



Análise e seleção de aspectos cinemáticos da marcha humana para classificação de indivíduos por faixa etária

Trabalho de Conclusão de Curso

Engenharia de Computação

REBECA SARAI AGUIAR DA SILVA
Orientador: Prof. Fernando Buarque
Coorientador: M.Sc. Rafael Caldas



Rebeca Sarai Aguiar da Silva

Análise e seleção de aspectos cinemáticos da marcha humana para classificação de indivíduos por faixa etária

Monografia apresentada como requisito para obtenção do diploma de Bacharel em Engenharia de Computação pela Escola Politécnica de Pernambuco – Universidade de Pernambuco.

Engenharia de Computação
Escola Politécnica de Pernambuco
Universidade de Pernambuco

Orientador: Prof. Fernando Buarque
Coorientador: M.Sc. Rafael Caldas

Recife - PE, Brasil
Dezembro de 2018

Rebeca Sarai Aguiar da Silva

Análise e seleção de aspectos cinemáticos da marcha humana para classificação de indivíduos por faixa etária/ Rebeca Sarai Aguiar da Silva. – Recife - PE, Brasil, Dezembro de 2018-

61 f. : il.

Orientador: Prof. Fernando Buarque

Trabalho de Conclusão de Curso – Engenharia de Computação

Escola Politécnica de Pernambuco

Universidade de Pernambuco, Dezembro de 2018.

1. Marcha. 2. Redes Neurais. 3. Self Organizing Maps. 4. Técnicas de Agrupamento. 5. K-Means. 6. Fuzzy C-Means. I. Prof. Fernando Buarque. II. Universidade de Pernambuco. III. Escola Politécnica. IV. Título

Este trabalho é dedicado à minha família.

MONOGRAFIA DE FINAL DE CURSO

Avaliação Final (para o presidente da banca)*

No dia 20 de dezembro de 2018, às 12:00 horas, reuniu-se para deliberar a defesa da monografia de conclusão de curso do discente **REBECA SARAI AGUIAR DA SILVA**, orientado pelo professor **Fernando Buarque de Lima Neto**, sob título **Análise e seleção de aspectos cinemáticos da marcha humana para classificação de indivíduos por faixa etária**, a banca composta pelos professores:

Diego José Rátiva Millan

Fernando Buarque de Lima Neto

Após a apresentação da monografia e discussão entre os membros da Banca, a mesma foi considerada:

Aprovada Aprovada com Restrições* Reprovada

e foi-lhe atribuída nota: 9.0 (*Nove*)

*(Obrigatório o preenchimento do campo abaixo com comentários para o autor)

O discente terá 30 dias para entrega da versão final da monografia a contar da data deste documento.



DIEGO JOSÉ RÁTIVA MILLAN

FERNANDO BUARQUE DE LIMA NETO

* Este documento deverá ser encadernado juntamente com a monografia em versão final.

Agradecimentos

Gostaria de agradecer a Deus, o Autor e Mantenedor de todas as coisas, No qual baseio minha fé, por ter sido meu refúgio indispensável nos mais delicados momentos que se passaram desde o início da minha graduação até aqui.

Também sou grata ao Prof. Dr. Fernando Buarque pela oportunidade e pelo suporte no pouco tempo que lhe coube, e ao M.Sc. Rafael Caldas pela sua paciência, correções e incentivos.

*“But I’m not sentimental
This skin and bones is a rental
And no one makes it out alive.”*
Switchfoot

Resumo

A análise da marcha a partir de informações coletadas por sensores de movimento se tornou comum nos últimos anos devido às suas contribuições para atenção a saúde. Várias técnicas de *machine learning* são aplicadas para simplificar a interpretação da marcha e desenvolver modelos de auxílio à decisão, isso também se tornou comum devido à disponibilidade generalizada de sensores para medir movimentos em smartphones, rastreadores fitness e smartwatches. Para auxiliar na avaliação de disfunções de marcha neste trabalho foram desenvolvidas duas técnicas híbridas baseadas em: self-organizing maps (SOM) e dois algoritmos de agrupamento, que no escopo deste trabalho foram o K-means e o Fuzzy C-means (FCM). Os algoritmos produzidos foram comparados com os algoritmos base a fim de avaliar a qualidade da classificação. Os experimentos utilizaram uma base coletada por sensores inerciais de 180 indivíduos divididos em seis grupos relacionados por faixa etária. Os resultados mostraram que o método SOM + FCM possui uma alta precisão (88%). Os resultados também mostraram a relevância da Cadência na classificação indivíduos por faixa etária.

Palavras-chave: Análise de Marcha, Redes Neurais, Self organizing Maps, Agrupamento, K-Means, Fuzzy C-Means.

Abstract

The gait analysis based on information collected by motion sensors has become common in recent years. Several techniques of machine learning are applied to simplify gait interpretation and to develop decision support models. This happens because of the widespread availability of sensors to measure movements on smartphones, fitness trackers and smartwatches. In order to support the evaluation of gait dysfunctions, two hybrid techniques were developed based on self-organizing maps (SOM) and two clustering algorithm, which in the scope of this work were K-means and Fuzzy C-means (FCM). The produced algorithms were compared with the base versions to assess the quality of the classification. The experiments used a database collected by inertial sensors of 180 individuals divided into six age groups. The results showed that the SOM + FCM method has a high precision (88 %) identifying with groups with significantly similar gait patterns. The results also showed the relevance of Cadence in the classification by age group.

Keywords: Gait analysis, Self organizing Maps, Neural Networks, Clustering, K-Means, Fuzzy C-Means, Feature Relevance.

Lista de ilustrações

Figura 1 – Fases da marcha	19
Figura 2 – Ilustração de um neurônio biológico	22
Figura 3 – Mapa de kohonen de duas dimensões	23
Figura 4 – Configuração do sistema de captura marcha	28
Figura 5 – Exemplo de heatmap das 7 variáveis de entrada	29
Figura 6 – Níveis de abstração de abordagens híbridas	30
Figura 7 – Distribuição dos dados dos grupos A e C para SOM + K-Means	35
Figura 8 – Heatmap dos padrões de aprendizado dos grupos A e C	35
Figura 9 – Distribuição dos dados dos grupos E e F para SOM + K-Means	36
Figura 10 – Heatmap dos padrões de aprendizado dos grupos E e F	36
Figura 11 – Heatmap com todas as sete entradas para classificação dos grupos A e C, utilizando os dados de somente um indivíduo	44
Figura 12 – Heatmap SOM + K-MEANS com todas as sete entradas para classificação dos grupos A e C	44

Lista de tabelas

Tabela 1 – Resultado do algoritmo SOM + K-Means para classificação dos Grupos A e C	34
Tabela 2 – Resultado do algoritmo SOM + K-Means para classificação dos Grupos E e F	36
Tabela 3 – Resultado do algoritmo SOM + FC-Means para classificação dos Grupos A e C	37
Tabela 4 – Resultado do algoritmo SOM + FC-Means para classificação dos Grupos E e F	37
Tabela 5 – Resultado do algoritmo SOM + K-Means para classificação dos Grupos A e C com 6 variáveis de entrada	38
Tabela 6 – Resultado do algoritmo SOM + K-Means para classificação dos Grupos E e F com 6 variáveis de entrada	38
Tabela 7 – Resultado do algoritmo SOM + FC-Means para classificação dos Grupos A e C com 6 variáveis de entrada	38
Tabela 8 – Resultado do algoritmo SOM + FC-Means para classificação dos Grupos E e F com 6 variáveis de entrada	39
Tabela 9 – Resultado do algoritmo SOM + K-Means para classificação dos Grupos A e C com 5 variáveis de entrada	40
Tabela 10 – Resultado do algoritmo SOM + K-Means para classificação dos Grupos E e F com 5 variáveis de entrada	41
Tabela 11 – Resultado do algoritmo SOM + FC-Means para classificação dos Grupos A e C com 5 variáveis de entrada	42
Tabela 12 – Resultado do algoritmo SOM + FC-Means para classificação dos Grupos E e F com 5 variáveis de entrada	43
Tabela 13 – Melhores Resultados com 4 variáveis de entrada	43
Tabela 14 – Resultados do algoritmo SOM + K-Means para classificação dos Grupos A e C com 2 variáveis de entrada	50
Tabela 15 – Resultados do algoritmo SOM + K-Means para classificação dos Grupos E e F com 2 variáveis de entrada	51
Tabela 16 – Resultados do algoritmo SOM + FC-Means para classificação dos Grupos A e C com 2 variáveis de entrada	52
Tabela 17 – Resultados do algoritmo SOM + FC-Means para classificação dos Grupos E e F com 2 variáveis de entrada	53
Tabela 18 – Resultados do algoritmo SOM + K-Means para classificação dos Grupos A e C com 3 variáveis de entrada	54

Tabela 19 – Resultados do algoritmo SOM + K-Means para classificação dos Grupos E e F com 3 variáveis de entrada	55
Tabela 20 – Resultados do algoritmo SOM + FC-Means para classificação dos Grupos A e C com 3 variáveis de entrada	56
Tabela 21 – Resultados do algoritmo SOM + FC-Means para classificação dos Grupos E e F com 3 variáveis de entrada	57
Tabela 22 – Resultados do algoritmo SOM + K-Means para classificação dos Grupos A e C com 4 variáveis de entrada	58
Tabela 23 – Resultados do algoritmo SOM + K-Means para classificação dos Grupos E e F com 4 variáveis de entrada	59
Tabela 24 – Resultados do algoritmo SOM + FC-Means para classificação dos Grupos A e C com 4 variáveis de entrada	60
Tabela 25 – Resultados do algoritmo SOM + FC-Means para classificação dos Grupos E e F com 4 variáveis de entrada	61

Lista de abreviaturas e siglas

IMU	<i>Inertial Measurement Unit</i>
SOM	<i>Self Organizing Maps</i>
RNA	Redes Neurais Artificiais
FCM	<i>Fuzzy C-Means</i>
MEMS	<i>Micro Electro Mechanical Systems</i>
OU-ISIR	<i>Osaka University-Institute of Scientific and Industrial Research</i>
BMU	<i>Best Matching Unit</i>
EMG	Eletromiografia
TP	<i>True Positives</i>
TN	<i>True Negatives</i>
FP	<i>False Positives</i>
FN	<i>False Negatives</i>
ACT	<i>Average Cycle Time</i>
Rst	<i>Temporal Stance Ratio</i>
Rsw	<i>Temporal Swing Ratio</i>
RssR	<i>Temporal Stance-Swing Ratio</i>
RssL	<i>Temporal Stance-Swing Ratio</i>
Rs	<i>Overall Symmetry Ratio</i>

