

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

THIAGO SIMÕES DIAS

**LUVA INSTRUMENTADA PARA RECONHECIMENTO DE PADRÕES DE
GESTOS EM LIBRAS**

DISSERTAÇÃO

CURITIBA

2020

THIAGO SIMÕES DIAS

**LUVA INSTRUMENTADA PARA RECONHECIMENTO DE PADRÕES DE GESTOS
EM LIBRAS**

Dissertação apresentada ao programa de Pós-graduação em Engenharia Elétrica e Informática Industrial da Universidade Tecnológica Federal do Paraná como requisito parcial para obtenção do grau de “Mestre em Ciências” – Área de Concentração: Engenharia Biomédica.
Orientador: Prof. Dr. Sérgio Francisco Pichorim

CURITIBA

2020

Dados Internacionais de Catalogação na Publicação

Dias, Thiago Simões

Luva instrumentada para reconhecimento de padrões de gestos em Libras [recurso eletrônico] / Thiago Simões Dias. -- 2020.

1 arquivo texto (95 f.): PDF; 3,22 MB.

Modo de acesso: World Wide Web

Título extraído da tela de título (visualizado em 4 jun. 2020)

Texto em português com resumo em inglês

Dissertação (Mestrado) - Universidade Tecnológica Federal do Paraná.

Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial, Curitiba, 2020

Bibliografia: f. 85-90.

1. Engenharia elétrica - Dissertações. 2. Língua brasileira de sinais. 3. Língua brasileira de sinais - Aparelhos e instrumentos. 4. Comunicação visual - Aspectos sociais. 5. Comunicação visual - Aparelhos e instrumentos. 6. Surdos - Meios de comunicação. 7. Língua de sinais - Aparelhos e instrumentos. I. Pichorim, Sérgio Francisco. II. Universidade Tecnológica Federal do Paraná. Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial. III. Título.

CDD: ed. 23 – 621.3

Biblioteca Central da UTFPR, Câmpus Curitiba

Bibliotecário: Adriano Lopes CRB-9/1429

TERMO DE APROVAÇÃO DE DISSERTAÇÃO

A Dissertação de Mestrado intitulada “**Luva Instrumentada para Reconhecimento de Padrões de Gestos em Libras**” defendida em sessão pública pelo(a) candidato(a) **Thiago Simões Dias**, no dia **24 de março de 2020**, foi julgada para a obtenção do título de Mestre em Ciências, Área de Concentração: **Engenharia Biomédica**, Linha de Pesquisa: Sensores e Instrumentos de Aplicação Biomédica, e aprovada em sua forma final, pelo Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial - CPGEI.

BANCA EXAMINADORA:

Prof. Dr. Sérgio Francisco Pichorim – Presidente – UTFPR

Prof. Dr. Carmen Caroline Rasera – UTFPR

Prof. Dr. César Janeczko – UTFPR

Prof. Dr. Marcos Santos Hara – IFPR

A via original deste documento encontra-se arquivada na Secretaria do Programa, contendo a assinatura da Coordenação após a entrega da versão corrigida do trabalho.

Curitiba, 24 de março de 2020.

Carimbo e Assinatura do(a) Coordenador(a) do Programa

AGRADECIMENTOS

Ao Senhor Jesus Cristo, em quem depusitei toda minha confiança e me salva todas as vezes que O invoco. Aos irmãos em Cristo que estiveram comigo nesse desafio.

A minha esposa Gabriella, por seu amor, suporte, cuidado e paciência. Aos meus pais, Edivaldo e Lucia, pelo cuidado, incentivo e orações. Aos meus irmãos, pela amizade e companheirismo.

Ao meu orientador Sérgio Francisco Pichorim, pela dedicação, paciência, confiança depositada em mim e sabedoria na condução desse trabalho.

Ao meu amigo José Jair Alves Mendes Júnior, pelo auxílio no desenvolvimento deste trabalho e por compartilhar sua experiência sobre o tema abordado.

Aos professores do CPGEI e colegas do mestrado.

RESUMO

DIAS, Thiago Simões. LUIVA INSTRUMENTADA PARA RECONHECIMENTO DE PADRÕES DE GESTOS EM LIBRAS. 95 f. Dissertação – Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial, Universidade Tecnológica Federal do Paraná. Curitiba, 2020.

Este trabalho apresenta o desenvolvimento de um sistema para reconhecer padrões de gestos da Língua Brasileira de Sinais (Libras). Este sistema é composto de uma luva instrumentada, sistema de aquisição, processamento e classificação por meio de Redes Neurais Artificiais (RNA). A luva desenvolvida possui cinco sensores flexíveis, dois sensores de contato e um sensor inercial (acelerômetro e giroscópio tri-axiais). Para coletar os dados referentes aos gestos executados por voluntários, duas versões de sistemas de aquisição de dados foram utilizadas: sistema de aquisição de dados com fio e sistema de aquisição de dados sem fio. No sistema com fio, cinco voluntários participaram da coleta de dados referente aos gestos das letras do alfabeto manual em Libras. Já com o sistema sem fio, dez voluntários participaram da coleta de dez diferentes palavras em Libras. Os dados coletados foram segmentados em três janelas (quantidades fixas de amostras do sinal) que representam as fases de construção, gesto e relaxamento dos gestos. Após a segmentação, cada janela segmentada foi submetida à extração de características para gerar um vetor de características. Para a classificação, o vetor de características foi dividido em 80% para treinamento e 20% para teste da RNA. A taxa média de acertos obtida para os gestos do alfabeto manual foi de 96,19% e a taxa média obtida para os gestos de palavras foi de 98,96%. Durante o desenvolvimento da pesquisa, algumas contribuições foram geradas por meio das análises realizadas, evidenciando o potencial do sistema para reconhecer gestos em Libras. As análises realizadas e discutidas no trabalho são relacionadas aos sensores, características, voluntários e quantidade de dados separados para treinamento da rede.

Palavras-chave: Reconhecimento de gestos, Luva instrumentada, Língua Brasileira de Sinais, Redes Neurais Artificiais.

ABSTRACT

DIAS, Thiago Simões. INSTRUMENTED GLOVE FOR RECOGNITION OF THE LIBRAS GESTURES PATTERNS. 95 f. Dissertação – Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial, Universidade Tecnológica Federal do Paraná. Curitiba, 2020.

This paper presents the development of a system to recognize gestures patterns of the Brazilian Sign Language (Libras). This system is composed of an instrumented glove, acquisition system, processing and classification by Artificial Neural Networks (RNA). The developed glove has five flex-sensors, two contact sensors and an inertial sensor (three-axis accelerometer and gyroscope). Two versions of data acquisition systems were used to collect the data regarding the gestures performed by volunteers: wired data acquisition system and wireless data acquisition system. In the wired system, five volunteers participated in the collection of data related to the characters of Libras alphabet. With the wireless system, ten volunteers participated in the collection of ten different words in Libras. The collected data were segmented in three windows (fixed amounts of signal samples) that represent the construction period, gesture period and relaxation period of the gestures. After the segmentation, each segmented window was submitted to the extraction of features to generate a vector of features. For classification, the vector of features was divided into 80% for training and 20% for testing of the RNA. The accuracy rate obtained for manual alphabet gestures was 96.19% and the accuracy rate obtained for word gestures was 98.96%. During the research processes, some contributions were generated through the performed analysis, evidencing the potential of the system to gestures recognize in Libras. The performed and discussed analyses in the work are related to sensors, characteristics, volunteers and amount of separate data for network training.

Keywords: Gesture Recognition, Instrumented Glove, Brazilian Sign Language, Artificial Neural Networks.

LISTA DE FIGURAS

Figura 1 - Luvas instrumentadas para reconhecer gestos da língua de sinais taiwanesa. A informação do gesto reconhecido é enviada para um celular, onde é apresentada.	21
Figura 2 - Dispositivo vestíveis para reconhecimento das letras do alfabeto na língua de sinais americana. Apresentando: a) Dispositivo vestível com sensores flexíveis com suporte de dedo, sensor inercial, microcontrolador e módulo Bluetooth e b) Sensores de pressão.....	22
Figura 3 - Luva com sensores flexíveis e inerciais para reconhecimento de Libras. Os números de 1 a 8 representam os sensores inerciais.	24
Figura 4 - Ilustração dos ossos e articulações presentes na mão humana.	25
Figura 5 - Ilustração dos principais movimentos realizados pelas mãos. Apresentando a mão direita realizando movimentos de flexão, extensão, oposição, reposição e abdução.	26
Figura 6 - Ilustração das regiões de curvatura dos dedos da mão direita, em que IF é a articulação interfalângiana, MCF é a metacarpofalângiana, IFD é a interfalângiana distal e IFP é a interfalângiana proximal.	27
Figura 7 – Alfabeto manual em Libras. Ilustração das 26 letras, divididas em 20 sinais estáticos e 6 dinâmicos.	29
Figura 8 – Gestos de palavras em Libras selecionados para o trabalho.	30
Figura 9 - Modelo de um neurônio artificial simples.	32
Figura 10 - Exemplo de rede neural com arquitetura: a) feedforward, b) realimentada e c) com estrutura reticulada.	33
Figura 11 - Exemplo de uma Rede Neural MLP com três camadas: camada de entrada com 4 dados, camada escondida com 3 neurônios e camada de saída com 1 neurônio.	34
Figura 12 – Momento de parada do processo de aprendizagem da rede neural. Nesse exemplo a parada antecipada ocorre na décima época de treinamento.	38
Figura 13 – Rachaduras em escala de micrométrica causadas no carbono devido à curvatura do sensor flexível.....	39
Figura 14 – Sensor flexível comercial, sendo a) foto do sensor, b) divisor de tensão utilizado para realizar a leitura do sensor e c) ilustração da curvatura do sensor.	40

Figura 15 – Estrutura de um acelerômetro MEMS, formada por massa, eletrodo e molas. A aceleração causa o deslocamento da massa e altera os valores de capacitâncias C1 e C2 medidos entre a massa e o eletrodo.	42
Figura 16 – Aceleração da gravidade apresentada nos eixos do acelerômetro. A saída do acelerômetro apresenta sempre ± 1 g no eixo que está perpendicular ao sentido da gravidade quando o dispositivo é colocado em uma superfície plana.....	43
Figura 17 – Estrutura de um giroscópio MEMS, que é formada por massa, eletrodos e molas. A força de Coriolis causa vibração na massa e altera os valores de capacitância medidos entre a massa e o eletrodo.	44
Figura 18 – Eixos de sensibilidade do giroscópio MEMS MPU6050.	45
Figura 19 – Sensores de contato feito de tecido condutivo costurado sobre a luva. .	46
Figura 20 - Luva Instrumentada com 5 sensores de flexão, 1 sensor inercial e 2 sensores de contato, em que (a) apresenta o posicionamento dos sensores, (b) apresenta a foto da luva e a imagem ampliada do tecido condutivo costurado sobre a luva e (c) apresenta foto da linha condutiva e do tecido condutivo.	48
Figura 21 – Processo utilizado na coleta de dados. Os dados referentes aos sensores da luva são lidos pelo microcontrolador Arduino e enviados, via comunicação serial, para o software LabVIEW no computador.....	50
Figura 22 – Processo utilizado na coleta de dados. Os dados referentes aos sensores da luva são lidos pelo microcontrolador e enviados sem fio para o software LabVIEW, no computador.	50
Figura 23 – Fluxograma das etapas de desenvolvimento da pesquisa.....	52
Figura 24 – Voluntário sentado de frente para a tela do computador e o painel que informa qual gesto deve ser realizado. O painel apresenta um aviso sonoro e visual, além dos gestos que devem ser realizados.	53
Figura 25 – Ilustração da segmentação dos dados coletados para as três primeiras letras do alfabeto. Os dados foram segmentados em janela de construção (Constr.), janela de gesto e janela de relaxamento (Relax.).	54
Figura 26 – Ilustração da etapa de segmentação e extração de características realizadas para gerar o vetor de características.....	55
Figura 27 – Validação cruzada k-partições utilizada para escolher a quantidade de neurônios para a rede neural.	57

Figura 28 – Sensor inercial posicionado na região superior do dorso da mão. As setas Ax, Ay e Az representam os sinais de aceleração e as setas Gx, Gy e Gz indicam os sinais de giro.	59
Figura 29 - Apresentação do comportamento dos sinais do sensor inercial durante a realização de algumas letras.....	59
Figura 30 – Transferidor de ângulos com haste adaptada. A haste adaptada auxilia a fixação dos sensores e permite o ajuste do sensor no ângulo desejado.	60
Figura 31 - Retas de relação entre grau de curvatura e tensão dos sensores flexíveis, em que (a) é o sensor flexível de 5,6 cm de comprimento ativo ($y = -0,0043x+1,428$) e (b) é o sensor flexível de 9,52 cm ($y = -0,0035x + 0,968$).	61
Figura 32 – Resultados obtidos para a classificação com redes neurais considerando as características de forma individual. Para os sensores de contato, apenas a moda foi considerada e para os sensores flexíveis e inercial, seis características foram consideradas.....	62
Figura 33 - Classificação da rede neural considerando os sensores de forma individual com três características (valor máximo, mínimo e média) para os sensores flexíveis e inercial e a moda para os sensores de contato.	63
Figura 34 – Matriz de confusão obtida para a classificação individual com os sensores flexíveis.	64
Figura 35 – Matriz de confusão obtida para a classificação individual com o acelerômetro.	65
Figura 36 – Matriz de confusão obtida para a classificação individual com o giroscópio.	66
Figura 37 – Matriz de confusão obtida para a classificação individual com o sensor de contato 1.....	67
Figura 38 – Matriz de confusão obtida para a classificação individual com o sensor de contato 2.....	67
Figura 39 - Classificação da rede neural considerando os sensores em grupo com as três características (valor máximo, mínimo e média) para os sensores flexíveis e inercial e a moda para os sensores de contato.	68
Figura 40 – Matriz de confusão para a classificação da rede neural com todos os sensores, utilizando três características (valor máximo, mínimo, média) para os sensores flexíveis e inercial e uma (moda) para os sensores de contato.	69

Figura 41 – Classificação da rede neural considerando os dados de quatro voluntários para o treinamento e de um voluntário para o teste.	70
Figura 42 – Classificação da rede neural com os dados divididos em diferentes proporções para treinamento e teste.....	71
Figura 43 – Resultados obtidos para a classificação com redes neurais considerando as características de forma individual. Para os sensores de contato, apenas a moda foi considerada e para os sensores flexíveis e inercial, seis características foram consideradas.....	72
Figura 44 – Classificação da rede neural considerando os sensores de forma individual com três características (média, valor máximo e valor mínimo) para os sensores flexíveis e inercial e a moda para os sensores de contato.	73
Figura 45 – Matriz de confusão obtida na classificação individual com os cinco sensores flexíveis.....	74
Figura 46 – Matriz de confusão obtida na classificação individual com apenas o acelerômetro.	75
Figura 47 – Matriz de confusão obtida na classificação individual com apenas o giroscópio.....	75
Figura 48 – Matriz de confusão obtida na classificação individual com o sensor de contato 1.....	76
Figura 49 – Matriz de confusão obtida na classificação individual com o sensor de contato 2.....	76
Figura 50 – Classificação obtida com a rede neural considerando os sensores em grupo com as três características (valor máximo, mínimo e média) para os sensores flexíveis e inercial e a moda para os sensores de contato.	77
Figura 51 – Matriz de confusão obtida na classificação da rede neural com o grupo que apresentou o melhor desempenho (flexíveis, acelerômetro e giroscópio), utilizando três características (valor máximo, mínimo, média) para os sensores flexíveis e inercial.	78
Figura 52 – Classificação da rede neural considerando os dados de nove voluntários para o treinamento e de um voluntário para o teste.	79
Figura 53 – Classificação obtida na rede neural com os dados divididos em diferentes proporções para treinamento e teste.....	80

LISTA DE TABELAS

Tabela 1 - Trabalhos que utilizam sensores flexíveis, inerciais e de contato posicionados nas mãos para reconhecimento de gestos em língua de sinais.	20
Tabela 2 – Informação de 20 bits que auxilia na discriminação dos gestos.	21

LISTA DE SIGLAS

dps	graus por segundo (do inglês <i>degrees-per-second</i>)
EMG	Eletromiografia
IA	Inteligência Artificial
IF	Interfalangiana
IFD	Interfalangiana distal
IFP	Interfalangiana proximal
I ² C	Circuito Inter-Integrado (do inglês <i>Inter-Integrated Circuit</i>)
KNN	K-Vizinhos Próximos (do inglês <i>K-Nearest Neighbors</i>)
LDA	Análise de Discriminantes Lineares (do inglês <i>Linear Discriminant Analysis</i>)
Libras	Língua Brasileira de Sinais
MCF	Metacarpofalângiana
MEMS	Sistemas Microeletromecânicos (do inglês <i>Microelectromechanical System</i>)
MLP	Perceptron de Múltiplas Camadas (do inglês <i>Multi-Layer Perceptron</i>)
RNA	Redes Neurais Artificiais
R ²	Coeficiente de correlação
SVM	Máquinas de Vetores de Suporte (do inglês <i>Support Vector Machine</i>)
USB	Barramento Serial Universal (do inglês <i>Universal Serial Bus</i>)

LISTA DE SÍMBOLOS

A	Área das placas do capacitor
a	Aceleração
C	Capacitância
d	Distância entre as placas do capacitor
D	Saída desejada (rótulo) da RNA
E_M	Erro quadrático médio da RNA
F	Força
g	Força da gravidade
$g()$	Função de ativação da RNA
K	Curtose
k	Constante elástica da mola
m	Massa
N	Número de elementos de um conjunto de dados
S	Assimetria
SD	Desvio padrão
u	Potencial de ativação da RNA
W_n	Pesos sinápticos da RNA
W_1	Conjunto de pesos entre a camada de entrada e a camada escondida da RNA
W_2	Conjunto de pesos entre a camada escondida e a camada de saída da RNA
X	Dados de entrada da RNA
x	Deslocamento da massa
X_i	Conjunto de dados
X_n	Sinais de entrada da RNA
y	Sinal de saída da RNA
Y_1	Vetor de saída da camada escondida da RNA
Y_2	Vetor de saída da camada de saída da RNA
δ_1	Gradiente local do erro da camada escondida da RNA

δ_2	Gradiente local do erro da camada de saída da RNA
ϵ_0	Permissividade dielétrica do vácuo
η	Taxa de aprendizagem da RNA
θ	Limiar de ativação (ou <i>bias</i>) da RNA
μ	Média
ρ	Número de amostras (exemplos) da RNA
Σ	Combinador linear da RNA

SUMÁRIO

1 INTRODUÇÃO	15
1.1 OBJETIVOS	16
1.1.1 Objetivo Geral	16
1.1.2 Objetivos Específicos	16
1.2 JUSTIFICATIVA	17
1.3 ORGANIZAÇÃO DA DISSERTAÇÃO	17
2 ESTADO DA ARTE	18
3 FUNDAMENTAÇÃO TEÓRICA	25
3.1 ANATOMIA DAS MÃOS	25
3.2 LÍNGUA DE SINAIS	27
3.2.1 Língua Brasileira de Sinais	28
3.3 REDES NEURAIS ARTIFICIAIS	31
3.3.1 Redes Perceptron Multicamadas	34
3.4 SENSORES	38
3.4.1 Sensores Flexíveis	38
3.4.2 Sensores Inerciais	40
3.4.3 Sensores de Contato	45
4 MATERIAIS E MÉTODOS	47
4.1 LUVA INSTRUMENTADA	47
4.2 SISTEMA DE AQUISIÇÃO DE DADOS	49
4.2.1 Sistema de aquisição de dados com fio	49
4.2.2 Sistema de aquisição de dados sem fio	50
4.4 METODOLOGIA DO PROCEDIMENTO EXPERIMENTAL	51
4.5 SEGMENTAÇÃO DOS DADOS E EXTRAÇÃO DE CARACTERÍSTICAS	53
4.7 CLASSIFICAÇÃO	56
5 RESULTADOS	58
5.1 COMPORTAMENTO DOS SENSORES	58
5.1.1 Análise do comportamento do sensor Inercial	58
5.1.2 Análise do comportamento dos sensores flexíveis	60
5.2 SISTEMA COM FIO – ALFABETO	61
5.2.1 Análise das características	61
5.2.2 Análise por tipo de sensor	62
5.2.3 Análise dos sensores em grupos	68
5.2.4 Análise dos voluntários	70
5.2.5 Análise da quantidade de dados para treinamento	71
5.3 SISTEMA SEM FIO – PALAVRAS	72
5.3.1 Análise das características	72
5.3.2 Análise do tipo de sensor	73
5.3.3 Análise dos sensores em grupos	77
5.3.4 Análise dos voluntários	78
5.3.5 Análise da quantidade de dados para treinamento	79
5.4 COMPARAÇÃO DOS RESULTADOS	80
6 CONSIDERAÇÕES FINAIS E CONCLUSÃO	82
REFERÊNCIAS	85
ANEXO A – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (TCLE)	91

1 INTRODUÇÃO

A deficiência auditiva atinge pessoas de todo o mundo. No Brasil, o censo demográfico de 2010 apontou que mais de 7 milhões de pessoas possuem alguma deficiência auditiva (IBGE, 2010). A surdez é definida como a dificuldade ou até mesmo a impossibilidade de ouvir e, dependendo do grau da surdez, a aquisição da linguagem oral pode ser comprometida (BIBLIOTECA VIRTUAL EM SAÚDE, 2017). Por isso, muitas dessas pessoas utilizam a língua de sinais para se comunicar. A língua de sinais é uma língua estruturada, com gramática própria e é específica para cada país (RAMOS, 2002).

No Brasil, os surdos se comunicam por meio da Língua Brasileira de Sinais (Libras), que é definida pela lei 10.436 de 2002 como o modo de comunicação e expressão da comunidade surda (BRASIL, 2002). Apesar de ser uma língua reconhecida por lei, atualmente poucas pessoas sabem se comunicar utilizando Libras. Para aprender Libras é necessário estudo e imersão na comunidade surda, assim como no processo de aprendizagem de qualquer idioma (BIBLIOTECA VIRTUAL EM SAÚDE, 2017).

No dia a dia dos surdos, a comunicação com os ouvintes que não dominam a língua de sinais é um grande desafio (CHEOK; OMAR; JAWARD, 2017). Algumas atividades simples, como pedir alguma informação para um ouvinte, pode se tornar uma difícil tarefa para o surdo. Os surdos podem enfrentar dificuldades de comunicação até mesmo dentro de casa com seus familiares (LOPES; LEITE, 2011). Nesse cenário, percebe-se a necessidade de desenvolvimento de soluções capazes de auxiliar a vida de pessoas surdas.

Atualmente, diversas pesquisas voltadas para o reconhecimento de gestos em língua de sinais estão sendo desenvolvidas com o uso de câmeras, análise dos sinais musculares e luvas instrumentadas. Com o avanço da tecnologia, cada vez mais tipos de materiais, sensores e sistemas de desenvolvimento estão acessíveis. Pode-se citar como exemplo a tecnologia de Sistemas Microeletromecânicos (MEMS), em que sensores inerciais pequenos e de baixo custo são desenvolvidos com o uso da microeletrônica (SHAEFFER, 2013).

A maioria das luvas instrumentadas possuem sensores de flexão, inerciais e de contato/toque para identificar os gestos em língua de sinais. As pesquisas existentes com essas luvas empregam diferentes formas de classificar os gestos em

língua de sinais. Porém existe a necessidade de analisar de forma mais profunda a contribuição de cada um dos sensores na identificação de gestos, bem como as etapas de processamento e classificação dos gestos em língua de sinais.

Neste trabalho foi desenvolvido um sistema para reconhecer gestos em Libras. Como parte deste sistema, foi construída uma luva com sensores flexíveis, de contato e inercial para identificar direção, movimento e configuração da mão durante a execução de gestos em Libras. Para realizar a coleta de dados, foi desenvolvido um sistema de aquisição de dados. Diferentes voluntários, vestindo a luva, executaram gestos de letras do alfabeto e de palavras em Libras. Os dados coletados foram segmentados e submetidos a extração de características, para serem classificados por meio de Redes Neurais Artificiais (RNA). A partir desses processos, diferentes análises foram realizadas, gerando uma série de contribuições relacionadas aos sensores utilizados, aos voluntários e à extração de características.

1.1 OBJETIVOS

1.1.1 Objetivo Geral

O objetivo geral deste trabalho consiste no desenvolvimento de um sistema composto de uma luva instrumentada e um sistema de aquisição dos sinais para a posterior classificação de padrões de gestos em Libras utilizando RNA.

1.1.2 Objetivos Específicos

- Desenvolvimento de uma luva instrumentada para identificar padrões de gestos em Libras;
- Analisar os sensores da luva;
- Realizar a coleta de dados utilizando a luva instrumentada com diferentes voluntários;
- Organizar os dados coletados e segmentar as diferentes classes de gestos;
- Realizar a extração de características dos sinais referentes aos gestos;
- Treinar uma RNA para classificar os dados;
- Analisar os resultados obtidos com a classificação em diferentes cenários.

1.2 JUSTIFICATIVA

Nota-se que há uma grande necessidade de soluções voltadas para pessoas com deficiência. Em 2015 foi criada a lei 13.146, denominada lei brasileira de inclusão da pessoa com deficiência, que visa melhorar a acessibilidade e inclusão de pessoas com deficiência na sociedade. Dentre os direitos estabelecidos por essa lei, pode-se citar o direito à habilitação e à reabilitação, que tem por objetivo desenvolver a autonomia e participação social do deficiente com as demais pessoas (BRASIL, 2015).

A utilização das mãos durante a comunicação é, de modo geral, uma forma natural de transmitir mensagens. Porém, para se comunicar com surdos que utilizam a língua de sinais, é necessário dominar essa língua. Desta forma, ferramentas e tecnologias voltadas para a comunicação entre surdos e ouvintes são muito importantes para melhorar a inclusão social da comunidade surda.

No Brasil, as pesquisas que utilizam luvas instrumentadas para reconhecimento de gestos em língua de sinais ainda não estão consolidadas. Dentro dessa linha de pesquisa, diversos métodos de construção, técnicas de processamento e classificação podem ser explorados. As pesquisas nessa área podem contribuir na criação de soluções voltadas para a inclusão e acessibilidade de pessoas com surdez.

1.3 ORGANIZAÇÃO DA DISSERTAÇÃO

Esse trabalho está organizado em 6 capítulos:

- No capítulo 2 os principais trabalhos que utilizam luva instrumentada para reconhecer língua de sinais são apresentados;
- No capítulo 3 são abordados os conceitos da anatomia das mãos, da língua de sinais e dos sensores utilizados no trabalho;
- O capítulo 4 descreve os materiais utilizados no desenvolvimento da luva e do sistema de aquisição de dados. Em seguida, é apresentada a metodologia, que envolve os processos de aquisição, segmentação e classificação dos gestos;
- O capítulo 5 apresenta os resultados obtidos na classificação dos gestos em diferentes cenários;
- O capítulo 6 apresenta as considerações finais e as contribuições geradas.

2 ESTADO DA ARTE

Neste capítulo, os trabalhos relacionados a sistemas automatizados de reconhecimento de padrões de gestos em língua de sinais são apresentados, abordando, principalmente, os trabalhos que utilizam sensores posicionados nas mãos para identificação de gestos manuais.

Os sistemas automatizados utilizados para reconhecimento de movimentos e gestos das mãos podem ser divididos em duas abordagens: baseada em imagem, que utiliza imagens estáticas ou vídeos, e a baseada em sensores, que utiliza diferentes tipos de sensores (CHEOK; OMAR; JAWARD, 2017).

A abordagem baseada em imagens pode ter seu processamento de imagem comprometido devido a alguns fatores como variação da luz, dependência de cor e problemas com o fundo da imagem. Por outro lado, a simplicidade é um fator positivo, pois é possível utilizar a câmera de um celular para realizar a captura de imagens ou vídeos (GHANEM; CONLY; ATHITSOS, 2017).

Nessa abordagem, diferentes métodos são propostos para realizar o reconhecimento de palavras e letras em língua de sinais. Alguns trabalhos utilizaram imagens e vídeos para reconhecer gestos em Libras (CARNEIRO et al., 2016; FILHO et al., 2016; TEODORO; BERNARDES; DIGIAMPIETRI, 2017) e na língua americana de sinais (SAHOO; ARI; GHOSH, 2018; XIE; HE; LI, 2018).

A abordagem baseada em sensores é uma alternativa que exige processamento simples e rápido para o reconhecimento de movimentos e gestos, porém, ao mesmo tempo, pode ser uma alternativa de custo elevado e desconfortável para o usuário (GHANEM; CONLY; ATHITSOS, 2017). Pode-se perceber que a maioria dos sistemas de reconhecimento de gestos para língua de sinais utilizam sensores inerciais. Em conjunto com os inerciais, outros tipos de sensores, como flexíveis e eletrodos para captação de Eletromiografia (EMG), são utilizados para identificar a configuração de mão utilizada na execução do gesto.

