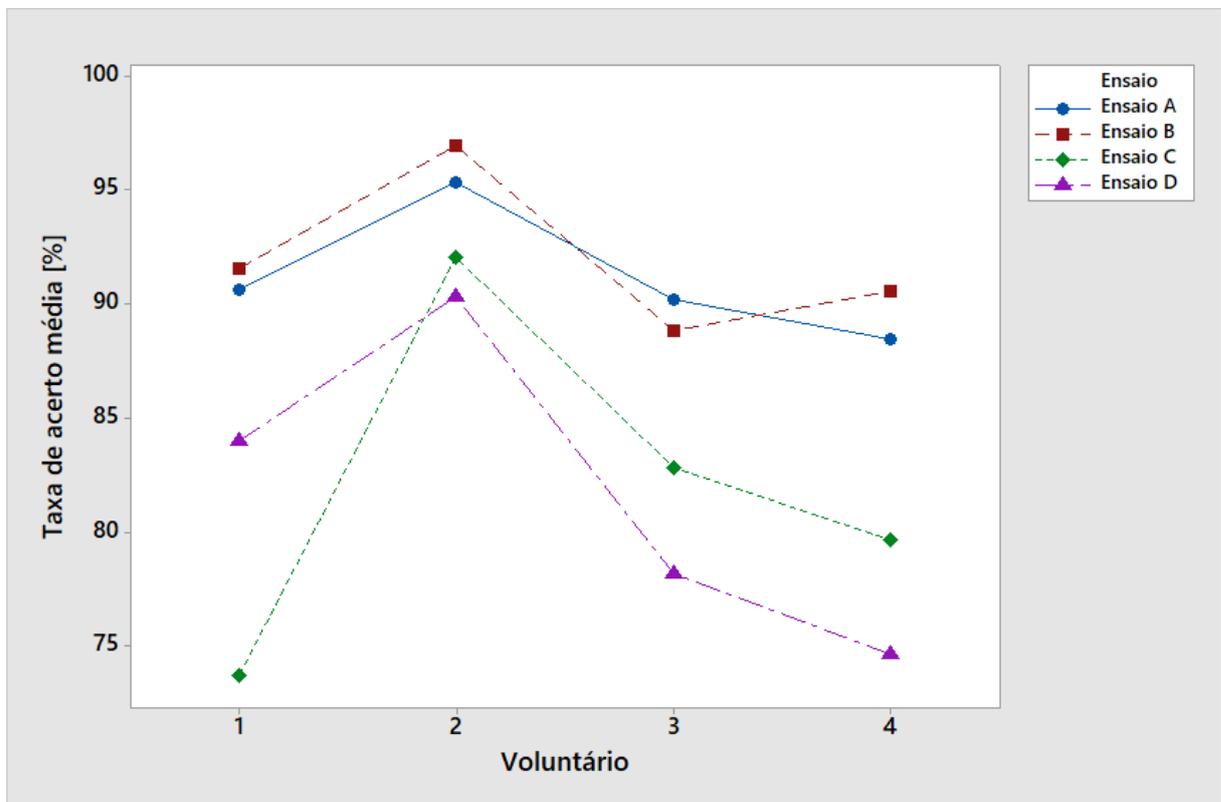


**Figura 48 Análise da interação entre movimento e tipo de ensaio.**

Como já mencionado, o tempo de duração dos exercícios difere de um tipo de ensaio para outro. Os ensaios que contêm mais repetições são, naturalmente, mais longos. Nestes exercícios, aumenta a probabilidade de o voluntário apresentar cansaço, tanto muscular como mental, o que pode afetar o modo como ele executa os movimentos. Tal fato pode ser considerado como uma das razões para a variabilidade constatada na taxa de acerto média dos movimentos em função do tipo de ensaio.

#### 4.2.9 Análise da Interação entre Tipo de Ensaio e Voluntário

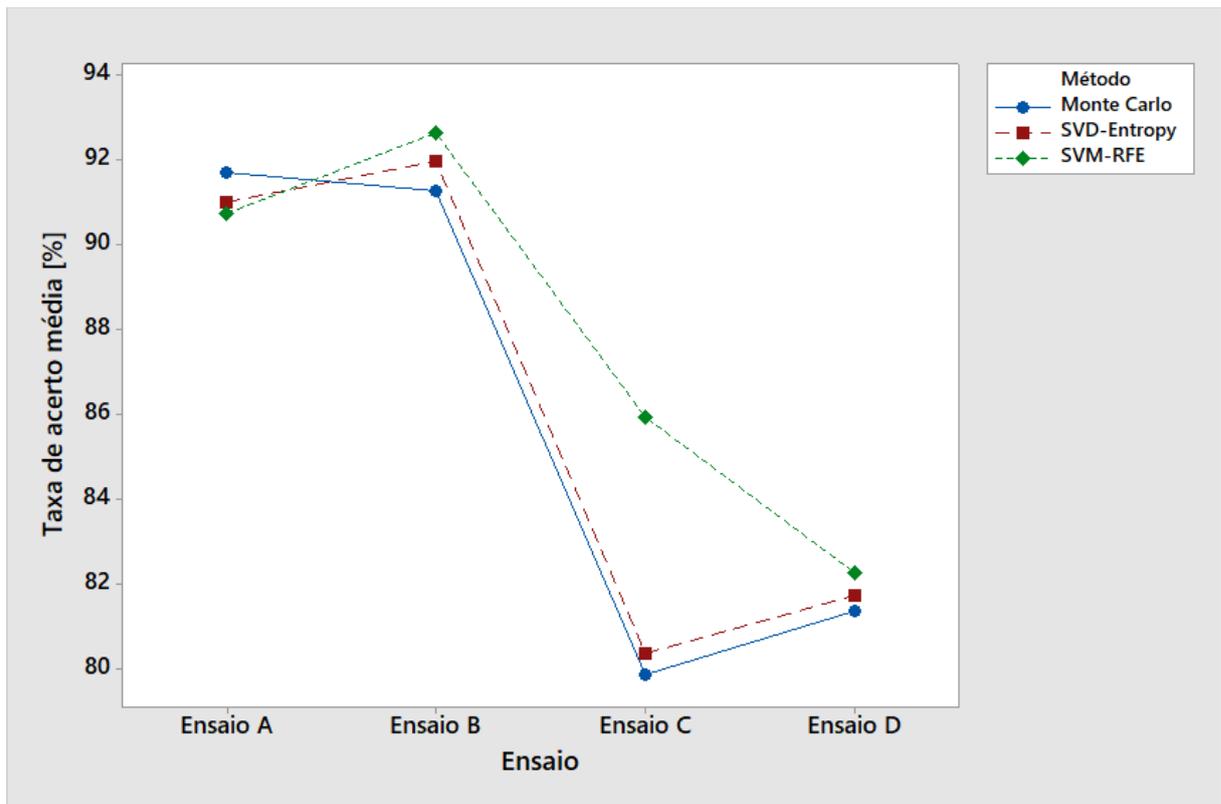
De acordo com os resultados da Tabela 4, a influência da interação entre tipo de ensaio e voluntário na taxa de acerto média foi constatada significativa, com um nível de confiança de 95%. A Figura 49 mostra o comportamento da taxa de acerto média em função dos voluntários para cada um dos ensaios realizados. A partir da análise desta figura, percebe-se comportamento análogo ao da Figura 48, ou seja, as curvas dos ensaios A e B estão separadas das dos ensaios C e D. Além disso, também aqui existem cruzamentos entre as curvas dos ensaios A e B e entre os ensaios C e D. Neste caso, a questão da duração dos ensaios também pode estar relacionada com a variabilidade observada na taxa de acerto média de cada sujeito em função do tipo de ensaio. Isto se explica pelo fato de o nível de cansaço apresentado durante os exercícios e, conseqüentemente, a influência deste fator no modo como o sujeito executa o movimento, varia de voluntário para voluntário.



**Figura 49** Análise da interação entre voluntário e tipo de ensaio.

#### 4.2.10 Análise da Interação Entre Tipo de Ensaio e Método de Seleção de Características

A Figura 50 apresenta a relação entre o fator controlável tipo de ensaio e a taxa de acerto média para os três métodos de seleção de características implementados. A influência da interação destes dois fatores controláveis na variável de resposta foi verificada significativa com nível de confiança de 95%, segundo resultados expostos na Tabela 4.



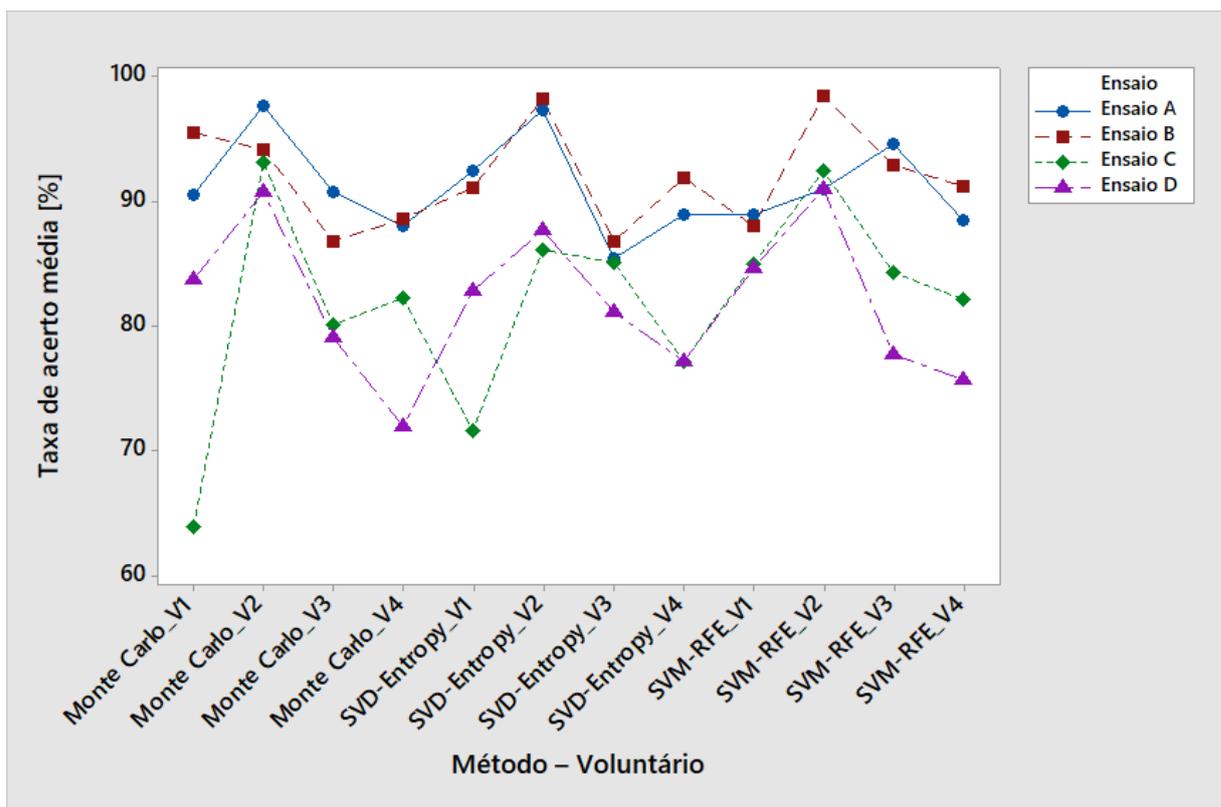
**Figura 50 Análise da interação entre método de seleção e tipo de ensaio.**

Da Figura 50 destaca-se o fato de que há apenas um ponto de cruzamento entre as três curvas. Em três dos ensaios (B, C e D) a ordem dos métodos de seleção em relação a taxa de acerto média obtida não se altera. Para o Ensaio A, o método SVM-RFE, que para os demais ensaios correspondeu à taxa de acerto mais alta, foi o responsável pelo valor percentual mais baixo com 90,7%. Em contrapartida, o algoritmo Monte Carlo foi o que resultou na maior taxa de acerto média para tal ensaio com 91,7%. Situação inversa observou-se nos demais ensaios, onde o algoritmo SVM-RFE apresentou as maiores taxas de acerto (92,6%, 85,9% e 82,3% para

os Ensaios B, C e D respectivamente) e Monte Carlo resultou nas menores (91,3%, 79,9% e 81,4% para os Ensaios B, C e D respectivamente). Tal resultado indica que, de acordo com o tipo de ensaio executado, existem métodos de seleção de características que proporcionam melhores taxas de acerto ao classificador, pelo menos para os ensaios e algoritmos de seleção analisados neste Projeto de Experimentos.

#### 4.2.11 Análise da Interação entre Voluntário, Método de Seleção e Tipo de Ensaio

A Figura 51 mostra a relação entre a taxa de acerto média para a interação entre voluntário, método e ensaio. Para melhor visualização dos resultados, foi plotada a curva da interação entre voluntário e método de seleção de características para cada um dos ensaios.



