

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

LUIGI DE LA CRUZ FAXINA

**SISTEMA DE INTERFACE HOMEM-MÁQUINA PARA RECONHECIMENTO DE
PADRÕES DE GESTOS EM LIBRAS**

CURITIBA

2024

LUIGI DE LA CRUZ FAXINA

**SISTEMA DE INTERFACE HOMEM-MÁQUINA PARA RECONHECIMENTO DE
PADRÕES DE GESTOS EM LIBRAS**

**Human-machine interface system for recognition of gesture patterns in
Brazilian Sign Language (Libras)**

Trabalho de Conclusão de Curso de Graduação
apresentado como requisito para obtenção do
título de Bacharel em Engenharia Eletrônica
do Curso de Bacharelado em Engenharia
Eletrônica da Universidade Tecnológica Federal
do Paraná.

Orientador: Prof. Dr. José Jair Alves Mendes
Junior

Coorientador: Me. Thiago Simões Dias

CURITIBA

2024



[4.0 Internacional](https://creativecommons.org/licenses/by/4.0/)

Esta licença permite compartilhamento, remixe, adaptação e criação a partir do trabalho, mesmo para fins comerciais, desde que sejam atribuídos créditos ao(s) autor(es). Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

LUIGI DE LA CRUZ FAXINA

**SISTEMA DE INTERFACE HOMEM-MÁQUINA PARA RECONHECIMENTO DE
PADRÕES DE GESTOS EM LIBRAS**

Trabalho de Conclusão de Curso de Graduação
apresentado como requisito para obtenção do
título de Bacharel em Engenharia Eletrônica
do Curso de Bacharelado em Engenharia
Eletrônica da Universidade Tecnológica Federal
do Paraná.

Data de aprovação: 28/junho/2024

José Jair Alves Mendes Junior
Título Prof. Dr.
Universidade Tecnológica Federal do Paraná

Carmen Caroline Rasera
Título Profa. Dra
Universidade Tecnológica Federal do Paraná

Renata Oliveira de Sousa
Título Profa. Dra
Universidade Tecnológica Federal do Paraná

Thiago Simões Dias
Título Me.
Universidade Tecnológica Federal do Paraná

CURITIBA

2024

Dedico este trabalho a meus pais, cujo amor e apoio sempre foram minha fonte de inspiração.

À minha família e amigos, por sua compreensão, paciência e incentivo ao longo dessa longa jornada. E a todas as pessoas que, de alguma forma, contribuíram para a realização deste trabalho.

AGRADECIMENTOS

Gostaria de expressar minha sincera gratidão ao meu orientador, Prof. Dr. José Jair Alves Mendes Junior, pela orientação, incentivo e sabedoria compartilhada ao longo deste trabalho. Agradeço também ao professor Prof. Me. Thiago Simões Dias pela dedicação, paciência e pelos valiosos ensinamentos que contribuíram significativamente para o desenvolvimento deste projeto. Ao Matheus Taborda Carneiro, pela imprescindível ajuda no laboratório BIOTECA(Laboratório de Biossinais e Tecnologias Assistivas), sua colaboração foi fundamental para o sucesso deste trabalho. À minha família - meu pai Adilson Faxina, minha mãe Rita Pilar e minhas irmãs, Mariana e Mayara - meu mais profundo reconhecimento por serem meu porto seguro, por seu amor, apoio e por servirem como minha fonte de inspiração constante. Aos meus amigos, verdadeiros companheiros de jornada, pela compreensão e incentivo ao longo dessa trajetória acadêmica. Por fim, estendo meu agradecimento a todas as pessoas que, de alguma forma, contribuíram para a realização deste trabalho.

Inteligência é a habilidade de se adaptar à
mudança.
(Stephen Hawking)

RESUMO

Este trabalho propõe o desenvolvimento de um sistema inovador para a captura e classificação de gestos em Língua Brasileira de Sinais (Libras) utilizando tecnologias de sensoriamento vestível e *Deep Learning* (Aprendizado profundo). A justificativa para essa pesquisa está na necessidade de oferecer alternativas eficazes para a comunicação em Libras, especialmente entre indivíduos com e sem deficiência auditiva. Os objetivos abrangem desde a implementação de luvas instrumentadas com uma variedade de sensores, incluindo sensores flexíveis, inerciais (acelerômetros e giroscópios) e uma *armband* Myo para aquisição de Eletromiografia de superfície (sEMG), além do desenvolvimento do sistema para captura de dados durante a execução dos gestos. Além disso, propõe-se a aplicação de modelos de *Deep Learning* para a classificação dos gestos. Os resultados obtidos revelam uma acurácia promissora na classificação dos gestos, alcançando 89,77% com o modelo gMLP e 92,21% com o *InceptionTime*. Esta abordagem destaca-se por sua inovação ao comparar e integrar diferentes tipos de sensores, como sEMG, sensores flexíveis e inerciais de acelerômetro e giroscópio, e pela eficácia demonstrada na aplicação de técnicas de *Deep Learning* na classificação de gestos em Libras.

Palavras-chave: eletromiografia de superfície; bracelete; luva instrumentada; deep learning; reconhecimento de padrões.

ABSTRACT

This work proposes the development of an innovative system for capturing and classifying gestures in Brazilian Sign Language (Libras) using wearable sensing technologies and Deep Learning. The justification for this research lies in the need to provide effective alternatives for communication in Libras, especially among individuals with and without hearing impairments. The objectives range from implementing instrumented gloves with a variety of sensors, including flexible and inertial sensors (such as accelerometers and gyroscopes), as well as the Armband Myo, to developing the system for data capture during gesture execution. Additionally, the application of Deep Learning models for gesture classification is proposed. The results obtained reveal a promising accuracy in gesture classification, reaching 89.77% with the gMLP model and 92.21% with InceptionTime. This approach stands out for its innovation in comparing and integrating different types of sensors, such as surface electromyography (EMG), flexible sensors, and inertial sensors of accelerometer and gyroscope, and for the demonstrated effectiveness in applying Deep Learning techniques in gesture classification in Libras.

Keywords: surface electromyography; armband; instrumented glove; deep learning; pattern recognition.

LISTA DE FIGURAS

Figura 1 – Luvas desenvolvidas para a identificação de gestos em Língua Brasileira de Sinais (Libras) nos estudos de: (a) Silvano <i>et al.</i> (2016), (b) Azevedo <i>et al.</i> (2017), (c) Santos, Ferreira e Alves (2019) e (d) Borges (2020).	18
Figura 2 – Alfabeto manual em Libras.	19
Figura 3 – Exemplos da mudança de significado dos sinais em Libras ao mudar um parâmetro (configuração da mão, movimento e localização respectivamente) em sua composição.	20
Figura 4 – Músculos superficiais do antebraço direito nas vistas superior (a) e posterior (b).	21
Figura 5 – Tipos de movimentos nos membros superiores: (a) Abdução, (b) Flexão, (c) Rotação.	23
Figura 6 – Sensor flexível 2,2": (a) Imagem do sensor, (b) Esquemático do funcionamento da flexão do sensor.	25
Figura 7 – Circuito divisão de tensão utilizado junto ao sensor flexível 2,2".	25
Figura 8 – MPU-6050.	26
Figura 9 – Eixo neural motor esquelético do sistema nervoso.	27
Figura 10 – Registros eletromiográficos dos músculos extensor e flexor longo do polegar durante flexão e extensão alternados da articulação distal do polegar humano.	27
Figura 11 – ESP32.	29
Figura 12 – Rede <i>Gated Multilayer Perceptron</i> (gMLP).	31
Figura 13 – Rede <i>InceptionTime</i>	32
Figura 14 – Fluxograma do sistema de aquisição de dados.	33
Figura 15 – Representação gráfica da luva instrumentada desenvolvida.	34
Figura 16 – <i>armband Myo™</i>	35
Figura 17 – Fluxograma do funcionamento do sistema de aquisição de dados.	37
Figura 18 – Voluntário utilizando as luvas instrumentadas desenvolvidas nesse projeto junto a <i>armband Myo™</i>	39
Figura 19 – Passos utilizados para o processamento e identificação dos movimentos.	41
Figura 20 – Realização da coleta de dados com um voluntário.	43

Figura 21 – Parte dos dados brutos dos sensores flexíveis e inerciais da luvas direita para a coleta da palavra “Agarrar”.	44
Figura 22 – Parte dos dados brutos dos sensores EMG de uma coleta da palavra “Agarrar”.	45
Figura 23 – Parte dos dados brutos dos sensores flexíveis e inerciais da luvas direita para a coleta da palavra “Escola”.	45
Figura 24 – Parte dos dados brutos dos sensores EMG de uma coleta da palavra “Escola”.	46
Figura 25 – Dados de um sensor de Acelerômetro e inserção de um limiar para identificar movimento ou repouso.	47
Figura 26 – Dados de um canal de EMG e inserção de um limiar para identificar movimento ou repouso.	47
Figura 27 – Dados de um canal de EMG e inserção de um limiar para identificar movimento ou repouso.	48
Figura 28 – Detecção de movimento em um sinal EMG.	48
Figura 29 – Amostra segmentada com dados de uma luva e do bracelete do gesto da palavra “agarrar”.	49
Figura 30 – Divisão da distribuição para a validação cruzada.	49
Figura 31 – Predição dinâmica InceptionTime. Comparação entre a primeira e a última época em um treinamento.	50
Figura 32 – Métricas de validação durante execução das 100 épocas com o modelo InceptionTime.	51
Figura 33 – Distribuições obtidas para as 10 repetições de validação cruzada.	52
Figura 34 – Matriz de confusão obtida durante o treinamento com o modelo <i>InceptionTime</i>.	53
Figura 35 – Importância das sensores (atributos de entrada do classificador) para o reconhecimento dos sinais.	54

LISTA DE TABELAS

Tabela 1 – Grau de liberdade e movimentos realizados em cada articulação dos membros superiores.	22
Tabela 2 – Médias das acurácias obtidas para todas as validações cruzadas de <i>Deep Learning</i> para os modelos gMLP e <i>InceptionTime</i>.	51
Tabela 3 – Maiores confusões entre palavras durante a classificação com o modelo de <i>InceptionTime</i>.	52

LISTA DE ABREVIATURAS E SIGLAS

Siglas

API	Interface de Programação de Aplicações
BLE	<i>Bluetooth Low Energy</i>
DNN	<i>Deep neural networks</i>
dps	<i>degrees per second</i>
DT	Decision Tree
EMG	Eletromiografia
g	Aceleração da gravidade
GL	Graus de liberdade
gMLP	<i>Gated Multilayer Perceptron</i>
GPIO	<i>General Purpose Input/Output</i>
IA	Inteligência Artificial
IBGE	Instituto Brasileiro de Geografia e Estatística
IoT	<i>Internet of Things</i>
KNN	k-Nearest Neighbor
Libras	Língua Brasileira de Sinais
LS	Linguagem de Sinais
ML	<i>Machine Learning</i>
MLP	<i>Multilayer Perceptron</i>
RNA	Rede Neural Artificial
sEMG	Eletromiografia de Superfície
SLR	<i>Sign Language Recognition</i>
tsai	<i>State-of-the-art Deep Learning library for Time Series and Sequences.</i>

SUMÁRIO

1	INTRODUÇÃO	14
1.1	Objetivos	15
1.1.1	Objetivo geral	15
1.1.2	Objetivos específicos	16
1.2	Justificativa	16
1.3	Estado da Arte	17
1.4	Estrutura do trabalho	18
2	REFERENCIAL TEÓRICO	19
2.1	Língua Brasileira de Sinais	19
2.1.1	Características anatômicas para aquisição dos sinais de Libras	20
2.2	Sistemas de reconhecimento de linguagens de sinais	23
2.2.1	Conceitos de sensoriamento e instrumentação	24
2.2.1.1	Sensores resistivos do tipo flexíveis	24
2.2.1.2	Inerciais	25
2.2.1.3	Eletromiografia	26
2.3	ESP32	28
2.4	Deep Learning	29
2.4.1	gMLP	30
2.4.2	InceptionTime	31
3	MATERIAIS E MÉTODOS	33
3.1	Materiais	33
3.1.1	Luva instrumentada	34
3.1.2	Armband Myo™	35
3.1.3	Sistema de aquisição e tratamento de dados	35
3.2	Métodos	36
3.2.1	Coleta de dados	37
3.2.2	Processamento e sincronização dos dados	39
3.2.3	Classificação com Deep Learning	40
4	RESULTADOS E DISCUSSÃO	43
4.1	Sistema físico e aquisição de dados	43

4.2	Sincronização e segmentação dos dados	45
4.3	Classificação com Deep Learning	49
4.4	Discussão	53
5	CONSIDERAÇÕES FINAIS	56
	REFERÊNCIAS	58

1 INTRODUÇÃO

Segundo dados de 2019 do Instituto Brasileiro de Geografia e Estatística (IBGE), coletados para a Pesquisa Nacional de Saúde, existem aproximadamente 11 milhões de pessoas com deficiência auditiva no Brasil, o que corresponde a 5,13% da população (IBGE. . . , 2019c). Destas, 21,55% não conseguem ouvir de modo algum ou apresentam muita dificuldade para ouvir. Para os portadores dessa deficiência, um dos principais meio de comunicação é a Linguagem de Sinais, no caso dos surdos brasileiros, a Língua Brasileira de Sinais, Libras.

A Libras é uma língua de modalidade gestual-visual que utiliza movimentos gestuais e expressões faciais como meio de comunicação (ALMEIDA, 2013). Em 2002, a publicação da Lei no 10.436/2002 reconheceu a Língua Brasileira de Sinais como meio legal de comunicação e expressão dos surdos. O objetivo do Governo Federal com essa lei foi aumentar a representatividade de cidadãos surdos brasileiros, profissionais intérpretes de Libras e de toda a comunidade surda (BRASIL, 2002). No entanto, apesar da lei existir há mais de 20 anos, as pessoas com deficiências auditivas ainda enfrentam muitas dificuldades para ter acesso à educação, mercado de trabalho e serviços básicos do dia a dia. Apenas 5,9% das pessoas com deficiências auditivas possuem o ensino superior finalizado e 79,6% não chegam a completar o ensino médio (IBGE. . . , 2019a). Na população de surdos em idade de trabalhar (com 14 anos ou mais de idade), apenas 25,4% se encontram atuando ativamente (IBGE. . . , 2019b).

Esses números refletem o problema relacionado à inclusão social e profissional de pessoas com deficiência auditiva, evidenciando a necessidade de desenvolver tecnologias que facilitem a comunicação entre falantes e não-falantes de Libras. O desenvolvimento de dispositivos que realizem a conversão da Libras para a Língua Portuguesa é essencial para melhorar a interação dessas pessoas na sociedade brasileira.

Nessa perspectiva, diversos trabalhos têm estudado a implementação de tecnologia assistiva com o desenvolvimento de sistemas para o reconhecimento de padrões de gestos envolvendo linguagens de sinais (AZEVEDO *et al.*, 2017; DINIZ; SOUZA; NEVES, 2019; BORGES, 2020; REZENDE, 2021; OLIVEIRA *et al.*, 2023). Estes dispositivos, chamados de Sistemas de Reconhecimento de Linguagem de Sinais, do inglês *Sign Language Recognition* (SLR), têm como objetivo auxiliar a vida das pessoas surdas em suas atividades diárias (MENDES JUNIOR, 2020).

No estudo sobre técnicas de reconhecimento de gestos manuais e de língua de sinais, Das *et al.* (2021) comentam que existem duas abordagens principais para o reconhecimento de linguagens de sinais: a utilização de sensores e a abordagem baseadas em visão computacional.

Os modelos baseados em visão são considerados mais abrangentes devido à sua capacidade de capturar uma gama maior de gestos e expressões através de imagens obtidas por câmeras de vídeo, incluindo câmeras estéreo que fornecem informações de profundidade (CHEOK; OMAR; JAWARD, 2019). Sem a necessidade de utilizar dispositivos vestíveis no

usuário, os estudos mostram ser um sistema eficaz (GHOTKAR *et al.*, 2012; AL-HAMMADI *et al.*, 2020). No entanto, essa abordagem apresenta desafios significativos, como a necessidade de grande processamento computacional para análise de imagens, sensibilidade a variações de iluminação e a complexidade de lidar com gestos rápidos ou ambientes com iluminação variável, obstruções e gestos sobrepostos (CHEOK; OMAR; JAWARD, 2019).

Abordagens baseadas em sensores requerem o uso de dispositivos como sensores inerciais, sensores flexíveis, sensores de contato, ultrassônicos, entre outros instrumentos utilizados para capturar o movimento, posição e velocidade das mãos (CHEOK; OMAR; JAWARD, 2019). Exemplos dessa abordagem podem ser encontradas nas luvas desenvolvidas por Silvano *et al.* (2016), Santos, Ferreira e Alves (2019), Dias (2020) e PIGNATARI *et al.* (2023). Este método mostrou-se válido na identificação de alguns tipos de sinais em Libras utilizando a luva instrumentada e rede neural para classificação (DIAS, 2020).

Outros estudos dentro da abordagem de utilização de sensores incluem a classificação de gestos através da análise da atividade elétrica muscular, captadas com o uso da Eletromiografia de Superfície (sEMG). Essa técnica foi utilizada por MENDES JUNIOR (2020) e por Amor, Ghoul e Jemni (2023) em seus respectivos trabalhos.

Tendo como meta a criação de um dispositivo que ajude na inclusão de brasileiros com deficiência auditiva, este trabalho envolve o aprimoramento do dispositivo desenvolvido por Dias (2020), incorporando a utilização de dados sEMG, como utilizado por MENDES JUNIOR (2020), para o reconhecimento de sinais de Libras. Este dispositivo é composto por duas luvas instrumentadas com sensores flexíveis e inerciais, uma *armband* comercial Myo™ e um sistema de captura simultânea dos dados gerados por todos esses dispositivos. Como aprimoramento do trabalho mencionado, há a inserção de mais de uma luva, permitindo a coleta dos sinais desenvolvidos pelas duas mãos, e os sinais eletromiográficos adquiridos na mão dominante. Os dados coletados foram analisados e tratados de forma que os sinais estejam sincronizados. Por fim, os dados foram inseridos em modelos de aprendizagem baseados em *Deep Learning* para verificar a capacidade de reconhecimentos dos sinais e avaliar a importância de cada sensor e técnica na identificação dos sinais de Libras.

1.1 Objetivos

Nesta seção, serão apresentados os objetivos do trabalho.

1.1.1 Objetivo geral

O objetivo geral deste projeto é desenvolver um sistema composto por duas luvas instrumentadas, um bracelete (*armband*) comercial Myo™ e um sistema para a aquisição dos dados

de diferentes naturezas fornecidos por todos os três dispositivos, com o objetivo de categorizar e reconhecer padrões de gestos em Libras utilizando diferentes modelos de *Deep Learning*.

1.1.2 Objetivos específicos

- Desenvolver um sistema de aquisição de sinais composto por duas luvas instrumentadas com sensores e comunicação *Bluetooth* e uma braçadeiras de sinais eletromiográficos;
- Integrar os dispositivos (*armband* e as luvas instrumentadas) para a aquisição dos sinais por meio de um sistema em *software*;
- Adquirir sinais dos sensores flexíveis e inerciais (acelerômetro e giroscópio) das luvas instrumentadas e eletromiográficos da *armband* para a criação um banco de dados;
- Processar e sincronizar os sinais presentes no banco de dados;
- Construir modelos de aprendizagem baseados em *Deep Learning* para classificar os dados coletados;
- Analisar e comparar a contribuição dos sensores para o reconhecimento de padrões de gestos em Libras.

1.2 Justificativa

Os estudos sobre sistemas de reconhecimento de gestos em Libras frequentemente se concentram em um único tipo de dispositivo, seja ele uma luva instrumentada, como demonstrado nos trabalhos de Silvano *et al.* (2016), Santos, Ferreira e Alves (2019), Dias (2020) e PIGNATARI *et al.* (2023), uma *armband* com sinais de sEMG, como utilizada por MENDES JUNIOR (2020) e por Amor, Ghouli e Jemni (2023), ou através de vídeos, como os trabalhos de Ghotkar *et al.* (2012), Al-Hammadi *et al.* (2020) e Zholshiyeva *et al.* (2024). No entanto, com o desenvolvimento do *hardware* de uma nova luva instrumentada e de um *software* capaz de capturar os dados da luva e de uma *armband* de sEMG de forma simultânea, emerge a oportunidade de realizar a fusão de diferentes tipos de sensores e grandezas para reconhecer padrões de gestos em Libras. Essa oportunidade é evidenciada pelo estudo de Kunapipat *et al.* (2018), no qual os pesquisadores obtiveram sucesso ao combinar uma *armband* comercial Myo™ com dados de uma luva instrumentada para classificar três gestos básicos: mão aberta, mão fechada e mão em repouso.

Assim, surge a oportunidade de testar essa abordagem utilizando dois tipos de dispositivos, na qual além de comparar a eficiência de cada instrumento para a classificação, também é possível explorar a possibilidade de reconhecer gestos mais complexos. Essa capacidade de

realizar análises mais detalhadas e complexas dos gestos em Libras é crucial para o desenvolvimento de sistemas de reconhecimento mais precisos e abrangentes, contribuindo assim para a melhoria da comunicação e inclusão de indivíduos surdos na sociedade.

