



UNIVERSIDADE FEDERAL DE UBERLÂNDIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA BIOMÉDICA

EBER LAWRENCE SOUZA GOUVEIA

SISTEMA SEMI-AUTÔNOMO BASEADO EM VISÃO  
NEUROMÓRFICA PARA CONTROLE DE  
PRÓTESES MIOELÉTRICAS

Uberlândia

2022

EBER LAWERENCE SOUZA GOUVEIA

SISTEMA SEMI-AUTÔNOMO BASEADO EM VISÃO  
NEUROMÓRFICA PARA CONTROLE DE  
PRÓTESES MIOELÉTRICAS

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Biomédica da Universidade Federal de Uberlândia, como parte dos requisitos para obtenção do grau de Mestre em Ciências.

Área de concentração: Engenharia Biomédica  
Linha de pesquisa: Engenharia de reabilitação e tecnologias assistivas

Orientador: Alcimar Barbosa Soares

---

Assinatura do orientador

Uberlândia

2022

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU  
com dados informados pelo(a) próprio(a) autor(a).

G719 Gouveia, Eber Lawrence Souza, 1995-  
2022 Sistema Semi-autônomo Baseado em Visão Neuromórfica  
para Controle de Próteses Mioelétricas [recurso  
eletrônico] / Eber Lawrence Souza Gouveia. - 2022.

Orientador: Alcimar Barbosa Soares.

Dissertação (Mestrado) - Universidade Federal de  
Uberlândia, Pós-graduação em Engenharia Biomédica.

Modo de acesso: Internet.

Disponível em: <http://doi.org/10.14393/ufu.di.2021.112>

Inclui bibliografia.

Inclui ilustrações.

1. Engenharia biomédica. I. Soares, Alcimar Barbosa,  
1965-, (Orient.). II. Universidade Federal de  
Uberlândia. Pós-graduação em Engenharia Biomédica. III.  
Título.

CDU: 62:61

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:  
Gizele Cristine Nunes do Couto - CRB6/2091  
Nelson Marcos Ferreira - CRB6/3074


**UNIVERSIDADE FEDERAL DE UBERLÂNDIA**

Coordenação do Programa de Pós-Graduação em Engenharia Biomédica  
 Av. João Naves de Ávila, 2121, Bloco 3N, Sala 115 - Bairro Santa Mônica, Uberlândia-MG, CEP 38400-902  
 Telefone: (34) 3239-4761 - www.ppgeb.feelt.ufu.br - ppegb@feelt.ufu.br


**ATA DE DEFESA - PÓS-GRADUAÇÃO**

Programa de Pós-Graduação em:	Engenharia Biomédica				
Defesa de:	Dissertação de Mestrado Acadêmico, 080, PPGEB				
Data:	vinte e nove de março de dois mil e vinte e dois	Hora de início:	14:10	Hora de encerramento:	16:15
Matrícula do Discente:	11912EBI003				
Nome do Discente:	Eber Lawrence Souza Gouveia				
Título do Trabalho:	Sistema semi-autônomo baseado em visão neuromórfica para controle de próteses mioelétricas				
Área de concentração:	Engenharia Biomédica				
Linha de pesquisa:	Engenharia de Reabilitação e Tecnologias Assistivas				
Projeto de Pesquisa de vinculação:	Neuromorphic Feedback. A Strategy to Enhance Prosthesis Embodiment and Performance				

Reuniu-se em web conferência pela plataforma Conferência Web em conformidade com a PORTARIA Nº36, DE 19 DE MARÇO DE 2020 da COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR - CAPES, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Engenharia Biomédica, assim composta: Professores Doutores: Adson Ferreira da Rocha - UnB; Márcio José da Cunha - FEELT/UFU; Alcimar Barbosa Soares - PPGEB/UFU orientador do candidato.

Iniciando os trabalhos o presidente da mesa, Dr. Alcimar Barbosa Soares, apresentou a Comissão Examinadora e o candidato, agradeceu a presença do público, e concedeu ao Discente a palavra para a exposição do seu trabalho. A duração da apresentação do Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor(a) presidente concedeu a palavra, pela ordem sucessivamente, aos(às) examinadores(as), que passaram a arguir o(a) candidato(a). Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o(a) candidato(a):

Aprovado.

Esta defesa faz parte dos requisitos necessários à obtenção do título de Mestre.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por **Marcio José da Cunha, Professor(a) do Magistério Superior**, em 29/03/2022, às 16:12, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Alcimar Barbosa Soares, Professor(a) do Magistério Superior**, em 29/03/2022, às 16:12, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Adson Ferreira da Rocha, Usuário Externo**, em 29/03/2022, às 16:12, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site [https://www.sei.ufu.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **3471393** e o código CRC **52BA2A33**.

SISTEMA SEMI-AUTÔNOMO BASEADO EM VISÃO  
NEUROMÓRFICA PARA CONTROLE DE  
PRÓTESES MIOELÉTRICAS

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Biomédica da Universidade Federal de Uberlândia, como parte dos requisitos para obtenção do grau de Mestre em Ciências.

Área de concentração: Engenharia Biomédica  
Linha de pesquisa: Engenharia de reabilitação e tecnologias assistivas

Orientador: Alcimar Barbosa Soares

---

Assinatura do orientador

Uberlândia, 29 de março de 2022

Banca Examinadora:

---

Alcimar Barbosa Soares – PhD (UFU)

---

Adson Ferreira Da Rocha – Dr (UnB)

---

Márcio José da Cunha – Dr (UFU)

Dedico este trabalho aos meus pais, pelo estímulo,  
carinho e compreensão.

## AGRADECIMENTOS

Aos meus pais Eber Pereira de Souza e Sandra Gouveia de Faria pelo apoio durante todo o meu percurso acadêmico e pessoal, estando sempre ao meu lado e prontos para me ajudar nos momentos em que mais precisei.

Em especial ao meu pai, que faleceu no dia 20 de junho de 2021 devido a complicações causadas pela COVID-19. Infelizmente ele não pode estar comigo até o fim desse trabalho, mas sem dúvidas esse trabalho não seria possível sem todo apoio e carinho dele durante todo o meu percurso até aqui. Embora a vida na terra seja passageira, seus ensinamentos permaneceram comigo eternamente.

A minha esposa Adrielle Nazar Moraes que mesmo nos momentos mais difíceis esteve ao meu lado, sendo o meu porto seguro e me motivando a seguir sempre em frente.

Ao Professor Dr. Alcimar Barbosa Soares pelo incentivo e orientação neste trabalho.

Ao laboratório de Engenharia Biomédica (BioLab) por fornecer o laboratório para pesquisas e experimentos durante todo o período de execução desse trabalho.

Aos meus companheiros de laboratório que contribuíram diretamente ou indiretamente para a realização desse trabalho. Agradeço também pelas inúmeras conversas filosóficas durante as madrugadas, que ajudaram o meu desenvolvimento tanto profissional quanto pessoal. Em especial ao Eduardo Borges Gouveia que foi meu companheiro de pesquisa durante todo o período de mestrado, contribuindo diretamente no desenvolvimento desse trabalho.

Aos voluntários que se disponibilizaram a participar dos experimentos propostos sem nenhum ganho financeiro.

Aos professores da Faculdade de Engenharia Elétrica (FEELT) por proporcionar conhecimentos em diversas áreas da engenharia.

A Universidade Federal de Uberlândia (UFU) por me proporcionar a oportunidade de graduação no curso de Engenharia Biomédica.

Ao CNPq, CAPES e FAPEMIG pelo apoio financeiro durante todo o período de execução desse projeto de pesquisa.

“If I have seen further, it is by standing  
on the shoulders of Giants.”

(Issac Newton)

## RESUMO

Nos últimos anos, diversos trabalhos têm buscado melhores formas de realizar o controle de próteses de membro superior. Porém, os métodos atuais de controle de próteses ainda estão longe de atingir as características do membro humano, como controlar diversos graus de liberdade de forma natural e automática. Além disso, os sistemas de controle baseados na abordagem tradicional de Controle Direto não podem fornecer muitos tipos de preensão, pois a complexidade de controle aumenta exponencialmente e, conseqüentemente, aumenta a carga cognitiva necessária do usuário. Observando tais limitações, este trabalho apresenta um sistema de controle semiautônomo, que utiliza uma câmera neuromórfica acoplada à prótese para selecionar o tipo de preensão e corrigir o ângulo do punho da prótese de acordo com a forma e ângulo do objeto almejado, respectivamente. Além disso, o sistema pode segmentar o objeto almejado dos demais contidos em uma única cena. Por fim, foram realizados dois experimentos *pick-and-place* com voluntários, sendo o primeiro utilizando o sistema de controle semiautônomo proposto e o segundo com um sistema de controle manual, baseado na comutação de movimentos por meio de blocos. Além disso, foi utilizado o questionário NASA-TLX para verificar o nível de carga de trabalho durante os experimentos de forma subjetiva.

**Palavras-chave:** Controle de próteses. Visão neuromórfica. Sistema de controle semiautônomo.

## ABSTRACT

In recent years, several studies have been looking for better ways for controlling upper-limb prostheses. However, current prosthesis control methods are still far from reaching the characteristics of the human limb, such as handling different degrees of freedom naturally and automatically. Furthermore, control systems based on the traditional Direct Control approach cannot provide many grasp types as the control complexity increases exponentially and consequently increases the required cognitive load of the user. Observing these limitations, this work presents a semi-autonomous control system, which uses a neuromorphic camera attached to the prosthesis for selecting the type of grasp and correct the angle of the prosthesis wrist according to the shape and angle of the desired object, respectively. Furthermore, the system can segment the desired object from the others contained in a single scene. Finally, two pick-and-place experiments were carried out with volunteers, the first using the proposed semi-autonomous control system and the second using a manual control system, based on switching movements through blocks. In addition, the NASA-TLX questionnaire was used to subjectively verify the level of workload during the experiments.

**Keywords:** Prosthesis control. Neuromorphic vision. Semi-autonomous control system.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Representação do experimento de Luigi Galvani em 1791. Quando aplicado um potencial em cima das pernas de uma rã, ocorria sua contração. ....	8
Figura 2 – Ilustração da estrutura fisiológica do sistema neuromotor. É possível ver as diversas unidades motoras, onde cada fibra muscular está conectada somente um neurônio motor. Entretanto, um único neurônio motor pode se conectar a diversas fibras musculares. ....	10
Figura 3 – Exemplos de posições no posicionamento dos eletrodos de eletromiografia de superfície. A forma do sinal e seu espectro de frequências variam de acordo com o lugar escolhido. ....	11
Figura 4 – Diagrama simplificado de um sistema de eletromiografia de superfície. ....	12
Figura 5 – Ilustração de um modelo de próteses acionado por tirantes. No exemplo em questão, a prótese usada consistia em uma garra mecânica. ....	17
Figura 6 – Ilustração de um modelo clássico de próteses eletrônicas. Nesse exemplo, o sistema se baseia no uso de eletromiografia de superfície para realizar o controle. ....	18
Figura 7 – Diagrama de blocos para ambos os tipos de controle (a) CD e (b) PR. ....	20
Figura 8 – Ilustração dos principais pontos em comum entre uma câmera convencional e o olho humano. ....	22
Figura 9 – Imagens obtidas a partir de oito diferentes tempos de exposição. ....	23
Figura 10 – Características do sinal obtido por cada um dos sensores contidos na matriz de sensores. (a) Resposta de tensão para um único sensor de uma câmera baseada em eventos. A amplitude do sinal é diretamente proporcional a luminosidade recebida. (b) Eventos gerados à medida que o sinal de tensão ultrapassa os limiares de ativação. ....	25
Figura 11 – Comparação entre as imagens de saída de uma câmera convencional e uma câmera baseada em eventos. (a) Reconstrução de uma imagem a partir de eventos obtidos por uma câmera DVS128. Representação dos frames e eventos de saída gerados por uma (b) câmera convencional e (c) câmera baseada em eventos ao longo do tempo. ....	26
Figura 12 – Exemplos de aplicações de câmeras baseadas em eventos. (a) Robotic Goalie desenvolvido para interceptar bolas em alta velocidade (3 ms) indo em direção ao gol. (b) Sistema de reconstrução de vídeos e imagens a partir de eventos obtidos de uma câmera DVS da Samsung. A imagem mostra três colunas de três exemplos dos eventos obtidos da câmera (esquerda), da reconstrução realizada pelo sistema desenvolvido (centro) e a imagem obtida	

por meio de uma câmera convencional (direita). (c) Sistema semiautônomo para controle de próteses com o auxílio de uma câmera baseada em eventos. (d) Robô com rápida resposta no controle do equilíbrio de um lápis a partir de visão estéreo. ....	28
Figura 13 – Comparação entre um (a) neurônio e um (b) perceptron. ....	30
Figura 14 – Estrutura simplificada de uma CNN. (a) Imagem de entrada, (b) camada Convolutacional, (c) camada de Pooling, (d) camada Fully-Connected e (e) saída da rede. ....	32
Figura 15 – Transformação de conjunto de dados usando PCA. (a) Conjunto de dados original contendo três dimensões. (b) Conjunto de dados obtidos a partir das duas PCs mais relevantes. ....	33
Figura 16 – Obtenção das PCs de um relógio posicionado em três orientações diferentes. A linha verde indica a PC 1 e a linha amarela representa a PC 2. O relógio foi posicionado em: (a) -72 (b) -45 e (c) 58 graus de acordo com a PC 1. ....	34
Figura 17 – Exemplo do uso da PCA para obtenção da orientação de múltiplos objetos. (a) e (b) câmera convencional. (c) e (d) câmera baseada em eventos. ....	35
Figura 18 – Exemplificação de uma paisagem topográfica com diferentes níveis e bacias de água. ....	37
Figura 19 – (a) Intensidade dos pixels de uma imagem ao longo do eixo X. (b) segmentação das áreas de acordo com a sua intensidade. À medida que o limiar sobe, novas barreiras são criadas. ....	37
Figura 20 – Câmera neuromórfica DVS128 desenvolvida pela empresa Inivation. ....	40
Figura 21 – Três imagens de uma mesma gravação, todavia cada uma com um número de FPS diferente. (a) 1000 FPS, (b) 100 FPS e (c) 10 FPS. ....	44
Figura 22 – Ambiente de gravação dos objetos. No exemplo em questão, foi realizada a gravação de uma laranja. ....	46
Figura 23 – Imagens obtidas por meio do processo de conversão de eventos desenvolvido. (a) Exemplos de objetos classificados como Tripod. (b) Exemplos de objetos classificados como Power. ....	47
Figura 24 – Diagrama simplificado do algoritmo de obtenção do tipo de preensão e ângulo do objeto. ....	48
Figura 25 – Diagrama expandido da etapa de segmentação do objeto. ....	49
Figura 26 – Diagrama expandido da etapa de classificação do tipo de preensão. ....	52

Figura 27 – Comparação entre os métodos de ajuste dos eixos na imagem segmentada de uma chave. (a) Imagem segmentada, (b) ajuste sem verificar o tamanho dos eixos, causando a distorção do objeto e (c) método de ajuste dos eixos da imagem proposto, evitando que ocorra a distorção do objeto. ....	53
Figura 28 – Modelo de CNN usado para classificar o tipo de preensão necessária de acordo com o formato do objeto contido na imagem. ....	55
Figura 29 – Diagrama expandido da etapa de obtenção da orientação do objeto.....	57
Figura 30 – Referência para o cálculo do ângulo do objeto. O vetor de referência é sempre na vertical, enquanto o vetor da PC1 pode estar em qualquer direção dentro da faixa de -90 até +90 graus. ....	59
Figura 31 – Diagrama das principais etapas do sistema de controle semiautônomo.....	61
Figura 32 – Fluxograma do sistema de controle semiautônomo. ....	62
Figura 33 – Prótese de mão direita Bebionic usada nos experimentos contidos nesse trabalho. ....	62
Figura 34 – Circuito de aquisição e condicionamento de sinal de sEMG usado nesse trabalho. O circuito possui um canal e seu amplificador de instrumentação é o INA118.....	63
Figura 35 – Adaptação do disco de encoder na prótese Bebionic. (a) Disco de encoder impresso na impressora 3D. (b) Visão inferior do encoder adaptado na prótese. (c) Visão superior/frontal do disco de encoder adaptado na prótese.....	66
Figura 36 - Diagrama das principais etapas do sistema de controle manual. (a) Aquisição do sinal de sEMG do usuário. (b) Extração de informações da envoltória do sinal de sEMG para efetuar o controle do sistema. (c) Bloco 1 do sistema, o qual possui as funções de rotação do punho da prótese. (d) Bloco 2 do sistema, o qual possui as funções relacionadas a preensão do objeto. (e) Prótese de mão Bebionic que será controlada. ....	67
Figura 37 - Fluxograma do sistema de controle manual.....	68
Figura 38 – Ilustração do ambiente de coleta durante a realização dos experimentos utilizando os sistemas de controle (a) semiautônomo e (c) manual. ....	71
Figura 39 – Evolução do modelo durante as 10 épocas do período de treinamento. As linhas contínuas indicam o desempenho do conjunto de treinamento, enquanto as tracejadas indicam o conjunto de teste para cada uma das partes da validação cruzada. (a) Acurácia por época e (b) Perda por época. ....	76

Figura 40 – Evolução do modelo durante as 30 épocas do período de treinamento. (a) Acurácia por época durante o treinamento (azul) e teste (vermelho). (b) Perda por época durante o treinamento (azul) e teste (vermelho). .....	77
Figura 41- Erro resultante do processo de validação do algoritmo de obtenção do ângulo. (a) Erro do ângulo obtido dos 30 objetos agrupados (a) pelas classes preensão e (b) pela angulação correta. ....	77
Figura 42 - Erro resultante do processo de validação do algoritmo de obtenção do ângulo das 7 angulações base agrupadas para cada um dos objetos pertencentes a classe (a) Tripod e (b) Power. ....	78
Figura 43 – Porcentagem de acertos e erros do classificador para as suas etapas do experimento usando o sistema de controle semiautônomo. Tarefa de preensão realizada (a) com mesmos objetos usados no treinamento do modelo (grupo 1 de objetos) e (b) com objetos nunca vistos na etapa de treinamento do modelo (grupo 2 de objetos). ....	79
Figura 44 – Ângulo esperado contra ângulo obtido, apresentando a reta ideal (cinza) onde o erro é igual a zero. Amostras obtidas dos usuários que usaram o sistema de controle (a) semiautônomo e (b) manual.....	80
Figura 45 – Erro do ângulo obtido para ambos os experimentos realizados usando o sistema de controle (a) semiautônomo e (b) manual. ....	81
Figura 46 – Tempo de execução da tarefa de preensão para ambos os experimentos realizados. (a) Todos os tempos de execução por sistema de controle. (b) Tempo execução por voluntário usando o sistema de controle semiautônomo. (b) Tempo execução por voluntário usando o sistema de controle manual.....	81
Figura 47 – Tempo de execução da tarefa de preensão para cada uma das 30 repetições realizadas. Todas as repetições executadas pelos voluntários do sistema de controle (a) semiautônomo e (b) manual.....	82
Figura 48 - Erro no ajuste do ângulo do punho pelo tempo de execução de tarefa para todas as repetições realizadas, apresentando a distribuição dos dados tanto no eixo vertical, quanto no eixo horizontal. ....	83
Figura 49 – Porcentagem de escorregamento dos objetos durante a transferência de posição do objeto. ....	83
Figura 50 – Média e desvio padrão das pontuações para cada fator de carga de trabalho obtidas por meio do questionário NASA-TLX. ....	84

Figura 51 – Comparação entre os dois sistemas de controle por meio da análise subjetiva dos 10 voluntários que participaram dos experimentos. \* significa diferença estatística entre os grupos ( $p = 0,05$ ).....85

## LISTA DE TABELAS

Tabela 1 – Divisão dos bits que contém as informações essenciais dos dados gravados pela câmara DVS128.....	41
Tabela 2 – Decodificação dos dados de endereço e polaridade contidos em um arquivo de gravação da DVS128. Na coluna de Posição, o “X” representa as posições onde as respectivas informações estão armazenadas.....	41
Tabela 3 – Acurácia e perda obtidas em cada uma das 5 partes da validação cruzada. ....	76

## LISTA DE ABREVIATURAS E SIGLAS

AM	Aprendizado de Máquina
ANN	Redes Neurais Artificiais – <i>Artificial Neural Network</i>
CD	Controle Direto
CNN	Rede Neural Convolucional – <i>Convolutional Neural Network</i>
DAVIS	Sensor de Visão de Pixel Dinâmico e Ativo – <i>Dynamic and Active Pixel Vision Sensor</i>
DoF	Graus de Liberdade – <i>Degrees of freedom</i>
DVS	Sensor de visão dinâmico – <i>Dynamic Vision Sensor</i>
EMG	Eletromiografia
FPS	Frames por segundo
LDA	Análise de Discriminante Linear – <i>Linear Discriminant Analysis</i>
MB	Desfoque de movimento – <i>Motion Blur</i>
MLP	Perceptron de multicamadas – <i>Multilayer Perceptron</i>
MUAP	Potencial de Ação da Unidade Motora - <i>Motor Unit Action Potential</i>
NN	Redes Neurais – <i>Neural Networks</i>
PC	Componente Principal – <i>Principal Component</i>
PCA	Análise de Componentes Principais – <i>Principal Component Analysis</i>
RNN	Redes Neurais Recorrentes – <i>Recurrent Neural Network</i>
RP	Reconhecimento de Padrões
sEMG	Eletromiografia de Superfície
SUS	Sistema Único de Saúde
SVM	Máquina de Vetores de Suporte – <i>Support Vector Machine</i>
UDP	Protocolo de datagrama do usuário – <i>User Datagram Protocol</i>

## SUMÁRIO

<b>1. INTRODUÇÃO.....</b>	<b>1</b>
<b>1.1 OBJETIVOS .....</b>	<b>4</b>
<b>1.2 ESTRUTURA DA DISSERTAÇÃO .....</b>	<b>5</b>
<b>2. ELETROMIOGRAFIA DE SUPERFÍCIE.....</b>	<b>7</b>
<b>2.1 HISTÓRIA .....</b>	<b>7</b>
<b>2.2 O SINAL ELETROMIOGRÁFICO .....</b>	<b>9</b>
2.2.1 ORIGEM FISIOLÓGICA .....	9
2.2.2 AQUISIÇÃO DE SINAIS .....	10
<b>3. PRÓTESES DE MEMBROS SUPERIORES .....</b>	<b>14</b>
<b>3.1 HISTÓRIA .....</b>	<b>14</b>
<b>3.2 TIPOS DE PRÓTESES DE MEMBROS SUPERIORES.....</b>	<b>16</b>
3.2.1 PASSIVAS .....	16
3.2.2 ATIVAS.....	17
<b>4. VISÃO BASEADA EM EVENTOS .....</b>	<b>22</b>
<b>4.1 CÂMERAS NEUROMÓRFICAS VERSUS CÂMERAS CONVENCIONAIS ...</b>	<b>24</b>
<b>4.2 PRINCIPAIS APLICAÇÕES CONTIDAS NA LITERATURA .....</b>	<b>27</b>
<b>5. VISÃO COMPUTACIONAL .....</b>	<b>29</b>
<b>5.1 REDES NEURAIAS CONVOLUCIONAIS .....</b>	<b>29</b>
5.1.1 ESTRUTURA DE UMA CNN.....	31
<b>5.2 ANÁLISE DE COMPONENTES PRINCIPAIS .....</b>	<b>33</b>
5.2.1 VISÃO COMPUTACIONAL E PCA: ORIENTAÇÃO DE OBJETOS.....	34
<b>5.3 SEGMENTAÇÃO DE IMAGENS – ALGORITIMO DE WATERSHED .....</b>	<b>36</b>
<b>6. METODOLOGIA.....</b>	<b>39</b>
<b>6.1 EXPERIMENTOS OFFLINE .....</b>	<b>39</b>
6.1.1 AQUISIÇÃO E CONDICIONAMENTO DE DADOS .....	39
6.1.1.1 <i>Dynamic Vision Sensor 128</i> .....	40
6.1.1.2 <i>Aquisição de eventos</i> .....	40
6.1.1.2.1 <b>Offline</b> .....	41
6.1.1.2.2 <b>Online</b> .....	42
6.1.1.3 <i>Conversão de eventos para quadros</i> .....	42
6.1.2 CONJUNTO DE DADOS .....	44

6.1.2.1	<i>Seleção de objetos e definição dos tipos de apreensão</i> .....	45
6.1.2.2	<i>Gravação do conjunto de objetos</i> .....	45
6.1.2.3	<i>Geração do conjunto de imagens</i> .....	47
6.1.3	<b>PROCESSAMENTO DE IMAGENS</b> .....	48
6.1.3.1	<i>Segmentação do objeto almejado</i> .....	48
6.1.3.1.1	<b>Binarização da imagem</b> .....	49
6.1.3.1.2	<b>Segmentação dos objetos contidos na cena</b> .....	50
6.1.3.1.3	<b>Seleção do objeto</b> .....	51
6.1.3.1.4	<b>Definição da região de interesse</b> .....	51
6.1.3.1.5	<b>Conversão da nuvem de pontos para imagem</b> .....	51
6.1.3.2	<i>CLASSIFICAÇÃO DO TIPO DE APREENSÃO</i> .....	52
6.1.3.2.1	<b>Ajuste dos eixos</b> .....	53
6.1.3.2.2	<b>Redimensionamento da imagem</b> .....	54
6.1.3.2.3	<b>Extração de atributos e classificação</b> .....	55
6.1.3.2.4	<b>Validação cruzada</b> .....	56
6.1.3.2.5	<b>Treinamento do modelo</b> .....	56
6.1.3.3	<i>Obtenção da orientação do objeto</i> .....	57
6.1.3.3.1	<b>Obtenção do autovetor, autovalor e centro de massa do objeto</b> .....	57
6.1.3.3.2	<b>Componentes principal e secundária</b> .....	58
6.1.3.3.3	<b>Cálculo do ângulo</b> .....	58
6.1.3.3.4	<b>Validação</b> .....	60
6.2	<b>EXPERIMENTOS EM TEMPO REAL</b> .....	60
6.2.1	<b>O SISTEMA DE CONTROLE SEMIAUTÔNOMO</b> .....	61
6.2.1.1	<i>Estrutura de controle e componentes do sistema</i> .....	61
6.2.1.1.1	<b>Prótese de mão</b> .....	62
6.2.1.1.2	<b>Intenção de movimento</b> .....	63
6.2.1.1.3	<b>Fases automatizadas</b> .....	64
6.2.2	<b>SISTEMA DE CONTROLE MANUAL</b> .....	67
6.2.2.1	<i>Estrutura de controle e componentes do sistema</i> .....	67
6.2.3	<b>EXPERIMENTO COM VOLUNTÁRIOS</b> .....	69
6.2.3.1	<i>Grupo amostral</i> .....	70
6.2.3.2	<i>Ambiente de coleta</i> .....	70
6.2.3.3	<i>Tarefa</i> .....	71

6.2.3.4	<i>Script de escolha de objetos</i> .....	72
6.2.3.5	<i>Questionário NASA-TLX</i> .....	73
<b>7.</b>	<b>RESULTADOS</b> .....	<b>75</b>
<b>7.1</b>	<b>EXPERIMENTOS OFFLINE</b> .....	<b>75</b>
7.1.1	VALIDAÇÃO CRUZADA DO MODELO DE CNN.....	75
7.1.2	VALIDAÇÃO DO ALGORITMO DE OBTENÇÃO DO ÂNGULO .....	77
<b>7.2</b>	<b>EXPERIMENTOS EM TEMPO REAL</b> .....	<b>78</b>
7.2.1	OBJETOS APRENDIDOS X OBJETOS NÃO APRENDIDOS .....	79
7.2.2	COMPARAÇÃO ENTRE OS SISTEMAS DE CONTROLE .....	79
7.2.2.1	<i>Ajuste do ângulo do punho</i> .....	80
7.2.2.2	<i>Tempo de execução</i> .....	81
7.2.2.3	<i>Escorregamento do objeto</i> .....	83
7.2.2.4	<i>Análise subjetiva – Questionário NASA-TLX</i> .....	84
<b>8.</b>	<b>DISCUSSÃO</b> .....	<b>86</b>
<b>8.1</b>	<b>EXPERIMENTOS OFFLINE</b> .....	<b>86</b>
<b>8.2</b>	<b>EXPERIMENTOS EM TEMPO REAL</b> .....	<b>88</b>
<b>8.3</b>	<b>COMPARAÇÕES COM O ESTADO DA ARTE</b> .....	<b>92</b>
<b>9.</b>	<b>CONCLUSÃO</b> .....	<b>95</b>
<b>9.1</b>	<b>TRABALHOS FUTUROS</b> .....	<b>96</b>

# Capítulo 1

---

## 1. INTRODUÇÃO

A mão humana é uma ferramenta de grande importância, pois permite ao ser humano realizar tarefas que exigem tanto força quanto precisão. Isso é possível graças à quantidade de Graus de Liberdade (DoF) que ela possui (21 DoF para a mão e 6 DoF para o punho) e ao posicionamento oposto do polegar em relação aos demais dedos (Cordella et al., 2016). Dessa forma, o ser humano é capaz de criar e manusear diferentes ferramentas essenciais para facilitar a execução de atividades complexas relacionadas ao dia a dia. Além disso, vários trabalhos enfatizam a importância das mãos humanas para a comunicação e interação social, de forma que elas são responsáveis por transmitir uma mensagem mesmo que de forma não verbal (Biddiss & Chau, 2007; Cordella et al., 2016; Lindner et al., 2010; Murray, 2005). Dessa forma, ao sofrer uma amputação da mão, o sujeito é impossibilitado de executar tarefas importantes, impactando diretamente suas atividades diárias, além de limitar a sua capacidade de trabalho e de comunicação.

As amputações de membros superiores são uma fatalidade que afeta diferentes tipos de pessoas, com causas variadas ao redor do mundo (Esquenazi, 2004; Gaine et al., 1997; Peixoto et al., 2017). A maioria das amputações em países em desenvolvimento são causadas por traumas resultantes de fraturas tratadas inadequadamente, acidentes de trabalho causados por máquinas motorizadas e acidentes de trânsito. No caso de países com conflitos recorrentes e recente história de guerra, traumas decorrentes de armas de fogo são as principais causas de amputações. Já nos países desenvolvidos, as complicações vasculares do diabetes são a principal causa das amputações, que podem ser agravadas pelo uso do tabaco (Bastos-Filho et al., 2014; Esquenazi, 2004; Gaine et al., 1997; Peixoto et al., 2017; Resnik et al., 2012; Wheaton, 2017).

De acordo com o banco de dados do Sistema Único de Saúde (SUS), cerca de 1480 novos casos de amputações de membros superiores são registrados por ano no Brasil. Contudo, os números totais podem ser ainda maiores, levando-se em conta que esses dados incluem apenas casos registrados pelo SUS (DATASUS, 2020). De fato, esses dados indicam uma necessidade de amparar as pessoas que sofreram algum tipo de amputação. Nesse contexto, o uso de próteses têm sido uma excelente abordagem para substituir o membro perdido, recuperando a destreza

ao realizar tarefas do dia a dia, e, conseqüentemente, aumentando a capacidade de interação social e autoestima do amputado.

Dada a demanda por recuperar as funções de membros perdidos, as próteses de membro superior sofreram diversas modificações e possuem uma longa história de desenvolvimento. Os primeiros modelos de prótese de membro superior consistiam em um gancho atuado por uma combinação de cordas operadas pelo usuário (Billock, 1986). Entretanto, o avanço da tecnologia nos últimos anos possibilitou embarcar motores diretamente na mão robótica, principalmente graças ao desenvolvimento de motores com alto torque e tamanhos cada vez menores. Além disso, a produção de novos tipos de materiais, possibilitou a construção de modelos cada vez mais leves e parecidos com a mão humana. Atualmente, existem diversos tipos de próteses de membros superiores, que variam desde modelos de baixo custo, produzidas principalmente com o auxílio de impressoras 3D (ten Kate et al., 2017), até modelos sofisticados, produzidos por empresas especializadas no desenvolvimento de dispositivos médicos, como a Ottobock (Ottobock, 2020) e Ossur (Össur Corporate, 2020).

Os modelos de próteses comerciais atuais são mecanicamente sofisticados, capazes de reproduzir os principais movimentos da mão humana. Apesar de não ser usada como prótese ainda, o modelo de mão robótica Shadow Hand desenvolvida pela Shadow Robot Company tem chamado atenção pelo número de DoF que possui (20 DoF), de forma que ela vem sendo utilizada para realizar tarefas de alta complexidade, como o manuseamento de objetos (ROS Components, 2020). Já os modelos Bebionic da Ottobock e i-Limb da Ossur, podem ser considerados os mais avançados, mesmo com um número de DoF reduzido em comparação com a Shadow Hand, já que seu uso é aplicado diretamente em próteses. Entretanto, mesmo com o avanço no sistema mecânico das próteses, um dos principais problemas enfrentados é a dificuldade de prover a interface entre a prótese e o sistema nervoso do amputado. Dessa forma, o usuário não consegue controlar todos os DoF da prótese, o que implica na falta de domínio sobre o membro biônico. Além disso, quanto maior o número de DoF a serem controlados, maior é a complexidade do sistema de controle, tornando o uso da prótese menos natural, cansativo e desgastante. Conseqüentemente, pode levar o amputado a desistir do uso desse tipo de prótese (Resnik et al., 2018; White et al., 2017).

Com o intuito de reduzir a complexidade no controle da prótese, os modelos comerciais, em sua maioria, são baseados no método de Controle Direto (CD). Entretanto, esse tipo de controle é pouco natural, pois utiliza apenas dois canais de Eletromiografia de superfície (sEMG), normalmente antagonistas, que controlam as duas direções dos motores da prótese. Além disso, esse método funciona a partir do chaveamento das preensões pré-programadas da

prótese, de forma que o usuário precisa realizar co-contrações subsequentes para chegar no tipo de preensão desejada (White et al., 2017). Para diminuir os desafios encontrados por usuários desse tipo de controle, alternativas foram desenvolvidas. Um exemplo que merece destaque é o método baseado em Reconhecimento de Padrões (RP), cuja popularidade aumentou significativamente nos últimos anos devido ao constante avanço nas áreas de Aprendizado de Máquina (AM) e Eletrônica Embarcada.

Em 2014, a empresa norte americana COAPT lançou o primeiro modelo comercial baseado em RP, fornecendo uma quebra de paradigma no controle de próteses de mão (Coapt, 2020). Entretanto, mesmo sendo o mais promissor, o modelo de RP ainda possui algumas limitações, como a falta de robustez gerada pela constante alteração nos padrões do sinal de sEMG, o que pode resultar em um mau funcionamento do sistema (Beaulieu et al., 2017). Nesse contexto, têm sido pesquisadas outras formas de amenizar os problemas relacionados à configuração espacial dos membros, combinando próteses mioelétricas com sensores inerciais (Hays et al., 2019; Osborn et al., 2016; Resnik et al., 2018). Porém, é necessário diminuir a complexidade do controle, principalmente porque existem muitos graus de liberdade para serem controlados. Pensando nisso, alguns trabalhos recentes trouxeram uma nova abordagem que se baseia em adicionar visão computacional no sistema de controle, buscando uma melhor relação entre o número de DoF e complexidade de controle (Došen et al., 2010; Došen & Popović, 2011; Ghazaei et al., 2017; Hays et al., 2019; Markovic et al., 2014, 2015).

Em 2010, Dosen et al introduziram a ideia da prótese de membro superior combinada com a visão computacional para automatizar a escolha da preensão (Došen et al., 2010). Nesse caso, uma câmera fotográfica foi colocada na mão protética e sua ativação foi controlada pelo usuário por meio de sinais EMG. Além disso, um ultrassom foi usado para obter informações de distância e estimar o tamanho do objeto. O objetivo principal era encontrar o melhor dos quatro tipos de preensão (ou seja, pinça, lateral, palmar e esférica) de acordo com a forma, tamanho e orientação do objeto. Outros trabalhos usaram um modelo semelhante, como Markovic et al (Markovic et al., 2014). Porém, neste caso, a câmera foi posicionada nos óculos que foram usados pelo voluntário. Além disso, eles usaram realidade aumentada para fornecer feedback sobre o tipo de preensão e o tamanho da abertura da mão, discriminando entre quatro tipos de preensão (palmar, lateral, tri-dígito e pinça).

Em 2017, Ghazaei et al seguiram a mesma linha baseada em visão computacional. No entanto, eles treinaram uma arquitetura baseada em Redes Neurais Convolucionais (CNN) para aprender o formato de diferentes objetos previamente rotulados e classificados entre quatro tipos de preensão, sendo elas: palmar neutro, palmar pronado, tri-dígito e pinça (Ghazaei et al.,

2017). Além disso, o modelo foi testado com um novo conjunto de objetos que não fizeram parte do treinamento da CNN para verificar se o modelo foi robusto para novos objetos que não pertenciam a nenhum grupo de imagens previamente rotuladas. Entretanto, apesar de toda sua robustez, o sistema proposto se mostrou pouco dinâmico e natural, pois não levou em conta DoF importantes como o de rotação do punho.

Hays et al elaboraram outra abordagem recentemente que utilizou câmeras baseadas em eventos para simplificar o controle do sistema (Hays et al., 2019). Essas câmeras baseadas em eventos são inspiradas na retina e funcionam de forma assíncrona de acordo com a diferença de luminosidade em cada pixel (Lichtsteiner et al., 2008). Esse tipo de abordagem reduz algumas etapas de processamento e torna o sistema mais otimizado. Além disso, os autores combinaram este sistema com sensores táteis nas pontas dos dedos da prótese para evitar o deslizamento do objeto. Apesar de terem demonstrado eficácia, os objetos escolhidos foram do conjunto de dados PHAM (Yerrabelli, 2018), que consistiam em um cubo, prisma e cilindro, para as respectivas preensões: palmar, tri-dígito e pinça. Dessa forma, não foi possível garantir a capacidade de generalização quando confrontado com objetos comuns do dia a dia. Além disso, assim como os outros trabalhos apresentados acima, não se levou em conta um ambiente mais realístico, que normalmente contém múltiplos objetos em uma única cena.

## 1.1 OBJETIVOS

Apesar das abordagens apresentadas anteriormente serem essenciais para melhorar o controle da prótese, elas ainda estão longe de proporcionar um modelo de controle simples e que se aproxime às estratégias similares àquelas utilizadas para controle de uma mão natural, permitindo que o usuário possa utilizar a prótese com baixo esforço cognitivo. Dessa forma, esse trabalho propõe uma nova abordagem para o controle de próteses de mão, que busca atacar três importantes problemas encontrados na literatura nesta área: 1) Classificar o tipo de preensão ideal para o objeto almejado de maneira robusta e genérica de acordo com sua forma (Tripod para objetos delicados e que necessitem de precisão ou Power para objetos maiores que necessitem de força); 2) Controlar a angulação do punho de acordo com a orientação do objeto almejado de forma gradual e não apenas para poucas posições pré-definidas; 3) Segmentar o objeto almejado dos demais objetos, possibilitando que o usuário possa escolher um item específico, mesmo em cenas com múltiplos objetos.

Para validação do modelo, é realizado um experimento em tempo real de *pick and place* em que os usuários deverão realizar a tarefa utilizando o sistema de controle semiautônomo proposto e, posteriormente, o sistema de controle manual tradicional, para comparação e avaliação.

## 1.2 ESTRUTURA DA DISSERTAÇÃO

Este trabalho está dividido em 9 capítulos. Neste primeiro capítulo foi apresentada uma breve introdução do tema e os objetivos que devem ser alcançados ao fim da execução do presente estudo.

O capítulo 2 apresenta uma introdução sobre eletromiografia de superfície pontuando alguns dos principais acontecimentos históricos. Em seguida, é feita uma descrição do processo fisiológico que dá origem ao sinal de sEMG e por fim é explicado todo o processo de aquisição e condicionamento do sinal de sEMG.

O capítulo 3 apresenta uma breve introdução as próteses de membros superiores, pontuando os principais pontos históricos para alcançar os modelos atuais. Na sequência, são apresentados tipos de próteses existentes atualmente e qual as aplicações mais comuns para cada um desses tipos.

O capítulo 4 introduz as câmeras neuromórficas, comparando primeiramente seu funcionamento com a visão biológica e posteriormente com as câmeras convencionais. Por último, são apresentados os principais trabalhos que utilizaram de câmeras neuromórficas no seu desenvolvimento.

O capítulo 5 apresenta um resumo das principais técnicas de visão computacional utilizadas nesse trabalho. Primeiramente, são explicadas as Redes Neurais Convolucionais, pontuando os principais fatos históricos que foram necessários para o seu desenvolvimento. Em seguida, é explicada a técnica de Análise de Componentes Principais, focando principalmente no seu uso para obtenção de angulação de objetos. Por último, é apresentada a técnica de Watershed para segmentação de objetos em imagens.