**Figura 51 Interação entre voluntário, método e tipo de ensaio.**

Analisando a Figura 51 nota-se que a interação entre os fatores controláveis voluntário, tipo de ensaio e método de seleção de características influenciaram a taxa de acerto média. Isso

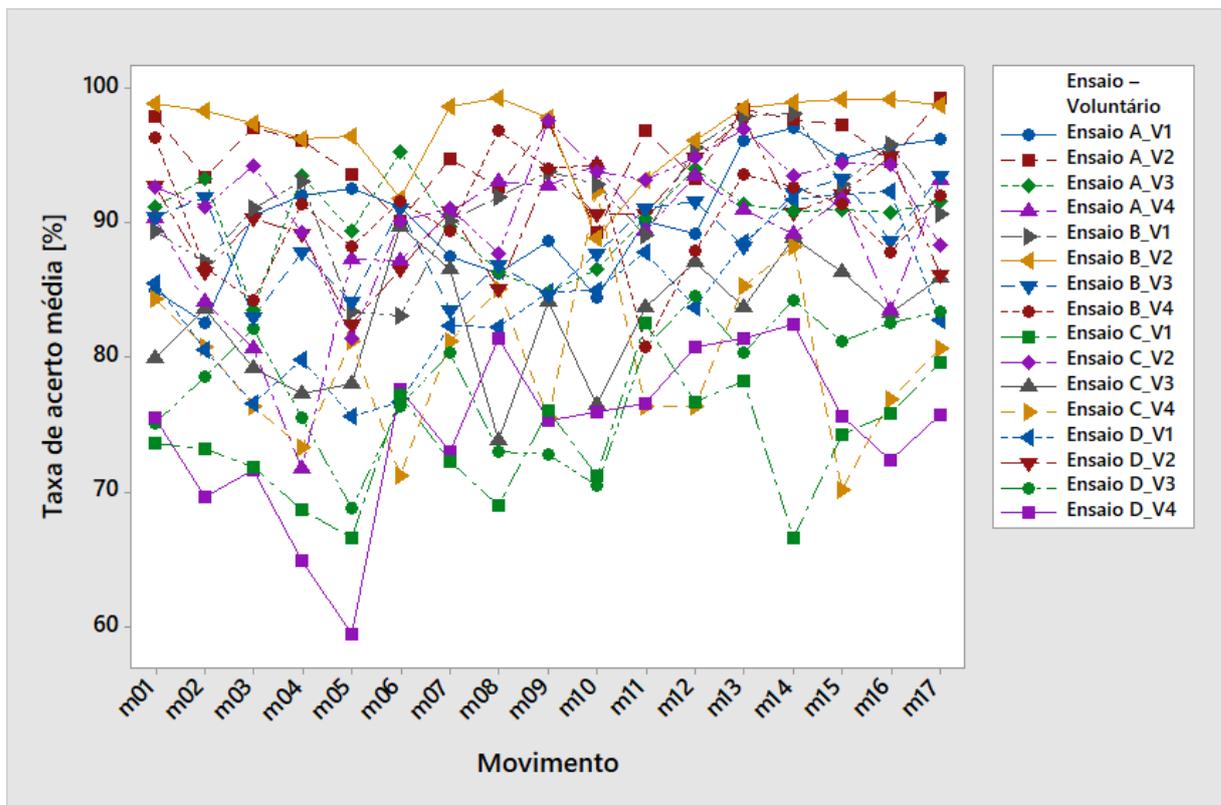
fica evidente, por exemplo, ao se analisar a curva referente ao Ensaio C. Ela registra a menor taxa de acerto média para o voluntário 1 empregando o método de seleção de características Monte Carlo. Porém, analisando-se ainda esta Figura 51, percebe-se que para o voluntário 2 e utilizando o método SVM-RFE, a taxa de acerto média obtida se eleva, ficando superior inclusive a do Ensaio A que, por sua vez, é o que possui as maiores taxas de acerto em boa parte das combinações método-voluntário.

Tal fato indica uma possível relação entre as características selecionadas pelos diferentes métodos de seleção empregados, com o padrão dos sinais de sEMG adquiridos para os movimentos executados. Além disso, deve-se considerar que a execução dos movimentos sofre a influência do tipo de ensaio que o voluntário está executando. Para ensaios mais demorados, o indivíduo pode apresentar sinais de cansaço, tanto físico quanto mental, o que pode alterar o modo como ele realiza os movimentos e, conseqüentemente, os sinais de sEMG adquiridos. Também convém ressaltar que a questão da aleatoriedade da apresentação dos movimentos nos ensaios C e D também contribuíram para alteração no padrão dos sinais de sEMG pelas razões já citadas na seção 4.2.3.

#### **4.2.12 Análise da Interação entre Voluntário, Movimento e Tipo de Ensaio**

A Figura 52 apresenta a relação entre a taxa de acerto média para a interação entre voluntário, método e ensaio. Aqui, foi adotada a mesma estratégia da seção 4.2.11 para a exposição da relação dos fatores controláveis. Deste modo, foi plotada a curva dos movimentos para cada combinação da interação entre voluntário e tipo de ensaio. De acordo como os resultados do Projeto de Experimentos expostos na Tabela 4, a interação entre os fatores controláveis movimento, voluntário e tipo de ensaio na taxa de acerto do classificador, demonstrou ser significativa, uma vez que o fator  $p$  encontrado foi de 0,0.

Analisando a Figura 52, percebe-se que as curvas correspondentes às combinações ensaio-voluntário se cruzam constantemente ao longo da variação dos movimentos, o que evidencia a influência de tais fatores na variável de resposta. Como já mencionado, cada ser humano apresenta aspectos fisiológicos individuais, o que explica a existência de movimentos que sejam mais acessíveis para alguns voluntários e de execução mais difícil para outros. Adicionando o fator tipo de ensaio, acrescenta-se para a análise a questão da duração do ensaio e a forma como os movimentos são apresentados ao voluntário. Como já é notório das seções anteriores, essas questões influenciaram significativamente a taxa de acerto média obtida. Deste modo, é perfeitamente compreensível que, para um dado movimento, se observem diferentes taxas de acerto média ao variar a combinação entre voluntário e tipo de ensaio, como está ilustrado na Figura 52.



**Figura 52** Interação entre voluntário, ensaio e movimento.

#### 4.2.13 Análise da Interação entre Voluntário, Método de Seleção e Movimento

A relação entre a taxa de acerto média e movimento para cada combinação de método de seleção e voluntário está representada na Figura 53. Foi constatada que a interação entre tais fatores controláveis influencia significativamente a taxa de acerto média, com nível de confiança de 95%. De fato, ao examinar a Figura 53, verifica-se uma série de cruzamentos entre as curvas de cada uma das combinações de método de seleção e voluntário, o que justifica o resultado obtido do Projeto de Experimentos. Logo, conclui-se que a escolha do algoritmo de seleção de características mais adequado depende do voluntário e dos movimentos realizados. Tal determinação é válida entre os sujeitos participantes da Base de Dados do IEE e em relação aos movimentos e métodos analisados neste trabalho.

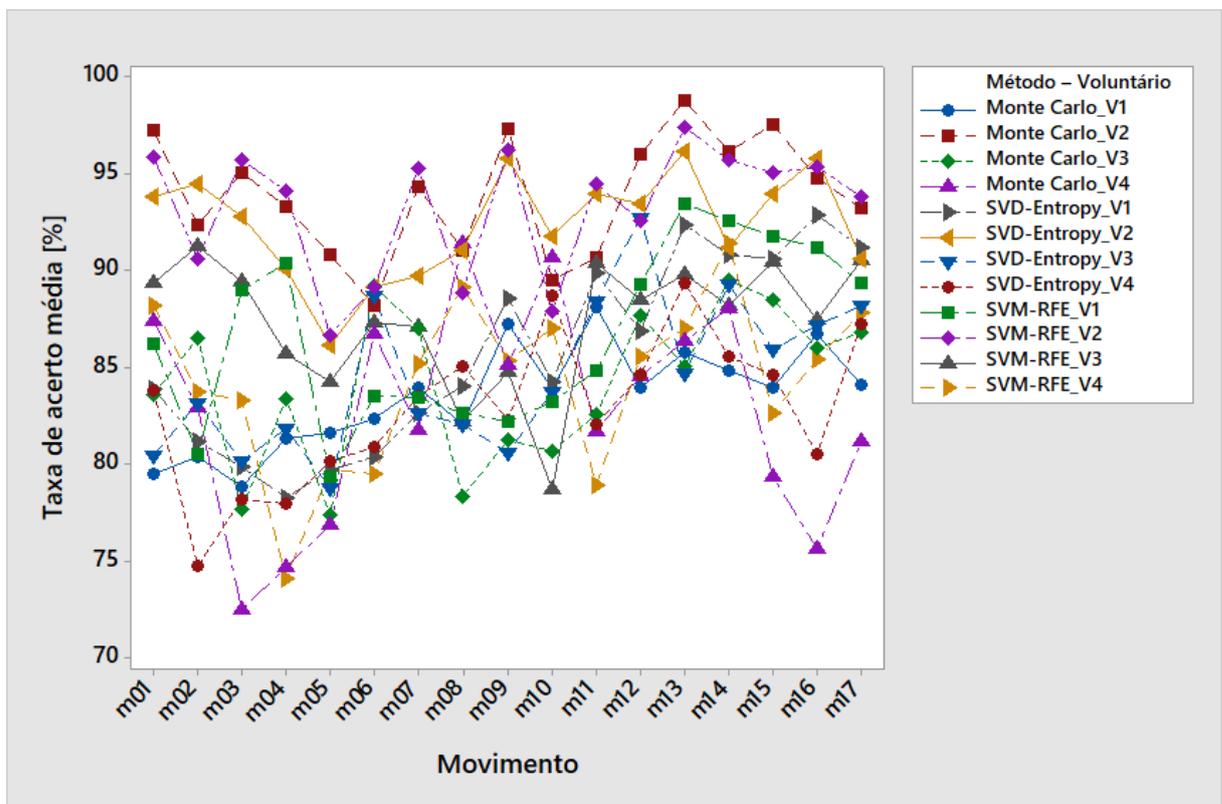


Figura 53 Análise da interação entre método de seleção, voluntário e movimento.

#### 4.2.14 Análise da Interação entre Tipo de Ensaio, Método de Seleção e Movimento

A Figura 54 mostra a relação entre movimento e a taxa de acerto média para todas as combinações de métodos de seleção e tipo de ensaio. De acordo com os resultados descritos na Tabela 4, a influência da interação entre tais fatores controláveis e a taxa de acerto média do classificador é significativa com nível de confiança de 95%. Contudo, o que se conclui dessa análise é que, a escolha do método de seleção de características que proporcionará ao classificador taxa de acerto média mais elevada dentre os três analisados neste trabalho, dependerá simultaneamente do tipo de ensaio (A, B, C ou D) e do movimento considerado (m01 a m17).

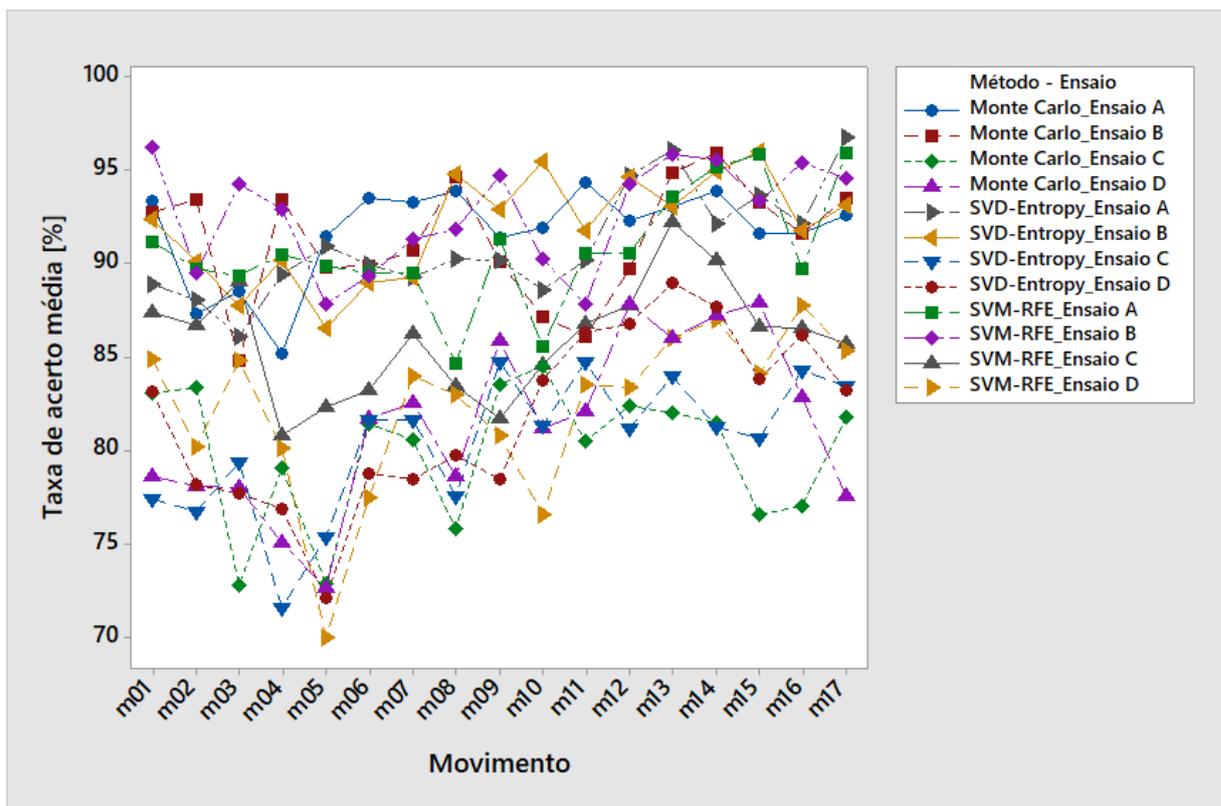


Figura 54 Análise da interação entre método de seleção, tipo de ensaio e movimento.

### 4.2.15 Análise da Interação entre Voluntário, Movimento, Método de Seleção e Tipo de Ensaio

A Figura 55 mostra um conjunto de gráficos relacionando todos os quatro fatores controláveis que apresentaram interferência significativa na taxa de acerto média da correspondente base de dados avaliada neste subcapítulo 4.2.

