

Wérikson Frederiko de Oliveira Alves

**Classificador de Gesto  
utilizando Distância Euclidiana**

Viçosa - MG, Brasil

2022

Wérikson Frederiko de Oliveira Alves

## **Classificador de Gesto utilizando Distância Euclidiana**

Monografia apresentada ao Departamento de Engenharia Elétrica do Centro de Ciências Exatas e Tecnológicas da Universidade Federal de Viçosa, para a obtenção dos créditos da disciplina ELT 402 – Projeto de Engenharia II – e cumprimento do requisito parcial para obtenção do grau de Bacharel em Engenharia Elétrica.

Universidade Federal de Viçosa  
Centro de Ciências Exatas e Tecnológicas  
Departamento de Engenharia Elétrica  
Núcleo de Especialização em Robótica

Orientadora: Ketia Soares Moreira  
Coorientador: Alexandre Santos Brandão

Viçosa - MG, Brasil


2022

Wérikson Frederiko de Oliveira Alves

## Classificador de Gesto utilizando Distância Euclidiana

Monografia apresentada ao Departamento de Engenharia Elétrica do Centro de Ciências Exatas e Tecnológicas da Universidade Federal de Viçosa, para a obtenção dos créditos da disciplina ELT 402 – Projeto de Engenharia II – e cumprimento do requisito parcial para obtenção do grau de Bacharel em Engenharia Elétrica.


Trabalho aprovado. Viçosa - MG, Brasil, 15 de dezembro de 2022:

Documento assinado digitalmente  
 KETIA SOARES MOREIRA  
Data: 20/12/2022 10:25:32-0300  
Verifique em <https://verificador.iti.br>

---

**Ketia Soares Moreira**


Orientadora

Documento assinado digitalmente  
 ALEXANDRE SANTOS BRANDAO  
Data: 22/12/2022 09:04:59-0300  
Verifique em <https://verificador.iti.br>

---

**Alexandre Santos Brandão**


Co-orientador

Documento assinado digitalmente  
 KEVIN BRAATHEN DE CARVALHO  
Data: 21/12/2022 18:48:10-0300  
Verifique em <https://verificador.iti.br>

---

**Kevin Braathen de Carvalho**

Convidado 1

Documento assinado digitalmente  
 VICTOR PELLANDA DARDENGO  
Data: 19/12/2022 17:28:42-0300  
Verifique em <https://verificador.iti.br>

---

**Victor Pellanda Dardengo**

Convidado 2

*A felicidade não está no fim da jornada,  
e sim em cada curva do caminho que percorremos para encontrá-la.*

# Agradecimentos

Primeiramente, eu gostaria de agradecer a deus por ter me permitido chegar onde estou hoje. Foi uma longa jornada, 5 anos de estudos, havendo momentos tranquilos e conturbados, mas no fim consegui me sobressair nos momentos mais difíceis.

Em segundo lugar, gostaria de agradecer a minha família que me apoiou a todo momento nessa jornada, me ajudando no que fosse possível. Gostaria de agradecer principalmente aos meus pais e meus irmãos, por sempre estarem me apoiando durante esse período.

Agradeço aos professores do curso que através dos seus ensinamentos me permitiram chegar onde estou hoje. Em especial, gostaria de agradecer à minha orientadora, Ketia, e ao meu co-orientador, Alexandre, que neste último ano estiveram me auxiliando no desenvolvimento deste trabalho, e me ajudaram durante o curso, contribuindo enormemente para a minha formação.

Agradeço também a todos os meus amigos, colegas de curso, e ao pessoal do NERO, que estiveram comigo, me apoiaram e me ajudaram nos momentos que eu mais precisei durante o curso. Agradeço, em especial, ao Celso e ao Hiago, duas pessoas que conheci no início dessa jornada, e que me ajudaram em vários momentos, contribuindo para que eu chegasse onde estou hoje.

*“Às vezes, não vemos o final do caminho no começo da jornada. Só nos resta segui-lo.”(Danielle Steel)*

# Resumo

Aplicações baseadas em reconhecimento de ações estão cada vez mais comuns no dia a dia. Isto porque possibilita ao usuário acionar equipamentos sem a necessidade de tocá-los e, além disso, pode ser realizado remotamente. Neste contexto, o objetivo deste trabalho é a criação de uma ferramenta de classificação de ações, pré-definidas, para o usuário em tempo real.

Esta ferramenta adquire os dados das ações dos usuários e os categoriza em um conjunto de 14 classes. Os marcadores registrados pelo Sistema de Rastreamento *OptiTrack* foram representados por duas pulseiras e um crachá, a fim de tornar o processo de classificação mais fácil e dinâmico.

Estas informações são analisadas por um programa desenvolvido no software MatLab, no qual indica, ao final, a ação classificada. Primeiramente, para a análise destas informações, foi necessário realizar um pré-processamento dos dados de entrada, composto de uma centralização espacial dos dados rastreados dos marcadores.

Em seguida, para reduzir a dimensão dos dados, foi aplicada a técnica de Análise das Componentes Principais (PCA) sendo extraídas às três principais componentes do movimento mais relevantes. Por fim, estas informações foram utilizadas como entradas do classificador.

Por sua simplicidade, utilizou-se o classificador por KNN, com  $k$  igual a 1, a qual se assemelha à Distância Euclidiana, o qual consiste na comparação da série de informações do conjunto de dados reduzidos da ação a ser classificada com as ações armazenadas na base de dados. Ao percorrer a base de dados, aquela em que se obter a menor diferença irá representar a classe da ação a ser classificada. A validação se deu de forma *offline* através do treinamento da ferramenta. Os resultados da matriz de confusão consolidaram a proposta. Ao final, conclui-se que o classificador obteve taxas de acerto superiores a 90% para 11 das 14 classes, e superior a 74% para as demais classes.

O principal benefício deste trabalho é formalizar um classificador simplista, de baixo custo computacional, que requer uma base de dados mínima e satisfatória para generalização das ações.

**Palavras-chave:** Reconhecimento de Ações; Robótica Interativa; PCA; *OptiTrack*; KNN; Distância Euclidiana.

# Abstract

Applications based on action recognition are becoming increasingly common in everyday life. This is because it allows the user to trigger equipments without the need to touch them and, moreover, it can be done remotely. In this context, the goal of this work is to create a tool to classify predefined actions for the user in real time.

This tool acquires the data from the users' actions and categorizes them into a set of 14 classes. The markers recorded by the *OptiTrack* Tracking System were represented by two wristbands and a badge in order to make the classification process easier and more dynamic.

This information is analyzed by a program developed in the MatLab software, which indicates, at the end, the classified action. First, for the analysis of this information, it was necessary to perform a pre-processing of the input data, composed of a spatial centering of the traced data of the markers.

Then, to reduce the dimensionality of the data, the Principal Component Analysis (PCA) technique was applied, and the three most relevant motion components were extracted. Finally, this information was used as inputs to the classifier.

For its simplicity, the KNN classifier was used, with  $k$  equal to 1, which is similar to the Euclidean Distance, and consists in comparing the series of information of the reduced data set of the action to be classified with the actions stored in the database. By going through the database, the one with the smallest difference will represent the class of the action to be classified. The validation took place *offline* by training the tool. The results of the confusion matrix consolidated the proposal. At the end, we conclude that the classifier obtained hit rates higher than 90

The main benefit of this work is to formalize a simplistic classifier, of low computational cost, that requires a minimal and satisfactory database for generalization of actions.

**Key-words:** Action Recognition; Interactive Robotics; PCA; *OptiTrack*; KNN; Euclidean Distance.



# Lista de ilustrações

Figura 1 – Componentes necessários para a realização do trabalho. . . . .	21
Figura 2 – Posicionamento das câmeras no ambiente de simulação. . . . .	22
Figura 3 – Pulseiras, Crachá e Corpos rígidos utilizados para realizar o rastreamento dos dados. . . . .	23
Figura 4 – Localização dos marcadores no usuário. . . . .	24
Figura 5 – Classes de gestos no conjunto de dados. . . . .	25
Figura 6 – Fluxograma com as etapas e respostas do classificador proposto. . . . .	29
Figura 7 – Painel de captura do software utilizado para as gravações das amostras. . . . .	31
Figura 8 – Exemplo de uma amostra coletada durante o treino. . . . .	32

# Lista de tabelas

Tabela 1 – Custo médio para realização do trabalho. . . . .	22
Tabela 2 – Matriz de confusão com o resultado da simulação 1, 88,57% de acerto.	33
Tabela 3 – Matriz de confusão com o resultado agregado após a simulação 1, 88,57% de acerto. . . . .	33
Tabela 4 – Matriz de confusão com o resultado da simulação 2, 87,14% de acerto.	34
Tabela 5 – Matriz de confusão com o resultado agregado após a simulação 2, 87,86% de acerto. . . . .	35
Tabela 6 – Matriz de confusão com o resultado da simulação 3, 95,71% de acerto.	36
Tabela 7 – Matriz de confusão com o resultado agregado após a simulação 3, 90,48% de acerto. . . . .	36
Tabela 8 – Matriz de confusão com o resultado da simulação 4, 90,00% de acerto.	37
Tabela 9 – Matriz de confusão com o resultado agregado após a simulação 4, 90,36% de acerto. . . . .	38
Tabela 10 – Matriz de confusão com o resultado da simulação 5, 97,14% de acerto.	39
Tabela 11 – Matriz de confusão com o resultado agregado após a simulação 5, 91,71% de acerto. . . . .	39
Tabela 12 – Matriz de confusão com o resultado da simulação 6, 92,86% de acerto.	40
Tabela 13 – Matriz de confusão com o resultado agregado após a simulação 6, 91,90% de acerto. . . . .	41
Tabela 14 – Matriz de confusão com o resultado da simulação 7, 97,14% de acerto.	42
Tabela 15 – Matriz de confusão com o resultado agregado após a simulação 7, 92,65% de acerto. . . . .	42
Tabela 16 – Matriz de confusão com o resultado da simulação 8, 90,00% de acerto.	43
Tabela 17 – Matriz de confusão com o resultado agregado após a simulação 8, 92,32% de acerto. . . . .	44
Tabela 18 – Matriz de confusão com o resultado da simulação 6, 91,43% de acerto.	45
Tabela 19 – Matriz de confusão com o resultado agregado após a simulação 6, 92,22% de acerto. . . . .	45
Tabela 20 – Matriz de confusão com o resultado da simulação 10, 95,71% de acerto.	46
Tabela 21 – Matriz de confusão com o resultado agregado após a simulação 10, 91,90% de acerto. . . . .	47
Tabela 22 – Resultado geral da validação, 92,57% de acerto . . . . .	48
Tabela 23 – Resumo dos resultados de cada simulação executada no trabalho. . . .	48

# Lista de abreviaturas e siglas

IHM	Interação Homem-Máquina
HRC	Human Robot Collaboration
PCA	Análise das Componentes Principais
LDA	Análise Discriminante Linear
RNA	Rede Neural Artificial
RGB-D	Red, Green, Blue - Depth
PDI	Processamento de Imagem
SRO	Sistema de Rastreamento OptiTrack
SVC	Sistema de Visão Computacional
KNN	K-Nearest Neighbors
SVM	Support Vector Machine

# Lista de símbolos

$T$	Marcador referente ao tórax
$L$	Marcador referente ao punho esquerdo
$R$	Marcador referente ao punho direito
$\in$	Pertence
$\mathbb{R}$	Conjuntos dos números reais
$ $	Tal que
$q$	Número do quadro analisado
$Q$	Total de quadros da amostra analisada
$x$	Coordenada espacial X
$y$	Coordenada espacial Y
$z$	Coordenada espacial Z
$VP$	Vetor de características do quadro atual
$MI$	Matriz de características da amostra atual
$f$	Número de quadros analisados na medição
$t$	Tempo de amostragem analisado
$V_m$	Velocidade mínima para se iniciar um movimento
$H_m$	Altura mínima para se iniciar um movimento
$d$	Número de quadros extras a serem armazenados
$DC_{aux}$	Matriz de dados centralizados
$DC$	Subconjunto da matriz de dados centralizados
$\lambda_i$	Importância da variável
$C$	Matriz de coeficiente da função PCA em uma amostra da base de dados
$G$	Subconjunto da matriz de coeficiente da função PCA em uma amostra da base de dados

$\bar{G}$	Subconjunto da matriz de coeficiente da função PCA na amostra analisada
$H$	Distância euclidiana entre as amostras comparadas
$Th$	Limiar de segurança para classificar um gesto
$Z$	Classe inconclusiva

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>14</b>
1.1	Organização textual	16
1.2	Objetivos e metas	17
<b>2</b>	<b>TRABALHOS RELACIONADOS</b>	<b>18</b>
2.1	Extração das informações espaciais	18
2.2	Reconhecimento de uma ação	19
<b>3</b>	<b>METODOLOGIA</b>	<b>21</b>
3.1	<b>Materiais Utilizados</b>	<b>21</b>
3.1.1	Orçamentos e disponibilidade de recursos	21
3.2	<b>Métodos</b>	<b>22</b>
3.2.1	Coleta de dados	22
3.2.2	Algoritmo de reconhecimento de ações	26
3.2.2.1	Parâmetros utilizados	26
3.2.2.2	Pré-Processamento dos dados	27
3.2.2.3	Classificação	28
3.2.3	Experimentos	30
<b>4</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>31</b>
4.1	<b>Coleta de dados</b>	<b>31</b>
4.2	<b>Classificação</b>	<b>31</b>
4.2.1	Simulação 1	32
4.2.2	Simulação 2	34
4.2.3	Simulação 3	35
4.2.4	Simulação 4	37
4.2.5	Simulação 5	38
4.2.6	Simulação 6	40
4.2.7	Simulação 7	41
4.2.8	Simulação 8	43
4.2.9	Simulação 9	44
4.2.10	Simulação 10	46
<b>5</b>	<b>CONSIDERAÇÕES FINAIS</b>	<b>49</b>
	<b>REFERÊNCIAS</b>	<b>51</b>

# 1 Introdução

A robótica é uma das áreas de pesquisa sobre tecnologia muito importante, que está sempre em desenvolvimento, sendo muitas vezes usada para auxiliar o ser humano em tarefas que operam de forma automática e repetidamente (CAMPOS, 2019). Desde o seu surgimento, as aplicações da robótica sempre foram extensas, começando na indústria e se espalhando para outras áreas, como, por exemplo, na saúde, segurança e na educação.

A medida em que os meios tecnológicos evoluem, é possível observar uma crescente em relação à Interação Homem-Máquina (IHM) a qual pode ser aplicada em diversos setores, como, por exemplo, controle de robôs, plataformas ou de salas e/ou ambientes fechados remotamente (MARTINELLI et al., 2022; WIEDERER et al., 2020; HAZRA; SANTRA, 2018). Dessa forma, pode-se dizer que a IHM se baseia em um sistema onde uma tarefa pode ser executada pela cooperação de um componente humano e de um componente de automação, podendo ela ser uma tarefa simples ou uma mais complexa.

