

UNIVERSIDADE ESTADUAL PAULISTA "JÚLIO DE MESQUITA FILHO"
FACULDADE DE CIÊNCIAS - CAMPUS BAURU
DEPARTAMENTO DE COMPUTAÇÃO
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

DANIEL RICARDO DOS SANTOS JANGUA

**SISTEMA DE RECONHECIMENTO BIOMÉTRICO BASEADO EM
MARCHA UTILIZANDO POSES 2D**

BAURU - SP
Fevereiro/2022

DANIEL RICARDO DOS SANTOS JANGUA

**SISTEMA DE RECONHECIMENTO BIOMÉTRICO BASEADO EM
MARCHA UTILIZANDO POSES 2D**

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 - SP
Fevereiro/2022

Santos Jangua, Daniel Ricardo dos.

Sistema de Reconhecimento Biométrico Baseado em Marcha Utilizando Poses 2D/ Daniel Ricardo dos Santos Jangua, 2022

62 f. : il.

Orientador: Prof. Dr. Aparecido Nilceu Marana

Monografia (Graduação)—Universidade Estadual Paulista. Faculdade de Ciências, Bauru, 2022

1. Biometria. 2. Reconhecimento de Marcha. 3. Estimação de Pose.
- I. Universidade Estadual Paulista. Faculdade de Ciências. II. Título.

Daniel Ricardo dos Santos Jangua

Sistema de Reconhecimento Biométrico Baseado em Marcha Utilizando Poses 2D

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 Ciência da Computação

Profa. Dra. Simone das Graças Domingues Prado

Universidade Estadual Paulista "Júlio de Mesquita Filho"
Faculdade de Ciências
Departamento de Ciência da Computação

Profa. Juliana da Costa Feitosa

Universidade Estadual Paulista "Júlio de Mesquita Filho"
Faculdade de Ciências
Departamento de Ciência da Computação

Bauru, _____ de _____ de _____.

Dedico esta monografia à minha família, amigos, colegas, professores e a todos aqueles que de alguma colaboraram para que eu pudesse chegar até aqui.

Agradecimentos

Agradeço imensamente à minha família, por todo apoio e incentivo para que eu trilhasse esse caminho, e tudo o que me ensinaram e me permitiram aprender.

Agradeço à todos os professores que me guiaram até este momento, não só os da graduação, mas também à todos que fizeram parte da minha formação desde o Ensino Básico. Agradeço em especial o Prof. Assoc. Aparecido Nilceu Marana, por toda orientação e aprendizado que me proporcionou durante a graduação tanto nas disciplinas que ministrou quanto nos projetos de pesquisa que desenvolvemos juntos.

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

E por fim, agradeço à todos os colegas e amigos de curso que fizeram parte desta caminhada, por todo apoio, companhia e momentos que compartilhamos.

Somewhere, something incredible is waiting to be known.

Carl Sagan

Resumo

Nas últimas décadas a biometria, utilizada para identificação de indivíduos por meio do estudo de suas características físicas e comportamentais, tem se tornado uma ferramenta cada vez mais presente no cotidiano da sociedade, sendo aplicada em cenários cada vez mais desafiadores. Apesar disso, ainda existe um importante desafio a ser superado: a identificação automática de indivíduos em vídeos de baixa resolução, à distância, de maneira sigilosa e não invasiva. Nesse cenário, em que os sistemas biométricos convencionais não se mostram eficazes, o reconhecimento de marcha ganha destaque. Entretanto, mesmo os métodos de reconhecimento de marcha considerados estado-da-arte ainda não apresentam a robustez adequada para serem utilizados em situações reais, para solucionar essa questão, pode-se adotar técnicas de multibiometria, que buscam combinar diferentes métodos biométricos em um único sistema. Neste trabalho, foram estudados e avaliados dois métodos de reconhecimento de marcha baseados em poses 2D e um método baseado em medidas antropométricas, bem como a fusão entre esses métodos utilizando multibiometria. O resultado dos experimentos mostra que, no geral, a fusão entre os métodos apresentam melhorias consideráveis em termos de acurácia. Por fim, foi desenvolvido um sistema biométrico baseado no reconhecimento de marcha utilizando a fusão de métodos que apresentou o melhor resultado.

Palavras-chave: Biometria; Reconhecimento de Marcha; Multibiometria; Estimação de Pose.

Abstract

In the last decades, biometrics, used to identify individuals through the study of their physical and behavioral characteristics, has become an increasingly present tool in society's daily life, being applied in increasingly challenging scenarios. Despite this, there is still an important challenge to be overcome: the automatic identification of individuals in low resolution videos, at a distance, in a confidential and non-invasive way. In this scenario, in which conventional biometric systems are not effective, gait recognition is highlighted. However, even state-of-the-art gait recognition methods still do not have the adequate robustness to be used in real situations, to solve this issue, multibiometric techniques can be adopted, which seek to combine different biometric methods into a single system. In this work, two gait recognition methods based on 2D poses and a method based on anthropometric measurements were studied and evaluated, as well as the fusion between these methods using multibiometrics. The results of the experiments show that, in general, the fusion between the methods presents considerable improvements in terms of accuracy. Finally, a biometric system based on gait recognition was developed using the fusion of methods that presented the best result.

Keywords: Biometrics; Gait Recognition; Multibiometrics; Pose Estimation.

Listas de figuras

Figura 1 – Características comumente extraídas em sistemas baseados em digitais, íris e face.	15
Figura 2 – Exemplo de aplicação do método OpenPose em uma imagem contendo múltiplas pessoas.	16
Figura 3 – Estrutura dos sistemas biométricos divididos em duas fases.	21
Figura 4 – Segmentação e aprimoramento de qualidade de uma face. À esquerda, a imagem capturada, e à direita, o resultado após a fase de pré-processamento.	22
Figura 5 – Diagrama de Voronoi representando o método 1-NN.	24
Figura 6 – Fase de autenticação de um sistema biométrico baseado em verificação.	25
Figura 7 – Fase de autenticação de um sistema biométrico baseado em identificação.	26
Figura 8 – Modelos de poses 2D usados pelos métodos OpenPose (a) e PifPaf (b).	27
Figura 9 – Funcionamento do método OpenPose.	28
Figura 10 – Arquitetura da CNN utilizada no OpenPose.	28
Figura 11 – <i>Pipeline</i> do PifPaf.	29
Figura 12 – Ciclo de marcha representado em uma sequência de quatro <i>frames</i> .	30
Figura 13 – Exemplo da aplicação dos detectores de pose OpenPose e PifPaf.	30
Figura 14 – Exemplo do uso de silhuetas para representar o ciclo de marcha.	31
Figura 15 – Diagrama de blocos do método proposto para reconhecimento de marcha.	32
Figura 16 – Fluxo de execução de um sistema biométrico baseado em multibiometria a nível de pontuação.	36
Figura 17 – Distribuição das comparações genuínas e impostoras de um sistema biométrico de verificação.	38
Figura 18 – Exemplo de uma curva ROC (TARxFAR).	39
Figura 19 – Exemplo de uma curva CMC.	41
Figura 20 – Exemplos de vídeos do CASIA Gait Dataset-A.	42
Figura 21 – Exemplos de variações de condições de vestimenta e transporte do CASIA Gait Dataset-B.	43
Figura 22 – Esquemas de configuração dos experimentos realizados.	44
Figura 23 – Curvas CMC - CASIA Gait Datset-A (média das três direções) - Distância Euclidiana.	47
Figura 24 – Curvas CMC - CASIA Gait Datset-A (média das três direções) - Distância Qui-quadrado.	48
Figura 25 – Curvas CMC - CASIA Gait Datset-B (apenas lateral).	48
Figura 26 – Principais componentes do módulo de sensor do sistema biométrico.	50
Figura 27 – Resumo da arquitetura do sistema biométrico.	51
Figura 28 – Circuito do módulo de sensor do sistema biométrico.	52

Figura 29 – Pinagem do módulo ESP32-CAM.	52
Figura 30 – Pinagem do módulo HC-SR501.	53
Figura 31 – Esquema de funcionamento do servidor HTTP para extração de características.	54
Figura 32 – Diagrama do banco de dados utilizado pelo sistema biométrico.	55
Figura 33 – Fluxo da fase de registro através da interface gráfica.	56
Figura 34 – Painel de monitoramento em tempo real da fase de autenticação.	57
Figura 35 – Exemplos de imagens capturadas para a base de dados do sistema.	58

Lista de tabelas

Tabela 1 – Acurárias rank-1 obtidas na base de dados CASIA Gait Dataset-A.	46
Tabela 2 – Acurárias rank-1 obtidas na base de dados CASIA Gait Dataset-A - Qui-quadrado.	47
Tabela 3 – Acurárias rank-1 obtidas na base de dados CASIA Gait Dataset-B.	47

Lista de abreviaturas e siglas

1-NN	<i>1-Nearest Neighbours</i>
AUC	<i>Area Under the Curve</i>
CMC	<i>Cumulative Match Characteristic</i>
CNN	<i>Convolutional Neural Network</i>
CPU	<i>Central Processing Unit</i>
EER	<i>Equal Error Rate</i>
FAR	<i>False Accept Rate</i>
FNIR	<i>False Negative Identification Rate</i>
FPIR	<i>False Positive Identification Rate</i>
FRR	<i>False Reject Rate</i>
FTP	<i>File Transfer Protocol</i>
GPU	<i>Graphical Processing Unit</i>
HTTP	<i>Hypertext Transfer Protocol</i>
IoT	<i>Internet of Things</i>
K-NN	<i>K-Nearest Neighbours</i>
LDA	<i>Linear Discriminant Analysis</i>
PAF	<i>Part Assotiation Fields</i>
PIF	<i>Part Intensity Fields</i>
ROC	<i>Receiver Operating Characteristic</i>
SGBD	Sistema Gerenciador de Banco de Dados
SQL	<i>Structured Query Language</i>
TAR	<i>True Accept Rate</i>
TPIR	<i>True Positive Identification Rate</i>
USB	<i>Universal Serial Bus</i>

Sumário

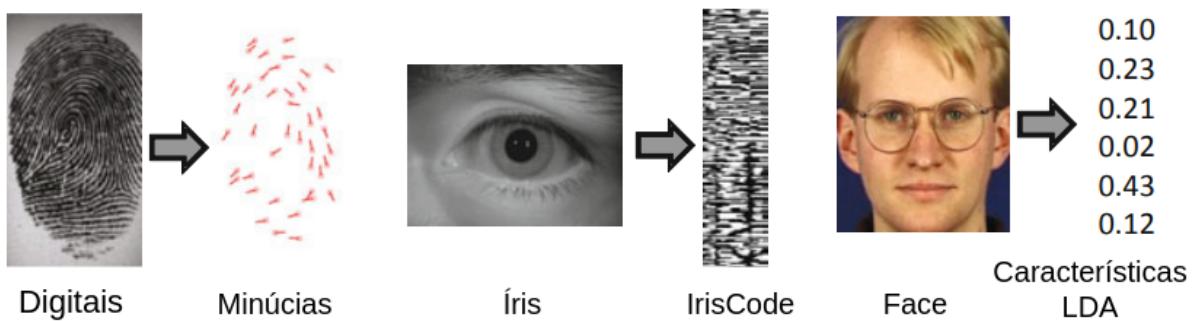
1	INTRODUÇÃO	15
1.1	Problema	17
1.2	Justificativa	17
1.3	Objetivos	18
1.4	Organização do Trabalho	19
2	FUNDAMENTAÇÃO TEÓRICA	20
2.1	Sistemas Biométricos	20
2.1.1	Sensores	20
2.1.2	Extração de características	21
2.1.3	Base de Dados	23
2.1.4	Classificação	23
2.1.4.1	Classificador K-NN	23
2.2	Funcionalidades Biométricas	24
2.2.1	Verificação	24
2.2.2	Identificação	25
2.3	Estimação de Pose	26
2.3.1	OpenPose	26
2.3.2	PifPaf	28
2.4	Reconhecimento de Marcha	29
2.4.1	Método Baseado nas Posições das Partes dos Membros	31
2.4.1.1	Estimação de Pose	31
2.4.1.2	Extração das Características	32
2.4.1.3	Histogramas	33
2.4.1.4	Classificação	33
2.4.2	Método Baseado em Histogramas de Movimento	33
2.5	Biometria Suave e Multibiometria	34
2.5.1	Método Baseado em Medidas Antropométricas	36
2.6	Métricas de Desempenho	37
2.6.1	Medidas para Sistemas de Verificação	37
2.6.2	Medidas para Sistemas de Identificação	40
3	METODOLOGIA	42
3.1	Bases de Dados	42
3.1.1	CASIA Gait Dataset-A	42
3.1.2	CASIA Gait Dataset-B	42

3.2	Experimentos	43
4	DESENVOLVIMENTO	45
4.1	Experimentos	45
4.2	Resultados e Discussão	46
5	SISTEMA BIOMÉTRICO	50
5.1	Sensor	51
5.2	Extração de Características	53
5.3	Base de Dados	54
5.4	Registro e Autenticação	55
5.5	Resultados e Discussão	57
6	CONCLUSÃO	59
	REFERÊNCIAS	60

1 Introdução

A biometria, que consiste no estudo estatístico de características físicas ou comportamentais dos seres humanos, também chamadas de traços biométricos ([JAIN; ROSS; NANDAKUMAR, 2011](#)), tem se tornado uma ferramenta cada vez mais essencial para a identificação de indivíduos na sociedade, seja para evitar fraudes, controlar acesso a áreas ou informações restritas, ou até mesmo identificar criminosos em ambientes públicos como portos e aeroportos. Com os avanços dos estudos no campo da biometria, essa identificação tende a ser cada vez mais rápida e acurada. Atualmente já existem alguns métodos biométricos consolidados e amplamente aplicados, que utilizam traços como digitais, íris e face para identificar os indivíduos, como as abordagens ilustradas na Figura 1.

Figura 1 – Características comumente extraídas em sistemas baseados em digitais, íris e face.



Fonte: Adaptado de [Jain, Ross e Nandakumar \(2011\)](#).

Apesar de todo o desenvolvimento na área, ainda resta um importante desafio a ser superado pela biometria: a identificação de pessoas em tempo real por meio de imagens de vídeo em baixa resolução (como câmeras de segurança), de forma discreta, à distância e sem a necessidade de cooperação do sujeito a ser identificado. Neste contexto, a utilização de técnicas clássicas de biometria como íris ou digitais se torna inviável, e o uso de características faciais muito difícil, uma vez que o desempenho desses métodos possui uma forte dependência em relação à qualidade de resolução das imagens capturadas, além de não apresentar bons resultados na presença de elementos como bonés, chapéus e óculos escuros, por exemplo ([ARANTES, 2010](#)). Em casos como esse, o uso de características da marcha pode ser a única opção viável ([WAN; LI; PHOHA, 2018](#)).

A marcha, ou o ato de caminhar, pode ser descrita como uma composição de comportamentos motores formando movimentos integrados do corpo humano, que se repetem periodicamente, formando o chamado "ciclo de marcha" ([ARANTES, 2010](#)). Este ciclo pode ser compreendido como um padrão de assinatura que, segundo estudos, possui unicidade suficiente para ser utilizado na identificação de indivíduos ([NIXON; CARTER, 2006](#)). Os métodos de

reconhecimento de marcha podem ser classificados em duas categorias: os baseados em modelo, que utilizam um modelo espaço-temporal do corpo humano para representar a caminhada e extraír informações de marcha (LIU et al., 2016) (LIMA; SCHWARTZ, 2019) (JANGUA; MARANA, 2020), e os livres de modelo, que se baseiam na análise da deformação da silhueta do indivíduo durante a caminhada para extraír informações de marcha (YU et al., 2007) (CHEN et al., 2009).

Os métodos de reconhecimento de marcha mais modernos, baseados em modelo, utilizam ferramentas de estimativa de pose para mapear a posição dos membros do indivíduo na cena. A estratégia adotada para estimativa de pose é um fator crítico para esse tipo de método, principalmente se tratando da detecção de múltiplas pessoas em uma imagem. Estimadores de pose como o OpenPose de Cao et al. (2018) e o PifPaf de Kreiss, Bertoni e Alahi (2019) adotam abordagens que buscam reduzir ao máximo o esforço computacional sem perder a confiança da detecção, utilizando estratégias *bottom-up* (de baixo para cima) baseadas em informações globais da imagem, o que proporciona um custo computacional invariável em relação à quantidade de pessoas na imagem. A Figura 2 ilustra o resultado obtido pelo estimador OpenPose (CAO et al., 2018) em uma imagem com múltiplas pessoas.

Figura 2 – Exemplo de aplicação do método OpenPose em uma imagem contendo múltiplas pessoas.



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

A multibiometria se baseia na premissa de que os seres humanos identificam uns aos outros baseando-se em múltiplas características biométricas. Esse processo pode ser visto como uma reconciliação de evidências em meio à múltiplas fontes de informação. Dessa forma, o objetivo da multibiometria é realizar a identificação biométrica baseando-se em múltiplos métodos ou traços biométricos fundidos em algum nível do sistema biométrico com o objetivo de melhorar o desempenho e aumentar a robustez desse sistema. Essa fusão pode ocorrer em nível de sensores, nível de características, nível de pontuação (do inglês, "score"), nível de classificação (do inglês, "ranking") ou nível de decisão (JAIN; ROSS; NANDAKUMAR, 2011).

De acordo com Jain, Dass e Nandakumar (2004), as características de biometria suave

(do inglês, "*soft biometrics*") são definidas como traços que proporcionam algum tipo de informação discriminativa sobre algum indivíduo, mas não suficiente para determinar sua identidade por si só. As técnicas de biometria suave são normalmente utilizadas em conjunto com métodos que utilizam traços biométricos com unicidade suficiente para identificação do indivíduo, fundidos por meio de técnicas de multibiometria, com a finalidade de acrescentar mais robustez ao sistema de identificação biométrica ([JAIN; DASS; NANDAKUMAR, 2004](#)).

1.1 Problema

Os métodos biométricos mais comuns como os baseados em reconhecimento de face, íris e digitais não apresentam características que permitem sua utilização em cenários em que se faz necessária a identificação à distância de indivíduos, sem necessidade de cooperação e em imagens de baixa resolução ([ARANTES, 2010](#)). Uma possível solução para esse problema é a utilização de características de marcha para realizar a identificação, uma vez que essa abordagem atende aos requisitos do problema ([WAN; LI; PHOHA, 2018](#)).

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, estas técnicas apresentam resultados mais robustos na classificação, uma vez que utilizam as características dinâmicas da marcha ([ARANTES, 2010](#)). 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.