Sumário

1	INTRODUÇÃO	15
1.1	Motivação e Caracterização do Problema	15
1.2	Objetivo Geral	16
1.3	Objetivos Específicos	17
1.4	Estrutura do Documento	17
2	FUNDAMENTAÇÃO TEÓRICA	18
2.1	Breve História do Estudo da Marcha	18
2.2	Marcha	18
2.3	Redes Neurais	21
2.3.1	Self Organizing Maps	22
2.4	Técnicas de Agrupamento	25
2.4.1	K-Means	25
2.4.2	Fuzzy C-Means	26
3	MÉTODOS	27
3.1	Aquisição dos Dados	27
3.2	Pré-Processamento	28
3.3	Extração de Características	28
3.4	Processamento	29
3.4.1	SOM + FCM	30
3.4.2	SOM + K-Means	31
3.5	Experimentos	32
3.6	Métricas	32
4	RESULTADOS	34
4.1	Cenário 2G7F	34
4.1.1	Resultados do algoritmo SOM + K-Means	34
4.1.2	Resultados do algoritmo SOM + FC-Means	36
4.2	Cenário 2G6F	37
4.2.1	Resultados do algoritmo SOM + K-Means	37
4.2.2	Resultados do algoritmo SOM + FC-Means	37
4.3	Cenário 2G5F	38
4.3.1	Resultados do algoritmo SOM + K-Means	39
4.3.2	Resultados do algoritmo SOM + FC-Means	39
4.4	Cenário 2G4F	39

4.5	Cenário 2G3F	39
4.6	Cenário 2G2F	40
4.7	Relevância das Características	40
5	CONCLUSÕES E TRABALHOS FUTUROS	45
	REFERÊNCIAS	46
	APÊNDICES	49
	APÊNDICE A – RESULTADOS DE TODOS OS EXPERIMENTOS	
	RESTANTES	50

1 Introdução

Este trabalho de conclusão de curso se propõe a estudar e desenvolver um classificador de indivíduos por faixa etária para auxiliar fisioterapeutas em suas atividades de avaliação de distúrbios de marcha. Este capítulo descreve a introdução da monografia, e está organizado em 4 seções. Na Seção 1.1 são descritas tanto a motivação para a execução deste trabalho e a caracterização do problema. Em seguida, nas Seções 1.2 e 1.3 são apresentados os objetivos gerais e específicos, bem como a proposta de solução do projeto. Por fim, a Seção 1.4 detalha a organização da monografia.

1.1 Motivação e Caracterização do Problema

A marcha humana se refere à maneira de andar de um indivíduo, esse movimento pode ser descrito como uma série de estados de transição, sendo dividido em fases de apoio e balanço (1). A pesquisa da análise de marcha humana data da década de 1970 (2) quando foi usada para fins médicos no diagnóstico precoce de vários distúrbios (3) (4). O monitoramento das características da marcha ao longo do tempo mostrou-se útil no diagnóstico precoce de doenças e suas complexidades. Mais recentemente, a análise da marcha tem sido usada para diferenciar um indivíduo dos outros sendo usada para fins de identificação biométrica (5).

O método mais natural de análise de marcha é a observação clínica, tal método exige a presença do paciente e do profissional de saúde e se baseia na expertise do avaliador. Uma sofisticação da análise observacional é a filmagem da marcha do indivíduo, permitindo ao especialista a possibilidade de repetir a observação várias vezes e comparar a situação do indivíduo antes e depois do tratamento. Esses métodos convencionais muitas vezes não apresentam alto índice de precisão, já que os resultados dependem exclusivamente do conhecimento e habilidade do médico, tornando às vezes, a conclusão subjetiva.

Com o avanço da tecnologia, aumentaram-se as pesquisas na área, e com isso vários centros médicos desenvolveram laboratórios unicamente para analisar os padrões do movimento humano. Tais laboratórios conseguem detectar alterações que não podem ser percebidas a olho nu, acrescentando muitas vantagens no estudo e análise da marcha. Atualmente, o recurso mais eficiente para a análise cinemática da marcha é o rastreamento optoeletrônico tridimensional. Nesse sistema, imagens são coletadas por câmeras filmadoras e repassadas para um *software* que reconstrói todo o movimento do paciente de forma tridimensional. A desvantagem desse método é o alto custo, e por isso é restrito a grandes centros tecnológicos, usados por uma pequena elite de profissionais.

O desejo de ter um resultado similar e de forma acessível fez com que a análise da marcha a partir de informações coletadas por sensores de movimento se tornasse um tópico de pesquisa ativo. A popularização de *smartphones* e *wearable devices* fez com que esses sensores se tornassem acessíveis à população, tornando a configuração para coleta de dados de análise de marcha não intrusiva e ubíqua.

A maioria desses sensores de movimento *wearable* utilizam sensores inerciais baseados em *Micro Electro Mechanical Systems* (MEMS). Esses sensores inerciais, também conhecidos como Unidades de Medida Inercial (IMUs), que medem a aceleração linear, a velocidade angular e o vetor de campo magnético em um sistema de coordenadas local tridimensional. Além da medição inercial de sinais, alguns dispositivos comercialmente disponíveis incorporam algoritmos que fornecem estimativas da orientação do sensor em relação a um sistema global de coordenadas fixas.

Esses sensores inerciais, compostos basicamente de acelerômetros e giroscópios, são utilizados de forma combinada, como IMUs. Devido ao seu tamanho pequeno, portabilidade e alto poder de processamento, as IMUs são amplamente utilizadas para análises complexas de movimento (6) (7). Um grande número de pesquisas têm sido conduzidas no desenvolvimento de sistemas de reconhecimento de marcha utilizando sensores inerciais (6) (8) (9). A desvantagem é que os dados coletados por sensores são representados de maneira complexa, mesmo após o pré-processamento (10). Por isso o uso de técnicas robustas e capazes de lidar com essa complexidade torna-se essencial na tarefa de identificar um indivíduo de outros.

Estudos prévios mostraram o uso de algoritmos adaptativos para a classificação de indivíduos por faixa etária, como self-organizing maps (SOM) e Fuzzy C-Means, comprovou-se que a técnica SOM não consegue distinguir pequenas mudanças no padrão de marcha. Também foi comprovado que o FCM apesar de apresentar melhores resultados, ainda apresenta um resultado insatisfatório na classificação por todas as faixas etárias, além de só conseguir atender aos requisitos entre alguns grupos etários (10). A proposta deste trabalho é construir um modelo híbrido que demonstre desempenho superior que os mesmos algoritmos separados e reconhecendo o estado padrão da marcha de indivíduos saudáveis para cada faixa etária. Podendo ser estendido no desenvolvimento de um sistema de avaliação do desempenho da marcha do paciente.

1.2 Objetivo Geral

O desenvolvimento deste trabalho tem como objetivo principal elaborar uma proposta de seleção de aspectos cinemáticos da marcha humana para classificação de indivíduos por faixa etária.

1.3 Objetivos Específicos

Assim, para que os objetivos principais fossem alcançados, as seguintes metas foram cumpridas:

- Implementar o algoritmo tradicional SOM;
- Implementar a técnica de agrupamento de dados tradicional K-Means;
- Implementar a técnica de agrupamento de dados tradicional FCM;
- Implementar a hibridização das técnicas mencionadas acima para construir um modelo confiável para classificação de indivíduos por faixa etária;
- Propor um modelo de extração de características a partir dos dados coletados por sensores inerciais;
- Construir um sistema de baixo custo, que exija baixo custo computacional para ser utilizado pelos profissionais em suas próprias clínicas e com seu próprio material;
- Prover resultados simples e intuitivos para os profissionais de fisioterapia.

1.4 Estrutura do Documento

Este trabalho está dividido em quatro capítulos, incluindo este que conta com uma introdução a respeito do tema e objetivos do projeto. Posteriormente, o Capítulo 2 traz um estudo a respeito da marcha e demais conceitos necessários para a compreensão do sistema desenvolvido neste trabalho.

No Capítulo 3 são descritas as etapas do processo que resultou no sistema proposto, Capítulo 4 apresenta os resultados obtidos. Finalmente, o Capítulo 5 encerra o trabalho com uma discussão a respeito do que foi desenvolvido, possíveis melhoramentos e novas ideias para o projeto.

2 Fundamentação Teórica

Neste capítulo são apresentados conceitos, definições matemáticas e estatísticas necessárias à compreensão deste trabalho.

2.1 Breve História do Estudo da Marcha

O estudo de como o corpo se movimenta é um campo de pesquisa antigo. Documentações antigas que datam de 1000 A.C, demonstram a prática do que hoje são consideradas artes marciais altamente relacionadas ao movimento como o Kung Fu e o Tai Chi (11). Diversas pesquisas também foram realizadas na Grécia antiga por Zeno de Elea, Hipócrates e Aristóteles com o intuito de entender melhor a conexão entre o movimento e o músculo (11) (12) (13). Tais estudos culminaram no nascimento do período anatômico durante o império romano, continuando e atraindo a atenção de grandes nomes como Leonardo da Vinci (14) e Andreas Vesalius (15) na parte anatômica e Galileo Galilei (13), René Descartes (13), Sir Isaac Newton (13) na parte do movimento.

Uma nova era na análise da marcha surgiu em 1836 em Leipzig com o trabalho dos irmãos Weber, Wilhelm e Eduard. Com uma formação multidisciplinar em física, matemática, anatomia e fisiologia, os irmãos Weber introduziram os fundamentos científicos da mecânica da marcha humana (13). Através de uma série de experimentos, eles formularam um modelo matemático que resume a mecânica da marcha humana (13). Eles mediram e relataram a postura, a fase de balanço, o movimento do tronco, duração do passo e o comprimento passo (16). Apesar de sua contribuição ser eminente, nem todas as suas teorias foram aceitas e muitas delas invalidadas, apesar disso, a partir dos irmãos Weber, começaram as pesquisas que culminaram nos conhecimentos e técnicas que hoje existem.

2.2 Marcha

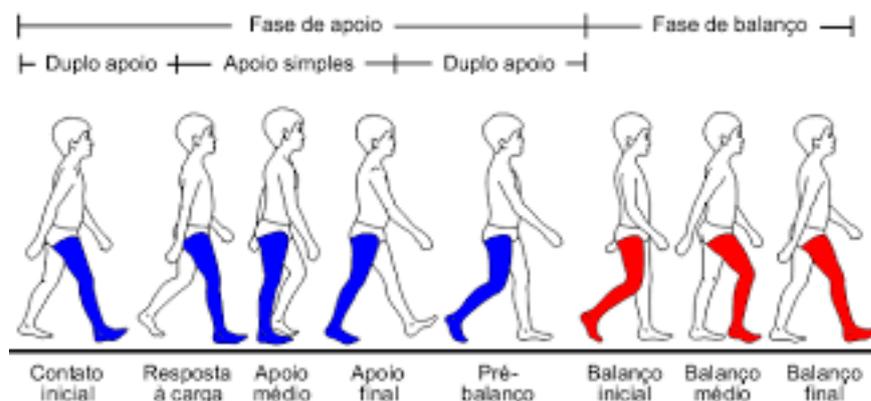
A marcha é uma atividade complexa que envolve o sistema nervoso central e periférico, e todo o sistema músculo-esquelético. É uma tarefa funcional que exige interações complexas e coordenação entre as articulações do corpo, especialmente da extremidade inferior do corpo humano. Antes do movimento em si, primeiro é necessário que haja um registro e ação do andar através do comando do sistema nervoso central e depois a sua transmissão para o sistema nervoso periférico. Em seguida, músculos se contraem e desenvolvem uma tensão para a geração de momentos e forças sobre as articulações sinoviais. Depois, acontece a regulação das forças e momentos nas articulações pelos seg-

mentos esqueléticos. E por fim os segmentos deslocam-se, procurando um andar funcional, promovendo as forças de reação no solo (17).

