



UNIVERSIDADE FEDERAL DO AMAZONAS - UFAM

FACULDADE DE TECNOLOGIA

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Reconhecimento de gestos de membros superiores utilizando
sensores de movimento e fotopletismografia

Marcos Negreiros Rylo

Manaus - AM

Abril 2023

Marcos Negreiros Rylo

Reconhecimento de gestos de membros superiores utilizando
sensores de movimento e fotopletismografia

Dissertação apresentado ao curso de Mestrado em Engenharia Elétrica, área de concentração Controle e Automação de Sistemas, do Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas.

Orientador: Vicente Ferreira de Lucena Júnior, Dr. –Ing

Coorientador: Renan Landau Paiva de Medeiros, Dr.

UNIVERSIDADE FEDERAL DO AMAZONAS - UFAM

FACULDADE DE TECNOLOGIA

Manaus - AM

Abril 2023

Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

R994r Rylo, Marcos Negreiros
Reconhecimento de gestos de membros superiores utilizando
sensores de movimento e fotopletismografia / Marcos Negreiros
Rylo . 2023
87 f.: il. color; 31 cm.

Orientador: Vicente Ferreira de Lucena Júnior
Coorientador: Renan Landau Paiva de Medeiros
Dissertação (Mestrado em Engenharia Elétrica) - Universidade
Federal do Amazonas.

1. Fotopletismografia. 2. dispositivos vestíveis. 3. reconhecimento
de gestos. 4. aprendizado de máquina. I. Lucena Júnior, Vicente
Ferreira de. II. Universidade Federal do Amazonas III. Título

MARCOS NEGREIROS RYLO

RECONHECIMENTO DE GESTOS DE MEMBROS SUPERIORES UTILIZANDO SENSORES DE MOVIMENTO E FOTOPLETISMOGRAFIA.

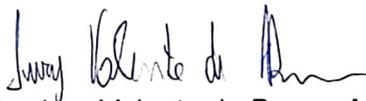
Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica na área de concentração Controle e Automação de Sistemas.

Aprovada em 28 de abril de 2023.

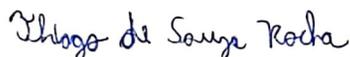
BANCA EXAMINADORA



Prof. Dr. Vicente Ferreira de Lucena Junior, Presidente
Universidade Federal do Amazonas



Prof. Dr. Iury Valente de Bessa, Membro
Universidade Federal do Amazonas



Prof. Dr. Thiago de Souza Rocha, Membro
Instituto de Ciência e Tecnologia-Sidia

Aos meus pais, por nunca terem medido esforços para me proporcionar um ensino de qualidade durante toda a minha vida.

Agradecimentos

Aos meus colegas de trabalho, em quem convivi durante os últimos anos, pelo companheirismo e pela troca de experiências assim como conselhos que me permitiram crescer não só como pessoa, mas também como profissional.

Ao professor Vicente Ferreira de Lucena Júnior, por ter sido meu orientador e ter desempenhado tal função com dedicação e pela paciência com a qual guiou o meu aprendizado.

*“A mente que se abre a uma nova ideia
jamais voltará ao seu tamanho original”*

Albert Einstein

Reconhecimento de gestos de membros superiores utilizando sensores de movimento e fotopleletismografia

Autor: Marcos Negreiros Rylo

Orientador: Vicente Ferreira de Lucena Júnior, Dr. –Ing

Resumo

Este trabalho visa avaliar as técnicas de aprendizagem de máquinas usando fotopleletismografia de baixa frequência associada a sensores de movimento de dispositivos vestíveis, tais como relógios inteligentes, no reconhecimento de gestos do pulso e dos dedos. Após a segmentação dos gestos baseado na identificação de artefatos de movimento no sinal de fotopleletismografia (PPG), modelos de classificação utilizando Máquinas de Vetores de Suporte, Florestas Aleatórias e XgBoost foram treinados utilizando atributos estatísticas extraídas de sinais PPG e sensores de movimento. O projeto aponta que frequências de 25 Hz são adequadas para o processo de reconhecimento, alcançando resultados de até 82% precisão e 82% de revocação.

Palavras-chave: Fotopleletismografia, dispositivos vestíveis, reconhecimento de gestos, aprendizado de máquina.

Reconhecimento de gestos de membros superiores utilizando sensores de movimento e fotopletismografia

Autor: Marcos Negreiros Rylo

Orientador: Vicente Ferreira de Lucena Júnior, Dr. –Ing

Abstract

This project aims to evaluate machine learning techniques using low frequency photoplethysmography coupled with motion sensors from wearable devices, such as smart watches, in recognizing wrist and finger gestures. After gesture segmentation based on the identification of motion artifacts in the photoplethysmography (PPG) signal, classification models using Support Vector Machines, Random Forests and XgBoost were trained using statistical attributes extracted from PPG signals and motion sensors. The project indicates that frequencies of 25 Hz are suitable for the recognition process, achieving results of up to 82% accuracy and 82% recall.

Keywords: photoplethysmography; wearable, gesture recognition, machine learning

Lista de ilustrações

Figura 1 – Técnicas para reconhecimento de gestos	20
Figura 2 – Tipos de Fotopletismografia	22
Figura 3 – Formato de Onda - Fotopletismografia	23
Figura 4 – Acelerometro do tipo piezoelétrico	26
Figura 5 – Representação de Hiperplano - SVM	29
Figura 6 – Representação de Hiperplano - SVM - Não linear	31
Figura 7 – Representação de uma árvore de decisão	32
Figura 8 – Fluxograma de um algoritmo de Árvore de Decisão	34
Figura 9 – Tipos de wavelet	38
Figura 10 – Tipos de wavelet	38
Figura 11 – Tipos de <i>wavelet</i>	39
Figura 12 – Decomposição <i>Wavelet</i> aplicada a Sinal PPG	40
Figura 13 – Classificação de métodos de seleção de atributos	42
Figura 14 – Proposta de gestos	52
Figura 15 – Fluxograma do algoritmo para reconhecimento de gestos	54
Figura 16 – Fluxograma do pré-processamento	54
Figura 17 – Sinal de PPG - Mão em Repouso	55
Figura 18 – Sinal de PPG - Mão em Repouso - Após Wavelet	56
Figura 19 – Determinação de Características para Segmentação	56
Figura 20 – Sinal de PPG - Mão em repouso - Gradiente	57
Figura 21 – Sinal de PPG - Mão em repouso - Identificação de Picos	57
Figura 22 – Segmentação de Gestos	58
Figura 23 – Sinal de PPG - Movimento	58
Figura 24 – Sinal de PPG - Movimento - Segmentado	59
Figura 25 – Comparação Classificadores	65
Figura 26 – Curvas de aprendizado	66
Figura 27 – Escalabilidade	66
Figura 28 – Curva de Aprendizado e Escalabilidade XGBoost	67

Figura 29 – Redução de Dimensionalidade - 7 Gestos - Florestas Aleatórias	69
Figura 30 – Matriz de Confusão - 7 Gestos - Florestas Aleatórias	70
Figura 31 – Matriz de Confusão - 6 Gestos - Florestas Aleatórias	71
Figura 32 – Matriz de Confusão - 7 Gestos - SVM	72
Figura 33 – Matriz de Confusão - 6 Gestos - SVM	73
Figura 34 – Matriz de Confusão - 7 Gestos - XGBoost	74
Figura 35 – Matriz de Confusão - 6 Gestos - XGBoost	75
Figura 36 – Matriz de Confusão - 5 Gestos - XGBoost	75
Figura 37 – Projeção para utilização de dispositivos inteligentes para <i>SmartHome</i>	76

Lista de tabelas

Tabela 1 – Validação Cruzada (K-fold onde $k = 5$)	35
Tabela 2 – Técnicas e Modelos utilizados no reconhecimento de gestos	49
Tabela 3 – Lista de Atributos	53
Tabela 4 – Parâmetros da Floresta Aleatória	60
Tabela 5 – Parâmetros da Máquina de Vetores de Suporte	60
Tabela 6 – Parâmetros XGBoost	61
Tabela 7 – Parâmetros selecionados para os classificadores RF e SVM	62
Tabela 8 – Resultado - Segmentação	63
Tabela 9 – Comparação Classificadores	64
Tabela 10 – Resultado - RF - 7 Gestos	67
Tabela 11 – Resultado - RF - 6 Gestos	68
Tabela 12 – Resultado - RF - 5 Gestos	68
Tabela 13 – Resultado - SVM - 7 Gestos	69
Tabela 14 – Resultado - SVM - 6 Gestos	70
Tabela 15 – Resultado - SVM - 5 Gestos	71
Tabela 16 – Resultado - XGBoost - 7 Gestos	72
Tabela 17 – Resultado - XGBoost - 6 Gestos	73
Tabela 18 – Resultado - XGBoost - 5 Gestos	74

Sumário

1	INTRODUÇÃO	14
1.1	Problema	16
1.2	Motivação	16
1.3	Objetivos	17
1.3.1	Objetivo Geral	17
1.3.2	Objetivos Específicos	18
1.4	Organização do documento	18
2	FUNDAMENTAÇÃO TEÓRICA	19
2.1	Reconhecimento de Gestos	19
2.2	Sensores	21
2.2.1	Fotopletoisomografia	21
2.2.1.1	Sinal	23
2.2.1.2	Configuração do Sensor	24
2.2.2	Sensores de movimento	25
2.3	Reconhecimento de Padrões	27
2.4	Máquinas de Vetores de Suporte	28
2.5	Métodos Baseados em Árvores	31
2.5.1	Métodos Baseados em Árvores	31
2.5.2	Florestas Aleatórias	33
2.5.3	<i>Extreme Gradient Boosting</i> (XGBoost)	34
2.6	Validação Cruzada	35
2.7	Mineração de Dados	36
2.7.1	Transformação de Dados	36
2.7.2	Suavização	36
2.7.2.1	Decomposição Wavelet	37
2.7.2.2	Transformada Discreta de Wavelet	39
2.7.3	Agregação	40

2.7.4	Normalização	41
2.8	Seleção de Atributos	41
3	TRABALHOS RELACIONADOS	44
4	DESENVOLVIMENTO	51
4.1	Materiais	51
4.1.1	Base de Dados	51
4.1.2	Definição do Protocolo de Gestos	51
4.2	Implementação	53
4.2.1	Transformação dos Dados	54
4.2.1.1	Suavização e Normalização dos Dados	54
4.2.1.2	Segmentação dos gestos	56
4.2.1.3	Extração de Atributos	59
4.2.2	Treinamento dos Modelos	59
4.2.2.1	Florestas Aleatórias	60
4.2.2.2	Máquina de vetores de suporte	60
4.2.2.3	<i>Extreme Gradient Boosting</i> (XGBoost)	61
4.2.2.4	Teste do Modelo	61
5	RESULTADOS	63
5.1	Segmentação dos Gestos	63
5.2	Classificação dos gestos	64
5.2.1	Florestas Aleatórias	67
5.2.2	Máquinas de Vetores de Suporte	68
5.2.3	<i>Extreme Gradient Boosting</i> (XGBoost)	70
5.3	Discussão	71
6	CONSIDERAÇÕES FINAIS	76
	Referências	82
	APÊNDICE A – PRIMEIRO APÊNDICE	86

A.1	Pseudocódigo - Limiar de FC e Distancia entre picos	86
A.2	Pseudocódigo - Segmentação	86

1 Introdução

Atualmente, os equipamentos vestíveis, do inglês *wearables* são uma categoria de dispositivos que devido aos avanços tecnológicos na área de sensores e conectividade apresentam diversas aplicabilidades e funcionalidades as quais permitem uma interação cada vez ampla do usuário com o ambiente.

Dentre os diversos dispositivos inteligentes disponíveis no mercado, podemos apresentar os *smartwatches*, ou relógios inteligentes, os quais apresentam diversas funções relacionadas a saúde e bem estar do usuário como a contagem de passos, a detecção de quedas, a identificação de distância percorrida, monitoramento de frequência cardíaca, entre outras. Estas informações são obtidas por meio dos dados coletados principalmente dos sensores, tais como, acelerômetros e/ou giroscópios, e dos sensores ópticos de fotopletismografia, do inglês *Photoplethysmogram* (PPG).

A fotopletismografia é uma técnica não invasiva baseada na detecção de alterações do fluxo sanguíneo por meio de absorção e reflexão de luzes de um ou mais diodos emissores de luz (LED). Estes sensores permitem a estimativa de informações como a da pressão sanguínea, de frequência cardíaca, de níveis de estresse e de oxigenação entre outros. Diversos trabalhos têm sido realizados na literatura para estimar as referidas informações tais como (HERAVI et al., 2014) e (LIU; PO; FU, 2017) que apresentam estudos da utilização do sinal de PPG para estimar valores de pressão sanguínea, já (SEVIL et al., 2005) aborda a utilização do PPG para estimar níveis de estresse, enquanto (BAGHA; SHAW, 2011) apresenta uma metodologia para obtenção de níveis de oxigenação e frequência cardíaca.

A incorporação do PPG representa para os *smartwatches* um salto na gama de aplicações as quais podem ser desenvolvidas, uma vez que o sinal de PPG apresenta informações a respeito de diversas características (ALLEN, 2007).

Um dos grandes problemas enfrentados ao utilizar o sinal de PPG em *smartwatches* é a suscetibilidade a contaminação por conta dos movimentos do usuário durante a realização de atividades. Essa contaminação se dá por conta da alteração do fluxo sanguíneo por diferentes

fatores, mas ocorre principalmente pelo movimento dos tendões e músculos que provocam deslocamentos no fluxo sanguíneo durante o movimento. Esse tipo de ruído no sinal de fotopleletismografia pode ser denominado artefatos de movimento, ou simplesmente MA. Entretanto há áreas de estudos que buscam exatamente a existência desses artefatos de movimento para realizar a identificação de padrões.

No campo de ciência de dados, o reconhecimento de padrões tem por objetivo a classificação de uma amostra em um determinado grupo de classes. Uma amostra pode ser uma imagem, sinais em forma de onda. O fluxo de dados em um processo de reconhecimento de padrões consiste basicamente em um sensor que obtém as observações as quais devem ser classificadas, um mecanismo de extração de características, que podem ser quantitativas ou qualitativas, e o algoritmo classificador.

Dentro do reconhecimento de padrões, um dos tópicos muito discutido na literatura é o reconhecimento de gestos (MITRA; ACHARYA, 2007). Este tema está intimamente ligado à visão computacional. Os gestos são movimentos do corpo, mãos, braços e cabeça, voluntários ou involuntários, que revelam estado psicológico ou intenção de exprimir ou realizar algo. Muitas abordagens têm sido estudadas utilizando sistemas de captura de imagens e algoritmos de visão computacional para interpretar a linguagem de sinais.

Os sistemas de rastreamento de gestos em dispositivos móveis são possíveis pela presença de sensores de movimento, como unidades de medida inercial (IMUs). Nesses dispositivos, a detecção de gestos depende da execução de gestos baseados em movimentos pelos usuários, que podem ser reconhecidos por esses sensores de movimento. Isso pode tornar a captura de sinal de gestos sutis ou de baixo movimento um desafio, pois pode ser difícil distingui-los de movimentos naturais ou ruído.

Para tratar o desafio mencionado, considerando a capacidade de extração de informações do PPG, anteriormente descrita, este trabalho propõe uma análise quanto à viabilidade da utilização de sensores de fotopleletismografia em conjunto com os sensores de movimento, como acelerômetros e giroscópios, para a geração de sinais que possam ser utilizados para o treinamento de um modelo de classificação. O intuito é desenvolver um modelo que possa reconhecer gestos finos para fins de interação com outros dispositivos, como sistemas de multimídia.

Para tal, serão estudados os desempenhos de diferentes algoritmos de classificação, sendo eles, *Random Forests*, Máquinas de Vetores de Suporte e *Extreme Gradient Boosting* (XgBoost).

1.1 Problema

Com base na literatura, podemos identificar que a utilização dos *smartwatches* como interface de manipulação do ambiente por parte do usuário ainda não foi amplamente explorada (MITRA; ACHARYA, 2007) É possível identificar técnicas que utilizam elementos/sensores adicionais aos relógios para realizar a coleta de sinais para executar a identificação e o reconhecimento de gestos.

Existem estudos exploratórios que avaliam a viabilidade de utilização de métodos isolados de fotopletismografia e de sensores de movimento, entretanto não foram encontrados durante a pesquisa estudos que combinem as duas técnicas de forma complementar.

Portanto, este trabalho se propõe a responder a seguinte questão: "A utilização combinada dos sensores de fotopletismografia com os sensores de movimento presentes em dispositivos inteligentes tais como os relógios inteligentes, permite a identificação de gestos finos do usuário?"

1.2 Motivação

Os *smartwatches* são dispositivos vestíveis que possibilitam uma maior interação do usuário com o ambiente, diversas funcionalidades já existentes permitem o monitoramento contínuo de sinais fisiológicos do usuário assim como o reconhecimento de atividades físicas, isso ocorre devido a existência de sensores específicos tais como os sensores que detectam movimento.

Entretanto não há aplicações consolidadas que permitam o reconhecimento de gestos executados pelo usuário para realizar comandos de forma remota para interagir com outros dispositivos. Algumas soluções trazidas na literatura fazem uso de dispositivos adicionais ao *smartwatch*, como em (SAPONAS et al., 2010) que apresenta um estudo utilizando eletromio-

grafia.

Os estudos até então utilizando os sensores de movimento, tendo como exemplo o *Serendipity* (WEN; ROJAS; DEY, 2016) e o *FinDroidHR* (ZHANG et al., 2018) apresentam resultados onde a utilização de sensores de movimento se torna difícil quando o usuário está em movimento, uma vez que componentes advindos da translação do usuário podem ocasionar ruídos indesejáveis ao sinal.

Considerando que o sinal advindo da Fotopletismografia é suscetível a variações advindas da movimentação do usuário, essa técnica poderia ser utilizada de forma complementar para a correta identificação dos gestos do usuário.

Por fim, destacamos a necessidade de estudo de métodos utilizando somente os elementos sensores disponíveis dentro dos dispositivos inteligentes para o reconhecimento de gestos e posteriormente a interação com outros dispositivos do ambiente.

1.3 Objetivos

Nesta seção, serão abordados os objetivos gerais e específicos pretendidos por esta pesquisa.

1.3.1 Objetivo Geral

Propor um método que a partir da utilização dos sensores de movimento e ópticos embarcados em dispositivos inteligentes vestíveis seja capaz de realizar a segmentação do sinal para a identificação de movimentos e a sua posterior classificação, isto sem alterar as características padrões dos dispositivos.

1.3.2 Objetivos Específicos

- Estudar utilização combinada dos sensores de movimento e dos sensores de fotopletismografia para o reconhecimento de gestos finos.
- Avaliar a capacidade da utilização de sensores para monitoramento de frequência cardíaca na segmentação do sinal de fotopletismografia para reconhecimento de gestos.
- Apresentar um estudo comparativo de algoritmos de aprendizado de máquina, Máquinas de Vetores de Suporte e *Random Forests* e XgBoost para a classificação de gestos utilizando sensores ópticos de fotopletismografia e sensores de movimento, como acelerômetros e giroscópio.

1.4 Organização do documento

Este trabalho está organizado em seções, estas são, Introdução, Fundamentação Teórica, Trabalhos relacionados, Desenvolvimento, Resultados e Considerações finais. Na introdução é apresentado ao leitor uma contextualização do problema proposto nesta dissertação tal como o objetivo geral e específico. A fundamentação teórica apresenta conceitos uma breve revisão de conceitos necessários ao leitor para que seja possível um melhor entendimento quanto as técnicas e métodos utilizados neste trabalho.

Na seção de trabalhos relacionados, apresentamos uma revisão da literatura com os autores mais relevantes no tema de reconhecimento de gestos encontrados durante o processo de revisão bibliográfica. No desenvolvimento apresentamos a implementação de nosso método para a realização da segmentação e reconhecimento de gestos utilizando as tecnologias presentes nos *smartwatches*. Por fim nas seções de Resultados e Considerações Finais apresentamos respectivamente os resultados objetivos por meio da implementação da nossa proposta de trabalho e apresentamos uma discussão comparando os resultados obtidos com os resultados das pesquisas apresentados na seção de trabalhos relacionados.

2 Fundamentação Teórica

Neste capítulo será apresentada uma revisão de alguns tópicos essenciais para o desenvolvimento desta dissertação, esta revisão inclui uma pesquisa em livros e artigos relevantes ao assunto para entendimento das técnicas de coleta, pré-processamento e pós-processamento para a execução desta pesquisa.

2.1 Reconhecimento de Gestos

O reconhecimento de gestos trata da capacidade de extrair informações pertinentes a movimentos realizados pelo usuário para fins de interação com outros dispositivos.

Atualmente, a maior parte das técnicas é baseada em visão computacional na qual a utilização de técnicas de processamento de imagens e análise de séries temporais permite a identificação e classificação de gestos e movimentos. Um bom exemplo é a identificação e classificação de gestos da linguagem de sinais, possibilitando a sua tradução automática. Com o desenvolvimento de outras tecnologias hoje podemos citar também as aplicações que envolvem realidade virtual (RAUTARAY; AGRAWAL, 2011) , robótica (TRIGUEIROS; RIBEIRO; REIS, 2014) e telemedicina (HONDA; WEBER; LUETH, 2007) .

Mitra e Achyarya, (MITRA; ACHARYA, 2007) dividem os os tipos possíveis de gestos em três categorias principais:

- Gestos Corporais.
- Expressões Faciais.
- Gestos dos membros superiores.

Os movimentos corporais estão relacionados principalmente com os movimentos de translação de todo o corpo. Diversas tecnologias já foram desenvolvidas para realizar a captura e

rastreamento dos movimentos, sendo as técnicas baseadas em sistemas que utilizam detecção de imagens e marcadores conectados aos usuários as mais comuns.

As expressões faciais incluem movimentos como o balançar da cabeça, tendo como exemplo os movimentos de afirmação e negação, movimento dos olhos e sobrancelhas, as expressões faciais também estão relacionadas às demonstrações de surpresa, felicidade, raiva e outros. Por fim, os gestos com membros superiores dão ênfase a movimentação de mãos e braços, para os quais podemos citar como exemplo de aplicação a linguagem de sinais.

O reconhecimento de gestos, conforme exposto anteriormente permite a utilização de diversas técnicas e métodos. A Figura 1 apresenta uma breve classificação das abordagens para aquisição dados em um fluxo padrão para o reconhecimento de gestos. Quanto a modelagem do problema, podem ser usadas abordagens estatísticas, visão computacional, processamento de imagem, entre outras. Temos como alguns exemplos a utilização de modelagem estatística, como Análise de Componentes Principais (PCA) ou *Hidden Markov Models* (HMM) (YAMATO; OHYA; ISHII, 1992).

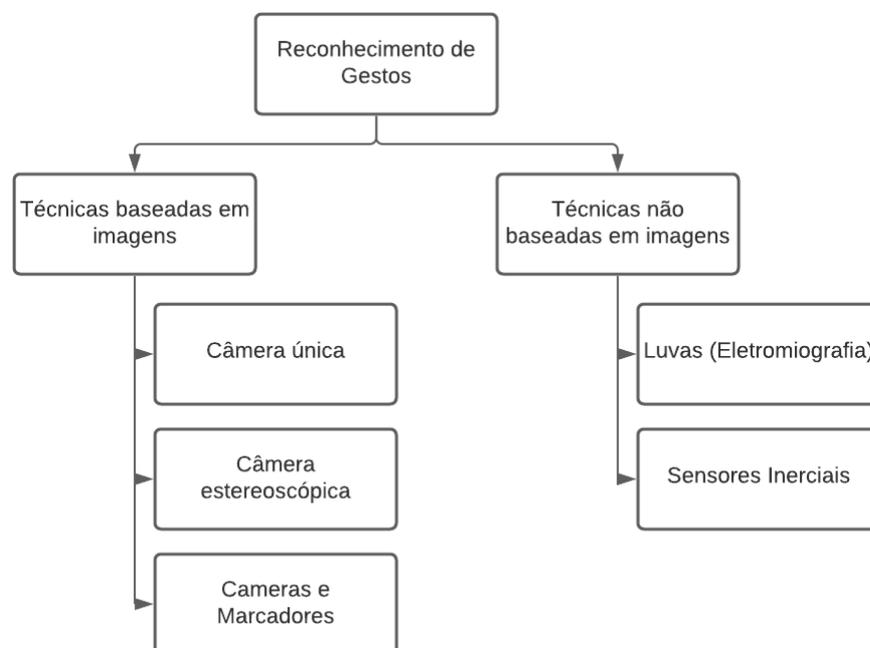


Figura 1 – Técnicas para reconhecimento de gestos

Técnicas utilizando visão computacional e técnicas de reconhecimento de padrões foram amplamente exploradas em Bowden et al. (BOWDEN et al., 2003) e Davis e Shah (DAVIS; SHAH, 1994).