1.3 Estado da Arte

O reconhecimento de padrões de gestos, tanto para identificar línguas de sinais em diferentes idiomas quanto para utilizar gestos como meio de interface homem-máquina, é um campo de pesquisa que tem recebido atenção crescente nas últimas décadas. Diversos estudos têm explorado diferentes abordagens tecnológicas para facilitar a detecção e reconhecimento de gestos, dentre as analisadas, podem ser mencionadas:

Li *et al.* (2024) propuseram uma luva equipada com fibra óptica de polímero para detectar a flexão dos dedos, alcançando alta precisão na detecção de gestos para a interação homem-máquina. Por outro lado, Al-Khazzar, Altaweel e Hussain (2024) utilizaram os sinais de sEMG e *Deep neural networks* (DNN) (Rede Neural Profunda em português) de quatro camadas para alcançarem uma precisão média de 95% no reconhecimento de gestos realizados com as mãos.

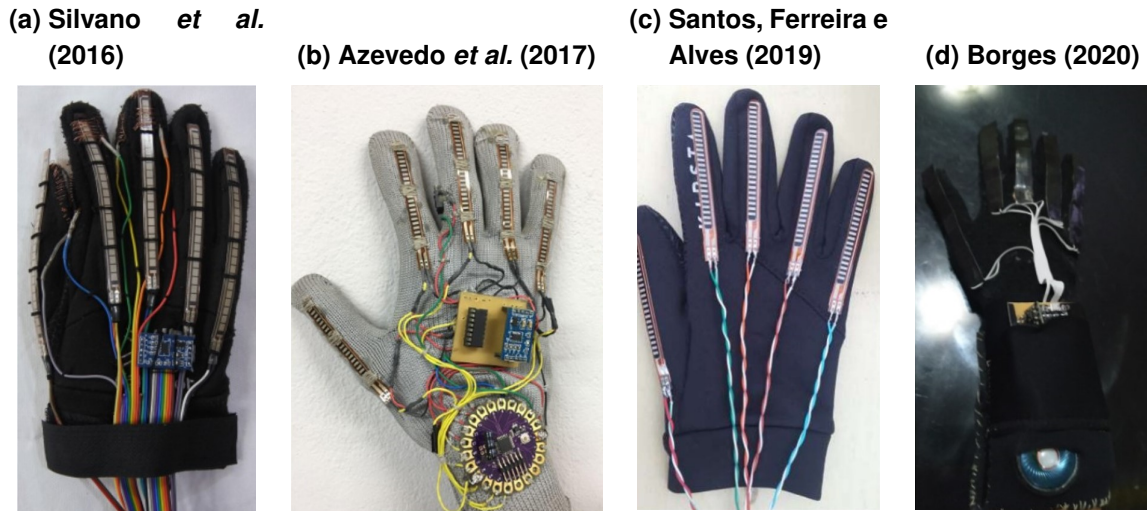
Na abordagem com utilização de luvas instrumentadas com sensores flexíveis e apresentando um sistema de reconhecimento de gestos dinâmicos em tempo real, Lee e Bae (2020) propuseram um algoritmo de detecção de gestos baseado em *Deep Learning* para identificar o início e o fim de sequências de gestos em um fluxo contínuo de dados. Os autores introduziram um algoritmo de simplificação de sequência e um algoritmo de reconhecimento de gestos para atuar com a variação de gestos.

Utilizando dois tipos de dispositivos ao mesmo tempo, Kunapipat *et al.* (2018) realizaram um estudo coletando dados de sEMG através da *armband* comercial Myo™ e dados da flexão dos dedos através de sensores de fibra óptica localizados em uma luva instrumentada. Os autores realizaram a classificação com sucesso de 3 tipos de movimentos das mãos (mão aberta, mão fechada e mão em repouso).

Voltado para identificação dos gestos da linguagem dos sinais, os estudos de Silvano *et al.* (2016), Azevedo *et al.* (2017), Santos, Ferreira e Alves (2019) e Borges (2020) com o desenvolvimento das luvas instrumentadas com sensores mostrados na Figura 1.

Outro estudo nessa mesma categoria é o desenvolvido por Dias (2020), que utiliza sensores inerciais, de contato e flexíveis para capturar os movimentos das mãos e dedos. A integração desses sensores com redes neurais possibilita a identificação precisa de gestos específicos em Libras. Dias (2020) demonstrou que uma luva equipada com esses sensores pode reconhecer gestos básicos com alta acurácia, utilizando algoritmos de aprendizagem de máquina para classificar os sinais.

Figura 1 – Luvas desenvolvidas para a identificação de gestos em Libras nos estudos de: (a) Silvano *et al.* (2016), (b) Azevedo *et al.* (2017), (c) Santos, Ferreira e Alves (2019) e (d) Borges (2020).



Fonte: Adaptado de Silvano *et al.* (2016), Azevedo *et al.* (2017), Santos, Ferreira e Alves (2019) e Borges (2020).

1.4 Estrutura do trabalho

Este trabalho está organizado na seguinte forma:

Capítulo 1 - Introdução: Apresenta uma visão geral do tema, os objetivos geral e específicos, a justificativa do estudo, o estado da arte e a estrutura do trabalho.

Capítulo 2 - Referencial Teórico: Aborda os principais conceitos e teorias que fundamentam o estudo, incluindo a Libras, biomecânica dos gestos, sistemas de reconhecimento de linguagens de sinais, tipos de sensores (flexíveis, inerciais e sEMG), microcontroladores e *Deep Learning*, com foco nos modelos gMLP e *InceptionTime*.

Capítulo 3 - Materiais e Métodos: Descreve os materiais utilizados, como a luva instrumentada, o *Myo Armband*, os métodos de comunicação *Bluetooth* e o *software* de aquisição de dados. Também detalha os métodos de coleta de dados, processamento e sincronização dos dados, e a classificação utilizando *Deep Learning*.

Capítulo 4 - Resultados: Apresenta e discute os resultados obtidos durante o estudo. Isso inclui os resultados do sistema físico, do sistema de aquisição de dados, do tratamento e sincronização dos dados, e da classificação com *Deep Learning*. Também inclui uma seção de discussões onde os resultados são analisados em profundidade.

Capítulo 5 - Considerações Finais: Fornece uma conclusão do trabalho, resumindo os principais achados, destacando as contribuições do estudo e sugerindo direções para pesquisas futuras.

2 REFERENCIAL TEÓRICO

Este capítulo descreve os principais conceitos teóricos sobre a Língua Brasileira de Sinais, os movimentos dos braços e mãos durante a execução dos gestos em Libras, o sinal eletromiográfico, a instrumentação e o processamento dos sinais.

2.1 Língua Brasileira de Sinais

A Língua Brasileira de Sinais, conhecida também pela sigla Libras, é um dos meios de comunicação e expressão oficial dos deficientes auditivos e sua comunidade no Brasil, sendo reconhecido pela lei 10.436 (BRASIL, 2002). Tal qual a Língua Portuguesa, a Libras possui suas próprias regras gramaticais, e a partir do momento que é reconhecida como língua ela deixa de ser considerada uma linguagem, perdendo o status de mímica e gesto (ALMEIDA, 2013).

As línguas de sinais diferem das línguas orais em termos de ordem e estrutura frásica. Portanto, em vez de simplesmente colocar um sinal após o outro como se faz com palavras em línguas orais, é crucial representar a informação visualmente, visto que os surdos dependem de memória visual (ALMEIDA, 2013).

Como descrito por Quadros e Karnopp (2009), a Libras ela é uma Língua que utiliza a modalidade gestual-visual, onde a informação linguística é transmitida através de gestos manuais realizados por quem está comunicando a mensagem e recebida pela visão de quem a interpreta. A estrutura da Língua Brasileira de Sinais é composta por alguns parâmetros básicos, como a configuração das mãos, o ponto de articulação ou localização, o movimento, a orientação/direcionalidade das mãos e a expressão facial e/ou corporal (MORAIS *et al.*, 2018). Além desses, outros parâmetros como datilologia (alfabeto manual, apresentado na Figura 2), verbos, frases e pronomes pessoais também são utilizados para definir um sinal (MENEZES; FEITOSA, 2015). Esses elementos, quando combinados, formam um sinal que possui significado semântico (MORAIS *et al.*, 2018).

Figura 2 – Alfabeto manual em Libras.



Fonte: Adaptado de Almeida (2013).

Durante a execução de um sinal em Libras, a alteração de qualquer um dos parâmetros mencionados pode mudar o significado do sinal. Um exemplo dessa alteração pode ser vista na Figura 3, a qual apresenta três tipos de mudança de parâmetros (configuração da mão, movimento e localização respectivamente) enquanto os outros parâmetros são mantidos constantes, resultando assim na mudança do significado dos sinais.

Como pode ser observado na Figura 3, mantendo o movimento e posição da mão mas alterando a sua configuração, pode se alterar o significado do sinal de “Pedra” para “Queijo”. O segundo exemplo representa uma mudança de significado do sinal de “Irmão” para “Igual” quando altera-se o movimento mas se mantém a posição e a configuração das mãos para esses sinais. No terceiro caso da Figura 3, tem-se dois sinais com a mesma configuração de mão e movimento, porém com uma diferença na posição: uma na direção da boca (representando o sinal de “Sábado”) e outra na direção da testa (representando “Aprender”).

Figura 3 – Exemplos da mudança de significado dos sinais em Libras ao mudar um parâmetro (configuração da mão, movimento e localização respectivamente) em sua composição.



Fonte: Adaptado de Quadros e Karnopp (2009).

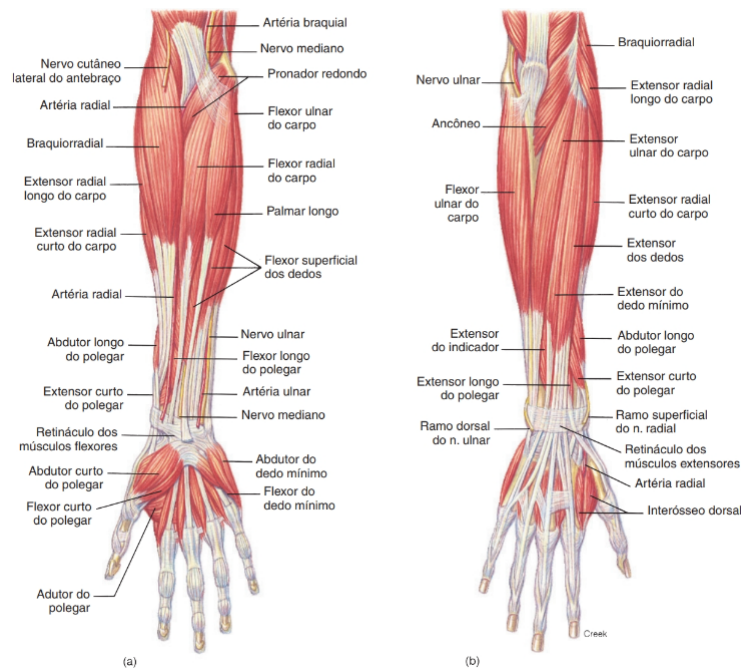
2.1.1 Características anatômicas para aquisição dos sinais de Libras

Ao realizar os sinais de uma palavra ou frase em Libras, os principais movimentos do corpo humano estão presentes nos membros superiores, como na mão, no pulso e no antebraço (QUADROS; KARNOPP, 2009).

Para que seja possível o corpo humano realizar esses movimentos, ele se utiliza de algumas estruturas corporais, entre elas o sistema esquelético, articulações e o sistema muscular (VANPUTTE; REGAN; RUSSO, 2016). Segundo Graaff (2003) o "sistema esquelético, diretamente ou indiretamente suporta e protege todos os sistemas do corpo". Além da função de auxiliar nos movimentos do corpo, também auxilia na sustentação, proteção de órgãos, no processo de formação das células sanguíneas (hematopoiese), armazenamento de gorduras e armazenamento de minerais (GRAAFF, 2003). Nos membros superiores, pode-se citar como membros do sistema esquelético os ossos nos braços (úmero), nos antebraços (ulna e rádio) e nas mãos (carpais, metacarpais e falanges).

As articulações são responsáveis por interligar dois ou mais ossos (VANPUTTE; REGAN; RUSSO, 2016), as quais possuem como principal função tornar possível a movimentação e a sustentação do corpo para promover a estabilidade (MANSOUR; FAGUNDES; ANTUNES, 2019). Para os sinais em Libras, os músculos do antebraço (Figura 4) possuem grande importância, uma vez que eles são responsáveis pelos movimentos nas articulações dos punhos, mãos e dedos, elementos fundamentais dos sinais de Libras (QUADROS; KARNOPP, 2009). Além da movimentação, os músculos atuam na produção de calor e sustentação do corpo (GRAAFF, 2003).

Figura 4 – Músculos superficiais do antebraço direito nas vistas superior (a) e posterior (b).



Fonte: Graaff (2003).

A partir destes elementos, o corpo humano é capaz de realizar diversos tipos de movimentos com diferentes Graus de liberdade (GL), ou seja, movimentos em diferentes planos espaciais, na qual cada articulação possui seus graus de liberdade (variando de 1 a 3 graus) e movimentos, como pode-se ver detalhadamente na Tabela 1. Existem 6 principais movimentos nos membros segundo Hamill e Knutzen (2008), sendo os pares flexão e extensão, abdução e adução e rotação medial e rotação lateral, ocorrendo em combinações variadas como apresentado na Figura 5.

A definição desses movimento são baseados em método universal de um sistema de planos e eixos imaginários, onde o corpo é dividido em 3 planos e 3 eixos. O plano sagital, também conhecido como anteroposterior, atravessa o corpo da frente para trás, dividindo-o em partes direita e esquerda. O plano frontal, ou coronal, corta o corpo de lado a lado, separando-o em partes anterior e posterior. O plano transversal, ou horizontal, cruza o corpo horizontalmente, dividindo-o em partes superior e inferior. Quanto aos eixos, eles correspondem a linhas imaginárias que atravessam o centro de uma articulação, em torno das quais uma parte do corpo se

movimenta. O eixo sagital (anteroposterior) atravessa a articulação de frente para trás. O eixo frontal, ou transverso, atravessa a articulação de lado a lado. Por fim, o eixo vertical, ou longitudinal, cruza a articulação de cima para baixo (MANSOUR; FAGUNDES; ANTUNES, 2019).

Conforme indicado por Hamill e Knutzen (2008), a extensão é um "movimento de retificação em que o ângulo relativo da articulação entre dois segmentos adjacentes aumenta à medida que a articulação retorna à posição zero ou posição de referência." Por outro lado, a abdução é o movimento que afasta um segmento da linha mediana do corpo. A rotação, por sua vez, refere-se ao movimento de um segmento em relação a um eixo vertical que o atravessa, de modo que a superfície anterior do segmento se move em direção à linha mediana do corpo, enquanto a superfície posterior se afasta dessa linha (HAMILL; KNUTZEN, 2008).

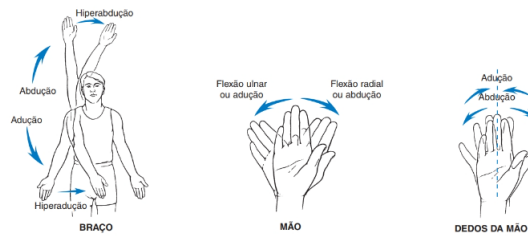
Tabela 1 – Grau de liberdade e movimentos realizados em cada articulação dos membros superiores.

Segmento	Articulação	GL	Movimentos
Braço	Ombro	3	Flexão, extensão, hiperextensão, abdução, adução, hiperabdução, hiperadução, abdução horizontal, adução horizontal, rotação med./lat., circundução.
Braço/ombro	Esternoclavicular	3	Elevação, depressão, abdução, adução (protração, retração), rotação.
Antebraço	Cotovelo	1	Flexão, extensão, hiperextensão.
	Radioulnar	1	Pronação, supinação.
Mão	Radiocarpal	2	Flexão, extensão, hiperextensão, flexão radial, flexão ulnar, circundução.
Dedos da mão	Metacarpofalângica	2	Flexão, extensão, hiperextensão, abdução, adução, circundução.
	Interfalângica	1	Flexão, extensão, hiperextensão.
Polegar	Carpometacarpal	2	Flexão, extensão, abdução, adução, oposição, circundução.
	Metacarpofalângica	1	Flexão, extensão.
	Interfalângica	1	-

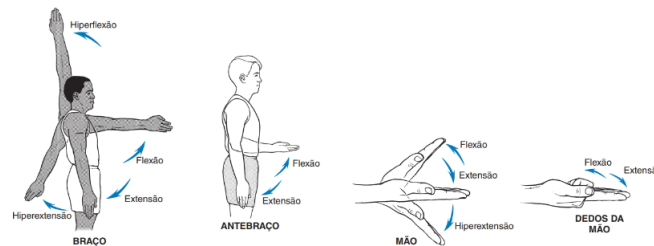
Fonte: Adaptado de Hamill e Knutzen (2008).

Figura 5 – Tipos de movimentos nos membros superiores: (a) Abdução, (b) Flexão, (c) Rotação.

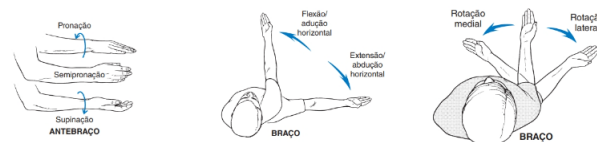
(a) Abdução



(b) Flexão



(c) Rotação



Fonte: Adaptado de Hamill e Knutzen (2008).

2.2 Sistemas de reconhecimento de linguagens de sinais

Os sistemas de reconhecimento de linguagens de sinais (SLR) têm o objetivo de facilitar a comunicação entre pessoas com deficiência auditiva, que utilizam a linguagem dos sinais, e pessoas falantes (MENDES JUNIOR, 2020). Para que esse objetivo seja alcançado, os sistemas SLR buscam desenvolver algoritmos e métodos para identificar corretamente uma sequência de sinais produzidos e compreender seu significado (COOPER; HOLT; BOWDEN, 2011). Nos últimos anos, uma quantidade significativa de pesquisa tem sido dedicada ao desenvolvimento de sistemas de reconhecimento de sinais para diversas Linguagem de Sinais (LS). Estas pesquisas têm concluído que a criação de tais sistemas representa um desafio multidisciplinar, envolvendo áreas como métodos de captura de gestos, classificadores de aprendizado de máquina, do inglês, *Machine Learning* (ML), compreensão de ações humanas e processamento de linguagem natural (AL-AHDAL; NOORITAWATI, 2012). Diversas técnicas foram desenvolvidas

para identificar a LS, sendo estas classificadas em duas abordagens principais: reconhecimento baseado em *hardware* e reconhecimento baseado em visão (CHEOK; OMAR; JAWARD, 2019). No contexto do reconhecimento baseado em *hardware*, os dados são coletados por meio de dispositivos físicos como luvas, pulseiras e sensores. Nessa modalidade, a posição da mão, sua velocidade e movimentação são registradas utilizando-se luvas equipadas com sensores específicos (DAS *et al.*, 2021). Já nas abordagens baseadas em visão, o rastreamento das mãos é realizado utilizando câmeras e processamento utilizando algoritmos (GHOTKAR *et al.*, 2012).

2.2.1 Conceitos de sensoriamento e instrumentação

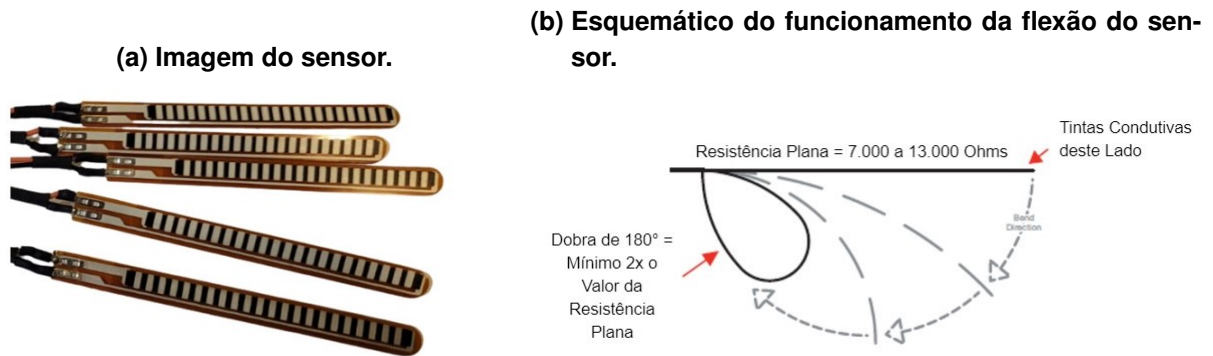
Os sensores são dispositivos utilizados para medir a variação de uma grandeza física (STEVAN; SILVA, 2015). Eles servem para informar um circuito eletrônico sobre um evento externo, ao qual o circuito deve responder ou a partir do qual deve iniciar uma determinada ação (WENDLING, 2010). Esses dispositivos são sensíveis a diversas formas de energia, como luminosa, térmica ou cinética, e seu objetivo é fornecer informações precisas sobre grandezas que precisam ser medidas, como temperatura, pressão, velocidade, corrente, aceleração, posição, entre outras (THOMAZINI; ALBUQUERQUE, 2020). Para capturar os dados das principais variações de medidas que ocorrem durante o movimentos dos gestos em Libras, como o movimento, posição e velocidade das mãos, são necessários o uso de sensores inerciais e sensores flexíveis, ou através da atividade elétrica muscular utilizando sEMG (CHEOK; OMAR; JAWARD, 2019; AMOR; GHOUL; JEMNI, 2023).