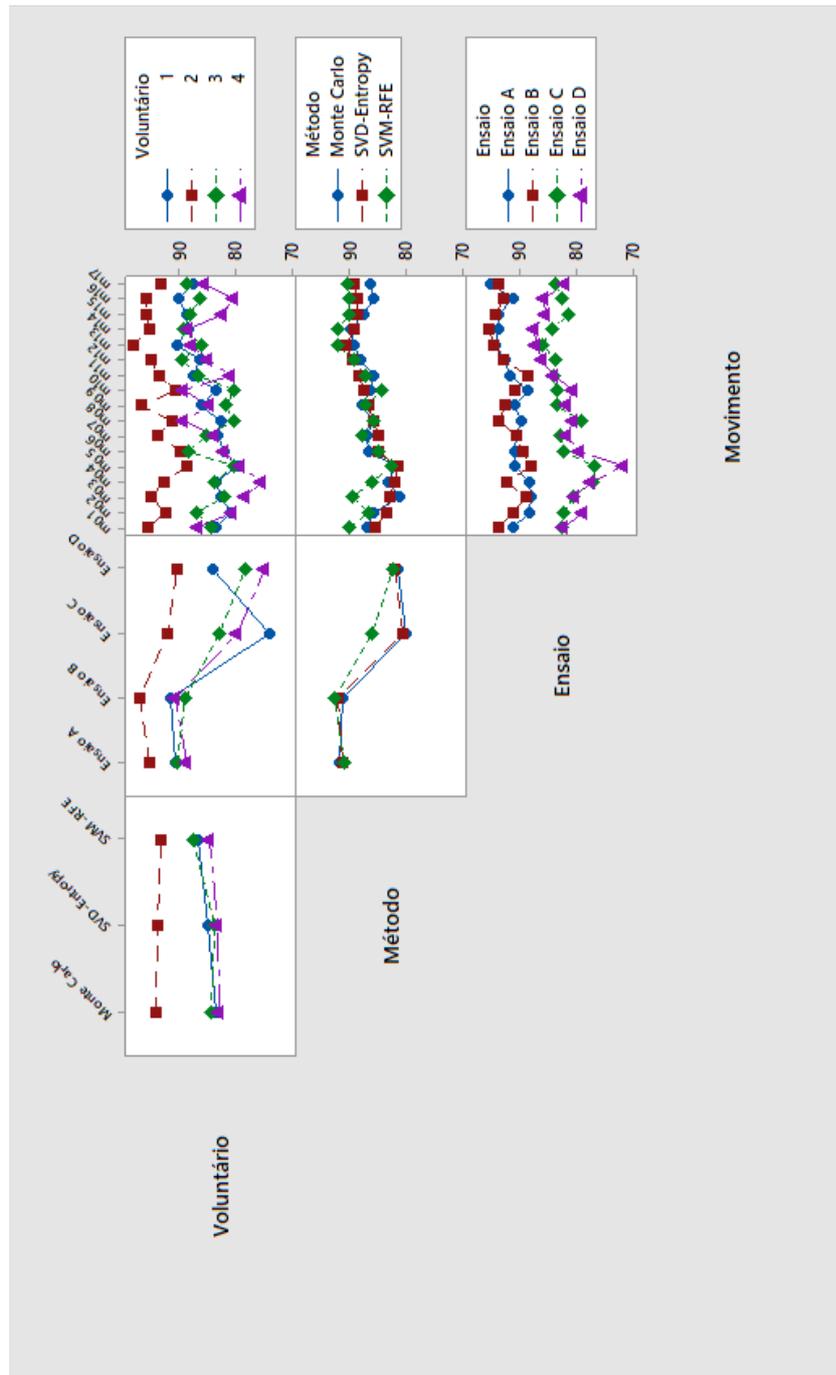


Figura 55 Interação entre voluntário, método, movimento e ensaio.

De acordo com a Tabela 4, a interação entre os quatro fatores controláveis analisados (voluntário, método, ensaio e movimento), influenciaram significativamente a taxa de acerto média, uma vez que fator  $p$  é menor do que 0,05. Uma vez que foi verificado que a interação dos fatores controláveis ensaio, método e voluntário impactaram na variável de reposta, é natural que se observe variação na taxa de acerto média ao se considerar também o fator movimento. Isso significa que, não está descartada a hipótese de que ocorra variabilidade na taxa de acerto média ao se alterar os quatro fatores simultaneamente.

#### **4.3 RESULTADOS DO PROJETO DE EXPERIMENTOS COM AS BASES NINAPRO E ENSAIO A DA BASE DO IEE**

Da mesma forma e pelas mesmas razões que o executado para o Projeto de Experimentos com a Base de Dados do IEE, todas as análises aqui expostas são baseadas na taxa de acerto média obtida para cada fator controlável e suas combinações. No decorrer deste capítulo serão abordados os resultados da influência de cada um dos fatores controláveis (movimento, método de seleção e base de dados), e também da interação de todas as combinações entre eles, na taxa de acerto média obtida. A Tabela 5 apresenta os resultados obtidos do correspondente Projeto de Experimentos com as Bases Ninapro e Ensaio A da Base do IEE com um intervalo de confiança de 95%.

Da Tabela 5 conclui-se que, dentre todos os fatores controláveis e das interações entre eles, apenas os fatores base de dados e movimento e a interação entre base e movimento apresentaram influência significativa na taxa de acerto média. Para todos os outros efeitos analisados, não foi observada significância nos resultados, visto que o valor do fator  $p$  superou a marca dos 5 por cento. Deste modo, o produto das interações entre método e movimento, método e base de dados, e entre os três fatores controláveis, não será abordado, pois não alteram de forma significativa (com um erro de 5%) a taxa de acerto média. Como no atual Projeto de

Experimentos não foi considerado voluntário como fator controlável, a análise da taxa de acerto média por indivíduo nas bases do Ninapro se dará em uma seção específica na sequência desta.

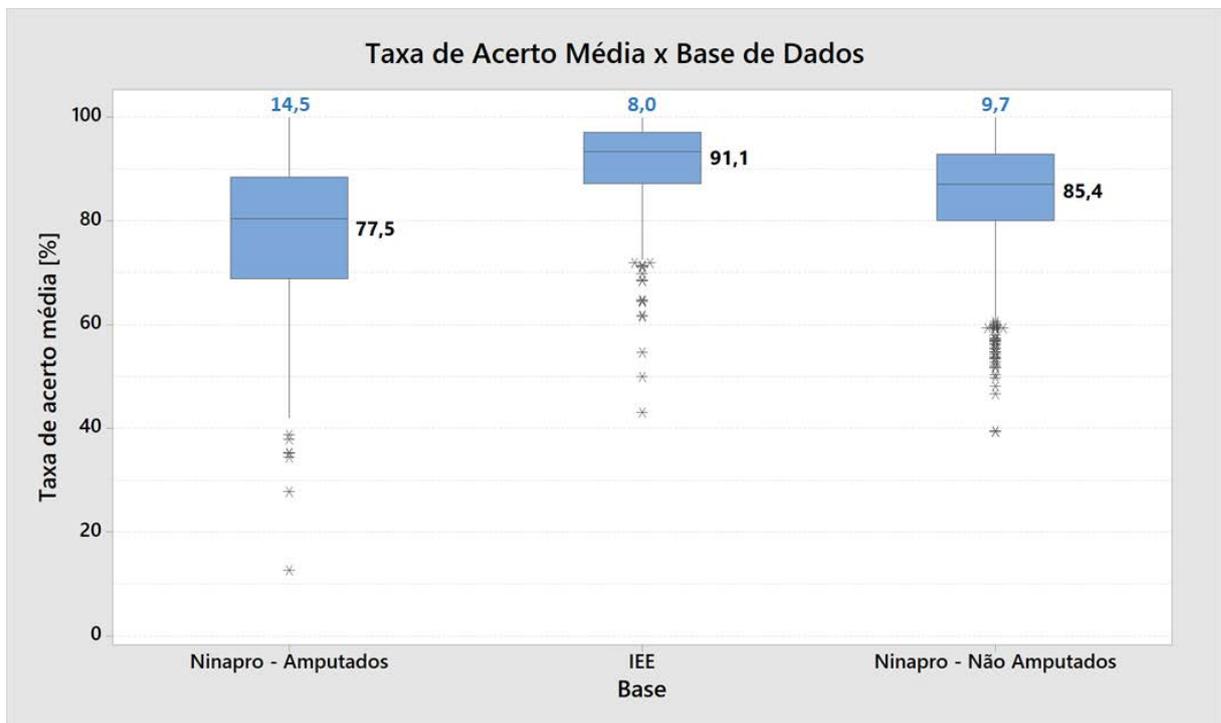
**Tabela 5 Análise de Variância do Projeto com as Bases Ninapro e Ensaio A da Base do IEE.**

Fonte de Variação	Soma dos Quadrados	Graus de Liberdade	Médias Quadradas	Fator F	Fator P	Resultado
Método de Seleção (fator A)	81	2	40,5	0,42	0,66	Não significativo
Base de Dados (fator B)	52264	2	26132,2	268,38	0,00	Significativo
Movimento (fator C)	14513	16	907,1	9,32	0,00	Significativo
Interação AB	130	4	32,5	0,33	0,86	Não significativo
Interação AC	1385	32	43,3	0,44	1,00	Não significativo
Interação BC	15324	32	478,9	4,92	0,00	Significativo
Interação ABC	2669	64	41,7	0,43	1,00	Não significativo
Erro	292984	3009	97,4			
Total	391046	3161				

De acordo com os resultados da análise de variância expostos na Tabela 5, não foi verificada significância na diferença observada entre as taxas de acerto média dos três métodos de seleção de características implementados, uma vez que o valor de  $p$  é 0,83, ou seja, maior do que 0,05. Este resultado confronta o obtido na análise da Base de Dados do IEE, realizada na seção 4.2.4. Logo, para os dados das Bases Ninapro não foi verificada significativa a influência do método de seleção de características. Isso levando em conta unicamente as taxas de acerto médias resultantes. Embora não possa ser descartada a consideração de que a inclusão de uma etapa de seleção de características no processo de classificação tenha contribuído de forma positiva para a taxa de acerto média obtida, como será discutido na seção 4.6 onde os resultados deste trabalho serão comparados com o de outras pesquisas na área.

### 4.3.1 Análise do Fator Controlável Base de Dados

A Figura 56 apresenta a média (em preto) e o desvio padrão (em azul) das taxas de acerto obtidas de cada uma das bases de dados analisadas. Da verificação dos resultados desta Tabela, destaca-se o fato de que o desvio padrão da taxa de acerto obtida com a Base Ninapro formada por dados de ensaios com voluntários amputados, foi consideravelmente maior do que o das outras duas, apresentando valor igual a 14,5% enquanto que as demais obtiveram desvio padrão de 9,7% (Ninapro – não amputados) e 8,0% (IEE). Essa superioridade neste resultado pode ser explicada pela existência de maior variabilidade entre os perfis de voluntários desta base em relação às demais, uma vez que os sujeitos apresentam diferenças nos níveis de amputação e tempo de perda do membro, além do fato de alguns já usarem prótese e outros não.



**Figura 56** Dispersão da taxa de acerto obtida para cada base de dados.

De acordo com a Tabela 5, o efeito do fator controlável base de dados na taxa de acerto média é estatisticamente significativo, visto que o fator  $p$  é igual a 0,00. A Figura 56 mostra a

relação entre tal fator e a variável de resposta. A partir da análise do referido gráfico, observa-se uma considerável variabilidade na taxa de acerto média quando se analisa as diferentes bases de dados. A Base de Dados do IEE foi a que obteve o maior resultado com 91,1%. Já a menor taxa de acerto média (77,5%) ocorreu quando se considerou a base do Ninapro que contém sinais adquiridos de voluntários com amputação transradial.

As Bases do Ninapro e a do IEE possuem diferenças em vários aspectos. Por mais que no presente trabalho procurou-se replicar mais fidedignamente possível o processo de aquisição implementado pela equipe do Ninapro, vários pontos acabaram divergindo entre os dois. Apenas para citar um exemplo, a equipe do Ninapro utilizou eletrodos *wireless* enquanto que na aquisição executada no Laboratório do IEE a conexão entre eletrodos e eletromiógrafo foi realizada via cabo. Além disso, o desconhecimento de informações a respeito da condução dos experimentos pela equipe do Ninapro, tais quais, o modo como os voluntários eram orientados a realizar os movimentos, também constituem fontes de diversidade entre as bases. Assim, era de se esperar certa variabilidade nos padrões dos sinais de sEMG e, conseqüentemente, nas taxas de acerto obtidas em ambas, explicando assim, a diferença de 5,7% nos resultados alcançados através da classificação dos sinais adquiridos no Laboratório do IEE e os dos voluntários não amputados do Ninapro. Além disso as três bases envolvem voluntários diferentes.

Além disso, entre as bases contendo sinais de indivíduos não amputados e a formada por voluntários com amputação, é compreensível que se observe diferença nas taxas de acerto média. Isso porque deve-se levar em conta que a aquisição de sinais de sEMG neste último grupo é mais difícil de se executar e não pode ser realizada da mesma forma do que no primeiro. O número de eletrodos e sua disposição dependem do grau da amputação e das condições fisiológicas dos músculos remanescentes no coto. Como por exemplo, em um dos voluntários não foi possível a utilização dos 12 eletrodos, sendo fixados somente 10. Além disso, a

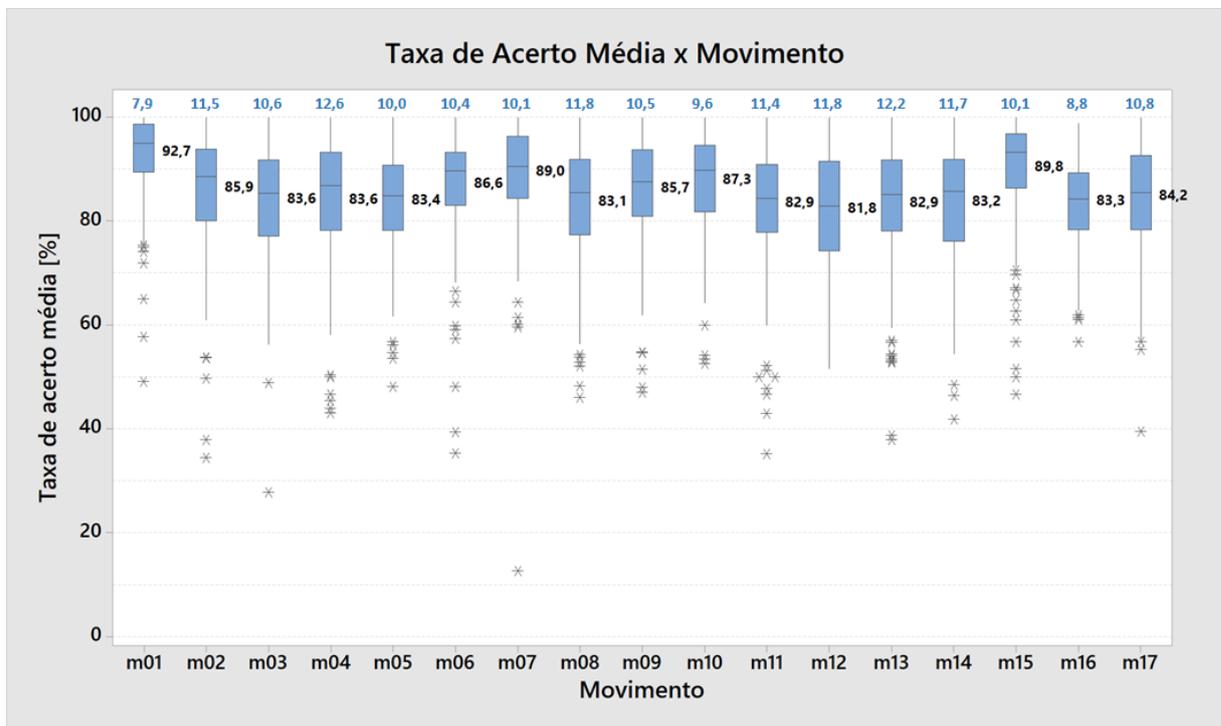
qualidade do sinal adquirido também decai consideravelmente quando se refere à pacientes sem a totalidade dos membros superiores, seja devido às limitações musculares seja pelo fato do sinal ser adquirido a partir de uma intenção de movimento apenas. Além disso, é importante ressaltar um dos principais pontos negativos na geração da Base Ninapro com relação aos amputados. Nestes voluntários outros músculos poderiam ter sido utilizados, como por exemplo, o trapézio que poderia auxiliar na melhoria da qualidade do sinal de sEMG e, por consequência, na taxa de acerto de alguns dos movimentos. Esta crítica é um dos motivos para a geração da Base de Dados do IEE que, evidentemente, necessita do aumento do número de voluntários, assim como, de voluntários com vários graus de amputação. Além disso, será possível avaliar a relação dos eletrodos posicionados de forma aleatória na correspondente taxa de acerto de determinados movimentos, ou seja, a relação entre posicionamentos aleatórios e posicionamento na região ventral de determinados músculos.