Mesmo as técnicas de reconhecimento de marcha consideradas estado-da-arte, ainda apresentam espaço para melhorias, uma vez que a robustez destes métodos, no geral, ainda não são suficientes para o funcionamento adequado em ambientes reais. O fluxo clássico de melhoria desses métodos se baseia na reconstrução dos mesmos alterando sua estrutura básica, o que gera retrabalho e demanda um esforço considerável, principalmente se tratando de abordagens estado-da-arte.

1.2 Justificativa

A biometria já é um elemento essencial para garantir a segurança dos indivíduos na sociedade. Entretanto as técnicas clássicas de biometria ainda não solucionam alguns desafios importantes que ainda representam obstáculos para que sistemas biométricos possam ser utilizados de forma mais 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).

O uso de métodos de reconhecimento de marcha baseados em modelo, como os propostos por Liu et al. (2016), Lima e Schwartz (2019) e Jangua e Marana (2020), mostram-se os mais adequados para a situação, uma vez que apresentam melhores resultados quando comparados com métodos livres de modelo (ARANTES, 2010) 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.

A utilização de técnicas de multibiometria permite a fusão entre diferentes métodos, compondo um único sistema biométrico, que busca unir as qualidades dos métodos fundidos de modo que suas deficiências sejam atenuadas, atribuindo uma maior robustez ao sistema como um todo. Entre as principais vantagens do uso da multibiometria, além da melhoria de acurácia estão o tratamento da não universalidade do método e maior tolerância à ruídos e falhas (ROSS; NANDAKUMAR; JAIN, 2006).

1.3 Objetivos

Objetivo Geral:

Implementar um sistema de reconhecimento de marcha para controle de acesso a áreas restritas utilizando poses 2D extraídas de vídeos capturados por uma única câmera de baixa resolução.

Objetivos Específicos:

- a) Estudar métodos de reconhecimento de marcha baseados em poses 2D.
- b) Implementar os métodos de reconhecimento de marcha escolhidos.
- c) Comparar o desempenho dos métodos na tarefa de reconhecimento biométrico utilizando dois estimadores de pose diferentes.
- d) Montar a estrutura física do sistema biométrico, utilizando a câmera de baixa resolução e um sensor de presença.
- e) Capturar imagens de pessoas previamente identificadas caminhando, utilizando a câmera de baixa resolução, para compor a base de dados do sistema biométrico.
- f) Desenvolver um sistema para computadores que receba uma nova captura de vídeo de algum indivíduo caminhando, aplique um dos métodos de reconhecimento de marcha selecionados e realize o reconhecimento biométrico do indivíduo.

1.4 Organização do Trabalho

Os capítulos seguintes deste trabalho são organizados da seguinte forma:

Capítulo 2: Apresenta os conceitos de biometria e reconhecimento de marcha que serviram como base para toda a metodologia desenvolvida, além de apresentar em detalhes cada um dos métodos que foram implementados.

Capítulo 3: Apresenta as bases de dados utilizadas para os experimentos, bem como a configuração adotada em cada experimento realizado.

Capítulo 4: Apresenta cada etapa do processo de desenvolvimento do trabalho em detalhes, além de apresentar e discutir os resultados obtidos pelos experimentos.

Capítulo 5: Apresenta os detalhes do desenvolvimento e estrutura do sistema biométrico construído, seguindo a lógica de arquitetura padrão apresentada no Capítulo 2.

Capítulo 6: Apresenta as conclusões obtidas pelo trabalho, tanto na etapa dos experimentos, quanto no desenvolvimento do sistema.

2 Fundamentação Teórica

A execução deste trabalho exige conhecimento de conceitos relacionados a sistemas biométricos, reconhecimento de marcha, estimação de pose 2D e multibiometria, portanto, nesta seção estão apresentados conceitos chave referentes às áreas de biometria e reconhecimento de marcha, incluindo informações acerca das técnicas utilizadas para as tarefas de reconhecimento, bem como os métodos de estimação de pose utilizados. Além disso, são abordados conceitos relacionados à microprocessadores e microcontroladores, com foco no equipamento utilizado para a construção do sistema biométrico proposto neste trabalho.

2.1 Sistemas Biométricos

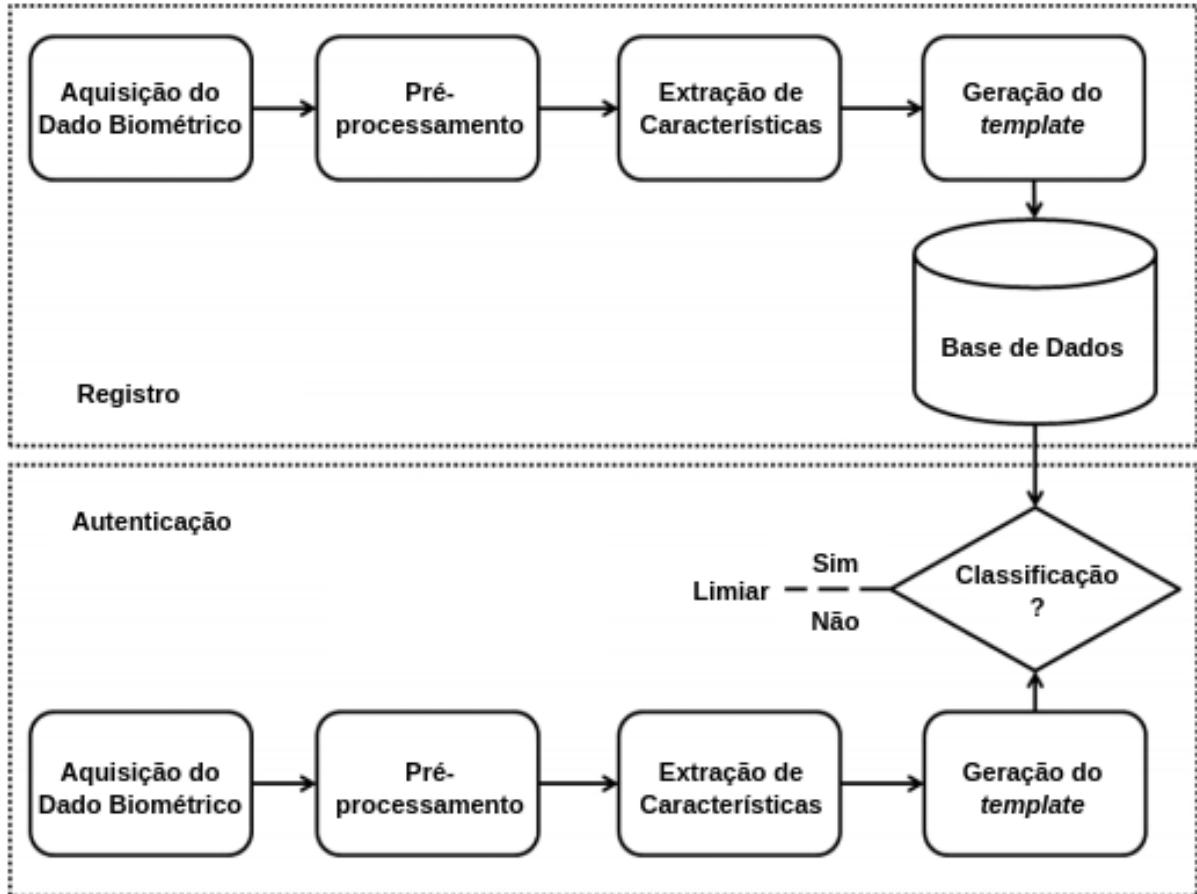
Um sistema biométrico utiliza medidas estatísticas de uma ou mais características de indivíduos, sejam elas físicas ou comportamentais, para reconhecê-los ou verificar sua identidade. Esses traços podem incluir características como digital, íris, face, voz, ou até mesmo a maneira como a pessoa caminha. Todo sistema biométrico possui duas fases, a fase de registro, e a de autenticação. Na fase de registro são coletadas amostras de características dos indivíduos e armazenadas em uma base de dados junto ao seu respectivo rótulo (identificação do indivíduo). Já na fase de autenticação, são comparados dados obtidos em tempo real com os contidos na base de dados para realizar o reconhecimento ([CHOWDHURY et al., 2018](#)). A Figura 3 ilustra a estrutura dos sistemas biométricos. Neste capítulo será abordada a construção de cada parte que compõe a estrutura básica de um sistema biométrico genérico. Essas partes podem ser divididas em: sensores, extração de características, base de dados e classificador ([JAIN; ROSS; NANDAKUMAR, 2011](#)).

2.1.1 Sensores

Os sensores compõem uma unidade de extrema importância para o desempenho de um sistema biométrico. Uma interface com o usuário adequada para o cenário em que o sistema está inserido é essencial para obter e registrar os dados biométricos brutos dos indivíduos com qualidade e confiabilidade. Segundo [Jain, Ross e Nandakumar \(2011\)](#), o *design* dessa interface é um fator crítico para uma implementação de sucesso, uma vez que devem ser consideradas questões como intuitividade, ergonomia, facilidade e rapidez além da qualidade das amostras extraídas dos indivíduos.

Na maioria dos sistemas biométricos os dados brutos das amostras são imagens bidimensionais, como é o caso de sistemas baseados em digitais, face, íris, etc. Para os dados baseados em imagens, os fatores que mais influenciam na qualidade das amostras são resolução,

Figura 3 – Estrutura dos sistemas biométricos divididos em duas fases.



Fonte: Adaptado de [Choudhury et al. \(2018\)](#).

taxa de quadros e sensibilidade da câmera. Outra questão importante que deve ser analisada no planejamento de um sensor são as características demográficas da região em que o sistema será aplicado (como idade, gênero e cultura), além de fatores como o custo, tamanho e durabilidade ([JAIN; ROSS; NANDAKUMAR, 2011](#)). Entre os sensores mais comumente utilizados por sistemas biométricos estão o leitor de digitais, a câmera e o leitor de íris.

2.1.2 Extração de características

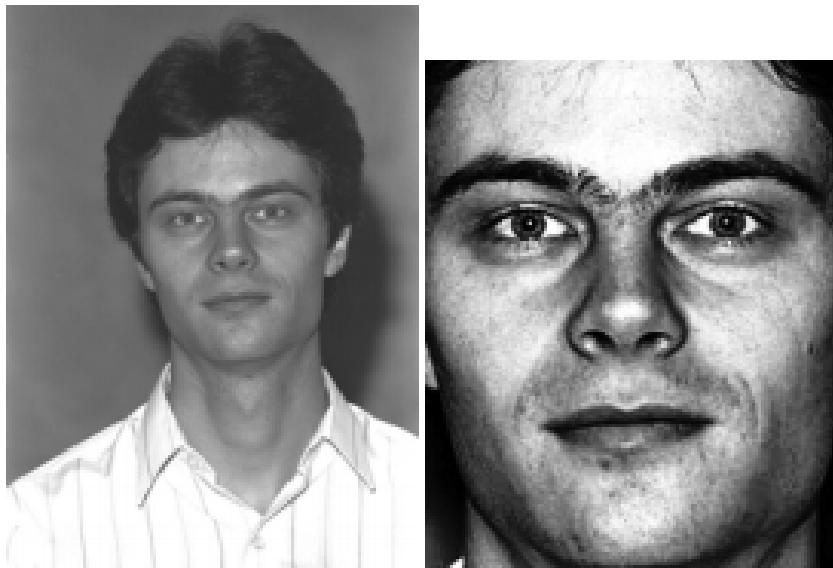
Normalmente os dados extraídos pelos sensores não podem ser usados diretamente no sistema biométrico, uma vez que podem apresentar certa quantidade de ruídos, variações de ângulo, iluminação, posição, e até mesmo a presença de informações que não colaboram para a identificação do indivíduo. Assim, os dados brutos são submetidos à uma operação de pré-processamento antes mesmo de realizar a extração de características. De acordo com [Jain, Ross e Nandakumar \(2011\)](#) há três passos que são comumente aplicados para o pré-processamento: avaliação da qualidade, segmentação e aprimoramento.

O primeiro passo do pré-processamento é a avaliação da qualidade da amostra. Nessa fase o sistema analisa se o dado coletado possui qualidade adequada para seguir com o

processamento, caso a resposta seja negativa, o sistema para o processamento e pode exigir que a amostra seja coletada novamente. Caso a resposta seja positiva, o sistema segue para a etapa de segmentação. Nesta etapa, o objetivo é separar a informação que será realmente usada do restante, que é considerado ruído de fundo (como por exemplo o recorte de faces em um sistema baseado em reconhecimento facial).

Por fim, os dados biométricos segmentados são submetidos à fase de aprimoramento, onde um algoritmo de aprimoramento de qualidade busca melhorar a qualidade dos dados reduzindo os ruídos gerados na captura do dado. Quando os dados biométricos são imagens, esse algoritmo normalmente é referente à suavização ou equalização de histograma da imagem. A Figura 4 mostra a imagem de uma face obtida (à esquerda) e o resultado (à direita) após a segmentação e aprimoramento de qualidade.

Figura 4 – Segmentação e aprimoramento de qualidade de uma face. À esquerda, a imagem capturada, e à direita, o resultado após a fase de pré-processamento.



Fonte: [Jain, Ross e Nandakumar \(2011\)](#).

Após o pré-processamento ocorre o processo de extração de características propriamente dito. O objetivo é gerar uma representação digital mais compacta e expressiva das características biométricas obtidas. Essa representação é chamada de *template* ("modelo", em português) e deve conter apenas a informação discriminatória essencial para o reconhecimento do indivíduo ([JAIN; ROSS; NANDAKUMAR, 2011](#)).

Essa fase é definida especificamente para cada tipo de sistema biométrico, podendo assumir diferentes tipos de representação para cada abordagem, como por exemplo: sistemas baseados em digitais utilizam um conjunto de pontos que representam as minúcias; íris é representada como um vetor binário que descreve a resposta binarizada de uma imagem de entrada para filtros de Gabor; e a face, comumente representada como um vetor de coeficientes obtidos pelo método de LDA (*Linear Discriminant Analysis* - "Análise Discriminante Linear",

em português). A Figura 1 mostra as características que normalmente são extraídas em sistemas baseados em digitais, íris e face, respectivamente.

Após a construção dos *templates* na fase de registro, os mesmos são guardados em uma base de dados do sistema biométrico. Na fase de autenticação esse *template* é recuperado na base e comparado com o que foi obtido em tempo real pelo sistema. Os dados guardados na base durante a fase de registro são chamados de *gallery* ("galeria", em português), já as obtidas no momento da autenticação são normalmente chamadas de *probe* ("sonda", em português) (JAIN; ROSS; NANDAKUMAR, 2011).

2.1.3 Base de Dados

A base de dados de um sistema biométrico funciona como um repositório de informações biométricas, armazenando os *templates* obtidos na fase de registro a partir dos dados brutos. O conjunto de características extraídas é armazenado juntamente com um rótulo ou alguma informação que identifique o indivíduo ao qual determinada amostra pertence, para futuras consultas na fase de autenticação.

Segundo Jain, Ross e Nandakumar (2011), uma importante decisão no projeto de um sistema biométrico é se a base de dados será centralizada ou distribuída. Armazenar todas as informações em um ponto central pode ser vantajoso de um ponto de vista de segurança do sistema, podendo garantir a segurança e integridade dos dados através do isolamento físico e controle de acesso restrito. Por outro lado, o comprometimento do ponto central resultaria na parada de todo o sistema.

2.1.4 Classificação

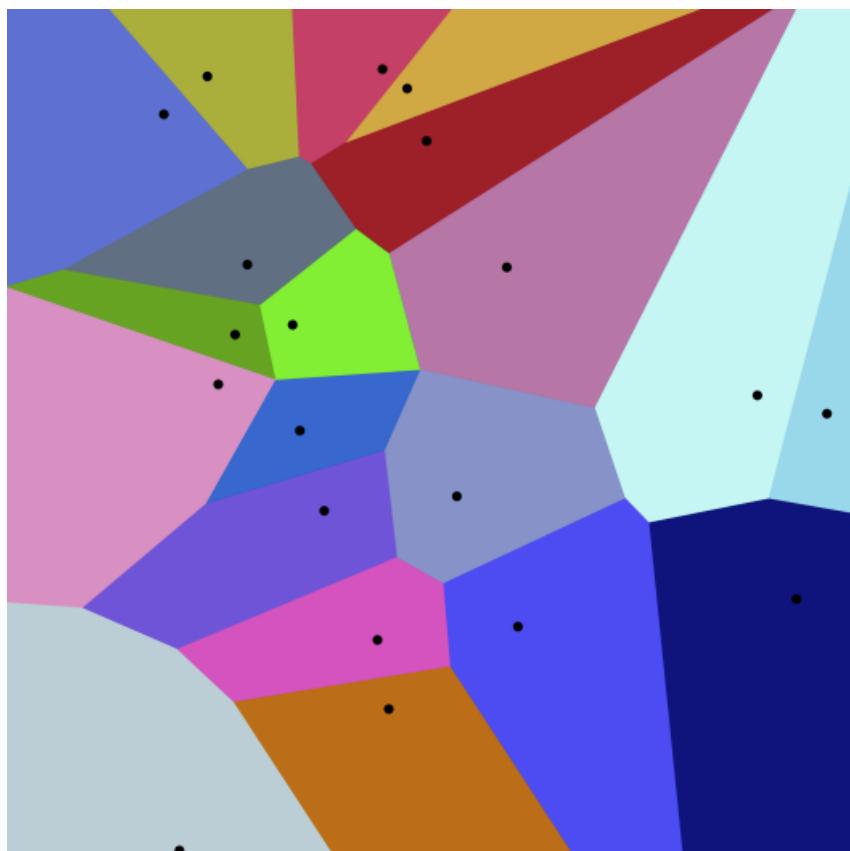
O objetivo do classificador em um sistema biométrico é comparar as amostras coletadas em tempo real (*probe*) com as amostras *gallery*, armazenadas na base de dados e gerar uma pontuação de correspondência (do inglês "*matching score*", ou apenas "*score*"), que é uma medida de similaridade entre as duas amostras, ou seja, quanto maior a pontuação, mais próximas estão as duas amostras. O classificador pode também calcular a dissimilaridade entre as amostras, nesse caso quanto menor o valor mais próximas elas estão. O classificador também inclui um módulo responsável pela tomada de decisão, baseando-se no *score* para validar a identidade do indivíduo ou reconhecê-lo (JAIN; ROSS; NANDAKUMAR, 2011).