Com a busca por uma técnica simples e que possa atender diversas áreas, várias técnicas de interação foram criadas e estudadas com o passar do tempo, visando permitir uma interação mais humana com a tecnologia. Dessa forma, tendo uma maior liberdade ao se realizar os movimento, ao evitar equipamentos tradicionais que se utilizam de componentes específicos e muitas vezes presos ao corpo do operador. Assim, a busca por uma interface de comunicação mais natural ocorre pela tentativa de deixar as pessoas cada vez mais à vontade ao interagir com uma máquina.

Com isto em mente, diversas formas de interfaces já foram criadas e estudadas, como, por exemplo, reconhecimento de voz (CAMPOS; MOUTINHO, 2020), periféricos e reconhecimento de gestos (CHEN et al., 2019). Entre estas, as aplicações baseadas em informações sobre o corpo humano, como, por exemplo, reconhecimento de gesto (baseadas na forma da mão, movimento do corpo ou de membros específicos do corpo), expressões e vozes (SOUSA et al., 2020; GORDILLO, 2018), têm sido usadas cada vez mais em diversas áreas (jogos, acessibilidade, reabilitações, esportes, treinamentos, etc.) (CASTRO; SAKATA; ZAINA, 2019; SOARES et al., 2018; PORCELLIS; SIEDLER; HOLZSCHUH, 2019).

Fala, movimento, gestos e emoções sincronizados geram vídeos que descrevem desejos de ações. Um “Gesto” é uma ação, que quando executada de forma voluntária representa um sinal de comunicação verbal, um comando. Ao criar padrões específicos de comando baseado em alguma parte do corpo, um humano e um robô poderão interagir remotamente, como, por exemplo, o controle de um veículo autônomo por meio de gestos. Por se tratar de uma máquina, a atenção na execução será elevada, e a resposta ao comando

imediate. Assim, essas máquinas poderão auxiliar o ser humano em tarefas específicas, realizando assim a colaboração entre humano e robô (HRC) (MARQUES; CARRIJO; MORAIS, 2020).

Segundo Basilio, Carvalho e Brandao (2019), devido à facilidade e a possibilidade de se trabalhar remotamente e as mais variadas aplicações, o reconhecimento de gestos vem ganhando mais destaque no meio acadêmico. Assim, neste trabalho os autores propõem um método de reconhecimento de ações com enfoque para aplicação em robótica social, utilizando-se de redes neurais simples com redução da dimensionalidade das entradas, treinadas em um banco de dados próprio com um total de onze classes de ações, sendo duas delas ações neutras. As ações foram representadas por *features*, coordenadas cartesianas das juntas do corpo, extraídas pelos sensores.

Outro exemplo de aplicação é o trabalho feito por Carvalho, Basílio e Brandão (2022), no qual os autores propuseram uma solução de um problema real, que pode ser levado para outros cenários mais complexos, ao se preservar seus atributos e a sua proposta de ser uma aplicação de Robótica Socialmente Assistiva. Ainda segundo este trabalho, a principal contribuição seria o *framework* de interação humano-robô, tendo como motivação o desenvolvimento de um aplicativo de Robótica Social Assistiva, com um sistema de implementação fácil de seguir e fornecendo um caminho claro para ser usado como uma abordagem mais prática, por exemplo, na descrição e na resposta de um classificador de ação para um agente do mundo real.

Mais um exemplo aplicação é o trabalho realizado por Carvalho et al. (2022), no qual os autores apresentaram uma solução para a teleoperação de uma equipe heterogênea de robôs móveis. Neste trabalho, os robôs são representados por um veículo aéreo não tripulado e um veículo terrestre não tripulado. Para executar a tarefa, padrões de gestos de alto nível foram feitos em uma estação remota, e um classificador de Rede Neural Artificial (RNA) usado. Como sensor, uma câmera RGB-D (Vermelho, Verde, Azul — Profundidade) gerou dados esqueléticos extraídos e treinados para serem identificados. Como resultado, o classificador usou de dados customizados para construir os padrões de gestos, permitindo o uso de gestos suaves e intuitivos para a operação de robôs móveis.

Com base nisto, a interação homem robô por meio de gestos viabiliza, por exemplo, que uma pessoa possa controlar um determinado ambiente fechado ou um veículo não tripulado sem precisar estar presente no local. Alinhado a isto, existem variadas formas de se fazer o controle de veículos autônomos, entretanto todos seguem a mesma estrutura, sendo esta, baseada no uso de um Sistema de Visão Computacional (SVC), da análise dos dados e por fim a definição do comando do veículo (BASILIO; CARVALHO; BRANDAO, 2019; CARVALHO; BASÍLIO; BRANDÃO, 2022; CARVALHO et al., 2022).

Por meio do SVC é possível extrair diversas informações a partir do processamento de um conjunto de imagens, as quais estarão armazenadas no banco de dados. Atualmente,



existem diversos meios e formas de se fazer este processamento de imagens que possibilitam a captura de informações, como, por exemplo, o uso de sensores como o *OptiTrack* (FERIGATO; LOUREIRO, 2018), a *Intel RealSense* (TEIXEIRA, 2021) e o *Kinect* (SANTOS et al., 2019), contudo, dependendo do investimento e/ou da aplicação, um pode ser mais vantajoso que o outro.

Técnicas de aprendizagem de máquinas são utilizadas para o reconhecimento de ações na identificação por gestos para o controle veículos autônomos. Para aplicar estas abordagens, muitas vezes é essencial avaliar uma série de variáveis registradas em um banco de dados, que pode ser bastante grande em algumas circunstâncias. Diante disto, novas abordagens de processamento de dados, como a Análise de Componentes Principais (PCA) e a Análise Discriminatória Linear (LDA), foram desenvolvidas e vem sendo cada vez mais utilizadas atualmente. Ao permitir a execução de redução de dimensionalidade, compactação de dados, extração de características e visualização de dados, estas abordagens visam facilitar e simplificar a interpretação e o processamento de dados (SAWHNEY et al., 2019; LAHAW; ESSAIDANI; SEDDIK, 2018; GOULART et al., 2019).

Ao longo do tempo foram desenvolvidas variadas técnicas para se realizar a classificação de ações e padrões, como as baseadas em *Artificial Neural Networks* (ANNs) (BRAGA, 2022; FONSECA, 2019), *K-nearest neighbor* (KNN) (ALCANTARA; FERREIRA; SANTOS, 2019; ANDRADE et al., 2021), *Support Vector Machine* (SVM) (SHEENA; NARAYANAN, 2021), *Euclidean Distance* (ALVES et al., 2022), *Vector Space* (DHARMALINGAM; PALANISAMY, 2018), entre outras. Essas técnicas utilizam os dados processados do banco de dados construído, e a partir do método aplicado/escolhido, realiza comparações entre a amostra analisada e o banco de dados, tendo como resposta final a ação classificada.

## 1.1 Organização textual

Este trabalho está dividido em 5 capítulos. O conceito e a área de aplicação são introduzidos no primeiro capítulo, com alguns exemplos de outros trabalhos com temas relacionados. O segundo capítulo analisa as etapas da realização do trabalho e apresenta uma série de trabalhos comparáveis que servem como base para o trabalho produzido. As metodologias empregadas em cada etapa do desenvolvimento do trabalho são descritas em detalhes no Capítulo 3. Os resultados são fornecidos no quarto capítulo. O capítulo final apresenta as considerações finais do trabalho.

## 1.2 Objetivos e metas

Tendo isto em vista, esse trabalho visa contribuir com uma solução prática para comando de veículos autônomos por comando de gestos em tempo real, no qual o operador irá permanecer em um local remoto e por meio de gestos pré-estabelecidos irá gerar um comando para um robô (protótipo de um veículo autônomo) executar. Assim, a implementação de um sistema de reconhecimento de gestos em tempo real foi o objetivo principal deste trabalho.

Em outras palavras, o objetivo, é criar uma ferramenta de reconhecimento de ações de fácil execução. Dessa forma, esta ferramenta conseguirá adquirir as informações de uma ação e discriminá-la em um grupo de 14 classes, em que se utilize poucos parâmetros de entrada e que possua uma técnica de processamento simples no qual o usuário realizará gestos, predefinidos, que estarão associados com ações de alto nível.

Para atingir o objetivo principal, foram estabelecidos objetivos específicos a serem alcançados, sendo estes:

- **Coletar as informações espaciais:** Ao estudar as principais técnicas utilizadas na classificação de padrões de movimentos e na extração de características dos sensores durante os gestos;
- **Processar os dados coletados:** Ao determinar as principais informações de cada gesto, visando facilitar e simplificar a análise dos dados nas próximas etapas.
- **Desenvolver a ferramenta de reconhecimento:** Ao implementar um classificador de gestos utilizando a ferramenta matemática para o reconhecimento de padrões de movimento;

## 2 Trabalhos relacionados

### 2.1 Extração das informações espaciais

A aquisição de dados e a extração de característica é onde determina os aspectos do corpo humano e sua ação a ser rastreada, através do uso de sensores. Alguns métodos de detecção utilizam de câmeras de profundidade ou dispositivos sobre a pele, como *magnetic field trackers*, luvas e roupas que utilizam rastreamento óptico, que captam as informações de posição (BASILIO; CARVALHO; BRANDAO, 2019).

Existem diversas técnicas para estimar movimento em sequência de imagens e reduzir a quantidades de dados a serem processadas. Este processamento digital de imagem (PDI) se dá pelo rastreio das informações relacionadas aos marcadores utilizados pelo usuário ao estabelecer uma correspondência da estrutura de uma imagem entre consecutivos quadros baseados nas suas características principais (JÚNIOR et al., 2020).

A utilização do movimento temporal de objetos sobre um plano de imagem é uma forte sugestão visual para entender a estrutura e movimento em três dimensões (JÄHNE; HAUSSECKER; GEISSLER, 1999). Conhecendo-se a base temporal entre duas imagens consecutivas, o campo vetorial de deslocamento pode ser inferido a partir do campo de velocidades.

Utilizando uma simples analogia, o próprio sistema visual humano usa o movimento visual para inferir propriedades do mundo real 3D baseando-se em um pequeno conhecimento prévio do mesmo, o que é um dos principais objetivos da visão computacional (TRUCCO; VERRI, 1998). Por outro lado, a câmera RGB-D usa de sensores para permitir uma análise detalhada do movimento. Assim, o resultado do reconhecimento de ações por robôs requer a utilização de diferentes dispositivos de imagem e técnicas de rastreamento (BASILIO; CARVALHO; BRANDAO, 2019).

Levando-se em consideração que, para se ter um reconhecimento de ação eficiente, diversos dispositivos podem ser usados, podendo eles variarem de situação para situação. Neste sentido, o sensor *OptiTrack* se torna uma excelente ferramenta para a coleta de dados, devido a sua alta precisão e pela quantidade de informação que se pode coletar. Em seu conjunto, o *OptiTrack* é composto por marcadores reflexivos e câmeras infravermelhas (que fazem a função de sensores de profundidade), capazes de inferir a posição tridimensional dos marcadores (MOTIVE, 2020).

Este sistema de rastreamento possibilita a calibração e a configurações de diversas funções que auxiliam a captura de informações, fornecimento de interfaces para captura e processamento de dados 3D, podendo ser gravado ou transmitido ao vivo para outros

programas (MOTIVE, 2020). Além destas funcionalidades, o *OptiTrack* pode ser utilizado em outras aplicações, como, por exemplo, produções visuais, estudo de movimentos de corpos, realidade virtual, animações e na robótica.

Um método de coleta de dados envolve o uso de marcadores, que podem ser ativos ou passivos. Os marcadores ativos usam diodos emissores de luz que estão situados no corpo do usuário para emitir sua própria luz, enquanto os marcadores passivos refletem apenas a luz infravermelha das câmeras. A câmera pode reconhecer os marcadores mais brilhantes e distingui-los de coisas menos brilhantes captando este sinal, emitido ou refletido, conforme a situação. Outra diferença entre eles é que os marcadores passivos não necessitam de cabos ou baterias, contudo requerem mais pré-processamento (VITAL et al., 2015; FURTADO et al., 2019).

## 2.2 Reconhecimento de uma ação

Um exemplo sobre o uso de técnicas de reconhecimento de ações é em Carvalho, Basílio e Brandão (2022), no qual os autores fizeram um trabalho que descreve a estrutura de interação de um humano e um robô, contando com a redução da dimensão das entradas do sistema, a fim de exigir um conjunto de dados menor para o treinamento de uma Rede Neural Artificial. Outra questão que precisa ser abordada é que, em algumas circunstâncias, o ruído pode impedir que as amostras de dados incluam movimentos ou gestos importantes para aplicações específicas. Em outras palavras, dependendo das ações utilizadas e do nível de complexidade da ação, controlar um veículo autônomo usando o reconhecimento da ação pode resultar em um excesso de informação que deve ser processada, o que pode se revelar impraticável em algumas circunstâncias.

A distribuição desigual de amostras para cada ação é outra questão que afeta a aquisição. Como resultado, diferentes aplicações exigem diferentes tamanhos de amostras, duração de tempo e duração da ação na coleta de dados. Além disso, para execução de comandos, o algoritmo deve reconhecer a ação motora humana. Para chegar a esse resultado é necessário obter a extração de características e a classificação (SHAHROUDY et al., 2016; CARVALHO; BASÍLIO; BRANDÃO, 2022).

Em geral, a identificação de um gesto é feito por meio de comparações entre as informações disponibilizadas em uma base de dados. Desta forma, um dos pontos principais em trabalhos de identificação e/ou reconhecimento é determinar a forma de se realizar estas comparações, podendo elas serem efetuadas por técnicas matemáticas simples ou mais complexas como: ANNs, KNN, SVM, *Euclidean Distance*, *Vector Space*, etc (FONSECA, 2019; ANDRADE et al., 2021; SHEENA; NARAYANAN, 2021; ALVES et al., 2022; DHARMALINGAM; PALANISAMY, 2018).

Com isto, a depender da aplicação trabalho, uma técnica pode ser melhor que outra,

no âmbito do poder computacional necessário para realizá-la. Por exemplo, a distância euclidiana é uma técnica simples e rápida que necessita de poucos dados de entrada e realiza as comparações entre estas informações para obter uma resposta (ALVES et al., 2022). Já as Redes Neurais Artificiais (RNA) são técnicas computacionais que apresentam um modelo matemático inspirado nos humanos. Neste caso, as entradas das redes são ponderadas pelos pesos e por meio do aprendizado supervisionado, saídas e entradas são correlacionadas e usadas para classificação, necessitando de um poder computacional maior, mais tempo para treino, entre outros (BRAGA, 2022; FONSECA, 2019; HAYKIN, 2001).