A marcha humana é considerada uma atividade cíclica, pois pode ser decomposta em uma série de eventos discretos, que começam com o contato inicial de um pé com o chão e termina com o contato subsequente desse mesmo pé com o chão. O ciclo de marcha é dividido em duas fases: a fase de apoio e a fase de balanço. A fase de apoio é a parte do ciclo em que o pé está em contato com o solo e corresponde a cerca de 60% do ciclo de marcha de um adulto normal (18). A fase de balanço é a parte em que o pé não está em contato com o solo e representa 40% do ciclo.

A fase de apoio subdivide-se em cinco eventos distintos: contato inicial (duplo-apoio inicial); resposta à carga (suporte simples); apoio médio (suporte simples); apoio terminal (suporte simples); pré-balanço (duplo-apoio).

Figura 1 – Fases da marcha



Fonte: extraído de (19).

Logo depois, começa a fase de balanço. Essa fase é caracterizada por três eventos distintos: balanço inicial; balanço médio; balanço terminal. Ao término da fase de balanço o ciclo de marcha está completo e pode ser reiniciado a qualquer momento. Além do ciclo de marcha, existem outras medidas lineares relacionadas ao passo ou a variáveis temporais. Alguns exemplos são: a cadência, o número de passos em um determinado intervalo de tempo, o comprimento do passo, o comprimento da passada e a velocidade de marcha (20). Essas variáveis de tempo e distância fornecem uma gama de informações sobre o padrão de marcha de um certo indivíduo. Por isso, qualquer desvio dessas variáveis dos valores normais provoca uma ligeira variação no padrão de marcha saudável. Mesmo que a marcha seja uma atividade comum e com fases bem definidas, existe uma ligeira variação no padrão de um indivíduo para outro.

A análise de marcha tem vital importância na compreensão e tratamento de

distúrbios da locomoção, podendo ser aplicada em pesquisas clínicas, fisioterapia, medicina esportiva, reabilitação e outros. Sua importância é tal que justificou a criação dos chamados Laboratórios de Movimento (21). O método mais natural de análise de marcha é a observação clínica, tal método só exige a presença do paciente e do profissional de saúde. Uma sofisticação da análise observacional é a filmagem da marcha do indivíduo. Permitindo ao especialista a possibilidade de repetir a observação várias vezes e comparar a situação do indivíduo antes e depois do tratamento. Esses métodos convencionais muitas vezes não apresentam alto índice de precisão, já que os resultados dependem exclusivamente do conhecimento e habilidade do médico, tornando assim a conclusão subjetiva. Várias outras abordagens também são adotadas na análise de marcha como: eletrogoniômetros, sensores ultra sônicos, sistemas de visão computacional para análise de imagens, marcadores subcutâneos dentre outros (22).

Atualmente, os métodos para a análise incluem a colocação de marcadores externos sobre o paciente, o monitoramento da marcha e o seu registro em vídeo, juntamente com o registro de sinais eletromiográficos (EMG). Em um instante obtêm-se os parâmetros gerais da marcha e realiza-se a análise cinemática, tal análise é feita por câmeras conectadas a um computador que definem os movimentos das articulações nos membros inferiores. Com os dados da análise cinemática e com a determinação das forças que cada pé exerce sobre o solo consegue-se calcular os momentos (13), gerando um relatório completo do estado da marcha do paciente. No entanto, o estudo dessa área ainda é restrita no Brasil, especialmente por causa dos altos custos relacionados a montagem de um centro de pesquisa.

Em um estudo abrangente, Muro-de-la-Herran *et al.* (2014) classificou sistemas convencionais de análise de marcha como *wearable* e *nonwearable*. Sistemas *nonwearable* são implementados no confinamento de um laboratório onde o movimento é capturado de forma bem definida e gravado através de vários sensores controlados por um estação computadorizada. Por outro lado, sistemas *wearable* libertam o clínico e o indivíduo do confinamento do laboratório, já que os sensores são montados no indivíduo, permitindo assim a gravação dos dados de marcha durante as atividades do dia-a-dia.

A popularização de *smartphones* e *wearable devices* fez com que a configuração para coleta de dados de análise de marcha se tornasse não intrusiva e ubíqua, com isso várias pesquisas foram desenvolvidas em torno de sistemas de reconhecimento de marcha (9). A maioria dos sensores de movimento *wearable* usam sensores inerciais baseados em *Micro Electro Mechanical Systems* (MEMS). Esses sensores inerciais, também conhecidos como Unidades de Medida Inercial (IMUs), incluem goniômetros, acelerômetros, sensores ultra-sônicos, dinamômetros, sensores de pressão plantar, entre outros e são capazes de medir a aceleração, a taxa angular e o vetor de campo magnético em um sistema de coordenadas local tridimensional.

Um conceito importante e necessário para o entendimento deste trabalho é o conceito de cinemática. Cinemática é o ramo da engenharia mecânica em que o movimento dos corpos é descrito sem considerar as forças responsáveis pelo movimento (23) (24). Na análise marcha, a cinemática se concentra no estudo do movimento entre as partes do corpo. Parâmetros cinemáticos incluem medições de ângulos articulares diartrodiais, deslocamentos, velocidades e acelerações. Além disso, os dados cinemáticos podem ser complementados com variáveis temporais e de passo que incluem cadência, velocidade de caminhada, tempo de passada, comprimento da passada, tempo do passo, comprimento do passo, período de suporte do membro único e período de suporte de duplo membro.

O uso mais frequente da análise da marcha é como auxílio na tomada de decisão, tal análise complementa as informações iniciais obtidas através de cuidadosos exames físicos. A interpretação dos dados coletados requer uma equipe multidisciplinar com experiência em engenharia, cinesiologia, fisioterapia e medicina. Os métodos empregados envolvem a apresentação de dados de forma compreensível e avaliação de estudos anteriores, para definir uma história natural da condição em estudo.

2.3 Redes Neurais

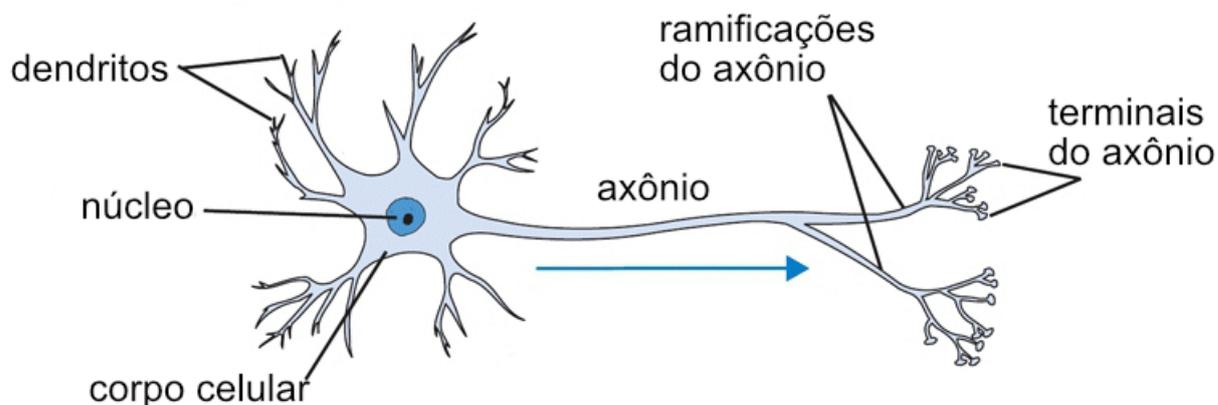
Segundo Kohonen *et al.* (1988) redes neurais artificiais (RNA) são redes paralelas massivamente interconectadas de elementos simples (geralmente adaptativos) com organizações hierárquicas que interagem com objetos do mundo real da mesma maneira que sistema nervoso biológico faz.

Da definição de Kohonen, percebe-se que o estudo de RNA baseia-se no sucesso de sistemas biológicos, que consistem em células nervosas muito simples e numerosas, que funcionam maciçamente em paralelo e possuem a capacidade de aprender. Um resultado deste procedimento de aprendizagem é a capacidade de generalizar e associar dados. Sendo assim, após o treinamento, uma rede neural pode encontrar soluções razoáveis para problemas semelhantes que não foi explicitamente treinada para resolver.

O cérebro humano é capaz de executar milhões de operações em paralelo e é formado por mais ou menos 10^{11} neurônios que continuamente se re-organizam ou são reorganizados por influências externas. Um neurônio nada mais é do que um interruptor com informações de entrada e saída como mostrado na Figura 2. O interruptor será ativado se houver estímulos suficientes de outros neurônios atingindo um limiar, resultando na saída de informação, um pulso que é enviado para outros neurônios.

As RNA funcionam de maneira análoga ao cérebro humano, também sendo constituídas de neurônios artificiais, que ao receberem os dados de entrada, têm uma função de ativação processando a entrada líquida que vai fazer o neurônio disparar, ou não, de acordo com um limiar.

Figura 2 – Ilustração de um neurônio biológico



Fonte: extraído de (25).

O processo de aprendizado de uma RNA pode acontecer por novas conexões entre neurônios, remoção de conexões, atualização de conexões, modificação do valor limiar dos neurônios, adicionando neurônios ou removendo neurônios (25). Essas mudanças podem ser alcançadas pela atualização dos pesos, tais atualizações são feitas iterativamente por um algoritmo de treinamento que segue um dos seguintes paradigmas:

- **Aprendizado Supervisionado:** é apresentado um conjunto de treino, consistindo de entradas e correspondentes saídas desejadas.
- **Aprendizado por Reforço:** para cada entrada apresentada, é produzida uma indicação (reforço) sobre a adequação das saídas correspondentes produzidas pela rede.
- **Aprendizado Não-Supervisionado:** a rede tem seus pesos ajustados sem a utilização de pares de entradas-saídas desejadas e sem indicações sobre a adequação de suas saídas correspondentes produzidas pela rede.