2.2.1.1 Sensores resistivos do tipo flexíveis

Os sensores flexíveis, visto na Figura 6 (a), podem ser encontrados em muitas aplicações, incluindo dobradiças em máquinas, sensores robóticos, sensores de portas e até mesmo úteis para medir efeitos torcionais ou como uma maneira de medir movimentos dos dedos (SPECTRA SYMBOL, 2014; BELL, 2014). Para realizar as medições, os sensores de flexão variam a sua resistência elétrica proporcionalmente a deformação mecânica a qual é submetido, essa deformação pode ser tanto de flexão quanto extensão, como pode ser visto na Figura 6 (b) (SPECTRA SYMBOL, 2014).

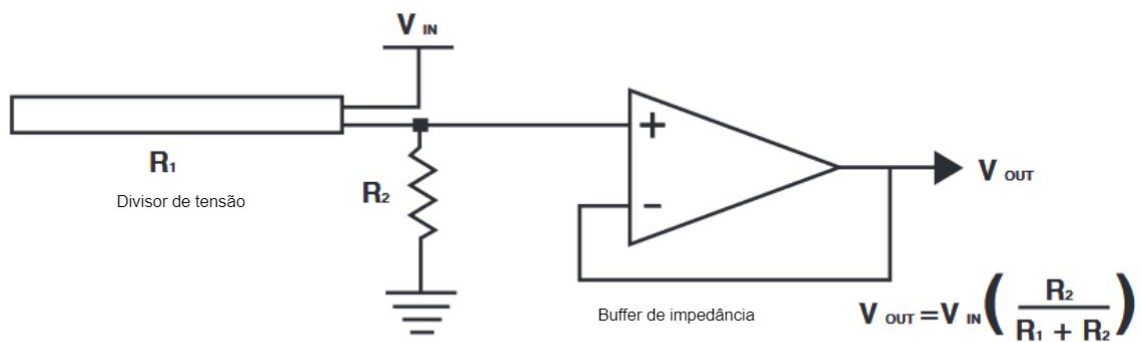
Como é possível observar no esquemático da Figura 7, para medir a variação de resistência do sensor ao ser submetido a uma deformação, utiliza-se um circuito básico divisor de tensão para cada sensor. Aplicando este sensor na luva instrumentada juntamente ao seu circuito de condicionamento citado, com a referência da mão em repouso e a medição da variação de resistência durante o movimento, é possível definir a intensidade de flexão ou extensão a que cada dedo foi submetido para configurar o sinal em Libras (SILVANO *et al.*, 2016).

Figura 6 – Sensor flexível 2,2": (a) Imagem do sensor, (b) Esquemático do funcionamento da flexão do sensor.



Fonte: Adaptado de SPECTRA SYMBOL (2014).

Figura 7 – Circuito divisão de tensão utilizado junto ao sensor flexível 2,2".



Fonte: Adaptado de SPECTRA SYMBOL (2014).

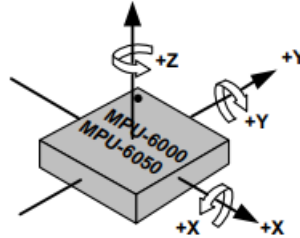
2.2.1.2 Inerciais

Sensores inerciais são projetados para capturar o movimento do próprio sensor ou do objeto ao qual estão conectados, detectando variações de velocidade, inclinação, vibração e outros parâmetros em vários eixos (BELL, 2014). Esses dispositivos consistem em um circuito integrado que incorpora um sistema de micromecânica baseado em células capacitivas microssinadas, junto com um condicionador de sinais, tudo encapsulado em um único chip (THOMAZINI; ALBUQUERQUE, 2020).

O sensor inercial MPU-6050, representado pela Figura 8, é capaz de medir variações de movimento em seis eixos, sendo três eixos referentes a um giroscópio e três eixos referentes a um acelerômetro (INVENSENSE'S, 2013). Isso permite a medição da taxa de rotação ou velocidade angular através do giroscópio e da aceleração linear através do acelerômetro. Para garantir o rastreamento preciso de movimentos rápidos e lentos, o sensor oferece uma faixa completa de giroscópio programável pelo usuário, com opções de ± 250 , ± 500 , ± 1000 e

± 2000 %seg (*degrees per second* (dps)) e um acelerômetro com faixa de escala programável de Aceleração da gravidade (g), com opções de $\pm 2g$, $\pm 4g$, $\pm 8g$ e $\pm 16g$. Além disso, o sensor possui um sensor de temperatura integrado e um oscilador no chip, com variação de $\pm 1\%$ na faixa de temperatura operacional (INVENSENSE'S, 2013).

Figura 8 – MPU-6050.



Fonte: Adaptado de InvenSense's (2013).

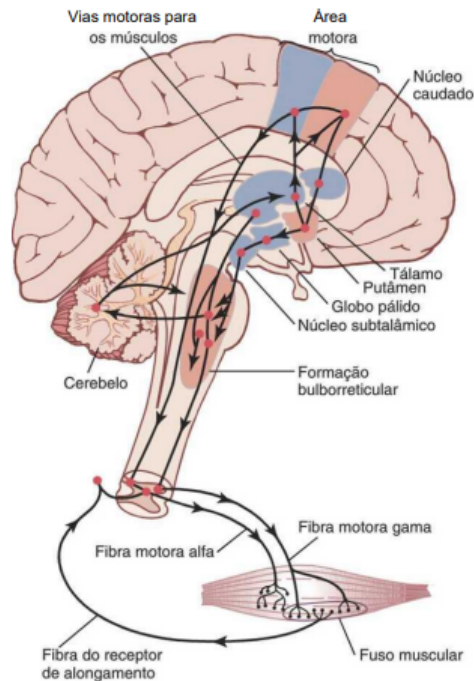
2.2.1.3 Eletromiografia

Quando uma pessoa executa um movimento, como os gestos associados à língua de sinais de Libras, ela emprega os sistemas esquelético, articular e muscular para comunicar-se (VANPUTTE; REGAN; RUSSO, 2016). A seguir, será explorado mais detalhadamente como esse movimento é realizado, desde o pensamento inicial da pessoa que deseja transmitir a mensagem até a execução dos movimentos necessários para a compreensão da mesma.

O Sistema Nervoso desempenha um papel crucial ao conectar o pensamento à execução e controle do movimento, por meio das funções motoras. Na Figura 9 é ilustrado a representação do neuroeixo motor “esquelético” do sistema nervoso, responsável pelo controle das contrações musculares esqueléticas (GUYTON *et al.*, 1971). Para realizar a contração muscular, tem-se o disparo do neurônio motor, seguida pela liberação do neurotransmissor acetilcolina na placa motora, resultando em um aumento da condutância ao Na^+ e K^+ na membrana da placa motora, gerando assim um potencial de ação nas fibras musculares. As fibras musculares são compostas por miofibrilas que contêm os filamentos de actina (finos) e miosina (grossos), esse potencial gera uma propagação da despolarização ao longo dos túbulos (BARRETT, 2013). Em seguida, há uma liberação de Ca^{2+} e difusão até os filamentos grossos e finos, culminando na formação de pontes cruzadas entre a actina e a miosina, o que resulta no deslizamento dos filamentos finos sobre os grossos, produzindo assim o movimento (BARRETT, 2013).

A junção neuromuscular é um ponto crítico onde ocorre a liberação da acetilcolina pelo neurônio motor somático, o que transforma um sinal químico em um sinal elétrico na fibra muscular (SILVERTHORN, 2017). Essa transformação é fundamental para a aquisição dos sinais elétricos que indicam a atividade muscular. A Eletromiografia (Eletromiografia (EMG)) é a técnica utilizada para captar esses sinais elétricos, medindo a diferença de potencial entre dois pontos ao longo de uma fibra muscular. A Figura 10 apresentada por Silverthorn (2017), traz um

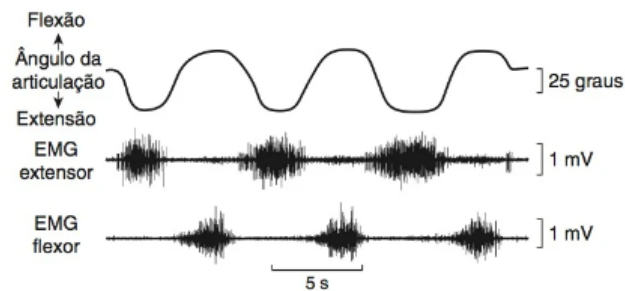
Figura 9 – Eixo neural motor esquelético do sistema nervoso.



Fonte: (GUYTON *et al.*, 1971).

exemplo de sinal de EMG durante flexão e extensão alternadas da articulação distal do polegar humano, registrando os músculos extensor e flexor longo do polegar.

Figura 10 – Registros eletromiográficos dos músculos extensor e flexor longo do polegar durante flexão e extensão alternados da articulação distal do polegar humano.



Fonte: Silverthorn (2017).

A Eletromiografia de superfície (sEMG) é a técnica não invasiva utilizada para medir através da pele a atividade elétrica dos músculos, utilizando eletrodos que captam os sinais musculares pela diferença de potencial entre eles (GARCIA; VIEIRA, 2011). Esses sinais transmitem informações sobre a ativação muscular, como, por exemplo, a intensidade da contração muscular, a manifestação mioelétrica da fadiga muscular e o recrutamento de unidades motoras (GARCIA; VIEIRA, 2011). Para garantir a precisão na detecção dos sinais sEMG, os eletrodos devem ser posicionados na direção das fibras musculares, pois o desalinhamento pode impactar significativamente as estimativas das variáveis EMG. Há uma necessidade de padronização da metodologia quanto à localização específica dos eletrodos, já que a variação na posição de-

les pode causar superestimação ou subestimação da amplitude, frequência média, frequência mediana e velocidade de condução das fibras musculares (MERLETTI; RAINOLDI; FARINA, 2001).

Os eletrodos recomendados para sEMG possuem superfícies de detecção compostas por duas barras paralelas, cada uma com 1,0 cm de comprimento e 1-2 mm de largura, separadas por uma distância de 1,0 cm, e operam com uma largura de banda de 20-500 Hz (LUCA, 1997). Em geral, a maior parte da energia dos sinais sEMG encontra-se na faixa de 10 a 200 Hz (FRIDLUND; CACIOPPO, 1986).

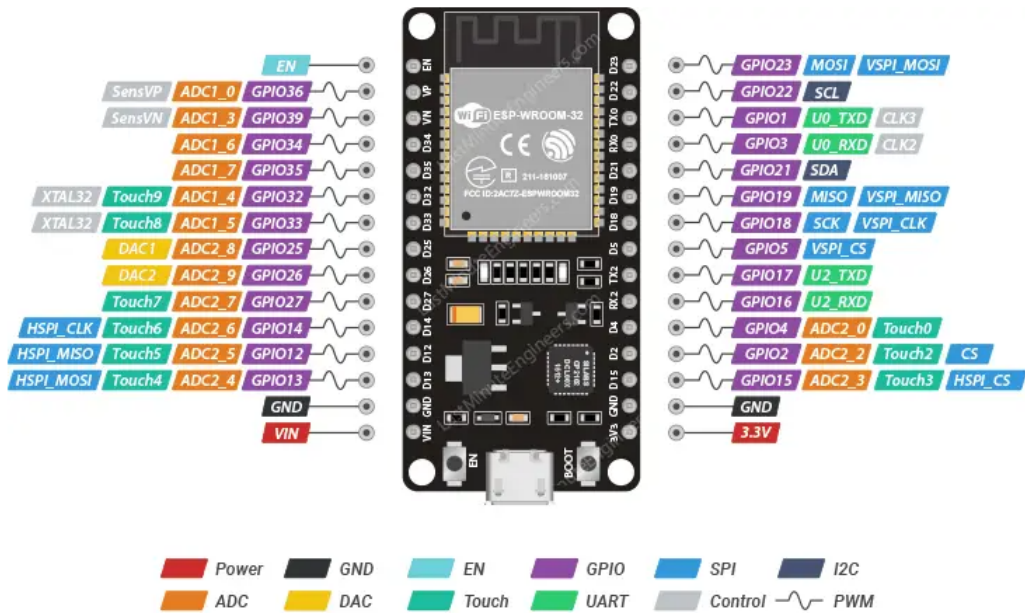
2.3 ESP32

Para realizar o desenvolvimento do projeto em um sistema embarcado, integrando as entradas de todos os sensores listados anteriormente e transmitindo os dados para uma central, é necessário utilizar um microcontrolador. O microcontrolador ESP32, desenvolvido pela Espressif Systems, destaca-se como uma escolha ideal devido à sua conectividade integrada via *Wi-Fi* e *Bluetooth*, que o torna altamente adequado para projetos de *Internet of Things* (IoT) (MAIER; SHARP; VAGAPOV, 2017). Além disso, o modelo utilizado ESP32 DevKit V1 conta com 30 portas de entrada/saída, *General Purpose Input/Output* (GPIO), como ilustrado na Figura 11, permitindo a conexão com uma ampla gama de sensores. Outra característica importante deste dispositivo é sua unidade de processamento dupla (*Dual Core*), permitindo a execução de tarefas em paralelo, uma em cada núcleo, sem perda de desempenho em comparação com sistemas que possuem apenas uma unidade de processamento (SYSTEMS, 2024).

O desenvolvimento na plataforma ESP32 proporciona diversas possibilidades de utilização. O dispositivo ESP32 pode ser programado a partir de qualquer sistema operacional, seja Windows, Linux ou MacOS. É possível utilizar a extensão (*add-on*) ESP32 para Arduino, a plataforma nativa Espressif IoT Development Framework, bem como o ambiente Python através do interpretador Micropython (BABIUCH; FOLTÝNEK; SMUTNÝ, 2019).

O microcontrolador é compatível com as especificações Bluetooth *Low Energy* (BLE), suportando múltiplas conexões, e é capaz de operar a velocidades de até 4 Mbps (SYSTEMS, 2024). Seus GPIOs incluem dois conversores analógico-digitais (ADCs) de 12 bits, oferecendo um total de 18 canais, com isso é possível conectar até 18 sensores a um único módulo ESP32 (ŠKRABA *et al.*, 2019). Esses ADCs podem ser configurados com resoluções de 9, 10 ou 12 bits e atenuações de 0dB, 6dB ou 11dB, permitindo ajustes para diferentes faixas de entrada (MAIER; SHARP; VAGAPOV, 2017).

Figura 11 – ESP32.



Fonte: Adaptado de (ESP32..., 2022).

2.4 Deep Learning

Em Ciência da Computação, a Inteligência Artificial (IA) é o campo que estuda “agentes inteligentes”, dispositivos que podem perceber o ambiente e tomar decisões para maximizar suas chances de alcançar seus objetivos (SHINDE; SHAH, 2018). Segundo Shinde e Shah (2018) o aprendizado é um aspecto essencial dessas máquinas, tornando o aprendizado de máquina (ML) um subcampo importante da IA.

Como definido por Goodfellow, Bengio e Courville (2016), o aprendizado de máquina é, essencialmente, uma forma de estatística aplicada, que coloca maior ênfase no uso de computadores para estimar estatisticamente funções complexas, com menor foco na prova de intervalos de confiança em torno dessas funções. Além disso, a maioria dos algoritmos de aprendizado de máquina pode ser classificada em duas categorias principais: aprendizado supervisionado e aprendizado não supervisionado, e entre os objetivo mais comuns para o uso delas pode-se citar a Classificação, Regressão, Transcrição, Tradução, Detecção de anomalias, entre outras (GOODFELLOW; BENGIO; COURVILLE, 2016).

Uma maneira confiável de validar o desempenho de modelos de ML é usando a técnica de divisão de treinamento/teste, onde parte dos dados é reservada para validação após o desenvolvimento do modelo, isso permite estimar imparcialmente como o modelo se comportaria em previsões reais (VABALAS *et al.*, 2019). Quando os dados são limitados, uma solução comum é a validação cruzada. Ao invés de treinar um único modelo fixo, a validação cruzada desenvolve vários modelos em diferentes partes dos dados de forma iterativa (BROWNE, 2000). Um método popular é o K-Fold, onde o conjunto de dados é dividido em várias amostras. Cada

modelo é treinado em uma amostra e validado nas outras, repetindo o processo até que todas as partes dos dados tenham sido usadas para validação. O desempenho do modelo é então avaliado calculando a média dos desempenhos de classificação de todas as amostras de validação (SICSÚ; SAMARTINI; BARTH, 2023).

Entre outras técnicas do ML, tem-se o *Deep Learning*, ou aprendizado profundo, que utiliza redes neurais artificiais com muitas camadas (DNN) para modelar e entender padrões complexos em grandes volumes de dados (ALZUBAIDI *et al.*, 2021). Além disso, elas geralmente contêm neurônios avançados em contraste com as redes neurais artificiais simples, ou seja, elas podem usar operações avançadas (por exemplo, convoluções) ou múltiplas ativações em um único neurônio, em vez de usar uma função de ativação simples (JANIESCH; ZSCHECH; HEINRICH, 2021). Este trabalho explora dois modelos de *deep learning* específicos: o Gated Multilayer Perceptron (gMLP) e o *InceptionTime*, ambos escolhidos por suas capacidades distintas e eficazes de lidar com dados sequenciais e séries temporais.

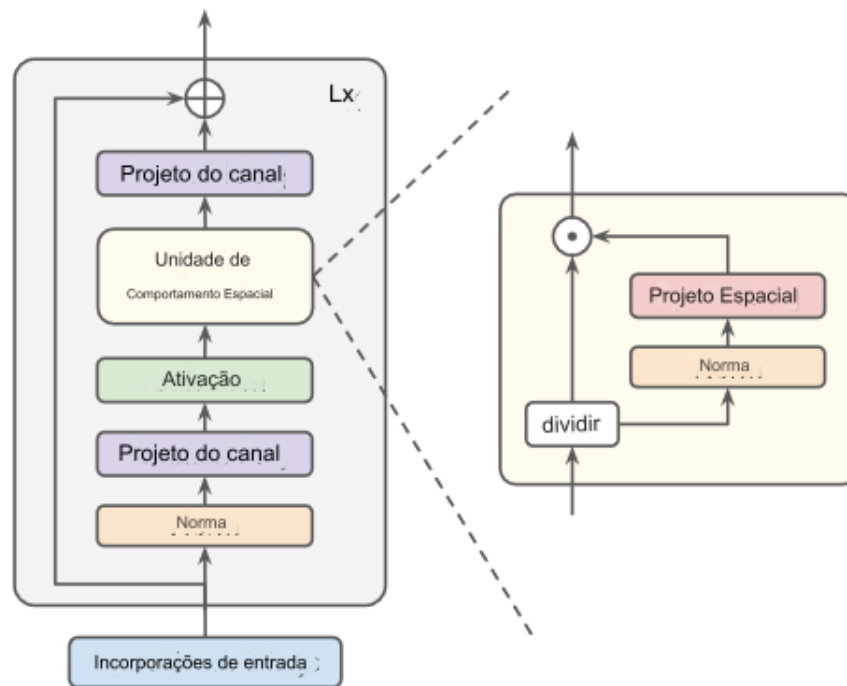
2.4.1 gMLP

O gMLP (*Gated Multilayer Perceptron*) é uma arquitetura de rede neural que se baseia em *perceptrons* multicamadas, onde cada camada está totalmente conectada à próxima, utilizando mecanismos de *gating* para capturar informações em dados sequenciais, sendo uma alternativa aos modelos baseados em *Transformers*. O gMLP utiliza projeções de canais e projeções espaciais com parametrização estática, sem depender de mecanismos de autoatenção, essas projeções espaciais funcionam de forma linear conforme exibido no esquemático da arquitetura presente na Figura 12 (LIU *et al.*, 2021).

O funcionamento do gMLP envolve o uso de camadas *Multilayer Perceptron* (MLP) com mecanismos de regulação (*gating*), que atuam como filtros controlando o fluxo de informações ao longo da rede. Esses mecanismos ajustam dinamicamente quais informações são passadas adiante, ajudando a regular o fluxo de informações e a capturar dependências de longo alcance em dados sequenciais (LIU *et al.*, 2021).