Outra técnica muito utilizado em problemas deste tipo é o algoritmo K-nearest neighbor, o qual consegue encontrar uma quantidade mínima de vizinhos, representado pela variável  $K$ , onde será calculado a distância entre cada atributo com relação aos  $K$  vizinhos. Assim, a classe de resposta que aparecer mais vezes dentre os  $K$  vizinhos será escolhida (LEITE et al., 2021). Um ponto interessante a se citar é que o método da *Euclidean Distance*, é uma forma alternativa do algoritmo de classificação KNN, no caso específico em que  $K = 1$ .

Generalizando, reconhecer padrões de gestos reveste-se de grande importância para aplicações industriais. Reconhecendo-se a posição e orientação dos corpos e aplicando-se técnicas matemáticas e computacionais, pode-se chegar a sistemas automatizados de reconhecimento e execução de comandos, possibilitando aplicação prática em diferentes esferas. O ramo de controle de veículos autônomos estuda como os robôs reconhecem, modelam e executam padrões (CHANDARANA et al., 2017; SCHULTE et al., 2022).

Assim, este trabalho teve como foco o uso da técnica PCA cuja função principal é encontrar o padrão dos dados analisados e assim encontrar a base de dados mais significativa para então retorná-lo como um subconjunto de informações, revelando a estrutura oculta no conjunto de dados e filtrando o ruído (KURITA, 2019). Além disto, este trabalho também teve como foco o uso do método da *Euclidean Distance* para realizar as classificações objetivando um baixo custo computacional.

## 3 Metodologia

### 3.1 Materiais Utilizados

Para a realização deste trabalho foi utilizado o Sistema de Rastreamento *OptiTrack* (SRO), o qual utiliza o *software Motive* e as câmeras RGB-D, para o rastreamento dos dados espaciais e em seguida, estes dados foram processados no classificador desenvolvido. Na Figura 1 são apresentadas as câmeras RGB-D e os marcadores passivos usados no trabalho. Além disso, também é apresentado a ilustração de um computador indicando os *softwares* usados para coleta de dados e desenvolvimento da ferramenta, sendo eles o *MatLab* e *Motive*.



Figura 1 – Componentes necessários para a realização do trabalho.

#### 3.1.1 Orçamentos e disponibilidade de recursos

O trabalho foi executado utilizando o *software* de simulação *MATLAB* e o SRO. Os principais recursos utilizados no trabalho estavam disponíveis no laboratório de robótica da Universidade Federal de Viçosa (UFV) e não possuíram nenhum gasto para sua utilização no trabalho. O preço médio dos componentes são apresentados na Tabela 1.

Tabela 1 – Custo médio para realização do trabalho.

<b>Materiais</b>	<b>Valor Unitário [U\$]</b>	<b>Aplicação</b>
4 Câmeras Prime x 13	3046	Mapeamentos dos dados
4 Câmeras Prime x 22	4717	Mapeamentos dos dados
Licença do <i>software Motive</i>	999	Acesso aos dados rastreados
Licença do software MatLab	2350	Desenvolvimento do classificador
<b>Preço total:</b>		U\$ 34402

\* Os valores informados foram levantados na data atual de escrita deste documento (30/08/2022).

## 3.2 Métodos

Trabalhar com reconhecimento de gestos por meio de técnicas de aprendizado de máquinas, exige comumente um abundante conjunto de dados. Este trabalho, entretanto, possui a proposta de interagir com um determinado ambiente através de gestos simples que possam ser realizados por diferentes usuários em tempo real. Assim, criou-se uma base de dados própria, onde cada tarefa está associada a uma classe diferente, com o intuito de realizar comandos próprios e ganhar um pouco mais de liberdade e flexibilidade.

### 3.2.1 Coleta de dados

O SRO é constituído por 8 câmeras infravermelhas posicionadas conforme na Figura 2 e para o rastreamento dos dados foram criados 3 “Corpos Rígidos” no *software Motive*, sendo estes: duas pulseiras compostas pelo modelo 3D impresso e por 3 marcadores, e um crachá composto por um cartão de identificação e 5 marcadores (neste caso retratado pela fita reflexiva).

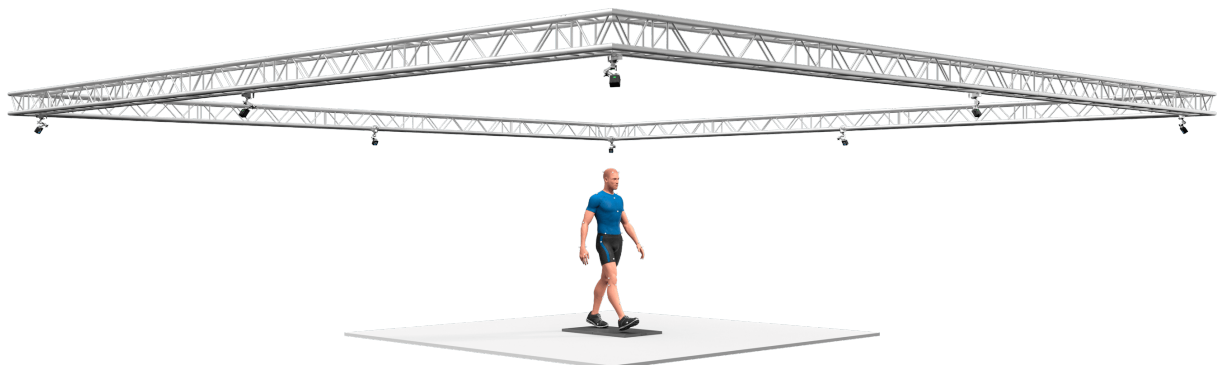


Figura 2 – (MOTIVE, 2020)

Na Figura 3, são apresentados os componentes utilizado durante os experimentos de validação, e como eles são vistos pelo *software*, sendo composta por um marcador,

posicionado na região do tórax (T), por uma pulseira no punho esquerdo (L) e por uma pulseira no punho direito (R). Vale ressaltar que para a construção dos “Corpos Rígidos” os quais serão os objetos a serem rastreados, é importante que os marcadores possuam posicionamentos diferentes (de forma aleatória) a fim de evitar que o software alterne entre os objetos rastreados, causando um erro na classificação.

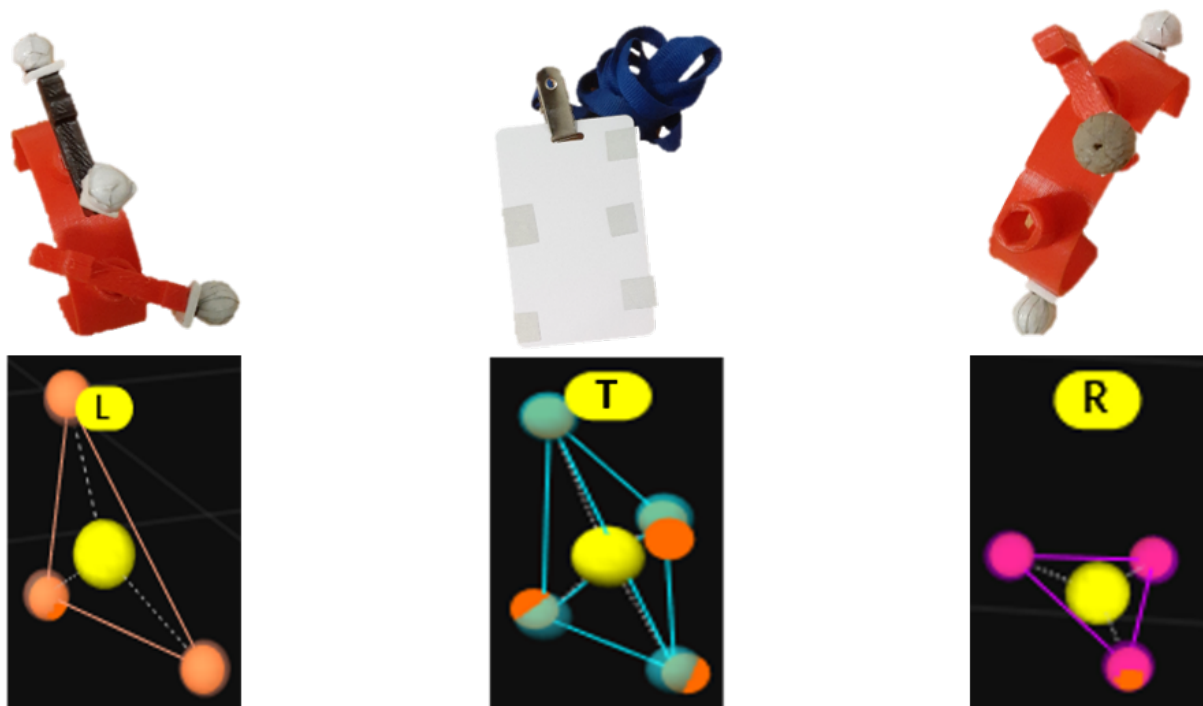


Figura 3 – Pulseiras, Crachá e Corpos rígidos utilizados para realizar o rastreamento dos dados.

Por meio do posicionamento das câmeras é possível se ter um grande espaço de trabalho, dessa forma podendo executar os gestos de diversas posições diferentes. Assim, a fim de tornar mais dinâmico e simples o processo de reconhecimento do gesto, foram utilizadas duas pulseiras que serão colocadas, uma em cada punho do usuário e um terceiro marcador na região do tórax, atuando como ponto de referência, para serem os marcadores rastreados pelo *OptiTrack*, conforme apresentado na Figura 4.

Para facilitar o mapeamento dos dados foi optado por realizar os mesmos gestos realizados por (CARVALHO et al., 2022) no qual se verificou serem simples e eficientes tanto para o usuário realizar quanto para o *software* rastreá-lo. Dessa forma, as classes utilizadas são apresentadas na Figura 5 e suas descrições são apresentadas a seguir:

- A) Levantar as duas mãos, à frente do corpo, e depois abaixar a direita;
- B) Girar os dois braços, terminando com eles na frente do peito, com a mão direita ligeiramente acima da esquerda;



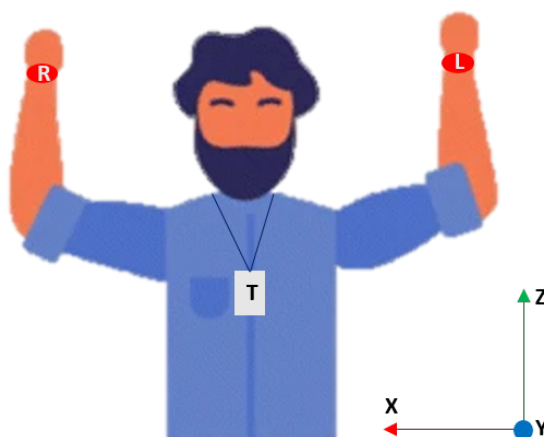


Figura 4 – Localização dos marcadores no usuário.

- C) Levantar as duas mãos, à frente do corpo, e depois abaixar a esquerda;
- D) Desenhar um círculo ao lado do rosto com a mão direita, enquanto mantém o braço esquerdo abaixado ao lado do corpo;
- E) Desenhar um círculo ao lado do rosto com a mão direita, enquanto levanta o braço esquerdo até a altura do tórax;
- F) Desenhar um círculo ao lado do rosto com a mão direita, enquanto levanta o braço esquerdo até a altura da cabeça;
- G) Desenhar um círculo ao lado do rosto com a mão esquerda, enquanto mantém o braço direito abaixado ao lado do corpo;
- H) Desenhar um círculo ao lado do rosto com a mão esquerda, enquanto levanta o braço direito até a altura do tórax;
- I) Desenhar um círculo ao lado do rosto com a mão esquerda, enquanto levanta o braço direito até a altura da cabeça;
- J) Cruzar os dois braços formando um X na frente do peito;
- K) Mover o braço direito para direita e depois deslizar para a esquerda até encostar no estômago;
- L) Levantar os dois braços até a altura do estômago e depois deslize a mão direita para a diagonal direita superior até a altura do ombro;
- M) Levantar os dois braços até a altura do estômago e depois deslize a mão esquerda para a diagonal esquerda superior até a altura do ombro;
- N) Mover o braço esquerdo para esquerda e depois deslizar para a direita até encostar no estômago.

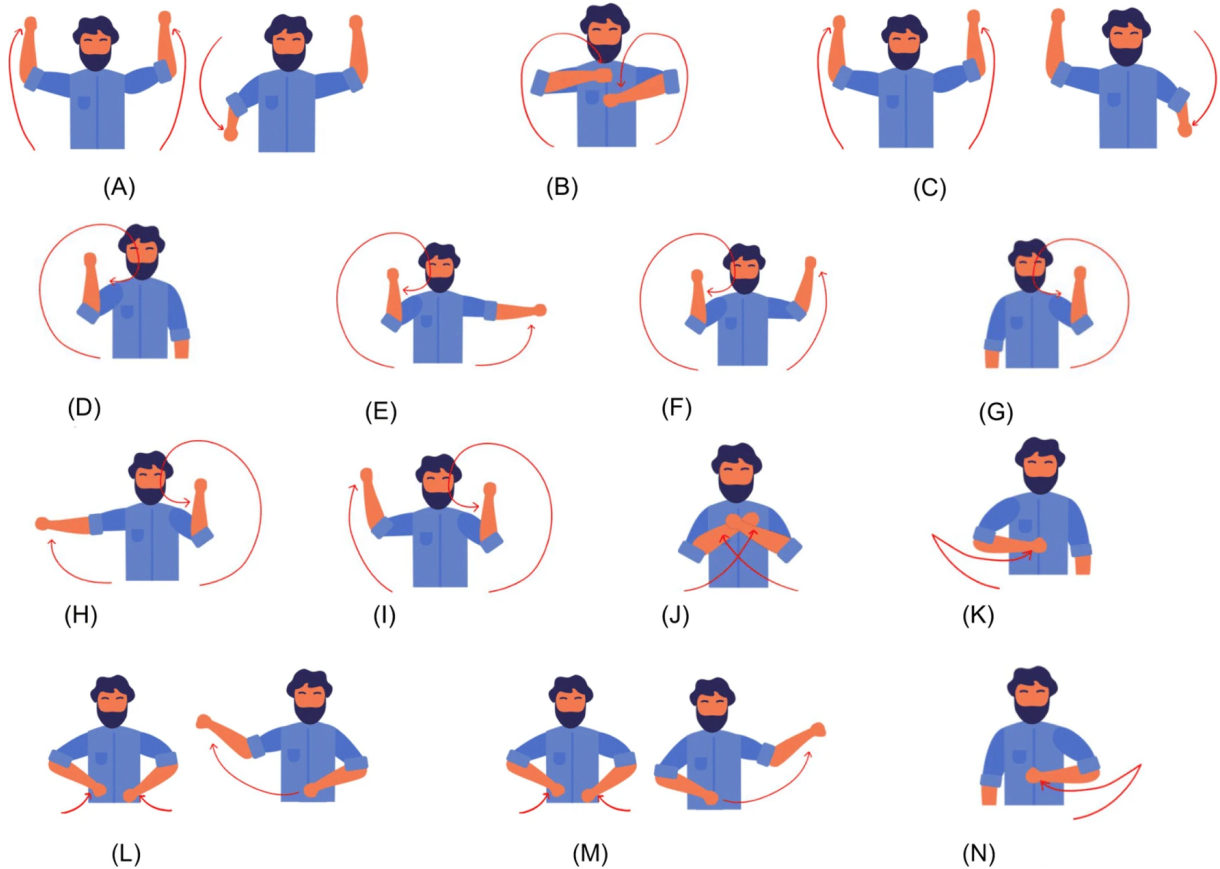


Figura 5 – (CARVALHO et al., 2022)

Utilizando o *OptiTrak* para extrair as informações dos marcadores, cada amostra foi capturada com uma taxa de 120 FPS (quadros por segundo). Cada quadro consiste em um vetor de características ( $VP$ ), Equação (3.1), o qual é composto por um marcador, posicionado na região do tórax ( $T$ ), por uma pulseira no punho esquerdo ( $L$ ) e por uma pulseira no punho direito ( $R$ ), onde cada marcador possui uma componente  $x$ ,  $y$  e  $z$ , sendo  $q$  o índice que indica qual o quadro atual. Ao se iniciar uma ação, é gerado um vetor de características a cada quadro, o qual é concatenado um após o outro, que juntos compõem a matriz de características ( $MI$ ).

$$VP_q = [T_q \quad L_q \quad R_q] \mid VP \in \mathbb{R}^{1 \times 9} \quad (3.1)$$

Matematicamente,  $T_q$ ,  $L_q$  e  $R_q \in \mathbb{R}^{1 \times 3}$  representam as coordenadas  $x$ ,  $y$  e  $z$  do tórax, do punho esquerdo e do punho direito, respectivamente, com  $q = 1, 2, \dots, Q$ , onde  $Q$  informa o último quadro. Em seguida, ao se iniciar uma ação, é gerado um vetor de características a cada quadro de duração, o qual é concatenado um após o outro, compondo assim a matriz de características ( $MI$ ). Ou seja, a amostra completa é armazenada como a concatenação subsequente da matriz de características, Equação (3.2). Após obter essa

matriz, o programa desenvolvido entra no segundo estágio, o qual realiza o reconhecimento do gesto.

$$MI = \begin{bmatrix} T_1 & L_1 & R_1 \\ T_2 & L_2 & R_2 \\ T_3 & L_3 & R_3 \\ \vdots & \vdots & \vdots \\ T_Q & L_Q & R_Q \end{bmatrix} \in \mathbb{R}^{Q \times 9} \quad (3.2)$$

### 3.2.2 Algoritmo de reconhecimento de ações

Ao realizar uma análise comparativa com outras bases de dados, em trabalhos semelhantes, percebe-se que muitas delas têm uma abundante quantidade de amostras para cada gesto. Contudo, como a proposta do trabalho é desenvolver um classificador simples, foi optado por começar a base de dados com 70 amostras iniciais, sendo 14 classes diferentes e 5 amostras para cada classe, podendo aumentar a base para os gestos que ficarem com resultado abaixo do esperado, visando se ter uma base de dados mínima, mas eficiente.

Desta forma, utilizando o *software Motive*, foram realizadas os rastreamento das posições espaciais de cada marcador, conforme mostrado na Figura 4, nas quais o início e o fim de cada ação foram considerados a posição de repouso, sendo esta definida quando uma pessoa está com os braços para baixo.

#### 3.2.2.1 Parâmetros utilizados

Para se detectar o início e o fim de um gesto, a abordagem escolhida foi a utilização de um gatilho duplo, no qual são avaliados a altura e a velocidade de ambas as pulseiras. Levando-se em conta que todos os movimentos necessitam que ao menos uma das pulseiras seja erguida em relação ao solo, foi escolhido a coordenada  $z$  de cada marcador para se analisar. Portanto, ao se analisar uma janela fixa de  $f$  quadros, resultando em uma amostragem de  $t$  milissegundos, é possível obter a altura e velocidade de cada pulseira naquele instante. Logo, quando a velocidade e a altura do mesmo alcançassem um determinado valor mínimo ( $V_m$  em metros por quadros e  $H_m$  em metros), foram armazenados os últimos  $d$  quadros anteriores ao acionamento de uma das condições e em seguida o quadro atual e parando somente quando ambas as condições não forem mais cumpridas.

A taxa de perda de dados foi investigada para frequências de 60 Hz, 120 Hz, 180 Hz e 240 Hz, a fim de escolher a frequência de coleta de informações que seria empregada. Além disso, o volume de dados aumenta com a frequência, sendo outro elemento significativo. A

fim de capturar as informações sem produzir uma quantidade excessiva de dados, estes dois pontos foram combinados para determinar um valor de 120 Hz.

Os valores de  $f$  e  $t$  são baseados na frequência de coleta de informações utilizada no *software Motive*, neste caso 120 Hz, portanto para uma amostragem de 100 milissegundos foram obtidos intervalos fixos de 12 quadros, e dessa forma o valor de  $d$  é definido como um a menos que  $f$ , 11 quadros. Os valores de  $V_m$  e  $H_m$  foram determinados ao considerar a altura média da posição de repouso dos possíveis usuários, sendo esta 0,95 metros, e a velocidade média do movimento dada a precisão milimétrica do SRO, sendo esta 0,001 metros por quadros.

### 3.2.2.2 Pré-Processamento dos dados

Assim, para realizar o reconhecimento da ação, foram coletadas as informações das duas pulseiras, ao longo do gesto. Já o marcador que está posicionado na região do tórax foi utilizado como referência, ao considerar que um mesmo gesto realizado em posições diferentes terá informações diferentes devido aos seus valores de coordenadas globais diferirem, dessa forma o marcador de referência permite que o gesto seja coletado com um referencial relativo em relação ao crachá. Além disso, outra função desse marcador é a correção de uma eventual movimentação do corpo em decorrência do movimento dos braços, ao realizar um gesto.

Dessa forma, foi obtida a distância relativa entre as pulseiras e o crachá, durante o gesto. Com isso, todas as informações de  $x$ ,  $y$  e  $z$ , foram centralizadas em torno de do marcador “T”. Para isto, adotou-se a Equação 3.3, onde  $MI$  é a matriz contendo todas as informações dos três marcadores referentes aquela amostra e  $DC_{aux}$  é o conjunto de dados centralizado.

$$DC_{aux} = MI - \begin{bmatrix} T_1 & T_1 & T_1 \\ T_2 & T_2 & T_2 \\ T_3 & T_3 & T_3 \\ \vdots & \vdots & \vdots \\ T_Q & T_Q & T_Q \end{bmatrix}, \text{ onde } DC_{aux} \in \mathbb{R}^{Q \times 9} \quad (3.3)$$

Com isto, após a centralização dos dados, são desconsideradas as informações relacionadas ao marcador do tórax. Isto é, após a operação anterior as informações das três primeiras colunas de  $DC_{aux}$  foram zeradas, e com isto diminuindo o número de variáveis a serem processadas nas próximas etapas, gerando um novo conjunto de dados:  $DC \in \mathbb{R}^{Q \times 6}$ .

Em seguida, aplicando a técnica da PCA em  $DC$  é possível descobrir qual a ordem de importância de cada variável para o gesto em análise. Ou seja, qual componente e qual pulseira é tem maior impacto naquele movimento. Esta importância é explicada através

das variâncias dos componentes principais, os autovalores da matriz de covariância dos dados, retornados como um vetor coluna, através do uso da técnica PCA.

Cada autovalor corresponde ao impacto de uma das variáveis de entrada na ação analisada, sendo que cada variável de entrada é dada por uma coordenada espacial ( $x$ ,  $y$  e  $z$ ) de cada pulseira, totalizando seis autovalores. Ao se determinar a ordem de importância das variáveis, torna-se vantajoso analisar apenas as de maior impacto na ação, reduzindo o número de variáveis a serem analisadas novamente.

Para isto, foi utilizado a função *PCA* presente no *software MATLAB*. Ao se passar os dados do gesto em análise, serão retornados os autovalores em ordem decrescente, conforme em (3.4) onde  $\lambda_i$  representa a importância de cada variável para aquela amostra.

$$\lambda_1 > \lambda_2 > \lambda_3 > \lambda_4 > \lambda_5 > \lambda_6 \quad (3.4)$$

Além disso, também será retornado o conjunto de dados  $C \in \mathbb{R}^{6 \times 6}$  que representa os coeficientes dos componentes principais, onde cada coluna da matriz contém os coeficientes para um componente principal e se encontram em ordem decrescente de variação de importância. A partir disto, uma classe de referência, aqui denominada  $\bar{G}$ , é obtida pela representação dos autovetores associados a estes autovalores do gesto analisado.

Dado esta informação, após aplicar a técnica na base de dados inicial, foi observado que os três primeiros autovalores descrevem mais de 90% das informações sobre o gesto. Logo, optou-se por escolher os três autovetores associados aos maiores autovalores para cada amostra ( $\lambda_1$ ,  $\lambda_2$  e  $\lambda_3$ ), diminuindo a quantidade de dados a serem processados nas próximas etapas do trabalho. Como foram escolhidas as três componentes mais importantes, ou seja, as variáveis que possuem maior significado para aquele gesto, foi obtido  $G \in \mathbb{R}^{6 \times 3}$ , que contém um conjunto de dados reduzido de  $C \in \mathbb{R}^{6 \times 6}$ .

### 3.2.2.3 Classificação

O processo de classificação é baseado na comparação de uma ação qualquer com outra ação de referência. Para realizar isto, são utilizados os dados contidos na matriz  $G \in \mathbb{R}^{6 \times 3}$  obtido da amostra de referência e da amostra em análise.

Para a construção da base dados, foram realizadas as leituras das 70 amostras iniciais e armazenando as informações sobre a PCA de cada uma em sua respectiva classe, ou seja, a base de dados é composta amostras com tamanho de  $6 \times 3$ .

O reconhecimento do gesto pode ser obtido por meio da comparação da série de informações do conjunto de dados reduzidos da ação a ser classificada com a base de dados. Ao percorrer a base de dados, alterando para diferentes classes, aquela em que se obter a menor diferença representa a classe da ação a ser classificada. Em outras palavras, a

classificação acontece comparando a classe corrente,  $G$ , com as classes da base de dados,  $\bar{G}$ , pelo cálculo da menor distância, conforme,

$$H = \sum_{i=1}^6 \sum_{j=1}^3 (\bar{G}_{i,j} - G_{i,j})^2. \quad (3.5)$$

Em seguida, para obter a classificação final de gesto, a amostra corrente deve possuir a menor distância euclidiana para toda a base de dados. Ou seja, do total de amostra que cada classe possui na base de dados, aquele em que se obter o menor valor para  $H$ , representa o gesto executado. Além disto, pensando em situações onde o usuário possa executar um gesto fora da base de dados ou acabar realizando o movimento errado, foi estabelecido um limiar de segurança,  $Th$ , igual a 1,5, no qual se o valor de  $H$  superar o mesmo, a amostra analisada será considerada “inconclusiva”,  $Z$ . Este valor de limiar foi determinado através da realização de alguns testes onde se analisou o valor médio de  $H$  para cada classe de ação executada.

A Figura 6 ilustra um fluxograma que resume as etapas do subsistema de classificação realizadas neste trabalho e seus respectivos resultados.

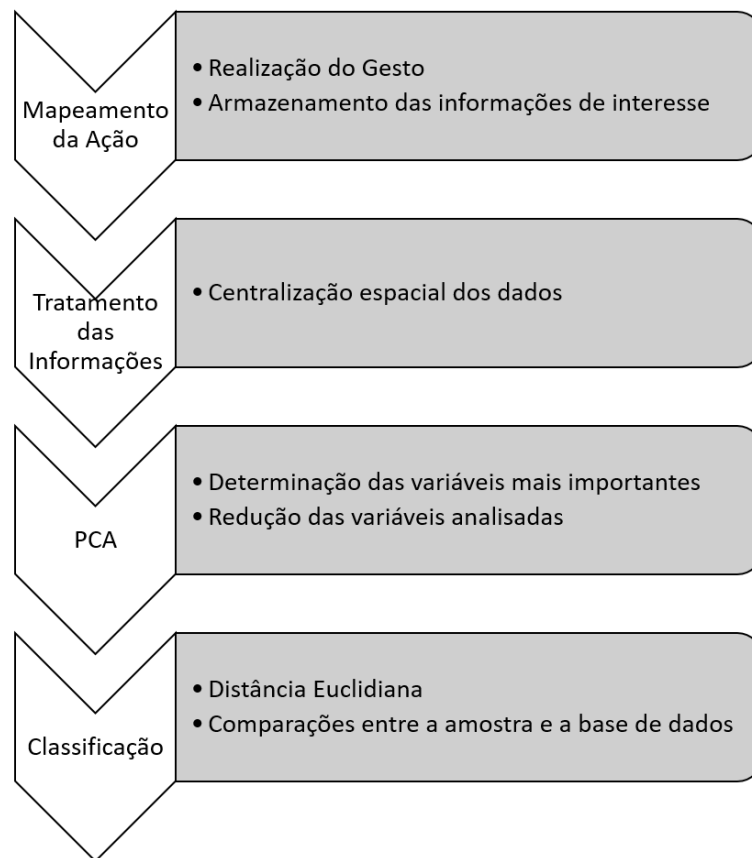


Figura 6 – Fluxograma com as etapas e respostas do classificador proposto.

### 3.2.3 Experimentos

Para utilizar a ferramenta de classificação, baseada na Equação (3.5), é necessário haver uma base de dados inicial para comparar com o gesto executado. Assim, foram gravados 70 gestos iniciais, sendo cinco de cada classe, onde cada gesto foi salvo em uma linha diferente na matriz que compõe a base de dados.

A validação da ferramenta desenvolvida foi feita ao realizar diversas simulações, nas quais foram realizadas os gestos pré-definidos por um único usuário e em seguida analisando se o classificador proposto realizou a classificação correta ou não, por meio da matriz de confusão. Com isto, para a fase de treino, foi realizado um treinamento dividido em dez etapas, onde em cada etapa seriam executados 70 gestos, cinco de cada classe de forma totalmente aleatória, totalizando 700 gestos de treino, sendo 50 para cada amostra, e assim analisando a eficiência da base para a etapa em questão.

Durante o treino, a ferramenta de classificação desenvolvida poderia retornar três repostas, “correta”, “errada” ou “inconclusiva”. Para cada amostra classificada corretamente, foi armazenado o seu respectivo conjunto de dados reduzido da sua PCA, matriz  $\tilde{\mathbf{G}}$ , e caso classificado erroneamente ou como inconclusivo, a amostra foi descartada. Ou seja, o tamanho da base de dados se deu em função da quantidade de amostras classificadas corretamente.

Um ponto muito importante em operações que envolvem ações humanas, é a repetibilidade da ação, ao se considerar que repetir um movimento de forma exatamente igual é uma tarefa árdua. Logo, ao incrementar a base de dados com várias amostras para cada gesto, possibilita que a ferramenta identifique as ações com uma pequena margem de variação no gesto (por exemplo, quando executado um pouco mais rápido ou mais lento, com os punhos em uma posição mais alta ou mais baixa).

Em outras palavras, a base de dados que iniciou com 5 amostras para cada classe, ao final do treinamento possuirá como tamanho máximo 55 amostras para cada classe, caso todas as classificações sejam corretas. Após isto, ao analisar os resultados de cada classe, para aquelas que não alcançaram o resultado esperado, poderia ser executado uma rodada extra de treino focada apenas nessas classes. Dessa forma, tendo uma base de dados adaptável, com tamanho mínimo, porém eficiente, evitando que o tempo de processamento aumente muito.

## 4 Resultados e Discussão

### 4.1 Coleta de dados

Para realizar as leituras dos gestos, foi utilizado o *software* Motive, no qual o *layout* de captura é apresentado na Figura 7. Nela, pode ser visto o espaço a ser mapeado e a presença de três câmeras na parte superior.

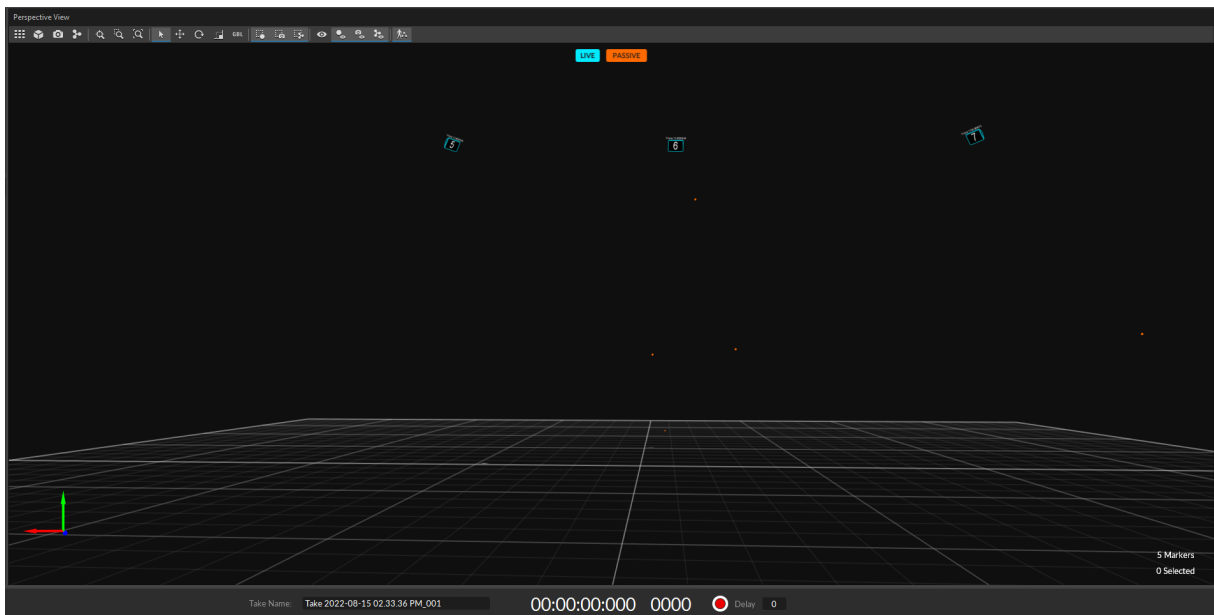


Figura 7 – Painel de captura do software utilizado para as gravações das amostras.

Na Figura 8 é apresentado uma das amostras da classe de gesto N utilizadas no trabalho, na qual  $Q = 3822$  quadros. Na parte superior da figura é mostrada o espaço de trabalho e o mapeamento dos marcadores em suas respectivas posições em determinado instante e na parte inferior é mostrado o gráfico do movimento em relação ao tempo de simulação. Sendo assim, ao analisar as duas partes simultaneamente, observa-se que a medida que o braço esquerdo do usuário se move, o valor no eixo corresponde, para o marcador *Left*, se altera também, mostrando o mapeamento daquele marcador.

### 4.2 Classificação

A validação do método proposto foi feita ao realizar diversas simulações, nas quais foram realizadas os gestos pré-definidos e em seguida analisando se o classificador proposto realizou a classificação correta ou não, por meio da matriz de confusão.



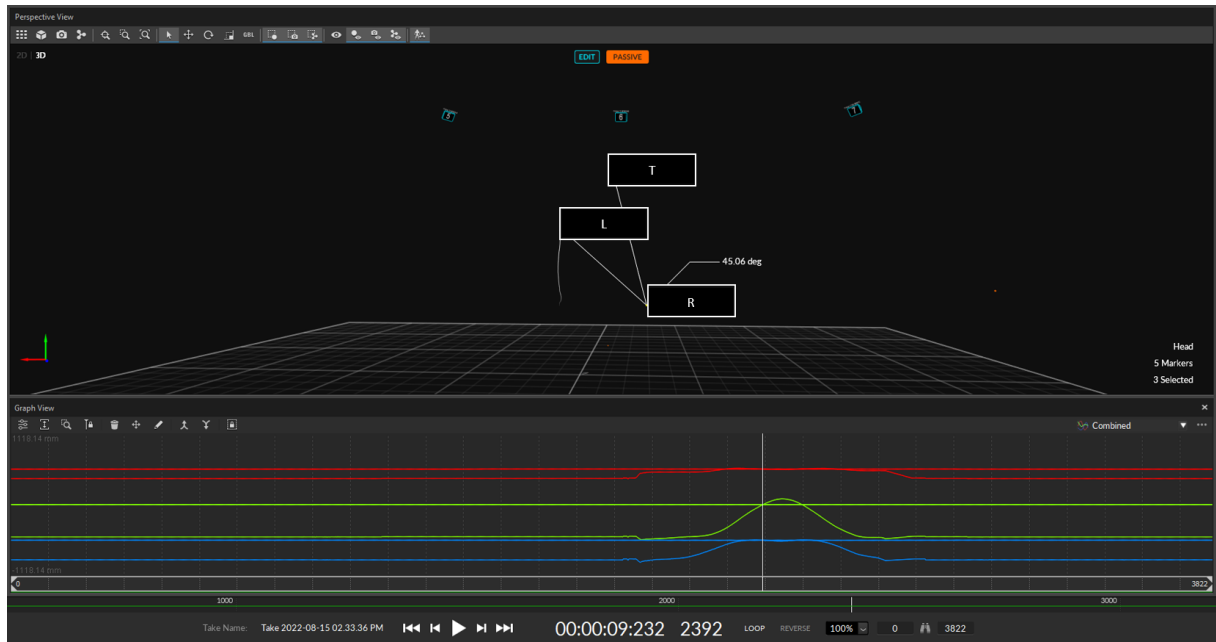


Figura 8 – Exemplo de uma amostra coletada durante o treino.

### 4.2.1 Simulação 1

Ao realizar a primeira rodada de simulação, foram obtidas os resultados apresentados na Tabela 2, em que é possível perceber que houveram nove classes com 100% de acerto (A, B, C, D, E, F, G, K e N), três com 80% de acerto (H, I e L), uma classe com 60% de acerto (M) e uma classe com 40% (J), uma amostra foi confundida com outra e sete amostras classificadas como inconclusivas, com isto, esta rodada de simulação obteve um desempenho médio de 88,57% de acerto em 70 amostras avaliadas.

Já a Tabela 3 retrata o resultado total do teste de validação, contudo como é a primeira simulação, o resultado acabou por ser igual ao da simulação. Além disto, pode-se perceber que as classes de piores resultados (abaixo de 80%) são as classes J e M, com 40% e 60%, respectivamente, de acerto individuais de um total de 5 amostras para cada uma.

Tabela 2 – Matriz de confusão com o resultado da simulação 1, 88,57% de acerto.

		Classe Previstas															FP
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	Z	
Classes Desejada	A	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	B	-	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	C	-	-	5	-	-	-	-	-	-	-	-	-	-	-	-	-
	D	-	-	-	5	-	-	-	-	-	-	-	-	-	-	-	-
	E	-	-	-	-	5	-	-	-	-	-	-	-	-	-	-	-
	F	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-	-
	G	-	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-
	H	-	-	-	-	-	-	-	4	-	-	-	-	-	-	1	1
	I	-	-	-	-	-	-	-	1	4	-	-	-	-	-	-	1
	J	-	-	-	-	-	-	-	-	-	2	-	-	-	-	3	3
	K	-	-	-	-	-	-	-	-	-	-	5	-	-	-	-	-
	L	-	-	-	-	-	-	-	-	-	-	-	4	-	-	1	1
	M	-	-	-	-	-	-	-	-	-	-	-	-	3	-	2	2
	N	-	-	-	-	-	-	-	-	-	-	-	-	-	5	-	-
	Z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	FN	-	-	-	-	-	-	-	1	-	-	-	-	-	-	7	88.57%

Tabela 3 – Matriz de confusão com o resultado agregado após a simulação 1, 88,57% de acerto.

		Classe Previstas															FP
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	Z	
Classes Desejada	A	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	B	-	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	C	-	-	5	-	-	-	-	-	-	-	-	-	-	-	-	-
	D	-	-	-	5	-	-	-	-	-	-	-	-	-	-	-	-
	E	-	-	-	-	5	-	-	-	-	-	-	-	-	-	-	-
	F	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-	-
	G	-	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-
	H	-	-	-	-	-	-	-	4	-	-	-	-	-	-	1	1
	I	-	-	-	-	-	-	-	1	4	-	-	-	-	-	-	1
	J	-	-	-	-	-	-	-	-	-	2	-	-	-	-	3	3
	K	-	-	-	-	-	-	-	-	-	-	5	-	-	-	-	-
	L	-	-	-	-	-	-	-	-	-	-	-	4	-	-	1	1
	M	-	-	-	-	-	-	-	-	-	-	-	-	3	-	2	2
	N	-	-	-	-	-	-	-	-	-	-	-	-	-	5	-	-
	Z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	FN	-	-	-	-	-	-	-	1	-	-	-	-	-	-	7	88.57%

### 4.2.2 Simulação 2

Após realizar a segunda rodada de simulação, foram obtidas os resultados apresentados na Tabela 4, em que é possível perceber que houveram sete classes com 100% de acerto (A, B, C, G, H, L e N), cinco com 80% de acerto (D, E, F, K e M), duas classes com 60% de acerto (I e J). Para esta simulação, nove gestos não foram classificados corretamente, sendo cinco confundidos e quatro inconclusivos. Com isto, esta rodada de simulação obteve um desempenho médio de 87,14% de acerto em 70 amostras avaliadas.

A Tabela 5 apresenta o resultado agregado do teste de validação, no qual as classes de piores resultados, nesta etapa, são as classes J, I e M, com 50%, 70% e 70%, respectivamente, de acerto individuais de um total de 10 amostras para cada uma. Além disto, nota-se que o desempenho geral dos testes é 87,86% de um total de 140 amostras avaliadas até esta simulação.

Tabela 4 – Matriz de confusão com o resultado da simulação 2, 87,14% de acerto.

		Classe Previstas															
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	Z	FP
Classes Desejada	A	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	B	-	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	C	-	-	5	-	-	-	-	-	-	-	-	-	-	-	-	-
	D	-	-	-	4	-	-	-	1	-	-	-	-	-	-	-	-
	E	-	-	-	-	4	-	-	1	-	-	-	-	-	-	-	-
	F	-	-	-	-	1	4	-	-	-	-	-	-	-	-	-	-
	G	-	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-
	H	-	-	-	-	-	-	-	5	-	-	-	-	-	-	-	-
	I	-	-	-	-	-	-	-	2	3	-	-	-	-	-	-	-
	J	-	-	-	-	-	-	-	-	-	3	-	-	-	-	2	-
	K	-	-	-	-	-	-	-	-	-	-	4	-	-	-	1	-
	L	-	-	-	-	-	-	-	-	-	-	-	5	-	-	-	-
	M	-	-	-	-	-	-	-	-	-	-	-	-	4	-	1	-
	N	-	-	-	-	-	-	-	-	-	-	-	-	-	5	-	-
	Z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	FN	-	-	-	-	1	-	-	4	-	-	-	-	-	-	4	87.14%

Tabela 5 – Matriz de confusão com o resultado agregado após a simulação 2, 87,86% de acerto.

		Classe Previstas															FP
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	Z	
Classes Desejada	A	10	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	B	-	10	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	C	-	-	10	-	-	-	-	-	-	-	-	-	-	-	-	-
	D	-	-	-	9	-	-	-	1	-	-	-	-	-	-	-	1
	E	-	-	-	-	9	-	-	1	-	-	-	-	-	-	-	1
	F	-	-	-	-	1	9	-	-	-	-	-	-	-	-	-	1
	G	-	-	-	-	-	-	10	-	-	-	-	-	-	-	-	-
	H	-	-	-	-	-	-	-	9	-	-	-	-	-	-	1	1
	I	-	-	-	-	-	-	-	3	7	-	-	-	-	-	-	3
	J	-	-	-	-	-	-	-	-	-	5	-	-	-	-	5	5
	K	-	-	-	-	-	-	-	-	-	-	9	-	-	-	1	1
	L	-	-	-	-	-	-	-	-	-	-	-	9	-	-	1	1
	M	-	-	-	-	-	-	-	-	-	-	-	-	7	-	3	3
	N	-	-	-	-	-	-	-	-	-	-	-	-	-	10	-	-
	Z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	FN	-	-	-	-	1	-	-	5	-	-	-	-	-	-	11	87.86%

### 4.2.3 Simulação 3

Na terceira rodada de teste, observou-se pela Tabela 6 que nenhum gesto executado foi confundido com outro gesto, e apenas três gestos foram classificados como inconclusivo, tendo um desempenho médio de 95,71% de acerto, sendo que houveram doze classes com 100% de acerto (A, B, C, D, E, F, H, I, J, K, L e N), uma classe com 80% de acerto (G), uma classe com 60% de acerto (M).

Nesta etapa, é possível notar que as classes de piores resultados são as classes J e M, com 66,66% e 66,66%, respectivamente, de acerto individuais de um total de 15 amostras para cada uma. Além disto, nota-se que o desempenho geral dos testes é 90,48% de um total de 210 amostras avaliadas até esta simulação. Um ponto interessante a se ressaltar é que após esta simulação a classe I não se encontra mais entre os resultados ruins, conforme pode ser visto na Tabela 7.

Tabela 6 – Matriz de confusão com o resultado da simulação 3, 95,71% de acerto.

		Classe Previstas															FP	
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	Z		
Classes Desejada	A	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	B	-	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	C	-	-	5	-	-	-	-	-	-	-	-	-	-	-	-	-	
	D	-	-	-	5	-	-	-	-	-	-	-	-	-	-	-	-	
	E	-	-	-	-	5	-	-	-	-	-	-	-	-	-	-	-	
	F	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-	-	
	G	-	-	-	-	-	-	4	-	-	-	-	-	-	-	-	1	1
	H	-	-	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-
	I	-	-	-	-	-	-	-	-	5	-	-	-	-	-	-	-	-
	J	-	-	-	-	-	-	-	-	-	5	-	-	-	-	-	-	-
	K	-	-	-	-	-	-	-	-	-	-	5	-	-	-	-	-	-
	L	-	-	-	-	-	-	-	-	-	-	-	5	-	-	-	-	-
	M	-	-	-	-	-	-	-	-	-	-	-	-	3	-	2	2	-
	N	-	-	-	-	-	-	-	-	-	-	-	-	-	5	-	-	-
	Z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	FN	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3	95.71%

Tabela 7 – Matriz de confusão com o resultado agregado após a simulação 3, 90,48% de acerto.

		Classe Previstas															FP	
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	Z		
Classes Desejada	A	15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	B	-	15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	C	-	-	15	-	-	-	-	-	-	-	-	-	-	-	-	-	
	D	-	-	-	14	-	-	-	1	-	-	-	-	-	-	-	-	1
	E	-	-	-	-	14	-	-	1	-	-	-	-	-	-	-	-	1
	F	-	-	-	-	1	14	-	-	-	-	-	-	-	-	-	-	1
	G	-	-	-	-	-	-	14	-	-	-	-	-	-	-	-	1	1
	H	-	-	-	-	-	-	-	14	-	-	-	-	-	-	-	1	1
	I	-	-	-	-	-	-	-	3	12	-	-	-	-	-	-	-	3
	J	-	-	-	-	-	-	-	-	-	10	-	-	-	-	-	5	5
	K	-	-	-	-	-	-	-	-	-	-	14	-	-	-	-	1	1
	L	-	-	-	-	-	-	-	-	-	-	-	14	-	-	1	1	
	M	-	-	-	-	-	-	-	-	-	-	-	-	10	-	5	5	
	N	-	-	-	-	-	-	-	-	-	-	-	-	-	15	-	-	
	Z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	FN	-	-	-	-	1	-	-	5	-	-	-	-	-	-	-	14	90.48%

#### 4.2.4 Simulação 4

Na rodada seguinte, observa-se pela Tabela 8 que houveram três gestos classificados erroneamente e quatro gestos classificados como inconclusivo. Nesta simulação, foi obtido um desempenho de 90% de acerto, havendo nove classes com 100% de acerto (A, B, C, D, E, G, J, L e N), três com 80% de acerto (F, K e M), duas classes com 60% de acerto (H e I).

Já o resultado agregado para esta etapa, Tabela 9, indica um desempenho geral de 90,36% de um total de 280 amostras executadas até esta simulação, onde as piores classes são as I, J e M, com 75%, 75% e 70%, respectivamente, de acerto individuais de um total de 20 amostras para cada uma. Com os resultados desta simulação, nota-se que a classe I voltou a ficar entre as piores classes.

Tabela 8 – Matriz de confusão com o resultado da simulação 4, 90,00% de acerto.

		Classe Previstas															FP		
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	Z			
Classes Desejada	A	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	B	-	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	C	-	-	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	D	-	-	-	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	E	-	-	-	-	5	-	-	-	-	-	-	-	-	-	-	-	-	-
	F	-	-	-	-	-	4	-	-	-	-	-	-	-	-	-	1	1	-
	G	-	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-	-	-
	H	-	-	-	-	-	-	2	3	-	-	-	-	-	-	-	-	-	2
	I	-	-	-	-	-	-	-	1	3	-	-	-	-	-	-	1	2	-
	J	-	-	-	-	-	-	-	-	-	5	-	-	-	-	-	-	-	-
	K	-	-	-	-	-	-	-	-	-	-	4	-	-	-	-	1	1	-
	L	-	-	-	-	-	-	-	-	-	-	-	5	-	-	-	-	-	-
	M	-	-	-	-	-	-	-	-	-	-	-	-	4	-	1	1	-	-
	N	-	-	-	-	-	-	-	-	-	-	-	-	-	5	-	-	-	-
	Z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	FN	-	-	-	-	-	-	2	1	-	-	-	-	-	-	-	4	90,00%	-

Tabela 9 – Matriz de confusão com o resultado agregado após a simulação 4, 90,36% de acerto.

		Classe Previstas															
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	Z	FP
Classes Desejada	A	20	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	B	-	20	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	C	-	-	20	-	-	-	-	-	-	-	-	-	-	-	-	-
	D	-	-	-	19	-	-	-	1	-	-	-	-	-	-	-	1
	E	-	-	-	-	19	-	-	1	-	-	-	-	-	-	-	1
	F	-	-	-	-	1	18	-	-	-	-	-	-	-	-	1	2
	G	-	-	-	-	-	-	19	-	-	-	-	-	-	-	1	1
	H	-	-	-	-	-	-	2	17	-	-	-	-	-	-	1	3
	I	-	-	-	-	-	-	-	4	15	-	-	-	-	-	1	5
	J	-	-	-	-	-	-	-	-	-	15	-	-	-	-	5	5
	K	-	-	-	-	-	-	-	-	-	-	18	-	-	-	2	2
	L	-	-	-	-	-	-	-	-	-	-	-	19	-	-	1	1
	M	-	-	-	-	-	-	-	-	-	-	-	-	14	-	6	6
	N	-	-	-	-	-	-	-	-	-	-	-	-	-	20	-	-
	Z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	FN	-	-	-	-	1	-	2	6	-	-	-	-	-	-	18	90.36%

#### 4.2.5 Simulação 5

Prosseguindo com a simulação, observa-se pela Tabela 10 que houve apenas dois gestos classificados como inconclusivo, enquanto houveram doze classes com 100% de acerto (A, B, C, D, F, G, H, I, K, L, M e N) e duas classes com 80% de acerto (E e J) obtendo um desempenho de 97,14% de acerto.

Pela Tabela 11, nota-se que o desempenho geral dos testes foi de 91,71% de um total de 350 amostras executadas até esta simulação. Após esta simulação, analisando os resultados agregados, as piores classes são novamente a J e M com 76% e 76%, respectivamente, de acerto individuais de um total de 25 amostras para cada uma, mostrando que a classe I está oscilando o número de acerto para cada simulação já feita.

Tabela 10 – Matriz de confusão com o resultado da simulação 5, 97,14% de acerto.

		Classe Previstas															FP
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	Z	
Classes Desejada	A	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	B	-	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	C	-	-	5	-	-	-	-	-	-	-	-	-	-	-	-	-
	D	-	-	-	5	-	-	-	-	-	-	-	-	-	-	-	-
	E	-	-	-	-	4	-	-	-	-	-	-	-	-	-	1	1
	F	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-	-
	G	-	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-
	H	-	-	-	-	-	-	-	5	-	-	-	-	-	-	-	-
	I	-	-	-	-	-	-	-	-	5	-	-	-	-	-	-	-
	J	-	-	-	-	-	-	-	-	-	4	-	-	-	-	1	1
	K	-	-	-	-	-	-	-	-	-	-	5	-	-	-	-	-
	L	-	-	-	-	-	-	-	-	-	-	-	5	-	-	-	-
	M	-	-	-	-	-	-	-	-	-	-	-	-	5	-	-	-
	N	-	-	-	-	-	-	-	-	-	-	-	-	-	5	-	-
	Z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	FN	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	97.14%

Tabela 11 – Matriz de confusão com o resultado agregado após a simulação 5, 91,71% de acerto.

		Classe Previstas															FP
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	Z	
Classes Desejada	A	25	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	B	-	25	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	C	-	-	25	-	-	-	-	-	-	-	-	-	-	-	-	-
	D	-	-	-	24	-	-	-	1	-	-	-	-	-	-	-	1
	E	-	-	-	-	23	-	-	1	-	-	-	-	-	-	1	2
	F	-	-	-	-	1	23	-	-	-	-	-	-	-	-	1	2
	G	-	-	-	-	-	-	24	-	-	-	-	-	-	-	1	1
	H	-	-	-	-	-	-	2	22	-	-	-	-	-	-	1	3
	I	-	-	-	-	-	-	-	4	20	-	-	-	-	-	1	5
	J	-	-	-	-	-	-	-	-	-	19	-	-	-	-	6	6
	K	-	-	-	-	-	-	-	-	-	-	23	-	-	-	2	2
	L	-	-	-	-	-	-	-	-	-	-	-	24	-	-	1	1
	M	-	-	-	-	-	-	-	-	-	-	-	-	19	-	6	6
	N	-	-	-	-	-	-	-	-	-	-	-	-	-	25	-	-
	Z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	FN	-	-	-	-	1	-	2	6	-	-	-	-	-	-	-	20



### 4.2.6 Simulação 6

Na sexta simulação, foi obtido 92,86% de acerto, sendo este constituído por dez classes com 100% de acerto (B, D, E, F, G, J, K, L, M e N), três classes com 80% de acerto (A, H e I), uma classe com 60% de acerto (C) e cinco amostras classificadas como inconclusivo, sendo estes resultados apresentados na Tabela 12.

O resultado agregado após esta simulação é apresentado na Tabela 13, com o desempenho geral dos testes de 91,90% de um total de 420 amostras executadas até esta simulação. Com os resultados desta simulação, observou-se que os resultados agregados de todas as classes tiveram desempenho individuais acima de 80% de acerto, de um total de 30 amostras para cada uma, indicando que a base de dados alcançou um número satisfatório de elementos para todas as classes.

Tabela 12 – Matriz de confusão com o resultado da simulação 6, 92,86% de acerto.

		Classe Previstas															FP	
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	Z		
Classes Desejada	A	4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1
	B	-	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	C	-	-	3	-	-	-	-	-	-	-	-	-	-	-	-	2	2
	D	-	-	-	5	-	-	-	-	-	-	-	-	-	-	-	-	-
	E	-	-	-	-	5	-	-	-	-	-	-	-	-	-	-	-	-
	F	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-	-	-
	G	-	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-	-
	H	-	-	-	-	-	-	-	4	-	-	-	-	-	-	-	1	1
	I	-	-	-	-	-	-	-	-	4	-	-	-	-	-	-	1	1
	J	-	-	-	-	-	-	-	-	-	5	-	-	-	-	-	-	-
	K	-	-	-	-	-	-	-	-	-	-	5	-	-	-	-	-	-
	L	-	-	-	-	-	-	-	-	-	-	-	5	-	-	-	-	-
	M	-	-	-	-	-	-	-	-	-	-	-	-	5	-	-	-	-
	N	-	-	-	-	-	-	-	-	-	-	-	-	-	5	-	-	-
	Z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	FN	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	5	92.86%

Tabela 13 – Matriz de confusão com o resultado agregado após a simulação 6, 91,90% de acerto.

		Classe Previstas															FP
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	Z	
Classes Desejada	A	29	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1
	B	-	30	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	C	-	-	28	-	-	-	-	-	-	-	-	-	-	-	2	2
	D	-	-	-	29	-	-	-	1	-	-	-	-	-	-	-	1
	E	-	-	-	-	28	-	-	1	-	-	-	-	-	-	1	2
	F	-	-	-	-	1	28	-	-	-	-	-	-	-	-	1	2
	G	-	-	-	-	-	-	29	-	-	-	-	-	-	-	1	1
	H	-	-	-	-	-	-	2	26	-	-	-	-	-	-	2	4
	I	-	-	-	-	-	-	-	4	24	-	-	-	-	-	2	6
	J	-	-	-	-	-	-	-	-	-	24	-	-	-	-	6	6
	K	-	-	-	-	-	-	-	-	-	-	28	-	-	-	2	2
	L	-	-	-	-	-	-	-	-	-	-	-	29	-	-	1	1
	M	-	-	-	-	-	-	-	-	-	-	-	-	24	-	6	6
	N	-	-	-	-	-	-	-	-	-	-	-	-	-	30	-	-
	Z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	FN	-	-	-	-	1	-	2	6	-	-	-	-	-	-	25	91.90%

#### 4.2.7 Simulação 7

A simulação seguinte obteve 97,14% de acerto, havendo doze classes com 100% de acerto (A, B, C, D, E, F, G, J, K, L, M e N), duas classes com 80% de acerto (H e I) e duas amostras classificadas como inconclusivo, resultados, este, apresentado na Tabela 14.

Nesta etapa, observa-se que novamente os resultados agregados de todas as classes tiveram desempenho individuais acima de 80% de acerto, de um total de 35 amostras para cada uma. O resultado agregado após esta simulação é apresentado na Tabela 15, com o desempenho geral dos testes de 92,65% de um total de 490 amostras executadas até esta simulação. Isto revela que com a base de dados atual, já é possível obter bons resultados para algumas classes e ótimos resultados para outras.

Tabela 14 – Matriz de confusão com o resultado da simulação 7, 97,14% de acerto.

		Classe Previstas															FP
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	Z	
Classes Desejada	A	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	B	-	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	C	-	-	5	-	-	-	-	-	-	-	-	-	-	-	-	-
	D	-	-	-	5	-	-	-	-	-	-	-	-	-	-	-	-
	E	-	-	-	-	5	-	-	-	-	-	-	-	-	-	-	-
	F	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-	-
	G	-	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-
	H	-	-	-	-	-	-	-	4	-	-	-	-	-	-	1	1
	I	-	-	-	-	-	-	-	-	4	-	-	-	-	-	1	1
	J	-	-	-	-	-	-	-	-	-	5	-	-	-	-	-	-
	K	-	-	-	-	-	-	-	-	-	-	5	-	-	-	-	-
	L	-	-	-	-	-	-	-	-	-	-	-	5	-	-	-	-
	M	-	-	-	-	-	-	-	-	-	-	-	-	5	-	-	-
	N	-	-	-	-	-	-	-	-	-	-	-	-	-	5	-	-
	Z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	FN	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	97.14%

Tabela 15 – Matriz de confusão com o resultado agregado após a simulação 7, 92,65% de acerto.

		Classe Previstas															FP
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	Z	
Classes Desejada	A	34	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1
	B	-	35	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	C	-	-	33	-	-	-	-	-	-	-	-	-	-	-	2	2
	D	-	-	-	34	-	-	-	1	-	-	-	-	-	-	-	1
	E	-	-	-	-	33	-	-	1	-	-	-	-	-	-	1	2
	F	-	-	-	-	1	33	-	-	-	-	-	-	-	-	1	2
	G	-	-	-	-	-	-	34	-	-	-	-	-	-	-	1	1
	H	-	-	-	-	-	-	2	30	-	-	-	-	-	-	3	5
	I	-	-	-	-	-	-	-	4	28	-	-	-	-	-	3	7
	J	-	-	-	-	-	-	-	-	-	29	-	-	-	-	6	6
	K	-	-	-	-	-	-	-	-	-	-	33	-	-	-	2	2
	L	-	-	-	-	-	-	-	-	-	-	-	34	-	-	1	1
	M	-	-	-	-	-	-	-	-	-	-	-	-	29	-	6	6
	N	-	-	-	-	-	-	-	-	-	-	-	-	-	35	-	-
	Z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	FN	-	-	-	-	1	-	2	6	-	-	-	-	-	-	27	92.65%

### 4.2.8 Simulação 8

A Tabela 16 apresenta os resultados da antepenúltima simulação, onde foi obtido 90,00% de acerto, sendo dez classes com 100% de acerto (A, B, C, E, F, G, H, K, L e N), duas classes com 80% de acerto (D e M), uma classe com 60%, uma classe com 40%, duas amostras foram confundidas e cinco amostras classificadas como inconclusivo.

Aqui, observa-se que devido às amostras confundidas e as inconclusivas, as classes I e J, caíram de rendimento, indo para 77,5% de acerto individuais, ambas, de um total de 40 amostras para cada uma. O resultado agregado após esta simulação é apresentado na Tabela 17, com o desempenho geral dos testes de 92,32% de um total de 560 amostras executadas até esta simulação. Isto revela que com a base de dados atual, já é possível obter bons resultados (acima de 75%) para algumas classes e ótimos resultados (acima de 90%) para outras.

Tabela 16 – Matriz de confusão com o resultado da simulação 8, 90,00% de acerto.

		Classe Previstas															
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	Z	FP
Classes Desejada	A	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	B	-	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	C	-	-	5	-	-	-	-	-	-	-	-	-	-	-	-	-
	D	-	-	-	4	-	-	-	-	-	-	-	-	-	-	1	1
	E	-	-	-	-	5	-	-	-	-	-	-	-	-	-	-	-
	F	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-	-
	G	-	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-
	H	-	-	-	-	-	-	-	5	-	-	-	-	-	-	-	-
	I	-	-	-	-	-	-	-	-	3	-	-	-	-	-	2	2
	J	-	2	-	-	-	-	-	-	-	2	-	-	-	-	1	3
	K	-	-	-	-	-	-	-	-	-	-	5	-	-	-	-	-
	L	-	-	-	-	-	-	-	-	-	-	-	5	-	-	-	-
	M	-	-	-	-	-	-	-	-	-	-	-	-	4	-	1	1
	N	-	-	-	-	-	-	-	-	-	-	-	-	-	5	-	-
	Z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	FN	-	2	-	-	-	-	-	-	-	-	-	-	-	-	5	90,00%

Tabela 17 – Matriz de confusão com o resultado agregado após a simulação 8, 92,32% de acerto.

		Classe Previstas															FP	
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	Z		
Classes Desejada	A	39	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1
	B	-	40	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	C	-	-	38	-	-	-	-	-	-	-	-	-	-	-	-	2	2
	D	-	-	-	38	-	-	-	1	-	-	-	-	-	-	-	1	2
	E	-	-	-	-	38	-	-	1	-	-	-	-	-	-	-	1	2
	F	-	-	-	-	1	38	-	-	-	-	-	-	-	-	-	1	2
	G	-	-	-	-	-	-	39	-	-	-	-	-	-	-	-	-	1
	H	-	-	-	-	-	-	2	35	-	-	-	-	-	-	-	3	5
	I	-	-	-	-	-	-	-	4	31	-	-	-	-	-	-	5	9
	J	-	2	-	-	-	-	-	-	-	31	-	-	-	-	-	7	9
	K	-	-	-	-	-	-	-	-	-	-	38	-	-	-	-	2	2
	L	-	-	-	-	-	-	-	-	-	-	-	39	-	-	-	1	1
	M	-	-	-	-	-	-	-	-	-	-	-	-	33	-	7	7	7
	N	-	-	-	-	-	-	-	-	-	-	-	-	-	40	-	-	-
	Z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	FN	-	2	-	-	1	-	2	6	-	-	-	-	-	-	-	32	92.32%

#### 4.2.9 Simulação 9

A penúltima simulação apresentou um desempenho médio de 91,43% de acerto. Isto é, houveram onze classes com 100% de acerto (A, C, D, E, F, G, H, K, L, M e N), uma classe com 80% de acerto (J), uma classe com 60% (B), uma classe com 40% (I), seis amostras inconclusivas. A Tabela 18 apresenta este resultado.

A Tabela 19 apresenta o resultado agregado do teste de validação, no qual as classes de piores resultados, nesta etapa, são as classes I e J, com 73,33% e 77,77%, respectivamente, de acerto individuais de um total de 45 amostras para cada uma. Além disto, nota-se que o desempenho geral dos testes é 92,22% de um total de 630 amostras avaliadas até esta simulação.

Tabela 18 – Matriz de confusão com o resultado da simulação 6, 91,43% de acerto.

		Classe Previstas															FP	
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	Z		
Classes Desejada	A	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	B	-	3	-	-	-	-	-	-	-	-	-	-	-	-	-	2	2
	C	-	-	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	D	-	-	-	5	-	-	-	-	-	-	-	-	-	-	-	-	-
	E	-	-	-	-	5	-	-	-	-	-	-	-	-	-	-	-	-
	F	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-	-	-
	G	-	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-	-
	H	-	-	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-
	I	-	-	-	-	-	-	-	-	2	-	-	-	-	-	-	3	3
	J	-	-	-	-	-	-	-	-	-	4	-	-	-	-	-	1	1
	K	-	-	-	-	-	-	-	-	-	-	5	-	-	-	-	-	-
	L	-	-	-	-	-	-	-	-	-	-	-	5	-	-	-	-	-
	M	-	-	-	-	-	-	-	-	-	-	-	-	5	-	-	-	-
	N	-	-	-	-	-	-	-	-	-	-	-	-	-	5	-	-	-
	Z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	FN	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	6	91.43%

Tabela 19 – Matriz de confusão com o resultado agregado após a simulação 6, 92,22% de acerto.

		Classe Previstas															FP	
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	Z		
Classes Desejada	A	44	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1
	B	-	43	-	-	-	-	-	-	-	-	-	-	-	-	-	2	2
	C	-	-	43	-	-	-	-	-	-	-	-	-	-	-	-	2	2
	D	-	-	-	43	-	-	-	1	-	-	-	-	-	-	-	1	2
	E	-	-	-	-	43	-	-	1	-	-	-	-	-	-	-	1	2
	F	-	-	-	-	1	43	-	-	-	-	-	-	-	-	-	1	2
	G	-	-	-	-	-	-	44	-	-	-	-	-	-	-	-	1	1
	H	-	-	-	-	-	-	2	40	-	-	-	-	-	-	-	3	5
	I	-	-	-	-	-	-	-	4	33	-	-	-	-	-	-	8	12
	J	-	2	-	-	-	-	-	-	-	35	-	-	-	-	-	8	10
	K	-	-	-	-	-	-	-	-	-	-	43	-	-	-	-	2	2
	L	-	-	-	-	-	-	-	-	-	-	-	44	-	-	-	1	1
	M	-	-	-	-	-	-	-	-	-	-	-	-	38	-	-	7	7
	N	-	-	-	-	-	-	-	-	-	-	-	-	-	45	-	-	-
	Z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	FN	-	2	-	-	1	-	2	6	-	-	-	-	-	-	-	38	92.22%

### 4.2.10 Simulação 10

A última simulação obteve um desempenho médio de 95,71% de acerto. A Tabela 20 apresenta este resultado em que houveram onze classes com 100% de acerto (A, B, D, E, G, H, J, K, L, M e N), três classes com 80% de acerto (C, F e I) e três amostras inconclusivas.

Por fim, analisando-se o resultado geral agregado, Tabela 21, observa-se que o desempenho geral obtido para 700 amostras foi de 92,57%, em que a pior classe obtida foi a I com 74% de acerto individual de um total de 50 amostras.

Tabela 20 – Matriz de confusão com o resultado da simulação 10, 95,71% de acerto.

		Classe Previstas																
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	Z	FP	
Classes Desejada	A	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	B	-	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	C	-	-	4	-	-	-	-	-	-	-	-	-	-	-	-	1	1
	D	-	-	-	5	-	-	-	-	-	-	-	-	-	-	-	-	-
	E	-	-	-	-	5	-	-	-	-	-	-	-	-	-	-	-	-
	F	-	-	-	-	-	4	-	-	-	-	-	-	-	-	-	1	1
	G	-	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-	-
	H	-	-	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-
	I	-	-	-	-	-	-	-	-	4	-	-	-	-	-	-	1	1
	J	-	-	-	-	-	-	-	-	-	5	-	-	-	-	-	-	-
	K	-	-	-	-	-	-	-	-	-	-	5	-	-	-	-	-	-
	L	-	-	-	-	-	-	-	-	-	-	-	5	-	-	-	-	-
	M	-	-	-	-	-	-	-	-	-	-	-	-	5	-	-	-	-
	N	-	-	-	-	-	-	-	-	-	-	-	-	-	5	-	-	-
	Z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	FN	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3	95.71%

Tabela 21 – Matriz de confusão com o resultado agregado após a simulação 10, 91,90% de acerto.

		Classe Previstas															FP	
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	Z		
Classes Desejada	A	49	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1
	B	-	48	-	-	-	-	-	-	-	-	-	-	-	-	-	2	2
	C	-	-	47	-	-	-	-	-	-	-	-	-	-	-	-	3	3
	D	-	-	-	48	-	-	-	1	-	-	-	-	-	-	-	1	2
	E	-	-	-	-	48	-	-	1	-	-	-	-	-	-	-	1	2
	F	-	-	-	-	1	47	-	-	-	-	-	-	-	-	-	2	3
	G	-	-	-	-	-	-	49	-	-	-	-	-	-	-	-	1	1
	H	-	-	-	-	-	-	2	45	-	-	-	-	-	-	-	3	5
	I	-	-	-	-	-	-	-	4	37	-	-	-	-	-	-	9	13
	J	-	2	-	-	-	-	-	-	-	40	-	-	-	-	-	8	10
	K	-	-	-	-	-	-	-	-	-	-	48	-	-	-	-	2	2
	L	-	-	-	-	-	-	-	-	-	-	-	49	-	-	-	1	1
	M	-	-	-	-	-	-	-	-	-	-	-	-	43	-	-	7	7
	N	-	-	-	-	-	-	-	-	-	-	-	-	-	50	-	-	-
	Z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	FN	-	2	-	-	1	-	2	6	-	-	-	-	-	-	-	41	92.57%

Ao realizar uma análise com mais detalhes sobre os resultados gerais, percebe-se que a maioria das classificações erradas se deve entre as classes que possuem componentes parecidos, por exemplo, G, H e I, nos quais, apenas a posição do braço direito varia entre eles. Já a maioria das classificações inconclusivas se deve principalmente ao fato do baixo valor estabelecido para o limiar ( $Th$ ), cuja função de evitar classificar possíveis gestos involuntários, errados ou indesejados foi cumprida. Ou seja, caso o valor de  $TH$  fosse maior, algumas amostras inconclusivas poderiam ter seu resultado alterado, podendo ser classificadas como corretas ou errada em vez de inconclusiva.

A Tabela 22 apresenta o resultado do classificador desenvolvido, em percentuais, de um total de 700 amostras classificadas, sendo 50 amostras de cada classe. Nesta tabela, cada linha indica a classe desejada, ou seja, a resposta correta para o gesto executado. Enquanto, cada coluna representa a classe classificada pela ferramenta de reconhecimento. A coluna  $FP$  (falso positivo) indica quantas amostras não foram classificadas corretamente e neste caso, a classe I, que possui o maior valor desta coluna, foi a que teve pior desempenho. Já linha  $FN$  (falso negativo) indica o quanto aquele gesto foi confundido com outro gesto, sendo que a classe H, apesar de ter tido um bom desempenho quando realizado o gesto H, algumas amostras foram classificadas como H quando não deveriam ser. O desempenho geral foi de 92,57%.

A Tabela 23 apresenta a evolução das principais informações acerca da base de dados, do resultado de cada simulação e do resultado agregado após cada simulação. Ao término da validação, a base de dados que iniciou com 70 amostras terminou com 718



Tabela 22 – Resultado geral da validação, 92,57% de acerto

	Classe Previstas															
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	Z	FP
Classes Desejada	A	98	-	-	-	-	-	-	-	-	-	-	-	-	2	2
	B	-	96	-	-	-	-	-	-	-	-	-	-	-	4	4
	C	-	-	94	-	-	-	-	-	-	-	-	-	-	6	6
	D	-	-	-	96	-	-	-	2	-	-	-	-	-	2	4
	E	-	-	-	-	96	-	-	2	-	-	-	-	-	2	4
	F	-	-	-	-	2	94	-	-	-	-	-	-	-	4	6
	G	-	-	-	-	-	-	98	-	-	-	-	-	-	2	2
	H	-	-	-	-	-	-	4	90	-	-	-	-	-	6	10
	I	-	-	-	-	-	-	-	8	74	-	-	-	-	18	26
	J	-	4	-	-	-	-	-	-	-	80	-	-	-	16	20
	K	-	-	-	-	-	-	-	-	-	-	96	-	-	4	4
	L	-	-	-	-	-	-	-	-	-	-	-	98	-	2	2
	M	-	-	-	-	-	-	-	-	-	-	-	-	86	14	14
	N	-	-	-	-	-	-	-	-	-	-	-	-	-	100	-
	Z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	FN	-	4	-	-	2,04	-	3,92	11,77	-	-	-	-	-	-	92,57

elementos ao todo (70 amostras iniciais mais as 648 amostras classificadas corretamente).

Tabela 23 – Resumo dos resultados de cada simulação executada no trabalho.

Simulação	Tamanho da base de dados antes da simulação	Resultado da simulação		Resultado agregado do teste	
		Acertos	Taxa de acerto (%)	Acertos	Taxa de acerto (%)
1	70	62	88,57	62	88,57
2	132	61	87,14	123	87,86
3	193	67	95,71	190	90,48
4	260	63	90,00	253	90,36
5	323	68	97,14	321	91,71
6	391	65	92,86	386	91,90
7	456	68	97,14	454	92,65
8	524	63	90,00	517	92,32
9	587	64	91,43	581	92,22
10	651	67	95,71	648	92,57
<b>Total após a última simulação</b>	<b>718</b>	<b>648</b>	-	-	-

## 5 Considerações finais

Neste trabalho foi proposto um método de classificação de gesto utilizando Distância Euclidiana baseado no rastreamento de pulseiras com marcações reflexivas, ao se analisar os dados pré-processados obtidos através da matriz de características da PCA, dos dados coletadas por meio do OptiTrack.

Ao observar as informações relacionadas as posições dos marcadores, foi possível concluir que, apesar de o baixo número de marcadores utilizados, a proximidade entre eles e apenas 3 “corpos rígidos”, a serem rastreados, o OptiTrack se mostrou ter um bom desempenho na coleta de dados ao não haver muitas perdas de referências e troca de posições entre os “corpos rígidos”.

Outro ponto importante observado, foi que a posição das câmeras teve grande contribuição na execução das simulações. Isto é, ao se considerar que as ambas as pulseiras e o crachá ficam apontados para direções diferentes e, que a pulseira pode ser apontada para uma nova direção durante a ação, foi necessário que todo o ambiente pudesse ser visualizado, e isto foi possível devido à configuração do ambiente de simulação.

A análise de componentes principais se mostrou também ser muito eficaz e prática, ao conseguir reduzir diversos conjuntos de informações, os quais possuíam tamanhos variados, facilitando a análise e comparações das informações. Com isto, evitando que o tempo de classificação aumentasse muito, a medida que a base de dados foi aumentando.

O classificador apresentou uma taxa de acerto média de 92,57% o qual pode ser considerado tendo um excelente desempenho, ao se considerar a simplicidade da técnica usada, o tamanho da base, e ao comparar com outros trabalhos como, por exemplo, em (CARVALHO et al., 2022) onde foi obtido uma taxa de acerto superior a 96%, e em (CARVALHO; BASÍLIO; BRANDÃO, 2022) no qual foram obtidos resultados acima de 96,7% de acerto, ambas utilizando ANNs.

As classes de melhor desempenho foram as N, L, G, A, K, E, D e B, todas acima de 95% de acerto, totalizando 8. Já as classes de pior desempenho foram as I, J e M, todas abaixo de 90%, totalizando 3. A classe inconclusiva se provou ser um recurso muito útil, para reduzir a quantidade de erros, ao se considerar momentos onde o usuário levante o punho sem querer ou para fazer alguma outra ação. Dessa forma, movimentos diferentes dos pré-estabelecidos serão classificados como inconclusivos, reforçando a viabilidade e a eficiência, do mesmo, devido à simplicidade e do pequeno banco de dados construído.

Outro ponto a se concluir é que com esta ferramenta desenvolvida é possível que outros usuários, com diferentes físicos (braços mais longos ou curtos), possam utilizá-las.

Apesar do tamanho dos braços diferirem, ao aplicar a técnica da PCA, foi possível extrair o padrão do movimento do gesto. Dessa forma, sendo necessário que o outro usuário faça algumas sequências de todos os gestos para atualizar a base de dados, de modo a aumentar a taxa de acerto do novo usuário.

Por fim, outra conclusão a ser comentada é que, com o bom desempenho do classificador, ao associar as classes com tarefas de alto nível (por exemplo, ações para um veículo autônomo), espera-se que todas as classes, com exceção da I, J e M, sejam classificadas corretamente, demonstrando que o trabalho alcançou o objetivo de ter um baixo custo computacional e um alto desempenho para a maioria das classes.

Para trabalhos futuros, pretende-se criar um padrão ouro, ao escolher uma série de amostras de cada classe como referência, para ser usada como comparação, avaliando a média e o desvio padrão de cada amostra (curva gaussiana). Outro ponto a se abordar em trabalhos futuros é a implementação de um limiar adaptativo, onde o valor de  $Th$  diminuiria à medida que um gesto fosse executado corretamente e adicionado ao banco de dados. Além destes, o processo de treinamento precisa ser melhorado. Isto pode ser feito aplicando a validação cruzada ao banco de dados e com vários conjuntos de amostras em uma ordem aleatória.

# Referências

- ALCANTARA, F. C. de; FERREIRA, L. F.; SANTOS, H. S. dos. Avaliação do knn para reconhecimento de placas veiculares no modelo brasileiro. *Perspectiv@s*, v. 15, n. 15, p. 6–9, 2019. 16
- ALVES, W. F. d. O. A. et al. Interação com o ambiente mediante classificação por distância euclidiana. *CBA-Congresso Brasileiro de Automatica*, 2022. 16, 19, 20
- ANDRADE, V. S. et al. Reconhecimento de faces utilizando padrão binário local e knn. *Revista Eletrônica de Sistemas de Informação e Gestão Tecnológica*, v. 11, n. 1, 2021. 16, 19
- BASILIO, V. T.; CARVALHO, K. B.; BRANDAO, A. Reconhecimento de ações por rna em aplicações de robótica social. *XIV Simpósio Brasileiro de Automação Inteligente, Ouro Preto/MG*, Campinas, 2019. Disponível em: <<https://proceedings.science/sbai-2019/papers/reconhecimento-de-acoes-por-rna-em-aplicacoes-de-robotica-social?lang=pt-br>>. 15, 18
- BRAGA, B. M. Deep learning para classificação e processamento de imagens. Universidade Estadual Paulista (Unesp), 2022. 16, 20
- CAMPOS, F. R. *A robótica para uso educacional*. [S.l.]: Editora Senac São Paulo, 2019. 14
- CAMPOS, G. A.; MOUTINHO, L. d. S. *DEEP: uma arquitetura para reconhecer emoção com base no espectro sonoro da voz de falantes da língua portuguesa*. 57 f. Monografia (Bacharelado em Ciência da Computação) — Universidade de Brasília, Brasília, 2020. Disponível em: <<https://bdm.unb.br/handle/10483/27583>>. 14
- CARVALHO, K. B. de; BASÍLIO, V. T.; BRANDÃO, A. S. Action recognition for educational proposals applying concepts of social assistive robotics. *Cognitive Systems Research*, Elsevier, v. 71, p. 1–8, 2022. 15, 19, 49
- CARVALHO, K. B. de et al. Gestures-teleoperation of a heterogeneous multi-robot system. *The International Journal of Advanced Manufacturing Technology*, Springer, v. 118, n. 5, p. 1999–2015, 2022. 15, 23, 25, 49
- CASTRO, E. M.; SAKATA, T. C.; ZAINA, L. A. Explorando o potencial da interação nui em um jogo de pensamento computacional para crianças. *Revista Brasileira de Informática na Educação*, v. 27, n. 02, p. 140, 2019. ISSN 2317-6121. Disponível em: <<http://ojs.sector3.com.br/index.php/rbie/article/view/v27n02140166>>. 14
- CHANDARANA, M. et al. A natural interaction interface for uavs using intuitive gesture recognition. In: SAVAGE-KNEPSHIELD, P.; CHEN, J. (Ed.). *Advances in Human Factors in Robots and Unmanned Systems*. Cham: Springer International Publishing, 2017. p. 387–398. ISBN 978-3-319-41959-6. 20
- CHEN, F. et al. Wristcam: A wearable sensor for hand trajectory gesture recognition and intelligent human–robot interaction. *IEEE Sensors Journal*, v. 19, n. 19, p. 8441–8451, 2019. 14

- DHARMALINGAM, S.; PALANISAMY, A. Vector space based augmented structural kinematic feature descriptor for human activity recognition in videos. *Etri Journal*, Wiley Online Library, v. 40, n. 4, p. 499–510, 2018. 16, 19
- FERIGATO, A.; LOUREIRO, M. A. Gesto corporal e conteúdo expressivo na construção da performance musical de alunos de harpa. In: *XXVIII Congresso da Associação Nacional de Pesquisa e Pós-Graduação em Música–Manaus*. [S.l.: s.n.], 2018. p. 1–7. 16
- FONSECA, M. G. B. d. *Redes neurais artificiais aplicadas à classificação de gestos da mão através de sinais eletromiográficos*. Dissertação (Mestrado em Engenharia Elétrica) — UNIVERSIDADE FEDERAL DA BAHIA, 2019. 16, 19, 20
- FURTADO, J. S. et al. Comparative analysis of optitrack motion capture systems. In: *Advances in Motion Sensing and Control for Robotic Applications*. [S.l.]: Springer, 2019. p. 15–31. 19
- GORDILLO, C. D. A. *Realce e Reconhecimento de Voz Contínua em Ambientes Adversos*. Tese (Doutorado) — PUC-Rio, 2018. 14
- GOULART, C. et al. Visual and thermal image processing for facial specific landmark detection to infer emotions in a child-robot interaction. *Sensors*, MDPI AG, v. 19, n. 13, p. 2844, Jun 2019. ISSN 1424-8220. Disponível em: <<http://dx.doi.org/10.3390/s19132844>>. 16
- HAYKIN, S. *Redes neurais: princípios e prática*. [S.l.]: Bookman Editora, 2001. 20
- HAZRA, S.; SANTRA, A. Robust gesture recognition using millimetric-wave radar system. *IEEE Sensors Letters*, v. 2, n. 4, p. 1–4, 2018. 14
- JÄHNE, B.; HAUSSECKER, H.; GEISSLER, P. *Handbook of computer vision and applications*. [S.l.]: Citeseer, 1999. v. 2. 18
- JÚNIOR, S. et al. Um modelo de visão computacional para análise biomecânica do valgismo dinâmico. Universidade Federal de Alagoas, 2020. 18
- KURITA, T. Principal component analysis (pca). In: \_\_\_\_\_. *Computer Vision: A Reference Guide*. Cham: Springer International Publishing, 2019. p. 1–4. ISBN 978-3-030-03243-2. Disponível em: <[https://doi.org/10.1007/978-3-030-03243-2\\_649-1](https://doi.org/10.1007/978-3-030-03243-2_649-1)>. 20
- LAHAW, Z. B.; ESSAIDANI, D.; SEDDIK, H. Robust face recognition approaches using pca, ica, lda based on dwt, and svm algorithms. In: *2018 41st International Conference on Telecommunications and Signal Processing (TSP)*. [S.l.: s.n.], 2018. p. 1–5. 16
- LEITE, A. de S. et al. Algoritmos de classificação aplicados ao controle interno em economicidade de combustíveis—um estudo de caso na polícia militar de pernambuco. *Revista de Engenharia e Pesquisa Aplicada*, v. 6, n. 5, p. 27–36, 2021. 20
- MARQUES, A. L.; CARRIJO, R. S.; MORAIS, A. S. de. Robótica colaborativa: Importância e desafios. 2020. ISSN 2596-2221. 15
- MARTINELLI, D. et al. *Sistema adaptativo para teleoperação de basemóvel através de reconhecimentos gestuais*. Dissertação (Mestrado em Engenharia Elétrica e Informática Industrial) — Universidade Tecnológica Federal do Paraná, Curitiba, 2022. Disponível em: <<http://repositorio.utfpr.edu.br/jspui/handle/1/28785>>. 14

- MOTIVE. *Motive Documentation - NaturalPoint Products*. [s.n.], 2020. Accessed: 2022-04-18. Disponível em: <[https://v22.wiki.optitrack.com/index.php?title=Motive\\_Documentation](https://v22.wiki.optitrack.com/index.php?title=Motive_Documentation)>. 18, 19, 22
- PORCELLIS, D.; SIEDLER, M.; HOLZSCHUH, M. Desenvolvimento e implantação de um jogo interativo voltado para crianças com distúrbios articulatorios. *Anais dos Workshops do Congresso Brasileiro de Informática na Educação*, v. 8, n. 1, p. 1409, 2019. ISSN 2316-8889. Disponível em: <<http://ojs.sector3.com.br/index.php/wcbie/article/view/9108>>. 14
- SANTOS, C. C. dos et al. Uso de transfer learning para o reconhecimento de gestos dinâmicos. In: *Congresso Brasileiro de Automática-CBA*. [S.l.: s.n.], 2019. v. 1, n. 1. 16
- SAWHNEY, S. et al. Real-time smart attendance system using face recognition techniques. In: *2019 9th International Conference on Cloud Computing, Data Science Engineering (Confluence)*. [S.l.: s.n.], 2019. p. 522–525. 16
- SCHULTE, J. et al. Autonomous human-vehicle leader-follower control using deep-learning-driven gesture recognition. *Vehicles*, v. 4, n. 1, p. 243–258, 2022. ISSN 2624-8921. Disponível em: <<https://www.mdpi.com/2624-8921/4/1/16>>. 20
- SHAHROUDY, A. et al. Ntu rgb+d: A large scale dataset for 3d human activity analysis. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.: s.n.], 2016. 19
- SHEENA, C. V.; NARAYANAN, N. K. Static gesture classification and recognition using hog feature parameters and k-nn and svm-based machine learning algorithms. In: BAJPAI, M. K.; SINGH, K. K.; GIAKOS, G. (Ed.). *Machine Vision and Augmented Intelligence—Theory and Applications*. Singapore: Springer Singapore, 2021. p. 157–166. ISBN 978-981-16-5078-9. 16, 19
- SOARES, V. N. et al. Análise das ações técnicas do goleiro de futebol: estudo preliminar. *RBFF-Revista Brasileira de Futsal e Futebol*, Instituto Brasileiro de Pesquisa e Ensino em Fisiologia do Exercício (IBPEFEX), v. 10, n. 38, p. 307–313, 2018. ISSN 1984-4956. 14
- SOUSA, A. L. de et al. Fourface: Uma ferramenta de reconhecimento de expressões faciais. *Brazilian Journal of Development*, v. 6, n. 10, p. 81667–81675, Oct. 2020. Disponível em: <<https://ojs.brazilianjournals.com.br/ojs/index.php/BRJD/article/view/18811>>. 14
- TEIXEIRA, F. J. M. *Interações gestuais para plataformas de streaming de vídeo*. Tese (Doutorado) — Universidade da Madeira (Portugal), 2021. 16
- TRUCCO, E.; VERRI, A. *Introductory techniques for 3-D computer vision*. [S.l.]: Prentice Hall Englewood Cliffs, 1998. v. 201. 18
- VITAL, J. et al. Tecnologias para a análise do movimento humano. In: R. Ruben, M. Vieira, C. Campos, H. Almeida, J. Siopa, P. Bártolo, & J. Folgado (Edits.), *6º Congresso Nacional de Biomecânica*. [S.l.: s.n.], 2015. p. 1–6. 19
- WIEDERER, J. et al. Traffic control gesture recognition for autonomous vehicles. In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. [S.l.: s.n.], 2020. p. 10676–10683. 14