2.2 Sensores

Sensores são dispositivos que respondem a estímulos, sejam físicos ou químicos, de modo que produzem um sinal que pode ser transformado em outra grandeza física para fins de medição ou monitoramento de uma variável em questão.

Diversos tipos de sensores estão presentes em nosso cotidiano, seja de forma aparente, como um sensor de temperatura instalado em um ambiente para demonstração ao usuário quanto as condições de temperatura do local, ou sensores de movimento instalados em dispositivos que permitem o acendimento de lâmpadas ou outros processos, (MöLLER, 2016).

Na área médica, por exemplo, um paciente pode ser equipado com sensores que medem constantemente funções biológicas específicas, como temperatura, pressão arterial, frequência cardíaca, eletrocardiograma (ECG), respiração e outros, (AWOTUNDE; CHAKRABORTY; FOLORUNSO, 2022).

2.2.1 Fotopletismografia

A técnica de fotopletismografia, do inglês *photoplethysmograph* (PPG), trata de uma técnica não invasiva que por intermédio de sensores óticos identifica alterações no volume sanguíneo. O volume se altera devido ao pulso de pressão, e é detectado a partir da iluminação da superfície da pele através de um diodo emissor de luz (LED) e, em seguida, mede-se a quantidade de luz transmitida ou refletida pelo fotodiodo (PELAEZ; VILLEGAS, 2007).

O comprimento de onda do LED usado no sensor tem um papel importantíssimo na medição do fluxo sanguíneo. Os sensores PPG operando em diferentes espectros e frequências podem ser usados na detecção de mudanças no sangue no fluxo sanguíneo. O sinal de tensão que é a saída do sensor PPG é diretamente proporcional ao volume de sangue fluindo através dos vasos sanguíneos. Pequenas mudanças no volume de sangue podem ser detectadas com precisão usando esta técnica. Os resultados mostram que quanto mais sangue flui através das artérias, a quantidade de luz que é refletida ou transmitida pelo sangue diminui e vice-versa (MORAES, 2018).

As leituras de um sensor de fotopleletismografia podem ser realizadas de duas formas. Sendo estes o método de transmissão e o método reflexivo, a diferença se dá no posicionamento dos fotodiodos. A figura 2 ilustra o conceito de ambos os métodos de transmissão, no método de transmissão, a fonte de luz e fotodiodo são colocados em lados opostos. Quando a pele está iluminada, a quantidade de luz que é absorvida é detectada pelo sensor presente no lado oposto da fonte de luz. No método de refletância, a fonte de luz e o fotodiodo são colocados no mesmo lado. Quando a pele está iluminada, a quantidade de luz que é refletida pelo sangue é detectada pelo sensor o qual se encontra do mesmo lado da fonte de luz.

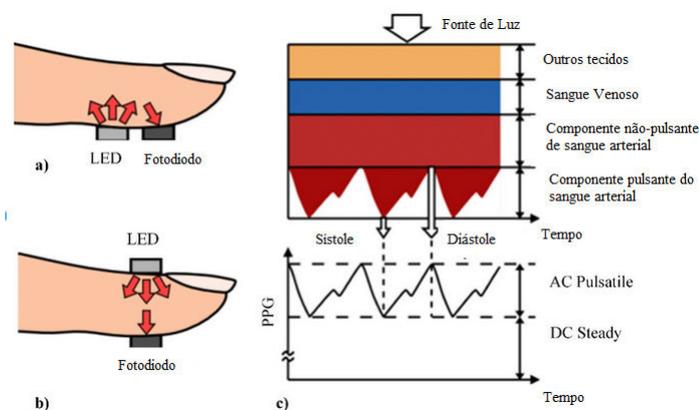


Figura 2 – Tipos de Fotopleletismografia

Fonte: (TAMURA et al., 2014)

O local de utilização do sensor afeta diretamente o resultado das leituras. Em ambientes ambulatoriais, os sensores de fotopleletismografia são comumente usados no dedo. No entanto, em alguns casos de doenças ou condições clínicas tal como casos de choque, hipotermia entre outros, o fluxo sanguíneo periférico pode ser reduzido, resultando em um PPG onde não é possível distinguir efetivamente a pulsação (BUDIDHA, 2015). Nesse caso, o PPG pode ser obtido por meio de um sensores na cabeça, sendo os locais mais comuns o ouvido, o septo nasal e a testa.

O sinal de PPG pode ser alterado por outras atividades fisiológicas, tal como a respiração, a hipovolemia e outras condições circulatórias (REISNER et al., 2008).

A respiração afeta o ciclo cardíaco, variando a pressão intrapleural, a pressão entre a parede torácica e os pulmões. Como o coração reside na cavidade torácica entre os pulmões, a pressão parcial da inspiração e da expiração influencia muito a pressão na veia cava e o enchimento do átrio direito. A movimentação voluntária ou involuntária de um usuário enquanto da utilização de sensores do tipo PPG irá acarretar a modificação dos sinais lidos pelo sensor.

Ghamari et. al define essas alterações como Artefatos de Movimento (GHAMARI; NAZERAN; SOLTANPUR, 2018).

2.2.1.1 Sinal

O sinal da fotopletismografia é uma forma de onda que apresenta as variações da radiação óptica. As variações de alta frequência, a parte 'AC', são causadas por mudanças na volume de sangue arterial a cada batimento cardíaco e variações de frequência mais baixas, a parte 'DC', são causadas por mudanças na outros componentes do tecido, como sangue venoso e capilar, tecido sem sangue, etc.

A Figura 2 apresenta de forma simplificada o caminho percorrido pela luz emitida até a identificação da parte AC. O sinal PPG exibe um padrão quase periódico que consiste em uma onda de pulso arterial para cada batimento cardíaco. Cada onda de pulso PPG consiste em duas fases distintas, como mostra a Figura 3, a anacrótica e a fases catacróticas, correspondendo aos membros ascendente e descendente, respectivamente.

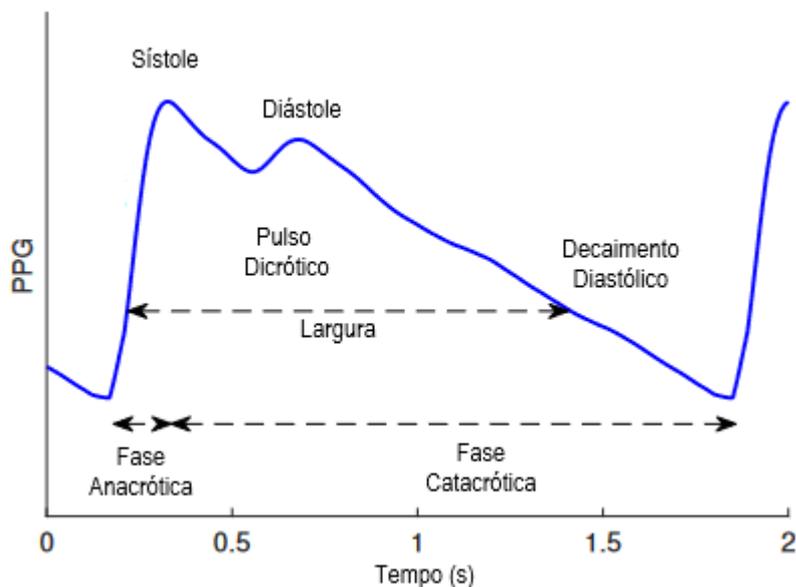


Figura 3 – Formato de Onda - Fotopletismografia

Fonte: (CHARLTON; MAROZAS, 2021)

A subida sistólica na fase anacrótica é causada pela expansão do sistema arterial devido ao fluxo sanguíneo. A taxa de expansão está ligada ao contratilidade do coração, e a amplitude

do pico sistólico está ligada ao volume sistólico. O pico diastólico é causado por reflexões de ondas, com sua localização e tempo influenciados pela rigidez arterial. O decaimento diastólico é determinado pela contração exponencial do sistema arterial devido à saída de sangue.

2.2.1.2 Configuração do Sensor

Para que um sistema de fotopletismografia possa fornecer dados de qualidade, com os menores níveis de ruído, é necessária que a implementação dos sensores sejam definidas quanto ao local de captação, comprimento de onda, intensidade de operação, taxa de amostragem e configurações geométricas.

As funções de monitoramento cardíaco são usualmente presentes em dispositivos vestíveis tais como os relógios inteligentes. Dentre as marcas disponíveis no mercado, a Samsung e a Apple apresentam em suas linhas *Galaxy Watch* e *Apple Watch* respectivamente sistemas de monitoramento de frequência cardíaca durante a execução de atividades ou repouso. Para essa atividade o principal é capacidade do sensor de identificar os picos sistólicos (TADI et al., 2016). Para tal, uma taxa de amostragem relativamente baixa pode ser usada conforme o teorema de amostragem de Nyquist, pois a largura de banda cardíaca é inferior a 5 Hz. Um comprimento de onda mais curto pode ser usado utilizando por exemplo a luz verde ou vermelha para adquirir o sinal PPG do tecido transcutâneo, uma vez que a penetração mais profunda fornecida por comprimentos de onda mais longos não é necessária. Por fim é possível janelas de 10 a 12 segundos o que é suficiente para medir a frequência cardíaca com precisão usando métodos de domínio do tempo e da frequência (BUDIDHA; KYRIACOU, 2018).

A forma de onda PPG fornece uma riqueza de informações sobre o funcionamento cardiovascular e pode ser usada para obter várias informações, como frequência cardíaca, frequência respiratória, saturação arterial de oxigênio e pressão arterial.

A aquisição simultânea de outras informações, como: eletrocardiograma (ECG), acelerometria e resposta galvânica da pele, pode aumentar o potencial de extração de informações. As medições de movimento fornecidas por acelerômetros podem ser usados para remover os chamados Artefatos de Movimento. No entanto, os acelerômetros não diferenciam entre acelera-

ção devido ao movimento e aceleração devido à gravidade. Esse problema pode ser resolvido utilizando o sinal de giroscópio para melhorar a remoção de interferência de movimento nos sinais de PPG.

2.2.2 Sensores de movimento

Os dispositivos inteligentes, como telefones, hoje são equipados com diversos tipos de sensores que permitem a detecção de movimentos simples para executar ações, como a rotação da câmera de um dispositivo, o acionamento de luzes e afins. Essas funções estão intimamente ligadas aos sistemas de movimento inercial. Estes sistemas são compostos principalmente por acelerômetros, giroscópicos e magnetômetros. Uma unidade de medição inercial funciona detectando a aceleração linear usando um ou mais acelerômetros e a taxa de rotação usando um ou mais giroscópios (IOSA et al., 2016). Atualmente diversos dispositivos vestíveis como os relógios inteligentes, ou os chamados *smartbands* incluem (IMUs) para medir e identificar atividades do dia-a-dia, como caminhadas, corridas, dentre outros. Alguns sistemas de jogos, como os controles remotos, usam IMUs para medir e rastrear o movimento.

O acelerômetro é um dispositivo que funciona para medir aceleração precisa do sistema. A aceleração é medida em relação a outro sistema em queda livre, de modo que esta está atrelada à sua sensação de peso. A utilização deste tipo de sensor é amplamente explorada em diversos segmentos na indústria usualmente temos a detecção e monitoramento de vibrações em sistemas mecânicos. No âmbito da saúde podemos encontrar acelerômetros e dispositivos que informam a inclinação de membros, tais como braços robóticos e outros.

Acelerômetros podem funcionar a partir de diversos fenômenos físicos. Atualmente com os avanços tecnológicos os sensores do tipo piezoelétricos, piezoresistivos e capacitivos têm sido amplamente utilizados em dispositivos cada vez menores. Estes sensores fornecem uma saída que é proporcional à aceleração a ser medida. Os acelerômetros são encontrados em diversos tamanhos, massas, sensibilidade, eixos de medição e faixas de amplitude e frequência. A Figura 4 ilustra de maneira simples o conceito de funcionamento de um acelerômetro do tipo piezoelétrico.

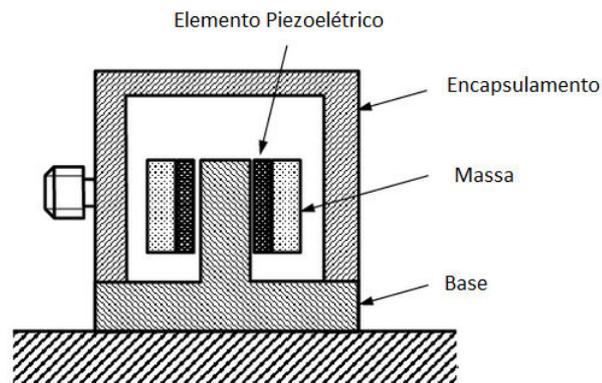


Figura 4 – Acelerômetro do tipo piezoelétrico

Fonte: (MONTALVÃO, 2015)

Se o conjunto massa estiver em repouso, o material piezoelétrico suporta o peso da massa e sofre uma compressão que resulta em uma voltagem nos terminais. Se o conjunto for acelerado, a massa irá comprimir ou não o elemento piezoelétrico. Desse modo, obtém-se uma voltagem nos terminais que deve ser proporcional à aceleração do sistema.

Outro sensor de movimento muito utilizado em dispositivos embarcados é o giroscópio, que consiste em um dispositivo cujo eixo de rotação mantém sempre a mesma direção quando não existirem forças que perturbem o sistema. Seu funcionamento se baseia no princípio da inércia, e ao contrário do acelerômetro que pode mensurar somente a aceleração com a qual um sistema está exposto, o giroscópio é capaz de identificar a direção na qual o sistema está sendo acelerado.

Diversos dispositivos inteligentes, tais como os *smartphones* e *smartwatches* apresentam estes tipos de sensores, sendo os giroscópio capazes de identificar exatamente a direção para qual o sistema está se movendo.

Essa associação de sensores é amplamente utilizada em sistemas de navegação de aeronaves. Também encontramos esses elementos em sistemas de entretenimento, como jogos eletrônicos que rastreiam o movimento do usuário para realizar a interação com o ambiente virtual.

2.3 Reconhecimento de Padrões

O reconhecimento de padrões é uma área de estudo a qual tem por principal objetivo a identificação e a classificação de um grupo de imagens, sinais ou parâmetros dentro de um grupo, sendo esses pré-determinados ou não. (THEODORIDIS; KOUTROUMBAS, 2009)

Um sistema de reconhecimento de padrões consiste de maneira geral em um elemento sensor, um módulo de processamento e o módulo de classificação. O elemento sensor irá obter as observações e leituras do ambiente, estas podem ou não passar por um tratamento específico, que serão classificadas. A partir dos dados coletados, o módulo de processamento, prossegue com a extração de características, essas características geralmente de natureza estatística das observações. E por fim o módulo de classificação, que irá empregar alguma técnica de aprendizado de máquinas para realizar a classificação das observações.

Os modelos de classificação podem ser agrupados três grupos, sendo eles aprendizado supervisionado, não supervisionado e por reforço. No aprendizado supervisionado a classificação irá ocorrer a partir do treinamento de um classificador com um banco de dados rotulado, ou seja, será apresentado ao algoritmo de classificação um conjunto de treinamento que já foi previamente classificado. Podemos citar como exemplo de métodos de aprendizado supervisionado, as Redes Neurais Artificiais, Máquinas de Vetores de Suporte, Classificadores Estatísticos, tal como Naive Bayes entre outros. (THEODORIDIS; KOUTROUMBAS, 2009)

No aprendizado não supervisionado, o sistema não recebe previamente um conjunto de treinamento estabelecendo então as classes dos padrões através de análise de padrões estatísticos. Por fim no aprendizado por reforço o algoritmo utiliza tentativa e erro em busca de uma solução para determinado problema, a cada tentativa o algoritmo recebe penalizações ou recompensas, sendo que o objetivo é sempre maximizar a recompensa total. (THEODORIDIS; KOUTROUMBAS, 2009)

2.4 Máquinas de Vetores de Suporte

Máquinas de vetores de suporte, do inglês, *Support Vector Machine* (SVM) são um conjunto de métodos de aprendizado supervisionado não probabilísticos que são empregados no processo de reconhecimento de padrões.

Dado um conjunto de entrada o SVM irá prever a qual classe os dados de entrada pertencem construindo um hiperplano ou conjunto de hiperplanos em um espaço de n-dimensões de forma a agrupar o maior número de pontos da mesma classe em um único lado do hiperplano enquanto que a distância dos pontos ao hiperplano é maximizada.

Esse algoritmo apresenta larga aplicação na categorização de textos, uma vez que a aplicação desta técnica pode reduzir significativamente a necessidade de instâncias de treinamento rotuladas. Alguns métodos de análise semântica superficial são baseados em máquinas de vetores de suporte.

A determinação do hiperplano é realizada por um subconjunto de pontos de ambas as classes, estes são denominados vetores de suporte. A equação que permite separar linearmente os pontos de cada classe, ou seja determinar o hiperplano é definido por 2.1

$$g(x) = w^T x + w_0 = 0 \quad (2.1)$$

Na equação acima, x_i , $i = 1, 2, \dots, N$ são os vetores de treinamento, este vetor contém informações pertinentes a ambas as classes. Para problemas linearmente separáveis, um perceptron realizaria a mesma atividade, entretanto as máquinas de vetores de suporte tentam maximizar a distância do hiperplano.

Podemos selecionar dois hiperplanos paralelos que separam as duas classes de dados, de forma que a distância entre elas seja a maior possível. A região delimitada por esses dois hiperplanos é chamada de "margem", como pode ser visto na Figura 5, e o hiperplano de margem máxima é o hiperplano que fica a meio caminho entre eles. Com um conjunto de dados normalizado ou padronizado, esses hiperplanos podem ser descritos pelas equações 2.2 e 2.3.

$$g(x) = w^T x - w_0 = 1 \quad (2.2)$$

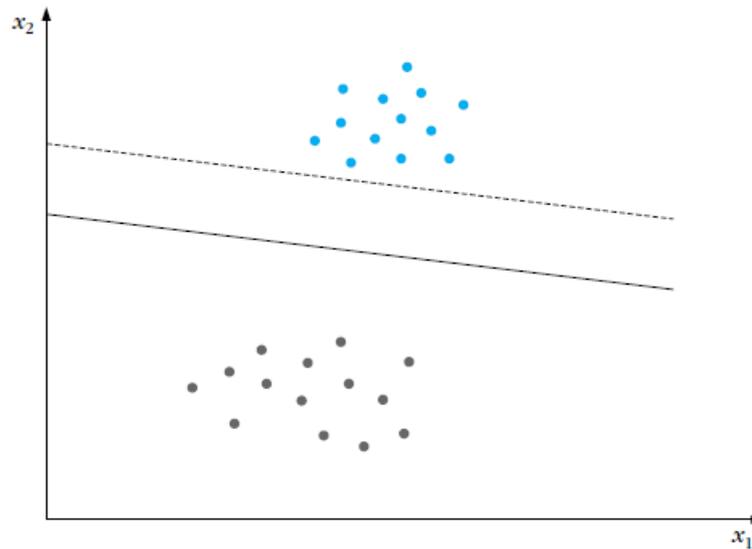


Figura 5 – Representação de Hiperplano - SVM

Fonte: (THEODORIDIS; KOUTROUMBAS, 2009)

$$g(x) = w^T x - w_0 = -1 \quad (2.3)$$

A distância euclidiana do hiperplano para os pontos que estão a sua margem é dada por:

$$\frac{y_j(wx_j + b)}{\|w\|} = \frac{1}{\|w\|} \quad (2.4)$$

Para otimizarmos o espaço de separação do hiperplano, ou seja, para alcançarmos um valor máximo na equação ??, temos que fazer com que $\|w\|$ tenha um valor mínimo. Para evitar que os pontos de dados caiam na margem, é adicionada a seguinte restrição, ver equações 2.5 e 2.6.

$$w^T x - w_0 \geq 1 \text{ se } y_i = 1 \quad (2.5)$$

$$w^T x - w_0 \leq 1 \text{ se } y_i = -1 \quad (2.6)$$

As equações 2.5 e 2.6 podem ser reescritas da seguinte forma:

$$y_i(w^T x - w_0) \geq 1, \text{ para todos } 1 \leq i \leq n \quad (2.7)$$

Para estender o SVM para casos em que os dados não são linearmente separáveis, a função de perda a ser utilizada é dada por:

$$\max(0, 1 - y_i(w^T x - w_0)) \quad (2.8)$$

Esta função é zero se a restrição for satisfeita, ou seja, se x_i estiver no lado correto da margem. Para dados no lado errado da margem, o valor da função é proporcional à distância da margem.

O objetivo da otimização é minimizar a seguinte equação:

$$\|w\|^2 + \lambda \left[\frac{1}{n} \sum_{i=1}^n \max(0, 1 - y_i(w^T x - w_0)) \right] \quad (2.9)$$

Na equação 2.9 o parâmetro λ realiza a compensação entre aumentar o tamanho da margem e garantir que x_i esteja no lado correto da margem. Assim, para valores suficientemente grandes de λ , ele se comportará de forma semelhante a 2.7, se os dados de entrada forem linearmente classificáveis, mas ainda aprenderá se uma regra de classificação é viável ou não.

O algoritmo de hiperplano de margem máxima deu origem ao classificador linear, posteriormente Corinna e Vapnik (CORTES; VAPNIK, 2009) sugeriram uma maneira de criar classificadores não lineares aplicando um *kernel*.

O algoritmo resultante é semelhante ao separador linear, entretanto cada produto escalar é substituído por uma função de *kernel* não linear. Isso permite que o algoritmo ajuste o hiperplano de margem máxima. Os tipos de *kernel* comumente utilizados: Polinomial (Homogêneo e não-homogêneo), Gaussiano e tangente hiperbólico. A Figura 6 apresenta o resultado da aplicação de um separador não linear.

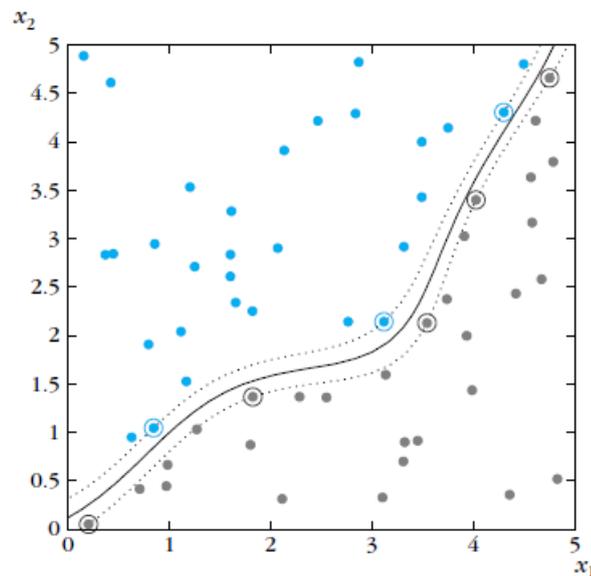


Figura 6 – Representação de Hiperplano - SVM - Não linear
Fonte: (THEODORIDIS; KOUTROUMBAS, 2009)

2.5 Métodos Baseados em Árvores

2.5.1 Métodos Baseados em Árvores

As técnicas utilizando árvores de decisão introduzidas inicialmente por (BREIMAN et al., 1984) têm como objetivo discriminar, em diferentes regiões, observações com base em suas características, realizando partições binárias recursivas que por fim se assemelham as ramificações de uma árvore. Esta técnica é utilizada como base para técnicas mais complexas.