Embora os MLPs tradicionais não sejam ideais para capturar dependências temporais ou sequenciais, o gMLP, com seus mecanismos de *gate*, é mais adequado para tarefas que envolvem dados sequenciais ou temporais, como séries temporais e processamento de linguagem natural. O gMLP é treinado usando técnicas padrão de aprendizado profundo e pode ser implementado usando estruturas populares de *Deep Learning* como PyTorch e TensorFlow (LIU *et al.*, 2021).

Figura 12 – Rede gMLP.



Fonte: Adaptado de Liu *et al.* (2021).

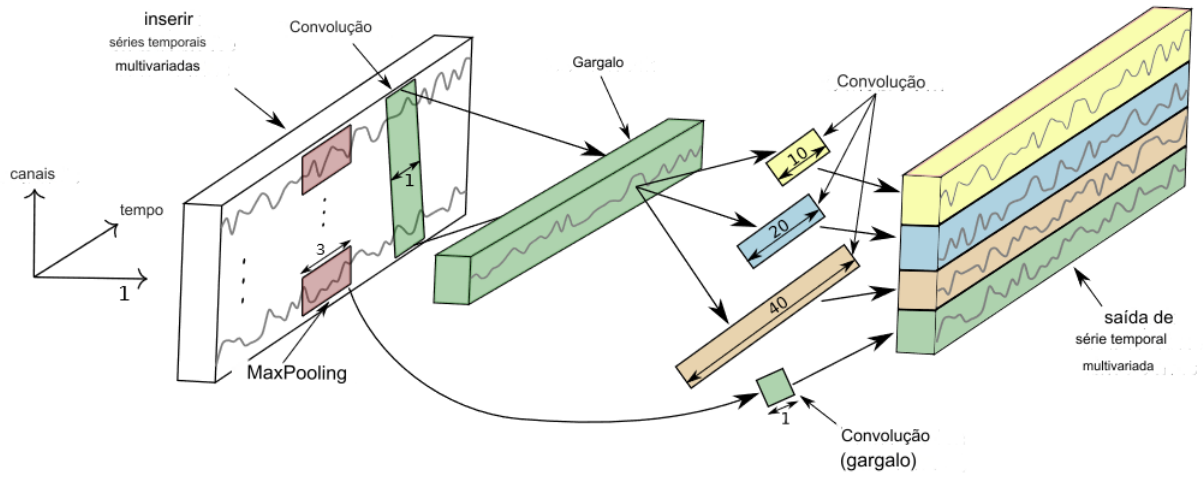
2.4.2 InceptionTime

O *InceptionTime* é um modelo de rede neural convolucional desenvolvido para a classificação de séries temporais. Ele é uma junção de 5 redes *Inception*, onde cada uma aplica filtros de diferentes comprimentos simultaneamente às séries temporais de entrada. Essa abordagem permite a extração automática de características relevantes de séries temporais longas e curtas. O modelo *InceptionTime* é projetado para superar algoritmos tradicionais de classificação de séries temporais, alcançando um desempenho superior ao explorar filtros maiores do que os utilizados em tarefas de visão computacional. Além disso, o *InceptionTime* é altamente escalável e eficaz para lidar com conjuntos de dados do mundo real, oferecendo uma abordagem inovadora e precisa para a classificação de séries temporais (FAWAZ *et al.*, 2020).

A arquitetura *InceptionTime*, exibida na Figura 13 é caracterizada pelo uso de módulos *Inception*, cada um composto por duas etapas principais: um "gargalo" que reduz a dimensionalidade inicial dos dados e convoluções paralelas de diferentes comprimentos para capturar características em várias escalas temporais. Além das convoluções, o modelo incorpora operações de *MaxPooling* para assegurar invariância a pequenas perturbações nos dados. Após a aplicação dos módulos *Inception*, é utilizada uma camada Agrupamento da Média Global para a agregação de características, seguida por uma função exponencial normalizada para a classificação final. A arquitetura foi refinada com base na eficiência computacional e na habilidade de

lidar com dependências temporais complexas, tornando-se especialmente útil para aplicações como a classificação de séries temporais multivariadas (FAWAZ *et al.*, 2020).

Figura 13 – Rede InceptionTime.



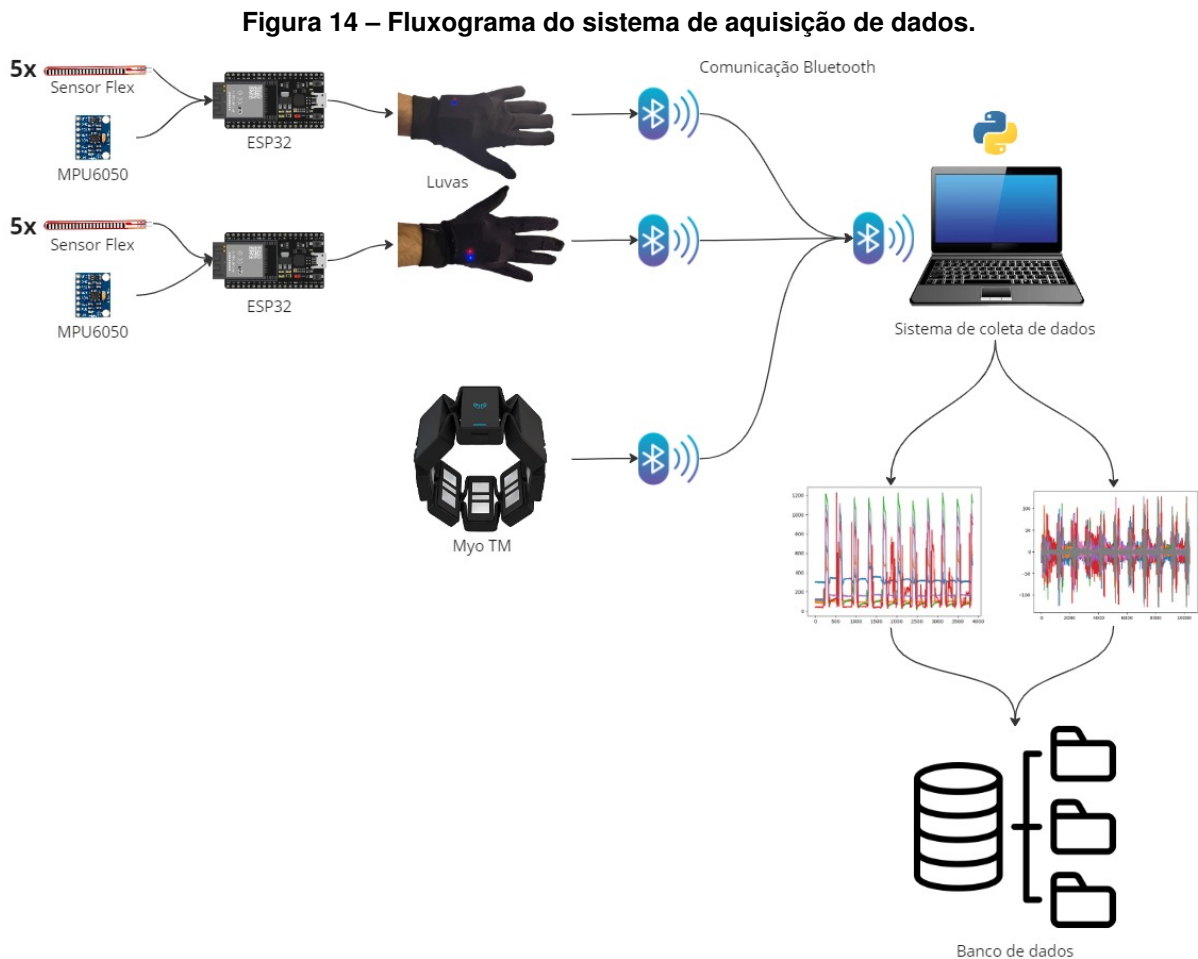
Fonte: Adaptado de Fawaz *et al.* (2020).

3 MATERIAIS E MÉTODOS

Após apresentar o objetivo deste trabalho e detalhar sua fundamentação teórica, pôde-se realizar o desenvolvimento das duas luvas instrumentadas com comunicação sem fio, além do sistema de aquisição de dados das luvas em paralelo aos dados da *armband* Myo™. Este capítulo apresenta os materiais e métodos utilizados para esse desenvolvimento.

3.1 Materiais

A Figura 14 apresenta o fluxograma do sistema de aquisição de dados desenvolvido neste trabalho. Nas próximas seções, serão detalhados cada um dos itens presentes na sua construção e comunicação entre os dispositivos.

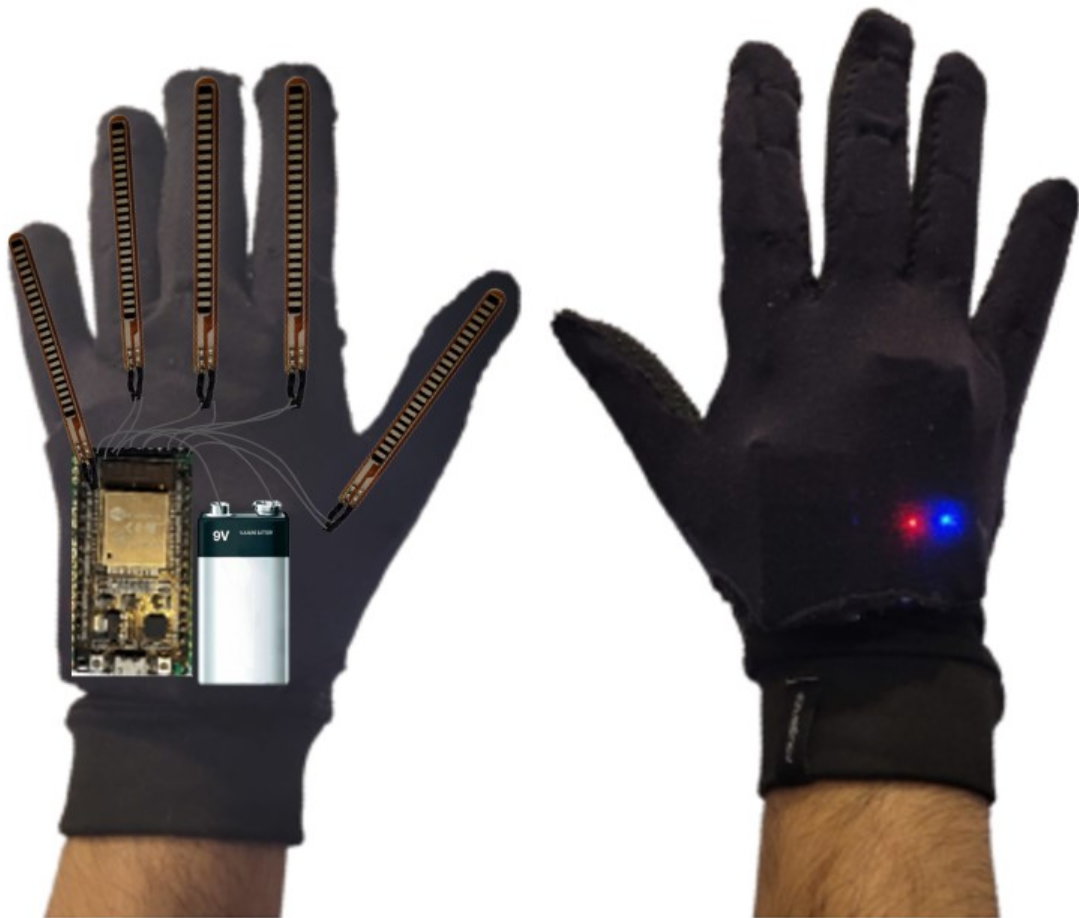


Fonte: Autoria própria (2024).

3.1.1 Luva instrumentada

Neste trabalho, desenvolveu-se duas luvas instrumentadas, uma para a mão esquerda e outra para a mão direita, utilizando como referência o trabalho de Dias (2020). Utilizou-se luvas com boa destreza, fabricadas com materiais de alta elasticidade. Foram costurados sensores flexíveis sobre a superfície de cada dedo de modo que envolvessem as principais articulações. Para os sensores flexíveis, foi utilizado um circuito divisor de tensão, conforme sugerido pelo fabricante (SPECTRA SYMBOL, 2014). Além disso, cada luva foi equipada com um sensor inercial MPU-6050 e conectada a um microcontrolador ESP32. Estes elementos foram unificados em uma placa para reduzir espaço, sendo esta placa fixada no topo da luva. Para a alimentação do sistema, utilizaram-se baterias de 9 V, fixadas ao lado da placa principal. Por fim, costurou-se uma nova camada de tecido para o acabamento da luva, servindo de suporte e proteção para todas as conexões do sistema. Na Figura 15, é exibida a representação gráfica da configuração interna do sistema na luva esquerda e, na luva direita, como ficou com a camada de tecido para o acabamento do vestível.

Figura 15 – Representação gráfica da luva instrumentada desenvolvida.



Fonte: Autoria própria (2024).

3.1.2 *Armband Myo™*

Desenvolvido e fabricado pela Thalmic Labs, a *armband Myo™*, exibida na Figura 16, foi lançada em 2015 com o objetivo de ser um dispositivo de fácil aquisição de sinais eletromiográficos para o desenvolvimento de interfaces homem-máquina. Utilizando oito sensores de eletromiografia de superfície (sEMG), a *armband* é capaz de transmitir os dados dos sinais elétricos capturados nos músculos do antebraço através de comunicação Bluetooth a uma taxa de 200 amostras por segundo (VISCONTI *et al.*, 2018).

Para isso, a Myo™ conta com um processador ARM Cortex M4 (120Mhz MK22FN1M), um chip *Bluetooth Low Energy* (BLE) NRF51822 para a transmissão via Bluetooth 4.0, duas baterias recarregáveis de lítio (3.7 V - 260mAh), porta *Universal Serial Bus* para o carregamento, sensor de movimento MPU-9150 de 9 eixos, motor de vibração para *feedback* ao usuário, além dos oito eletrodos do tipo seco feitos em aço inoxidável e oito amplificadores operacionais ST 78589, um para cada eletrodo (RAWAT; VATS; KUMAR, 2016; VISCONTI *et al.*, 2018).

A *armband* foi projetada para caber em antebraços com uma circunferência na faixa de 19 a 34 cm, devido ao material flexível de sua estrutura e demais componentes que a facilitam o bom posicionamento dos eletrodos (ThalmicLabs, 2018). Apesar de sua fabricação ter sido descontinuada, estudos recentes continuam utilizando esse dispositivo devido a sua Interface de Programação de Aplicações (API) pública onde é possível capturar os dados brutos fornecidos pelos sensores presentes na *armband* com uma taxa de amostragem de 200 Hz e resolução de 8 bits.(KUNAPIPAT *et al.*, 2018; AMOR; GHOU; JEMNI, 2023; CARNEIRO, 2023; AL-KHAZZAR; ALTAWHEEL; HUSSAIN, 2024).

Figura 16 – *armband Myo™*.



Fonte: Adaptado de Thalmic Labs Inc.

3.1.3 Sistema de aquisição e tratamento de dados

Para realizar a comunicação do *hardware* de cada Luva instrumentada, desenvolveu-se um código utilizando a linguagem de programação “C”. O código desenvolvido para executar no microcontrolador ESP32 é capaz de realizar a leitura dos sensores flexíveis e dos sensores

inerciais, empacotar os dados em um *buffer* e os transmitir via serial utilizando comunicação Bluetooth.

Por sua vez, também foi realizado um programa para adquirir os dados captados pela *armband Myo™*, utilizando a linguagem de programação Python, executando em um dispositivo com conexão Bluetooth. Conforme apresentado na Figura 14, foi desenvolvido um programa que acessa as portas seriais Bluetooth para receber os dados das luvas instrumentadas e se conectar às funções da API da Myo™. Após receber os dados, o sistema os converte em bases compiladas no formato “CSV”. As etapas de funcionamento do programa de aquisição de dados são apresentadas na Figura 17. Para iniciar a captura de uma amostra dos sinais de Libras, foi desenvolvido um menu inicial, onde é possível escolher qual dispositivo deve ser utilizado para cada coleta de dados (luva esquerda, luva direita, *armband Myo™* ou o conjunto desses dispositivos). Também é possível determinar o tempo que o sistema vai receber os dados e definir o nome e descrição da coleta.

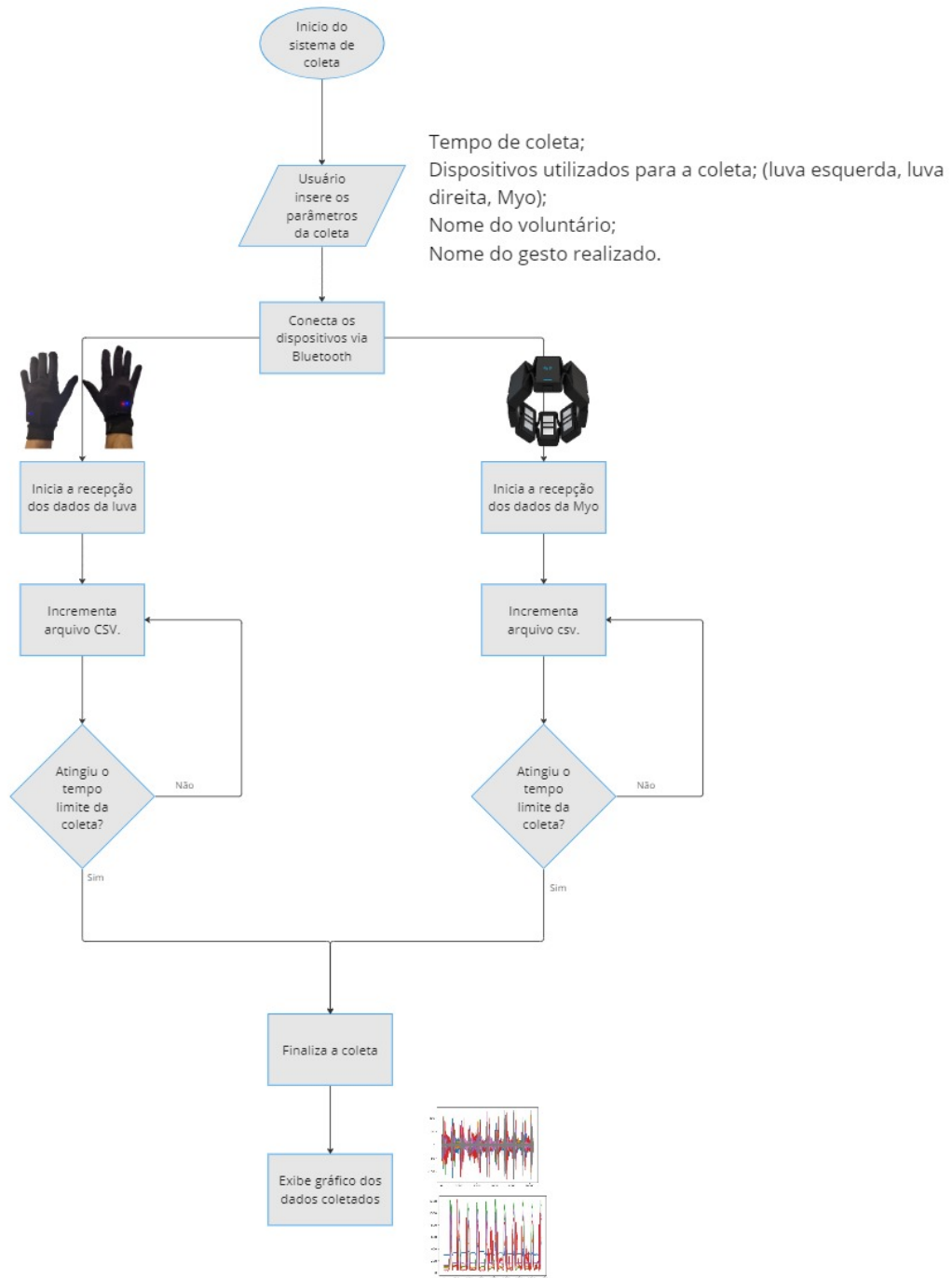
Ao coletar os dados com os três dispositivos (duas luvas e o bracelete), funções de *threads* em Python são utilizadas para permitir o paralelismo das funções que recebem os dados. Dessa forma, um processo não impacta o outro, permitindo capturar as informações disponibilizadas pelos sensores de cada dispositivo. Após finalizar a coleta da amostra, o sistema gera gráficos para uma pré-visualização dos dados adquiridos de cada sensor, facilitando a identificação por parte do usuário se todos os dados foram coletados corretamente.

Com os dados coletados salvos em arquivos no formato “CSV”, o processamento dos sinais foi realizado no ambiente do Google Colab PRO, um serviço na nuvem que executa o Jupyter Notebook com acesso a GPUs com 16GB e 51GB de RAM. Por meio da linguagem Python e as bibliotecas “numpy”, “pandas”, “matplotlib” e “scipy”, os arquivos foram transformados em *dataframes* para facilitar o processamento. Os gestos foram segmentados em arquivos únicos. Com os arquivos processados e rotulados, diferentes modelos de *Deep Learning* do pacote *State-of-the-art Deep Learning library for Time Series and Sequences*. (tsai) foram utilizados para a classificação dos sinais em Libras. Os modelos utilizados foram gMLP - *Gated Multilayer Perceptron* (LIU *et al.*, 2021) e *InceptionTime* (FAWAZ *et al.*, 2020).