Os sinais de EMG têm sido bastante empregados. Por exemplo, utilizando o dispositivo comercial *armband Myo*, alguns trabalhos têm a proposta de classificar palavras e letras em Libras (ABREU et al., 2016; MENDES JUNIOR et al., 2019) e na língua americana de sinais (JANE; SASIDHAR, 2018; WU; SUN; JAFARI, 2016).

Os sensores flexíveis e inerciais são utilizados em diversos trabalhos que envolvem reconhecimento de gestos da língua de sinais de diferentes países. Na

pesquisa realizada não foi encontrado um produto comercial específico para aplicações de reconhecimento de gestos da língua de sinais. Existem luvas instrumentadas comerciais utilizadas para aplicações de realidade virtual e animação, como a *CyberGlove II*, que possui sensores flexíveis (CYBERGLOVE, 2017). O valor da *CyberGlove* não é informado pelo fabricante. A Tabela 1 apresenta os trabalhos publicados recentemente sobre o uso dos sensores flexíveis e inerciais. Os trabalhos considerados mais relevantes, listados na Tabela 1 são apresentados a seguir.

Shaheen e Mehmood (2018) propõem reconhecer as letras e números da língua de sinais do Paquistão utilizando sensores flexíveis (um sobre cada dedo), acelerômetro e giroscópio de três eixos e sensor de contato (feito com folha de cobre). Os sensores flexíveis são utilizados para a leitura da curvatura dos dedos; o sensor de contato é utilizado para diferenciar abdução (afastamento lateral) dos dedos indicador e médio e o acelerômetro e o giroscópio são utilizados para detectar a orientação da mão e movimentos, respectivamente.

Para o experimento, 10 voluntários realizaram 36 gestos (26 letras do alfabeto e 10 números). Cada gesto foi executado 20 vezes, com tempo de duração de aproximadamente 20 segundos e, por meio de um microcontrolador, os dados foram mapeados. Para o mapeamento dos dados, a orientação e movimento da mão, a média dos valores dos sensores flexíveis para cada gesto e a abdução dos dedos indicador e médio foram considerados. A orientação e o movimento da mão foram definidos por meio de valores de limiar definidos pelo autor. A taxa média de classificação para os 36 gestos obtida com o mapa de gestos foi de 93,4%.

Kau e seus colaboradores (2015) desenvolveram um sistema para reconhecer algumas palavras da língua de sinais taiwanesa. O sistema portátil (formado por duas luvas) possui 5 sensores de flexão em conjunto com 1 sensor inercial (acelerômetro e giroscópio de 3 eixos) instalados em cada luva. Os sensores flexíveis são utilizados para realizar a leitura da postura dos dedos, o acelerômetro é utilizado para monitorar a orientação da palma da mão e o giroscópio é utilizado para monitorar a trajetória do gesto realizado pela mão. Um detalhe importante nesse trabalho é a simplicidade do processamento, em que o sistema realiza a leitura dos sensores (referente as duas mãos) e verifica os seguintes parâmetros: postura dos dedos, orientação da palma da mão e trajetória do gesto. A partir de valores de limiar, esses parâmetros são codificados em uma informação de 20 bits que auxilia na discriminação dos gestos, conforme apresenta a Tabela 2.

Tabela 1 - Trabalhos que utilizam sensores flexíveis, inerciais e de contato posicionados nas mãos para reconhecimento de gestos em língua de sinais.

País	Objetivo	Quantidade de Sinais	Voluntários	Sensores	Classificador	Taxa de Acerto	Referência
Língua de Sinais do Paquistão	Letras e números (uma mão)	36	10	Flexíveis, Inercial e de contato (folha de cobre)	Mapa de gestos	93,4%	(SHAHEEN; MEHMOOD, 2018)
Língua de Sinais Americana	Letras (uma mão)	26	5	Inercial	Análise de Discriminantes Lineares (LDA)	85%	(ABUALOLA et al., 2016)
Língua de Sinais de Taiwan	Palavras (duas mãos)	5	5	Flexíveis e Inercial	Mapa de gestos	94,56%	(KAU et al., 2015)
Língua de Sinais	Números (uma mão)	10	10	Flexíveis	K-Vizinhos Próximos (KNN)	85%	(MONTALVO et al., 2018)
Língua de Sinais Americana	Letras (uma mão)	26	---	Flexíveis e Inercial	<i>Dinamic Time Warping e Nearest Mapping</i>	96,5%	(JANI; KOTAK; ROY, 2018)
Língua de Sinais Americana	Letras (uma mão)	26	12	Flexíveis, inercial e de contato (pressão)	Máquinas de Vetores de Suporte (SVM)	98,2%	(LEE; LEE, 2018)
Língua de Sinais Americana	Letras (uma mão)	---	---	Flexíveis	Mapa de gestos	---	(JADHAV; TRIPATHY, 2018)
Língua de Sinais Americana	Palavras (uma mão)	9	1	Flexíveis e acelerômetro	---	---	(AMBAR et al., 2018)
Língua de Sinais Americana	Letras (duas mãos)	---	---	Flexíveis	---	---	(PUROHIT; PATIDAR; KUSHWAH, 2017)
Língua de Sinais Indiana	Palavras (duas mãos)	---	---	Flexíveis e Inercial	Mapa de gestos	---	(REWARI et al., 2018)
Língua de Sinais da Indonésia	Letras e números (uma mão)	36	---	Flexíveis e Inercial	Mapa de gestos	91%	(HAQ; SUWARDIYANTO; HUDA, 2018)
Língua de Sinais	Letras (uma mão)	---	---	Flexíveis e Inercial	Redes Neurais	95,8%	(SILVA et al., 2017)

Fonte: Autoria própria.

Tabela 2 – Informação de 20 bits que auxilia na discriminação dos gestos.

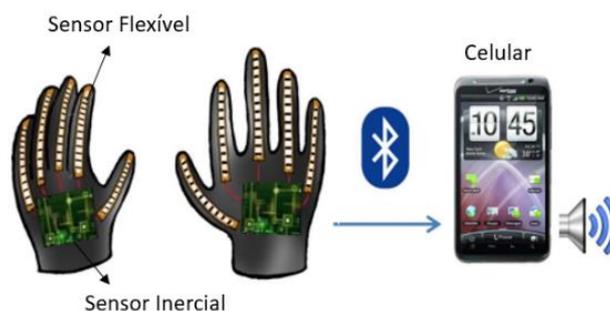
Postura		Orientação da palma		Trajetória do movimento	
10 bits		6 bits (duas mãos)		4 bits (duas mãos)	
Mão esquerda	5 bits	baixo	001	Nenhum	00
		cima	110		
		trás	010	Horário	01
Mão direita	5 bits	frente	101		
		direita	100		
		esquerda	011		

Fonte: Adaptado de (KAU et al., 2015).

Cada dedo recebe apenas dois estados (1 bit): curvado ou não curvado. Dessa forma, para os dez dedos (referente as duas mãos) têm-se a informação de 10 bits. Para determinar a orientação da palma da mão esquerda, verifica-se se ela está voltada para cima, para baixo ou para trás. Já para determinar a orientação da mão direita, verifica-se se sua orientação está voltada para frente, para esquerda ou para a direita. Assim, a orientação da palma das mãos possui a informação de 6 bits. Por fim, para determinar a trajetória da mão esquerda e da mão direita, tem-se uma informação de 4 bits, onde verifica-se se houve rotação da mão e se é no sentido horário ou anti-horário.

Para a classificação, 5 palavras da língua de sinais taiwanesa foram selecionadas. Cada palavra foi executada 50 vezes por 5 voluntários. Para considerar se o gesto é válido, o sistema monitora a postura das mãos. Se ela não se alterar por um tempo definido, então é provável que é a postura que o voluntário executou. O sistema alcançou 94,56% de taxa média de acertos. A informação do gesto classificado é enviada via Bluetooth para ser apresentada em texto ou áudio em um celular, conforme ilustra a Figura 1.

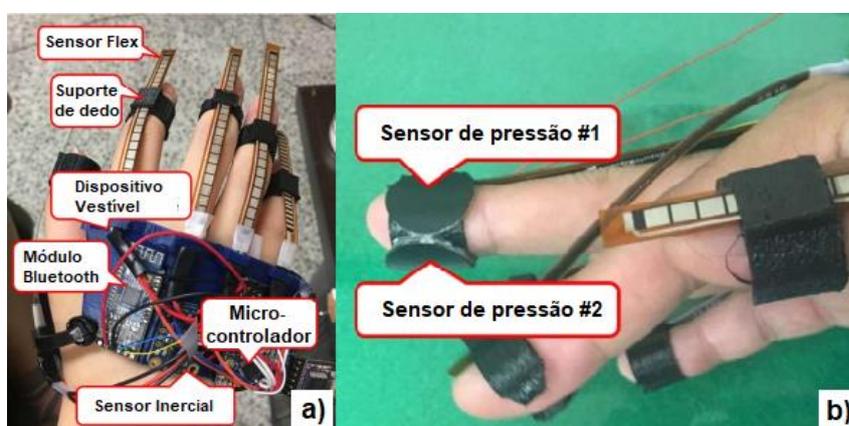
Figura 1 - Luvas instrumentadas para reconhecer gestos da língua de sinais taiwanesa. A informação do gesto reconhecido é enviada para um celular, onde é apresentada.



Fonte: Adaptado de (KAU et al., 2015).

Em Lee e Lee (2018), para reconhecer as letras do alfabeto na língua de sinais americana, foi desenvolvido um dispositivo (impresso em 3D) que fixa os sensores na mão do voluntário que irá executar os gestos. Esse dispositivo, apresentado na Figura 2, possui 5 sensores flexíveis, 1 sensor inercial e, para aumentar a taxa de classificação, 2 sensores de pressão foram adicionados ao sistema. Um sistema microcontrolado é responsável por realizar a leitura dos sensores e a classificação dos gestos.

Figura 2 - Dispositivo vestíveis para reconhecimento das letras do alfabeto na língua de sinais americana. Apresentando: a) Dispositivo vestível com sensores flexíveis com suporte de dedo, sensor inercial, microcontrolador e módulo Bluetooth e b) Sensores de pressão.



Fonte: Adaptado de (LEE; LEE, 2018).

Os sensores de pressão são utilizados para identificar a abdução lateral e a sobreposição dos dedos indicador e médio. Os valores lidos de cada um dos sensores flexíveis são normalizados pelo método *z-score* e separados, de acordo com esses valores, em três regiões de curvatura: não curvado, parcialmente curvado ou totalmente curvado. Esse método de normalização é um diferencial que, de forma simples, visa amenizar a variação dos valores dos sensores causados pelos diferentes tamanhos de mão. Já os valores lidos do sensor inercial (aceleração e giro) são utilizados para calcular os ângulos de inclinação em relação aos eixos *x*, *y* e *z* pelo método do filtro complementar. Para determinar se ocorreu algum movimento da mão durante o gesto, o desvio padrão desses ângulos é calculado.

A coleta de dados foi feita com 12 voluntários. Cada voluntário executou 20 vezes cada gesto (letras do alfabeto), com tempo de duração de aproximadamente 10 segundos. Para montar o vetor de características, foi utilizado uma janela deslizante de 3 segundos para cada 10 segundos de dados coletados. Um aspecto importante

nesse trabalho é a definição do tempo de duração da janela deslizante. Esse tempo de 3 segundos foi definido para conseguir incluir a leitura de gestos que possuem movimento. Dessa forma, foi calculada a média das 10 características (normalização dos 5 sensores flexíveis, desvio padrão dos 3 ângulos e 2 sensores de pressão) no período de 3 segundos.

Para classificar 28 padrões de gestos foram considerados: 26 letras do alfabeto, um padrão de gesto neutro e um padrão de gesto inválido. O padrão de gesto inválido é definido pelas classificações que alcançam menos de 50% de confiança. Para a classificação, o algoritmo Máquinas de Vetores de Suporte (SVM) foi utilizado. O sistema alcançou 98,2% de taxa média de acertos com a utilização dos sensores de pressão e 65,7% sem a utilização dos sensores de pressão. A letra classificada é transmitida via Bluetooth para um dispositivo móvel, que apresenta visualmente e reproduz em áudio.

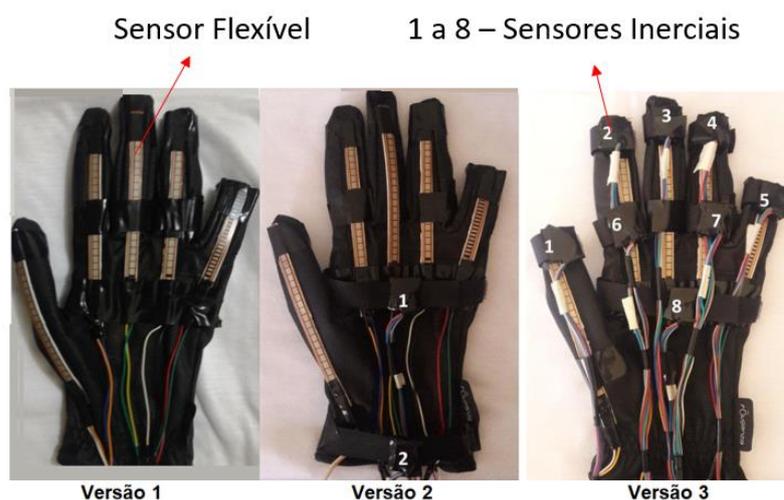
Silva e seus colaboradores (2017) propõem o uso de uma luva com 5 sensores flexíveis e diferentes quantidades de sensores inerciais (acelerômetro e giroscópio) para reconhecer letras e números do alfabeto em língua de sinais. No trabalho não está especificado a qual país a língua de sinais utilizada pertence. Os sensores inerciais foram instalados em diferentes pontos da luva, de acordo com cada versão, conforme apresenta a Figura 3. Na versão 1, a luva possui apenas 5 sensores flexíveis e nenhum sensor inercial. Na versão 2, a luva possui 5 sensores flexíveis e 2 sensores inerciais. Já na versão 3, a luva possui 5 sensores flexíveis e 8 sensores inerciais. Os sensores flexíveis são utilizados para identificar curvaturas dos dedos e os sensores inerciais para identificar movimentos da mão.

A coleta de dados com usuários da língua de sinais foi feita por meio de um sistema embarcado conectado a um computador, porém não foi informado a quantidade de usuários nem o tempo de coleta para cada gesto. Os gestos coletados (100 exemplos de cada) foram letras e números, porém, para a classificação, os gestos que possuem configuração de mão e movimentos parecidos foram removidos. Para montar a matriz de entrada de dados do classificador, os valores dos 5 sensores flexíveis e dos eixos x, y e z do acelerômetro e do giroscópio foram considerados.

A versão 3 obteve o melhor desempenho (taxa média de acertos de 95,8%), seguido da versão 2 (89,6%) e da versão 1 (62,0%). Para a classificação, todas as versões utilizaram Redes Neurais Artificiais (RNA). Esse trabalho contribuiu mostrando que, aumentando o número de sensores inerciais, a taxa de classificação

de gestos pode aumentar, porém o processamento também aumenta. Além disso, não foi utilizado nenhum sensor de contato, que pode contribuir bastante na discriminação dos gestos que possuem abdução dos dedos.

Figura 3 - Luva com sensores flexíveis e inerciais para reconhecimento de Libras. Os números de 1 a 8 representam os sensores inerciais.



Fonte: Adaptado de (SILVA et al., 2017).

Observando os trabalhos correlatos e considerando que um sistema de reconhecimento de gestos deve ser projetado para surdos utilizarem no dia a dia, este trabalho propõe uma técnica de processamento diferente dos trabalhos existentes. O sistema proposto realiza a segmentação de um padrão de pré-gesto, de um padrão de gesto e de um padrão de pós gesto. Após a segmentação, a extração de características dos dados segmentados é realizada e para a classificação, uma RNA é treinada com dados de diferentes voluntários.

Além da técnica de processamento, esse trabalho visa analisar vários aspectos, como: contribuição de cada tipo de sensor na luva, análise das características que melhor discriminam os gestos, contribuição dos sensores em grupo, análise da classificação por voluntário e análise da quantidade de dados separada para treinamento da rede neural.

Neste capítulo foram apresentados os trabalhos correlatos referentes a sistemas automatizados de reconhecimento de padrões de gestos em língua de sinais. No próximo capítulo será apresentada a fundamentação teórica, na qual são abordados os conceitos relacionados à anatomia das mãos, língua de sinais e sensores utilizados no trabalho.

3 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, a fundamentação teórica é apresentada abordando os principais conceitos da anatomia das mãos, da língua de sinais e das tecnologias utilizadas nesse trabalho.

3.1 ANATOMIA DAS MÃOS

Os gestos manuais em língua de sinais são executados com as mãos. Assim, para melhor compreensão deste trabalho, é importante a apresentação dos principais ossos, articulações e movimentos relacionados às mãos. A Figura 4 ilustra os ossos e articulações da mão, bem como o nome de cada dedo, denominado de polegar, indicador, médio, anelar e mínimo.

Figura 4 - Ilustração dos ossos e articulações presentes na mão humana.



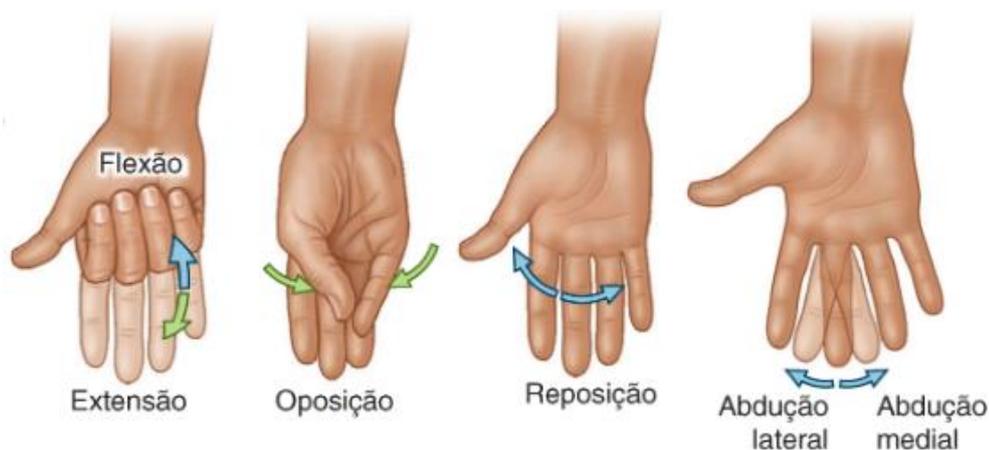
Fonte: Adaptado de (BONTRAGER; LAMPIGNANO, 2015).

As mãos e os punhos são integrantes dos membros superiores do corpo e cada mão e punho juntos possuem 27 ossos, divididos em 3 regiões. A primeira região é o carpo, que é formada pelos ossos do punho; a segunda região é o metacarpo,

formada pelos ossos da palma da mão e, por fim, a terceira região são as falanges, formada pelos ossos do dedo polegar (distal e proximal) e dos demais dedos (distal, medial e proximal). No dedo polegar estão localizadas as articulações metacarpofalangiana (MCF) e interfalangiana (IF). E nos demais dedos estão localizadas as articulações metacarpofalangiana (MCF), interfalangiana proximal (IFP) e interfalangiana distal (IFD) (BONTRAGER; LAMPIGNANO, 2015).

A mão humana possui uma estrutura que possibilita a realização de diversos movimentos em relação aos dedos, como: flexão, extensão, oposição, reposição e abdução, conforme ilustra a Figura 5.

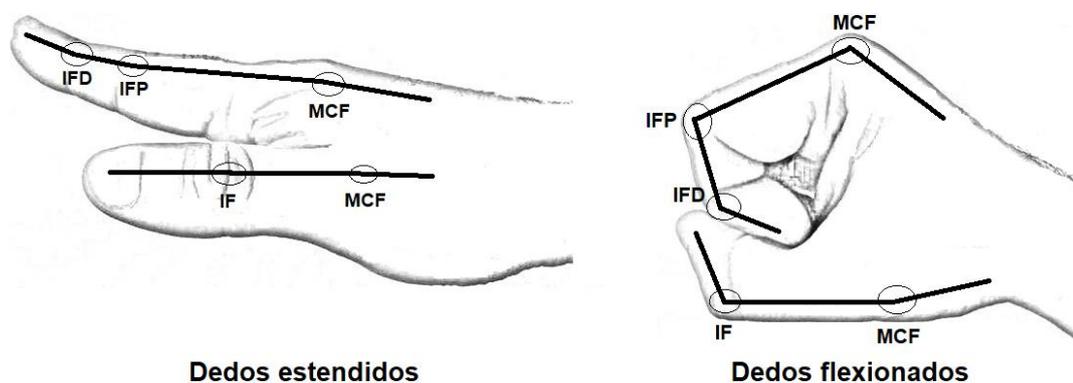
Figura 5 - Ilustração dos principais movimentos realizados pelas mãos. Apresentando a mão direita realizando movimentos de flexão, extensão, oposição, reposição e abdução.



Fonte: Adaptado de (MOOREL; DALLEY; AGUR, 2018).

A flexão e oposição dos dedos é possível devido a presença de articulação entre os ossos; e cada articulação possui um grau de curvatura, que são os ângulos. Conforme ilustra a Figura 6, no dedo polegar existe a formação de dois ângulos na região das articulações IF e MCF; já nos demais dedos existe a formação de três ângulos na região das articulações MCF, IFP e IFD. Se, por exemplo, a mão estiver fechada (todos os dedos curvados), os valores desses ângulos estarão entre aproximadamente 20 e 90°; se os dedos da mão estiverem estendidos, os valores desses ângulos serão próximos de 0°.

Figura 6 - Ilustração das regiões de curvatura dos dedos da mão direita, em que IF é a articulação interfalângiana, MCF é a metacarpofalângiana, IFD é a interfalângiana distal e IFP é a interfalângiana proximal.



Fonte: Autoria Própria.

Monitorando os valores desses ângulos, pode-se identificar a configuração de mão que está sendo executada na realização dos gestos em língua de sinais e, desta forma, discriminar uma grande quantidade de gestos. Porém, como o tamanho dos ossos e articulações são diferentes para cada pessoa, na execução de um mesmo gesto, os valores dos ângulos formados na região das articulações variam. Por isso, nesse trabalho, para discriminar gestos em Libras, dados de gestos coletados com diferentes voluntários foram utilizados para treinar uma rede neural.

3.2 LÍNGUA DE SINAIS

A partir da necessidade natural de comunicação entre pessoas que não utilizam a comunicação oral, surgiram os sistemas linguísticos que são denominados língua de sinais (QUADROS, 2008). A língua de sinais é adquirida pela criança surda de forma espontânea e é considerada uma língua de modalidade espaço-visual, pois os sinais são realizados com o uso das mãos, da face e de outras partes do corpo, e a informação é recebida pelos olhos (BAKER, 2011; DIZEU; CAPORALI, 2005; QUADROS; KARNOPP, 2007). A língua de sinais é diferente para cada país e possui gramática própria, organizada e complexa, o que torna possível transmitir qualquer assunto (RAMOS, 2002).

Os surdos não são necessariamente mudos, pois as causas que levam à deficiência auditiva não estão diretamente ligadas às causas da deficiência da fala (ALMEIDA, 2013). A afonia, que pode ser absoluta ou parcial, é um distúrbio da linguagem falada causada por fatores fisiológicos. Quando a causa tem origem na

deficiência auditiva, a pessoa possui deficiência auditiva e também da fala (POLLMANN, 1994).

3.2.1 Língua Brasileira de Sinais

No Brasil, o início da língua de sinais é marcado pela criação do Instituto Nacional de Educação dos Surdos (Ines), em 26 de setembro de 1857. Nessa época, apesar da existência do Ines, o uso da língua de sinais chegou a ser proibido em sala de aula, ficando permitida apenas a língua oral. Em 1980, com base nas pesquisas da professora Lucinda Ferreira Brito, o bilinguismo foi finalmente iniciado no Brasil (GOLDFEALD, 2002). A lei nº 10.436 determina a Libras como meio de comunicação e expressão, originário de comunidades de pessoas surdas do Brasil (BRASIL, 2002).

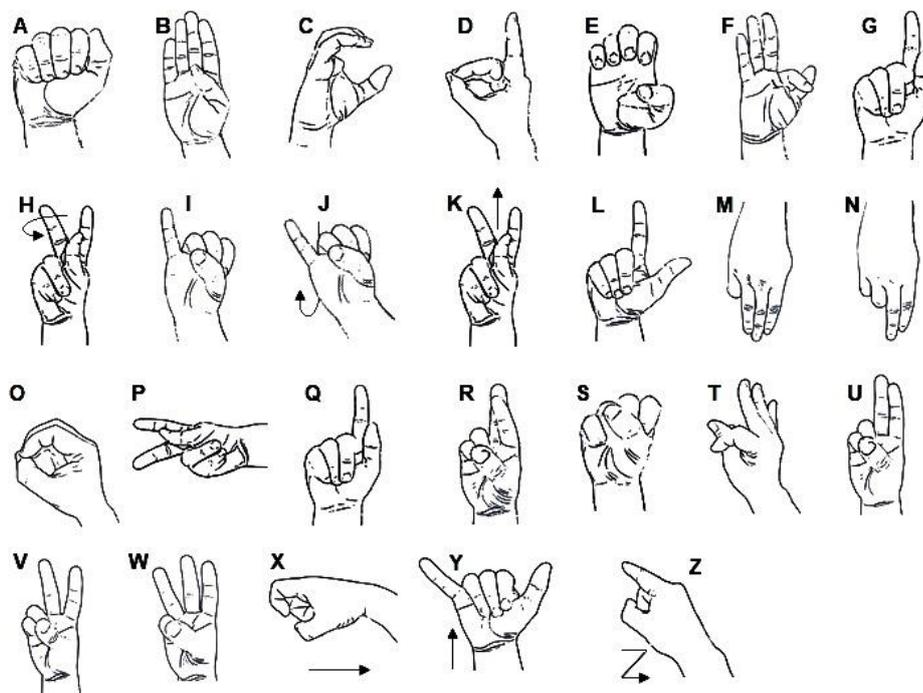
Os sinais em Libras podem ser executados de forma manual (com o uso das mãos) e de forma não manual (sem o uso das mãos). A forma manual pode ser executada com apenas uma das mãos ou com as duas mãos. Já a forma não manual utiliza a face para executar os sinais (JEPSEN et al., 2015). Na forma não manual, as expressões faciais e corporais podem representar diferentes significados, como afirmação, negação, interrogação ou emoção (KUMADA, 2016).

Os sinais manuais em Libras são estruturados a partir de cinco parâmetros e a partir da combinação deles, diferentes sinais podem ser realizados. O primeiro parâmetro é a configuração das mãos, que representa o formato das mãos durante a realização do gesto. O segundo parâmetro é a localização, que se refere a qual região do corpo o sinal é executado. O terceiro parâmetro é o movimento, que é o percurso de movimento que os dedos ou o punho realizam durante o gesto. O quarto parâmetro é a orientação das palmas das mãos, e indica a direção que a palma da mão aponta durante o gesto. O quinto e último parâmetro refere-se a traços não manuais, nos quais a expressão facial, o movimento do corpo e do olhar são considerados (PEREIRA et al., 2011).

O alfabeto manual é geralmente utilizado para soletrar nomes de pessoas e de lugares e palavras que não têm sinal definido (STROBEL; FERNANDES, 1998). Para a execução do alfabeto manual, apenas uma mão é utilizada, a velocidade de soletração é similar ao da fala e a localização do sinal é na região peitoral (PEREIRA et al., 2011). O alfabeto manual possui um total de 26 sinais, sendo que 6 possuem o parâmetro movimento, denominados dinâmicos, e 20 não possuem o parâmetro

movimento, denominados estáticos, como ilustrado na Figura 7. As letras com movimento são as marcadas com uma seta que indica uma ação.

Figura 7 – Alfabeto manual em Libras. Ilustração das 26 letras, divididas em 20 sinais estáticos e 6 dinâmicos.



Fonte: Autoria Própria.