Deste modo, o resultado do Projeto de Experimentos concernente ao impacto do fator base de dados na taxa de acerto média é explicável através das razões supracitadas. Além disso, tal conclusão também foi corroborada na pesquisa de (CENE, 2016), onde análise similar foi realizada, inclusive com as mesmas bases do Ninapro utilizadas no presente trabalho.

#### **4.3.2 Análise do Fator Controlável Movimento**

Observando os resultados da análise das variâncias expostos na Tabela 5, a hipótese de que o fator controlável movimento influencia significativamente a taxa de acerto média do classificador não foi descartada, uma vez que o valor de  $p$  é 0,00. Tal resultado confirma o que foi discutido na seção 4.2 onde o fator movimento também mostrou-se influenciar significativamente a taxa de acerto média no Projeto de Experimentos com a Base de Dados do IEE.

O valor médio e o desvio padrão da taxa de acerto obtida para cada movimento considerado, estão descritos (em preto e em azul respectivamente) na Figura 57. Analisando os dados desta Figura, observa-se a ocorrência de maior variabilidade na caracterização do movimento m13, resultando em valor de desvio padrão de 12,2%. Em contrapartida, as menores dispersões em torno da taxa de acerto média foi verificada para os movimentos m01 e m16, com valores de desvio padrão de 7,9% e 8,8% respectivamente.



**Figura 57** Relação entre o fator controlável movimento com a taxa de acerto média.

A partir da Figura 57, observa-se uma variação máxima de 10,9% entre as taxas de acerto média de cada um dos movimentos. Os movimentos com maior taxa de acerto média foram o m01, m07 e m15 (92,7%, 89,0% e 89,8% respectivamente). Em contrapartida, os movimentos que apresentaram as menores taxas de acerto média foram o m11, m12 e o m13 (82,9%, 81,8% e 82,9% respectivamente). Convém destacar que os dois últimos estavam entre os que apresentaram os melhores resultados quando se analisou a Base do IEE. Esse fato

evidencia que o fator base de dados influencia significativamente as taxas de acerto média obtidas. Por conseguinte, tal diferença nos resultados pode ser explicada por distinções entre os protocolos de aquisição e também pelas características dos voluntários formadores das três bases consideradas. Logo, a análise das taxas de acerto média em relação a cada um dos movimentos em si fica vinculada à base de dados a qual se está considerando. Porém, o que se pode concluir da Figura 57 é que os movimentos m01, m07 e m15 são os que foram melhor caracterizados, independentemente da base de dados. Já os movimentos m12 e o m13 apresentaram padrões menos discriminantes nas Bases do Ninapro em relação a realizada na Base do IEE, possivelmente devido a divergências nos processos de aquisição e nas características dos voluntários recrutados em cada uma delas. Contudo, deve-se levar em conta que nesta análise considerou-se apenas os ensaios referentes ao Ensaio A da Base do IEE.

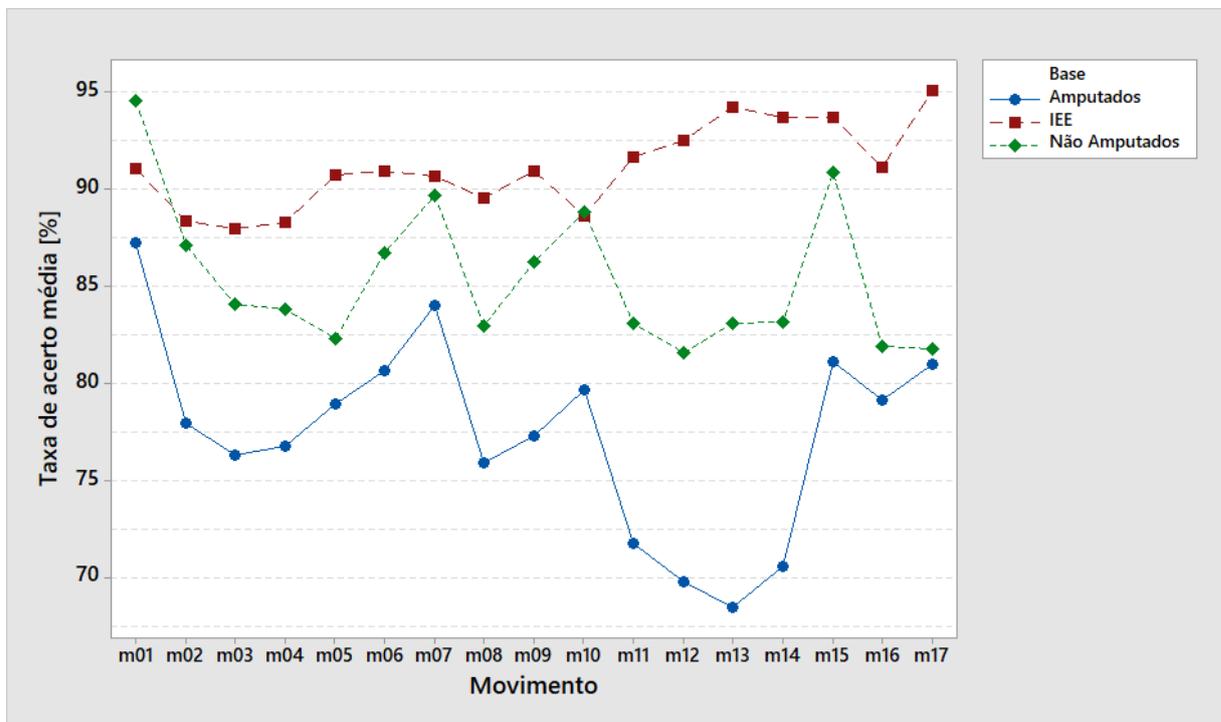
#### **4.3.3 Análise da Interação entre Base de Dados e Movimento**

Segundo exposto na Tabela 5, o efeito da interação entre os fatores controláveis base de dados e movimento na taxa de acerto média é estatisticamente significativo, uma vez que o fator  $p$  é igual a 0,00. O referido gráfico mostra a relação entre a referida interação e a variável de resposta. A partir da análise da Figura 58, nota-se que, como esperado, a base de dados do Ninapro que contém sinais de sEMG adquiridos de voluntários com amputação transradial, foi a que apresentou as menores taxas de acerto para todos os movimentos. Em relação às outras duas, observa-se que a curva de ambas se cruzam em vários pontos, indicando uma relação de dependência entre a variável de resposta e a interação entre base e movimentos.

Analisando as três curvas, observa-se que as que se referem às Bases Ninapro apresentam o mesmo padrão em relação a taxa de acerto média associada a cada movimento. Logo, a diferença se encontra na comparação entre as Bases Ninapro e a do IEE. Em relação a esta última, as maiores taxas de acerto média foram obtidas para os movimentos m13, m14,

m15 e m17 (94,2%, 93,7%, 93,7% e 95,1 respectivamente), enquanto que, para as primeiras, resultaram dos movimentos m01, m07 e m15 (87,2%, 84,0% e 81,1% respectivamente para a base de amputados; 94,6%, 89,7% e 90,9% respectivamente para a base de não amputados).

Tal diferença entre bases fica evidente ao se considerar que os movimentos m13 e m17, os quais estão entre os que apresentaram as maiores taxas de acerto para a Base do IEE, são os que obtiveram as menores taxas para as Bases do Ninapro formadas por voluntários com e sem amputação respectivamente. Esse fato, por sua vez, pode ser explicado por possíveis diferenças nos protocolos de aquisição utilizados na formação de ambas as Bases como por exemplo, a forma como os voluntários foram instruídos para a realização dos movimentos, tanto na questão da dinâmica quanto na da força aplicada durante sua execução. Desta forma, diferenças nos padrões de sinal para os mesmos movimentos podem ter sido obtidas entre as Bases Ninapro e a do IEE.



**Figura 58** Relação entre a interação entre base de dados e movimento com a taxa de acerto média.

#### **4.4 ANÁLISE DA TAXA DE ACERTO MÉDIA POR VOLUNTÁRIO PARA AS BASES NINAPRO**

A Figura 59 mostra os gráficos com a taxa de acerto média por voluntário das Bases Ninapro com sinais oriundos de indivíduos com os membros intactos e com amputação. Analisando o referido gráfico, observa-se uma variabilidade de 19,1% nas taxas de acerto média máxima e mínima em relação aos voluntários sem amputação. Essa diferença no desempenho do classificador entre indivíduos pode ser explicada através dos mesmos fatores explicitados na seção 4.2.2. Como já mencionado, o ser humano possui atributos fisiológicos particulares, que podem vir a facilitar ou dificultar a realização dos movimentos, bem como, influenciar na qualidade do sinal adquirido. Além disso, os indivíduos participantes do experimento podem apresentar fadiga muscular durante os ensaios, dependendo do condicionamento físico de cada um, o que impacta consideravelmente a aquisição de sinais de sEMG.

Outro fator que está associado diretamente ao resultado de classificação é a questão do treinamento aplicado aos voluntários antes de iniciar a aquisição. Quanto maior o tempo de duração e o aproveitamento deste período por parte do sujeito para a resolução de eventuais dúvidas e instruções adicionais, mais familiarizado com os movimentos ele ficará e, por conseguinte, melhor será a reprodução dos movimentos no momento da realização do ensaio. Deste modo, reduz-se as chances de o voluntário cometer equívocos tais como, confusão de movimentos, esquecimento, execução na hora errada, dentre outros.

Já considerando o gráfico referente aos dados dos voluntários amputados da Figura 59, nota-se uma variabilidade na taxa de acerto média de 28,5%, ou seja, maior do que a observada entre os sujeitos intactos. Neste caso, diferentemente de quando se analisa a base formada por indivíduos intactos, existe a influência de fatores adicionais nos resultados de classificação. Como exemplo pode-se citar: tempo em que ocorreu a amputação, grau da amputação (altura do membro em que foi realizada), uso de prótese, tempo de uso e tipo de prótese, aspectos

fisiológicos da musculatura remanescente no coto, experiência de uso do sistema proposto de ensaios, dentre outras. Deste modo, é natural se obter maior variabilidade na taxa de acerto entre os voluntários com amputação transradial.

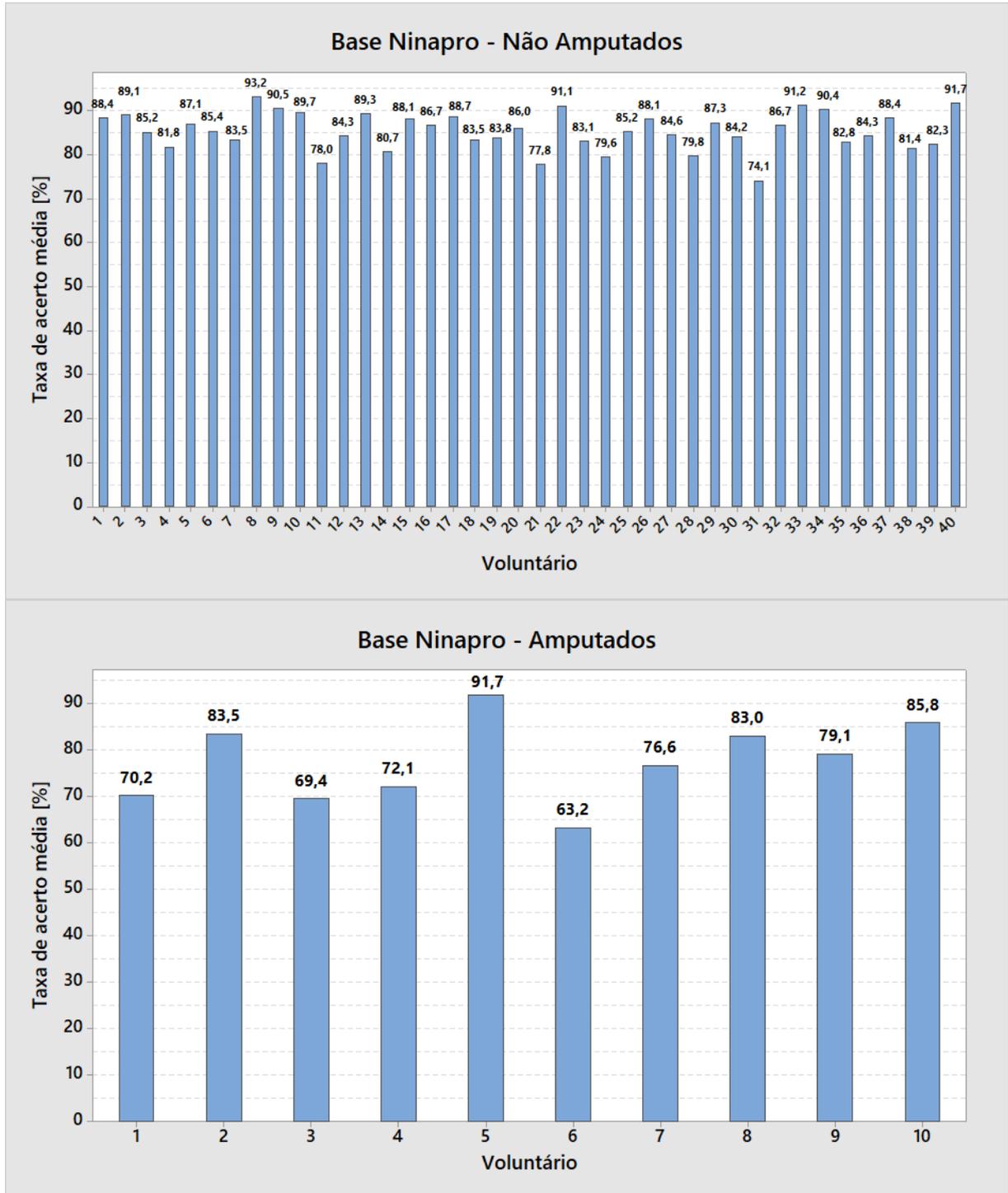


Figura 59 Relação entre voluntário e a taxa de acerto média para as Bases Ninapro.