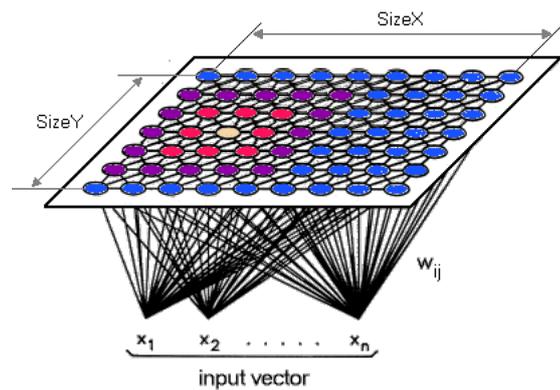
2.3.1 Self Organizing Maps

Teuvo Kohonen desenvolveu nos anos oitenta a *Self-Organizing feature Map*, brevemente referido como *self organizing maps* ou SOMs (26). Essa técnica é um modelo matemático abstrato inicialmente inspirado no córtex cerebral humano. A percepção visual humana e o cérebro compõem o sistema de cognição, a informação visual é processada tanto na retina quanto no cérebro, mas acredita-se que a maioria do processamento é feito

na retina, como extração de linhas, ângulos, curvas, contrastes, cores e movimento, nesse modelo a retina é considerada um tipo complexo de rede neural.

Essa técnica é um método de classificação de aprendizado completamente não supervisionado. Transformando uma entrada de alta dimensão (N dimensões) em um grid de células de baixa dimensão (dimensões G) definindo a topologia da rede.

Figura 3 – Mapa de kohonen de duas dimensões



Fonte: Obtido em <<http://bit.ly/2KUfByp>> (Acessado em 29/11/2018) .

A saída da rede é um grid de nós conectados às entradas através de um conjunto de pesos, representado na Figura 3. Esse algoritmo difere dos algoritmos de clustering tradicionais em duas características significativas, sendo a primeira a maneira pela qual os grupos são definidos. Embora o resultado no final da análise do SOM seja alguma forma de agrupamento de dados, ao contrário de um algoritmo de agrupamento, a metodologia básica do SOM não está primariamente relacionada ao agrupamento de dados ou à identificação de agrupamentos. A rede tenta encontrar nodes (nós) ou pontos no espaço de medição que são representativos as observações próximas e, quando consideradas em conjunto, descrevem a função de distribuição multidimensional do conjunto de dados.

Ao ser apresentada ao vetor de entrada, a rede calcula a ativação de cada nó usando:

$$\alpha_{i,j,k} = W_{i,j} P_k \quad (2.1)$$

Onde, $\alpha_{i,j,k}$ e $W_{i,j}$ são o valor de ativação e o peso, conectando P_k ao nó. P_k é mapeado ao nó com o maior valor de ativação. Cada neurônio da camada topológica é completamente conectado com a camada de entrada, os vetores de peso dessas conexões formam o protótipo associado a cada neurônio. Em cada etapa de treinamento, um vetor de amostra x do conjunto de dados de entrada é escolhido e uma medida de similaridade é calculada entre ele e todos os vetores de peso do mapa. A *best-matching unit* (BMU)

é a unidade cujo vetor de peso tem a maior semelhança com a amostra de entrada. A semelhança é geralmente definida por meio de uma medida de distância, tipicamente euclidiana:

$$d = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2.2)$$

Depois que o elemento de processamento vencedor é escolhido, o vetor de pesos é atualizado de acordo com a lei de aprendizagem usada. SOM difere da aprendizagem competitiva básica de modo que, em vez de adaptar apenas o elemento de processamento vencedor, os vizinhos do elemento de processamento vencedor também são adaptados (26). A propriedade de auto-organização da rede é baseada na vizinhança do elemento vencedor. (Hecht-Nielsen, 1990). Durante o treinamento os pesos são atualizados da seguinte forma:

$$x_{t+1} = x_i(t) + h_{ci}(t) \cdot [x(t) - x_i(t)] \quad (2.3)$$

Onde t é o tempo, $x(t)$ é o vetor de entrada randomicamente desenhado dos dados de entrada e $h_{ci}(t)$ é o centro da vizinhança em torno da unidade vencedora c no momento t . Esse último termo é uma função não crescente do tempo e da distância da unidade i da BMU, geralmente é composta por dois componentes: a função da taxa de aprendizagem $\alpha(t)$ e a função de vizinhança:

$$h_{ci}(t) = \alpha(t) \cdot h(\|r_c - r_i\|, t) \quad (2.4)$$

onde r_i é a localização da unidade i no grid.

À medida que a aprendizagem prossegue e novos vetores de entrada são dados ao mapa, a taxa de aprendizado $\alpha(t)$ gradualmente diminui para zero, de acordo com o tipo de função de taxa de aprendizagem especificada.

O algoritmo SOM realiza os seguintes passos (26):

1. Aleatoriamente seleciona um vetor de entrada;
2. Encontra a similaridade entre o vetor de entrada e os neurônios de acordo com a Equação 2.2;
3. Move o neurônio na direção do vetor de pesos com a menor distância de acordo com a Equação 2.3;
4. Repete esse procedimento até que o número máximo de iterações seja atingido.

2.4 Técnicas de Agrupamento

2.4.1 K-Means

O algoritmo K-Means é frequentemente utilizado na literatura de agrupamento de dados, sempre sendo utilizado como base de comparação com outros algoritmos. Trata-se de um algoritmo de agrupamento genérico bastante simples, porém bastante eficiente e de baixo custo computacional, tornando-o bastante popular. O algoritmo convencional é descrito abaixo (27) (28).

Para solucionar o problema de agrupamento o algoritmo K-Means utiliza uma função de similaridade para agrupar instâncias semelhantes. A similaridade entre duas instâncias é muitas vezes mensurada através de uma função de distância. A versão clássica do algoritmo K-Means utiliza como função de similaridade a função de distância euclidiana, Equação 2.2.

Primeiramente o algoritmo inicializa cada um dos k centróides c_k de forma aleatória, em seguida cada instância x_i é adicionada ao grupo cuja distância ao centróide c_k é mínima. Uma das etapas principais do algoritmo é o recálculo dos centróides de acordo:

$$c_k = \frac{1}{n_k} \sum_{i=1}^{n_k} x_i^k \quad (2.5)$$

A execução do algoritmo chega ao fim quando um critério de parada é atingido. O critério de parada normalmente é modelado como uma função de erro. Diferentes implementações do K-Means utilizam diferentes critérios de parada, como por exemplo estagnação de centróides ou quando nenhuma instância x_i mudar de grupo (16). A versão tradicional do algoritmo KMeans utiliza como função a ser minimizada a Equação 2.6, que representa a soma dos quadrados das distâncias entre cada uma das instâncias x_i e seus respectivos centróides.

$$e = \sum_{k=1}^K \sum_{i=1}^{n_k} |x_i^k - c_k|^2 \quad (2.6)$$

O algoritmo K-Means utilizado neste projeto realiza as seguintes ações:

1. Seleciona K pontos para inicializar o centróides;
2. Para cada entrada na base de dados, atribui a entrada ao grupo cuja a centróide é mínima;
3. Atualiza os centróides de acordo com a Equação 2.5;
4. Repete esse procedimento até que o critério de para seja atingido.

2.4.2 Fuzzy C-Means

Fuzzy c-means é um tipo de algoritmo de clustering, em que cada ponto de dados pode pertencer a mais de um cluster. A participação de cada dado em cada cluster representa o grau em que o dado pertence a esse cluster, representada pela função de minimização descrita na equação 2.7.

$$J_m = \sum_{I=1}^D \sum_{I=1}^N u_{ij}^m \|x_i - c_j\|^2 \quad (2.7)$$

onde D é número de dados de entrada, N é o número de clusters, m é valor que controla o grau de *fuzziness*. x_i é o i th elemento de entrada, c_j é o centro do j th cluster e u_{ij} é o grau de pertencimento de x_i no cluster j .

Por fim, os dados serão atribuídos ao cluster de forma que sua participação nesse cluster seja a maior. Assim, os itens no mesmo cluster são sempre semelhantes em comparação com itens em clusters diferentes. Clusters são identificados por características de similaridade, tais recursos podem ser distância e conectividade. O FCM executa as seguintes etapas durante o processo de agrupamento:

1. Aleatoriamente inicializa os valores de pertencimento dos clusters, u_{ij} , de acordo com a Equação 2.8;
2. Calcula o centro dos clusters;
3. Atualiza os valores de u_{ij} de acordo com a Equação 2.9;
4. Calcula a função objetivo J_m , apresentada na Equação 2.7;
5. Repete esses procedimentos até que a condição de parada seja atingida.

$$c_j = \frac{\sum_{I=1}^D u_{ij}^m x_i}{\sum_{I=1}^D u_{ij}^m} \quad (2.8)$$

$$u_{ij}^m = \frac{1}{\sum_{k=1}^N \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}} \quad (2.9)$$

3 Métodos

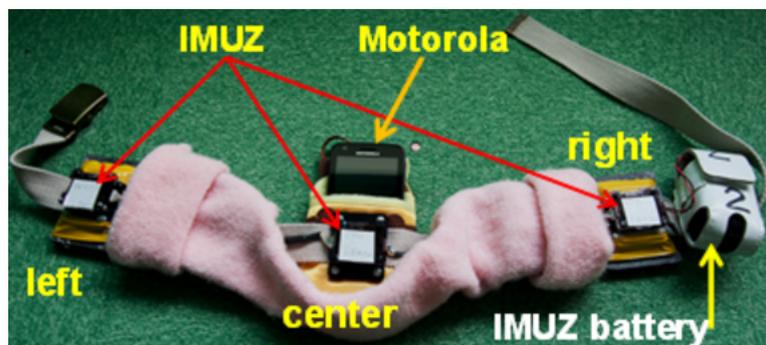
Para endereçar o problema foram implementados algoritmos híbridos entre os algoritmos Self organizing maps (SOM), Fuzzy C-Means (FCM) e K-Means (26) (29). Os mesmos foram comparados com experimentos realizados com o algoritmo tradicional do FCM (10). Além disso foram realizadas análises de desempenho entre os algoritmos propostos, destacando o algoritmo mais apropriado na classificação de indivíduos. Pretendeu-se com isso a elaboração de um modelo que possa ser aplicado por profissionais de fisioterapia em suas próprias clínicas como meio de auxiliar no diagnóstico de disfunções de marcha. Todos os algoritmos foram implementados na linguagem de programação Python.

3.1 Aquisição dos Dados

Os dados utilizados foram coletados da base de dados de marcha OU-ISIR, que destina-se a auxiliar pesquisadores na área geral de desenvolvimento, teste e avaliação de algoritmos para reconhecimento de marcha. A base consiste de indivíduos saudáveis, de ambos os gêneros e de idades variadas, totalizando 180 amostras divididas em 6 grupos de acordo com a faixa etária. Incluindo indivíduos de menos de 10 anos de idade; de 10-19 anos de idade; 20-29 anos de idade; 30-39 anos de idade; 40-49 anos de idade e indivíduos acima de 50 anos de idade, mas não maior que 75 anos de idade.