As árvores de decisão também denominadas CART (*Classification and Regression Trees*), aplicam a segmentação do espaço, p -dimensional, criado por variáveis p , de forma a gerar regiões disjuntas de classificação.

Cada elemento do domínio da classificação é denominado classe. Uma árvore de decisão é uma árvore na qual cada nó é rotulado com um recurso de entrada. Os ramos vindos de um nó rotulado com um recurso de entrada são rotulados com cada um dos valores possíveis do recurso de destino, ou o ramo leva a um nó de decisão subordinado em um recurso de entrada diferente.

Cada folha da árvore é rotulada com uma classe ou uma distribuição de probabilidade sobre as classes, significando que o conjunto de dados foi classificado pela árvore em uma classe específica ou em uma distribuição de probabilidade particular.

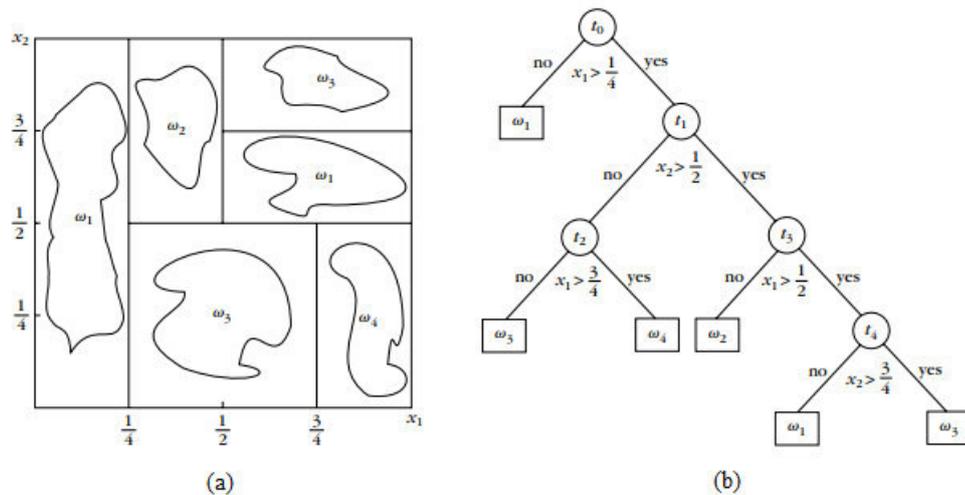


Figura 7 – Representação de uma árvore de decisão

Fonte: (THEODORIDIS; KOUTROUMBAS, 2009)

Na figura 7(a) temos um exemplo de um espaço bidimensional dividido em diversas regiões, na figura 7(b), é apresentada uma proposta de árvore de decisão é possível observar que nas árvores de decisão temos sucessivas partições binárias que são baseadas em limiares que caracterizam as regiões de classificação.

Em cada nó, o conjunto de perguntas a serem feitas deve ser decidido. Cada questão corresponde a um binário específico dividido em dois nós descendentes. Cada nó, t , está associado a um subconjunto específico X_t do conjunto de treinamento X . A divisão de um nó é equivalente à divisão do subconjunto X_t em dois subconjuntos descendentes disjuntos, X_{tY} conjunto que representa as respostas positivas, X_{tN} conjunto de respostas negativas.

Os critérios de divisão devem ser escolhidos de forma a alcançar a melhor divisão do conjunto de candidatos. Para que a metodologia de crescimento da árvore, do nó raiz até as folhas, faça sentido, cada divisão deve gerar subconjuntos homogêneos de classe em comparação com o subconjunto X_t do ancestral. Isso representa que os vetores de treinamento em cada um dos novos subconjuntos mostram uma preferência maior por uma classe específica, enquanto os dados em X_t são mais igualmente distribuídos entre as classes.

Se o algoritmo dividir o nó de forma que os pontos que pertencem a ω_1, ω_2 classes formem o subconjunto X_{tY} , e os pontos de ω_3, ω_4 classes formem o subconjunto X_{tN} , então os novos subconjuntos são mais homogêneos em comparação com X_t . A equação 2.10 calcula o nível de homogeneidade.

$$I(t) = - \sum_{i=1}^M P(\omega_i|t) \log_2 P(\omega_i|t) \quad (2.10)$$

Na equação acima, M representa o número total de classes, e \log_2 representa o logaritmo na base 2. A equação 2.10 apresenta o nível de entropia do conjunto X_t , podemos inferir que o nível de impureza atingirá seu máximo, se todas as probabilidades forem iguais a $1/M$, se tivermos somente uma classe teremos também que o nível de impureza será nulo. O decréscimo do nível de impureza é dado pela equação 2.11

$$\Delta I(t) = I(t) - \frac{N_{tY}}{N_t} I(t_Y) - \frac{N_{tN}}{N_t} I(t_N) \quad (2.11)$$

Onde N_{tY} representa o número de amostras que foram atribuídos a determinada classe com o valor positivo, $I(t_Y)$ e $I(t_N)$ são respectivamente as impurezas dos nós t_Y e t_N .

A parada de subdivisões é determinada por um critério de parada, usualmente esse critério é um limiar T onde após sucessivas divisões o valor máximo de $\Delta I(t)$ é menor que T .

Por fim, a atribuição de uma classe, ou seja, quando um nó é determinado como uma folha, a equação utilizada é apresentada na equação 2.12.

$$j = \operatorname{argmax}_i P(\omega_i|t) \quad (2.12)$$

A figura 8 apresenta um fluxograma ilustrando os passos a serem seguidos em uma árvore de decisão.

2.5.2 Florestas Aleatórias

Random Forest (RF), ou Floresta Aleatória, é uma técnica de aprendizado de máquina usada para resolver problemas de regressão e classificação. Esta técnica utiliza aprendizagem por conjunto, que é uma técnica que combina muitos classificadores para fornecer soluções para problemas complexos. Um algoritmo de RF consiste na avaliação de diversas árvores de decisão.

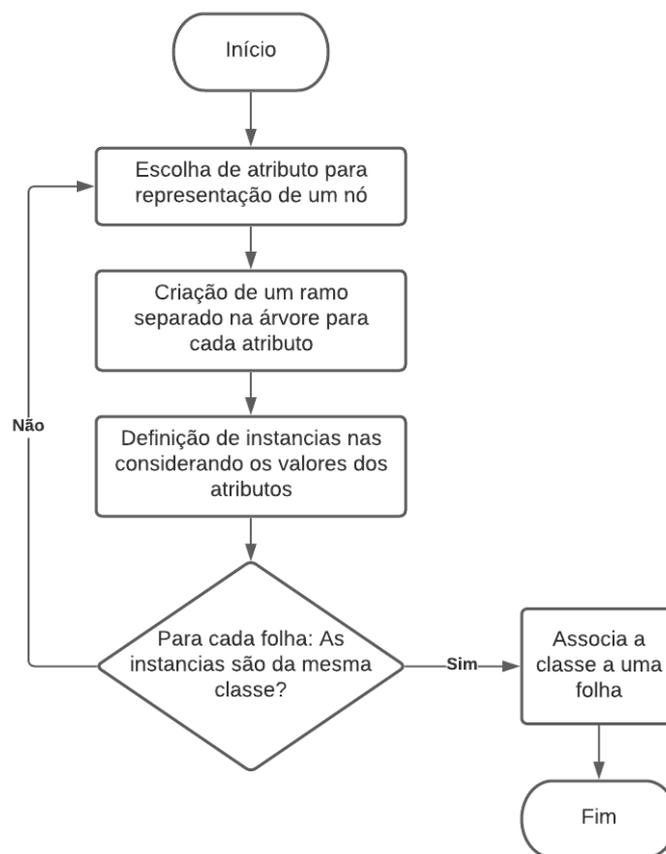


Figura 8 – Fluxograma de um algoritmo de Árvore de Decisão

Fonte: Própria

Para tarefas de classificação, a saída da floresta aleatória é a classe selecionada pela maioria das árvores. Os algoritmos de RF usam a ideia de ensacamento em conjunto com a seleção de recursos aleatórios (BREIMAN, 2001). O recurso a ser dividido em cada nó é selecionado como o melhor entre um conjunto de recursos F escolhidos aleatoriamente, onde F é um parâmetro definido pelo usuário. Esta aleatoriedade introduzida tem um efeito substancial na melhoria do desempenho.

2.5.3 *Extreme Gradient Boosting (XGBoost)*

Este é um método baseado em árvores amplamente utilizado dentre os algoritmos com aumento de gradiente (CHEN; GUESTRIN, 2016). Entretanto o que difere este método dos demais métodos de árvores é a utilização de um aumento de gradiente para minimizar a perda durante o processo de adição de novos modelos.

O aumento de gradiente ou *Gradient Boosting* é uma técnica de aprendizado de máquinas que cria um modelo em forma de um conjunto de modelos de previsão fracos que usualmente são árvores de decisão. Um modelo de árvores gradientes é construído de forma faseada assim como outros métodos de *boosting*, mas ele generaliza os outros métodos, permitindo a otimização de uma função de perda arbitrária.

2.6 Validação Cruzada

No processo de aprendizado de máquina podemos utilizar de diversas formas para treinar nossos modelos, de forma genérica o objetivo das etapas de treinamento é reduzir as taxas de erro do teste, de forma a assegurar uma maior precisão do modelo.

Dependendo da quantidade disponível de dados para aprendizado e teste podemos utilizar algumas metodologias para aplicar no treinamento de nossos modelos.

O objetivo da validação cruzada é definir um conjunto de dados para testar o modelo durante a fase de treinamento. Ele envolve a partição do conjunto de dados de treinamento em subconjuntos, onde um subconjunto é mantido para testar o desempenho do modelo. Esse conjunto de dados é chamado de conjunto de dados de validação. A tabela 1, ilustra a validação cruzada que empregaremos neste trabalho, os dados disponíveis serão divididas em 5 grupos

Treinamento				Teste
Pasta 1	Pasta 2	Pasta 3	Pasta 4	Pasta 5
1	2	3	4	5
5	1	2	3	4
4	5	1	2	3
3	4	5	1	2
2	3	4	5	1

Tabela 1 – Validação Cruzada (K-fold onde k = 5)

O método de validação cruzada é uma boa alternativa para detectar sobre-ajuste, *overfit-*

ting, ou seja, a quando um modelo se ajusta muito bem ao conjunto de dados, mas se mostra ineficaz para prever novos resultados.

2.7 Mineração de Dados

2.7.1 Transformação de Dados

No processo de mineração de dados, é comum a realização de processos onde os dados são transformados ou consolidados em formas apropriadas para mineração. Podemos citar a suavização, que funciona para remover o ruído dos dados, a agregação, onde são realizadas as operações onde são extraídas informações como por exemplo a média de um determinado segmento de sinal. A normalização é outro exemplo, onde os dados do atributo são dimensionados de modo a se enquadrarem em um pequeno intervalo, como 0 a 1 ou -1 a 1. (HAN; PEI; KAMBER, 2011) E por fim a construção de atributos, onde novos atributos são construídos e adicionados do conjunto de atributos fornecido para ajudar no processo de mineração, esse processo pode ser realizado utilizando a combinação de atributos anteriormente obtidos por meio da agregação.

2.7.2 Suavização

A suavização é uma forma de limpeza dos dados onde é possível realizar a remoção de componentes ou dados indesejados ao modelo, podendo realizada de diversas formas. O primeiro passo na limpeza de dados é a detecção de discrepâncias estas podem ser causadas por vários fatores, incluindo formulários de entrada de dados mal projetados que têm muitos campos opcionais, erro humano na entrada de dados, erros deliberados. Discrepâncias também pode surgir de representações de dados inconsistentes e do uso inconsistente de códigos. Erros em dispositivos de instrumentação que registram dados e erros de sistema são outra fonte de discrepâncias.

Sensores dos mais diversos tipos estão suscetíveis aos mais diversos tipos de ruídos, sejam

eles advindos de limitações internas de hardware ou software. Esses ruídos quando introduzidos em um fluxo de dados podem produzir efeitos indesejáveis, como técnica de tratamento e limpeza podemos apontar os filtros e as técnicas de decomposição de sinais. (HAN; PEI; KAMBER, 2011)

2.7.2.1 Decomposição Wavelet

Decomposição Wavelet é uma ferramenta de análise e síntese de sinais assim como a transformada de Fourier. No que tange ao campo de tratamento de sinais, a aplicação desse tipo de transformada incluem supressão de sinais, supressão de ruído; e compactação. (HAN; PEI; KAMBER, 2011)

O grande diferencial desta técnica é a capacidade de analisar sinais cujo espectro variam no tempo. Para muitos tipos de sinais, a análise de Fourier é extremamente útil, entretanto a transformação no domínio da frequência acarreta em uma perda da informação temporal do sinal.

No caso de sinais fisiológicos, tal como o sinal advindo da fotopletismografia, estes apresentam características transitórias ou não-estacionárias. Portanto se faz necessária a utilização de uma técnica que permita avaliar essas mudanças ao longo do tempo, tornando a decomposição *wavelet* uma excelente candidata para a análise deste tipo de sinal.

A decomposição wavelet é uma técnica amplamente utilizada na mineração de dados quando se trata de processamento de sinais fisiológicos. Na literatura temos como exemplo os trabalhos de (JAAFAR; ROZALI, 2017), (LI; CHUNG, 2013) e (SAREEN et al., 2008), com destaque para do trabalho de Sareen et al. que utiliza a técnica no processo de extração de atributos.

O janelamento temporal utilizado na decomposição *wavelet* aplica as menores larguras ao componentes de alta frequência, e as maiores larguras para os componentes de baixa frequência. A figura 9 apresenta uma breve comparação entre o método de onduleta e o método de Fourier com janelamento.

Uma onduleta, do inglês *wavelet*, é uma forma de onda a qual possui uma duração

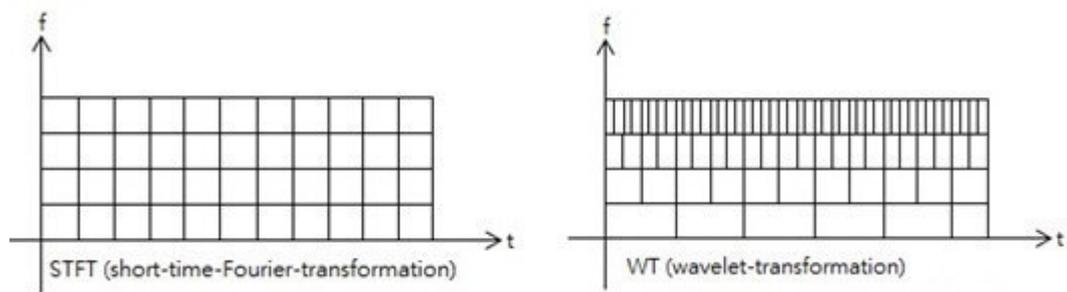


Figura 9 – Tipos de wavelet

Fonte: (GANESAN et al., 2003)

limitada e apresenta valor de média igual a zero. Em uma comparação com uma senoide que possui um tamanho ilimitado no tempo e o comportamento contínuo, ou seja, sem alterações abruptas, as onduletas apresentam irregularidades e são assimétricas, ver figura 10

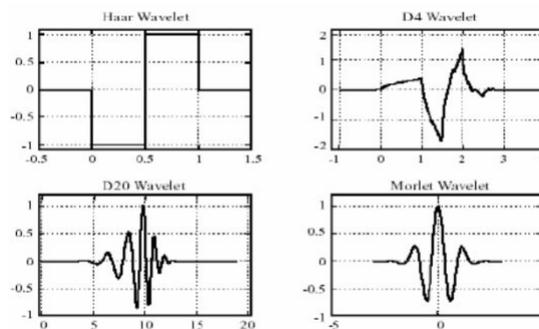


Figura 10 – Tipos de wavelet

Fonte: (PERERA; RAJAPAKSE; JAYASINGHE, 2007)

A aplicação da transformada de Fourier irá decompor um sinal em ondas senoidais em um determinado espectro de frequências. Na decomposição *wavelet*, é gerada uma série de sinais em versões deslocadas e escalonadas da *wavelet* original. Sinais com mudanças abruptas são potencialmente melhor analisados com uma onda irregular quando comparados a uma análise de uma senoide. De forma geral, podemos classificar a análise por *waveles* em duas categorias, análise contínua e discreta. Na transformada contínua, CWT, há um grande esforço computacional para o cálculo de escalas, o que pode ser questionável uma vez que para determinadas análises, somente algumas escalas são passíveis de análise, para solucionar esse problema temos então a transformada discreta de *wavelet*, DWT.

2.7.2.2 Transformada Discreta de Wavelet

Decomposição *wavelet* para um número muito grande de escalas é uma tarefa que gera uma carga elevada de processamento e consequentemente um grande volume de dados. Uma estratégia para reduzir o esforço computacional é a escolha de apenas um sub-conjunto de escalas e posições específicas para as quais os coeficientes devem ser calculados.

A decomposição discreta divide o sinal em componentes de alta e baixa Escala, ou maior e menor largura respectivamente. Os componentes de alta são chamadas aproximação, estas correspondem as frequências mais baixas, os elementos de alta frequência, responsáveis por alterações abruptas no sinal são definidos como detalhes, ver figura 11.

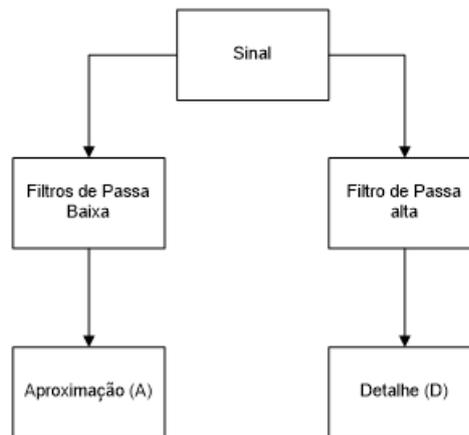


Figura 11 – Tipos de *wavelet*

Fonte: Própria

Matematicamente a discretização de um sinal por meio da técnica de DWT é representada pela equação 2.13

$$\Psi_{m,n}(t) = \frac{1}{\sqrt{a_0^m}} \Psi\left(\frac{t - nb_0 a_0^m}{a_0^m}\right), m, n \in \mathbb{Z}, a_0 > 1, b_0 \neq 0 \quad (2.13)$$

Onde os parâmetros escala (a), translação (b) são discretizados para $a = a_0^m$ e $b = nb_0 a_0^m$, onde a_0 e b_0 são amostras do intervalo e m e n são os parâmetros de translação. É possível notar uma correlação entre as discretizações da escala e o tempo. À medida que o valor da escala aumenta, a largura da *wavelet* também aumenta, diminuindo o número de amostras no tempo.

Por fim o sinal resultante de uma DWT pode ser representado por:

$$DWT_{m,n} = \frac{1}{a_0^{m/2}} \sum_n f(t) \psi \left(\frac{t - nb_0}{a_0^m} \right) \quad (2.14)$$

Onde é o valor de $DWT_{m,n}$ é conhecido como coeficiente *wavelet* ou coeficiente de detalhe na posição (m, n) da matriz de representação. A figura 12 apresenta o resultado da decomposição *wavelet* aplicada ao sinal de PPG. É possível observar a separação dos chamados artefatos de movimento, componentes de alta frequência, sem a perda de características importantes do sinal inicial.

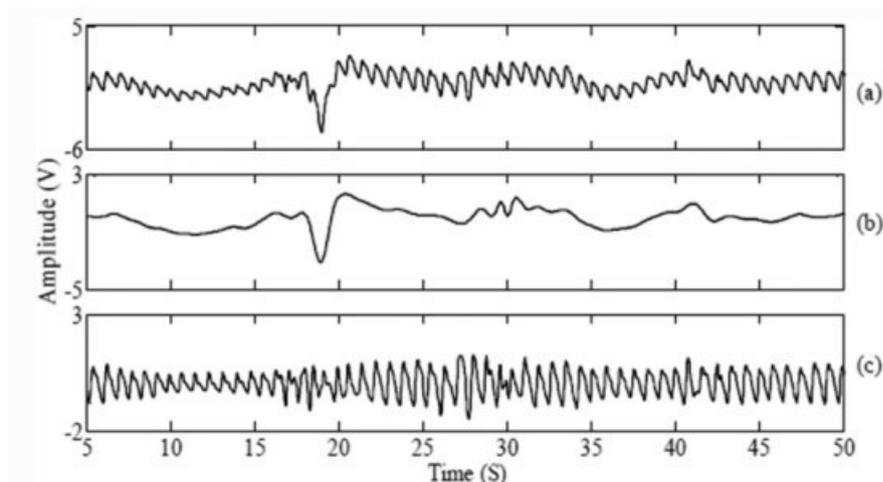


Figura 12 – Decomposição *Wavelet* aplicada a Sinal PPG

Fonte: (POLLREISZ; TAHERINEJAD, 2019)

Na figura 12 (a), é apresentado o sinal original de PPG, nota-se a existência de componentes de alta frequência. A partir da aplicação da transformada discreta de *wavelet* temos como resultante os itens (b) e (c) onde (b) representa as escalas de baixa frequência separada dos denominados detalhes, estes apresentados no item (c).

2.7.3 Agregação

De acordo com (HAN; PEI; KAMBER, 2011) a agregação e a generalização servem como formas de redução de dados. Técnicas de redução de dados podem ser aplicadas para obter uma representação reduzida do conjunto de dados que é muito menor em volume, mas mantém características relevantes que representam os dados originais. Essa etapa ajuda principalmente a reduzir esforços computacionais durante o processo de treinamento de um classificador.

2.7.4 Normalização

Um atributo é normalizado escalando seus valores de modo que sua variação estejam compreendidas dentro de um intervalo, como 0,0 a 1,0. A normalização é particularmente útil para algoritmos de classificação envolvendo redes neurais ou que envolvam a otimização de distancias, podemos citar como exemplos os métodos não supervisionados de formação de *clusters*. No caso de algoritmos de retro propagação, a normalização dos valores de entrada para cada atributo medido ajudará na redução de tempo da fase de aprendizado. Para métodos baseados em distância, a normalização ajuda a evitar que atributos com intervalos inicialmente grandes superem atributos com intervalos inicialmente menores como atributos binários. Algoritmos de classificação que utilizam planos de separação necessitam de processos de normalização uma vez que envolvem cálculos de distancia para que sejam estabelecidos os hiperplanos de separação.

Podemos citar como método de escalonamento a utilização de normalização de mínimo e máximo, onde é realizada uma transformação linear do conjunto de dados inicial. Essa técnica irá realizar o mapeamento de cada valor do conjunto para valores correspondentes dentro do intervalo mínimo e máximo. A equação 2.15 representa este processo:

$$Amostra = \frac{v - \min_A}{\max_A - \min_A} \quad (2.15)$$

A normalização de mínimo e máximo, preserva as relações entre os valores de dados originais.

2.8 Seleção de Atributos

A seleção de atributos ou variáveis é um processo de redução do número de variáveis de entrada que serão aplicadas em um modelo preditivo. A redução do número de variáveis de entrada é necessária para reduzir o custo computacional da modelagem de forma que o processamento seja otimizado. Isso se torna importante principalmente quando temos aplicações embarcadas, onde os dispositivos não possuem tanto poder computacional para realizar etapas de pré-processamento complexas.

Os métodos de seleção de recursos baseados em estatísticas envolvem a avaliação do relacionamento entre cada variável de entrada e a variável de saída. A partir de métricas, tal como o precisão, acuraria e *recall* nos casos dos modelos de classificação, é possível identificar quais variáveis de entrada possuem uma maior relevância com os resultados. Esses métodos podem ser rápidos e eficazes, embora a escolha de medidas estatísticas dependa do tipo de dados das variáveis de entrada e saída.

Existem dois tipos principais de métodos para a seleção de atributos, como mostrado na figura 13. Sendo estas técnicas divididas principalmente em supervisionadas e não supervisionadas, os métodos supervisionados podem ser divididos em *wrapper*, filtro e intrínseco. Quando o resultado é ignorado durante a eliminação dos atributos, a técnica não é supervisionada. As técnicas supervisionadas usam a variável de destino para o cálculo de métricas que permitem avaliar a relevância deste atributo.