Ainda considerando o gráfico referente aos dados dos voluntários amputados da Figura 59, a maior taxa de acerto obtida foi de 91,7% com o voluntário 05. Esse resultado é interessante, visto que, ele está entre os melhores, inclusive, da base de dados dos indivíduos com membros intactos, sendo a segunda maior taxa. Todavia, vale ressaltar que tal sujeito teve apenas 10% do membro retirado enquanto que o voluntário 03, por exemplo, que é o que possui a segunda menor taxa de acerto, sofreu amputação de 70% do membro superior. Logo, a musculatura remanescente no primeiro é maior, possibilitando a fixação dos eletrodos em locais mais próximos possíveis da disposição original implementada em voluntários sem amputação, e, conseqüentemente, a obtenção de padrões de movimentos mais próximas das conseguidas em voluntários intactos.

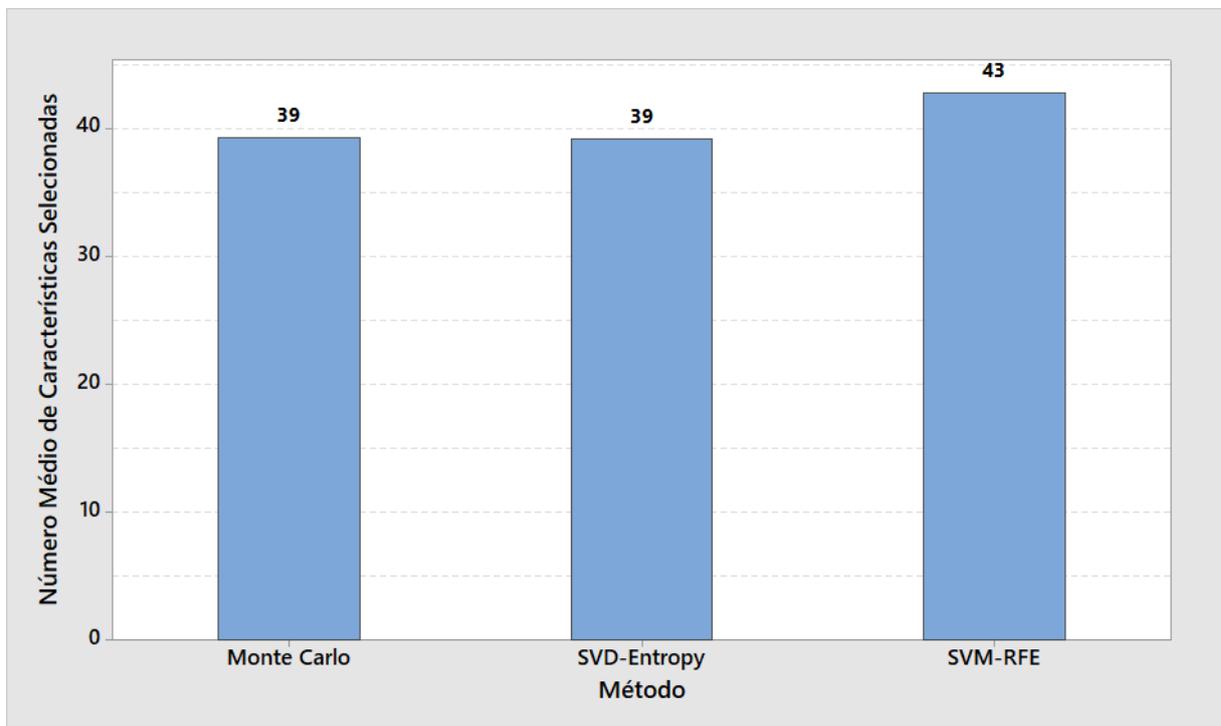
Além disso, o voluntário 05 é usuário de prótese cinemática, ou seja, ele está mais habituado a executar certos movimentos do que indivíduos que não utilizam prótese, ou só a usam para fins estéticos. Os voluntários 02, 08 e 10 apresentaram resultados interessantes, ficando próximos ao valor médio da base dos indivíduos intactos. Contudo, ao analisar os dados dos referidos sujeitos, verifica-se que o 08 e o 10 são usuários de prótese com controle mioelétrico. Para um usuário utilizar uma prótese desta natureza, é necessário que ele seja submetido por um considerável período de fisioterapia e adaptação. Logo, a musculatura remanescente nestes indivíduos permanece ativa, possibilitando a aquisição de sinais de sEMG de qualidade superior aos que não se enquadram nesta condição.

Por outro lado, ainda analisando a Figura 59, observa-se que os voluntários 03 e 06 foram os que apresentaram os menores resultados de classificação. Entretanto, deve-se levar em consideração que estes são os sujeitos que apresentam o maior grau de amputação dentre todos os 10. O voluntário 03 teve 70% do membro superior direito retirado e o 06 sofreu amputação de 60% do seu membro superior esquerdo. Devido a isso, a alocação dos eletrodos nestes indivíduos foi prejudicada, impactando nos padrões de sinal de sEMG obtidos para cada

movimento. Inclusive, para o 06 não foi possível a fixação de todos os 12 eletrodos, sendo utilizados somente 10.

#### 4.5 ANÁLISE DO NÚMERO DE CARACTERÍSTICAS SELECIONADAS

A Figura 60 mostra a relação entre o número médio de características selecionadas por cada um dos algoritmos implementados. Tal valor consiste na média global de todos os ensaios realizados por todos os voluntários pertencentes as três bases analisadas. Para a verificação da significância da diferença observada entre os três métodos, foi realizado novo Projeto de Experimentos. Nesse caso, considerou-se apenas o fator controlável método de seleção, sendo a variável de resposta o número médio de características selecionadas. Como resultado, obteve-se um fator  $p$  igual a 0,00. Logo, verificou-se a influência do algoritmo de seleção no número de características selecionadas, com nível de confiança de 95%.



**Figura 60** Relação entre o método de seleção de características e o número médio de características selecionadas.

Analisando a Figura 60, observa-se que o algoritmo SVM-RFE resultou em um valor médio maior de características selecionadas por ensaio em comparação aos outros dois (43 características por ensaio, enquanto que os métodos Monte Carlo e *SVD Entropy* selecionaram 39 características). Isso significa uma redução de 9,3% no número de atributos do conjunto de dados. Tal fato explicita pequena vantagem no uso de tais métodos em relação ao SVM-RFE, visto que não foi identificada diferença significativa entre as taxas de acerto média dos três algoritmos quando se analisou as Bases do Ninapro conjuntamente ao Ensaio A da Base do IEE. Os algoritmos Monte Carlo e *SVD Entropy* alcançam resultados similares em termos de taxa de acerto ao SVM-RFE, porém a partir de um banco de dados caracterizado por um número menor de atributos.

Por consequência, a tarefa do classificador, neste caso, fica menos complexa e, conseqüentemente, diminui o tempo de processamento. Considerando-se que tempo de processamento também está associado a desempenho, pode-se dizer que os algoritmos Monte Carlo e *SVD Entropy* possibilitam desempenho superior do classificador do que o SVM-RFE, porém obtendo a mesma taxa de acerto média, isso para as bases de dados analisadas no presente trabalho. Todavia, considerando-se unicamente a Base de Dados do IEE e seus 4 tipos de Ensaio, o algoritmo SVM-RFE apresentou superioridade em relação aos demais no concernente a taxa de acerto média do classificador. Logo, para os ensaios da referida base, a afirmação de qual algoritmo de seleção resulta em melhor desempenho do classificador deve ser tomada pesando-se tempo de processamento e taxa de acerto obtida.

#### **4.6 COMPARAÇÃO DOS RESULTADOS COM TRABALHOS RELACIONADOS**

Ao longo do desenvolvimento deste estudo, uma série de trabalhos na área de caracterização de sinais de sEMG para a classificação de movimentos do segmento mão-braço

foram estudados. A Tabela 6 apresenta uma seleção das pesquisas mais relevantes e que agregaram de forma mais significativa para a realização do presente trabalho.

**Tabela 6 Trabalhos relacionados à classificação de movimentos a partir de sinais de sEMG.**

Autor	Classificador	Características	Algoritmo de seleção de características	Número de Movimentos	Número de Voluntários	Taxa de Acerto [%]
(ZHAI <i>et al.</i> , 2016)	SVM	Espectrograma do sinal	-	49	40	75,7
(RIILLO <i>et al.</i> , 2014)	Redes Neurais	RMS, Amplitude de <i>Willison</i> , Média	-	5	21	87,3 86,6
(ABDELMASEEH ; CHEN; STASHUK, 2016)	Mínima distância entre trajetórias	Distância entre trajetórias	-	40	40	89,0
(ZHANG <i>et al.</i> , 2017)	LDA	VMA, CZ, CFO, MIS	-	6	10	96,3
(LI <i>et al.</i> , 2017)	LDA	VMA, CZ, CFO, MIS	-	4	4	72,0 91,7
(CENE, 2016)	Regressão Logística	RMS, Frequência mediana, Variância	-	17	40 não amputados 10 amputados	74,2 61,9
(YAN; LIU, 2013)	SVM	Energia dos subespaços de <i>Wavelet</i>	<i>Sequencial Forward Selection, Backward Elimination, PCA, RST</i>	4	30	95,0
Este trabalho	RELM	13 no total conforme seção 3.2	SVM-RFE, Monte Carlo, SVD <i>Entropy</i>	17	40 não amputados 10 amputados	85,4 77,5

Na pesquisa de (ZHAI *et al.*, 2016), foi abordada a questão da caracterização baseada em espectrograma de sinais de sEMG para a classificação de movimentos do segmento mão-braço. O trabalho consistiu em um estudo comparativo entre tal caracterização e as puramente temporais. Nesse caso, como característica do domínio do tempo foi utilizado o valor RMS do

signal e, para a tarefa de classificação, o algoritmo SVM. A análise do espectrograma se deu através da aplicação da Transformada rápida de Fourier (*Fast Fourier Transform* – FFT) em janelas temporais do sinal de eletromiografia e, para redução de dimensionalidade, selecionou-se somente os 100 componentes principais por meio da implementação do algoritmo *Principal Component Analysis* (PCA). Como resultado, foi obtida maior taxa de acerto quando se aplicou características baseadas no espectrograma dos sinais de sEMG do que quando se considerou somente o valor RMS, destacando a propriedade que as características extraídas no domínio da frequência possuem, de discriminar padrões de movimentos do segmento mão-braço. Analisando o resultado de classificação, a taxa de acerto média alcançada por movimento por voluntário da Base 2 Ninapro, foi de 75,7%. Esse valor é consideravelmente inferior ao atingido no presente trabalho (85,4%).

Já o trabalho de (RIILLO *et al.*, 2014) está direcionado para a etapa de pré-processamento do sinal de eletromiografia, com o objetivo de aumentar a capacidade de discriminação das características extraídas. Sendo assim, dois algoritmos foram avaliados e comparados entre si. Um deles é o *Common Spatial Pattern* (CSP), que consiste em um algoritmo supervisionado o qual maximiza a variância de uma classe específica e minimiza para as demais através da aplicação de filtros espaciais na matriz de dados. Tal método tem por finalidade o de elevar a discriminação entre as classes e é aplicado antes da etapa de extração de características. Já o outro algoritmo implementado foi o PCA, com a finalidade de selecionar as características mais relevantes. Este último foi implementado, naturalmente, após a etapa de extração de características. A melhor taxa de acerto obtida com o PCA foi de 87,3% e ocorreu quando se utilizou a combinação das características RMS e amplitude de Willison. Já para CSP o melhor resultado foi de 86,6, com a combinação de características RMS, amplitude de Willison e média. Para ambos os casos se realizou a etapa de classificação através de Redes Neurais Artificiais. Não foi verificada significância estatística na diferença observada entre os

métodos de pré-processamento avaliados. Entretanto, a principal contribuição desta pesquisa está na determinação de um sistema de classificação para a tarefa de reconhecimento de padrões de movimentos a partir de sinais de eletromiografia, sustentada pelos resultados interessantes reportados. Todavia, comparando tais resultados aos obtidos neste trabalho para voluntários sem amputação, observa-se relativa superioridade dos obtidos na presente análise quando se considera a Base do IEE Ensaio A (91,1%) e resultados similares quando se analisa a Base Ninapro (85,4%). Porém, convém ressaltar o fato de que são considerados um número maior de movimentos neste trabalho (17 contra apenas 5 de (RIILLO *et al.*, 2014)) o que torna a tarefa de classificação ainda mais complexa. Por outro lado, o número de características extraídas e de eletrodos utilizados na aquisição é menor em (RIILLO *et al.*, 2014). Logo, devido a essas divergências nas duas propostas não se pode determinar superioridade de um método em relação ao outro.