3.2 Métodos

Nesta seção, são apresentados os métodos utilizados para a coleta de dados, processamento dos mesmos e por fim para a classificação dos sinais em Libras.

Figura 17 – Fluxograma do funcionamento do sistema de aquisição de dados.



Fonte: Autoria própria (2024).

3.2.1 Coleta de dados

Os métodos de coleta de dados seguiram as diretrizes do Comitê de Ética em Pesquisa envolvendo Seres Humanos da Universidade Tecnológica Federal do Paraná, tendo sido aprovados mediante parecer nº. 5701386.

Para criar o banco de dados a ser analisado, foram selecionados 7 voluntários, sendo 2 mulheres e 5 homens, para a coleta de amostras de dados da realização dos sinais em

Libras, utilizando as duas luvas e a *armband Myo™* para todos os sinais. Foram selecionadas 50 palavras para que os 7 usuários realizassem os sinais de Libras, repetindo 8 vezes cada um deles. Os 50 sinais foram escolhidos com base no estudo de Faisal *et al.* (2022) e representam diferentes formas de movimentos e posições das mãos. Além disso, estes sinais permitem a criação de sentenças usando a língua de sinais.

Quadro 1 – Palavras utilizadas para a realização dos sinais em Libras.

#	Palavra	#	Palavra	#	Palavra	#	Palavra	#	Palavra
1	Agarrar	11	Como	21	Ir	31	Parar	41	Remédio
2	Agora	12	Contar	22	Legal	32	Partir	42	Saber
3	Amar	13	Desculpa	23	Leite	33	Pegar	43	Sentar
4	Apagar	14	Dever	24	Limpar	34	Perder	44	Sentir
5	Assustar	15	Doce	25	Livro	35	Poder	45	Tchau
6	Atrasar	16	Doente	26	Médico	36	Por Favor	46	Tempo
7	Avião	17	Escola	27	Melhor	37	Porquê	47	Trabalhar
8	Boliche	18	Eu	28	Molhar	38	Quê?	48	Você
9	Bom	19	Gelo	29	Noite	39	Querer	49	Você Mesmo
10	Cachorro	20	Gostar	30	Ouvir	40	Rápido	50	Zero

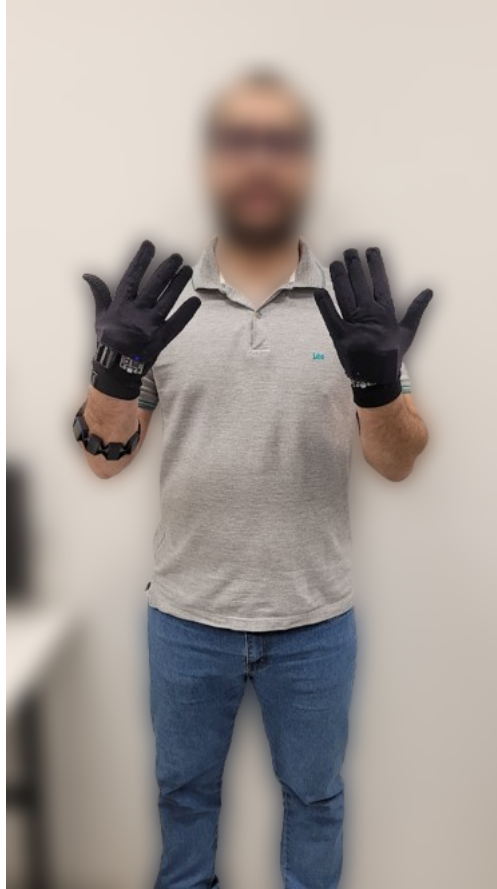
Fonte: Autoria própria (2024).

Como dito por Merletti, Rainoldi e Farina (2001) existe a necessidade de padronização da metodologia quanto à localização específica dos eletrodos, assim, um padrão foi definido para as coletas: o vestimento da *armband Myo™* foi realizado no braço dominante do voluntário, com o sensor principal (identificado através de um *led*) voltado para a parte frontal do antebraço, cerca de 4 centímetros abaixo do cotovelo. Esse procedimento pode ser visto na Figura 18.

As coletas de dados foram realizadas com um voluntário de cada vez, e para cada uma das 50 palavras, os voluntários tiveram 52 segundos corridos para executar as 10 repetições, das quais 8 foram aproveitadas para o estudo. Como a maioria dos voluntários não conheciam os movimentos em Libras, a cada mudança de palavra para a coleta, um vídeo com o exemplo de como o sinal deveria ser executado era reproduzido em *loop* em um monitor à frente do voluntário. Assim, cada voluntário executou o mesmo sinal 10 vezes seguidas, intervalando com momentos em repouso após cada execução. A cada 5 segundos, a *armband Myo™* vibrava de modo que o usuário soubesse que deveria iniciar a repetição do gesto novamente. Ao fim dos 52 segundos, a *armband* vibrava de modo mais forte e contínuo, indicando para o voluntário o fim da coleta para aquela palavra. Esse processo de coleta se repetiu entre todos os usuários para as 50 palavras selecionadas.

Com isso, foi possível a criação do banco de dados com 2.800 amostras de sinais em Libras, distribuídos em 5.600 arquivos no formato “CSV” (metade dos arquivos com dados dos 8

Figura 18 – Voluntário utilizando as luvas instrumentadas desenvolvidas nesse projeto junto a *armband* Myo™.



Fonte: Autoria própria (2024).

canais EMG da *armband*, e a outra metade com dados dos 10 sensores flexíveis e dois sensores inerciais).

3.2.2 Processamento e sincronização dos dados

Com o banco de dados coletado, foi necessário realizar o processamento dos dados brutos a fim de sincronizar os dados das luvas com os dados da *armband*, além de segmentar cada uma das 8 repetições de cada palavra. Uma necessidade para esse processamento estava ligada ao fato de que a taxa de amostragem do sistema da luva não era fixa. O sistema de aquisição utilizava os dados coletados pela *armband* como referência para controlar o tempo de coleta. Dessa forma, como os dados das luvas não eram controlados durante cada captura, o volume de dados recebidos delas oscilava entre as coletas. Em média, a frequência da aquisição dos sensores da luva foi de 60 amostras/segundo, enquanto que a frequência de amostragem da *armband* foi fixa em 200 amostras/segundo. Essa taxa de amostragem se deve ao próprio funcionamento da *armband*, não sendo possível fazer alteração. Com isso, o dispositivo desenvolvido possui mais dados no arquivo da *armband* em relação ao arquivo de aquisição das luvas.

O processo utilizado para sincronizar os dados foi através da identificação dos movimento realizado com os dados das luvas e da *armband* de forma isolada antes realizar a união como pode ser visto na Figura 19, tendo em vista que os dados das duas luvas foram sincronizados no momento da captura. Através de um limiar *threshold* inserido analisando os dados de um canal de EMG e de um canal de acelerômetro da luva direita, foi possível deixar os dados da coleta binários, com o valor de 1 para movimento e 0 para repouso. Para remover o ruído, foi utilizado um filtro, utilizando o uma função em Python de filtro Butterworth, definido com ordem 1 e uma frequência de corte de 0,0015 Hz para os dados EMG e 0,005 Hz para os dados da Luva. Por padrão, isso cria um filtro passa-baixa, permitindo identificar os momentos de realização do sinal e de e repouso. Com isso, os sinais em Libras foram identificados e segmentados para os *dataframes* das luvas e da *armband*. Com os dados dos movimentos segmentados entre cada repetição, foi realizada a união dos dados dos dispositivos utilizando o mesmo método utilizado no primeiro processo, unindo e ordenando os dados através do *timestamp*, e preenchendo os dados nulos com o valor do último dado coletado.

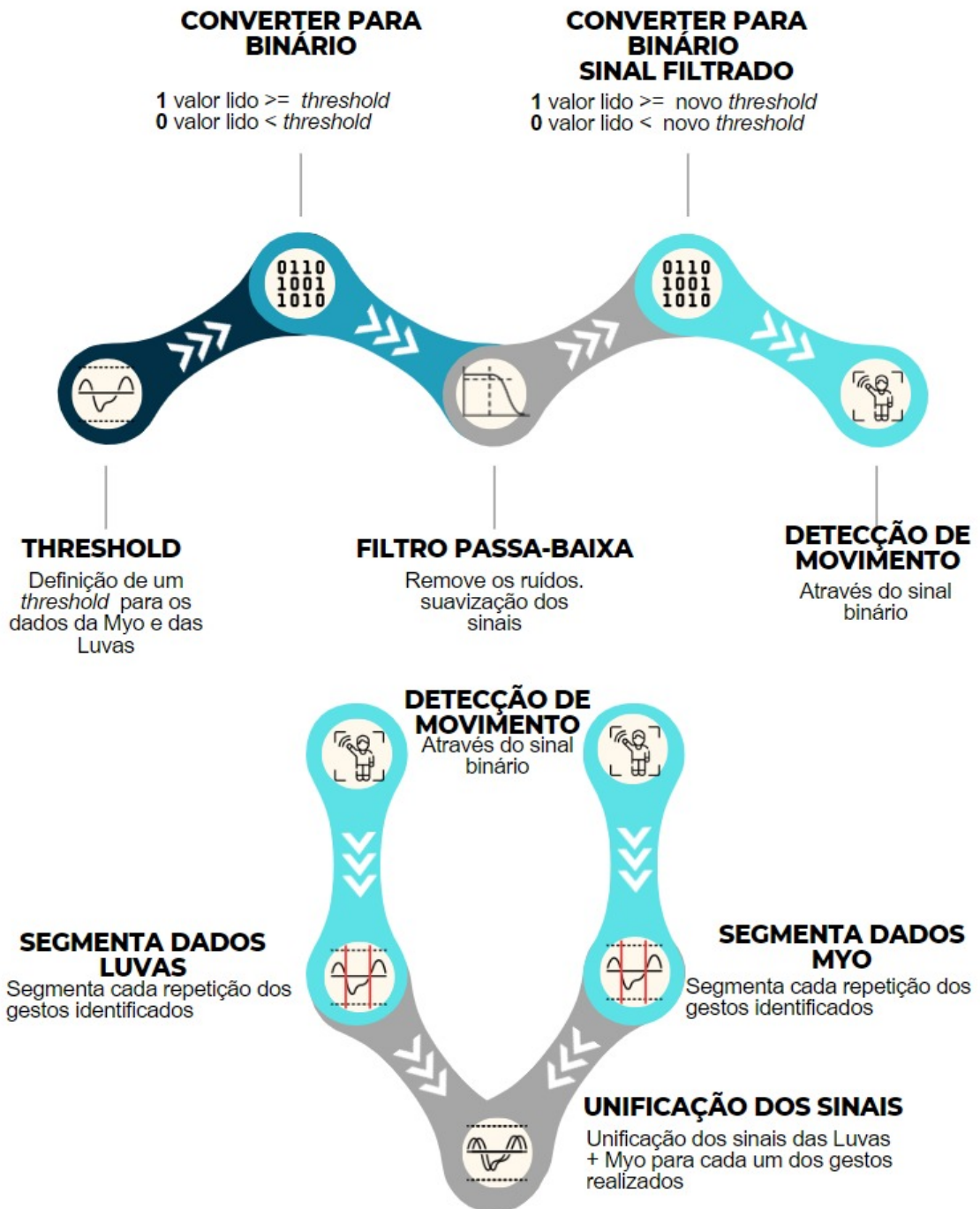
Os arquivos dos sensores EMG foram unidos com os dados das luvas, utilizando o campo *timestamp* como chave. Assim, cada *timestamp* representa uma linha da base de dados. Quando dois arquivos tinham exatamente o mesmo *timestamp*, os dados foram incrementados na mesma linha, mas em colunas diferentes. Após juntar os dados e ordenar através do *timestamp*, foi necessário preencher as colunas que tinham dados nulos. Para isso, foi utilizado o método de inserir nos campos nulos o mesmo valor do último dado coletado não nulo, método de preenchimento *forward-fill* fornecido pela biblioteca Pandas (PANDAS, 2024). Com isso, foi possível obter um arquivo com os dados da luva e da *armband* sincronizados para cada repetição dos voluntários em cada uma das 50 palavras.

3.2.3 Classificação com Deep Learning

Para a classificação dos sinais em Libras, foram utilizados dois modelos de *Deep Learning*: o gMLP (*Gated Multilayer Perceptron*) e o InceptionTime. Ambos os modelos foram treinados e testados utilizando os dados segmentados e sincronizados obtidos na etapa de coleta e processamento. Cada gesto foi representado por uma sequência de sinais capturados pelos sensores flexíveis e inerciais das luvas, bem como pelos sinais EMG da *armband*.

Para avaliar o desempenho dos modelos de *Deep Learning*, foi empregado o método de validação cruzada *k-fold*. Este método é amplamente utilizado para validar a robustez e a generalização dos modelos em conjuntos de dados limitados. O método *k-fold cross-validation* consiste em dividir o conjunto de dados em *k* subconjuntos (ou “folds”) de tamanho aproximadamente igual. Em cada rodada de validação, um dos *k* subconjuntos é utilizado como conjunto de teste, enquanto os outros *k-1* subconjuntos são combinados para formar o conjunto de treinamento. Este processo é repetido *k* vezes, de modo que cada subconjunto é utilizado exatamente

Figura 19 – Passos utilizados para o processamento e identificação dos movimentos.



Fonte: Autoria própria (2024).

uma vez como conjunto de teste. Os resultados das k iterações são agregados para fornecer uma estimativa final da performance do modelo.

Neste trabalho, utilizou-se k igual a 5 para a validação cruzada (5-fold cross-validation) e a cada validação os grupo de treinamento e de teste foram gerados aleatoriamente respeitando a proporção 80% e 20% respectivamente. Estes grupos gerados foram aplicados para treinar os dois modelos paralelamente. O processo foi repetido dez vezes para proporcionar uma validação estatística com maior robustez, permitindo assim uma avaliação mais confiável e precisa do

desempenho dos modelos. Dessa forma, buscou-se minimizar a variação e garantir que os resultados obtidos fossem representativos e consistentes.

Os resultados da classificação foram avaliados com base em métricas como Acurácia, Perda de Treinamento, Perda de Validação e Área Sob a Curva ROC, permitindo uma análise comparativa detalhada entre os modelos gMLP e InceptionTime.

4 RESULTADOS E DISCUSSÃO

Nesta seção, são apresentados e discutidos os resultados obtidos nas diversas etapas do projeto, incluindo a análise do sistema físico, a eficácia dos algoritmos de coleta, o resultado da sincronização dos dados e a comparação entre diferentes modelos de *Deep Learning* para a classificação dos sinais em Libras.

4.1 Sistema físico e aquisição de dados

O sistema físico desenvolvido, conforme apresentado na Figura 20, composto pelas duas luvas vestíveis instrumentadas com os sensores flexíveis e os sensores inerciais MPU6050 conectados e controlados pelo microcontrolador ESP32. Os sensores da luva foram alimentados por uma bateria de 9 V e os dados foram transmitidos via Bluetooth, sendo capaz de realizar amostragem de dados em uma taxa média de 60 amostras/segundo, capturando os dados da flexão e extensão dos dedos, além dos movimentos de aceleração, rotação e velocidade angular durante a execução dos gestos.

Figura 20 – Realização da coleta de dados com um voluntário.

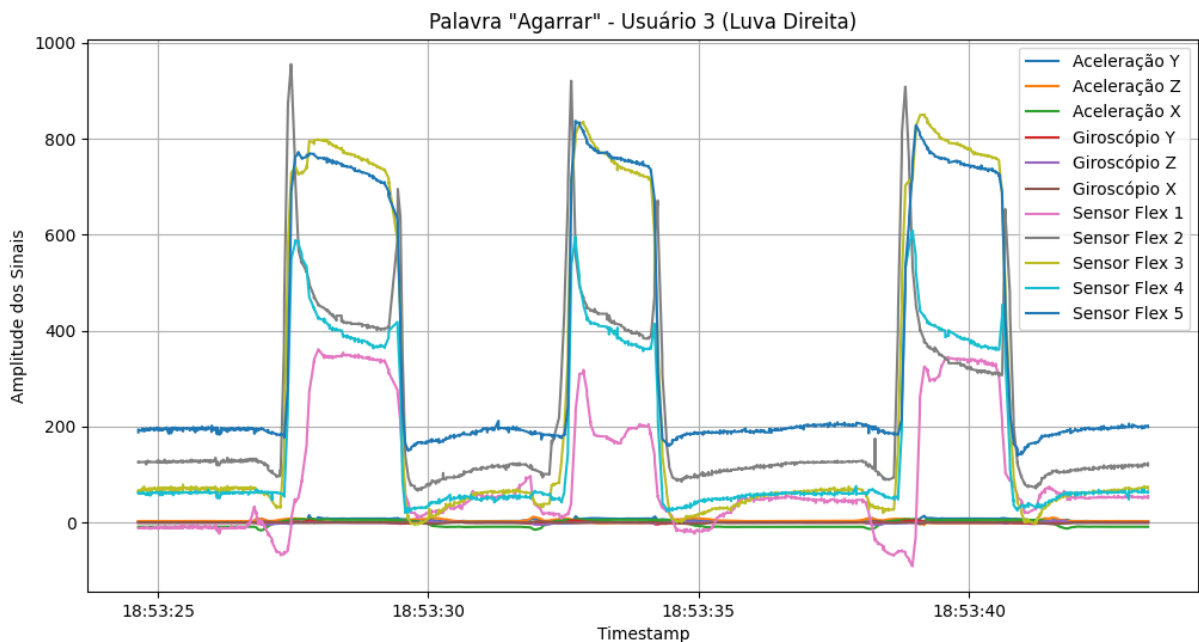


Fonte: Autoria própria (2024).

O sistema de aquisição de dados implementado foi capaz de capturar os movimentos desenvolvidos pelo usuário de forma parametrizada, uma vez que é possível editar os parâmetros como tempo de coleta e seleção dos equipamentos (luva da mão esquerda, luva da mão direita e *armband Myo™*). Utilizando as *threads* em Python, o sistema foi capaz de realizar a amostragem de dados de forma paralela. Cada coleta realizada resultou em dois arquivos de dados. O primeiro arquivo continha os dados dos 8 canais EMG da *armband Myo™*. O segundo arquivo foi composto por dados dos 10 sensores flexíveis (5 de cada luva) e 12 dados referentes aos sensores inerciais (6 de cada luva) sendo 3 de acelerômetro em cada eixo X, Y e Z e 3 do giroscópio para os mesmos eixos. A taxa de amostragem da *armband* foi de 200 amostras/segundo enquanto da luva em média 60 amostras/segundo.

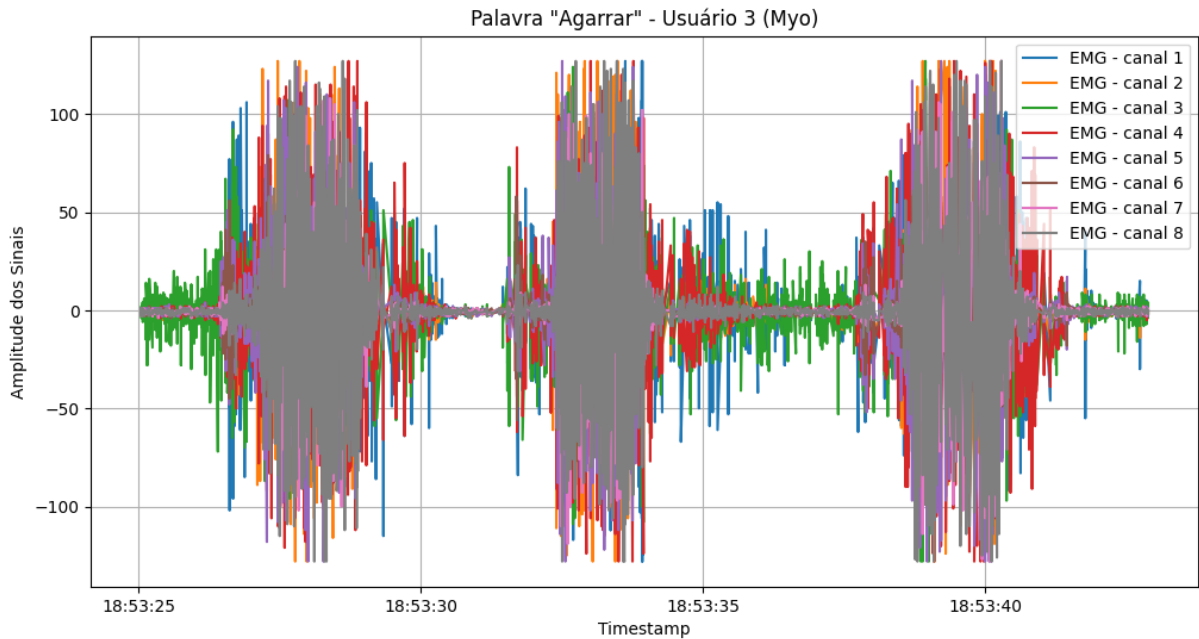
Para exemplificar os sinais obtidos pelos sensores, as Figuras 21 e 22 apresentam os dados dos sensores flexíveis, dos inerciais e de EMG de uma das aquisições para o sinal que representa a palavra “agarrar”, enquanto as Figuras 23 e 24 apresentam os sinais para a palavra “escola”. Como pode ser visto, os sinais obtidos pelos sensores flexíveis possuem um aspecto mais suavizados que os sinais de EMG, os quais possuem uma característica aspecto ruidoso. Para este sinal, pode-se perceber que foram recrutados movimentos relacionados a todos os sensores flexíveis, pois estes estão localizados sobre todos os dedos e refletem a atuação do sinal de Libras para agarrar (durante este movimento, os dedos se fecham progressivamente, como se estivessem envolvendo um objeto). O mesmo pode ser observado pelo sinal de EMG, que apresenta o recrutamento dos canais referentes a musculatura correspondente a esse sinal.