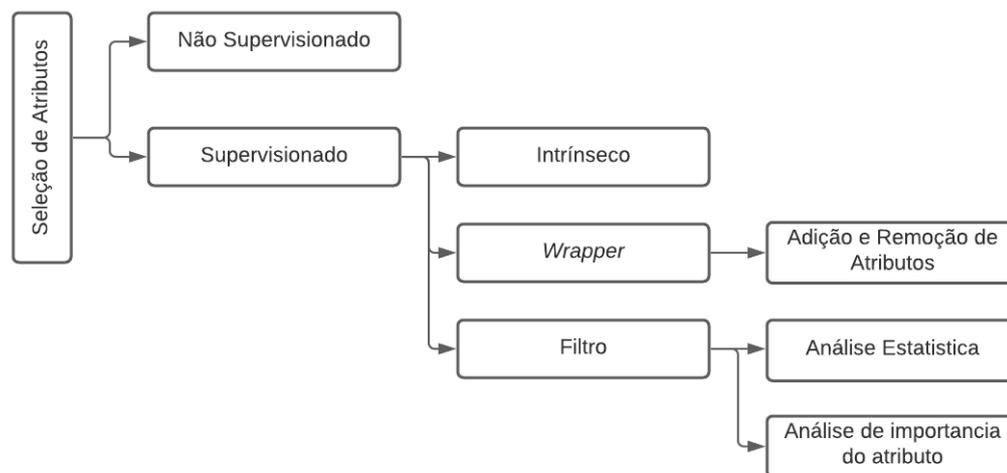


Figura 13 – Classificação de métodos de seleção de atributos

Fonte: Própria

Os métodos de *wrapper* avaliam vários modelos usando procedimentos que adicionam e/ou removem as variáveis para encontrar a combinação ideal que maximiza o desempenho do modelo (KUHN; JOHNSON, 2013). Os métodos de seleção de recursos de filtro usam técnicas estatísticas para avaliar a relação entre cada variável de entrada e a variável de destino, e essas pontuações são usadas como base para selecionar as variáveis que serão usadas no modelo.

A seleção de *features* também está relacionada com técnicas de redução dimensional, uma vez que ambos os métodos buscam reduzir a quantidade de variáveis de entrada para um

modelo. A diferença é que a seleção de recursos seleciona recursos para manter ou remover do conjunto de dados, enquanto a redução de dimensionalidade cria uma projeção dos dados, resultando em recursos de entrada inteiramente novos.

3 Trabalhos Relacionados

Nesta seção, apresentamos pesquisas selecionadas por apresentarem os objetivos similares e/ou metodologias que tenham semelhança com a deste projeto. Foi realizada uma busca sistemática de publicações nas seguintes plataformas de pesquisa IEEE, Springer, ACM, Elsevier e ScienceDirect.

A a busca de semelhanças entre sinais e parâmetros representa uma parte do processo de identificações de padrões. Em destaque podemos colocar os sistemas baseados em visão computacional, a sua utilização varia desde identificação de letras e algarismos a identificação de movimentos seja de pessoas, veículos, ou demais objetos. Além das técnicas baseadas em imagem, outros dispositivos podem ser utilizados para gerar sinais que possam ser analisados e posteriormente classificados, tal como as unidades de medida inerciais. Estes dispositivos geralmente imbuídos de acelerômetros e giroscópios apresentam uma frequência de amostragem superior a 100hz, isso permite um sinal com uma maior quantidade de detalhes. Entretanto existem limitações na utilização de sensores inerciais quando o movimento possui uma amplitude reduzida, pois os padrões se tornam mais complexos para segmentação e classificação.

A aplicação de sistemas de reconhecimento de padrões voltada para gestos é abordada por Perrault onde o autor apresenta um sistema onde permite ao usuário interagir com um dispositivo vestível por meio de funções simples como controle de telas ou simulação de toque do usuário. (PERRAULT et al., 2013).

Outro método bastante utilizado, é eletromiografia. Esta técnica não-invasiva verifica a atividade elétrica das membranas das células musculares. O monitoramento ao longo do tempo da tensão elétrica é utilizada para realizar o reconhecimento dos gestos. Essa técnica foi abordada por Zhang et al (ZHANG et al., 2018) onde é proposto a utilização de um conjunto de eletrodos em uma pulseira, os resultados de precisão alcançados foram de 99%, a pulseira com os eletrodos permite que diversos pontos de contato sejam avaliados apresentando uma cobertura mais ampla dos tecidos do pulso. (WEIGEL et al., 2017) apresenta uma abordagem similar onde propõe a utilização de sensores em contato direto com a pele do usuário para realizar o monitoramento de alterações elétricas.

A utilização de diferentes tipos de sensores também já foi apresentada por Lu et al. (LU et al., 2014), onde o autor utiliza acelerômetros para captar movimentos amplos e sensores de eletromiografia para movimentos de menor amplitude, o método proposto pelo autor utiliza um classificador linear Bayesiano e um algoritmo de *Dynamic Time Warping* aprimorado. Com o dispositivo usado no antebraço, o usuário pode manipular um telefone celular usando 19 gestos predefinidos ou até mesmo personalizados. Chan et al. (CHAN et al., 2013) apresenta um método para coleta e reconhecimento de gestos utilizando a posição dos dedos, onde, por meio de um sensor de efeito Hall conectado ao dispositivo identifica gestos realizados a curtas distâncias do dispositivo. Este método pode alcançar uma exatidão de até 93%. O método apesar de apresentar elevada exatidão foca principalmente na atividade de rastreamento da posição dos dedos atuando principalmente como uma forma de *touchpad*.

Além de técnicas que envolvam a eletromiografia ou a detecção de campos magnéticos, foram propostos também, métodos que utilizam outros sistemas como os sistemas de captação de áudio e a fotopletismografia por exemplo.

Nandakumar et al. (NANDAKUMAR et al., 2016) e Gupta et al. (GUPTA et al., 2012) utilizam o sistema de captação de áudio como um sonar para realizar a detecção de movimentos dos dedos do usuário. Em ambos os trabalhos a mão do usuário não é equipada com nenhum dispositivo, os sensores são usualmente equipados em dispositivos externos tal como *smartphones* e computadores. Gupta et al. Foca principalmente na detecção de gestos mais robustos, para isso o autor gera um sinal acústico inaudível que sofre uma mudança de frequência quando é refletido pela mão do usuário. Nandakumar et al. Apresenta uma proposta similar, entretanto apresenta uma abordagem inovadora que utiliza uma técnica de modulação comumente usada em comunicação sem fio chamada *Orthogonal Frequency Division Multiplexing* (OFDM). O trabalho mostra que o método consegue rastrear os dedos em um ambiente bidimensional com uma precisão média de 8 mm, usando microfones e alto-falantes integrados.

Utilizando a fotopletismografia (PPG) podemos citar Zhang et al. (ZHANG et al., 2018) onde o autor utilizou os sensores de PPG de um relógio inteligente para realizar a identificação de um conjunto de gestos previamente determinado, apresentando uma comparação quanto a utilização de sensores de movimento. Em seu trabalho o autor utiliza dois classificadores, as Máquinas de vetores de suporte e a técnica de Random Forest para realizar a tarefa de

classificação dos gestos, os mesmos parâmetros são utilizados para processar os dados oriundos do sensor de PPG e dos sensores de movimento. De acordo com o autor, os modelos que utilizam sensores de movimento, apresentam desempenho superior ao modelo baseado em fotoplestismografia, exceto quando o usuário não está em movimento de translação.

Wen et al. (WEN; ROJAS; DEY, 2016), apresenta em seu trabalho uma nova técnica de reconhecimento de gestos usando sensores de movimento integrados (acelerômetro e giroscópio) em *smartwatches* prontos para uso. O método proposta calcula a magnitude de todos os eixos combinados, para compensar as diferentes orientações do dispositivo. As features utilizadas no modelo incluem 7 recursos estatísticos de uma janela deslizante de 1 segundo: média, desvio padrão desvio padrão, máximo, mínimo, quantil. Em seguida, uma transformação rápida de Fourier (FFT) da janela produz 25 bandas de potência, e mantemos as 10 bandas inferiores, que representam a frequência de 0 a 10 Hz. Como um único gesto leva cerca de 0,4s a 0,8s para ser realizado, o que gera um pico entre 1 2Hz, essa faixa de banda é suficiente para capturar as características de frequência dos gestos. Esse processo é replicado para cada eixo, resultando em um 357 atributos para cada sinal.

O autor enfatiza a determinação de um conjunto de gestos que não possua movimentos conflitantes, ou seja, gestos que não contenham componentes comuns em outros gestos. Por fim esses atributos são utilizados para treinar um modelos de aprendizado de máquinas de forma supervisionada, sendo que as máquinas de vetores de suporte obtiveram os melhores resultados. O autor não aponta em seu artigo um processo de redução de dimensionalidade ou seleção de atributos. O precisão do modelo no reconhecimento de gestos finos como beliscar, bater e esfregar os dedos, atingiu uma média 87% de acurácia.

O trabalho de Wen et al. apresenta também a necessidade de a necessidade de uma metodologia adequada para a coleta dos gestos, ele propõe que sejam utilizadas janelas de 3 segundos para a execução de cada gesto, isso facilitará a segmentação do banco de dados e o pós-processamento.

O trabalho de Wen et al. apresenta também a necessidade de a necessidade de uma metodologia adequada para a coleta dos gestos, ele propõe que sejam utilizadas janelas de 3 segundos para a execução de cada gesto, isso facilitará a segmentação do banco de dados e o pós-processamento.

Em (BOUKHECHBA et al., 2019), os resultados apresentam evidências da eficácia de modelos utilizando redes neurais convolucionais e recorrentes para prever diferentes atividades humanas, como por exemplo, caminhada, corrida, salto. Os resultados também destacam a importância de extrair sinais de artefatos de movimento e como usá-los para entender o comportamento humano. Trabalhos futuros pretendem estudar o efeito da combinação de dados PPG com outros sensores de baixa energia, como acelerômetro e giroscópio.

No trabalho de (SUBRAMANIAN; SAVUR; SAHIN, 2020), é apresentado que a utilização de sinais de PPG podem alcançar resultados semelhantes a técnicas que utilizam eletromiógrafos. A pulseira foi utilizada num processo de comparação de quatro gestos. As evidências desta pesquisa mostram que, apesar de uma menor quantidade de sensores e uma taxa de amostragem menor, os classificadores são capazes de aprender os gestos em sinais PPG e os resultados são comparáveis a um classificador que foi treinado com dados de um eletromiógrafo. Os modelos treinados com sinais de PPG podem classificar um conjunto de 4 gestos com as mãos com uma precisão de 93%.

O autor aponta também que a utilização de sensores de movimento podem ampliar positivamente os resultados (um acelerômetro pode ser usado com sinais PPG em um esquema de fusão de dados para identificar a orientação da mão).

(ZHANG et al., 2018) apresenta um modelo que foi capaz de reconhecer com 90.55% de precisão e um *recall* de 90.73% um total de 10 gestos, considerando situações onde há movimento do usuário.

O método proposto utiliza em seu pré-processamento uma etapa de suavização do sinal, onde é utilizada a Decomposição Wavelet Discreta. Em seguida o módulo de *Gradient-based Movement Detection (GMD)* onde o sinal é convertido em uma curva de gradiente, onde o autor irá identificar atributos necessários para a segmentação dos gestos.

O módulo de extração de atributos coleta um total de 138 atributos que incluem características extraídas do domínio do tempo e frequência. A seleção de atributos é realizada por meio dos algoritmos de Seleção Sequencial de Atributo, Importancia de Features (*Random Forest*) e ReliefF *feature selection*.

Dentre as oportunidades de estudos posteriores, é apontada a utilização de abordagens

de aprendizado profundo, como redes neurais convolucionais, para classificar gestos a partir do espectrograma de PPG para necessariamente realizar a extração de atributos. Trabalhos futuros podem introduzir novos atributos para ajudar os classificadores atuais.

Em uma abordagem utilizando sensores de movimento e sensores de PPG, (ZHAO et al., 2019) realiza a classificação dos gestos de punho e mão para realizar a identificação de gestos relacionados a linguagem de sinal americana utilizando métodos de aprendizado profundo, a técnica conseguiu alcançar precisão de aproximadamente 89% utilizando somente sensores de PPG. O autor apresentou também as limitações da utilização separada de sensores de movimento, onde o classificador obteve resultado de 83% de acurácia e 84% de *recall*.

No trabalho de (LEE; CUONG; WEI, 2017), os autores propõem uma série de sensores, incluindo sensores posicionados nos dedos indicador e polegar para realizar a identificação de gestos finos. Um total de seis gestos foram avaliados, dentre eles gestos já conhecidos dentre os usuários de *smartphones*, como os movimentos de pinça para realização de aproximação e afastamento em fotos. A acurácia média obtida foi de 93.19%.

Dentre os artigos citados nesta pesquisa, o trabalho de (LING et al., 2021), apresenta maior correlação com esta proposta, o autor aponta que os sensores de fotopletismografia são mais adequados para a implementação no processo de reconhecimento de gestos em dispositivos vestíveis, devido às suas vantagens de ser insensível ao ruído do movimento e reduzir o carga de treinamento do usuário.

Além de técnicas que envolvam a eletromiografia ou a detecção de campos magnéticos, foram propostos também, métodos que utilizam por exemplo os sistemas de captação de áudio, Nandakumar et al. (NANDAKUMAR et al., 2016) e Gupta et al. (GUPTA et al., 2012) utilizam o dispositivo como um sonar para realizar a detecção de movimentos dos dedos do usuário.

O autor aponta também que um número maior de canais podem aumentar a precisão do reconhecimento baseado em PPG. A combinação de PPG no espectro vermelho e acelerômetros obtém o melhor desempenho do classificador as métricas de precisão são $96,5 \pm 2,5\%$ no cenário de caminhada; $95,4 \pm 2,3\%$ no cenário de *jogging* e $92,9 \pm 3,6\%$ no cenário de corrida.

Autor	Dispositivo / Amostragem	Sensores	Segmentação	Classificador	Gestos / Resultados
BOUKHECHBA et al. , 2019	Huawei Watch 2 Smartwatch Amost. 100Hz	PPG	Janela Deslizante	DNN	5 Movimentos de Corpo 0.78 F1 Score
SUBRAMANIAN; SAVUR; SAHIN., 2020	Pulseira com 3 Sensores de PPG Amost. 60Hz	PPG, IMU	Janela Deslizante (1s)	SVM, <i>Gradient Boosting</i>	4 Gestos 93% Precisão
ZHANG et al. , 2018	Galaxy Watch 3 (Tizen) Amost. 100Hz	PPG	PPG e Análise de Ruidos	SVM	10 gestos 90.55% Precisão 90.73% Recall
WEN; ROJAS; DEY, , 2016	Galaxy Watch 3 (Tizen) Amost. 100Hz	IMU	Janela Deslizante (1s)	k-NN	5 Gestos 87% F1 Score
ZHAO et al. , 2019	Pulseira com 2 Sensores de PPG Amots. 100Hz	PPG	<i>Short-time energy e Dynamic Time Warping (DTW)</i>	<i>Gradient Boosting, DNN (ResNet), Transfer Learning</i>	9 Gestos 90% Precisão
LEE; CUONG; WEI. , 2017	Pulseira com 9DOF	IMU	<i>Dynamic Time-Warping</i>	CNN, SVM	6 Gestos 93.19% Precisão
LING et al., 2020	Pulseira com 3 Sensores de PPG Amots. 100Hz	PPG	PPG e Análise de Ruidos	SVM, CNN, <i>Long Short-Term Memory (LSTM)</i>	14 Gestos 96.5% Precisão
Proposta	Galaxy Watch 4 PPG Amost. 25Hz IMU Amost. 100Hz	PPG, IMU	PPG e Análise de Ruidos	RF, SVM e XgBoost	7 Gestos

Tabela 2 – Técnicas e Modelos utilizados no reconhecimento de gestos

A tabela 2 apresenta um comparativo entre as técnicas de aquisição de sinal e dos algoritmos de classificação que foram utilizados, nota-se a vasta utilização de algoritmos baseados em árvores e máquinas de vetores de suporte. Podemos observar também que as soluções atualmente implementadas usualmente propõe novos dispositivos, com uma disposição diferente dos sensores e frequências de amostragem superiores a 60Hz no caso de (SUBRAMANIAN; SAVUR; SAHIN, 2020).

Para a tarefa de segmentação, podemos citar como os métodos mais utilizados o *Dynamic Time-Warping (DTW)*, *Sliding Windows* e análise de PPG e ruídos de movimento. O algoritmo de DTW é utilizado principalmente para encontrar o alinhamento entre duas séries temporais, desta forma é possível encontrar padrões entre medições de eventos com diferentes fases e períodos. As janelas deslizantes *Sliding Windows*, dividem uma série temporal em intervalos de tempo determinados e analisam dentro desse intervalo se há ou não a presença de uma característica conhecida, tal como pontos de maior ou menor energia.

Na análise de PPG e ruídos de movimento, as abordagens apresentadas utilizam os artefatos de movimento de forma a criar balizas para determinar o início e fim dos eventos.

Por fim este trabalho irá adotar os métodos dispostos na última linha da tabela 2, um dos maiores diferenciais nessa proposta é a utilização de um sensor de fotopleismografia operando a 25Hz. O método de segmentação se baseará na análise de artefatos de movimento no sinal de PPG.

4 Desenvolvimento

A proposta desta dissertação é abordar a utilização de métodos de aprendizado de máquina e reconhecimento de padrões para realizar a identificação e a classificação de movimentos adquiridos por sensores disponíveis em *smartwatches*. Os sensores que serão considerados neste processo serão os sensores de Fotopletismografia, e os sensores de Movimento, acelerômetros e giroscópios, estes são sensores largamente utilizados em diversos dispositivos inteligentes.

4.1 Materiais

Nesta seção é apresentado o processo de definição do protocolo de execução dos gestos para a criação do banco de dados experimental para segmentação e classificação de gestos.

4.1.1 Base de Dados

Como etapas de desenvolvimento temos o processo de criação do banco de dados, onde foi definido um conjunto de gestos assim como um protocolo para a execução e coleta. Após a consolidação do banco de dados, foi realizado o pré-processamento dos dados. Nessa etapa foram aplicadas técnicas para realizar a preparação do banco de dados para aplicação nos modelos de classificação, sendo esta etapa fundamental para a definição dos atributos e características que serão usados no modelo.

4.1.2 Definição do Protocolo de Gestos

Para este projeto foi realizada a coleta dos gestos utilizando os dispositivos inteligentes na mão dominante do usuário, foram considerados 8 gestos, os quais foram executados pelos participantes 20 vezes.

Os gestos considerados nesta proposta, ver figura 14, são gestos que consideram principalmente a movimentação dos punhos e falanges, uma vez que o objetivo é avaliar a viabilidade de utilizar os dados coletados pela técnica de fotopletismografia de punho no processo de identificação de gestos.

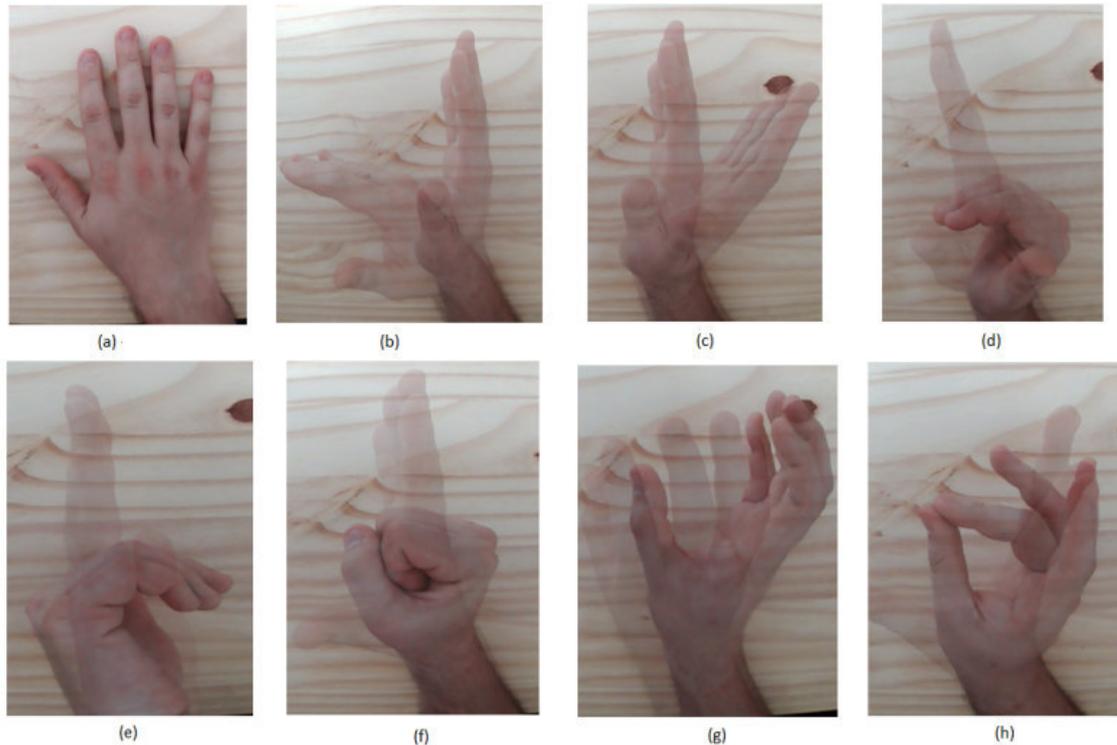


Figura 14 – Proposta de gestos

Fonte: Própria

Na figura 14, (a) representa a mão em repouso (gesto 1), (b) gesto 2, (c) gesto 3, (d) gesto 4, (e) gesto 5, (f) gesto 6, (g) gesto 7, (h) gesto 8.

Os movimentos utilizam tecidos, musculares ou fibrosos (tendões), próximos ao local em que o relógio geralmente é posicionado (Punho). A partir da ativação muscular, o corpo entende que aquela região necessita de maior volume sanguíneo desta forma sendo mais irrigada, e com isso os vasos sanguíneos apresentam variações mais significantes em sua espessura, sendo fato importante para a aquisição do sinal do sensor de PPG. A utilização da mão dominante do usuário permite um maior controle e precisão dos movimentos.

Cada gesto foi executado em janelas temporais de 3 segundos. Para o processo de coleta foi utilizado um relógio inteligente, *Galaxy Watch 4*, o mesmo é equipado com um Acelerômetro triaxial, Giroscópio triaxial, Sensor óptico de Taxa de Batimento Cardíaco (Fotopletismografia).

Os sensores de movimento realizam aquisições a 100Hz, o sensor de Fotopletismografia tem uma taxa de 25Hz.

4.2 Implementação

A figura 15 apresenta um fluxograma simplificado com as etapas do algoritmo proposto para a segmentação e reconhecimento de gestos. Os dados previamente coletados, foram pré-processados a fim de remover ruídos e artefatos de movimento para garantir uma correta segmentação dos gestos. Em um segundo passo, um algoritmo foi utilizado para identificar a janela de movimento assim como segmentar cada movimento de forma a criar um vetor específico para cada amostra, criando um banco de dados com segmentos de sinal onde estão localizados os dados relativos aos gestos executados por cada participante. Para cada segmento serão calculados atributos no domínio da frequência e do tempo para os sinais de PPG, Acelerômetros e Giroscópios. A tabela 3 apresenta uma relação de atributos os quais foram extraídos nesta fase.

Atributo	Domínio	Descrição
mean_	Tempo	Média
std_	Tempo	Desvio padrão
max_	Tempo	Valor máximo do sinal
min_	Tempo	Valor mínimo do sinal
median_	Tempo	Mediana
itq_range_	Tempo	Intervalo interquartil
neg_count_	Tempo	Quantidade de amostras na porção negativa
pos_count_	Tempo	Quantidade de amostras na porção positiva
above_mean_	Tempo	Quantidade de amostras na porção acima da média
peak_count_	Tempo	Quantidade de picos
skewness_	Tempo	Skewness
kurtosis_	Tempo	Kurtosis
energy	Tempo	Energia
meanPSD_	Frequência	Média da Densidade Espectral de Potência
stdPSD_	Frequência	Desvio Padrão da Densidade Espectral de Potência
maxPSD_	Frequência	Valor máximo da Densidade Espectral de Potência
minPSD_	Frequência	Valor mínimo da Densidade Espectral de Potência
medianPSD_	Frequência	Mediana da Densidade Espectral de Potência
entropy	Frequência	Entropia
skewnessPSD_	Frequência	Skewness da Densidade Espectral de Potência
kurtosisPSD_	Frequência	Kurtosis da Densidade Espectral de Potência

Tabela 3 – Lista de Atributos

Foi aplicado um algoritmo de seleção, que a partir de um modelo treinado com todos os atributos classifica os N atributos mais importantes para o processo de classificação. O modelo com os atributos selecionados foi retreinado a fim de avaliação de desempenho.