Por outro lado, o trabalho de (ABDELMASEEH; CHEN; STASHUK, 2016) aborda uma metodologia diferente para a tarefa de discriminação de movimentos a partir de sinais de sEMG. O algoritmo proposto não faz uso das características e classificadores tradicionais para o aprendizado. O processo de classificação é realizado através do cálculo da mínima distância entre as trajetórias de ativação dos sinais de sEMG associados aos movimentos, onde o conceito de trajetória de ativação corresponde ao segmento do sinal pré-processado (retificado, filtrado e normalizado) referente à execução do movimento. O método utilizado para o cálculo do afastamento entre trajetórias foi o *Multidimensional Dynamic Time Warping* (MD-DTW), o qual faz o alinhamento entre duas trajetórias através da busca pela Menor Distância Euclidiana entre elas. Deste modo, para uma dada trajetória de teste (sem uma classe associada a ela), determina-se a classe mais provável como sendo a que resultar na menor distância obtida entre todas as existentes no conjunto de treinamento. Outro objetivo da pesquisa foi o de desenvolver um algoritmo para detecção do instante de tempo de início de atividade muscular baseado no

valor RMS do sinal de sEMG pré-processado. Os resultados expostos na Tabela 6 correspondem às taxas de acerto média dos 40 voluntários da Base 2 Ninapro para 40 movimentos distintos. O de 89,0% foi alcançado considerando os rótulos das classes disponibilizados pela equipe do Ninapro, já o de 86,7% foi originado a partir da utilização da segmentação obtida através do método de detecção de atividade muscular implementado pelos autores. Comparando tais resultados ao obtido no presente trabalho para a mesma Base 2 Ninapro, observa-se superioridade dos resultados de (ABDELMASEEH; CHEN; STASHUK, 2016) quando este utilizou os rótulos das classes disponibilizadas pela equipe do Ninapro. Por outro lado, com a utilização da segmentação obtida através do método de detecção de atividade muscular, percebe-se superioridade do resultado obtido neste trabalho (85,4%). Porém, deve-se levar em conta que, para o cálculo da distância entre trajetórias, considerou-se o sinal do movimento por inteiro, enquanto que neste trabalho a tarefa de classificação baseou-se em janelas temporais. Por conseguinte, o método proposto em (ABDELMASEEH; CHEN; STASHUK, 2016) fica restrito a aplicações que podem ser controladas por comandos discretos. Referindo-se a próteses, por exemplo, ele pode ser utilizado na decodificação do movimento pretendido e bloquear ou habilitar o dispositivo.

Em contrapartida, a pesquisa de (ZHANG *et al.*, 2017b) está centrada na aplicação de algoritmos de pós-processamento para melhorar a taxa de acerto do classificador de movimentos em análises em tempo real. A proposta do trabalho consistiu na avaliação de três métodos distintos de detecção de atividade muscular, de modo a habilitar ou desabilitar a prótese durante a execução de um movimento ou repouso respectivamente. São eles: Valor Médio Absoluto, Operador de Energia *Teager-Kaiser*, e *Mechanomyogram* (MMG). Para tanto, foram adquiridos e analisados sinais de eletromiografia de 10 voluntários (8 intactos e 2 apresentando amputação) durante a execução de 6 movimentos distintos, repetidos por 4 vezes. Para a etapa de classificação, foram extraídas 4 características (valor médio absoluto, número

de cruzamentos por zero, comprimento da forma de onda, e mudanças de inclinação do sinal) em janelas temporais de 150 ms com deslocamentos de 100 ms, as quais alimentaram o classificador baseado no método *Linear Discriminant Analysis* (LDA). Assim, o algoritmo proposto consistiu em avaliar em tempo real, dentro das janelas temporais supracitadas, se há a presença de atividade muscular. Através deste resultado, o algoritmo decidiu se, em caso de o voluntário estar em repouso, é mantido tal estado ou se a prótese é habilitada e o movimento da saída do classificador é enviado à prótese, ou, em caso de execução de movimento, se a última classe de movimento determinada pelo classificador é mantida ou se o movimento enviado à prótese é atualizado com nova saída do classificador ou ainda se a prótese é desabilitada e o usuário volta ao estado de repouso. Os resultados obtidos desta análise mostraram que o uso de tais métodos de pós-processamento diminuiu consideravelmente os erros de classificação durante a execução dos movimentos em relação ao uso direto da saída do classificador na prótese (redução de cerca de 80%). No que se refere ao estado de repouso, não houve alteração significativa na taxa de erro observada. Por consequência, a grande contribuição desta pesquisa está na confirmação do ganho que se tem com a introdução de algoritmos de detecção de atividade muscular para a tarefa de classificação de movimentos em tempo real. Essa etapa consistiu em um passo futuro do presente trabalho visando o objetivo do desenvolvimento de uma prótese com maior número de graus de liberdade. A taxa de acerto indicada na Tabela 6 está relacionada à análise *offline*, realizada sem o uso dos referidos métodos de pós-processamento. Ela é superior à alcançada neste trabalho, porém a comparação é injusta visto que as bases de dados diferem em diversos pontos, tais quais, número de eletrodos, número de repetições e a quantidade de movimentos que nesse trabalho é de apenas 6.

Em seu estudo (LI *et al.*, 2017) agrega a análise de sinais de eletroencefalografia (EEG) aos de eletromiografia na tarefa de classificação de movimentos do segmento mão-braço. Para este fim, foi realizada a aquisição de sinais de sEMG e EEG de 4 voluntários com amputação

transradial durante a execução de 4 movimentos com 10 repetições. Os sinais de eletromiografia foram obtidos a partir de 32 eletrodos monopolares de superfície, e os de eletroencefalografia através de 64 eletrodos. Para o processo de classificação foram extraídas 4 características no domínio do tempo (valor médio absoluto, número de cruzamentos por zero, comprimento da forma de onda, e mudanças de inclinação do sinal), as quais foram utilizadas para a tarefa de classificação baseada em LDA. Foram realizadas uma série de treinamentos utilizando apenas os sinais de sEMG, após apenas os de EEG e, por fim, para um conjunto de combinações de ambos, variando-se o número de um e de outro. Assim, avaliou-se qual a combinação de eletrodos apresenta o melhor resultado em termos de taxa de acerto. Desta análise, concluiu-se que a alternativa mais adequada é o uso conjunto de todos os 32 sinais de eletromiografia e os 64 de eletroencefalografia. Com essa configuração chegou-se a uma taxa de acerto média de 91,7% enquanto que, quando se utilizou somente os sinais de sEMG atingiu-se a marca dos 77%. Já a classificação realizada somente a partir dos 64 canais de EEG resultou na taxa de acerto média de 75,1%. Logo, ficou evidente a melhoria obtida na taxa de acerto do classificador com o acréscimo da análise de EEG a de eletromiografia na caracterização de movimentos do segmento mão-braço. Conseqüentemente, tal estratégia se torna uma boa alternativa para o seguimento das pesquisas nessa área. Inclusive para o do presente trabalho, visto que a análise da Base Ninapro composto por sinais adquiridos de voluntários amputados resultou em uma taxa de acerto média de 77,5%, similar à obtida nesta pesquisa quando se utilizou somente os sinais de sEMG (77,0%).

Já no trabalho de (CENE, 2016), foi avaliado a influência dos fatores eletrodos (número e posicionamento), características, e movimento na taxa de acerto média do classificador de movimentos. A análise foi feita considerando-se as Bases 2 e 3 Ninapro e em sinais de sEMG adquiridos experimentalmente de 10 voluntários com membros intactos durante a execução dos mesmos 17 movimentos que constituem as duas primeiras. A etapa de classificação baseou-se

no algoritmo Regressão Logística. Os resultados mostraram que os três fatores analisados impactam de forma significativa na taxa de acerto média, o que ressalta a importância de se avaliar a questão dos eletrodos, movimentos e características, adicionalmente à etapa de classificação. Tal resultado acaba justificando a realização deste estudo, visto que o foco aqui está justamente na etapa de seleção de características. No concernente à taxa de acerto, (CENE, 2016) reporta taxas de acerto média de 74,2% e 61,9% para as Bases Ninapro de voluntários intactos e com amputação respectivamente. Neste estudo obteve-se resultados iguais a 85,4% e 77,5% para as mesmas Bases. Essa diferença observada nos resultados indica que a introdução da etapa de seleção de características no processo de classificação levou a uma melhora na caracterização dos movimentos. Porém vale ressaltar o fato de que, na referida pesquisa, as características foram extraídas no segmento de sinal que compreende toda a execução do movimento, e não em janelas temporais de tamanhos fixos como o realizado aqui. Além disso, os classificadores também são diferentes.

Enfim, a pesquisa de (YAN; LIU, 2013) analisa a tarefa de classificação de movimentos a partir de sinais de sEMG sob o ponto de vista da etapa de seleção de características. Nele, quatro métodos de seleção de características são implementados e comparados entre si: dois do tipo *wrapper*, e dois baseados em filtro. Os primeiros, por serem algoritmos supervisionados, utilizam classificador baseado em Redes Neurais sendo a escolha das características realizada através dos métodos de seleção sequencial para frente (*sequential forward selection*) e eliminação para trás (*backward elimination*). Já os baseados em filtros consistem nos algoritmos PCA e *Rough Set Theory* (RST). Para a análise, foram considerados sinais de sEMG adquiridos de 30 voluntários durante execução de 4 diferentes movimentos. A caracterização do sinal foi realizada através do cálculo da energia de cada um dos 16 subespaços originados da aplicação da transformada de *Wavelet* até o 4º nível. Assim, para cada amostra do sinal 16 características foram extraídas. O classificador utilizado para avaliar o melhor grupo de

características selecionadas foi o SVM. Desta análise, verificou-se que os algoritmos baseados em filtro resultaram em conjuntos menores de características e resultaram em melhores taxas de acerto dos que os do tipo *wrapper*. Dentre os métodos PCA e RST, este último alcançou a melhor taxa de acerto (95,0%) e com menor número de características (9 contra 10 selecionadas através da aplicação da PCA). A abordagem proposta em (YAN; LIU, 2013) difere da realizada neste estudo uma vez que, aqui, os algoritmos de seleção de características analisados foram empregados para a determinação de um ranking, sendo a seleção em si executada por um método do tipo *wrapper* sequencial para frente. Além disso, o número de movimentos e de voluntários considerados também são distintos. Logo, comparações de resultados de ambos os trabalhos se torna injusta.

## 5 CONCLUSÃO

No presente trabalho avaliou-se três algoritmos de seleção de características, desenvolvidos para a aplicação em problemas de seleção de genes para detecção de câncer, na tarefa de caracterização de movimentos a partir de sinais de sEMG. São eles: SVM-RFE, Monte Carlo *Feature Selection*, e *SVD Entropy*. Para tanto, foram analisados dados oriundos das Bases 2 e 3 Ninapro, além de sinais adquiridos no Laboratório do IEE. Como principal resultado pode-se verificar que a introdução de tal etapa de pré-processamento no processo de classificação proporcionou promissoras taxas de acerto, sendo superiores inclusive aos resultados de outros trabalhos que utilizaram as mesmas bases Ninapro, descritos na seção 4.6. Tal resultado indica a aplicabilidade dos referidos métodos de seleção de características também no processamento de sinais de sEMG para classificação de movimentos do segmento mão braço.

A partir dos resultados obtidos nos projetos de experimentos desenvolvidos, não foi verificada diferença significativa na taxa de acerto média entre os três métodos de seleção de características quando se analisou as Bases Ninapro adicionalmente ao Ensaio A da Base do IEE. Porém, através do Projeto de Experimentos com a Base de Dados do IEE, onde os sinais de todos os quatro tipos de ensaios foram considerados, verificou-se que o método de seleção de características influenciou de forma significativa a taxa de acerto do classificador. O algoritmo SVM-RFE apresentou a maior taxa de acerto média com 87,9%.

Por outro lado, ao se considerar o número de características como variável de resposta, observou-se que o algoritmo SVM-RFE selecionou um número médio maior de características por ensaio do que os outros dois métodos (43 enquanto que Monte Carlo e *SVD Entropy* resultaram em 39). Assim, os métodos *SVD Entropy* e Monte Carlo proporcionaram conjuntos de dados representados por um número menor de atributos ao classificador em relação ao SVM-RFE. Redução essa de 9,3%. Conseqüentemente, a tarefa de classificação fica menos complexa, reduzindo o tempo de processamento. Como tempo de processamento também está associado

a desempenho, pode-se dizer que os métodos SVD *Entropy* e Monte Carlo demonstraram superioridade em relação ao SVM-RFE para os dados analisados das Bases Ninapro e Ensaio A da Base do IEE. Entretanto, considerando-se os sinais adquiridos nos quatro tipos de ensaio realizados no Laboratório do IEE, a afirmação de qual algoritmo de seleção resulta em melhor desempenho do classificador depende da ponderação entre taxa de acerto e tempo de processamento, visto que, nestes dados, o algoritmo SVM-RFE apresentou superioridade frente aos demais em relação à taxa de acerto do classificador.

Em relação à Base de Dados no IEE, os resultados de classificação obtidos indicaram que a qualidade dos sinais de sEMG estão coerentes aos do Ninapro, visto que as taxas de acerto média de ambas as bases ficaram próximas, sendo a do IEE superior inclusive (para os Ensaios A e B). Este fato valida o protocolo de aquisição e a etapa de pré-processamento dos sinais de sEMG seguidos no desenvolvimento da Base do IEE. Além de também abrir a possibilidade de futura publicação para que outros pesquisadores possam utilizá-la.

Analisando os resultados de classificação obtidos, observa-se que as taxas de acerto alcançadas no presente trabalho estão coerentes com as obtidas nos estudos apresentados na Tabela 6. Dentre as pesquisas supracitadas, a que mais se aproxima deste trabalho é a desenvolvida por (CENE, 2016), onde foram analisadas as mesmas Bases 2 e 3 Ninapro. Os resultados superiores deste trabalho indicam que a introdução de uma etapa de seleção de características melhora a representação dos movimentos através de sinais de sEMG.

Finalmente, o presente trabalho visou a abordagem do problema de caracterização de movimentos do segmento mão-braço a partir de sinais de sEMG com a introdução da etapa de seleção de características, muitas vezes negligenciada nesta área. Os resultados alcançados nesta análise mostraram que esta forma de enxergar o problema é promissora e justificam a continuidade do trabalho neste caminho, visando o objetivo global de proporcionar futuras

próteses com maior número de graus de liberdade a pacientes que sofreram amputações transradial.