No capítulo 6 são apresentados os materiais e métodos necessários para a realização desse trabalho, explicando cada etapa dos experimentos offline e em tempo real.

O capítulo 7 apresenta os resultados obtidos tanto para os experimentos offline quanto para os experimentos em tempo real, comparando os dois sistemas de controle desenvolvidos.

No capítulo 8 é apresentada a discussão dos resultados obtidos, pontuando todos os pontos positivos e negativos de ambos os sistemas de controle.

O capítulo 9 apresenta a conclusão desse estudo, indicando os principais achados que foram alcançados ao fim do trabalho. Além disso, são apresentados os principais horizontes para trabalhos futuros, que têm o potencial de preencher as limitações existentes no presente estudo.

# Capítulo 2

---

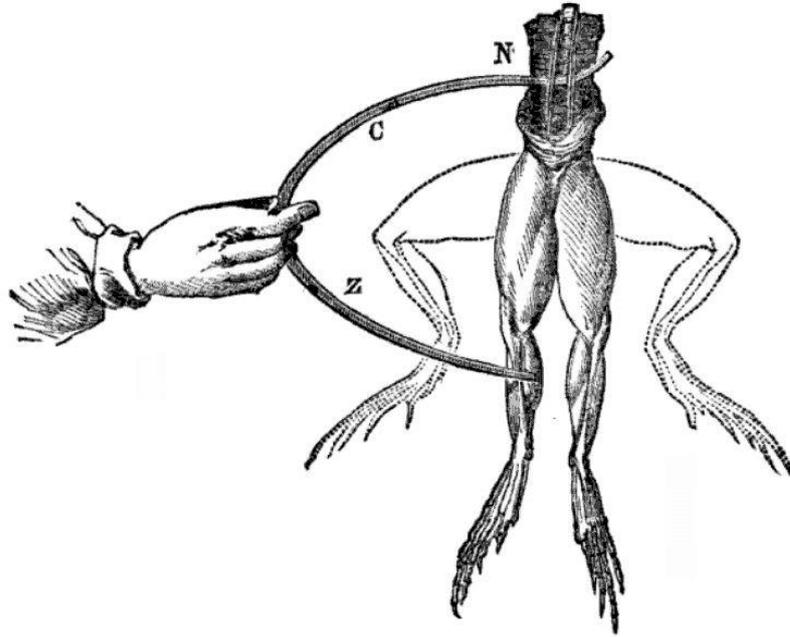
## 2. ELETROMIOGRAFIA DE SUPERFÍCIE

A análise de sinais fisiológicos é uma excelente forma de compreender o funcionamento do corpo humano (Basmajian, 1967). Por meio desses sinais, é possível mensurar aspectos importantes para fornecer um diagnóstico detalhado do estado atual de um indivíduo (Cavalcanti & Vieira, 2011). Em específico, a musculatura esquelética tem sido amplamente estudada, pois constitui cerca de 40% do corpo humano e é o atuador responsável pelo movimento do corpo por meio da sua contração (Hall & Guyton, 2017). Assim, a possibilidade de estudar a ativação dos músculos esqueléticos, por meio do registro dos potenciais elétricos produzidos durante as contrações musculares, é de particular relevância na ciência do esporte e na medicina de reabilitação (Basmajian, 1967; Cavalcanti & Vieira, 2011; Cram, 2003; Criswell, 2011; Hall & Guyton, 2017). Estudo esse que é denominado de eletromiografia (EMG).

### 2.1 HISTÓRIA

O termo eletromiografia foi introduzido pelo fisiologista francês Étienne-Jules Marey (Reaz et al., 2006). Entretanto, a relação entre eletricidade e contração muscular foi observada pela primeira vez por Luigi Galvani em 1791 (Basmajian, 1967). Em um dos seus experimentos mais conhecidos, ele despolarizou os músculos das pernas de uma rã tocando-os com hastes de metal (Figura 1). Esse experimento foi fundamental para despertar a curiosidade de diversos outros grandes pesquisadores como Emil du Bois-Reymond e Étienne-Jules Marey, levando a descobertas fundamentais para o avanço da EMG. Além disso, essa descoberta é geralmente reconhecida por representar o nascimento da neurofisiologia, tornando Galvani um dos principais pesquisadores nessa área.

Figura 1 – Representação do experimento de Luigi Galvani em 1791. Quando aplicado um potencial em cima das pernas de uma rã, ocorria sua contração.



Fonte: David Ames Wells, *The science of common things: a familiar explanation of the first* (página 290)

Quase meio século depois, em 1849, em um de seus experimentos mais famosos, Emil du Bois-Reymond forneceu a primeira evidência de atividade elétrica nos músculos humanos durante uma contração voluntária. Utilizando um galvanômetro, ele notou desvios muito pequenos, mas muito consistentes e previsíveis, sempre que o sujeito flexionava uma mão ou um braço. Entretanto, pelo fato de a tecnologia dessa época ser muito limitada, realizar a aquisição da atividade elétrica dos músculos não era uma tarefa trivial. Dessa forma, o primeiro experimento mais preciso foi realizado apenas em 1920 pelos fisiologistas norte-americanos Herbert Spencer Gasser e Joseph Erlanger, que usaram um osciloscópio de raios catódicos recém-inventado para mostrar os sinais dos músculos. Esse experimento foi de grande importância para a época, de forma que lhes rendeu o Prêmio Nobel em 1944.

Na primeira metade do século XIX, a EMG tinha sua aplicação mais comum voltada para diagnósticos clínicos apontando possíveis problemas neurológicos e neuromusculares. Além disso, a EMG também foi amplamente usada em muitos tipos de laboratórios de pesquisa, incluindo aqueles envolvidos em biomecânica, controle motor, fisiologia neuromuscular, distúrbios do movimento, controle postural e fisioterapia. Porém, uma das aplicações mais consolidadas ainda estava por vir.

No início dos anos 1950, o matemático norte-americano Norbert Wiener observou que os sinais biológicos que controlavam as ações perdidas na amputação ainda estavam presentes no corpo do amputado (Alter, 1966). Dessa forma, ele sugeriu que alguns desses sinais poderiam

ser usados para controlar uma prótese, em vez de usar movimentos corporais brutos. Dando origem a área de controle de próteses baseada em sinais de EMG. Em seguida, no final dos anos 1950 e início dos anos 1960, vieram os primeiros trabalhos que se baseavam em sinais de EMG para o controle de próteses (BATTYE et al., 1955; KOBRINSKII, 1960; Kobrinsky, 1960).

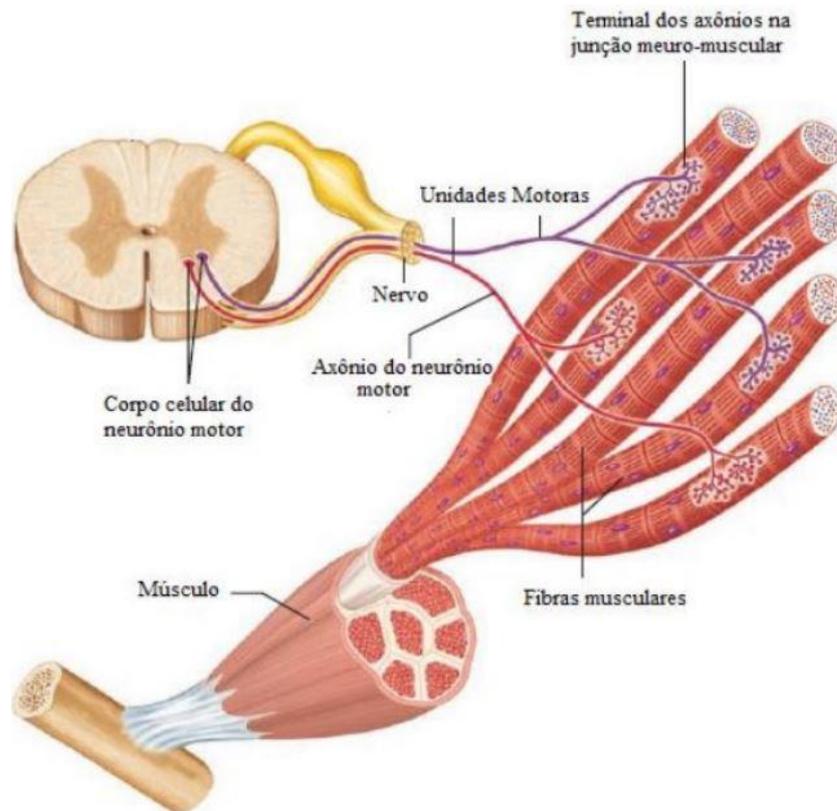
## 2.2 O SINAL ELETROMIOGRÁFICO

### 2.2.1 ORIGEM FISIOLÓGICA

As contrações voluntárias produzidas pelos músculos esqueléticos são resultantes de comandos originados no cérebro, de forma que ele envia sinais eletroquímicos por meio do sistema nervoso para o neurônio motor que inerva várias fibras musculares (Hall & Guyton, 2017). Cada neurônio motor possui um potencial de membrana de cerca de  $-70$  mV. Dessa forma, quando há a intenção de realizar um movimento, uma despolarização dos neurônios motores se inicia, de forma que se o potencial na célula for maior que o limiar de despolarização ( $-50$  mV), um potencial de ação é desencadeado ao longo do neurônio, se espalhando por toda membrana muscular e conseqüentemente, gerando a contração dos músculos inervados por esse neurônio.

Um neurônio muscular é capaz de inervar várias fibras musculares (Figura 2), de forma que o conjunto composto por um único neurônio motor e fibras musculares supridas por seus ramos axonais dão origem as entidades funcionais do sistema neuromuscular, denominada como unidade motora (UM) (De Luca, 2002). Uma vez que um neurônio motor é descarregado, potenciais de ação são gerados em suas junções neuromusculares e então se propagam ao longo de todas as fibras musculares, em direção às regiões dos tendões. A soma desses potenciais é denominada potencial de ação da unidade motora (MUAP) e é responsável pela contração muscular (Cavalcanti & Vieira, 2011).

Figura 2 – Ilustração da estrutura fisiológica do sistema neuromotor. É possível ver as diversas unidades motoras, onde cada fibra muscular está conectada somente um neurônio motor. Entretanto, um único neurônio motor pode se conectar a diversas fibras musculares.



Fonte: Mambrito e De Luca, 1983.

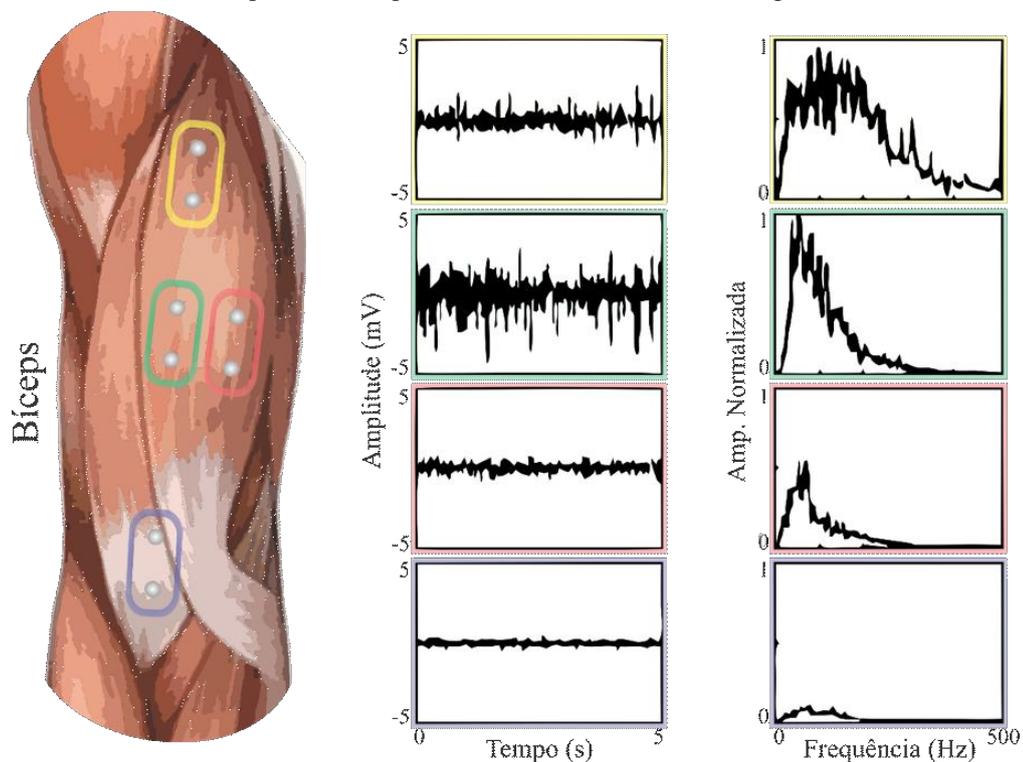
### 2.2.2 AQUISIÇÃO DE SINAIS

A energia gerada pelos músculos tem um valor muito pequeno, de forma que sua amplitude é da casa de microvolts ( $\mu\text{V}$ ). Assim, é necessário usar instrumentos muito sofisticados e sensíveis para amplificar este sinal de forma que possa ser visto, filtrando qualquer tipo de artefatos presentes no sinal. Entretanto, os primeiros experimentos com sEMG eram bastante complicados devido a tecnologia pouco avançada da época. Um dos principais problemas enfrentados acontecia devido a facilidade de interferência e contaminação do sinal por meio de outras fontes de energia eletromagnética presentes no ambiente de gravação. Felizmente, com os avanços na eletrônica e a introdução do amplificador diferencial em 1950, foi possível realizar a aquisição de sinais de sEMG mais robustos a ruídos e interferências externas. Esse acontecimento foi de grande importância para elevar a sEMG para outro patamar, expandindo seus horizontes ao migrar do campo da pesquisa para a área clínica. Assim, a sEMG

teve suas primeiras aplicações em biofeedback, quiropraxia, fisioterapia, medicina física, neurologia e urologia (Criswell, 2011).

Como visto na seção 2.2.1, o sinal eletromiográfico tem origem no cérebro a partir da intenção de movimento. Entretanto, esse sinal se desloca por meio de potenciais de ação até as fibras musculares onde ocorre a contração dos músculos, e é exatamente nessa região que serão posicionados os eletrodos diferenciais. Esses eletrodos são responsáveis pela transdução da energia química presente nos músculos para energia elétrica presente em todo o circuito de aquisição. Pelo fato de o sinal ter uma amplitude muito baixa, a etapa de posicionamento dos eletrodos é de extrema importância. A Figura 3 ilustra possíveis posicionamentos dos eletrodos na superfície de um determinado músculo. Assim, o posicionamento mais indicado é na região do ventre do músculo, de forma que são captados o maior número de potenciais de ação e conseqüentemente, obtendo uma maior amplitude do sinal. Todo o processo do melhor posicionamento é descrito por Davis no livro *Manual of surface electromyography* (Davis, 1959).

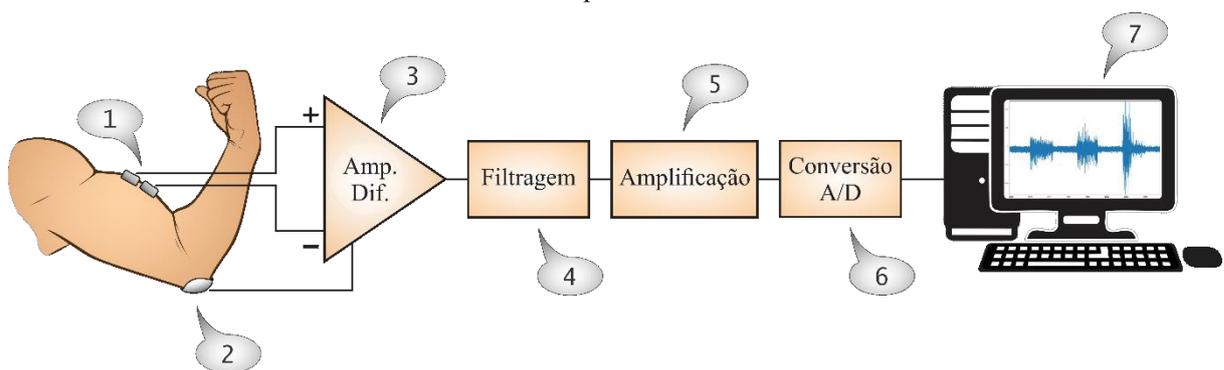
Figura 3 – Exemplos de posições no posicionamento dos eletrodos de eletromiografia de superfície. A forma do sinal e seu espectro de frequências variam de acordo com o lugar escolhido.



A Figura 4 mostra o diagrama que exemplifica as principais etapas da aquisição da sEMG. Após a transdução realizada pelos eletrodos, o sinal passa por um amplificador diferencial, que é fundamental para a remoção de artefatos causados por interferências eletromagnéticas. Na

amplificação diferencial, três eletrodos são necessários: dois eletrodos de captação dos sinais e um eletrodo de referência. Os eletrodos de captação dos sinais são colocados sobre os músculos, com o eletrodo de referência simplesmente fazendo um bom contato em algum lugar do corpo. Ambos os eletrodos de sinais registram a energia gerada pelos músculos em um instante específico de tempo à medida que a fibra muscular se despolariza. Os eletrodos estão posicionados com cerca de dois centímetros entre eles. Essa configuração gera uma defasagem no tempo no sinal captado por cada eletrodo, pois ambos registram os potenciais de ação da fibra em instantes diferentes de tempo, o que faz com que esses sinais tenham formas de onda diferentes. Dessa forma, quando passados pelo circuito diferencial, apenas o que é único do sinal captado em cada eletrodo é mantido, descartando a maior parte dos artefatos gerados pelo ambiente. Além disso, a etapa de amplificação diferencial é responsável pela primeira amplificação do sinal, que passa a ter uma amplitude de pico a pico na casa de milivolts (mV).

Figura 4 – Diagrama simplificado das unidades componentes de um sistema para coleta de eletromiografia de superfície.



Após a etapa da amplificação diferencial, o sinal é passado por um filtro passa-baixa e um passa-alta, resultando em um filtro passa-faixa. Esses filtros são essenciais para reduzir o espectro de frequência de modo a restar apenas a banda do sinal eletromiográfico, que está contida entre 15 e 400 hertz (Hz) (Cavalcanti & Vieira, 2011). O filtro passa-baixa deve remover todos os artefatos de alta frequência ocasionados por outros componentes do circuito, como osciladores. Já o filtro passa-alta é responsável pela remoção de níveis de tensão contínua que não foram totalmente removidos na etapa de amplificação diferencial, de forma que esses níveis de tensão podem afetar drasticamente o sinal final à medida que o sinal for amplificado. Além disso, os filtros passa-alta são responsáveis por eliminar as baixas frequências geradas por artefatos biológicos, pois são altamente inconstantes e podem alterar a forma de onda desejada.

A próxima etapa consiste na utilização de um filtro-notch, que é responsável pela remoção de artefatos gerados pela rede local. A frequência de oscilação da rede varia de acordo com cada lugar, de modo que no Brasil a frequência é de 60 Hz, porém em outros lugares esse valor pode ser de 50 Hz. Dessa forma, circuitos de aquisição construídos para funcionar no Brasil, devem ter um filtro-notch responsável por atenuar a frequência de 60 Hz do espectro de frequência do sinal eletromiográfico.

Por fim, o ganho do amplificador é um estágio crucial em sistemas de aquisição de sinais de sEMG. A amplificação é importante para que a amplitude dos sinais de sEMG detectados corresponda à faixa dinâmica do conversor A/D. Normalmente, a faixa dinâmica do conversor A/D em sistemas eletromiográficos varia de  $\pm 2,5$  V a  $\pm 10$  V. Além disso, um outro ponto que complementa a necessidade da amplificação do sinal antes da etapa de digitalização é sua resolução em bits. Normalmente a resolução de um conversor A/D utilizado na eletromiografia varia de 10 a 16 bits, de forma que quanto maior a resolução, maior é a capacidade do conversor A/D de discriminar pequenas variações do sinal. Dessa forma, se a amplitude do sinal não estiver de acordo com a faixa dinâmica do conversor A/D, boa parte da sua capacidade de discriminação será inutilizada, evitando com que o sinal digitalizado esteja representando fielmente o sinal fisiológico.

# Capítulo 3

---

## 3. PRÓTESES DE MEMBROS SUPERIORES

Como já apontado, a mão humana é um dos membros mais complexos do corpo humano, pois possui diversos DoF e é capaz de combinar destreza e força. Assim, quando um sujeito sofre uma amputação, recuperar suas funções básicas e estética passa a ser uma tarefa muito complicada. Ao longo dos anos, essa necessidade de recuperar tais funções, resultou no aparecimento das próteses de membros superiores, que com o avanço da tecnologia se aproximam cada vez mais de um modelo ideal. As próteses são dispositivos artificiais que substituem uma parte do corpo ausente, que pode ser perdida por trauma, doença ou condição presente ao nascer (distúrbio congênito), o que faz com que um dispositivo protético seja de grande importância para restaurar as funções básicas e estética do membro perdido. Dessa forma, a evolução tanto da mecânica quanto do controle é essencial para tornar esses modelos cada vez mais próximos do membro humano.

### 3.1 HISTÓRIA

Os primeiros modelos de próteses foram registrados no Egito em uma prótese ocular cerca de 2900 a.C. (Pine et al., 2015). Em 2000, o professor Nerlich e seu grupo encontraram um outro modelo de próteses junto a uma múmia em uma tumba também no Egito. Nesse caso, a prótese encontrada era de um dedo polegar do pé direito (Nerlich et al., 2000). Já as próteses de membros superiores tiveram seus primeiros registros em 77 d.C. pelo filósofo romano Pliny em seu livro “*Naturalis Historia*”, onde descreveu o uso de uma prótese de mão pelo general romano Marcus Sergius para segurar seu escudo durante a Segunda Guerra Púnica (218-201 a.C.).

Depois de mais de mil anos, nos séculos XV e XVI surgiram novos registros da invenção e uso de próteses de membros superiores. Primeiro em 1505, onde o cavaleiro alemão Götz von Berlichingen perdeu sua mão direita após um disparo de canhão contra sua espada. Esse caso ficou bastante famoso pela arquitetura da prótese desenvolvida, que foi feita de ferro de forma que seus dedos podiam ser flexionados e estendidos passivamente nas articulações

metacarpofalangianas, interfalangianas proximais e interfalangianas distais, bem como a articulação interfalangiana do polegar (Putti, 2005). Tempos depois, em 1517, o historiador italiano Paolo Giovio registrou que o pirata turco Horuk Barbarossa perdeu sua mão direita na Batalha da Bugia contra a Espanha. Dessa forma foi desenvolvido uma prótese de ferro, permitindo que ele continuasse em batalha (Zuo & Olson, 2014). Já em 1622, um caso parecido voltou a acontecer na Batalha de Fleury, onde o duque Christian de Brunswick que teve sua mão substituída por uma prótese desenvolvida por um artesão holandês após perdê-la em combate (Putti, 2005). Em todos os casos apresentados até agora, as próteses foram desenvolvidas para combate, de modo que o usuário voltou a ser capaz de segurar seu escudo, espada ou outra arma de batalha.

Por volta de 1600, o médico cirurgião italiano Giovanni Tommaso Minadoi registrou dois casos de indivíduos amputados que utilizavam próteses. O primeiro tinha perdido as duas mãos, mas, com próteses, podia colocar/tirar o chapéu, amarrar/desamarrar a bolsa e até mesmo assinar o próprio nome. O segundo, que teve sua mão direita amputada, pode realizar muitas ações voluntárias usando uma mão de ferro (Putti, 2005).

Após anos de evolução, em 1818, o dentista alemão Peter Baliff introduziu o conceito de próteses acionadas por meio de movimentos do corpo (acionamento por tirantes), recuperando alguns dos DoF básicos da mão humana, como a abertura e fechamento da mão (Meier, 2004). Esse sistema proporcionou pela primeira vez ao usuário uma prótese de mão com movimentos fluidos, em vez de apenas um objeto estranho anexado ao corpo. Em 1948, esse modelo de próteses foi melhorado com o uso de cabo Bowden ao invés das tiras volumosas normalmente utilizadas. Mesmo com um avanço constante, esses modelos de próteses acionadas por tirantes ainda se baseiam na utilização de cabos Bowden até os dias de hoje.

Outro acontecimento importante que ocorreu no final da década de 40 foi a introdução dos primeiros modelos de próteses elétricas controladas por meio de sinais de EMG, pelo estudante de física Reinhold Reiter em 1948 e pelo matemático Norbert Wiener em 1950. Essa descoberta inspirou novos pesquisadores da área a se basear nessa abordagem, como Kobrinski em 1960 (Kobrinsky, 1960). Após 70 anos desde os primeiros experimentos, esse modelo de prótese é amplamente utilizado nos dias de hoje devido principalmente ao avanço tecnológico em campos como da medicina, microeletrônica, mecânica e materiais.

## 3.2 TIPOS DE PRÓTESES DE MEMBROS SUPERIORES

Existem diversas formas de dividir os tipos de próteses existentes atualmente. Entretanto, a divisão mais vista na literatura divide as próteses de membros superiores em duas categorias: passivas e ativas (Cordella et al., 2016). Além disso, cada uma das categorias possui suas subcategorias. As próteses passivas são subdivididas em cosméticas e funcionais. Já as próteses ativas são subdivididas de acordo com o seu acionamento, que pode ocorrer por meio de tirantes ou eletronicamente. As categorias e subcategorias de próteses mencionadas serão descritas de forma detalhada nos tópicos subsequentes.

### 3.2.1 PASSIVAS

As próteses passivas receberam esse nome pois elas não possuem nenhum tipo de atuação. Assim, esses modelos são estáticos, de forma que qualquer alteração necessária na forma da prótese seja feita de forma manual pelo usuário. Como dito anteriormente, as próteses passivas são divididas em cosméticas e funcionais de acordo com a sua finalidade.

- **Cosméticas:**

Em sua maioria, as próteses cosméticas possuem apenas uma função: restituir a aparência estética do membro perdido pelo usuário. Em alguns casos, restaurar a aparência física é mais relevante para o usuário do que recuperar as funcionalidades do membro perdido, tornando um modelo cosmético mais adequado. Um outro motivo frequente é dado quando recuperar as funcionalidades do membro perdido é muito complicado, sendo que em alguns casos chega a ser impossível. Dessa forma, a aparência física pode se tornar um ponto chave para recuperar a autoestima do usuário.

- **Funcionais:**

Por outro lado, as próteses funcionais possuem boa parte das características das cosméticas. Porém, elas são construídas para facilitar a execução de tarefas simples. Em períodos de guerra, diversos modelos de próteses funcionais foram desenvolvidos para que combatentes conseguissem segurar seu escudo ou espada. Na seção 3.1 foram apresentados

alguns exemplos práticos do uso de próteses funcionais que ocorreram em batalhas marcantes da história humana.

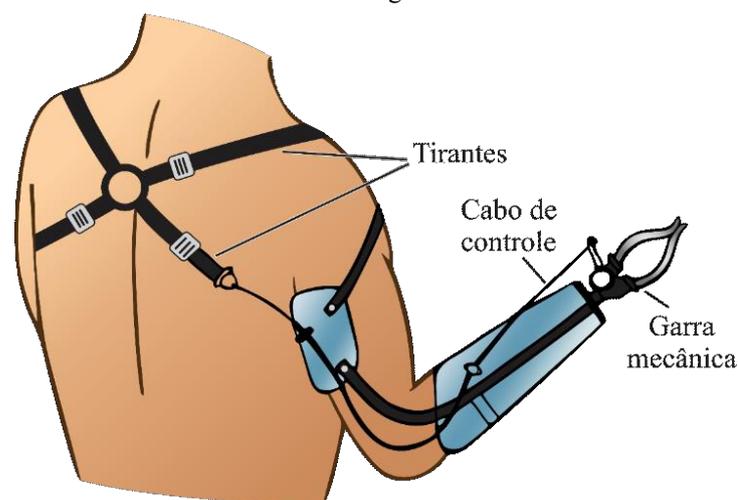
### 3.2.2 ATIVAS

Diferentemente das passivas, as próteses ativas fornecem alimentação externa por meio de motores e, conseqüentemente, possuem capacidade de ação. Embora ofereçam maior desempenho e funcionalidade, eles representam o sistema com a maior complexidade (Windrich et al., 2016). As próteses ativas são divididas em dois grupos principais de acordo com seu tipo de ativação: por tirantes ou eletronicamente.

- Por tirantes:

O acionamento de próteses por tirantes foi o pioneiro na categoria de próteses ativas. Esses modelos têm seu controle baseado na utilização de cabos presos ao membro sadio do amputado por meio de tirantes ( Figura 5). Ao realizar os movimentos de abdução ou extensão do ombro (articulação glenoumeral) o cabo de controle é tensionado por meio de um sistema mecânico (Billock, 1986). Por sua vez, esse tensionamento implica na abertura da prótese de forma proporcional, ou seja, quanto maior for a tensão no cabo, maior será a abertura da prótese.

Figura 5 – Ilustração de um modelo de próteses acionado por tirantes. No exemplo em questão, a prótese usada consistia em uma garra mecânica.



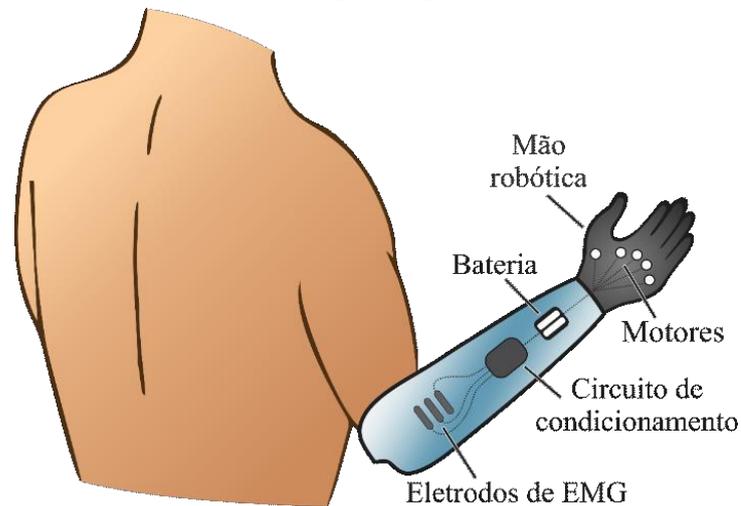
Embora em alguns casos eles sejam muito úteis, os modelos de próteses acionados por tirantes são limitados em relação a DoF, podendo realizar apenas a abertura/fechamento da prótese. Conseqüentemente, o usuário consegue realizar apenas tarefas de prensão que não

necessitem de muita precisão. Além disso, toda a energia usada para atuar o sistema vem do próprio usuário, gerando fadiga muscular para o usuário de acordo com o tempo de uso (Cordella et al., 2016).

- Eletrônica:

Em contrapartida, os modelos próteses eletrônicas trouxeram uma nova perspectiva para a área. Principalmente por aprimorar aspectos importantes do sistema, como a forma de atuação, que graças ao uso de fontes de alimentação externas deixou de ser feita por meio de movimentos brutos realizados pelo ombro e passou a ser efetuada por meio de motores elétricos (Figura 6).

Figura 6 – Ilustração de um modelo clássico de próteses eletrônicas. Nesse exemplo, o sistema se baseia no uso de eletromiografia de superfície para realizar o controle.



Essa nova abordagem ocasionou em modelos de próteses sofisticados e diversos DoF, de modo que cada dedo da prótese passou a ter suas junções atuadas de forma individual. De fato, ao se aumentar o número de DoF é possível alcançar um mecanismo mais próximo da mão humana. Entretanto, como consequência desse aumento dos DoF, a complexidade do sistema de controle também aumenta que, por sua vez, implica em uma maior dificuldade no controle do sistema pelo usuário. Dessa forma, levando em conta que os sinais dos músculos estão diretamente ligados à contração muscular, a sEMG passou a ser um forte e promissor sinal de controle para as próteses eletrônicas.

No século XX, a sEMG ganhou muita força em aplicações médicas e, em conjunto com os grandes avanços obtidos nas áreas de eletrônica, foi possível a construção de circuitos de condicionamento cada vez mais robustos e compactos. Em meados dos anos 40, alguns pesquisadores começaram os primeiros trabalhos de controle de próteses baseados em sinais de

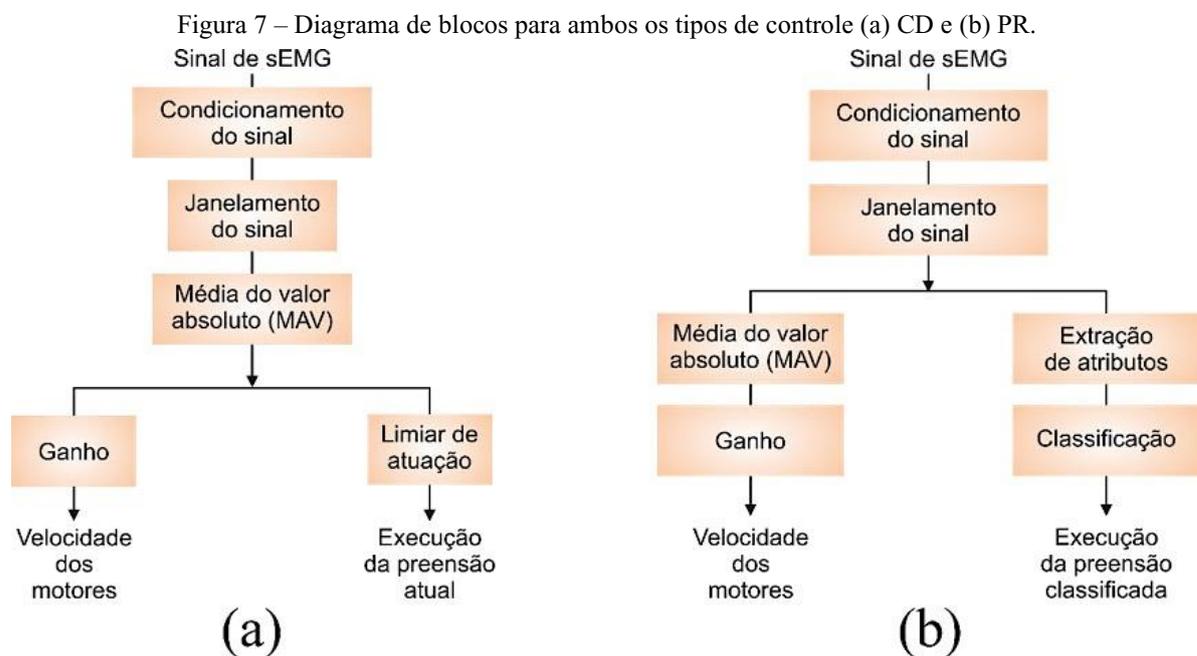
sEMG, dando origem as próteses eletromiográficas. Até os dias de hoje essa forma de controle é a mais utilizada pelos grupos de pesquisa e modelos de próteses comerciais.

Apesar dos inúmeros avanços tecnológicos nos últimos anos, os modelos de próteses eletromiográficas ainda possuem limitações importantes, como a dificuldade de controlar os diversos graus de liberdade da prótese com poucos sinais de sEMG. Ao se analisar a fisiologia da mão humana, percebe-se que existem milhares de fibras musculares, que se agrupam formando 20 músculos no antebraço e 30 músculos na mão. Por sua vez, esses músculos conseguem atuar todos os 27 DoF da mão/punho trabalhando em conjunto ou individualmente. Por outro lado, embora forneça uma boa representação do sinal responsável pela contração muscular, a sEMG não consegue captar o sinal de cada músculo individualmente, já que os sinais de sEMG são a somatória dos sinais gerados pelas fibras musculares próximas ao eletrodo, de forma que quanto mais próxima a fibra estiver do eletrodo, maior é sua amplitude (Hall & Guyton, 2017). Conseqüentemente, os modelos de próteses eletromiográficas comerciais se apoiam em formas de controle mais simplificadas, que consigam ter um bom funcionamento mesmo com poucos sinais de sEMG.

Em sua grande maioria, os modelos comerciais usam o CD, que se baseia na utilização de dois canais de eletromiografia, normalmente antagonistas, que são responsáveis pelo chaveamento do tipo de preensão da prótese de mão (Figura 7). Esse controle possui pontos positivos, como a troca entre tipos de preensão de forma rápida e acurada, já que as preensões são pré-programadas no controlador do sistema (White et al., 2017). Dessa forma, mesmo que pouco natural, esse tipo de controle é bastante intuitivo para o usuário. Em contrapartida, a complexidade do controle aumenta drasticamente à medida que o número de DoF aumenta, pois o usuário possivelmente precisará realizar diversas contrações até chegar na preensão desejada. Os modelos de próteses comerciais baseadas nesse modelo possuem blocos de preensões limitados, de forma que normalmente o usuário pode fazer o chaveamento de apenas dois tipos de preensão por bloco.

Em consequência dos avanços na área de AM, o controle baseado em RP surgiu com a promessa de tornar natural o controle da prótese para o usuário. Embora a sEMG forneça apenas uma representação global da ativação dos músculos, é possível encontrar padrões nas ativações que são responsáveis por um determinado tipo de preensão em um membro sadio. Portanto, mesmo após sofrer uma amputação, o indivíduo consegue enviar comandos de uma determinada preensão para os músculos remanescentes, que possibilita um modelo de AM aprender os padrões do sinal de sEMG desses músculos. Conseqüentemente, essa abordagem permite que o usuário realize mais tipos de preensões em comparação com o CD.

Diferentemente do CD, um sistema de controle baseado em RP normalmente precisa de mais canais de eletrodos espalhados ao longo dos músculos remanescentes do antebraço, pois é possível alcançar uma maior representatividade da atividade muscular responsável por um tipo de preensão. Esses músculos normalmente são aqueles pertencentes ao antebraço que são os responsáveis pelos movimentos dos dedos como o flexor carpo ulnar, flexor carpo radial, extensor carpo ulnar, extensor carpo radial, extensor longo do polegar e palmar longo. Entretanto, esse aumento do número de canais gera um aumento na complexidade do sistema, já que serão necessários mais circuitos de condicionamento e um microcontrolador mais avançado.



O controle baseado em RP esteve apenas na área da pesquisa por muitos anos, já que no início possuía algumas limitações, principalmente em seu hardware, que dependia de máquinas grandes e potentes para rodar o modelo de aprendizagem. Entretanto, com o avanço da tecnologia na área de eletrônica embarcada, novos modelos de microcontroladores capazes de rodar modelos de AM capazes de proporcionar uma classificação rápida e robusta foram desenvolvidos. Além disso, a constante alteração nos padrões do sinal de sEMG foi um fator limitante para o surgimento de modelos comerciais de controle de próteses baseado em RP, podendo ocasionar no mau funcionamento do sistema (Resnik et al., 2018). Dessa forma, durante muitos anos, muitos pesquisadores da área de controle de próteses concentraram seus esforços em alcançar um modelo de classificação suficientemente robusto à mudança constante dos padrões da sEMG.

Em 2014, o primeiro modelo comercial baseado em RP foi lançado pela empresa COAPT. Com essa inovação, os modelos baseados em RP saíram apenas da área de pesquisa, dando o primeiro passo para uma nova etapa do controle de próteses de membros superiores. Recentemente, a Ottobock, uma das maiores empresas da área de próteses comerciais lançou a MyoPlus, que é o seu próprio modelo de controle de próteses baseado em RP (Ottobock, 2021).

# Capítulo 4

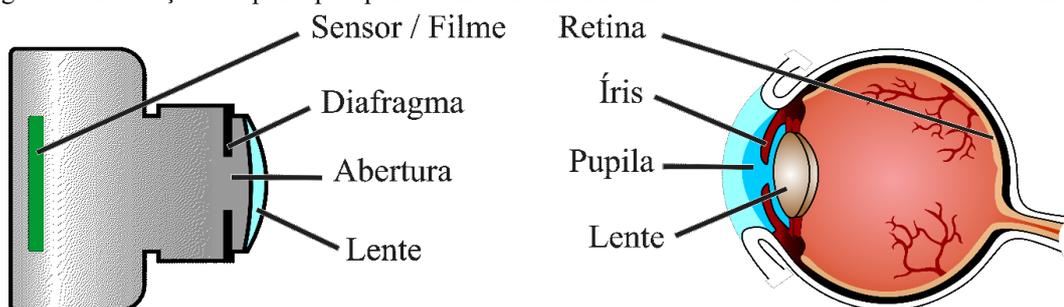
---

## 4. VISÃO BASEADA EM EVENTOS

A visão é reconhecida como um dos sentidos de maior importância, já que muitos sistemas biológicos usam a visão como sua forma mais proeminente de coletar informações sobre seu ambiente (Faugeras, 1993). Sabendo disso, o ser humano buscou durante anos formas de conseguir reproduzir a visão, desde as primeiras câmeras obscuras até as sofisticadas câmeras existentes nos dias de hoje, que estão anexadas a praticamente todos os telefones celulares, computadores portáteis e estabelecimentos comerciais.