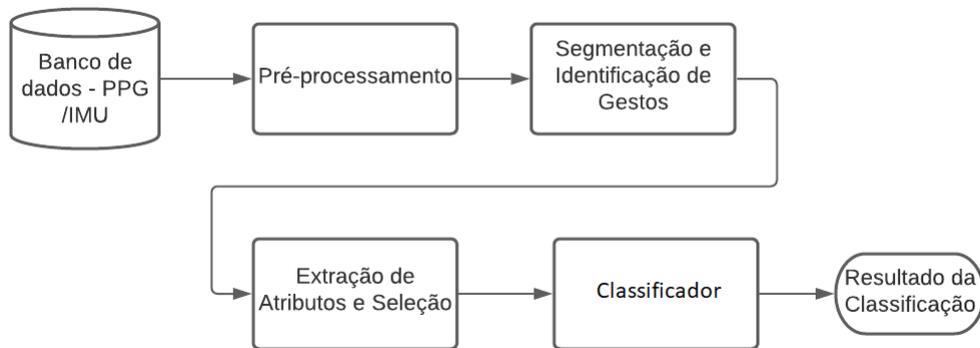


Figura 15 – Fluxograma do algoritmo para reconhecimento de gestos
 Fonte: Própria

4.2.1 Transformação dos Dados

4.2.1.1 Suavização e Normalização dos Dados

O fluxograma da figura 16 apresenta as rotinas aplicadas durante o processo de pré-processamento do sinal. A normalização e a filtragem do sinal é realizada para todos os vetores de movimento de cada participante.

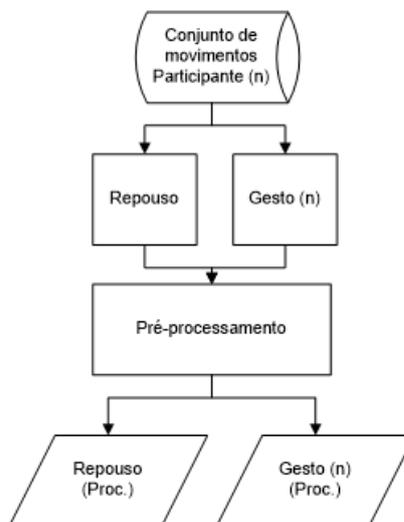


Figura 16 – Fluxograma do pré-processamento
 Fonte: Própria

A fase de pré-processamento compreende a normalização, a suavização sinal e o cálculo de gradiente. O sinal de PPG é obtido por 4 Canais, sendo esses dispostos geometricamente no relógio, de forma a garantir que ao menos um canal faça o reconhecimento de alterações no fluxo sanguíneo. Inicialmente é realizada uma média de todos os canais e o resultado por fim é normalizado utilizando a técnica de mínimo e máximo, conforme apresentado anteriormente. A normalização também foi realizada para os sinais provenientes dos acelerômetros e dos giroscópios. Para o processo de suavização foi utilizada a técnica de decomposição wavelet, para o sinal de PPG foi aplicado uma onduleta *Daubechies, db1*, e no processo de reconstrução não foram considerados os detalhes, ou seja as componentes de alta frequência. Esta técnica é amplamente utilizada no processamento de dados fisiológicos, como em (ZHANG et al., 2018) e (ZHAO et al., 2019). As figuras 17 e 18 apresentam os sinais antes e pós tratamento, nota-se que o formato original da onda não foi alterado. Devido as configurações do sensor, principalmente a frequência de aquisição, o sinal de PPG não apresenta contaminação severa de componentes de alta frequência.

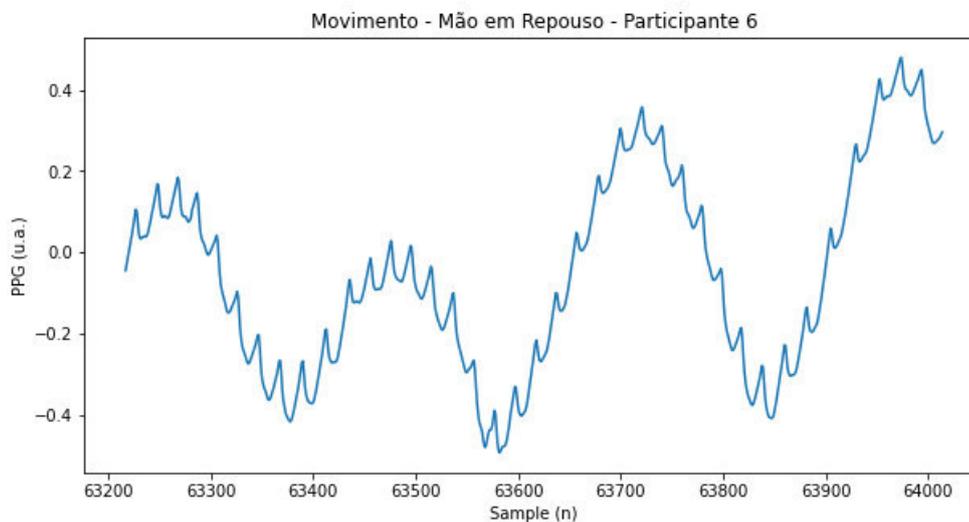


Figura 17 – Sinal de PPG - Mão em Repouso

Fonte: Própria

Por fim foi realizado o cálculo de gradiente do sinal. O gradiente é obtido usando diferenças de segunda ordem, o gradiente é a diferença entre os dois valores finais em cada extremidade da matriz.

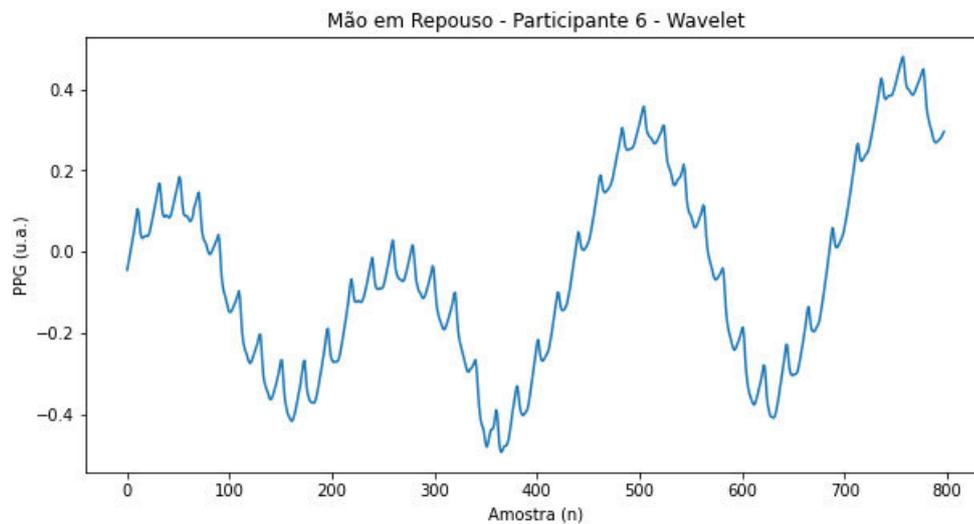


Figura 18 – Sinal de PPG - Mão em Repouso - Após Wavelet

Fonte: Própria

4.2.1.2 Segmentação dos gestos

A extração dos segmentos dos gestos é o primeiro passo no processo de reconhecimento e classificação dos movimentos. Durante o protocolo cada participante realizou a repetição de um movimento 20 vezes em intervalos de 3 segundos. Para que o sinal possa ser utilizado pelo classificador se faz necessário o processamento e a segmentação de cada gesto. A figura 19, apresenta uma breve descrição dos passos empregados durante essa etapa.



Figura 19 – Determinação de Características para Segmentação

Fonte: Própria

Cada participante apresenta diferentes condições fisiológicas, dentre elas podemos destacar como fator principal a frequência cardíaca. Esta característica pode ser facilmente identificada por meio do sinal proveniente da mão em repouso. Como pode ser visto na figura 20, após o pré-processamento, temos um vetor com o gradiente de sinal de PPG.

Podemos observar a existência de um sinal periódico presente no gradiente de sinal, que é frequência cardíaca do usuário em situação de repouso. Na etapa de identificação de limiar é realizada a identificação de todos esses picos em uma janela temporal de 10 segundos, ou seja uma janela de 10 segundos percorre todo sinal, identificando todos os pontos máximos. A partir dessa identificação podemos extrair um valor médio da amplitude máxima do sinal (HR). Além

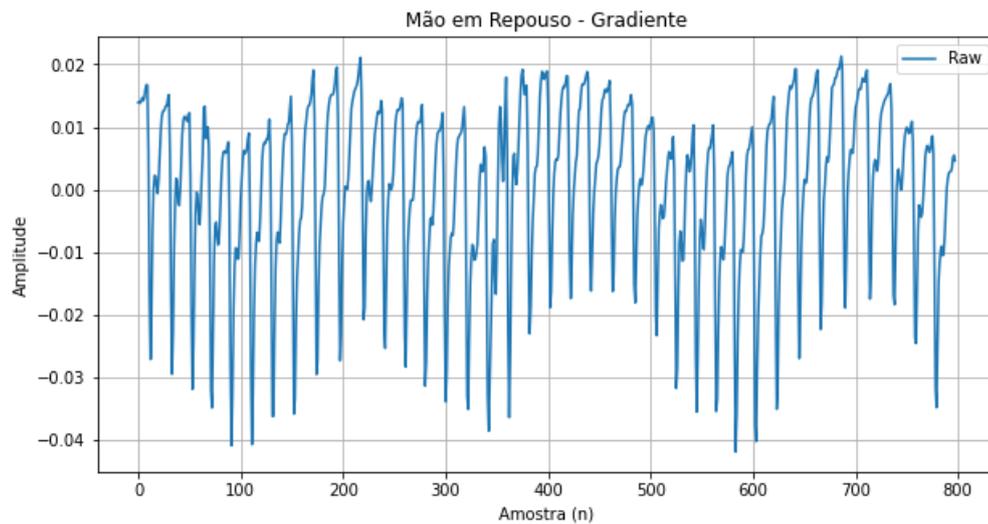


Figura 20 – Sinal de PPG - Mão em repouso - Gradiente

Fonte: Própria

da amplitude do sinal, outro fator importante é a distância entre picos, ou seja a distância entre um batimento e outro. O valor de (PPD), *peak-to-peak distance*, é obtido através das médias de distância entre dois picos consecutivos. A Figura 21 apresenta o resultado da identificação dos picos de HR.

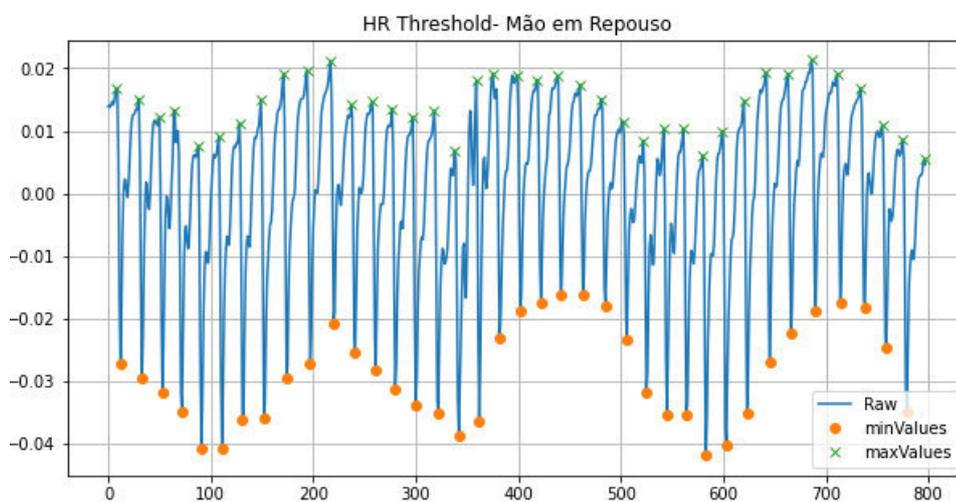


Figura 21 – Sinal de PPG - Mão em repouso - Identificação de Picos

Fonte: Própria

As informações de amplitude e distância média entre as pulsações da mão em repouso, foram utilizadas para buscar no vetor de movimentos, os pontos que representam o início e o fim de cada gesto.

O fluxograma da figura 22, resume as etapas realizadas durante o processo de segmentação de gestos.

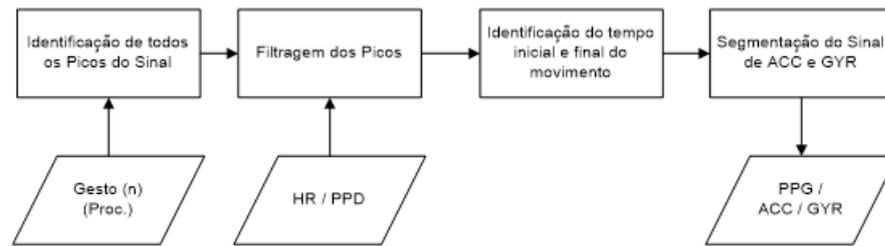


Figura 22 – Segmentação de Gestos

Fonte: Própria

Para cada vetor de movimentos, foi realizada a identificação de todos os picos, assim como foi feito na identificação do HR e PPD, entretanto podemos notar na Figura 23 um comportamento diferente do sinal da mão em repouso. Essa diferença se dá principalmente devido às variações na quantidade de luz refletida pelos tecidos. De acordo com diferentes movimentos e os níveis de ativação muscular o sinal de PPG apresentará distorções, este ruído é denominado como artefato de movimento.

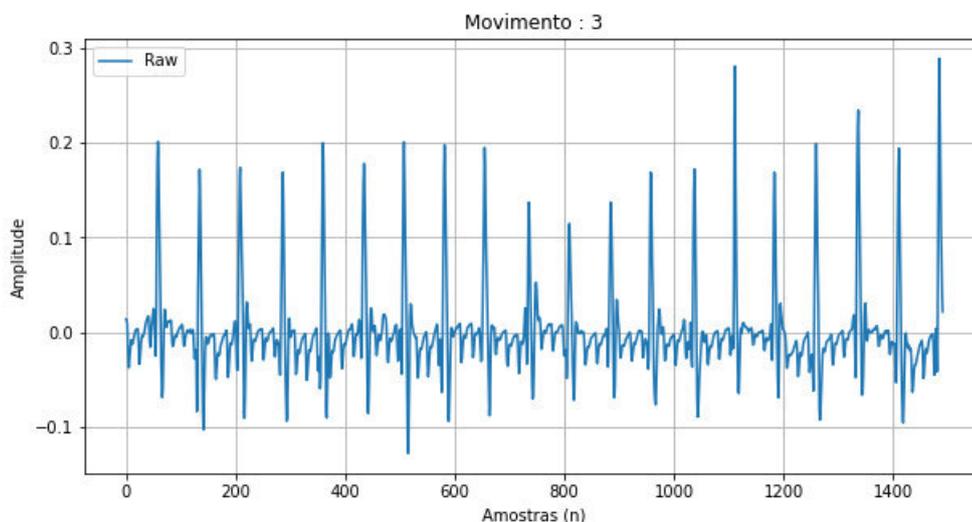


Figura 23 – Sinal de PPG - Movimento

Fonte: Própria

A partir da identificação de todos os picos e comparando o comportamento de cada pico, amplitude e distância entre picos, com os dados obtidos na etapa anterior, temos uma identificação do início e fim de cada gesto, ver figura 24.

Os pontos verdes da figura 24, representam os picos selecionados como pontos de separação entre o começo e o término de cada gesto. Cada ponto apresenta um índice temporal o qual é utilizado para localizar em um buffer, os dados de acelerômetro e giroscópio para cada intervalo. O Apêndice A.1 e A.2 apresenta um pseudocódigo para implementação do algoritmo

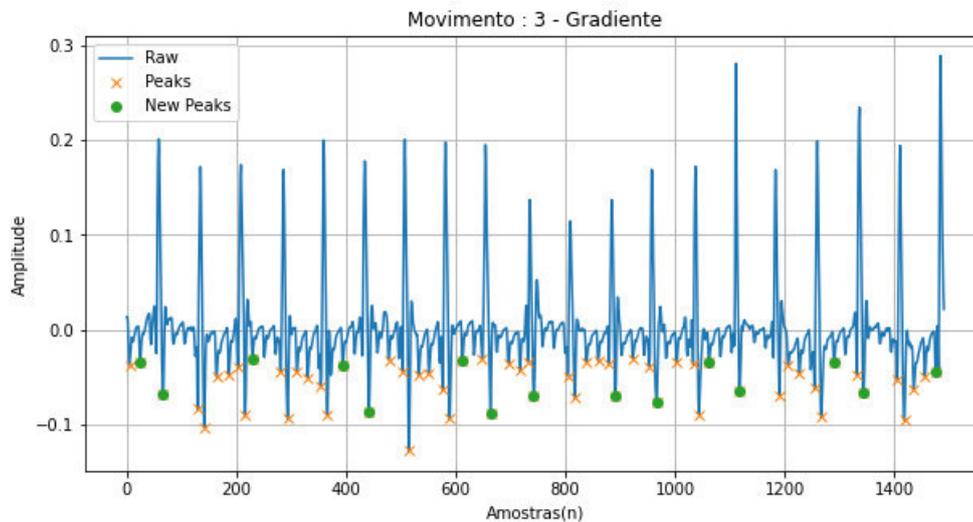


Figura 24 – Sinal de PPG - Movimento - Segmentado

Fonte: Própria

de segmentação.

4.2.1.3 Extração de Atributos

Nesta etapa, os vetores contendo informações do PPG, Acelerômetros e Giroscópios para cada gesto e para cada participante são utilizados para o cálculo dos atributos listados na tabela 3. A extração de 21 características foi realizada para cada vetor separadamente, ou seja, para o PPG e para cada componente x, y e z do acelerômetro e giroscópio resultando em um total de 147 atributos para cada gesto.

4.2.2 Treinamento dos Modelos

No processo de treinamento dos modelos, de Florestas Aleatórias, Máquinas de Vetores de Suporte e XGBoost, os parâmetros ótimos para os classificadores foram obtidos utilizando um método de varredura para diversos parâmetros iniciais dos classificadores. Esse método realiza o treinamento e a verificação dos resultados para cada conjunto de parâmetros que foram determinados para a fase de treinamento. Durante o processo de treinamento, foi utilizado o método de validação cruzada com 5 divisões, para avaliação do desempenho dos respectivos

modelos nesta etapa.

4.2.2.1 Florestas Aleatórias

Para o processo de treinamento utilizando o algoritmo de florestas aleatórias, a varredura levou em consideração a otimização dos seguintes parâmetros, ver tabela 4. Onde *min_samples_leaf* representa o número mínimo de amostras necessárias para estar em um nó, o número estimadores e os critérios, ou seja a função que avalia a qualidade das divisões.

Parâmetro	Valores Propostos	Valores Escolhidos
Critério	Gini, Entropia	Gini
<i>min_samples_leaf</i>	1,2,4,8	2
<i>n_estimators</i>	10,50,100,1000	100

Tabela 4 – Parâmetros da Floresta Aleatória

Após o processo de treinamento e seleção de atributos, temos que o melhor classificador foi obtido utilizando 27 atributos, apresentados na tabela 7.

4.2.2.2 Máquina de vetores de suporte

Foi aplicado um processo de varredura no treinamento do modelo baseado em máquina de vetores de suporte. A tabela 5 apresenta os valores propostos para o processo de otimização e os valores escolhidos após treinamento.

Parâmetro	Valores Propostos	Valores Escolhidos
C	1,1000,2000,5000,10000,100000	5000
Kernel	Polinomial, RBF e Sigmoidal	Polinomial
Grau do Classificador	2,3,4	2
Gama	0.1, 1, 10	0.1

Tabela 5 – Parâmetros da Máquina de Vetores de Suporte

Os atributos que foram selecionados como mais relevantes no processo de classificação totalizaram 31 atributos, são eles, ver tabela 7:

Na tabela 5, C representa o parâmetro de regularização, Kernel onde podemos avaliar a utilização de modelo polinomial, RBF e Sigmoidal, o Grau do Classificador e por fim o valor de Gama. É válido ressaltar que o valor de gama só é aplicável quando o Kernel utilizado é polinomial. No processo de classificação utilizando máquinas de vetores de suporte, foram selecionados os 31 atributos mais relevantes, ver Tabela 7.

4.2.2.3 *Extreme Gradient Boosting* (XGBoost)

O modelo baseado em XGBoost foi treinado de forma similar aos modelos de Florestas aleatórias e Máquinas de Vetores de Suporte. Os hiperparâmetros selecionados durante o processo de treinamento são apresentados na tabela 6.

Parâmetro	Valores Propostos	Valores Escolhidos
Profundidade Máxima	3,4,5,6,7	6
Taxa de Aprendizado	0.3, 0.2, 0.1, 0.05	0.3
Gamma	0, 0.1, 0.3, 0.5, 0.7, 0.9, 1	0

Tabela 6 – Parâmetros XGBoost

A taxa de aprendizado especifica a rapidez com que o modelo se ajusta aos erros residuais. A profundidade máxima especifica quão fundo os nós de decisão da árvore podem ir, e γ são utilizados para ajustar a regularização realizada pelo algoritmo de XGBoost. Para um total de 7 gestos o modelo baseado em XGBoost utilizou 24 Atributos, ver tabela 7.

Da seleção de atributos aplicada nos três modelos utilizados, é possível observar os dados provenientes dos sensores de movimento, acelerômetro e giroscópio, tiveram maior relevância no processo de classificação. Isso já era esperado, uma vez que as frequências de aquisição destes sinais são superiores quando comparados com a aquisição dos dados da fotopletiografia.

4.2.2.4 Teste do Modelo

Para que possamos visualizar a capacidade de generalização dos dados, uma parte dos dados do banco, foi separada somente para testes, ou seja essas amostras não foram utilizadas nas

Classificador	Atributos
Florestas Aleatórias	std_y, std_z, energy_z, std_gyrx, max_gyrx, min_gyrx, median_abv_gyrx, itq_range_gyrx, skewness_gyrx, energy_gyrx, stdPSD_gyrx, maxPSD_gyrx, entroS_gyrx, std_gyry, max_gyry, min_gyry, skewness_gyry, energy_gyry, stdPSD_gyry, maxPSD_gyry, entroS_gyry, std_gyrz, max_gyrz, min_gyrz, skewness_gyrz, meanPSD_gyrz, medianPSD_gyrz
SVM	std_x, itq_range_x, std_y, median_abv_y, itq_range_y, stdPSD_y, maxPSD_y, std_z, median_abv_z, itq_range_z, energy_z, stdPSD_z, maxPSD_z, std_gyrx, max_gyrx, min_gyrx, skewness_gyrx, energy_gyrx, stdPSD_gyrx, maxPSD_gyrx, entroS_gyrx, std_gyry, max_gyry, min_gyry, energy_gyry, stdPSD_gyry, maxPSD_gyry, entroS_gyry, std_gyrz, min_gyrz, medianPSD_gyrz
XGBoost	stdPSD, maxPSD, pos_count_x, max_y, skewness_psd_y, std_gyrx, max_gyrx, min_gyrx, itq_range_gyrx, energy_gyrx, stdPSD_gyrx, maxPSD_gyrx, std_gyry, max_gyry, min_gyry, skewness_gyry, maxPSD_gyry, entroS_gyry, std_gyrz, min_gyrz, skewness_gyrz, meanPSD_gyrz, stdPSD_gyrz, minPSD_gyrz

Tabela 7 – Parâmetros selecionados para os classificadores RF e SVM

etapas de treinamento e validação. O conjunto de dados para teste representa 30% do montante total de amostras incluindo todos os gestos de todos os participantes.