2.1.4.1 Classificador K-NN

O classificador de *K* vizinhos mais próximos, também chamado de K-NN (*K-Nearest Neighbours*), é uma das abordagens mais comuns em aplicações de reconhecimento de padrões. O método busca definir a classe de um objeto baseado nos *k* vizinhos mais próximos de sua posição no espaço de características (COVER; HART, 1967). Para definir a proximidade entre

as amostras, utilizam-se funções de distância, podendo variar de acordo com o escopo de cada aplicação. A variação mais simples desse classificador é considerar apenas o vizinho mais próximo para definir o rótulo do objeto, definindo $k = 1$, comumente chamado de 1-NN (*1-Nearest Neighbour*). Segundo [Mitchel \(1997\)](#), para este caso, o espaço de características do problema pode ser representado pelo diagrama de Voronoi. Na Figura 5, cada cor representa a região mais próxima de cada ponto. Utilizando o 1-NN, cada amostra seria classificada de acordo com a região delimitada pelo ponto que define a região.

Figura 5 – Diagrama de Voronoi representando o método 1-NN.



Fonte: [MYSID e CYP \(2022\)](#).

2.2 Funcionalidades Biométricas

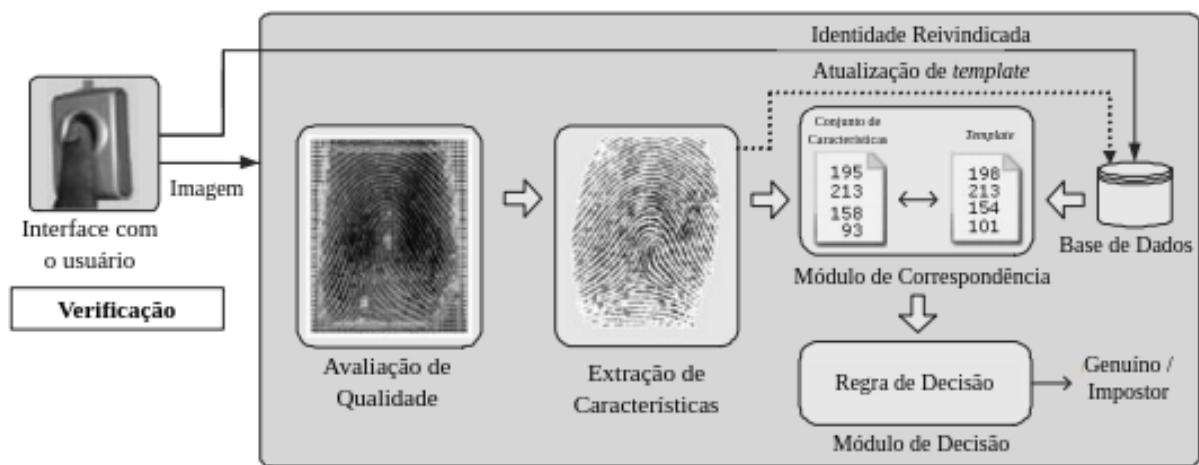
Um sistema biométrico pode oferecer duas funcionalidades distintas, a verificação e a identificação. A verificação busca validar uma identidade que o indivíduo em questão está reivindicando, enquanto a função de identificação busca atribuir uma identidade para um indivíduo detectado. A seguir são apresentadas cada uma delas.

2.2.1 Verificação

Na verificação um usuário reivindica uma determinada identidade, e cabe ao sistema decidir se essa reivindicação é genuína ou não através da análise dos traços biométricos do

indivíduo. Nesse caso, a amostra coletada é comparada apenas ao *template* correspondente à identidade que foi reivindicada pelo usuário. Caso a amostra coletada e o *template* possuam um alto grau de similaridade o sistema aceita a reivindicação como genuína. Caso contrário, ela é rejeitada e o usuário é considerado um impostor. Sistemas de verificação são comumente aplicados em cenários em que o objetivo é restringir o acesso a certos serviços para pessoal não autorizado (JAIN; ROSS; NANDAKUMAR, 2011). A Figura 6 mostra o funcionamento de um sistema de verificação na fase de autenticação.

Figura 6 – Fase de autenticação de um sistema biométrico baseado em verificação.



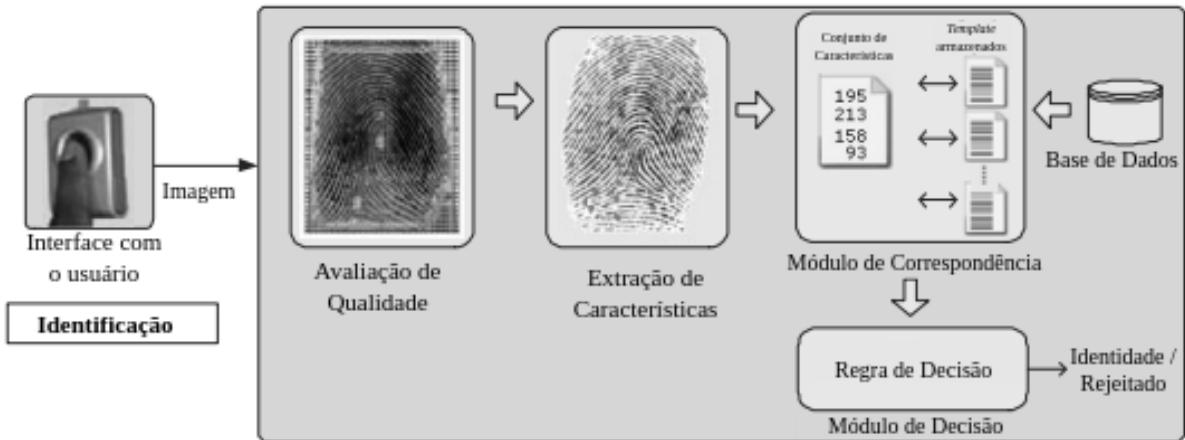
Fonte: Adaptado de Jain, Ross e Nandakumar (2011).

2.2.2 Identificação

De acordo com Jain, Ross e Nandakumar (2011), a funcionalidade de identificação pode ainda ser dividida em duas classes, a identificação positiva e a negativa. Na identificação positiva o sistema busca responder a pergunta "Você é alguém conhecido pelo sistema?", determinando a identidade do indivíduo com base em um conjunto de identidades conhecidas. Já a identificação negativa tenta responder a pergunta "Você é quem diz não ser?", considerando que o usuário está negando sua identidade. O propósito dos sistemas de identificação negativa é impedir que uma pessoa utilize múltiplas identidades.

Tanto na identificação positiva quanto na negativa a amostra coletada do usuário é comparada com os *templates* de todas as pessoas da base de dados, e o sistema retorna como saída a identidade da pessoa cujo *template* mais se assemelha à amostra ou uma resposta indicando que a pessoa analisada não está registrada na base de dados. A Figura 7 mostra o funcionamento de um sistema de identificação na fase de autenticação.

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



Fonte: [Jain, Ross e Nandakumar \(2011\)](#).

2.3 Estimação de Pose

Estimação de pose 2D de múltiplas pessoas em tempo real é um importante componente em aplicações de visão computacional, que permite que máquinas tenham uma compreensão de pessoas em imagens e vídeos, e pode ser definida como o problema de localizar pontos-chave de partes do corpo dos indivíduos presentes na cena ([CAO et al., 2018](#)).

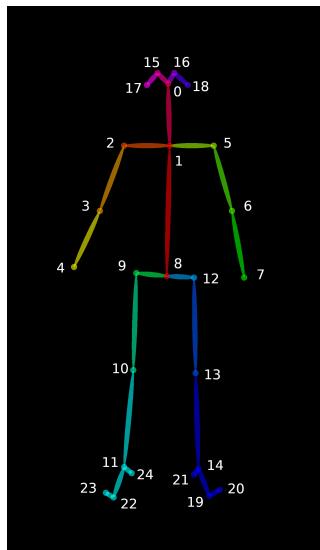
As abordagens para estimação de pose podem ser agrupadas em métodos *bottom-up* (de baixo para cima) e *top-down* (de cima para baixo). Os métodos *top-down* utilizam um detector de pessoas e em seguida um detector de pose para cada uma das detecções da imagem, o que faz com que a complexidade do método 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, essa abordagem possui vantagens em relação ao tempo de execução, uma vez que utiliza as informações globais da imagem de maneira direta ([CAO et al., 2018](#)).

Os dois métodos utilizados neste projeto de pesquisa, OpenPose ([CAO et al., 2018](#)) e PifPaf ([KREISS; BERTONI; ALAHI, 2019](#)), utilizam a estratégia *bottom-up*. Enquanto no OpenPose é possível utilizar uma opção de formato de saída com um modelo formado por 25 pontos chaves (Figura 8a), incluindo cabeça, tronco e pés, o modelo do PifPaf conta com 17 pontos chaves (Figura 8b), e não possui opções para detectar informações dos pés.

2.3.1 OpenPose

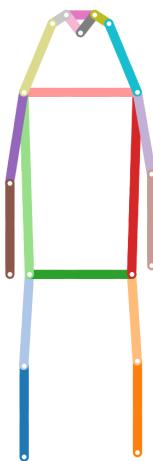
O OpenPose ([CAO et al., 2018](#)) é uma abordagem em tempo real para detecção de poses 2D de múltiplas pessoas em uma imagem, capaz de realizar as detecções com alta acurácia e um bom desempenho computacional. É o primeiro sistema de código aberto para detecção de pose 2D em tempo real que inclui corpo, pés, mãos e face. Diferente de muitas abordagens

Figura 8 – Modelos de poses 2D usados pelos métodos OpenPose (a) e PifPaf (b).



(a) Modelo do OpenPose.

Fonte: Cao et al. (2018).



(b) Modelo do PifPaf.

Fonte: Kreiss, Bertoni e Alahi (2019).

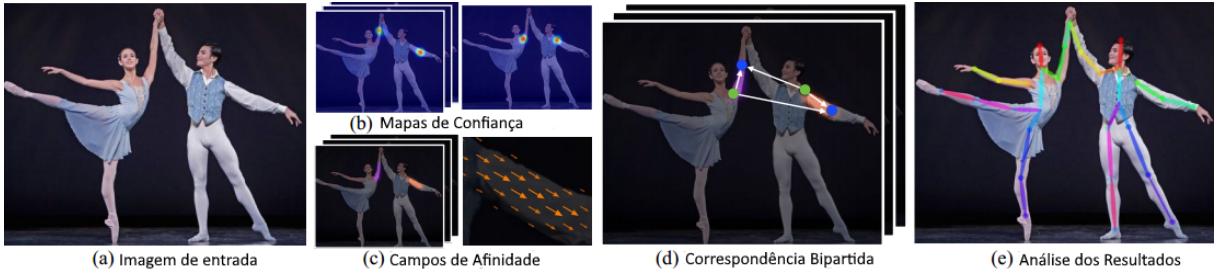
focadas em estimar a pose de indivíduos, essa aplicação trata a imagem inicialmente de forma global, com uma abordagem *bottom-up*, em contraste às abordagens anteriores que inicialmente detectavam cada indivíduo na imagem e suas respectivas poses um por vez, ocasionando um grande aumento do esforço computacional em função da quantidade de pessoas na imagem.

O método ganha performance e mantém a acurácia ao utilizar uma abordagem diferente chamada de campos de afinidade de 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 redes neurais convolucionais (CNN - *Convolutional Neural Network*) 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 probabilidade de ocorrer e gerar poses não válidas vão sendo eliminadas iterativamente até atingir o resultado final, com os pontos conectados e as poses detectadas ([CAO et al., 2018](#)). A Figura 9 ilustra o funcionamento do método.

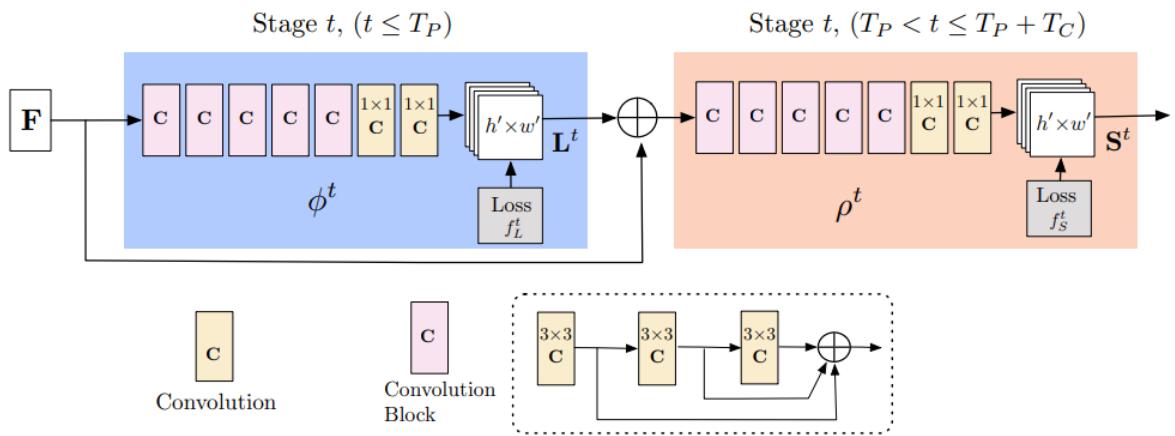
A Figura 10 mostra a arquitetura da CNN que prediz os campos de afinidade (que codificam a associação entre as partes) iterativamente, em azul e, em bege a detecção dos mapas de confiança. As previsões são refinadas iterativamente a cada estágio $t \in \{1, \dots, T\}$, com supervisão intermediária em cada estágio (CAO et al., 2018).

Figura 9 – Funcionamento do método OpenPose.



Fonte: [Cao et al. \(2018\)](#).

Figura 10 – Arquitetura da CNN utilizada no OpenPose.



Fonte: [Cao et al. \(2018\)](#).

2.3.2 PifPaf

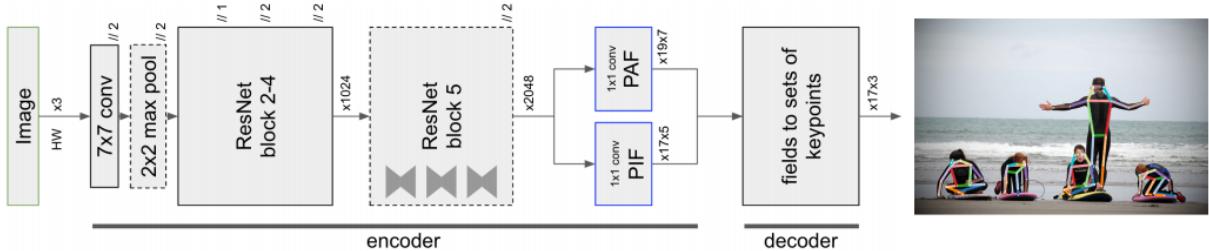
O PifPaf ([KREISS; BERTONI; ALAHI, 2019](#)) é um método de estimativa de poses 2D em tempo real, capaz de reconhecer poses de múltiplas pessoas presentes em uma imagem. Este método foi pensado para ser aplicado no ramo da navegação autônoma, tendo como principal foco a detecção de pessoas à longas distâncias, baixas resoluções e alto nível de oclusão, com o objetivo de aumentar a segurança no trânsito de veículos autônomos e robôs de entrega.

Nesta abordagem, são usadas dois tipos de redes neurais para desempenhar dois papéis importantes na detecção das poses. Para cada parte do corpo presente na imagem uma rede prediz a confiança, a posição precisa e o tamanho. Essa informação é chamada de PIF (*Part Intensity Fields* - Mapas de Intensidade de Partes) ([KREISS; BERTONI; ALAHI, 2019](#)) 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* - Campos de Associação de Partes), graças à esta abordagem, são obtidos resultados que superam as abordagens anteriores 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, 2019). A Figura 11 ilustra o *pipeline* do método, 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, 2019).

Figura 11 – Pipeline do PifPaf.



Fonte: Kreiss, Bertoni e Alahi (2019).

2.4 Reconhecimento de Marcha

Marcha pode ser definida como comportamentos motores compostos por movimentos repetitivos e integrados do corpo humano que formam um padrão de movimentos corporais que se repetem a cada ciclo (ARANTES; GONZAGA, 2011). Estudos anteriores nas áreas da Medicina e Psicologia mostraram que a marcha humana possui diversos componentes que podem ser usados para identificar um indivíduo (ARANTES; GONZAGA, 2011). Todos esses trabalhos indicam que cada indivíduo possui uma estrutura musculoesquelética única, fazendo com que o reconhecimento de marcha humana para identificação biométrica seja factível.

Segundo Arantes e Gonzaga (2011), o reconhecimento de marcha possui propriedades únicas que outros métodos biométricos não possuem: (i) a marcha humana pode ser capturada à distância e em baixas resoluções, (ii) o reconhecimento de marcha pode ser performado com instrumentação simples, como uma única câmera ou um acelerômetro, por exemplo, (iii) o reconhecimento da marcha humana não depende da cooperação do indivíduo para ser executado, (iv) a marcha é muito difícil de ser imitada e (v) funciona mesmo com oclusão parcial de partes do corpo do indivíduo.

De acordo com Nixon e Carter (2006), um ciclo de marcha é o intervalo de tempo entre instâncias sucessivas de contato inicial entre os pés e o chão, e cada perna possui dois períodos, uma fase de postura, quando o pé está em contato com o solo e uma fase de balanço, quando o pé está fora do chão e movendo-se para frente rumo ao próximo passo. A Figura 12 ilustra um ciclo de marcha resumido em quatro quadros de um vídeo retirado da base de dados pública CASIA Gait Dataset-A (WANG et al., 2003).

Os métodos utilizados para reconhecimento de marcha podem ser classificados em dois grupos de acordo com sua abordagem: a baseada em modelo e a livre de modelo. A

Figura 12 – Ciclo de marcha representado em uma sequência de quatro frames.



Fonte: Adaptado de [Wang et al. \(2003\)](#).

abordagem baseada em modelo incorpora informações da estrutura do corpo humano por meio de modelos do corpo humano pré-treinados. Métodos baseados em modelo variam muito em termos de complexidade e informações extraídas, além disso, podem ser utilizados diversos tipos de modelo, mas os mais comuns atualmente são baseados em detectores de pose robustos, como o OpenPose ([CAO et al., 2018](#)) e o PifPaf ([KREISS; BERTONI; ALAHI, 2019](#)). A Figura 13 ilustra, da esquerda para a direita, a aplicação do OpenPose (meio) e do PifPaf (direita), respectivamente, em um quadro de um vídeo. A principal vantagem do uso de um modelo, em detrimento da perda de performance computacional, é o ganho de robustez, reduzindo os erros induzidos por ruídos e obstruções nas imagens ([JAIN; ROSS; NANDAKUMAR, 2011](#)).