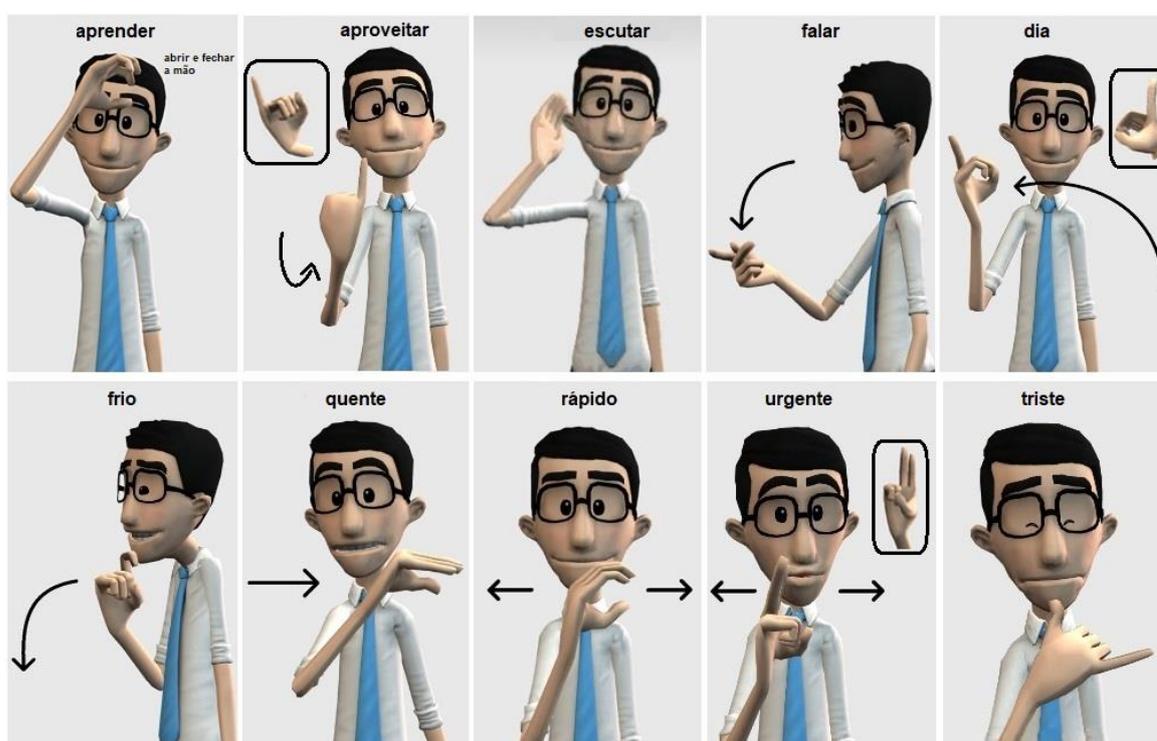
No alfabeto em Libras, com exceção dos traços não manuais, todos os parâmetros citados acima podem ser identificados. As 6 letras que possuem o parâmetro movimento são: 'H', 'J', 'K', 'X', 'Y' e 'Z'. Dentre essas letras, existe o movimento do tipo linear e o movimento do tipo rotacional, em que ambos são executados pela mão como um todo. O movimento das letras 'K', 'X', 'Y' e 'Z' são do tipo linear e com direções diferentes. Já o movimento das letras 'H' e 'J' são do tipo rotacional e com direções diferentes também. A direção que a mão aponta durante a realização do gesto também é muito importante em sua identificação. As letras 'N' e 'U' são idênticas e o único parâmetro que as diferencia é a direção que a mão aponta.

Em todas as letras do alfabeto pode-se perceber algum tipo de flexão, extensão, oposição ou abdução dos dedos, que altera a configuração da mão no momento do sinal. Na letra 'E', por exemplo, todos os dedos ficam flexionados; na letra 'C' todos os dedos ficam parcialmente flexionados e na letra 'B' apenas o dedo polegar permanece flexionado, enquanto os demais dedos ficam estendidos. A

oposição dos dedos pode ser percebida, na letra 'M', porém, nesse caso, existe a combinação de oposição e flexão dos dedos polegar e mínimo. A abdução dos dedos ocorre em diversas letras. Na letra 'U', por exemplo, existe a abdução dos dedos indicador e médio. Existe também um tipo de sobreposição que ocorre entre os dedos, que é o caso da letra 'R', em que o dedo médio fica posicionado sobre o dedo indicador e ambos ficam estendidos.

Na língua de sinais, existem sinais diferentes que são executados da mesma forma, porém apenas algum parâmetro é diferente, como a localização na qual o sinal é feito (MORAIS et al., 2018). Nos gestos de palavras em Libras, os cinco parâmetros citados anteriormente podem ser identificados, inclusive os traços não manuais, que se refere a expressão facial, movimento do corpo e olhar. A Figura 8 apresenta dez palavras em Libras frequentemente utilizadas no dia a dia do surdo.

Figura 8 – Gestos de palavras em Libras selecionados para o trabalho.



Fonte: Adaptado de (HAND TALK, 2020).

Pode-se observar que alguns desses gestos possuem a mesma configuração de mãos de alguma letra do alfabeto, porém a orientação, a localização e/ou o tipo de movimento são diferentes. O sinal 'triste', por exemplo, apresenta traços não manuais de tristeza e possui a mesma configuração de mão da letra 'Y'. O sinal 'frio', apresenta

movimento de tremor com a cabeça, expressão facial de sofrimento e possui a mesma configuração de mão da letra 'X', com movimento do tipo linear em frente a região do queixo. O sinal 'aprender' apresenta movimento apenas dos dedos: flexão e extensão durante a execução do gesto.

3.3 REDES NEURAIS ARTIFICIAIS

As RNA constituem uma área da Inteligência Artificial (IA) de grande desenvolvimento e consistem em modelos computacionais com sistema de processamento paralelo capazes de adaptar, aprender, generalizar, agrupar e organizar dados. Uma RNA é composta pela conexão entre unidades de processamento que se comunicam e são inspiradas em neurônios biológicos (LIMA; PINHEIRO; SANTOS, 2014).

A RNA geralmente é implementada com sistemas eletrônicos ou programada em computador. Os neurônios (que são as unidades de processamento) armazenam conhecimento experimental pelo denominado processo de aprendizagem. No processo das RNA, as conexões entre os neurônios artificiais, que são denominados de pesos sinápticos, são utilizadas para armazenar o conhecimento. O processo de aprendizagem é realizado por meio de um procedimento denominado algoritmo de aprendizagem, que modifica os valores dos pesos da rede de forma ordenada (HAYKIN, 2007).

O modelo da RNA mais simples, constituído por apenas um neurônio artificial, é formado por um conjunto de sinais de entrada X_n , pesos sinápticos W_n , combinador linear Σ , limiar de ativação (ou *bias*) θ , potencial de ativação u , função de ativação $g()$ e sinal de saída y . O conjunto de sinais de entrada é proveniente do meio externo e representa as variáveis da aplicação. Os pesos sinápticos são utilizados para ponderar cada variável do sinal de entrada. O combinador linear combina todas as variáveis de entrada já ponderadas para produzir o potencial de ativação. O limiar de ativação é a fronteira utilizada para gerar um valor de disparo para produzir o potencial de ativação. O potencial de ativação é o valor da diferença entre o resultado do combinador linear e o limiar de ativação. Por fim, a função de ativação limita a saída do neurônio em um intervalo conhecido (SILVA; SPATTI; FLAUZINO, 2010).

A Figura 9 apresenta o modelo de um neurônio simples, em que cada um dos sinais de entrada, proveniente do meio externo, é multiplicado pelo seu respectivo

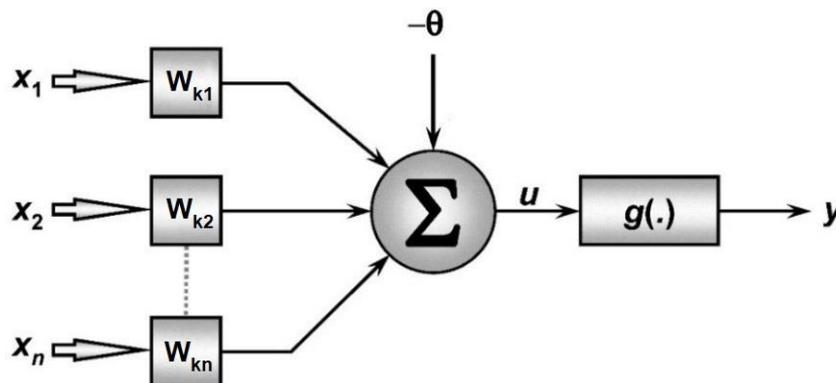
peso sináptico e subtraído do limiar de ativação. O resultado passa pelo combinador linear, que realiza a soma ponderada e, depois de subtraído do limiar de ativação, produz o potencial de ativação (u), dado por

$$u = \sum_{i=1}^n W_{ki} \cdot X_{ki} - \theta . \quad (1)$$

Por fim, o resultado apresentado pela equação 1 passa pela função de ativação, que produz a saída (y), ou seja

$$y = g(u) . \quad (2)$$

Figura 9 - Modelo de um neurônio artificial simples.

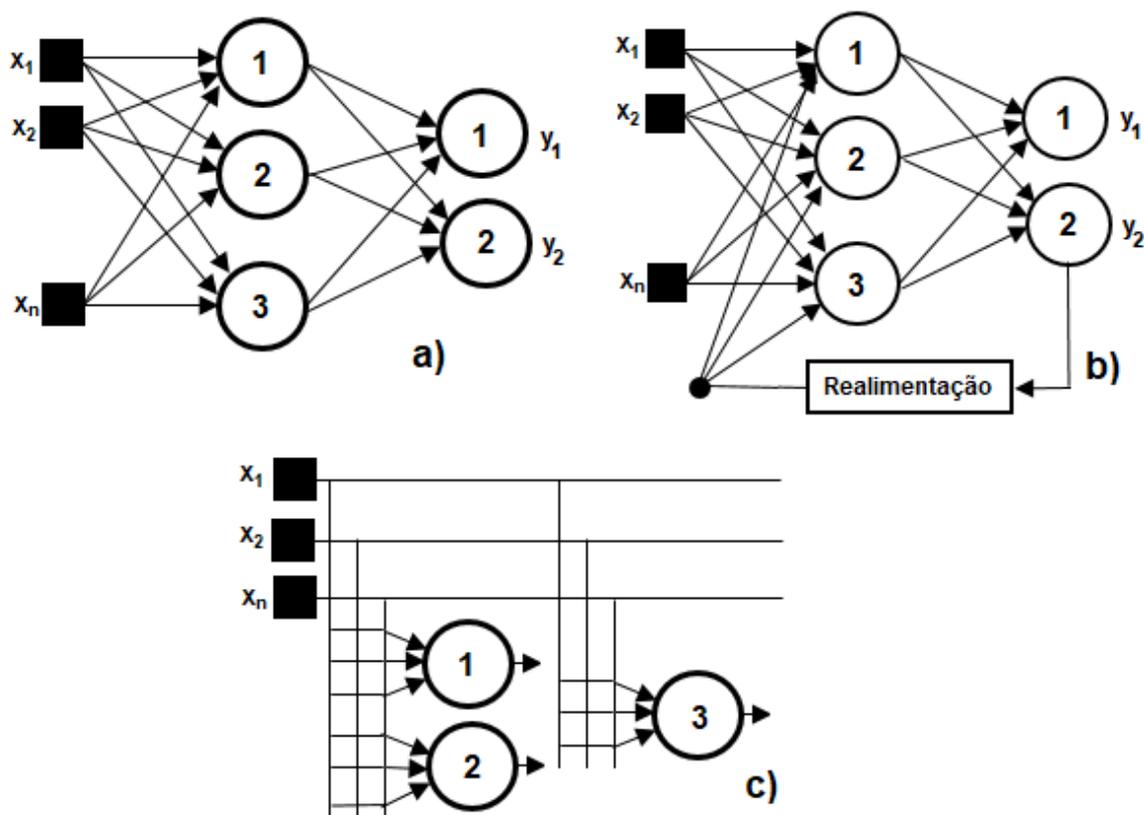


Fonte: Adaptado de (SILVA; SPATTI; FLAUZINO, 2010).

A arquitetura de uma RNA consiste na forma como os neurônios estão arranjados e interligados. De forma simples, uma rede neural pode ser dividida em três partes: camada de entrada, camada intermediária e camada de saída. Quando o fluxo de informações percorre a rede em uma única direção, desde a camada de entrada até a camada de saída, a rede é denominada *feedforward*. Uma das principais aplicações para as redes *feedforward* é o reconhecimento de padrões. Quando o fluxo de informações da camada de saída é retroalimentado a alguma camada anterior, a rede é denominada recorrente ou realimentada e é utilizada em aplicações que envolvem sistemas variantes em relação ao tempo, como previsão de séries temporais. E quando os sinais de entrada são inseridos em todos os neurônios, a rede é denominada rede com estrutura reticulada e podem ser utilizadas em diversas

aplicações como otimização de sistemas e grafos (SILVA; SPATTI; FLAUZINO, 2010). A Figura 10 apresenta exemplos de redes do tipo *feedforward*, realimentada e com estrutura reticulada.

Figura 10 - Exemplo de rede neural com arquitetura: a) *feedforward*, b) realimentada e c) com estrutura reticulada.



Fonte: Autoria Própria.

As redes neurais são capazes de generalizar soluções, ou seja, aprender a partir dos dados de entrada e relacioná-los à saída. Desta forma, uma rede neural treinada recebe os dados de entrada e é capaz de apresentar uma resposta próxima da desejada. Para realizar o treinamento da rede a fim de que ela alcance seus objetivos, é necessário que o algoritmo de treinamento ajuste os valores dos pesos sinápticos e dos limiares de ativação até alcançar a generalização da solução (SILVA; SPATTI; FLAUZINO, 2010).

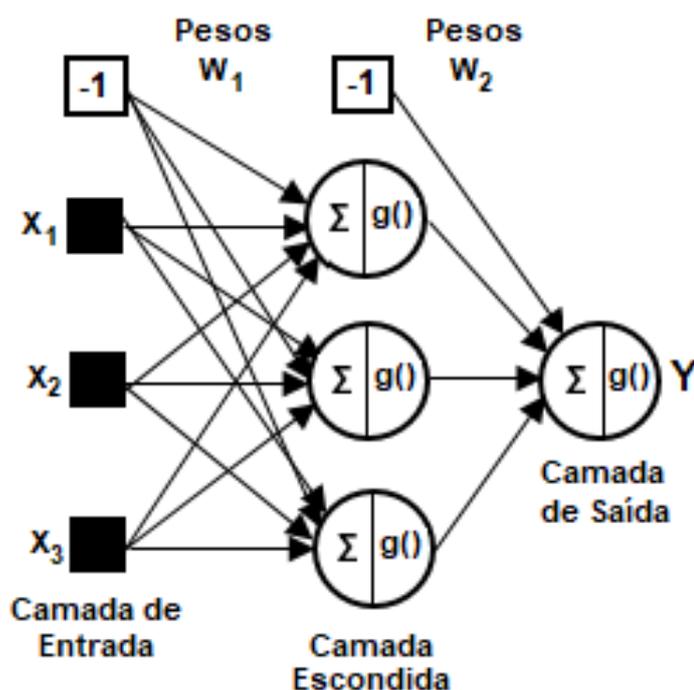
A apresentação completa de um conjunto de treinamento durante o processo de aprendizagem é denominada de época de treinamento. O tipo de aprendizado de uma rede pode ser supervisionado ou não supervisionado. No aprendizado

supervisionado, para cada exemplo de entrada, a saída é conhecida (rotulada). E no aprendizado não supervisionado, o algoritmo de treinamento não utiliza exemplos rotulados para ajustar os pesos sinápticos (FACELI et al., 2011; HAYKIN, 2007).

3.3.1 Redes Perceptron Multicamadas

As Redes Neurais Perceptron de Múltiplas Camadas (MLP) do tipo *feedforward* fazem parte de uma importante classe de redes neurais com treinamento supervisionado e são utilizadas em diversas aplicações como: reconhecimento de padrões, aproximação de funções, entre outras. Tradicionalmente, elas apresentam uma camada de entrada, uma ou mais camadas intermediárias (escondida), e uma camada de saída (HAYKIN, 2007; SILVA; SPATTI; FLAUZINO, 2010). A Figura 11 apresenta o exemplo de uma rede neural MLP formada por: uma camada de entrada com 4 entradas, uma camada escondida com 3 neurônios e uma camada de saída com 1 neurônio. Os limiares de ativação possuem valor -1 em suas entradas e as ligações entre as camadas são os conjuntos de pesos W_1 (que possui 12 pesos) e W_2 (que possui 4 pesos).

Figura 11 - Exemplo de uma Rede Neural MLP com três camadas: camada de entrada com 4 dados, camada escondida com 3 neurônios e camada de saída com 1 neurônio.



Fonte: Autoria Própria.

No processo de treinamento da MLP é utilizado o algoritmo *backpropagation*, que é baseado em dois passos: passo para frente (*forward*) e passo para trás (*backward*). No passo *forward*, a resposta da rede é obtida a partir da propagação dos sinais referentes às amostras de treinamento inseridas na rede. Neste passo, os pesos iniciais da rede permanecem inalterados. No passo *backward*, o erro para cada saída da rede, calculado a partir da comparação entre a resposta da rede (obtida na fase *forward*) e a resposta da rede desejada, é utilizado para fazer os ajustes dos pesos sinápticos e limiares de todos os neurônios da rede. Para minimizar o erro de treinamento um método bastante utilizado é o gradiente descendente, que deriva a função de ativação das camadas da rede. Desta forma, o processo de treinamento, utilizando o algoritmo *backpropagation*, busca, a cada iteração, o ajuste dos pesos sinápticos e limiares dos neurônios, a fim de minimizar a soma dos erros (HAYKIN, 2007; SILVA; SPATTI; FLAUZINO, 2010).

Na etapa *forward*, como o fluxo de informações é para frente, deve-se primeiro calcular o vetor de saída da camada escondida Y_1 e em seguida o vetor de saída da camada de saída Y_2 . Para calcular Y_1 deve-se realizar a multiplicação entre os dados de entrada X e o vetor de pesos W_1 transposto, e o resultado deve passar pela derivada da função de ativação $g'()$. Da mesma forma, Y_2 é calculado por meio da multiplicação entre o vetor Y_1 e o vetor de pesos W_2 transposto, e o resultado deve passar pela derivada da função de ativação $g'()$. Os vetores Y_1 e Y_2 , respectivamente, são encontrados por

$$Y_1 = g'(X \cdot W_1^T) , \quad (3)$$

$$Y_2 = g'(Y_1 \cdot W_2^T) . \quad (4)$$

Na etapa *backward*, para realizar o ajuste dos pesos utilizando o método do gradiente, deve-se calcular o gradiente local do erro da camada de saída, dado por

$$\delta_2 = (D - Y_2) \cdot g'(Y_1 \cdot W_2^T) \quad (5)$$

em que δ_2 é o gradiente local do erro da camada de saída; e D é a saída desejada.

Em seguida, a matriz de pesos W_2 ajustada é calculada por

$$W_2 = W_2 + \eta \cdot \delta_2 \cdot Y_2 \quad (6)$$

em que η é a taxa de aprendizagem.

Da mesma forma, pode-se calcular o gradiente local do erro da camada escondida δ_1 e o ajuste dos pesos da camada de escondida W_1 , respectivamente, por

$$\delta_1 = (\delta_2 \cdot W_2) \cdot g'(X \cdot W_1^T) \quad , \quad (7)$$

$$W_1 = W_1 + \eta \cdot \delta_1 \cdot X \quad . \quad (8)$$

Para realizar o cálculo do gradiente, deve-se conhecer a derivada da função de ativação utilizada na rede MLP. Por isso, a função de ativação deve ser totalmente diferenciável. Dentre as principais funções totalmente diferenciáveis utilizadas na rede MLP, pode-se citar: função logística, tangente hiperbólica, gaussiana e linear (HAYKIN, 2007; SILVA; SPATTI; FLAUZINO, 2010).

Como critério de parada, deve-se calcular o erro quadrático médio E_M , que é utilizado para avaliar o desempenho global do algoritmo *backpropagation*. Dado um conjunto de dados para treinamento, ρ representa o número de amostras (exemplos) do conjunto, D é o valor desejado para cada amostra (rótulo) e Y_2 é o valor de saída da rede para cada amostra, determina-se o erro por

$$E_M = \frac{1}{2 \cdot \rho} \sum_{k=1}^{\rho} \Sigma (D - Y_2)^2 \quad . \quad (9)$$

As etapas *forward* e *backward* devem ser executadas até o erro quadrático médio entre duas épocas sucessivas alcançar um valor suficientemente pequeno pré-determinado, denominado precisão ε . Outra opção é especificar uma quantidade de épocas para que o treinamento seja finalizado (SILVA; SPATTI; FLAUZINO, 2010).

Na prática, do total de amostras disponíveis para a rede, 60 a 90% são utilizadas para o processo de aprendizagem, denominado subconjunto de

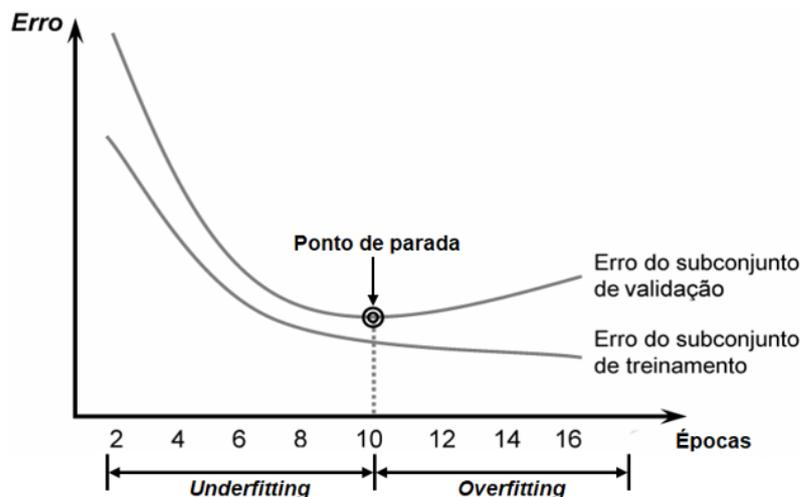
treinamento. As amostras restantes, denominado subconjunto de teste, são utilizadas para verificação da generalização e validação da rede (SILVA; SPATTI; FLAUZINO, 2010). A escolha da topologia da rede é geralmente feita de forma empírica e com apenas uma camada escondida (FACELI et al., 2011). A partir da escolha de algumas topologias candidatas, pode-se utilizar a técnica de validação cruzada para selecionar a melhor. Dentre os métodos de validação cruzada, destaca-se a validação cruzada k -partições, que foi utilizada neste trabalho.

Nesse método, o conjunto total de amostras é dividido em k -partições, sendo $k-1$ para o subconjunto de treinamento e o restante para o subconjunto de teste. Desta forma, a rede neural com a topologia candidata é treinada k vezes a fim de que todas as partições tenham sido utilizadas como subconjunto de teste. Esse processo deve ser executado para todas as topologias candidatas, e a topologia que apresentar o mais alto desempenho global é considerada a melhor (SILVA; SPATTI; FLAUZINO, 2010).

A situação de *overfitting* é caracterizada pela memorização excessiva da rede. Nesta situação, a rede não generaliza corretamente, ou seja, não consegue determinar uma saída de forma adequada frente a uma entrada não apresentada durante a fase de treinamento. Isso pode ocorrer devido ao aumento incontrolado de neurônios e de camadas intermediárias. Nesta condição, a rede apresenta erro quadrático pequeno durante a fase de aprendizado e erro quadrático elevado durante a fase de teste (generalização). Por outro lado, a quantidade insuficiente de neurônios pode levar a rede à situação de *underfitting*, que significa que a rede não aprendeu o suficiente. Nesta condição, a rede apresenta erro quadrático elevado para a fase de aprendizado e de teste (SILVA; SPATTI; FLAUZINO, 2010).

A fim de encontrar a generalização adequada da rede, pode-se utilizar o procedimento de parada antecipada na implementação da técnica de validação cruzada. Esse procedimento consiste em checar o erro do subconjunto de teste durante a etapa de aprendizagem. Quando o erro quadrático do subconjunto de teste aumentar entre épocas sucessivas, a etapa de aprendizagem é encerrada. Desta forma, a parada antecipada procura superar o *underfitting* e evitar a condição de *overfitting* (SILVA; SPATTI; FLAUZINO, 2010). A Figura 12 ilustra o procedimento de parada antecipada, onde pode-se ver que o ponto de parada é quando ocorre o menor erro no subconjunto de validação.

Figura 12 – Momento de parada do processo de aprendizagem da rede neural. Nesse exemplo a parada antecipada ocorre na décima época de treinamento.



Fonte: Adaptado de (SILVA; SPATTI; FLAUZINO, 2010).

3.4 SENSORES

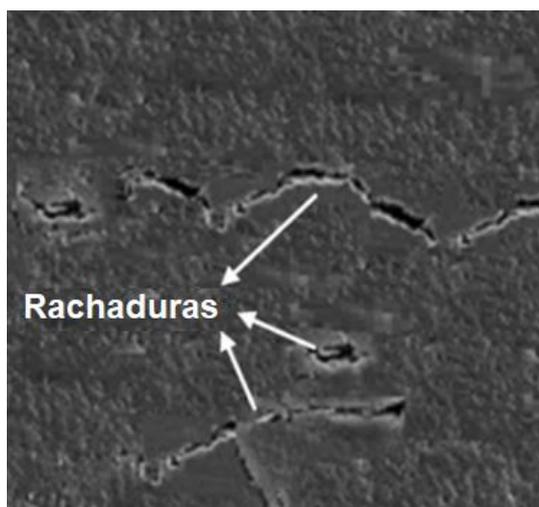
Os gestos manuais em língua de sinais são executados com as mãos, então, para realizar a leitura destes gestos, é necessário utilizar alguns tipos de sensores eletrônicos. Para a detecção da flexão e extensão dos dedos, os sensores flexíveis podem ser utilizados. Alguns gestos são realizados com a mão voltada para cima e outros, com ela voltada para baixo, e dessa forma, um acelerômetro pode ser empregado para detectar a direção da aceleração da gravidade e, por conseguinte, saber a posição da mão. Algumas letras têm característica dinâmica, ou seja, acelerômetro e ou giroscópios podem ser usados para determinar as acelerações e as velocidades desses movimentos. Por fim, algumas letras são caracterizadas por terem os dedos sobrepostos, ou mesmo afastados, por esse motivo sensores de toque ou contato podem ajudar no reconhecimento. A seguir uma descrição detalhada destes sensores eletrônicos será apresentada.

3.4.1 Sensores Flexíveis

O sensor flexível pode ser encontrado em diferentes tipos de sistemas vestíveis, como reconhecimento de gestos, jogos e reabilitação das mãos. O valor de sua resistência elétrica varia de forma proporcional ao grau de flexão em que ele é submetido (RASHID; HASAN, 2018). O sensor flexível é formado por um substrato de

plástico flexível que possui uma camada de tinta impressa. Essa tinta é composta por elementos de carbono resistivo e a flexão do sensor provoca nessa tinta pequenas rachaduras, que se separam de acordo com o grau de flexão, causando a modificação do padrão de condutividade desse material (SAGGIO; ORENGO, 2018). Desta forma, a resistência elétrica medida no material condutivo é proporcionalmente alterada (RASHID; HASAN, 2018). O material condutivo (tinta) utilizado nos sensores flexíveis é composto por partículas de elementos condutores, como carbono ou prata, misturadas com agentes de ligação (ABAS; RAHMAN, 2016). A Figura 13 apresenta a imagem de rachaduras causadas no material do sensor (carbono) que tendem a se separar quando o sensor é submetido à curvatura, causando variação no valor da resistência elétrica medida.

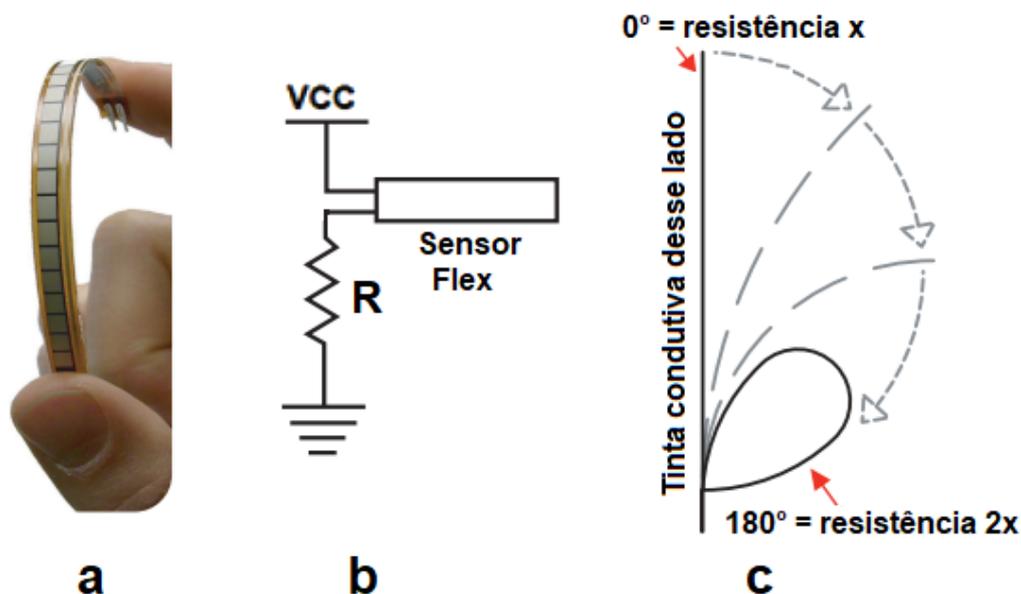
Figura 13 – Rachaduras em escala de micrométrica causadas no carbono devido à curvatura do sensor flexível.