Para avaliação de desempenho utilizaremos métricas, precisão, revocação (*recall*) e F1-score. A precisão e a revocação são métricas de desempenho que se aplicam a dados obtidos a partir de um espaço amostral. A precisão representa a fracção de instâncias classificadas de forma correta dentre as instâncias disponíveis. A revocação é a fracção de instâncias que foram identificadas. Por fim o F1-Score é a exatidão do modelo, este valor é calculado a partir da precisão e da recuperação do teste.

5 Resultados

Nesta seção, serão abordados os resultados obtidos do desenvolvimento deste projeto, esta seção será dividida em resultados do processo de segmentação de gestos e o processo de classificação.

5.1 Segmentação dos Gestos

O algoritmo de segmentação de gestos, conforme apresentado anteriormente, é responsável por identificar dentro de uma janela temporal, a existência de um movimento. Esse processo se dá pela verificação do limiar de frequência cardíaca e pela distância média entre pontos de cada indivíduo, previamente calculados por meio de técnicas de fotopleletismografia.

Do banco de dados coletado originalmente de 20 indivíduos, dos quais puderam ser aproveitados os dados de 16 participantes, totalizando 2.240 gestos. Os dados de 4 participantes foram descartados devido à inconsistência na gravação das series temporais, essas inconsistências se apresentam principalmente como perda de sinal durante o processo de gravação do protocolo.

A tabela 8 apresenta os resultados do processo de segmentação dos gestos.

Descrição	Quantidade Identificada	Diferença	Percentual
Gesto 2	254	66	79%
Gesto 3	303	17	95%
Gesto 4	249	71	78%
Gesto 5	259	61	81%
Gesto 6	307	13	96%
Gesto 7	311	9	97%
Gesto 8	268	52	84%
Total	1951	289	87%

Tabela 8 – Resultado - Segmentação

O algoritmo de segmentação obteve taxas de segmentação superiores a 78% para todos os gestos coletados. Tendo alcançado o maior índice para o gesto número 7, que representa

	Qtd. Gestos	Precisão	Recall	F1-Score
F.A	5	0.922	0.93	0.922
	6	0.875	0.875	0.877
	7	0.82	0.821	0.819
SVM	5	0.918	0.916	0.916
	6	0.883	0.882	0.882
	7	0.801	0.802	0.801
XgBoost	5	0.929	0.928	0.928
	6	0.895	0.891	0.892
	7	0.8377	0.832	0.833

Tabela 9 – Comparação Classificadores

a abertura e o fechamento da mão. Isso demonstra que quanto maior o nível de ativação dos músculos e tendões, maior será o ruído no sinal de PPG e consequentemente será mais evidente ao algoritmo o que são efetivos.

O aproveitamento do algoritmo em segmentar os gestos foi de 87%, isso significa que de todo o banco de dados previsto, 87% dos dados foram segmentados em gestos e os atributos puderam ser extraídos para cada segmento.

5.2 Classificação dos gestos

O processo de treinamento de modelos foi realizado considerando três tipos de classificadores, as florestas aleatórias, as máquinas de vetores de suporte e o método baseado em árvore XgBoost. A Tabela 9 apresentam um resumo dos resultados obtidos pelos algoritmos de classificação propostos, para diferentes combinações de gestos nas etapas de treinamento.

É possível observar que a redução de gestos no conjunto de classificação resulta em um aumento das métricas do algoritmo, tendo os classificadores alcançado valores de aproximadamente 93% de precisão considerando um conjunto de apenas 5 gestos e um mínimo de 80% para todo o conjunto de testes. Podemos notar também um desempenho superior das florestas aleatórias e para o XgBoost, ambas alcançaram precisão de 82% e 83.7% e *Recall* de 82.1% e 83.2% para um conjunto de 7 gestos, maior que as máquinas de vetores de suporte para o mesmo conjunto.

O modelo utilizando XgBoost teve um desempenho superior aos demais classificadores

em todos os conjuntos de gestos, sendo que no conjunto considerando todos os gestos propostos neste trabalho ele foi 1.37% superior ao segundo melhor classificador, o algoritmo baseado em Florestas aleatórias. A figura 25 apresenta o desempenho de de forma detalhada um comparativo entre todos os classificadores.

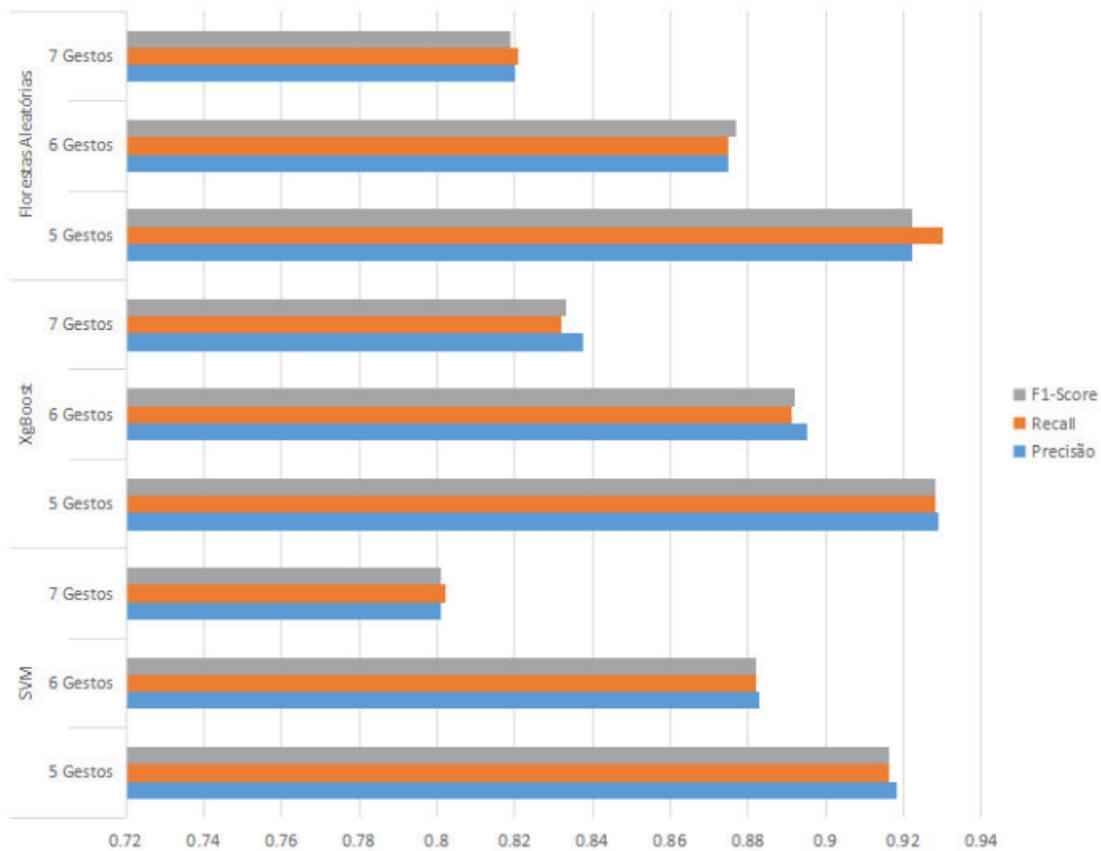


Figura 25 – Comparação Classificadores

Fonte: Própria

A figura 26 apresenta o F1 Score obtidas da validação cruzada do treinamento dos modelos de SVM e FA, nestes testes foram avaliados o impacto do tamanho do conjunto de treinamento, onde podemos observar que para ambos os classificadores, temos um aumento gradual dos scores à medida que aumenta do tamanho do conjunto de treino. A Figura 27 apresenta o tempo necessário para treinamento dos modelos de acordo com a quantidade de amostras no conjunto de treinamento para cada modelo experimentado. Podemos observar que o desempenho de algoritmos baseados em SVM, apresentam uma redução de aproximadamente 50% no tempo de treinamento.

A figura 28 apresenta o resultado da mesma avaliação para o modelo de XgBoost, tal como nas florestas aleatórias, podemos verificar que o modelo ainda apresenta um perfil

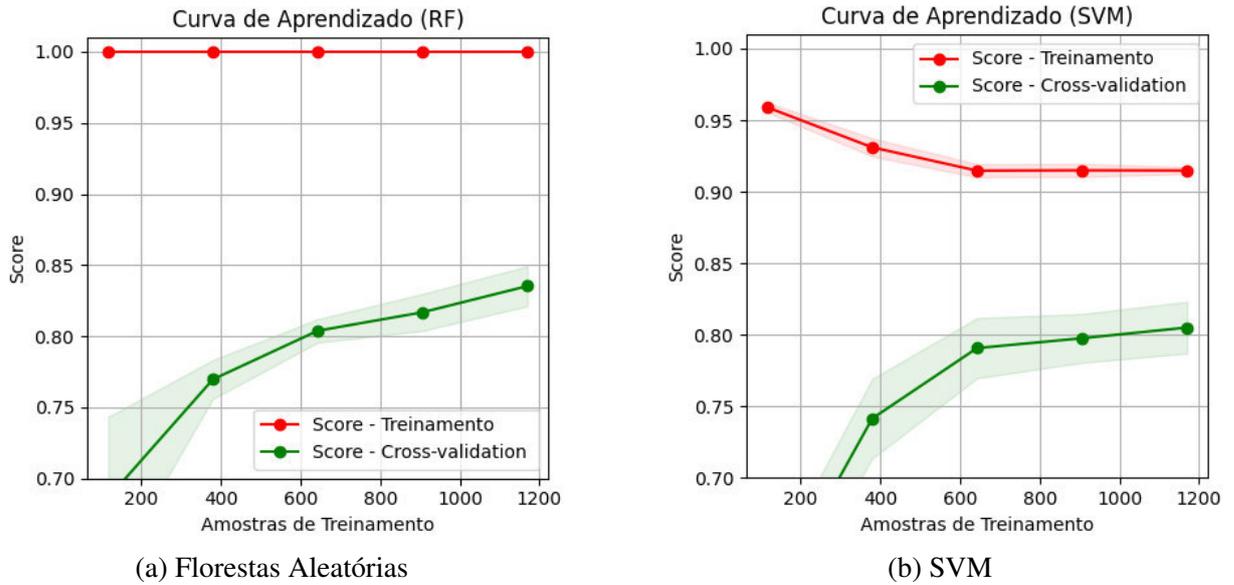


Figura 26 – Curvas de aprendizado

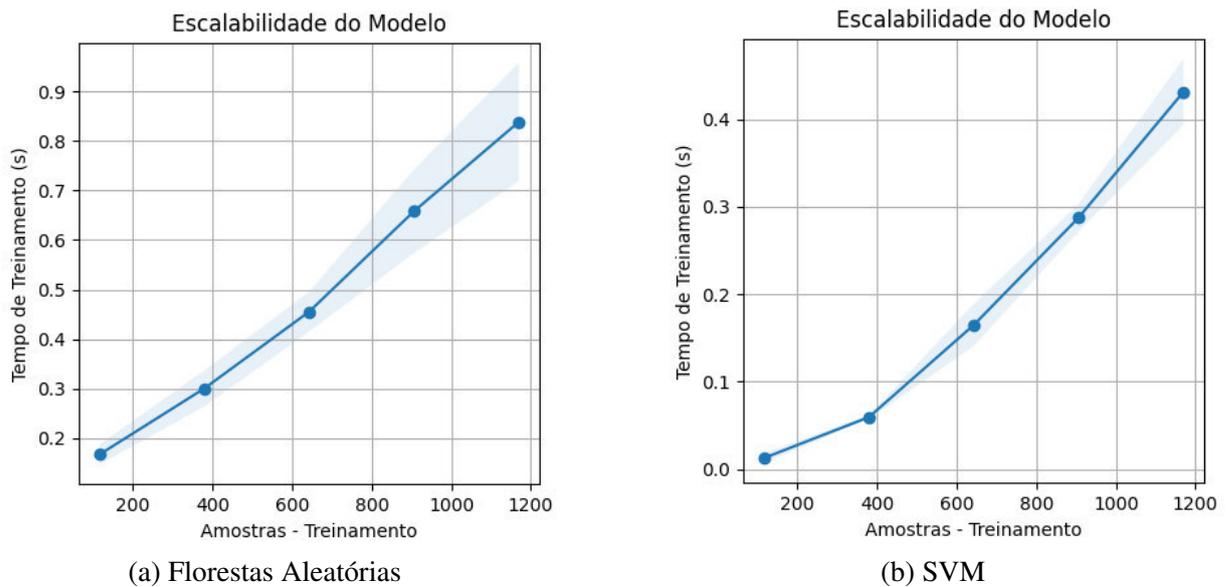


Figura 27 – Escalabilidade

ascendente para o número de elementos no conjunto de treinamento. Quando comparamos o tempo de treinamento do algoritmo de XgBoost com os modelos de Florestas Aleatórias e Máquinas de Vetores de Suporte, notamos que o tempo de treinamento se torna 10 vezes maior, resultado esperado uma vez que os critérios de otimização tornam o algoritmo com maior custo computacional.

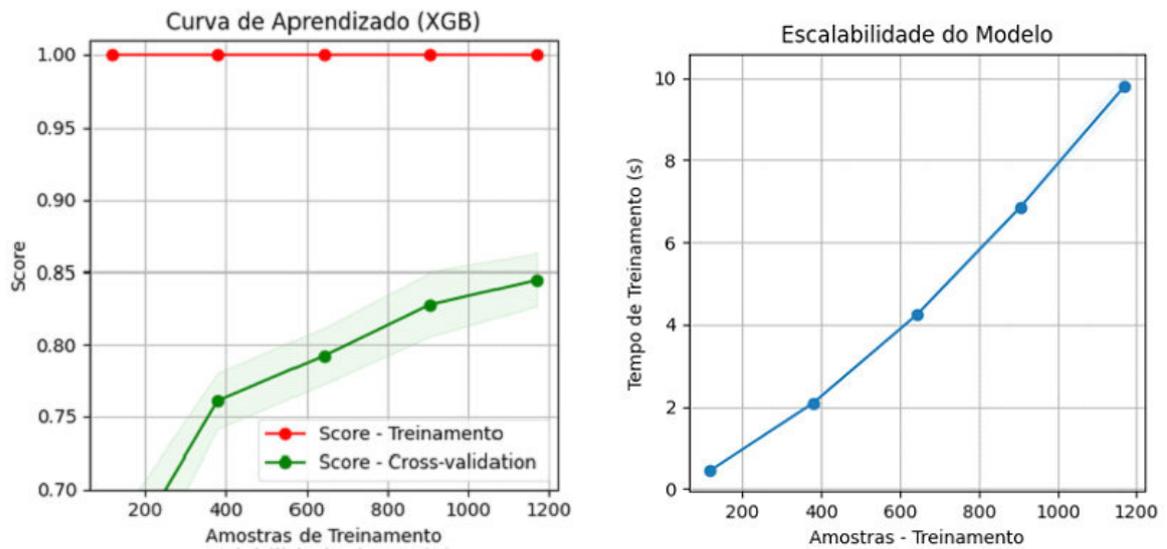


Figura 28 – Curva de Aprendizado e Escalabilidade XGBoost

Fonte: Própria

5.2.1 Florestas Aleatórias

As florestas aleatórias apresentaram uma capacidade de classificação superior ao das máquinas de vetores de suporte para conjuntos com 5 e 8 gestos, para visualizarmos a capacidade de generalização deste modelo, aplicamos os dados de teste previamente segmentados do conjunto inicial de dados. As tabelas 10, 11, 12 apresentam os resultados detalhados do desempenho do classificador para este conjunto.

Descrição	Precisão	Recall	F1 - Score	Amostras
Gesto 2	0,96	0,96	0,96	70
Gesto 3	0,95	0,94	0,94	79
Gesto 4	0,76	0,80	0,78	60
Gesto 5	0,71	0,60	0,65	65
Gesto 6	0,95	0,99	0,97	77
Gesto 7	0,72	0,72	0,72	80
Gesto 8	0,69	0,74	0,71	57
Média	0,82	0,82	0,82	-

Tabela 10 – Resultado - RF - 7 Gestos

Na tabela 10 podemos observar que o gesto que apresenta maior taxa de erro é o Gesto 5 alcançando uma precisão de 0.71 e uma revocação de 0.65. Observando a matriz de confusão da figura 30 podemos observar que o gesto 5 e o gesto 7 apresentam conflitos, sendo que 26% dos Movimentos definidos realmente como 5 foi classificado como movimento 7.

Descrição	Precisão	Recall	F1 - Score	Amostras
Gesto 2	0,92	0,96	0,94	70
Gesto 3	0,91	0,91	0,91	74
Gesto 4	0,85	0,8	0,83	66
Gesto 6	0,94	0,97	0,96	77
Gesto 7	0,9	0,83	0,87	78
Gesto 8	0,73	0,78	0,75	58
Média	0,875	0,875	0,876	-

Tabela 11 – Resultado - RF - 6 Gestos

Fazendo a exclusão do movimento 5 dos conjuntos de treinamento e de teste, temos um incremento da precisão geral do classificador, ver tabela 11, tendo o resultado médio um acréscimo de 5%. Na matriz de confusão, ver figura 31 podemos notar que com a remoção do gesto conflitante a taxa de classificação correta do movimento 7 subiu aproximadamente 10%. Por fim, a caráter de análise, removendo o gesto com menor taxa de precisão e de revocação da fase de treinamento e teste, temos um aumento da precisão geral do classificador para 0.922 e revocação para 0.930, ver tabela 12.

Descrição	Precisão	Recall	F1 - Score	Amostras
Gesto 2	0,88	0,95	0,91	60
Gesto 3	0,98	0,92	0,95	87
Gesto 4	0,86	0,93	0,89	55
Gesto 6	0,96	0,99	0,97	76
Gesto 7	0,93	0,86	0,89	78
Média	0,922	0,93	0,922	-

Tabela 12 – Resultado - RF - 5 Gestos

A figura 29 demonstra que para alcançar a precisão de 83% para o conjunto de 7 gestos a configuração ótima de atributos é de 27 atributos. Podemos notar também que o aumento do número de atributos a partir deste valor não irá aumentar a precisão do modelo.

5.2.2 Máquinas de Vetores de Suporte

O algoritmo utilizando máquinas de vetores de suporte obteve um desempenho inferior no processo de qualificação envolvendo 7 e 5 gestos quando comparado aos modelos de Florestas Aleatórias e XgBoost, entretanto os resultados apresentados no conjunto de 6 gestos foram 0,8%

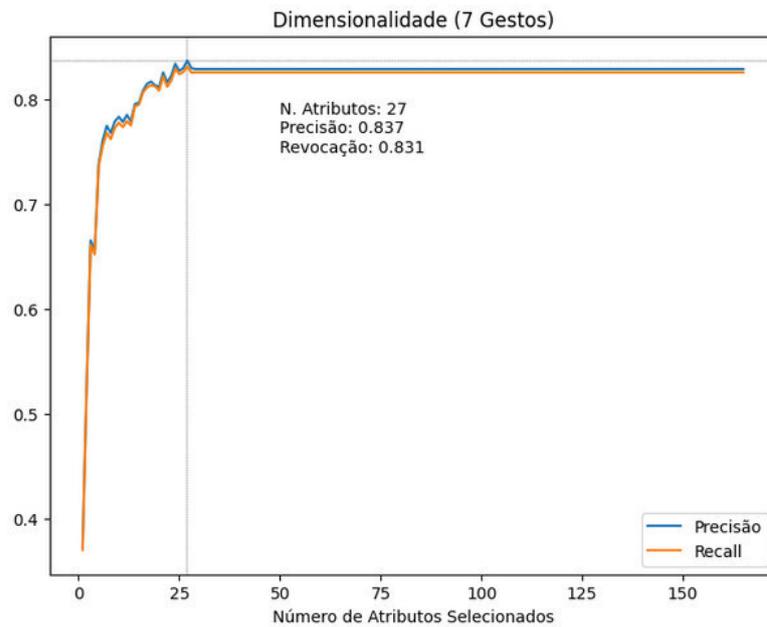


Figura 29 – Redução de Dimensionalidade - 7 Gestos - Florestas Aleatórias

Fonte: Própria

superiores aos alcançados pelas Florestas Aleatórias.

As tabelas 13, 14 e 15 apresentam os resultados para os classificadores considerando respectivamente 7, 6 e 5 gestos. De forma parecida com o método de florestas aleatórias, o gesto 5 apresenta conflitos com o gesto 7, a sua remoção resulta em um aumento de 8% na precisão. As figuras 32 e 33 apresentam a matriz de confusão dos classificadores florestas aleatórias com 6 gestos e máquinas de vetores de suporte com 7 gestos respectivamente.

Descrição	Precisão	Recall	F1 - Score	Amostras
Gesto 2	0,89	0,96	0,92	70
Gesto 3	0,93	0,94	0,93	79
Gesto 4	0,8	0,78	0,79	60
Gesto 5	0,57	0,58	0,58	65
Gesto 6	0,95	0,96	0,95	77
Gesto 7	0,76	0,64	0,69	80
Gesto 8	0,65	0,7	0,67	57
Média	0,79	0,79	0,79	-

Tabela 13 – Resultado - SVM - 7 Gestos

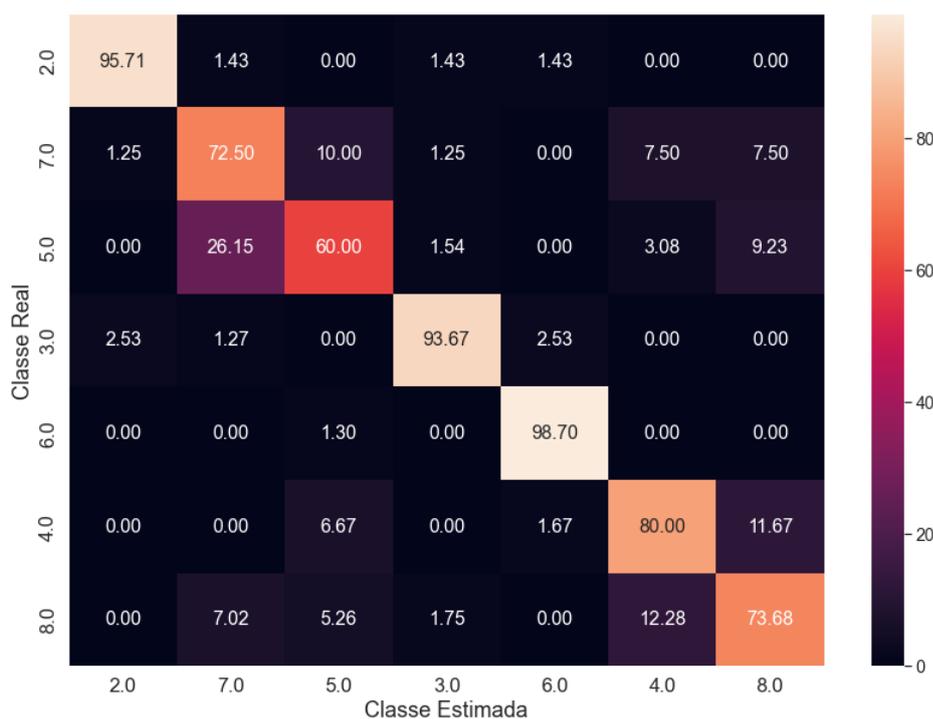


Figura 30 – Matriz de Confusão - 7 Gestos - Florestas Aleatórias
 Fonte: Própria

Descrição	Precisão	Revocação	F1 - Score	Amostras
Gesto 2	0,96	0,94	0,95	70
Gesto 3	0,87	0,91	0,89	74
Gesto 4	0,82	0,82	0,82	66
Gesto 6	0,96	0,96	0,96	77
Gesto 7	0,91	0,79	0,85	78
Gesto 8	0,71	0,81	0,76	58
Média	0,87	0,87	0,871	-

Tabela 14 – Resultado - SVM - 6 Gestos

5.2.3 Extreme Gradient Boosting (XGBoost)

O modelo utilizando baseado em árvores com gradiente descendente XgBoost alcançou as melhores métricas quando comparado com os demais classificadores. Em destaque este algoritmo alcançou um resultado de 83.7% de precisão no conjunto de sete gestos, resultado este aproximadamente 1.7% superior ao classificador usando Florestas Aleatórias. As tabelas 16, 17 e 18 apresentam os resultados de classificação do modelo em conjuntos de 7, 6 e 5 gestos. Podemos observar uma diferença da ordem de 5% na classificação do gesto número 5 quando comparado com os classificadores usando F.A e SVM.