Figura 13 – Exemplo da aplicação dos detectores de pose OpenPose e PifPaf.



Fonte: Elaborado pelo autor.

Por outro lado, as abordagens livres de modelo são mais computacionalmente eficientes, uma vez que não depende de um detector de poses. Nessa abordagem são extraídas características baseando-se no movimento das silhuetas dos indivíduos em função do tempo. A Figura 14 mostra como a silhueta do indivíduo pode ser usada para representar o seu ciclo de marcha. As silhuetas do indivíduo, extraídas em cada quadro da imagem são utilizadas para gerar uma assinatura de silhueta, que pode ser usada para identificar o indivíduo ([JAIN; ROSS; NANDAKUMAR, 2011](#)).

De acordo com [Jain, Ross e Nandakumar \(2011\)](#), os maiores desafios encontrados em sistemas biométricos baseados em reconhecimento de marcha são relacionados à influência que fatores como vestimenta, tipo de calçado, superfície de caminhada, velocidade de caminhada e direção em relação à câmera, que exercem influência direta sobre a performance da classificação. Esses fatores ainda dificultam a aplicação e validação desses métodos em ambientes reais, por conta disso os métodos de reconhecimento de marcha são normalmente testados em bases de

Figura 14 – Exemplo do uso de silhuetas para representar o ciclo de marcha.



Fonte: [Jain, Ross e Nandakumar \(2011\)](#).

dados que consistem em ambientes controlados.

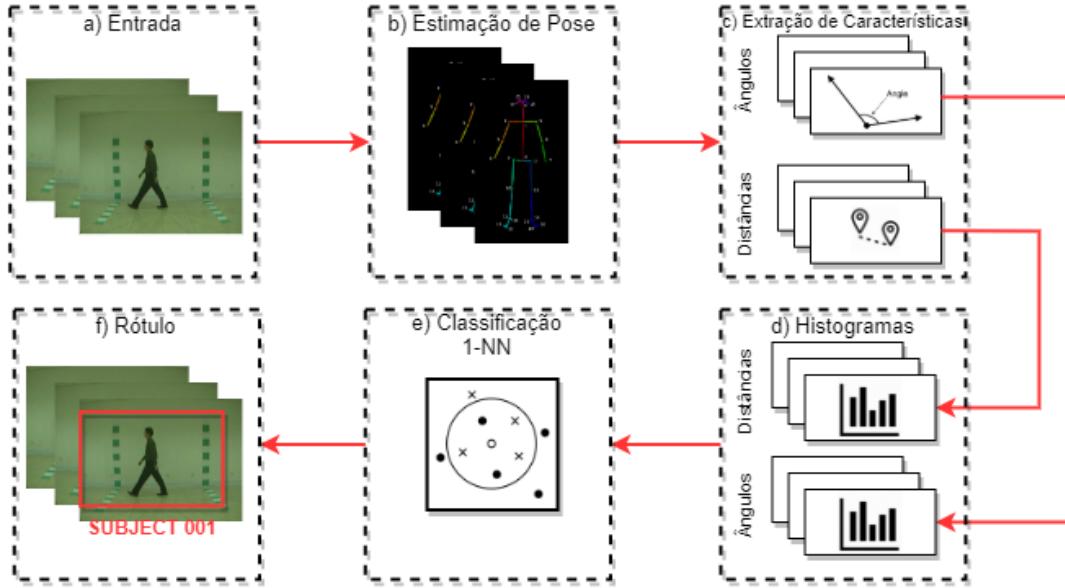
2.4.1 Método Baseado nas Posições das Partes dos Membros

Como o método é baseado na estimação de poses em vídeos dos indivíduos caminhando, a primeira etapa do método consiste no uso do OpenPose ([CAO et al., 2018](#)) para extrair as poses 2D dos indivíduos em cada quadro do vídeo de entrada. Em seguida, para cada quadro, as coordenadas dos pontos de juntas detectados são utilizados para calcular a angulação de cada parte dos membros em relação ao eixo horizontal da imagem e a distância entre o ponto que representa o pescoço e a reta que passa pelos dois pontos que delimitam determinada parte de membro. Essas informações são usadas para construir, para cada parte de membro, dois histogramas (um para os ângulos e outro para as distâncias) que representam o comportamento do mesmo durante toda a sequência de caminhada, e que são usados como vetores de características de marcha. Por fim, os dois vetores de características são usados por um classificador 1-NN (do inglês *Nearest Neighbour* - vizinho mais próximo) para determinar a identidade do indivíduo presente no vídeo de entrada. A Figura 15 ilustra o método passo-a-passo.

2.4.1.1 Estimação de Pose

A estimação de pose é o primeiro passo do método, nele o OpenPose ([CAO et al., 2018](#)) é aplicado no vídeo de entrada e, para cada *frame* do vídeo, obtém-se as coordenadas dos pontos-chave do corpo humano de acordo com o formato mostrado na Figura 8a. A saída do algoritmo corresponde a um arquivo em formato JSON, que contém, para cada quadro do vídeo, as coordenadas verticais e horizontais para cada ponto.

Figura 15 – Diagrama de blocos do método proposto para reconhecimento de marcha.



Fonte: Elaborado pelo autor.

2.4.1.2 Extração das Características

Depois da estimativa de pose, os pontos mais importantes para a análise de marcha (aqueles que codificam mais informação) são separados para serem usados nesta etapa, a saber: 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13 e 14, que formam os membros superiores e inferiores. Veja na Figura 8a o modelo de pose e os pontos utilizados neste trabalho. Para cada parte de membro uma sequência é construída, formada pela angulação entre a parte de membro e o eixo horizontal em cada frame. Para este cálculo, 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 φ , entre o vetor w e o eixo horizontal $(x_2, y_2) = (1, 0)$ pode ser obtido utilizando a Equação 2.1.

$$\varphi = \arccos \frac{x_1 * x_2 + y_1 * y_2}{\sqrt{x_1^2 + y_1^2} * \sqrt{x_2^2 + y_2^2}} \quad (2.1)$$

De forma semelhante 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 (ponto 1 na Figura 8a) 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 as Equações 2.2 e 2.3.

$$Proj_w v = \left(\frac{v \cdot w}{\|w\|^2} \right) * w \quad (2.2)$$

$$d = \|v - Proj_w v\| \quad (2.3)$$

2.4.1.3 Histogramas

Após calcular as sequências de ângulos e distâncias, para cada parte de membro são construídos dois histogramas, um para os ângulos e um para as distâncias, que representam o comportamento global da parte de membro em questão durante a sequência de marcha completa. Ambos histogramas são construídos com 16 *bins* cada, um valor encontrado empiricamente. Como neste método são considerados 8 partes de membros, para um vídeo são construídos 16 histogramas, 8 de cada tipo. Por fim, todos os histogramas de ângulo são concatenados, formando um único vetor de características, o mesmo é feito com os histogramas de distância. Assim, o descritor de marcha é formado por dois vetores contendo 128 características cada.

2.4.1.4 Classificação

Para o reconhecimento da marcha um classificador 1-NN é usado. Esse classificador funciona de maneira muito simples, são calculadas as distâncias entre o descritor obtido do vídeo a ser classificado (*probe*) e todos os descritores dos vídeos rotulados (*gallery*), sendo o rótulo daquele vídeo que apresentar a menor distância será dado ao vídeo de entrada. Neste caso, a distância entre duas amostras d_d é dada pela média aritmética entre a distância dos vetores de ângulos (d_1) e dos vetores de distâncias (d_2), ou seja, $d_d = (d_1 + d_2)/2$. Não é necessário normalizar as distâncias d_1 e d_2 para fazer esse cálculo, uma vez que os histogramas já são normalizados.

2.4.2 Método Baseado em Histogramas de Movimento

Este método proposto por [Lima e Schwartz \(2019\)](#) apresenta uma abordagem que utiliza inicialmente um método para a detecção das poses dos indivíduos, em seguida utiliza as coordenadas dos pontos chave de cada quadro para construir sinais que representam a posição de um ponto chave qualquer em função do ponto que representa o pescoço. Além desses sinais são analisadas as taxas de falhas de detecção de cada membro, guardando em uma estrutura separada os índices dos membros que possuem uma taxa de erro maior que uma tolerância γ .

Na sequência, calcula-se os Histogramas de Movimento, que capturam informação estática da marcha, dividindo os valores dos sinais em intervalos para representar a marcha. Em seguida realiza-se a comparação desses Histogramas para a classificação, utilizando a distância Euclidiana para o cálculo da semelhança entre as amostras.

Após extraídas as coordenadas do corpo pelo método de estimação de pose, obtemos como retorno para cada quadro t do vídeo i , onde b é o índice do ponto chave, $P_{b,t}^i$, sendo que $P_{b,t}^i$ é um par de coordenadas (x, y) , com coordenadas não negativas, exceto em casos de falha de detecção, que é retornado o par $(-1, -1)$.

Para esse método, dos 18 pontos retornados, utiliza-se 13, que são pescoço, pulsos, cotovelos, quadris, joelhos, tornozelos e ombros. Como as coordenadas inválidas geradas por

falha de detecção podem interferir no resultado, cria-se o índice de ruído N^i para o vídeo i , que guarda o índice de todos os membros que possuem um percentual de coordenadas inválidas maior que γ .

Para a construção dos Sinais S^i , que para cada ponto b usado possui duas linhas, representando as coordenadas x e y . Assim, S^i possui 24 linhas e um número de colunas igual à quantidade de quadros analisados no vídeo. O sinal é criado levando em consideração a posição do ponto relativa ao pescoço ($b = 1$), calculada como:

$$S_{2b-3,t}^i = \begin{cases} -1, & P_{b,t}^i \cdot x = -1 \\ \frac{P_{b,t}^i \cdot x - P_{1,t}^i \cdot x}{\max_j P_{j,t}^i \cdot y - P_{1,t}^i \cdot y}, & P_{b,t}^i \cdot x \neq -1 \end{cases} \quad (2.4)$$

$$S_{2b-2,t}^i = \begin{cases} -1, & P_{b,t}^i \cdot y = -1 \\ \frac{P_{b,t}^i \cdot y - P_{1,t}^i \cdot y}{\max_j P_{j,t}^i \cdot y - P_{1,t}^i \cdot y}, & P_{b,t}^i \cdot y \neq -1 \end{cases} \quad (2.5)$$

Para que os sinais sejam invariantes em relação à distância da pessoa no vídeo, usa-se a distância vertical entre o pescoço e o pé que está no chão como denominador em ambos os casos.

Após a obtenção dos sinais, o próximo passo é a construção dos Histogramas de Movimento H^i . Os valores de S^i são divididos em intervalos em $(-1,1]$, e cada ocorrência de um valor de S^i incrementa sua posição correspondente no Histograma (LIMA; SCHWARTZ, 2019). Em seguida os valores de H^i são divididos pelo número de quadros analisados, tornando os Histogramas invariantes em relação à quantidade de quadros de cada vídeo. O número de intervalos utilizados no Histograma deve ser definido de forma a obter a melhor performance do método, assim como o outro parâmetro γ .

Para performar o reconhecimento de marcha utilizando os Histogramas e a distância Euclidiana, obtemos H^{p*} e H^{g*} de duas amostras removendo os pontos que possuem índices em N^p ou N^g , então H^{p*} e H^{g*} são vetorizados e calculamos sua distância normalizando em relação ao número de linhas de H^{g*} , segundo a equação 2.6.

$$D_{p,g} = \frac{\| \text{vec}(H^{p*}) - \text{vec}(H^{g*}) \|}{\sqrt{\text{num_linhas}(H^{g*})}} \quad (2.6)$$

2.5 Biometria Suave e Multibiometria

De acordo com Jain, Dass e Nandakumar (2004), 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. Esses traços podem ser classificados em dois grupos: discretos e contínuos. Os traços como gênero,

etnia, cor dos olhos e tom de pele são por natureza classificados como variáveis discretas, enquanto características como peso, altura e antropometria (medida dos membros) são variáveis contínuas.

Apesar de individualmente não apresentarem um bom desempenho na tarefa de reconhecimento biométrico, as características suaves podem ser utilizadas para aprimorar a performance de sistemas biométricos tradicionais ([JAIN; DASS; NANDAKUMAR, 2004](#)), como reconhecimento facial, digitais ou até mesmo reconhecimento de marcha. Um exemplo da usabilidade dessas características no aprimoramento de sistemas biométricos primários pode ser a seguinte situação: existem 3 indivíduos registrados em um sistema biométrico baseado em digitais, sendo eles A (180 cm de altura, gênero masculino), B (170 cm de altura, gênero feminino) e C (160 cm de altura, gênero masculino). Ao coletar uma amostra X da digital do indivíduo A, e ela ser comparada com as amostras registradas na base de dados, o sistema fique em um empasse, atribuindo pontuações muito próximas entre os indivíduos A e B, impossibilitando que o indivíduo A seja corretamente reconhecido. Essa situação poderia ser contornada ao utilizar um sistema biométrico secundário, capaz de identificar o gênero e a altura do indivíduo da amostra X como gênero masculino e altura 178 cm, nesse caso ao realizar uma combinação apropriada de todas essas informações, o sistema conseguiria definir com um alto grau de confiabilidade que a amostra X corresponde ao indivíduo A.

Nesse contexto, um fator importante que deve ser analisado é a maneira como essas informações podem ser combinadas de forma a levar a um resultado correto. Para isso devem ser analisados diversos fatores envolvendo a maneira como essas informações são representadas e a distribuição dessas características, respostas que podem ser obtidas por meio do estudo da multibiometria.

O termo multibiometria pode ser explicado como a fusão de diferentes tipos de informações em um único sistema de reconhecimento biométrico ([ROSS; NANDAKUMAR; JAIN, 2006](#)). Entre as principais vantagens do uso dessa técnica é possível citar: (i) a melhora significativa da acurácia dependendo das informações e da metodologia da fusão realizada, (ii) o tratamento de questões como a não universalidade do método biométrico, (iii) maior tolerância à dados ruidosos capturados pelos sensores, por possuir mais de um extrator de características e diferentes representações, e (iv) a maior tolerância à falhas, uma vez que podem existir mais de um algoritmo ou sensores trabalhando simultaneamente.

A utilização de múltiplas fontes de informação pode ser realizada em seis cenários diferentes: multi-sensores, multi-algoritmos, multi-unidades (múltiplas instâncias de um identificador biométrico), multi-amostras, multi-modalidades (diferentes identificadores biométricos) ou híbridos, que integram combinações dos cinco cenários anteriores ([ROSS; NANDAKUMAR; JAIN, 2006](#)).

As fusões podem ser realizadas em cinco níveis distintos, antes ou após o casamento (*matching*). Os níveis de fusão antes do casamento são os seguintes:

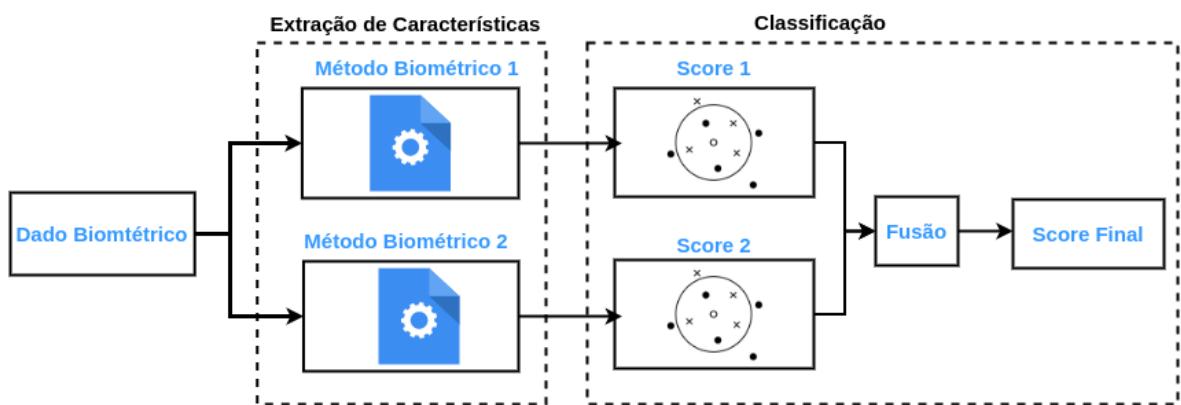
- Nível de sensor: combinação dos dados brutos;
- Nível das características: combinação entre os conjuntos de características extraídos.

Os níveis de fusão após o casamento são os seguintes:

- Nível de pontuação (*score*): combinação entre as pontuações de diferentes algoritmos;
- Nível de classificação (*rank*): as classificações obtidas nas saídas de cada sistema são reordenadas gerando um único *ranking*;
- Nível de decisão: as decisões de diferentes sistemas biométricos são ponderadas para gerar uma decisão única.

Para este trabalho, foi utilizada a fusão em nível de pontuação (*score*). A Figura 16 ilustra o funcionamento de um sistema biométrico que realiza a fusão de dois métodos em nível de *score*. A fusão ocorre na fase de classificação, e após utilizar um classificador para cada método, gerando duas pontuações distintas, uma combinação é feita entre elas para gerar um resultado único.

Figura 16 – Fluxo de execução de um sistema biométrico baseado em multibiometria a nível de pontuação.



Fonte: Elaborado pelo autor.

2.5.1 Método Baseado em Medidas Antropométricas

O trabalho desenvolvido por Tavares et al. (2020) propõe um método baseado em características de biometria suave em conjunto com um estágio de pré-processamento. O método pode ser dividido em cinco passos: subtração do fundo da imagem, pré-processamento, estimativa de pose, extração de características, e por fim, identificação do indivíduo.

No primeiro passo do método, a imagem de entrada tem seu fundo removido por meio de uma técnica de subtração de *background*, eliminando os elementos do fundo que poderiam

interferir no processo de detecção da pose. Em seguida, no segundo passo utiliza-se uma filtragem morfológica na imagem resultante do processo de subtração de fundo. É utilizado o filtro morfológico de abertura para eliminar os ruídos de alta frequência presentes na imagem.

No terceiro estágio do método, aplica-se nas imagens filtradas um estimador de poses 2D, como o OpenPose ([CAO et al., 2018](#)) ou o PifPaf ([KREISS; BERTONI; ALAHI, 2019](#)) com o objetivo de obter a posição de cada membro do indivíduo na cena. A extração das características suaves ocorre no passo quatro do método, onde o vetor de características é gerado. O vetor de características possui dimensão 8, e é formado pelas medidas dos comprimentos dos braços, antebraços, coxas e pernas do indivíduo. Por fim, um classificador 1-NN é utilizado para comparar os vetores de características obtidos, utilizando duas funções de distância: Euclidiana e Manhattan.