As câmeras convencionais e os olhos possuem diversos pontos em comum em seus princípios de funcionamento. No olho, o nível de luz recebida é controlado de acordo com a abertura do músculo da íris. Dessa forma, o olho é capaz de se adaptar em ambientes tanto com pouca e muita luz. Em seguida, essa luz é refratada pela córnea e por fim chega na retina, que é um tecido sensível a luz. Assim como o olho, as câmeras convencionais possuem a habilidade de formar imagens por meio da luz recebida, de modo que os níveis de luz recebidos são definidos de acordo com a abertura da lente. Após essa etapa, a luz passa pelas lentes até chegarem no sensor de imagem da câmera.

Figura 8 – Ilustração dos principais pontos em comum entre uma câmera convencional e o olho humano.



Comparando ambas as estruturas do olho e das câmeras convencionais, podem-se observar diversas similaridades, de modo que o sensor sensível a luz, o diafragma, a abertura da lente e as lentes da câmera possuem as mesmas funções básicas da retina, da íris, da pupila e da lente do olho respectivamente (Figura 8). Entretanto, quando se adiciona a grandeza de tempo, as formas biológica e eletrônica diferem em pontos importantes. A visão humana é contínua, de forma que os olhos estão recebendo um fluxo de informações a todo momento. Por

outro lado, as câmeras convencionais não conseguem fornecer um fluxo de informação contínuo no tempo, mas sim um conjunto de imagens do ambiente geradas sequencialmente, dando origem a um vídeo.

Existem dois termos importantes que devem ser levados em conta ao se gravar um vídeo. O primeiro é o tempo de exposição, que consiste no tempo em que o sensor fica exposto à luz, dando origem a uma imagem (também chamada de quadro quando se trata de vídeos). Dessa forma, quanto menor for o tempo de exposição, menor são os efeitos de borrão ou *motion blur* (MB) presentes nos quadros contidos no vídeo. Por outro lado, ao se reduzir o tempo de exposição o sensor recebe menor quantidade de luz, o que afeta diretamente a qualidade da imagem. O segundo termo é a taxa de quadros, que é a quantidade de quadros contidos em um determinado período. A forma mais comum de se medir a taxa de quadros se dá por meio de quadros por segundo, ou do inglês, *frames per second* (FPS).

As câmeras atuais possuem números de FPS diferentes de acordo com o modelo, de forma que os valores mais comuns são de 24 e 30 FPS. Tomando como exemplo um vídeo gravado a 30 FPS, tem-se que o intervalo entre o início da formação de um quadro para o próximo é de 33 ms. Embora um vídeo gravado a 30 FPS pareça contínuo aos olhos humanos, parte das informações contidas nesses intervalos são perdidas, ou distorcidas pelo efeito de MB. Na Figura 9 é ilustrado esse efeito, o qual ocorre quando um objeto ou algo em movimento se desloca em alta velocidade na frente da câmera, fazendo com que ele apareça distorcido na mesma direção em que ocorreu o movimento.

Figura 9 – Imagens obtidas a partir de oito diferentes tempos de exposição.



De fato, os problemas apresentados acima afetam algumas aplicações do campo de visão computacional, como rastreamento de objetos em altas velocidades. Assim, as câmeras neuromórficas surgiram com uma proposta bioinspirada capaz de resolver algumas das limitações das câmeras convencionais (Lichtsteiner et al., 2008).

O modelo pioneiro de câmera baseada em eventos a se destacar foi o *Dynamic Vision Sensor 128* (DVS128), desenvolvido pelo pesquisador suíço Patrick Lichtsteiner e colaboradores em 2008 (Lichtsteiner et al., 2008). No entanto, a câmera DVS128 e suas variantes descartam informações de intensidade absoluta que são úteis para o reconhecimento e classificação de objetos. Por conta disso, em 2014 foi lançado o modelo *Dynamic and Active Pixel Vision Sensor* (DAVIS), que segue os mesmos princípios assíncronos da DVS128 combinados com a captação de eventos de forma síncrona (Brandli et al., 2014). Como resultado, as câmeras DAVIS possibilitaram um modelo híbrido, capaz de combinar as vantagens de ambos os tipos de câmeras.

O impacto das câmeras baseadas em eventos foi tão grande que recentemente grandes empresas resolveram direcionar parte dos seus esforços para essa área. A Samsung vem sendo uma das empresas mais engajadas, de modo que em 2017 publicaram seu primeiro modelo em modelo de câmera neuromórfica (Son et al., 2017). Esse modelo possui uma resolução espacial de 640x480 pixels, amenizando uma das principais limitações das câmeras baseadas em eventos lançadas anteriormente.

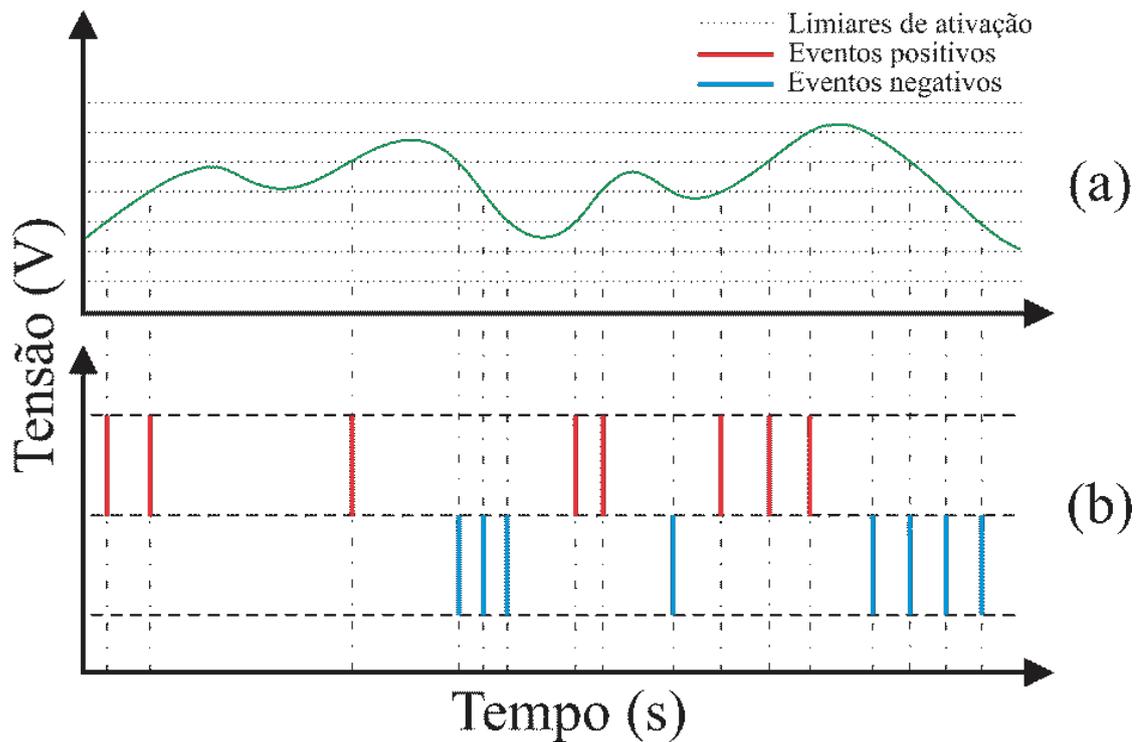
#### 4.1 CÂMERAS NEUROMÓRFICAS VERSUS CÂMERAS CONVENCIONAIS

As câmeras neuromórficas, ou também conhecidas como câmeras baseadas em eventos possuem esse nome pois seu princípio de funcionamento é inspirado na morfologia da visão biológica. Os primeiros trabalhos relacionados a reprodução do sistema visual humano foram desenvolvidos pela estudante de neurociência norte-americana Misha Mahowald e seu orientador Carver Mead em 1991 (Mahowald & Mead, 1991). Tal trabalho deu origem a novas pesquisas na área, de forma que anos depois foram lançados os primeiros modelos de câmera baseada em eventos (Brandli et al., 2014; Lichtsteiner et al., 2008; Posch et al., 2011; Son et al., 2017).

Diferentemente das câmeras convencionais, que geram frames a uma determinada taxa padrão, as câmeras baseadas em eventos funcionam de maneira assíncrona de acordo com as mudanças de luminosidade da cena em cada um dos pixels que formam a matriz de sensores

(Gallego et al., 2020). Ou seja, cada pixel pode gerar um evento (*spike*) de acordo com a luminosidade recebida individualmente. A Figura 10 mostra a resposta da tensão ao longo do tempo em um único sensor. Nota-se que a cada vez que o sinal de tensão ultrapassa um limiar, um evento é gerado, que pode ser positivo ou negativo (Figura 10b).

Figura 10 – Características do sinal obtido por cada um dos sensores contidos na matriz de sensores. (a) Resposta de tensão para um único sensor de uma câmera baseada em eventos. A amplitude do sinal é diretamente proporcional a luminosidade recebida. (b) Eventos gerados à medida que o sinal de tensão ultrapassa os limiares de ativação.

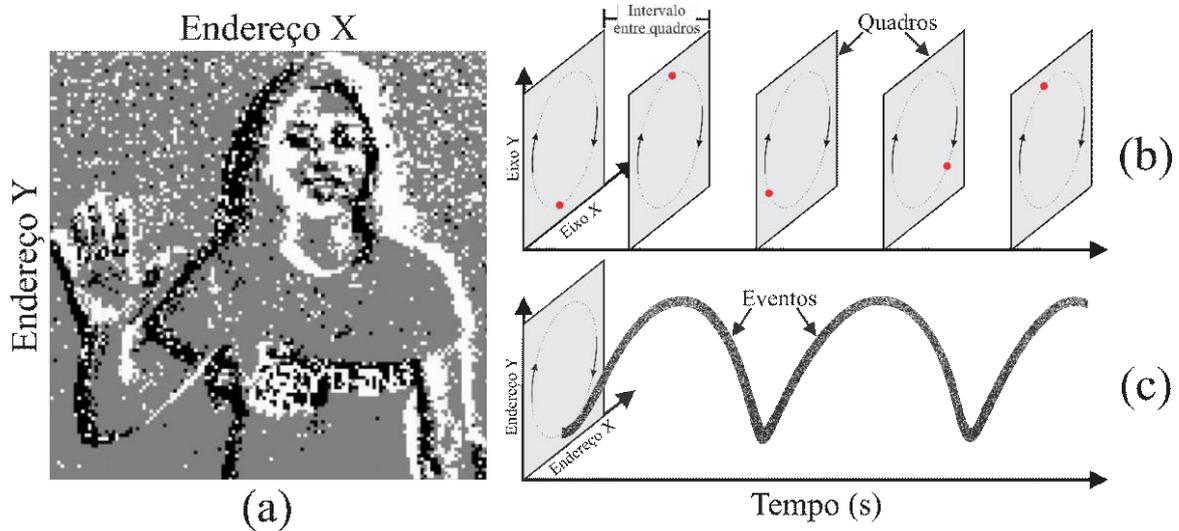


Embora nesses sistemas baseados em eventos a ideia de frame não exista, podemos fazer algumas comparações com as câmeras convencionais em relação a velocidade de aquisição de informação, de forma que se consideramos cada evento gerado em um instante de tempo como um novo frame, temos que uma câmera baseada em eventos pode alcançar um milhão de FPS. Para exemplificar isso, a Figura 11b e a Figura 11c ilustram a aquisição de uma imagem ao longo do tempo para ambas os tipos de câmera. Pode-se notar que existem espaços entre os quadros gerados em uma câmera padrão onde as informações da cena são perdidas ou captadas de forma distorcida. Por outro lado, a aquisição de uma câmera baseada em eventos é praticamente constante.

Essas câmeras não possuem uma escala gradual de cores, pois são baseadas apenas em eventos. Ou seja, caso aconteça uma mudança na luminosidade recebida por um pixel suficientemente grande para ultrapassar o limiar de disparo, será gerado um evento de valor

unitário no respectivo instante de tempo, que pode ser positivo ou negativo, de acordo com o aumento ou diminuição dessa luminosidade recebida. Dessa forma, a representação gráfica em 2D de uma gravação realizada por uma câmera baseada em eventos possui apenas três cores que representam os pixels onde: ocorreram os eventos positivos, ocorreram os eventos negativos e não geraram nenhum evento (Figura 11a). Por outro lado, a representação gráfica em 3D é mais interessante, pois o eixo do tempo passa a ser representado e os pixels onde não há eventos deixam de ser representados (Figura 11c).

Figura 11 – Comparação entre as imagens de saída de uma câmera convencional e uma câmera baseada em eventos. (a) Reconstrução de uma imagem a partir de eventos obtidos por uma câmera DVS128. Representação dos frames e eventos de saída gerados por uma (b) câmera convencional e (c) câmera baseada em eventos ao longo do tempo.



Os eventos são captados por cada um dos sensores situados na matriz de pixels e enviados para fora da câmera usando um barramento de saída digital compartilhado, normalmente usando leitura de representação de evento de endereço, ou do inglês *address-event representation* (AER) (Liu et al., 2014). São enviadas quatro informações da matriz de pixels: endereço X e Y, tempo T e polaridade P. A partir desses dados, é possível fazer a reconstrução de uma imagem por meio do agrupamento de eventos durante um período determinado.

As câmeras baseadas em eventos possuem algumas vantagens em comparação com as câmeras convencionais. Primeiro, essas câmeras possuem uma resolução temporal muito alta e baixa latência (ambas na ordem de microssegundos), de forma que o efeito de MB comum em câmeras convencionais é evitado. Além disso, elas possuem uma faixa dinâmica muito alta (140 dB vs. 60 dB de câmeras convencionais). Por último, elas têm um baixo consumo de energia devido a sua forma de aquisição, que funciona de maneira assíncrona e apenas quando há alteração significativa na luminosidade recebida, evitando aquisição de informações

redundantes (Lichtsteiner et al., 2008). Portanto, as câmeras de eventos têm um grande potencial para robótica e aplicações vestíveis em cenários desafiadores para câmeras padrão, como alta velocidade e alta faixa dinâmica.

## 4.2 PRINCIPAIS APLICAÇÕES CONTIDAS NA LITERATURA

Embora recente, a abordagem neuromórfica mostrou ser eficiente em diversas aplicações, principalmente naquelas voltadas para a área da robótica, como em rastreamento de objetos. Em 2013 o engenheiro norte americano Tobias Delbruck publicou um trabalho que chamou muita atenção devido a capacidade do sistema de efetuar o rastreamento de objetos em alta velocidade (Delbruck & Lang, 2013). O trabalho desenvolvido por ele consistiu em uma espécie de robô goleiro, que a partir do uso de uma câmera DVS128 foi capaz de rastrear bolas em alta velocidade e impedir que elas entrassem dentro do gol. O sistema foi capaz de rastrear e corrigir a posição do goleiro em um período de 3 ms (Figura 12a).

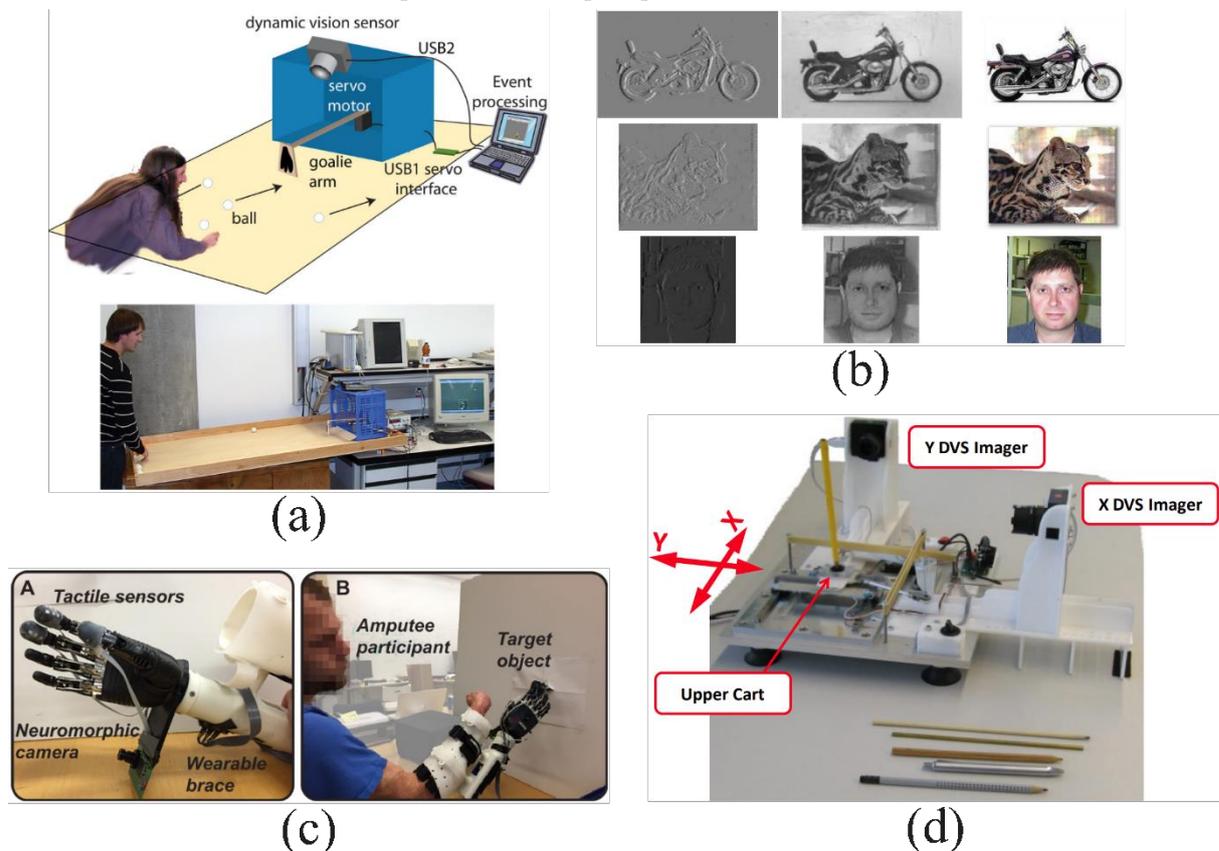
Outro trabalho com aplicação direta na área da robótica foi desenvolvido pelo pesquisador Jorg Conradt juntamente com outros nomes de peso, como Patrick Lichtsteiner e Tobias Delbruck (Conradt et al., 2009). Dessa vez, eles desenvolveram um sistema composto por um par de câmeras DVS128 para equilibrar um lápis. Desse modo, o sistema foi capaz de reajustar rapidamente a posição da base em que o lápis foi colocado à medida que ele pendia para uma direção (Figura 12d). Embora o uso de duas câmeras aumente o custo computacional, a combinação delas gera uma visão estéreo, que permite que o robô tenha uma noção de profundidade da cena.

Em 2019, o pesquisador Henri Rebecq juntamente com outros pesquisadores usou a Samsung DVS combinada com um modelo de redes neurais convolucionais recorrentes para efetuar a reconstrução de vídeos de objetos em alta velocidade (Rebecq et al., 2019). Como resultado, os vídeos reconstruídos pelo sistema desenvolvido alcançaram incríveis 4800 FPS (Figura 12b). Além disso, eles compararam o vídeo obtido por meio do sistema reconstrução desenvolvido por eles com outros vídeos da mesma cena gravado por meio de diferentes tipos de câmeras comerciais.

Recentemente, alguns grupos de pesquisadores ampliaram os horizontes para as aplicações de câmeras neuromórficas, como para sistemas semiautônomos de controle de próteses. Em 2019, o pesquisador Mark Hays e colaboradores do departamento de engenharia biomédica da Johns Hopkins publicaram um sistema responsável por automatizar as etapas de

seleção de preensão e angulação do punho protético (Hays et al., 2019). O principal objetivo desse trabalho consistiu em reduzir a complexidade do controle de sistema de próteses de membros superiores (Figura 12c).

Figura 12 – Exemplos de aplicações de câmeras baseadas em eventos. (a) Robotic Goalie desenvolvido para interceptar bolas em alta velocidade (3 ms) indo em direção ao gol. (b) Sistema de reconstrução de vídeos e imagens a partir de eventos obtidos de uma câmera DVS da Samsung. A imagem mostra três colunas de três exemplos dos eventos obtidos da câmera (esquerda), da reconstrução realizada pelo sistema desenvolvido (centro) e a imagem obtida por meio de uma câmera convencional (direita). (c) Sistema semiautônomo para controle de próteses com o auxílio de uma câmera baseada em eventos. (d) Robô com rápida resposta no controle do equilíbrio de um lápis a partir de visão estéreo.



Fonte: (a) (Delbruck & Lang, 2013), (b) (Rebecq et al., 2019), (c) (Hays et al., 2019), (d) (Conradt et al., 2009)

Atualmente, as câmeras neuromórficas já mostraram que possuem um grande potencial em aplicações de visão computacional e robótica em cenários desafiadores atualmente inacessíveis às câmeras tradicionais (Gallego et al., 2020). Entretanto, a comparação direta entre essas câmeras e as convencionais estão longe de serem justas, uma vez que as câmeras convencionais já estão consolidadas tanto no campo das pesquisas quanto em sistemas comerciais.

# Capítulo 5

---

## 5. VISÃO COMPUTACIONAL

Como descrito no capítulo 4, existem diferentes tipos de câmeras capazes de gerar imagens próximas ao que a nossa visão nos fornece. Por sua vez, essas imagens são de grande importância para o ser humano. Por exemplo, para trazer à tona um momento vivido no passado em que tal imagem foi capturada. No entanto, com o constante avanço da tecnologia, as informações captadas por uma câmera passaram a ter outras características relevantes, de forma que ao combinar essas imagens com recursos computacionais tornou-se possível fornecer visão para máquinas, dando origem a uma das áreas mais estudadas atualmente: a visão computacional. Consequentemente, essa ferramenta foi utilizada para desenvolver diversas aplicações capazes de facilitar algumas atividades importantes presentes no dia a dia. Entre elas, se destacam o reconhecimento de faces, imagens médicas, carros autônomos e rastreamento de objetos (Szeliski, 2010).

Foram utilizadas três importantes técnicas presentes na área de visão computacional para a realização desse trabalho, sendo elas as Redes Neurais Convolucionais, a Análise de Componentes Principais e a segmentação de imagens por meio do algoritmo de Watershed. Desse modo, nos subtópicos seguintes essas técnicas serão apresentadas e discutidas individualmente.

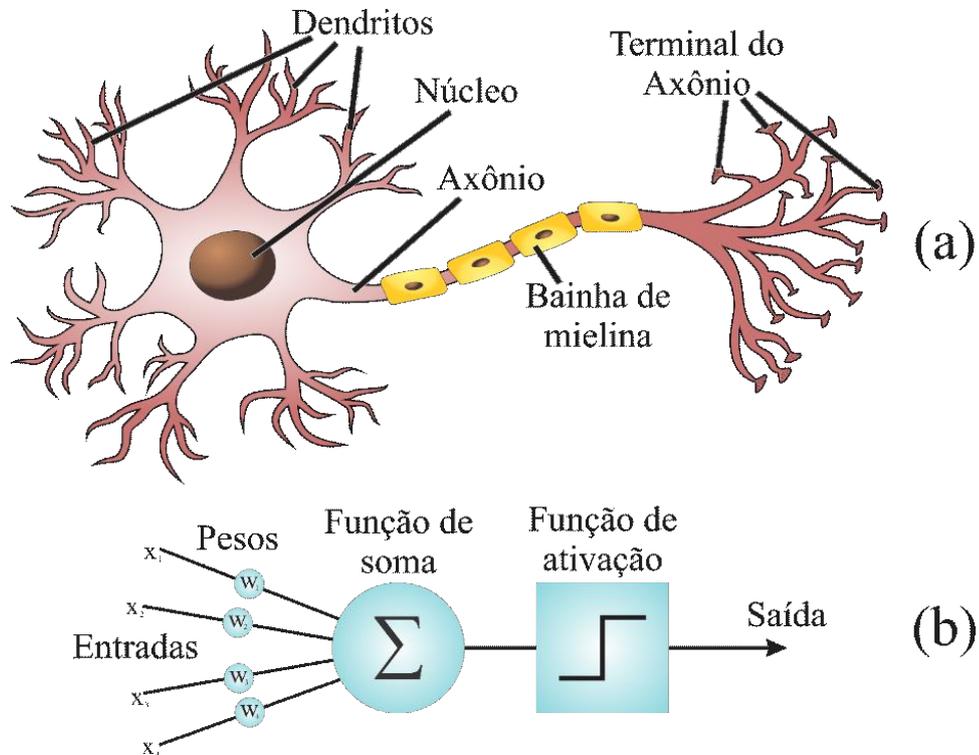
### 5.1 REDES NEURAIAS CONVOLUCIONAIS

Criar sistemas inteligentes capazes de tomar decisões de acordo com dados recebidos foi um grande desafio do último século. De fato, muitos modelos matemáticos e estatísticos foram desenvolvidos, os quais são populares até os dias de hoje. Entre eles, o perceptron, desenvolvido pelo psicólogo norte americano Frank Rosenblatt em 1958 (Rosenblatt, 1958). Inspirado no funcionamento do sistema nervoso humano, Rosenblatt tinha como objetivo desenvolver um modelo matemático que possuísse as principais funções de um neurônio.

A Figura 13 mostra uma comparação entre um neurônio com um perceptron. Observando a figura, podemos ver que assim como um neurônio, o perceptron recebe várias entradas cada

uma com um peso, de modo que quanto maior o peso, maior é a relevância dessa entrada. Assim, o perceptron é capaz de gerar um pulso de saída ou não de acordo com as suas entradas.

Figura 13 – Comparação entre um (a) neurônio e um (b) perceptron.



Assim como um neurônio, um único perceptron não é capaz de realizar tarefas de classificação complexas, mas sim problemas lineares. Entretanto, quando combinados em múltiplas camadas eles se tornam uma ferramenta muito poderosa, capaz de resolver problemas de alta complexidade. Essa combinação de vários perceptrons é chamada de Perceptrons de Multicamadas (MLP).

O modelo do perceptron foi muito importante para diversos modelos de Redes Neurais (NN), entre elas têm-se as Redes Neurais Artificiais (ANN), Redes Neurais Recorrentes (RNN), Redes Neuras de Eventos (SNN) e Redes Neurais Convolucionais (CNN). Porém, esses modelos de NN não obtiveram um crescimento muito grande até no fim do século passado. O principal motivo do pouco crescimento desses modelos foi devido à sua alta complexidade, exigindo um custo computacional maior que os computadores da época conseguiam fornecer. Dessa forma, era mais viável a utilização de modelos de AM mais simples, como as Máquinas de Vetores de Suporte (SVM) e Análises Discriminantes Lineares (LDA). A situação foi se alterando à medida que computadores cada vez mais potentes eram desenvolvidos, desencadeando o grande potencial existente nos modelos de NN.

Os modelos de NN têm sido de grande importância na área de visão computacional. Entretanto, a etapa de extração de atributos de uma imagem foi um grande desafio durante muitos anos. Com sua primeira aparição em 1998 no trabalho desenvolvido pelo cientista de computação francês Yann LeCun, as CNN foram as que obtiveram o maior destaque na área de visão computacional (LeCun et al., 1998). Esse destaque é devido principalmente a sua estratégia de automatização da etapa de extração de atributos de uma imagem. Assim, todo o sistema de aprendizagem passou a ser feito a partir de dados brutos de uma imagem.

### 5.1.1 ESTRUTURA DE UMA CNN

As camadas hierárquicas são um dos aspectos mais interessantes da estrutura de uma CNN, pois são diretamente inspiradas pelas noções clássicas das células da neurociência visual, de modo que seus níveis de hierarquia são similares ao LGN – V1 – V2 – V4 – IT na via ventral do córtex visual (Lecun et al., 2015). Assim, os recursos de nível superior são obtidos pela composição dos de nível inferior. Ou seja, o conjunto de bordas formam um contorno. Por sua vez, um conjunto de contornos formam a estrutura de um objeto, e um conjuntos de objetos formam uma cena contida em uma imagem.

A estrutura de uma CNN é composta por três tipos de camadas: Convolutacional, Pooling e Fully-Connected.

- Convolutacional:

O objetivo principal da convolução é extrair atributos da imagem de entrada. A convolução preserva a relação espacial entre pixels, aprendendo recursos de imagem usando pequenos quadrados de dados de entrada. A camada de convolução é o bloco de construção central da CNN. Ele carrega a parte principal da carga computacional da rede. Esta camada realiza um produto escalar entre duas matrizes, onde uma matriz representa uma parte da imagem de entrada e a outra matriz é o conjunto de parâmetros aprendíveis, também chamada de kernel ou filtro. A matriz de saída é a junção de todos os produtos escalares obtidos após o kernel percorrer toda a imagem.

- Pooling:

A camada de Pooling pode ser resumida como a camada responsável por reduzir a dimensionalidade da camada de entrada. Isso diminui a quantidade necessária de computação e pesos. Além disso, o processo de Pooling fornece alguma invariância de translação, o que significa que um objeto seria reconhecível independentemente de onde aparecesse no quadro.

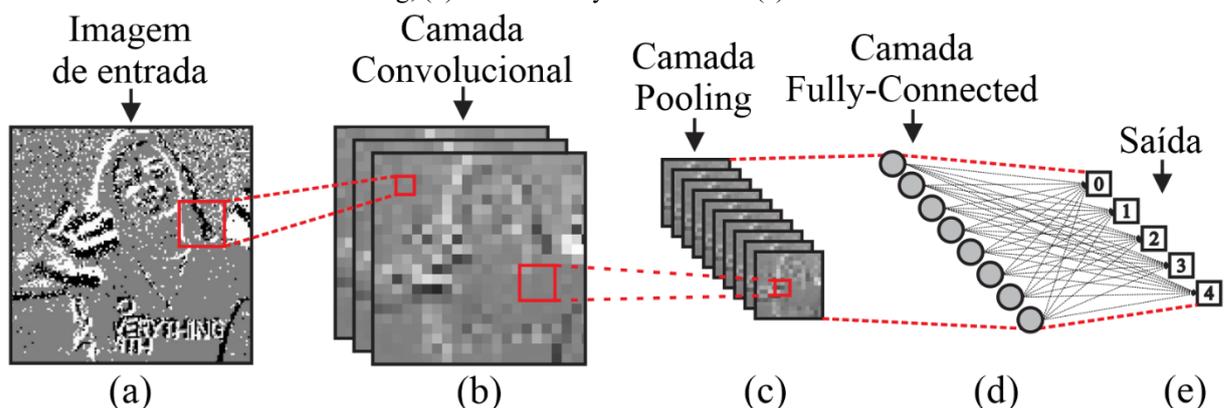
A operação de agrupamento é processada em cada fatia da representação individualmente. Existem várias funções de agrupamento, como a média da vizinhança retangular e uma média ponderada com base na distância do pixel central. No entanto, o processo mais popular é o Pooling máximo, que relata a saída máxima da vizinhança.

- Fully-Connected:

A saída das camadas convolucionais e de Pooling representam recursos de alto nível da imagem de entrada. Ou seja, basicamente são responsáveis pela extração de atributos dos dados de entrada. Já a camada Fully-Connected, tem como objetivo usar esses recursos para classificar a imagem de entrada em várias classes com base no conjunto de dados de treinamento.

O termo “Fully-Connected” implica que todos os neurônios da camada anterior estão conectados a todos os neurônios da próxima camada. Normalmente, a camada Fully-Connected é uma MLP tradicional que usa uma função de ativação de softmax na camada de saída. Entretanto, outras técnicas de aprendizado supervisionado como o SVM e a LDA também podem ser usados.

Figura 14 – Estrutura simplificada de uma CNN. (a) Imagem de entrada, (b) camada Convolutacional, (c) camada de Pooling, (d) camada Fully-Connected e (e) saída da rede.



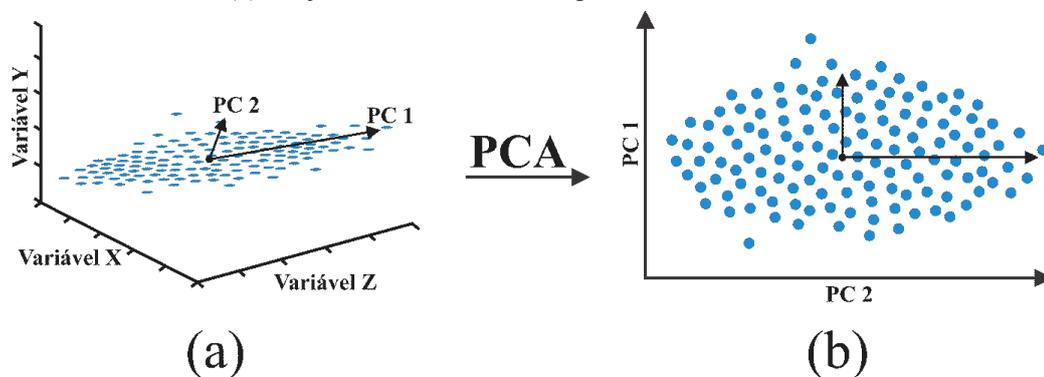
Embora a junção de pelo menos uma camada de cada tipo seja suficiente para formar uma CNN, os modelos mais comuns possuem várias camadas de cada um dos tipos, aprimorando significativamente o modelo.

## 5.2 ANÁLISE DE COMPONENTES PRINCIPAIS

Desenvolvida pelo matemático inglês Karl Pearson em 1901, a Análise de Componentes Principais, ou do inglês *Principal Component Analysis* (PCA) é uma ferramenta matemática de suma importância quando se trabalha com conjuntos de dados (Pearson, 1901). Sua principal aplicação consiste na redução na dimensionalidade de um conjunto de dados que possui muitas variáveis inter-relacionadas, enquanto retém o máximo possível da variação presente no conjunto de dados (Jolliffe, 1986).

Com o uso da PCA é possível identificar padrões em um certo conjunto multidimensional de dados, buscando reduzir a quantidade de informações redundantes e destacar os eixos de maior variabilidade. Dessa forma, os dados originais são reorganizados em um novo conjunto de variáveis, os componentes principais (PCs), que não são correlacionados e são ordenados de forma que os primeiros retenham a maior parte da variação presente em todas as variáveis originais (Figura 15).

Figura 15 – Transformação de conjunto de dados usando PCA. (a) Conjunto de dados original contendo três dimensões. (b) Conjunto de dados obtidos a partir das duas PCs mais relevantes.



Cada CP é responsável por uma porcentagem da variabilidade total dos dados contidos no conjunto original. Dessa forma, o número de PCs sempre será menor ou igual ao número de variáveis contidas no conjunto de dados originais, e a combinação de todas as PCs representa 100% da variabilidade do conjunto original. Entretanto, em grande parte das vezes, poucas PCs são suficientes para representar o conjunto de dados originais, de modo que as PCs de menores variabilidades podem ser desconsideradas.

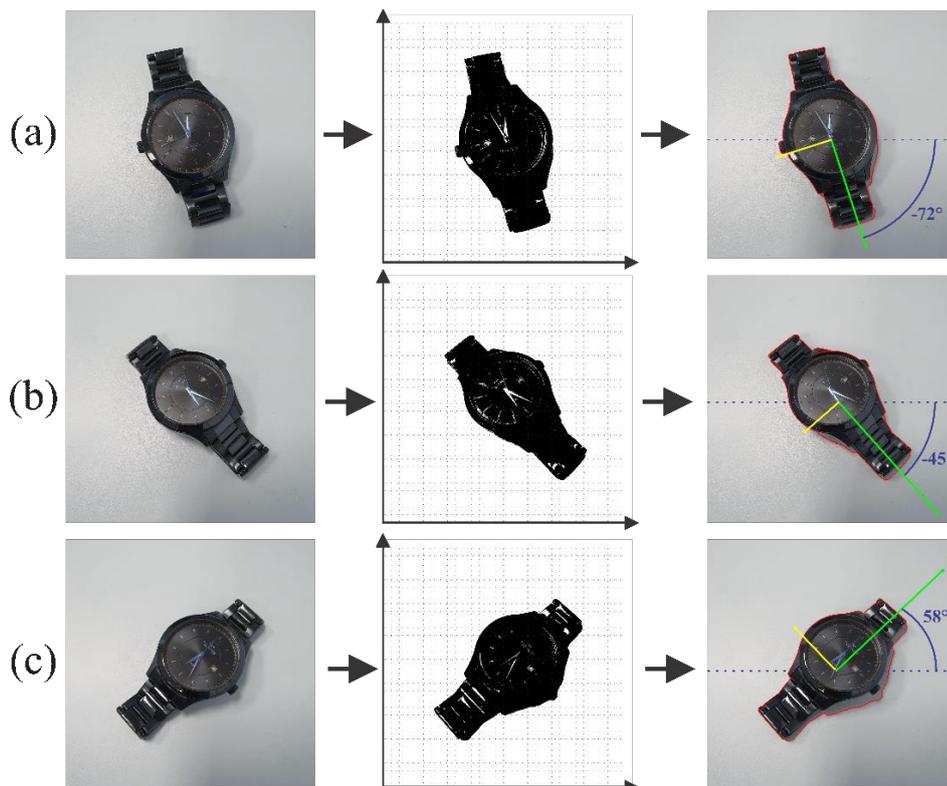
Nos últimos anos, a PCA esteve muito presente na área de ML, pois muitas vezes os conjuntos de atributos originais são extensos e parcialmente redundantes. Levando em conta que algoritmos de ML tem seu desempenho diretamente afetado pela variabilidade dos dados e tamanho do conjunto de dados, a técnica de PCA se torna uma ferramenta útil para reduzir o

número de variáveis mesmo mantendo uma porcentagem alta da variabilidade total do conjunto de dados original.

### 5.2.1 VISÃO COMPUTACIONAL E PCA: ORIENTAÇÃO DE OBJETOS

Embora a aplicação mais comum da PCA seja para redução de dimensionalidade de conjunto de dados multidimensionais, é possível utilizá-la para obter a orientação de objetos em imagens e vídeos. Essa abordagem pode ser interessante em aplicações da área de visão computacional em que a orientação do objeto contido na imagem é relevante (Yi & Marshall, 2000).

Figura 16 – Obtenção das PCs de um relógio posicionado em três orientações diferentes. A linha verde indica a PC 1 e a linha amarela representa a PC 2. O relógio foi posicionado em: (a) -72 (b) -45 e (c) 58 graus de acordo com a PC 1.



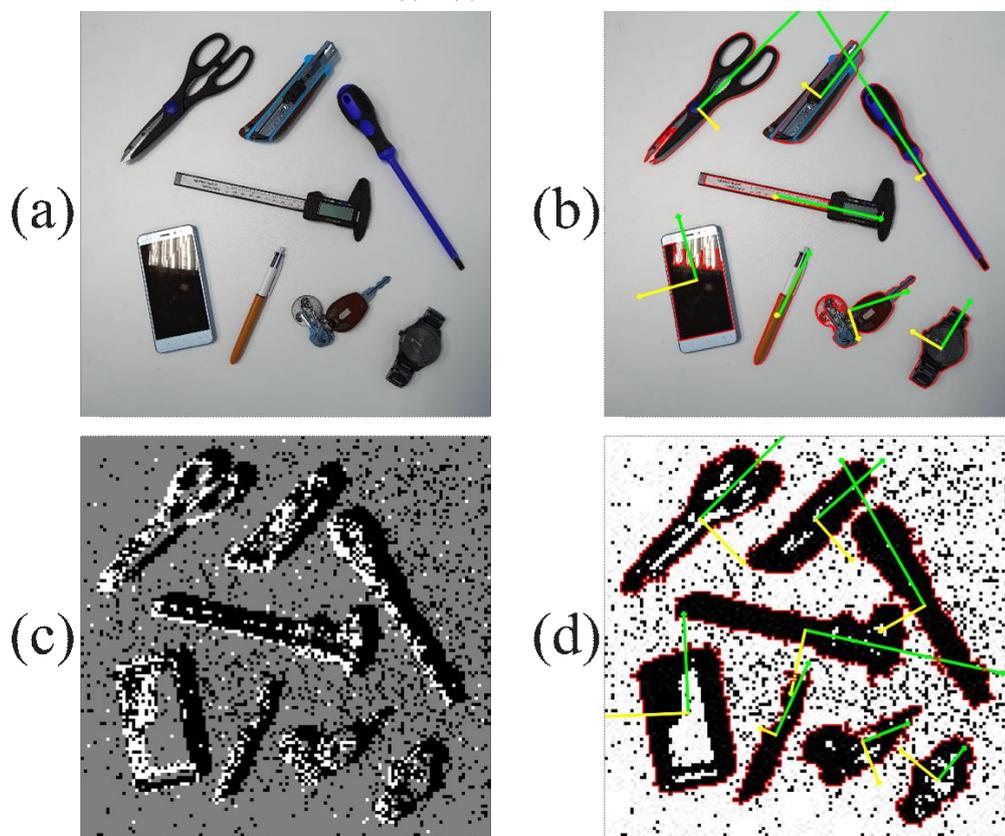
Considere que existe um objeto em uma imagem que estamos interessados em obter sua orientação de acordo com seu lado mais longo, de modo que cada dimensão corresponde a um eixo da imagem onde o objeto está inserido (Figura 16). O primeiro passo para obter a orientação do objeto é transformar a imagem para preto e branco, fazendo com que a imagem se torne apenas o objeto formado por uma nuvem de pontos. Ao observar a segunda coluna da

Figura 16, é possível notar que temos um conjunto de dados de duas dimensões. A partir desses dados, é possível obter os dois eixos de maior variabilidade do conjunto de dados por meio da utilização da PCA.

Os novos eixos obtidos como resultado da execução da PCA consistem em dois vetores chamados autovetores, que são também as PCs do conjunto de dados. O tamanho de cada autovetor é codificado em um autovalor correspondente, que indica o quanto os dados variam ao longo da PC. O início dos vetores próprios é o centro de todos os pontos no conjunto de dados.

Como descrito no capítulo 4, uma câmera baseada em eventos nos fornece uma imagem binária se desconsiderarmos as polaridades dos eventos. Ou seja, as imagens geradas por essas câmeras já fornecem uma nuvem de pontos, evitando boa parte do processamento necessário para deixar a imagem apta para a utilização da PCA. A Figura 17 ilustra o resultado da obtenção da orientação dos objetos contidos em uma imagem gerada por ambas as câmeras convencional e baseada em eventos.

Figura 17 – Exemplo do uso da PCA para obtenção da orientação de múltiplos objetos. (a) e (b) câmera convencional. (c) e (d) câmera baseada em eventos.



Apesar de existirem outros métodos de para obter a orientação de um objeto, por exemplo um modelo de CNN, a técnica de PCA se mostra muito eficaz. Isso fica mais evidente quando

se trabalha com aplicações que necessitam de um processamento rápido, e conseqüentemente de modelos mais simples. Nesses casos, o uso de uma CNN pode afetar diretamente no desempenho do sistema já que seu processamento gera um alto custo computacional.

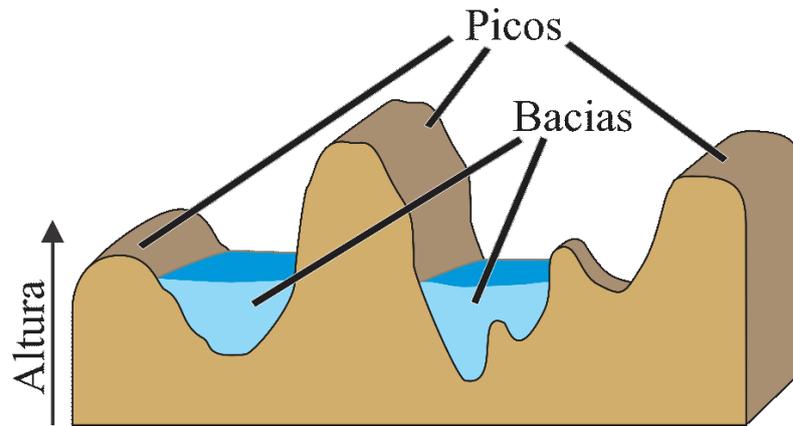
### 5.3 SEGMENTAÇÃO DE IMAGENS – ALGORITMO DE WATERSHED

No processamento digital de imagens e na visão computacional, a segmentação da imagem é o processo de separar uma imagem em vários segmentos de acordo com o seu conteúdo. O objetivo da segmentação é extrair informações de determinadas partes de uma imagem para possibilitar uma análise posterior mais simples e precisa dessa região de interesse (ROI). Em outras palavras, a segmentação da imagem é o processo de atribuir um rótulo a cada pixel de uma imagem de forma que os pixels com o mesmo rótulo compartilhem certas características.

Dentre as técnicas existentes, o algoritmo de Watershed é um dos algoritmos de segmentação mais clássicos, que tem suas origens na morfologia matemática. O conceito geral foi introduzido pelos matemáticos franceses Digabel e Christian Lantuejoul em 1978 e aprimorado por Serge Beucher, outro matemático francês em 1979 (Beucher & Lantuejoul, 1979; Digabel & Lantuejoul, 1978). Um grande avanço na aplicabilidade foi alcançado em 1991 por Luc Vincent e Pierre Soille, que apresentaram um algoritmo mais rápido e mais preciso do que os anteriores (Vincent et al., 1991). Desde então, seu uso se estende para diversas aplicações relacionadas ao campo de visão computacional, como no processamento de imagens médicas, ciência de materiais e rastreamento de objetos (Atta-Fosu et al., 2016; Pham et al., 2000; Waggoner et al., 2013; Wang, 1998; Yang et al., 2006).

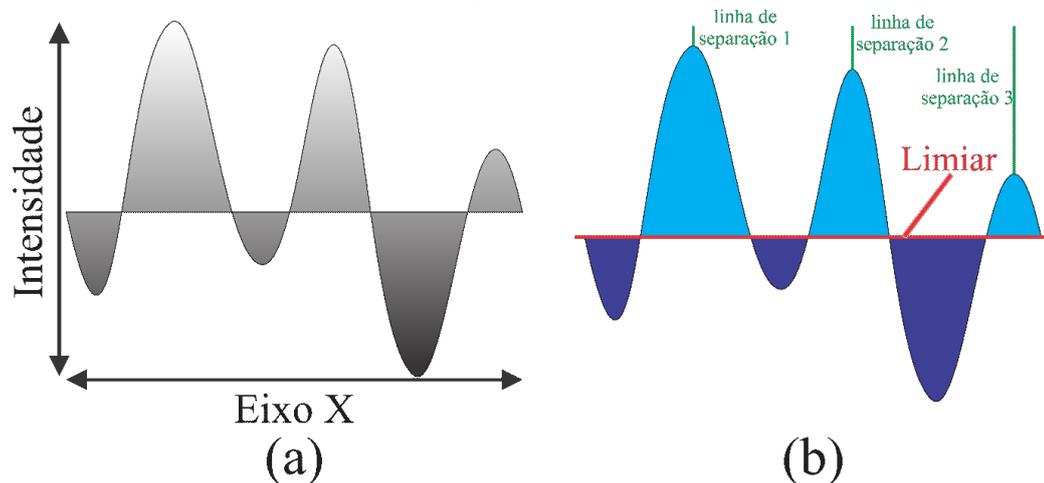
O algoritmo de Watershed baseia-se na representação de uma imagem em tons de cinza, de modo que cada tom de cinza representa um nível. Para exemplificar isso, considere uma paisagem topográfica, a qual possui diferentes vales e picos ao longo de suas dimensões (Figura 18). Agora, considere que começamos a encher todos os vales isolados com água. À medida que o nível de água sobe, cada um dos picos locais vai sendo coberto pela água de acordo com sua altura, fazendo com que as águas contidas em cada um dos vales se misturem. Para evitar que isso ocorra, é preciso que sejam construídas barreiras delimitadoras nos locais onde a água se fundirá, as quais serão usadas para determinar os limites de cada um dos segmentos.

Figura 18 – Exemplificação de uma paisagem topográfica com diferentes níveis e bacias de água.



Assim como no exemplo acima, a segmentação por meio do algoritmo de Watershed tem o objetivo de separar diferentes áreas de acordo com seus respectivos níveis. Nesse caso, os tons de cinza da imagem equivalem aos picos e vales da paisagem topográfica, de forma que quanto mais próximo do tom de branco maior é o nível desse segmento (Figura 19a). Na Figura 19b, é possível ver que à medida que aumentamos o limiar em vermelho, eventualmente os segmentos de serão mesclados. Portanto, precisamos colocar as linhas de separação entre os seguimentos.

Figura 19 – (a) Intensidade dos pixels de uma imagem ao longo do eixo X. (b) segmentação das áreas de acordo com a sua intensidade. À medida que o limiar sobe, novas barreiras são criadas.



Na Figura 19 é mostrado um exemplo simples e com poucas variações. Porém, ao se trabalhar com imagens reais, vários pontos de mínimos locais são criados, onde todos os outros pixels que apresentam valores compreendidos em um gradiente descendente que termina em um ponto de mínimo local são agrupados e formam uma região que são separados por linhas de separação (Preim & Botha, 2014). Dessa forma, ao atribuir um rótulo de bacia ou bacia

hidrográfica a cada pixel, as imagens com ruído resultarão em várias pequenas regiões, o que gera um fenômeno de supersegmentação.

Na seção 5.2.1 deste capítulo, foi apresentada uma aplicação da PCA para obter a orientação de um objeto contido em uma imagem. Porém, até então não foi comentado que no exemplo apresentado na Figura 16 e Figura 17 foi utilizado um algoritmo de Watershed para fazer a segmentação desses objetos. As delimitações obtidas após a utilização do algoritmo estão representadas pela linha vermelha que contorna cada um dos objetos.

# Capítulo 6

---

## 6. METODOLOGIA

Neste capítulo, é feita uma descrição detalhada dos materiais e métodos que foram utilizados para a realização deste trabalho. A metodologia foi dividida em duas etapas principais, sendo elas a etapa de experimentos offline e a etapa de experimentos em tempo real.

Deste ponto em diante, para melhor facilitar o entendimento dos diversos cálculos e processamentos realizados, será utilizado o termo “imagem” para descrever operações em uma imagem individual, e o termo “quadros” para descrever operações em um fluxo temporal de imagens (vídeo).

### 6.1 EXPERIMENTOS OFFLINE

A etapa de experimentos offline é dividida em três fases. Primeiro, foram desenvolvidos os scripts em C++ e Python para efetuar a aquisição e o condicionamento dos dados recebidos da câmera neuromórfica. Na sequência, foi criado um conjunto de dados com a mesma câmera neuromórfica, contendo objetos frequentemente usados em atividades do dia a dia. Por último, foi desenvolvido um algoritmo em Python capaz de segmentar o objeto principal de uma cena com vários objetos, classificar o tipo de preensão necessária de acordo com a forma do objeto e ajustar a orientação do punho da prótese de acordo com a angulação do objeto.

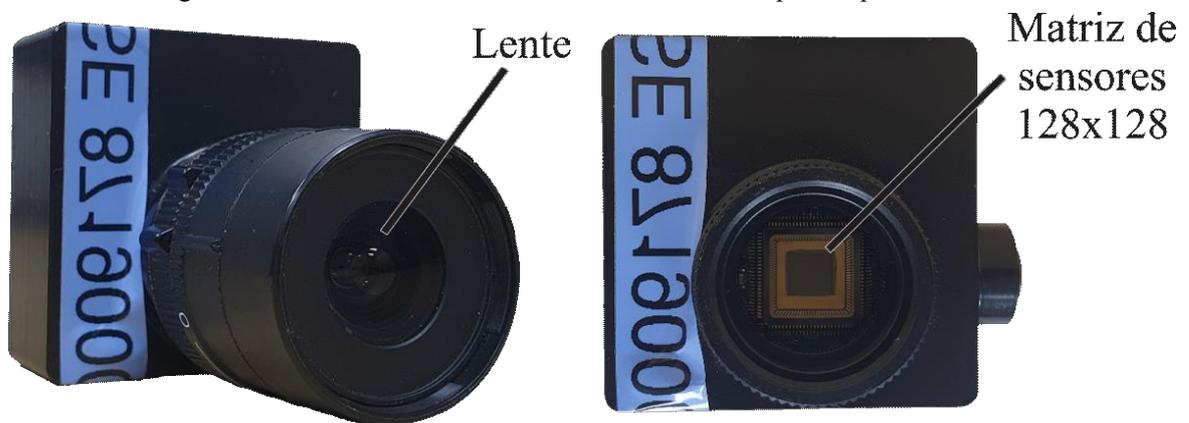
#### 6.1.1 AQUISIÇÃO E CONDICIONAMENTO DE DADOS

A etapa de aquisição e condicionamento é essencial para as etapas subsequentes, pois é responsável por todo o pré-processamento necessário para transformar os dados brutos recebidos da câmera neuromórfica em quadros, os quais serão utilizados em todas as demais etapas. Dessa forma, essa etapa foi subdividida em três fases: seleção da câmera neuromórfica, aquisição de eventos e conversão de eventos para quadros. Cada uma dessas fases está descrita de forma detalhada nos tópicos posteriores.

### 6.1.1.1 *Dynamic Vision Sensor 128*

A câmera escolhida para esse trabalho foi a *Dynamic Vision Sensor 128* (DVS128). Essa câmera foi uma das pioneiras na área de visão baseada em eventos, de modo que tem sido utilizada em grande parte dos trabalhos relacionados a essa área. Ela é composta por uma matriz de 128 por 128 pixels, resultando em um total de 16384 sensores, os quais podem disparar de modo assíncrono a cada 15  $\mu$ s. A Figura 20 mostra a câmera DVS128 que foi utilizada nos experimentos presentes nesse trabalho.

Figura 20 – Câmera neuromórfica DVS128 desenvolvida pela empresa Inivation.



A DVS128 possui dimensões externas de 50x56x66 mm, enquanto sua matriz de sensores tem dimensões de 15x15 mm. Embora suas dimensões externas sejam relativamente grandes, esse modelo de câmera neuromórfica é amplamente capaz de cumprir com o objetivo dentro do sistema proposto.

### 6.1.1.2 *Aquisição de eventos*

Grande parte do processo de aquisição dos dados da câmera DVS128 foi desenvolvido pela empresa suíça Inivation e disponibilizada gratuitamente em sua página do GitHub (*Inivation*, n.d.). Eles desenvolveram uma série de bibliotecas escritas em C++ (cAER) e Java (jAER) capazes de configurar e receber os dados da câmera DVS128. A estrutura de cada um dos eventos recebidos é composta por 16 bits, que armazenam as informações de polaridade, endereço X, endereço Y e eventos externos, e 32 bits que armazenam o tempo que ocorreu o evento. Dessa forma, cada evento é armazenado em 48 bits, ou 6 bytes (Tabela 1).

Tabela 1 – Divisão dos bits que contém as informações essenciais dos dados gravados pela câmera DVS128.

<b>Bits</b>	<b>Significado</b>	<b>Descrição</b>
16-47	Tempo	Tempo que ocorreu o evento ( $\mu$ s)
15	Evento externo	Evento externo detectado no pino IN
14-8	Endereço Y	Endereço Y do evento. (0, 0) no canto esquerdo inferior da tela.
7-1	Endereço X	Endereço X do evento. (0, 0) no canto esquerdo inferior da tela.
0	Polaridade	Mudança de luminosidade: '1' significa aumento (ON), '0' significa diminuição (OFF).

Existem dois modos de ler os dados obtidos da câmera DVS128, que são por meio de uma gravação (offline) ou em tempo real (online).

#### 6.1.1.2.1 Offline

Os dados gravados por meio da câmera DVS128 são salvos em um formato de arquivo específico denominado de AEDAT. Assim, foi necessário desenvolver um script em Python para decodificar os dados da gravação contidos nos arquivos.

A estrutura dos dados contidos no arquivo segue a mesma codificação da Tabela 1, de modo que os dados precisam ser decodificados para obter as informações geradas pela câmera. A Tabela 2 mostra a decodificação feita por meio do uso de uma máscara de bits e o deslocamento dos bits de acordo com a posição da sua respectiva informação, isto é, endereço X, endereço Y e polaridade.

Tabela 2 – Decodificação dos dados de endereço e polaridade contidos em um arquivo de gravação da DVS128. Na coluna de Posição, o “X” representa as posições onde as respectivas informações estão armazenadas.

	<b>Posição</b>	<b>Máscara</b>	<b>Deslocamento</b>	<b>Resultado</b>
Endereço Y	0XXX XXXX 0000 0000	0111 1111 0000 0000	8	0 até 127
Endereço X	0000 0000 XXXX XXX0	0000 0000 1111 1110	1	0 até 127
Polaridade	0000 0000 0000 000X	0000 0000 0000 0001	0	0 ou 1

A partir dos valores obtidos dos primeiros bytes de cada evento, pode-se utilizar a Equação 1 para obter o resultado de cada uma das informações.

$$E \& m \gg S = V \quad (1)$$

Onde  $E$  é a informação para decodificar,  $m$  é a sua máscara,  $S$  é o seu deslocamento e  $V$  é o valor decodificado, que pode ser entre 0 e 127 para os endereços  $X$  e  $Y$ , e 0 ou 1 para a polaridade.

Após a decodificação completa do arquivo, obtêm-se quatro vetores de tamanhos iguais, respectivos aos valores de polaridade, endereço  $X$ , endereço  $Y$  e tempo. Dessa forma, a combinação das posições de mesmo índice dos quatro vetores indica um evento gerado pela câmera.

Agora, com os quatro vetores decodificados, os dados estão prontos para a etapa de conversão de eventos para quadros descrita na seção 6.1.1.3.

#### 6.1.1.2.2 *Online*

Para ler os dados em tempo real, foram utilizadas as bibliotecas de aquisição de dados em C++ desenvolvidas pela empresa Inivation. Após receber esses dados em C++, foi necessário enviá-los para um script em Python, onde serão realizados os processamentos posteriores.

Para enviar as informações contidas dentro do script em C++ para o novo script em Python, foi utilizada a técnica de Socket com o protocolo UDP (do inglês *User Datagram Protocol*). Esse protocolo é comumente utilizado na transmissão de informações dentro de uma mesma máquina ou entre máquinas diferentes, onde a velocidade no envio das informações é mais importante do que qualquer outro parâmetro. As principais aplicações estão envolvidas no envio de dados de áudio e vídeo.

Levando em consideração que nesse trabalho foi escolhido um modelo de câmera que se baseia principalmente na velocidade na aquisição de informações, o protocolo UDP é ideal para a transmissão dessas informações entre diferentes linguagens de programação.

A partir desse ponto, praticamente todos os processamentos serão realizados em Python, pois é uma linguagem de alto nível muito eficiente e, além disso possui uma vasta quantidade de módulos de processamento de dados disponíveis.

#### 6.1.1.3 *Conversão de eventos para quadros*

Como descrito no capítulo 4, os dados obtidos por meio das câmeras neuromórficas não são síncronos e baseados em quadros como nas câmeras convencionais, mas sim baseados em um fluxo de eventos gerados em cada um dos pixels contidos na matriz de sensores de luminosidade. Dessa forma, após receber os dados de ambas as formas online e offline, foi

realizada a conversão dos eventos obtidos da câmera DVS128 para quadros, possibilitando o uso futuro de ferramentas importantes de processamento de imagens como o algoritmo de Watershed, CNN e PCA.

Como descrito na Equação 2, o processo de formação de um quadro foi baseado no acúmulo de eventos durante um período pré-determinado.

$$Q(t) = \sum_{n=t_i}^{t_f} E \quad (2)$$

Onde  $Q$  é o quadro gerado,  $E$  são os eventos obtidos da câmera DVS128,  $t_i$  é o tempo inicial e  $t_f$  é o tempo final do acúmulo dos eventos, de modo que a subtração de  $t_f$  por  $t_i$  gera o período de formação de um quadro  $t$ . Dessa forma, a formação de um quadro não depende do número de eventos gerados, mas sim do período em que os eventos foram acumulados.

Inicialmente,  $Q$  é um vetor de duas dimensões (128x128), contendo apenas valores iguais a 127. Como os valores do vetor são de 8 bits, tem-se que o valor mínimo possível é 0 e o valor máximo é 255. Dessa forma, o valor de 127 é equivalente a posição neutra do quadro, onde não existem eventos, enquanto os valores de 0 e 255 representam os eventos negativos e positivos, respectivamente. Após criar o vetor inicial, os eventos existentes no período  $t$  são inseridos na matriz de acordo com o seu endereço  $X$  e  $Y$  e polaridade, conforme descrito na Equação (3).

$$\text{Para cada } E \text{ em } L \rightarrow Q_{(x=E_x, y=E_y)} = \begin{cases} 0, & \text{se } E_p = 0 \\ 255, & \text{se } E_p = 1 \end{cases} \quad (3)$$

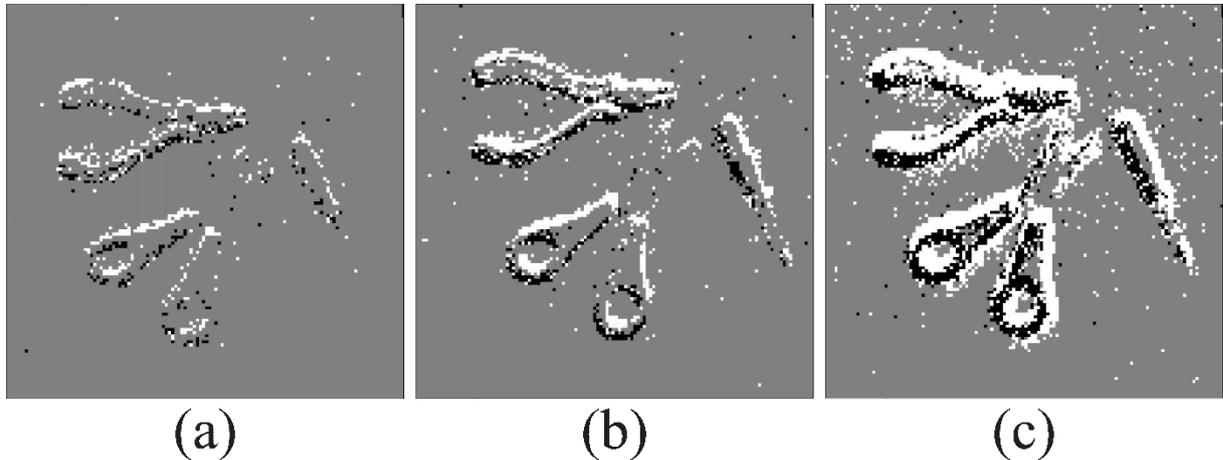
Onde  $E_p$ ,  $E_x$  e  $E_y$ , são respectivamente, os valores da polaridade, endereço  $X$  e endereço  $Y$  do evento  $E$  contido no vetor de eventos  $L$ .

O período de acúmulo de eventos pode ser comparado com o tempo de exposição de uma câmera convencional. Embora seja similar à taxa de FPS de um vídeo, o período em que os eventos serão acumulados é definido de acordo com a aplicação, possibilitando o aumento ou diminuição do número de FPS de uma gravação.

A possibilidade de ajustar o número de FPS dependendo da aplicação resulta em uma troca de benefícios e malefícios, de forma que ao se aumentar o FPS tem-se como resultado uma menor quantidade eventos por quadro; todavia, tem-se uma redução na sobreposição de dados e no intervalo entre quadros, permitindo a análise dos dados em alta velocidade. Por outro lado, ao se diminuir o FPS tem-se como resultado um maior número de dados redundantes; porém, existem mais dados em um mesmo quadro, possibilitando uma análise mais precisa das

informações contidas em cada quadro. A Figura 21 apresenta a mesma gravação, todavia com diferentes números de FPS, exemplificando os prós e contras descritos acima.

Figura 21 – Três imagens de uma mesma gravação, todavia cada uma com um número de FPS diferente. (a) 1000 FPS, (b) 100 FPS e (c) 10 FPS.



Para aplicações que exigem um tempo de resposta baixo, como rastreamento de objetos em altas velocidades, é uma boa escolha ter um número de FPS mais alto, evitando que o algoritmo perca o objeto na cena. Entretanto, o tempo de resposta necessário para esse trabalho não precisa necessariamente ser muito baixo, pois estamos interessados em informações como forma e orientação do objeto. Assim, foi escolhido um período  $t$  de 40 ms, que é equivalente a uma taxa de 25 FPS.

### 6.1.2 CONJUNTO DE DADOS

Para realizar a etapa de classificação do tipo de preensão, é necessária a utilização de um conjunto de dados. Entretanto, o baixo número de conjunto de dados coletados a partir de câmeras neuromórficas disponíveis é uma das suas principais desvantagens em relação às câmeras convencionais.

Devido aos conjuntos de dados serem limitados, é preciso coletar novos dados, de modo que seja criado um conjunto de dados que satisfaça as necessidades da aplicação em questão. Dessa forma, a primeira etapa desse trabalho consistiu na construção de um novo conjunto de dados, o qual tenha gravações de alguns dos principais objetos comumente utilizados em atividades do dia a dia. Esse conjunto de dados foi utilizado posteriormente para o treinamento e validação do modelo de aprendizado utilizado.

### ***6.1.2.1 Seleção de objetos e definição dos tipos de preensão***

Antes de descrever os objetos que foram escolhidos para fazer parte desse conjunto de dados, é importante explicar como cada um foi rotulado, de modo que esses rótulos são fundamentais para entender a etapa de classificação do tipo de preensão e os experimentos em tempo real usando o sistema de controle proposto.

Existem duas classes de saída, sendo elas Tripod ou Power. A classe Tripod compreende objetos pequenos e delicados. Ou seja, para realizar a preensão desses objetos é preciso ter mais sutileza do que força, usando apenas os dedos polegar, indicador e médio da mão. Por outro lado, a classe Power compreende objetos grandes. Isto é, para realizar a preensão desses objetos é preciso ter mais pontos de contato com o objeto, usando todos os dedos da mão.

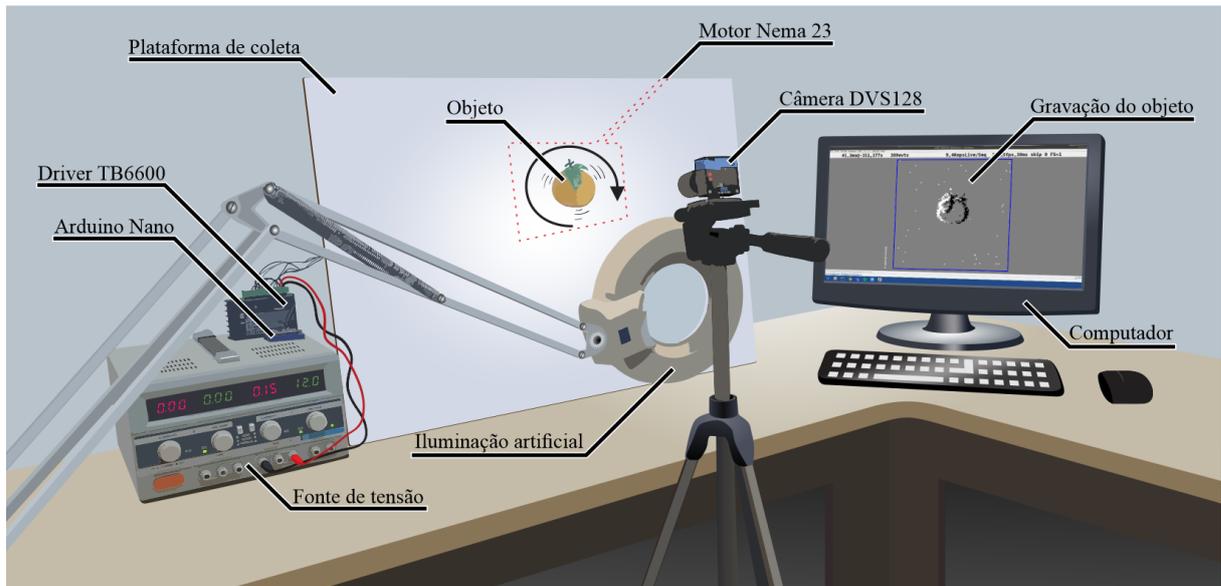
Existe uma grande variedade de objetos com os quais temos contato ao longo da nossa vida. Entretanto, alguns desses objetos estão presentes constantemente no nosso dia a dia, exercendo uma função importante e, em alguns casos, indispensável. Dessa forma, foram escolhidos 30 objetos comuns em atividades do dia a dia na construção do conjunto de dados presente nesse trabalho. Assim, foram rotulados 15 objetos como preensão Power e 15 como preensão Tripod. Todos os objetos e seus respectivos rótulos contidos no conjunto de dados criado estão apresentados na Tabela A1 do apêndice B.

### ***6.1.2.2 Gravação do conjunto de objetos***

Foi desenvolvido um cenário para realizar a gravação dos objetos selecionados. Posteriormente, essas gravações serão convertidas em imagens, que por sua vez, formarão o conjunto de dados usados para o treinar o modelo de CNN proposto.

A Figura 22 ilustra o cenário criado e os principais materiais utilizados na realização da gravação. Primeiramente, um motor de passo Nema 23 foi posicionado na parte de trás da plataforma de coleta. Em seguida, o objeto foi anexado diretamente em seu eixo por meio de uma fita de velcro. Dessa forma, o objeto rotaciona sempre que o motor for acionado. Por último, a câmera DVS128 foi posicionada na mesma altura do objeto e com uma distância horizontal de 30 cm. Após a montagem da estrutura de coleta, iniciou-se de fato a gravação dos objetos. Cada um dos 30 objetos selecionados foi gravado durante uma volta completa do eixo do motor, equivalente a 20 segundos.

Figura 22 – Ambiente de gravação dos objetos. No exemplo em questão, foi realizada a gravação de uma laranja.



Foi escolhido um motor de passo para essa etapa pois eles fornecem informações importantes como o ângulo de cada passo, podendo então ajustar facilmente a velocidade de rotação. Em específico, o motor Nema 23 possui um ângulo de passo de 1,8 graus. Entretanto, esse número pode ser reduzido quando combinado com o driver TB6600, que foi ajustado para reduzir o ângulo de passo em 16 vezes. Levando em conta que uma volta completa é equivalente a 360 graus, foi possível então calcular o número de passos do motor (Equação (4)). O motor foi alimentado com uma fonte de tensão contínua ajustada para 12 volts, enquanto seu acionamento foi feito por meio de um Arduino Nano e um driver TB6600.

$$\omega = \frac{V_c * R_p}{A_p} \quad (4)$$

Onde  $V_c$  é uma volta completa em graus,  $A_p$  é o ângulo de passo do motor Nema 23 e  $R_p$  é o número de vezes que esse ângulo foi reduzido. A partir desses valores pode-se então encontrar o número de micro passos  $\omega$  que, para esse experimento, foi igual a 3200 passos.

Com o número de micro passos em mãos, foi desenvolvido um script para ajustar o tempo necessário para que o motor realizasse um passo. Sabendo que foi definido um tempo de 20 segundos para uma volta completa do objeto, tem-se que o tempo gasto por passo é igual ao tempo de uma volta completa pelo número total de passos, isto é, 6,25 ms.

Além do movimento de rotação do objeto, foram feitos pequenos movimentos circulares com a câmera DVS128. Esses movimentos foram feitos de modo manual por meio do suporte da câmera.

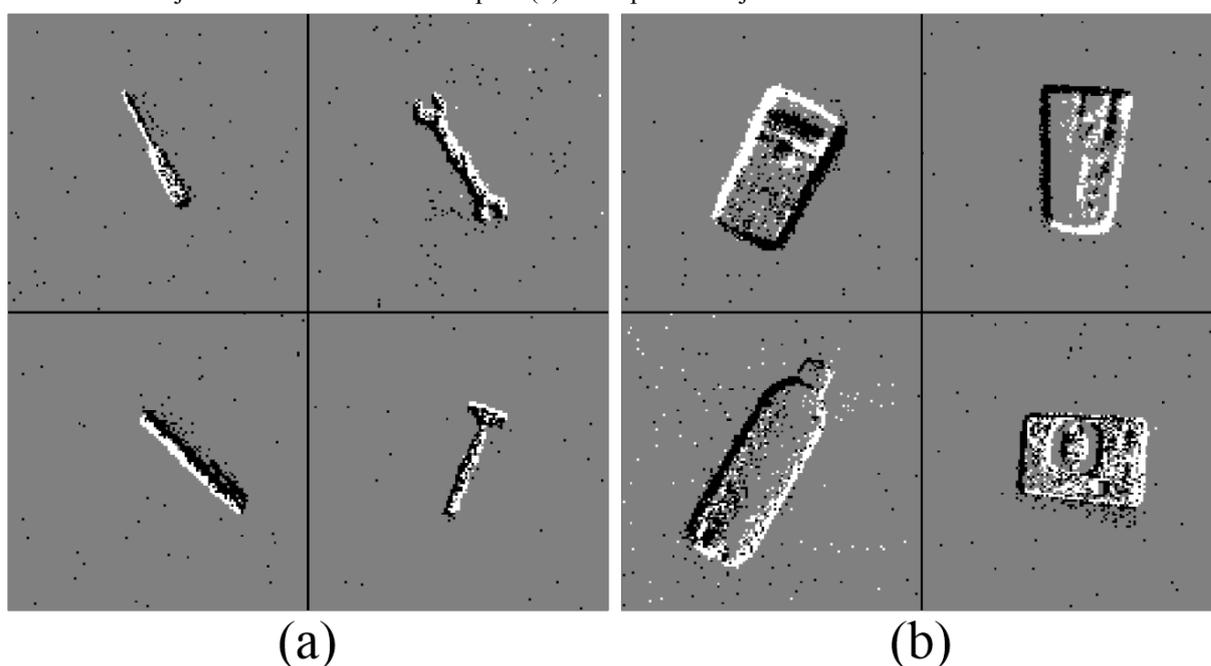
### 6.1.2.3 Geração do conjunto de imagens

Após a gravação do conjunto de objetos proposto de acordo com o protocolo descrito na seção 6.1.2.2, foi necessário realizar transformação dessas gravações em imagens, de forma que elas serão utilizadas no treinamento do modelo de ML apresentado na seção 6.1.3.2.3.

Essa etapa foi feita utilizando o método de conversão de eventos para quadros descrita na seção 6.1.1.3, de modo que cada uma das imagens é resultante de um período de 40 ms de sua respectiva gravação.

Ao final dessa conversão, foram obtidas 9000 imagens no total, sendo 300 pertencentes a cada um dos 30 objetos previamente definidos e apresentados na seção 6.1.2.1. Dessa forma, ao dividir as imagens de acordo com sua classe, tem-se que ambas as classes dos tipos de preensão Tripod e Power possuem 4500 imagens. A Figura 23 apresenta algumas das imagens pertencentes ao conjunto de dados obtidas após a conversão.

Figura 23 – Imagens obtidas por meio do processo de conversão de eventos desenvolvido. (a) Exemplos de objetos classificados como Tripod. (b) Exemplos de objetos classificados como Power.



Em seguida, o conjunto de dados foi dividido em dois grupos de imagens. O primeiro grupo é composto por todas as imagens dos 10 primeiros objetos de cada classe contidos na Tabela A1 apresentada no apêndice A. Como cada objeto possui 300 imagens, esse grupo é composto por 6000 imagens. Esse grupo foi posteriormente usado para o treinamento do modelo de CNN proposto.

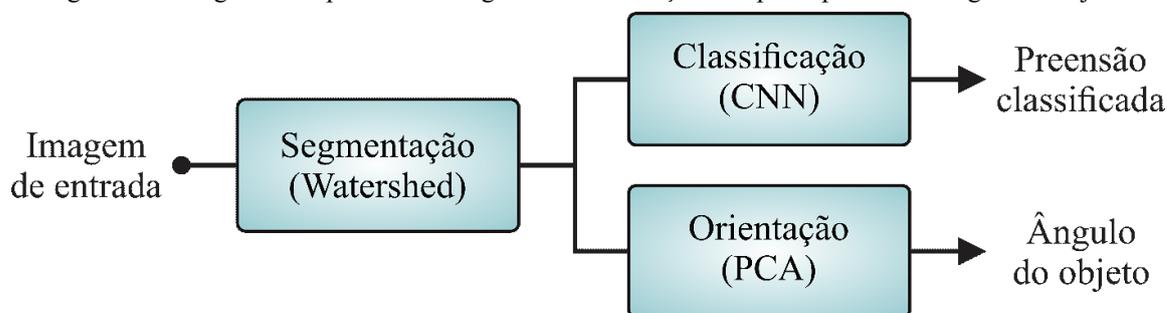
Por outro lado, o segundo grupo é composto por todas as imagens dos 5 últimos objetos de cada classe contidos na Tabela A1, ou seja, 3000 imagens. Diferentemente do primeiro grupo, o modelo de CNN proposto não teve contato com as imagens contidas no segundo grupo durante seu treinamento. Isso será fundamental no experimento em tempo real, pois será feita uma análise de robustez e capacidade de generalização do sistema para objetos jamais vistos.

### 6.1.3 PROCESSAMENTO DE IMAGENS

A etapa de processamento de imagens corresponde à maior parte do sistema proposto. Nela, estão contidas as subetapas de segmentação, classificação da forma e obtenção da orientação do objeto. A Figura 24 mostra o diagrama simplificado de todas as subetapas contidas na etapa de processamento de imagens. Cada um dos blocos apresentados no diagrama possui estágios importantes, os quais são explicados detalhadamente nas próximas subseções.

Ao fim da etapa de processamento de imagens, espera-se obter a classificação do tipo de preensão necessária para segurar o objeto e o ângulo do objeto, de modo que esses resultados serão utilizados no sistema em tempo real de controle de próteses semiautônomo proposto.

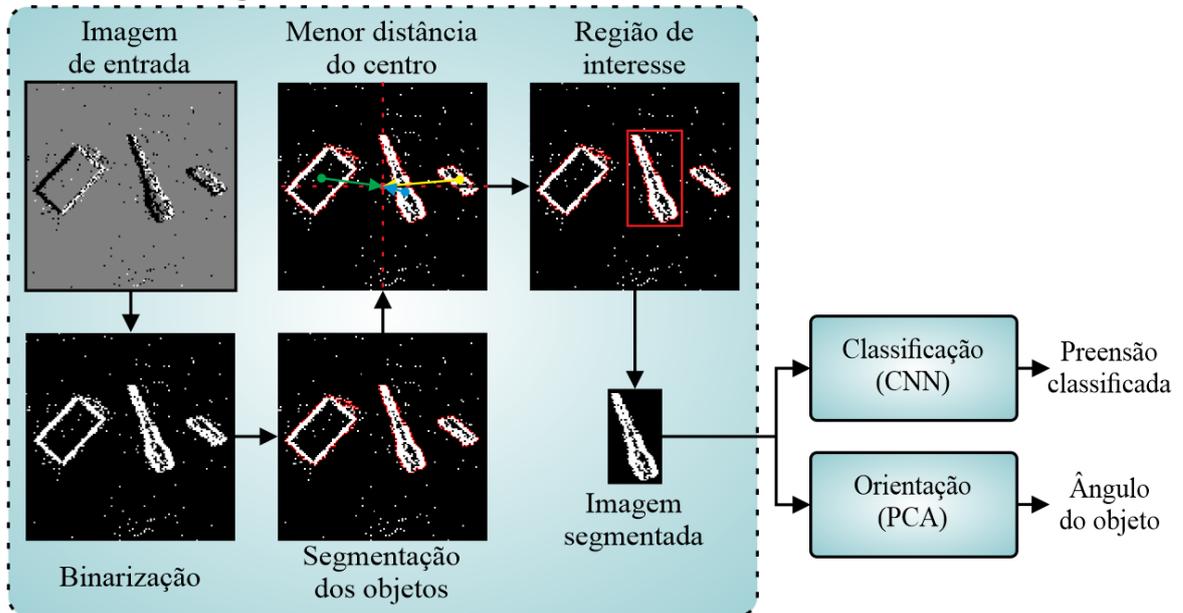
Figura 24 – Diagrama simplificado do algoritmo de obtenção do tipo de preensão e ângulo do objeto.



#### 6.1.3.1 Segmentação do objeto almejado

A segmentação do objeto consiste no processamento necessário para fornecer a imagem de entrada para ambas as etapas de classificação da forma e obtenção da orientação do objeto. Essa etapa está subdividida em 5 estágios, sendo eles: a binarização da imagem de entrada, segmentação dos objetos contidos na cena, a seleção do objeto mais próximo ao centro da imagem, a definição da região de interesse e o recorte da imagem de acordo com os limites encontrados (Figura 25). Os procedimentos realizados para a conclusão de cada um desses estágios serão explicados nos tópicos subsequentes.

Figura 25 – Diagrama expandido da etapa de segmentação do objeto.  
Segmentação (Watershed)



#### 6.1.3.1.1 Binarização da imagem

A binarização da imagem é primeira etapa de todo o processamento e é fundamental para as próximas etapas. Como descrito na seção 5.3, a segmentação de imagens usando o algoritmo de Watershed é feita de acordo com os níveis de tom de cinza da imagem, de modo que, se a imagem for colorida, é necessário convertê-la para tons de cinza.