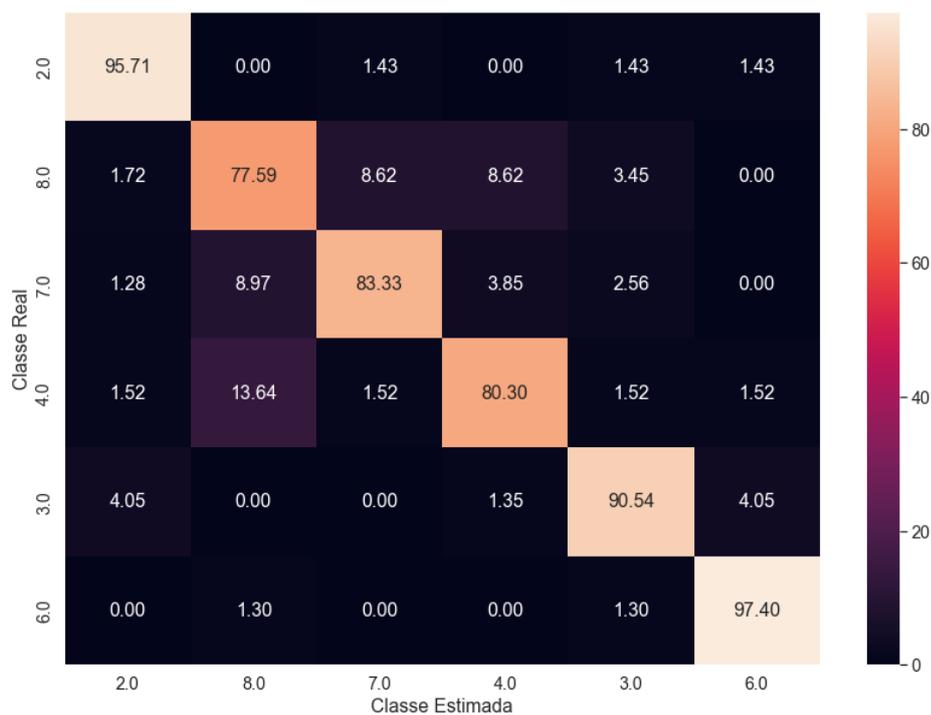


Figura 31 – Matriz de Confusão - 6 Gestos - Florestas Aleatórias
 Fonte: Própria

Descrição	Precisão	Revocação	F1 - Score	Amostras
Gesto 2	0,92	0,95	0,93	60
Gesto 3	0,96	0,91	0,93	87
Gesto 4	0,81	0,93	0,86	55
Gesto 6	0,97	0,99	0,98	76
Gesto 7	0,94	0,87	0,91	78
Média	0,92	0,93	0,922	-

Tabela 15 – Resultado - SVM - 5 Gestos

Nas figuras 34, 35 e 36 temos a matriz de confusão com os resultados da classificação do conjunto de teste, notamos comportamento similar aos demais classificadores, com a remoção dos gestos conflitantes, temos um aumento da precisão e da revocação do modelo.

5.3 Discussão

Comparando os resultados obtidos nesta pesquisa com as soluções já apresentadas anteriormente na literatura, ver tabela 2, podemos observar que o classificador com 5 gestos teve um desempenho superior quando comparado a técnica apresentada por (SUBRAMANIAN;

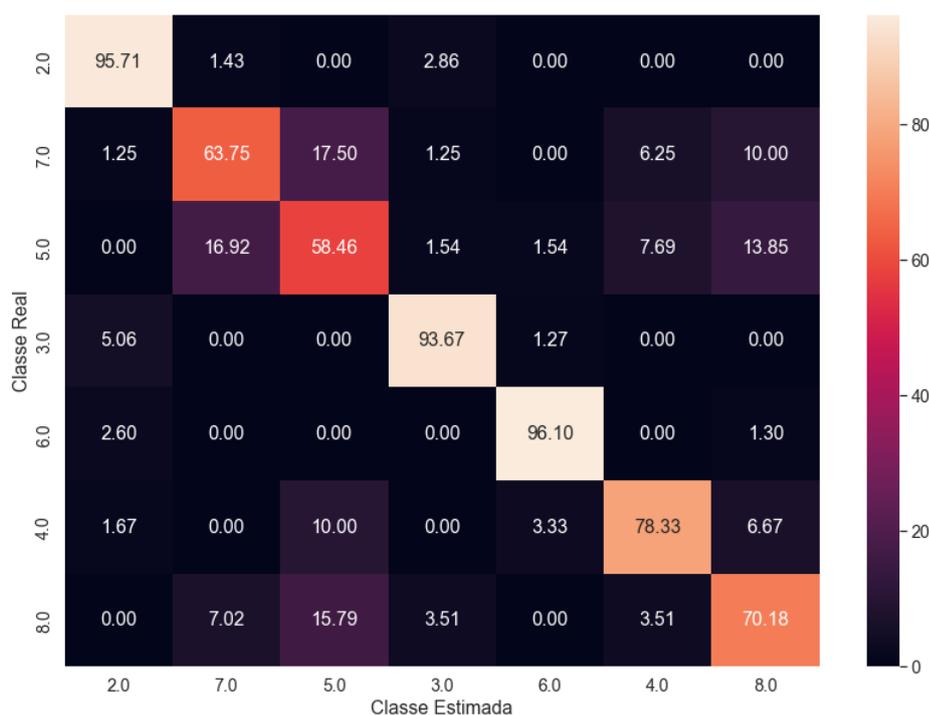


Figura 32 – Matriz de Confusão - 7 Gestos - SVM

Fonte: Própria

Descrição	Precisão	Revocação	F1 - Score	Amostras
Gesto 2	0,89	0,95	0,92	58
Gesto 3	0,96	0,94	0,95	93
Gesto 4	0,81	0,81	0,81	57
Gesto 5	0,75	0,68	0,71	68
Gesto 6	0,97	0,98	0,98	64
Gesto 7	0,76	0,76	0,76	80
Gesto 8	0,78	0,82	0,8	68
Média	0,85	0,85	0,85	-

Tabela 16 – Resultado - XGBoost - 7 Gestos

SAVUR; SAHIN, 2020), que utilizou uma pulseira equipada com 3 sensores PPG coletando amostras em 60Hz e teve uma precisão de 93% com um conjunto de 4 gestos.

Wen et ai. (WEN; ROJAS; DEY, 2016) apresenta uma proposta semelhante a proposta deste projeto apresentado a utilização de um SmartWatch coletando informações dos sensores de movimento a uma taxa de amostragem de 100Hz, o resultado obtido para um conjunto de 6 gestos foi de 87%, neste projeto o conjunto de 6 gestos obteve uma precisão de 89.5% com o XgBoost, 88.3% com o SVM e 87.5% com as floretas aleatórias.

Podemos observar também que os resultados deste projeto quando comparados com o

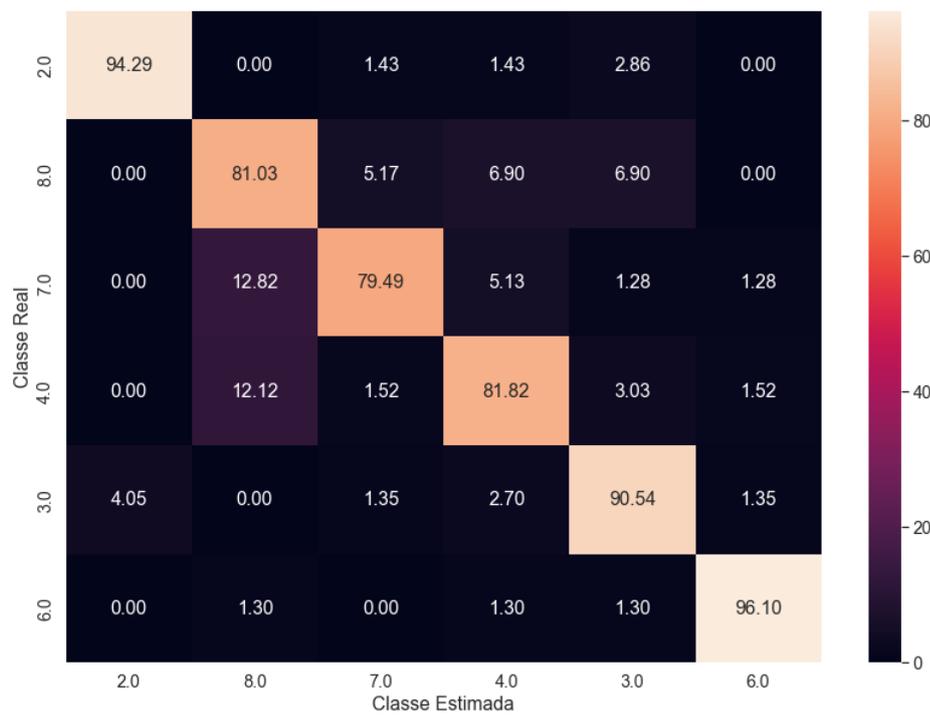


Figura 33 – Matriz de Confusão - 6 Gestos - SVM

Fonte: Própria

Descrição	Precisão	Revocação	F1 - Score	Amostras
Gesto 2	0,9	0,97	0,93	59
Gesto 3	0,98	0,93	0,95	89
Gesto 4	0,83	0,85	0,84	59
Gesto 6	0,97	0,99	0,98	71
Gesto 7	0,84	0,88	0,86	77
Gesto 8	0,82	0,75	0,78	68
Média	0,89	0,895	0,89	-

Tabela 17 – Resultado - XGBoost - 6 Gestos

trabalho de Lee e Cuong (LEE; CUONG; WEI, 2017), para um conjunto de 6 gestos apesar de inferiores, demonstram que mesmo com dispositivos com uma menor quantidade de sensores e utilizando algoritmos de aprendizado máquina de menor complexidade é possível alcançar métricas aproximadas.

Apesar dos resultados promissores com os conjuntos de até 7 gestos, há limitações em detrimento da taxa de amostragem do sensor de PPG. (ZHAO et al., 2019) e (LING et al., 2021) mencionam que há chances de melhoria dos resultados se utilizadas taxas mais altas de amostragem de PPG, o que pode proporcionar um sinal com mais informações tanto para a etapa de segmentação como para a extração de atributos. Podemos notar também os resultados

Descrição	Precisão	Revocação	F1 - Score	Amostras
Gesto 2	0,89	0,95	0,92	57
Gesto 3	0,95	0,96	0,96	79
Gesto 4	0,89	0,9	0,9	62
Gesto 6	0,99	0,99	0,99	83
Gesto 7	0,91	0,84	0,87	75
Média	0,926	0,928	0,928	-

Tabela 18 – Resultado - XGBoost - 5 Gestos

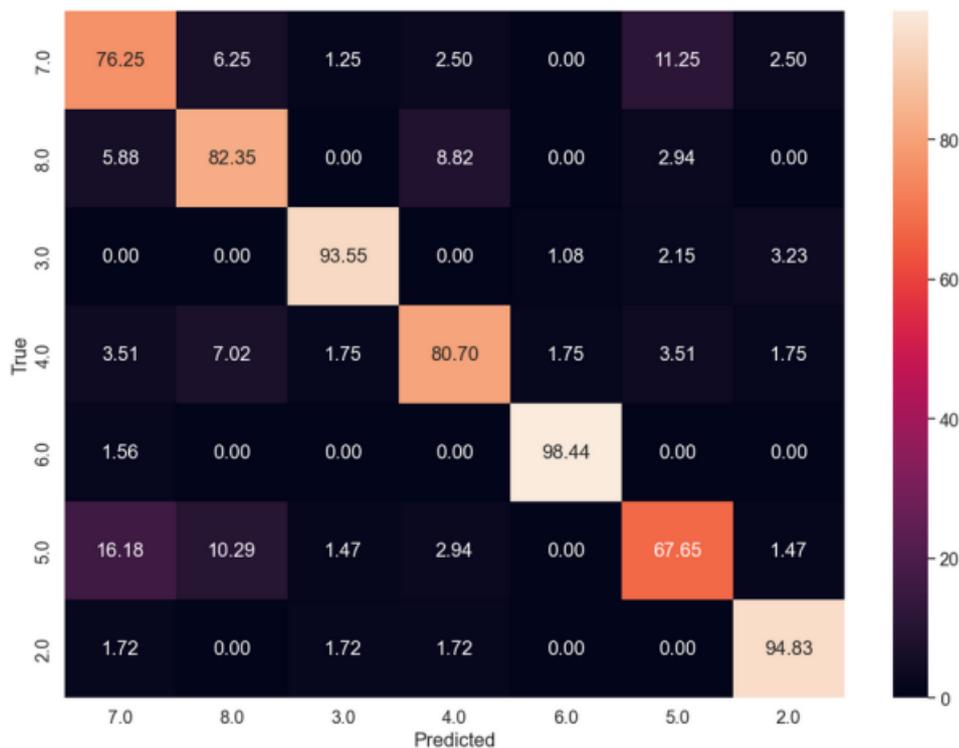


Figura 34 – Matriz de Confusão - 7 Gestos - XGBoost

Fonte: Própria

positivos em ambos trabalhos quanto a utilização de *Deep Learning*, técnicas mais avançadas de como por exemplo *Long Short-Term Memory (LSTM)* aumentam a capacidade de classificação de conjunto de gestos.

Entretanto devemos ressaltar que a proposta deste trabalho é a viabilidade de utilização de dispositivos vestíveis, tal como os relógios inteligentes para o a execução de segmentação e classificação de gestos. Em dispositivos deste tipo existem limitações de processamento que podem inviabilizar a implementação de técnicas mais complexas. A proposta apresentada por essa pesquisa considera soluções que tenham o menor impacto seja para o processamento ou para o consumo de energia do dispositivo.

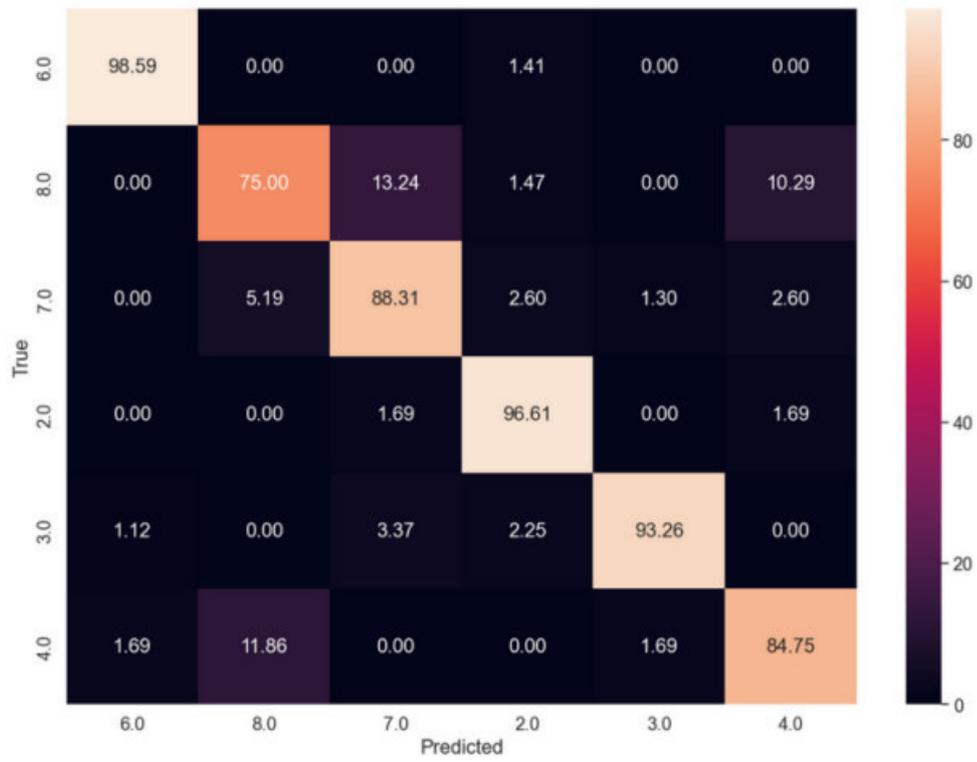


Figura 35 – Matriz de Confusão - 6 Gestos - XGBoost
 Fonte: Própria

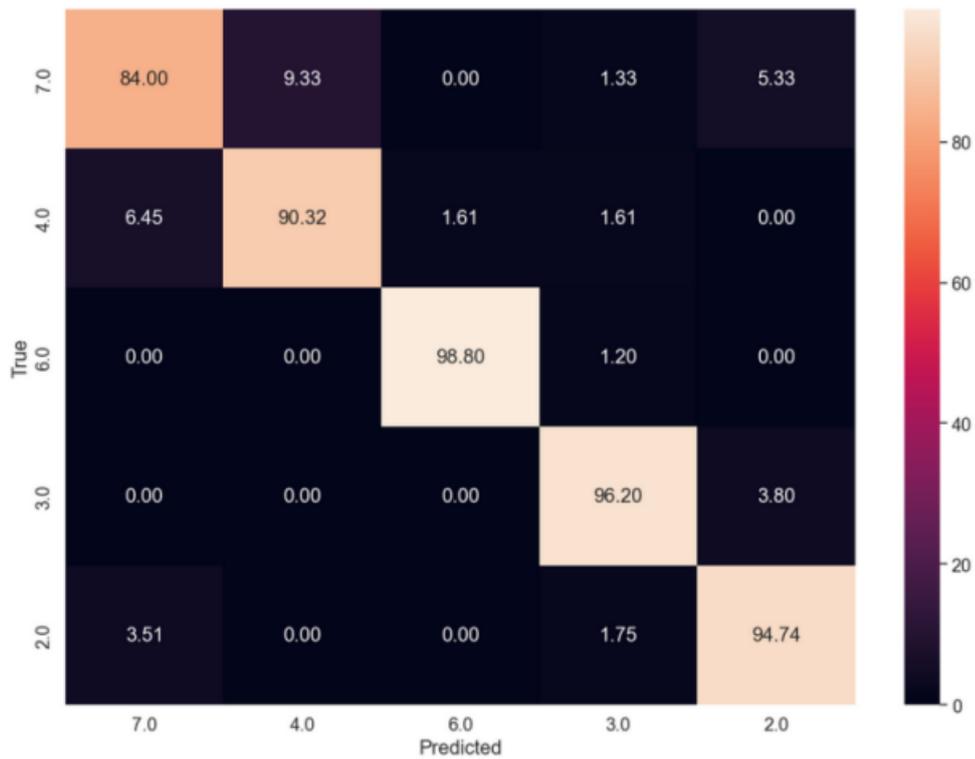


Figura 36 – Matriz de Confusão - 5 Gestos - XGBoost
 Fonte: Própria

6 Considerações finais

Vivemos em uma época onde diversos setores da sociedade estão passando por uma transformação digital, sendo possível ver essa transformação em diversos setores, como na medicina, transporte, segurança. Um dos pontos que mais merece destaque nesse processo de modernização é a Internet das Coisas, que representa a interconexão digital de objetos cotidianos com a internet. De acordo com a figura 37, é previsto um aumento de mais de 100% na quantidade de usuários até 2027 no que tange o mercado de Casas Inteligentes.

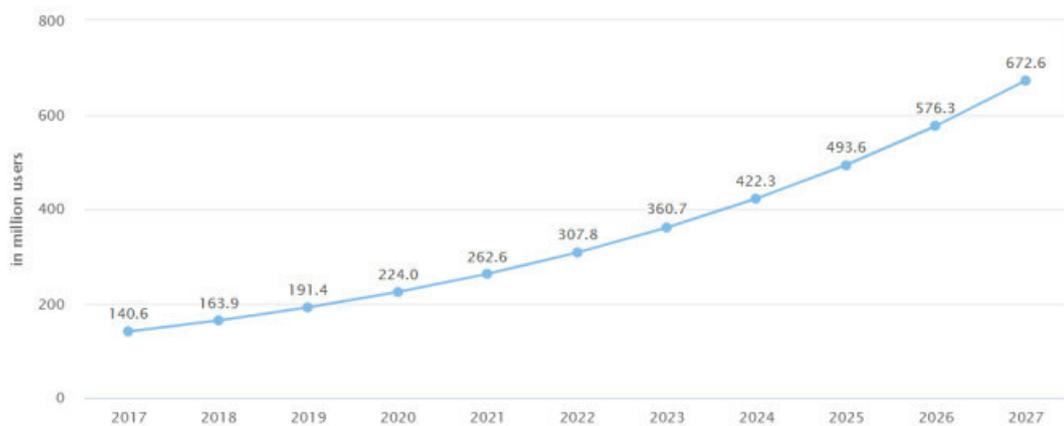


Figura 37 – Projeção para utilização de dispositivos inteligentes para *SmartHome*

Fonte: (STATISTA, 2022)

Um dispositivo inteligente é um dispositivo geralmente conectado a redes ou até mesmo a outros dispositivos, utilizando diferentes protocolos sem fio, como Bluetooth, NFC, Wi-Fi, 5G entre outros. O termo dispositivo inteligente também pode se referir a um dispositivo que exibe algumas propriedades de computação ubíqua, incluindo embora não necessariamente - inteligência artificial.

Alguns dispositivos inteligentes que estão presentes na vida cotidiana dos usuários são os relógios e pulseiras inteligentes. Além das funcionalidades relacionadas ao monitoramento de aspectos fisiológicos do usuário, alguns destes dispositivos permitem a interação do usuário com o ambiente de diferentes maneiras, podemos citar por exemplo os pagamentos por aproximação.

Durante os últimos anos a utilização dos relógios inteligentes ganhou destaque de acordo com a *The Business Research Company* (The Business Research Company, 2021), O crescimento

deve-se principalmente à retomada das atividades das empresas e à adaptação à nova normalidade enquanto se recuperavam do impacto do COVID-19. O mercado de relógios inteligentes consiste na venda dos dispositivos inteligentes e serviços relacionados. Hoje estes dispositivos são empregados principalmente para aplicações no campo da medicina, esportes, jogos e assistência pessoal.

Estes dispositivos contam com uma gama de sensores embarcados, dentre os quais podemos citar os sensores de movimento, estes amplamente utilizados em outros dispositivos inteligentes para a interação do usuário com o próprio aparelho. Os relógios e as *smartbands* por sua vez apresentam um sistema de medição de fluxo sanguíneo que utiliza a fotopletismografia para realizar a mensuração de informações tais, como frequência cardíaca, oxigenação do sangue, pressão arterial, entre outros.

O monitoramento do fluxo sanguíneo entretanto por meio da fotopletismografia apresenta alguns desafios, principalmente quanto aos erros de leitura ou ruídos ocasionados em detrimento do movimento do usuário. Estes ruídos quando presentes no sinal de fotopletismografia são denominados de Artefatos de Movimento. Esses artefatos representam um problema para a determinação da frequência cardíaca durante a execução de atividades físicas. Diante desse cenário temos o principal questionamento a ser respondido por este trabalho, que é a viabilidade de um relógio inteligente equipado com sensores de movimento e um sensor de fotopletismografia de realizar a detecção e classificação de gestos, será possível utilizar os artefatos de movimento como mecanismo para detecção de movimento?

Conforme apresentado na seção 3, alguns estudos utilizam os sensores de movimento e os sensores de fotopletismografia para realizar a classificação de gestos, sejam gestos finos, ou sejam focados na movimentação de punhos e falanges, ou gestos mais amplos como movimento do braço como um todo. Entretanto uma grande parcela desses estudos introduzem dispositivos desenvolvidos utilizando métodos os quais não estão disponíveis em dispositivos inteligentes, podemos destacar os eletrodos associados aos sistemas de eletromiografia ou sensores de fotopletismografia com múltiplos sensores dispostos em diferentes posições e frequências de aquisição superiores a 100hz.