Os dados de marcha foram adquiridos por quatro sensores. Três IMUs, cada qual incluindo um acelerômetro triaxial e um quarto acelerômetro KXTF9 triaxial incorporado em um smartphone. A configuração da IMU consistiu em: $\pm 4g$ para acelerômetros e $\pm 500\text{deg/s}$ para giroscópios, além disso todos os quatro sensores funcionaram a 100 Hz. A IMU central e o smartphone foram colocados na parte central da cintura dos participante, os sensores esquerdo e direito foram colocados em ambos os lados da cintura respectivos, a configuração do sistema está representada na Figura 4. Durante a fase de teste, os indivíduos realizaram uma caminhada por um caminho plano, ao mesmo tempo, as IMUs capturavam uma sequência de sinais, correspondendo ao giroscópio de 3 eixos e aos sinais do acelerômetro de 3 eixos (10). Os indivíduos foram agrupados em seis grupos de acordo com a faixa etária: Grupo A consiste em indivíduos com menos de 10 anos de idade; Grupo B, 10-19 anos de idade; Grupo C, 20-29; Grupo D, 30-39; Grupo E, 40-49; Grupo F, indivíduos acima de 50 anos de idade.

Figura 4 – Configuração do sistema de captura marcha



Fonte: Extraído de (30)

3.2 Pré-Processamento

Com os dados coletados pelos sensores inerciais ainda é necessário que alguns dados sejam extraídos e calculados para que as informações necessárias para a análise de marcha estejam completas, como por exemplo: o tempo médio de ciclo (**ACT**), o tempo médio de fase de apoio e o tempo médio de balanço. Para extrair os parâmetros restantes a partir dos dados coletados pelos sensores Caldas *et al.* (2016) aplicaram nos dados um filtro passa-baixo e uma frequência de corte de 8Hz, para suavizar o sinal e minimizar os ruídos.

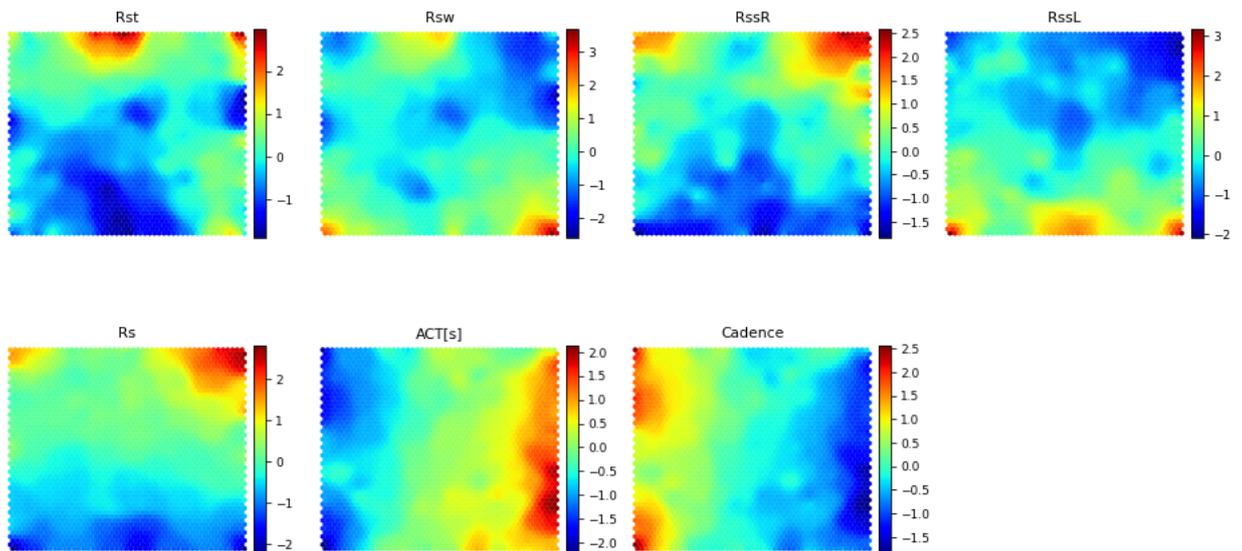
Com esses dados suavizados e sem ruídos foi possível calcular as relações de simetria da marcha, a **ACT** e a **Cadência** (30). Extraído por fim, todos os dados a serem utilizados como entrada dos algoritmos. Sobre cada indivíduo, foram coletadas sete variáveis temporais (comumente chamadas de recursos, do inglês *features*) inicialmente consideradas relevantes. Cinco delas são variáveis de simetria: **Rst**; **Rsw**; **RssR** e **RssL**; **Rs**. Duas delas variáveis cinemáticas: **ACT** e **Cadência**. Cada variável representa um aspecto diferente extraído da análise de marcha. A simetria entre duas pernas na fase de apoio (**Rst**), e de balanço (**Rsw**). A razão geral de simetria (**Rs**). Por fim, as simetrias em cada uma das pernas durante a fase de apoio são (**RssR**, **RssL**). As variáveis cinemáticas são o tempo médio de ciclo (**ACT**) e o número de passos sobre o tempo (**Cadência**).

3.3 Extração de Características

O objetivo dessa etapa é representar o conjunto de dados iniciais exclusivamente com as características necessárias para distinguir um padrão de outro. Para realizar essa tarefa a técnica SOM foi utilizada. Após a coleta e o pré-processamento dos dados, todas as sete características foram passadas como entrada do algoritmo SOM (31).

A saída da rede é um *heatmap* representando o padrão de aprendizado da rede a respeito de cada recurso de entrada. A partir desse gráfico pode-se concluir que alguns recursos são mais relevantes para o processo de aprendizado da rede que outros.

Figura 5 – Exemplo de heatmap das 7 variáveis de entrada



Fonte: Autoria própria

O mapa de calor da Figura 5, representa uma grid hexagonal com um tamanho de mapa de 50x50. O mapa de calor é feito de múltiplos "nós", cada nó pode representar várias amostras de entrada. O azul claro identifica áreas onde o valor do mapa é 0, revelando a presença de nós vazios no mapa. A cor vermelha identifica concentrações de indivíduos de um grupo, enquanto a cor azul escuro identifica o outro grupo. A análise do heatmap gerado pela técnica SOM permite que diferentes características sejam selecionadas de acordo com os grupos a serem classificados. Por exemplo, valores de *Cadência* no canto superior esquerdo só podem pertencer a um dos grupos. No entanto, os valores do canto inferior direito podem pertencer ao grupo oposto.

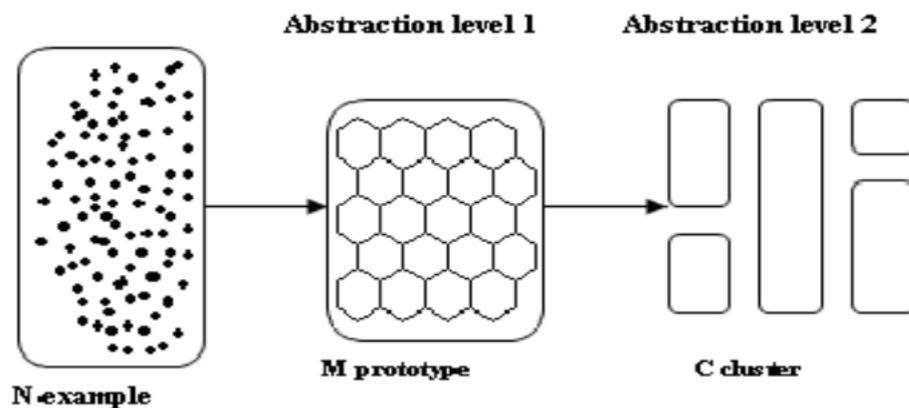
A fim de comprovar os resultados obtidos pela análise preliminar dos heatmaps todos as combinações de experimentos foram realizados e os resultados serão discutidos no Capítulo 4.

3.4 Processamento

Para construir um modelo híbrido, várias técnicas são geralmente combinadas em dois estágios. O primeiro estágio é baseado em técnicas de agrupamento ou de classificação,

que podem ser usadas para pré-processar os dados. Em seguida, a saída do primeiro estágio (ou seja, os dados processados) é usada para construir o modelo de previsão ou agrupamento.

Figura 6 – Níveis de abstração de abordagens híbridas



Fonte: Extraído de (32)

Esta monografia explora o uso de SOM seguido por algoritmos de agrupamento, que neste caso são o FC-Means e o K-Means. Os dados de entrada são usados no primeiro nível de clustering usando a técnica SOM. O grid de saída é fornecido a um dos algoritmos de clustering FC-Means/K-menas para processamento adicional. Tais algoritmos são usados para otimizar os clusters com base no clustering grosseiro realizado pela técnica SOM. O uso de um algoritmo de clustering recebendo como entrada os vetores de saída da técnica SOM, ao invés de agrupar diretamente os dados, é considerada uma abordagem computacionalmente eficaz [25].

Nesse trabalho, dois algoritmos híbridos foram desenvolvidos para classificar indivíduos por faixa etária.

3.4.1 SOM + FCM

O algoritmo SOM realiza seu processo adaptativo:

1. Aleatoriamente seleciona um vetor de entrada;
2. Encontra a similaridade entre o vetor de entrada e os neurônios de acordo com a Equação 2.2;
3. Move o neurônio na direção do vetor de pesos com a menor distância de acordo com a Equação 2.3;

4. Repete esse procedimento até que o número máximo de iterações seja atingido.

O FCM utiliza o mapa de saída do SOM como entrada e realiza a operação de agrupamento:

1. Aleatoriamente inicializa os valores de pertencimento dos clusters, u_{ij} , de acordo com a Equação 2.8;
2. Calcula o centro dos clusters;
3. Atualiza os valores de u_{ij} de acordo com a Equação 2.9;
4. Calcula a função objetivo J_m , apresentada na Equação 2.7;
5. Repete esses procedimentos até que a condição de parada seja atingida.

3.4.2 SOM + K-Means

SOM realiza seu processo adaptativo:

1. Aleatoriamente seleciona um vetor de entrada;
2. Encontra a similaridade entre o vetor de entrada e os neurônios de acordo com a Equação 2.2;
3. Move o neurônio na direção do vetor de pesos com a menor distância de acordo com a Equação 2.3;
4. Repete esse procedimento até que o número máximo de iterações seja atingido.

K-Means utiliza o mapa de saída do SOM como entrada e realiza a operação de agrupamento:

1. Utiliza o grid de saída da SOM como valores iniciais;
2. Escolhe-se um número de clusters desejados K ;
3. Para cada entrada na base de dados, atribui a entrada ao grupo cuja a centróide é mínima;
4. Atualiza os centróides de acordo com a Equação 2.5;
5. Repete esse procedimento até que o critério de para seja atingido.

3.5 Experimentos

Uma série de experimentos foi realizada de forma a confirmar a eficiência dos algoritmos desenvolvidos. Os experimentos se basearam em outros anteriores realizados por Caldas *et al.* (2016), onde antes da realização dos experimentos, uma série de testes estatísticos foram realizados.