Para calcular a distância D entre dois vídeos *gallery* e *probe*, que possuem n e m quadros respectivamente, os m vetores do vídeo de *probe* são comparados com os n vetores do vídeo de *gallery*. A distância D então é definida como a média entre as k menores distâncias calculadas. Segundo os experimentos realizados pelo autor, os melhores resultados foram obtidos com $k = 11$.

2.6 Métricas de Desempenho

A ciência do reconhecimento biométrico se apoia em duas importantes premissas chamadas de singularidade e permanência, no contexto dos traços biométricos utilizados. Um traço biométrico é dito singular se ele é suficiente para diferenciar quaisquer duas pessoas no mundo baseando-se no identificador. Já um traço biométrico permanente significa que ele se conserva inalterado com o tempo durante toda a vida do indivíduo ([JAIN; ROSS; NANDAKUMAR, 2011](#)).

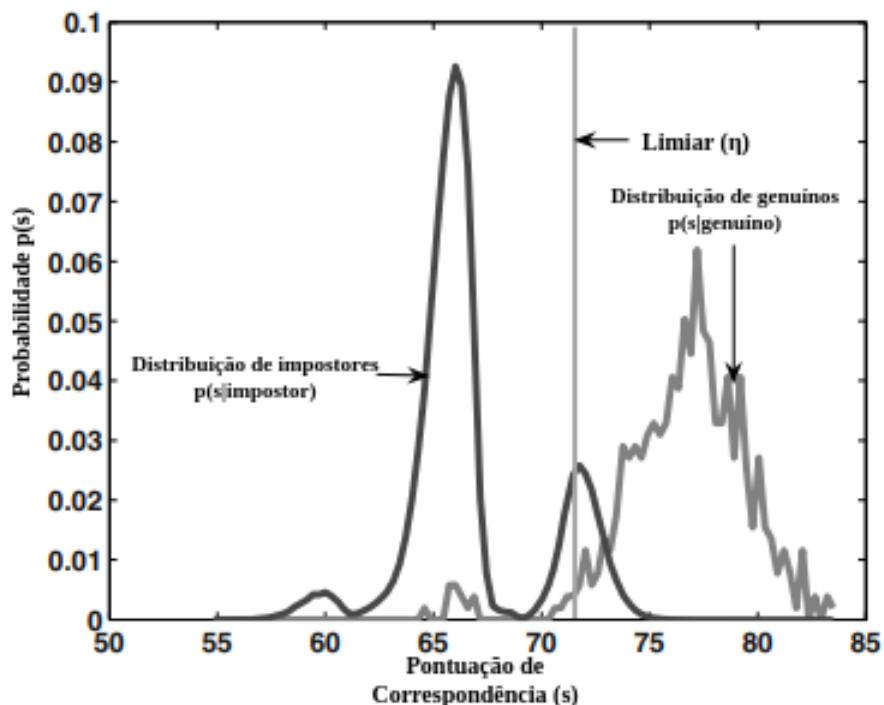
Todavia, segundo [Jain, Ross e Nandakumar \(2011\)](#), essas duas premissas não são garantidas na prática, uma vez que características físicas, apesar de raras, podem não ser únicas, e muitas delas, assim como características comportamentais, sofrem mudanças conforme o indivíduo envelhece. Dessa forma, fica claro que sistemas biométricos não são perfeitos, e por mais robustos que sejam estão sujeitos à falhas, assim, se fazem necessárias medidas de desempenho que avaliem a performance dos sistemas biométricos nas tarefas de verificação e identificação para nos dizer o quanto bom determinado sistema é na execução de seu propósito.

2.6.1 Medidas para Sistemas de Verificação

No contexto dos sistemas de verificação, duas medidas são comumente utilizadas para a análise do desempenho: FAR (*False Accept Rate* - Taxa de Falsa Aceitação) e FRR (*False Reject Rate* - Taxa de Falsa Rejeição). Uma comparação é chamada de genuína se é medida a similaridade entre duas amostras correspondentes a um mesmo indivíduo, e é chamada de

impostora caso a comparação seja entre amostras de duas pessoas diferentes. Os sistemas de verificação realizam sua decisão baseada em um limiar η , caso a similaridade seja maior que η o usuário é aceito, caso contrário ele é rejeitado. Dessa forma, FRR pode ser definida como a fração de *scores* genuínos menores que η , e FAR a fração de *scores* impostores maiores ou iguais ao limiar η (JAIN; ROSS; NANDAKUMAR, 2011). A Figura 17 exemplifica a distribuição dos *scores* genuínos e impostores de um sistema assim como o limiar η . Nesse caso, podemos calcular a FAR como a área sob a curva das comparações impostoras do lado direito do limiar, e FRR como a área sob a curva das comparações genuínas do lado esquerdo do limiar, assim como definido nas Equações 2.7 e 2.8, respectivamente.

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



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

$$FAR(\eta) = p(s \geq \eta | \text{impostor}) = \int_{\eta}^{\infty} p(s | \text{impostor}) ds \quad (2.7)$$

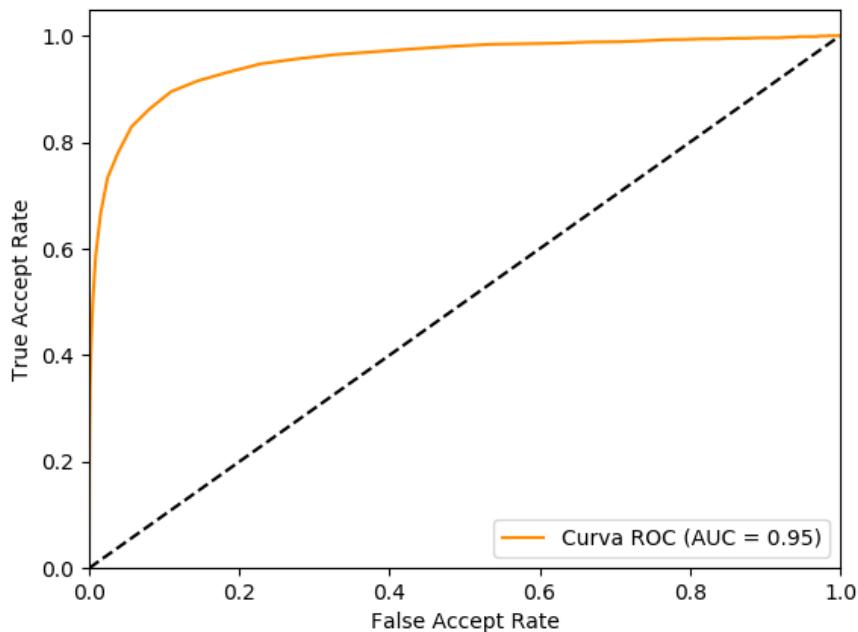
$$FRR(\eta) = p(s < \eta | \text{genuíno}) = \int_{-\infty}^{\eta} p(s | \text{genuíno}) ds \quad (2.8)$$

Em muitos casos ao invés de usar a FRR, utiliza-se uma variação dela chamada de TAR (*True Accept Rate* - Taxa de Aceitação Verdadeira), que representa a porção das comparações genuínas que foram corretamente aceitas, ou seja, seu *score* foi maior que o limiar η . A TAR pode ser obtida de acordo com a Equação 2.9.

$$TAR(\eta) = p(s \geq \eta | \text{genuíno}) = 1 - FRR(\eta) \quad (2.9)$$

Como o limiar de um sistema biométrico pode variar de acordo com o nível de segurança requerido para diferentes aplicações, as taxas TAR e FAR são calculadas para diferentes valores de limiar e representados em uma curva conhecida como ROC (*Receiver Operating Characteristic* - Característica Operacional do Receptor) plotada utilizando TARxFAR, como mostra a Figura 18. Segundo [Jain, Ross e Nandakumar \(2011\)](#), a melhor maneira de comparar dois sistemas biométricos é através de suas curvas ROC. Se a TAR de um sistema A é constantemente maior que a TAR de um outro sistema B para os valores correspondentes de FAR, isso significa que o sistema A possui uma performance superior à do B. Entretanto, se as duas curvas de intersectam em algum ponto, isso significa que em alguns pontos operacionais (valores de FAR) o sistema A é melhor, e em outros B é o melhor.

Figura 18 – Exemplo de uma curva ROC (TARxFAR).



Fonte: Elaborado pelo autor.

Para esses casos também é possível fazer a análise pela AUC (*Area Under the Curve* - Área sob a Curva) da curva ROC, que nesse caso quanto maior (mais próximo de 1) melhor é o desempenho do sistema. A performance de um sistema biométrico também pode ser analisada por meio de valores únicos. Um exemplo é a medida de EER (*Equal Error Rate* - Taxa de Erro Igual), que é referente a um ponto na curva ROC em que FAR e FRR se igualam, nesse caso, um menor valor de EER indica uma performance melhor.

2.6.2 Medidas para Sistemas de Identificação

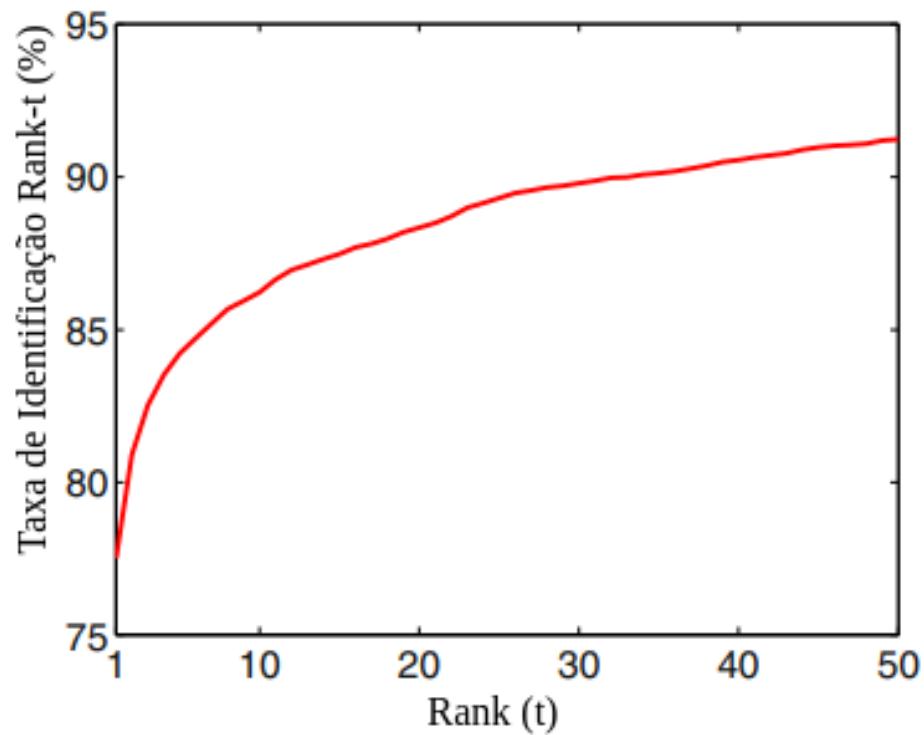
Assim como no caso dos sistemas de verificação, os sistemas de identificação estão sujeitos à dois tipos de erros. Supondo que um sistema biométrico de identificação contendo N pessoas registradas retorne como saída uma conjunto contendo t identidades ($1 \leq t \ll N$), o *rank* de identificação é definido como a classificação correta da identidade do usuário nas t primeiras correspondências retornadas.

Uma identificação falsa positiva ocorre quando o sistema retorna uma identidade para um usuário que não está de fato registrado na base de dados, a taxa de ocorrência desse erro é chamada de FPIR (*False Positive Identification Rate* - Taxa de Identificação Falsa Positiva). A taxa de identificação falsa negativa (FNIR - *False Negative Identification Rate*) se refere ao cenário em que o usuário é registrado no sistema, mas no processo de identificação sua identidade correta não está contida no conjunto das t identidades retornadas.

Outra medida relacionada à FNIR é a TPIR (*True Positive Identification Rate* - Taxa de Identificação Verdadeira Positiva), que é a porção de indivíduos registrados no sistema que quando analisados sua identidade verdadeira é retornada entre o conjunto de identidades da saída. Essa taxa pode ser relacionada com a FNIR da seguinte maneira: $FNIR = 1 - TPIR$. O TPIR também pode ser chamado de taxa de identificação *rank-t*, e o valor particular de TPIR para $t = 1$ é chamado de acurácia *rank-1*, que é uma das métricas mais utilizadas para avaliar o desempenho de sistemas de identificação ([JAIN; ROSS; NANDAKUMAR, 2011](#)).

A taxa de identificação *rank-t* (R_t) também pode ser calculada para diferentes valores de t , formando a curva CMC (*Cumulative Match Characteristic* - Característica de Correspondência Cumulativa), que plota R_t em função de $t = 1, 2, \dots, N$, onde N é o número de usuários registrados no sistema. A Figura 19 mostra um exemplo da curva CMC.

Figura 19 – Exemplo de uma curva CMC.



Fonte: Adaptado de [Jain, Ross e Nandakumar \(2011\)](#).

3 Metodologia

Neste capítulo é apresentada a metodologia adotada para a obtenção dos resultados apresentados nos diferentes experimentos de reconhecimento biométrico baseado em marcha realizados neste trabalho, bem como as bases de dados utilizadas e o protocolo utilizado no decorrer dos experimentos.

3.1 Bases de Dados

Para avaliar os desempenhos dos métodos estudados e implementados foram utilizadas duas bases de dados públicas: CASIA Gait Dataset-A ([WANG et al., 2003](#)), composta por vídeos de indivíduos (homens e mulheres) caminhando em diferentes ângulos e situações em um ambiente externo, e o CASIA Gait Dataset-B ([YU; TAN; TAN, 2006](#)), composta por imagens de pessoas caminhando em 11 ângulos diferentes (variando de 0 até 180 graus) em um ambiente interno.

3.1.1 CASIA Gait Dataset-A

O CASIA Gait Dataset-A, criado em 2001, inclui imagens de 20 indivíduos caminhando em um ambiente externo, cada um com 12 sequências de vídeo, 4 sequências para cada ângulo de visão da câmera: 0, 45 e 90 graus, que representam a visão lateral, oblíqua e frontal do indivíduo caminhando, respectivamente. A Figura 20 mostra exemplos dos ângulos de visão presentes da base de dados.

Figura 20 – Exemplos de vídeos do CASIA Gait Dataset-A.



(a) Visão lateral.

(b) Visão frontal.

(c) Visão oblíqua.

Fonte: [Wang et al. \(2003\)](#).

3.1.2 CASIA Gait Dataset-B

A CASIA Gait Database B é uma grande base de dados de vídeo para análise de marcha, em múltiplos ângulos de visão e em diferentes situações adversas que poderiam

causar interferência na forma como a pessoa normalmente caminha, como no caso em que o indivíduo carrega uma mochila nas costas. A base foi criada em 2005 e conta com imagens de 124 indivíduos caminhando em três variações diferentes, sendo elas: normal, com casaco e carregando uma mochila. Para cada variação existem 11 ângulos de visão e quatro variações de velocidade em cada ângulo, cada vídeo varia entre 4 e 7 segundos de duração com uma taxa de 24 quadros por segundo ([YU; TAN; TAN, 2006](#)). A Figura 21 mostra as variações em relação ao uso de roupas e mochilas.

Figura 21 – Exemplos de variações de condições de vestimenta e transporte do CASIA Gait Dataset-B.



(a) Normal.

(b) Usando um casaco.

(c) Carregando uma mochila.

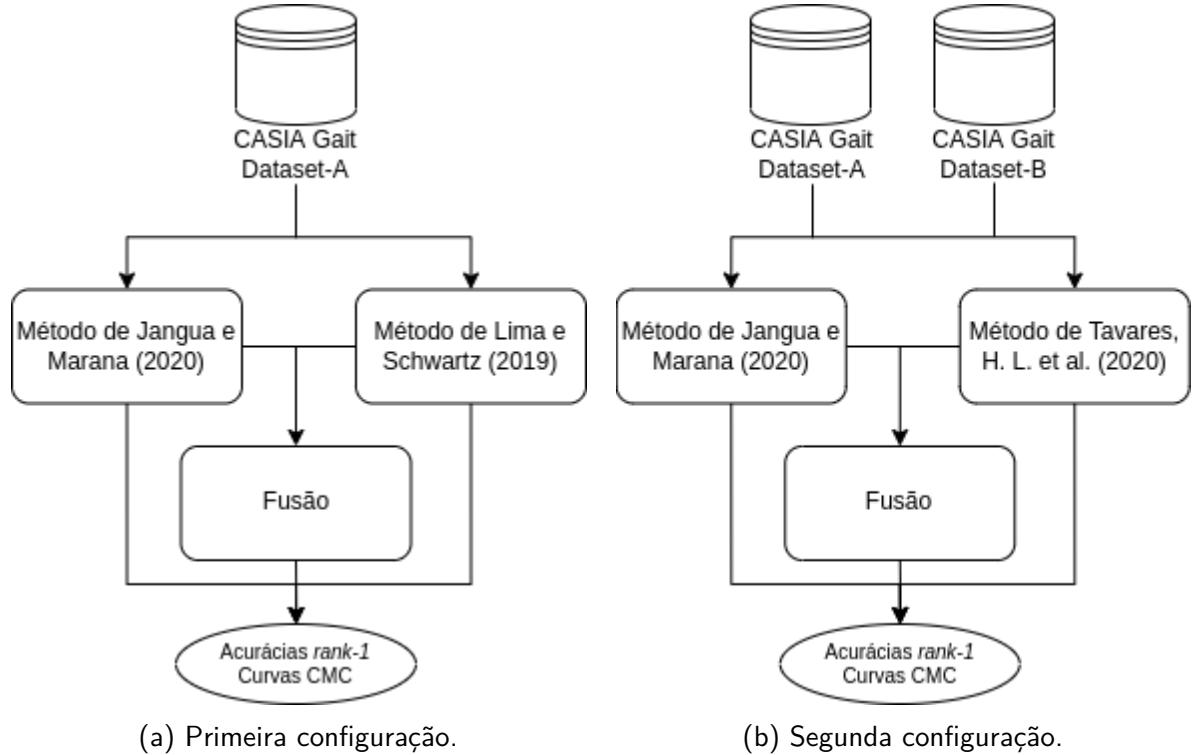
Fonte: [Yu, Tan e Tan \(2006\)](#).