Os relógios inteligentes e *smartbands* dos mais diversos modelos e fabricantes apresentam em de forma geral um sensor do tipo MEMs que é responsável pela aquisição de sinais de

posição utilizando acelerômetros e giroscópios e um sensor óptico para realizar o monitoramento da frequência cardíaca utilizando a técnica de fotopleletismografia. Os sensores ópticos nesses dispositivos apresentam taxas de aquisição de aproximadamente 25hz, isso se dá devido a necessidade de redução do consumo de energia do dispositivo assim como pelo teorema de amostragem de Nyquist onde a taxa de amostragem deve ser maior que o dobro da componente de maior frequência a ser analisada no sinal medido.

O projeto apresentado teve como objetivo realizar o reconhecimento de gestos utilizando dispositivos comerciais, ou seja, sem a utilização de sensores adicionais e/ou alterações nas funcionalidades padrões do dispositivo de forma a responder a questão, será possível um dispositivo inteligente tal como os relógios ou *smartbands* realizar a identificação de gestos? Para isso foi realizado um amplo estudo nos campos da fisiologia de forma entender a relação entre os movimentos e a sua correlação com as alterações no sinal de fotopleletismografia. Foi criado um conjunto de 7 gestos, estes focados na interação do usuário com outros dispositivos, tal como um equipamento multimídia. Foi estabelecido um protocolo para a coleta de amostras de execução destes gestos, onde cada usuário reproduziu cada movimento 20 vezes. Esse processo gerou um banco de dados com um total de 2240 amostras.

Os dados passaram por um processo de pré-processamento, onde foram aplicados algoritmos para realizar a segmentação dos gestos e posteriormente a extração de atributos. O algoritmo de segmentação se baseou no sinal de fotopleletismografia, e do banco de dados inicial de 2240 amostras foi possível realizar a separação de 1951 amostras, ou seja um aproveitamento de 87

Esse resultado indica que um sensor de fotopleletismografia, operando em frequências nominalmente utilizadas para detecção de frequência cardíaca, neste caso de 25hz, é capaz de ser utilizado no processo de identificação de artefatos de movimento, para que assim seja possível a segmentação e classificação dos gestos. A segmentação é o primeiro passo no processo de reconhecimento de gestos. Os movimentos que foram segmentados por essa etapa apresentam informações do sensor de fotopleletismografia e sensores de movimento, sendo as amostras submetidas a um processo de extração de atributos. Essa fase representa uma etapa onde podemos explorar quais os atributos teriam maior impacto no processo de classificação. A tabela 3 apresenta todos os atributos que foram então obtidos. No processo de treinamento dos classificadores, foi utilizado a estratégia de *cross-validation*, devido ao tamanho e as características dos *datasets*.

A tabela 9 apresenta um comparativo entre todos os classificadores apresentados neste projeto. O desempenho do algoritmo baseado no XgBoost obteve o maior resultado na tarefa de classificação no conjunto de 7 gestos, alcançando um precisão de 83.7%, revocação de 83.2 e F1 Score de 0.833, resultados superiores quando comparados com o segundo melhor classificador, o modelo utilizando Florestas Aleatórias. Realizando uma análise da matriz de confusão para os classificadores, figuras 30, 32 e 34 podemos observar que em todos os modelos existe um conflito entre a classificação de 2 gestos, o gesto 5 e o 7.

A partir das tabelas o gesto com a menor taxa de precisão, de acordo com as tabelas 10, 13 e 16 o gesto 5, apresenta os menores valores de F1 Score sendo de 0.65 para as Florestas aleatórias, 0.58 para o algoritmo de máquinas de vetores de suporte e 0.71 para o algoritmo utilizando XgBoost.

Realizando o treinamento e a classificação em um novo conjunto de gestos, excluindo o gesto 5, obtivemos um incremento na capacidade de classificação de todos os classificadores, resultado já esperado uma vez que foi diminuído a quantidade de grupos de classificação. Os resultados são respectivamente, 87,5% de precisão com o classificador F.A. e 88,2% para a máquina de vetores de suporte e 89.5% para o modelo utilizando o XgBoost. Podemos notar que para esse conjunto de dados, as máquinas de vetores de suporte tiveram desempenho superior as florestas aleatórias, entretanto o Xgboost teve um resultado 1.3% superior ao segundo colocado.

Podemos observar que a remoção de um gesto conflitante resultou em um acréscimo de aproximadamente 5% na capacidade de classificação dos algoritmos. Isso demonstra a importância da definição do conjunto de gestos assim como o protocolo de coleta, uma vez que esse tipo de conflito é causado devido a similaridade do movimento dos tendões e músculos que são ativados durante a execução do movimento.

Seguindo a mesma filosofia de remoção do gesto com menor índice de classificação correta, e reduzindo o conjunto de movimentos para 5 movimentos, temos que o classificadores alcançam pontuações de F1 superiores a 0.9. Entretanto a vantagem do XgBoost para os demais classificadores é reduzida, tendo resultados muito próximos entre todos os classificadores.

No processo de treinamento também foi avaliado a quantidade de atributos mínimos necessários para realizar a classificação das amostras. Utilizando um método supervisionado,

chegamos a um número de 27 atributos para as florestas aleatórias e 31 para a SVM e 24 atributos para o XgBoost, isso representa aproximadamente 18% do número total de atributos disponíveis.

Fazendo uma breve análise dos atributos apresentados na tabela 6, podemos notar que para ambos os classificadores, os atributos estão principalmente relacionados aos sensores de movimento, considerando os três eixos. Isso nos mostra que apesar da fotopletismografia ser capaz de identificar os segmentos onde ocorre o movimento, o sinal, devido a baixa taxa de aquisição, não carrega informações suficientemente relevantes no processo de classificação dos gestos.

As figuras 25, 26 e 27 apresentam também uma possibilidade de incremento na precisão e revocação dos classificadores, uma vez que para todos os classificadores, podemos observar que o incremento no conjunto de treinamento permitiria uma melhora nos resultados. O custo computacional do algoritmo desenvolvido neste projeto também demonstra que o tempo de treinamento necessário é relativamente baixo para os modelos de Florestas Aleatórias e Máquinas de Vetores de suporte o que permite de forma não custosa a escalabilidade do modelo. Entretanto as técnicas de otimização utilizadas pelo algoritmo de XgBoost aumentam a complexidade do modelo e conseqüentemente seu tempo de treinamento é superior aos demais em no mínimo 10 vezes.

Por fim, em uma retrospectiva dos objetivos deste projeto podemos apresentar que todos foram alcançados. É viável a utilização dos sensores de fotopletismografia já existentes em *smartwatches* para a segmentação dos movimentos de pulso. Isso é possível em detrimento da existência dos artefatos de movimento. O uso combinado de atributos provenientes dos acelerômetros e giroscópios com os segmentos obtidos pela fotopletismografia permite a sua classificação por meio de modelos de aprendizado de máquinas. Esses resultados mostram que a utilização dos *smartwatches* como uma forma de interação com um ambiente inteligente é viável, seja para tarefas simples ou até mesmo em processos de interpretação de linguagem de sinais conforme apresentado na literatura.

O presente projeto também foi publicado nas seguintes conferencias, 12th International Conference on Consumer Electronics (ICCE-Berlin), sob o título de "Gesture recognition of wrist motion using low-frequency PPG." (RYLO et al., 2022), e The 12th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare

(ICTH 2022), (RYLO; LANDAU; JR, 2022).

Podemos ressaltar também para trabalhos futuros a avaliação de métodos de classificação mais complexos, tal como técnicas utilizando redes neurais profundas. A utilização de outros tipos de atributos, como por exemplo atributos no domínio de tempo-frequência, incluindo também a agregação de atributos para gerar outras informações que possam apresentar relevância no processo de classificação. Uma outra possibilidade para lidar com a grande quantidade de atributos selecionados é a utilização de técnicas de redução de dimensionalidade, tal como *Principal Component Analysis* (PCA) ao invés de somente técnicas de seleção de atributos.

Considerando a natureza do problema que lida com séries temporais, poderia ser utilizado modelos tal como as *Long short-term memory* (LSTM), entretanto seria de extrema importância a expansão do banco de dados uma vez que Redes Neurais tendem a precisar de conjunto de dados maiores para alcançar resultados satisfatórios em seus processos de treinamento e validação.

É possível também um estudo mais aprofundado das técnicas de segmentação do sinal de fotopletiografia por meio de métodos baseados em inteligência artificial.

Referências

- ALLEN, J. Photoplethysmography and its application in clinical physiological measurement. *Physiological Measurement*, v. 28, p. R1–39, 04 2007. 14
- AWOTUNDE, J. B.; CHAKRABORTY, C.; FOLORUNSO, S. O. A secured smart healthcare monitoring systems using blockchain technology. In: _____. *Intelligent Internet of Things for Healthcare and Industry*. Cham: Springer International Publishing, 2022. p. 127–143. ISBN 978-3-030-81473-1. Disponível em: <https://doi.org/10.1007/978-3-030-81473-1_6>. 21
- BAGHA, S.; SHAW, L. A real time analysis of ppg signal for measurement of spo2 and pulse rate. *International journal of computer applications*, International Journal of Computer Applications, 244 5 th Avenue,# 1526, New ... , v. 36, n. 11, p. 45–50, 2011. 14
- BOUKHECHBA, M.; CAI, L.; WU, C.; BARNES, L. Actippg: Using deep neural networks for activity recognition from wrist-worn photoplethysmography (ppg) sensors. In: . [S.l.: s.n.], 2019. 47
- BOWDEN, R.; ZISSERMAN, A.; KADIR, T.; BRADY, M. Vision based interpretation of natural sign languages. 06 2003. 20
- BREIMAN, L. Random forests. *Machine Learning*, v. 45, n. 1, p. 5–32, Oct 2001. ISSN 1573-0565. Disponível em: <<https://doi.org/10.1023/A:1010933404324>>. 34
- BREIMAN, L.; FRIEDMAN, J. H.; OLSHEN, R. A.; STONE, C. J. *Classification and Regression Trees*. Belmont, CA: Wadsworth International Group, 1984. 31
- BUDIDHA, K. Investigation of photoplethysmography and arterial blood oxygen saturation from the ear-canal and the finger under conditions of artificially induced hypothermia. In: . [S.l.: s.n.], 2015. v. 2015, p. 7954–7957. 22
- BUDIDHA, K.; KYRIACOU, P. A. In vivo investigation of ear canal pulse oximetry during hypothermia. *Journal of Clinical Monitoring and Computing*, v. 32, n. 1, p. 97–107, Feb 2018. ISSN 1573-2614. Disponível em: <<https://doi.org/10.1007/s10877-017-9975-4>>. 24
- CHAN, L.; LIANG, R.-H.; TSAI, M.-C.; CHENG, K.-Y.; SU, C.-H.; CHEN, M.; CHENG, W.-H.; CHEN, B.-Y. Fingerpad: Private and subtle interaction using fingertips. *ACM User Interface Software and Technology Symposium (ACM UIST 2013)*, 10 2013. 45
- CHARLTON, P.; MAROZAS, V. Wearable photoplethysmography devices. In: KYRIACOU, P.; ALLEN, J. (Ed.). *Photoplethysmography*. [S.l.]: Elsevier, 2021. 23
- CHEN, T.; GUESTRIN, C. Xgboost: A scalable tree boosting system. In: . [S.l.: s.n.], 2016. p. 785–794. 34
- CORTES, C.; VAPNIK, V. Support-vector networks. *Chem. Biol. Drug Des.*, v. 297, p. 273–297, 01 2009. 30
- DAVIS, J.; SHAH, M. Visual gesture recognition. *Vision, Image and Signal Processing, IEE Proceedings -*, v. 141, p. 101 – 106, 05 1994. 20

- GANESAN, R.; DAS, T.; SIKDER, A.; KUMAR, A. Wavelet-based identification of delamination defect in cmp (cu-low k) using nonstationary acoustic emission signal. *Semiconductor Manufacturing, IEEE Transactions on*, v. 16, p. 677 – 685, 12 2003. 38
- GHAMARI, M.; NAZERAN, H.; SOLTANPUR, C. A review on wearable photoplethysmography sensors and their potential future applications in health care. v. 4, p. 195–202, 08 2018. 23
- GUPTA, S.; MORRIS, D.; PATEL, S.; TAN, D. Soundwave: using the doppler effect to sense gestures. 05 2012. 45, 48
- HAN, J.; PEI, J.; KAMBER, M. *Data Mining: Concepts and Techniques*. Elsevier Science, 2011. (The Morgan Kaufmann Series in Data Management Systems). ISBN 9780123814807. Disponível em: <<https://books.google.com.br/books?id=pQws07tdpjoC>>. 36, 37, 40
- HERAVI, Y.; AMIN, M.; KEIVAN, M.; SIMA, J. A new approach for blood pressure monitoring based on ecg and ppg signals by using artificial neural networks. *International Journal of Computer Applications*, v. 103, n. 12, p. 36–40, 2014. 14
- HONDA, Y.; WEBER, S.; LUETH, T. Stability analysis for tactile pattern based recognition system for hand gestures. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, v. 2007, p. 4033–6, 02 2007. 19
- IOSA, M.; PICERNO, P.; PAOLUCCI, S.; MORONE, G. Wearable inertial sensors for human movement analysis. *Expert Review of Medical Devices*, v. 13, 06 2016. 25
- JAAFAR, R.; ROZALI, M. A. A. Estimation of breathing rate and heart rate from photoplethysmogram. In: *2017 6th International Conference on Electrical Engineering and Informatics (ICEEI)*. [S.l.: s.n.], 2017. p. 1–4. 37
- KUHN, M.; JOHNSON, K. *Applied predictive modeling*. New York, NY: Springer, 2013. ISBN 9781461468493 1461468493 1461468485 9781461468486. Disponível em: <<http://www.amazon.com/Applied-Predictive-Modeling-Max-Kuhn/dp/1461468485/>>. 42
- LEE, B. G.; CUONG, T.; WEI, C. Smart hand device gesture recognition with dynamic time-warping method. In: . [S.l.: s.n.], 2017. p. 216–219. ISBN 978-1-4503-5430-1. 48, 73
- LI, G.; CHUNG, W.-Y. Detection of driver drowsiness using wavelet analysis of heart rate variability and a support vector machine classifier. *Sensors*, v. 13, n. 12, p. 16494–16511, 2013. ISSN 1424-8220. Disponível em: <<https://www.mdpi.com/1424-8220/13/12/16494>>. 37
- LING, Y.; CHEN, X.; RUAN, Y.; ZHANG, X.; CHEN, X. Comparative study of gesture recognition based on accelerometer and photoplethysmography sensor for gesture interactions in wearable devices. *IEEE Sensors Journal*, PP, p. 1–1, 05 2021. 48, 73
- LIU, M.; PO, L.-M.; FU, H. Cuffless blood pressure estimation based on photoplethysmography signal and its second derivative. *International Journal of Computer Theory and Engineering*, IACSIT Press, v. 9, n. 3, p. 202, 2017. 14
- LU, Z.; XIANG, C.; LI, Q.; ZHANG, X.; ZHOU, P. A hand gesture recognition framework and wearable gesture-based interaction prototype for mobile devices. *Human-Machine Systems, IEEE Transactions on*, v. 44, p. 293–299, 04 2014. 45

- MITRA, S.; ACHARYA, T. Gesture recognition: A survey. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, v. 37, p. 311 – 324, 06 2007. 15, 16, 19
- MONTALVÃO, D. Sensors and signal processing - chapter 5 of the book "mechatronics: Fundamentals and applications". In: _____. [S.l.: s.n.], 2015. p. 125–220. ISBN 9781482239317. 26
- MORAES, J. Advances in photoplethysmography signal analysis for biomedical applications. *Sensors*, v. 18, p. 1–26, 2018. 21
- MÖLLER, D. *Guide to Computing Fundamentals in Cyber-Physical Systems*. [S.l.: s.n.], 2016. ISBN 978-3-319-25176-9. 21
- NANDAKUMAR, R.; IYER, V.; TAN, D.; GOLLAKOTA, S. Fingerio: Using active sonar for fine-grained finger tracking. In: . [S.l.: s.n.], 2016. p. 1515–1525. 45, 48
- PELAEZ, E. A.; VILLEGAS, E. Led power reduction trade-offs for ambulatory pulse oximetry. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, v. 2007, p. 2296–9, 02 2007. 21
- PERERA, N.; RAJAPAKSE, A.; JAYASINGHE, R. On-line discrete wavelet transform in empty environment and applications in protection relaying. 01 2007. 38
- PERRAULT, S.; LECOLINET, E.; EAGAN, J.; GUIARD, Y. Watchit: Simple gestures and eyes-free interaction for wristwatches and bracelets. In: . [S.l.: s.n.], 2013. p. 1451–1460. 44
- POLLREISZ, D.; TAHERINEJAD, N. Detection and removal of motion artifacts in ppg signals. *Mobile Networks and Applications*, Aug 2019. ISSN 1572-8153. Disponível em: <<https://doi.org/10.1007/s11036-019-01323-6>>. 40
- RAUTARAY, S.; AGRAWAL, A. Interaction with virtual game through hand gesture recognition. 12 2011. 19
- REISNER, A.; SHALTIS, P.; MCCOMBIE, D.; ASADA, H.; WARNER, D.; WARNER, M. Utility of the Photoplethysmogram in Circulatory Monitoring. *Anesthesiology*, v. 108, n. 5, p. 950–958, 05 2008. ISSN 0003-3022. Disponível em: <<https://doi.org/10.1097/ALN.0b013e31816c89e1>>. 22
- RYLO, M.; LANDAU, R.; JR, V. L. Gesture recognition of wrist motion based on wearables sensors. *Procedia Computer Science*, v. 210, p. 181–188, 11 2022. 81
- RYLO, M.; SILVA, W.; LANDAU, R.; JR, V. L. Gesture recognition of wrist motion using low-frequency ppg. In: . [S.l.: s.n.], 2022. p. 1–4. 80
- SAPONAS, T.; TAN, D.; MORRIS, D.; TURNER, J.; LANDAY, J. Making muscle-computer interfaces more practical. In: . [S.l.: s.n.], 2010. v. 2, p. 851–854. 16
- SAREEN, M.; ABHINAV; PRAKASH, P.; ANAND, S. Wavelet decomposition and feature extraction from pulse signals of the radial artery. In: *2008 International Conference on Advanced Computer Theory and Engineering*. [S.l.: s.n.], 2008. p. 551–555. 37
- SEVIL, M.; RASHID, M.; ASKARI, M. R.; SAMADI, S.; HAJIZADEH, I.; CINAR, A. Psychological stress detection using photoplethysmography. *IEEE Transactions on intelligent transportation systems*, v. 6, n. 2, p. 156–166, 2005. 14

- STATISTA. *Smart Home - Worldwide | Statista Market Forecast* — *statista.com*. 2022. <<https://www.statista.com/outlook/dmo/smart-home/worldwide#smart-homes>>. [Accessed 22-Feb-2023]. 76
- SUBRAMANIAN, K.; SAVUR, C.; SAHIN, F. Using photoplethysmography for simple hand gesture recognition. In: . [S.l.: s.n.], 2020. p. 307–312. 47, 50, 72
- TADI, M. J.; LEHTONEN, E.; HURNANEN, T.; KOSKINEN, J.; ERIKSSON, J.; PÄNKÄÄLÄ, M.; TERÄS, M.; KOIVISTO, T. A real-time approach for heart rate monitoring using a hilbert transform in seismocardiograms. *Physiological measurement*, IOP Publishing, v. 37, n. 11, p. 1885, 2016. 24
- TAMURA, T.; MAEDA, Y.; SEKINE, M.; YOSHIDA, M. *Wearable Photoplethysmographic Sensors—Past and Present*. 2014. 22
- The Business Research Company. *Smart Watch Global Market Report 2021: COVID-19 Growth and Change to 2030*. [S.l.], 2021. 76
- THEODORIDIS, S.; KOUTROUMBAS, K. *Pattern Recognition, Fourth Edition*. [S.l.]: Academic Press, 2009. Hardcover. ISBN 9781597492720. 27, 29, 31, 32
- TRIGUEIROS, P.; RIBEIRO, F.; REIS, L. Generic system for human-computer gesture interaction. In: . [S.l.: s.n.], 2014. 19
- WEIGEL, M.; NITTALA, A.; OLWAL, A.; STEIMLE, J. Skinmarks: Enabling interactions on body landmarks using conformal skin electronics. In: . [S.l.: s.n.], 2017. p. 3095–3105. 44
- WEN, H.; ROJAS, J. R.; DEY, A. Serendipity: Finger gesture recognition using an off-the-shelf smartwatch. In: . [S.l.: s.n.], 2016. p. 3847–3851. 17, 46, 72
- YAMATO, J.; OHYA, J.; ISHII, K. Recognizing human action in time-sequential images using hidden markov model. In: . [S.l.: s.n.], 1992. J76-D-II, p. 379 – 385. ISBN 0-8186-2855-3. 20
- ZHANG, Y.; GU, T.; LUO, C.; KOSTAKOS, V.; SENEVIRATNE, A. Findroidhr: Smartwatch gesture input with optical heartrate monitor. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, v. 2, p. 1–42, 03 2018. 17, 44, 45, 47, 55
- ZHAO, T.; LIU, J.; WANG, Y.; LIU, H.; CHEN, Y. Towards low-cost sign language gesture recognition leveraging wearables. *IEEE Transactions on Mobile Computing*, PP, p. 1–1, 12 2019. 48, 55, 73

APÊNDICE A – Primeiro apêndice

Nesta seção será apresentado informações complementares a metodologia deste projeto.

A.1 Pseudocódigo - Limiar de FC e Distancia entre picos

O pseudocódigo para obtenção de algumas variáveis do processo são apresentados na tabela abaixo, as funções complementares CalcGradient, Diff e FindPeaks foram implementadas a partir dos pacotes Numpy e Scipy, estas podem ser substituídas respectivamente por, np.gradient, np.diff e scipy.signal.find_peaks.

Limiar de FC e Distancia entre picos	
Entrada:	Série Temporal do usuário em Descanso (<i>hr_ts</i>)
Saída:	Limiar de Frequencia Cardiaca (<i>hr_tresh</i>) e Distancia entre Picos (<i>ppd</i>)
	<i>shift_window</i> = 10
	<i>hr_ts_gradient</i> = CalcGradient(<i>hr_ts</i>)
Processo:	<i>peaks, val</i> = FindPeaks(Inv(<i>hr_ts</i>), height=0, distance= <i>shift_window</i>)
	<i>distances</i> = Diff(<i>peaks</i>)
	<i>hr_thres</i> = média(<i>peaks</i>)
	<i>ppd</i> = Média(<i>distances</i>)

A.2 Pseudocódigo - Segmentação

O pseudocódigo para a segmentação dos gestos é apresentado na tabela abaixo. A variável *idx_seg_gesto* é uma lista onde cada item possui dois valores sendo estes o tempo inicial e final de um determinado segmento.

Segmentação

Entrada: Série Temporal PPG (hr_ts), Série Temporal IMU (imu_ts),
Limiar de frequência cardíaca (hr_thres) e Distância Média entre Picos (ppd)

Saída: Objeto com as séries temporais de todos os
gestos segmentados (seg_gesto_ppg, seg_gesto_imu)

```
shift_window = 10
```

```
seg_gesto = []
```

```
hr_ts = WaveletFilter(hr_ts)
```

```
imu_ts = WaveletFilter(imu_ts)
```

```
hr_ts_gradient = CalcGradient(hr_ts)
```

```
peaks, val = FindPeaks(Inv(hr_ts_gradient), height=hr_thres, distance=shift_window)
```

```
distances = Diff(peaks)
```

Processo: i = 0

```
para cada item em distances:
```

```
    segmento = Diff( distances[item] <ppd )
```

```
    se segmento <ppd:
```

```
        idx_seg_gesto[i] = segmento
```

```
        i++
```

```
para cada item em idx_seg_gesto:
```

```
    inicio = idx_seg_gesto[item][0]
```

```
    fim = idx_seg_gesto[item][1]
```

```
    seg_gesto_ppg[item] = hr_ts_gradient[inicio:fim]
```

```
    seg_gesto_imu[item] = imu_ts[inicio:fim]
```