Os testes de Kolmogorov-Smirnov e *Student's t-test* foram aplicados para checar a normalidade da distribuição dos dados e as semelhanças entre os grupos etários. Os resultados mostraram que os dados apresentam uma distribuição normal. Os resultados também mostraram que os grupos A e C (0-10 e 20-29) apresentam uma diferença significativa quando comparados um com outro, em seis dos sete parâmetros disponíveis. De maneira semelhante, porém contrária, os grupos E e F (40-49 e >50) não apresentaram diferença significativa quando comparados entre si, exceto em um dos sete parâmetros disponíveis (10).

Os experimentos foram realizados levando em consideração os grupos de idade e a quantidade de características disponíveis. Como entrada dos algoritmos pode-se usar alguma variação das cinco variáveis de simetria (Rst, Rsw, RssR, RssL e Rs) e duas variáveis cinemáticas (ACT e Cadência) dos 180 sujeitos aleatoriamente selecionados. Inicialmente, todos os seis grupos de idade foram comparados, juntamente com os sete parâmetros de entrada coletados pelas IMUs. Em seguida, experimentos foram realizados com os grupos de maior e menor similaridade, ou seja, grupos A e C e grupos E e F. Em relação as abordagens híbridas, os algoritmos admitem qualquer combinação de recursos de entrada, assim, foram realizados experimentos variando de dois a sete parâmetros de entrada para os dois grupos, com o objetivo de validar a ideia anterior de que algumas características são mais importantes que outras no processo de classificação. A abrangência de cenários resultou em 704 experimentos.

Com base em experimentos anteriores e preliminares, foi estabelecido as seguintes configurações: Tamanho do mapa - 50x50; Número de iterações - 2000; Base de dados de treinamento e teste com 80% e 20% das amostras, respectivamente (10).

3.6 Métricas

A precisão foi avaliada usando as métricas Precisão (P), *Recall* (R) e *F1-score* (F1), Acurácia (A), definidos em função de positivos verdadeiros (TP), falsos positivos (FP), e Falsos Negativos (FN) (33) (34).

$$Precisão(P) = \frac{TP}{(TP + FP)} \quad (3.1)$$

$$\text{Recall}(R) = \frac{TP}{(TP + FN)} \quad (3.2)$$

$$\text{F1-score}(F1) = \frac{2PR}{(P + R)} \quad (3.3)$$

$$\text{Acurácia}(A) = \frac{TP + TN}{(TP + FP + TN + FN)} \quad (3.4)$$

4 Resultados

Este capítulo tem como objetivo apresentar os resultados obtidos com os dois algoritmos híbridos produzidos (SOM + FCM e SOM + K-Means) para o problema de classificação de indivíduos por faixa etária a partir de dados coletados por sensores inerciais. Os resultados estão organizados em cenários e foram comparados com os resultados dos experimentos reproduzidos por Wang e Caldas (2018). Por fim, dois aspectos serão analisadas: a eficiência dos algoritmos na tarefa de classificação e a relevância dos recursos disponíveis, a fim de avaliar os resultados obtidos pela análise preliminar dos heatmaps discutidos na Seção 3.3.

4.1 Cenário 2G7F

Nas próximas seções, será realizada uma análise das métricas sobre os resultados obtidos pelos experimentos que realizaram a classificação de somente dois grupos, tendo como entradas as sete características disponíveis.

Os resultados a seguir refletem os resultados entre os grupos A e C realizados por (10) (30) considerando apenas duas variáveis de entrada. Ambos os resultados apresentam um alto índice de precisão na classificação dos grupos A e C, em contrapartida os resultados encontrados entre os grupos E e F apresentaram uma melhora significativa nos critérios de avaliação.

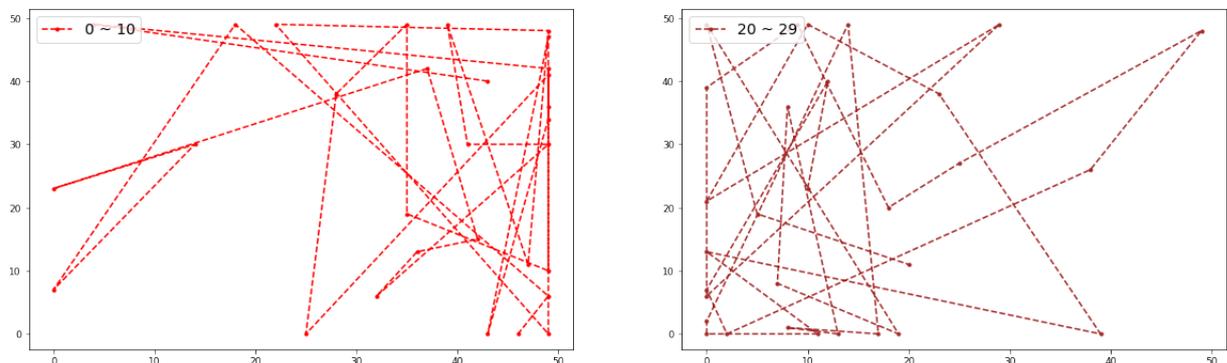
4.1.1 Resultados do algoritmo SOM + K-Means

O algoritmo desmonstrou um índice acima de 80% na classificação de dois grupos: A e C, como descrito na Tabela 1. Tais grupos apresentam uma diferença significativa quando comparados um com o outro, concluindo-se que a rede obteve um desempenho bom em separar grupos diferentes. Ao analisar a distribuição dos dados de entrada no mapa de saída do algoritmo na Figura 7, nota-se uma clara distribuição de grupos. O mapa de calor da Figura 8 representa os padrões de aprendizado da rede neural SOM em relação aos recursos de entrada disponíveis, os mapas de calor serão abordados detalhadamente na Seção 4.7.

Tabela 1 – Resultado do algoritmo SOM + K-Means para classificação dos Grupos A e C

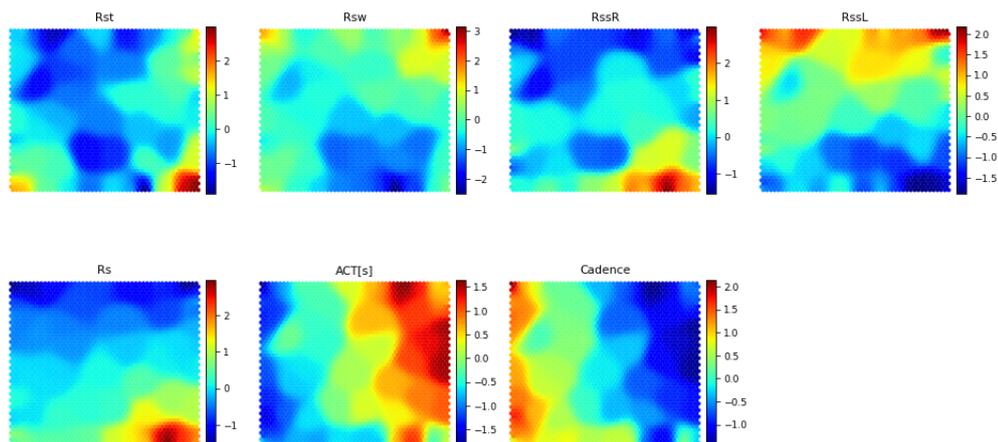
Parâmetros de entrada	Precisão	Recall	F1-Score	Acurácia
ACT[s] Cadência Rs RssR Rst Rsw RssL	0,8571	0,8000	0,8000	0,7916

Figura 7 – Distribuição dos dados dos grupos A e C para SOM + K-Means



Fonte: Autoria própria

Figura 8 – Heatmap dos padrões de aprendizado dos grupos A e C



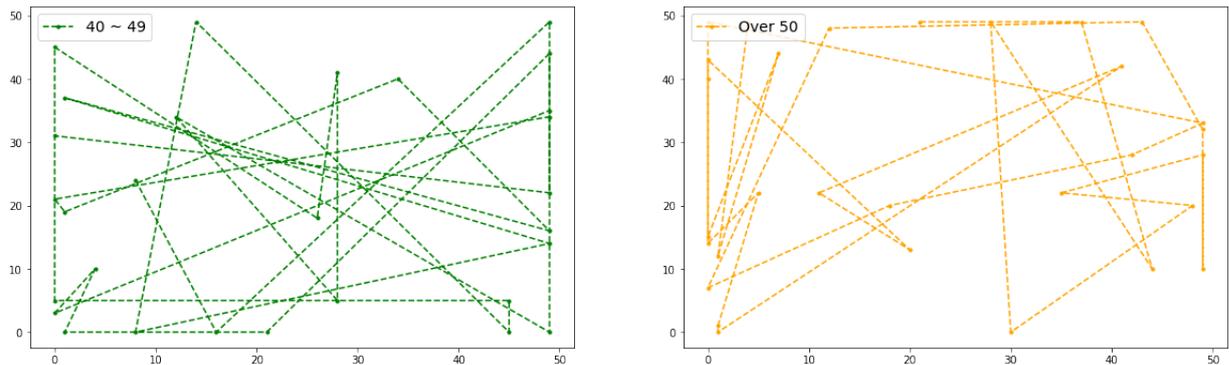
Fonte: Autoria própria

Na classificação dos grupos E e F, o algoritmo apresentou índices menores como descrito na Tabela 2. No entanto, como esses grupos não apresentam uma diferença significativa estatisticamente, pode-se concluir que o algoritmo também apresenta um bom resultado. Diferentemente do experimento com os grupos A e C, a distribuição dos dados de entrada representada na Figura 9 não mostra uma clara divisão entre os grupos. O mapa de calor da Figura 10 representa os padrões de aprendizado da rede neural SOM em relação aos recursos de entrada disponíveis, os mapas de calor serão abordados detalhadamente na Seção 4.7.