No caso deste trabalho, as imagens já são formadas em tons de cinza, pois contêm apenas valores de 0 para pixels que representam os eventos negativos (preto), 127 para pixels sem eventos (cinza) e 255 para pixels que representam os eventos positivos (branco). Entretanto, um mesmo objeto é representado por ambos os pixels brancos e pretos, o que pode dividir o objeto e conseqüentemente afetar a performance da segmentação. Dessa forma, foi necessário fazer a binarização da imagem, usando a equação (5) na imagem de entrada.

$$\text{Para cada } p \text{ em } S \rightarrow p = \begin{cases} 0, & \text{se } p = 127 \\ 255, & \text{se } p = 0 \end{cases} \quad (5)$$

Onde  $S$  é a imagem de entrada e  $p$  é cada um dos pixels contidos na imagem. Assim, todos os pixels sem eventos passaram a ser pretos, e todos os pixels com ambos os eventos positivos e negativos passaram a ser brancos.

### *6.1.3.1.2 Segmentação dos objetos contidos na cena*

Para realizar a segmentação dos objetos, foi utilizado o algoritmo de Watershed, que consiste em um processo de segmentação morfológico e regional. Para isso, é usada uma abordagem de análise de imagem, considerando uma topografia de superfície baseada em tons de cinza, onde o valor de cinza do pixel representa a elevação da paisagem. Todo o processo de segmentação por meio do algoritmo de Watershed foi explicado de forma detalhada na Seção 5.3.

Levando em conta que o algoritmo de Watershed precisa de uma imagem de duas dimensões para ser usado, não é possível usá-lo em um fluxo de eventos obtidos por meio de uma câmera neuromórfica, reforçando a importância da etapa anterior de agrupamento de eventos.

Um evento isolado gerado por uma câmera neuromórfica carrega consigo poucas informações, pois não consegue representar nada além de um ponto em um espaço de 3 dimensões. Todavia, um conjunto de eventos conseguem representar formas complexas de diferentes tipos de objetos contidos em uma cena.

Levando em conta que o objetivo dessa etapa é justamente obter a forma dos objetos gravados, os quadros formados a partir dos eventos obtidos da câmera DVS128 conseguem prover essas informações com maestria. Além disso, ao se usar o algoritmo de Watershed nesses quadros, é possível extrair a nuvem de pontos e a localização de todos os objetos contidos na cena. Portanto, abordagens morfológicas podem ser uma metodologia interessante para ser aplicada em registros neuromórficos.

No entanto, existem muitos eventos que não são úteis espalhados pela imagem devido a pequenas mudanças de luminosidade recebida pelos sensores. Esses ruídos presentes nas gravações neuromórficas causam uma supersegmentação (explicada na Seção 5.3). Para resolver este problema, foram executados três passos para reduzir os problemas causados devido a essa supersegmentação. Primeiro, foi usado um filtro de limiar para nuvem de pontos muito pequenas, eliminando aquelas que possuíam áreas muito pequenas. Em seguida, foi criada uma caixa delimitadora retangular cobrindo o contorno de diferentes regiões, mesclando as nuvens de pontos presentes dentro das caixas delimitadoras que se cruzam. Por fim, as nuvens de pontos presentes dentro das caixas delimitadoras com centroides próximos também foram mescladas de acordo com um limite de distância, evitando então que um mesmo objeto se separe em mais de uma nuvem de pontos.

### 6.1.3.1.3 Seleção do objeto

Como uma única cena pode conter mais de um objeto, foi determinado que o objeto selecionado deverá ser sempre aquele que estiver mais próximo ao centro da imagem. Dessa forma, foi feito o cálculo do módulo dos vetores entre cada um dos centros de massa dos objetos contidos na cena e o centro da imagem, usando a Equação 6.

$$|\vec{v}| = |\vec{OC}| = |\vec{C} - \vec{O}| = \sqrt{(C_x - O_x)^2 + (C_y - O_y)^2} \quad (6)$$

Onde  $\vec{C}$  é o vetor com origem no ponto zero da imagem até o seu centro,  $\vec{O}$  é o vetor com origem no ponto zero da imagem até o centro de massa do objeto e  $|\vec{v}|$  é o módulo do vetor resultante da subtração dos vetores  $\vec{C}$  e  $\vec{O}$ , obtendo então a distância entre o centro de massa do objeto e o centro da imagem. O cálculo do  $|\vec{v}|$  é feito para todos os objetos contidos na cena, de modo que o objeto que possuir o menor valor de  $|\vec{v}|$  será o selecionado.

### 6.1.3.1.4 Definição da região de interesse

No tópico 6.1.3.1.2 foi descrito o processo de obtenção da nuvem de pontos de todos os objetos contidos na cena. Agora, será definida a região de interesse (ROI) de acordo com a nuvem de pontos do objeto selecionado.

Para obter a ROI, é necessário encontrar ambos os pontos de mínimo e máximo valor do objeto. Em seguida, é preciso encontrar o vetor resultante entre os pontos de mínimo e máximo como descrito na Equação 7.

$$\vec{u} = \overrightarrow{P_{min}P_{max}} = P_{max} - P_{min} = \langle P_{xmax} - P_{xmin}, P_{ymax} - P_{ymin} \rangle \quad (7)$$

Onde  $\vec{u}$  é o vetor resultante dos pontos de mínimo e máximo do objeto. Por fim, temos que o módulo do primeiro e segundo índice do vetor indicam respectivamente a largura e a altura da ROI.

### 6.1.3.1.5 Conversão da nuvem de pontos para imagem

Nas seções anteriores foram descritos os processamentos necessários para a obtenção da nuvem de pontos de todos os objetos contidos na cena, selecionar o mais próximo ao centro e

definir a ROI. Dessa forma, são obtidas todas as informações necessárias para recriar a imagem do objeto de forma separada.

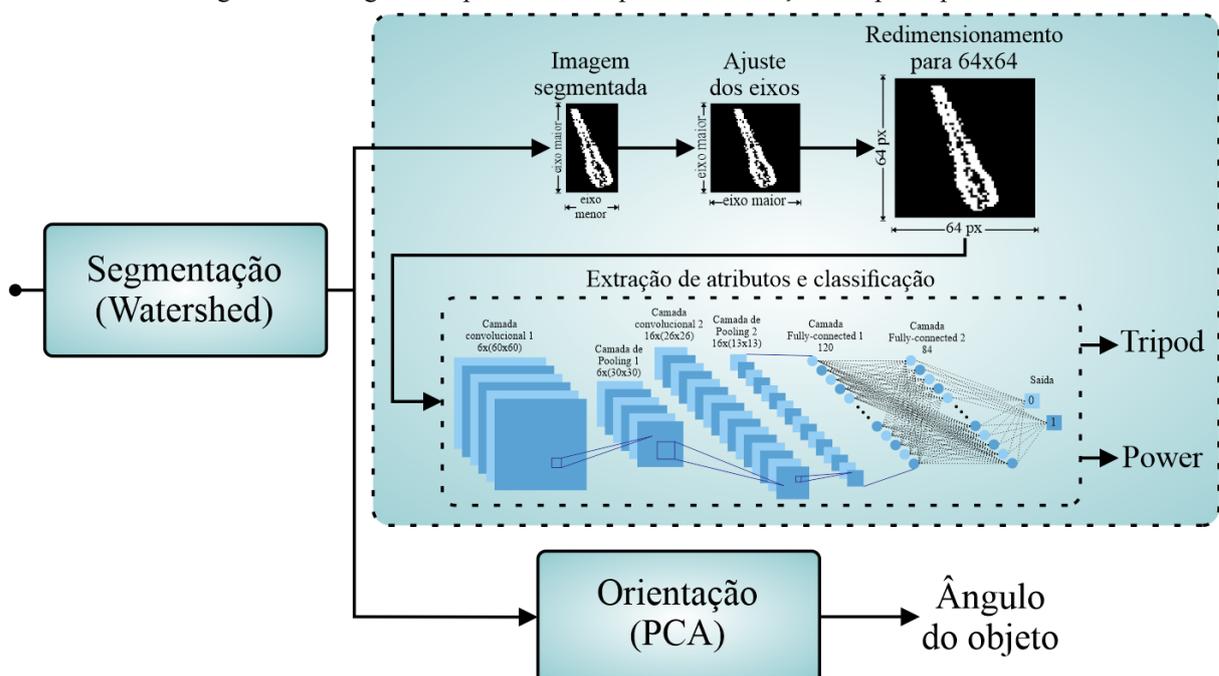
Nessa fase, a nuvem de pontos é inserida em uma nova matriz de zeros com o mesmo formato da ROI determinada, gerando então uma nova imagem contendo apenas o objeto selecionado.

Ao longo do desenvolvimento desse trabalho, cogita-se efetuar apenas o recorte da imagem original, em vez de se recriar uma nova imagem a partir da nuvem de pontos e uma matriz de zeros. Entretanto, alguns pixels na vizinhança do objeto ainda poderiam estar ativos, podendo influenciar diretamente nas etapas posteriores.

### 6.1.3.2 CLASSIFICAÇÃO DO TIPO DE PREENSÃO

Após a obtenção da imagem do objeto segmentado, são iniciadas as duas etapas de classificação do tipo de prensão e obtenção da orientação do objeto, as quais ocorrem de forma paralela. Nessa seção, estão descritas as principais etapas necessárias para obter a classificação do tipo de prensão, sendo elas: o ajuste dos eixos, o redimensionamento, a extração de atributos e a classificação (Figura 26). Como resultado, essa etapa fornece o tipo de prensão, classificado de acordo com o formato do objeto, podendo ser Tripod ou Power.

Figura 26 – Diagrama expandido da etapa de classificação do tipo de prensão



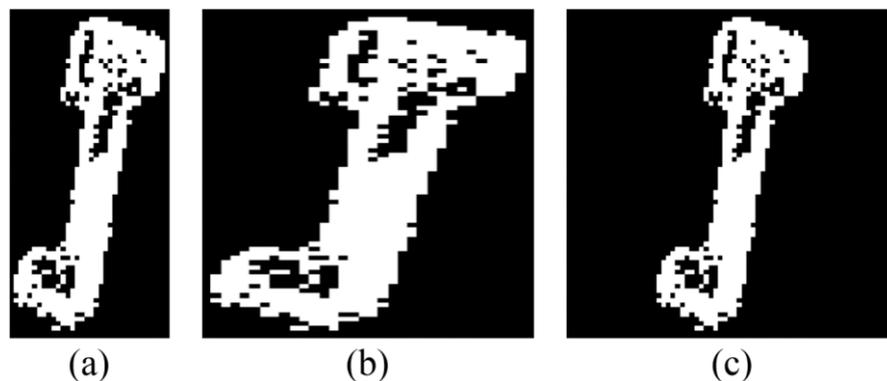
### 6.1.3.2.1 Ajuste dos eixos

As imagens de entrada de uma CNN são padronizadas, de modo que todas as imagens usadas para ambas as etapas de treinamento e teste devem possuir obrigatoriamente formatos iguais. Entretanto, o formato de cada imagem de saída da etapa de segmentação varia de acordo com o objeto segmentado. Dessa forma, é necessário padronizar essas imagens antes de usá-las no modelo de CNN desenvolvido.

As imagens de entrada de uma CNN são padronizadas, de modo que todas as imagens usadas para ambas as etapas de treinamento e teste devem possuir obrigatoriamente formatos iguais. Entretanto, o formato de cada imagem de saída da etapa de segmentação varia de acordo com o objeto segmentado. Dessa forma, é necessário padronizar essas imagens antes de usá-las no modelo de CNN desenvolvido.

O primeiro passo para resolver esse problema é efetuar o ajuste dos eixos, tornando a imagem quadrada independentemente do formato do objeto segmentado. Uma abordagem possível seria redimensionar a imagem segmentada para o tamanho predefinido da entrada da CNN (Figura 27b). Entretanto, essa abordagem acaba distorcendo o objeto contido na imagem, afetando o principal propósito da etapa de classificação que é obter o tipo de prensão de acordo com o formato do objeto. Dessa forma, a abordagem escolhida foi acrescentar pixels nas extremidades do menor eixo da imagem, tornando a imagem quadrada sem alterar o formato do objeto (Figura 27c).

Figura 27 – Comparação entre os métodos de ajuste dos eixos na imagem segmentada de uma chave. (a) Imagem segmentada, (b) ajuste sem verificar o tamanho dos eixos, causando a distorção do objeto e (c) método de ajuste dos eixos da imagem proposto, evitando que ocorra a distorção do objeto.



Para realizar esse procedimento, são necessárias duas fases. Primeiro, é necessário verificar a quantidade de pixels de ambos os eixos X e Y. Em seguida, o eixo com o menor número de pixels é igualado ao eixo com o maior número de pixels, adicionando a mesma

quantidade de pixels nas extremidades. Assim, apesar da mudança realizada no formato da imagem, o objeto continuará no seu centro.

Para melhor compreensão dessa etapa, pode-se analisar o exemplo apresentado na Figura 26. No caso, tem-se uma imagem retangular, onde o eixo Y é maior do que o eixo X. Após o processamento, o eixo X passou a ter o mesmo número de pixels que o eixo Y, tornando a imagem quadrada sem distorcê-la. Entretanto, caso o eixo X da imagem fosse maior o procedimento seria invertido, igualando o eixo Y ao eixo X.

#### 6.1.3.2.2 Redimensionamento da imagem

Ao final da transformação da imagem retangular em uma imagem quadrada, é necessário redimensionar a imagem para um formato padrão, que será o mesmo usado na entrada do modelo de CNN desenvolvido.

Como descrito na seção 6.1.1.3, as imagens construídas a partir dos eventos gerados pela câmera DVS128 possui formato de 128x128 pixels. Após o processo de tornar a imagem segmentada quadrada, tem-se que a imagem passa a ter dimensões muito menores do que a imagem original. Entretanto, esse formato varia de acordo com as dimensões do objeto almejado. Assim, foi utilizado um tamanho de 64x64 pixels, que é exatamente a metade do tamanho da imagem original.

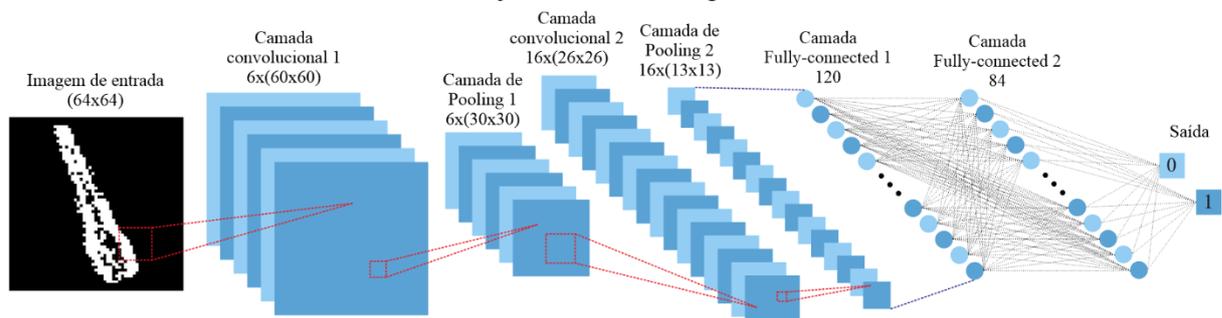
Para isso, foi utilizado a função de redimensionamento (*resize*) da biblioteca OpenCV, que é capaz de alterar as dimensões de uma imagem por meio de um método de interpolação. A função de redimensionamento de imagem usada possui diferentes formas de realizar a interpolação da imagem. Pelo fato de a imagem possuir apenas pixels brancos e pretos, é importante que ela não altere esse aspecto ao aumentar ou diminuir o seu tamanho. Dessa forma, foi utilizada a interpolação por meio dos vizinhos mais próximos, que mantém os níveis de cores existentes na imagem original.

O formato de 64x64 foi escolhido por conta de dois motivos. Primeiramente, o próprio objeto já vai ter pelo menos um dos seus eixos com um valor próximo a 64 pixels, evitando que a imagem seja fortemente alterada. Além disso, levando-se em conta que a etapa de classificação possui uma grade quantidade de cálculos que serão realizados de acordo com o tamanho da imagem, tem-se que quanto menor o seu tamanho, menor será a quantidade de cálculos necessários, reduzindo então o custo computacional dessa etapa. Por outro lado, ao se reduzir drasticamente o tamanho da imagem, perde-se informações importantes dela. Dessa forma, foi definido de forma empírica o tamanho de 64x64.

### 6.1.3.2.3 Extração de atributos e classificação

A etapa de classificação é composta por duas subetapas, sendo a primeira relacionada a extração de atributos e a segunda relacionada a classificação. Todo o modelo foi construído em Python utilizando a biblioteca Keras. A Figura 28 apresenta a arquitetura do modelo construído, desde as camadas de extração de atributos até as camadas de classificação.

Figura 28 – Modelo de CNN usado para classificar o tipo de prensão necessária de acordo com o formato do objeto contido na imagem.



Como descrito na seção anterior, a imagem de entrada da rede é binária e possui tamanho de 64x64. A primeira etapa é responsável pela extração de atributos da imagem, a qual passa por uma camada convolucional com kernels de tamanho 3x3, obtendo 6 filtros resultantes da convolução. Em seguida, os filtros têm seu tamanho reduzido por meio de uma camada de Pooling, usando o valor máximo para cada janela de tamanho 2x2. Na sequência, é usada mais uma camada convolucional, obtendo 16 filtros resultantes da convolução. Em seguida, os filtros resultantes têm seu tamanho reduzido outra vez por meio da segunda camada de Pooling. Por último, os filtros passam por uma camada de achatamento e então estão prontos para entrar na etapa de classificação.

A etapa de classificação é composta por duas camadas Fully-connected, sendo a primeira composta por 120 perceptrons e a segunda por 84. Por último, tem-se então a camada de saída, que é composta por 2 perceptrons, de modo que o resultado pode ser 0 para objetos que requerem o tipo de prensão Tripod ou 1 para os objetos que requerem o tipo de prensão Power.

#### 6.1.3.2.4 Validação cruzada

Com o intuito de verificar a capacidade de generalização e a robustez do modelo construído, foi realizada uma validação cruzada, usando todo conjunto de dados para treinamento e teste do modelo. Para isso, foram usadas as 6000 imagens pertencentes ao grupo 1 do conjunto de imagens. Em seguida, foi definido a divisão dessas imagens em 5 partes iguais. Ou seja, cada parte continha 1200 imagens.

Um ponto importante para se destacar é o balanceamento das amostras, que garante que cada uma das partes contenha a mesma quantidade de imagens de ambas as classes de preensão Tripod e Power, ou seja, cada parte possui 600 imagens relativas a cada classe. Esse balanceamento é importante, pois evita que o modelo seja enviesado durante o treinamento para seja qual for a classe.

Na sequência, quatro partes foram usadas para treinar o modelo, enquanto a parte remanescente foi usada para testar o modelo. Esse processo se repetiu até que cada uma das partes tenha sido usada na etapa de treinamento quatro vezes e na etapa de teste uma vez. É importante destacar que o modelo foi resetado no início do treinamento de cada uma das partes.

O treinamento do modelo foi feito durante 10 épocas para cada uma das 5 partes. Ou seja, o modelo foi treinado com todo o conjunto de imagens de treinamento 10 vezes. Além disso, os pesos da rede foram recalculados a cada 300 imagens. Levando em conta que existiam 4800 imagens de treinamento para cada parte, tem-se que os pesos do modelo foram recalculados 16 vezes durante uma época.

#### 6.1.3.2.5 Treinamento do modelo

Após efetuar a validação cruzada do modelo, foi realizado o seu treinamento final, que foi utilizado nos experimentos com o sistema de controle semiautônomo desenvolvido (detalhados na seção 6.2).

Para isso, foram utilizados todo o conjunto de imagens pertencentes ao grupo 1 de objetos para treinar o modelo, ou seja, 6000 imagens. As outras 3000 imagens pertencentes ao grupo 2 de objetos não foram incluídas na etapa de treinamento do modelo em nenhum momento. Entretanto, elas foram usadas para testar a robustez e capacidade de generalização do modelo de CNN.

Diferentemente da etapa da validação cruzada, o treinamento do modelo foi feito durante 30 épocas para todas as 6000 imagens de treinamento. Já os pesos da rede foram recalculados

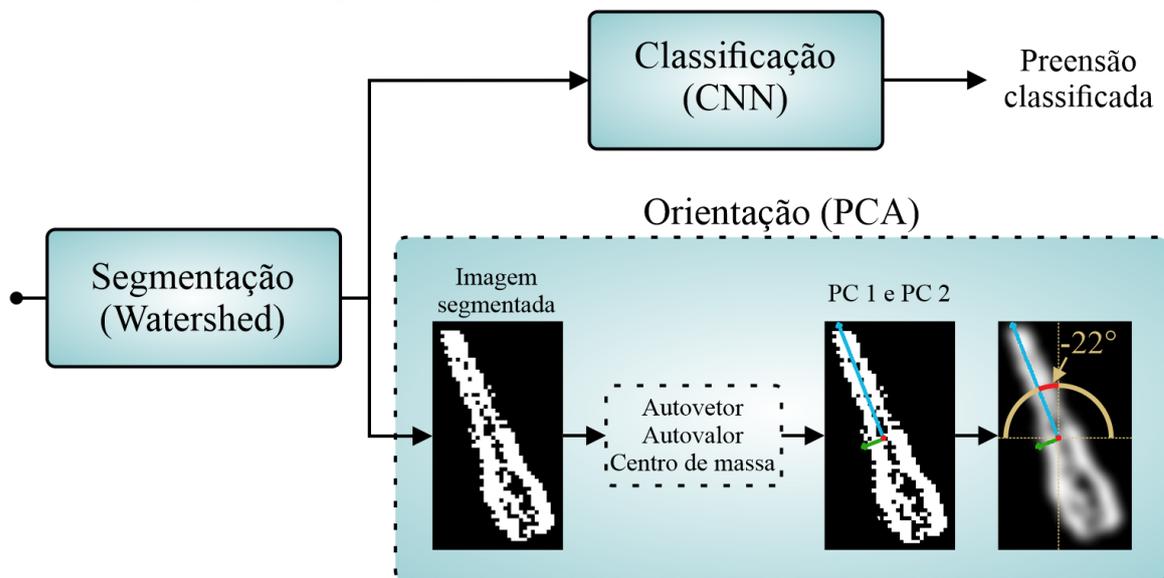
a cada 400 imagens. Levando em conta que existiam 6000 imagens de treinamento no total, tem-se que os pesos do modelo foram recalculados 15 vezes durante uma época.

### 6.1.3.3 Obtenção da orientação do objeto

A última etapa do processamento de imagens consiste na obtenção da orientação do objeto na cena. Como dito anteriormente, essa etapa ocorre paralelamente a etapa de classificação. Assim, ao final dessa etapa espera-se obter o ângulo do objeto segmentado, possibilitando realizar as correções da angulação do punho da prótese de mão no sistema de controle em tempo real proposto na seção 6.2.1.

Como ilustrado na Figura 29, a etapa de obtenção da orientação do objeto é dividida em três fases principais, sendo elas: obtenção do autovetor, autovalor e centro de massa do objeto, cálculo dos vetores que representam as PCs 1 e 2, e por fim, cálculo do ângulo do objeto a partir dos vetores encontrados. Além disso, foi feita uma validação do método de obtenção da orientação proposto com o intuito de garantir que o ângulo calculado está de acordo com o ângulo real do objeto. Todas essas etapas estão descritas detalhadamente nas seções subsequentes.

Figura 29 – Diagrama expandido da etapa de obtenção da orientação do objeto.



#### 6.1.3.3.1 Obtenção do autovetor, autovalor e centro de massa do objeto

A primeira fase após obter a imagem do objeto segmentada consiste em obter o autovetor, autovalor e centro de massa do objeto. Todos esses termos foram explicados da seção 5.2, onde foram abordados os principais pontos relacionados a técnica de PCA.

Levando em consideração que os principais procedimentos relacionados ao tratamento da imagem já foram realizados na etapa de segmentação, torna-se muito mais fácil obter as informações resultantes da PCA. Assim, para a realização dessa etapa foi utilizada a função de PCA da biblioteca OpenCV, que usa a imagem segmentada como entrada e retorna o ponto de centro de massa do objeto e os valores de autovetor e autovalor respectivos as PCs 1 e 2.

#### 6.1.3.3.2 Componentes principal e secundaria

Com o centro de massa do objeto e os valores de autovetor e autovalor de cada uma das PCs, é necessário realizar o cálculo dos vetores que representam essas PCs utilizando a Equação 8. Dessa forma, esses vetores são resultados da multiplicação do autovetor pelo seu respectivo autovalor, transladando o resultado para o ponto do centro de massa do objeto. A partir desses vetores, torna-se possível realizar o cálculo do ângulo do objeto.

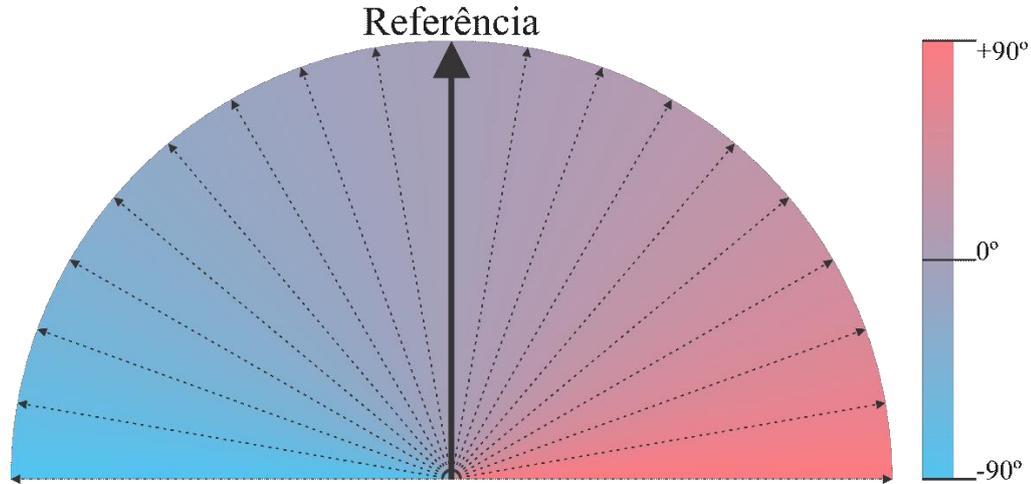
$$\overrightarrow{B_{PC}} = \langle \overrightarrow{v}_x * V + c_x, \overrightarrow{v}_y * V + c_y \rangle \quad (8)$$

Onde  $\overrightarrow{B_{PC}}$  é o vetor resultante,  $\vec{v}$  é o autovetor,  $V$  é o autovalor e  $c$  é o centro de massa do objeto. Esse cálculo é feito para cada uma das PCs separadamente. Para visualizar esses vetores, pode-se analisar o terceiro retângulo da Figura 29, que mostra o centro de massa, o vetor da PC1 e o vetor da PC2, nas cores vermelha, azul e verde, respectivamente.

#### 6.1.3.3.3 Cálculo do ângulo

Após obter os vetores das PCs, é feito o cálculo da orientação do objeto. Para isso, foi necessário definir um vetor de referência, isto é, o ângulo zero. Assim, o ângulo do objeto será sempre calculado a partir da diferença do vetor resultante da PC1 com o vetor de referência definido. A Figura 30 ilustra a faixa possível de angulação do objeto, sendo o vetor de referência o ângulo de 0 graus. Assim, quanto mais o vetor da PC1 do objeto estiver para esquerda do vetor de referência, mais negativo ele vai ser, até o máximo de menos 90 graus. O mesmo acontece para o lado direito, porém o ângulo passa a ser positivo.

Figura 30 – Referência para o cálculo do ângulo do objeto. O vetor de referência é sempre na vertical, enquanto o vetor da PC1 pode estar em qualquer direção dentro da faixa de -90 até +90 graus.



Na Figura 30 nota-se que a variação total é de 180 graus. Entretanto, é possível que a PC1 do objeto esteja nos dois quadrantes inferiores da circunferência, que não estão representados na figura. Nesses casos, a PC1 é espelhada diagonalmente. Isto é, caso a direção do vetor esteja no quadrante inferior esquerdo da circunferência, ele será espelhado para o quadrante superior direito. O motivo dessa abordagem é que na tarefa de prensão analisada nesse trabalho, não importa a direção do vetor, mas sim a angulação entre o objeto e o vetor de referência definido. Dessa forma, é possível analisar a angulação do objeto usando apenas nos quadrantes superiores.

Após a obtenção do vetor da PC1 devidamente obtido e corrigido quando necessário, é finalmente realizado o cálculo do ângulo existente entre os vetores de referência e PC1. Como descrito na Equação 9, tem-se que o ângulo entre dois vetores é igual ao arco cosseno do produto escalar desses vetores, dividido pela multiplicação dos seus módulos.

$$\theta = \cos^{-1} \left( \frac{\overrightarrow{v_{ref}} \cdot \overrightarrow{B_{PC1}}}{|\overrightarrow{v_{ref}}| * |\overrightarrow{B_{PC1}}|} \right) \quad (9)$$

Onde  $\overrightarrow{v_{ref}}$  é o vetor de referência,  $\overrightarrow{B_{PC1}}$  é o vetor resultante da PC1 e  $\theta$  é o ângulo entre os dois vetores. Assim, após obter o ângulo  $\theta$ , tem-se então a angulação do objeto almejado.

Caso a subtração dos módulos dos vetores das PCs 1 e 2 seja menor que 10% do módulo do vetor de maior amplitude, o ângulo do objeto não é calculado. Essa exceção ocorre para impedir que o algoritmo tente encontrar um ângulo para objetos que não possuem um lado significativamente maior que o outro, tais como bolas e cubos.

#### 6.1.3.3.4 Validação

O sistema de obtenção do ângulo atual do objeto foi testado utilizando o mesmo mecanismo de rotação descrito na seção 6.1.2.2, já que é possível ajustar o ângulo de maneira mais precisa com o motor de passo.

O processo de validação foi feito da forma descrita a seguir. O objeto foi anexado ao eixo do motor, posicionando seu lado maior na vertical. Em seguida, foi gerado um ângulo aleatório podendo ser qualquer número inteiro de -90 até 90. Esse processo se repetiu 5 vezes para cada um dos objetos contidos no conjunto de objetos desenvolvido, exceto para a bola e a laranja. Pelo fato desses dois objetos possuírem lados iguais ou muito próximos, o sistema não efetuou o cálculo do ângulo (conforme explicado na seção anterior), ocasionando então a exclusão desses objetos.

Após realizar esse procedimento para todos os objetos, foi então realizado o cálculo do erro, que é o módulo da diferença do ângulo desejado e o ângulo obtido pelo sistema (Equação (10)).

$$E_a = |A_d - A_o| \quad (10)$$

Onde  $A_d$  é o ângulo desejado,  $A_o$  é o ângulo obtido. Assim, é possível obter o erro  $E_a$  por meio da subtração desses dois ângulos.

## 6.2 EXPERIMENTOS EM TEMPO REAL

Na etapa de experimentos online, foram desenvolvidos dois sistemas de controle de próteses usando apenas um canal de sEMG.

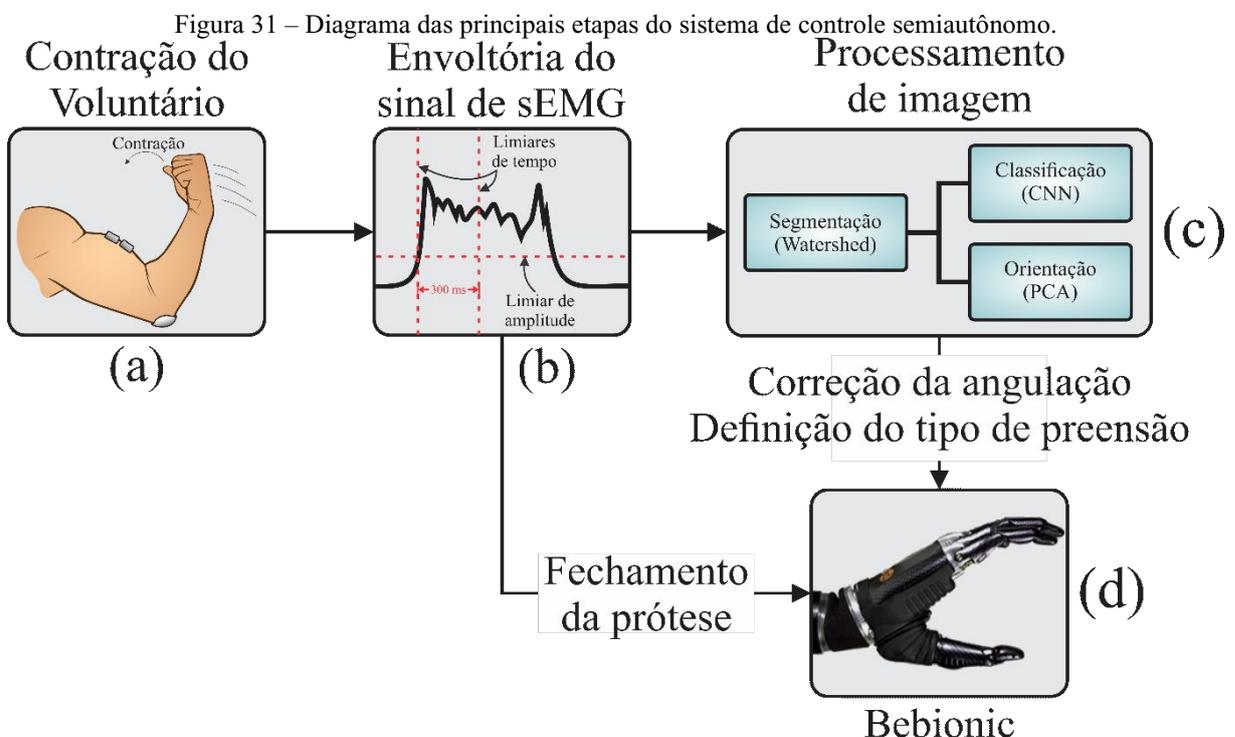
O primeiro sistema, que é o principal foco desse trabalho, baseia-se na utilização de uma câmera neuromórfica para automatizar as etapas de escolha do tipo de preensão e correção da angulação do punho, de modo que o usuário tenha como função realizar apenas o fechamento e abertura da prótese.

Por outro lado, o segundo sistema baseia-se no chaveamento de blocos para a realização de uma tarefa de preensão, de modo que o usuário tenha que realizar todos as etapas do controle manualmente, incluindo o ajuste da angulação do punho e definição do tipo de preensão.

Ambos os sistemas foram testados com voluntários, com o intuito de verificar e comparar seus desempenhos em termos de seleção do tipo de preensão, estimativa da orientação do objeto, tempo de execução da tarefa e carga de trabalho.

### 6.2.1 O SISTEMA DE CONTROLE SEMIAUTÔNOMO

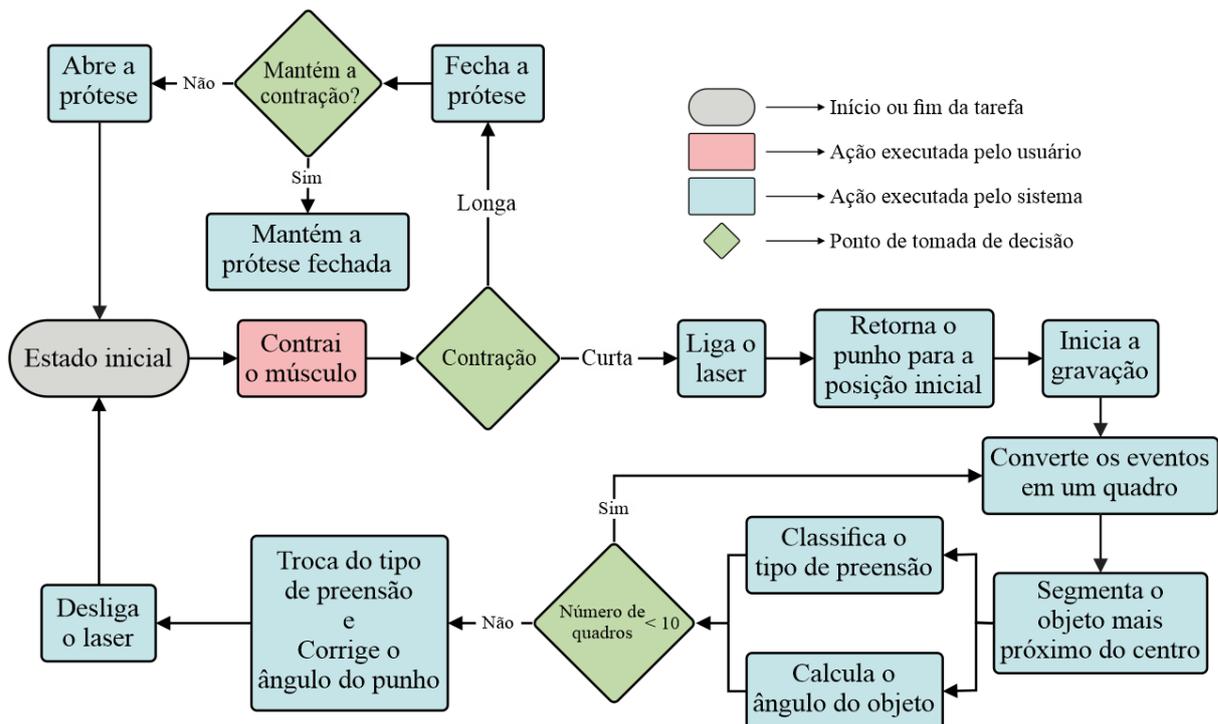
Após o desenvolvimento de todo o script de processamento de imagens, foi desenvolvido o sistema de controle semiautônomo, com objetivo de definir a angulação correta do punho e selecionar o tipo de preensão necessária a partir dos resultados da etapa de processamento de imagens. A Figura 31 apresenta o diagrama geral do sistema proposto. Como se observa, o controle geral da operação é prerrogativa do usuário, que define, por meio da ativação da musculatura do coto (no caso de usuários amputados), o início do processo como um todo. Descrevemos, a seguir, os diversos elementos deste sistema.



#### 6.2.1.1 Estrutura de controle e componentes do sistema

Nessa sub-seção, apresentamos toda a estrutura de controle e os principais materiais utilizados para o desenvolvimento do sistema de controle e do hardware responsável pela interface entre o voluntário e a prótese de mão (bypass). A Figura 32 apresenta o fluxograma relativo ao controle da prótese, de modo que os blocos vermelhos são ações realizadas pelo usuário e os blocos azuis são ações realizadas pelo sistema.

Figura 32 – Fluxograma do sistema de controle semiautônomo.



### 6.2.1.1.1 Prótese de mão

Para a realização desse trabalho, foi utilizada uma prótese de mão direita Bebionic desenvolvida pela Ottobock. Essa prótese é um dos modelos comerciais mais consolidados atualmente no mercado. Além disso, ela também tem sido utilizada em um amplo número de estudos relacionados a controle de próteses, já que apesar de muito sofisticado, o controle da Bebionic é bastante simples. A Figura 33 mostra a prótese Bebionic utilizada em todos os experimentos em tempo real contidos nesse trabalho.

Figura 33 – Prótese de mão direita Bebionic usada nos experimentos contidos nesse trabalho.



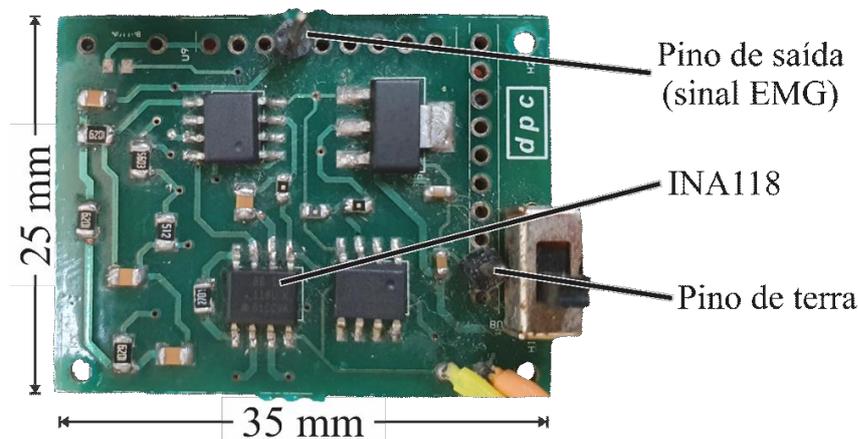
Por padrão, a Bebionic utiliza o controle DC, que é o mais comum em modelos comerciais. Dessa forma, foram necessárias diversas adaptações para utilizar a Bebionic no sistema proposto.