Figura 21 – Parte dos dados brutos dos sensores flexíveis e inerciais da luva direita para a coleta da palavra “Agarrar”.



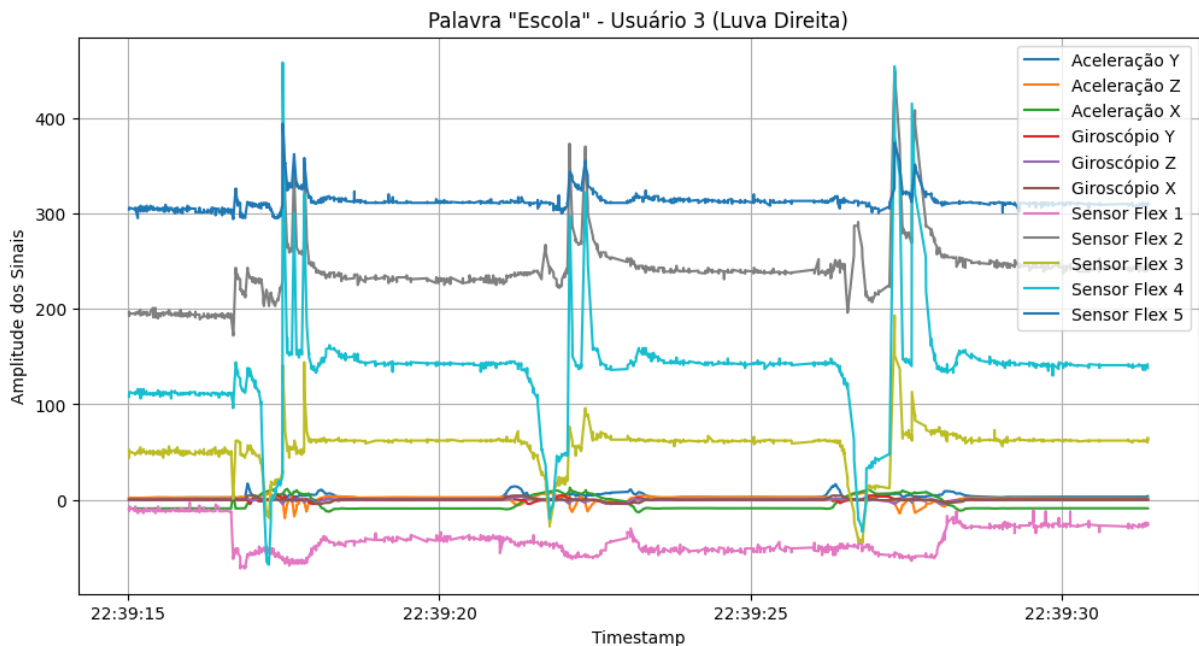
Fonte: Autoria própria (2024).

Figura 22 – Parte dos dados brutos dos sensores EMG de uma coleta da palavra “Agarrar”.



Fonte: Autoria própria (2024).

Figura 23 – Parte dos dados brutos dos sensores flexíveis e inerciais da luva direita para a coleta da palavra “Escola”.

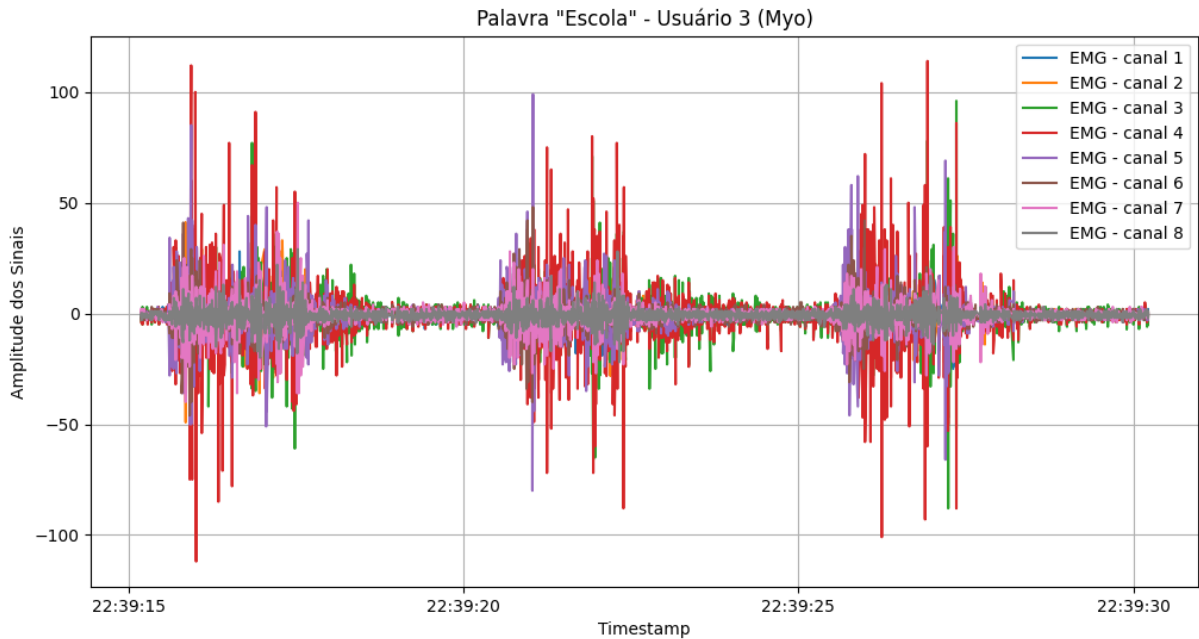


Fonte: Autoria própria (2024).

4.2 Sincronização e segmentação dos dados

Por meio dos métodos descritos na Seção 3.2.2, identificando os momentos da realização do sinal e de repouso, os dados adquiridos do sistema de aquisição foram segmentados e sincronizados. Primeiramente, foram definidos os limiares das luvas e da *armband* que atuam

Figura 24 – Parte dos dados brutos dos sensores EMG de uma coleta da palavra “Escola”.



Fonte: Autoria própria (2024).

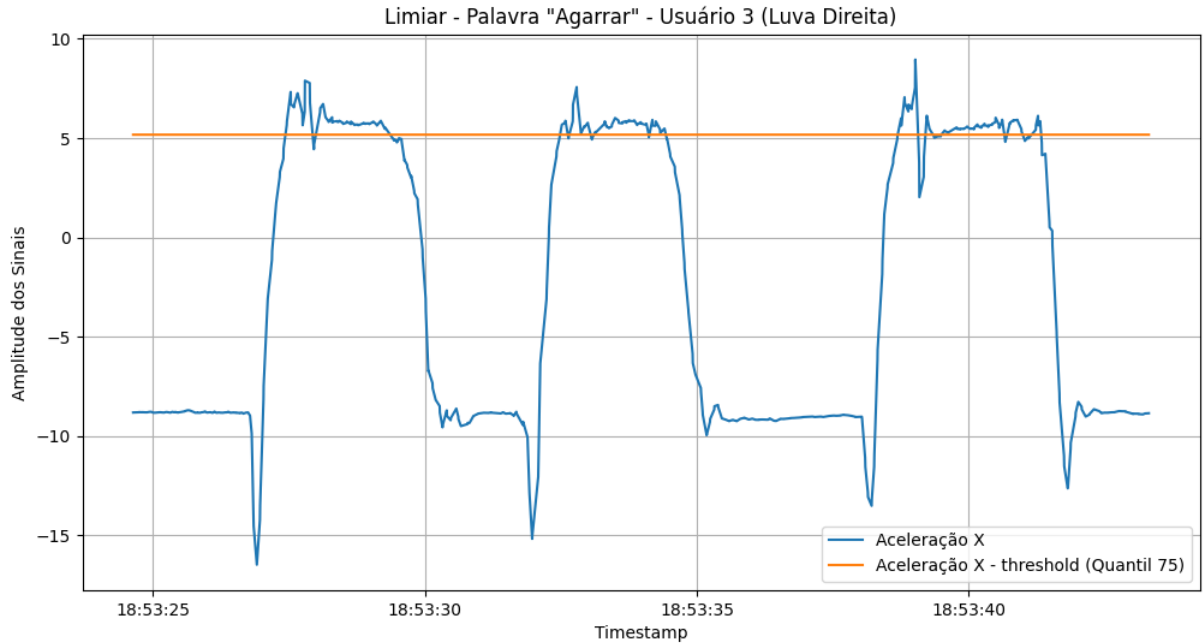
na detecção entre movimento e repouso. O valor do limiar das luvas foi definido dinamicamente para cada usuário com o uso dos sinais do eixo X do sensor de acelerômetro, utilizado o quartil de 75% desses dados. Esta porcentagem foi definida de forma empírica. Para o limiar de EMG, um canal foi utilizado para a definição através do desvio padrão. Como apresentado nas Figuras 25 e 26, o sinal referente a palavra 'agarrar' possui diferenças definidas entre a execução do sinal e o repouso. Após esta definição do limiar, foram identificados quando cada um dos sinais ultrapassou o limiar, sendo apresentado um resultado na Figura 27 para o exemplo do sinal EMG da palavra 'agarrar'.

Uma vez que os sinais resultantes da primeira etapa possuem pontos nos quais há várias sobreposições entre sinal e repouso, esta resultante foi inserida no filtro passa-baixa para diminuir os ruídos, conforme apresentado na Figura 27. Pode-se notar que com a aplicação do filtro, foi feita a suavização dos sinais, permitindo que estes pudessem ser identificados e segmentados. Assim, os sinais suavizados foram mais uma vez passados por um limiar, que foi definido com o quartil de 50% tanto para os dados das luvas quanto para os dados da *armband*.

A Figura 28 apresenta a interposição entre os instantes obtidos com a saída do filtro passa-baixa, na detecção de passagem pelo limiar, e os sinais de EMG de entrada. Nota-se que o ruído na identificação foi eliminado e que os sinais conseguem ser separados entre ativação e repouso.

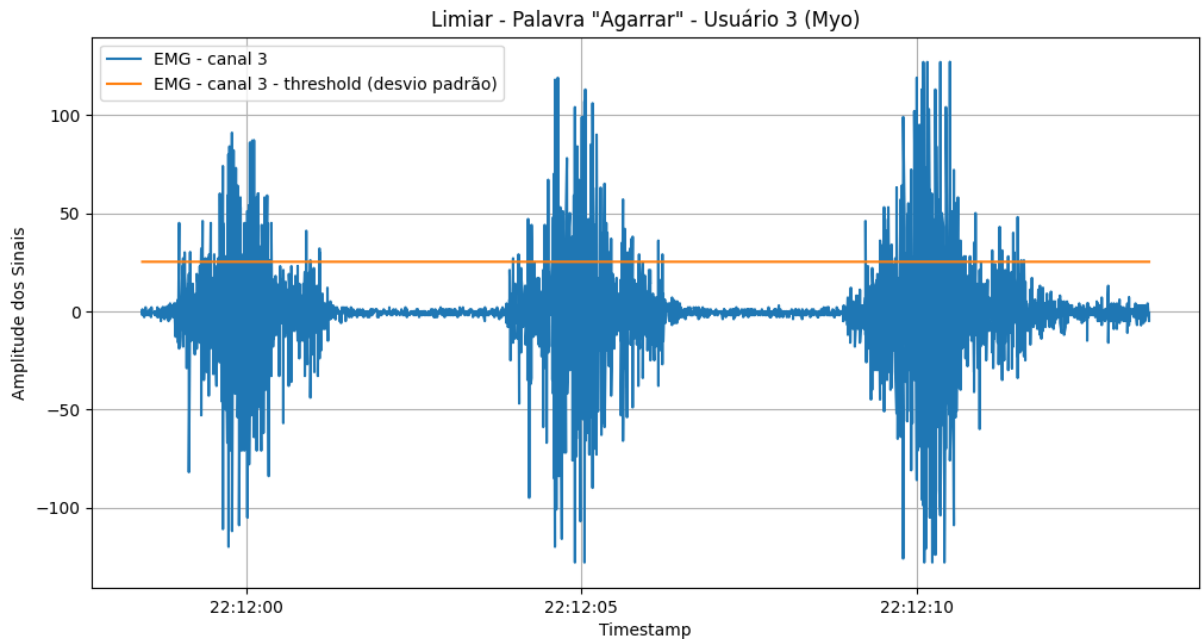
Assim, os dados das luvas e os dados de EMG foram segmentados para cada repetição de sinal em Libras identificado em cada um dos dois sistemas. Essa segmentação foi realizada salvando em novos arquivos somente os dados quando a resultante do filtro foi igual a “1”, ou

Figura 25 – Dados de um sensor de Acelerômetro e inserção de um limiar para identificar movimento ou repouso.



Fonte: Autoria própria (2024).

Figura 26 – Dados de um canal de EMG e inserção de um limiar para identificar movimento ou repouso.

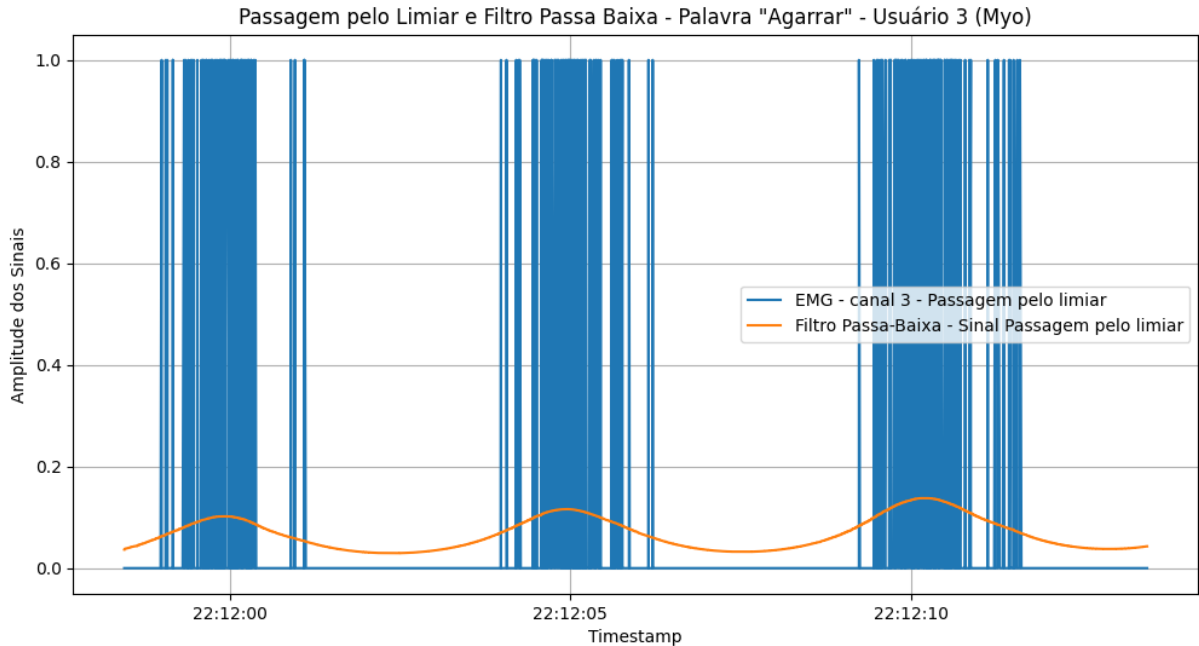


Fonte: Autoria própria (2024).

seja, o sinal em Libras estava sendo executado, com isso os dados coletados durante o repouso foram descartados.

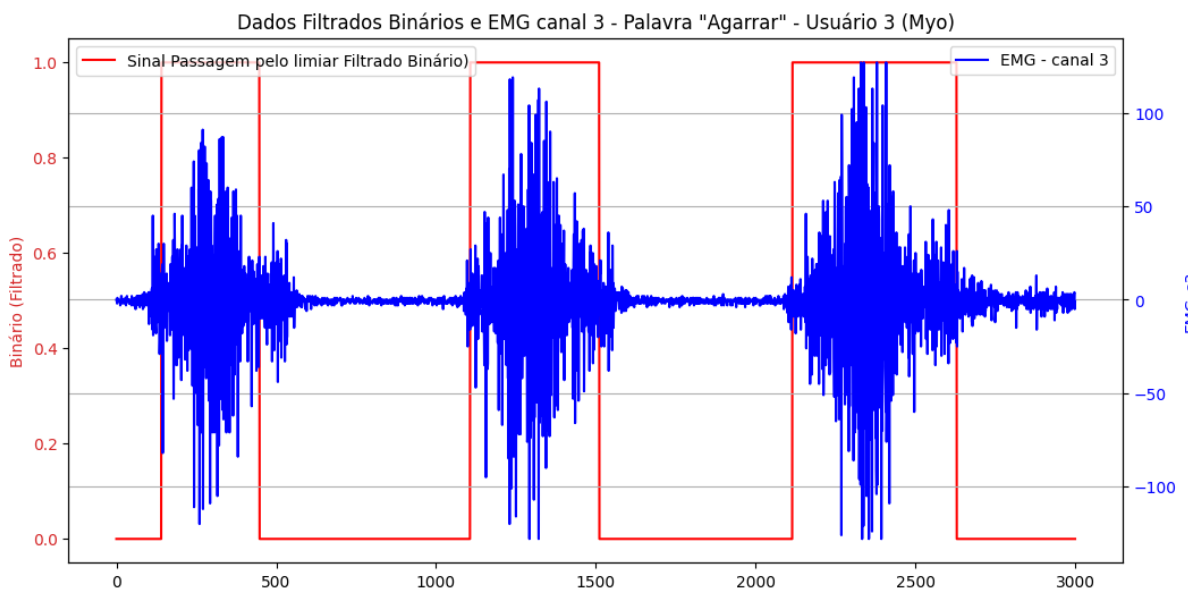
Por fim, com os dados de cada repetição segmentados, foi possível realizar a união entre os dados das Luvas e os dados da *armband* equivalentes. Para isso os dados foram sincronizados em relação ao *timestamp*, e quando havia algum valor nulo ele era preenchido pelo último

Figura 27 – Dados de um canal de EMG e inserção de um limiar para identificar movimento ou repouso.



Fonte: Autoria própria (2024).

Figura 28 – Detecção de movimento em um sinal EMG.

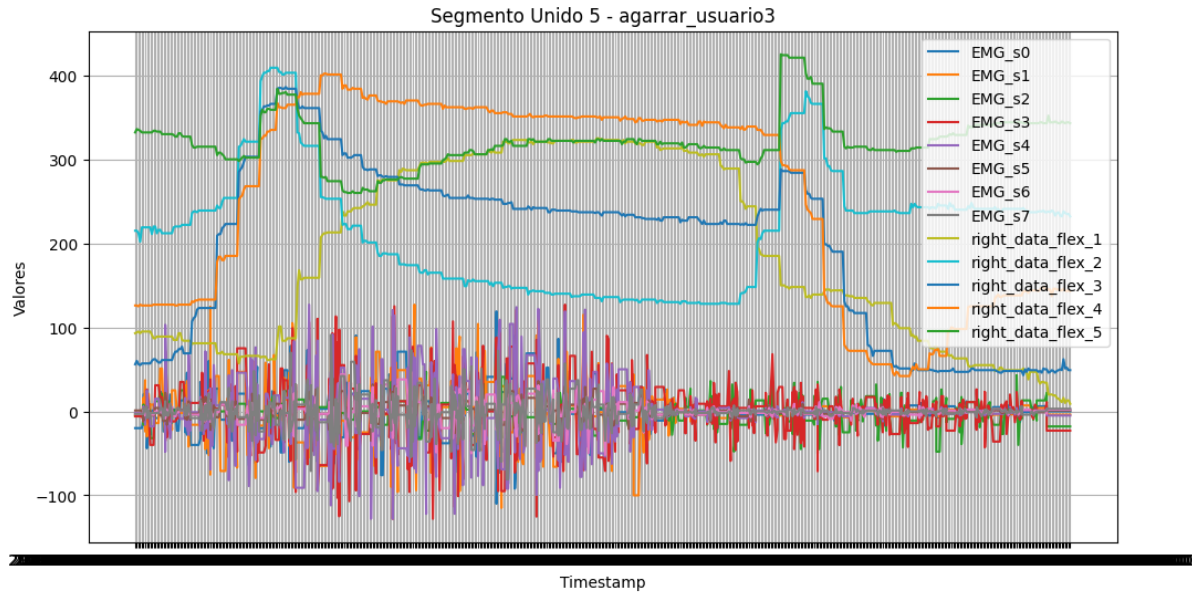


Fonte: Autoria própria (2024).

valor recebido. Para exemplificar essa ação, a Figura 29 mostra os dados dos sensores flexíveis da mão direita e os 8 canais de EMG da *armband*. Nota-se que quando há ativação muscular, há também a resposta do sensor flexível. Em comparação com o EMG, pode-se notar o efeito do preenchimento de amostras, fazendo com que os sinais flexíveis sejam mantidos por um determinado valor por mais tempo. Porém, isso não impacta o processo de processamento e classificação, já que a resposta do sensor flexível é mais lenta em relação ao sinal de EMG. Isso ocorre porque a resposta do sensor flexível depende das ações humanas e estas podem ocorrer

em uma frequência média de 5 Hz, conforme descrito por Huang e Griffin (2006). Esta frequência é significativamente inferior à taxa de amostragem de ao 200 Hz (ou amostras/segundo) da *armband Myo™*.