Tabela 2 – Resultado do algoritmo SOM + K-Means para classificação dos Grupos E e F

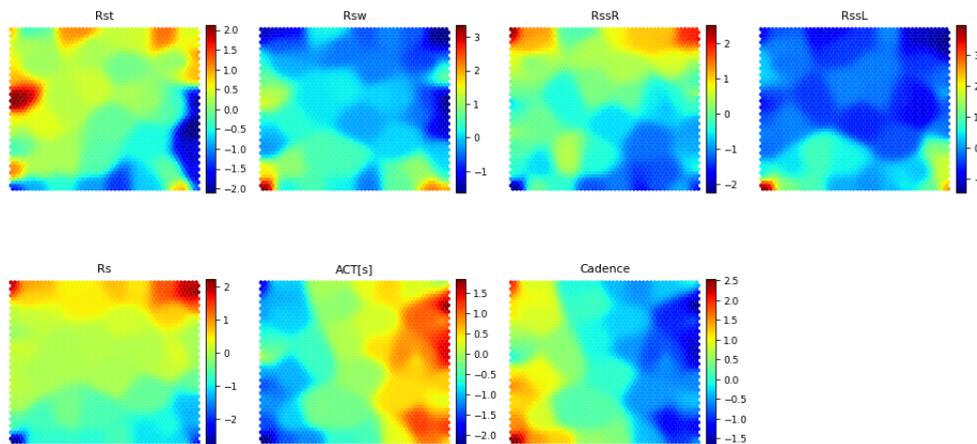
Parâmetros de entrada						Precisão	Recall	F1-Score	Acurácia	
ACT[s]	Cadência	Rs	RssR	Rst	Rsw	RssL	0,7083	0,7000	0,7000	0,6969

Figura 9 – Distribuição dos dados dos grupos E e F para SOM + K-Means



Fonte: Autoria própria

Figura 10 – Heatmap dos padrões de aprendizado dos grupos E e F



Fonte: Autoria própria

4.1.2 Resultados do algoritmo SOM + FC-Means

Observando as Tabelas 3 e 1 pode-se constatar que no caso dos grupos A e C o resultado é exatamente o mesmo. Conclui-se novamente que ambos os algoritmos possuem habilidades de generalização parecida, especialmente quando os grupos são significativamente diferentes. Realizando a mesma comparação com as Tabelas 4 e 2

percebe-se que o algoritmo SOM + K-Means é um pouco melhor na tarefa de classificar indivíduos parecidos. As Figuras 7 e 9 também refletem a distribuição de grupos para a técnica SOM + FC-Means.

Tabela 3 – Resultado do algoritmo SOM + FC-Means para classificação dos Grupos A e C

Parâmetros de entrada	Precisão	Recall	F1-Score	Acurácia
ACT[s] Cadência Rs RssR Rst Rsw RssL	0,8571	0,8000	0,8000	0,7916

Tabela 4 – Resultado do algoritmo SOM + FC-Means para classificação dos Grupos E e F

Parâmetros de entrada	Precisão	Recall	F1-Score	Acurácia
ACT[s] Cadência Rs RssR Rst Rsw RssL	0,6190	0,6000	0,6000	0,5899

4.2 Cenário 2G6F

Nas próximas seções, será realizada uma análise das métricas sobre os resultados obtidos pelos experimentos que realizaram a classificação de somente dois grupos. Os grupos escolhidos foram os com maior e menor índice de similaridade e por fim recebendo como entradas seis características. Todas as seis combinações possíveis foram usadas como entrada dos algoritmos.

Da mesma forma que na Seção 4.1 os resultados a seguir também refletem os resultados entre os grupos A e C realizados por [10] [30] considerando apenas duas variáveis de entrada. Ambos os resultados apresentam um alto índice de precisão na classificação dos grupos A e C, em contrapartida os resultados encontrados entre os grupos E e F apresentaram uma melhora significativa (considerando o melhor resultado).

4.2.1 Resultados do algoritmo SOM + K-Means

Nas Tabelas 5 e 6 estão descritos os resultados dos experimentos realizados. Percebe-se que o algoritmo apresenta maiores índices na classificação dos grupos A e C e que entre os grupos E e F. Na classificação dos grupos A e C percebe-se uma relação entre a **Cadência** e os melhores resultados. Entre os grupos E e F percebe-se que uma sequência de entradas parece apresentar melhor resultado que as outras, a sequência sem o valor de **Rs**.

4.2.2 Resultados do algoritmo SOM + FC-Means

Nas Tabelas 7 e 17 estão descritos os resultados dos experimentos realizados. Assim como na Seção 4.1 os resultados de ambos os algoritmos são muito parecidos. No

Tabela 5 – Resultado do algoritmo SOM + K-Means para classificação dos Grupos A e C com 6 variáveis de entrada

Parâmetros de entrada	Precisão	Recall	F1-Score	Acurácia
ACT[s] Cadência Rs RssL RssR Rst	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência Rs RssL RssR Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência Rs RssL Rst Rsw	0,8571	0,8000	0,8000	0,7916
Cadência Rs RssL RssR Rst Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Rs RssL RssR Rst Rsw	0,3809	0,4000	0,4000	0,3750
ACT[s] Cadência RssL RssR Rst Rsw	0,8571	0,8000	0,8000	0,7916

Tabela 6 – Resultado do algoritmo SOM + K-Means para classificação dos Grupos E e F com 6 variáveis de entrada

Parâmetros de entrada	Precisão	Recall	F1-Score	Acurácia
ACT[s] Cadência Rs RssL RssR Rst	0,6190	0,6000	0,6000	0,5899
ACT[s] Cadência Rs RssL RssR Rsw	0,6000	0,6000	0,6000	0,5993
ACT[s] Cadência Rs RssL Rst Rsw	0,6190	0,6000	0,6000	0,5833
Cadência Rs RssL RssR Rst Rsw	0,5000	0,5000	0,5000	0,4949
ACT[s] Rs RssL RssR Rst Rsw	0,6000	0,6000	0,6000	0,6000
ACT[s] Cadência RssL RssR Rst Rsw	0,7777	0,6000	0,6000	0,5238

experimento entre os grupos A e C obteve-se o maior índice de precisão deste trabalho 88%.

Tabela 7 – Resultado do algoritmo SOM + FC-Means para classificação dos Grupos A e C com 6 variáveis de entrada

Parâmetros de entrada	Precisão	Recall	F1-Score	Acurácia
ACT[s] Cadência Rs RssL RssR Rst	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência Rs RssL RssR Rsw	0,8824	0,8800	0,8800	0,8798
ACT[s] Cadência Rs RssL Rst Rsw	0,8571	0,8000	0,8000	0,7916
Cadência Rs RssL RssR Rst Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Rs RssL RssR Rst Rsw	0,3809	0,4000	0,4000	0,3750
ACT[s] Cadência RssL RssR Rst Rsw	0,8571	0,8000	0,8000	0,7916

4.3 Cenário 2G5F

Nas próximas seções, será realizada uma análise das métricas sobre os resultados obtidos pelos experimentos que realizaram a classificação de somente dois grupos. Os grupos escolhidos foram os com maior e menor índice de similaridade e por fim recebendo como entradas cinco características. Todas as 21 combinações possíveis foram usadas como entrada dos algoritmos.

Tabela 8 – Resultado do algoritmo SOM + FC-Means para classificação dos Grupos E e F com 6 variáveis de entrada

Parâmetros de entrada	Precisão	Recall	F1-Score	Acurácia
ACT[s] Cadência Rs RssL RssR Rst	0,6000	0,6000	0,6000	0,6000
ACT[s] Cadência Rs RssL RssR Rsw	0,6000	0,6000	0,6000	0,6000
ACT[s] Cadência Rs RssL Rst Rsw	0,6190	0,6000	0,6000	0,5833
Cadência Rs RssL RssR Rst Rsw	0,6000	0,6000	0,6000	0,6000
ACT[s] Rs RssL RssR Rst Rsw	0,6000	0,6000	0,6000	0,6000
ACT[s] Cadência RssL RssR Rst Rsw	0,7777	0,6000	0,6000	0,5238

Os dados referentes as Tabelas 9 e 11 mostram que existem uma clara dependência entre os grupos A e C e a variável cinemática *Cadência*. Todos os experimentos que utilizaram a *Cadência* obtiveram alto índice de precisão e acurácia. Em relação a classificação dos grupos E e F, nas Tabelas 10 e 12 pode-se perceber que os experimentos que possuíam *ACT[s] Rs Rsw* em suas entradas apresentaram um índice de precisão maior que os demais. Também se aplica a este cenário a mesma comparação com os trabalhos passado realizada na Seção 4.1 e 4.2.

4.3.1 Resultados do algoritmo SOM + K-Means

Nas Tabelas 9 e 10 estão descritos os resultados dos experimentos realizados.

4.3.2 Resultados do algoritmo SOM + FC-Means

Nas Tabelas 11 e 12 estão descritos os resultados dos experimentos realizados.

4.4 Cenário 2G4F

Os dados referentes aos experimentos que utilizaram os mesmos dois grupos e quarto características como entrada, não difere muito dos resultados da Seção 4.3. Experimentos com as 36 combinações possíveis foram realizados, porém na Tabela 13 só estão os quatro melhores. A tabela completa com os 144 experimentos se encontra no Apêndice A.

4.5 Cenário 2G3F

Os resultados dos experimentos com três recursos mantiveram-se no mesmo padrão dos resultados da Seção 4.4 e 4.3. Todos os experimentos que utilizaram a variável *Cadência* - *Cadência* obtiveram maiores índices de precisão e acurácia. A Tabela completa com o 136 experimentos se encontra no Apêndice A.

Tabela 9 – Resultado do algoritmo SOM + K-Means para classificação dos Grupos A e C com 5 variáveis de entrada

SOM + FKMEANS				
5 Combinações de parâmetros de entrada e experimentos entre grupo A e grupo C.				
Parâmetros de Entrada	Precisão	Recall	Acurácia	F1-score
ACT[s] Cadência Rs RssL RssR	0,8571	0,8000	0,8000	0,7916
Rs RssL RssR Rst Rsw	0,3809	0,4000	0,4000	0,375
ACT[s] Cadência Rs RssR Rst	0,8571	0,8000	0,8000	0,7916
Cadência Rs RssL RssR Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Rs RssL RssR Rsw	0,3809	0,4000	0,4000	0,375
ACT[s] Cadência RssL RssR Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Rs RssR Rst Rsw	0,5000	0,5000	0,5000	0,4949
Cadência Rs RssR Rst Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência RssR Rst Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Rs RssL Rst Rsw	0,2916	0,3000	0,3000	0,2929
Cadência Rs RssL RssR Rst	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência Rs RssL Rst	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência RssL Rst Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência Rs Rst Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência RssL RssR Rst	0,8571	0,8000	0,8000	0,7916
Cadência Rs RssL Rst Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência Rs RssR Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência Rs RssL Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Rs RssL RssR Rst	0,5000	0,5000	0,5000	0,4949
ACT[s] RssL RssR Rst Rsw	0,4000	0,4000	0,4000	0,4000
Cadência RssL RssR Rst Rsw	0,8571	0,8000	0,8000	0,7916

4.6 Cenário 2G2F

Os resultados dos experimentos com três recursos mantiveram-se no mesmo padrão dos resultados da Seção 4.4 e 4.3. Todos os experimentos que utilizaram a variável *Cadência* obtiveram maiores índices de precisão e acurácia. A Tabela completa com os 84 experimentos se encontra no Apêndice A.