3.2 Experimentos

Os experimentos realizados podem ser divididos em duas configurações diferentes em relação ao seu principal objetivo. Em ambos os casos, foram utilizadas técnicas de multibiometria pra realizar a fusão entre dois métodos diferentes, e confrontar o desempenho individual de cada método com o obtido pela fusão entre eles. Para todos os casos, foram utilizadas como medida de desempenho a acurácia *rank-1* e a curva CMC. A primeira configuração dos experimentos (Figura 22a) avaliou o desempenho dos métodos de [Jangua e Marana \(2020\)](#) e [Lima e Schwartz \(2019\)](#), bem como a fusão entre eles, utilizando um classificador 1-NN e a função de distância Euclidiana. A segunda configuração (Figura 22b) avaliou o desempenho do método de [Jangua e Marana \(2020\)](#) quando combinado com o método baseado em medidas antropométricas de [Tavares et al. \(2020\)](#), utilizando o classificador 1-NN e a função de distância Qui-quadrado.

Para avaliar o desempenho dos métodos de [Lima e Schwartz \(2019\)](#) e o de [Jangua e Marana \(2020\)](#), e da fusão entre eles, foram realizados experimentos utilizando a base de dados CASIA Gait Dataset-A ([WANG et al., 2003](#)). O experimento foi realizado em três etapas: (i) avaliação do desempenho do método de [Lima e Schwartz \(2019\)](#), (ii) avaliação do desempenho do método de [Jangua e Marana \(2020\)](#) e por fim (iii) fusão em nível de *score* entre os dois métodos e avaliação do desempenho da fusão. Em todos os experimentos foi utilizado o classificador 1-NN com função de distância Euclidiana. Para a análise dos resultados

Figura 22 – Esquemas de configuração dos experimentos realizados.



Fonte: Elaborado pelo autor.

foram utilizadas as acurárias rank-1 de cada método, além de suas curvas CMC. Em todos os três estágios desse experimento foram utilizados todos os vídeos da base de dados, calculando a acurácia em cada ângulo de câmera separadamente (frontal, lateral e oblíquo).

Para avaliar o desempenho da utilização do método baseado em características de biometria suave, de [Tavares et al. \(2020\)](#), em conjunto com o método de [Jangua e Marana \(2020\)](#), foram utilizadas as bases de dados CASIA Gait Dataset-A ([WANG et al., 2003](#)) e CASIA Gait Dataset-B ([YU; TAN; TAN, 2006](#)). Assim como o anterior, este experimento foi dividido em três etapas principais: a avaliação individual de cada um dos dois métodos, e a avaliação da fusão entre eles. Nas três etapas, foi utilizado o classificador 1-NN associado à função de distância Qui-quadrado e foram utilizados como indicadores de desempenho as acurárias *rank-1* e as curvas CMC. Para a base da CASIA Gait Dataset-A ([WANG et al., 2003](#)) foram utilizados todos os vídeos, calculando a acurácia em cada ângulo de câmera separadamente, já na CASIA Gait Dataset-B ([YU; TAN; TAN, 2006](#)), foram utilizados todos os vídeos em situação de marcha normal, no ângulo de câmera lateral.

4 Desenvolvimento

Este trabalho foi desenvolvido em duas etapas. Na primeira, foram realizados experimentos envolvendo alguns dos métodos de reconhecimento de marcha e de biometria suave, bem como a fusão entre eles. E na segunda, foi desenvolvido um sistema biométrico completo baseado em reconhecimento de marcha, utilizando o método que obteve os melhores resultados nos experimentos.

4.1 Experimentos

Durante a primeira fase do desenvolvimento, foram realizados experimentos utilizando métodos de reconhecimento de marcha e um método de reconhecimento baseado em medidas antropométricas, utilizando os protocolos e as bases de dados apresentados no Capítulo 3.

Tanto a implementação dos métodos quanto os experimentos foram realizados utilizando a linguagem de programação *Python*, escolhida por ser uma linguagem de fácil acesso, que oferece um conjunto amplo de ferramentas para aprendizagem de máquina, ciência de dados e computação científica, além de contar com uma ampla documentação.

Todos os métodos biométricos estudados, tanto os baseados em reconhecimento de marcha (Capítulo 2.4) quanto o baseado em medidas antropométricas (Capítulo 2.5) utilizam uma abordagem baseada em modelo, que depende de uma fase de pré-processamento para a extração de características, que compreende a aplicação de um método de estimativa de pose, para obter as poses 2D dos indivíduos nas imagens das bases de dados. Como foi abordado no Capítulo 2.3, atualmente existem duas opções de estimadores de pose de código aberto, o *OpenPose* (CAO et al., 2018) e o *PifPaf* (KREISS; BERTONI; ALAHI, 2019). Para o desenvolvedor deste trabalho, foi escolhido o método *OpenPose*, baseando-se em algumas premissas: (i) o método *PifPaf* não possui como principal foco a detecção exata da pose da pessoa, apresentando uma estrutura de esqueleto com menos pontos-chave e consequentemente codificando uma menor quantidade de informações espaço-temporais do corpo do indivíduo (KREISS; BERTONI; ALAHI, 2019), (ii) em questões de custo computacional, ambos os métodos apresentam requisitos semelhantes (CAO et al., 2018) (KREISS; BERTONI; ALAHI, 2019), e (iii) o método *OpenPose* detecta 25 pontos-chave, incluindo pontos relacionados aos pés (CAO et al., 2018), enquanto o *PifPaf* detecta apenas 17 pontos-chave (KREISS; BERTONI; ALAHI, 2019).

Para a implementação dos métodos biométricos, foi utilizada, principalmente, a biblioteca *NumPy* (HARRIS et al., 2020). Essa biblioteca oferece acesso a uma gama de operações que envolvem matrizes e vetores através da linguagem *Python*, utiliza diversas operações de

otimização por meio de técnicas de vetorização e *broadcasting*, além de facilitar a leitura do código, uma vez que utiliza uma sintaxe que se assemelha às expressões matemáticas. Já para a realização dos experimentos, principalmente para o cálculo das funções de distância e visualização das curvas CMC, foram utilizados os pacotes *Scikit-learn* (PEDREGOSA et al., 2011) e *Matplotlib* (HUNTER, 2007). *Scikit-learn* é um pacote que conta com diferentes métodos de aprendizagem de máquina, além de diferentes funções de distância pré-implementadas. Já a biblioteca *Matplotlib*, oferece poderosas ferramentas para visualização de dados em gráficos 2D.

4.2 Resultados e Discussão

O primeiro experimento realizado buscou avaliar o desempenho dos métodos de Lima e Schwartz (2019) e Jangua e Marana (2020), ambos métodos de reconhecimento de marcha baseados em análise de poses 2D extraídas utilizando o estimador de poses OpenPose (CAO et al., 2018). Além disso, foi realizada a fusão entre os dois métodos, a nível de score, calculando a média aritmética simples entre as distâncias calculadas entre as mesmas duas amostras por cada um dos dois métodos. A Tabela 1 mostra as acurárias rank-1 obtidas pelos métodos no Dataset-A em cada uma das três direções. Foi utilizado o classificador 1-NN, com a função de distância Euclidiana. É possível notar que apesar dos dois métodos já apresentarem bons resultados individualmente, a fusão entre os dois métodos proporciona uma performance ainda melhor na tarefa de identificação biométrica, obtendo 100% de acurácia no melhor caso para o ângulo frontal.

Tabela 1 – Acurárias rank-1 obtidas na base de dados CASIA Gait Dataset-A.

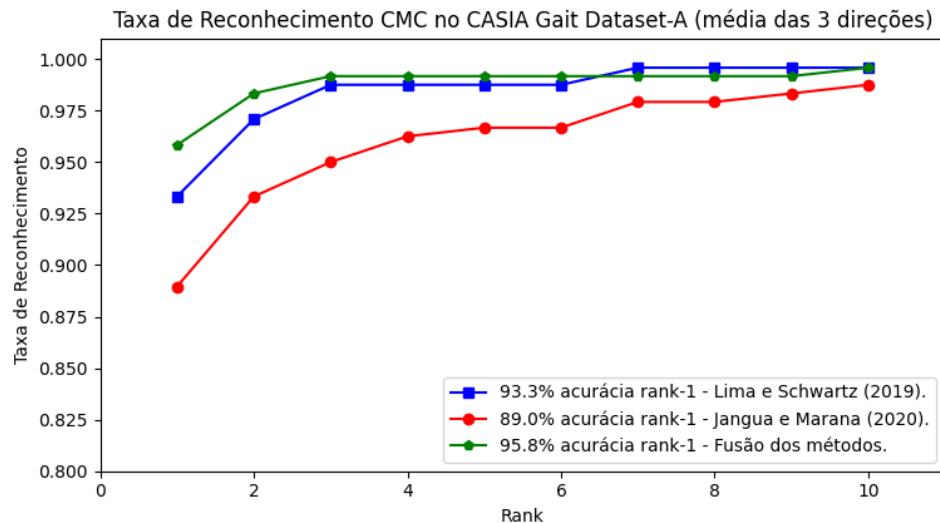
Método	Lateral	Oblíqua	Frontal
Jangua e Marana (2020) - Euclidiana	86.25%	86.25%	96.25%
Lima e Schwartz (2019) - Euclidiana	85.00%	97.50%	97.50%
Fusão - Euclidiana	90.00%	97.50%	100.00%

Fonte: Elaborado pelo autor.

A Figura 23 mostra as curvas CMC obtidas individualmente pelos dois métodos e a curva CMC obtida com a fusão no nível de pontuação. Observando tais curvas, fica ainda mais evidente a melhoria de desempenho proporcionada pela aplicação de técnicas de multibiometria. Mesmo utilizando a taxa de reconhecimento média das três angulações de câmera do CASIA Gait Dataset-A a fusão entre os dois métodos apresenta uma taxa de reconhecimento de 95.8% de acurácia em *rank-1*.

O segundo experimento avaliou os métodos de Tavares et al. (2020) e Jangua e Marana (2020), e verificou o impacto causado nesse desempenho pela fusão desses dois métodos. Diferente do método de Jangua e Marana (2020), o método de Tavares et al. (2020) não é baseado em reconhecimento de marcha, mas utiliza as medidas antropométricas extraídas

Figura 23 – Curvas CMC - CASIA Gait Dataset-A (média das três direções) - Distância Euclidiana.



Fonte: Elaborado pelo autor.

de poses 2D para identificar os indivíduos. A fusão entre os dois métodos foi realizada em nível de *score*, realizando a média entre as distâncias obtidas. As Tabelas 2 e 3 mostram as acurárias *rank-1* obtidas pelos métodos nas bases Dataset-A, em cada uma das três direções, e Dataset-B na direção lateral, utilizando a função de distância Qui-quadrado e o classificador 1-NN. É possível notar que o uso complementar de características de biometria suave, como é o caso das medidas antropométricas, proporciona uma melhora significativa no desempenho na tarefa de identificação biométrica.

Tabela 2 – Acurárias rank-1 obtidas na base de dados CASIA Gait Dataset-A - Qui-quadrado.

Método	Lateral	Oblíqua	Frontal
Jangua e Marana (2020) - Qui-quadrado	87.50%	92.50%	95.00%
Tavares et al. (2020) - Qui-quadrado	85.00%	91.25%	88.75%
Fusão - Qui-quadrado	95.00%	97.50%	100.00%

Fonte: Elaborado pelo autor.

Tabela 3 – Acurárias rank-1 obtidas na base de dados CASIA Gait Dataset-B.

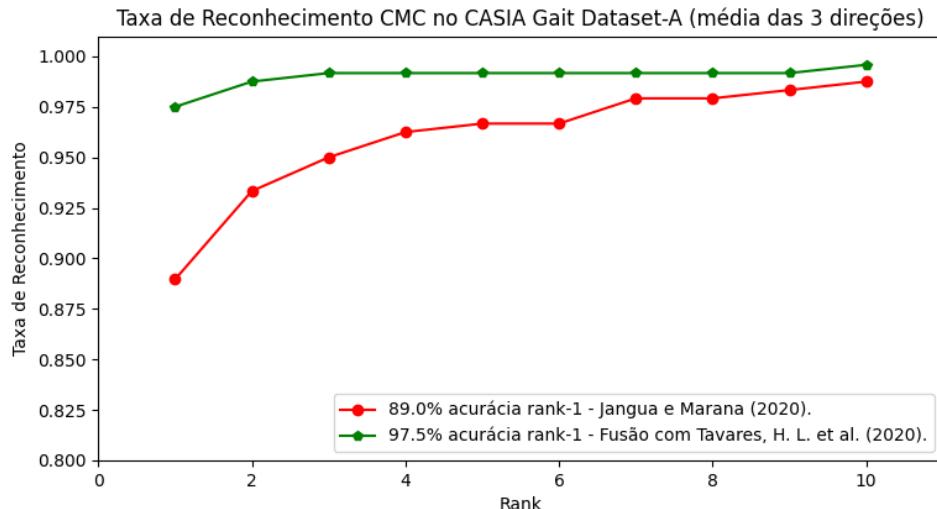
Método	Lateral
Jangua e Marana (2020) - Qui-quadrado	94.22%
Tavares et al. (2020) - Qui-quadrado	86.96%
Fusão - Qui-quadrado	97.18%

Fonte: Elaborado pelo autor.

A Figura 24 mostra a curva CMC obtida individualmente pelo método de Jangua e Marana (2020) e a curva CMC obtida com a fusão no nível de pontuação com o método de

Tavares et al. (2020), ambos utilizando a função de distância Qui-quadrado com o classificador 1-NN na base de dados CASIA Gait Dataset-A.

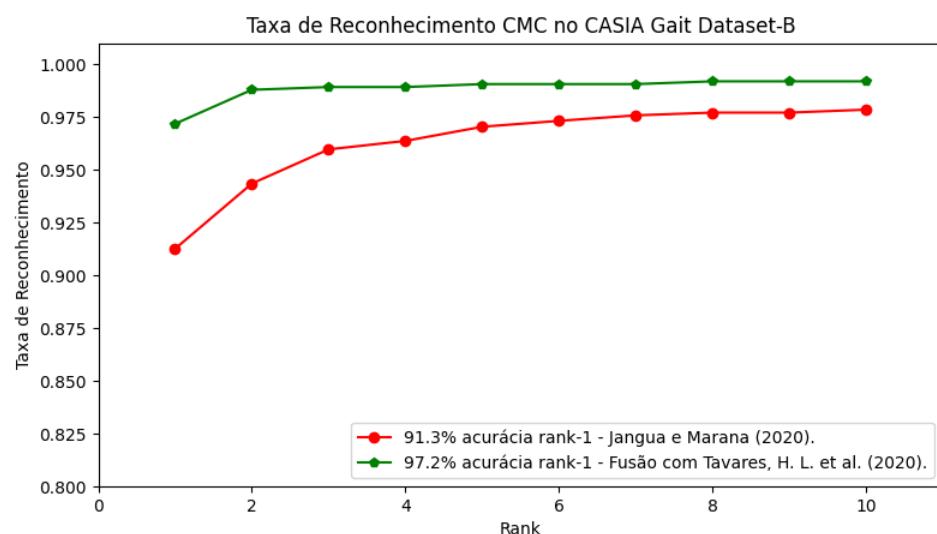
Figura 24 – Curvas CMC - CASIA Gait Datset-A (média das três direções) - Distância Qui-quadrado.



Fonte: Elaborado pelo autor.

Já na Figura 25, são exibidas as curvas CMC do método de Jangua e Marana (2020) e sua fusão com o de Tavares et al. (2020) sob as mesmas condições anteriores, mas utilizando as imagens da base de dados CASIA Gait Dataset-B, na visão lateral apenas.

Figura 25 – Curvas CMC - CASIA Gait Datset-B (apenas lateral).



Fonte: Elaborado pelo autor.

Observando tais curvas, fica ainda mais evidente a melhoria de desempenho proporcionada pela aplicação de técnicas de multibiometria utilizando características suaves. A fusão

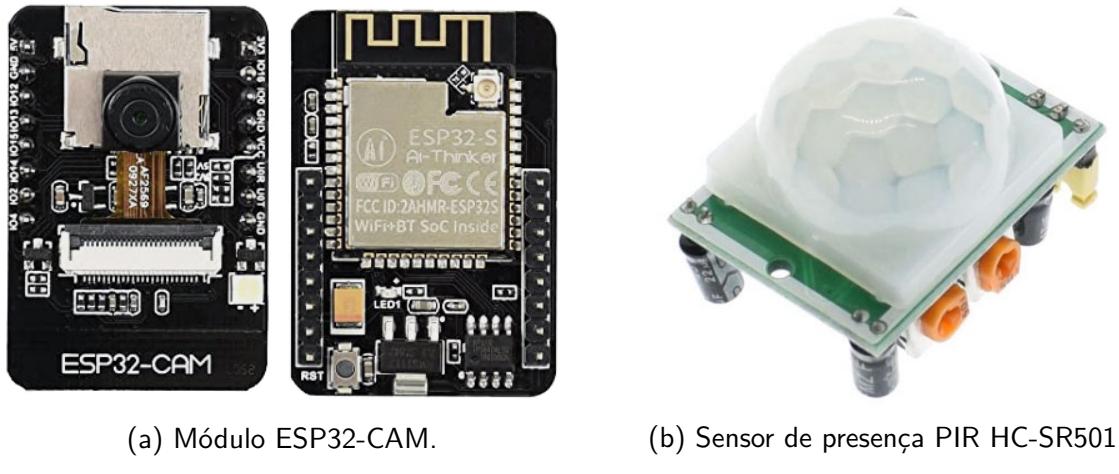
entre os dois métodos apresenta uma taxa de reconhecimento de 97.5% de acurácia em *rank-1*, e 97.2% no Dataset-B.

5 Sistema Biométrico

A primeira etapa do projeto consistiu no estudo teórico acerca dos métodos de reconhecimento de marcha baseados em poses 2D, bem como os conceitos teóricos em que eles se baseiam. Em seguida, alguns dos métodos estudados foram escolhidos para serem implementados e aplicados na tarefa de reconhecimento biométrico. Esses métodos foram avaliados utilizando o estimador de pose OpenPose ([CAO et al., 2018](#)). Para os experimentos foram utilizadas as bases de dados públicas CASIA Gait Dataset-A ([WANG et al., 2003](#)) e CASIA Gait Dataset-B ([YU; TAN; TAN, 2006](#)).