Figura 29 – Amostra segmentada com dados de uma luva e do bracelete do gesto da palavra “agarrar”.

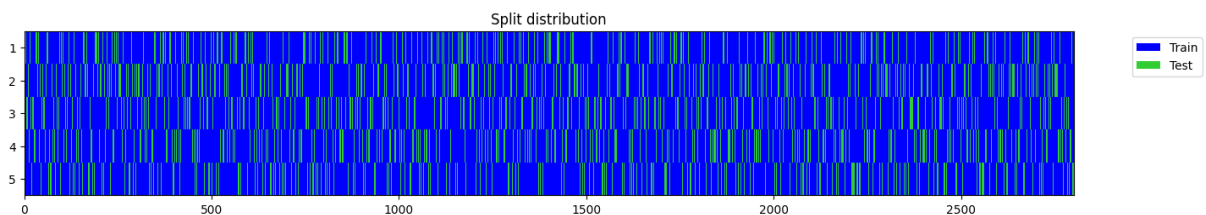


Fonte: Autoria própria (2024).

4.3 Classificação com Deep Learning

Após realizar a sincronização e segmentação de cada gesto, foram empregados os modelos gMLP e *InceptionTime* de *Deep Learnig* para a classificação. Os dados foram divididos 10 vezes utilizando validação cruzada com 5 subconjuntos independentes em treinamentos de 100 épocas, utilizando 20% dos dados para teste e 80% para treinamento dos modelos, usando a função “get_split” da biblioteca “tsai”, esta função separa de forma aleatória os dados de entrada em grupos de treino e teste. A Figura 30 apresenta uma das 10 distribuições obtida para cada um dos 5 subconjuntos de teste utilizados na validação cruzada.

Figura 30 – Divisão da distribuição para a validação cruzada.

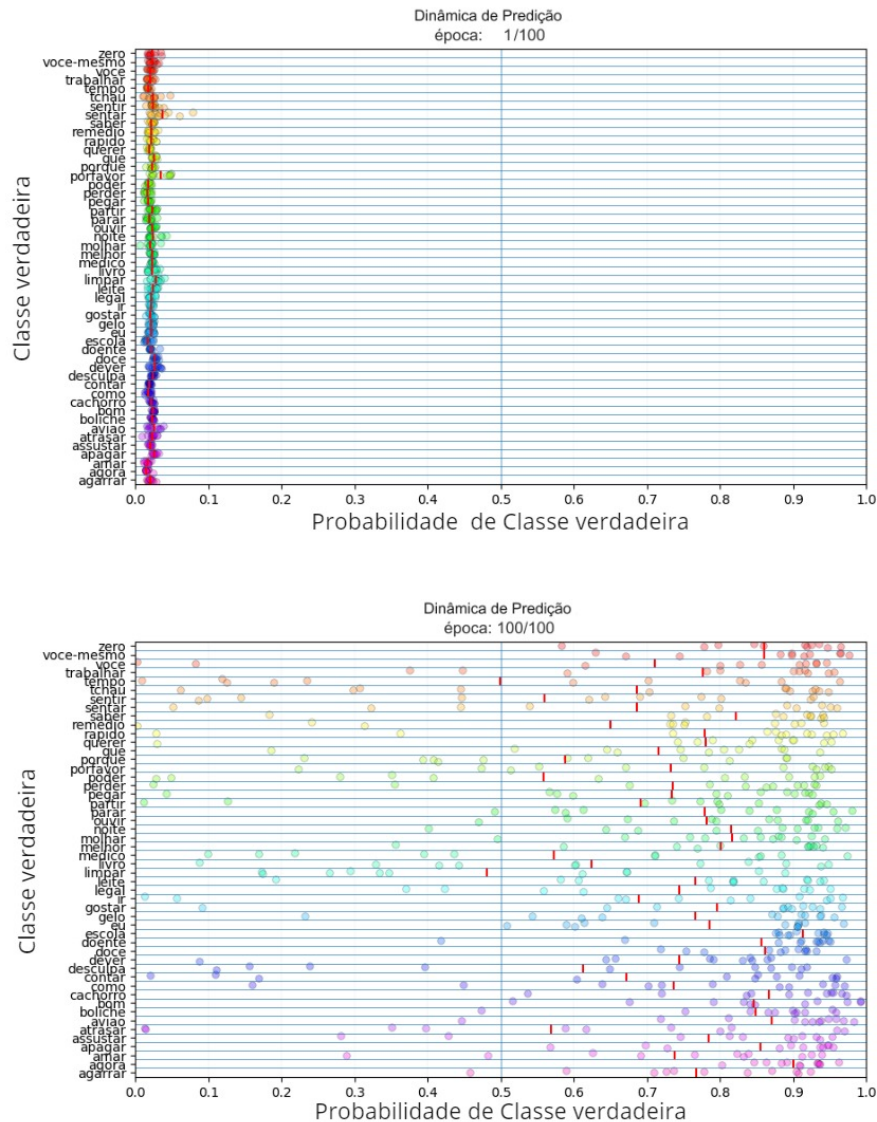


Fonte: Autoria própria (2024).

Conforme o treinamento avançava ao longo das 100 épocas, a probabilidade de acerto das classes foi monitorada, conforme ilustrado na Figura 31. Ela compara o instante inicial do

treinamento com o instante final, proporcionando uma visão intuitiva do progresso do modelo. Conforme pode ser observado, a probabilidade de acerto das classes é alterada ao decorrer das épocas, saindo da probabilidade a priori (aproximadamente igual a 1/50) para valores próximos das acurácias obtidas.

Figura 31 – Predição dinâmica InceptionTime. Comparação entre a primeira e a última época em um treinamento.

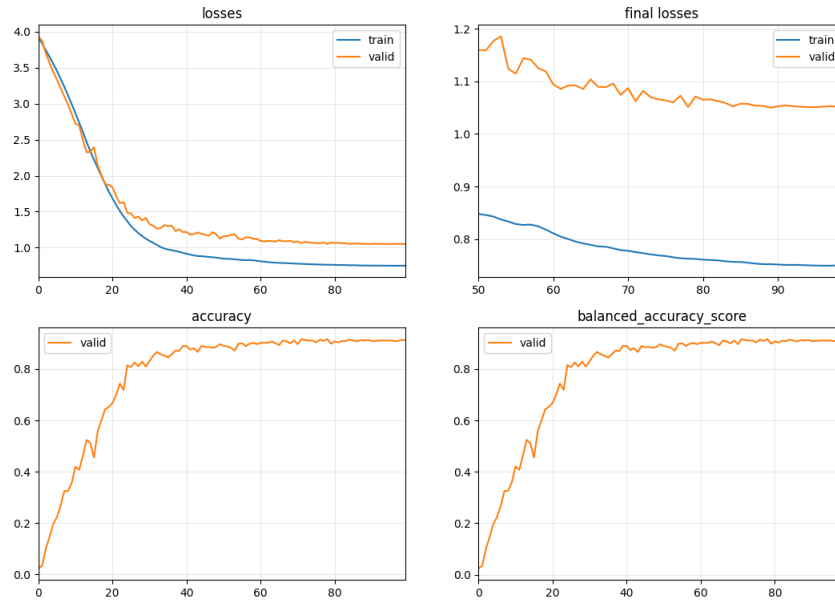


Fonte: Autoria própria (2024).

Após as 100 épocas de cada treinamento e 10 repetições de validações cruzadas com os 5 subconjuntos, alcançou-se acurácia média de 89,77%, perda de treinamento de 0,639 e perda de validação de 1,004 com o modelo gMLP. Com o modelo *InceptionTime*, a acurácia média obtida foi de 92,21% com perda de treinamento e perda de validação 0,68 e 0,95 respectivamente, representadas na Figura 32. Percebeu-se que o modelo *InceptionTime* se saiu superior para este caso de uso, uma vez que além de apresentar uma acurácia média superior e uma menor perda de validação, o que indica que é mais preciso e generaliza melhor do que

o modelo gMLP, ele também utilizou menos unidades de processamento, levando em média 2 minutos, utilizando a GPU T4 do Google Colab, para realizar o treinamento de cada época, frente aos 11 minutos levados pelo modelo gMLP.

Figura 32 – Métricas de validação durante execução das 100 épocas com o modelo InceptionTime.



Fonte: Autoria própria (2024).

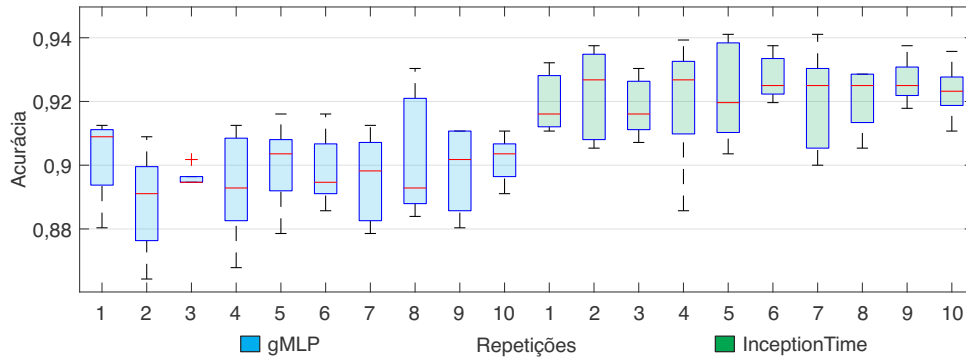
Para confirmar os resultados verificados nas métricas da Figura 32, a Figura 33 apresenta as distribuições das acurácias obtidas para cada uma das 10 repetições, tanto para o modelo gMLP quanto para o *InceptionTime*. As médias dos resultados obtidos para todas as validações pode ser visto na Tabela 2. Além disso, foi utilizado um teste estatístico para verificar as distribuições obtidas pelos dois classificadores possuíam diferenças significantes. A partir do teste de Wilcoxon, verificou-se que as duas distribuições possuem diferença estatística em um intervalo de 95% de confiança ($p < 0.05$), ou seja, as duas distribuições são estatisticamente diferentes.

Tabela 2 – Médias das acurácias obtidas para todas as validações cruzadas de *Deep Learning* para os modelos gMLP e *InceptionTime*.

Repetição	Acurácia gMLP	Acurácia <i>InceptionTime</i>
1	90,21%	91,96%
2	88,82%	92,25%
3	89,61%	91,82%
4	89,36%	92,00%
5	90,00%	92,29%
6	89,86%	92,75%
7	89,57%	92,00%
8	90,29%	92,07%
9	89,82%	92,64%
10	90,18%	92,32%

Fonte: Autoria própria (2024).

Figura 33 – Distribuições obtidas para as 10 repetições de validação cruzada.



Fonte: Autoria própria (2024).

Uma vez que modelo *InceptionTime* apresentou os melhores resultados, a Figura 34 apresenta a Matriz de Confusão obtida com este modelo, sendo possível identificar os gestos com maior taxa de erro de classificação. Pode ser observado pela diagonal principal que o modelo apresentou alta acurácia para todos os gestos. Nota-se que, em alguns casos, houveram poucas confusões (ou seja, classificações errôneas) acima de 1. Este casos estão detalhados na Tabela 3, a qual resume as confusões ocorridas entre os respectivos pares de gestos.

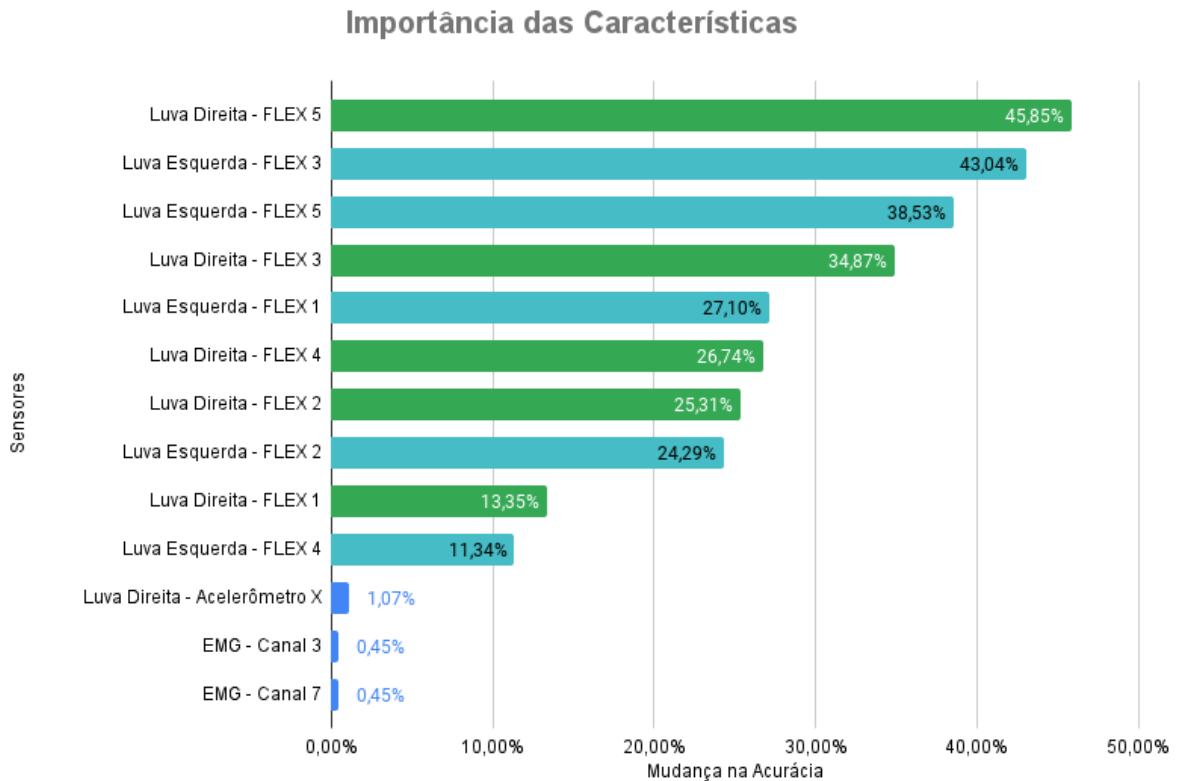
Tabela 3 – Maiores confusões entre palavras durante a classificação com o modelo de *InceptionTime*.

Sinal original	Sinal interpretado	Quantidade de confusões
“limpar”	“parar”	3
“cachorro”	“boliche”	2
“eu”	“querer”	2
“pegar”	“leite”	2
“querer”	“amar”	2
“saber”	“boliche”	2
“tempo”	“sentar”	2

Fonte: Autoria própria (2024).

Utilizando a biblioteca dos modelos, foi possível validar a importância de cada sensor e canal de EMG, neste caso, sendo considerado como atributo de entrada do classificador ou *feature*, fornecida aos modelos. Observou-se que, ao inserir todos os dados brutos com uma normalização dos mesmos, as *features* provenientes dos sensores flexíveis foram identificadas como as mais importantes para o desempenho do modelo, entre elas a com maiores relevância as originárias nos sensores flexíveis 3 e 5 que estão localizadas nos dedos médios e mínimos respectivamente de cada mão. O resultado demonstra que mesmo não sendo a mão dominante, os dados provenientes da mão esquerda tem grande importância para o aumento da acurácia do modelo como pode ser visto na Figura 35.

Figura 35 – Importância das sensores (atributos de entrada do classificador) para o reconhecimento dos sinais.



Fonte: Autoria própria (2024).

Na fase de classificação, os modelos gMLP e *InceptionTime* foram treinados utilizando uma validação cruzada com 5-fold e repetida dez vezes, garantindo assim uma robusta avaliação estatística. Os resultados demonstraram que o modelo *InceptionTime* apresentou uma acurácia média superior de 92,21% em comparação aos 89,77% do modelo gMLP, além de uma menor perda de validação, indicando melhor capacidade de generalização (Tabela 2). Este desempenho superior, aliado ao menor tempo de treinamento por época, torna o *InceptionTime* mais eficiente e eficaz para o uso em reconhecimento de gestos em Libras.

A análise da matriz de confusão (Figura 34) revelou que, embora houvesse confusões entre alguns gestos, estas foram relativamente poucas, com no máximo três erros por par de gestos. Os gestos mais confundidos estão listados na Tabela 3, evidenciando a necessidade de ajustes finos ou maior volume de dados para essas classes específicas.

A avaliação dinâmica das probabilidades de acerto para cada classe (Figura 31) confirmou a consistência do modelo *InceptionTime* ao longo das diferentes repetições e subconjuntos de validação cruzada. Além disso, a análise da importância das *features* (Figura 35) destacou a relevância dos dados provenientes dos sensores flexíveis, especialmente dos dedos médios e mínimos de ambas as mãos, para o desempenho do modelo. Este achado é coerente com a literatura exposta no referencial teórico, que indica a significativa contribuição dos movimentos dos dedos para a interpretação de gestos.

O Quadro 2 apresenta a comparação com os resultados obtidos neste trabalho com trabalhos semelhantes na literatura, descritos na Seção 1.3. Pode-se notar certas similaridades e diferenças significativas. Em particular, o uso do método de *Deep Learning* apresentou uma acurácia compatível com a de outros métodos, como Decision Tree (DT), k-Nearest Neighbor (KNN) e Rede Neural Artificial (RNA), com a vantagem de não requerer a extração de características dos sinais para utilizar o modelo. Esta abordagem simplifica o processo e potencialmente aumenta a eficiência.

A utilização de duas luvas e sensores EMG contribuiu significativamente para a precisão do modelo, permitindo a realização de gestos mais complexos que aqueles considerados nos estudos de Silvano *et al.* (2016) e Santos, Ferreira e Alves (2019). A luva de Borges (2020) alcançou uma acurácia muito alta de 99,97%, porém utilizou apenas seis sensores e um número reduzido de gestos. A limitação a um único voluntário também pode afetar a generalização dos resultados. Em contraste, o presente trabalho envolveu mais voluntários e um conjunto mais diversificado de gestos, proporcionando uma avaliação mais robusta e aplicável em cenários reais.

Em síntese, o uso de duas luvas e sensores EMG, aliado ao método de *Deep Learning*, permitiu alcançar uma acurácia competitiva, sem a necessidade de extração de características, e com capacidade para reconhecer gestos mais complexos. Estes fatores destacam a eficácia e a inovação deste estudo em relação aos trabalhos existentes.

Quadro 2 – Comparação com outros estudos relacionados a Luvas tradutoras de Libras.

Autor	Sensores	Voluntários	Sinais	Classificador	Acurácia
Silvano <i>et al.</i> (2016)	6	6	26	DT	90,38%
Santos, Ferreira e Alves (2019)	6	-	36	KNN	95%
Borges (2020)	6	1	22	RNA	99,97%
Este trabalho	20	7	50	<i>InceptionTime</i>	92,21%

Fonte: Autoria própria (2024).

5 CONSIDERAÇÕES FINAIS

Este trabalho teve como objetivo principal desenvolver um sistema para a captura e classificação de gestos em Libras utilizando tecnologias de sensoriamento vestível e *Deep Learning*. O referencial teórico abordou conceitos de Libras, sensoriamento vestível, processamento de sinais e classificação de dados utilizando técnicas de *Deep Learning*. Para alcançar esse objetivo, foram utilizadas luvas instrumentadas com sensores flexíveis e inerciais, juntamente com a *Armband Myo*, para captura de dados durante a execução dos gestos.

Através da implementação do sistema de aquisição de dados e da aplicação de modelos de *Deep Learning*, foi possível obter resultados promissores na classificação dos gestos em Libras, atingindo uma acurácia média de 92,21% com o modelo *InceptionTime*. As tecnologias e ferramentas utilizadas demonstraram ser eficazes na captura e processamento dos dados, permitindo uma análise detalhada dos movimentos realizados durante a execução dos gestos.