Fonte: Adaptado de (ABAS; RAHMAN, 2016).

Os sensores flexíveis comerciais utilizados nesse trabalho, conforme apresenta a Figura 14(a), são desenvolvidos pela *Spectra Symbol* e estão disponíveis em dois modelos: o modelo com faixa de medição de 9,52 cm e o de 5,6 cm de comprimento. O valor de resistência do sensor varia de acordo com o grau de curvatura que o sensor é submetido. Para realizar a leitura do sensor, pode-se utilizar um circuito divisor de tensão simples (Figura 14(b)). Conforme ilustra a Figura 14(c), o valor de resistência elétrica medido quando o corpo do sensor está totalmente curvado (180°) é aproximadamente o dobro do valor quando o sensor não está curvado (0°).

Figura 14 – Sensor flexível comercial, sendo a) foto do sensor, b) divisor de tensão utilizado para realizar a leitura do sensor e c) ilustração da curvatura do sensor.



Fonte: Adaptado de (SPECTRA SYMBOL, 2014).

3.4.2 Sensores Inerciais

Nos últimos anos, a tecnologia MEMS impulsionou a entrada de sensores inerciais em produtos eletrônicos de consumo, como *smart phones*, *tablets* e *videogames* (SHAEFFER, 2013). A tecnologia MEMS consiste em um processo de fabricação de circuitos integrados que combina elementos mecânicos, sensores, atuadores e microeletrônica em um substrato comum de silício (SILVA, 2007). Essa tecnologia é responsável pela redução das dimensões, produção em grandes quantidades e diminuição de custo de muitos sensores (MAENAKA, 2008). Com essa tecnologia sensores como o acelerômetro e o giroscópio podem estar encapsulados em um único substrato de silício (SHAEFFER, 2013).

3.4.2.1 Acelerômetro

O acelerômetro é um sensor que apresenta sensibilidade à aceleração, à vibração ou ao choque. Seu princípio básico pode ser explicado pela segunda lei de Newton, em que a ação da aceleração em uma massa obtém uma força como resultado (BALBINOT; BRUSAMARELLO, 2019). Conforme o princípio da lei de Hooke, essa força, quando aplicada em uma mola, gera um deslocamento da massa

que é proporcional à aceleração (CUARELLI, 2013). A segunda lei de Newton e a lei de Hooke, respectivamente, são dadas por

$$F = m \cdot a \quad , \quad (10)$$

$$x = \frac{F}{k} \quad (11)$$

em que:

F é a força (em N);

m é a massa (em Kg);

a é a aceleração (em m/s²);

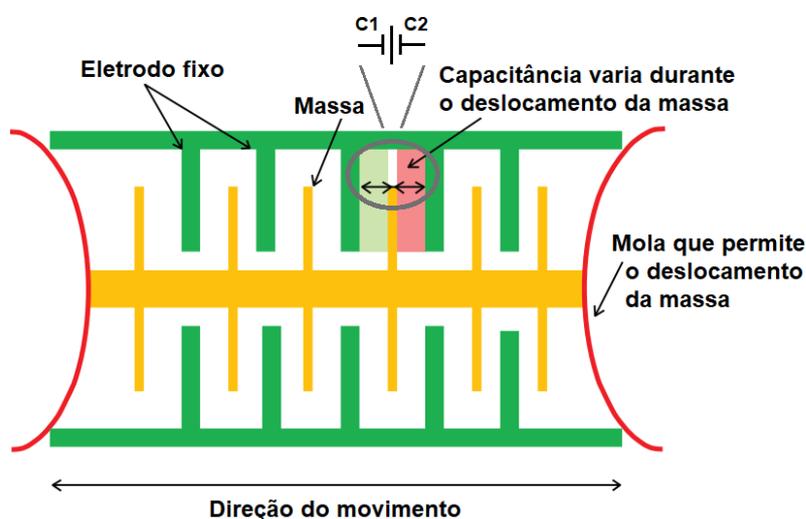
x é o deslocamento da massa (em m); e

k é a constante elástica da mola (em N/m).

A medição é geralmente realizada em unidade g (força da gravidade) e 1g equivale a aproximadamente 9,81 m/s² (BALBINOT; BRUSAMARELLO, 2019). Existem acelerômetros com diferentes tecnologias, como piezoelétrico, piezoresistivos, capacitivos e os MEMS (BALBINOT; BRUSAMARELLO, 2019). No capacitivo, por exemplo, a distância entre placas do capacitor influencia no valor da capacitância e, por isso, pode ser utilizado para fazer a leitura do deslocamento da massa que é causado devido a alguma aceleração.

O acelerômetro de tecnologia MEMS é formado por uma estrutura composta por massa, eletrodos fixos e molas (que permitem mobilidade à massa). Conforme ilustra a Figura 15, no MEMS capacitivo, a aceleração provoca o deslocamento da massa e, conseqüentemente, altera os valores de capacitância medidos entre a massa e o eletrodo (MAENAKA, 2008). Esses valores de capacitância são utilizados para produzir um sinal de saída que é proporcional à aceleração (AMERINI et al., 2016).

Figura 15 – Estrutura de um acelerômetro MEMS, formada por massa, eletrodo e molas. A aceleração causa o deslocamento da massa e altera os valores de capacitâncias C1 e C2 medidos entre a massa e o eletrodo.



Fonte: adaptado de (MASOUD et al., 2019).

A equação para determinar o valor da capacitância de um capacitor é dada por

$$C = \epsilon_0 \cdot \frac{A}{d} , \quad (12)$$

em que:

C é a capacitância (em F);

ϵ_0 é a permissividade dielétrica do vácuo (em F/m);

A é a área das placas (m²); e

d é a distância entre as placas.

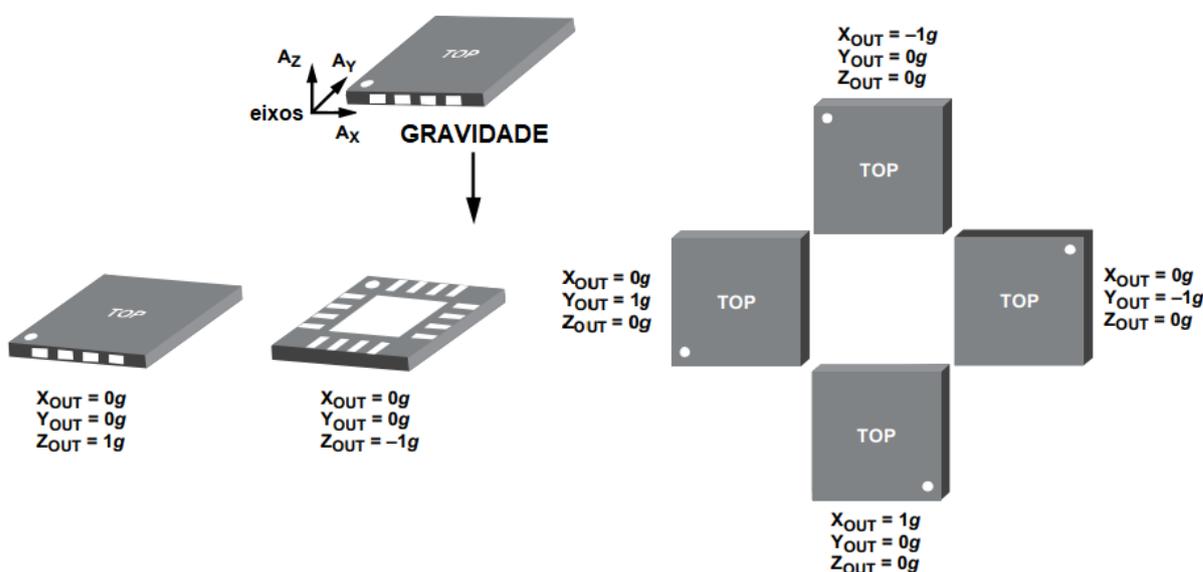
A distância entre as placas de um capacitor influencia no valor da capacitância. Por isso, na ausência de aceleração, não há deslocamento da massa e os valores de capacitância C1 e C2 são muito próximos. Na presença de aceleração, ocorre o deslocamento da massa e os valores de capacitância de C1 e C2 mudam.

Um exemplo de dispositivos MEMS comerciais é o MPU6050 (utilizado nesse trabalho) que possui acelerômetro e giroscópio tri-axiais em um mesmo módulo. O acelerômetro possui uma massa de prova para cada eixo e o deslocamento causado na massa devido a aceleração é detectado por sensores capacitivos. O protocolo de comunicação I²C (do inglês, *Inter-Integrated Circuit*, Circuito Inter-Integrado) é

utilizado para comunicação e o sinal de saída é digitalizado por meio de um conversor analógico-digital de 16 bits. O sinal de saída é fornecido em g e sua escala pode ser ajustada em: ± 2 , ± 4 , ± 8 ou ± 16 g (INVENSENSE, 2013). Se, por exemplo, a escala for ajustada para trabalhar na faixa ± 8 g, então o dispositivo será capaz de realizar a leitura no intervalo de aproximadamente $-78,48$ a $+78,48$ m/s².

Os sinais do acelerômetro indicam a aceleração dinâmica e a aceleração estática de um corpo. A aceleração estática é causada devido a projeção da força gravitacional da terra sobre os eixos do acelerômetro, quando o corpo a ser medido está em repouso. Já a aceleração dinâmica é causada quando o corpo está em movimento, então a saída do acelerômetro apresenta a mistura da aceleração dinâmica com a componente gravitacional (CUARELLI, 2013). Conforme ilustra a Figura 16, se o dispositivo estiver alinhado à uma superfície plana, a saída do eixo perpendicular ao plano irá apresentar 1 ou -1 g, que equivale à aceleração da gravidade da terra (9,8 m/s²), e os demais eixos irão apresentar 0 g.

Figura 16 – Aceleração da gravidade apresentada nos eixos do acelerômetro. A saída do acelerômetro apresenta sempre ± 1 g no eixo que está perpendicular ao sentido da gravidade quando o dispositivo é colocado em uma superfície plana.



Fonte: Adaptado de (ANALOG DEVICES, 2009).

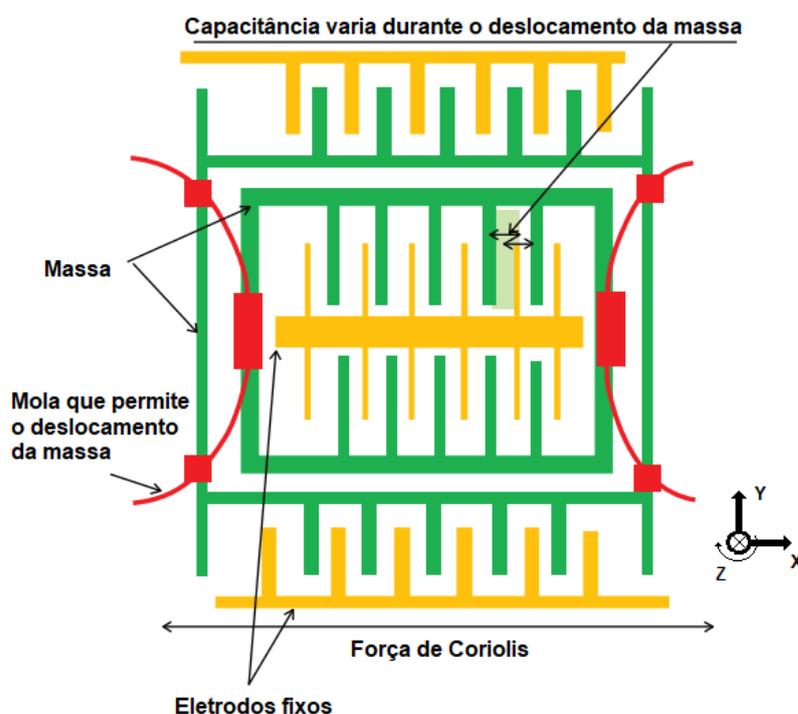
3.4.2.2 Giroscópio

O giroscópio é um dispositivo utilizado para medir a velocidade angular e, por não possuir referência fixa, pode ser utilizado para medir mudanças de estado

(STEVAN JR; SILVA, 2015). De acordo com Shaeffer (2013), a unidade de medida geralmente utilizada é dps (do inglês, *degrees-per-second*, graus por segundo - °/s).

A estrutura de um giroscópio MEMS é parecida com a do acelerômetro MEMS, pois possui massa, eletrodos fixo e molas (que permite o deslocamento da massa) (MASOUD et al., 2019). Porém, no giroscópio, a velocidade angular exercida sobre a massa faz surgir a força de Coriolis, que causa a vibração da massa e é detectada por uma estrutura sensível à variação da capacitância (AMERINI et al., 2016). Desta forma, uma rotação no eixo z do giroscópio, por exemplo, causa o deslocamento da massa no eixo x. A Figura 17 ilustra a estrutura interna de um giroscópio MEMS.

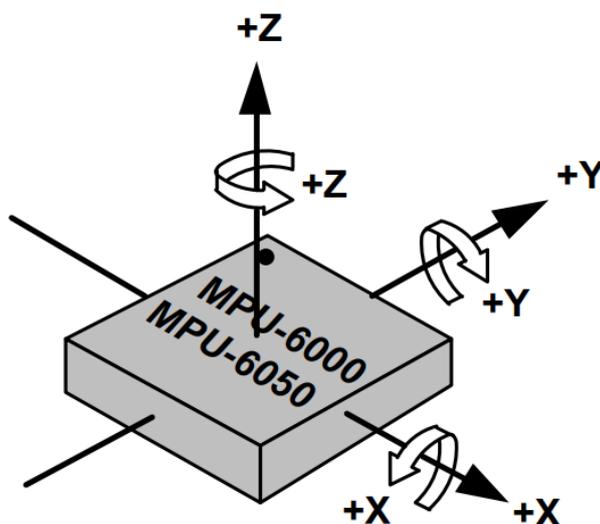
Figura 17 – Estrutura de um giroscópio MEMS, que é formada por massa, eletrodos e molas. A força de Coriolis causa vibração na massa e altera os valores de capacitância medidos entre a massa e o eletrodo.



Fonte: Adaptado de (MASOUD et al., 2019).

O giroscópio presente no dispositivo MEMS MPU6050 detecta rotação em 3 eixos de sensibilidade (x, y e z), conforme apresenta a Figura 18. A vibração da massa, causada devido à força de Coriolis, é detectada por meio de sensores capacitivos e o sinal resultante é amplificado, demodulado e filtrado. Sua saída é digitalizada por um conversor analógico digital de 16 bits e o protocolo de comunicação utilizado é o I²C. A unidade do sinal de saída é °/s e sua escala pode ser ajustada em ±250, ±500, ±1000 ou ±2000 °/s (INVENSENSE, 2013).

Figura 18 – Eixos de sensibilidade do giroscópio MEMS MPU6050.



Fonte: (INVENSENSE, 2013).

3.4.3 Sensores de Contato

Em luvas instrumentadas, sensores de contato feito com tecido condutivo podem ter a função de identificar a abdução ou sobreposição entre dois dedos. Para realizar a mesma função, outros tipos de sensores podem ser utilizados. Em Lee e Lee (2018) sensores de pressão foram utilizados e em Shaheen e Mehmood (2018) sensores feitos com folha de cobre foram utilizados. O tecido condutivo possui a vantagem de ser leve, fino e flexível.

O tecido condutivo Medtex P180, desenvolvido pela *Shieldex*, é formado por 94% de nylon, 6% de elastômero e é revestido com prata pura (99% de pureza). A espessura do tecido é de aproximadamente 0,55 mm e seu peso é de aproximadamente 183 g/m². Esse tecido é utilizado em diferentes aplicações, como roupas médicas e material base para curativos (SHIELDEX, 2018). Como esse tecido possui elastômero e é revestido de prata, ele apresenta características elásticas e baixo valor de resistência elétrica. Desta forma, ele pode ser utilizado como chave de contato em um circuito elétrico. A possibilidade de ser costurado em outros tecidos (procedimento realizado nesse trabalho) torna-se uma boa opção para aplicação com sistemas vestíveis. A Figura 19 apresenta uma foto do tecido costurado sobre a luva nos dedos indicador e médio.

Figura 19 – Sensores de contato feito de tecido condutivo costurado sobre a luva.



Fonte: Aatoria Própria.

Neste capítulo foram apresentados os conceitos relacionados à anatomia das mãos, redes neurais, língua de sinais e de alguns sensores utilizados no trabalho. Desta forma, no próximo capítulo serão apresentados os materiais empregados no desenvolvimento do sistema, bem como a metodologia utilizada nos processos de aquisição, processamento e classificação dos gestos.

4 MATERIAIS E MÉTODOS

Este capítulo apresenta os materiais utilizados no desenvolvimento da luva instrumentada e do sistema de aquisição de dados (versão com fio e versão sem fio). Em seguida, é apresentada a metodologia utilizada na coleta de dados, no processamento e na classificação dos dados.

4.1 LUVA INSTRUMENTADA

Para realizar o reconhecimento dos gestos, desenvolveu-se uma luva instrumentada, que é uma versão aprimorada do trabalho de Dias e Pichorim (2018). Nessa versão aprimorada, foram incluídos na luva o sensor inercial e os sensores de contato. A luva possui características elásticas (fabricada com tecido de Neoprene e Poliéster) e é equipada com diferentes tipos de sensores. A Figura 20 (a) ilustra por meio de um desenho como os sensores estão posicionados na luva e a Figura 20 (b) apresenta uma foto da luva. Na região superior de cada dedo foi posicionado um sensor flexível, totalizando cinco sensores flexíveis; na região do metacarpo do dorso da mão foi posicionado um sensor inercial (acelerômetro e giroscópio tri-axiais) e na região da falange distal e proximal dos dedos indicador e médio foram costurados sensores de contato, sendo um em cada região.

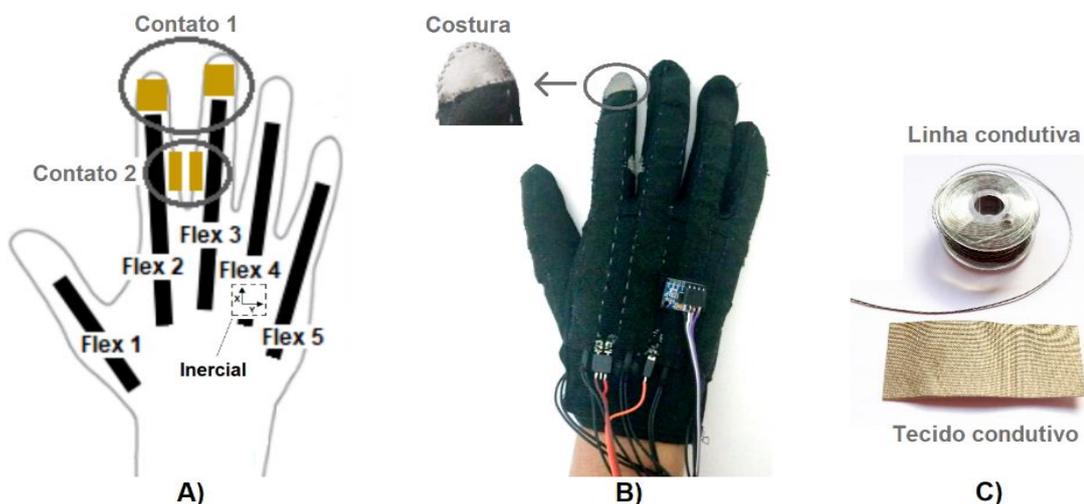
Para costurar os pontos de tecido condutivo sobre a luva, utilizou-se uma linha de costura condutiva. A linha condutiva também foi utilizada para realizar a conexão entre o tecido e os cabos elétricos que levam a informação para o sistema de leitura. A Figura 20 (c) apresenta foto da linha condutiva e do tecido condutivo. Esse procedimento foi realizado a fim de evitar excesso de cabos elétricos ligados à luva, que podem atrapalhar a execução dos gestos.

Os pontos de contato são utilizados para diferenciar sobreposição e abdução lateral dos dedos indicador e médio. Quando, ocorre a abdução lateral dos dedos indicador e médio, ocorre também a abertura e fechamento dos contatos e, dessa forma, o sistema de leitura dos sensores pode identificar.

Os sensores flexíveis, desenvolvidos pela empresa Spectra Symbol, são utilizados para monitorar a flexão dos dedos. Para monitorar as articulações dos dedos indicador, médio, anelar e mínimo, quatro sensores com faixa de medição de 9,52 cm de comprimento foram utilizados. Cada um desses dedos citados possui três

articulações (IFD, IFP e MCF), porém devido ao tamanho dos sensores, apenas duas são monitoradas: IFP e MCF. Para monitorar o dedo polegar que possui apenas duas articulações (IF e MCF), um sensor com faixa de medição de 5,6 cm de comprimento foi utilizado. Circuitos divisores de tensão foram utilizados para realizar a leitura dos sensores e possuem tensão de alimentação de 3,3 V. Para o divisor de tensão do sensor de 5,6 cm, o resistor 47 k Ω foi utilizado e para os sensores de 9,52 cm, resistores de 10 k Ω foram utilizados. Os valores dos resistores foram escolhidos seguindo os trabalhos de Lee e Lee (2018) e Rewari e seus colaboradores (2018).

Figura 20 - Luva Instrumentada com 5 sensores de flexão, 1 sensor inercial e 2 sensores de contato, em que (a) apresenta o posicionamento dos sensores, (b) apresenta a foto da luva e a imagem ampliada do tecido condutivo costurado sobre a luva e (c) apresenta foto da linha condutiva e do tecido condutivo.



Fonte: Autoria Própria.

O sensor inercial MEMS MPU6050, que possui acelerômetro e giroscópio tri-axiais, foi posicionado na região superior do dorso da mão e é utilizado para identificar a direção e os movimentos lineares e de giro que a mão executa durante a realização dos gestos. O sensor MPU6050 é alimentado com tensão de 3,3 V e o protocolo I²C foi utilizado para comunicação. As escalas de sensibilidade do acelerômetro e do giroscópio foram configuradas para trabalhar com ± 16 g e ± 2000 °/s, respectivamente, que são as sensibilidades máximas do sensor.

O custo total da luva desenvolvida é de R\$ 423,11. O sensor inercial MPU6050 custa R\$ 17,90. O sensor flexível com faixa de medição de 5,6 cm de comprimento custa R\$ 44,93. Cada sensor flexível com faixa de medição de 9,52 cm de comprimento tem o valor de R\$ 80,07. O valor do tecido condutivo de 50 mm² é de R\$ 29,00. E, por fim, dois metros de linha de costura condutiva custa R\$ 11,00.

4.2 SISTEMA DE AQUISIÇÃO DE DADOS

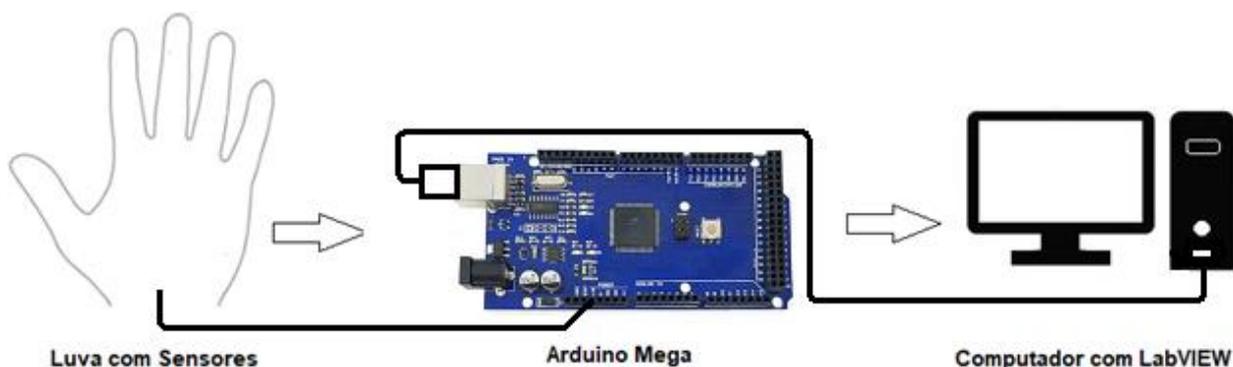
Nessa seção, os dois tipos de sistemas de aquisição (com fio e sem fio) utilizados para enviar os dados da luva instrumentada para o computador são apresentados. O sistema de aquisição com fio foi utilizado para coletar dados durante a realização de gestos do alfabeto em Libras. Já o sistema de aquisição sem fio foi utilizado para coletar dados durante a realização de alguns gestos de palavras em Libras.

4.2.1 Sistema de aquisição de dados com fio

Para realizar a coleta dos dados da luva instrumentada com um sistema com fio, foi utilizado um microcontrolador ATmega2560, que possui protocolo de comunicação I²C e 54 portas de entrada/saída, sendo 16 para leitura de entradas analógicas (ARDUINO, 2020). Para a leitura dos sensores flexíveis foram utilizadas 5 portas analógicas de 10 bits de resolução. Para a leitura do sensor inercial, foram utilizadas duas portas digitais com comunicação I²C. Já para a leitura dos pontos de contato, foram utilizadas duas portas digitais. Os sinais de todos os sensores são amostrados pelo microcontrolador em uma frequência de 50 Hz e enviados para um computador com *software* LabVIEW. A comunicação entre LabVIEW e o Arduino foi realizada utilizando comunicação serial com taxa de transmissão de 19200 bauds, sendo apropriada para essa aplicação sem perda de pacotes e de atraso na sincronização dos dados recebidos pelo sistema de aquisição. A alimentação de todo o sistema foi realizada por meio da porta do Barramento Serial Universal (USB) do computador. A Figura 21 apresenta o processo utilizado para a coleta de dados.

No Arduino, os dados lidos de todos os sensores são organizados em uma mensagem de texto do tipo *string* separada por letras, em que cada letra representa um sensor e o valor após à letra corresponde ao valor lido do sensor. Por exemplo, na mensagem “A200B100C150D250E164aX149aY40aZ1 gX78gY139gZ89R1U0”, as letras A, B, C, D e E indicam os valores dos cinco sensores flexíveis, que variam de 0 a 1023; as letras aX, aY, aZ, gX, gY e gZ indicam os valores brutos dos eixos do acelerômetro (que variam de 0 a 32768) e do giroscópio (que variam de 0 a 32800), respectivamente; já as letras R e U indicam os valores dos sensores de contato, que são de 1 bit cada.

Figura 21 – Processo utilizado na coleta de dados. Os dados referentes aos sensores da luva são lidos pelo microcontrolador Arduino e enviados, via comunicação serial, para o software LabVIEW no computador.

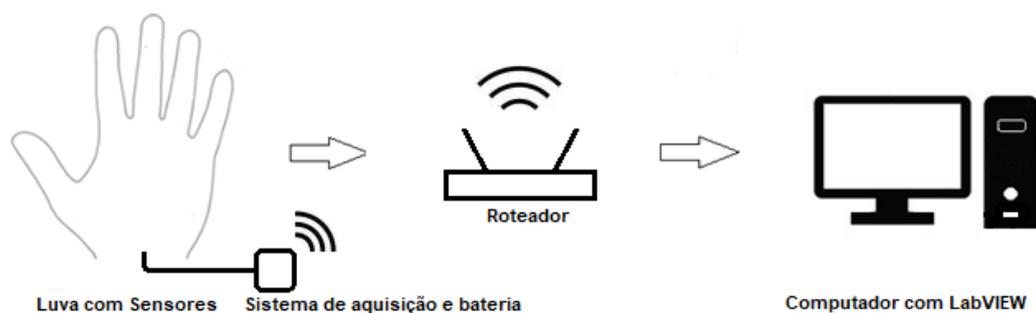


Fonte: Autoria Própria.

4.2.2 Sistema de aquisição de dados sem fio

Para realizar a coleta dos dados com um sistema sem fio, foi utilizado o módulo Esp32 NodeMCU da empresa Espressif, que possui acesso a rede Wi-Fi com padrão 802.11 b/g/n, com taxa de transmissão de até 150 Mbps (ESPRESSIF, 2019). Esse módulo possui um microcontrolador de 32 bits, protocolo de comunicação I²C e 34 portas programáveis de entrada/saída, em que 18 dessas portas são canais de entrada analógica de 12 bits. A Figura 22 apresenta o sistema sem fio utilizado para a coleta de dados.