## 6 PROPOSTAS PARA TRABALHOS FUTUROS

A partir dos resultados obtidos, verificou-se que o algoritmo proposto no presente trabalho é promissor e justifica o seguimento de trabalhos nesse sentido, como já mencionado no Capítulo 5. Dentre os possíveis trabalhos que podem dar sequência a esta pesquisa, os principais são:

- Em relação à Base de Dados desenvolvida, seria interessante a aquisição de sinais de sEMG de voluntários com amputação, uma vez que, eles são os indivíduos aos quais a pesquisa está sendo direcionada. Adicionalmente, com essa expansão da Base de Dados, seria possível a comparação desta com a Base 3 do Ninapro também;
- Ainda sobre a Base de Dados do IEE, o protocolo de aquisição poderia ser alterado de modo a tornar os ensaios menos cansativos para o voluntário. Como mencionado na seção 4.2.3 os ensaios ficaram muito longos, principalmente o C e o D, impactando na caracterização dos movimentos para estes ensaios. Deste modo, caberia avaliar a redução do tempo de repouso entre a realização de movimentos ou do tempo de execução dos movimentos. Além disso, pode-se repensar a necessidade dos ensaios com 10 repetições, visto que, as taxas de acerto obtidas para os ensaios A e C ficaram próximas das alcançadas em B e D, respectivamente. Em outras palavras, quando se alterou o número de repetições de 6 para 10 se observou pequena variação na taxa de acerto média, tanto para os ensaios sequenciais quanto para os aleatórios;
- Com o objetivo de melhorar a caracterização dos movimentos considerados, a proposta sugerida na pesquisa de (LI *et al.*, 2017) se apresenta como uma boa alternativa. Em sua pesquisa, (LI *et al.*, 2017) conseguiu melhorar as taxas de acerto do seu classificador de movimentos com o uso de sinais de EEG

adicionalmente aos de sEMG. Por conseguinte, seria interessante a introdução da análise de sinais de EEG ao método proposto no presente trabalho;

- Neste trabalho, a questão das características que foram selecionadas em cada ensaio não foi abordada. Entretanto, essa análise é interessante de ser realizada, visto que, a partir dela pode-se obter a informação da frequência em que cada característica é escolhida e, assim, avaliar quais são as mais selecionadas pelo classificador. Além disso, é possível verificar se existe relação entre as características selecionadas e voluntário, o que seria interessante, em caso afirmativo, para futuras aplicações do algoritmo em próteses;
- Visando executar a tarefa de classificação de movimentos em tempo real, o trabalho de (ZHANG *et al.*, 2017b) implementa três algoritmos para detecção de atividade muscular, os quais proporcionaram bons resultados. Assim, como trabalho futuro, tais algoritmos poderiam ser testados no método proposto neste estudo, uma vez que, para o desenvolvimento de prótese com controle mioelétrico é necessário que se realize a classificação de movimentos em tempo real.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ABDELMASEEH, M.; CHEN, T. W.; STASHUK, D. W. Extraction and classification of multichannel electromyographic activation trajectories for hand movement recognition. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, New York, v. 24, n. 6, p. 662-673, Jun. 2016.
- ABHISHEK; SINGH, S. A novel algorithm to preprocess cancerous gene expression dataset for efficient gene selection. In: INTERNATIONAL CONFERENCE FOR CONVERGENCE IN TECHNOLOGY (I2CT), 2., 2017, Mumbai. **Anais...** New York: IEEE, 2017. p. 632-635.
- ADRIAN, E. D.; BRONK, D. W. The discharge of impulses in motor nerve fibres. **The Journal of Physiology**, Malden, v. 67, n. 2, p. 9-151, Mar. 1929.
- AGGARWAL, C. C.; REDDY, C. K. **DATA Clustering Algorithms and Applications**. New York: Taylor & Francis, 2014.
- AKUSOK, A. et al. High-performance extreme learning machines: a complete toolbox for big data applications. **IEEE Access**, New York, v. 3, p. 1011-1025, Jun. 2015.
- ALBASHISH, D. et al. Multi-scoring feature selection method based on svm-rfe for prostate cancer diagnosis. In: INTERNATIONAL CONFERENCE ON ELECTRICAL ENGINEERING AND INFORMATICS, 5., 2015, Bali. **Anais...** New York: IEEE, 2015. p. 682-686.
- ALTER, O.; BROWN, P. O.; BOTSTEIN, D. Singular value decomposition for genome-wide expression data processing and modeling. **Proceedings of the National Academy of Sciences**, New York, v. 97, n. 18, p. 10101-10106, Aug. 2000.
- AN, S.; WANG, J.; WEI, J. Local-nearest-neighbors-based feature weighting for gene selection. **IEEE/ACM Transactions on Computational Biology and Bioinformatics**, New York, v. PP, n. 99, p. 1-12, Jun. 2017.
- ARJUNAN, S.; KUMAR, D. Decoding subtle forearm flexions using fractal features of surface electromyogram from single and multiple sensors. **Journal of NeuroEngineering and Rehabilitation**, London, v. 7, n. 1, Oct. 2010.
- ATZORI, M. et al. Building the Ninapro database: A resource for the biorobotics community. In: PROCEEDINGS OF THE IEEE RAS AND EMBS INTERNATIONAL CONFERENCE ON BIOMEDICAL ROBOTICS AND BIOMECHATRONICS, 4., 2012, Rome. **Anais...** New York: IEEE, 2012. p. 1258-1265.
- ATZORI, M. et al. A benchmark database for myoelectric movement classification. **Transactions on neural systems and rehabilitation engineering**, Piscataway, v. 23, n. 1, p. 73-83, Jun. 2013.
- ATZORI, M. et al. Electromyography data for non-invasive naturally-controlled robotic hand prostheses. **Scientific data**, Basingstoke, v. 1, p. 1-13, Nov. 2014.
- BACKES, A. R.; BRUNO, O. M. Técnicas de estimativa da dimensão fractal: Um estudo comparativo. **INFOCOMP Journal of Computer Science**, São Paulo, v. 4, n. 3, Jan. 2005.
- BARTLETT, P. L. The sample complexity of pattern classification with neural networks: The size of the weights is more important than the size of the network. **IEEE Transactions on Information Theory**, New York, v. 44, n. 2, p. 525-536, Mar. 1998.
- BASPINAR, U.; VAROL, H. S.; YILDIZ, K. Classification of hand movements by using artificial neural network. In: INTERNATIONAL SYMPOSIUM ON INNOVATIONS IN INTELLIGENT SYSTEMS AND APPLICATIONS, 1., 2012, Trabzon. **Anais...** New York: IEEE, 2012. p. 1-4.
- BISMAR, T. A. et al. Interrogation of ERG gene rearrangements in prostate cancer identifies a prognostic 10-gene signature with relevant implication to patients' clinical outcome. **BJU International**, Hoboken, v. 113, n. 2, p. 309-319, Feb. 2014.

- BOSER, B. E.; GUYON, I. M.; VAPNIK, V. N. A training algorithm for optimal margin classifiers. In: PROCEEDINGS OF THE FIFTH ANNUAL WORKSHOP ON COMPUTATIONAL LEARNING THEORY, 5., 1992, Pittsburgh. **Anais...** New York: ACM, 1992. p. 144-152.
- CENE, V. H. **Desenvolvimento de um projeto de experimentos para a caracterização de sinais mioelétricos através do uso de regressão logística**. 2016. 137 p. Dissertação (Mestrado em engenharia) – Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2016.
- CENE, V. H. **Proposta de classificador redundante híbrido para a caracterização de sinais mioelétricos baseado em extreme learning machines**. 2017. Qualificação (Doutorado em engenharia) – Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2017.
- CENE, V. H.; BALBINOT, A. Optimization of features to classify upper - limb movements through semg signal processing. **Brazilian Journal of Instrumentation and Control**, Curitiba, v. 4, n. 3, p. 14-20, 2016.
- CENE, V. H.; FAVIEIRO, G.; BALBINOT, A. Using non - iterative methods and random weight networks to classify upper - limb movements through sEMG signals. In: PROCEEDINGS OF THE ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY, 39., 2017, Jeju. **Anais...** New York: IEEE, 2017. p. 2047-2050.
- CHANG, C.-C.; LIN, C.-J. LIBSVM: A library for support vector machines. **ACM Transactions on Intelligent Systems and Technology (TIST)**, New York, v. 2, n. 3, 2011.
- DAVID, H. A. Early sample measures of variability. **Statistical Science**, Beachwood, v. 13, n. 4, p. 368-377, 1998.
- DE LUCA, C. J. The use of surface electromyography in biomechanics. **Journal of Applied Biomechanics**, Champaign, v. 13, p. 135-163, May 1997.
- DENG, W.; ZHENG, Q.; CHEN, L. Regularized extreme learning machine. In: IEEE SYMPOSIUM ON COMPUTATIONAL INTELLIGENCE AND DATA MINING, 1., 2009, Nashville, **Anais...** New York: IEEE, 2009. p. 389-395.
- DRAMÍŃSKI, M. et al. Monte Carlo feature selection for supervised classification. **Bioinformatics**, Oxford, v. 24, n. 1, p. 110-117, Nov. 2008.
- FARINA, D.; NEGRO, F. Accessing the neural drive to muscle and translation to neurorehabilitation technologies. **IEEE Reviews in Biomedical Engineering**, New York, v. 5, p. 3-14, jan. 2012.
- FAVIEIRO, G. W. **Desenvolvimento De Um Sistema Neuro-Fuzzy Para Análise De Sinais Mioelétricos Do Segmento Mão-Braço**. 2012. 147 p. Dissertação (Mestrado em engenharia) – Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2012.
- FU, R. et al. A new feature selection method based on relief and SVM-RFE. In: INTERNATIONAL CONFERENCE ON SIGNAL PROCESSING (ICSP), 12., 2014, Hangzhou, **Anais...** New York: IEEE, 2014. p. 1363-1366.
- GENG, Y. et al. Real-time Classification of Forearm Movements Based on High Density Surface Electromyography. In: IEEE INTERNATIONAL CONFERENCE ON REAL-TIME COMPUTING AND ROBOTICS, 1., 2017, Okinawa, **Anais...** New York: IEEE, 2017. p. 246-251.
- GUYON, I. et al. Gene selection for cancer classification using support vector machines. **Machine Learning**, Berlin, v. 46, n. 1-3, p. 389-422, Jan. 2002.
- GUYTON, A. C.; HALL, J. E. **Fundamentos de fisiologia**. 12. ed. São Paulo: Elsevier, 2011.
- HALL, J. E.; GUYTON, A. C. **Tratado de fisiologia médica**. 12. ed. Rio de Janeiro: Elsevier, 2011.

- HAN, F. et al. A gene selection method for microarray data based on binary pso encoding gene-to-class sensitivity information. **IEEE/ACM transactions on computational biology and bioinformatics**, New York, v. 14, n. 1, p. 85-96, Feb. 2017.
- HAYKIN, S. **Redes neurais: Princípios e prática**. 2. ed. Porto Alegre: Bookman, 2001.
- HEARST, M. A. et al. Support vector machines. **IEEE Intelligent Systems and their Applications**, New York, v. 13, p. 18-28, 1998.
- HUANG, G.-B. et al. Extreme learning machine for regression and multiclass classification. **IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics**, New York, v. 42, n. 2, p. 513-529, Apr. 2012.
- HUANG, G.-B.; ZHU, Q.-Y.; SIEW, C.-K. Extreme learning machine: a new learning scheme of feedforward neural networks. In: PROCEEDINGS OF THE INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS, 1., 2004, Budapest, **Anais...** New York: IEEE, 2004. p. 985-990.
- HUANG, H. et al. EMG pattern recognition using decomposition techniques for constructing multiclass classifiers. In: PROCEEDINGS OF THE IEEE RAS AND EMBS INTERNATIONAL CONFERENCE ON BIOMEDICAL ROBOTICS AND BIOMECHATRONICS, 6., 2016, Singapore, **Anais...** New York: IEEE, 2016. p. 1296-1301.
- ISON, M.; ARTEMIADIS, P. The role of muscle synergies in myoelectric control: trends and challenges for simultaneous multifunction control. **Journal of Neural Engineering**, Bristol, v. 11, n. 5, p. 1-22, Oct. 2014.
- KARABULUT, D. et al. Comparative evaluation of EMG signal features for myoelectric controlled human arm prosthetics. **Biocybernetics and Biomedical Engineering**, Amsterdã, v. 37, n. 2, p. 326-335, Mar. 2017.
- KIM, S. Margin-maximized redundancy-minimized SVM-RFE for diagnostic classification of mammograms. In: IEEE INTERNATIONAL CONFERENCE ON BIOINFORMATICS AND BIOMEDICINE WORKSHOPS, BIBMW, 1., 2011, Atlanta, **Anais...** New York: IEEE, 2011. p. 562-569.
- KONRAD, P. The ABC of EMG. n. April, USA: Noraxon USA, 2005.
- LI, X. et al. A motion-classification strategy based on sEMG-EEG signal combination for upper-limb amputees. **Journal of NeuroEngineering and Rehabilitation**, London, v. 14, n. 1, Jan. 2017.
- LORENA, A. C.; DE CARVALHO, A. C. P. L. F. Uma introdução às support vector machines. **Revista de Informática Teórica e Aplicada**, Porto Alegre, v. 14, n. 2, p. 43-67, Jan. 2007.
- MCCULLOCH, W.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **Bulletin of Mathematical Biophysics**, New York, v. 5, p. 115-133, Dec. 1943.
- MERLETTI, R.; DE LUCA, C. J. New techniques in surface electromyography. **Computer Aided EMG and Expert Systems**, New York, p. 115-124, 1989.
- MOURA, K. O. A.; FAVIEIRO, G. W.; BALBINOT, A. Support vectors machine classification of surface electromyography for non-invasive naturally controlled hand prostheses. In: PROCEEDINGS OF THE ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY, EMBS, 38., 2016, Orlando, **Anais...** New York: IEEE, 2016. p. 788-791.
- MOURA, K. O. A. **Proposta de um modelo adaptativo baseado em sensores virtuais e tvarna na caracterização de sinais mioelétricos do segmento mão-braço**. 2016. Qualificação (Doutorado em engenharia) – Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2016.
- MYUNG, I. J. Tutorial on maximum likelihood estimation. **Journal of Mathematical Psychology**, Amsterdã, v. 47, p. 90-100, Nov. 2003.
- NATIONAL INSTRUMENTS. **High-Accuracy M Series Multifunction DAQ for USB - 18-Bit, up to 625 kS/s, up to 32 Analog Inputs**: Product Specification. 2014. 17 p. Disponível