#### 6.2.1.1.2 *Intenção de movimento*

A primeira etapa do sistema em tempo real proposto consiste em compreender a intenção de movimento do usuário. Para isso, foi utilizado um canal de sEMG para adquirir a envoltória do sinal eletromiográfico do usuário. Por sua vez, essa envoltória será essencial para definir quando o usuário deseja iniciar uma tarefa de prensão semiautomatizada ou simplesmente realizar o fechamento da prótese.

O dispositivo de sEMG utilizado nesse trabalho foi desenvolvido em 2015 no Laboratório de Engenharia Biomédica (BioLab) para outras aplicações relacionadas ao controle de próteses mioelétricas. A Figura 34 apresenta o dispositivo de sEMG usado, o qual possui um único canal de sEMG e uma área de 22 por 35 mm.

Figura 34 – Circuito de aquisição e condicionamento de sinal de sEMG usado nesse trabalho. O circuito possui um canal e seu amplificador de instrumentação é o INA118.



Os eletrodos de sinal do dispositivo de sEMG foram posicionados em cima do músculo flexor radial do carpo do voluntário, enquanto o eletrodo de referência foi posicionado em cima da região do epicôndilo medial do úmero. Na sequência, o sinal analógico do usuário é convertido para digital por meio do conversor A/D de 8 bits de um Arduino Nano, sendo que a taxa de amostragem foi definida em 1000 Hz. Com o sinal digitalizado, é feita a retificação do sinal, possibilitando então extrair a envoltória do sinal por meio de uma função de média móvel. Essa função de média móvel reduz a frequência do sinal para 20 Hz. Por fim, foi removido o offset do sinal, de modo que o sinal basal se mantenha o mais próximo de zero.

Como apresentado nos primeiros blocos de tomada de decisão da Figura 32, existem dois limiares importantes do sistema de controle, sendo eles o limiar de amplitude e o limiar de tempo. O primeiro limiar é a amplitude que a envoltória do sinal de sEMG do usuário precisa ultrapassar para poder passar para as próximas etapas do sistema. Na sequência temos o limiar de tempo, que indica se o usuário quer realizar apenas o fechamento da prótese (caso o tempo de contração ultrapasse o limiar de tempo) ou iniciar o sistema de ajuste do tipo de preensão e angulação do punho da prótese (caso não ultrapasse o limiar de tempo).

O limiar de amplitude foi ajustado para 20% da força máxima do usuário. Ou seja, cada usuário terá um limiar diferente de acordo com a amplitude máxima da envoltória do seu sinal de sEMG, possibilitando um ajuste mais subjetivo. Já o limiar de tempo foi definido em 300 ms independentemente do voluntário. Esse tempo foi encontrado de empiricamente, de modo que não seja muito curto, tornando impossível que o usuário entre no modo automático e nem muito longo, adicionando um atraso muito grande no sistema.

#### *6.2.1.1.3 Fases automatizadas*

Existem dois caminhos após compreender a intensão de movimento do usuário por meio do seu sinal de sEMG. O primeiro e mais simples, consiste em realizar o fechamento da prótese. Essa tarefa pode ser executada a qualquer instante, desde que o sinal de sEMG do usuário se mantenha acima do limiar de amplitude por mais de 500 ms (contração longa). De fato, o sistema tem um atraso igual ao limiar de tempo na resposta do movimento. Porém, esse atraso é recompensado quando levamos em conta que diferentemente de outros sistemas de controle existentes, o sistema proposto possui apenas um canal de eletromiografia, simplificando significativamente a sua estrutura.

O segundo caminho possível e mais complexo consiste no ajuste do tipo de preensão e angulação do punho da prótese. O sistema executa essa tarefa desde que o sinal de sEMG do usuário não se mantenha acima do limiar de amplitude por mais de 500 ms (contração curta). Entretanto, algumas ações são necessárias para alcançar o objetivo final. Dessa forma, após ocorrer uma contração curta, um laser posicionado junto a lente da câmera é ativado, servindo como pista visual para o usuário conseguir direcionar o centro da câmera para o objeto almejado. O laser permanecerá ligado até o fim da tarefa de preensão.

Após ligar o laser, o punho da prótese retorna para a posição inicial caso esteja em uma outra posição. Essa posição inicial foi definida mantendo a palma da prótese sempre virada para

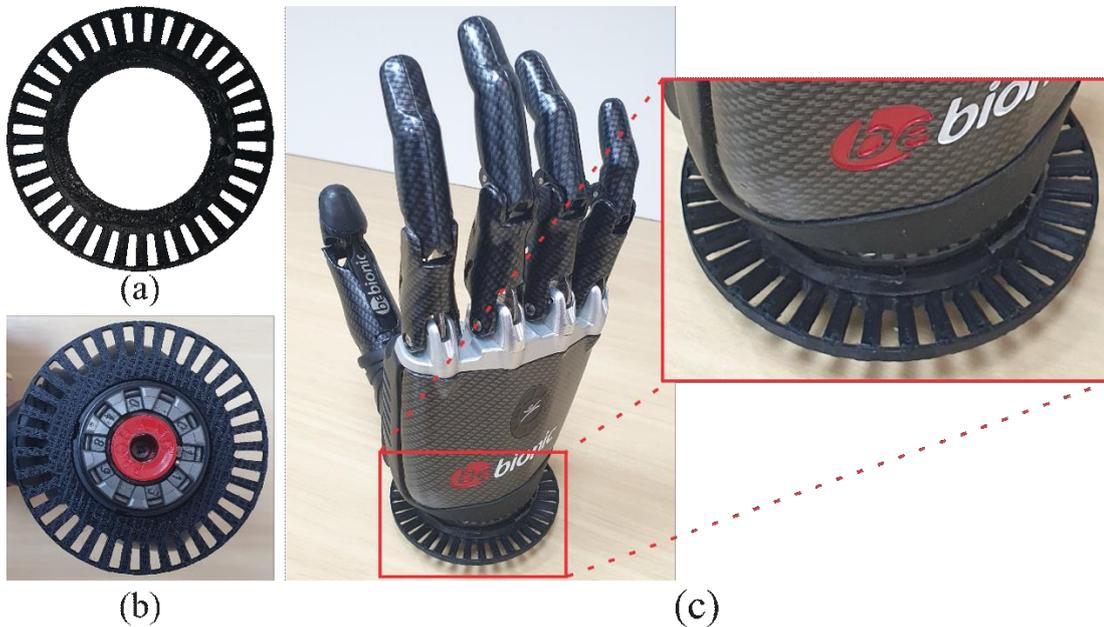
o lado esquerdo, podendo então rotacionar 90 graus para esquerda (palma para baixo) e 90 graus para a direita (palma para cima).

Em seguida, inicia-se a gravação da cena, adquirindo eventos da câmera DVS128 e convertendo-os em um quadro. A partir do quadro formado, a etapa de processamento de imagens descrita na seção 6.1.3 é efetuada, realizando primeiro a segmentação do objeto almejado e, na sequência, classificando o tipo de preensão e obtendo a angulação de acordo com o seu formato e sua orientação respectivamente. Esse processo se repete por 20 vezes, de modo que tanto o tipo de preensão quanto angulação do punho finais são alcançados por meio da média dos dez resultados parciais. Como cada quadro possui 40 ms, o processo de formação de 20 quadros é equivalente a 800 ms, que é o tempo total necessário para executar essa fase do sistema.

Por último, os resultados do tipo de preensão e angulação da prótese obtidos são enviados para o Arduino Nano por meio de uma porta serial. Esses resultados então são usados para trocar o tipo de preensão e angulação do punho atuais da prótese. Caso o tipo de preensão classificado seja diferente do que está configurado na prótese, o sistema altera automaticamente para o tipo de preensão classificado. Caso o tipo de preensão seja igual tanto na prótese quanto na classificação, o tipo de preensão é mantido inalterado.

Foi necessário adaptar um disco de encoder na prótese para corrigir a angulação do punho da prótese, já que não havia uma forma de verificar quantos graus e para que lado a prótese rotacionou (Figura 35). Tal desafio ocorreu pelo fato de a prótese utilizar um motor de corrente contínua (MCC) para realizar a rotação do punho. Assim, utilizar um disco de encoder passa a ser a melhor opção para obter essas informações.

Figura 35 – Adaptação do disco de encoder na prótese BeBionic. (a) Disco de encoder impresso na impressora 3D. (b) Visão inferior do encoder adaptado na prótese. (c) Visão superior/frontal do disco de encoder adaptado na prótese.



O disco de encoder foi desenvolvido utilizando o AutoCAD, que consiste em uma plataforma de desenvolvimento de objetos 3D. Em seguida, ele foi impresso em uma impressora 3D, utilizando um filamento preto de Acrilonitrila butadieno estireno (ABS). Esse plástico é um dos materiais mais comuns para filamentos usados em impressões 3D, podendo ser facilmente adquirido. Dessa forma, ele se encaixou perfeitamente às necessidades do experimento proposto.

O disco de encoder possui um total de 36 ciclos. Como uma volta inteira possui 360 graus, tem-se que a resolução do disco de encoder é igual a 10 graus. Entretanto, foi considerado um passo para ambas as mudanças de estado geradas pelas bordas de subida e descida, fazendo com que cada passo indique uma rotação de 5 graus.

Levando em conta que o valor recebido do script de processamento de imagens em Python fornece o valor de -90 até 90 graus, foi necessário realizar uma normalização para converter graus em passos. Essa normalização foi feita utilizando a Equação 11.

$$P = \frac{G * N_p}{360} \quad (11)$$

Onde  $G$  é o resultado em graus obtido,  $N_p$  é igual ao número de passos do disco de encoder e  $P$  é o resultado convertido para passos. Levando em conta que o ângulo recebido pode variar

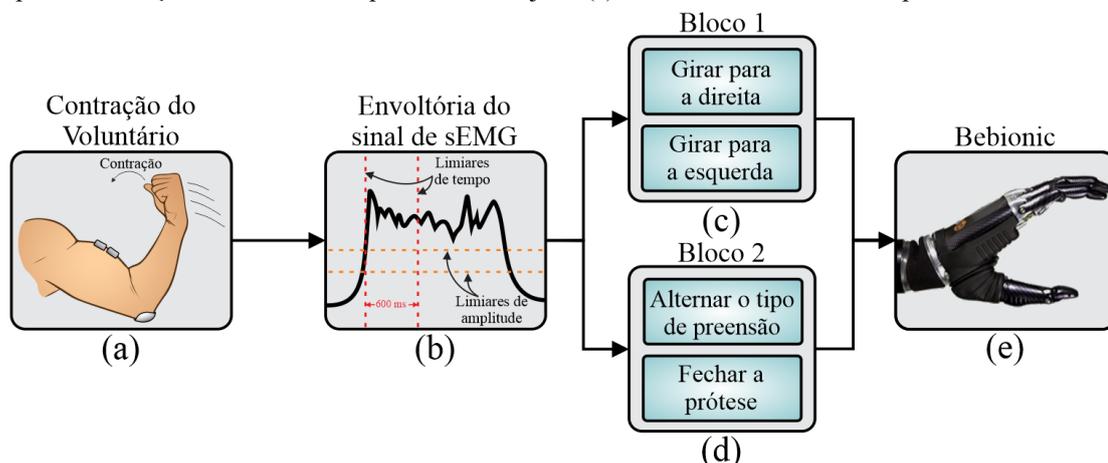
de -90 até 90 graus, após converter de graus para passos, a prótese pode rotacionar 18 passos no sentido horário e 18 passos no sentido anti-horário.

O peso total do sistema de controle semiautônomo foi igual 1174 gramas, de modo que o peso da câmera DVS128 mais o seu suporte de encaixe no bypass foi igual a 126 gramas.

## 6.2.2 SISTEMA DE CONTROLE MANUAL

O sistema de controle manual é o segundo sistema desenvolvido no presente trabalho. Diferentemente do sistema de controle semiautônomo, no sistema de controle manual o usuário precisa passar por todas as etapas de controle, realizando o ajuste do ângulo e alternando o tipo de preensão de acordo com a angulação e formato do objeto almejado. A Figura 36 apresenta o diagrama geral do sistema controle manual, desde a intenção de movimento do voluntário até o controle do fechamento da prótese de mão. Esse diagrama será posteriormente detalhado, explicando os pontos chave do sistema.

Figura 36 - Diagrama das principais etapas do sistema de controle manual. (a) Aquisição do sinal de sEMG do usuário. (b) Extração de informações da envoltória do sinal de sEMG para efetuar o controle do sistema. (c) Bloco 1 do sistema, o qual possui as funções de rotação do punho da prótese. (d) Bloco 2 do sistema, o qual possui as funções relacionadas a preensão do objeto. (e) Prótese de mão Bebionic que será controlada.

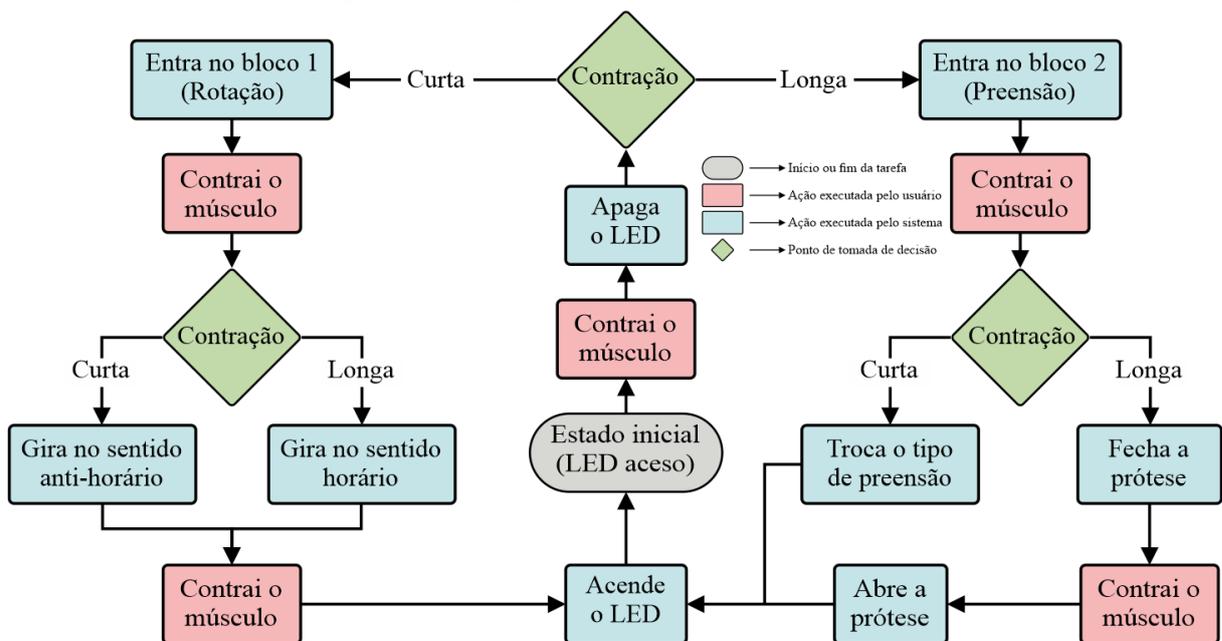


### 6.2.2.1 Estrutura de controle e componentes do sistema

O sistema de controle manual herdou grande parte dos materiais usados no sistema de controle semiautônomo. Entretanto, esse sistema de controle não possui as partes responsáveis por automatizar as etapas de ajuste da orientação e tipo de preensão da prótese de acordo com o formato e angulação do objeto almejado descrita na seção 6.1.3. Dessa forma, não foi necessário o uso da câmera DVS128 e do laser presentes no sistema de controle semiautônomo.

Embora a estrutura física de ambos os sistemas seja similar, as suas formas de controle são bastante distintas. O sistema de controle manual é dividido em blocos, de modo que cada bloco é composto por sub-blocos. A Figura 37 apresenta o fluxograma relativo as etapas de controle do sistema de controle manual.

Figura 37 - Fluxograma do sistema de controle manual.



Nesse sistema de controle, o usuário possui total controle do sistema, de modo que as ações realizadas pelo sistema são exclusivamente respostas aos comandos fornecidos. Como o usuário possui apenas um sinal de controle, a estratégia desse sistema de controle é a navegação por meio de blocos. Assim, os blocos de tomada de decisão representados por meio de losangos verdes na Figura 37 indicam as entradas em cada um dos blocos de acordo com o período de contração, que pode ser curta ou longa.

Para entender como é feita a entrada em cada um dos blocos, é preciso entender a definição de contração curta e contração longa de acordo com o tempo que o sinal de sEMG do usuário se mantém acima dos limiares de amplitude. De acordo com a Figura 36b, é possível notar a existência de dois limiares de amplitude, sendo eles o limiar inferior e limiar superior. Uma contração começa quando o sinal de sEMG do usuário é maior que o limiar de amplitude superior e termina quando é menor que o limiar de amplitude inferior. Dessa forma, o tempo de contração continua incrementando enquanto o sinal de sEMG oscilar acima do limiar inferior de amplitude. Para determinar quando é uma contração curta ou longa, foi definido um limiar de tempo de 500 ms, de modo que se o tempo de contração for menor que 500 ms a contração é curta, caso contrário ela é longa. Após compreender a definição de contração curta e longa, pode-se entrar nas etapas apresentadas no fluxograma da Figura 37.

Foi anexado um LED no by-pass, responsável por indicar quando o sistema está no estado inicial. Sabendo que o sistema se encontra no estado inicial por meio do feedback visual gerado pelo LED aceso, o usuário pode então escolher entrar em um dos dois blocos possíveis de acordo com seu tempo de contração.

- Contração curta (bloco 1):

No caso de uma contração curta, o sistema entra no bloco 1, que possui as funções de rotação do punho da prótese. Em seguida, já dentro do bloco 1, o usuário deve definir a direção da rotação do punho por meio de uma nova contração. Assim, caso seja realizada uma contração curta a prótese irá girar no sentido anti-horário, e caso seja realizada uma contração longa a prótese irá girar no sentido horário. A partir do momento que a prótese começa a girar, o usuário pode dar o comando de parada por meio de uma contração curta, acendendo novamente o LED e retornando o sistema para o estado inicial.

- Contração longa (bloco 2):

No caso de uma contração longa, o sistema entra no bloco 2, que possui as funções relacionadas ao tipo de preensão da prótese. Dentro do bloco 2, o usuário pode executar duas ações por meio de uma nova contração. Caso o usuário realize uma contração curta, o tipo de preensão atual é trocado, ou seja, se o tipo de preensão atual for Tripod será alterado para Power, caso contrário, será alternado para Tripod. Em seguida, e em seguida é emitido um alerta sonoro (beep), o LED acende novamente e o sistema volta para o estado inicial. Por outro lado, caso o usuário realize uma contração longa, irá ocorrer o fechamento da prótese, que se manterá fechada até que o usuário realize uma nova contração, acendendo novamente o LED e retornando o sistema para o estado inicial.

O peso total do sistema de controle manual foi igual 1048 gramas, já que a câmera DVS128 mais o seu suporte de encaixe no bypass não foram necessários nesse sistema.

### 6.2.3 EXPERIMENTO COM VOLUNTÁRIOS

Para demonstrar ambos os sistemas de controle funcionando e posteriormente realizar comparações entre eles, foi realizado um experimento com voluntários onde eles precisavam

executar uma tarefa de *pick and place*, ou seja, pegar um objeto que está em uma determinada posição e levá-lo para outra. Essa seção apresenta todo o procedimento necessário para a realização do experimento. Primeiramente, apresenta-se como foram escolhidos os voluntários para a execução desse trabalho. Em seguida, é descrita a tarefa a ser executada pelos voluntários. Por último, é descrito o funcionamento do script de escolha de objetos contidos na tarefa realizada pelo voluntário.

### **6.2.3.1 Grupo amostral**

Foram selecionados 10 voluntários destros, já que a prótese utilizada nesse trabalho é de mão direita. Os voluntários foram divididos aleatoriamente em dois grupos iguais, de modo que cada grupo possuiu 5 voluntários. Assim, os voluntários do primeiro grupo realizaram o experimento usando o sistema de controle semiautônomo e os voluntários do segundo grupo realizaram o experimento usando o sistema de controle manual. Os voluntários possuíam idades entre 22 e 31 anos, de modo que a idade média foi de  $25 \pm 2,43$ . Todos os voluntários se declararam saudáveis e aptos a participar do experimento proposto, sendo que nenhum deles possuíam experiência no controle de próteses mioelétricas. Esta pesquisa foi aprovada pelo Comitê de Ética em Pesquisa (CEP) da Universidade Federal de Uberlândia (UFU), com parecer (CAAE: 60910916.0.0000.5152).

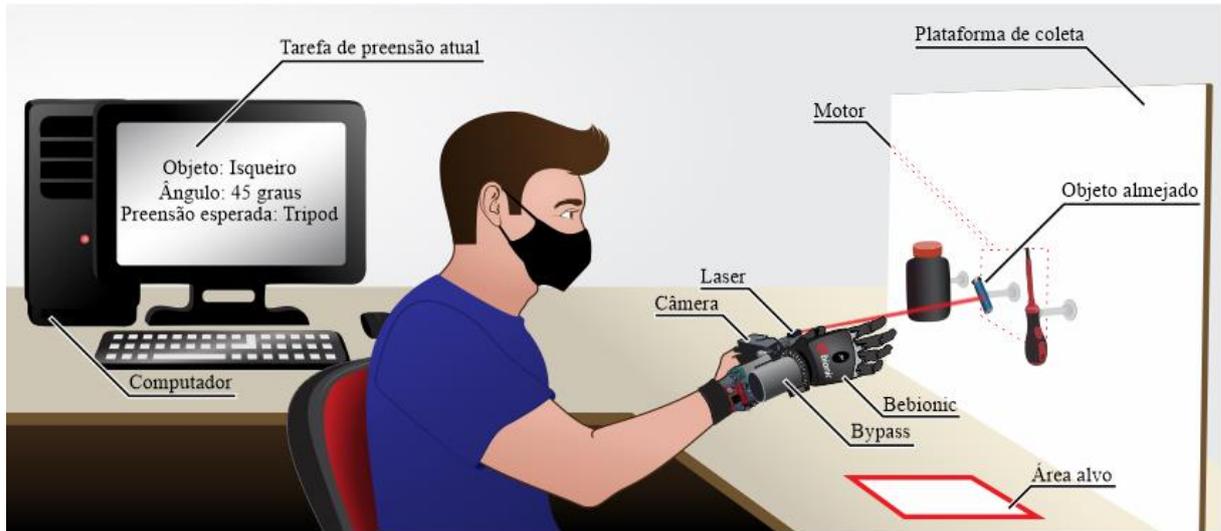
### **6.2.3.2 Ambiente de coleta**

O ambiente de coleta consiste na estrutura criada para realizar o experimento proposto. A Figura 38 ilustra as principais partes do ambiente de coleta para ambos os experimentos com o sistema de controle semiautônomo e manual, contendo o aparato necessário para a realização dos experimentos.

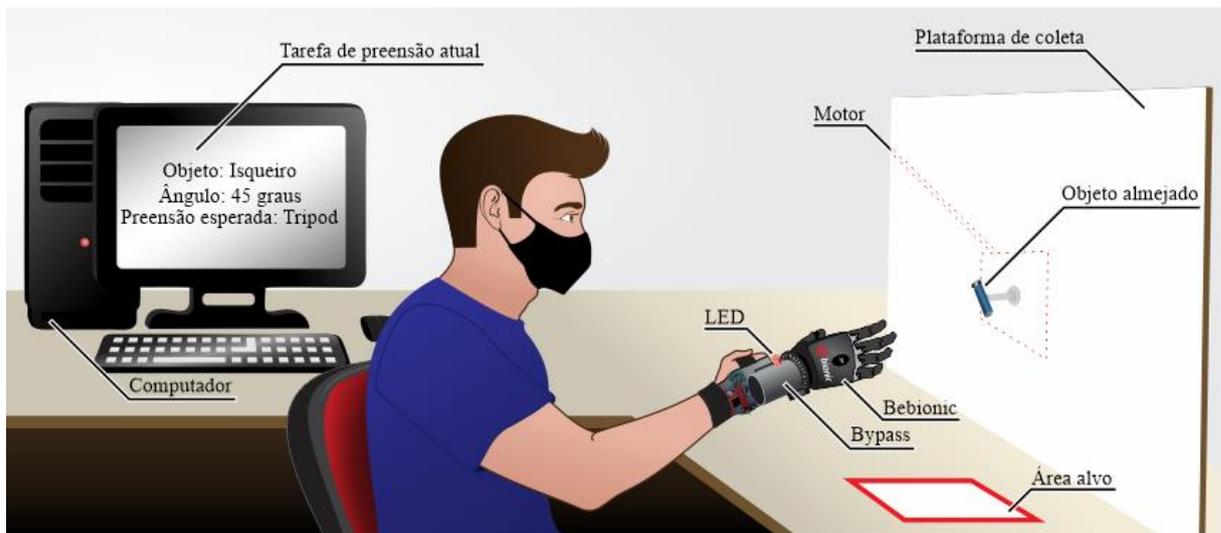
Primeiramente, foi desenvolvida uma plataforma para posicionar os objetos que serão capturados pelo voluntário. Essa plataforma possui três bases, de modo que é possível encaixar um objeto em cada uma delas. Os objetos são fixados nessas bases por meio de uma fita de velcro, que permite encaixar e remover o objeto da plataforma facilmente. Além disso, com a fita de velcro é possível posicionar o objeto na plataforma em diferentes angulações.

O voluntário foi requisitado a sentar de maneira confortável e de forma ereta em uma cadeira. Em seguida, o bypass desenvolvido foi fixado no braço do voluntário. Com o intuito de evitar fadiga, foi posicionado um suporte para o voluntário apoiar seu braço durante os intervalos da tarefa proposta.

Figura 38 – Ilustração do ambiente de coleta durante a realização dos experimentos utilizando os sistemas de controle (a) semiautônomo e (c) manual.



(a)



(b)

A última parte do ambiente de coleta consiste no computador usado para executar os processamentos necessários. Além disso, esse mesmo computador foi usado para rodar o script de escolha de objetos (detalhado na seção 6.2.3.4). O computador utilizado possui um processador Intel® Core™ i7-8700K, 16 GB de memória RAM e uma placa de vídeo NVIDIA GeForce RTX 2070.

### 6.2.3.3 Tarefa

A tarefa é composta por três fases. Na primeira fase, o voluntário foi solicitado a realizar a preensão dos mesmos objetos usados no treinamento do modelo de CNN. Essa fase consiste em 5 blocos. Por sua vez, cada bloco possui 6 tarefas de preensão, de modo que o voluntário

pegue aleatoriamente 6 dos 20 objetos existentes (3 objetos definidos com Tripod e 3 objetos definidos como Power). Assim, o voluntário terá realizado 30 preensões no final dessa fase. No caso do sistema de controle manual, foi solicitado que o voluntário ajustasse o ângulo da prótese o mais próximo possível do ângulo que o objeto almejado foi posicionado. Entretanto, não foi especificado qual era exatamente esse ângulo, de modo que o voluntário usou apenas o seu feedback visual para efetuar essa correção.

A estrutura da segunda fase segue os mesmos princípios da primeira. Entretanto, o voluntário foi solicitado a realizar a preensão de novos objetos, os quais não foram usados na etapa de treinamento do modelo de CNN. Assim como na primeira, essa fase consiste em 5 blocos. Cada bloco é composto por 5 tarefas de preensão, de modo que o voluntário pegue 4 dos 10 objetos novos. Assim, o voluntário terá realizado 20 preensões no final dessa fase. Como essa fase visa analisar o desempenho do modelo de classificação desenvolvido, ela ocorreu apenas para o experimento do sistema de controle semiautônomo.

A terceira fase consiste na avaliação subjetiva do voluntário sobre as duas primeiras fases usando o questionário NASA-TLX apresentado na seção 6.2.3.5. O voluntário foi solicitado a preencher o questionário, respondendo atentamente cada uma das perguntas. O pesquisador responsável pela coleta esteve disponível durante todo o processo, respondendo as possíveis dúvidas do voluntário sobre as perguntas contidas no questionário. Essa fase ocorreu para ambos os experimentos dos sistemas de controle semiautônomo e manual.

#### **6.2.3.4 *Script de escolha de objetos***

A ordem que os objetos são dispostos para que o voluntário realizasse a preensão é gerada de forma aleatória por meio de um script em Python. Assim, mesmo que o número de preensões seja menor que o número total de objetos, todos os objetos têm a mesma probabilidade de serem escolhidos. Além disso, o script gera a orientação em graus que cada um dos objetos deve ser colocado. Essa orientação é gerada de forma aleatória em um intervalo de -90 até 90 graus. Por último, no experimento realizado com o sistema de controle semiautônomo, o script indica mais dois objetos para ficarem nas outras duas áreas disponíveis da plataforma usada para encaixar os objetos.

Como ilustrado na Figura 38, foi usado um motor de passo na parte de trás da plataforma para ajustar com mais precisão o ângulo do objeto. Primeiramente, o objeto atual é posicionado verticalmente de acordo com seu maior lado. Em seguida, o motor ajusta automaticamente o objeto para o mesmo ângulo que foi definido pelo script.

### 6.2.3.5 *Questionário NASA-TLX*

Ao final dos experimentos, os voluntários foram requisitados a preencher o questionário *Task Load Index* (NASA-TLX). Esse questionário foi desenvolvido no ano de 1988 pela *National Aeronautics and Space Administration* (NASA) com o intuito de obter uma análise subjetiva da carga de trabalho exigida para uma tarefa específica (Hart & Staveland, 1988).

Essa análise é feita a partir de seis fatores, sendo eles: demanda mental, demanda física, demanda temporal, performance, esforço e frustração. Utilizando as respostas fornecidas pelo voluntário, é possível realizar uma análise geral da tarefa ou separadamente de acordo com cada um dos fatores, podendo compreender em quais pontos o voluntário foi mais exigido.

No Apêndice 2 está anexado o questionário NASA-TLX utilizado para coletar as informações de carga de trabalho fornecidas por cada um dos voluntários. Além disso, todos os fatores pertencentes ao questionário estão descritos de maneira detalhada na **Error! Reference source not found.** contida no apêndice, auxiliando o voluntário a compreendê-los para então preencher o questionário de forma apropriada.

O questionário NASA-TLX é feito em duas etapas. Na primeira delas, os fatores são colocados em pares, de modo que o voluntário deve escolher o fator que foi mais impactante para a sua experiência de carga de trabalho durante a tarefa realizada. Por exemplo, para o par contendo os fatores “demanda mental” e “demanda física”, o voluntário deve assinalar qual dos dois fatores foi mais impactante durante a tarefa realizada.

Na segunda parte do questionário, são feitas seis perguntas, onde cada uma está diretamente relacionada com um fator individualmente. Para responder uma pergunta, o voluntário deve assinalar um ponto da escala de classificação, que vai de 0 a 20. Cada escala de classificação possui pontos finais que descrevem a escala, por exemplo, os pontos finais que descrevem a escala do fator “esforço” são “muito baixa” para o limite inferior e “muito alta” para o limite superior. Ou seja, caso o voluntário entenda que a tarefa demandou muito esforço, ele deve assinalar um ponto da escala mais próximo do limite superior.

Após obter o questionário respondido, é necessário quantificar os resultados fornecidos por cada um dos voluntários. Dessa forma, tem-se que os resultados do questionário NASA-TLX são obtidos usando os escores de ponderação da escolha pareada e os escores individuais para cada fator de carga de trabalho (Equação 12).

$$\bar{x} = \frac{\sum_{i=1}^6 (x_i * w_i)}{\sum_{i=1}^6 w_i} \quad (12)$$

Onde  $x$  é a classificação do fator de carga de trabalho e  $w$  é o peso resultante da soma das ocorrências do fator de carga de trabalho da etapa de escolha pareada.

# Capítulo 7

---

## 7. RESULTADOS

Nesse capítulo estão apresentados os resultados obtidos por meio dos experimentos realizados no presente estudo. Esses resultados estão divididos em duas partes, sendo elas relativas aos experimentos offline e em tempo real. Por sua vez, os experimentos offline são referentes aos experimentos de validação do sistema de processamento de imagens descrito na seção 6.1.3. Por outro lado, os resultados alcançados por meio experimentos em tempo real são relativos à avaliação dos dois sistemas de controle desenvolvidos, os quais foram testados com voluntários.

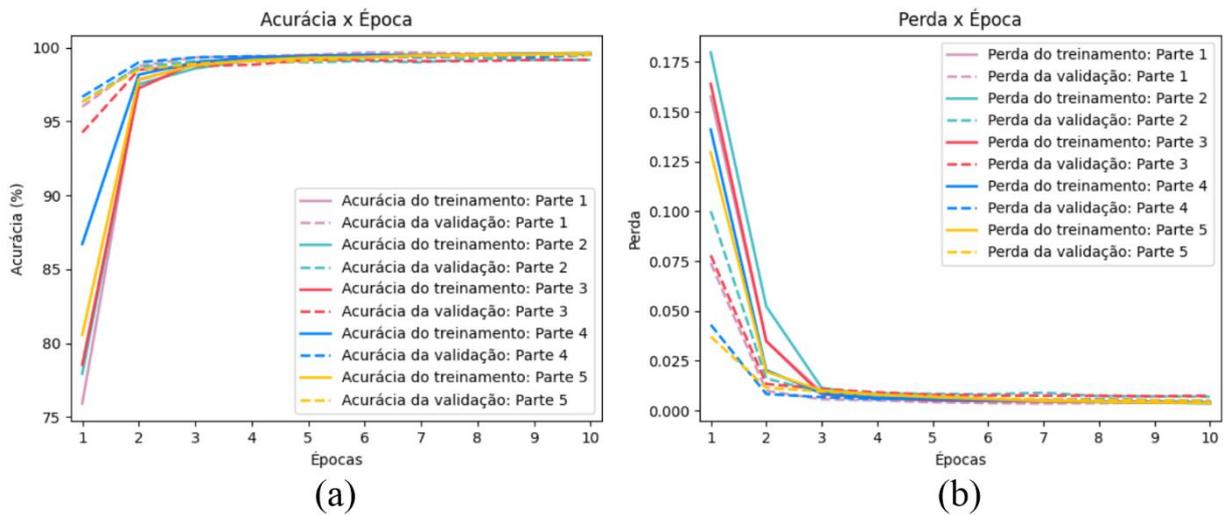
### 7.1 EXPERIMENTOS OFFLINE

O primeiro passo para os experimentos offline foi construir o conjunto de imagens formadas a partir de gravações de objetos por meio da câmera neuromórfica DVS128. Como já foi apresentado na seção 6.1.2.3, foram obtidas 9000 imagens dos 30 objetos selecionados, os quais são comumente usados em tarefas do dia a dia.

#### 7.1.1 VALIDAÇÃO CRUZADA DO MODELO DE CNN

Seguindo a metodologia apresentada na seção 6.1.3.2.4, foi feita a validação cruzada do modelo de CNN proposto, verificando o quão bom é sua performance para diferentes combinações do conjunto de imagens desenvolvido. A Figura 39 apresenta a evolução da acurácia e perda durante as 10 épocas de treinamento.

Figura 39 – Evolução do modelo durante as 10 épocas do período de treinamento. As linhas contínuas indicam o desempenho do conjunto de treinamento, enquanto as tracejadas indicam o conjunto de teste para cada uma das partes da validação cruzada. (a) Acurácia por época e (b) Perda por época.



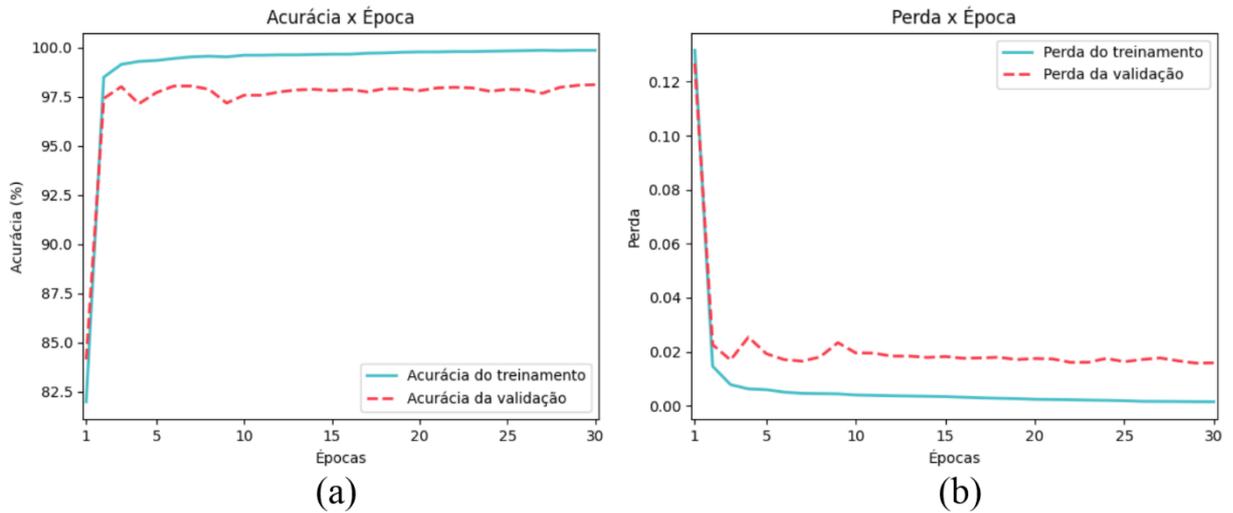
A Tabela 3 apresenta a acurácia obtida para cada uma das 5 partes da validação cruzada. A acurácia média foi de  $99,3167\% \pm 0,1616\%$ , enquanto a perda média foi de  $0,0059 \pm 0,0016$ .

Tabela 3 – Acurácia e perda obtidas em cada uma das 5 partes da validação cruzada.

Parte	Acurácia (%)	Perda
1	99,4167	0,0046
2	99,1667	0,0072
3	99,0833	0,0082
4	99,4167	0,0045
5	99,5000	0,0048
Média $\pm \sigma$	$99,3167 \pm 0,1616$	$0,0059 \pm 0,0016$

Na sequência da validação cruzada, o modelo foi novamente treinado para ser usado posteriormente no sistema de controle semiautônomo. A Figura 40 apresenta o desempenho da classificação do modelo durante as 30 épocas de treinamento. Ao fim do treinamento, o modelo obteve uma acurácia de  $97,8999\%$  e perda de  $0,0177$  para o conjunto de teste, composto pelas imagens obtidas do grupo 1 de objetos.

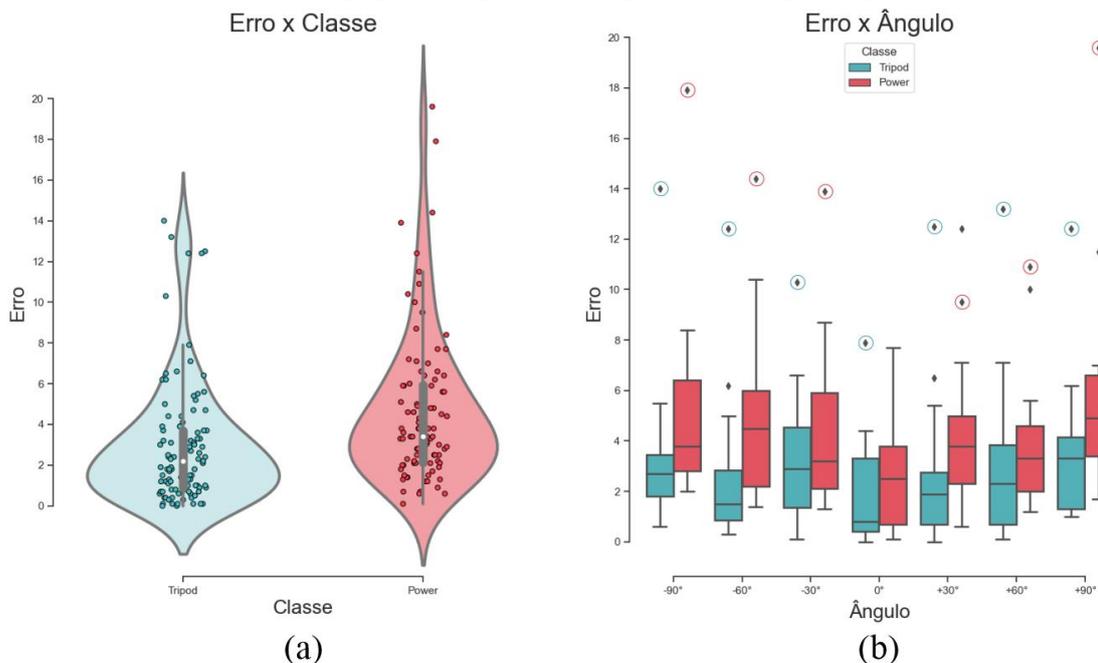
Figura 40 – Evolução do modelo durante as 30 épocas do período de treinamento. (a) Acurácia por época durante o treinamento (azul) e teste (vermelho). (b) Perda por época durante o treinamento (azul) e teste (vermelho).