Figura 22 – Processo utilizado na coleta de dados. Os dados referentes aos sensores da luva são lidos pelo microcontrolador e enviados sem fio para o software LabVIEW, no computador.



Fonte: Autoria Própria.

Para realizar a leitura dos sensores flexíveis, cinco portas analógicas do microcontrolador foram utilizadas. Para a leitura do sensor inercial, duas portas digitais com comunicação I²C foram utilizadas. Já para realizar a leitura dos pontos de contato, duas portas digitais foram utilizadas. Os sinais dos sensores são amostrados a uma

frequência de 50 Hz e organizados em uma mensagem de texto do tipo *string*. Essa mensagem possui o mesmo processo de organização utilizado no sistema de aquisição com fio, descrito na seção anterior. Para enviar os dados para o computador, o Esp32 se conecta a uma rede Wi-Fi (fornecida por um roteador com frequência de 2,4 GHz) cria um servidor *web* local e envia os dados dos sensores para esse servidor. O computador com *software* LabVIEW (conectado à mesma rede Wi-Fi) realiza a leitura dos sensores acessando o servidor *web* e armazena os dados lidos em arquivo de texto. A alimentação do sistema de aquisição e dos sensores foi realizada por meio de uma bateria de 9 V.

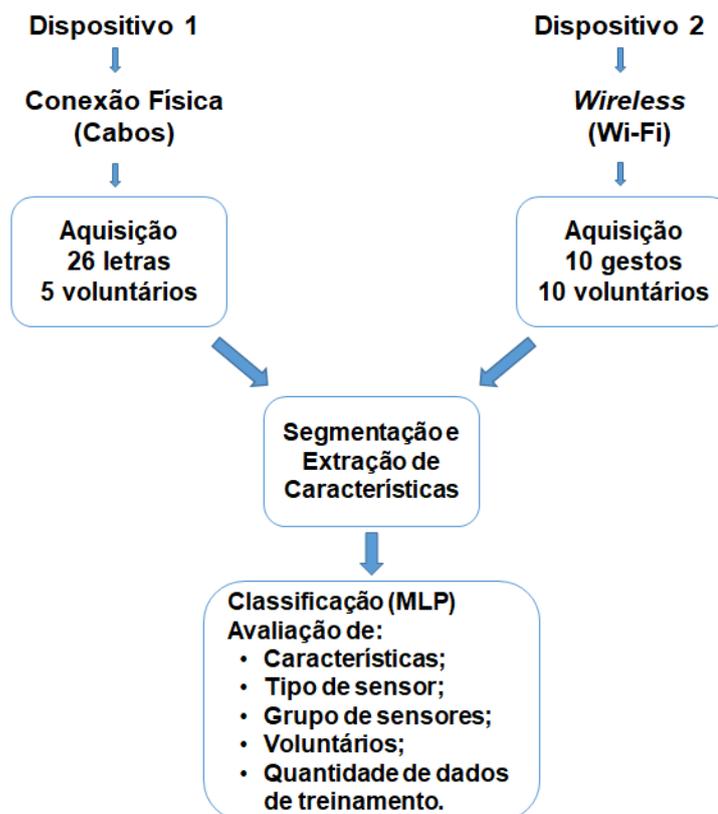
4.4 METODOLOGIA DO PROCEDIMENTO EXPERIMENTAL

O experimento foi aprovado e está dentro dos padrões autorizados pelo Comitê de Ética de Pesquisa em Seres Humanos da Universidade Tecnológica Federal do Paraná - CAAE 30162814.5.0000.5547 (Anexo A). Os gestos em Libras foram coletados dos voluntários e guardados em arquivos de texto para serem tratados (segmentação e extração de características) e classificados. Dois diferentes procedimentos de coleta foram realizados. No primeiro, as letras do alfabeto manual em Libras foram coletadas dos voluntários, utilizando a luva com sistema com fio (versão 1). No segundo procedimento de coleta, alguns gestos de palavras em Libras foram coletados dos voluntários, utilizando a luva com sistema sem fio (versão 2).

Após a etapa de coleta de dados com as duas versões, os dados a serem classificados são submetidos à segmentação e extração de características. Todas as etapas de desenvolvimento do trabalho são descritas na Figura 23. Com o objetivo de conhecer e avaliar sistema de forma mais aprofundada, uma série de análises são realizadas durante as etapas de desenvolvimento e processamento:

- As características que apresentam melhores resultados na classificação;
- Os tipos de gestos que cada sensor consegue identificar;
- Contribuição dos sensores de forma individual e em grupo na classificação;
- Avaliação de cada voluntário na classificação, utilizando o método *leave-one-subject-out*;
- O impacto que a quantidade de dados separada para treinamento da rede neural causa na classificação.

Figura 23 – Fluxograma das etapas de desenvolvimento da pesquisa.



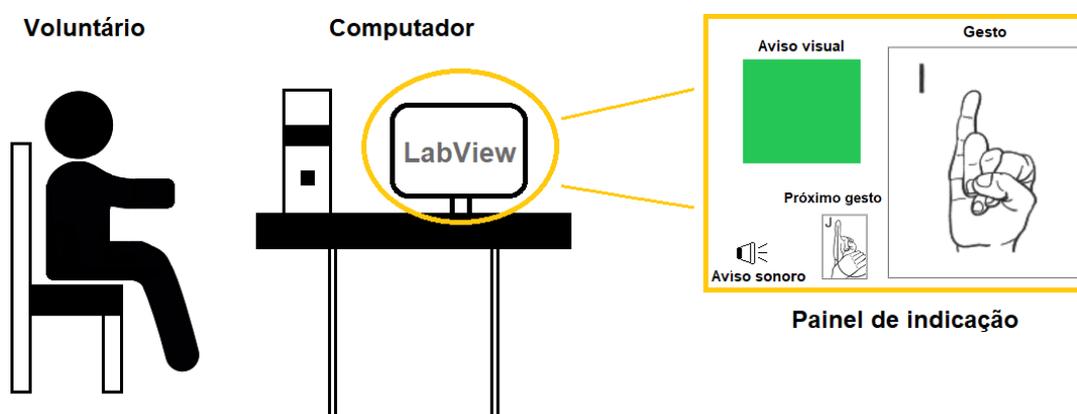
Fonte: Autoria Própria.

Para a realização do experimento utilizando o sistema com fio, foram coletados dados de 5 voluntários (3 homens e 2 mulheres), não usuários de Libras, com idade média de 24 anos e desvio padrão de 1,30. Esses voluntários realizaram os 26 padrões de letras do alfabeto (Figura 7) por 15 vezes sem pausa entre as repetições. A duração da coleta de dados para cada voluntário foi de aproximadamente 20 minutos.

Para o experimento com o sistema sem fio, 10 voluntários (5 homens e 5 mulheres), não usuários de Libras, executaram padrões de 10 palavras em Libras (Figura 8). Para a escolha dessas palavras, um especialista em Libras foi consultado e dez palavras frequentemente utilizadas na comunicação em Libras foram selecionadas. A idade média dos voluntários é de 24,8 anos e o desvio padrão de 3,70. Cada voluntário executou de forma sequencial 20 vezes cada palavra, com uma pausa de aproximadamente 2 minutos entre a 10ª e a 11ª repetição. Considerando o tempo de descanso, o tempo de duração da coleta para cada voluntário foi de 11 minutos.

Toda as coletas foram realizadas no período da manhã ou da tarde e, antes de seu início, cada voluntário teve liberdade para realizar os gestos até se familiarizar com o sistema. Os voluntários, sentados de frente para a tela do computador, são guiados por meio de um painel criado no LabView, conforme ilustra a Figura 24. Por meio de aviso sonoro e visual, o painel, de forma sequencial, indica o gesto que deve ser realizado e após o tempo de 1,3 s, indica que a execução do gesto deve ser terminada, totalizando um tempo 2,6 s entre o gesto a ser executado e o próximo.

Figura 24 – Voluntário sentado de frente para a tela do computador e o painel que informa qual gesto deve ser realizado. O painel apresenta um aviso sonoro e visual, além dos gestos que devem ser realizados.

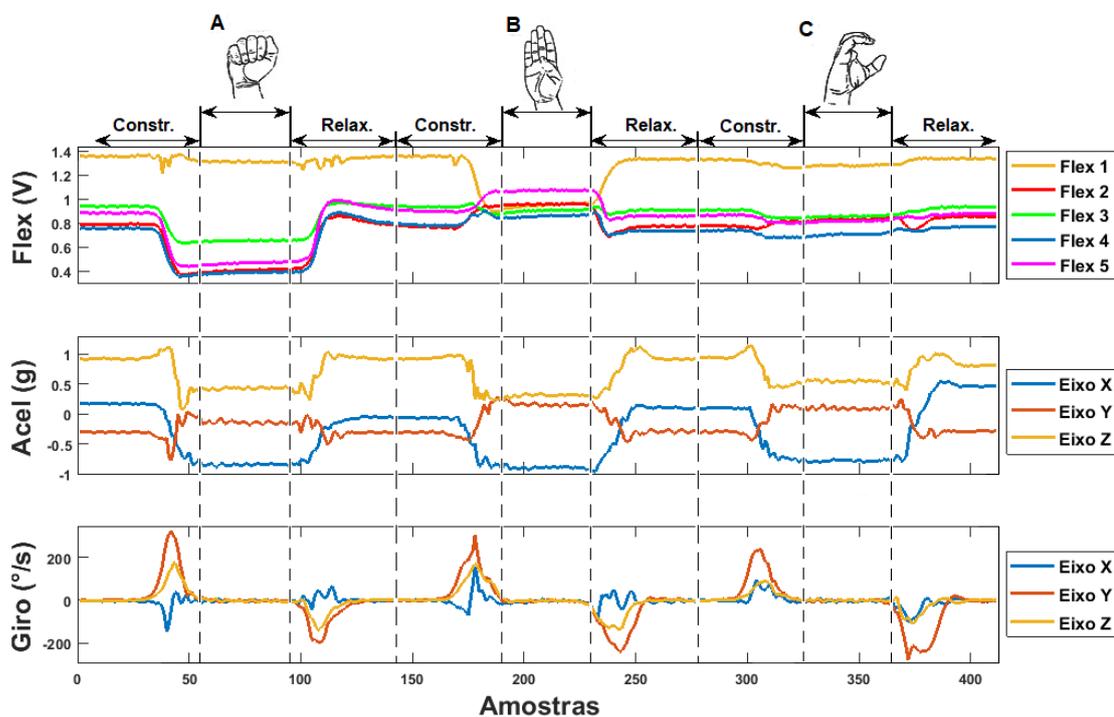


Fonte: Autoria Própria.

4.5 SEGMENTAÇÃO DOS DADOS E EXTRAÇÃO DE CARACTERÍSTICAS

Para realizar a identificação dos gestos, os dados coletados foram segmentados em janelas de amostras, que representam as diferentes classes e movimentos. Conforme ilustra a Figura 25, cada classe de gesto foi segmentada em três janelas, denominadas de fase de construção, gesto e fase de relaxamento. A fase de construção representa o momento em que o voluntário recebe a ordem para executar o gesto e inicia a execução do mesmo; o gesto representa o momento em que o voluntário executou o padrão de gesto indicado e a fase de relaxamento é o momento em que o voluntário recebe a ordem para parar de executar o gesto. O tempo de duração de cada janela de construção e de relaxamento dos gestos é de 0,9 s (45 amostras) e o tempo de duração de cada janela de gesto é de 0,8 s (40 amostras).

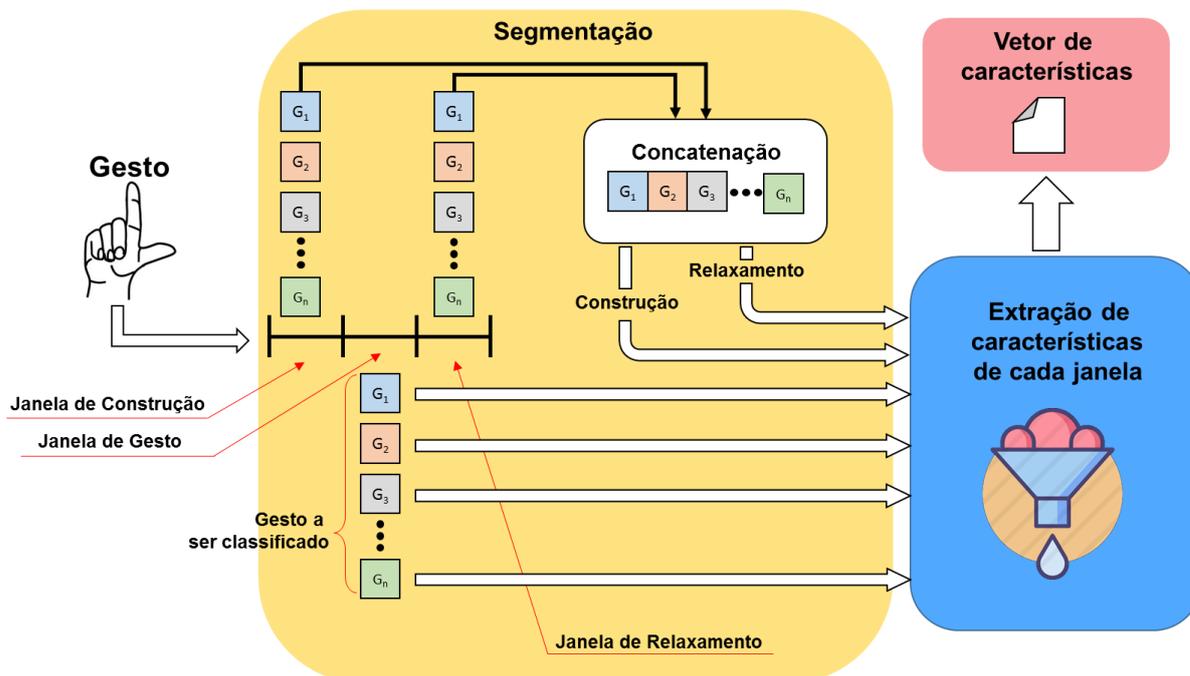
Figura 25 – Ilustração da segmentação dos dados coletados para as três primeiras letras do alfabeto. Os dados foram segmentados em janela de construção (Constr.), janela de gesto e janela de relaxamento (Relax.).



Fonte: Autoria Própria.

A extração de características foi realizada para cada janela segmentada, porém, antes disso, todas as janelas de construção e de relaxamento, referentes a cada repetição dos gestos, foram concatenadas. Nos dados obtidos no procedimento de coleta utilizando o sistema com fio, realizou-se a extração de características de 26 janelas de gesto e de 2 janelas concatenadas, totalizando 28 janelas a cada repetição do alfabeto. Desta forma, a realização dos gestos, pelos 5 voluntários e com 15 repetições, resultou em 2100 exemplos, que correspondem às 28 classes (26 gestos, construção e relaxamento). Da mesma forma, nos dados obtidos no procedimento de coleta utilizando o sistema sem fio, foi realizada a extração de característica de 10 janelas de gesto e de 2 janelas concatenadas, totalizando 12 janelas a cada repetição. No total, 2400 exemplos de gestos foram obtidos com 10 voluntários realizando 20 repetições. A Figura 26 ilustra o processo de segmentação e extração de característica realizado.

Figura 26 – Ilustração da etapa de segmentação e extração de características realizadas para gerar o vetor de características.



Fonte: Autoria Própria.

Como as saídas dos sensores de contato fornecem valor digital de 1 bit, para realizar a extração de características, apenas a moda foi extraída. Já para os demais sensores (flexíveis e inercial), seis características no domínio do tempo foram analisadas: moda, valor máximo, valor mínimo, média, desvio padrão, assimetria e curtose.

A moda é o valor que aparece com mais frequência em um conjunto de dados (OLIVEIRA, 2017). Dentro de um conjunto de dados, o valor máximo é o maior valor e o valor mínimo é o menor valor. A média (μ) é caracterizada pela razão entre soma de todos os valores (X_i) do conjunto e o número (N) de elementos deste conjunto (OLIVEIRA, 2017), dada por

$$\mu = \frac{1}{N} \sum_{i=1}^N X_i . \quad (13)$$

O desvio padrão (SD) representa a aproximação entre os valores do conjunto de dados e sua média (OLIVEIRA, 2017), expresso por

$$SD = \sqrt{\frac{1}{N-1} \sum_{i=1}^N |Xi - \mu|^2} . \quad (14)$$

O parâmetro (S) serve para caracterizar a assimetria dos dados em relação a um ponto central (NIST/SEMATECH, 2012). Quando a assimetria é negativa, significa que os dados estão concentrados mais à esquerda da média e quando a assimetria é positiva, os dados estão concentrados mais à direita da média. Em uma distribuição normal (perfeitamente simétrica), por exemplo, o valor de assimetria é zero (MATHWORKS, 2020). A assimetria é expressa como

$$S = \frac{\frac{1}{N} \sum_{i=1}^N (Xi - \mu)^3}{SD^3} . \quad (15)$$

A curtose (K) é utilizada para definir o grau de achatamento dos dados em relação a uma distribuição normal (NIST/SEMATECH, 2012). Assim, quanto maior for o achatamento dos dados, maior é o valor da curtose, que é calculada por

$$K = \frac{\frac{1}{N} \sum_{i=1}^N (Xi - \mu)^4}{\left(\frac{1}{N} \sum_{i=1}^N (Xi - \mu)^2\right)^2} . \quad (16)$$

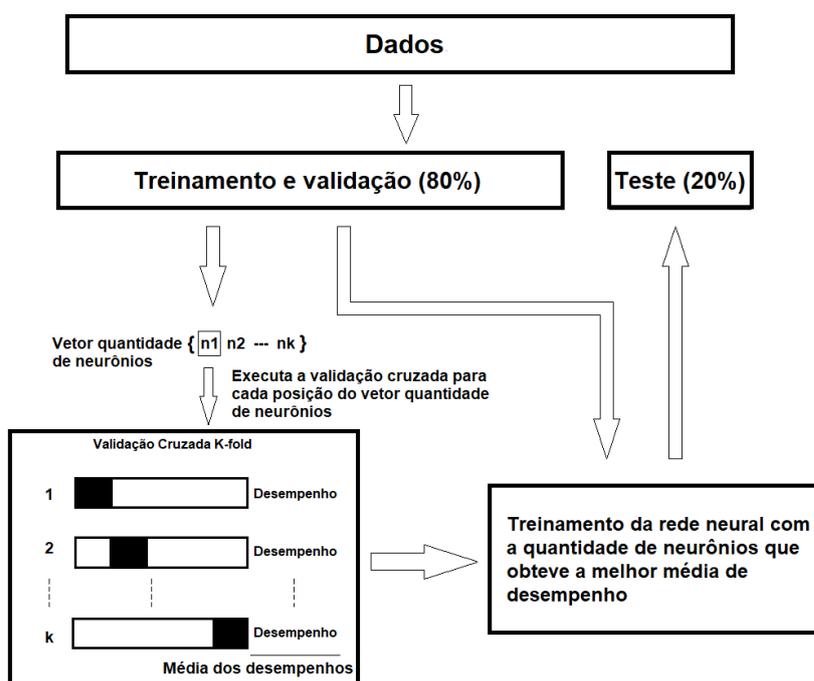
4.7 CLASSIFICAÇÃO

O classificador utilizado nesse trabalho foi a RNA com topologia MLP. O ambiente de programação utilizado foi o Matlab e todas as classificações ao longo do trabalho foram padronizadas com os mesmos parâmetros, com exceção do número de neurônios da camada escondida, que foi definida pela validação cruzada. A rede foi configurada com uma camada escondida, a função de transferência do tipo tangente hiperbólica foi utilizada na camada de entrada e na camada de saída da rede e os critérios de parada adotados foram: parada antecipada (*early stopping*) ou quando alcançar 1000 épocas de treinamento. Antes de serem submetidos à classificação, os dados foram normalizados pelo método de *z-score*.

O método de validação cruzada k -partições com k igual a 5 foi utilizado para analisar o desempenho da rede com diferentes quantidades de neurônios na camada escondida. Estes valores foram escolhidos empiricamente utilizando método de Fletcher-Gloss, que aponta uma faixa de valores a partir da quantidade de entradas e saídas da rede (SILVA; SPATTI; FLAUZINO, 2010). A fim de aumentar a área de busca e avaliar o comportamento da rede, decidiu-se aumentar em aproximadamente 10% os limites dessa faixa de valores.

Os dados para a classificação foram separados em 80% para treinamento e validação e 20% para teste final. A quantidade de neurônios na camada escondida foi escolhida com base na média de desempenho obtida pela validação cruzada, conforme ilustra a Figura 27. Para acelerar o processo de treinamento da rede, o algoritmo de gradiente conjugado escalado foi utilizado.

Figura 27 – Validação cruzada k -partições utilizada para escolher a quantidade de neurônios para a rede neural.



Fonte: Autoria Própria.

Neste capítulo foram apresentados os materiais empregados no desenvolvimento do sistema e a metodologia utilizada para coletar, processar e classificar os gestos. No próximo capítulo serão apresentados os resultados obtidos na análise do comportamento dos sensores e na classificação dos gestos, na qual diferentes cenários foram analisados.

5 RESULTADOS

Neste capítulo serão apresentados e discutidos os resultados obtidos na análise do comportamento dos sensores e na classificação dos gestos em diferentes cenários. Conforme explicado anteriormente e resumido na Figura 23, o experimento deste trabalho foi realizado utilizando duas versões de sistemas de aquisição de dados: sistema com fio e sistema sem fio. No sistema com fio, a coleta de dados foi realizada com 5 voluntários, que executaram de forma sequencial os 26 padrões de gestos do alfabeto em Libras. No sistema sem fio, os dados foram coletados de 10 voluntários durante a execução de 10 padrões de gestos de palavras em Libras.

5.1 COMPORTAMENTO DOS SENSORES

Nessa seção, o comportamento dos sinais de saída dos sensores de flexão e inercial (acelerômetro e giroscópio tri-axiais) foram analisados.

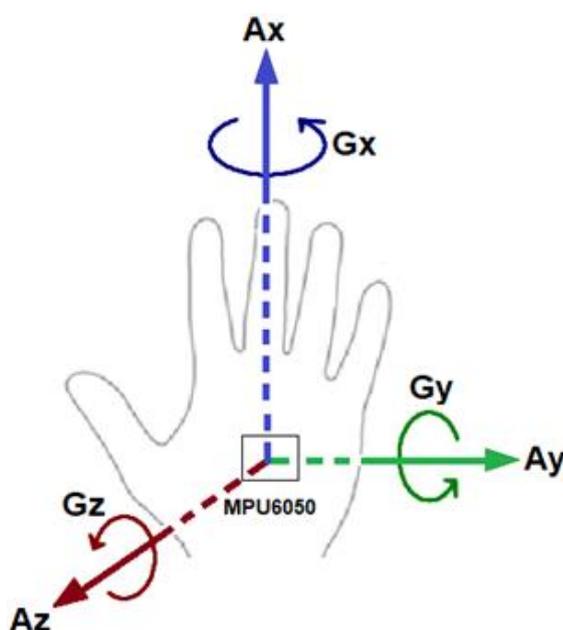
5.1.1 Análise do comportamento do sensor Inercial

Para conhecer o comportamento do sensor inercial MPU6050, as saídas correspondentes aos eixos X, Y e Z do acelerômetro e do giroscópio foram analisadas com o sensor posicionado na região superior do dorso da mão durante a realização de alguns movimentos. Quando a palma da mão está voltada para fora do corpo, os eixos do sensor ficam orientados conforme ilustra a Figura 28. Os sinais de aceleração dos eixos X, Y e Z do acelerômetro são indicados pelas setas Ax, Ay e Az, respectivamente. Já os sinais de giro dos eixos X, Y e Z do giroscópio são indicados pelas setas Gx, Gy e Gz, respectivamente.

O acelerômetro é capaz de detectar os gestos que possuem movimentos de aceleração em algum dos seus eixos e também identifica a posição da mão em relação à gravidade, ou seja, a orientação espacial da mesma. Da mesma forma, o giroscópio detecta os gestos que possuem movimento de giro. De modo geral, pode-se perceber que em qualquer movimento, seja apenas de giro ou de aceleração, ocorrem alterações nos sinais do acelerômetro e também no do giroscópio. Conforme apresenta a Figura 29, na letra 'l', que é uma letra estática, não houve alteração nos sinais do acelerômetro e do giroscópio, contudo pode-se observar a gravidade

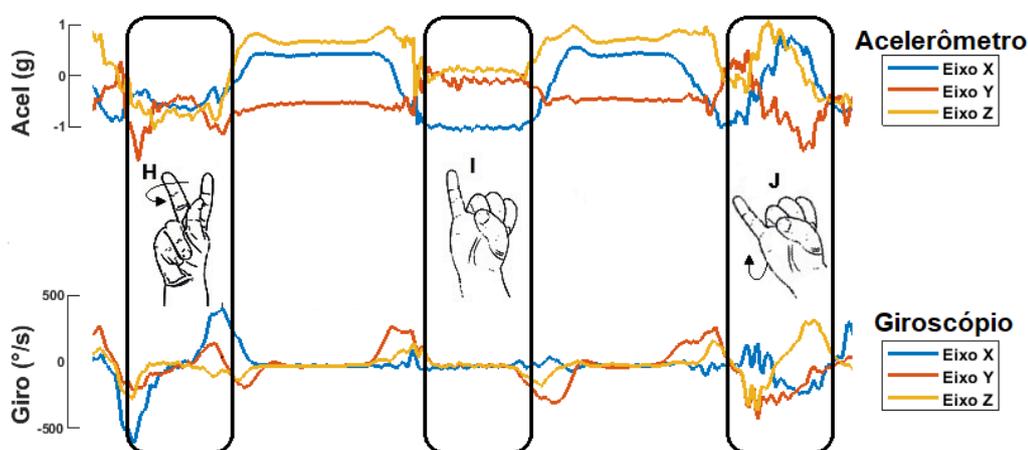
gerando o valor de aproximadamente -1 no eixo X ($A_x \approx -1g$). Já nos casos das letras 'H' e 'J', apesar de serem movimentos de giro, os sinais do acelerômetro também são modificados. Desta forma, para cada gesto dinâmico existe um padrão diferente, referente aos sinais do acelerômetro e do giroscópio. Nesse trabalho, por meio dos processos de segmentação, extração de características e classificação, esses padrões foram reconhecidos.

Figura 28 – Sensor inercial posicionado na região superior do dorso da mão. As setas A_x , A_y e A_z representam os sinais de aceleração e as setas G_x , G_y e G_z indicam os sinais de giro.



Fonte: Autoria Própria.

Figura 29 - Apresentação do comportamento dos sinais do sensor inercial durante a realização de algumas letras.



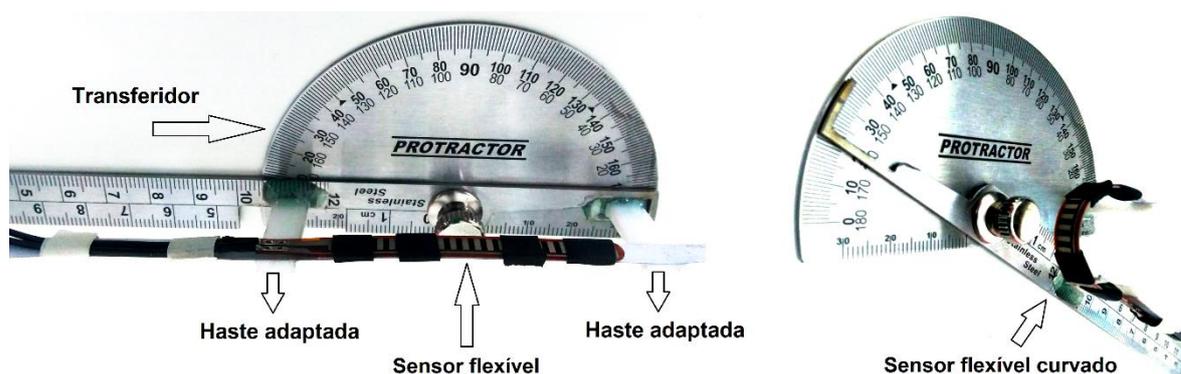
Fonte: Autoria Própria.

5.1.2 Análise do comportamento dos sensores flexíveis

Para conhecer o comportamento dos sensores flexíveis e determinar a relação entre os valores obtidos em tensão e o grau de curvatura deles, uma análise de correlação entre essas duas variáveis foi realizada. Para realizar a medição da curvatura de cada sensor, foi utilizado um transferidor de ângulos com haste adaptada, conforme apresenta a Figura 30. A haste adaptada auxilia na fixação do sensor e permite que o sensor seja ajustado no grau de curvatura desejado. Para realizar a medição dos valores de tensão, os mesmos microcontrolador e circuito divisor de tensão, descritos no capítulo 4.1, foram utilizados.

Por meio do transferidor de ângulos com haste adaptada, o grau de curvatura do sensor foi ajustado nos ângulos de 0 a 175° e as medidas de tensão foram registradas a cada 5°. Esse procedimento foi repetido quatro vezes para obter o gráfico referente à relação entre a tensão e o grau de curvatura do sensor flexível de 5,6 cm e de 9,52 cm, conforme apresenta a Figura 31. O coeficiente de correlação (R^2) encontrado para os sensores de 5,6 e 9,52 cm é de 0,9868 e 0,9800, respectivamente. Esses valores indicam correlação forte entre a tensão de saída e o grau de curvatura dos sensores.

Figura 30 – Transferidor de ângulos com haste adaptada. A haste adaptada auxilia a fixação dos sensores e permite o ajuste do sensor no ângulo desejado.

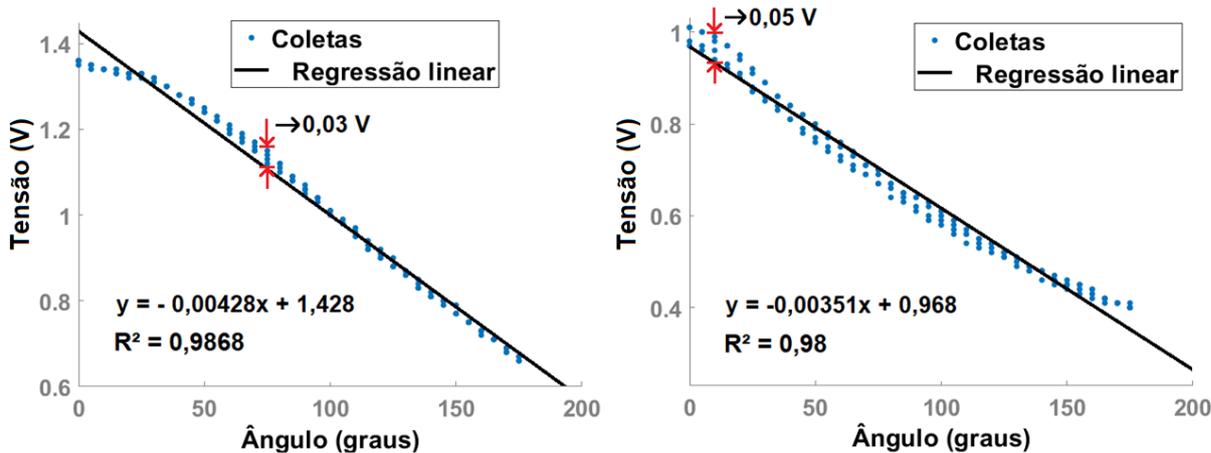


Fonte: Autoria Própria.

As maiores diferenças entre os pontos ocorrem em 75 graus para o sensor de 5,6 cm, que corresponde a 0,03 V e em 10 graus para o sensor de 9,52 cm, que corresponde a 0,05 V. Com essas retas de regressão, pode-se estimar os valores do grau de curvatura a partir das tensões medidas com erro máximo de 0,078 V para o

sensor do polegar e 0,056 V para os 4 sensores dos dedos. As sensibilidades de 3,5 e 4,3 mV/° foram determinados para os 4 sensores dos dedos e o sensor do polegar, respectivamente.

Figura 31 - Retas de relação entre grau de curvatura e tensão dos sensores flexíveis, em que (a) é o sensor flexível de 5,6 cm de comprimento ativo ($y = -0,0043x + 1,428$) e (b) é o sensor flexível de 9,52 cm ($y = -0,0035x + 0,968$).



(a) Sensor flexível de 5,6 cm

(b) Sensor flexível de 9,52 cm

Fonte: Autoria Própria.

5.2 SISTEMA COM FIO – ALFABETO

Nessa seção, as análises realizadas referentes à classificação dos sinais das letras do alfabeto em Libras são apresentadas.

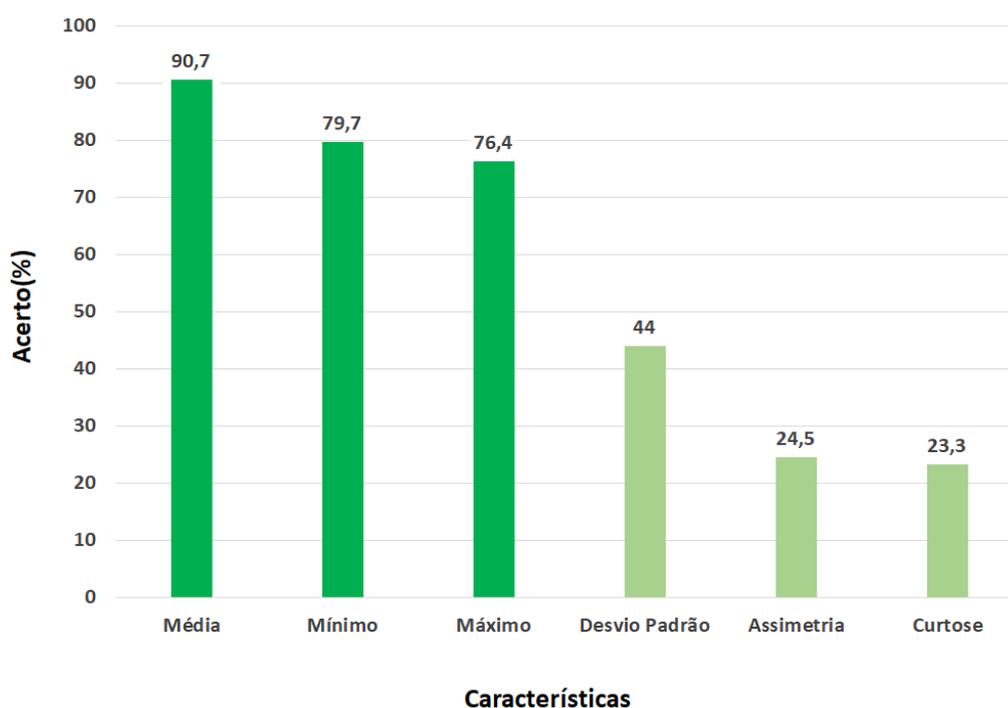
5.2.1 Análise das características

Como os valores de saída dos sensores de contato são digitais, apenas a moda foi extraída. Nos sensores flexíveis e inercial, para conhecer e selecionar as características que apresentam os melhores resultados na discriminação dos gestos, seis características no domínio do tempo (citadas no capítulo 4) foram analisadas.

Para essa análise, a rede neural MLP foi utilizada para realizar a classificação de forma individual para cada uma dessas seis características e a moda para os sensores de contato. As quantidades de neurônios utilizadas na camada escondida da rede neural foram: 37 neurônios (média), 33 neurônios (valor mínimo), 34

neurônios (valor máximo), 37 neurônios (desvio padrão), 34 neurônios (assimetria) e 34 neurônios (curtose). Esses valores foram escolhidos com base no melhor desempenho obtido na validação cruzada. Os resultados estão apresentados na Figura 32.

Figura 32 – Resultados obtidos para a classificação com redes neurais considerando as características de forma individual. Para os sensores de contato, apenas a moda foi considerada e para os sensores flexíveis e inercial, seis características foram consideradas.



Fonte: Autoria própria.

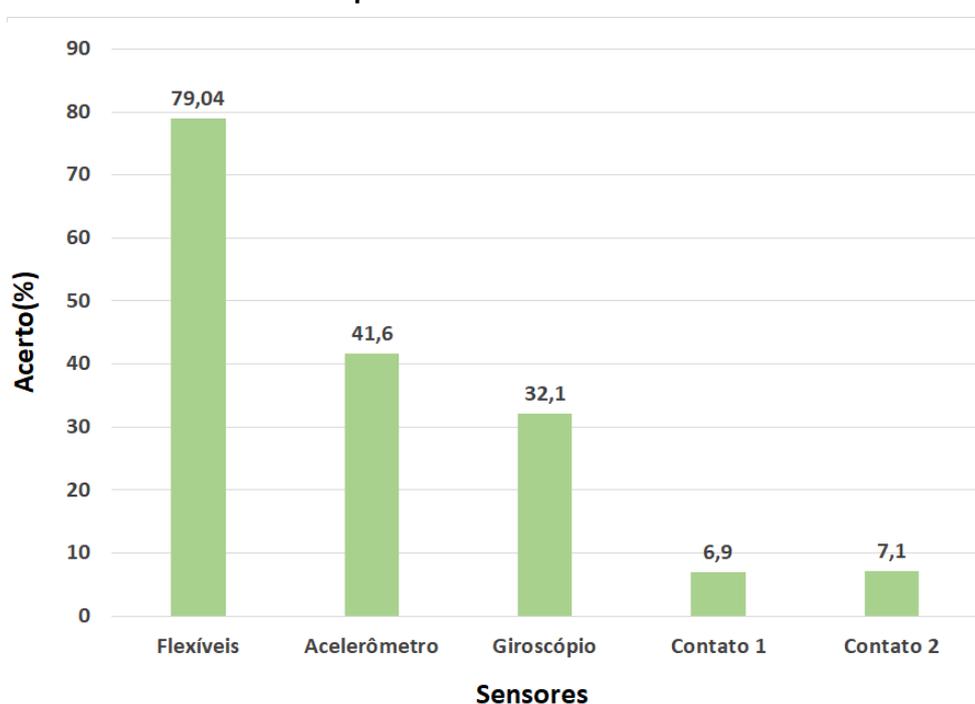
As características que obtiveram os melhores resultados foram: média (90,7%), valor máximo (76,4%) e valor mínimo (79,7%). As demais características (desvio padrão, assimetria e curtose) apresentaram taxas de acerto menores que 60%. Essas três características que obtiveram os melhores desempenhos individuais foram utilizadas para avaliar a contribuição de cada sensor na classificação, de forma individual e em grupos.

5.2.2 Análise por tipo de sensor

Para conhecer a contribuição de cada sensor na classificação, a análise de desempenho individual foi realizada considerando a moda para os sensores de

contato e as três melhores características (média, valor máximo e valor mínimo), obtidos no teste anterior, para os sensores flexíveis e inercial. As quantidades de neurônios utilizadas na camada escondida da rede neural foram: 35 neurônios (sensores flexíveis), 37 neurônios (acelerômetro), 35 neurônios (giroscópio), 30 neurônios (contato 1) e 18 neurônios (contato 2). A Figura 33 apresenta o desempenho obtido na classificação para cada sensor. Os sensores flexíveis apresentaram a maior taxa de acerto, seguidos do sensor inercial (acelerômetro e giroscópio) e dos sensores de contato.

Figura 33 - Classificação da rede neural considerando os sensores de forma individual com três características (valor máximo, mínimo e média) para os sensores flexíveis e inercial e a moda para os sensores de contato.



Fonte: Autoria própria.

Na matriz de confusão, obtida na classificação individual com os sensores flexíveis (Figura 34), pode-se observar que os sensores flexíveis auxiliam na discriminação das diferentes configurações de mão. Nota-se que as configurações de mãos parecidas foram bastante confundidas, como são os casos das letras 'H', 'K', 'N' e 'P'. Da mesma forma, em alguns casos, a letra 'R' foi confundida com as letras 'N', 'U' e 'V'.

Figura 34 – Matriz de confusão obtida para a classificação individual com os sensores flexíveis.

		Classe real																										Constr.	Relax.		
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z				
Classe predita	A	11																													
	B		15																												
	C			14														1													
	D				9								1																2		
	E					15						1	1				1										1				
	F						12															2									
	G				1																										
	H							14					1									2							2		
	I							1	13												3	1									
	J									14	4																	1			
	K											10																			
	L												0																		
	M									1			1	13						1									3		
	N														15																
	O															13					4		3			2		1			
	P																13														
	Q																	5													
	R				1																	1	10					2			
	S	4																													
	T																														
	U																														
	V																														
	W																														
	X																														
	Y																														
	Z																														
Constr.																															
Relax.																															

Fonte: Autoria Própria

O sensor inercial (acelerômetro e giroscópio) detecta a orientação da mão e auxilia no reconhecimento dos gestos que possuem movimento linear e movimento angular. Na classificação dos gestos utilizando apenas o acelerômetro, conforme apresenta a matriz de confusão (Figura 35), foi possível reconhecer letras que possuem movimento linear ('X', 'Y' e 'Z'), letras em que a orientação da mão é voltada para o lado ou para cima ('M', 'N', 'P' e 'Q') e os padrões de construção e relaxamento dos gestos (Constr. e Relax.). Apesar de a classificação ter sido realizada apenas com o acelerômetro, também foi possível reconhecer letras que possuem movimento angular ('H', 'J'). Nota-se que o classificador confundiu as letras que possuem orientações ou movimentos parecidos, como são os casos das letras 'M' e 'N' e 'H' e 'J'.

Figura 35 – Matriz de confusão obtida para a classificação individual com o acelerômetro.

		Classe real																										Constr.	Relax.	
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z			
Classe predita	A	0		1	2		1																				1			
	B	3	7	5	1	6	1	2		1										3		3								
	C			0															1											
	D				0		1																							
	E		1			1																				1				
	F	1				1	2									2						2								
	G	2		1	4		3	4				4	1			3			1	2	3				2					
	H	3					1		13																			4		
	I		5	4		4	4					2						1			4	2		5	1	1				
	J											14																		
	K				2		2	2					0			3						1								
	L				1			3					1	7					1		3	1	2	1						
	M														5	3			1											
	N													2	4												1			
	O				1			4					4				4				1									
	P	1											3	2	3	1		15		4	7				1	5				
	Q															6			14											
	R																			0										
	S																				0									
	T												1									0			2					
	U																						0							
	V	2	1	2	2	3								4					1		4	6	12	5						
	W																							0						
	X			1	1								1	3												8				
	Y	2			1					3		2				2			1		1			2			9			
	Z			1											1							1			1		15			
Constr.																														
Relax.	1	1												1	1				1						1				15	

Fonte: Autoria Própria.

A Figura 36 apresenta a matriz de confusão obtida na classificação realizada apenas com o giroscópio. Foi possível reconhecer as letras 'H' e 'J' (que possuem movimentos de giro) e, por possuírem movimentos parecidos, percebe-se que o classificador, em alguns casos, confundiu elas. Os padrões de construção e de relaxamento dos gestos obtiveram taxas de acerto próximas das obtidas com apenas o acelerômetro. Embora essa classificação tenha sido realizada utilizando apenas o giroscópio, as letras que possuem movimentos lineares ('X', 'Y' e 'Z') e orientação voltada para o lado ou para baixo ('M', 'N', 'P' e 'Q') foram reconhecidas, porém nota-se uma diminuição na taxa de reconhecimento das letras 'P' e 'Q'. A letra 'K', que possui movimento linear, obteve baixa taxa de acerto, porém, nota-se que em alguns casos ela foi confundida com as letras 'X' e 'Y', que também possuem movimento linear.

Figura 36 – Matriz de confusão obtida para a classificação individual com o giroscópio.

		Classe real																										Constr.	Relax.
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		
Classe predita	A	3	2	1	4	1	2	2			2							1		2	2	2	4			1			
	B	1	3	2		2			2																				
	C			0																									
	D	2	1		1			6				2	1		4	1				4					2				
	E	1	2		1	2	2											1	1	3	4	3	2			1			
	F	1					0					1																	
	G						0																						
	H			1					15		2						2									1			
	I	1	6	1	2	2	1			8									4	1		1			1				
	J										12																		
	K				1					1		1									1				1				
	L										1	1			1						1		2						
	M											1		6	5			2				1					1		
	N												1	1	4			3											
	O	2		6	5	3	4	2		1		1	1			10				2	5	1		2	2				
	P			1			1						2		2		5	1	2						1	3			
	Q							1				1		4	1		4	4								1			
	R																		0										
	S																			0									
	T																					0							
	U																						0						
	V						1						1					1	1					1					
	W	1	1			2	1	1					3						2	1			2	5	3				
	X			1			1	1					3	2	1		1	2			1	1				7			
	Y	2		2	1	3	2	2		3			4	1	1	2			1	5	1		3	1	3	1			
	Z													1	1			1	2			4		1		2			
Constr.	1									1															1				
Relax.																									2				

Fonte: Autoria Própria.

Os sensores de contato são utilizados para detectar diferenças como sobreposição (sensor de contato 1) e abdução lateral (sensor de contato 2) dos dedos indicador e médio. A Figura 37 apresenta a matriz de confusão obtida na classificação dos gestos utilizando apenas o sensor de contato 1. Pode-se observar que quase todos os exemplos da letra 'R' (padrão de sobreposição dos dedos indicador e médio) foram reconhecidos, pois é o único padrão de gesto que realiza o fechamento desse ponto de contato. A letra 'B' foi reconhecida como o gesto que possui o contato 1 aberto, porém poderia ser outra letra que possui esse contato aberto.

Já na classificação com apenas o sensor de contato 2, foi possível classificar todos os exemplos da letra 'V' (padrão de abdução, que é um contato aberto) corretamente, conforme apresenta a matriz de confusão, na Figura 38. Percebe-se que, dentre as letras que possuem o padrão de contato fechado, o classificador escolheu a letra 'S'.

Figura 37 – Matriz de confusão obtida para a classificação individual com o sensor de contato 1.

		Classe real																										Constr.	Relax.	
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z			
Classe predita	A	0																												
	B	15	15	15	15	15	15	15	15	15	15	15	15	14	15	13	15	15	1	15	15	15	15	15	15	15	15	15		
	C			0																										
	D				0																									
	E					0																								
	F						0																							
	G							0																						
	H								0																					
	I									0																				
	J										0																			
	K											0																		
	L												0																	
	M													0																
	N														0															
	O															0														
	P																0													
	Q																	0												
	R														1	2			0	14										
	S																			0										
	T																				0									
	U																					0								
	V																						0							
	W																							0						
	X																								0					
	Y																									0				
	Z																										0			
Constr.																											0			
Relax.																												0		

Fonte: Autoria Própria.

Figura 38 – Matriz de confusão obtida para a classificação individual com o sensor de contato 2.

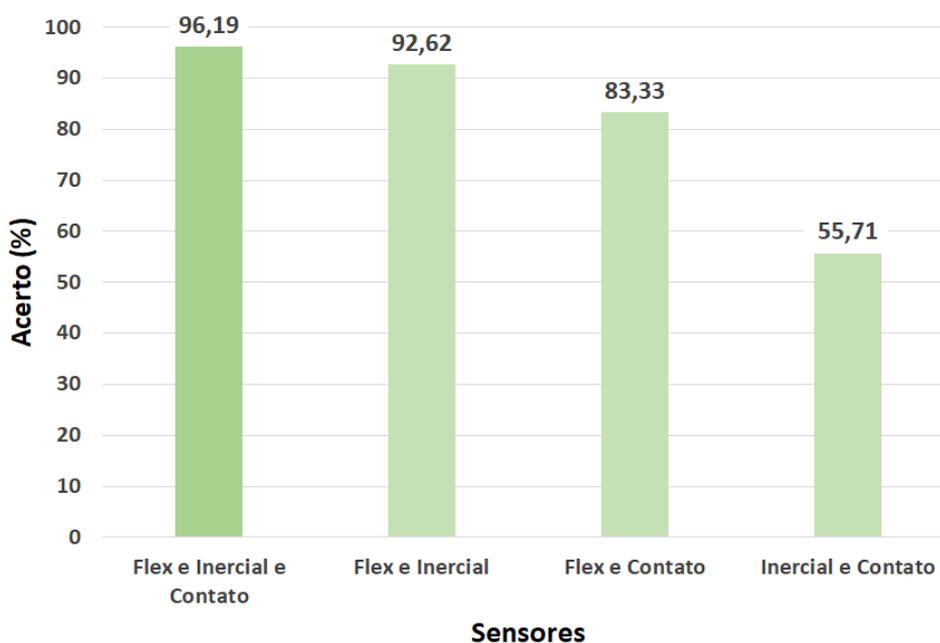
		Classe real																										Constr.	Relax.		
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z				
Classe predita	A	0																													
	B		0																												
	C			0																											
	D				0																										
	E					0																									
	F						0																								
	G							0																							
	H								0																						
	I									0																					
	J										0																				
	K											0																			
	L												0																		
	M													0																	
	N														0																
	O															0															
	P																0														
	Q																	0													
	R																		0												
	S	15	13	13	1	15	1		1	15	15		14	15	15		7	15	15		15					14		6	2		
	T																				0										
	U																					0									
	V		2	2	14		14	15	14			15	15	1			15	8			15			15		15	15	1	15	9	13
	W																							0							
	X																								0						
	Y																									0					
	Z																										0				
Constr.																											0				
Relax.																												0			

Fonte: Autoria Própria.

5.2.3 Análise dos sensores em grupos

Os resultados obtidos na análise dos sensores de forma individual mostraram que todos os sensores são importantes na classificação das letras do alfabeto manual, cada um contribui no reconhecimento de algum tipo de gesto. A análise de desempenho em grupo de sensores foi realizada para conhecer o desempenho de diferentes grupos de sensores na classificação. Os sensores foram organizados em 3 grupos: Flexíveis, Inerciais e de Contato. Para cada combinação de sensores a rede neural foi treinada com uma quantidade de neurônios na camada escondida: 53 neurônios (flex e inercial), 37 neurônios (flex e contato), 37 neurônios (inercial e contato) e 60 neurônios (flex, inercial e contato). A Figura 39 apresenta os resultados obtidos nessa análise.

Figura 39 - Classificação da rede neural considerando os sensores em grupo com as três características (valor máximo, mínimo e média) para os sensores flexíveis e inercial e a moda para os sensores de contato.



Fonte: Autoria Própria.

O grupo que possui todos os sensores (flexíveis, inercial e contato) foi o que alcançou o melhor desempenho (taxa de acerto de 96,19%). Na classificação com apenas os sensores flexíveis e inercial (sem os sensores de contato), o desempenho diminuiu em 3,57 pontos percentuais. Isso mostra que todos os sensores são importantes na classificação.

A Figura 40 apresenta a matriz de confusão obtida na classificação com todos os sensores (grupo que apresentou o melhor desempenho). Os padrões de gestos que obtiveram as menores taxas de acerto (86,6%) foram as letras 'A', 'K', 'R' e 'S'. A letra 'A' foi confundida com a letra 'S'; a letra 'K' foi confundida com as letras 'L' e 'P'; a letra 'R' foi confundida com 'U' e a letra 'S' foi confundida com a letra 'A'. Nota-se que em todos esses casos as letras possuem configurações de mão parecidas. Todas as letras que possuem movimento angular ('H' e 'J') e movimento linear ('X', 'Y' e 'Z'), com exceção da letra 'K', foram classificadas com taxa de 100% de acerto. As letras que possuem orientação de mão voltada para cima ou para baixo ('M', 'N', 'P' e 'Q') obtiveram taxas de acerto entre 93,3 e 100%.

Figura 40 – Matriz de confusão para a classificação da rede neural com todos os sensores, utilizando três características (valor máximo, mínimo, média) para os sensores flexíveis e inercial e uma (moda) para os sensores de contato.

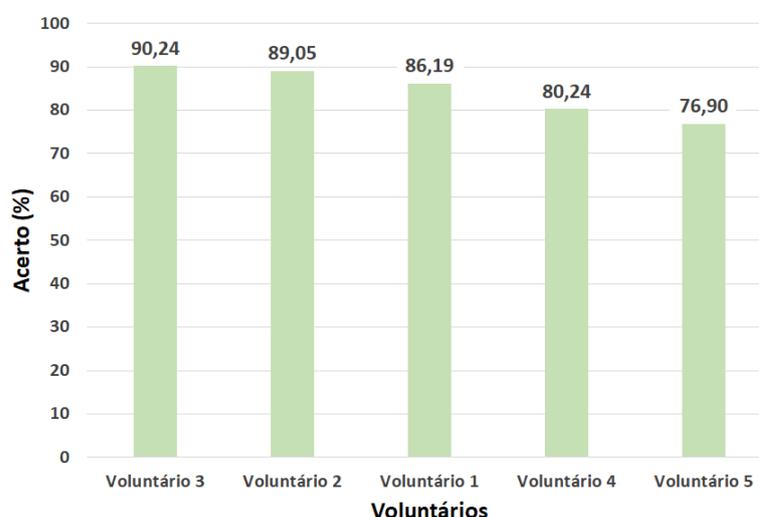
		Classe real																										Constr.	Relax.
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		
Classe predita	A	13																		2									
	B		15																										
	C			14													1												
	D				14																								
	E			1	1	15																							
	F						14																1						
	G							14																					
	H								15																				
	I									15																			
	J										15																		
	K											13							1										
	L											1	15																
	M													15															
	N														14														
	O															14													
	P											1					14												
	Q																	14											
	R																		15										
	S	2																	13										
	T																			13									
	U																					14							
	V																						15						
	W																							15					
	X																								15				
	Y																									15			
	Z																										15		
Constr.																											15		
Relax.																												15	

Fonte: Autoria Própria.

5.2.4 Análise dos voluntários

Por fim, para analisar o desempenho da rede para cada voluntário, o método de *leave-one-subject-out* foi aplicado. Nele, um voluntário é deixado de fora na etapa de treinamento e validação do classificador e na etapa de teste, apenas os dados do voluntário deixado de fora são utilizados. A classificação foi realizada com a rede neural MLP, utilizando a moda para os sensores de contato e as três características (média, valor mínimo e valor máximo) para os sensores flexíveis e inercial. Os dados de 4 voluntários foram separados para treinamento e validação e de 1 voluntário para teste. Para cada voluntário a rede neural foi treinada com uma quantidade de neurônios na camada escondida: 56 neurônios (voluntário 1), 57 neurônios (voluntário 2), 60 neurônios (voluntário 3), 57 neurônios (voluntário 4) e 60 neurônios (voluntário 5). A Figura 41 apresenta o resultado obtido na classificação para cada um dos voluntários.

Figura 41 – Classificação da rede neural considerando os dados de quatro voluntários para o treinamento e de um voluntário para o teste.



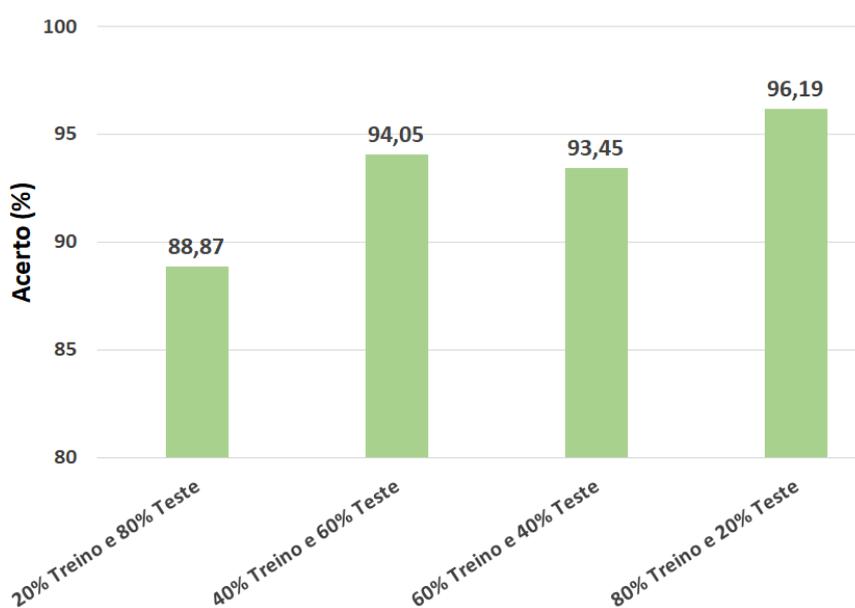
Fonte: Autoria Própria.

A maior taxa de acerto foi obtida pelo voluntário 3 (90,24%) e a menor taxa, pelo voluntário 5 (76,90%), que representa uma diferença de 13,34 pontos percentuais. Observou-se durante a realização dos gestos que alguns voluntários possuem costumes e, por isso, executam alguns gestos de maneira um pouco diferente da indicada no painel. Por isso, acredita-se que esse é o motivo que causa essa diferença 13,34 pontos percentuais na taxa de acerto na classificação.

5.2.5 Análise da quantidade de dados para treinamento

Em todas as análises realizadas anteriormente, os dados para a classificação foram divididos em 80% para treinamento e validação e 20% para teste final. Para analisar o desempenho da classificação em relação a quantidade de dados treinados, os dados foram divididos nas seguintes proporções para treinamento e teste (porcentagem de dados para treinamento / porcentagem de dados para teste): 20/80, 40/80, 60/40 e 80/20. As quantidades de neurônios utilizadas na camada escondida da rede neural na classificação considerando essas diferentes proporções foram: 59 neurônios (20/80), 60 neurônios (40/80), 59 neurônios (60/40) e 60 neurônios (80/20). Para a classificação, a moda foi utilizada para os sensores de contato e as três características (média, valor mínimo e valor máximo) para os sensores flexíveis e inercial. A Figura 42 apresenta os resultados obtidos na classificação com diferentes proporções de dados para treinamento e teste. Pode-se perceber uma tendência positiva da taxa de acerto em relação ao aumento da quantidade de dados para treinamento. Desta forma, acredita-se que aumentando a quantidade de dados coletados, pode-se obter melhores resultados na classificação.

Figura 42 – Classificação da rede neural com os dados divididos em diferentes proporções para treinamento e teste.



Fonte: Autoria Própria.

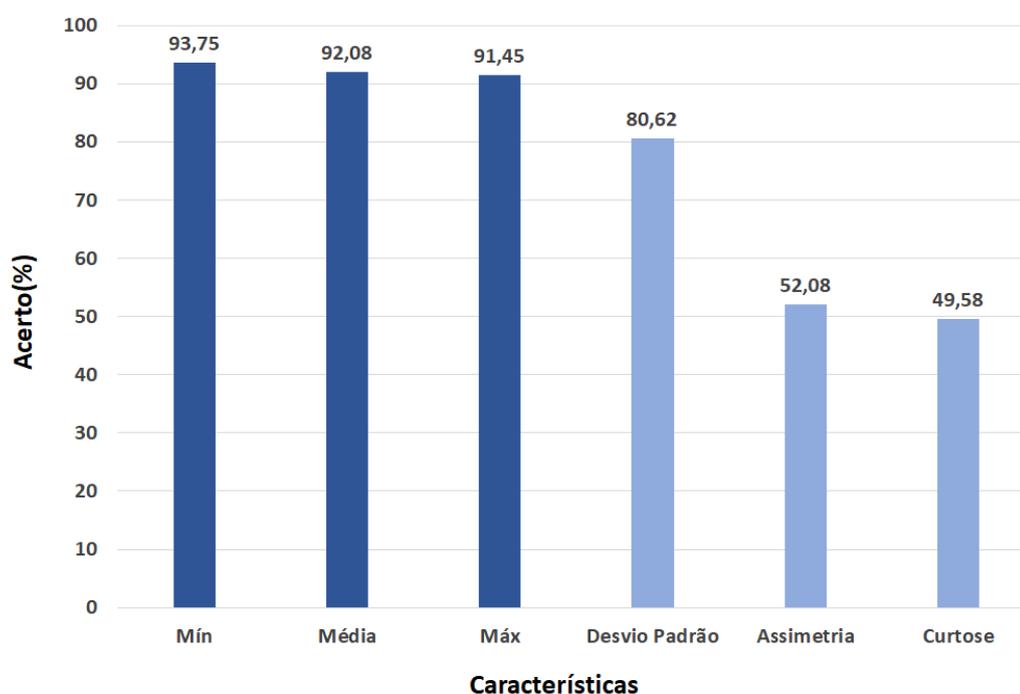
5.3 SISTEMA SEM FIO – PALAVRAS

Nessa seção, as análises realizadas referentes à classificação dos sinais de palavras em Libras são apresentadas.

5.3.1 Análise das características

De forma similar ao método de análise utilizado no sistema com fio, para selecionar as características que melhor contribuem na discriminação dos gestos, seis características foram analisadas: média, valor máximo, valor mínimo, desvio padrão, assimetria e curtose. Os sensores de contato, por possuírem saída digital, apenas a moda foi extraída. Para avaliar o desempenho de cada característica, a rede neural MLP foi utilizada para realizar a classificação, conforme apresenta a Figura 43. As quantidades de neurônios utilizados na camada escondida da rede neural foram: 22 (valor mínimo), 21 (média), 21 (valor máximo), 20 (desvio padrão), 22 (assimetria) e 20 (curtose).

Figura 43 – Resultados obtidos para a classificação com redes neurais considerando as características de forma individual. Para os sensores de contato, apenas a moda foi considerada e para os sensores flexíveis e inercial, seis características foram consideradas.



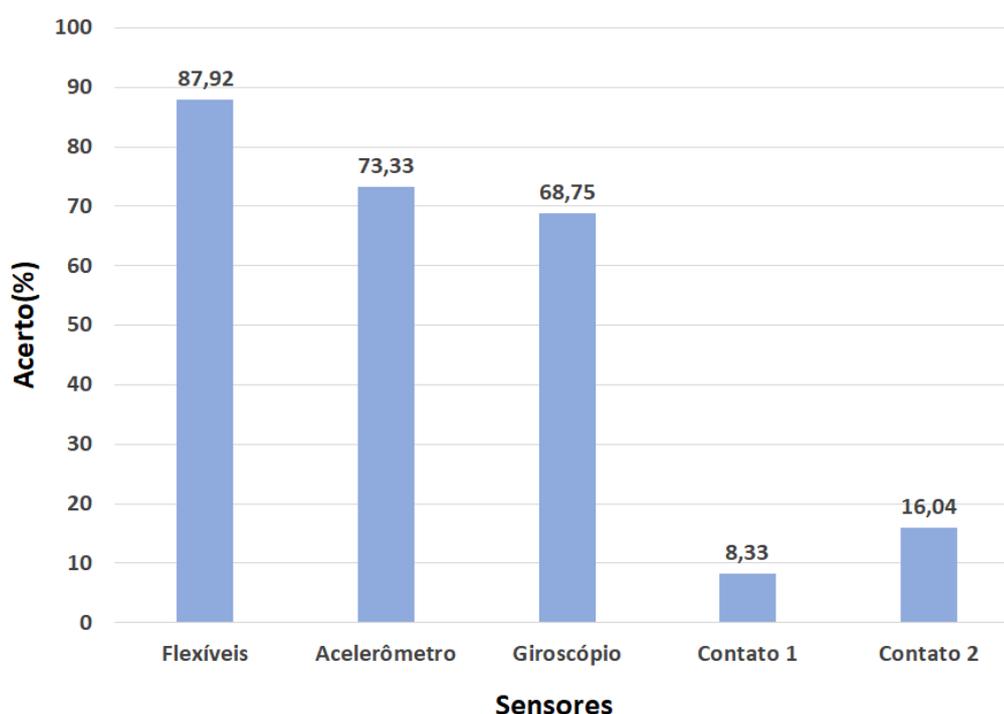
Fonte: Autoria Própria.

As melhores taxas de acerto foram obtidas com as características valor mínimo (93,75%), média (92,08%) e valor máximo (91,45%). Por isso, essas características foram selecionadas para serem utilizadas nas análises, que serão apresentadas a seguir.

5.3.2 Análise do tipo de sensor

Com as características selecionadas na análise anterior (média, valor máximo e valor mínimo) e a moda (para os sensores de contato), a classificação considerando apenas um tipo de sensor por vez foi realizada utilizando a rede neural MLP. Apenas os sensores de contato, que apesar de serem do mesmo tipo, foram classificados de forma separada: sensor de contato 1 e sensor de contato 2. As quantidades de neurônios utilizadas na camada escondida da rede neural foram: 22 neurônios (sensores flexíveis), 22 neurônios (acelerômetro), 22 neurônios (giroscópio), 22 neurônios (contato 1) e 20 neurônios (contato 2). A Figura 44 apresenta os resultados obtidos.

Figura 44 – Classificação da rede neural considerando os sensores de forma individual com três características (média, valor máximo e valor mínimo) para os sensores flexíveis e inercial e a moda para os sensores de contato.



Fonte: Autoria Própria.

A matriz de confusão obtida na classificação dos sensores flexíveis é apresentada na Figura 45. As palavras ‘escutar’, ‘quente’ e ‘rápido’ foram confundidas entre si. O mesmo ocorreu com as palavras ‘aproveitar’ e ‘triste’ e também com a palavra ‘dia’, que foi confundida com a palavra ‘frio’. Nota-se que em todos esses casos, as palavras possuem configurações de mão quase parecidas.

Figura 45 – Matriz de confusão obtida na classificação individual com os cinco sensores flexíveis.

Classe real

Classe predita	Classe real												
	Aprender	Aproveitar	Escutar	Falar	Dia	Frio	Quente	Rápido	Urgente	Triste	Construção	Relaxamento	
Aprender	35						1				2		
Aproveitar		34							1	8			
Escutar			31				6	4					
Falar				40									
Dia					31								
Frio						9	40						
Quente	1		6				30			1			
Rápido	1		3				3	36					
Urgente									38				
Triste		6								31			
Construção	3								1		38	2	
Relaxamento												38	

Fonte: Autoria Própria.

Na matriz de confusão da classificação utilizando apenas o acelerômetro, apresentada na Figura 46, nota-se que os gestos que possuem movimentos parecidos foram confundidos entre si, como são os casos das palavras ‘falar’ e ‘frio’ e também das palavras ‘rápido’ e ‘urgente’. Foi possível identificar os padrões de movimentos das classes ‘construção’ e ‘relaxamento’, porém, algumas vezes, essas classes também foram confundidas entre si.

Na classificação utilizando apenas o giroscópio (Figura 47), as palavras ‘falar’ e ‘frio’ e também as palavras ‘rápido’ e ‘urgente’ foram confundidas entre si. Nota-se que essas palavras possuem uma mistura de movimento linear e de giro e, por isso, foi possível identificar esses gestos tanto com acelerômetro, como com giroscópio. Da mesma forma, as classes ‘construção’ e ‘relaxamento’ foram confundidos entre si.

Figura 46 – Matriz de confusão obtida na classificação individual com apenas o acelerômetro.

Classe real

	Aprender	Aproveitar	Escutar	Falar	Dia	Frio	Quente	Rápido	Urgente	Triste	Construção	Relaxamento
Aprender	38					2						
Aproveitar		26		2	1	1					2	
Escutar			40									
Falar		3		26	1	6	1		1			
Dia					28	3	2	2	4			
Frio	1	3		9		20	1		1			
Quente		2		1	7	7	33	3	1	1	2	
Rápido						1		22	11		1	
Urgente				2	1		2	10	20			
Triste		1			1					38		
Construção	1	4						2		1	27	6
Relaxamento		1			1		1	1	2		8	34

Classe predita

Fonte: Autoria Própria.

Figura 47 – Matriz de confusão obtida na classificação individual com apenas o giroscópio.

Classe real

	Aprender	Aproveitar	Escutar	Falar	Dia	Frio	Quente	Rápido	Urgente	Triste	Construção	Relaxamento
Aprender	8										2	
Aproveitar		36										
Escutar	14		31			3	1		2	4		
Falar				21	1	15						
Dia			1	1	38			1	1	3		
Frio		1		13		22	1					
Quente	7		3	3	1		31	3	1	3		
Rápido	1			1			2	17	5	1		
Urgente	4						1	17	27			
Triste		1	5	1			4	2	2	28		
Construção	2								2		34	3
Relaxamento	4	2								1	4	37

Classe predita

Fonte: Autoria Própria.

O sensor de contato 1 é utilizado para detectar sobreposição entre os dedos indicador e médio, porém nenhum dos padrões das palavras utilizadas no treinamento possui esse tipo de sobreposição. Na classificação utilizando apenas o sensor de contato 1, o sensor não foi útil para identificar nenhum gesto específico. Pode-se notar na matriz de confusão obtida (Figura 48) que a palavra 'quente' foi identificada como o padrão de gesto que não possui sobreposição entre os dedos, porém poderia ter sido qualquer outro gesto.

Figura 48 – Matriz de confusão obtida na classificação individual com o sensor de contato 1.

		Classe real											
		Aprender	Aproveitar	Escutar	Falar	Dia	Frio	Quente	Rápido	Urgente	Triste	Construção	Relaxamento
Classe predita	Aprender	0											
	Aproveitar		0										
	Escutar			0									
	Falar				0								
	Dia					0							
	Frio						0						
	Quente	39	38	39	40	40	40	38	40	40	38	40	40
	Rápido								0				
	Urgente									0			
	Triste	1	2	1				2			2		
	Construção											0	
	Relaxamento												0

Fonte: Autoria Própria.

Na classificação utilizando apenas o sensor de contato 2, foi possível identificar a palavra 'falar', que possui padrão de abdução (contato aberto) e a palavra 'triste', que possui o padrão de contato fechado. A Figura 49 apresenta a matriz de confusão obtida na classificação utilizando apenas o sensor de contato 2.

Figura 49 – Matriz de confusão obtida na classificação individual com o sensor de contato 2.

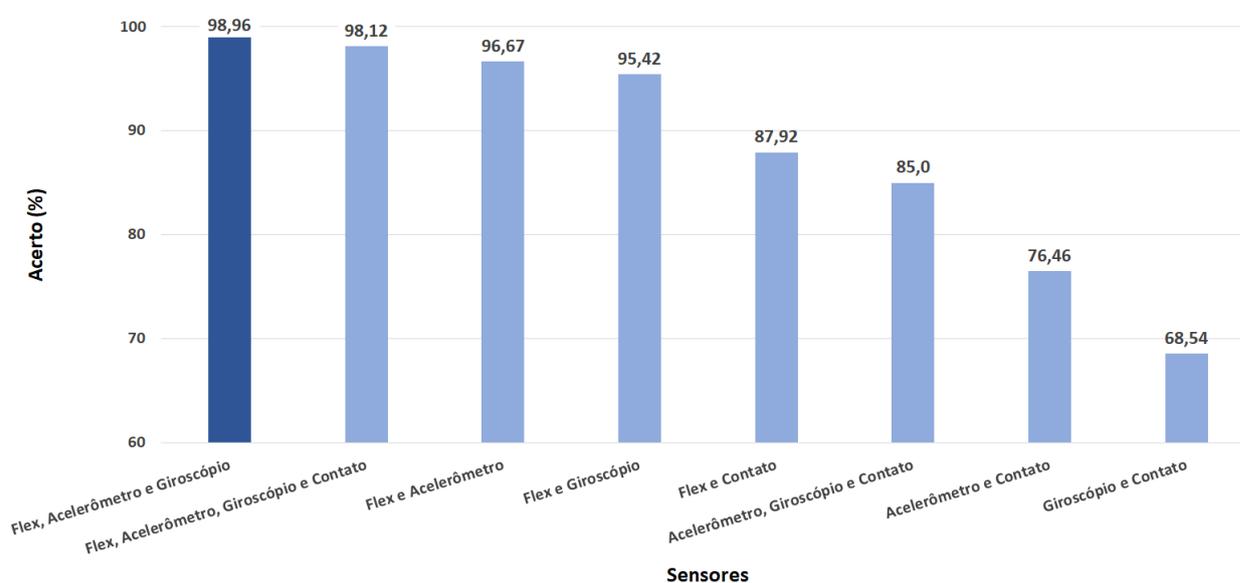
		Classe real											
		Aprender	Aproveitar	Escutar	Falar	Dia	Frio	Quente	Rápido	Urgente	Triste	Construção	Relaxamento
Classe predita	Aprender	0											
	Aproveitar		0										
	Escutar			0									
	Falar	13	3	1	40	39	40	6	8	4	3	12	15
	Dia					0							
	Frio						0						
	Quente							0					
	Rápido								0				
	Urgente									0			
	Triste	27	37	39		1		34	32	36	37	28	25
	Construção											0	
	Relaxamento												0

Fonte: Autoria Própria.

5.3.3 Análise dos sensores em grupos

Para conhecer o desempenho dos sensores em grupo, diferentes combinações de sensores foram utilizadas para classificação. Os sensores foram organizados em 4 grupos: flexíveis, acelerômetro, giroscópio e de contato. Em cada combinação de sensores a rede neural foi treinada com uma determinada quantidade de neurônios: 19 neurônios (flex e acelerômetro), 20 neurônios (flex e giroscópio), 34 neurônios (flex e contato), 18 neurônios (acelerômetro e contato), 22 neurônios (giroscópio e contato), 35 neurônios (acelerômetro, giroscópio e contato), 53 neurônios (flex, acelerômetro e giroscópio) e 52 neurônios (flex, acelerômetro, giroscópio e contato). A Figura 50 apresenta a taxa de classificação obtida para cada grupo de sensores. O grupo com os sensores flexíveis, acelerômetro e giroscópio foi o que obteve o melhor desempenho (98,96%). Isso mostra que os sensores de contato não foram úteis para auxiliar na identificação dos sinais das palavras utilizadas nessa análise.

Figura 50 – Classificação obtida com a rede neural considerando os sensores em grupo com as três características (valor máximo, mínimo e média) para os sensores flexíveis e inercial e a moda para os sensores de contato.



Fonte: Autoria Própria.

A matriz de confusão obtida na classificação com o grupo que apresentou o melhor desempenho (flexíveis, acelerômetro e giroscópio) é apresentada na Figura 51. Pode-se observar que quase todos os exemplos foram classificados corretamente.

Figura 51 – Matriz de confusão obtida na classificação da rede neural com o grupo que apresentou o melhor desempenho (flexíveis, acelerômetro e giroscópio), utilizando três características (valor máximo, mínimo, média) para os sensores flexíveis e inercial.

Classe real

	Aprender	Aproveitar	Escutar	Falar	Dia	Frio	Quente	Rápido	Urgente	Triste	Construção	Relaxamento
Aprender	40								1			
Aproveitar		39										
Escutar			40									
Falar				40								
Dia					40							
Frio		1				40						
Quente							39		1			
Rápido							1	40		1		
Urgente									38			
Triste										39		
Construção											40	
Relaxamento												40

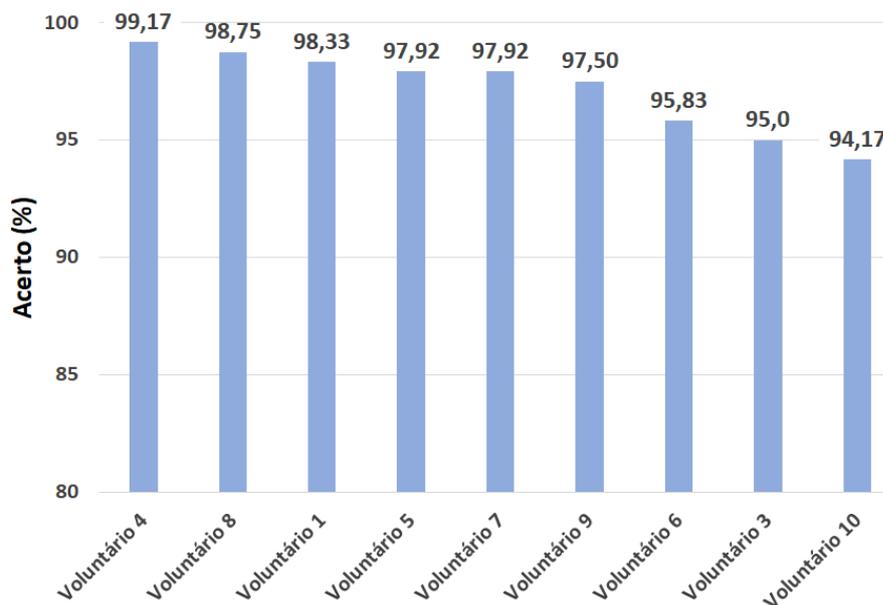
Classe predita

Fonte: Autoria Própria.

5.3.4 Análise dos voluntários

Nessa seção, o desempenho de cada voluntário é analisado utilizando o método *leave-one-subject-out*. Para isso, os dados de 9 voluntários são utilizados para treinamento e os dados de 1 voluntário são utilizados para teste final. Ou seja, 90% dos dados são para treinamento e validação e 10% são para teste final. Para a classificação, o grupo com sensores flexíveis e inercial foram utilizados (melhor grupo). As características média, valor mínimo e valor máximo foram utilizadas para os sensores flexíveis e inercial. Para cada voluntário a rede neural foi treinada com uma determinada quantidade de neurônios na camada escondida: 53 neurônios (voluntário 1), 52 neurônios (voluntário 2), 51 neurônios (voluntário 3), 52 neurônios (voluntário 4), 50 neurônios (voluntário 5), 51 neurônios (voluntário 6), 53 neurônios (voluntário 7), 53 neurônios (voluntário 8), 52 neurônios (voluntário 9) e 51 neurônios (voluntário 10). A Figura 52 apresenta a classificação obtida para cada voluntário.

Figura 52 – Classificação da rede neural considerando os dados de nove voluntários para o treinamento e de um voluntário para o teste.



Fonte: Autoria Própria.

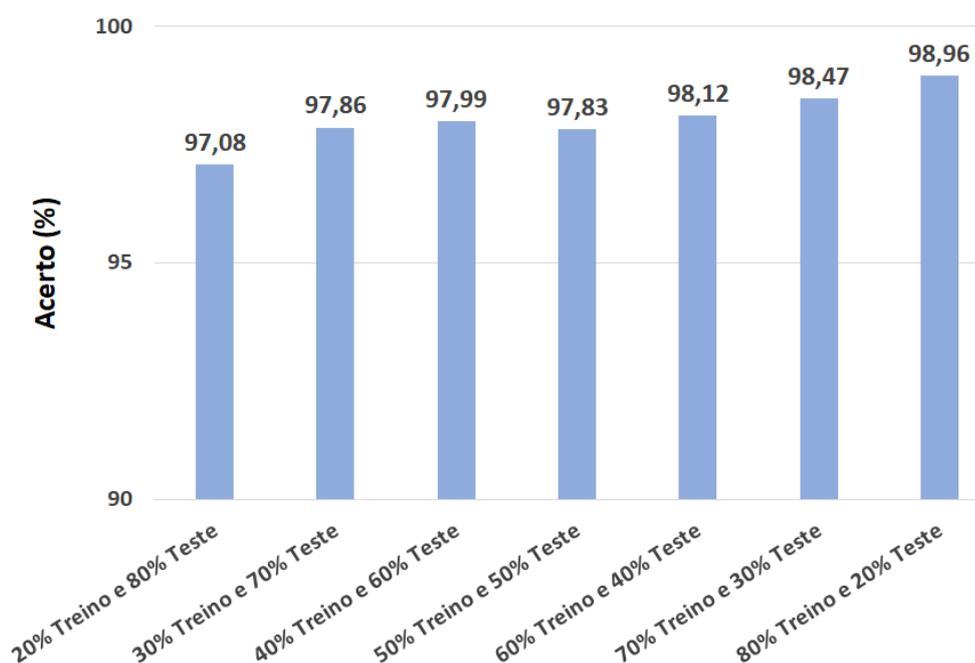
Os voluntários 4 e 8 obtiveram as melhores taxas de acerto: 99,17% e 98,75%. O voluntário 10 obteve a pior taxa de acerto (94,17%). Existem vários motivos que podem ter levado o voluntário 10 a obter o desempenho mais baixo: voluntário errou alguns gestos durante a execução, teve dificuldade para executar os gestos ou não tem precisão na execução dos gestos (variabilidade).

5.3.5 Análise da quantidade de dados para treinamento

Para essa análise, a classificação da rede foi avaliada com diferentes quantidades de dados para treinamento e teste. Os dados foram divididos nas seguintes proporções para treinamento e teste (porcentagem de dados para treinamento / porcentagem de dados para teste): 20/80, 30/70, 40/60, 50/50, 60/40, 70/30 e 80/20. As quantidades de neurônios utilizadas na camada escondida da rede na classificação considerando essas diferentes proporções foram: 53 neurônios (20/80), 50 neurônios (30/70), 53 neurônios (40/60), 53 neurônios (50/50), 53 neurônios (60/40), 50 neurônios (70/30) e 53 neurônios (80/20). O grupo de sensores que apresentou o melhor desempenho na análise de grupos (sensores flexíveis e inercial) foi utilizado e as características utilizadas foram: valor médio, valor máximo e valor mínimo.

Os resultados obtidos na classificação com os dados divididos em diferentes proporções são apresentados na Figura 53. Pode-se observar que existe uma tendência de aumento da taxa de acertos em relação à quantidade de dados para treinamento. Com esse resultado, acredita-se que quanto maior a quantidade de dados para treinamento, mais a rede pode aprender.

Figura 53 – Classificação obtida na rede neural com os dados divididos em diferentes proporções para treinamento e teste.



Fonte: Autoria Própria.

Ao longo deste capítulo foram apresentados e discutidos os resultados obtidos nas análises de comportamento dos sensores e na classificação das letras do alfabeto e dos gestos de palavras em diferentes cenários.

5.4 COMPARAÇÃO DOS RESULTADOS

Lee e Lee (2018) utilizaram SVM para classificar as 26 letras do alfabeto da Língua Americana de Sinais. Na primeira versão foram utilizados cinco sensores flexíveis e um sensor inercial (acelerômetro e giroscópio). A taxa média de acerto obtida foi de 65,7%. Já na segunda versão, utilizando cinco sensores flexíveis, um inercial e dois sensores de contato (pressão), a taxa de acerto foi de 98,2%. No trabalho de Shaheen e Mehmood (2018) cinco sensores flexíveis, um sensor inercial

e dois sensores de contato foram utilizados para classificar 26 letras e 10 números da Língua de Sinais do Paquistão. Para a classificação, um mapa de gestos foi utilizado e a taxa média de acerto obtida foi de 93,4%.

Esses trabalhos podem ser comparados com a classificação das 26 letras do alfabeto em Libras no cenário de grupos, realizada no presente trabalho, em que, com cinco sensores flexíveis e um sensor inercial, obteve-se uma taxa média de acerto de 92,62%. Já com a utilização de cinco sensores flexíveis, um sensor inercial e dois sensores de contato, a taxa média de acertos foi de 96,19%.

Abualola e seus colaboradores (2016) utilizaram seis sensores inerciais para classificar 26 letras da Língua de Sinais Americana. A taxa média de acerto obtida foi de 85% com o classificador LDA. Esse trabalho se assemelha com a classificação realizada no presente trabalho, para as 26 letras do alfabeto em Libras, no cenário de sensores individuais. Utilizando apenas um acelerômetro, obteve-se uma taxa média de acerto de 41,6%. Já utilizando apenas um giroscópio, a taxa média de acerto foi de 32,1%.

Kau e seus colaboradores (2015) utilizaram mapa de gestos para classificar cinco palavras da Língua de Sinais de Taiwan. Como essas palavras são executadas com as duas mãos, em cada mão foram utilizados cinco sensores flexíveis e um sensor inercial. A taxa média de acertos alcançada foi de 94,56%. Pode-se comparar esse resultado com o obtido no presente trabalho, na classificação de 10 palavras em Libras no cenário de grupos. Utilizando 5 sensores flexíveis e um sensor inercial, obteve-se uma taxa média de acerto de 98,96%. Uma das diferenças entre os trabalhos é que as palavras em Libras selecionadas no presente trabalho são executadas com apenas uma das mãos.

Comparando-se os resultados dos trabalhos relacionados com os resultados obtidos nesse trabalho, nota-se diferenças na taxa de classificação. Em alguns casos as taxas de acerto obtidas no presente trabalho foram superiores aos trabalhos comparados, e em alguns casos as taxas de acerto foram inferiores. Acredita-se que essas diferenças ocorrem devido a quantidade de sensores e de voluntários utilizados na coleta de dados, além do classificador utilizado e do método de segmentação e de extração de característica dos sinais. Além disso, os gestos da língua de sinais possuem variações para cada país, fator que também pode ter contribuído para as diferentes taxas de classificação obtidas.

6 CONSIDERAÇÕES FINAIS E CONCLUSÃO

Esta pesquisa apresentou um sistema desenvolvido para classificação de padrões de gestos em Libras utilizando redes neurais. Para realizar a leitura dos gestos, foi desenvolvida uma luva que possui acelerômetro, giroscópio, sensores flexíveis e de contato. Para a coleta de dados, duas versões de sistemas de aquisição de dados foram desenvolvidas e, então, o trabalho foi dividido em dois cenários. No primeiro, foi utilizado um sistema de aquisição de dados com fio e no segundo, foi utilizado um sistema de aquisição de dados sem fio. Após a aquisição, os dados foram submetidos a segmentação e extração de características.

Para conhecer o comportamento dos sensores instalados na luva, foram analisadas as saídas do acelerômetro, do giroscópio e dos sensores flexíveis. Foi possível identificar que existe relação entre os valores de tensão e o grau de curvatura dos sensores flexíveis. Os sensores de 5,6 cm e de 9,52 cm apresentaram o grau de correlação de 0,9868 e 0,9800, respectivamente, entre a tensão de saída e o grau de curvatura dos sensores. Em seguida, foi possível identificar, por meio de uma análise nos sinais dos três eixos do acelerômetro e do giroscópio, que existe um padrão diferente para cada tipo de movimento executado em Libras (inclinação, movimento linear e movimento de giro).

No cenário 1, utilizando a luva com o sistema de aquisição com fio, cinco voluntários executaram as letras do alfabeto em Libras. Foi possível identificar que a média, o valor mínimo e o valor máximo são as características que apresentam os melhores resultados na classificação. Foi possível identificar que todos os sensores (inercial, flexíveis e de contato) são importantes na classificação dos padrões dos sinais do alfabeto em Libras. Com os sensores flexíveis foi possível identificar todas as configurações de mão dos gestos do alfabeto, com exceção das configurações que envolvem sobreposição e abdução lateral dos dedos indicador e médio. O acelerômetro mostrou-se útil na identificação da orientação da mão e de movimentos lineares. O giroscópio mostrou-se útil na identificação da orientação da mão e de movimentos de giro. Já os sensores de contato foram úteis para identificar sobreposição e abdução lateral dos dedos indicador e médio. O grupo com todos os sensores foi o que obteve a melhor taxa de acerto 96,19%.

Na análise utilizando o método *leave-one-subject-out*, os resultados apontaram que a taxa de acerto varia para cada voluntário. Acredita-se que é devido

às diferentes formas e costumes de execução dos gestos. Por fim, a análise da quantidade de dados para treinamento mostrou que existe uma tendência de aumento da taxa de acerto, conforme ocorre o aumento da quantidade de dados separados para treinamento.

No cenário 2, utilizando a luva com o sistema de aquisição sem fio, 10 voluntários executaram 10 gestos de palavras em Libras. As características que apresentaram os melhores resultados na classificação foram: valor mínimo, valor máximo e média. Na análise individual por tipos de sensores, foi possível identificar que os sensores de contato não são úteis na discriminação desses dez gestos de palavras. Com os sensores flexíveis, foi possível discriminar as configurações de mão dos gestos, porém alguns gestos que possuem configurações de mão quase parecidas foram confundidos. O acelerômetro e o giroscópio foram úteis para discriminar os gestos que possuem uma mistura de movimento linear e de giro. A classificação utilizando o acelerômetro, giroscópio e os sensores flexíveis foi a que apresentou a melhor taxa de acerto 98,96%.

Na análise utilizando o método *leave-one-subject-out*, alguns voluntários obtiveram melhores taxas de acerto do que outros. Acredita-se que a forma de executar os gestos (variações para o mesmo gesto) e os erros cometidos durante a fase de coleta contribuíram para obter esses resultados. Por isso, com um treinamento prévio dos voluntários ou com a participação de voluntários fluentes em Libras na coleta de dados, os erros de execução dos gestos poderão ser menores. A análise da quantidade de dados para treinamento apresentou tendência de aumento da taxa de acerto, de acordo com o aumento da quantidade de dados para treinamento. Isso indica que quanto maior a quantidade de dados para treinamento, mais a rede adquire conhecimento.

Por fim, o método proposto se mostrou válido na identificação de alguns tipos de gestos em Libras utilizando a luva instrumentada e rede neural para classificação. Para trabalhos futuros, mais sensores podem ser incluídos na luva com o objetivo de melhorar a classificação dos gestos e pode-se sugerir algumas melhorias:

- Utilizar as duas mãos para o reconhecimento de gestos;
- Coletar dados com fluentes em Libras;
- Aumentar o tamanho do banco de dados;
- Aumentar o vocabulário de gestos;

- Realizar a classificação de frases;
- Realizar a classificação de forma *online*;
- Analisar a extração de características por tipo de sensor;

Com o desenvolvimento dessa pesquisa, dois trabalhos foram apresentados em eventos (um resumo e um artigo) e um trabalho foi submetido e está sendo avaliado:

DIAS, T. S.; PICHORIM, S. F. Luva com Sensores Flexíveis para Reconhecimento de Libras. In: XXVI CONGRESSO BRASILEIRO DE ENGENHARIA BIOMÉDICA, 2018, Armação de Búzios, RJ. **Anais...** 2018.

DIAS, T. S.; MENDES JUNIOR, J. J. A. ; PICHORIM, S. F. Luva instrumentada para reconhecimento de padrões de gestos em Libras. In: XII Simpósio de Engenharia Biomédica - IX Simpósio de Instrumentação e Imagens Médicas, 2019, Uberlândia. **Anais...** 2019.

Título do artigo submetido: *Analysis of Sensors in the Classification of the Brazilian Sign Language.*

Evento: *42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society in conjunction with the 43rd Annual Conference of the Canadian Medical and Biological Engineering Society.*

REFERÊNCIAS

- ABAS, B.; RAHMAN, K. Fabrication of flex sensors through direct ink write technique and its electrical characterization. **Applied Physics A**, v. 122, n. 972, 2016.
- ABREU, J. G.; TEIXEIRA, J. M.; FIGUEIREDO, L. S.; TEICHRIEB, V. Evaluating Sign Language Recognition Using the Myo Armband. In: XVIII SYMPOSIUM ON VIRTUAL AND AUGMENTED REALITY (SVR), 2016, Gramado. **Anais... IEEE**, 2016.
- ABUALOLA, H.; GHOTHANI, H. A.; EDDIN, A. N.; ALMOOSA, N.; POON, K. Flexible gesture recognition using wearable inertial sensors. In: 2016 IEEE 59TH INTERNATIONAL MIDWEST SYMPOSIUM ON CIRCUITS AND SYSTEMS (MWSCAS), 2016, Abu Dhabi. **Anais... IEEE**, 2016.
- ALMEIDA, W. G. **Introdução à Língua Brasileira de Sinais**. 1. ed. Ilhéus: UAB/UESC, 2013.
- AMBAR, R.; FAI, C. K.; WAHAB, M. H. A.; JAMIL, M. M. A.; MARADZI, A. A. Development of a Wearable Device for Sign Language Recognition. **Journal of Physics: Conference Series**, v. 1019, n. 1, 2018.
- AMERINI, I.; BESTAGINI, P.; BONDI, L.; CALDELLI, R.; CASINI, M.; TUBARO, S. Robust smartphone fingerprint by mixing device sensors features for mobile strong authentication. **Electronic Imaging**, v. 2016, n. 8, p. 1-8, 2016.
- ANALOG DEVICES. **Accelerometer ADXL335**. 2009. Disponível em: <<https://www.analog.com/en/products/adxl335.html>>. Acesso em: 15 dez. 2019.
- ARDUINO. **ARDUINO MEGA 2560**. 2020. Disponível em: <<https://store.arduino.cc/usa/mega-2560-r3>>. Acesso em: 15 fev. 2020.
- BAKER, A.; BOGAERDE, B. V. D.; PFAU, R.; SCHERMER, T. **The Linguistics of Sign Language: An Introduction**. Amsterdam: John Benjamins Publishing Company, 2011.
- BALBINOT, A.; BRUSAMARELLO, V. J. **Instrumentação e Fundamentos de Medidas**. 3. ed. Rio de Janeiro: LTC, 2019.
- BIBLIOTECA VIRTUAL EM SAÚDE. **Surdez**. 2017. Disponível em: <<https://bvsmms.saude.gov.br/dicas-em-saude/2506-surdez>>. Acesso em: 1 fev. 2020.
- BONTRAGER, K. L.; LAMPIGNANO, J. P. **Tratado de Posicionamento Radiográfico e Anatomia Associada**. Rio de Janeiro: Elsevier, 2015.
- BRASIL. Lei nº 10.436, de 24 de abril de 2002. **Diário Oficial da União**, Brasília, DF, 25 de abril de 2002. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/2002/l10436.htm>. Acesso em: 02 fev. 2020.

BRASIL. Lei nº 13.146, de 6 de julho de 2015. **Diário Oficial da União**, Brasília, DF, 7 de julho de 2015. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/l13146.htm>. Acesso em: 02 fev. 2020.

CARNEIRO, S. B.; SANTOS, E. D. F. M.; BARBOSA, T. M. G. A.; FERREIRA, J. O.; ALCALÁ, S. G. S.; DA ROCHA, A. F. Static gestures recognition for Brazilian Sign Language with kinect sensor. In: 2016 IEEE SENSORS 2016, Orlando, USA. **Anais... IEEE**, 2016.

CHEOK, M. J.; OMAR, Z.; JAWARD, M. H. A review of hand gesture and sign language recognition techniques. **International Journal of Machine Learning and Cybernetics**, v. 10, n. 1, p. 131–153, 2017.

CUARELLI, G. **Estimação de parâmetros da marcha humana a partir de medições feitas por acelerômetros e giroscópios e uso de mecanismo de referência**. 2013. Dissertação (Mestrado Profissional em Automação e Controle de Processos) - Instituto Federal de Educação, Ciência e Tecnologia de São Paulo, São Paulo, 2013.

CYBERGLOVE. **CyberGlove II: Overview**. 2017. Disponível em: <<http://www.cyberglovesystems.com/cyberglove-ii>>. Acesso em: 27 mar. 2020.

DIAS, T. S.; PICHORIM, S. F. Luva com Sensores Flexíveis para Reconhecimento de Libras. In: XXVI CONGRESSO BRASILEIRO DE ENGENHARIA BIOMÉDICA, 2018, Armação de Búzios, RJ. **Anais...** 2018.

DIZEU, L. C. T. de B.; CAPORALI, S. A. A língua de sinais constituindo o surdo como sujeito. **Educação & Sociedade**, v. 26, n. 91, p. 583–597, 2005.

ESPRESSIF. **Esp32-WROOM-32 Datasheet**. 2019. Disponível em: <<https://www.espressif.com/en/esp-wroom-32/resources>>. Acesso em: 7 dez. 2019.

FACELI, K.; LORENA, A. C.; GAMA, J.; DE CARVALHO, A. C. P. L. F. **Inteligência Artificial: Uma Abordagem de Aprendizado de Máquina**. 1. ed. Rio de Janeiro: LTC, 2011.

FILHO, C. F. C.; SANTOS, B. L.; SOUZA, R. S.; SANTOS, J. R.; COSTA, M. G. F. A new method for recognizing hand configurations of Brazilian gesture language. In: 38TH ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY (EMBC), 2016, Orlando, FL, USA. **Anais... IEEE**, 2016.

GHANEM, S.; CONLY, C.; ATHITSOS, V. A Survey on Sign Language Recognition Using Smartphones. In: THE 10TH INTERNATIONAL CONFERENCE, 2017, Island of Rhodes, Grécia. **Proceedings...** ACM Press, 2017.

GOLDFEALD, M. **A criança surda: Linguagem e cognição numa perspectiva sociointeracionista**. São Paulo: Plexus, 2002.

HAND TALK. **Aplicativo de celular Hand Talk: versão 2.7.15**. Disponível em: <<https://www.handtalk.me/br/Aplicativo>>. Acesso em: 20 fev. 2020.

HAQ, E. S.; SUWARDIYANTO, D.; HUDA, M. Indonesian Sign Language Recognition Application For Two-Way Communication Deaf-Mute People. In: 2018 3RD INTERNATIONAL CONFERENCE ON INFORMATION TECHNOLOGY, INFORMATION SYSTEM AND ELECTRICAL ENGINEERING (ICITISEE), Yogyakarta, Indonésia, 2018. **Anais... IEEE**, 2019.

HAYKIN, S. **Redes Neurais: princípios e prática**. 2. ed. Porto Alegre: Bookman, 2007.

IBGE. **Censo Demográfico de 2010**. 2010. Disponível em: <<https://www.ibge.gov.br/estatisticas/sociais/populacao/9662-censo-demografico-2010.html?edicao=9749&t=destaques>>. Acesso em: 5 fev. 2020.

INVENSENSE. **MPU-6000 and MPU-6050 Product Specification: Revision 3.4**, 2013.

JADHAV, D.; TRIPATHY, A. Gesture Aided Speech for Deaf and Mute. In: 2018 3RD INTERNATIONAL CONFERENCE FOR CONVERGENCE IN TECHNOLOGY (I2CT) 2018, Pune, Índia. **Anais... IEEE**, 2018.

JANE, S. P. Y.; SASIDHAR, S. Sign Language Interpreter: Classification of Forearm EMG and IMU Signals for Signing Exact English. In: IEEE 14TH INTERNATIONAL CONFERENCE ON CONTROL AND AUTOMATION (ICCA), Anchorage, USA, 2018. **Anais... IEEE**, 2018.

JANI, A. B.; KOTAK, N. A.; ROY, A. K. Sensor Based Hand Gesture Recognition System for English Alphabets Used in Sign Language of Deaf-Mute People. In: 2018 IEEE SENSORS 2018, New Delhi, Índia. **Anais... IEEE**, 2018.

JEPSEN, J. B.; De Clerck, G.; Lutalo-Kiingi, S.; McGregor, W. B. **Sign Languages of the world: A Comparative Handbook**. Berlim: De Gruyter Mouton, 2015.

KAU, L.; SU, W.; YU, P.; WEI, S. A real-time portable sign language translation system. 2015. In: IEEE 58TH INTERNATIONAL MIDWEST SYMPOSIUM ON CIRCUITS AND SYSTEMS (MWSCAS), Fort Collins, USA, 2015. **Anais... IEEE**, 2015.

KUMADA, K. M. O. **Libras – língua brasileira de sinais**. Londrina: Editora e Distribuidora Educacional S.A, 2016.

LEE, B. G.; LEE, S. M. Smart Wearable Hand Device for Sign Language Interpretation System With Sensors Fusion. **IEEE Sensors Journal**, v. 18, n. 3, p. 1224–1232, 2018.

LIMA, I.; PINHEIRO, C. A. M.; SANTOS, F. A. O. **Inteligência Artificial**. 1. ed. Rio de Janeiro: Elsevier, 2014.

LOPES, M. A. C.; LEITE, L. P. Concepções de surdez: a visão do surdo que se comunica em língua de sinais. **Revista Brasileira de Educação Especial**, v. 17, n. 2, p. 305–320, ago. 2011.

MAENAKA, K. MEMS inertial sensors and their applications. In: 5TH INTERNATIONAL CONFERENCE ON NETWORKED SENSING SYSTEMS, 2008, Kanazawa, Japão. **Anais... IEEE**, 2008.

MASOUD, M.; JARADAT, Y.; MANASRAH, A.; JANNOD, I. Sensors of Smart Devices in the Internet of Everything (IoE) Era: Big Opportunities and Massive Doubts. **Journal of Sensors**, v. 2019, 2019.

MATHWORKS. **Documentation: skewness**. 2020. Disponível em: <www.mathworks.com/help/stats/skewness.html>. Acesso em: 5 jan. 2020.

MENDES JUNIOR, J. J. A.; FREITAS, M. L. B.; STEVAN JR, S. L.; PICHORIM, S. F. Recognition of Libras Static Alphabet with MyoTM and Multi-Layer Perceptron. In: XXVI Brazilian Congress on Biomedical Engineering, 2018, Búzios, RJ. **Proceedings...** Springer, 2019.

MONTALVO, P. D. R.; TRUJILLO, P. G.; BOSMEDIANO, E. F.; GARCÍA, J. C.; POTOSI, S. O.; PEREIRA, H. B.; ORDÓÑEZ, D. H. P. Sign Language Recognition Based on Intelligent Glove Using Machine Learning Techniques. In: 2018 IEEE THIRD ECUADOR TECHNICAL CHAPTERS MEETING (ETCM) 2018, Cuenca. **Anais...** IEEE, 2018.

MOOREL, K. L.; DALLEY, A. F.; AGUR, A. M. R. **Anatomia Orientada para Clínica**. [s.l.] : Grupo GEN, 2018.

MORAIS, C. E. L.; PLINSKI, R. R. K.; MARTINS, G. P. T. C.; SZULCZEWSKI, D. M. **Libras**. 2. ed. Porto Alegre: SAGAH, 2018.

NIST/SEMATECH. **e-Handbook of Statistical Methods**. 2012. Disponível em: <<https://www.itl.nist.gov/div898/handbook/index.htm>>. Acesso em: 14 set. 2019.

OLIVEIRA, F. E. M. **Estatística e Probabilidade: com ênfase em exercícios resolvidos e propostos**. 3. ed. [s.l.]: LTC, 2017.

PEREIRA, M. C. C.; CHOI, D.; VIEIRA, M. I.; GASPAR, P.; NAKASATO, R. **Libras: Conhecimento além dos sinais**. São Paulo: Pearson, 2011.

POLLMANN, M. C. F. A patologia da fala no tratamento ortodôntico. **Repositório Aberto da Universidade do Porto**, ACTAS SPODF, n. 3, p. 77-86 1994.

PUROHIT, K.; PATIDAR, K.; KUSHWAH, R. S. A Wearable Hand Gloves Gesture Detection based on Flex Sensors for disabled People. **International Journal for Research in Applied Science and Engineering Technology**, [s. l.], v. 5, n. 8, p. 1825–1833, 2017.

QUADROS, R. M. **Educação de Surdos: A Aquisição da Linguagem**. Porto Alegre: Artmed, 2008.

QUADROS, R. M.; KARNOPP, L. B. **Língua de Sinais Brasileira: Estudos Lingüísticos**. Porto Alegre: Artmed, 2007.

RAMOS, C. R. **LIBRAS: A Língua de Sinais dos Surdos Brasileiros**. Rio de Janeiro: Arara Azul, 2002.

RASHID, A.; HASAN, O. Wearable technologies for hand joints monitoring for rehabilitation: A survey. **Microelectronics Journal**, v. 88, p. 173–183, 2018.

REWARI, H.; DIXIT, V.; BATRA, D.; HEMA, N. Automated Sign Language Interpreter. In: 2018 ELEVENTH INTERNATIONAL CONFERENCE ON CONTEMPORARY COMPUTING (IC3), Noida, Índia, 2018. **Anais...** 2018.

SAGGIO, G.; ORENGO, G. Flex sensor characterization against shape and curvature changes. **Sensors and Actuators A: Physical**, v. 273, p. 221–231, 2018.

SAHOO, J. P.; ARI, S.; GHOSH, D. K. Hand gesture recognition using DWT and F-ratio based feature descriptor. **IET Image Processing**, v. 12, n. 10, p. 1780–1787, 2018.

SHAEFFER, D. K. MEMS inertial sensors: A tutorial overview. **IEEE Communications Magazine**, v. 51, n. 4, p. 100-109, 2013.

SHAHEEN, H.; MEHMOOD, T. Talking Gloves: Low-Cost Gesture Recognition System for Sign Language Translation. In: 2018 IEEE REGION TEN SYMPOSIUM (TENSYP), 2018, Sydney, Australia. **Anais...** IEEE, 2018.

SHIELDEX. **Medtex P180**, V Technical Textiles, 2018.

SILVA, B. C. R.; FURRIEL, G. P.; PACHECO, W. C.; BULHOES, J. S. Methodology and comparison of devices for recognition of sign language characters. In: 2017 18TH INTERNATIONAL SCIENTIFIC CONFERENCE ON ELECTRIC POWER ENGINEERING (EPE), 2017, Kouty nad Desnou, Czech Republic. **Anais...** IEEE, 2017.

SILVA, I. N. D.; SPATTI, D. H.; FLAUZINO, R. A. **Redes Neurais Artificiais: para engenharia e ciências aplicadas**. São Paulo: Artliber, 2010.

SILVA, M. B. F. **Modelagem de chaves MEMS para aplicações em RF**. 2007. Dissertação (Mestrado em Engenharia Elétrica) - Escola Politécnica da Universidade de São Paulo, São Paulo, 2007.

SPECTRA SYMBOL. **FLEX SENSORS DATA SHEET**, 2014. Disponível em: <<https://www.spectrasymbol.com/product/flex-sensors/>>. Acesso em: 6 fev. 2020.

STEVAN JR, S. L.; SILVA, R. A. **Automação e Instrumentação Industrial com Arduino: Teoria e Projetos**. 1. ed. São Paulo: Érica, 2015.

STROBEL, K. L.; FERNANDES, S. **Aspectos Linguísticos de Libras**. Curitiba: SEED/DEE, 1998.

TEODORO, B. T.; BERNARDES, J.; DIGIAMPIETRI, L. A. Skin Color Segmentation and Levenshtein Distance Recognition of BSL Signs in Video. In: 30TH SIBGRAPI CONFERENCE ON GRAPHICS, PATTERNS AND IMAGES (SIBGRAPI) 2017, Niteroi, 2017. **Anais...** IEEE, 2017.

WU, J.; SUN, L.; JAFARI, R. A Wearable System for Recognizing American Sign Language in Real-Time Using IMU and Surface EMG Sensors. **IEEE Journal of Biomedical and Health Informatics**, v. 20, n. 5, p. 1281–1290, 2016.

XIE, B.; HE, X.; LI, Y. RGB-D static gesture recognition based on convolutional neural network. **The Journal of Engineering**, v. 2018, n. 16, p. 1515–1520, 2018.

ANEXO A – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (TCLE)

Título da pesquisa: Processamento de sinais biológicos utilizando fusão de sensores

Pesquisador: José Jair Alves Mendes Junior, UTFPR – Campus Ponta Grossa – Centro de Instrumentação nas Engenharias Biomédica e Esportiva (CIEBE).
Telefones: (42) 999967000

Pesquisador Responsável: Sérgio Francisco Pichorim

Local de realização da pesquisa: Centro de Instrumentação nas Engenharias Biomédica e Esportiva (CIEBE), UTFPR – Ponta Grossa – PR

A) INFORMAÇÕES AO PARTICIPANTE

1. Apresentação da pesquisa.

Você está sendo convidado a participar de uma pesquisa que avaliará padrões envolvendo reconhecimento de movimentos dos membros superiores e inferiores utilizando aparelhos de eletromiografia e sensores dinâmicos. Estes equipamentos, que não são invasivos nem causam dor, fazem o registro de sinais de movimentação dos membros por meio de sensores alocados na superfície da pele. Os eletrodos de eletromiografia e os sensores dinâmicos são posicionados nas regiões de interesse da leitura do movimento para coletar dados da resposta muscular durante a realização dos mesmos. A partir destes dados, serão investigados padrões de movimentação (gestos, passos, entre outros). Esse estudo deve contribuir para a investigação de detecção de gestos para tecnologias assistivas, análise de fadiga muscular e de análises postural e de marcha.

2. Objetivos da pesquisa.

O objetivo geral deste trabalho é analisar a técnica de fusão de sensores utilizando as informações provenientes de sensores não invasivos para análise de padrões.

3. Participação na pesquisa.

Sua participação na pesquisa será uma única vez e durará em torno de 30 minutos. Caso o convidado seja do sexo masculino, solicita-se que neste dia retire os pêlos das regiões em que serão alocados os eletrodos. Você ficará em uma sala em temperatura ambiente enquanto serão realizados as aquisições. Serão posicionados os sensores e será pedido para que se realize uma série de movimentos, como por exemplo gestos

com as mãos e andar uma quantidade de passos. Após a coleta, o equipamento será retirado.

4. Confidencialidade.

A sua privacidade será respeitada, ou seja, o seu nome ou qualquer outro dado, informação ou elemento que possa de qualquer forma lhe identificar, será mantido em sigilo. Os pesquisadores se responsabilizam pela guarda e confidencialidade das informações.

5. Riscos e Benefícios.

5a) Riscos: Os dispositivos desenvolvidos, por si só, não apresentam riscos para os indivíduos em que serão usados. Isso porque serão observadas as normas de segurança e operação de equipamentos biomédicos (como norma ANSI/AAMI ES1-1993, que trata sobre os limites de segurança da corrente elétrica em equipamentos eletro-médicos), além de se tratar de sistemas não invasivos.

Outros riscos que podem acontecer são de origem cotidiana, como por exemplo, o indivíduo tropeçar, colidir ou se engasgar. Entretanto, o ambiente será preparado para o experimento, onde todos os cuidados serão tomados. Todos os experimentadores serão orientados para evitar qualquer tipo de risco. Em caso de acidente, o Serviço de Atendimento Móvel de Urgência (Samu) será solicitado.

No momento da montagem dos sensores, poderá existir risco por constrangimento. Para evitar esse tipo de risco, o voluntário não precisará responder alguma pergunta caso não ache conveniente. Além disso, poderá optar ser auxiliado por um experimentador do sexo masculino ou feminino. Se o voluntário não se sentir à vontade poderá suspender o teste em qualquer momento.

5b) Benefícios: Dentre os benefícios desta pesquisa, o primeiro de destaque é a possibilidade dos voluntários do teste terem acesso aos seus dados individuais dos parâmetros coletados. O voluntário terá benefícios culturais ao conhecer um laboratório de pesquisa e ao conhecer ferramentas tecnológicas. Além disso, obterá uma explicação sobre alguns conceitos da pesquisa e poderá fazer perguntas relacionadas à engenharia biomédica, esportiva, eletrônica, elétrica e da computação. Como um benefício social, o voluntário será convidado para um feedback ao final da pesquisa sobre quais foram os resultados e desenvolvimentos que foram alcançados com a sua ajuda.

6. Critérios de inclusão e exclusão.

6a) Inclusão: Indivíduos de sexo masculino e feminino, com idade maior que 18 anos, sem a presença de nenhuma limitação física, psicológica ou fisiológica que impeça a execução do roteiro experimental proposto.

6b) Exclusão: Voluntários que não atenderam às solicitações que possibilitam a execução do experimento, como tricotomia da região estudada. Voluntários que não possam participar do experimento em sua íntegra ou que por algum motivo decidam abandonar o experimento durante a coleta de dados. Voluntários que apresentaram problemas com relação à montagem dos sensores como: aqueles que se sentirem desconfortáveis com os sensores ou, que por motivo pessoal ou técnico, possam alterar os resultados. Voluntários que apresentaram testes mal executados, com dificuldades técnicas ou de compreensão. Voluntários que no dia do teste apresentarem indisposição física, psicológica ou fisiológica que impeça a execução do roteiro experimental proposto.

7. Direito de sair da pesquisa e a esclarecimentos durante o processo.

É assegurada à você assistência durante toda a pesquisa, bem como é garantido o seu livre acesso a todas as informações e esclarecimentos adicionais sobre o estudo e suas consequências, enfim, tudo o que você queira saber antes, durante e depois da sua participação. Pode também se recusar a participar do estudo, ou se retirar a qualquer momento sem qualquer constrangimento ou penalização.

Você pode assinalar o campo a seguir, para receber o resultado desta pesquisa, caso seja de seu interesse :

() quero receber os resultados da pesquisa (email para envio : _____)

() não quero receber os resultados da pesquisa

8. Ressarcimento e indenização.

Como você não terá gastos ao participar desta pesquisa, não haverá ressarcimento. De igual maneira, caso ocorra algum dano decorrente da sua participação no estudo, será devidamente indenizado, conforme determina a lei.

ESCLARECIMENTOS SOBRE O COMITÊ DE ÉTICA EM PESQUISA:

O Comitê de Ética em Pesquisa envolvendo Seres Humanos (CEP) é constituído por uma equipe de profissionais com formação multidisciplinar que está trabalhando para assegurar o respeito aos seus direitos como participante de pesquisa. Ele tem por

objetivo avaliar se a pesquisa foi planejada e se será executada de forma ética. Se você considerar que a pesquisa não está sendo realizada da forma como você foi informado ou que você está sendo prejudicado de alguma forma, entre em contato com o Comitê de Ética em Pesquisa envolvendo Seres Humanos da Universidade Tecnológica Federal do Paraná (CEP/UTFPR). Av. Sete de Setembro, 3165, Bloco N, Térreo, Bairro Rebouças, CEP 80230-901, Curitiba-PR, telefone: (41) 3310-4494, e-mail: coep@utfpr.edu.br.

B) CONSENTIMENTO

Declaro que li e entendi todas as informações presentes neste Termo de Consentimento Livre e Esclarecido e tive a oportunidade de discutir as informações deste termo. Todas as minhas perguntas foram respondidas e eu estou satisfeito com as respostas. Entendo que receberei uma via assinada e datada deste documento e que outra via assinada e datada será arquivada pela pesquisadora responsável pelo estudo.

Enfim, tendo sido orientado quanto a todo o teor aqui mencionado e compreendido a natureza e o objetivo do já referido estudo, manifesto meu livre consentimento em participar, estando totalmente ciente de que não há nenhum valor econômico a receber ou a pagar, por minha participação.

Após reflexão e um tempo razoável, eu decidi, livre e voluntariamente, participar deste estudo.

Concordo que o material e as informações obtidas relacionadas a minha pessoa possam ser publicados em aulas, congressos, eventos científicos, palestras ou periódicos científicos. Porém, não devo ser identificado por nome ou qualquer outra forma.

Nome
completo: _____

RG: _____ Data _____ de
Nascimento: ___/___/___ Telefone: _____

Endereço:

CEP: _____ Cidade: _____ Estado: _____

Assinatura: _____ Data: ___/___/_____

Eu declaro ter apresentado o estudo, explicado seus objetivos, natureza, riscos e benefícios e ter respondido da melhor forma possível às questões formuladas.

Assinatura

pesquisador: Data: ___/___/___

(ou seu representante)

Nome

completo: _____

Para todas as questões relativas ao estudo ou para se retirar do mesmo, poderão se comunicar com _____, via e-mail: _____ ou telefone: _____.