A segunda etapa baseou-se na construção do sistema biométrico propriamente dito, partindo da sua estrutura física. Para isso, foi utilizado um módulo chamado ESP32-CAM, que conta com acessos à *bluetooth*, *Wi-Fi*, armazenamento interno por cartão de memória, microprocessador e uma câmera integrada OV2640 com resolução de 2 MP. A Figura 26a ilustra o módulo ESP32-CAM.

Figura 26 – Principais componentes do módulo de sensor do sistema biométrico.



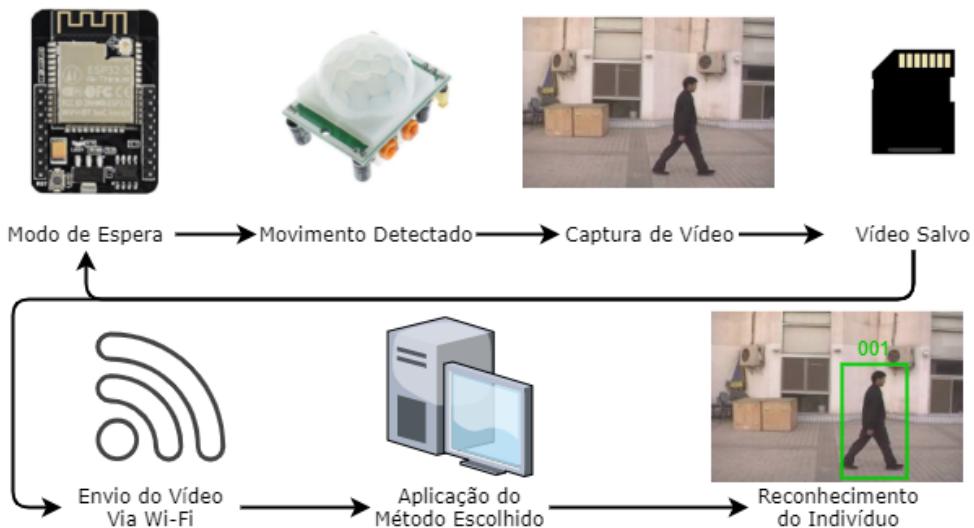
Fonte: [Pieters \(2021\)](#) e [Reis \(2021\)](#).

Em conjunto com o módulo ESP32-CAM foi utilizado um sensor de presença PIR HC-SR501. O sensor PIR servirá para ativar a captura de vídeo assim que alguém entrar no ambiente protegido pelo sistema. A Figura 26b mostra o sensor HC-SR501.

Uma vez montada a estrutura física do sistema, o ESP32-CAM foi utilizado para capturar sequências de vídeo de pessoas previamente identificadas caminhando, que compuseram a base de dados do sistema biométrico. Após montar a base de dados, foi realizado o desenvolvimento de um sistema para computadores em linguagem *Python*. Esse sistema recebe os arquivos de vídeo capturados em tempo real pelo ESP32-CAM via FTP (*File Transfer Protocol* - Protocolo de Transferência de Arquivos) através da rede *Wi-Fi* local. Em seguida, aplica um dos métodos

selecionados, que foi implementado separadamente em um servidor HTTP (*Hypertext Transfer Protocol* - Protocolo de Transferência de Hipertexto). Por fim, compara com os indivíduos registrados na base de dados e realiza o reconhecimento do indivíduo no vídeo capturado. A Figura 27 ilustra a estrutura básica do sistema biométrico. A seguir, são apresentados os módulos desenvolvidos que integram o sistema biométrico construído, seguindo a mesma estrutura apresentada no Capítulo 2.1.

Figura 27 – Resumo da arquitetura do sistema biométrico.



Fonte: Elaborado pelo autor.

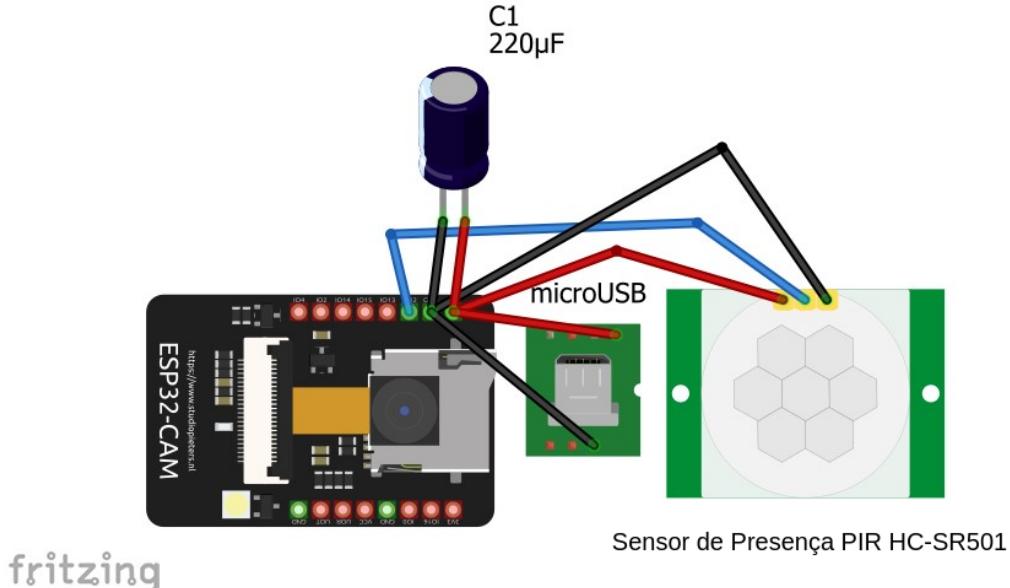
5.1 Sensor

O módulo de sensor do sistema biométrico construído é composto por uma estrutura de *hardware* muito simples e barata, comparada com os sensores comumente utilizados para coleta biométrica, como câmeras de alta resolução e leitores de digitais. O módulo é composto principalmente pela placa de desenvolvimento ESP32-CAM, que conta com uma câmera OV2640 integrada e um sensor de presença infra-vermelho HC-SR501. A Figura 28 mostra como foi montado o circuito do módulo sensor.

A entrada USB (*Universal Serial Bus* - Porta Serial Universal) é responsável pela alimentação em 5V do circuito, ligado em paralelo com um capacitor (C1) de $220\mu F$, para evitar possíveis ruídos e oscilações. A saída de sinal do sensor de presença é ligado ao pino 12 de entrada e saída do ESP32-CAM, que foi programado para iniciar uma gravação de 15 segundos sempre que o estado do pino muda de estado de baixo (0V) para alto (5V). As pinagens da placa ESP32-CAM e do sensor HC-SR501 são apresentadas nas Figuras 29 e 30, respectivamente.

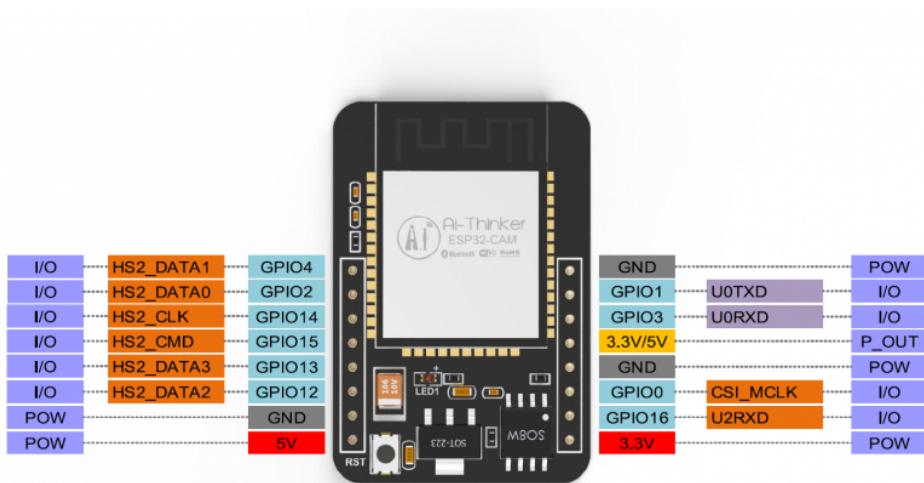
O ESP32-CAM é uma placa de desenvolvimento de baixo custo e de tamanho reduzido. Esta placa é comumente utilizada em projetos e soluções de Internet das Coisas (IoT - *Internet*

Figura 28 – Circuito do módulo de sensor do sistema biométrico.



Fonte: Adaptado de [Zah \(2022\)](#).

Figura 29 – Pinagem do módulo ESP32-CAM.

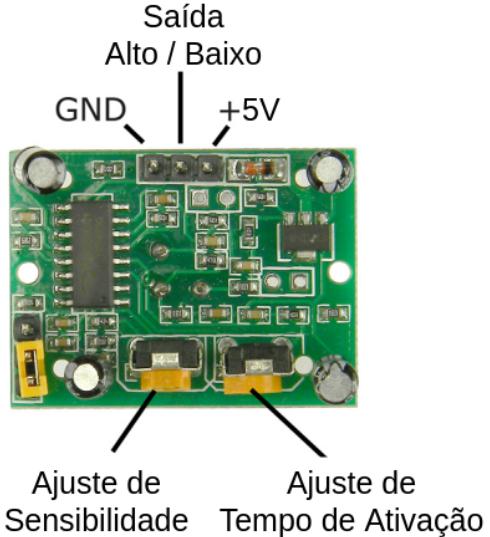


Fonte: Adaptado de [Ai-Thinker \(2022\)](#).

of Things) ([AI-THINKER, 2022](#)). O módulo possui integração com WiFi e Bluetooth, além de contar com duas Unidades Centrais de Processamento (CPU - *Central Processing Unit*) 32-bit LX6 de alta performance. De acordo com [Ai-Thinker \(2022\)](#), o ESP32-CAM conta ainda com memória interna de 520KB SRAM, suporte para câmera OV2640 e suporte para cartões de memória.

O HC-SR501 é baseado em tecnologia infra-vermelho, e é um módulo automático de alta sensibilidade, alta confiabilidade e baixa voltagem, capaz de detectar movimento em um raio de até 7 metros e em um ângulo de até 120 graus em relação ao sensor. É amplamente utilizado em soluções de IoT voltada para casas inteligentes como lâmpadas e ventiladores

Figura 30 – Pinagem do módulo HC-SR501.



Fonte: Adaptado de [ETC \(2022\)](#).

automáticos, por exemplo. Além da pinagem de ligação básica, é possível realizar o ajuste da sensibilidade (de 3 até 7 metros) e do tempo de ativação do sensor (3 segundos até 300 segundos) ([ETC, 2022](#)). Toda a programação necessária para o módulo ESP32-CAM foi realizada utilizando um repositório chamado [ESP32-CAM-Video-Recorder¹](#), no qual foram realizados apenas alguns ajustes como taxa de quadros do vídeo, qualidade da imagem e tempo de gravação. Além do funcionamento básico para o sensor, o repositório também contém a implementação do servidor FTP utilizado para obter os vídeos capturados remotamente.

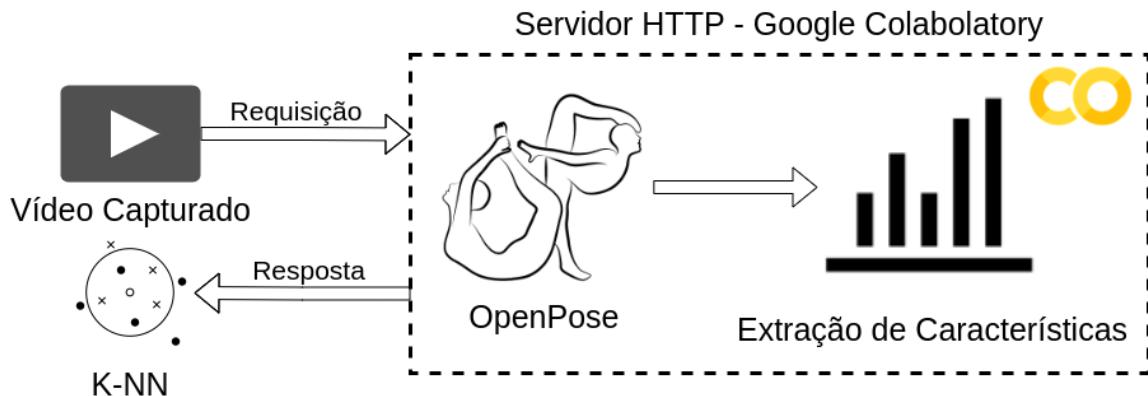
5.2 Extração de Características

A fase de extração de características do sistema biométrico se baseia na aplicação de um dos métodos estudados nos Capítulos [2.4](#) e [2.5](#). De acordo com os resultados obtidos pelos experimentos no Capítulo [4.2](#), foi escolhida uma das variações dos métodos que obteve os resultados mais robustos na tarefa de reconhecimento biométrico. Analisando os resultados, a abordagem selecionada foi a fusão entre os métodos de [Jangua e Marana \(2020\)](#) e [Tavares et al. \(2020\)](#), cujos resultados são apresentados nas Tabelas [2](#) e [3](#), e Figuras [24](#) e [25](#).

Para ambos os métodos que compõem a fusão, é necessário realizar a extração das poses 2D em cada vídeo capturado, na fase de pré-processamento. A execução do método de estimativa de pose somado ao processo de extração das características de cada um dos dois métodos demanda um maior poder de processamento, e para uma execução em tempo real necessita de uma placa de aceleração gráfica. Como não possuímos o *hardware* necessário para essa execução em tempo real, optamos por realizar este processamento em nuvem, em uma

¹ Disponível em <<https://github.com/jameszah/ESP32-CAM-Video-Recorder>>. Acesso em 1 de Fevereiro de 2022.

Figura 31 – Esquema de funcionamento do servidor HTTP para extração de características.



Fonte: Elaborado pelo autor.

instância de *notebook*² do *Google Colaboratory*, implementando o modelo como um serviço por meio de um servidor HTTP. A Figura 31 ilustra o funcionamento do módulo de extração de características.

O *Google Colaboratory* (ou simplesmente *Colab*) é um projeto que visa disseminar o ensino e pesquisa em aprendizado de máquina. O *Colab* é baseado em *notebooks Jupyter*, uma ferramenta de código aberto que integra linguagens interpretadas, bibliotecas e ferramentas para visualização. Um *notebook* pode ser executado localmente ou em nuvem, e é composto por múltiplas células executadas individualmente que possuem um código que gera saídas independentes, normalmente incluindo texto, tabelas, gráficos e imagens. Essa tecnologia facilita a replicação e compartilhamento de experimentos científicos (CARNEIRO et al., 2018).

Além da estrutura de execução baseada em *notebooks*, o *Colab* proporciona ambientes de execução pré-configurados para Python com todos os pacotes essenciais de inteligência artificial e aprendizagem de máquina instalados. São oferecidos ambientes de execução padrão, baseados em CPU e ambientes que contam com aceleração de GPU (*Graphical Processing Unit* - Unidade de Processamento Gráfico) configurados para utilizar ferramentas de paralelização, essenciais na melhoria de performance para algoritmos de aprendizado de máquina (CARNEIRO et al., 2018). 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 de tempo, variando de 6 até 12 horas.

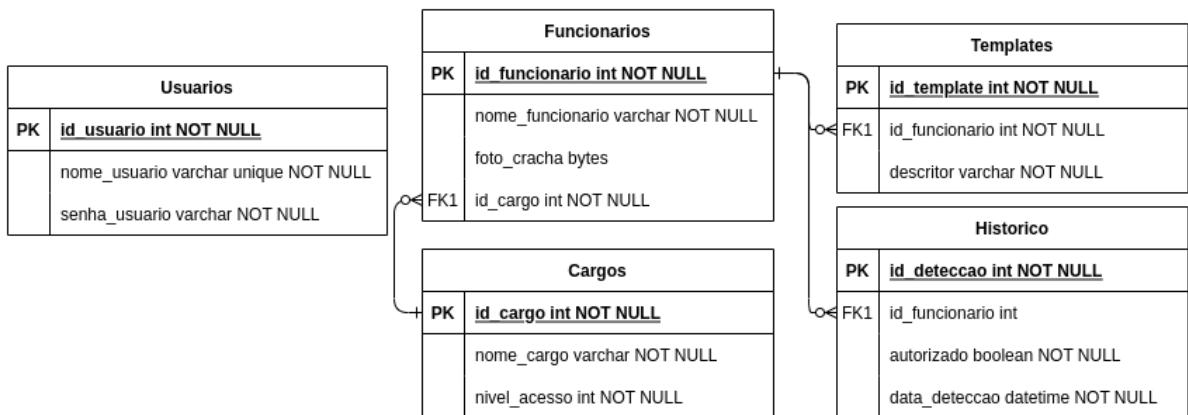
5.3 Base de Dados

O módulo da base de dados do sistema se resumiu em um banco de dados estrutural baseado em linguagem SQL (*Structured Query Language* - Linguagem de Consulta Estruturada),

² Disponível em <<https://colab.research.google.com/drive/1hnnBH3erEO1PkM4gvypsJyegoNAFjpW?usp=sharing>>. Acesso em 3 de Fevereiro de 2022.

utilizando o Sistema Gerenciador de Banco de Dados (SGBD) *SQLite*. A estrutura do banco simulou um mini-mundo de uma empresa que busca realizar o controle de acesso em áreas restritas, monitoradas por uma câmera. Neste cenário fictício, cada funcionário possui um cargo na empresa, que possui um nível de acesso associado à ele. Cada sala teria um nível de acesso mínimo para ser acessada. Dessa forma, ao identificar um funcionário com nível de acesso inferior ao permitido para determinado ambiente, um alerta de segurança deve ser disparado para uma central de monitoramento, e devem ser salvos os *logs* contendo o funcionário, a data e hora da detecção, bem como se o acesso foi negado ou autorizado. Além disso, é importante armazenar os *templates* adicionados na fase de registro para cada funcionário, para que sejam recuperados no momento da autenticação. A Figura 32 ilustra como foi feita a modelagem do banco utilizado pelo sistema biométrico. Além dos dados dos funcionários, cargos, seus respectivos *templates* e o histórico de detecções, também foi criada uma tabela para guardar as credenciais dos usuários do sistema.

Figura 32 – Diagrama do banco de dados utilizado pelo sistema biométrico.



Fonte: Elaborado pelo autor.

O *SQLite* é um sistema gerenciador de banco de dados SQL, que diferente da grande maioria dos SGBDs, não necessita de um processo servidor rodando separadamente. Ao invés disso, o *SQLite* lê e escreve arquivos comuns no disco de maneira direta, além de ser multiplataforma. Além disso, seu código fonte é de domínio público, e possui aplicações em diversas linguagem como C, C++, Python e Java. Essas características fazem do *SQLite* uma escolha popular para soluções simples de bancos de dados, sem perder a confiabilidade e segurança dos SGBDs mais comuns ([BHOSALE; PATIL; PATIL, 2015](#)).

5.4 Registro e Autenticação

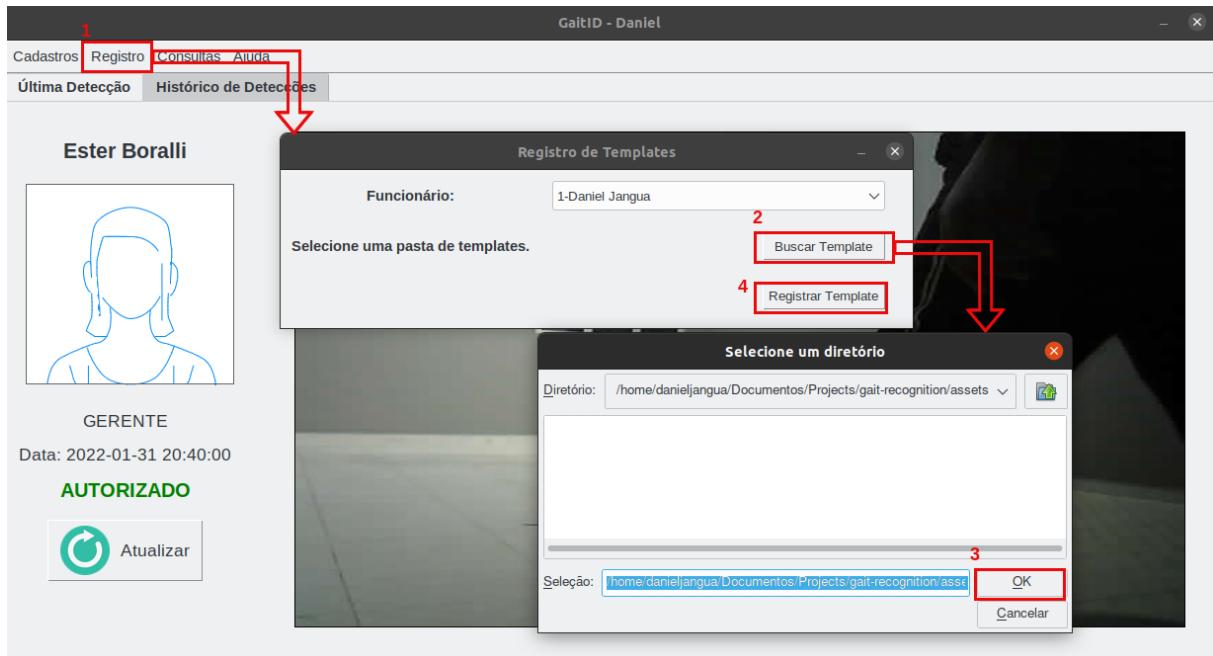
Todo sistema biométrico segue um fluxo de execução que pode ser dividido em dois momentos: o registro e a autenticação. Assim como foi abordado no Capítulo 2.1, a fase de registro se baseia no cadastro da biometria dos indivíduos na base de dados, os dados brutos são capturados pelo sensor, as características são extraídas e então são salvas na base de dados

na forma de um *template*. Já a fase de autenticação acontece em tempo real, onde os dados são coletados sem a informação prévia acerca da identidade do indivíduo, e baseando-se nos *templates* registrados na base de dados, o sistema deve determinar a identidade do sujeito.

No sistema biométrico construído, tanto a fase de autenticação quanto a de registro, assim como outras funcionalidades essenciais para a usabilidade do sistema, são realizadas por meio de uma interface gráfica com o usuário. A interface foi desenvolvida em linguagem *Python*, utilizando a biblioteca *Tkinter* ([PYTHON, 2022](#)), um pacote padrão do *Python* para interface com o usuário, que oferece suporte para diversas plataformas como *Windows*, *Linux* e *MacOS*.

A fase de registro do sistema pode ser resumida em três passos: (i) obtenção do vídeo do indivíduo caminhando, (ii) extração das características e (iii) armazenamento do *template* na base de dados. Todo esse processo é feito através da interface gráfica do sistema. Na tela inicial, deve-se navegar até a opção de "Registro", e então será aberta uma tela onde é preciso informar o funcionário que será associado ao *template*, bem como o caminho do vídeo capturado. Em seguida, ao clicar em "Registrar", o vídeo é enviado ao servidor HTTP via requisição, onde será pré-processado e as características serão extraídas. O *template* relacionado ao vídeo será enviado de volta no corpo da resposta HTTP. Por fim, o sistema salva automaticamente os descritores na tabela de *templates* da base de dados. A Figura 33 ilustra o fluxo que deve ser seguido para realizar o registro de um *template* para determinado funcionário.

Figura 33 – Fluxo da fase de registro através da interface gráfica.



Fonte: Elaborado pelo autor.

Já a fase de autenticação, é um processo em tempo real, que executa constantemente junto ao programa principal, em uma linha de execução concorrente (*thread*). Essa *thread* lista todos os arquivos do servidor FTP hospedado pelo ESP32-CAM aguardando por novas

capturas de vídeo. Ao receber uma nova captura, a *thread* realiza o *download* do vídeo, o envia para o servidor HTTP através de uma requisição e obtém os descritores na resposta do servidor. Em seguida, utiliza o classificador 1-NN para determinar a identidade do indivíduo caminhando no vídeo, com base nos *templates* previamente registrados no banco de dados. Por fim, é feita uma verificação entre o nível de acesso do indivíduo identificado e o nível de acesso permitido no ambiente que foi acessado, é definido se o acesso foi autorizado ou negado, e caso tenha sido negado, um alerta é enviado ao sistema principal. Sempre que uma autenticação é realizada, o sistema principal atualiza os seus registros, e exibe na tela principal os dados da última detecção, como o nome do funcionário, cargo, foto de crachá, o vídeo capturado, bem como o estado do acesso (autorizado ou negado). A Figura 34 ilustra o painel de monitoramento do sistema para a fase de autenticação.

Figura 34 – Painel de monitoramento em tempo real da fase de autenticação.



Fonte: Elaborado pelo autor.

Além das funcionalidades biométricas básicas, o sistema permite a realização do cadastro de novos usuários, de cargos, e de funcionários, bem como a consulta dos dados dos funcionários e dos cargos já cadastrados. Além disso, no painel principal do sistema, é possível acessar a aba de "Histórico de Detecções" onde são listadas todas as últimas detecções realizadas.

5.5 Resultados e Discussão

Para realizar a validação do sistema biométrico, foi necessário povoar a base de dados do sistema. Para isso, foram coletadas imagens de pessoas caminhando, reunidas em um pequeno conjunto de dados, seguindo o mesmo padrão da base de dados CASIA Gait Dataset-A ([WANG et al., 2003](#)), capturando sequências de marcha apenas com o ângulo de visão frontal.

Foram capturados imagens de 5 pessoas diferentes, com 6 vídeos cada, utilizando a câmera integrada ao módulo ESP32-CAM.

Figura 35 – Exemplos de imagens capturadas para a base de dados do sistema.



Fonte: Elaborado pelo autor.

Após a finalização da captura das imagens, entre os 6 vídeos capturados de cada funcionário, foram extraídos e registrados *templates* de 4 vídeos na base de dados do sistema. Os outros dois vídeos foram utilizados para validação. Para todos os casos, tanto com o indivíduo caminhando de frente, quanto de costas, o método de reconhecimento de marcha aplicado foi acurado o suficiente para reconhecer corretamente todos os funcionários cadastrados. Dessa forma, foi possível verificar que o método de Jangua e Marana (2020) associado ao de Tavares et al. (2020) apresenta uma robustez suficiente para funcionar em ambientes menos controlados, como foi o caso das imagens capturadas. Apesar disso, o ambiente ainda não simula uma situação totalmente real, uma vez que segue alguns protocolos, como o ângulo fixo de caminhada e a restrição de existir apenas uma pessoa por vez na cena.

As principais vantagens do método baseado em marcha, em contraste aos métodos como reconhecimento facial, por exemplo, está no baixo custo de implementação e alta escalabilidade, uma vez que, para este caso, foi utilizado um *hardware* de baixo custo, com uma câmera de baixa resolução e vídeos em baixa taxa de quadros por segundo, o que causaria problemas para performar métodos de reconhecimento facial, necessitando de sensores melhores e mais caros. No caso de uma empresa com muitas salas, não seria viável a escalabilidade de um sistema de reconhecimento facial com câmeras de alta resolução em cada sala devido ao alto custo de implementação. Todo o código do sistema criado, desde a interface gráfica em *Python*, a programação do módulo ESP32-CAM, a estrutura de criação do banco de dados em *SQLite*, o *notebook* do servidor HTTP, até a base de dados coletada, foram armazenados em um repositório público no *Github*³.

³ Disponível em <<https://github.com/Daniel-Jangua/gait-recognition>>. Acesso em 9 de Fevereiro de 2022.

6 Conclusão

Este trabalho foi desenvolvido em duas partes principais. A primeira parte se baseou em um estudo aprofundado dos temas que permeiam o tema Biometria, focando no estudo de métodos para reconhecimento de marcha. Além disso, foram estudados temas como biometria suave e multibiometria, bem como o efeito do uso dessas técnicas em conjunto com os métodos de reconhecimento de marcha. A partir dos estudos realizados, foram escolhidos alguns métodos biométricos baseados em reconhecimento de marcha e biometria suave, e foram realizados experimentos acerca da implementação da fusão entre esses métodos.

A primeira solução apresentada baseada em multibiometria foi a fusão de dois métodos de reconhecimento de marcha baseados em poses 2D. Os métodos escolhidos foram os de [Jangua e Marana \(2020\)](#) e o de [Lima e Schwartz \(2019\)](#), que geraram resultados que corroboraram a ideia de [Ross, Nandakumar e Jain \(2006\)](#) acerca da possibilidade de atribuir maior robustez à sistemas biométricos por meio da fusão entre diferentes sistemas. O resultado desses experimentos, mostraram que o uso dos dois métodos unidos apresenta melhor desempenho na tarefa de identificação biométrica do que ambos os métodos individualmente, evidenciando que eles se complementam, e suas vantagens individuais se sobreponem às suas deficiências.

A segunda solução também utilizou os conceitos de multibiometria, mas ao invés de utilizar dois métodos com o mesmo propósito, o foco se voltou para o uso de características de biometria suave, como as medidas antropométricas, em conjunto com métodos de reconhecimento de marcha, a fim de verificar as possíveis melhorias obtidas por essa técnica. Os experimentos foram executados utilizando o método de [Jangua e Marana \(2020\)](#) como base, e o de [Tavares et al. \(2020\)](#), baseado em biometria suave. Mais uma vez os resultados apresentaram melhorias expressivas na acurácia do método original, mostrando que as técnicas de biometria suave podem ser fortes aliadas na busca por métodos de reconhecimento mais robustos ([JAIN; DASS; NANDAKUMAR, 2004](#)).

A última etapa do desenvolvimento deste trabalho, se baseou na construção de um sistema biométrico completo, utilizando a fusão de métodos que obteve os resultados mais robustos nos experimentos, que no caso foi a união entre os métodos de [Jangua e Marana \(2020\)](#) e [Tavares et al. \(2020\)](#). Para a validação do sistema foram capturadas imagens de pessoas caminhando para compor a base de dados do sistema. Entre todas as pessoas registradas, o método foi capaz de reconhecer todas corretamente, apresentando acurácia e robustez mesmo em cenários menos controlados e com sensores de baixo custo, o que corrobora as ideias de [Arantes e Gonzaga \(2011\)](#), acerca das vantagens do uso de reconhecimento de marcha em detrimento dos métodos clássicos de identificação biométrica.

Referências

- AI-THINKER. *ESP32-CAM Camera Development Board*. 2022. Disponível em: <<https://docs.ai-thinker.com/en/esp32-cam>>. Acesso em: 1 de Fevereiro de 2022.
- ARANTES, M. *Método de reconhecimento da marcha humana por meio da fusão das características do movimento global*. Tese (Doutorado) — Universidade de São Paulo, 2010.
- ARANTES, M.; GONZAGA, A. Human gait recognition using extraction and fusion of global motion features. *Multimedia Tools and Applications*, p. 655–675, 2011. Disponível em: <<https://doi.org/10.1007/s11042-010-0587-y>>.
- BHOSALE, S.; PATIL, M. T.; PATIL, M. P. Sqlite: Light database system. *International Journal of Computer Science and Mobile Computing*, v. 4, n. April, p. 882–885, 2015.
- CAO, Z.; HIDALGO, G.; SIMON, T.; WEI, S.-E.; SHEIKH, Y. Openpose: realtime multi-person 2d pose estimation using part affinity fields. *arXiv preprint arXiv:1812.08008*, 2018.
- CARNEIRO, T.; NOBREGA, R. V. M. D.; NEPOMUCENO, T.; BIAN, G.-B.; ALBUQUERQUE, V. H. C. D.; FILHO, P. P. R. Performance analysis of google colaboratory as a tool for accelerating deep learning applications. *IEEE Access*, v. 6, p. 61677–61685, 2018.
- CHEN, C.; LIANG, J.; ZHAO, H.; HU, H.; TIAN, J. Frame difference energy image for gait recognition with incomplete silhouettes. *Pattern Recognition Letters*, 2009.
- CHOUDHURY, B.; THEN, P.; ISSAC, B.; RAMAN, V.; HALDAR, M. K. A survey on biometrics and cancelable biometrics systems. *International Journal of Image and Graphics*, 2018.
- COVER, T.; HART, P. Nearest neighbor pattern classification. *IEEE transactions on information theory*, v. 13, n. 1, p. 21–27, 1967.
- ETC. *HC-SR501 PIR MOTION DETECTOR*. 2022. Disponível em: <<https://datasheetspdf.com/pdf-file/775434/ETC/HC-SR501/1>>. Acesso em: 1 de Fevereiro de 2022.
- HARRIS, C. R.; MILLMAN, K. J.; WALT, S. J. van der; GOMMERS, R.; VIRTANEN, P.; COURNAPEAU, D.; WIESER, E.; TAYLOR, J.; BERG, S.; SMITH, N. J.; KERN, R.; PICUS, M.; HOYER, S.; KERKWIJK, M. H. van; BRETT, M.; HALDANE, A.; RÍO, J. F. del; WIEBE, M.; PETERSON, P.; GÉRARD-MARCHANT, P.; SHEPPARD, K.; REDDY, T.; WECKESSER, W.; ABBASI, H.; GOHLKE, C.; OLIPHANT, T. E. Array programming with NumPy. *Nature*, Springer Science and Business Media LLC, v. 585, n. 7825, p. 357–362, set. 2020. Disponível em: <<https://doi.org/10.1038/s41586-020-2649-2>>.
- HUNTER, J. D. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, IEEE COMPUTER SOC, v. 9, n. 3, p. 90–95, 2007.
- JAIN, A. K.; DASS, S. C.; NANDAKUMAR, K. Can soft biometric traits assist user recognition? In: JAIN, A. K.; RATHA, N. K. (Ed.). *Biometric Technology for Human Identification*. SPIE, 2004. v. 5404, p. 561 – 572. Disponível em: <<https://doi.org/10.11117/12.542890>>.

- JAIN, A. K.; ROSS, A. A.; NANDAKUMAR, K. *Introduction to biometrics*. [S.I.]: Springer Science & Business Media, 2011.
- JANGUA, D. R. S.; MARANA, A. N. A new method for gait recognition using 2d poses. In: XVI Workshop de Visão Computacional (WVC 2020). [S.I.: s.n.], 2020.
- KREISS, S.; BERTONI, L.; ALAHI, A. Pipaf: Composite fields for human pose estimation. *arXiv preprint arXiv:1903.06593*, 2019.
- 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. *British Machine Vision Conference - BMVC*, n. 82, p. 82.1–82.12, 2016.
- MITCHEL, T. M. *Machine Learning*. [S.I.]: McGraw-Hill Science/Engineering/Math, 1997.
- MYSID; CYP. *Euclidean Voronoi Diagram*. 2022. Disponível em: <https://commons.wikimedia.org/wiki/File:Euclidean_Voronoi_diagram.svg>. Acesso em: 14 de Janeiro de 2022.
- NIXON, M. S.; CARTER, J. N. Automatic recognition by gait. *Proceedings of the IEEE*, IEEE, v. 94, n. 11, p. 2013–2024, 2006.
- PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V. et al. Scikit-learn: Machine learning in python. *Journal of machine learning research*, v. 12, n. Oct, p. 2825–2830, 2011.
- PIETERS, A. *ESP32 CAM – Troubleshooting*. 2021. Disponível em: <<https://www.studiopieters.nl/esp32-cam-troubleshooting/>>. Acesso em: 4 de Maio de 2021.
- PYTHON, S. F. *Tkinter - Python interface to Tcl/Tk*. 2022. Disponível em: <<https://docs.python.org/3/library/tkinter.html>>. Acesso em: 8 de Fevereiro de 2022.
- REIS, F. dos. *Como funciona um Sensor de Movimento PIR – Passive Infra-red*. 2021. Disponível em: <<http://www.bosontreinamentos.com.br/elettronica/como-funciona-um-sensor-de-movimento-pir-passive-infrared/>>. Acesso em: 4 de Maio de 2021.
- ROSS, A.; NANDAKUMAR, K.; JAIN, A. K. *Handbook of Multibiometrics*. USA: Springer, 2006.
- TAVARES, H. L.; NETO, J. B. C.; PAPA, J. P.; COLOMBO, D.; MARANA, A. N. People identification based on soft biometrics features obtained from 2d poses. *9th Brazilian Conference on Intelligent Systems (BRACIS-2020)*, 2020.
- 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 trans Pattern Analysis and Machine Intelligence(PAMI)*, v. 25, n. 12, p. 1505–1518, 2003.
- YU, S.; TAN, D.; HUANG, K.; TAN, T. Reducing the effect of noise on human contour in gait recognition. *Internat. Conf. on Biometrics*, 2007.

YU, S.; TAN, D.; TAN, T. A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. In: *18th International Conference on Pattern Recognition (ICPR'06)*. [S.l.: s.n.], 2006. v. 4, p. 441–444. ISSN 1051-4651.

ZAH, J. *ESP32-CAM-Video-Recorder*. GitHub, 2022. Repositório do GitHub. Disponível em: <<https://github.com/jameszah/ESP32-CAM-Video-Recorder>>. Acesso em: 1 de Fevereiro de 2022.