Dessa forma, os resultados obtidos confirmam a hipótese de que o modelo de *Deep Learning InceptionTime* é adequado para a tarefa de reconhecimento de gestos em Libras, não apenas pela sua maior precisão, mas também pela eficiência computacional. A implementação deste modelo pode, portanto, contribuir significativamente para o desenvolvimento de sistemas de tradução automática de Libras, facilitando a comunicação entre surdos e ouvintes e promovendo a inclusão social.

No entanto, algumas dificuldades foram encontradas ao longo do desenvolvimento do trabalho. Questões relacionadas à sincronização dos dados das luvas com os dados da *armband* e à segmentação dos gestos foram identificadas e exigiram soluções específicas para garantir a precisão e confiabilidade dos resultados.

Em relação às vantagens, o sistema desenvolvido apresenta uma abordagem inovadora ao juntar diferentes tipos de sensores para a classificação de gestos em Libras, oferecendo uma alternativa e mostrando quais são os sensores que mais tem impacto durante essa classificação. Além disso, o uso de tecnologias vestíveis permite uma maior mobilidade e liberdade de movimento durante a execução dos gestos.

Por outro lado, algumas limitações foram observadas, como a necessidade de calibração e ajustes frequentes nas luvas instrumentadas, bem como a dependência de uma conexão *Bluetooth* estável para a transmissão dos dados.

No geral, o trabalho contribui para a comunidade acadêmica e profissional ao explorar novas abordagens para a comunicação em Libras e ao fornecer *insights* sobre o uso de tecnologias vestíveis e *Deep Learning* para essa finalidade.

Em conclusão, os objetivos propostos foram alcançados, e os resultados obtidos demonstram o potencial das tecnologias utilizadas no desenvolvimento do sistema. No entanto, há espaço para melhorias e otimizações futuras, especialmente no que diz respeito à sincronização dos dados e à robustez do sistema em diferentes ambientes e condições de uso.

Quanto às perspectivas futuras, é relevante continuar explorando formas de aprimorar o sistema, como a investigação de novos algoritmos de classificação e a realização de testes em um número maior de usuários fluentes em Libras. Essas iniciativas podem contribuir para uma maior precisão e confiabilidade do sistema, ampliando assim seu potencial impacto na sociedade.

REFERÊNCIAS

- AL-AHDAL, M. E.; NOORITAWATI, M. T. Review in Sign Language Recognition Systems. **2012 IEEE Symposium on Computers & Informatics (ISCI)**, p. 52–57, mar. 2012. Conference Name: 2012 IEEE Symposium on Computers & Informatics (ISCI) ISBN: 9781467316866 9781467316859 9781467316842 Place: Penang, Malaysia Publisher: IEEE. Disponível em: <http://ieeexplore.ieee.org/document/6222666/>.
- AL-HAMMADI, M. *et al.* Hand Gesture Recognition for Sign Language Using 3DCNN. **IEEE Access**, v. 8, p. 79491–79509, 2020. ISSN 2169-3536. Disponível em: <https://ieeexplore.ieee.org/document/9078786/>.
- AL-KHAZZAR, A.; ALTAWHEEL, Z.; HUSSAIN, J. Using deep neural networks in classifying electromyography signals for hand gestures. **IAES International Journal of Artificial Intelligence (IJ-AI)**, v. 13, p. 217, mar. 2024.
- ALMEIDA, W. G. Introdução à língua brasileira de sinais. **Letras Vernáculas**, 2013. Disponível em: <https://portalidea.com.br/cursos/1759e1ccb7d31282d8c60d083262898f.pdf>.
- ALZUBAIDI, L. *et al.* Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. **Journal of Big Data**, v. 8, n. 1, p. 53, mar. 2021. ISSN 2196-1115. Disponível em: <https://doi.org/10.1186/s40537-021-00444-8>.
- AMOR, A. B. H.; GHOUL, O. E.; JEMNI, M. Sign Language Recognition Using the Electromyographic Signal: A Systematic Literature Review. **Sensors**, v. 23, n. 19, p. 8343, jan. 2023. ISSN 1424-8220. Number: 19 Publisher: Multidisciplinary Digital Publishing Institute. Disponível em: <https://www.mdpi.com/1424-8220/23/19/8343>.
- AZEVEDO, M. R. S. de *et al.* LuBraS: Uma Arquitetura de comunicação Libras-Língua Portuguesa Utilizando Controle de Voz. **Mostra Nacional de Robótica**, 2017. Disponível em: https://www.researchgate.net/profile/Carlos-Pantoja-3/publication/321020682_LuBraS_Uma_Arquitetura_de_comunicacao_Libras-Lingua_Portuguesa_Utilizando_Controlde_de_Voz/links/5a0b1de4458515e4827478cc/LuBraS-Uma-Arquitetura-de-comunicacao-Libras-Lingua-Portuguesa-Utilizando-Controlde-de-Voz.pdf.
- BABIUCH, M.; FOLTÝNEK, P.; SMUTNÝ, P. Using the ESP32 Microcontroller for Data Processing. *In: 2019 20th International Carpathian Control Conference (ICCC)*. [s.n.], 2019. p. 1–6. Disponível em: <https://ieeexplore.ieee.org/document/8765944>.
- BARRETT, K. E. **Ganong fisiología médica (24a)**. [S.l.]: McGraw Hill Mexico, 2013.
- BELL, C. **Beginning Sensor Networks with Arduino and Raspberry Pi**. [S.l.]: Apress, 2014. Google-Books-ID: G5kQAAwAAQBAJ. ISBN 978-1-4302-5825-4.
- BORGES, R. M. Trabalho de Conclusão de Curso (graduação), **Luva inteligente para traduzir libras em língua portuguesa usando método de inteligência artificial**. Faculdade UnB Gama: Universidade de Brasília, 2020. Accepted: 2020-11-25T21:23:32Z. Disponível em: <https://bdm.unb.br/handle/10483/26021>.
- BRASIL, L. d. D. Lei nº 10.436, de 24 de abril de 2002. dispõe sobre a língua brasileira de sinais-libras e dá outras providências. **Diário Oficial da União**, Poder Legislativo Brasília, DF, p. 23–23, 2002.

- BROWNE, M. W. Cross-Validation Methods. **Journal of Mathematical Psychology**, v. 44, n. 1, p. 108–132, mar. 2000. ISSN 0022-2496. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0022249699912798>.
- CARNEIRO, M. T. Trabalho de Conclusão de Curso (graduação), **IDENTIFICAÇÃO DE GESTOS POR SINAIS ELETROMIOGRÁFICOS DE SUPERFÍCIE UTILIZANDO APRENDIZADO DE MÁQUINA**. Curitiba, PR: Universidade Tecnológica Federal do Paraná, 2023.
- 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, jan. 2019. ISSN 1868-8071, 1868-808X. Disponível em: <http://link.springer.com/10.1007/s13042-017-0705-5>.
- COOPER, H.; HOLT, B.; BOWDEN, R. Sign Language Recognition. *In: Visual Analysis of Humans*. [S.l.: s.n.], 2011. p. 539–562. ISBN 978-0-85729-996-3. Journal Abbreviation: Visual Analysis of Humans.
- DAS, S. *et al.* A Review on Sign Language Recognition (SLR) System: ML and DL for SLR. *In: 2021 IEEE International Conference on Intelligent Systems, Smart and Green Technologies (ICISSGT)*. [s.n.], 2021. p. 177–182. Disponível em: <https://ieeexplore.ieee.org/abstract/document/9719411>.
- DIAS, T. S. **Luva instrumentada para reconhecimento de padrões de gestos em Libras**. 2020. Dissertação (Mestrado) — Universidade Tecnológica Federal do Paraná, 2020.
- DINIZ, I.; SOUZA, W.; NEVES, R. Tecnologia assistiva: Luva para interpretação de letras do alfabeto de Libras para deficiente auditivo e com comunicação com celular. *In: . Bauru, SP: Copyright © Canal 6 Editora, 2019. p. 433. ISBN 978-85-7917-514-5.*
- ESP32 Pinout Reference. 2022. Disponível em: <https://lastminuteengineers.com/esp32-pinout-reference/>.
- FAISAL, M. A. A. *et al.* Exploiting domain transformation and deep learning for hand gesture recognition using a low-cost dataglove. **Scientific Reports**, v. 12, n. 1, p. 21446, dez. 2022. ISSN 2045-2322. Publisher: Nature Publishing Group. Disponível em: <https://www.nature.com/articles/s41598-022-25108-2>.
- FAWAZ, H. I. *et al.* InceptionTime: Finding AlexNet for Time Series Classification. **Data Mining and Knowledge Discovery**, v. 34, n. 6, p. 1936–1962, nov. 2020. ISSN 1384-5810, 1573-756X. ArXiv:1909.04939 [cs, stat]. Disponível em: <http://arxiv.org/abs/1909.04939>.
- FRIDLUND, A. J.; CACIOPPO, J. T. Guidelines for human electromyographic research. **Psychophysiology**, v. 23, n. 5, p. 567–589, set. 1986. ISSN 0048-5772.
- GARCIA, M. C.; VIEIRA, T. M. M. Surface electromyography: Why, when and how to use it. **Revista andaluza de medicina del deporte**, v. 4, n. 1, p. 17–28, 2011. Publisher: Centro Andaluz de Medicina del Deporte. Disponível em: <https://www.redalyc.org/pdf/3233/323327665004.pdf>.
- GHOTKAR, A. S. *et al.* Hand gesture recognition for Indian Sign Language. *In: 2012 International Conference on Computer Communication and Informatics*. Coimbatore, India: IEEE, 2012. p. 1–4. ISBN 978-1-4577-1583-9 978-1-4577-1580-8 978-1-4577-1582-2. Disponível em: <http://ieeexplore.ieee.org/document/6158807/>.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. [S.l.]: MIT Press, 2016. <http://www.deeplearningbook.org>.

GRAAFF, K. M. V. D. Anatomia humana. *In: Anatomia humana*. [s.n.], 2003. p. 840–840. Disponível em: <https://pesquisa.bvsalud.org/portal/resource/pt/dan-3388>.

GUYTON, A. C. *et al.* **Tratado de fisiologia médica**. [S.l.]: Interamericana México, 1971. v. 2011.

HAMILL, J.; KNUTZEN, K. M. **Bases biomecânicas do movimento humano**. Manole São Paulo, 2008. Disponível em: https://www.academia.edu/download/96580223/Bases_Biomecânicas_do_Movimento_Humano_4a_edicao_2_.pdf.

HUANG, Y.; GRIFFIN, M. J. Effect of voluntary periodic muscular activity on nonlinearity in the apparent mass of the seated human body during vertical random whole-body vibration. **Journal of Sound and Vibration**, v. 298, n. 3, p. 824–840, dez. 2006. ISSN 0022-460X. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0022460X06004743>.

IBGE Pesquisa Nacional de Saúde - Tabela 8220: Pessoas com deficiência auditiva, por nível de instrução e situação do domicílio (2019). 2019. Disponível em: <https://sidra.ibge.gov.br/tabela/8220>.

IBGE Pesquisa Nacional de Saúde - Tabela 8222: Pessoas com deficiência auditiva, por condição em relação à força de trabalho na semana de referência e situação do domicílio (2019). 2019. Disponível em: <https://sidra.ibge.gov.br/tabela/8222>.

IBGE Pesquisa Nacional de Saúde - Tabela 8223: Pessoas de cinco anos ou mais de idade que referiram dificuldade permanente para ouvir, por conhecimento da Língua Brasileira de Sinais - Libras e grau de dificuldade para ouvir. 2019. Disponível em: <https://sidra.ibge.gov.br/tabela/8223>.

INVENSENSE'S. **MPU-6000 and MPU-6050 Product Specification**. 2013. Disponível em: <https://invensense.tdk.com/products/motion-tracking/6-axis/mpu-6050/>.

JANIESCH, C.; ZSCHECH, P.; HEINRICH, K. Machine learning and deep learning. **Electronic Markets**, v. 31, p. 685–695, 2021. Disponível em: <https://consensus.app/papers/machine-learning-learning-janiesch/4a48516645025c5ca1ac20aba6d5a0b6/>.

KUNAPIPAT, M. *et al.* Sensor-assisted EMG data recording system. *In: 2018 15th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*. Chiang Rai, Thailand: IEEE, 2018. p. 772–775. ISBN 978-1-5386-3555-1. Disponível em: <https://ieeexplore.ieee.org/document/8619880/>.

LEE, M.; BAE, J. Deep Learning Based Real-Time Recognition of Dynamic Finger Gestures Using a Data Glove. **IEEE Access**, v. 8, p. 219923–219933, 2020. ISSN 2169-3536. Disponível em: <https://ieeexplore.ieee.org/document/9264164/>.

LI, J. *et al.* Plastic-Optical-Fiber-Enabled Smart Glove for Machine-Learning-Based Gesture Recognition. **IEEE Transactions on Industrial Electronics**, v. 71, n. 4, p. 4252–4261, abr. 2024. ISSN 0278-0046, 1557-9948. Disponível em: <https://ieeexplore.ieee.org/document/10130769/>.

LIU, H. *et al.* **Pay Attention to MLPs**. arXiv, 2021. ArXiv:2105.08050 [cs] version: 2. Disponível em: <http://arxiv.org/abs/2105.08050>.

LUCA, C. J. D. The use of surface electromyography in biomechanics. **Journal of applied biomechanics**, v. 13, n. 2, p. 135–163, 1997. Publisher: Human Kinetics, Inc. Disponível em: <https://journals.humankinetics.com/view/journals/jab/13/2/article-p135.xml>.

MAIER, A.; SHARP, A.; VAGAPOV, Y. Comparative analysis and practical implementation of the ESP32 microcontroller module for the internet of things. *In: 2017 Internet Technologies and Applications (ITA)*. [s.n.], 2017. p. 143–148. Disponível em: <https://ieeexplore.ieee.org/document/8101926>.

MANSOUR, N. R.; FAGUNDES, D. S.; ANTUNES, M. D. *Cinesiologia e biomecânica*. [s.n.], 2019. Disponível em: <https://integrada.minhabiblioteca.com.br/reader/books/9788595028616/pageid/0>.

MENDES JUNIOR, J. J. A. **Estudo da influência das etapas de segmentação, extração de características e classificação do alfabeto em linguagem brasileira de sinais a partir de sinais eletromiográficos de superfície**. 2020. Tese (Doutorado), 2020. Publisher: Universidade Tecnológica Federal do Paraná.

MENEZES, J. E. S. A. de; FEITOSA, C. R. de S. Língua Brasileira de Sinais (LIBRAS). *História*, v. 9, p. 3, 2015. Disponível em: https://www.kufunda.net/publicdocs/Livro_Linguagem%20Brasileira%20de%20Sinais_Libras.PDF.

MERLETTI, R.; RAINOLDI, A.; FARINA, D. Surface Electromyography for Noninvasive Characterization of Muscle. *Exercise and Sport Sciences Reviews*, v. 29, n. 1, p. 20, jan. 2001. ISSN 0091-6331. Disponível em: https://journals.lww.com/acsm-essr/fulltext/2001/01000/surface_electromyography_for_noninvasive.5.aspx.

MORAIS, C. E. L. d. *et al.* **Libras**. [S.l.]: Porto Alegre: SAGAH, 2018.

OLIVEIRA, I. A. D. *et al.* Análise De Luvas Transceptoras Para Auxílio Na Tradução Bidirecional Em Língua De Sinais. *In: Pensando a educação profissional, técnica e tecnológica*. 1. ed. Editora Científica Digital, 2023. p. 172–199. ISBN 9786553603059. Disponível em: <http://www.editoracientifica.org/articles/code/220809654>.

PANDAS. **pandas-dev/pandas: Pandas**. Zenodo, 2024. Disponível em: <https://zenodo.org/records/10957263>.

PIGNATARI, R. F. *et al.* PROTOTIPAÇÃO DE UMA LUVA PARA RECONHECIMENTO DE GESTOS UTILIZANDO SENSORES INTEGRADOS E INTELIGÊNCIA ARTIFICIAL. *Revista ft*, 2023. Disponível em: <https://revistaft.com.br/prototipacao-de-uma-luva-para-reconhecimento-de-gestos-utilizando-sensores-integrados-e-inteligencia-artificial/>.

QUADROS, R. M. d.; KARNOPP, L. B. **Língua de Sinais Brasileira: Estudos Lingüísticos**. [S.l.]: Artmed Editora, 2009. Google-Books-ID: _EJvIxL7Cd0C. ISBN 978-85-363-1174-6.

RAWAT, S.; VATS, S.; KUMAR, P. Evaluating and exploring the MYO ARMBAND. *In: 2016 International Conference System Modeling & Advancement in Research Trends (SMART)*. [s.n.], 2016. p. 115–120. Disponível em: <https://ieeexplore.ieee.org/abstract/document/7894501>.

REZENDE, T. M. **Reconhecimento automático de sinais da Libras : desenvolvimento da base de dados MINDS-Libras e modelos de redes convolucionais**. jul. 2021. Tese (Doutorado), jul. 2021. Accepted: 2022-03-03T19:47:41Z Publisher: Universidade Federal de Minas Gerais. Disponível em: <https://repositorio.ufmg.br/handle/1843/39785>.

SANTOS, H. S. D.; FERREIRA, L. F.; ALVES, P. G. L. Luva Tradutora Da Língua Brasileira De Sinais. 2019.

SHINDE, P. P.; SHAH, S. A Review of Machine Learning and Deep Learning Applications. *In: 2018 Fourth International Conference on Computing Communication Control and*

Automation (ICCUBEA). [s.n.], 2018. p. 1–6. Disponível em: <https://ieeexplore.ieee.org/abstract/document/8697857>.

SICSÚ, A. L.; SAMARTINI, A.; BARTH, N. L. **Técnicas de machine learning**. [S.l.]: Editora Blucher, 2023. ISBN 9786555063974.

SILVANO, F. G. *et al.* **ALFALUVA: Luva detectora de sinais em LIBRAS para auxílio na alfabetização**. Curitiba, PR, 2016.

SILVERTHORN, D. U. **Fisiologia humana: uma abordagem integrada**. [S.l.]: Artmed editora, 2017.

SPECTRA SYMBOL. **Flex Sensor Potentiometer Series DATA SHEET**. 2014. Disponível em: <https://www.spectrasymbol.com/resources/sensor-info-downloads>.

STEVAN, J.; SILVA, R. A. **Automação e instrumentação industrial com arduino: teoria e projetos**. Sao Paulo: Erica, 2015.

SYSTEMS, E. **ESP32 Series Datasheet v4.5**. 2024. Disponível em: <https://www.espressif.com/en/products/socs/esp32>.

ThalmicLabs. **Tech Specs | Myo Battery Life, Dimensions, Compatibility, and More**. 2018. Disponível em: <https://web.archive.org/web/20180721191357/https://www.myo.com/techspecs>.

THOMAZINI, D.; ALBUQUERQUE, P. U. B. D. **Sensores industriais : Fundamentos e aplicações**. [S.l.]: Saraiva Educação S.A., 2020. Google-Books-ID: 1qgPEAAAQBAJ. ISBN 978-85-365-3324-7.

VABALAS, A. *et al.* Machine learning algorithm validation with a limited sample size. **PLOS ONE**, v. 14, n. 11, p. e0224365, nov. 2019. ISSN 1932-6203. Publisher: Public Library of Science. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0224365>.

VANPUTTE, C.; REGAN, J.; RUSSO, A. **Anatomia e Fisiologia de Seeley-10ª Edição**. McGraw Hill Brasil, 2016. Disponível em: https://books.google.com/books?hl=pt-BR&lr=&id=_vW0DAAAQBAJ&oi=fnd&pg=PR1&dq=Anatomia+e+Fisiologia+de+Seeley+Cinnamon+Vanputte%3B+Reganm+Jennifer%3B+Andre&ots=cJSV9IZwME&sig=RqH-siWNMs2K0gejZxbpHbUFTAM.

VISCONTI, P. *et al.* Technical Features and Functionalities of Myo Armband: An Overview on Related Literature and Advanced Applications of Myoelectric Armbands Mainly Focused on Arm Prostheses. **International Journal on Smart Sensing and Intelligent Systems**, v. 11, n. 1, p. 1–25, jan. 2018. Disponível em: <https://sciendo.com/article/10.21307/ijssis-2018-005>.

WENDLING, M. Sensores. **Universidade Estadual Paulista. São Paulo**, v. 2010, p. 20, 2010. Disponível em: <https://www.academia.edu/download/55868161/4---sensores-v2.0.pdf>.

ZHOLSHIYEVA, L. *et al.* Human-machine interactions based on hand gesture recognition using deep learning methods. **International Journal of Electrical and Computer Engineering (IJECE)**, v. 14, n. 1, p. 741–748, fev. 2024. ISSN 2722-2578. Number: 1. Disponível em: <https://ijece.iaescore.com/index.php/IJECE/article/view/33858>.

ŠKRABA, A. *et al.* Prototype of Group Heart Rate Monitoring with ESP32. *In: 2019 8th Mediterranean Conference on Embedded Computing (MECO)*. [s.n.], 2019. p. 1–4. ISSN: 2637-9511. Disponível em: <https://ieeexplore.ieee.org/abstract/document/8760150>.