- em: <<http://www.ni.com/datasheet/pdf/en/ds-24>>. Acesso em: 14 jan. 2018.
- OKUN, O. Survey of Novel Feature Selection Methods for cancer classification. In: ELLOUMI, M., ZOMAYA, A. **Biological knowledge discovery handbook: preprocessing, mining, and postprocessing of biological data**. New Jersey: Wiley - IEEE Computer Society, 2013. p. 379-398.
- ORTIZ-CATALAN, M.; BRÅNEMARK, R.; HÅKANSSON, B. BioPatRec: a modular research platform for the control of artificial limbs based on pattern recognition algorithms. **Source Code for Biology and Medicine**, London, v. 8, Abr. 2013.
- POZZO, M.; FARINA, D.; MERLETTI, R. Electromyography: detection, processing, and applications. In: MOORE, J., ZOURIDAKIS, G. **Biomedical technology and devices handbook**. New York: CRC Press, 2004. p. 70-135.
- QUINLAN, J. R. Induction of decision trees. **Machine Learning**, Boston, v. 1, n. 1, p. 81-106, 1986.
- REAZ, M. B. I.; HUSSAIN, M. S.; MOHD-YASIN, F. Techniques of EMG signal analysis: detection, processing, classification and applications (Correction). **Biological Procedures Online**, Waterloo, v. 8, n. 1, p. 163-163, Mar. 2006.
- 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**, Oxford, v. 14, p. 117-125, Nov. 2014.
- ROKACH, L.; CHIZI, B.; MAIMON, O. Feature Selection by Combining Multiple Methods. **Advances in Web Intelligence and Data Mining**. Berlin, v. 23, p. 295-304, 2006
- ROSENFALCK, P. **Intra- and extracellular potential fields of active nerve and muscle fibres A physico-mathematical analysis of different models**. Copenhagen: Akademisk Forlag, (D.B.K), 1969.
- SERRE, D. **Matrices: theory and applications**. New York: Springer-Verlag, 2002.
- STAUDE, G.; WOLF, W. Objective motor response onset detection in surface myoelectric signals. **Medical Engineering and Physics**, Amsterdam, v. 21, n. 6-7, p. 449-467, Jul. 1999.
- SUAREZ, J. N.; SALCEDO, A. ID3 and k-means based methodology for internet of things device classification. In: INTERNATIONAL CONFERENCE ON MECHATRONICS, ELECTRONICS AND AUTOMOTIVE ENGINEERING (ICMEAE), 1., 2017, Cuernavaca, **Anais...** New York: IEEE, 2017. p. 129-133.
- TANG, J.; ALELYANI, S.; LIU, H. Feature selection for classification: a review. In: AGGARWAL, C. (Ed.). **Data classification: algorithms and applications**. New York, USA: Chapman & Hall/CRC, 2015. p. 38-49.
- TORTORA, G. J.; DERRICKSON, B. **Corpo humano : fundamentos de anatomia e fisiologia**. 8. ed. Porto Alegre: Grupo A - Artmed, 2012.
- TOSIN, M. C. et al. SEMG feature selection and classification using SVM-RFE. In: PROCEEDINGS OF THE ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY, EMBS, 39., 2017, Jeju **Anais...** New York: IEEE, 2017. p. 390-393.
- VARSHAVSKY, R. et al. Novel unsupervised feature filtering of biological data. **Bioinformatics**, Oxford, v. 22, n. 14, p. 507-13, Jul. 2006.
- WANG, J. et al. Improved SVM-RFE Feature Selection Method for Multi-SVM Classifier. In: INTERNATIONAL CONFERENCE ON ELECTRICAL AND CONTROL ENGINEERING (ICECE), 1., 2011, Yichang, **Anais...** New York: IEEE, 2011. p. 1592-1595.
- XIAO, Z. G.; MENON, C. Performance of Forearm FMG and sEMG for Estimating Elbow, Forearm and Wrist Positions. **Journal of Bionic Engineering**, Beijing, v. 14, n. 2, p. 284-295, Abr. 2017.
- XIE, R. et al. Monte carlo simulation based gene selection for ovarian cancer chemotherapy response prediction with microarray data. In: INTERNATIONAL CONFERENCE ON

- BIOINFORMATICS AND BIOMEDICAL ENGINEERING, 5., 2011, Wuhan, **Anais...** New York: IEEE, 2011. p. 1-4.
- YAN, Z.; LIU, Z. The study on feature selection strategy in EMG signal recognition. In: INTERNATIONAL CONFERENCE ON COMPLEX MEDICAL ENGINEERING (ICME), 1., 2013, Beijing, **Anais...** New York: IEEE, 2013. p. 711-716.
- YIN, J. et al. Improving the performance of SVM-RFE on classification of pancreatic cancer data. In: IEEE INTERNATIONAL CONFERENCE ON INDUSTRIAL TECHNOLOGY (ICIT), 1., 2016a, Taipei, **Anais...** New York: IEEE, 2016. p. 956-961.
- YIN, Z. et al. A novel SVM-RFE based biomedical data processing approach: basic and beyond. In: ANNUAL CONFERENCE OF THE IEEE INDUSTRIAL ELECTRONICS SOCIETY, 42., 2016b, Florence, **Anais...** New York: IEEE, 2016. p. 7143-7148.
- ZHAI, X. et al. Short latency hand movement classification based on surface EMG spectrogram with PCA. In: ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY (EMBC), 38., 2016, Orlando, **Anais...** New York: IEEE, 2016. p. 327-330.
- ZHANG, S. et al. Novel gene selection method for breast cancer intrinsic subtypes from two large cohort study. In: IEEE INTERNATIONAL CONFERENCE ON BIOINFORMATICS AND BIOMEDICINE (BIBM), 1., 2017a, Kansas City, **Anais...** New York: IEEE, 2017. p. 2198-2203.
- ZHANG, X. et al. Improving the robustness of electromyogram-pattern recognition for prosthetic control by a postprocessing strategy. **Frontiers in Neurorobotics**, Lausanne, v. 11, Sep. 2017b.
- ZHOU, D. et al. Bacterial memetic algorithm based feature selection for surface emg based hand motion recognition in long-term use. In: IEEE SYMPOSIUM SERIES ON COMPUTATIONAL INTELLIGENCE, 1., 2016, Athens, **Anais...** New York: IEEE, 2016. p. 1-7.
- ZHOU, X.; TUCK, D. P. MSVM-RFE: Extensions of SVM-RFE for multiclass gene selection on DNA microarray data. **Bioinformatics**, Oxford, v. 23, n. 9, p. 1106-1114, May. 2007.
- ZOU, K. et al. ID3 decision tree in fraud detection application. In: INTERNATIONAL CONFERENCE ON COMPUTER SCIENCE AND ELECTRONICS ENGINEERING, 1., 2012, Hangzhou, **Anais...** New York: IEEE, 2012. p. 399-402.

## **ANEXO A: TERMO LIVRE E ESCLARECIDO**

### **Termo de Consentimento Livre Esclarecido**

Nome do Projeto: Análise de Algoritmos de Seleção de Características para Classificação de Sinais de sEMG do Segmento Mão-Braço

Pesquisador Responsável: Dr. Alexandre Balbinot.

Pesquisadores/Assistentes/Alunos: Me. Vinícius Cene e Eng. Maurício Tosin.

As informações contidas nesta folha, fornecidas por Alexandre Balbinot têm por objetivo firmar acordo escrito com o(a) Voluntário(a) para participação da pesquisa acima referida, autorizando sua participação com pleno conhecimento da natureza dos procedimentos a que ele(a) será submetido(a).

1. Natureza da Pesquisa: Esta pesquisa tem como finalidade o desenvolvimento de uma plataforma de tecnologia assistiva para avaliação de sinais mioelétricos do segmento mão-braço por métodos matemáticos para possibilitar futuramente o controle de dispositivos auxiliares na reabilitação de amputados.
2. Envolvimento na Pesquisa: ao participar dessa pesquisa o senhor(a) auxiliará o desenvolvimento das pesquisas na área de instrumentação biomédica. O senhor(a) tem liberdade de se recusar a participar e ainda de recusar a continuar participando em qualquer fase da pesquisa, sem lhe ocasionar qualquer prejuízo. Sempre que quiser poderá pedir mais informações sobre a pesquisa através do telefone dos pesquisadores e do projeto.
3. Protocolo experimental: Inicialmente será realizada uma breve anamnese (entrevista), onde apenas os pesquisadores terão conhecimento dos dados informados. Posteriormente, será posicionado no braço e antebraço do Voluntário eletrodos de superfície, um vídeo será rodado informando os movimentos que devem ser realizados. Durante esse procedimento, estaremos adquirindo, através de um eletromiógrafo, os sinais mioelétricos do Voluntário, sendo esse o dado utilizado nesta pesquisa.
4. Riscos e desconfortos: Os procedimentos adotados nesta pesquisa obedecem aos Critérios da Ética em Pesquisa com Seres Humanos conforme Resolução nº.196/96 do Conselho Nacional de Saúde. O procedimento adotado pode causar uma leve fadiga muscular.
5. Confidencialidade: as informações neste estudo são estritamente confidenciais. Os dados do Voluntário serão identificados com um código e não com o nome. Somente os pesquisadores e o orientador desta pesquisa terão conhecimento dos dados.
6. Benefícios: Ao participar desta pesquisa você não terá nenhum benefício direto. Entretanto, esperamos que este estudo contribua com informações e elementos importantes à literatura e ao desenvolvimento da tecnologia assistiva, onde o pesquisador se compromete a divulgar os resultados obtidos.
7. Garantia de acesso: em qualquer etapa do estudo, você terá acesso aos profissionais responsáveis pela pesquisa para esclarecimento de eventuais dúvidas. O Pesquisador Responsável é o Dr. Eng. Alexandre Balbinot, que pode ser encontrado no endereço Av. Osvaldo Aranha 103 sala 206 D, POA/RS (UFRGS) Telefone 3308 3326.
8. Despesas e compensações: não há despesas pessoais para o participante em qualquer fase do estudo, incluindo exames e consultas. Também não há compensação financeira relacionada à

sua participação. Se existir qualquer despesa adicional, ela será absorvida pelo orçamento da pesquisa.

9. Liberdade de recusar ou retirar consentimento: Você tem a liberdade de retirar seu consentimento a qualquer momento e deixar de participar do estudo sem penalização.

10. Os dados obtidos são de livre uso do grupo de pesquisadores, ou seja, podem ser publicados e utilizados no desenvolvimento de trabalhos científicos. Porém em nenhum momento serão tornados públicos os dados pessoais do Voluntário.

Após estes esclarecimentos, solicitamos o seu consentimento de forma livre para permitir sua participação nesta pesquisa.

### Consentimento Livre e Esclarecido

Eu \_\_\_\_\_ portador do RG \_\_\_\_\_, de forma livre e esclarecida, manifesto meu consentimento em participar da pesquisa em participar nesse estudo. Ficaram claros para mim quais são os propósitos do estudo, os procedimentos a serem realizados, seus desconfortos e riscos, as garantias de confidencialidade e de esclarecimentos permanentes. Ficou claro também que minha participação é isenta de despesas. Concordo voluntariamente em participar deste estudo e poderei retirar o meu consentimento a qualquer momento, antes ou durante o mesmo, sem penalidades ou prejuízo ou perda de qualquer benefício que eu possa ter adquirido, ou no meu atendimento neste Serviço. Declaro que recebi cópia deste termo de consentimento, e autorizo a realização da pesquisa e a divulgação dos dados obtidos neste estudo.

**Local e data** \_\_\_\_\_, \_\_\_\_/\_\_\_\_/\_\_\_\_\_

\_\_\_\_\_  
Assinatura do Participante da Pesquisa

\_\_\_\_\_  
Assinatura do Pesquisador Responsável – Dr. Alexandre Balbinot.

Pesquisador Principal: Alexandre Balbinot – Tel:

Demais pesquisadores: Vinícius Horn Cene – Tel:

Maurício Tosin – Tel:

**ANEXO B: ANAMNESE****Identificação**

---

**Nome:****RG:****Idade:** **Idade em que ocorreu a amputação:****Altura:** **Peso:** **Sexo:** **Cor:****Naturalidade:****Profissão :** **Profissão antes da amputação:****Estado Civil:****Escolaridade:****Contato**

---

**Endereço:****Bairro:** **CEP:****Cidade:** **Estado:****Telefone:****e-mail:****Anamnese**

---

**Queixa Principal:****HMA:**

HMP:

Data(s) da(s) amputação(s):

Já usou prótese?

Sim       Não

Quanto tempo (meses)?      Quantas?

Tipo(s) da(s) prótese(s)?

## Anamnese

Doenças pré-existent:

Apresenta LER?

Medicamentos em uso:

Antecedentes cirúrgicos:

Hábitos de vida:

Tabagismo       Etilismo       Drogas Ilícitas – Quais?

## Etiologia da Amputação

*Vascular*

Quanto tempo durou a enfermidade?

Apresenta sinais ou sintomas de doenças obstrutiva?

Sim       Não

Quais?

Arteriosclerose obliterante

Dor isquêmica de repouso

Alterações em anexos epidérmicos

Microangiopatia diabética

Alterações de temperatura e cor da pele       Parasitismo

Traumática

Imediata       Eletiva

Como foi o acidente?

*Neoplasia*

Duração da enfermidade?

Condição da patologia:     metástase       complicações

Tratamento utilizado:     Quimioterapia     Radioterapia

*Infecciosa*

Qual a patologia?

Qual o agente infeccioso?

*Congênita*

Causa:

*Acidente por animais:*

Qual animal?

Nível da Amputação

MSD

desarticulação do ombro

transumeral proximal

transumeral medial

]transumeral distal

] desarticulação do cotovelo

] transradial proximal

] transradial medial

] transradial distal

]desarticulação de punho

]MSE

] desarticulação do ombro

] transumeral proximal

] transumeral medial

]transumeral distal

] desarticulação do cotovelo

] transradial proximal

] transradial medial

] transradial distal

]desarticulação de punho

Observações:

## Avaliação dos Cotos

---

### MEMBRO FANTASMA

Paciente apresenta sensação de membro fantasma?

]Sim  ]Não

Qual a sensação fantasma?

Pressão  Dormência  Formigamento  Posição do membro  Temperatura

Paciente apresenta dor fantasma?

Sim  Não

Sensação:

Disparo doloroso  Queimação  Aperto  Câimbra

Complicações intercorrências decorrentes da amputação:

Sim  Não

Quais?

do próprio coto  Sensoriais  Clínicas  Neurológicas  Psicológicas

## Canais Usados

Canal 1- Bíceps - Flexão do antebraço

Canal 2- Flexor Radial do Carpo - Flexão do punho e abdução da mão

Canal 3- Flexor Ulnar do Carpo - Flexão do punho e adução da mão

Canal 4- Extensor dos dedos - extensão das falanges

Canal 5- Pronador Redondo - Pronação

Canal 6- Braquioradial - Supinação

Canal 7- Palmar Longo - flexão da mão

Canal 8- Extensor Ulnar do Carpo - Extensão e adução da mão

Observações: