

**UNIVERSIDADE ESTADUAL PAULISTA "JÚLIO DE MESQUITA FILHO"**

**FACULDADE DE CIÊNCIAS - CAMPUS BAURU**

**DEPARTAMENTO DE COMPUTAÇÃO**

**BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

**LUIS HENRIQUE MORELLI**

**SISTEMA DE RECONHECIMENTO BIOMÉTRICO DE MÚLTIPLAS  
PESSOAS BASEADO EM MÉTODOS DE ESTIMAÇÃO DE POSE E  
RASTREAMENTO DE OBJETOS**

**BAURU**

**JANEIRO/2023**

LUIS HENRIQUE MORELLI

**SISTEMA DE RECONHECIMENTO BIOMÉTRICO DE MÚLTIPLAS  
PESSOAS BASEADO EM MÉTODOS DE ESTIMAÇÃO DE POSE E  
RASTREAMENTO DE OBJETOS**

Trabalho de Conclusão de Curso do Curso  
de Bacharelado em Ciência da Computação  
da Universidade Estadual Paulista "Júlio  
de Mesquita Filho", Faculdade de Ciências,  
Campus Bauru.

Orientador: Prof. Dr. Aparecido Nilceu Marana

BAURU  
JANEIRO/2023

M842s      Morelli, Luis Henrique

             Sistema de Reconhecimento Biométrico de Múltiplas Pessoas Baseado em  
Métodos de Estimação de Pose e Rastreamento de Objetos / Luis Henrique  
Morelli. -- Bauru, 2023

             79 p. : il., tabs.

             Trabalho de conclusão de curso (Bacharelado - Ciência da Computação) -  
Universidade Estadual Paulista (Unesp), Faculdade de Ciências, Bauru

             Orientador: Aparecido Nilceu Marana

             1. Biometria. 2. Estimação de Pose. 3. Rastreamento e Reidentificação de  
Objetos. 4. Reconhecimento de Marcha. 5. Medidas Antropométricas. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Ciências, Bauru.

Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

Luis Henrique Morelli

# **Sistema de Reconhecimento Biométrico de Múltiplas Pessoas Baseado em Métodos de Estimação de Pose e Rastreamento de Objetos**

Trabalho de Conclusão de Curso do Curso de Bacharelado em Ciência da Computação da Universidade Estadual Paulista "Júlio de Mesquita Filho", Faculdade de Ciências, Campus Bauru.

Banca Examinadora

---

**Prof. Dr. Aparecido Nilceu Marana**

Orientador

Universidade Estadual Paulista "Júlio de Mesquita Filho"

Faculdade de Ciências

Departamento de Computação

---

**Prof<sup>a</sup>. Dr<sup>a</sup>. Simone das Graças Domingues Prado**

Universidade Estadual Paulista "Júlio de Mesquita Filho"

Faculdade de Ciências

Departamento de Computação

---

**Prof. Dr. João Paulo Papa**

Universidade Estadual Paulista "Júlio de Mesquita Filho"

Faculdade de Ciências

Departamento de Computação

Bauru, 19 de janeiro de 2023.

*Dedico este trabalho a toda a minha família, ao meu cachorrinho Dante, aos meus professores e amigos, e a todos que estiveram presentes durante esta longa jornada.*

# Agradecimentos

Agradeço, primeiramente, aquele ou aquilo que denominados Deus, força criadora de todo universo e das leis que o regem, por conceber sobre mim o dom da vida e a oportunidade de experienciar todas as emoções provenientes dela.

Agradeço imensamente minha família, em especial aos meus pais, Luis Fabiano Morelli e Marta Mônica da Silva Morelli, e minha irmã, Monique Vitória Morelli, por todo amor, carinho, apoio, sustento, lições, motivação e demais substantivos cabíveis. Vocês são minha base e meu tudo, um simples agradecimento não é suficiente para mensurar o que sinto por vocês.

Agradeço a todos os meus professores do ensino fundamental, do ensino médio e da graduação por pavimentarem a estrada do conhecimento para mim. Em especial, agradeço ao Prof. Dr. Aparecido Nilceu Marana pela oportunidade de trabalharmos juntos em dois projetos de iniciação científica e neste trabalho de conclusão de curso, por todo aprendizado e orientação durante este período.

Agradeço ao CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico, pela bolsa PIBIC-UNESP concedida como financiamento para realização do projeto de Iniciação Científica intitulado Identificação Biométrica de Múltiplas Pessoas em Vídeos Utilizando Métodos de Rastreamento e de Estimação de Poses, a qual contemplou parte deste trabalho também.

Agradeço aos ex-alunos Daniel Ricardo dos Santos Jangua e Henrique Leal Tavares, pois sem as pesquisas desenvolvidas por eles este trabalho não seria possível.

Agradeço ao ex-aluno Gustavo de Rosa pela construção do modelo de TCC em LaTeX.

Por fim, agradeço a todos os amigos e colegas de curso, por tornarem todos os momentos, dentro e fora da universidade, mais leves e memoráveis. Em especial, agradeço os meus amigos de turma Davi Augusto Neves Leite, Giovani Candido e Luiz Fernando Merli de Oliveira Sementille por enfrentarem todos os desafios da graduação ao meu lado sempre.

*"It's the questions we can't answer that teach us the most. They teach us how to think."*

Patrick Rothfuss

# Resumo

A biometria se tornou uma importante ferramenta para identificação de indivíduos nas mais diversas áreas, usada para evitar fraudes e aumentar a segurança dos cidadãos na sociedade. A marcha tem sido proposta como característica biométrica, visto que humanos possuem formas distintas de caminhar. Ela é particularmente importante para a identificação de pessoas à distância. Na maioria das pesquisas encontradas na literatura, os métodos de reconhecimento de marcha limitam-se a cenários em que há apenas um indivíduo caminhando, quando, em situações mais realistas, é comum existirem várias pessoas. Os métodos atuais de reconhecimento de marcha não podem ser aplicados em cenários multimarcha. Sendo um problema mais desafiador, para atingir tal objetivo, são necessários métodos de detecção e rastreamento de pessoas. O DeepSORT é um método de rastreamento de múltiplos objetos que utiliza métricas de associação profunda e vem obtendo bons resultados na tarefa. Este trabalho propõe a construção de um sistema biométrico capaz de identificar pessoas através de sua marcha e das medidas antropométricas, empregando métodos de estimação de poses e rastreamento e reidentificação de objetos. Para tal efeito, um método de reconhecimento de múltiplas pessoas é proposto. O método foi implementado e avaliado em duas bases de dados públicas, a CASIA Gait Dataset-A e a SMVDU-Multi-Gait, além de uma base de dados privada, coletada para validação do sistema biométrico que foi construído incorporando o método proposto. Os resultados obtidos mostram que essa abordagem é adequada em cenários de caminhada individual e em grupo, apresentando bom desempenho na identificação de indivíduos e robustez para configurar como solução de baixo orçamento em sistemas visuais inteligentes.

**Palavras-chave:** Biometria; Estimação de Poses; Rastreamento e Reidentificação de Objetos; Reconhecimento de Marcha; Medidas Antropométricas.



# Abstract

Biometrics has become an important tool for identifying individuals in many different areas, used to prevent fraud and increase the security of citizens in society. Gait has been proposed as a biometric feature, since humans have distinctive ways of walking. It is particularly important for identifying people at a distance. In most of the research found in the literature, gait recognition methods are limited to scenarios where there is only one individual walking, when in more realistic situations it is common to have several people. Current gait recognition methods cannot be applied in multi-gait scenarios. Being a more challenging problem, to achieve such a goal, people detection and tracking methods are needed. DeepSORT is a multi-object tracking method that uses deep association metrics and has been achieving good results in the task. Based on this study, this paper proposes the construction of a biometric system capable of identifying people through their gait and anthropometric measurements, employing methods for pose estimation and object tracking and re-identification. To this end, a multi-person gait recognition method is proposed. The method was implemented and evaluated on two public databases, CASIA Gait Dataset-A and SMVDU-Multi-Gait, in addition to a private database, collected for validation of the biometric system that was built incorporating the proposed method. The results obtained show that this approach is suitable in individual and group walking scenarios, presenting good performance in identifying individuals and robustness to configure as a low-budget solution in visual intelligent systems.

**Keywords:** Biometrics; Pose Estimation; Object Tracking and Re-identification; Gait Recognition; Anthropometric Measurements.

# Lista de Figuras

Figura 1 – Exemplos de características do corpo usadas para reconhecimento biométrico.	16
Figura 2 – Ciclo da marcha e suas divisões.	17
Figura 3 – Exemplos de cenários multimarcha do MOT17.	18
Figura 4 – Arquitetura básica de um sistema biométrico genérico.	24
Figura 5 – Estrutura do sistema biométrico dividida em duas fases: registro e autenticação.	24
Figura 6 – Diagrama de Voronoi representando o método 1NN.	27
Figura 7 – Fase de autenticação de um sistema biométrico baseado em verificação.	28
Figura 8 – Fase de autenticação de um sistema biométrico baseado em identificação.	29
Figura 9 – Diagrama de blocos do método de Jangua e Marana (2020) para reconhecimento de marcha.	31
Figura 10 – Fluxo de execução de um sistema biométrico baseado em multibiometria a nível de pontuação.	34
Figura 11 – Modelos de pose 2D usado pelos métodos OpenPose e OpenPifPaf.	36
Figura 12 – <i>Pipeline</i> do método OpenPose.	37
Figura 13 – Arquitetura da CNN utilizada no OpenPose.	37
Figura 14 – <i>Pipeline</i> do método OpenPifPaf.	38
Figura 15 – Exemplo de aplicação do método DeepSORT.	39
Figura 16 – <i>Pipeline</i> do método DeepSORT.	41
Figura 17 – Distribuição das comparações genuínas e impostoras de um sistema biométrico de verificação.	43
Figura 18 – Exemplos de curvas ROC associadas as suas respectivas distribuições de comparação entre genuínos e impostores.	44
Figura 19 – Exemplo de uma curva CMC.	45
Figura 20 – <i>Pipeline</i> do método proposto.	48
Figura 21 – Medidas antropométricas obtidas das poses 2D dos indivíduos, utilizadas pelo método de Tavares, Marana e Neto (2021).	51
Figura 22 – Exemplos de vídeos da CASIA Gait Dataset-A.	55
Figura 23 – Exemplos de vídeos da SMVDU Multi-Gait em situação de marcha simples.	55
Figura 24 – Exemplos de vídeos da SMVDU Multi-Gait em situação multimarcha.	55
Figura 25 – Curvas CMC obtidas pelas distâncias Euclidiana e Qui-Quadrado para os diferentes descritores na CASIA Gait Dataset-A.	57
Figura 26 – Curvas CMC obtidas pelas distâncias Euclidiana e Qui-Quadrado para os diferentes descritores na SMVDU-Multi-Gait com estimador de poses OpenPose.	59

Figura 27 – Curvas CMC obtidas pelas distâncias Euclidiana e Qui-Quadrado para os diferentes descritores na SMVDU-Multi-Gait com estimador de poses OpenPifPaf. . . . .	60
Figura 28 – Resumo da arquitetura e do <i>pipeline</i> do sistema biométrico. . . . .	63
Figura 29 – Pinagem do módulo de gravação ESP32-CAM. . . . .	64
Figura 30 – Pinagem do módulo sensor HC-SR501. . . . .	64
Figura 31 – Circuito do módulo sensor do sistema biométrico. . . . .	65
Figura 32 – Esquema de funcionamento do servidor HTTP para aplicação do método. . . . .	66
Figura 33 – Diagrama do banco de dados utilizado pelo sistema biométrico. . . . .	67
Figura 34 – Exemplos de imagens dos vídeos capturados para a base de dados do sistema em situação de marcha simples. . . . .	68
Figura 35 – Exemplos de imagens dos vídeos capturados para a base de dados do sistema em situação multimarcha. . . . .	68
Figura 36 – Fluxo da fase de registro. . . . .	70
Figura 37 – Painel de monitoramento em tempo real da fase de autenticação. . . . .	71
Figura 38 – Curvas CMC obtidas na base de dados coletada para o sistema. . . . .	72

# Lista de Tabelas

Tabela 1 – Acurácia Rank-1 - CASIA Gait Dataset-A. . . . .	57
Tabela 2 – Acurácia Rank-1 - SMVDU-Multi-Gait com o estimador de poses OpenPose. . . . .	59
Tabela 3 – Acurácia Rank-1 - SMVDU-Multi-Gait com estimador de poses OpenPifPaf. . . . .	61
Tabela 4 – Acurácia Rank-1 - Base de Dados do Sistema. . . . .	72

# Lista de Abreviaturas e Siglas

1NN	1-Nearest Neighbor
AUC	Area Under the Curve
CMC	Cumulative Match Characteristic
CNN	Convolutional Neural Network
DBMS	Database Management System
DeepSORT	Simple Online Realtime Tracking with a Deep Association Metric
DNA	Ácido Desoxirribonucleico
FAR	False Accept Rate
FTP	File Transfer Protocol
FNIR	False Negative Identification Rate
FPIR	False Positive Identification Rate
FRR	False Reject Rate
GAR	Genuine Accept Rate
GPU	Graphics Processing Unit
GUI	Graphical User Interface
HTTP	Hypertext Transfer Protocol
IoT	Internet of Things
KNN	K-Nearest Neighbors
PAF <sup>1</sup>	Part Affinity Field
PAF <sup>2</sup>	Part Association Field
PIF	Part Intensity Field
ROC	Receiving Operating Characteristic
SORT	Simple Online Realtime Tracking

SQL	Structured Query Language
SRAM	Static Random Access Memory
TAR	True Accept Rate
TPIR	True Positive Identification Rate
USB	Universal Serial Bus

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>16</b>
<b>1.1</b>	<b>Problema</b>	<b>18</b>
<b>1.2</b>	<b>Justificativa</b>	<b>19</b>
<b>1.3</b>	<b>Objetivos</b>	<b>20</b>
1.3.1	Objetivo Geral	20
1.3.2	Objetivos Específicos	20
<b>1.4</b>	<b>Organização do Trabalho</b>	<b>21</b>
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>22</b>
<b>2.1</b>	<b>Biometria</b>	<b>22</b>
<b>2.2</b>	<b>Sistemas Biométricos</b>	<b>23</b>
2.2.1	Sensores	24
2.2.2	Extração de Características	25
2.2.3	Bases de Dados	26
2.2.4	Classificação	26
2.2.4.1	Classificador KNN	26
2.2.5	Funcionalidades	27
2.2.5.1	Verificação	28
2.2.5.2	Identificação	28
2.2.6	Reconhecimento de Marcha	29
<b>2.3</b>	<b>Método Baseado nas Posições das Partes de Membros</b>	<b>30</b>
2.3.1	Multimarcha	31
<b>2.4</b>	<b>Biometrias Suaves</b>	<b>32</b>
2.4.1	Método Baseado em Medidas Antropométricas	32
<b>2.5</b>	<b>Multibiometria</b>	<b>33</b>
<b>2.6</b>	<b>Estimação de Poses</b>	<b>35</b>
2.6.1	OpenPose	36
2.6.2	OpenPifPaf	37
<b>2.7</b>	<b>Rastreamento de objetos</b>	<b>38</b>
2.7.1	DeepSORT	39
<b>2.8</b>	<b>Métricas de Desempenho</b>	<b>41</b>
2.8.1	Medidas de Desempenho para Sistemas de Verificação	42
2.8.2	Medidas de Desempenho para Sistemas de Identificação	44
<b>2.9</b>	<b>Funções de distância</b>	<b>46</b>
2.9.1	Distância Euclidiana	46

2.9.2	Distância Qui-Quadrado . . . . .	46
<b>3</b>	<b>MÉTODO PROPOSTO . . . . .</b>	<b>47</b>
3.1	Estimação de Poses . . . . .	48
3.2	Rastreamento e Reidentificação . . . . .	49
3.3	Associação entre Poses e <i>Bounding Boxes</i> . . . . .	49
3.4	Extração das Características . . . . .	50
3.4.1	Descritores de Marcha Baseados no Método de Jangua e Marana (2020) . .	50
3.4.2	Descritores Antropométricos Baseados no Método de Tavares, Marana e Neto (2021) . . . . .	50
3.5	Geração de <i>Templates</i> . . . . .	51
3.6	Classificação . . . . .	52
<b>4</b>	<b>RESULTADOS EXPERIMENTAIS . . . . .</b>	<b>53</b>
4.1	Experimentos . . . . .	53
4.1.1	Bases de Dados . . . . .	54
4.1.1.1	CASIA Gait Dataset-A . . . . .	54
4.1.1.2	SMVDU-Multi-Gait . . . . .	55
4.1.2	Métricas . . . . .	56
4.2	Resultados e Discussão . . . . .	56
4.2.1	OpenPose . . . . .	58
4.2.2	OpenPifPaf . . . . .	60
<b>5</b>	<b>SISTEMA BIOMÉTRICO . . . . .</b>	<b>62</b>
5.1	Sensores . . . . .	63
5.2	Servidor HTTP . . . . .	65
5.3	Bases de Dados . . . . .	66
5.4	Registro e Autenticação . . . . .	69
5.5	Resultados e Discussão . . . . .	72
<b>6</b>	<b>CONCLUSÃO . . . . .</b>	<b>74</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>76</b>



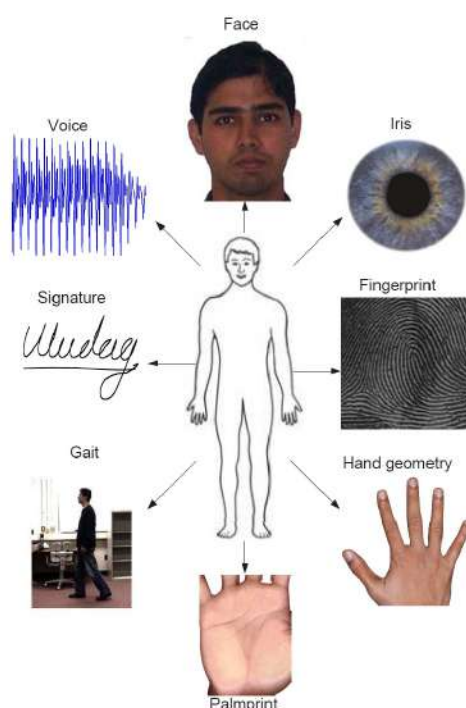
# 1 Introdução

O conjunto de atributos associados a uma pessoa constitui sua identidade pessoal (JAIN; ROSS; NANDAKUMAR, 2011). A capacidade de identificar indivíduos a partir de suas características únicas, baseando-se em seus aspectos pessoais, vem se destacando como uma ferramenta crucial para a sociedade, permitindo controlar e restringir o acesso a determinadas áreas ou informações, identificar criminosos em ambientes públicos movimentados, como aeroportos e portos, e evitar fraudes, impedindo falsificações de identidade e dados.

Segundo Jain, Ross e Nandakumar (2011), o reconhecimento biométrico é a ciência que permite a identificação de indivíduos através do estudo estatístico das características físicas e comportamentais dos seres humanos, denominados traços biométricos. Como os identificadores biométricos são inerentes a cada sujeito, dificultam sua manipulação, compartilhamento, ou esquecimento, corroborando para o avanço dos estudos neste campo.

Atualmente, métodos biométricos que empregam traços como digitais, íris e face estão inseridos em bancos e órgãos públicos, requerendo boas taxas de performance e confiabilidade, incentivando pesquisas que intuem o aprimoramento destas identificações, potencializando sua rapidez e acurácia. A Figura 1 mostra algumas das características biométricas utilizadas em pesquisas e aplicações.

Figura 1 – Exemplos de características do corpo usadas para reconhecimento biométrico.



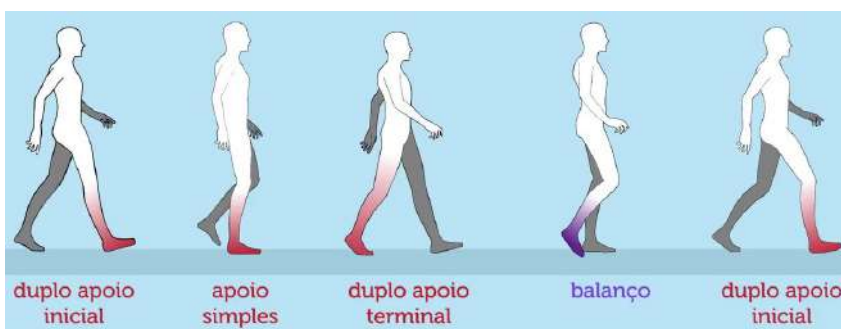
Fonte: Jain, Ross e Nandakumar (2011).

Contudo, mesmo com todo o desenvolvimento desta área, pesquisadores almejam superar o desafio de identificação de pessoas em tempo real, por meio de imagens de vídeo capturadas por câmeras de segurança em baixas resoluções, com posicionamentos desfavoráveis, em ambientes não controlados, de maneira discreta, à distância e sem a necessidade de cooperação dos alvos da detecção.

No contexto descrito, os métodos elencados anteriormente, compondo técnicas clássicas da biometria, se tornam inviáveis, haja vista a dificuldade no uso de características faciais e táteis, fazendo-se necessárias coletas de dados com alta qualidade, acarretando a queda do desempenho caso contrário. Ademais, os resultados são impactados pela presença de elementos oclusores, tais como bonés, chapéus e óculos escuros, como expõem Arantes e Gonzaga (2011). A utilização de características da marcha humana abarca uma das únicas opções possíveis nestes cenários (WAN; LI; PHOHA, 2018).

O reconhecimento humano baseado em marcha é um traço biométrico comportamental emergente para vigilância e monitoramento em sistemas visuais inteligentes (SINGH et al., 2018). A marcha tem sido proposta como característica biométrica, haja vista que humanos possuem uma maneira peculiar de caminhar (WANG et al., 2003), com padrões distintos, compreendendo uma assinatura que possui unicidade suficiente para servir como fator de identificação (NIXON; CARTER, 2006), facilmente interpretável a partir de um ponto de vista biomecânico. Ela pode ser definida por comportamentos motores formados por movimentos repetitivos e integrados do corpo humano, se repetindo em ciclos (ARANTES; GONZAGA, 2011), como ilustra a Figura 2.

Figura 2 – Ciclo da marcha e suas divisões.



Fonte: Marques e Vaz (2018).

Esta abordagem vem ganhando evidência nos últimos anos pois, diferentemente de outros tipos de biometria, como face, impressão digital e íris, pode ser coletada à distância e não requer a cooperação do indivíduo (LIMA; MELO; SCHWARTZ, 2021). Portanto, abarca uma das únicas opções de biometria possíveis para superar o desafio de identificação de pessoas em tempo real, por meio de imagens de vídeo capturadas por câmeras de segurança em baixas resoluções, com posicionamentos desfavoráveis, em ambientes não controlados, de maneira discreta, à distância e sem a necessidade de cooperação dos alvos da detecção (WAN; LI; PHOHA, 2018).

Em sua maioria, as pesquisas encontradas na literatura tratam sobre os métodos de reconhecimento de marcha limitados a cenários em que há apenas um indivíduo caminhando. Contudo, em situações mais realistas, considerando identificações em tempo real, em ambientes tais quais aeroportos, *shoppings* e estações de metrô, é comum as pessoas andarem em grupos ou existir várias pessoas em um mesmo ambiente. Esta situação é denominada multimarcha, exemplificada na Figura 3.

Figura 3 – Exemplos de cenários multimarcha do MOT17.



(a) Câmera frontal em um *shopping*.



(b) Câmera elevada em uma rua.

Fonte: Milan et al. (2016).

O objetivo do reconhecimento multimarcha é a identificação de determinado indivíduo quando ele está caminhando com indivíduos diferentes simultaneamente (CHEN et al., 2018). Complicações advindas dessa configuração incluem uma pessoa ocluindo outra ou sendo ocluída por algum objeto estático. É um problema muito mais desafiador, requerendo métodos precisos de detecção e rastreamento de pessoas para produzir bons resultados (DELGADO-ESCANO et al., 2020).

O rastreamento de vários objetos tem sido um objetivo de longa data em visão computacional (ZHANG et al., 2021). O objetivo é estimar trajetórias para objetos de interesse apresentados em vídeos. A correta resolução deste problema pode imediatamente beneficiar diversas aplicações, desde análises de vídeo inteligentes, até interação humano-computador e reconhecimento de atividade humana. Houve um progresso notável na detecção e reidentificação de objetos, dentre eles, um método estado-da-arte para rastreamento de múltiplos objetos utilizando métrica de associação profunda, o DeepSORT (WOJKE; BEWLEY; PAULUS, 2017).

## 1.1 Problema

Os métodos biométricos mais comuns, que aplicam técnicas clássicas, tais quais as biometrias baseadas em reconhecimento de digitais, íris e face, não apresentam características que permitem sua implementação em cenários onde as imagens de vídeo capturadas por câmeras de segurança possuem baixas resoluções e apresentem ângulos e posicionamentos desfavoráveis, estejam situadas em ambientes não controlados e precisem realizar a identificação de maneira

discreta, à distância e sem a necessidade de cooperação dos alvos da detecção (ARANTES; GONZAGA, 2011).

As técnicas de reconhecimento de marcha baseadas em modelo apresentam uma maior demanda de poder de processamento computacional em relação à abordagem livre de modelo, por outro lado apresentam resultados mais robustos na classificação, uma vez que utilizam as características dinâmicas da marcha (ARANTES; GONZAGA, 2011). O equilíbrio entre o custo computacional e acurácia do método se mostra um fator crítico de sucesso para implementação de um sistema biométrico baseado em marcha em ambientes reais.

Todavia, apesar de apresentarem resultados e performances satisfatórias, estes métodos compreendem situações com um único participante caminhando, quando, na realidade, por muitas vezes, as pessoas costumam caminhar acompanhadas de outras, sejam familiares, ou mesmo fazer parte de multidões que frequentam um mesmo local (CHEN et al., 2018).

Outrossim, na visão dos autores Chen et al. (2018), o comportamento de andar de uma pessoa tende a variar quando caminha com diferentes pessoas, introduzindo desafios para os métodos tradicionais de reconhecimento de marcha em caso de várias pessoas presentes em um mesmo cenário.

## 1.2 Justificativa

A biometria é um elemento essencial para garantir a segurança dos indivíduos em sociedade. Entretanto, as técnicas clássicas de biometria ainda não solucionam alguns desafios importantes, representando obstáculos para que sistemas biométricos possam ser utilizados de forma eficiente em ambientes reais. A utilização de novas abordagens se faz necessária para caminhar na direção do desenvolvimento desses sistemas. Tendo em vista essas necessidades, o reconhecimento de marcha se mostra uma ótima solução, uma vez que: (i) pode ser feito à distância; (ii) apresenta bom desempenho mesmo em imagens de baixa resolução; (iii) pode ser feito sem a cooperação do indivíduo; e (iv) o desempenho não é alterado quando os traços biométricos mais comuns, como face, íris ou digitais, estão ocultos nas imagens (WAN; LI; PHOHA, 2018).

A utilização de métodos baseados em modelos para reconhecimento de marcha, como os propostos por Liu et al. (2016), Lima e Schwartz (2019) e Jangua e Marana (2020), mostraram-se os mais adequados para a situação, visto que apresentam melhores resultados quando comparados com métodos livres de modelo, vide Arantes e Gonzaga (2011), e a demanda por poder computacional não se mostra mais um problema tão grande nos dias de hoje. Além disso, esses métodos utilizam estimadores de pose robustos, capazes de realizar a detecção de múltiplas pessoas e estimar suas poses em tempo real com alta confiabilidade, mesmo na presença de obstruções e variações de vestimenta.

Como a multimarcha difere da marcha simples, onde apenas um elemento aparece em cena, os métodos aplicados para esta não podem ser empregados para a multimarcha. Contudo, em essência, o reconhecimento de marcha com várias pessoas visa descobrir características que não se alteram ao caminhar com demais pessoas, ou seja, aspectos estáveis e invariantes, porém de baixo nível (CHEN et al., 2018). Ainda segundo Chen et al. (2018), como o reconhecimento de marcha de múltiplas pessoas é uma extensão da marcha simples, métodos tradicionais de reconhecimento de marcha são valiosos e podem ser empregados em situações multimarcha.

## 1.3 Objetivos

Este trabalho tem por objetivo realizar estudos na área de biometria, com enfoque no reconhecimento de marcha humana, intuindo alcançar bom desempenho na tarefa de reconhecimento multimarcha através da alteração de métodos propostos na literatura para marcha simples, com auxílio de biometrias suaves, fusão de características biométricas e rastreadores de objetos. Para tal efeito, o objetivo geral e os objetivos específicos são listados abaixo.

### 1.3.1 Objetivo Geral

Implementar um sistema de reconhecimento capaz de realizar a identificação biométrica em situações de marcha de múltiplos indivíduos, utilizando descritores de marcha e antropométricos, para o controle do acesso a áreas restritas, utilizando estimação de poses e rastreamento dos participantes, extraídos de vídeos capturados por uma única câmera de baixa resolução.

### 1.3.2 Objetivos Específicos

- a) Estudar os conceitos básicos sobre a Biometria e os sistemas biométricos;
- b) Estudar métodos de reconhecimento de marcha baseados em poses 2D;
- c) Estudar métodos de rastreamento e reidentificação de objetos;
- d) Implementar os métodos de reconhecimento de marcha escolhidos;
- e) Implementar o método de rastreamento de objetos escolhido;
- f) Comparar o desempenho dos métodos na tarefa de reconhecimento biométrico utilizando dois estimadores de pose diferentes;
- g) Montar a estrutura física do sistema biométrico, utilizando a câmera de baixa resolução e um sensor de presença;

- h) Capturar imagens de pessoas previamente identificadas caminhando, sozinhas e acompanhadas umas das outras, utilizando a câmera de baixa resolução, para compor a base de dados do sistema biométrico;
- i) Desenvolver um sistema para computadores que receba uma nova captura de vídeo de indivíduos caminhando, aplique o método de reconhecimento de marcha selecionado, conjuntamente com o método baseado em medidas antropométricas, e realize a identificação dos elementos detectados.

## 1.4 Organização do Trabalho

A divisão do trabalho em capítulos foi estruturada da seguinte maneira:

Capítulo 2: Apresenta os conceitos básicos da biometria e dos sistemas biométricos, com enfoque nos métodos que realizam a identificação através das características biométricas de medidas antropométricas e da marcha. Estes conceitos embasam a metodologia desenvolvida.

Capítulo 3: Compreende o *pipeline* do método proposto pelo trabalho.

Capítulo 4: Apresenta as etapas de desenvolvimento do trabalho em detalhes, demonstrando os resultados obtidos pelos experimentos, as bases de dados onde o método foi empregado e avaliado e os protocolos adotados no decorrer dos experimentos, discorrendo sobre suas causas e fundamentando as decisões para a construção do sistema biométrico.

Capítulo 5: Compreende o detalhamento das etapas de construção do sistema biométrico desenvolvido, descrevendo sua estrutura física e seus módulos lógicos, embasados na esquemática apresentada no Capítulo 2 e nas determinações dos experimentos realizados na seção 4.1.

Capítulo 6: Apresenta a conclusão do trabalho, dissertando a partir dos resultados experimentais obtidos no Capítulo 4 e do desempenho atingido pelo sistema biométrico elaborado.

## 2 Fundamentação Teórica

Neste capítulo são apresentados, inicialmente, conceitos básicos relacionados à biometria e sistema biométricos, abrangendo conhecimentos e componentes necessários para a construção de sistemas capazes de realizar as tarefas de verificação e identificação de indivíduos.

Uma das componentes chaves de um sistema biométrico é a biometria associada a ele. Portanto, são expostos o reconhecimento por marcha humana e por biometrias suaves, como também maneiras de utilizar estes dois tipos de traços biométricos na composição de um sistema multibiométrico.

A extração das características biométricas discutidas necessita da aplicação de métodos capazes de indicar a posição do corpo humano nas sequências de imagens que compõem os dados biométricos. Para tal finalidade, comenta-se sobre métodos de estimação de poses, algoritmos apropriados para essa tarefa.

A presença de múltiplas pessoas em um cenário dificulta a identificação de indivíduos, haja vista que, apesar de ser possível estimar as poses de vários indivíduos simultaneamente, não há como saber quais pontos pertencem a quem. Por esse motivo, conhecimentos de rastreamento e reidentificação de objetos auxiliam o reconhecimento biométrico nessas situações.

Por fim, são evidenciadas as métricas de desempenho de sistemas biométricos de verificação e identificação, como são formuladas e o que representam, bem como funções de distância empregadas para aquisição da pontuação de similaridade calculada pelos classificadores.

### 2.1 Biometria

A tarefa principal quando se fala em gerenciamento de identidade é o estabelecimento de um vínculo entre indivíduo e identidade pessoal. Isso diz respeito a capacidade de determinar a identidade de uma pessoa ou verificar a declaração de identidade de um sujeito sempre que necessário. O processo descrito é nomeado reconhecimento de pessoas (JAIN; ROSS; NANDAKUMAR, 2011).

Segundo Jain, Ross e Nandakumar (2011), existem essencialmente três métodos para o reconhecimento de uma pessoa: (i) métodos baseados em conhecimento, aqueles que são dependentes do conhecimento exclusivo de determinado indivíduo sobre informações secretas pré-cadastradas em um sistema, como senhas, número de identificação pessoal ou chave de criptografia; (ii) métodos baseados em posse, havendo a necessidade de o indivíduo estar em posse de um validador externo, como documento de identidade, passaporte ou habilitação; e (iii) métodos biométricos, estabelecendo uma relação da identidade de um indivíduo com seus

traços, físicos ou comportamentais, inerentes. A essa última, é dado o nome de reconhecimento biométrico. O reconhecimento biométrico, ou simplesmente biometria, oferece uma forma natural e mais confiável de solucionar problemas de reconhecimento de pessoas, haja vista que os identificadores biométricos são intrínsecos de cada sujeito e são mais difíceis de serem manipulados, compartilhados ou esquecidos (JAIN; ROSS; NANDAKUMAR, 2011).

Formalmente, Jain, Ross e Nandakumar (2011) definem o reconhecimento biométrico como a ciência que estabelece a identidade de um indivíduo com base nas características físicas e/ou comportamentais de uma pessoa de forma totalmente automatizada ou semiautomatizada, requerendo o auxílio de um sistema para interagir com o indivíduo procurando validação, conhecido como usuário.

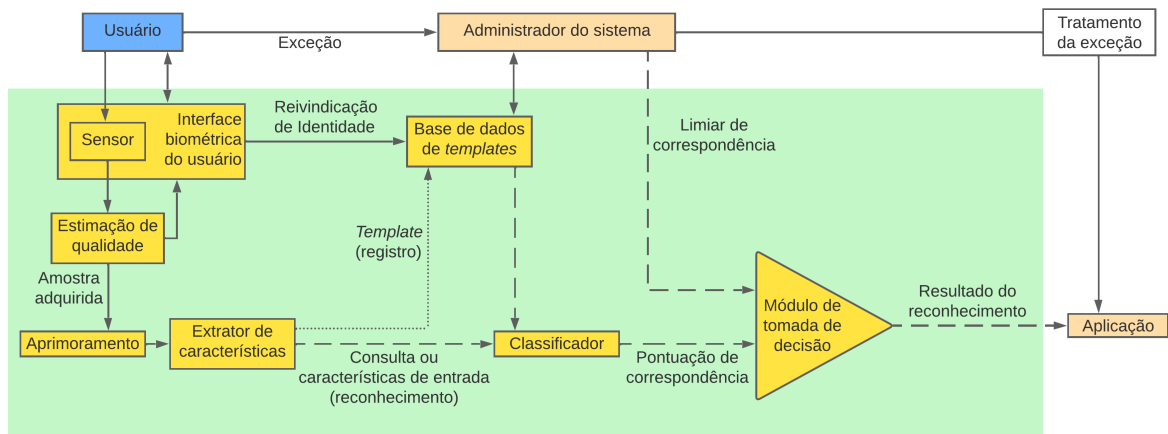
## 2.2 Sistemas Biométricos

Os sistemas biométricos estão se tornando cada vez mais importantes, pois fornecem meios mais confiáveis e eficientes de verificação de identidade. São um conjunto de diversas tecnologias com a função de extração, processamento, armazenamento, recuperação e comparação de dados biológicos, com o intuito de identificar e autenticar indivíduos (MARCONDES, 2020). Eles medem uma ou mais características físicas ou comportamentais, incluindo impressão digital, rosto, íris, retina, voz, assinatura, marcha, odor ou informações do DNA de um indivíduo para determinar ou verificar sua identidade (JAIN; ROSS; NANDAKUMAR, 2011). Segundo o autor, essas características recebem diferentes termos, como traços, indicadores, identificadores ou modalidades. Durante o curso deste trabalho, essas nomenclaturas serão intercambiáveis.

Choudhury et al. (2018) esclarecem que os sistemas biométricos são divididos em duas fases: (i) registro, responsável por efetuar as coletas das amostras de dados biométricos dos usuários, armazenando-as em uma base de dados junto a identificação do indivíduo, denominada rótulo; e (ii) autenticação, responsável por comparar os dados obtidos em tempo real em uma nova requisição de identificação aos dados armazenados na base de dados, realizando o reconhecimento do usuário. Portanto, as fases compreendem três etapas de operação: (i) captura de amostra biométrica da pessoa; (ii) extração de um conjunto de características relevantes da amostra capturada; e (iii) comparação do conjunto de identificadores extraídos em relação ao modelo, denominado *template*, definido no banco de dados. A autenticação compara os dados obtidos em tempo real e os dados armazenados na base de dados para realização do reconhecimento (CHOUDHURY et al., 2018). A Figura 4 demonstra os blocos da construção básica de um sistema biométrico genérico, enquanto a Figura 5 ilustra a divisão do sistema em suas fases de registro e autenticação.

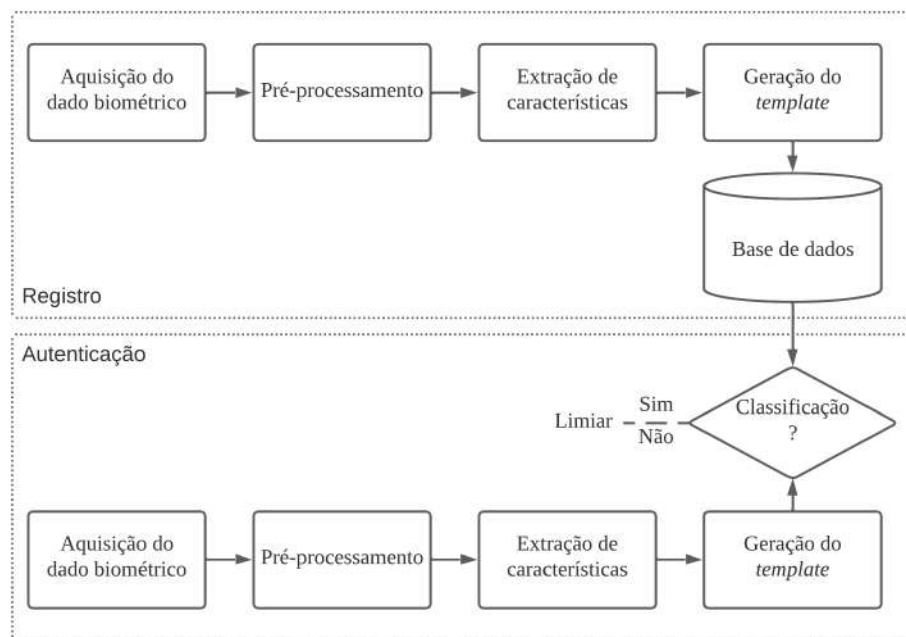


Figura 4 – Arquitetura básica de um sistema biométrico genérico.



Fonte: Adaptado de Jain, Ross e Nandakumar (2011).

Figura 5 – Estrutura do sistema biométrico dividida em duas fases: registro e autenticação.



Fonte: Adaptado de Choudhury et al. (2018).

A seguir, são descritos as componentes físicas e lógicas que compõem uma estrutura básica de um sistema biométrico genérico, sendo elas: (i) sensores; (ii) extração de características; (iii) bases de dados; e (iv) classificação.

### 2.2.1 Sensores

Os sensores são leitores biométricos responsáveis por realizar a interface com o usuário, medindo e registrando seus dados biométricos na base de dados do sistema (JAIN; ROSS; NANDAKUMAR, 2011). Os autores comentam que para a maioria das modalidades biométricas,

os indicadores são adquiridos na forma de imagens bidimensionais, como é o caso de sistemas que captam impressões digitais, face, íris e marcha. Existem exceções, tais quais a voz, registrada como um sinal de amplitude unidimensional, pressão da caneta, posição e velocidade para assinaturas online e para sistemas de base química, odor e DNA, *Deoxyribonucleic Acid*, do inglês, Ácido Desoxirribonucleico.

Portanto, a qualidade dessas detecções gera impactos, tanto positivos quanto negativos, no desempenho do sistema. Para imagens, por exemplo, fatores como resolução, taxa de quadros e sensibilidade da câmera desempenham um importante papel para determinar a qualidade da captura da modalidade (JAIN; ROSS; NANDAKUMAR, 2011).

### 2.2.2 Extração de Características

Habitualmente, os traços auferidos pelos sensores são submetidos a operações de pré-processamento antes que qualquer tipo de recurso para reconhecimento biométrico seja extraído, devido a possibilidade de os indicadores apresentarem informações que acabam prejudicando a identificação, vide ruídos, variações de ângulo, excesso ou falta de iluminação e posicionamento desfavorável (JAIN; ROSS; NANDAKUMAR, 2011).

As etapas que abarcam tal processo, esmiuça Jain, Ross e Nandakumar (2011), são três: (i) avaliação da qualidade, havendo a necessidade de acesso a amostra coletada, verificando sua adequação para as demais etapas do processo. Se a qualidade não for suficiente, pode-se tentar readquirir os dados do usuário ou acionar uma falha, alertando o administrador a demanda por intervenção; (ii) segmentação, com o objetivo de separar os dados biométricos do ruído de fundo, informações que não auxiliam o processo de reconhecimento, como acontece na separação do contorno do corpo de um indivíduo do cenário em volta dele; e (iii) aprimoramento de qualidade, visando aumentar a qualidade da imagem de acordo com os parâmetros da aplicação empregada, bem como reduzir os ruídos ainda presentes. É o caso de suavizações e equalizações de histogramas aplicados para minimização de ruídos introduzidos pela câmera ou variações de iluminação (JAIN; ROSS; NANDAKUMAR, 2011).

Após a certificação da qualidade e ajustes dos dados, eles estão prontos para terem suas características extraídas. A extração de características é o processo de geração de uma representação digital compacta e expressiva, contendo apenas as informações discriminatórias essenciais para o reconhecimento da pessoa, compondo um *template* (JAIN; ROSS; NANDAKUMAR, 2011).

Este *template* pode ser extraído de uma única amostra biométrica ou ser gerado pelo processamento de várias amostras adquiridas durante a inscrição, como é o caso das impressões digitais. Ademais, sistemas de reconhecimento baseado em face e marcha podem armazenar vários *templates* de um indivíduo, cada um correspondendo a uma pose diferente em relação à câmera.

Construídos os *templates* do usuário na fase de registro, eles são guardados na base de dados do sistema biométrico. No momento da autenticação, os *templates* são recuperados do banco de dados e comparado com o conjunto de características extraídas da nova modalidade adquirida para aquele usuário. Jain, Ross e Nandakumar (2011) definem os *templates armazenados* na base como *gallery*, enquanto aqueles captados em tempo real são chamados, normalmente, de *probe*.

### 2.2.3 Bases de Dados

O banco de dados do sistema biométrico atua como um repositório de informações biométricas, armazenando os *templates* provenientes da fase de registro a partir dos dados biométricos medidos pelos sensores, juntamente com informações pessoais de identidade, caracterizando o usuário (JAIN; ROSS; NANDAKUMAR, 2011).

### 2.2.4 Classificação

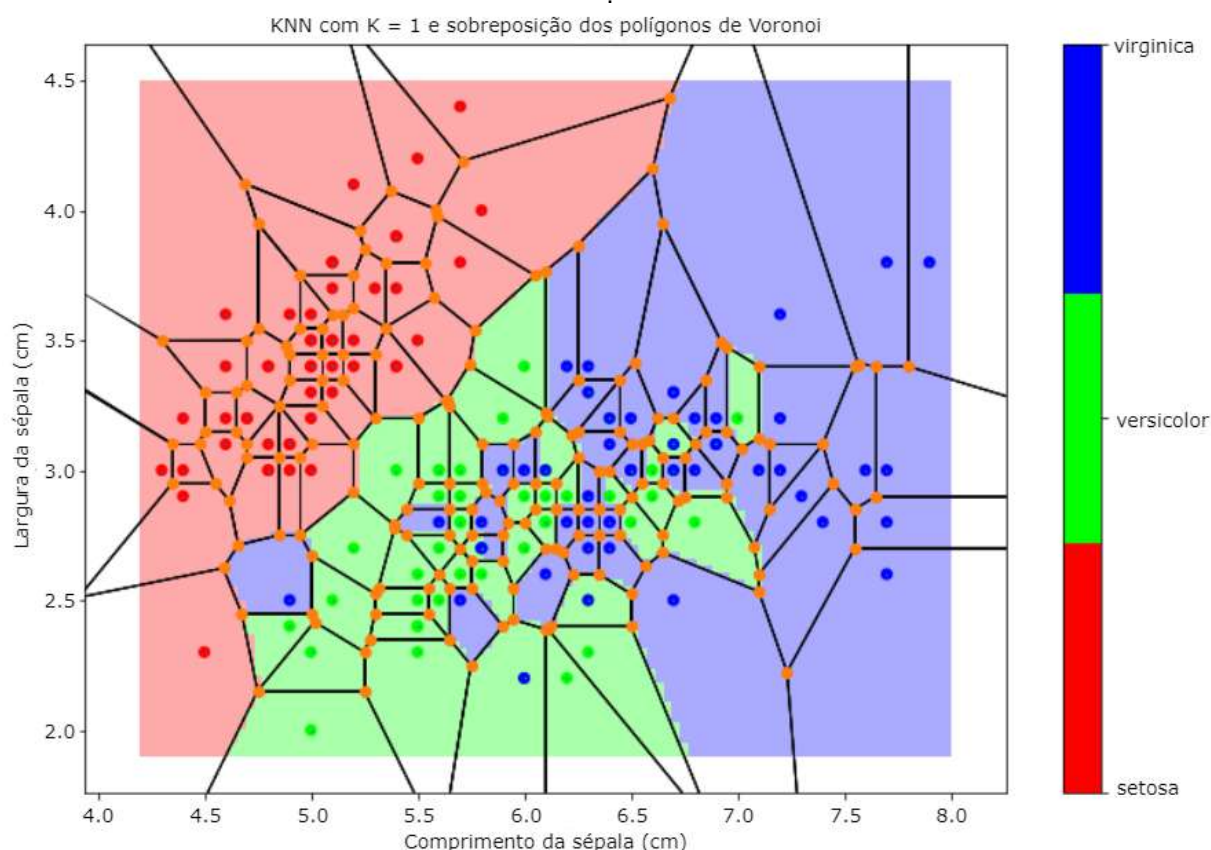
O classificador tem por objetivo a comparação das novas amostras coletadas, as *probes*, com as amostras armazenadas na base de dados do sistema biométrico, a *gallery*, a fim de gerar uma pontuação de correspondência utilizada para medir a similaridade entre duas amostras. Assim, quanto maior for a pontuação, mais similares serão as duas amostras comparadas. Analogamente, o classificador também é capaz de calcular a dissimilaridade entre amostras apenas invertendo a lógica anterior, isto é, quanto menor o valor, mais similares são, e vice-versa. Incorporado ao classificador, está o módulo de tomada de decisão, considerando as pontuações como critério de escolha e validação da identidade do indivíduo em seu reconhecimento (JAIN; ROSS; NANDAKUMAR, 2011).

#### 2.2.4.1 Classificador KNN

O algoritmo KNN, *K-Nearest Neighbors*, do inglês, K-ésimo Vizinho mais Próximo, é um algoritmo de classificação cuja estratégia é tentar classificar objetos levando em consideração os K vizinhos mais próximos de sua posição no espaço de características (COVER; HART, 1967). Assim, é um método fundamentado em métricas, cuja distância define a classe do objeto, contudo, elas podem variar de acordo com o escopo de cada aplicação. Configura uma das abordagens mais simples e comuns em aplicações de reconhecimento de padrões. Cover e Hart (1967) exprimem que o KNN é um algoritmo de classificação *lazy*, não envolvendo construção de modelos prévios. O novo exemplo é rotulado com base na quantidade de votos dos k exemplos mais próximos sendo que cada voto tem a distância como peso (WETTSCHERECK; AHA; MOHRI, 1997).

A abordagem mais simples para o método KNN localiza somente o vizinho mais próximo, considerando, portanto,  $K = 1$ . Nestes casos o método é conhecido por 1NN e pode-se representar o espaço de características pelo diagrama de Voronoi (MITCHELL, 1997), ilustrado na Figura 6, em que cada cor representa a região que se encontra mais próxima de um ponto. Ao aplicar o 1NN, cada uma das amostras seria classificada de acordo com o ponto representante de uma região delimitada.

Figura 6 – Diagrama de Voronoi representando o método 1NN.



Fonte: Preda (2020).

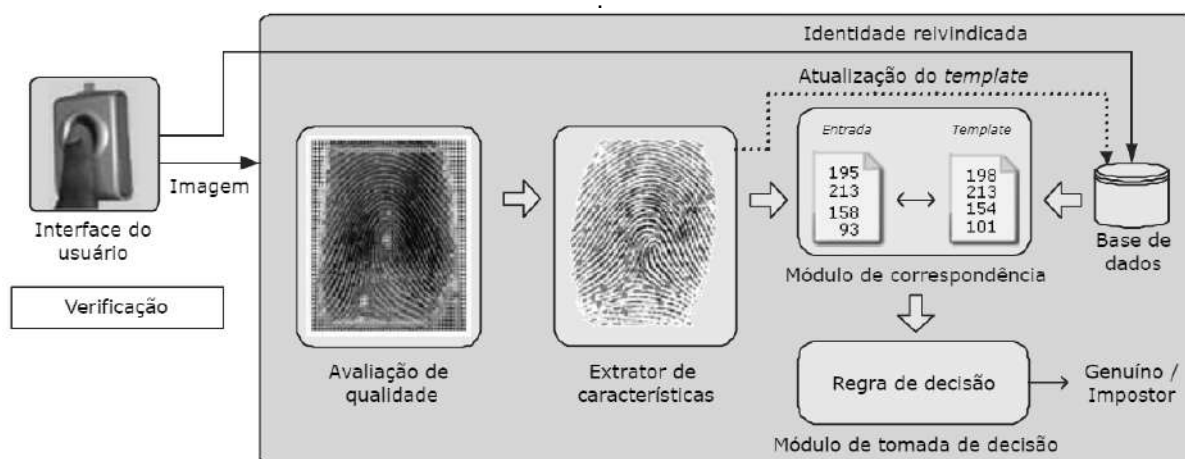
### 2.2.5 Funcionalidades

Segundo Jain, Ross e Nandakumar (2011), um sistema biométrico pode fornecer duas funcionalidades distintas: (i) verificação, que busca validar a identidade reivindicada por determinado indivíduo; e (ii) identificação, intentando atribuir a identidade correta para um indivíduo detectado. Maiores detalhes sobre cada funcionalidade são fornecidos nas subseções abaixo.

### 2.2.5.1 Verificação

Na verificação, um usuário apresenta alguma forma de identificador, por meio de um número de identificação pessoal, um nome de usuário ou um *token*, reivindicando uma identidade, e sua biometria é coletada. O sistema detecta as medidas, extrai as características, compara os dados de entrada com os dados registrados no banco de dados para a identidade daquele sujeito, acarretando uma correspondência um para um, e determina se a reivindicação é genuína ou não através da análise desses traços. Se a entrada do usuário e o modelo da identidade tiverem alto grau de similaridade, então é aceita como genuína. Caso contrário, é rejeitada e o usuário é considerado um impostor (JAIN; ROSS; NANDAKUMAR, 2011). A Figura 7 exibe o passo a passo da fase de autenticação de um sistema cuja funcionalidade é de verificação.

Figura 7 – Fase de autenticação de um sistema biométrico baseado em verificação.



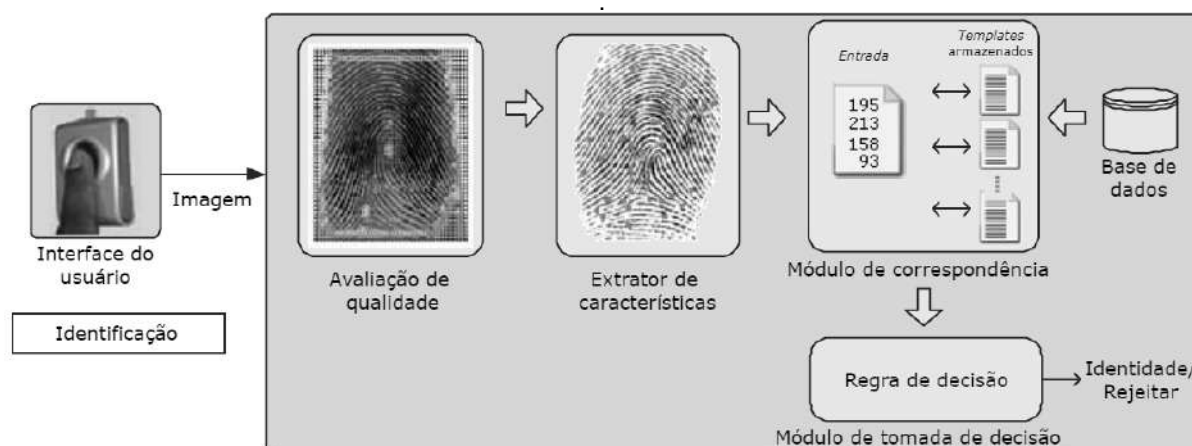
Fonte: Adaptado de Jain, Ross e Nandakumar (2011).

### 2.2.5.2 Identificação

Existem duas maneiras possíveis de aproveitar a funcionalidade de identificação no sistema biométrico, apontam Jain, Ross e Nandakumar (2011). A primeira delas é a identificação positiva, com o propósito de determinar a identidade do usuário a partir de um conjunto conhecido de identidades cadastradas no sistema biométrico. Neste contexto, o usuário não reivindica explicitamente uma identidade, sendo responsabilidade do sistema responder à pergunta "Você é alguém conhecido pelo sistema?".

A segunda maneira contrasta com a primeira, respondendo à pergunta "Você é quem diz que não é?", designada identificação negativa. Este tipo de identificação realiza uma triagem, com a intenção de impedir que uma única pessoa utilize múltiplas identidades, evitando que uma pessoa pleiteie diferentes benefícios sob diferentes personalidades. A Figura 8 demonstra as etapas da fase de autenticação de um sistema cuja funcionalidade é de identificação.

Figura 8 – Fase de autenticação de um sistema biométrico baseado em identificação.



Fonte: Adaptado de Jain, Ross e Nandakumar (2011).

Em ambos os casos, a entrada biométrica do usuário é comparada com todos os *templates* de todas as pessoas cadastradas no banco de dados, atribuindo a identidade cujo modelo apresentar maior grau de semelhança com o dado biométrico obtido ou indicando que o usuário não está cadastrado no sistema (JAIN; ROSS; NANDAKUMAR, 2011).

## 2.2.6 Reconhecimento de Marcha

A marcha, ou o ato de caminhar, pode ser definida por comportamentos motores compostos por movimentos repetitivos e integrados do corpo humano (ARANTES; GONZAGA, 2011). Trata-se de um padrão de movimentos corporais que se repetem a cada ciclo. Pesquisas realizadas nas últimas décadas mostram que os humanos têm padrões distintos e próprios de marcha (NIXON; CARTER, 2006) e que ela apresenta diversas características que permitem a diferenciação de indivíduos (ARANTES; GONZAGA, 2011). Por conseguinte, pode ser compreendida como uma característica biométrica importante, haja vista certa indicação de que a musculatura esquelética de cada pessoa é única, tornando a identificação biométrica por meio da marcha humana factível.

Alguns aspectos únicos da marcha, alegam Arantes e Gonzaga (2011), são: (i) a capacidade de ser capturada à grandes distância e em baixas resoluções; (ii) a ausência da necessidade de equipamentos complexos, performando bem com instrumentações simples, tais quais um único acelerômetro ou câmera; (iii) a ausência da dependência da cooperação dos alvos de detecção para sua execução; (iv) a dificuldade em copiar e imitar os trejeitos da marcha de outros indivíduos; e (v) se sobressai em situações onde ocorrem oclusões parciais dos membros do corpo do sujeito.

Ao caminhar, o corpo se locomove através de uma sequência de repetições de movimentos que o impulsionam para a frente enquanto mantém a postura estável, simultaneamente. Partindo da posição de repouso, quando ambos os membros propulsores, nossos pés, estão em

contato com o chão, quando o avanço é iniciado, apenas um dos membros serve como fonte de apoio, enquanto o outro avança para a nova posição. Posteriormente, membros invertem os papéis, estabelecendo o ciclo da marcha (PERRY et al., 2005). A Figura 2 exemplifica o ciclo da marcha e sua divisão em etapas.

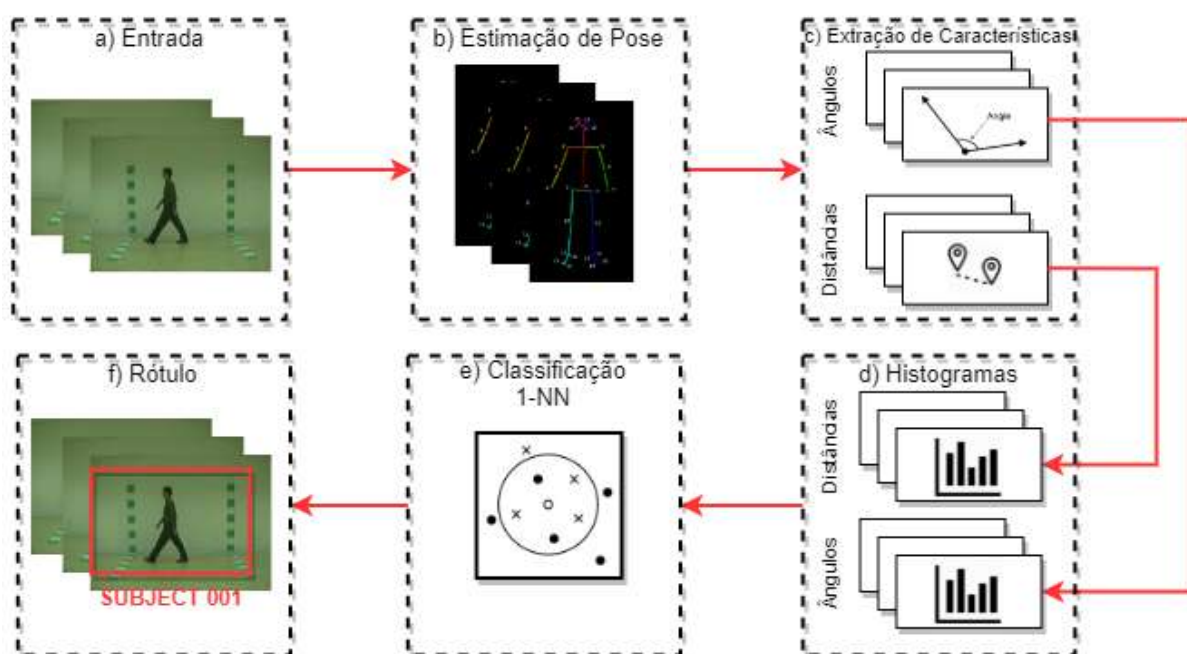
Atualmente, na visão de Chen et al. (2018), os estudos acerca do reconhecimento de marcha podem ser divididos em duas abordagens diferentes: (i) métodos baseados em modelo, os quais visam modelar os parâmetros das articulações humanas, isto é, incorporam informações sobre a estrutura do corpo através de modelos pré-treinados (JAIN; ROSS; NANDAKUMAR, 2011); e (ii) métodos livres de modelo, cuja tática consiste na captura de informações discriminantes a partir das silhuetas dos indivíduos em função do tempo. Elas geram uma assinatura de silhueta, utilizada na identificação do indivíduo (JAIN; ROSS; NANDAKUMAR, 2011).

Os métodos baseados em modelo costumam ser mais custosos computacionalmente, levando em conta que a maioria dos estudos recentes são alicerçados em detectores de pose robustos, a exemplo de aplicações como o OpenPose (CAO et al., 2018) e o OpenPifPaf (KREISS; BERTONI; ALAHI, 2021). Apesar das abordagens livres de modelo se saírem melhor em eficiência computacional e menos complexas, as principais vantagens do uso de modelos se encontra na redução de erros induzidos por obstruções e ruídos em imagens graças ao ganho de robustez, apontam Jain, Ross e Nandakumar (2011).

## 2.3 Método Baseado nas Posições das Partes de Membros

O trabalho desenvolvido por Jangua e Marana (2020) propõe um método baseado em descritores de marcha através da estimação de poses em vídeos de indivíduos caminhando. Primeiramente, o método utiliza o estimador de pose OpenPose (CAO et al., 2018) para extração das poses dos indivíduos de cada quadro do vídeo de entrada. Logo após, para cada quadro, manipula-se as coordenadas dos pontos das juntas detectadas para o cálculo do ângulo de cada parte dos membros em relação ao eixo horizontal da imagem. Da mesma forma, são extraídas as distâncias entre o ponto que representa o pescoço e a reta formada por dois pontos que determinam uma parte de membro. As informações são, então, utilizadas na confecção de dois histogramas, um para cada característica extraída, formando o descritor do comportamento de caminhada do indivíduo durante a sequência, usados como vetores de características de marcha. Por último, ambos os vetores de características são fornecidos a um classificador 1NN para obtenção da pontuação de similaridade e determinação da identidade do indivíduo detectado. A Figura 9 ilustra passo-a-passo o método descrito.

Figura 9 – Diagrama de blocos do método de Jangua e Marana (2020) para reconhecimento de marcha.



Fonte: Jangua e Marana (2020).

### 2.3.1 Multimarcha

Para Chen et al. (2018), estes estudos alcançaram resultados satisfatórios, com bons desempenhos em casos em que pessoas caminham sozinhas, apresentando boa tolerância em diferentes covariáveis e situações adversas. No entanto, considerando situações cotidianas mais realistas, pessoas geralmente caminham acompanhadas de seus membros da família, amigos, colegas de trabalho ou em multidões (CHEN et al., 2018), alterando seu comportamento de andar de maneira substancialmente maior que mudanças causadas por fatores interferentes tradicionais, tais quais vestimenta, acessórios como bolsa, bonés e óculos, configurando um desafio para o reconhecimento. Durante o curso deste trabalho, a marcha alterada por caminhar acompanhado de alguém receberá o nome de multimarcha, enquanto a marcha individual corriqueira será denominada marcha simples.

O objetivo do reconhecimento multimarcha é o de identificar pessoas enquanto caminham com outras pessoas (CHEN et al., 2018). Cunado, Nixon e Carter (1997) declaram que a marcha de um indivíduo é influenciada pela formação de seus ossos, músculos e seus hábitos, sendo difícil de ser modificada depois de formada, tornando a multimarcha factível a partir da existência características comuns entre marcha simples e multimarcha (CHEN et al., 2018).

Ainda que métodos convencionais aplicados a marcha simples não possam ser utilizados para multimarcha, Chen et al. (2018) apontam que a essência deste está na descoberta de características estáveis e invariantes de marcha simples de alto nível fundamentadas nas



características de baixo nível observáveis, que mantenham seu comportamento e não se alterem ao andar acompanhado. Estes pesquisadores complementam que os atributos, uma vez aprendidos, são utilizados para a identificação de cada pessoa e permitem a adaptação de técnicas de marcha simples para o cenário multimarcha, uma vez que está é uma extensão da primeira (CHEN et al., 2018).

## 2.4 Biometrias Suaves

Os primeiros sistemas biométricos desenvolvidos com finalidade de identificação de criminosos, num contexto histórico, eram respaldados em três conjuntos diferentes de características, esclarecem Jain, Ross e Nandakumar (2011). O primeiro conjunto era o das medidas antropométricas, tais quais peso, altura e comprimento dos braços. Outrossim, a aparência e descrições morfológicas da cor da pele, olhos e anomalias únicas de um sujeito, como a presença de tatuagens, marcas peculiares e cicatrizes compunham traços diferenciáveis e capazes de identificar alguém.

As biometrias suas podem ser classificadas em dois grupos: (i) discretos, tais quais a cor dos olhos, a etnia e o tom de pele; e (ii) contínuas, como a medida dos membros, conhecida como antropometria, o peso e a altura (JAIN; ROSS; NANDAKUMAR, 2011).

Os traços de biometria suave são características que codificam alguma informação sobre o indivíduo, mas não possuem distintividade e permanência suficientes para sozinhos determinarem a identidade do indivíduo (JAIN; ROSS; NANDAKUMAR, 2011). Contudo, mesmo diante de suas limitações, as biometrias suaves vêm se destacando e despertando o interesse dos estudiosos da área de Biometria, pois possuem a vantagem permitirem sua integração à sistemas biométricos tradicionais, intuindo melhora no desempenho e acurácia do reconhecimento (DANTCHEVA et al., 2010).

### 2.4.1 Método Baseado em Medidas Antropométricas

O trabalho desenvolvido por Tavares, Marana e Neto (2021) detalha um método que utiliza como traços biométricos as medidas antropométricas. Ele é dividido em cinco etapas, dentre elas: (i) subtração do fundo da imagem; (ii) filtragem morfológica; (iii) estimação de pose; (iv) extração de características; e (v) identificação.

As duas primeiras etapas compreendem o estágio de pré-processamento do método de reconhecimento. Na primeira delas, o fundo da imagem de entrada é removido por meio da técnica de subtração de background, eliminando ruídos, os elementos que podem interferir na detecção das poses. Posteriormente, é aplicado na imagem um filtro morfológico, resultante do processo de subtração do fundo, eliminando os ruídos de alta frequência da imagem resultante.

A próxima etapa é a de estimação de poses, com a finalidade de determinar as posições de todos os membros de cada indivíduo descoberto nas imagens pré-processadas, seguida da extração de características suaves, ocasionando no descritor antropométrico. O descritor é um vetor de características de 8 dimensões, formando pelas medidas dos comprimentos dos úmeros, direito e esquerdo, das ulnas, direita e esquerda, do fêmur, direito e esquerdo, e das tíbias, direita e esquerda.

A última etapa diz respeito a classificação. Um classificador 1NN é empregado para a comparação dos vetores de características gerados, através das funções de distância Euclidiana e City Block. O autor exprime que os melhores resultados são obtidos quando  $K = 1$ .

## 2.5 Multibiometria

Os seres humanos reconhecem uns aos outros com base em evidências apresentadas por múltiplas características biométricas, em conjunto com detalhes contextuais, como por exemplo o conhecimento prévio da presença de alguém em determinado momento e local (JAIN; ROSS; NANDAKUMAR, 2011). O processo de reconhecimento pode ser compreendido como uma reconciliação de indícios pertencentes a múltiplas fontes de informação, apontam Jain, Ross e Nandakumar (2011). Cada fonte, por conta própria, poder ser utilizada por conta própria, mas nem sempre atuam de forma confiável para a realização do reconhecimento biométrico. Todavia, a consolidação das informações apresentadas por cada uma das pistas que compõem o todo pode resultar em uma determinação mais precisa para verificação da identidade de um indivíduo.

Analogamente ao descrito acima, Jain, Ross e Nandakumar (2011) concluem que sistema biométricos podem ser projetados para reconhecer uma pessoa com base nas informações adquiridas de diferentes traços biométricos. Surge, assim, os sistemas multibiométricos.

A multibiometria, formalmente, pode ser definida como a fusão de diferentes tipos de informações em um único sistema de reconhecimento biométrico (ROSS; NANDAKUMAR; JAIN, 2011). Sua principal motivação é a melhora da acurácia de tais sistemas, apostando, em primeiro lugar, no aumento de dimensionalidade do espaço de características e na redução de sobreposições entre distribuições das características de diferentes indivíduos. Jain, Ross e Nandakumar (2011) evidenciam que a combinação de várias fontes biométricas é mais exclusiva para um indivíduo que amostrar biométricas únicas dele.

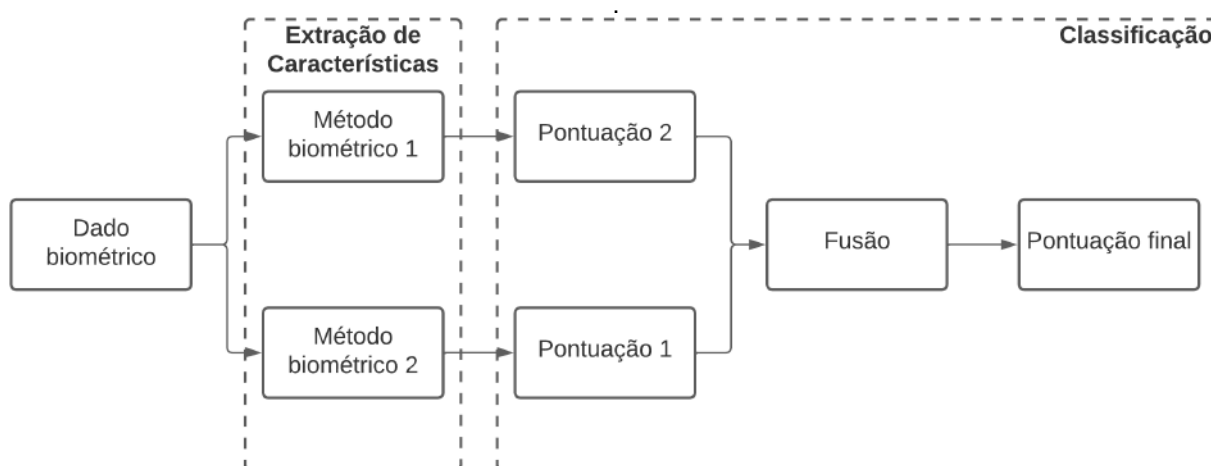
Prosseguem apontando que, em segundo lugar, a presença de ruídos, imprecisões e desvios inerentes, causados por fatores como o envelhecimento, em um subconjunto de características podem ser compensados através das informações discriminatórias fornecidas pelas outras biometrias.

De acordo com Jain, Ross e Nandakumar (2011), a fusão de dois tipos de biometrias distintas pode transcorrer em três diferentes níveis do processo de autenticação de um sistema biométrico:

1. Nível de características: os vetores de características obtidos para cada tipo de biometria após a extração das características podem ser combinados, por exemplo por concatenação, gerando um novo vetor de características de maior dimensionalidade para representar o indivíduo. Esse aumento de dimensão que ocorre com os dados pode ser um aspecto negativo em determinar situações. Nestes casos, técnicas de redução de dimensionalidade podem ser aplicadas para construção de representações mais compactas;
2. Nível de pontuação: as pontuações de similaridade dos descritores analisados para cada tipo de biometria são calculadas e combinadas por meio de operações de soma, máximo, divisões, multiplicações e outras regras definidas pelo sistema, gerando uma nova pontuação de similaridade utilizada na autenticação do usuário;
3. Nível de decisão: funciona com base em um sistema de votos. As pontuações de similaridade dos descritores analisados são calculadas para cada tipo de biometria e definem uma identidade para o indivíduo. A decisão final sobre qual identidade atribuir ao usuário é dada pela maioria de votos, havendo possibilidade de conceder pesos específicos aos pareceres de cada modalidade.

A Figura 10 explica a fusão de dois métodos biométricos em nível de pontuação.

Figura 10 – Fluxo de execução de um sistema biométrico baseado em multibiometria a nível de pontuação.



Fonte: Elaborada pelo autor.

## 2.6 Estimação de Poses

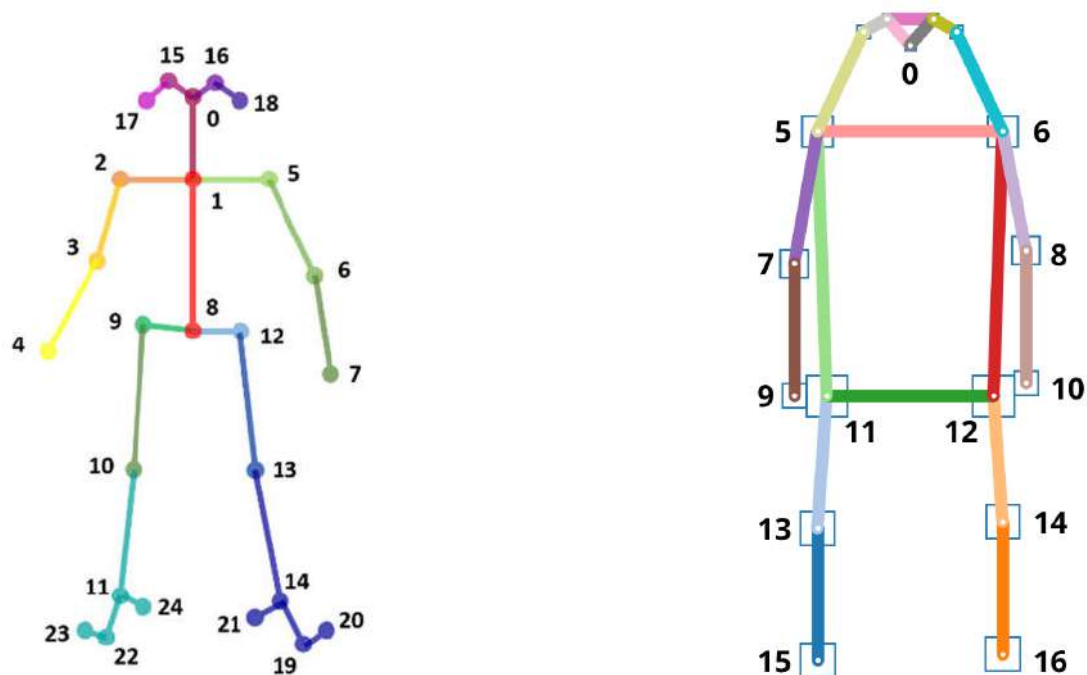
A estimação de pose 2D pode ser definida como o problema de localização de articulações do corpo humano (TOSHEV; SZEGEDY, 2014), dada uma imagem ou vídeo, permitindo que máquinas tenham compreensão das pessoas presentes nos *frames*. Com a localização das articulações, é possível calcular medidas e ângulos de diferentes partes do corpo e utilizá-las como características biométricas para a identificação da pessoa. A estimativa de pose serve também como uma ferramenta fundamental para várias aplicações práticas, como a reidentificação de pessoas (SU et al., 2017) e o reconhecimento de ações humanas (VARGES DA SILVA; MARANA, 2020).

Existem diferentes abordagens para se tratar o problema de estimação de poses 2D, podendo ser agrupadas em métodos *bottom-up* e *top-down*. Contudo, ambas compartilham os mesmos desafios: (i) cada imagem pode conter um número desconhecido de pessoas em diferentes posições e escalas; (ii) as interações entre as pessoas na cena causam interferência devido ao contato ou oclusão, dificultando a associação entre as partes do corpo para gerar a pose; e (iii) a complexidade computacional tende a aumentar de acordo com o número de pessoas na cena.

Os métodos *top-down* utilizam um detector de pessoas e, em seguida, um detector de pose para cada uma das detecções da imagem, o que faz com que a complexidade do método seja proporcional ao número de pessoas detectadas. Os métodos *bottom-up* estimam a posição de cada junção do corpo humano na imagem, e em seguida as associam para formar poses únicas. Esta possui vantagens em relação ao tempo de execução, uma vez que utiliza as informações globais da imagem de maneira direta (CAO et al., 2018).

Dentre os métodos existentes, dois se destacam dos demais devido ao seu desempenho e bons resultados na tarefa de estimação de poses: o (i) Openpose (CAO et al., 2018), permitindo utilizar uma opção de formato de saída com um modelo formado por 25 pontos-chave, vide Figura 11(a), incluindo cabeça, tronco e pés; e o (ii) OpenPifPaf, permitindo a utilização de 17 pontos-chave, 13 deles evidenciadas na Figura 11(b), e não possui opções para detectar informações dos pés. Ambos utilizam a estratégia *bottom-up*, em que estimam a posição de cada junção do corpo humano na imagem e depois as associam para formar poses únicas.

Figura 11 – Modelos de pose 2D usado pelos métodos OpenPose e OpenPifPaf.



(a) Modelo do OpenPose.

Fonte: Cao et al. (2018).

(b) Modelo do OpenPifPaf.

Fonte: Kreiss, Bertoni e Alahi (2021).

### 2.6.1 OpenPose

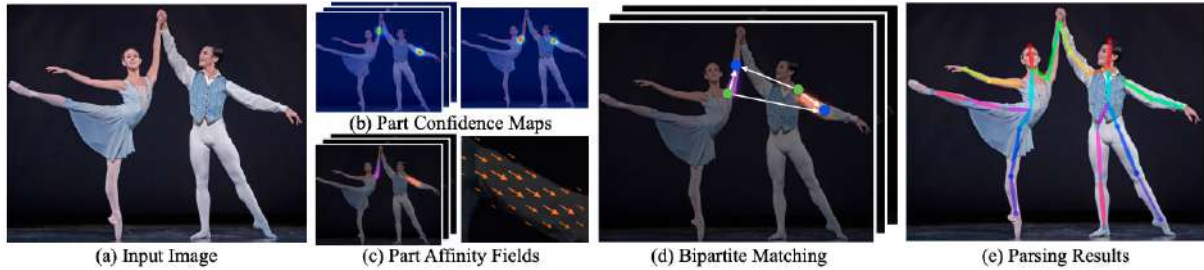
O método OpenPose<sup>1</sup> (CAO et al., 2018) é uma abordagem em tempo real para detecção de pose 2D de múltiplas pessoas em uma imagem, capaz de realizar as detecções com alta acurácia e um bom desempenho computacional, incluindo corpo, pés, mãos e face. O método ganha performance e mantém a acurácia ao utilizar uma abordagem diferente chamada de PAF, *Part Affinity Fields*, do inglês, Campos de Afinidade entre as Partes, ou campos de afinidade, que mapeiam a posição e orientação de membros presentes no domínio da imagem através de um conjunto de vetores de duas dimensões, juntamente com os mapas de confiança, que representam a confiança de existir uma parte de corpo em determinado pixel da imagem. A partir dessa informação global codificada é possível adotar uma estratégia de detecção e associação "gulosa", capaz de produzir resultados precisos por um baixo custo computacional (CAO et al., 2018).

No processo de detecção de pose os campos de afinidade são melhorados iterativamente juntamente com os mapas de confiança por meio de duas CNNs, *Convolutional Neural Networks*, do inglês, Redes Neurais Convolucionais, interligadas, uma para os campos de afinidade e outra para os mapas de confiança, que ao trabalharem juntas melhoram o resultado iterativamente. Em seguida, as partes começam a ser associadas, de maneira que as associações com menor

<sup>1</sup> Disponível em: <<https://github.com/CMU-Perceptual-Computing-Lab/openpose>>. Acesso em 8 de janeiro de 2023.

probabilidade de ocorrer e gerar poses não válidas vão sendo eliminadas iterativamente até atingir o resultado, com os pontos conectados e as poses detectadas (CAO et al., 2018). A Figura 12 ilustra o *pipeline* do método OpenPose.

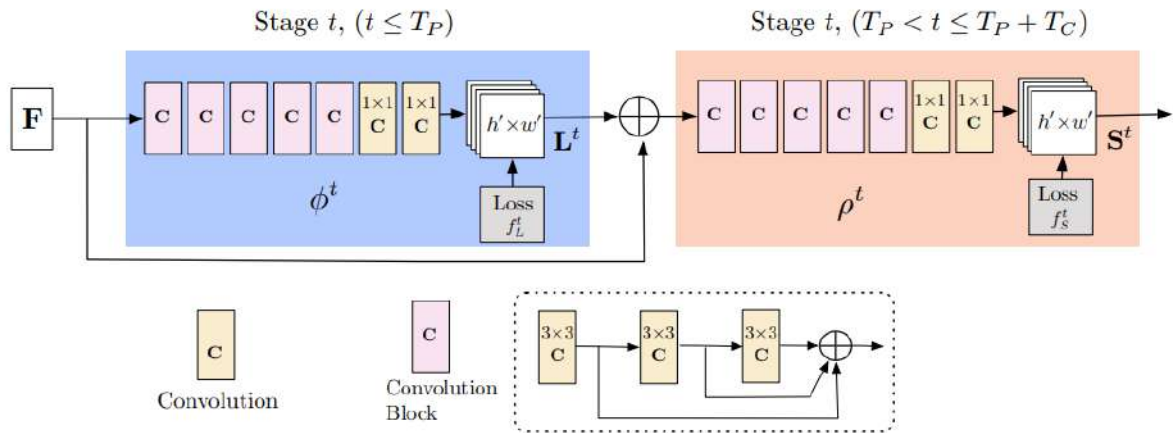
Figura 12 – *Pipeline* do método OpenPose.



Fonte: Cao et al. (2018).

A Figura 13 mostra a arquitetura da CNN que prediz os campos de afinidade (que codificam a associação entre as partes) iterativamente, em azul e, em bege a detecção dos mapas de confiança. As predições são refinadas iterativamente a cada estágio  $t \in \{1, \dots, T\}$ , com supervisão intermediária em cada estágio (CAO et al., 2018).

Figura 13 – Arquitetura da CNN utilizada no OpenPose.



Fonte: Cao et al. (2018).

## 2.6.2 OpenPifPaf

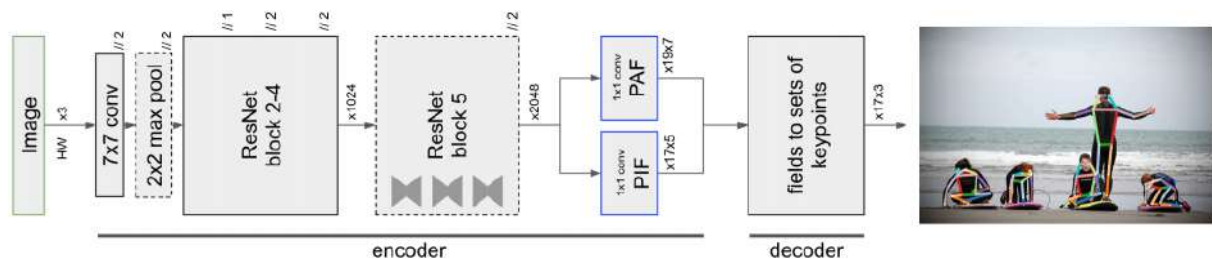
O método OpenPifPaf<sup>2</sup> (KREISS; BERTONI; ALAHI, 2021) é um método de reconhecimento de poses 2D em tempo real capaz de reconhecer poses de múltiplas pessoas presentes em uma imagem. Este método foi pensado para ser aplicado no ramo da navegação autônoma, tendo como principal foco a detecção de pessoas à longas distâncias, baixas resoluções e alto nível de oclusão com o objetivo de aumentar a segurança no trânsito de veículos autônomos e robôs de entrega.

<sup>2</sup> Disponível em: <<https://github.com/openpifpaf/openpifpaf>>. Acesso em 8 de janeiro de 2023.

Nesta abordagem, são usados dois tipos de rede neural para desempenhar dois papéis importantes na detecção das poses. Para cada parte de corpo presente na imagem uma rede prediz a confiança, a posição precisa e o tamanho. Essa informação é chamada de PIF, *Part Intensity Fields*, do inglês, Campos de Intensidade entre as Partes (KREISS; BERTONI; ALAHI, 2021), semelhante aos mapas de confiança utilizados no OpenPose (CAO et al., 2018).

A outra rede utilizada trata da associação entre as partes encontradas pelos campos de intensidade, é chamada de PAF, *Part Association Fields*, do inglês, Campos de Associação entre as Partes, graças à este passo do método, são obtidos resultados que superam as abordagens anteriores quando em imagens de baixa resolução e com diversas oclusões, pois para associar as partes e formar poses humanas completas os campos de associação de partes utilizam uma regressão baseada na função de perda de Laplace, que incorpora uma noção maior de incerteza (KREISS; BERTONI; ALAHI, 2021). A Figura 14 ilustra o *pipeline* do método OpenPifPaf, dada a imagem de entrada a rede neural codifica a imagem nos campos PIF e PAF e em seguida são decodificados formando as poses para cada pessoa na imagem, compostas por pares de coordenadas das junções de cada pessoa acompanhadas de um grau de confiança (KREISS; BERTONI; ALAHI, 2021).

Figura 14 – *Pipeline* do método OpenPifPaf.



Fonte: Kreiss, Bertoni e Alahi (2021).

## 2.7 Rastreamento de objetos

O rastreamento de múltiplos objetos visa prever trajetórias de múltiplos alvos em sequências de vídeo, permitindo sua implementação, muitas vezes de teor crítico, em uma gama de aplicações, desde carros autônomos até análise inteligente de vídeos (ZHANG et al., 2021). Pode ser entendido como um problema de associação de dados, onde o objetivo é associar as detecções de *frames* em sequências de vídeos. Para auxiliar o processo de associação de dados, rastreadores usam vários métodos para modelar o movimento e aparecimento de objetos na cena (BEWLEY et al., 2016).

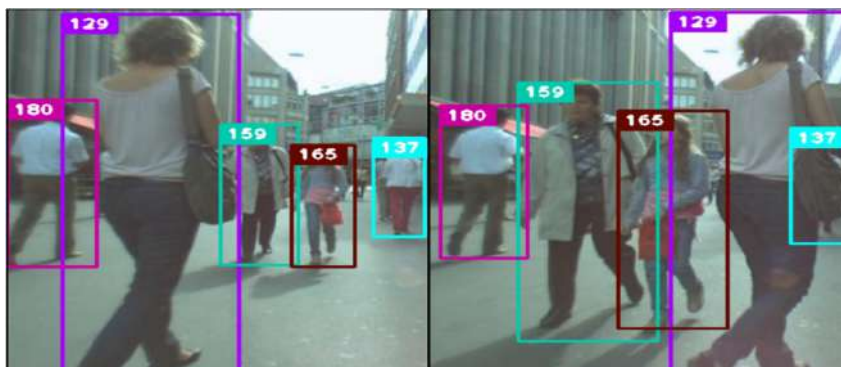
Os sistemas modernos de rastreamento de múltiplos objetos geralmente seguem o paradigma de rastreamento por detecção (WANG et al., 2019). Nele, as trajetórias dos objetos são geralmente encontradas em um problema de otimização global que processa todo lote de

vídeo de uma só vez. Este paradigma é dividido em duas etapas: (i) uma etapa de detecção, na qual os alvos presentes em cada um dos *frames* do vídeo são localizados; e (ii) uma etapa de associação, onde os alvos detectados são atribuídos e conectados a trajetórias existentes. No entanto, devido ao processamento em lote, esses métodos não são aplicáveis em cenários *online* onde uma identidade de destino deve estar disponível a cada *frame* (WOJKE; BEWLEY; PAULUS, 2017).

Existem métodos mais tradicionais capazes de fazer associações *frame a frame*, seja por uma hipótese de estado único, geradas pela ponderação individual e medidas por suas probabilidades de associação, ou com todas as hipóteses possíveis rastreadas e esquemas de poda deve aplicados para tratabilidade computacional. Apesar de apresentarem bons resultados, todavia, o desempenho desses métodos acarreta grande custo computacional e complexidade de implementação.

Para sanar as necessidades de detecções *online* de maneira eficiente, surgiram os métodos SORT (BEWLEY et al., 2016) e DeepSORT (WOJKE; BEWLEY; PAULUS, 2017). São estruturas simples que executam filtragem de Kalman no espaço das imagens e associam os dados quadro a quadro utilizando o método Húngaro com uma métrica de associação capaz de medir sobreposições de *bounding boxes*. Essas abordagens simples obtiveram resultados e desempenhos favoráveis em altas taxas de *frames*. A Figura 15 ilustra o rastreamento dos indivíduos em diferentes frames de vídeo.

Figura 15 – Exemplo de aplicação do método DeepSORT.



Fonte: Wojke, Bewley e Paulus (2017).

### 2.7.1 DeepSORT

O DeepSORT<sup>3</sup>, *Simple Online and Realtime Tracking with a Deep Association Metric*, desenvolvido por Wojke, Bewley e Paulus (2017), é uma extensão do algoritmo SORT, *Simple Online and Realtime Tracking*, de Bewley et al. (2016), que emprega uma abordagem pragmática e simples, utilizando uma combinação de técnicas tais quais o Filtro de Kalman e o Método Húngaro, para solucionar o problema de rastreamento de múltiplos objetos, focando em associar

<sup>3</sup> Disponível em: <[https://github.com/nwojke/deep\\_sort](https://github.com/nwojke/deep_sort)>. Acesso em 8 de janeiro de 2023.



objetos de maneira eficiente para aplicações em tempo real (BEWLEY et al., 2016). A acurácia e precisão de rastreamento observados nos resultados experimentais obtidos por ele competem com os demais métodos de rastreamento, atingindo boa performance em altas taxas de *frames* e redução dos na troca de identidade entre objetos rastreados observadas no SORT (WOJKE; BEWLEY; PAULUS, 2017).

Este algoritmo adota uma metodologia de rastreamento de hipótese única, com filtragem de Kalman recursiva e associação *frame* a *frame* dos dados (WOJKE; BEWLEY; PAULUS, 2017). Wojke, Bewley e Paulus (2017) explicam que o método toma como base as quatro componentes principais da abordagem proposta pelo SORT, incluindo algumas modificações:

- Detecção: a componente de detecção é baseada em CNNs e utiliza o FrRCNN, Faster Region Convolutional Neural Network, um *framework* que opera em duas fases: (i) fase de extração de características e proposição de possíveis regiões contendo objetos; e (ii) fase de classificação dos objetos detectados nas regiões obtidas pelo primeiro processo.
- Modelo de estimação: A representação do movimento incumbida de propagar a identidade atribuída a uma entidade ao próximo *frame* é descrita por um modelo de objeto, definido em um espaço 8-dimensional contendo os parâmetros apresentados na Equação 2.1.

$$M = [x, y, \gamma, h, \dot{x}, \dot{y}, \dot{\gamma}, \dot{h}] \quad (2.1)$$

As quatro primeiras variáveis referem-se à posição central  $(x, y)$ , o comprimento  $\gamma$  e a altura  $h$  da *bounding box*, enquanto os demais parâmetros indicam as respectivas velocidades em coordenadas da imagem associadas as variáveis anteriores. Emprega-se um filtro de Kalman padrão com movimento uniforme e um modelo de velocidade linear que considera as coordenadas  $(u, v, \gamma, h)$  da *bounding box* como observações diretas do estado do objeto.

- Associação dos dados: As associações entre os estados previsto pelo filtro de Kalman e as medidas obtidas em um novo *frame* são feitas pela construção de um problema de atribuição, resolvido pelo Método Húngaro. São integradas informações de movimento e aparência através da combinação de suas respectivas métricas, empregando a distância Mahalanobis e a similaridade por Cosseno. Combinadas, ambas as métricas se complementam, fornecendo informações sobre possíveis localizações de objetos com base em seu movimento e considerando informações de aparência úteis para recuperação de identidade após longos períodos de oclusão.

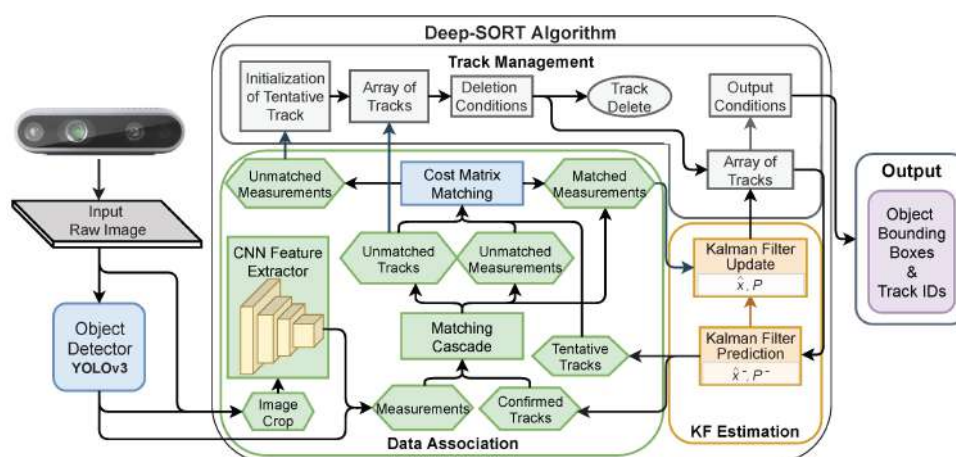
Quando um objeto é ocluído por longos períodos, as previsões subsequentes do filtro de Kalman aumentam a incerteza associada a localização do objeto. Portanto, é introduzido uma correspondência em cascata, dando prioridade aos objetos vistos mais frequentemente.

- Criação e eliminação de identidade de rastreamento: Para cada *bounding box* conta-se o número de *frames* desde sua última associação bem-sucedida. Esse contador é incrementado durante a previsão do filtro de Kalman e resetada para zero quando a *bounding box* é associada a uma detecção. O contador da *bounding box* que exceder uma idade máxima pré-definida  $A_{max}$  indica que o objeto associado a ela deixou a cena, tendo suas informações deletadas.

Novas hipóteses são iniciadas para cada detecção que não pôde ser associada a uma *bounding box* existente, sendo classificadas como provisórias durante os três primeiros *frames* em que são identificadas. É esperado que o algoritmo associe corretamente as métricas da modelagem para cada um desses três *frames*, isto é, detecte o objeto e preveja sua trajetória. As *bounding boxes* que não corresponderem a estes requisitos durante o tempo de prova são apagadas, enquanto as demais são adicionadas a lista de rastreamento.

Em suma, o DeepSORT gera coordenadas dos cantos superior esquerdo e inferior direito de uma *bounding box* associada a um objeto detectado em vídeo, contendo-o nesse intervalo de coordenadas, *frame a frame*. Cada *bounding box* recebe uma identidade de rastreamento exclusiva, um número inteiro positivo único que identifica a entidade atrelada a ela. Essa identidade permanece constante durante toda a sequência de *frames* ou enquanto o elemento permanecer em cena, compondo o rastreamento e permitindo a reidentificação posteriormente. A Figura 16 ilustra o *pipeline* do método DeepSORT.

Figura 16 – Pipeline do método DeepSORT.



Fonte: Pereira et al. (2022).

## 2.8 Métricas de Desempenho

Mesmo os sistemas biométricos mais robustos não são imunes a falhas. A ciência do reconhecimento biométrico é fundamentada em duas premissas importantes quanto as

características biométricas, segundo Jain, Ross e Nandakumar (2011). A primeira premissa é a da exclusividade, atribuída a um traço biométrico quando este é suficiente para a diferenciação de quaisquer dois indivíduos. A segunda premissa é a da permanência, atribuída a um identificador que se conserva o mesmo durante todo o tempo de duração da vida de um indivíduo.

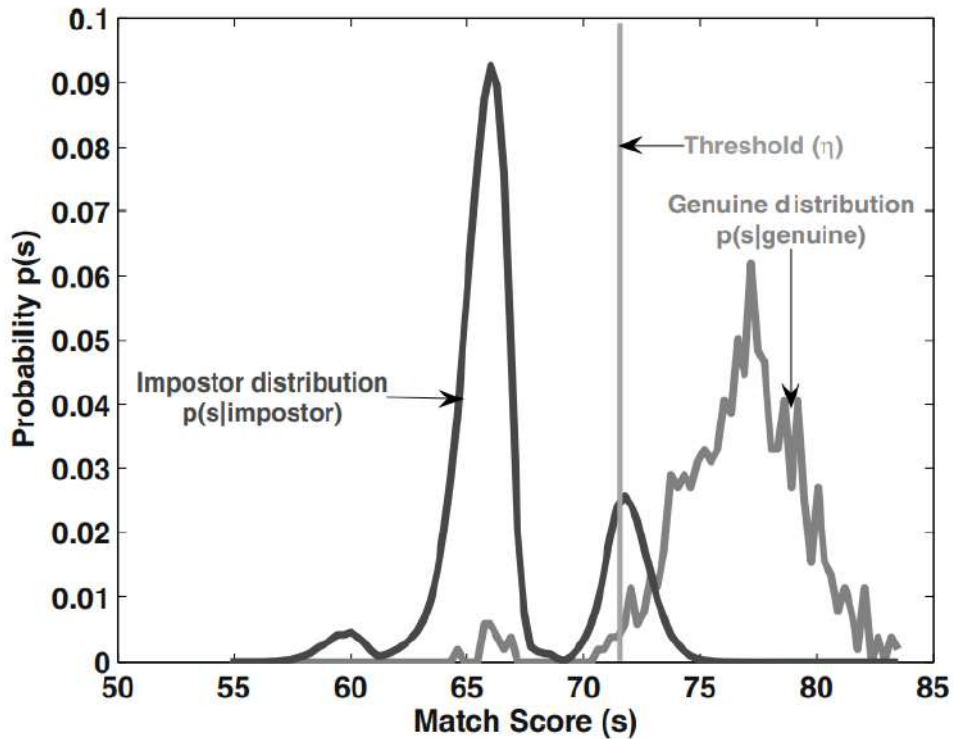
No entanto, ambas as premissas não podem ser garantidas na prática, haja vista que características comportamentais sofrem com mudanças conforme o envelhecimento de uma pessoa, bem como características físicas podem não ser únicas, a exemplo de indivíduos gêmeos (JAIN; ROSS; NANDAKUMAR, 2011). Assim, métricas de avaliação de desempenho capazes de medir a performance de sistemas biométricos auxiliam os administradores e arquitetos de sistemas a determinar a corretude e acurácia de seus métodos.

### 2.8.1 Medidas de Desempenho para Sistemas de Verificação

A análise de desempenho de sistemas baseados em verificação é efetuada, comumente, por duas medidas: (i) a FAR, *False Accept Rate*, do inglês Taxa de Falsa Aceitação; e (ii) a FRR, *False Reject Rate*, do inglês, Taxa de Falsa Rejeição. Ao medir a similaridade das amostras pertencentes a um mesmo indivíduo, a comparação é denominada genuína. Do contrário, a comparação impostora ocorre quando se compara amostras pertencentes a sujeitos diferentes (JAIN; ROSS; NANDAKUMAR, 2011). Seguindo as explicações do autor, pode-se definir a FRR como a fração de pontuações de similaridade genuínas menores que um limiar  $\eta$ . A FAR ocorre na situação inversa, sendo a fração das pontuações impostoras maiores ou iguais a  $\eta$ .

A distribuição das pontuações impostoras e genuínas, separadas pelas limiar  $\eta$ , de um sistema biométrico de verificação é representada pela Figura 17. A partir da representação gráfica da distribuição, é possível calcular a FRR como a área sob a curva das comparações genuínas, exibida do lado esquerdo do limiar. Ele é verdadeiro para a FAR, considerando o lado direito do limiar, levando em conta as comparações impostoras (JAIN; ROSS; NANDAKUMAR, 2011). A Equação 2.2 e a Equação 2.3 representam os cálculos para a FRR e para a FAR, respectivamente.

Figura 17 – Distribuição das comparações genuínas e impostoras de um sistema biométrico de verificação.



Fonte: Jain, Ross e Nandakumar (2011).

$$FRR(\eta) = p(s < \eta \mid \text{genuíno}) = \int_{-\infty}^{\eta} p(s \mid \text{impostor}) ds \quad (2.2)$$

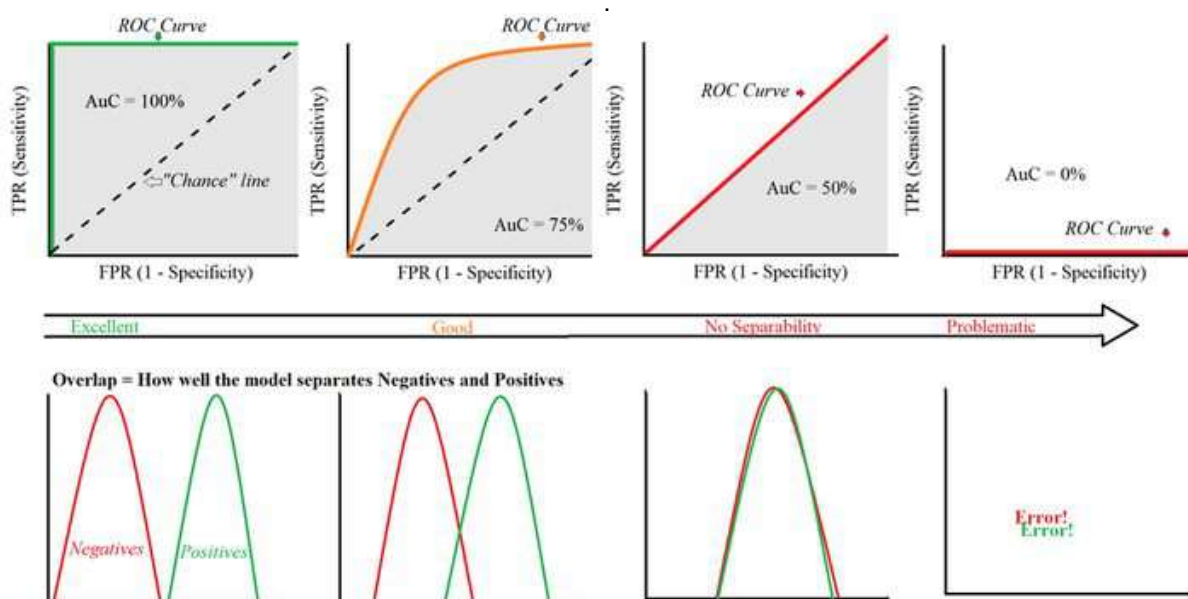
$$FAR(\eta) = p(s \geq \eta \mid \text{impostor}) = \int_{\eta}^{\infty} p(s \mid \text{impostor}) ds \quad (2.3)$$

A GAR, *Genuine Accept Rate*, do inglês, Taxa de Aceitação Genuína, ou ainda TAR, *True Accept Rate*, do inglês, Taxa de Aceitação Verdadeira é uma alternativa para a FRR. Essa forma é definida como a fração das comparações genuínas que foram corretamente aceitas, apresentando pontuação de similaridade maior que o limiar  $\eta$ . A GAR pode ser calculada de acordo com a Equação 2.4 (JAIN; ROSS; NANDAKUMAR, 2011).

$$GAR(\eta) = p(s \geq \eta \mid \text{genuíno}) = 1 - FFR(\eta) \quad (2.4)$$

Jain, Ross e Nandakumar (2011) mencionam ser o nível de segurança requerido por uma aplicação o responsável por determinar o limiar  $\eta$  de um sistema biométrico, o qual acaba sofrendo variações para os diferentes sistemas de verificação. Assim, para os diferentes valores de limiar, as taxas GAR e FAR são calculadas e representadas em uma curva, nomeada ROC, *Receiver Operating Characteristic*, do inglês, Característica Operacional do Receptor. A Figura 18 exibe exemplos de curvas ROC. O eixo FPR corresponde ao GAR, enquanto TPR ao FAR, com suas distribuições representadas nos gráficos logo abaixo cada curva ROC.

Figura 18 – Exemplos de curvas ROC associadas às suas respectivas distribuições de comparação entre genuínos e impostores.



Fonte: Glen (2019).

Destarte, se existem dois sistemas, o A e o B, o primeiro com uma GAR maior e valores de FAR correspondentes para ambos, isto significa que o sistema A apresenta melhor performance que o B. Não obstante, se há uma intersecção das curvas em algum ponto, o sistema B pode performar melhor que o A para certos valores de FAR.

Nestes cenários, também é possível comparar o desempenho desses sistemas estimando a área sob a curva ROC, intitulada AUC, *Area Under the Curve*, do inglês, Área sob a Curva, que varia entre 0.5 e 1. O valor 1 é um indicativo de que o sistema não possui erros (JAIN; ROSS; NANDAKUMAR, 2011).

A performance de um sistema biométrico pode, também, ser analisada por meio de valores únicos, a exemplo da medida EER, *Equal Error Rate*, do inglês, Taxa de Erro Igual. Essa métrica faz referência a um ponto da curva ROC em que a FAR e a FRR se igualam. Quanto menor este valor, melhor a performance do sistema (JAIN; ROSS; NANDAKUMAR, 2011).

## 2.8.2 Medidas de Desempenho para Sistemas de Identificação

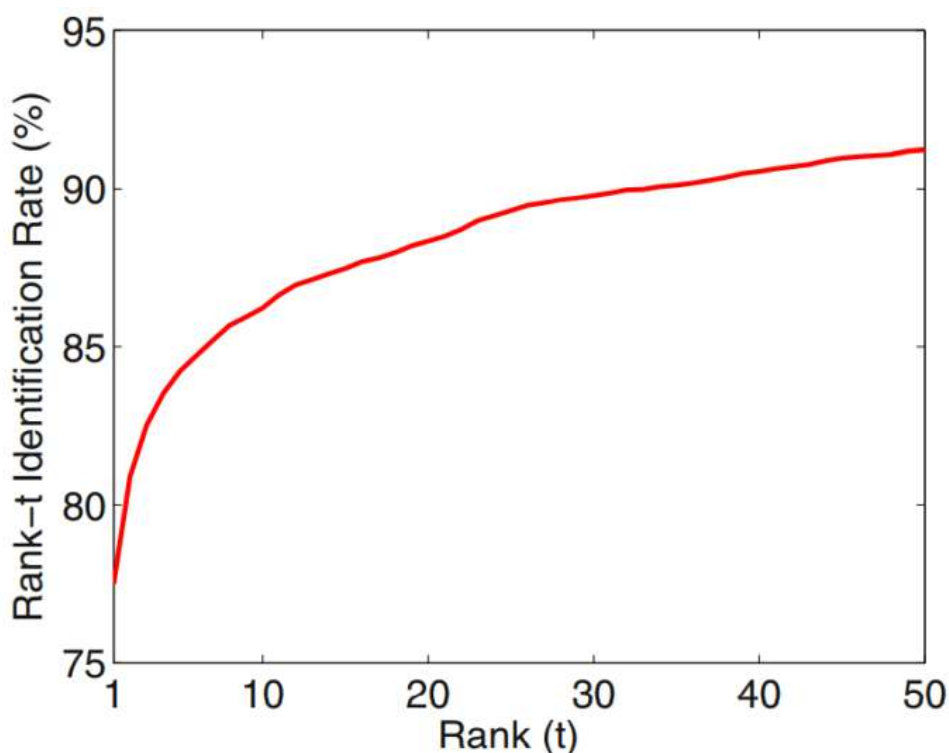
De maneira similar, os sistemas de identificação compreendem dois tipos erros, passíveis de quantificação por meio de medidas de desempenho. Para tal, tendo em conta um sistema biométrico de identificação com  $N$  pessoas registradas no banco de dados que retorna um conjunto de  $t$  identidades, sendo  $1 \leq t \ll N$ , o rank de identificação é definido como a classificação correta da identidade do usuário nas  $t$  primeiras correspondências mais similares (JAIN; ROSS; NANDAKUMAR, 2011).

Como comentam Jain, Ross e Nandakumar (2011), o primeiro erro é mensurado pela taxa de ocorrência FPIR, *False Positive Identification Rate*, do inglês, Taxa de Identificação Falsa Positiva, ocorrendo em ocasiões em que o sistema retorna uma identidade para um usuário não registrado na base de dados. Quando o usuário está registrado no sistema, porém, no processo de identificação, sua identidade não é encontrada dentre as  $t$  identidades retornadas, faz-se referência a FNIR, *False Negative Identification Rate*, do inglês, Taxa de Identificação Falsa Negativa.

Outra medida relacionada a FNIR é a TPIR, *True Positive Identification Rate*, do inglês, Taxa de identificação Verdadeira Positiva. Ela mede a fração de indivíduos registrados no sistema que tem sua identidade verdadeira retornada no conjunto de identidades de saída, se relacionando com a FNIR da seguinte maneira:  $FNIR = 1 - TPIR$  (JAIN; ROSS; NANDAKUMAR, 2011). Ainda segundo os autores, o TPIR pode ser chamado de taxa de identificação rank- $t$ . Um caso particular da taxa rank- $t$  é quando o  $t = 1$ , recendo a nomenclatura de acurácia *rank-1*, uma métrica bastante utilizada na avaliação de desempenho de sistemas de identificação.

O cálculo de diferentes valores de  $t$  para a taxa identificação rank- $t$  fornece a curva CMC, *Cumulative Match Characteristic*, do inglês, Característica de Correspondência Cumulativa. É uma curva que plota a taxa de identificação rank- $t$  em função de  $t = 1, 2, \dots, N$ , com  $N$  sendo o número de usuários registrados no sistema (JAIN; ROSS; NANDAKUMAR, 2011). A Figura 19 fornece um exemplo de curva CMC.

Figura 19 – Exemplo de uma curva CMC.



Fonte: Jain, Ross e Nandakumar (2011).

## 2.9 Funções de distância

Uma função de distância é definida dentro de um espaço métrico, podendo ser empregada para medir a similaridade entre vetores de características biométricas. Segue sua definição formal: Dado um conjunto  $S$  e três pontos,  $p, q$  e  $r \in S$ , uma função de distância em  $S$  é definida por  $d : S \times S \rightarrow \mathbb{R}$  com as seguintes propriedades:

1. Simetria:  $d(p, q) = d(q, p)$ ;
2. Não negatividade:  $d(p, q) > 0$  se  $p \neq q$  e  $d(p, q) = 0$  se  $p = q$ ;
3. Desigualdade triangular:  $d(p, q) \leq d(p, r) + d(q, r)$ .

Dentre as funções de distâncias existente, as subseções seguintes evidenciam a função de distância Euclidiana e a função de distância Qui-Quadrado.

### 2.9.1 Distância Euclidiana

Considerando dois pontos,  $p$  e  $q$ , pertencentes a um espaço vetorial  $n$ -dimensional, isto é,  $p = \{p_1, p_2, \dots, p_n\}$  e  $q = \{q_1, q_2, \dots, q_n\}$ , a distância Euclidiana  $D$  entre os pontos  $p$  e  $q$  é dada pela Equação 2.5.

$$D(p, q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (2.5)$$

### 2.9.2 Distância Qui-Quadrado

Considerando dois pontos,  $p$  e  $q$ , pertencentes a um espaço vetorial  $n$ -dimensional, isto é,  $p = \{p_1, p_2, \dots, p_n\}$  e  $q = \{q_1, q_2, \dots, q_n\}$ , a distância Qui-Quadrado  $\chi^2$  entre os pontos  $p$  e  $q$  é dada pela Equação 2.6.

$$\chi^2(p, q) = \frac{1}{2} \sum_{i=1}^n \frac{(p_i - q_i)^2}{p_i + q_i} \quad (2.6)$$

Vale ressaltar que a distância Qui-Quadrado não é uma métrica linear, realçando grandes discrepâncias entre os vetores de características de entrada, sendo amplamente utilizada para a comparação de histogramas (YANG et al., 2015).

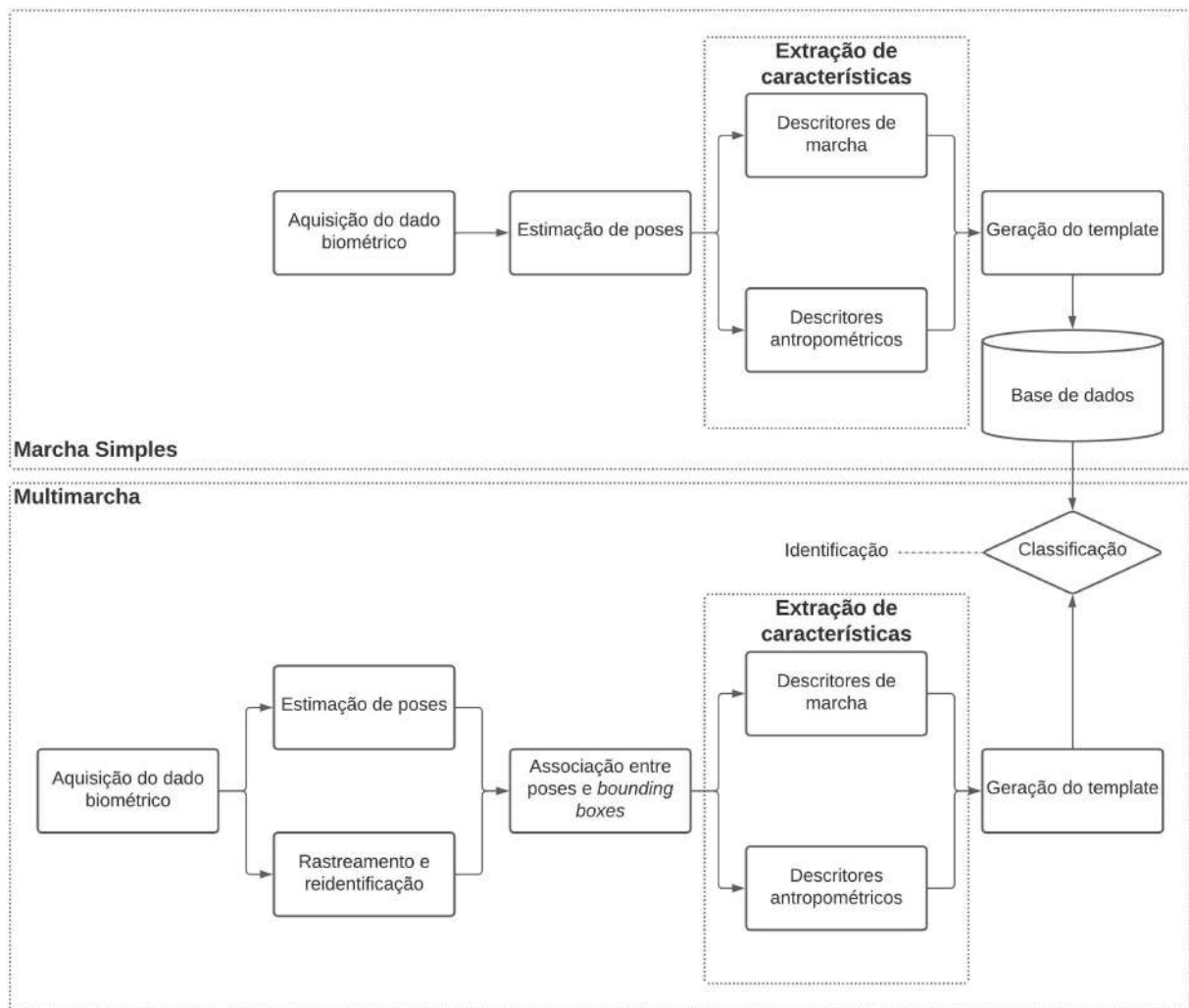
### 3 Método Proposto

Neste capítulo é apresentado o método proposto pelo trabalho. Ele une modificações das abordagens de Jangua e Marana (2020) para reconhecimento biométrico a partir de descritores de marcha e de Tavares, Marana e Neto (2021) para reconhecimento biométrico a partir de descritores antropométricos, aliados à métodos de estimação de poses e métodos de rastreamento de objetos para composição do esqueleto humano e determinação de sua localização em cada *frame* do vídeo capturado como dado biométrico.

A Figura 20 exibe as etapas do método, desde a aquisição do dado biométrico, em forma de vídeo, da marcha simples de um indivíduo para extração dos descritores de marcha e antropométricos, utilizados na composição da base de dado, até a classificação dos indivíduos em condições de multimarcha, onde lhes é atribuída uma identificação, são rastreados e associados aos pontos-chave da estimação de pose.



Figura 20 – Pipeline do método proposto.



Fonte: Elaborada pelo autor.

As subseções seguintes abordam com mais detalhes as etapas demonstradas na Figura 20.

### 3.1 Estimação de Poses

Já em posse dos dados biométricos, o primeiro passo do método é a estimativa de pose. Nele, método estimador escolhido é aplicado ao dado biométrico recebido como entrada, neste caso um vídeo da caminhada do indivíduo, e, para cada *frame*, são obtidas coordenadas dos pontos-chave do esqueleto humano, em concordância com o formato demonstrado na Figura 11(a), se o método estimador for o OpenPose, ou na Figura 11(b), caso o método estimador for o OpenPifPaf. O algoritmo tem como saída um arquivo em formato JSON, contendo, para cada *frame*, as coordenadas verticais e horizontais de cada ponto.

## 3.2 Rastreamento e Reidentificação

O rastreamento e reidentificação ocorrem conjuntamente com a estimação de poses quando os vídeos de entrada apresentam mais de um indivíduo. Nesse passo, o DeepSORT (WOJKE; BEWLEY; PAULUS, 2017) é aplicado ao vídeo de entrada e, para cada *frame*, obtém-se as *bounding boxes* da localização do indivíduo. A *bounding box* é um retângulo definido por dois pontos,  $P1 = (x_{P1}, y_{P1})$  e  $P2 = (x_{P2}, y_{P2})$ , indicando seu começo e fim, em pixels. A saída do algoritmo corresponde a um arquivo em formato CSV, que contém, para cada quadro do vídeo, a localização das *bounding boxes*, bem como uma identificação numérica única para cada entidade detectada.

## 3.3 Associação entre Poses e *Bounding Boxes*

As coordenadas do corpo e sua localização são atribuídas para todo elemento detectado em vídeo, sendo necessário associar cada conjunto de pontos-chave detectados com a identificação numérica única correta atribuída a cada indivíduo. Este problema caracteriza a etapa de associação. Nela, são analisadas as coordenadas dos pontos-chave e as coordenadas das *bounding boxes* de cada uma das entidades, *frame a frame*.

Para toda parte de membro, calcula-se a distância euclidiana de sua coordenada até as extremidades das *bounding boxes* encontradas para cada uma das entidades detectadas no vídeo, resultando, portanto, em quatro distâncias: (i)  $D_1$ , a distância do membro até o canto superior direito da *bounding box*, (ii)  $D_2$ , a distância do membro até o canto superior esquerdo da *bounding box*, (iii)  $D_3$ , a distância do membro até o canto inferior direito da *bounding box*, e (iv)  $D_4$ , a distância do membro até o canto inferior esquerdo da *bounding box*. O cálculo das distâncias é feito para todas as *bounding boxes* que compõe o quadro.

A distância de maior valor entre as quatro é escolhida, concatenando-a a um vetor de distâncias daquela sequência de vídeo, junto da identificação atribuída para a entidade que pertence àquela *bounding box*. Dentre os elementos desse vetor, é escolhido aquele com a menor distância, indicando que os pontos-chaves dos quais o membro selecionado pertence diz respeito ao indivíduo cuja distância entre a coordenada do pontos-chave daquele membro e a *bounding box* é mínima. Assim, concatena-se os pontos-chave a um vetor para esse sujeito, gerando um vetor bidimensional, em que a primeira dimensão corresponde a um indivíduo  $x$ , enquanto a segunda dimensão armazena todos os pontos-chave.

### 3.4 Extração das Características

A extração de características separa os pontos-chave mais importantes para a análise, ou seja, aqueles indicativos para a biometria escolhida, codificando uma maior quantidade de informações. Estes pontos são o mesmo para a formação dos descritores de marcha e antropométricos, e formam os membros superiores, os braços e antebraços, e os membros inferiores, as coxas e as pernas, do esqueleto do indivíduo, totalizando oito medidas.

Como o OpenPifPaf não apresenta anotações para o ponto do pescoço, sua utilização foi substituída pelo ponto do nariz quando aplicado.

Os cálculos para a formação de cada descritor são apresentados nas subseções a seguir.

#### 3.4.1 Descritores de Marcha Baseados no Método de Jangua e Marana (2020)

Dados dois pontos  $P_1 = (x_{P1}, y_{P1})$  e  $P_2 = (x_{P2}, y_{P2})$  que formam uma parte de membro, representa-se essa parte como um vetor bidimensional  $w = (x_1, y_1)$ , onde  $(x_1, y_1) = P_1 - P_2$ . Assim, o ângulo  $\varphi$ , entre o vetor  $w$  e o eixo horizontal  $(x_2, y_2) = (1, 0)$  pode ser obtido utilizando a Equação 3.1.

$$\varphi = \arccos \frac{x_1 * x_2 + y_1 * y_2}{\sqrt{x_1^2 + y_1^2} * \sqrt{x_2^2 + y_2^2}} \quad (3.1)$$

Analogamente à sequência de ângulos, a sequência de distâncias para cada parte de membro é formada pela distância  $d$  entre a reta que a parte do corpo está contida e o ponto do pescoço em cada *frame*. Considerando o vetor  $v = P_{neck} - P_2$ , onde  $P_{neck}$  é o ponto do pescoço e  $P_2$  sendo ainda um ponto que forma uma parte de membro, essa distância pode ser calculada utilizando a Equação 3.2 e a Equação 3.3.

$$Proj_w v = \left( \frac{v \cdot w}{\|w\|^2} \right) * w \quad (3.2)$$

$$d = \|v - Proj_w v\| \quad (3.3)$$

#### 3.4.2 Descritores Antropométricos Baseados no Método de Tavares, Marana e Neto (2021)

Os descritores antropométricos são calculados e gerados a partir do cálculo da distância entre as duas articulações que formam um membro do corpo do indivíduo. Por exemplo, para

o cálculo do comprimento do antebraço direito, basta calcular a distância entre o ponto do ombro direito e o ponto do cotovelo direito.

A função de distância Euclidiana é empregada para o cálculo das distâncias. Todos os comprimentos obtidos são normalizados, dividindo sua medida pela distância entre as articulações do pescoço e do quadril, que formam o tronco.

Vale ressaltar que o esqueleto estimado pelo OpenPifPaf não possui anotações para a posição do pescoço e do quadril. Portanto, para efetuar a normalização, são calculadas as distâncias, para cada *frame*, entre o nariz e o quadril direito e o nariz e o quadril esquerdo, sendo utilizada a maior distância entre essas duas. A Figura 21 exibe as medidas antropométricas extraídas pelo método.

Figura 21 – Medidas antropométricas obtidas das poses 2D dos indivíduos, utilizadas pelo método de Tavares, Marana e Neto (2021).



Fonte: Tavares, Marana e Neto (2021).

### 3.5 Geração de *Templates*

Logo após a extração das características, são construídos três histogramas, um para os ângulos, outro para as distâncias e o último para os comprimentos antropométricos, representando o comportamento global destes traços durante a sequência da marcha. Os histogramas são construídos com 16 *bins* cada, valor encontrado empiricamente nas pesquisas de Jangua e Marana (2020) e Tavares, Marana e Neto (2021). Cada histograma é gerado a partir da concatenação de 8 histogramas, um para cada parte de membro considerada, formando um único vetor de características. Estes histogramas intermediários são normalizados, dividindo sua medida pela quantidade de medidas de cada traço.

Assim, o descritor de marcha é constituído por dois vetores compostos de 128 características, enquanto o descritor antropométrico é constituído por um único vetor de 128 características.

### 3.6 Classificação

A classificação é realizada aplicando um classificador 1NN. O algoritmo opera da seguinte forma: primeiramente, são calculadas, a partir de uma função de distância, as pontuações de similaridade, entre o novo descritor obtido na detecção e os descritores rotulados armazenados na base de dados. O rótulo que apresentar menor distância, isto é, maior similaridade com o descritor para classificação, é atribuído ao indivíduo.

O cálculo da distância entre duas amostras depende do descritor escolhido. Para os descritores de marcha, a distância  $d_m$  de uma amostra a outra é dada pela média aritmética entre a distância dos vetores dos ângulos  $d_1$  e dos vetores de distâncias  $d_2$ , dada pela expressão  $d = (d_1 + d_2)/2$ . Para os descritores antropométricos,  $d_a$  é a distância obtida ao aplicar a função de distância nas amostras. Por último, para a fusão dos descritores, a distância  $d_f$  é dada pela expressão  $d_f = (d_m + d_a)/2$ . Nenhuma das distâncias precisa ser normalizada, visto que os histogramas foram normalizados na etapa anterior.

## 4 Resultados Experimentais

Este capítulo compreende a primeira etapa do trabalho, apresentando os resultados obtidos pelos métodos propostos, com a finalidade de determinar o melhor método a ser aplicado no sistema biométrico, segundo as métricas apresentadas na seção 2.8.

### 4.1 Experimentos

A implementação do método proposto e a realização dos experimentos foram implementadas na linguagem de programação Python. Atualmente, a linguagem Python é uma das mais utilizadas para estudos e pesquisas em aprendizado de máquina, ciência de dados, reconhecimento de padrões e computação científica, de fácil aprendizagem e acesso, contando com uma documentação ampla e didática.

Como visto na explicação do Capítulo 3, o método proposto é fundamentado em métodos de reconhecimento biométrico baseados em modelo. Portanto, uma das etapas que compreendem o método para permitir que a extração de características e a geração dos *templates* para classificação seja efetuada é a de estimação de pose, fornecendo as localizações de pontos de interesse nas imagens das bases de dados. Os dois métodos de código aberto escolhidos para este trabalho foram o OpenPose (CAO et al., 2018) e o OpenPifPaf (KREISS; BERTONI; ALAHI, 2021). Os dois algoritmos possuem utilidade prática bem definida e apresentaram bons resultados quando empregados na tarefa de reconhecimento, basta ver os resultados dos trabalhos de Jangua e Marana (2020) e Tavares, Marana e Neto (2021). Via de regra, para todos os experimentos realizados, são analisadas a performance do método em diferentes bases de dados utilizando os descritores de marcha e antropométricos, assim como a fusão desse dois a nível de pontuação para as duas funções de distância, Euclidiana e Qui-Quadrado, mencionadas no Capítulo 2, fazendo uso de um classificador 1NN.

O poder de processamento demandado para execução do método de estimação de pose em tempo real é alto e requer o auxílio de aceleração gráfico. Portanto, o emprego de uma GPU, *Graphics Processing Unit*, do inglês, Unidade de Processamento Gráfico é fundamental. Contudo, tal *hardware* não estava disponível para a realização dos experimentos. Assim, optou-se por executar o processamento em nuvem, com o auxílio de uma instância do Google Colaboratory.

O Colab, como é comumente chamado, é um projeto com a intenção de disseminar a pesquisa e o ensino em aprendizagem de máquina. Para tal, baseia-se em *IPython*, uma ferramenta *Open Source* com o objetivo de criar um ambiente abrangente para computação interativa e exploratória (PÉREZ; GRANGER, 2007). O *IPython* é um *shell* de comandos para

computação interativa em várias linguagens de programação, originalmente desenvolvido para a linguagem de programação Python. Além disso, os ambientes de execução e desenvolvimento são pré-configurados com pacotes essenciais para algoritmos de visão computacional, ciência de dados e aprendizado de máquina instalados.

Um Colab Notebook pode ser executado localmente ou em nuvem, sendo composto por múltiplas células executadas individualmente. Cada uma dessas células pode hospedar um código ou um texto em formato *markdown*. No caso dos códigos, eles normalmente incluem tabelas, gráficos, textos e imagens. Essa tecnologia facilita a replicação e compartilhamento de experimentos científicos (CARNEIRO et al., 2018). O autor explica que a infraestrutura da máquina virtual que executa o código é hospedada pela plataforma de computação em nuvem Google Cloud, e é alocada por um período limitado, variando de 6 até 12 horas.

Para a visualização gráfica das curvas CMC, foi utilizado o pacote Matplotlib (HUNTER, 2007). O Matplotlib é um pacote gráfico que permite integração com GUI, *Graphical User Interface*, do inglês, Interface Gráfica do Usuário, para desenvolvimento de aplicativos, suporta diferentes plataformas, oferece rasterização e vetorização de imagens com altíssima qualidade e fornece suporte para expressões matemáticas, trabalhado de forma interativa com o *shell* (CARNEIRO et al., 2018).

#### 4.1.1 Bases de Dados

Para o desenvolvimento destes experimentos foram utilizadas duas bases de dados públicas: a CASIA Gait Dataset-A (WANG et al., 2003) e a SMVDU Multi-Gait (SINGH et al., 2019). Ambas as bases são compostas por vídeos de indivíduos (homens e mulheres) caminhando em diferentes ângulos e situações.

##### 4.1.1.1 CASIA Gait Dataset-A

A CASIA Gait Dataset-A, criada em 2001, inclui imagens de 20 indivíduos caminhando em um ambiente externo, cada um com 12 sequências de imagem, variando em quantidade entre 37 e 127 imagens por sequência, sendo 4 sequências para cada ângulo de visão da câmera: 0, 45 e 90 graus, que representam as direções lateral, oblíqua e frontal da marcha de um indivíduo, respectivamente. A Figura 22 mostra exemplos dos ângulos de visão presentes da base de dados (WANG et al., 2003).

Figura 22 – Exemplos de vídeos da CASIA Gait Dataset-A.



Fonte: Wang et al. (2003).

#### 4.1.1.2 SMVDU-Multi-Gait

A SMVDU-Multi-Gait Dataset (SINGH et al., 2019) inclui vídeos de 20 indivíduos caminhando em um ambiente externo, cada um com 6 sequências de vídeo, variando o ângulo de visão da câmera em 0, 45 e 90 graus, representando as visões lateral, oblíqua e frontal de caminhada, respectivamente. A base de dados é dividida em duas categorias de dados, a SMVDU-Single-Gait e a SMVDU-Multi-Gait. Ambas as categorias são compostas por vídeos de indivíduos caminhando em diferentes ângulos e situações, individualmente ou em grupo. A Figura 23 e a Figura 24 mostram exemplos de quadros de vídeos desta base de dados, em diferentes ângulos de visão.

Figura 23 – Exemplos de vídeos da SMVDU Multi-Gait em situação de marcha simples.



Fonte: Singh et al. (2019).

Figura 24 – Exemplos de vídeos da SMVDU Multi-Gait em situação multimarcha.



Fonte: Singh et al. (2019).

A base de dados possui o objetivo de facilitar a análise do impacto do padrão da marcha enquanto se caminha sozinho ou acompanhado, facilitar a reconstrução da região de oclusão e a geração de características de marcha das entidades, possibilitando seu reconhecimento



e permitindo estudar o efeito da variação da visão da câmera em diferentes ângulos em um cenário de marcha de múltiplas pessoas.

#### 4.1.2 Métricas

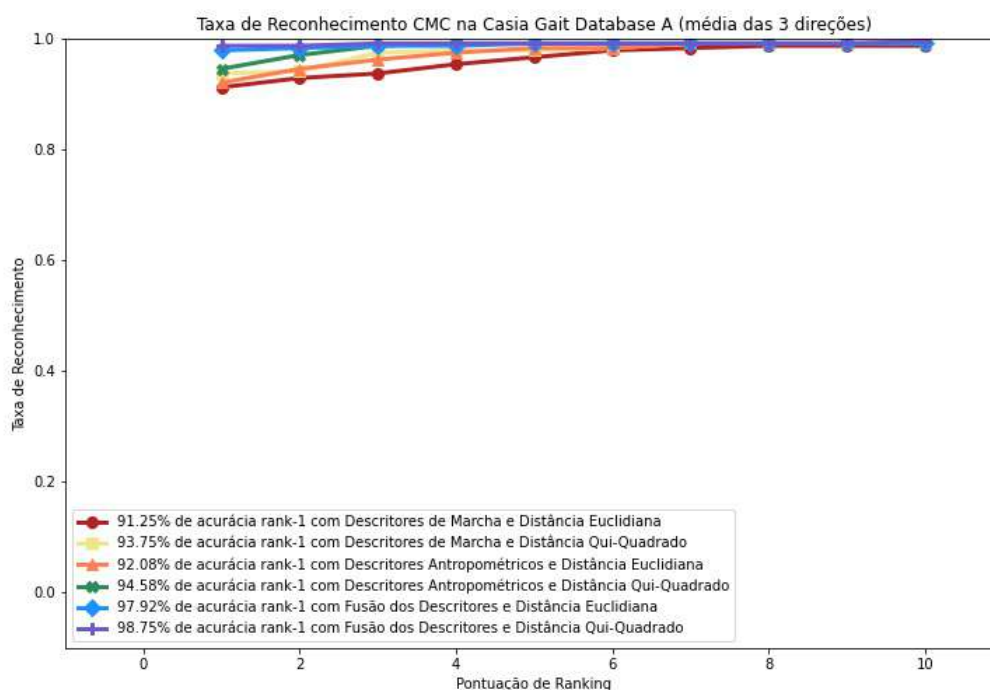
O Capítulo 2 apresenta as métricas de desempenho para sistemas biométricos com as funcionalidades de verificação e identificação. Os processos experimentais executados por este trabalho têm por objetivo a identificação de múltiplas pessoas em situação de marcha, portanto os avaliadores de desempenho empregados as medidas para tal finalidade, compreendo as curvas CMC, bem como a situação particular da taxa rank-t, acurácia rank-1.

### 4.2 Resultados e Discussão

O primeiro experimento foi direcionado a CASIA Gait Dataset-A (WANG et al., 2003) utilizando o método proposto, com o estimador de pose OpenPose, com a finalidade de determinar a corretude do método em situações de marcha simples. Para medir o desempenho dos descritores, foi empregada a curva CMC, usando uma precisão média obtida por meio da média das taxas de reconhecimento para as três diferentes direções de marcha: lateral ( $90^\circ$  em relação a posição da câmera), oblíqua ( $45^\circ$  em relação a posição da câmera) e frontal ( $0^\circ$  em relação a posição da câmera), totalizando 240 sequências de caminhada, 80 para cada direção.

A Figura 25 mostra os resultados desse experimento. Pode-se observar que a distância Qui-Quadrado apresentou melhor desempenho em relação a distância Euclidiana para todos os descritores considerados em ranks menores que 6, obtendo resultados muito próximos de 6 em diante. Estes corroborando outros estudos que indicam que a função Qui-Quadrado é uma boa métrica para comparação de histogramas, segundo Yang et al. (2015). Para mais, dentre os descritores analisados, ficou evidente que a fusão de características de marcha e antropométricas apresentou o melhor desempenho para ambas as funções de distância, em consonância com o exposto no Capítulo 2 de que a combinação de fontes biométricas é mais exclusiva para uma pessoa que amostras isoladas de cada biometria, melhorando o desempenho do método.

Figura 25 – Curvas CMC obtidas pelas distâncias Euclidiana e Qui-Quadrado para os diferentes descritores na CASIA Gait Dataset-A.



Fonte: Elaborada pelo autor.

A Tabela 1 mostra os valores para acurácia rank-1 obtidos pelos métodos de Lima e Schwartz (2019), Jangua e Marana (2020), Tavares, Marana e Neto (2021) e pelo método proposto neste trabalho para cada direção de marcha na CASIA Gait Dataset-A.

Tabela 1 – Acurácia Rank-1 - CASIA Gait Dataset-A.

Método	Lateral	Oblíquo	Frontal
Lima e Schwartz (2019)	<b>92.50%</b>	<b>96.25%</b>	<b>97.50%</b>
Jangua e Marana (2020) - Euclidiana	80.00%	87.50%	96.25%
Jangua e Marana (2020) - Qui-Quadrado	87.50%	92.50%	95.00%
Tavares, Marana e Neto (2021) - Euclidiana	78.75%	85.00%	90.00%
Tavares, Marana e Neto (2021) - Qui-Quadrado	85.00%	91.25%	88.75%
Descritores de Marcha - Euclidiana	87.50%	88.75%	97.50%
Descritores de Marcha - Qui-Quadrado	92.50%	90.00%	98.75%
Descritores Antropométricos - Euclidiana	91.25%	91.25%	93.75%
Descritores Antropométricos - Qui-Quadrado	92.50%	95.00%	96.25%
Fusão dos Descritores - Euclidiana	96.25%	<b>97.50%</b>	<b>100.00%</b>
Fusão dos Descritores - Qui-Quadrado	<b>98.75%</b>	<b>97.50%</b>	<b>100.00%</b>

Fonte: Elaborada pelo autor.

Pode-se observar que o comportamento do método proposto para situações de marcha simples foi como esperado e sua performance compete com aquelas dos métodos consolidados para reconhecimento baseado em poses, apresentando os melhores resultados entre todos os

métodos para todas as três direções de marcha. Novamente, os resultados mais robustos são obtidos para função de distância Qui-Quadrado e fusão dos descritores biométricos.

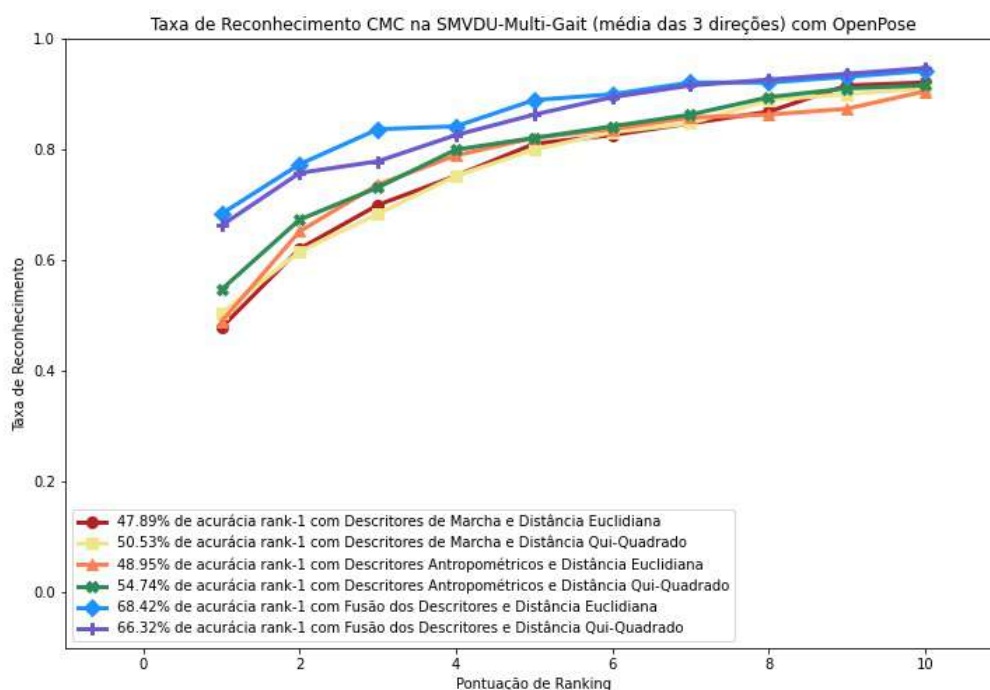
O segundo experimento foi conduzido na SMVDU-Multi-Gait com a finalidade de medir o desempenho do método proposto em situações multimarcha, sua real aplicação. Para tal efeito, utilizou-se a curva CMC, usando uma precisão média obtida através da média das taxas de reconhecimento para as três diferentes direções de marcha: lateral, oblíqua e frontal.

Outrossim, o experimento visou definir qual estimador de pose apresentou resultados mais robustos nas configurações adotadas para o método, confrontando os valores alcançados nos testes de reconhecimento para o OpenPose (CAO et al., 2018) e o OpenPifPaf (KREISS; BERTONI; ALAHI, 2021). Assim, o experimento foi dividido em duas etapas.

#### 4.2.1 OpenPose

A primeira etapa desse experimento aplica o método OpenPose como estimador de poses na composição do fluxo da aplicação do método proposto. A Figura 26 mostra os resultados desse experimento. Pode-se observar que, assim como ocorreu no primeiro experimento, os resultados mais robustos foram obtidos para a fusão dos descritores biométricos. Todavia, a distância Euclidiana apresentou os melhores resultados para pontuações de ranking menores que 6, diferentemente do experimento anterior, em que a função Qui-Quadrado foi superior. Para o rank 6 e acima, a taxa de reconhecimento é praticamente a mesma.

Figura 26 – Curvas CMC obtidas pelas distâncias Euclidiana e Qui-Quadrado para os diferentes descritores na SMVDU-Multi-Gait com estimador de poses OpenPose.



Fonte: Elaborada pelo autor.

A Tabela 2 mostra os valores para acurácia rank-1 obtidos pelo método proposto para cada direção de marcha na SMVDU-Multi-Gait, considerando o estimador de poses OpenPose.

Tabela 2 – Acurácia Rank-1 - SMVDU-Multi-Gait com o estimador de poses OpenPose.

Método	Lateral	Oblíquo	Frontal
Descritores de Marcha - Euclidiana	34.72%	48.57%	70.83%
Descritores de Marcha - Qui-Quadrado	37.50%	57.14%	77.08%
Descritores Antropométricos - Euclidiana	29.17%	52.86%	68.75%
Descritores Antropométricos - Qui-Quadrado	34.72%	58.57%	62.50%
Fusão dos Descritores - Euclidiana	<b>51.39%</b>	71.43%	<b>89.58%</b>
Fusão dos Descritores - Qui-Quadrado	<b>51.39%</b>	<b>72.86%</b>	79.17%

Fonte: Elaborada pelo autor.

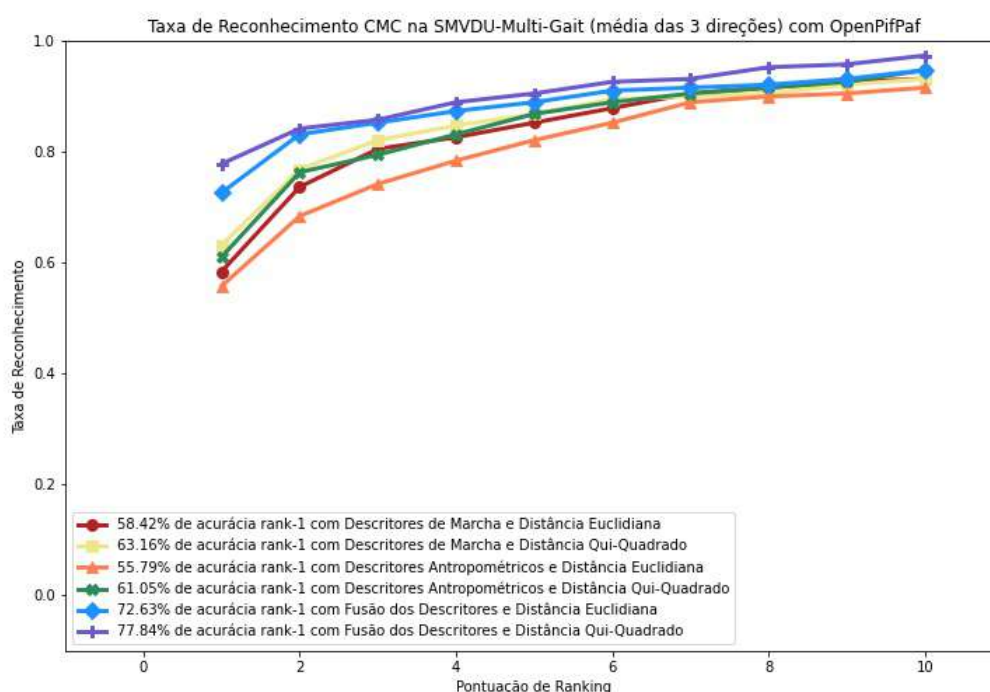
Outra vez, pode-se notar que a fusão dos descritores com a função de distância Qui-Quadrado lograram performances mais eficientes quando comparadas aos descritores de marcha ou descritores antropométricos apenas, para as direções lateral e oblíqua. Sem embargo aos resultados anteriores, para direções frontais de marcha, a configuração com o melhor valor inclui a fusão dos descritores com a função de distância Euclidiana, atingindo 89.59% de acurácia. Este resultado é um indicativo de que os histogramas gerados para direções de marcha frontais pelo método OpenPose possuem grandes discrepâncias, mesmo para um mesmo indivíduo,

acentuadas quando se aplica a função de distância Qui-Quadrado, com a função de distância Euclidiana se sobressaindo nesses casos.

#### 4.2.2 OpenPifPaf

A segunda etapa desse experimento aplica o método OpenPifPaf como estimador de poses na composição do fluxo da aplicação do método proposto. A Figura 27 mostra os resultados desse experimento. Em harmonia com o primeiro experimento realizado, o destaque em desempenho é atribuído a fusão de descritores com a função de distância Qui-Quadrado, permanecendo, para todos os 10 ranks considerados, unanimemente melhor.

Figura 27 – Curvas CMC obtidas pelas distâncias Euclidiana e Qui-Quadrado para os diferentes descritores na SMVDU-Multi-Gait com estimador de poses OpenPifPaf.



Fonte: Elaborada pelo autor.

A Tabela 3 mostra os valores para acurácia rank-1 obtidos pelo método proposto para cada direção de marcha na SMVDU-Multi-Gait, considerando o estimador de poses OpenPifPaf.

Tabela 3 – Acurácia Rank-1 - SMVDU-Multi-Gait com estimador de poses OpenPifPaf.

<b>Método</b>	<b>Lateral</b>	<b>Oblíquo</b>	<b>Frontal</b>
Descritores de Marcha - Euclidiana	51.39%	54.49%	64.58%
Descritores de Marcha - Qui-Quadrado	48.61%	65.71%	72.92%
Descritores Antropométricos - Euclidiana	52.78%	62.86%	60.42%
Descritores Antropométricos - Qui-Quadrado	63.80%	61.43%	64.58%
Fusão dos Descritores - Euclidiana	68.06%	<b>75.71%</b>	75.00%
Fusão dos Descritores - Qui-Quadrado	<b>75.00%</b>	<b>75.71%</b>	<b>85.42%</b>

Fonte: Elaborada pelo autor.

Uma vez mais ficou evidente a prevalência da fusão dos descritores biométricos com a função de distância Qui-Quadrado perante as demais configurações do método, contemplando as melhores acurácias calculadas para todas as três direções de marcha consideradas.

A partir dos experimentos conduzidos, constatou-se que o método apresentou bons resultados na tarefa de reconhecimento de pessoas em multimarcha. O método apresentou maior eficácia utilizando a fusão dos descritores com a função de distância Qui-Quadrado em todos os testes realizados, tanto para o estimador de pose OpenPose, com exceção da direção de marcha frontal na base SMVDU-Multi-Gait, cujos resultados mais expressivos foram indicados pela distância Euclidiana, quanto para o OpenPifPaf.

Considerando os valores de acurácia rank-1 de ambos os estimadores de pose, foi possível concluir que a configuração que atingiu os resultados mais robustos na tarefa de identificação de múltiplas pessoas compreende como método estimador de poses o OpenPifPaf, o algoritmo DeepSORT para o rastreamento e reidentificação de pessoas, a fusão dos descritores de marcha e antropométricos como traços biométricos e a função de distância Qui-Quadrado para o cálculo da pontuação de similaridade. Apesar de apresentar resultados levemente inferiores que o método OpenPose para direção de marcha frontal, o método OpenPifPaf compensa os 4% a menos de precisão observados nesta situação com um aumento de 25% na acurácia em comparação ao OpenPose para direção lateral e 3% para a direção oblíqua.

## 5 Sistema Biométrico

Os capítulos anteriores abordaram os conceitos que fundamentam o método biométrico de reconhecimento de marcha de múltiplas pessoas baseado em métodos de estimação de pose e rastreamento de objetos, o *pipeline* de tal método, com detalhamento de cada etapa sugerida, e os resultados obtidos na aplicação deste método em situações de marcha simples e multimarcha. Embasado nos que foi exposto e observado nestes capítulos, a etapa de implementação do sistema biométrico é apresentada.

Este capítulo descreve a construção do sistema biométrico, seguindo a esquemática da construção básica de um sistema genérico, apresentada na seção 2.2. Partindo de sua estrutura física, especificam-se os sensores utilizados para aquisição dos dados biométricos utilizados com a interface do usuário. Logo após, são apresentadas as componentes lógicas do sistema, como a estimação de poses, o rastreamento e reidentificação, a associação e a extração de características, que ocorrem como serviço por meio de um servidor HTTP, a composição da base de dados, os mecanismos de classificação de amostras e os resultados obtidos em sua construção.

Haja vista o exposto na seção 2.8 quanto a escolha das métricas para o reconhecimento com base em marcha e medidas antropométricas, o sistema tem como funcionalidade a identificação dos indivíduos presentes nos vídeos capturados.

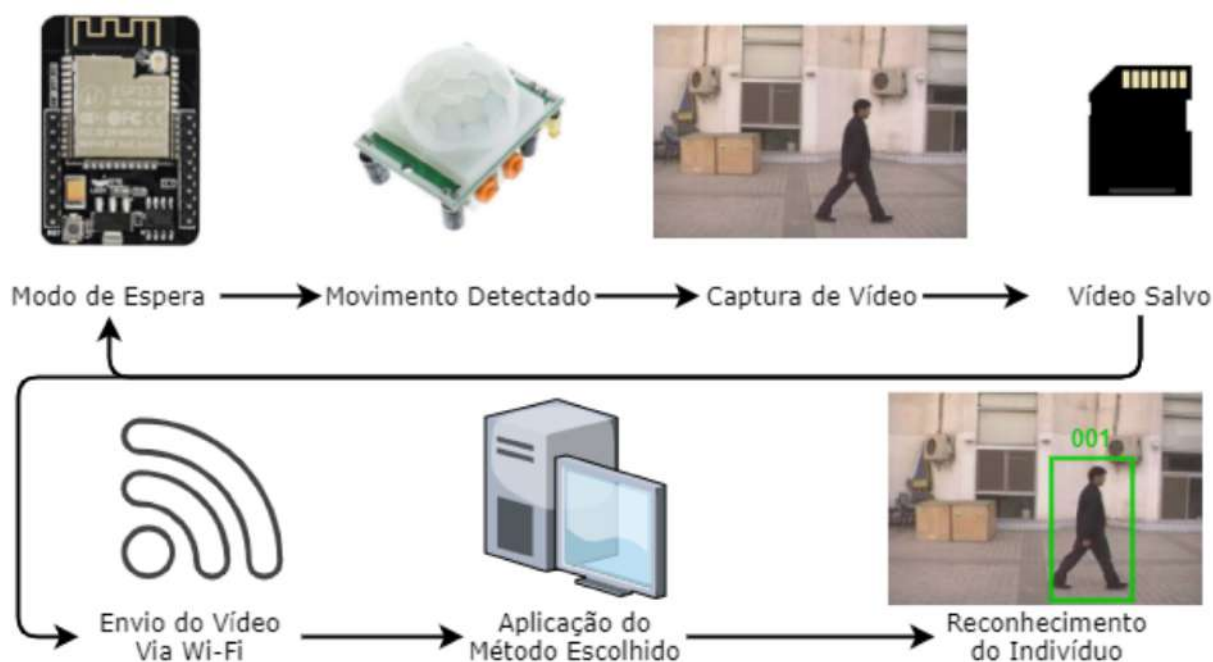
O sistema biométrico descrito neste capítulo tomou como base o sistema implementado por Jangua e Marana (2022)<sup>1</sup> como trabalho de conclusão de curso. Dentre as adaptações que permitiram a funcionalidade de reconhecimento de múltiplas pessoas estão: (i) substituição do estimador de pose OpenPose e integração do OpenPifPaf ao sistema, visando aumento na acurácia do sistema para cenários multimarcha, como visto no Capítulo 4; (ii) adição do módulo rastreador incorporando o método DeepSORT, responsável por detectar pessoas, definir sua trajetória e reidentificá-las perante a atribuição de uma identidade única; (iii) alterações nos módulos extrator de características e classificador, seguindo o exposto na seção 3.4 e na seção 3.6, respectivamente; e (iv) expansão das funcionalidades de inferência, capacitando-a na realização da classificação para dois indivíduos ou mais.

Outrossim, modificações menores na interface de usuário foram realizadas a fim de atribuir uma nova identidade visual para o aplicativo, apelidado de Biométrica. A Figura 28 exhibe o *pipeline* e a estrutura do sistema biométrico.

---

<sup>1</sup> Disponível em: <<https://github.com/Daniel-Jangua/gait-recognition>>. Acesso em 14 de novembro de 2022.

Figura 28 – Resumo da arquitetura e do *pipeline* do sistema biométrico.



Fonte: Jangua e Marana (2022).

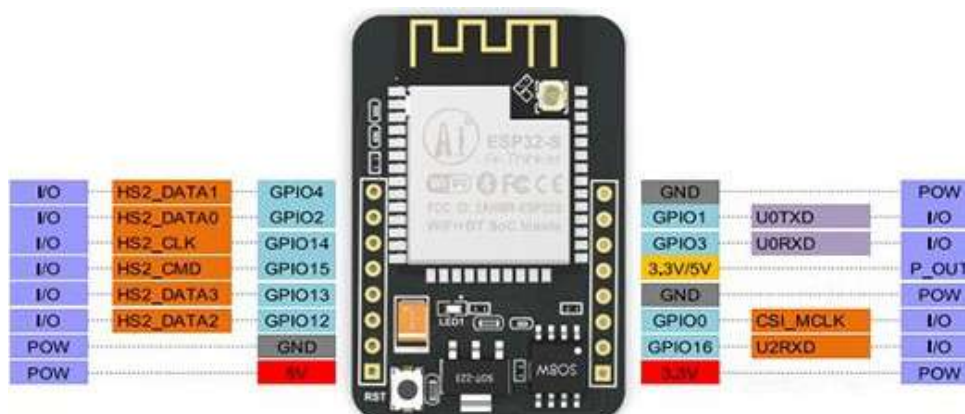
## 5.1 Sensores

O módulo dos sensores do sistema biométrico é composto por um hardware simples e barato se comparado com os sensores comumente utilizados para a coleta biométrica, que dispõem de câmeras com alta resolução. A captura dos vídeos é realizada pelo microcontrolador ESP32-CAM, ativada por um sensor de presença infravermelho HC-SR501, que dispara um sinal elétrico ao captar movimentos, iniciando a gravação.

O ESP32-CAM é um microcontrolador de baixo custo e de tamanho reduzido. Este módulo é geralmente utilizado em projetos e soluções de IoT, *Internet of Things*, do inglês, Internet das Coisas. Ele possui integração com Bluetooth e WiFi, além de possuir duas CPUs, *Central Processing Unit*, do inglês, Unidade Central de Processamento, de 32-bits com alta performance. AI-THINKER (2022) detalha, ainda, o suporte para câmera OV2640 e para cartões de memória microSD, além de uma memória interna de 520KB SRAM. A pinagem da placa de desenvolvimento ESP32-CAM é exibida na Figura 29.



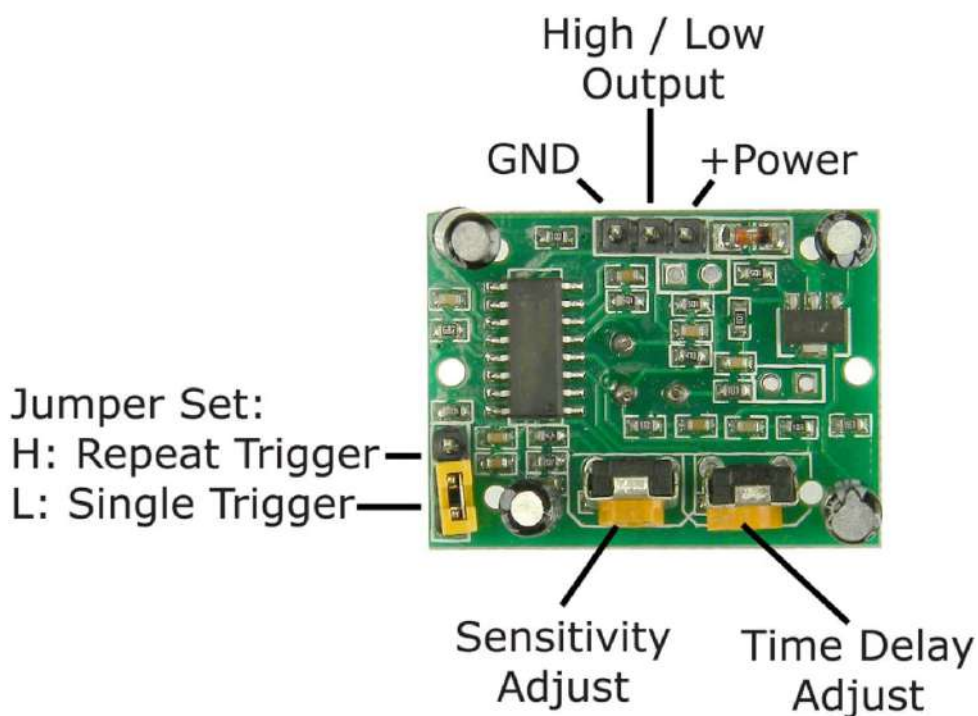
Figura 29 – Pinagem do módulo de gravação ESP32-CAM.



Fonte: AI-THINKER (2022).

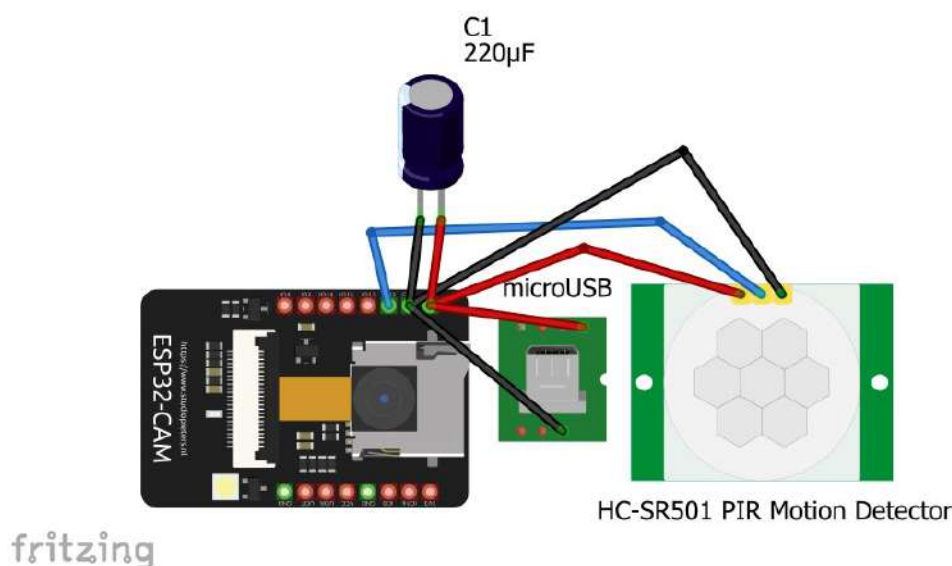
O HC-SR501 é um módulo automático de alta sensibilidade e baixa voltagem que desfruta de tecnologia infravermelha, capaz de detectar movimento em um raio de até 7 metros e em um ângulo de até 120 graus em relação ao sensor. Assim como o ESP32-CAM, é amplamente utilizado em soluções de IoT, permite o ajuste da sensibilidade da detecção, variando entre raios de 3 a 7 metros, bem como do tempo de ativação do sensor, demorando entre 3 e 300 segundos (ETC, 2022). A pinagem do sensor HC-SR501 é exibida na Figura 30, enquanto a Figura 31 explicita a montagem do circuito.

Figura 30 – Pinagem do módulo sensor HC-SR501.



Fonte: ETC (2022).

Figura 31 – Circuito do módulo sensor do sistema biométrico.



Fonte: Zah (2022).

Uma entrada USB, *Universal Serial Bus*, do inglês, Porta Serial Universal, é responsável pela alimentação do circuito, ligado em paralelo com um capacitor de 220  $\mu\text{F}$ , evitando possíveis ruídos e oscilações. A saída de sinal do sensor de presença é ligada ao pino 12 de entrada e saída do microcontrolador, o qual foi programado para iniciar uma captura de 10 a 15 segundos com a mudança de estado baixo para alto no pino. A programação que permite o funcionamento do microcontrolador e sua integração com o sensor, em conjunto com a implementação de um servidor FTP, *File Transfer Protocol*, do inglês, Protocolo de Transferência de Arquivos, utilizado para obtenção remota dos vídeos capturados, encontra-se no repositório do GitHub chamado ESP32-CAM-Video-Recorder<sup>2</sup>. Alguns ajustes tiveram que ser efetuados, dentre eles a taxa de quadros do vídeo, a qualidade da imagem e o tempo de gravação.

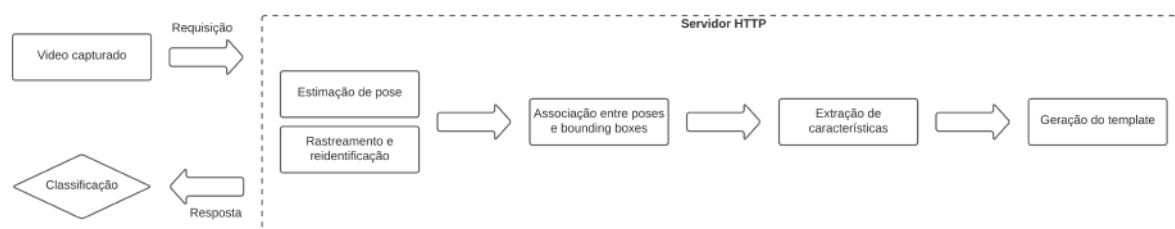
## 5.2 Servidor HTTP

Os módulos de pré-processamento, associação de poses e *bounding boxes*, extração de características e geração de *templates*. O pré-processamento compreende as etapas de estimação de poses e rastreamento e reidentificação de objetos. Os métodos empregados na fase de pré-processamento demandam um maior poder de processamento, havendo necessidade de utilização de uma GPU para o processamento em tempo real, como havia sido mencionado no Capítulo 4. Portanto, na ausência de tais ferramentas, optou-se por realizar o processamento em nuvem, fazendo uso de uma instância de notebook do Colab. Assim, o modelo e todas as etapas mencionadas anteriormente são implementados como um serviço por meio de um servidor HTTP, *Hypertext Transfer Protocol*, do inglês, Protocolo de Transferência Hipertexto.

<sup>2</sup> Disponível em: <<https://github.com/jameszah/ESP32-CAM-Video-Recorder>>. Acesso em 14 de novembro de 2022.

A aplicação faz uma requisição para o servidor, enviando as sequências de imagens que compõem um vídeo de detecção. O servidor aceita a requisição da aplicação, realiza as etapas listadas e retorna os *templates*, que são os descritores de marcha e antropométricos, como resposta. A Figura 32 ilustra o funcionamento do modelo como um serviço através do servido HTTP<sup>3</sup>.

Figura 32 – Esquema de funcionamento do servidor HTTP para aplicação do método.



Fonte: Elaborada pelo autor.

## 5.3 Bases de Dados

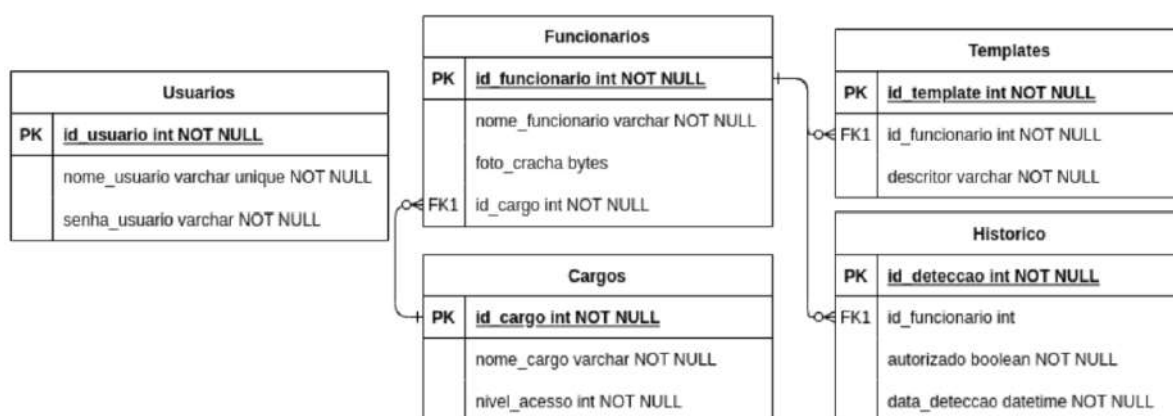
O módulo da base de dados do sistema segue a modelagem feita por Jangua e Marana (2022), se resumindo em um banco de dados estrutural baseado em linguagem SQL, Structured Query Language, do inglês, Linguagem de Consulta Estruturada, usando o DBMS, Database Management System, do inglês, Sistema de Gerenciamento de Banco de Dados, SQLite. O SQLite é um DBMS SQL de código aberto, que, diferentemente da maioria dos gerenciadores de banco de dados, não precisa de um processo de servidor separado rodando em paralelo para leitura e escrita de documentos, armazenando e acessando arquivos no disco local do computador (BHOSALE; PATIL; PATIL, 2015). Nas palavras dos autores, o formato do arquivo de banco de dados é multiplataforma, possuindo aplicações em diversas linguagens de programação, tais quais Python, C, C++ e Java, abarcando uma escolha popular para soluções simples de bases de dados, com a segurança e a confiabilidade fornecida pelos DBMSs mais conhecidos.

O banco de dados simula estruturalmente um minimundo de uma empresa em busca de realizar o controle de acesso a áreas restritas, monitoradas por uma câmera de baixa resolução. O cenário fictício é composto por funcionários, cada qual exercendo um cargo na empresa. Os cargos possuem níveis de acesso distintos associados a eles, utilizados para averiguar o acesso a determinada área. Cada área tem um nível de acesso mínimo para que se possa acessá-la. Assim, no caso de um funcionário com nível de acesso inferior ao permitido para determinado ambiente ser identificado andando por aquele espaço, o aplicativo dispara um alerta de segurança para uma central de monitoramento, sendo dever do administrador tomar

<sup>3</sup> Disponível em: <<https://colab.research.google.com/drive/198-USYmTpdAXCbcoD4PCf20TCUynosvk?usp=sharing>>. Acesso em 9 de janeiro de 2023

as providências cabíveis para esta situação. O software deve ser capaz, ainda, de armazenar logs das detecções com as seguintes informações: (i) nome do funcionário, (ii) data e hora da detecção, (iii) veredito do sistema, isto é, se o acesso foi autorizado ou negado. Para mais, considerando a etapa de autenticação de um sistema biométrico, os *templates* referentes a todo funcionário da empresa devem ser salvos no banco de dados para efetuar a classificação. A Figura 33 expõe a modelagem do banco utilizado pelo sistema.

Figura 33 – Diagrama do banco de dados utilizado pelo sistema biométrico.



Fonte: Jangua e Marana (2022).

O banco de dados do sistema foi povoado com sequências de imagens de pessoas caminhando, reunidas em um pequeno conjunto de dados com os mesmos padrões de gravação da base de dados SMVDU-Multi-Gait (SINGH et al., 2019). Foram capturados 42 vídeos de 3 pessoas diferentes, dentre eles 4 vídeos em situação de marcha simples para cada indivíduo, com ângulos de visão lateral e frontal, com 12 vídeos ao todo cadastrados e armazenados como *gallery*, e 10 vídeos em situação de multimarcha para cada dupla de indivíduos, com ângulos de visão lateral e frontal, caminhando para a mesma direção e em direções opostas, totalizando 30 vídeos, compondo a *probe*, utilizados para validação do sistema, comparados a *gallery* na classificação. As gravações foram feitas utilizando a câmera integrada ao módulo ESP32-CAM. A Figura 34 exhibe exemplos de imagens dos vídeos em que os participantes caminham sozinhos, enquanto a Figura 35 demonstra exemplos em que os participantes caminham em duplas.

Figura 34 – Exemplos de imagens dos vídeos capturados para a base de dados do sistema em situação de marcha simples.



(a) Visão frontal.



(b) Visão traseira.



(c) Visão lateral (direita para a esquerda).



(d) Visão lateral (esquerda para a direita).

Fonte: Elaborada pelo autor.

Figura 35 – Exemplos de imagens dos vídeos capturados para a base de dados do sistema em situação multimarcha.



(a) Visão frontal e traseira.



(b) Visão frontal.



(c) Visão lateral (ambos os sentidos).



(d) Visão lateral (da direita para a esquerda).

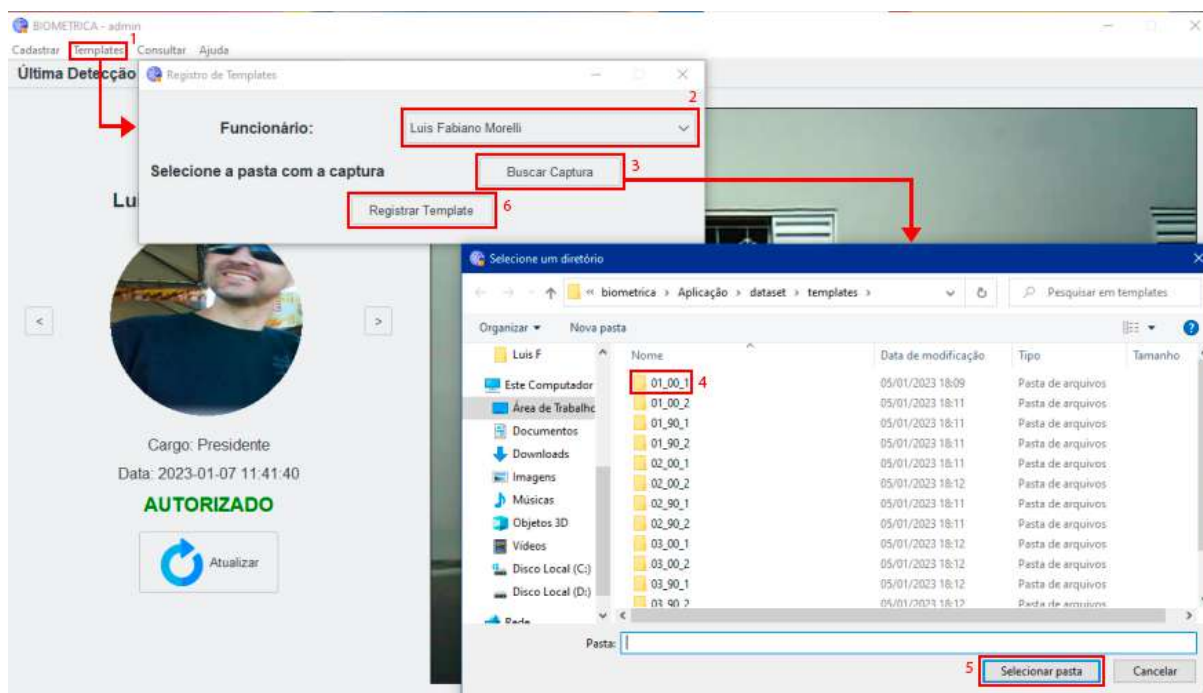
Fonte: Elaborada pelo autor.

## 5.4 Registro e Autenticação

A seção 2.2 explica o fluxo de execução de um sistema biométrico, dividido fundamentalmente em dois momentos: (i) o registro e (ii) a autenticação. Ambos os estágios, bem como as demais funcionalidades implementadas pelo sistema, interagem com o usuário por meio de uma interface gráfica. A interface foi criada em linguagem Python, com o auxílio da biblioteca Tkinter. O Tkinter é um *framework* leve integrado a biblioteca padrão do Python, reunindo componentes de interface gráfica para as aplicações GUI, *Graphical User Interface*, do inglês, Interface Gráfica do Usuário (PYTHON, 2022). O autor esclarece que a biblioteca é multiplataforma, portanto o mesmo código funciona no Windows, MacOS e Linux. Seus elementos visuais são renderizados usando elementos nativos de cada sistema operacional, portanto, aplicativos programados com Tkinter parecem pertencer à plataforma em que são executados.

A Figura 36 exemplifica o fluxo de uso da aplicação para realizar o armazenamento de um *template* para um funcionário da empresa fictícia, constituindo a etapa de registro. Partindo da tela inicial, deve-se clicar na opção "Templates", indicado pelo número 1 na figura, localizada na barra de menus do *software*. Uma nova janela será aberta, apresentando ao usuário os campos que necessitam ser preenchidos para continuar com o registro do *template*, sendo: (i) funcionário associado ao *template*, escolhido através da listagem da caixa de combinação do aplicativo, indicada pelo número 2 na figura, e (ii) caminho para o diretório que contém a captura do indivíduo caminhando sozinho, indicado pelo número 3 na figura. Para efetuar a seleção do diretório, uma nova tela será aberta, permitindo a escolha do caminho para a pasta com a captura, indicado na figura pelo número 4. Após selecionada, clicando no botão com o número 5 da figura, deve-se clicar no botão "Registrar Template". O vídeo será enviado para o servidor HTTP, processado, e o servidor retornará os descritores biométricos como resposta. Por fim, o *template* é armazenado na tabela de *templates* do banco de dados.

Figura 36 – Fluxo da fase de registro.



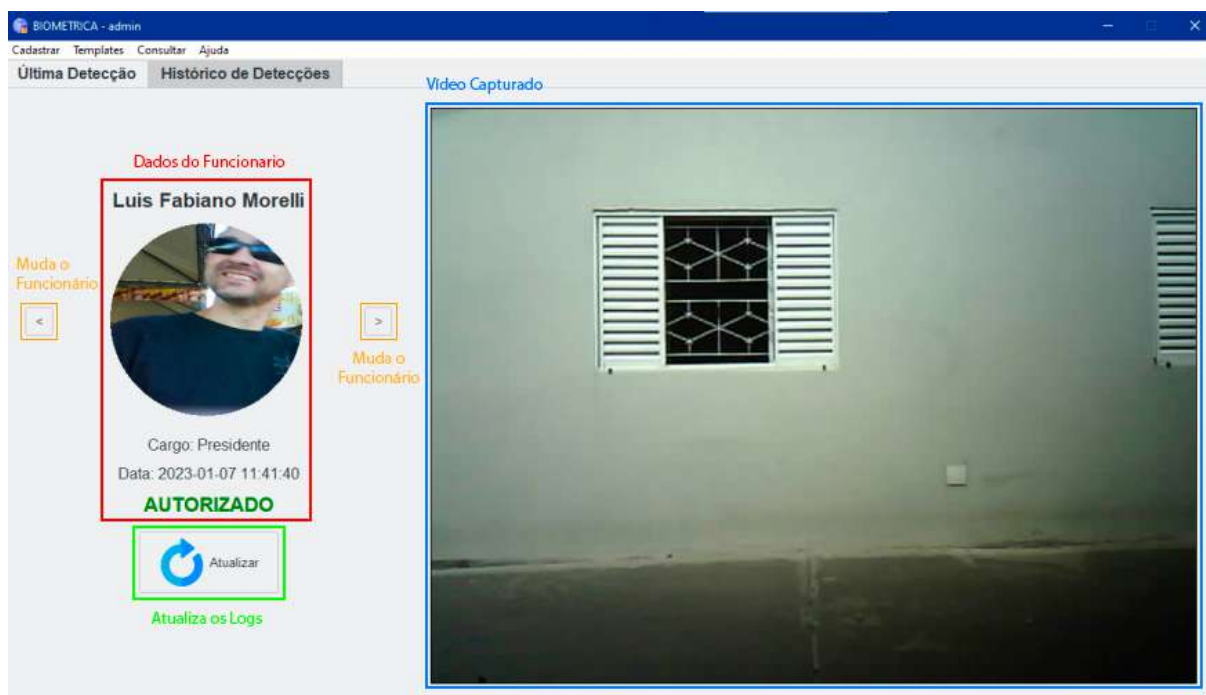
Fonte: Elaborada pelo autor.

A etapa de autenticação é um processo executado em uma *thread*, do inglês, linha de execução concorrente, ao programa principal, em tempo real, continuamente. A *thread* é responsável por listar todos os arquivos de vídeo capturados pelo módulo ESP32-CAM armazenado em seu servidor FTP, faz o download de cada arquivo guardado nele e os envia ao servidor HTTP através de uma requisição, retornando os descritores como resposta do servidor. O descritor recebido do servidor é comparado a todos os descritores da marcha simples de cada indivíduo, aplicando um classificador 1NN para determinação da identidade dos indivíduos captados em vídeo.

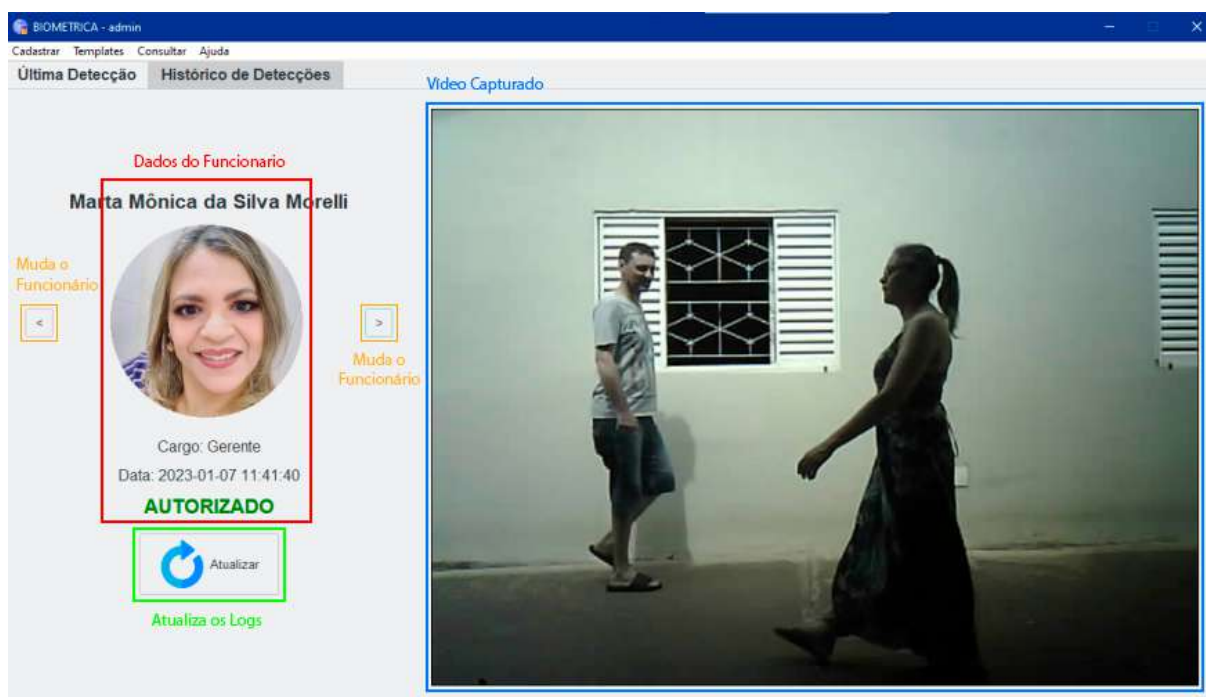
Ao final da etapa de identificação, após a identidade ser atribuída ao indivíduo, faz-se uma verificação do nível de acesso de seu cargo com o permitido pela área que acessou, definindo se o acesso é concedido ou negado. Um alerta de segurança é enviado ao sistema principal para os casos em que há incongruência com os níveis de acesso e este é negado. Independentemente do resultado da verificação, a última tarefa da autenticação é a de atualizar os registros de detecção do sistema, sendo exibido na tela principal da aplicação todos os dados referentes aos indivíduos detectados, tais quais o nome completo do funcionário, cargo que exerce na empresa, foto do crachá, data e hora da detecção e a captura disparada por seu acesso ao ambiente em questão, assim como o estado deste acesso. É possível visualizar as informações dos demais funcionários detectados clicando em setas ao lado da foto de crachá. Essas informações também podem ser visualizadas na aba "Histórico de Detecções", listando todas as identificações feitas pelo sistema até então. A Figura 37 demonstra o painel de monitoramento do sistema para o estágio de autenticação.



Figura 37 – Painel de monitoramento em tempo real da fase de autenticação.



(a) Dados do funcionário Luis Fabiano Morelli.



(b) Dados da funcionária Marta Morelli após o botão "muda o funcionário" ser pressionado.

Fonte: Elaborada pelo autor.

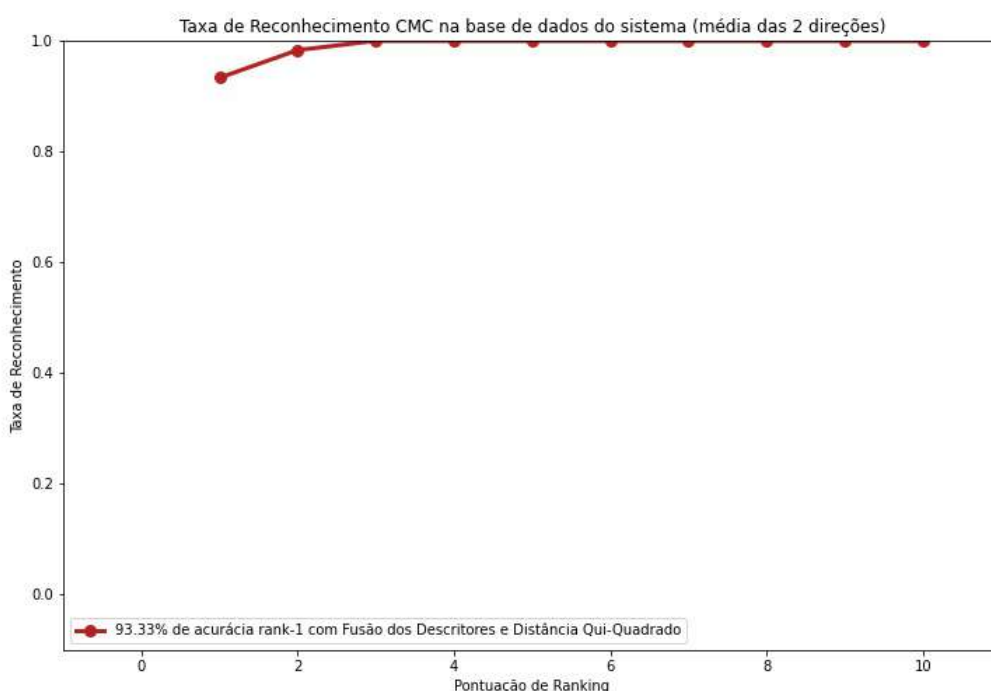
Ademais, complementando as funcionalidades biométricas citadas, o sistema é capaz de cadastrar novos funcionários, cargos e usuários, além de permitir consultar os dados dos cargos e dos funcionários cadastrados no banco de dados.



## 5.5 Resultados e Discussão

Para validar o sistema biométrico, um experimento foi conduzido, direcionado aos 30 vídeos da multimarcha em duplas dos indivíduos que compõem a base de dados do sistema. A Figura 38 mostra os resultados de tal experimento, respeitando configurações parecidas aquelas citadas na seção 4.1 para obtenção da média das acurácias rank-1 obtidas para os ângulos de visão lateral e frontal. Pode-se observar que, no geral, o método apresentou bons resultados, com acurácia média de 93.33%, reconhecendo corretamente os funcionários cadastrados na maioria dos vídeos capturados.

Figura 38 – Curvas CMC obtidas na base de dados coletada para o sistema.



Fonte: Elaborada pelo autor.

A Tabela 4 mostra os valores para acurácia rank-1 obtidos pelo método para as direções de marcha lateral e frontal na base de dados criada para o minimundo descrito.

Tabela 4 – Acurácia Rank-1 - Base de Dados do Sistema.

Visão	Acurácia
Lateral	97.22%
Frontal	87.50%
Geral	93.33%

Fonte: Elaborada pelo autor.

Observou-se que a análise de cada ângulo de visão individualmente resultou em bons resultados de acurácia para o método, indicando que ele foi acurado no reconhecimento dos

funcionários nas diferentes situações em que pode ser detectado. Outrossim, a direção lateral obteve os melhores resultados. Isso deve-se ao fato de apresentar dois vídeos a mais para classificação. Ainda, em consonância com Kreiss, Bertoni e Alahi (2021), demonstra que o OpenPifPaf desempenha bem em baixas resoluções e em cenas desordenadas e oclusas.

Os resultados apontam que o método proposto, utilizando o estimador de pose OpenPifPaf, o rastreador de objetos DeepSORT, como descritor a fusão de características de marcha e medidas antropométricas e função de distância Qui-Quadrado apresenta resultados robustos para operar em ambientes menos controlados. Contudo, os vídeos capturados não simulam uma situação real fielmente, haja vista que segue alguns protocolos, tais quais ângulos fixos para a caminhada e restrição de não existirem mais que duas pessoas caminhando simultaneamente nas cenas.

Para mais, o baixo custo de implementação, sendo necessário apenas uma câmera de baixa resolução, que grava vídeos com baixa taxa de quadros por segundo, e a alta capacidade de escalabilidade do sistema configuram as principais vantagens observadas. Os códigos da interface gráfica, da programação do módulo ESP32-CAM, da estrutura modelada para o banco de dados em SQLite, o Colab *Notebook* para o servidor HTTP e a base de dados criada para o sistema biométrico estão armazenados em um repositório público do GitHub<sup>4</sup>.

---

<sup>4</sup> Disponível em: <<https://github.com/Xilsu/biometrica>>. Acesso em 9 de janeiro de 2023.

## 6 Conclusão

Em suma, este trabalho cumpriu com os objetivos proposto para seu desenvolvimento. O Capítulo 2 demonstra os conhecimentos obtidos nos estudos propostos, compreendendo temas como biometria, com enfoque em sistemas biométricos e identificação de pessoas através de métodos para reconhecimento de marcha e biometrias suaves, como as medidas antropométricas. Ademais, o estudo de técnicas e vantagens da implementação de multibiometria em sistemas biométricos contribuiu para os resultados experimentais, compreendendo o fator mais influente na decisão da configuração do método proposto. Ainda, o estudo aprofundado dos métodos de Jangua e Marana (2020) e Tavares, Marana e Neto (2021) se mostrou essencial para entendimento dos conceitos práticos envolvidos na implementação de sistemas e obtenção de descritores biométricos, possibilitando as alterações que tornaram a identificação multimarcha possível.

Os resultados dos experimentos de validação do método proposto por este trabalho mostraram que ele é aplicável tanto para situações de marcha simples quanto multimarcha, alcançando bons resultados na tarefa de identificação de indivíduos em ambos os casos. O uso da fusão dos descritores biométricos de marcha e antropométricos sobressaiu-se quando comparada a ambos os descritores individualmente neste contexto, solidificando o exposto de que a união de características biométricas de um indivíduo aumenta a unicidade dos traços biométricos, com suas vantagens individuais se sobrepondo às suas deficiências. Para tais experimentos, também se notou que o estimador de pose OpenPifPaf performou a tarefa de estimação dos pontos-chave do esqueleto humano melhor que o OpenPose para a base de dados SMVDU-Multi-Gait, constituindo a escolha de estimador de pose para o sistema biométrico implementado.

Por fim, o último estágio de desenvolvimento compreendeu a construção de um sistema biométrico completo, desde a criação do circuito físico dos sensores para captura dos dados biométricos, até a programação da lógica do sistema, responsável pela extração de características biométricas, classificação e identificação dos indivíduos captados. O método de reconhecimento biométrico escolhido para compor o sistema emprega o estimador de pose OpenPifPaf, o rastreador de objetos DeepSORT, a fusão dos descritores de marcha e antropométricos como traço biométrico e a função de distância Qui-Quadrado para o cálculo da pontuação de similaridade, e obteve resultados robusto para corroborar com sua implementação para soluções de segurança em empresas com baixo custo envolvido.

Trabalhos futuros poderão focar em testar o sistema biométrico em ambientes mais ambiciosos, expandindo os protocolos adotados para utilização de maior quantidade de ângulos variados, considerando direções oblíquas de marcha, tais quais  $30^\circ$  e  $40^\circ$ , bem como áreas em que pessoas estão aglomeradas ou, ao menos, existam mais de duas pessoas caminhando juntas.

## Referências

- AI-THINKER. *ESP32-CAM camera development board*. 2022. Disponível em: <https://docs.ai-thinker.com/en/esp32-cam>. Acesso em: 14 de novembro de 2022.
- ARANTES, M.; GONZAGA, A. Human gait recognition using extraction and fusion of global motion features. *Multimedia Tools and Applications*, v. 55, p. 655–675, Dezembro 2011. ISSN 13807501.
- BEWLEY, A.; GE, Z.; OTT, L.; RAMOS, F.; UPCROFT, B. Simple online and realtime tracking. *Proceedings - International Conference on Image Processing, ICIP*, IEEE Computer Society, v. 2016-August, p. 3464–3468, Fevereiro 2016. Disponível em: <http://arxiv.org/abs/1602.00763>. Acesso em: 20 de janeiro de 2023.
- BHOSALE, S.; PATIL, M.; PATIL, P. International journal of computer science and mobile computing sqlite: Light database system. *International Journal of Computer Science and Mobile Computing*, v. 44, p. 882–885, Abril 2015.
- CAO, Z.; HIDALGO, G.; SIMON, T.; WEI, S. E.; SHEIKH, Y. Openpose: Realtime multi-person 2d pose estimation using part affinity fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE Computer Society, v. 43, p. 172–186, Dezembro 2018. ISSN 19393539. Disponível em: <https://arxiv.org/abs/1812.08008v2>. Acesso em: 20 de janeiro de 2023.
- CARNEIRO, T.; NÓBREGA, R. V. M. D.; NEPOMUCENO, T.; BIAN, G.-B.; ALBUQUERQUE, V. H. C. D.; FILHO, P. P. R. Performance analysis of google colab as a tool for accelerating deep learning applications. *IEEE Access*, v. 6, p. 61677–61685, 2018.
- CHEN, X.; WENG, J.; LU, W.; XU, J. Multi-gait recognition based on attribute discovery. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 40, n. 7, p. 1697–1710, 2018.
- CHOUDHURY, B.; THEN, P.; ISSAC, B.; RAMAN, V.; HALDAR, M. A survey on biometrics and cancelable biometrics systems. *International Journal of Image and Graphics*, v. 18, p. 1850006, Janeiro 2018.
- COVER, T. M.; HART, P. E. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, v. 13, p. 21–27, 1967. ISSN 15579654.
- CUNADO, D.; NIXON, M. S.; CARTER, J. N. Using gait as a biometric, via phase-weighted magnitude spectra. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Springer Verlag, v. 1206, p. 93–102, 1997. ISSN 16113349. Disponível em: <https://link.springer.com/chapter/10.1007/BFb0015984>. Acesso em: 20 de janeiro de 2023.
- DANTCHEVA, A.; VELARDO, C.; D'ANGELO, A.; DUGELAY, J. L. Bag of soft biometrics for person identification. *Multimedia Tools and Applications*, v. 51, p. 739–777, Janeiro 2010. ISSN 13807501.

DELGADO-ESCANO, R.; CASTRO, F. M.; CÓZAR, J. R.; MARÍN-JIMÉNEZ, M. J.; GUIL, N. Mupeg—the multiple person gait framework. *Sensors*, v. 20, n. 5, 2020. ISSN 1424-8220. Disponível em: <<https://www.mdpi.com/1424-8220/20/5/1358>>. Acesso em: 20 de janeiro de 2023.

ETC. *HC-SR501 DETECTOR Datasheet pdf - MOTION DETECTOR. Equivalent, Catalog*. 2022. Disponível em: <<https://datasheetspdf.com/pdf/775434/ETC/HC-SR501/1>>. Acesso em: 17 de novembro de 2022.

GLEN, S. *ROC Curve Explained in One Picture - DataScienceCentral.com*. 2019. Disponível em: <<https://www.datasciencecentral.com/roc-curve-explained-in-one-picture/\%7D>>. Acesso em: 20 de janeiro de 2023.

HUNTER, J. D. Matplotlib: A 2d graphics environment. *Computing in Science Engineering*, v. 9, n. 3, p. 90–95, 2007.

JAIN, A. K.; ROSS, A. A.; NANDAKUMAR, K. *Introduction to Biometrics*. [S.l.]: Springer Publishing Company, Incorporated, 2011. ISBN 0387773258.

JANGUA, D. R. dos S.; MARANA, A. N. A new method for gait recognition using 2d poses. In: *XVI Workshop de Visão Computacional (WVC 2020)*. [S.l.: s.n.], 2020.

JANGUA, D. R. dos S.; MARANA, A. N. *Sistema de Reconhecimento Biométrico Baseado em Marcha Utilizando Poses 2D*. 2022. Disponível em: <<https://dco-unesp-bauru.github.io/tcc-bcc-2021-2/trabalho/2022/03/01/DanielRSJ.html>>. Acesso em: 20 de janeiro de 2023.

KREISS, S.; BERTONI, L.; ALAHI, A. *OpenPifPaf: Composite Fields for Semantic Keypoint Detection and Spatio-Temporal Association*. arXiv, 2021. Disponível em: <<https://arxiv.org/abs/2103.02440>>. Acesso em: 20 de janeiro de 2023.

LIMA, V.; MELO, V.; SCHWARTZ, W. Simple and efficient pose-based gait recognition method for challenging environments. *Pattern Analysis and Applications*, v. 24, p. 1–11, Maio 2021.

LIMA, V. C. D.; SCHWARTZ, R. Gait recognition using pose estimation and signal processing. *Iberoamerican on Pattern Recognition - CIARP*, 2019.

LIU, D.; YE, M.; LI, X.; ZHANG, F.; LIN, L. Memory-based gait recognition. In: . [S.l.: s.n.], 2016. p. 82.1–82.12.

MARCONDES, J. S. *Biometria, Sistema Biométrico: O que é, Como Funciona?* Gestão de Segurança Privada, 2020. Disponível em: <<https://gestaodesegurancaprivada.com.br/biometria-sistema-biometrico-o-que-e-como-funciona/#Sistema-biometria>>. Acesso em: 20 de janeiro de 2023.

MARQUES, C.; VAZ, D. *Marcha Normal – ciclos e fases da marcha*. 2018. Vídeo do Youtube. Disponível em: <[https://www.youtube.com/watch?v=Qjj84musSpk&t=234s&ab\\_channel=CaedUFMG](https://www.youtube.com/watch?v=Qjj84musSpk&t=234s&ab_channel=CaedUFMG)>. Acesso em: 23 de dezembro de 2022.

MILAN, A.; LEAL-TAIXE, L.; REID, I.; ROTH, S.; SCHINDLER, K. *MOT16: A Benchmark for Multi-Object Tracking*. arXiv, 2016. Disponível em: <<https://arxiv.org/abs/1603.00831>>. Acesso em: 20 de janeiro de 2023.

MITCHELL, T. *Machine Learning*. [S.l.]: McGraw-Hill Education, 1997. (McGraw-Hill international editions - computer science series). ISBN 9780070428072.

NIXON, M. S.; CARTER, J. N. Automatic recognition by gait. *Proceedings of the IEEE*, v. 94, n. 11, p. 2013–2024, 2006.

PEREIRA, R.; CARVALHO, G.; GARROTE, L.; NUNES, U. J. Sort and deep-sort based multi-object tracking for mobile robotics: Evaluation with new data association metrics. *Applied Sciences*, v. 12, n. 3, 2022. ISSN 2076-3417. Disponível em: <<https://www.mdpi.com/2076-3417/12/3/1319>>. Acesso em: 20 de janeiro de 2023.

PÉREZ, F.; GRANGER, B. E. IPython: a system for interactive scientific computing. *Computing in Science and Engineering*, IEEE Computer Society, v. 9, n. 3, p. 21–29, Maio 2007. ISSN 1521-9615. Disponível em: <<https://ipython.org>>. Acesso em: 20 de janeiro de 2023.

PERRY, J.; SCHONEBERGER, B.; ARAUJO, A. G. N.; FREITAS, C. D. de. *Análise de marcha*. Manole, 2005.

PREDA, G. *KNN model interpretation for k=1 using Voronoi polygons*. Kaggle, 2020. Disponível em: <<https://www.kaggle.com/questions-and-answers/127759>>. Acesso em: 20 de janeiro de 2023.

PYTHON. *tkinter — Python interface to Tcl/Tk*. 2022. Python Software Foundation. Disponível em: <<https://docs.python.org/3/library/tkinter.html>>. Acesso em: 8 de janeiro de 2023.

ROSS, A. A.; NANDAKUMAR, K.; JAIN, A. K. *Handbook of Multibiometrics*. 1st. ed. [S.l.]: Springer Publishing Company, Incorporated, 2011. ISBN 1441935479.

SINGH, J. P.; ARORA, S.; JAIN, S.; SINGH, U. P. Vision-based gait recognition: A survey. *IEEE Access*, v. 6, p. 70497–70527, 2018.

SINGH, J. P.; ARORA, S.; JAIN, S.; SOM, U. P. S. A multi-gait dataset for human recognition under occlusion scenario. In: *2019 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT)*. [S.l.: s.n.], 2019. v. 1, p. 1–6.

SU, C.; LI, J.; ZHANG, S.; XING, J.; GAO, W.; TIAN, Q. *Pose-driven Deep Convolutional Model for Person Re-identification*. arXiv, 2017. Disponível em: <<https://arxiv.org/abs/1709.08325>>. Acesso em: 20 de janeiro de 2023.

TAVARES, H. L.; MARANA, A. N.; NETO, J. B. C. *Identificação de pessoas baseada em características antropométricas e de marcha extraídas de poses 2D*. 2021. Disponível em: <<https://repositorio.unesp.br/handle/11449/214430>>. Acesso em: 20 de janeiro de 2023.

TOSHEV, A.; SZEGEDY, C. Deeppose: Human pose estimation via deep neural networks. In: *2014 IEEE Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2014. p. 1653–1660.

VARGES DA SILVA, M.; MARANA, A. N. Human action recognition in videos based on spatiotemporal features and bag-of-poses. *Applied Soft Computing*, v. 95, p. 106513, 2020. ISSN 1568-4946. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S156849462030452X>>. Acesso em: 20 de janeiro de 2023.

- WAN, C.; LI, W.; PHOHA, V. V. A survey on gait recognition. *ACM Computing Surveys*, v. 51, n. 89, 2018.
- WANG, L.; TAN, T.; NING, H.; HU, W. Silhouette analysis-based gait recognition for human identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE Computer Society, Los Alamitos, CA, USA, v. 25, n. 12, p. 1505–1518, Dezembro 2003. ISSN 1939-3539.
- WANG, Z.; ZHENG, L.; LIU, Y.; LI, Y.; WANG, S. *Towards Real-Time Multi-Object Tracking*. arXiv, 2019. Disponível em: <<https://arxiv.org/abs/1909.12605>>. Acesso em: 20 de janeiro de 2023.
- WETTSCHERECK, D.; AHA, D. W.; MOHRI, T. A review and empirical evaluation of feature weighting methods for a class of lazy learning algorithms. *Artificial Intelligence Review*, Kluwer Academic Publishers, v. 11, p. 273–314, 1997.
- WOJKE, N.; BEWLEY, A.; PAULUS, D. Simple online and realtime tracking with a deep association metric. *Proceedings - International Conference on Image Processing, ICIP*, IEEE Computer Society, v. 2017-September, p. 3645–3649, Março 2017. ISSN 15224880. Disponível em: <<https://arxiv.org/abs/1703.07402v1>>. Acesso em: 20 de janeiro de 2023.
- YANG, W.; XU, L.; CHEN, X.; ZHENG, F.; LIU, Y. Chi-squared distance metric learning for histogram data. *Mathematical Problems in Engineering*, Hindawi Publishing Corporation, v. 2015, p. 1–12, Abril 2015. ISSN 15635147. Disponível em: <[https://www.researchgate.net/publication/276127503\\_Chi-Squared\\_Distance\\_Metric\\_Learning\\_for\\_Histogram\\_Data](https://www.researchgate.net/publication/276127503_Chi-Squared_Distance_Metric_Learning_for_Histogram_Data)>. Acesso em: 20 de janeiro de 2023.
- ZAH, J. *ESP32-CAM-Video-Recorder*. 2022. Repositório do Github. Disponível em: <<https://github.com/jameszah/ESP32-CAM-Video-Recorder>>. Acesso em: 14 de novembro de 2022.
- ZHANG, Y.; WANG, C.; WANG, X.; ZENG, W.; LIU, W. FairMOT: On the fairness of detection and re-identification in multiple object tracking. *International Journal of Computer Vision*, Springer Science and Business Media LLC, v. 129, n. 11, p. 3069–3087, Setembro 2021. Disponível em: <<https://doi.org/10.1007/s11263-021-01513-4>>. Acesso em: 20 de janeiro de 2023.