### 7.1.2 VALIDAÇÃO DO ALGORITMO DE OBTENÇÃO DO ÂNGULO

Em seguida, foi realizada a validação do sistema de obtenção do ângulo do objeto gravado. A Figura 41 apresenta os resultados obtidos por meio do processo de validação do algoritmo de obtenção do ângulo descrito na seção 6.1.3.3.

Figura 41- Erro resultante do processo de validação do algoritmo de obtenção do ângulo. (a) Erro do ângulo obtido dos 30 objetos agrupados (a) pelas classes preensão e (b) pela angulação correta.



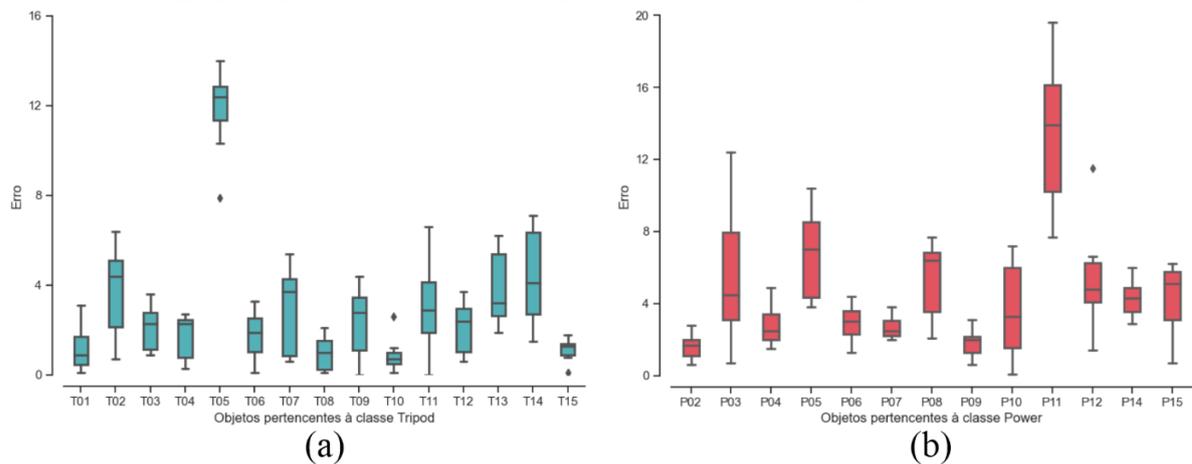
O erro médio para a classe Tripod foi de  $2,9324 \pm 2,9776$  graus, enquanto para a classe Power foi de  $4,7187 \pm 3,9062$  graus. Um teste não-paramétrico Mann-Whitney U revelou

diferença estatística entre os erros referentes aos ângulos obtidos dos 30 objetos separados pelas classes prensão Tripod e Power ( $p \approx 10^{-6}$ ).

Os pontos fora da curva (*outliers*) apresentados na Figura 41b indicam dois objetos específicos, sendo eles chave Allen (vermelho) e caneca (azul) pertencentes as classes Tripod e Power respectivamente. Se analisados separadamente, tem-se que o erro médio foi de  $11,8143 \pm 1,9075$  graus para a chave Allen e  $13,4143 \pm 4,0336$  graus para a caneca.

A Figura 42 apresenta o erro obtido por objeto. Na Figura 42a estão apresentados o erro de todos os 15 objetos pertencentes a classe Tripod. Por outro lado, a Figura 42b apresenta o erro de 13 dos 15 objetos pertencentes a classe Power, já que os objetos bola e laranja foram descartados por não possuírem um lado significativamente maior que o outro.

Figura 42 - Erro resultante do processo de validação do algoritmo de obtenção do ângulo das 7 angulações base agrupadas para cada um dos objetos pertencentes a classe (a) Tripod e (b) Power.



## 7.2 EXPERIMENTOS EM TEMPO REAL

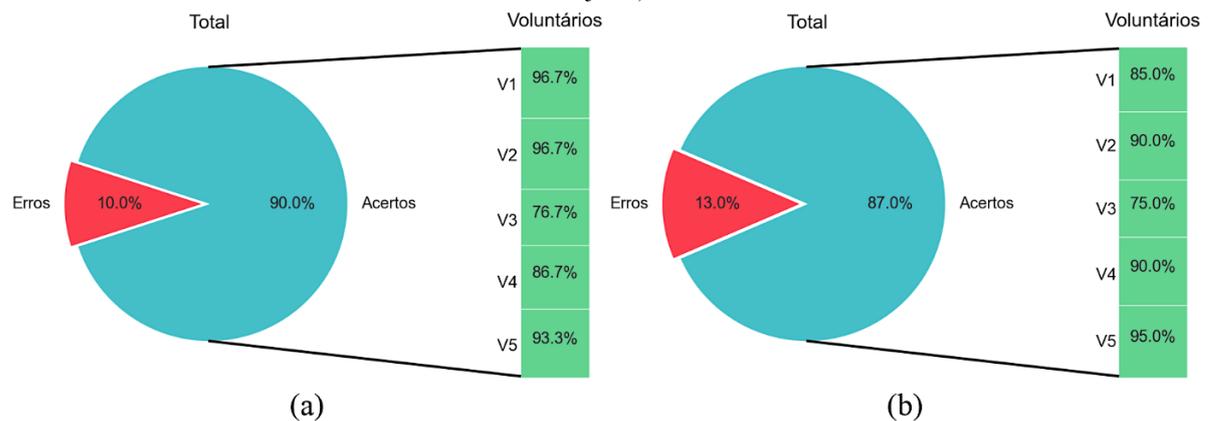
Os experimentos com os sistemas de controle em tempo real foram realizados após a validação do sistema de processamento de imagens. No primeiro experimento, 5 voluntários usaram o sistema de controle semiautônomo, enquanto no segundo outros 5 voluntários usaram o sistema de controle manual. Ambos os experimentos seguiram o protocolo descrito na seção 6.2.3.

Foi cronometrado o tempo total de execução das tarefas de prensão para cada um dos voluntários. Isso é, os períodos de ajuste do bypass no braço do voluntário e preenchimento do formulário NASA-TLX não fazem parte desse tempo. Dessa forma, tem-se que o tempo total de execução das tarefas de prensão foi de  $64,2 \pm 5,7$  e  $55,4 \pm 2,06$  minutos para os sistemas de controle semiautônomo e manual, respectivamente.

### 7.2.1 OBJETOS APRENDIDOS X OBJETOS NÃO APRENDIDOS

A primeira parte dos resultados consiste na análise de robustez do sistema de controle em tempo real quando confrontado com objetos que não fizeram parte do treinamento do seu modelo de classificação. A Figura 43 apresenta a porcentagem de acertos e erros obtidos do experimento usando o sistema de controle semiautônomo para ambos os conjuntos de imagens referentes aos grupos de objetos 1 e 2.

Figura 43 – Porcentagem de acertos e erros do classificador para as suas etapas do experimento usando o sistema de controle semiautônomo. Tarefa de prensão realizada (a) com mesmos objetos usados no treinamento do modelo (grupo 1 de objetos) e (b) com objetos nunca vistos na etapa de treinamento do modelo (grupo 2 de objetos).



A partir dos dados de acurácia dos 5 voluntários para as duas etapas, foi realizado um teste estatístico de Wilcoxon, o qual não foi possível rejeitar a hipótese nula que as distribuições são diferentes, não revelando diferença estatística entre as duas etapas de experimentos com objetos aprendidos e não aprendidos pelo modelo de classificação ( $p = 0,625$ ).

### 7.2.2 COMPARAÇÃO ENTRE OS SISTEMAS DE CONTROLE

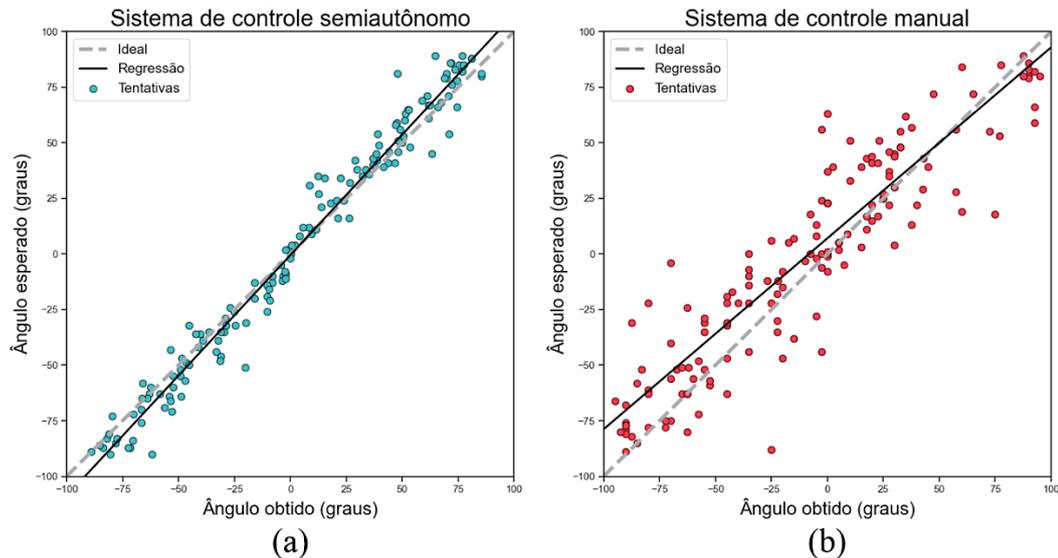
Nessa seção estão apresentados os resultados obtidos dos experimentos realizados com voluntários usando os sistemas de controle semiautônomo e manual. Entre os resultados quantitativos obtidos, estão o erro no ajuste do ângulo do punho, tempo de execução e escorramento do objeto em cada uma das 30 tarefas de prensão realizadas por cada um dos 10 voluntários, sendo 5 voluntários por sistema de controle. Por fim, serão apresentados os

resultados qualitativos obtidos da análise subjetiva da tarefa realizada pelos voluntários por meio do questionário NASA-TLX.

### 7.2.2.1 Ajuste do ângulo do punho

A Figura 44 apresenta a regressão linear (preto) das amostras de ângulo obtidas dos usuários que usaram o sistema de controle semiautônomo e manual. Além disso, as figuras apresentam a reta ideal (cinza), onde o erro obtido é igual a zero. No caso da bola, que é um objeto pertencente ao grupo 1 de objetos, o erro foi considerado sempre igual a zero, já que possui dimensões iguais e o ângulo do punho não interfere na tarefa de prensão.

Figura 44 – Ângulo esperado contra ângulo obtido, apresentando a reta ideal (cinza) onde o erro é igual a zero. Amostras obtidas dos usuários que usaram o sistema de controle (a) semiautônomo e (b) manual.



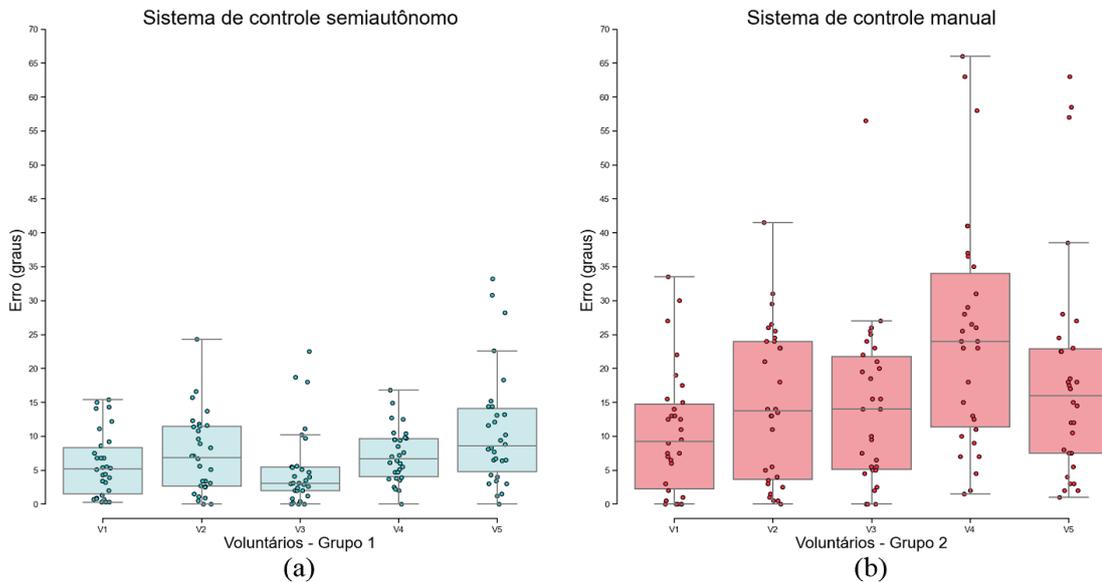
Os valores de inclinação e interceptação da reta obtidos pela regressão linear foram respectivamente 1,08 e -0,52 os dados dos voluntários do sistema de controle semiautônomo, e 0,84 e 8,27 para os dados dos voluntários do sistema de controle manual.

Além disso, foram calculados o erro quadrático médio (MSE) e R-quadrado a partir da reta ideal (45 graus), isto é, quando o erro é igual a zero. Assim, o valor do MSE foi de 91,49 graus para os voluntários do sistema de controle semiautônomo e 479,53 graus para os voluntários do sistema de controle manual. Por outro lado, o valor do R-quadrado foi de 0,96 para os voluntários do sistema de controle semiautônomo e 0,82 para os voluntários do sistema de controle manual.

A Figura 45 apresenta a distribuição do erro absoluto entre o ângulo esperado e o ângulo obtido para as 30 tarefas de prensão realizada por cada um dos 10 voluntários que realizaram

os experimentos. O valor médio do erro foi de  $7,28 \pm 6,23$  graus para os voluntários que usaram o sistema de controle semiautônomo e  $17,64 \pm 16,88$  graus para os voluntários que usaram o sistema de controle manual.

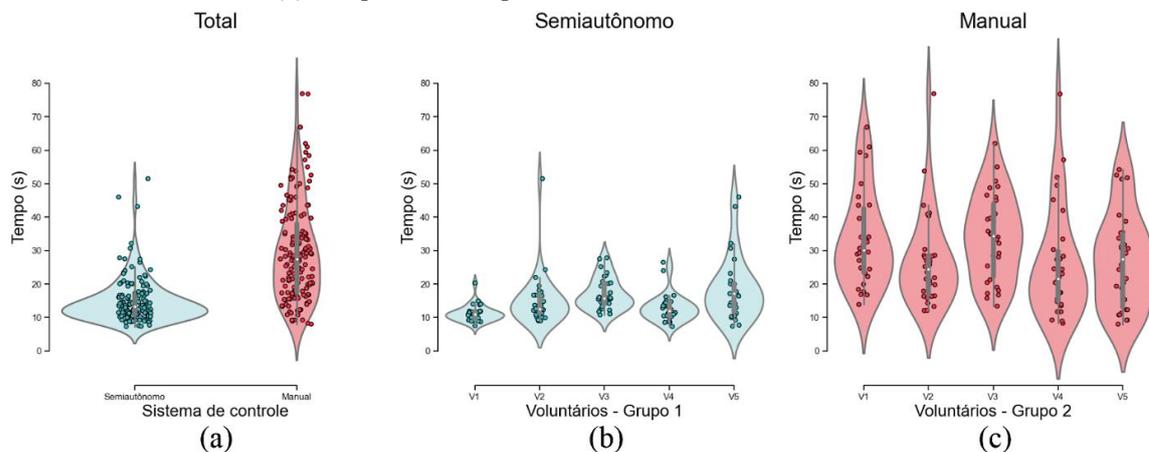
Figura 45 – Erro do ângulo obtido para ambos os experimentos realizados usando o sistema de controle (a) semiautônomo e (b) manual.



### 7.2.2.2 Tempo de execução

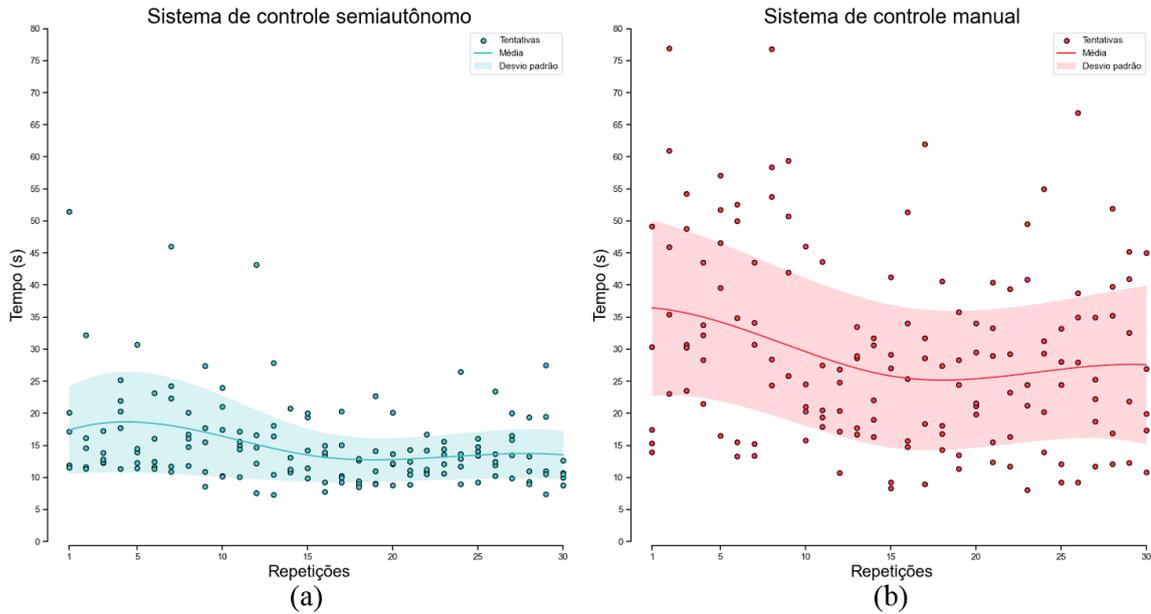
A Figura 46 apresenta os dados de tempo de execução da tarefa de prensão. Foram realizadas 300 repetições, sendo 150 por sistema de controle e 30 por voluntário. O tempo de execução médio foi de  $14,95 \pm 6,73$  segundos para o sistema de controle semiautônomo e  $29,41 \pm 14,60$  segundos para o sistema de controle manual.

Figura 46 – Tempo de execução da tarefa de prensão para ambos os experimentos realizados. (a) Todos os tempos de execução por sistema de controle. (b) Tempo execução por voluntário usando o sistema de controle semiautônomo. (c) Tempo execução por voluntário usando o sistema de controle manual.



A Figura 47 apresenta os dados do tempo de execução de tarefa ao longo das 30 repetições realizadas. Além disso, a figura apresenta a curva de tendência gerada por meio da média móvel dos dados com uma janela de 5 repetições.

Figura 47 – Tempo de execução da tarefa de prensão para cada uma das 30 repetições realizadas. Todas as repetições executadas pelos voluntários do sistema de controle (a) semiautônomo e (b) manual.

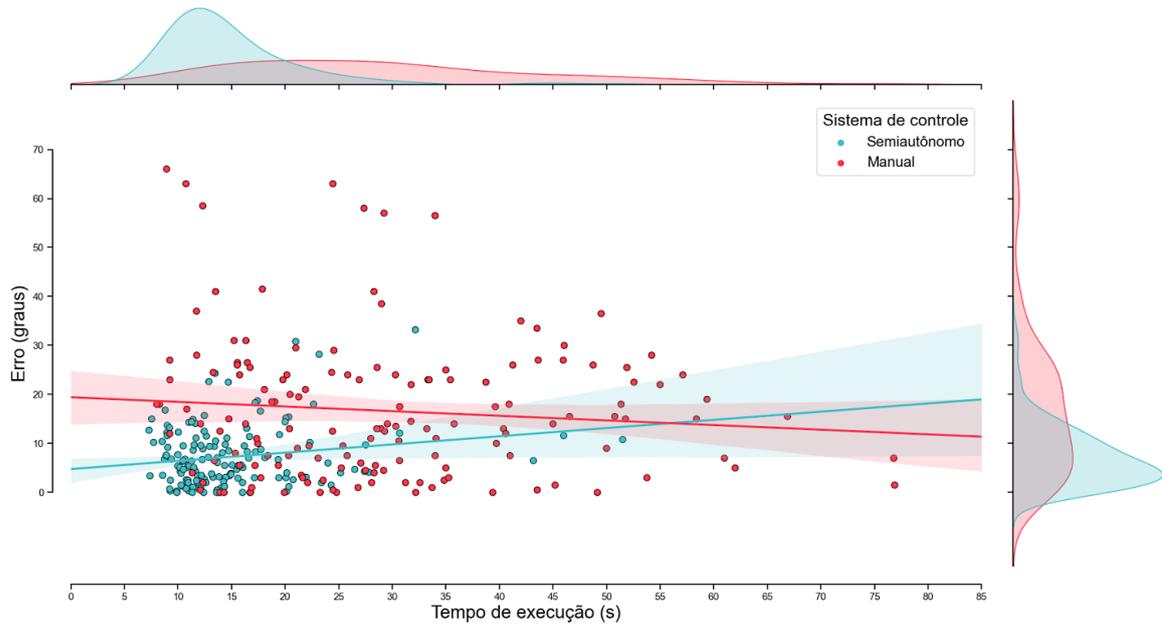


A partir dos dados da média móvel, tem-se que a média para o sistema de controle semiautônomo foi de  $17,37 \pm 6,78$  segundos para as 5 primeiras repetições e  $13,52 \pm 3,79$  segundos para as 5 últimas repetições, resultando em uma diminuição de 3,85 segundos no valor médio do tempo de execução da tarefa de prensão.

Por outro lado, tem-se que a média para o sistema de controle manual foi de  $36,44 \pm 13,65$  segundos para as 5 primeiras repetições e  $27,56 \pm 12,35$  segundos para as 5 últimas repetições, resultando em uma diminuição de 8,87 segundos no valor médio do tempo de execução da tarefa de prensão.

A Figura 48 apresenta o gráfico de dispersão para os valores do erro no ajuste do ângulo do punho e tempo de execução de tarefa para todas as repetições realizadas, destacando quais são referentes a cada sistema de controle. Além disso, são apresentadas as regressões lineares para dos dados para ambos os sistemas de controle.

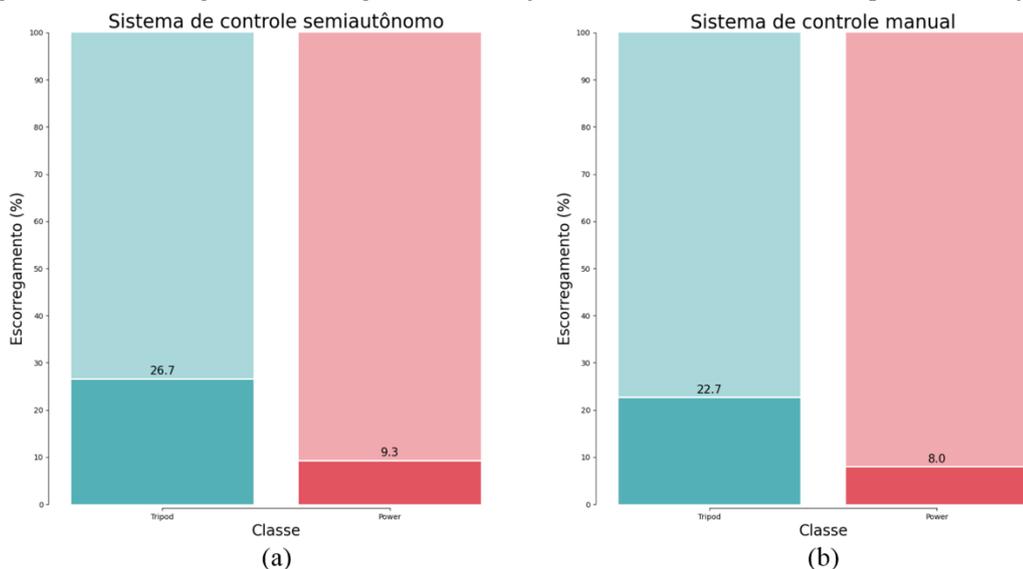
Figura 48 - Erro no ajuste do ângulo do punho pelo tempo de execução de tarefa para todas as repetições realizadas, apresentando a distribuição dos dados tanto no eixo vertical, quanto no eixo horizontal.



### 7.2.2.3 Escorregamento do objeto

A Figura 49 apresenta a porcentagem de objetos que escorregaram durante a transferência de lugar do objeto almejado durante a tarefa de prensão. No total, foram realizadas 75 repetições por classe para cada um dos sistemas de controle.

Figura 49 – Porcentagem de escorregamento dos objetos durante a transferência de posição do objeto.



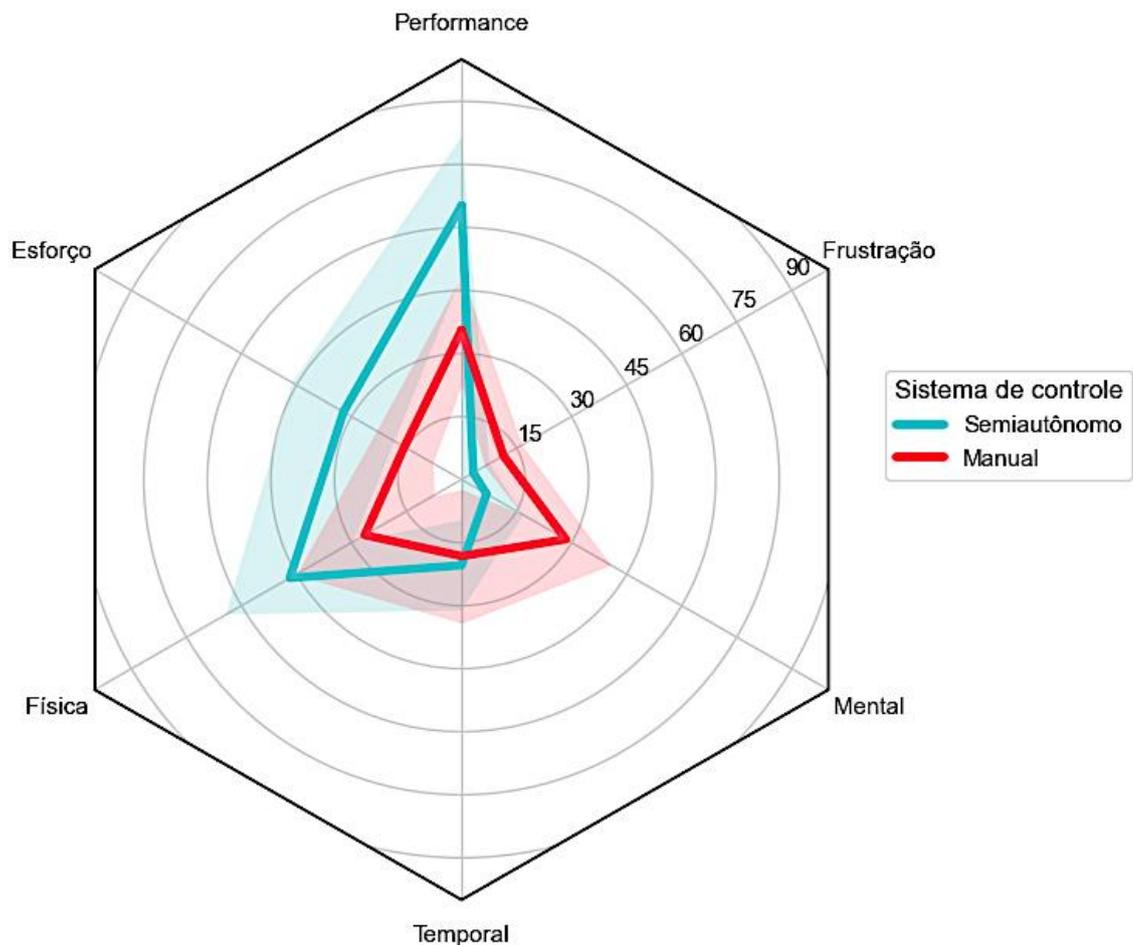
Foram realizadas 75 repetições da tarefa de prensão por classe para cada sistema de controle. Assim, 20 objetos da classe Tripod e 7 objetos da classe Power caíram durante a transferência de posição usando o sistema de controle semiautônomo, enquanto 17 objetos da

classe Tripod e 6 objetos da classe Power caíram durante a transferência de posição usando o sistema de controle manual.

#### 7.2.2.4 Análise subjetiva – Questionário NASA-TLX

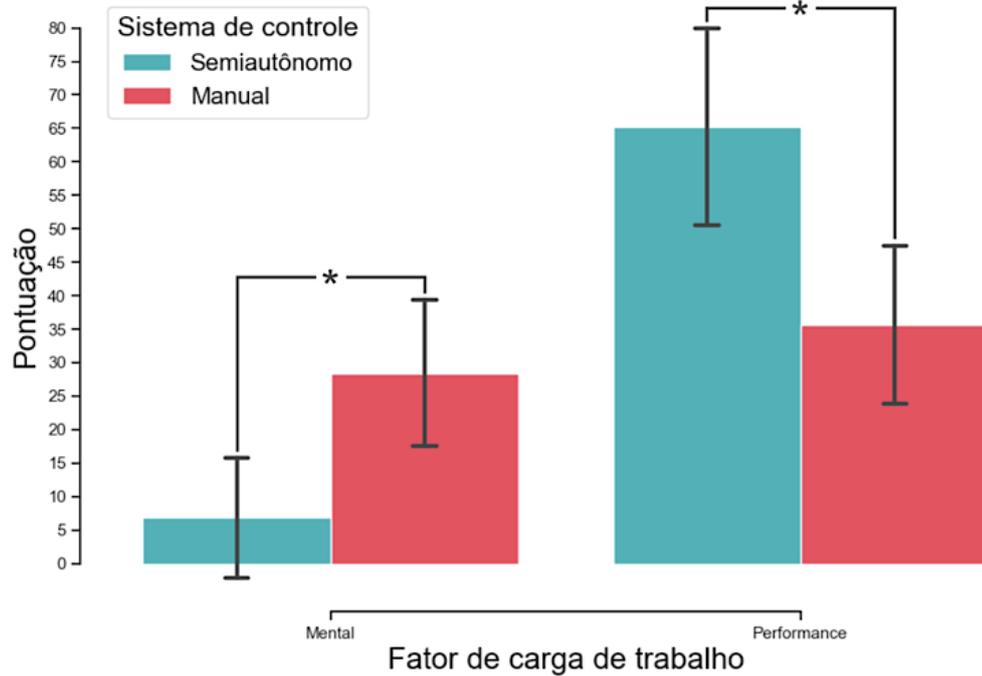
A Figura 50 apresenta os resultados obtidos por meio do questionário NASA-TLX, o qual foi preenchido por todos os 10 voluntários que participaram dos experimentos presentes nesse trabalho. Na figura, estão apresentadas as pontuações dos 6 fatores de carga de trabalho existentes no questionário NASA-TLX, que foram obtidos por meio da Equação (12).

Figura 50 – Média e desvio padrão das pontuações para cada fator de carga de trabalho obtidas por meio do questionário NASA-TLX.



Os fatores de esforço mental e performance foram analisados separadamente devido a sua importância no presente estudo. A Figura 51 apresenta esses fatores, que estão divididos de acordo com o sistema de controle usado. Tem-se que a média dos voluntários que realizaram o experimento com o sistema de controle semiautônomo foi de  $6,8 \pm 10,03$  pontos para o fator de carga mental e  $65,2 \pm 16,39$  pontos para o fator de performance.

Figura 51 – Comparação entre os dois sistemas de controle por meio da análise subjetiva dos 10 voluntários que participaram dos experimentos. \* significa diferença estatística entre os grupos ( $p = 0,05$ ).



O teste não-paramétrico Mann-Whitney U revelou diferença estatística entre as pontuações médias obtidas para cada sistema de controle, tanto para o fator de carga mental, quanto para o fator de performance ( $p = 0,05$  e  $n = 5$ ).

# Capítulo 8

---

## 8. DISCUSSÃO

Nesse capítulo, realizamos a análise dos resultados obtidos por meio dos experimentos realizados, ressaltando os principais pontos encontrados durante tais experimentos. Em primeiro lugar, estão apresentadas as discussões resultantes dos experimentos offline, referentes principalmente a validação dos algoritmos de processamento de imagens. Em seguida, são apresentadas as discussões resultantes aos experimentos em tempo real, comparando o sistema de controle semiautônomo proposto com outro sistema de controle baseado em chaveamento do sinal de controle do usuário (manual). Por fim, estão apresentadas as considerações finais em relação ao sistema proposto, comparando com os principais trabalhos pertencentes ao estado da arte no controle de próteses de membros superiores.

### 8.1 EXPERIMENTOS OFFLINE

Por meio dos experimentos offline, foi possível validar ambas as etapas de classificação do tipo de prensão e obtenção do ângulo do objeto do sistema de controle semiautônomo proposto.

O modelo de CNN construído tem como inspiração o modelo LeNet-5 desenvolvido pelo pesquisador Yann LeCun (LeCun et al., 1998). O LeNet-5 foi desenvolvido para classificar imagens binárias de números contidas no conjunto de dados MNIST. Desse modo, essa estrutura de rede vai de encontro com o objetivo da etapa de classificação desse trabalho.

A validação cruzada da classificação do tipo de prensão mostrou que o sistema foi capaz de classificar ambas as classes Tripod e Power independentemente do conjunto de treinamento e teste usados, obtendo uma acurácia média de  $99,3167 \pm 0,1616$  e perda de  $0,0059 \pm 0,0016$  para todas as partes de validação.

Esse método de validação tem sido amplamente utilizado para validar modelos de ML, pois todas as amostras contidas no conjunto de dados são utilizadas tanto no treinamento quanto no teste, possibilitando então que o modelo seja mais fielmente avaliado perante diferentes combinações das amostras disponíveis.

A partir dos dados de treinamento apresentados na Figura 39, pode-se notar que a acurácia obtida pelo modelo subiu rapidamente para mais de 99% para ambos os conjuntos de treinamento e teste. Isso ocorreu principalmente por dois motivos. O primeiro está relacionado a quantidade de amostras usadas no treinamento do modelo, isto é, 4800 amostras.

O segundo motivo, que está diretamente ligado ao primeiro, foi pelo modelo ser binário, ou seja, ele possui apenas duas classes de saída (Tripod e Power). A combinação desses dois fatores resulta em um treinamento relativamente simples, capaz de convergir para uma acurácia alta em poucas épocas de treinamento.

Embora a tarefa de classificar entre duas classes seja relativamente simples, deve-se notar que cada classe não foi de fato homogênea, já que o conjunto de imagens de cada classe foi composto por 20 objetos diferentes. O principal motivo dessa variedade de objetos por classe é que o objetivo dessa etapa não foi classificar qual era o objeto contido na cena, mas sim qual seria o tipo de preensão mais adequado para o seu formato. Essa heterogeneidade dos dados pode dificultar significativamente a performance do modelo quando ele for confrontado com imagens de objetos nunca vistos.

Um dos pontos mais importantes do presente estudo é apresentar um sistema que seja robusto para diferentes tipos de objetos, seja eles usados no treinamento ou não. Dessa forma, a Figura 40 apresentou a acurácia do sistema durante seu treinamento para ambos os conjuntos de imagens extraídas dos grupos de objetos 1 e 2 (Tabela A1 – Apêndice A), de modo que as imagens referentes ao grupo 2 de objetos não foram em nenhum momento usadas para o treinamento do modelo.

Embora a acurácia do modelo ter sido 1,4168% menor para as imagens de objetos não usadas em seu treinamento em relação à acurácia média da validação cruzada, pode-se considerar que o modelo se adaptou bem, pois foi capaz de alcançar uma acurácia de 97,89% e perda de 0,0177 para as 3000 imagens novas pertencentes ao grupo 2 de objetos.

Após a validação do modelo de CNN desenvolvido para classificar o tipo de preensão necessária de acordo com o formato do objeto, foi feita a validação do algoritmo de obtenção do ângulo do objeto. A validação mostrou que o sistema conseguiu obter bons resultados para as duas classes de objetos, com um erro médio de  $2,93 \pm 2,98$  graus para a classe Tripod e  $4,72 \pm 3,91$  graus para a classe Power.

A partir desses resultados, pode-se notar que a classe de preensão Power teve um erro médio de 1,79 graus maior que a classe de preensão Tripod. A razão dessa diferença no erro médio se deu pelas proporções dos objetos contidos em cada classe. Isto é, objetos definidos como Tripod possuem em sua maioria, um lado muito menor que o outro, de modo que as PCs

extraídas por meio da técnica de PCA se ajustam muito bem aos lados do objeto. Por outro lado, os objetos definidos como Power normalmente são maiores e com lados parecidos, fazendo com que as PCs apareçam nas diagonais desses objetos e não nos seus lados como o esperado.

Outro ponto que chama atenção em relação aos resultados alcançados da validação é o alto desvio padrão encontrado para ambas as classes. Como discutido anteriormente, cada uma das classes Tripod e Power possuem uma grande variedade de objetos, de forma que cada um deles tem suas peculiaridades. Dessa forma, o algoritmo de obtenção do ângulo do objeto se adaptou melhor para alguns tipos de objetos do que para outros.

Dentre os 30 objetos selecionados, dois se destacaram negativamente, sendo eles a chave Allen e a caneca. A chave Allen, que faz parte do grupo de objetos pertencentes a classe de preensão Tripod teve um erro médio 4,16 vezes maior que o erro médio do seu grupo. Esse erro tão elevado ocorreu devido ao seu formato em “L”, que faz com que as PCs extraídas apareçam em suas diagonais, e não no seu menor lado.

Já a caneca, que pertence a classe de preensão Power, teve um erro médio 2,86 vezes maior que o erro médio do seu grupo. O motivo por trás desse aumento do erro foi devido a sua alça, que torna os lados do objeto muito parecidos dependendo do ângulo de gravação. Conseqüentemente, as PCs extraídas tendem a ficar em um meio termo dos lados do objeto, aumentando o erro do ângulo obtido.

## 8.2 EXPERIMENTOS EM TEMPO REAL

Com todas as etapas do sistema de processamento de imagens validado e funcionando, foi então realizado os experimentos com voluntários. Primeiramente, assim como nos experimentos offline, foi analisado quão bem o sistema de controle semiautônomo performou para objetos não usados em seu treinamento. Dessa vez, esse experimento foi feito em tempo real com 5 voluntários.

Como apresentado na Figura 43, os resultados gerais foram muito próximos, sendo o erro médio para grupo de objetos usados no treinamento 3% menor em relação ao grupo de objetos não usados no treinamento. Isso reafirma a capacidade de generalizar diferentes objetos, usando apenas o seu formato para diferenciar o tipo de preensão mais apropriada.

Em relação aos sistemas de controle analisados, tem-se que ambos foram desenvolvidos ao longo desse trabalho e posteriormente testados com dois grupos diferentes compostos por 5 voluntários. O desenvolvimento do sistema de controle semiautônomo foi o foco principal desse

trabalho, pois é nesse sistema que estão as contribuições científicas propostas. Por outro lado, o sistema de controle manual foi desenvolvido com o intuito de efetuar uma comparação entre os dois sistemas, apontando as principais limitações atuais no controle de próteses baseado em blocos de movimentos ou CD.

Por meio dos resultados exibidos na Figura 44, pode-se observar que os dados do sistema de controle manual estão mais dispersos em relação ao sistema de controle semiautônomo. Essa afirmação é comprovada por meio do cálculo do R-quadrado e MSE, os quais foram muito maiores para o sistema de controle manual.

Para explicar isso, tem-se que no sistema de controle manual o ângulo do punho da prótese era definido pelo próprio voluntário. Isto é, o voluntário observava o ângulo do objeto almejado e em sequência corrigia ele mesmo o ângulo do punho, com o auxílio apenas da sua visão. Em contrapartida, o sistema de controle semiautônomo corrigia automaticamente a angulação do punho de acordo com as PCs extraídas do objeto almejado, de modo que o voluntário não tinha a preocupação de definir qual o melhor ângulo do punho da prótese para realizar a preensão do objeto almejado.

Outro ponto interessante a se destacar é o tempo de execução de tarefa, que foi consideravelmente menor para os voluntários que usaram o sistema de controle semiautônomo. Novamente, nota-se que no sistema de controle manual, o voluntário precisava realizar todas as etapas do controle de próteses manualmente. Isso é, o voluntário precisava corrigir o ângulo do punho para o mais adequado e alterar o tipo de preensão (se necessário) de acordo com o ângulo e formato do objeto almejado. Essa sequência de processos fez com que os voluntários que usaram o sistema de controle manual precisassem de mais tempo para realizar a mesma tarefa de preensão que os voluntários que usaram o sistema de controle semiautônomo.

Embora os voluntários que usaram o sistema de controle manual tenham demorado em média mais tempo para realizar uma tarefa de preensão, tem-se uma redução significativa no tempo de execução médio nas 5 últimas repetições em relação as 5 primeiras. Essa redução foi de 8,87 segundos para o sistema de controle manual e 3,85 segundos para o sistema de controle semiautônomo, indicando um aprendizado dos voluntários ao longo do experimento.

Assim como todo sistema, o usuário leva algum tempo para aprender como usá-lo. Entretanto, o sistema de controle manual apontou uma redução no tempo de 2,3 vezes maior que o sistema de controle semiautônomo. Uma boa justificativa é que o sistema de controle manual é um sistema mais complexo, de modo que a experiência em seu uso reduza consideravelmente o tempo de execução. Em contrapartida, o sistema de controle semiautônomo não exige que o usuário entenda e aprenda executar cada etapa de uma tarefa de

preensão, pois o sistema executa essas ações automaticamente. Conseqüentemente, a diferença entre tempo de execução no início e final do treinamento não tende a ser tão grande quanto para o sistema de controle manual.

A Figura 48 apresentou o gráfico de dispersão para os dados de erro pelo tempo de execução. Nesse gráfico, é interessante observar ambas as regressões lineares dos sistemas de controle. A reta de regressão obtida do sistema de controle manual indica que o erro tende a ser menor para as tentativas que tiveram um tempo de execução maior. Isso pode indicar que quanto mais rápido a tarefa de preensão for executada, maior será a imprecisão na correção do ângulo do punho da prótese. Já para o sistema de controle semiautônomo a tendência parece ser oposta, indicando que o erro tende a ser maior para as tentativas que tiveram um tempo de execução maior.

A Figura 49 apontou a porcentagem de escorregamentos durante os experimentos. Por meio dos resultados, nota-se que em ambos os sistemas de controle, a porcentagem de escorregamento dos objetos pertencentes a classe de preensão Tripod foi bem maior que para os objetos pertencente a classe Power.

Como descrito na seção 6.1.2.1, os objetos rotulados como Tripod são mais delicados e precisam de uma precisão maior na sua preensão. É importante lembrar que ambos os sistemas desenvolvidos não possuem feedback tátil. Além disso, a prótese não faz parte do corpo do voluntário, evitando que ele tenha a propriocepção do membro biônico. Esses fatores dificultam a capacidade de um indivíduo realizar tarefas que requerem uma alta precisão, como é o caso das tarefas de preensão de objetos rotulados com Tripod.

Cada voluntário foi requisitado a preencher o formulário NASA-TLX após concluir as 30 repetições da tarefa de preensão. Esse questionário é uma ferramenta muito importante para analisar a carga de trabalho de um voluntário para uma determinada tarefa. No caso dos experimentos desse trabalho, o questionário se encaixa muito bem, já que um dos objetivos propostos foi desenvolver um sistema de controle que diminua a carga cognitiva do usuário durante uma tarefa de preensão.

Todos os 6 fatores existentes no questionário têm sua importância. Entretanto, os fatores de carga mental e performance possuem maior relevância para o sistema de controle proposto, já que o principal objetivo proposto nesse estudo foi desenvolver um sistema de controle que permita uma alta performance, e ao mesmo tempo, diminua a carga cognitiva do usuário durante uma tarefa de preensão.

A partir dos resultados apresentados na Figura 50, observa-se que o fator de carga mental teve na média 21,6 pontos a menos para os voluntários que usaram o sistema de controle

semiautônomo. Já para o fator de performance, nota-se que ele teve uma média de 29,6 pontos a mais para os voluntários que usaram o sistema de controle semiautônomo. De acordo com esses dados, nota-se que o sistema de controle semiautônomo tende a reduzir a carga cognitiva do voluntário, mesmo mantendo uma alta performance.

Apesar de não serem o foco desse trabalho, os fatores de esforço e carga física também chamaram muito a atenção. Isso porque na média, suas pontuações foram maiores para os voluntários que usaram o sistema de controle semiautônomo. A partir disso, existem dois motivos que podem ter contribuído para esses fatores serem maiores para o sistema semiautônomo, sendo eles o peso de cada sistema e o tempo total de execução das tarefas de preensão.

Primeiramente, observa-se que o peso do sistema de controle semiautônomo possui 126 gramas a mais que o sistema de controle manual. Esse peso extra é referente à câmera mais o seu suporte, usado para encaixá-la no bypass. A princípio, esse peso extra não é um problema. Entretanto, existem dois fatores que o tornam importantes na hora de analisar os resultados. Como o bypass foi adaptado para encaixar o sistema rente ao braço do voluntário, nota-se que a articulação do cotovelo é responsável por sustentar todo o peso do sistema. Assim, surge um torque, que é diretamente proporcional ao tamanho do braço de alavanca, isto é, o braço do voluntário. Além disso, é importante ressaltar que foram realizadas 30 repetições por voluntário, ocasionando em uma fadiga natural da musculatura do voluntário à medida que o experimento avançou.

O tempo total de execução das tarefas de preensão é outro ponto que contribui em uma maior pontuação para os fatores de esforço e carga cognitiva. Apesar de cada tarefa de preensão usando o sistema de controle semiautônomo ter um tempo médio consideravelmente menor, o seu tempo total de execução das tarefas de preensão foi maior. Esse aumento no tempo total é devido a etapa extra que o experimento usando o sistema de controle semiautônomo possuiu, referente à análise do sistema para objetos não usados no treinamento do seu modelo de classificação.

Além desses dois fatores apresentados acima, é importante lembrar que o questionário NASA-TLX é feito em duas etapas, sendo que na primeira etapa, os fatores são colocados uns contra os outros, de modo que o voluntário precisa indicar qual dos fatores foi o mais impactante. Dessa forma, mesmo não recebendo uma classificação muito ruim na segunda etapa do questionário, alguns fatores se sobressaem pelo fato de terem sido selecionados mais vezes na etapa de um contra um.

### 8.3 COMPARAÇÕES COM O ESTADO DA ARTE

Os sistemas de controle de próteses de membros superiores pertencentes no estado da arte são baseados em CD ou RP. Como foi apresentado no capítulo 3, ambas as abordagens de controle possuem pontos positivos e negativos. Sistemas de controle baseados em CD são uma boa alternativa para poucos DoF. Todavia, à medida que o número de DoF aumenta, a complexidade do controle também aumenta. Conseqüentemente, o controle passa a exigir mais esforço cognitivo do usuário, de modo que o controle se torna pouco natural e difícil de ser executado.

Por outro lado, sistemas de controle de próteses baseados em RP tendem a tornar o controle mais natural, pois se baseia na extração de características dos músculos remanescentes do usuário para controlar a prótese. Assim, o usuário é capaz de usar os mesmos padrões de ativação que eram usados no membro biológico para controlar o membro biônico. Entretanto, desafios como a mudança constante dos padrões do sinal de sEMG de acordo com a posição do membro e deslizamento dos eletrodos fazem com que esse sistema seja pouco robusto durante um uso prolongado.

A ideia de automatizar etapas do controle de próteses surgiu com o intuito de combater algumas dessas limitações existentes das abordagens de controle atuais. O uso de sensores para controlar a força e câmeras para definir o tipo de preensão ideal são estratégias existentes na literatura e possuem um grande potencial (Došen et al., 2010; Došen & Popović, 2011; Ghazaei et al., 2017; Hays et al., 2019; Markovic et al., 2014; Nakagawa-Silva et al., 2019).

Nesse trabalho exploramos o uso de uma câmera baseada em eventos para automatizar as etapas de escolha do tipo de preensão e correção da angulação do punho de acordo com o formato e ângulo do objeto, respectivamente. Além disso, o sistema desenvolvido é capaz de segmentar o objeto almejado de outros objetos próximos, tornando o sistema robusto para cenas complexas.

Ghazaei et al desenvolveu uma estratégia similar, usando um modelo de CNN para classificar em 4 diferentes tipos de preensão, sendo elas pinça, tripod, palmar com o punho neutro e palmar com o punho pronado (Ghazaei et al., 2017). Todavia, a angulação do punho possuía apenas duas posições predefinidas, dependendo se era palmar com o punho neutro ou pronado.

Em contrapartida, o sistema proposto no presente estudo apresentou um ajuste gradual da angulação do punho, dependendo do ângulo do objeto. Assim, tem-se que ao combinar dois

tipos de preensão com a rotação gradual do punho tem-se um sistema bastante complexo, que pode ter inúmeras combinações possíveis para se adaptar à posição espacial do objeto almejado.

Outra diferença é o tipo de câmera usado para extrair informações do objeto almejado. Como é de se esperar, a maioria dos trabalhos usam câmeras convencionais baseadas em quadros para tal tarefa. Contudo, o aparecimento das câmeras neuromórficas baseadas em eventos expandiu os horizontes para o campo de visão computacional, de modo que elas podem ser uma ferramenta muito útil dependendo da aplicação em questão.

Como seu nome diz, as câmeras neuromórficas se baseiam na morfologia do sistema neural, fazendo como que a aquisição de dados seja assíncrona à medida que os seus sensores de luminosidade se despolarizam. Tais características fazem com que seu consumo energético seja menor em relação as câmeras convencionais, tornando seu uso apropriado para sistemas embarcados, como é o caso de sistemas de controle de próteses. Além disso, o fato dessas câmeras fornecerem apenas as informações das bordas do objeto foi outro ponto crucial na sua escolha para esse trabalho, pois reduziu significativamente a quantidade de processamentos necessários para extrair informações do formato do objeto.

O trabalho desenvolvido por Hays et al foi o primeiro trabalho a usar uma câmera baseada em eventos para obter informações do objeto almejado e, conseqüentemente, definir o seu tipo de preensão ideal (Hays et al., 2019). Contudo, o conjunto de objetos se restringiu a apenas 3 formatos geométricos e pouco usuais em tarefas do dia a dia. Por outro lado, o sistema de controle proposto no presente estudo abrangeu uma maior variedade de objetos comuns em atividades do dia a dia.

O campo de visão neuromórfica tem crescido consideravelmente nos últimos anos. Tal crescimento é resultante de grandes empresas de tecnologia começarem a lançar modelos de câmeras neuromórficas comercialmente, como é o caso a Samsung que lançou um modelo de sensor baseado em eventos em 2017 (Son et al., 2017). Embora seja um crescimento importante, elas estão longe de alcançar a popularidade das câmeras convencionais. Conseqüentemente, existem poucos conjuntos de dados obtidos por câmeras neuromórficas existentes na literatura.

Essa falta de bons conjuntos de dados impactou diretamente o presente estudo, que por conta disso teve como objetivo base construir um novo conjunto de dados obtidos por meio da câmera DVS128. Esse conjunto de dados foi crucial para tornar o modelo de classificação robusto para diversas formas de objetos, indicando se o tipo de preensão ideal para objeto é mais delicado, precisando de poucos dedos (Tripod) ou mais forte, necessitando usar todos os dedos da mão (Power).

Conseguir separar objetos por meio de visão computacional não é uma tarefa trivial, podendo ser impossível em alguns casos. Em ambientes reais, existem uma grande variedade de objetos em uma única cena, aumentando a dificuldade de conseguir definir o que é o objeto almejado e o que não é. Tal problema reforça a ideia de que sistemas de controle que se utilizam de visão computacional precisam necessariamente serem robustos para cenas complexas.

O os resultados obtidos por meio do questionário NASA-TLX é outro ponto importante para ressaltar. Como dito anteriormente, o principal objetivo por trás da automatização de etapas do controle de próteses é a redução do esforço cognitivo do usuário durante o controle. A partir disso, realizar uma comparação do esforço cognitivo do usuário com outro sistema ajuda a compreender as melhorias e limitações do sistema proposto.

# Capítulo 9

---

## 9. CONCLUSÃO

O presente estudo apresentou uma nova abordagem para o controle de próteses de membros superiores capaz de reduzir a complexidade do controle. Ao final dos experimentos, nota-se que o sistema de controle semiautônomo foi superior na maioria dos aspectos avaliados, demonstrando seu potencial para facilitar o uso das próteses de membros superiores, com redução dos esforços físico e mental para controle das mesmas.

Embora as próteses mioelétricas não sejam um conceito novo, ainda existem diversos aspectos para esses dispositivos se tornarem ideais, tentando reduzir ao máximo a taxa de abandono dos usuários durante o período de adaptação. Dessa forma, esse trabalho apresentou uma forma alternativa de tornar as próteses mioelétricas de membros superiores mais fáceis de serem controladas por meio da automatização algumas das etapas do controle usando informações do ambiente capturadas por uma câmera neuromórfica.

Apesar de ser uma tecnologia consideravelmente nova, as câmeras neuromórficas vem apresentando excelentes resultados quando o intuito é reduzir o fluxo de informações a serem processadas, focando apenas naquelas que são realmente importantes para a aplicação. Ao longo desse trabalho, foi demonstrado diversas vantagens desse tipo de câmera e como elas vão de encontro com os objetivos do trabalho proposto, tornando as câmeras neuromórficas uma excelente alternativa para o desenvolvimento de sistemas de controle de próteses semiautônomos.

Ao final dos experimentos foi notável a diferença dos sistemas apresentados, de modo que o sistema semiautônomo proposto foi superior em diversos aspectos, tais como tempo de execução e erro no ajuste do ângulo. Entretanto, como esses sistemas foram desenvolvidos para que usuários com amputação transradial possam usá-lo, é de extrema importância dar uma atenção especial para a análise subjetiva do usuário. Dessa forma, o questionário NASA-TLX foi indispensável nesse trabalho, de modo que mesmo em suas subjetividades, foi possível constatar que no geral o sistema proposto foi superior em fatores muito relevantes como Performance, Frustração e Mental.

Por fim, o presente estudo colaborou com o desenvolvimento de um novo conjunto de dados formado por objetos pertencentes ao nosso dia a dia gravados por uma câmera

neuromórfica. Tal conjunto de dados pode contribuir para futuras pesquisas correlacionadas, as quais podem utilizá-lo tanto para comparações com o sistema semiautônomo desenvolvido nesse trabalho quanto para outras aplicações relacionadas com visão neuromórfica.

## 9.1 TRABALHOS FUTUROS

Embora os objetivos principais tenham sido cumpridos, existem pontos que podem ser melhorados, possibilitando o sistema ter uma performance ainda melhor. O principal passo é executar esse mesmo experimento com mais voluntários, principalmente com voluntários que possuam amputação transradial, de modo que seja possível averiguar as vantagens do sistema proposto com pessoas que de fato irão se beneficiar com o desenvolvimento do mesmo.

Outro ponto é que o sistema de controle proposto possui dois tipos de preensão. Entretanto, esse número pode ser maior, acrescentando outros tipos de preensão, por exemplo a pinça, que usa apenas os dedos polegar e indicador e é usado para tarefas de extrema precisão.

Além disso, pode ser uma boa estratégia adicionar os DoF de pronação/supinação e adução/abdução do punho, possibilitando movimentos em todas as direções. Por último, tem-se que acrescentar um controle de força de preensão junto com feedback tátil para o usuário no sistema de controle proposto pode aumentar bastante sua usabilidade.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Alter, R. (1966). *Bioelectric control of prostheses*.
- Atta-Fosu, T., Guo, W., Jeter, D., Mizutani, C. M., Stopczynski, N., & Sousa-Neves, R. (2016). 3D clumped cell segmentation using curvature based seeded watershed. *Journal of Imaging*. <https://doi.org/10.3390/jimaging2040031>
- Basmajian, J. V. (1967). *Muscles Alive: Their Functions Revealed by Electromyography*. Queen's University.
- Bastos-Filho, T., Kumar, D., & Arjunan, S. P. (2014). Devices for Mobility and Manipulation for People with Reduced Abilities. In *Devices for Mobility and Manipulation for People with Reduced Abilities*. <https://doi.org/10.1201/b16870>
- BATTYE, C. K., NIGHTINGALE, A., & WHILLIS, J. (1955). The use of myo-electric currents in the operation of prostheses. *The Journal of Bone and Joint Surgery. British Volume*. <https://doi.org/10.1302/0301-620x.37b3.506>
- Beaulieu, R. J., Masters, M. R., Betthausen, J., Smith, R. J., Kaliki, R., Thakor, N. V., & Soares, A. B. (2017). Multi-Position Training Improves Robustness of Pattern Recognition and Reduces Limb-Position Effect in Prosthetic Control. *Journal of Prosthetics and Orthotics*. <https://doi.org/10.1097/JPO.000000000000121>
- Beucher, S., & Lantuejoul, C. (1979). Use of Watersheds in Contour Detection. In *International Workshop on Image Processing: Real-time Edge and Motion Detection/Estimation*.
- Biddiss, E., & Chau, T. (2007). Upper limb prosthesis use and abandonment: A survey of the last 25 years. *Prosthetics and Orthotics International*, 31(3), 236–257. <https://doi.org/10.1080/03093640600994581>
- Billock, J. N. (1986). Upper Limb Prosthetic Terminal Devices: Hands Versus Hooks. *Clinical Prosthetics and Orthotics*, 10(2), 57–65.
- Brandli, C., Berner, R., Yang, M., Liu, S. C., & Delbruck, T. (2014). A 240 × 180 130 dB 3 μs latency global shutter spatiotemporal vision sensor. *IEEE Journal of Solid-State Circuits*. <https://doi.org/10.1109/JSSC.2014.2342715>
- Cavalcanti, M. A., & Vieira, T. M. (2011). Surface electromyography: Why, when and how to use it. *Rev Andal Med Deporte*, 17–28. <https://doi.org/10.36104/amc.2018.1400>
- Coapt. (2020). *Coapt, LLC*. <https://coaptengineering.com/>
- Conradt, J., Cook, M., Berner, R., Lichtsteiner, P., Douglas, R. J., & Delbruck, T. (2009). A pencil balancing robot using a pair of AER dynamic vision sensors. *Proceedings - IEEE*

*International Symposium on Circuits and Systems.*

<https://doi.org/10.1109/ISCAS.2009.5117867>

- Cordella, F., Ciancio, A. L., Sacchetti, R., Davalli, A., Cutti, A. G., Guglielmelli, E., & Zollo, L. (2016). Literature review on needs of upper limb prosthesis users. *Frontiers in Neuroscience*, 10(MAY), 1–14. <https://doi.org/10.3389/fnins.2016.00209>
- Cram, J. R. (2003). The history of surface electromyography. *Applied Psychophysiology Biofeedback*, 28(2), 81–91. <https://doi.org/10.1023/A:1023802407132>
- Criswell, E. (2011). Cram's Introduction to Surface Electromyography, 2nd Edition. *Medicine & Science in Sports & Exercise*. <https://doi.org/10.1249/01.mss.0000399576.80711.7d>
- DATASUS. (2020). *Number of upper-limb amputations registred on SUS*. <http://tabnet.datasus.gov.br/cgi/tabcgi.exe?sih/cnv/qiuf.def>
- Davis, J. F. (1959). Manual of Surface Electromyography. *Aerospace Medical Laboratory, Wright Air Development Center, Air Research and Development Command, U.S. Air Force*, 59-184.
- Delbruck, T., & Lang, M. (2013). Robotic goalie with 3 ms reaction time at 4% CPU load using event-based dynamic vision sensor. *Frontiers in Neuroscience*, 7(7 NOV), 1–7. <https://doi.org/10.3389/fnins.2013.00223>
- Digabel, H., & Lantuéjoul, C. (1978). Iterative algorithms. *Proc. 2nd European Symp. Quantitative Analysis of Microstructures in Material Science, Biology and Medicine*, 19(7), 8.
- Došen, S., Cipriani, C., Kostić, M., Controzzi, M., Carrozza, M. C., & Popović, D. B. (2010). Cognitive vision system for control of dexterous prosthetic hands: Experimental evaluation. *Journal of NeuroEngineering and Rehabilitation*, 7(1), 1–14. <https://doi.org/10.1186/1743-0003-7-42>
- Došen, S., & Popović, D. B. (2011). Transradial Prosthesis: Artificial Vision for Control of Prehension. *Artificial Organs*, 35(1), 37–48. <https://doi.org/10.1111/j.1525-1594.2010.01040.x>
- Esquenazi, A. (2004). Amputation rehabilitation and prosthetic restoration. From surgery to community reintegration. In *Disability and Rehabilitation*. <https://doi.org/10.1080/09638280410001708850>
- Faugeras, O. (1993). *Three-dimensional computer vision: a geometric viewpoint*. MIT press.
- Gaine, W. J., Smart, C., & Bransby-Zachary, M. (1997). Upper limb traumatic amputees: Review of prosthetic use. *Journal of Hand Surgery: European Volume*, 22(1), 73–76. [https://doi.org/10.1016/S0266-7681\(97\)80023-X](https://doi.org/10.1016/S0266-7681(97)80023-X)

- Gallego, G., Delbruck, T., Orchard, G. M., Bartolozzi, C., Taba, B., Censi, A., Leutenegger, S., Davison, A., Conratt, J., Daniilidis, K., & Scaramuzza, D. (2020). Event-based Vision: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1. <https://doi.org/10.1109/TPAMI.2020.3008413>
- Ghazaei, G., Alameer, A., Degenaar, P., Morgan, G., & Nazarpour, K. (2017). Deep learning-based artificial vision for grasp classification in myoelectric hands. *Journal of Neural Engineering*, 14(3). <https://doi.org/10.1088/1741-2552/aa6802>
- Hall, John E., & Guyton, A. C. (2017). Guyton and Hall: Textbook of Medical Physiology. *Surgical Neurology International*. [https://doi.org/10.4103/sni.sni\\_327\\_17](https://doi.org/10.4103/sni.sni_327_17)
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology* (Vol. 52, pp. 139–183). Elsevier.
- Hays, M., Osborn, L., Ghosh, R., Iskarous, M., Hunt, C., & Thakor, N. V. (2019). Neuromorphic vision and tactile fusion for upper limb prosthesis control. *International IEEE/EMBS Conference on Neural Engineering, NER, 2019-March*, 981–984. <https://doi.org/10.1109/NER.2019.8716890>
- Inivation. (n.d.). Retrieved February 16, 2021, from <https://inivation.github.io/inivation-docs/>
- Jolliffe, I. T. (1986). *Principal Components in Regression Analysis*. [https://doi.org/10.1007/978-1-4757-1904-8\\_8](https://doi.org/10.1007/978-1-4757-1904-8_8)
- KOBRINSKII, A. E. (1960). *Use of Biocurrents for Control Purposes*.
- Kobrinisky, A. (1960). Bioelectric control systems. *Radio USSR (In Russian)*, 11, 37–39.
- Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. In *Nature*. <https://doi.org/10.1038/nature14539>
- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*. <https://doi.org/10.1109/5.726791>
- Lichtsteiner, P., Posch, C., & Delbruck, T. (2008). A  $128 \times 128$  120 dB 15  $\mu$ s latency asynchronous temporal contrast vision sensor. *IEEE Journal of Solid-State Circuits*, 43(2), 566–576. <https://doi.org/10.1109/JSSC.2007.914337>
- Lindner, H. Y. N., Nätterlund, B. S., & Hermansson, L. M. N. (2010). Upper limb prosthetic outcome measures: Review and content comparison based on International Classification of Functioning, Disability and Health. In *Prosthetics and Orthotics International*. <https://doi.org/10.3109/03093641003776976>

- Liu, S. C., Delbruck, T., Indiveri, G., Whatley, A., & Douglas, R. (2014). Event-based neuromorphic systems. In *Event-Based Neuromorphic Systems*. <https://doi.org/10.1002/9781118927601>
- Mahowald, M., & Mead, C. (1991). The Silicon Retina. *Scientific American*, 264, 76–83. <https://doi.org/10.1038/scientificamerican0591-76>
- Markovic, M., Dosen, S., Cipriani, C., Popovic, D., & Farina, D. (2014). Stereovision and augmented reality for closed-loop control of grasping in hand prostheses. *Journal of Neural Engineering*, 11(4). <https://doi.org/10.1088/1741-2560/11/4/046001>
- Markovic, M., Dosen, S., Popovic, D., Graimann, B., & Farina, D. (2015). Sensor fusion and computer vision for context-aware control of a multi degree-of-freedom prosthesis. *Journal of Neural Engineering*, 12(6). <https://doi.org/10.1088/1741-2560/12/6/066022>
- Meier, R. H. (2004). History of arm amputation, prosthetic restoration, and arm amputation rehabilitation. *Functional Restoration of Adults and Children with Upper Extremity Amputation*, 1–7.
- Murray, C. D. (2005). The Social Meanings of Prosthesis Use. *Journal of Health Psychology*. <https://doi.org/10.1177/1359105305051431>
- Nakagawa-Silva, A., Thakor, N. V., Cabibihan, J.-J., & Soares, A. B. (2019). A Bio-Inspired Slip Detection and Reflex-Like Suppression Method for Robotic Manipulators. *IEEE Sensors Journal*, 19(24), 12443–12453. <https://doi.org/10.1109/JSEN.2019.2939506>
- Nerlich, A. G., Zink, A., Szeimies, U., & Hagedorn, H. G. (2000). Ancient Egyptian prosthesis of the big toe. In *Lancet*. [https://doi.org/10.1016/S0140-6736\(00\)03507-8](https://doi.org/10.1016/S0140-6736(00)03507-8)
- Osborn, L., Kaliki, R. R., Soares, A. B., & Thakor, N. V. (2016). Neuromimetic Event-Based Detection for Closed-Loop Tactile Feedback Control of Upper Limb Prostheses. *IEEE Transactions on Haptics*. <https://doi.org/10.1109/TOH.2016.2564965>
- Össur Corporate. (2020). *Össur*. <https://corporate.ossur.com/corporate/>
- Ottobock. (2020). *Ottobock*. <https://www.ottobock.com/en/>
- Ottobock. (2021). *Myo Plus*. <https://www.ottobock-export.com/en/prosthetics/upper-limb/solution-overview/myo-plus-mustererkennung/index.html>
- Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*. <https://doi.org/10.1080/14786440109462720>
- Peixoto, A. M., Zimpel, S. A., Oliveira, A. C. A. de, Monteiro, R. L. S., & Carneiro, T. K. G. (2017). Prevalência de amputações de membros superiores e inferiores no estado de

- Alagoas atendidos pelo SUS entre 2008 e 2015. *Fisioterapia e Pesquisa*. <https://doi.org/10.1590/1809-2950/17029524042017>
- Pham, D. L., Xu, C., & Prince, J. L. (2000). Current methods in medical image segmentation. *Annual Review of Biomedical Engineering*. <https://doi.org/10.1146/annurev.bioeng.2.1.315>
- Pine, K. R., Sloan, B. H., & Jacobs, R. J. (2015). Clinical ocular prosthetics. In *Clinical Ocular Prosthetics*. <https://doi.org/10.1007/978-3-319-19057-0>
- Posch, C., Matolin, D., & Wohlgenannt, R. (2011). A QVGA 143 dB dynamic range frame-free PWM image sensor with lossless pixel-level video compression and time-domain CDS. *IEEE Journal of Solid-State Circuits*. <https://doi.org/10.1109/JSSC.2010.2085952>
- Preim, B., & Botha, C. (2014). Image Analysis for Medical Visualization. In *Visual Computing for Medicine*. <https://doi.org/10.1016/b978-0-12-415873-3.00004-3>
- Putti, V. (2005). Historical Prostheses. *The Journal of Hand Surgery: British & European Volume*, 30(3), 310–325. <https://doi.org/https://doi.org/10.1016/j.jhsb.2005.01.001>
- Reaz, M. B. I., Hussain, M. S., & Mohd-Yasin, F. (2006). Techniques of EMG signal analysis: Detection, processing, classification and applications. *Biological Procedures Online*. <https://doi.org/10.1251/bpo115>
- Rebecq, H., Ranftl, R., Koltun, V., & Scaramuzza, D. (2019). High speed and high dynamic range video with an event camera. In *arXiv*. <https://doi.org/10.1109/tpami.2019.2963386>
- Resnik, L., Huang, H. H., Winslow, A., Crouch, D. L., Zhang, F., & Wolk, N. (2018). Evaluation of EMG pattern recognition for upper limb prosthesis control: A case study in comparison with direct myoelectric control. *Journal of NeuroEngineering and Rehabilitation*, 15(1). <https://doi.org/10.1186/s12984-018-0361-3>
- Resnik, L., Meucci, M. R., Lieberman-Klinger, S., Fantini, C., Kelty, D. L., Disla, R., & Sasson, N. (2012). Advanced upper limb prosthetic devices: Implications for upper limb prosthetic rehabilitation. In *Archives of Physical Medicine and Rehabilitation*. <https://doi.org/10.1016/j.apmr.2011.11.010>
- ROS Components. (2020). *Shadow Dexterous*. <https://www.roscomponents.com/en/robotic-hands/117-shadow-dexterous-robotic-hand.html>
- Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*. <https://doi.org/10.1037/h0042519>
- Son, B., Suh, Y., Kim, S., Jung, H., Kim, J. S., Shin, C., Park, K., Lee, K., Park, J., Woo, J., Roh, Y., Lee, H., Wang, Y., Ovsiannikov, I., & Ryu, H. (2017). A 640×480 dynamic vision sensor with a 9µm pixel and 300Meps address-event representation. *Digest of Technical*

- Papers - IEEE International Solid-State Circuits Conference.*  
<https://doi.org/10.1109/ISSCC.2017.7870263>
- Szeliski, R. (2010). *Computer vision: algorithms and applications*. Springer Science & Business Media.
- ten Kate, J., Smit, G., & Breedveld, P. (2017). 3D-printed upper limb prostheses: a review. In *Disability and Rehabilitation: Assistive Technology*.  
<https://doi.org/10.1080/17483107.2016.1253117>
- Vincent, L., Vincent, L., & Soille, P. (1991). Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. <https://doi.org/10.1109/34.87344>
- Waggoner, J., Zhou, Y., Simmons, J., De Graef, M., & Wang, S. (2013). 3D materials image segmentation by 2D propagation: A graph-cut approach considering homomorphism. *IEEE Transactions on Image Processing*. <https://doi.org/10.1109/TIP.2013.2284071>
- Wang, D. (1998). Unsupervised video segmentation based on watersheds and temporal tracking. *IEEE Transactions on Circuits and Systems for Video Technology*.  
<https://doi.org/10.1109/76.718501>
- Wheaton, L. A. (2017). Neurorehabilitation in upper limb amputation: understanding how neurophysiological changes can affect functional rehabilitation. In *Journal of NeuroEngineering and Rehabilitation*. <https://doi.org/10.1186/s12984-017-0256-8>
- White, M. M., Zhang, W., Winslow, A. T., Zahabi, M., Zhang, F., Huang, H., & Kaber, D. B. (2017). Usability Comparison of Conventional Direct Control Versus Pattern Recognition Control of Transradial Prostheses. *IEEE Transactions on Human-Machine Systems*, 47(6), 1146–1157. <https://doi.org/10.1109/THMS.2017.2759762>
- Windrich, M., Grimmer, M., Christ, O., Rinderknecht, S., & Beckerle, P. (2016). Active lower limb prosthetics: A systematic review of design issues and solutions. In *BioMedical Engineering Online*. <https://doi.org/10.1186/s12938-016-0284-9>
- Yang, X., Li, H., & Zhou, X. (2006). Nuclei segmentation using marker-controlled watershed, tracking using mean-shift, and Kalman filter in time-lapse microscopy. *IEEE Transactions on Circuits and Systems I: Regular Papers*. <https://doi.org/10.1109/TCSI.2006.884469>
- Yerrabelli, R. (2018). *PHAM: Prosthetic hand assessment measure PHAM: PROSHETIC HAND ASSESSMENT MEASURE* (Issue August 2017).
- Yi, W., & Marshall, S. (2000). Principal component analysis in application to object orientation. *Geo-Spatial Information Science*, 3(3), 76–78.

Zuo, K. J., & Olson, J. L. (2014). The evolution of functional hand replacement: From iron prostheses to hand transplantation. In *Canadian Journal of Plastic Surgery*. <https://doi.org/10.1177/229255031402200111>

## APÊNDICE A – TABELA DE OBJETOS GRAVADOS

Tabela A1 – Lista de objetos contidos no conjunto de dados gerado, o qual possui 15 objetos referentes ao tipo de preensão Tripod e 15 objetos referentes ao tipo de preensão Power.

	<b>Preensão Tripod</b>	<b>Preensão Power</b>	
<b>Grupo 1</b>	01	Barbeador	Bola
	02	Bateria	Caixa
	03	Caneta	Calculadora
	04	Chave	Copo
	05	Chave Allen	Câmera
	06	Faca	Estojo de óculos
	07	Isqueiro	Garrafa
	08	Lapiseira	HD externo
	09	Pen drive	Lata de bebida
	10	Pincel	Pote
<b>Grupo 2</b>	11	Chave de fenda	Caneca
	12	Chave fixa	Carteira
	13	Colher	Laranja
	14	Escova de dentes	Lâmpada
	15	Parafuso	Celular

## APÊNDICE B – QUESTIONÁRIO NASA TSX ADAPTADO

O questionário abaixo é uma adaptação do NASA-TLX, que é amplamente utilizado para avaliar a carga de trabalho de um indivíduo durante uma determinada tarefa. Esse questionário foi utilizado no presente estudo para avaliar o sistema proposto por meio de uma análise subjetiva dos voluntários.

Tabela B1 – Detalhamento de cada um dos fatores de carga de trabalho pertencentes ao questionário NASA-TLX.

Fatores de carga de trabalho	Definição
<p><b>Nível de demanda mental</b> (baixa/alta)</p>	<p>Quanta atividade mental e perceptiva foi necessária (por exemplo, pensar, decidir, calcular, lembrar, olhar, pesquisar etc.)? A tarefa era fácil ou exigente, simples ou complexa, perdoadora ou exigente?</p>
<p><b>Nível de demanda física</b> (baixa/alta)</p>	<p>Quanta atividade física foi necessária (por exemplo, empurrar, puxar, girar, controlar, ativar etc.)? A tarefa era fácil ou exigente, lenta ou rápida, lenta ou extenuante, repousante ou laboriosa?</p>
<p><b>Nível de demanda temporal</b> (baixa/alta)</p>	<p>Quanta pressão de tempo você sentiu devido à taxa ou ritmo em que as tarefas ou elementos da tarefa ocorreram? O ritmo era lento e lento ou rápido e frenético?</p>
<p><b>Nível de Performance</b> (boa/ruim)</p>	<p>Quão bem-sucedido você acha que foi ao cumprir os objetivos da tarefa definidos pelo experimentador (ou por você mesmo)? Quão satisfeito você ficou com seu desempenho no cumprimento dessas metas?</p>
<p><b>Nível de Esforço</b> (baixa/alta)</p>	<p>Quão duro você teve que trabalhar (mental e fisicamente) para atingir seu nível de desempenho?</p>
<p><b>Nível de Frustração</b> (baixa/alta)</p>	<p>Quão inseguro, desanimado, irritado, estressado e aborrecido você se sentiu durante a tarefa?</p>

Tabela B2 – Parte 1 do questionário NASA-TLX, colocando sempre dois fatores um contra o outro.

Para cada par, escolha o fator que foi mais impactante para a sua experiência da carga de trabalho na tarefa que você executou recentemente:		
1	<input type="checkbox"/> Demanda temporal	<input type="checkbox"/> Demanda mental
2	<input type="checkbox"/> Performance	<input type="checkbox"/> Demanda mental
3	<input type="checkbox"/> Demanda mental	<input type="checkbox"/> Esforço
4	<input type="checkbox"/> Demanda temporal	<input type="checkbox"/> Esforço
5	<input type="checkbox"/> Demanda Física	<input type="checkbox"/> Performance
6	<input type="checkbox"/> Performance	<input type="checkbox"/> Demanda temporal
7	<input type="checkbox"/> Esforço	<input type="checkbox"/> Demanda Física
8	<input type="checkbox"/> Demanda mental	<input type="checkbox"/> Demanda Física
9	<input type="checkbox"/> Performance	<input type="checkbox"/> Frustração
10	<input type="checkbox"/> Esforço	<input type="checkbox"/> Performance
11	<input type="checkbox"/> Frustração	<input type="checkbox"/> Esforço
12	<input type="checkbox"/> Frustração	<input type="checkbox"/> Demanda mental
13	<input type="checkbox"/> Demanda Física	<input type="checkbox"/> Demanda temporal
14	<input type="checkbox"/> Demanda Física	<input type="checkbox"/> Frustração
15	<input type="checkbox"/> Demanda temporal	<input type="checkbox"/> Frustração

Agora você será apresentado a uma série de escalas de classificação.

Para cada uma das seis escalas, avalie a tarefa que você executou recentemente cruzando a localização da escala que corresponde à sua experiência. Cada linha possui dois pontos finais que descrevem a escala.

Considere suas respostas cuidadosamente ao distinguir entre as diferentes condições de tarefa e considere cada uma individualmente.

Demanda mental – Quão mentalmente exigente foi a tarefa? / Quanta atividade mental e perceptiva você gastou para esta tarefa?



Muito baixa

Muito alta

Demanda Física – Quão fisicamente exigente foi a tarefa? / Quanta atividade física você gastou para esta tarefa?



Muito baixa

Muito alta

Demanda temporal – Quão apressado ou apressado foi o ritmo da tarefa? / Quanta pressão de tempo você sentiu para concluir esta tarefa?



Muito baixa

Muito alta

Performance – Quão bem-sucedido você foi em realizar o que foi solicitado a fazer? / Quão bem-sucedido você acha que teve em cumprir os objetivos da tarefa?



Muito boa

Muito ruim

Esforço – Quão duro você teve que trabalhar para atingir seu nível de desempenho?



Muito baixa

Muito alta

Frustração – Quão inseguro, desanimado, irritado, estressado e aborrecido você estava durante esta tarefa?



Muito baixa

Muito alta

**APÊNDICE C – ARTIGOS PUBLICADOS**

1 – **Lawrence Souza, Eber,** & Barbosa Soares, Alcimar. (2019). Analysis of Electrode Positioning Influence on the Forearm for the Classification of Different Hand/Wrist Movements. Zenodo. <https://zenodo.org/record/3460952>.

2 – Pereira R. da Costa, Ana Clara, **Lawrence S, Eber,** & Barbosa Soares, Alcimar. (2019). Redes Neurais Artificiais na Classificação de Movimentos de Mão por Sinais Eletromiográficos. Zenodo. <https://zenodo.org/record/3461294>.

3 - Roza Gonçalves, Camila, Gustavo Sampaio Fernandes, Ítalo, Torres Cresto, Ana Carolina, Nakagawa Silva, Andrei, **Lawrence Souza Gouveia, Eber,** Ricardo de Jesus Oliveira, Sérgio, & Barbosa Soares, Alcimar. A 3D Joint Position Matching Protocol for Post-Stroke Proprioceptive Assessment. Biomedical Journal of Scientific & Technical Research.

4 – Gouveia, Eduardo, Vasconcelos, Lucas, **Souza Gouveia, Eber Lawrence,** Costa, Vinicius, Nakagawa, Andrei, & Barbosa Soares, Alcimar. An Object Tracking Using a Neuromorphic System Based on Standard RGB Cameras.

5 – Gouveia, Eduardo, **Souza Gouveia, Eber Lawrence,** Costa, Vinicius, Nakagawa, Andrei, & Barbosa Soares, Alcimar. Classification of Objects Using Neuromorphic Camera and Convolutional Neural Networks.

6 – **Souza Gouveia, Eber Lawrence,** Borges Gouveia, Eduardo, Costa, Vinicius, Nakagawa, Andrei, & Barbosa Soares, Alcimar. Neuromorphic Vision-aided Semi-autonomous System for Prosthesis Control.

7 – Gouveia, Eduardo; Tavares, Gustavo; Almada, Leandra; Silva, Andrei; **Gouveia, Eber Lawrence;** Cunha, Marcio; Lamounier, Edgard; Soares, Alcimar. (2021). Object detection using sparse data representation with convolutional neural networks for event-based cameras.

8 – **Souza Gouveia, Eber Lawrence,** Borges Gouveia, Eduardo, & Barbosa Soares, Alcimar. Neuromorphic Visually Aided Control System for Upper Limb Dexterous Prostheses. **(SUBMETIDO)**