4.7 Relevância das Características

Uma das propostas deste trabalho é a análise da relevância das variáveis de entrada, com o objetivo de gerar a representação do conjunto de dados iniciais exclusivamente com as características necessárias para distinguir um padrão de outro. Para isso a técnica SOM foi utilizada, preservando a topologia dos dados nos mapas produzidos. A partir do grid de saída da rede neural SOM, um mapa de calor é produzido para cada uma das variáveis de entrada do problema, esses mapas mostram que variáveis apresentaram maior importância durante o processo de aprendizado da rede. Para um experimento com informações de

Tabela 10 – Resultado do algoritmo SOM + K-Means para classificação dos Grupos E e F com 5 variáveis de entrada

SOM + FKMEANS				
5 Combinações de parâmetros de entrada e experimentos entre grupos E e grupo F.				
Parâmetros de Entrada	Precisão	Recall	Acurácia	F1-score
ACT[s] Cadência Rs RssL RssR	0,6000	0,6000	0,6000	0,6000
Rs RssL RssR Rst Rsw	0,6000	0,6000	0,6000	0,6000
ACT[s] Cadência Rs RssR Rst	0,2222	0,4000	0,4444	0,2857
Cadência Rs RssL RssR Rsw	0,5000	0,5000	0,5000	0,4949
ACT[s] Rs RssL RssR Rsw	0,7777	0,6000	0,6000	0,5238
ACT[s] Cadência RssL RssR Rsw	0,2222	0,4000	0,4444	0,2857
ACT[s] Rs RssR Rst Rsw	0,7777	0,6000	0,6000	0,5238
Cadência Rs RssR Rst Rsw	0,6559	0,6400	0,6400	0,6305
ACT[s] Cadência RssR Rst Rsw	0,2222	0,4000	0,4444	0,2857
ACT[s] Rs RssL Rst Rsw	0,7083	0,7000	0,7000	0,6969
Cadência Rs RssL RssR Rst	0,5000	0,5000	0,5000	0,4949
ACT[s] Cadência Rs RssL Rst	0,5000	0,5000	0,5000	0,4505
ACT[s] Cadência RssL Rst Rsw	0,2222	0,4000	0,4444	0,2857
ACT[s] Cadência Rs Rst Rsw	0,2222	0,4000	0,4444	0,2857
ACT[s] Cadência RssL RssR Rst	0,2222	0,4000	0,4444	0,2857
Cadência Rs RssL Rst Rsw	0,6190	0,6000	0,6000	0,5833
ACT[s] Cadência Rs RssR Rsw	0,5000	0,5000	0,5000	0,4505
ACT[s] Cadência Rs RssL Rsw	0,6190	0,6000	0,6000	0,5833
ACT[s] Rs RssL RssR Rst	0,4000	0,4000	0,4000	0,4000
ACT[s] RssL RssR Rst Rsw	0,6000	0,6000	0,6000	0,6000
Cadência RssL RssR Rst Rsw	0,6190	0,6000	0,6000	0,5833

somente um indivíduo, o mapa de calor gerado pela rede é igual ao mapa apresentado na Figura 11, esse mapa de calor contém um grid com todas as células possuindo o mesmo valor para cada recurso de entrada. Como apenas um indivíduo é usado, não existem impecilhos na classificação e os próprios valores de entrada representam os valores de indentificação da faixa etária do indivíduo.

Ao aplicar a mesma abordagem aos 60 indivíduos pertencentes aos grupos A e C o mapa resultante se transforma. A Figura 12 representa um heatmap gerado pela técnica SOM + K-Means para classificação dos grupos A e C, ou seja, grupos significativamente diferentes. No mapa de calor da Figura 12, o azul claro identifica áreas onde o valor do mapa é 0, revelando a presença de nós vazios no mapa. A cor vermelha identifica concentrações de indivíduos de um grupo, enquanto a cor azul escuro identifica o outro grupo. A análise do *heatmap* gerado pela técnica SOM permite que diferentes características sejam selecionadas de acordo com os grupos a serem classificados. Desse *heatmap* pode-se extrair as seguintes hipóteses:

Tabela 11 – Resultado do algoritmo SOM + FC-Means para classificação dos Grupos A e C com 5 variáveis de entrada

SOM + FKMEANS				
5 Combinações de parâmetros de entrada e experimentos entre grupo A e grupo C.				
Parâmetros de Entrada	Precisão	Recall	Acurácia	F1-score
ACT[s] Cadência Rs RssL RssR	0,8571	0,8000	0,8000	0,7916
Rs RssL RssR Rst Rsw	0,6190	0,6000	0,5833	0,6000
ACT[s] Cadência Rs RssR Rst	0,8571	0,8000	0,8000	0,7916
Cadência Rs RssL RssR Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Rs RssL RssR Rsw	0,3809	0,4000	0,4000	0,375
ACT[s] Cadência RssL RssR Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Rs RssR Rst Rsw	0,6000	0,6000	0,6000	0,6000
Cadência Rs RssR Rst Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência RssR Rst Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Rs RssL Rst Rsw	0,4000	0,4000	0,4000	0,4000
Cadência Rs RssL RssR Rst	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência Rs RssL Rst	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência RssL Rst Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência Rs Rst Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência RssL RssR Rst	0,8571	0,8000	0,8000	0,7916
Cadência Rs RssL Rst Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência Rs RssR Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência Rs RssL Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Rs RssL RssR Rst	0,5000	0,5000	0,5000	0,4949
ACT[s] RssL RssR Rst Rsw	0,6190	0,6000	0,6000	0,5833
Cadência RssL RssR Rst Rsw	0,8571	0,8000	0,8000	0,7916

1. Por causa da grandes concentração de nós vermelhos $ACT[s]$ e $Cadência$ podem ser características determinantes de um dos dois grupos classificados.
2. Rst e Rsw apresentam muitos nós vazios no mapa. Pode-se concluir que não são características determinantes no trabalho de classificação.
3. Rs e $RssR$ podem ser características determinantes do grupo azul.
4. $ACT[s]$ e $Cadência$ podem ser fatores determinantes de qualquer um dos grupos.

Resultados reais baseados nos experimentos:

1. Os resultados com $ACT[s]$ são inconclusivos. No entanto, os resultados com $Cadência$ mostram uma clara relevância na classificação dos grupos A e C. Todos os experimentos com essa variável conseguiram altos índices de precisão.
2. Os resultados dos experimentos apoiam essa hipótese, a rede SOM não consegue extrair muitas informações dessas variáveis.

Tabela 12 – Resultado do algoritmo SOM + FC-Means para classificação dos Grupos E e F com 5 variáveis de entrada

SOM + FKMEANS				
5 Combinações de parâmetros de entrada e experimentos entre grupo E e grupo F.				
Input Parameters	P	R	A	F1
ACT[s] Cadência Rs RssL RssR	0,6000	0,6000	0,6000	0,6000
Rs RssL RssR Rst Rsw	0,6000	0,6000	0,6000	0,6000
ACT[s] Cadência Rs RssR Rst	0,2222	0,4000	0,4444	0,2857
Cadência Rs RssL RssR Rsw	0,5000	0,5000	0,5000	0,4949
ACT[s] Rs RssL RssR Rsw	0,7777	0,6000	0,6000	0,5238
ACT[s] Cadência RssL RssR Rsw	0,2222	0,4000	0,4444	0,2857
ACT[s] Rs RssR Rst Rsw	0,5000	0,5000	0,5000	0,4545
Cadência Rs RssR Rst Rsw	0,5000	0,5000	0,5000	0,4949
ACT[s] Cadência RssR Rst Rsw	0,2222	0,4000	0,4444	0,2857
ACT[s] Rs RssL Rst Rsw	0,7083	0,7000	0,7000	0,6969
Cadência Rs RssL RssR Rst	0,5000	0,5000	0,5000	0,4949
ACT[s] Cadência Rs RssL Rst	0,5000	0,5000	0,5000	0,4505
ACT[s] Cadência RssL Rst Rsw	0,2222	0,4000	0,4444	0,2857
ACT[s] Cadência Rs Rst Rsw	0,1875	0,3000	0,3000	0,2307
ACT[s] Cadência RssL RssR Rst	0,2222	0,4000	0,4444	0,2857
Cadência Rs RssL Rst Rsw	0,6190	0,6000	0,6000	0,5833
ACT[s] Cadência Rs RssR Rsw	0,5000	0,5000	0,5000	0,4949
ACT[s] Cadência Rs RssL Rsw	0,5405	0,5400	0,5400	0,5383
ACT[s] Rs RssL RssR Rst	0,4000	0,4000	0,4000	0,4000
ACT[s] RssL RssR Rst Rsw	0,6000	0,6000	0,6000	0,6000
Cadência RssL RssR Rst Rsw	0,5000	0,5000	0,5000	0,4505

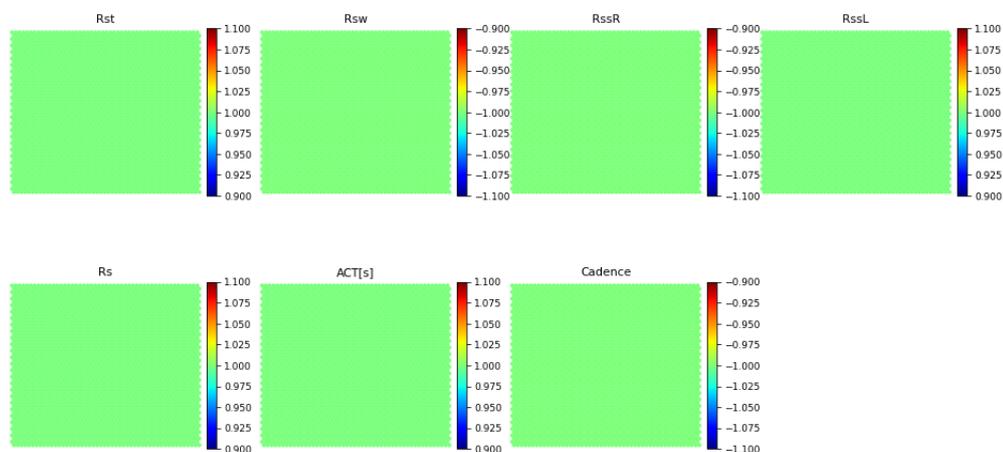
Tabela 13 – Melhores Resultados com 4 variáveis de entrada

Algoritmo	Grupo	Entrada	P	R	F1-Score	A
SOM + K-MEANS	A e C	Cadência Rs RssR Rsw	0,8571	0,8000	0,8000	0,7916
SOM + K-MEANS	E e F	Cadência Rs RssR Rsw	0,7777	0,6000	0,6000	0,5238
SOM + FC-MEANS	A e C	Cadência Rs RssR Rst	0,8571	0,8000	0,8000	0,7916
SOM + FC-MEANS	E e F	ACT[s] Rs RssR Rsw	0,7777	0,6000	0,6000	0,5238

3. Nesse caso, *Rs* e *RssR*, são características relevantes na classificação dos grupos E e F. Essa afirmação se mostrou verdade em alguns casos.
4. Os resultados com *ACT[s]* são inconclusivos. Porém o resultados com *Cadência* apontam que essa características só é relevante para os grupos A e C.

Tal análise pode facilitar os testes de novos algoritmos e evitar a realização de testes que de toda forma não trariam informações relevantes. No entanto, esse tipo análise fica inválida na medida que a quantidade de recursos aumenta, sendo necessário o uso de

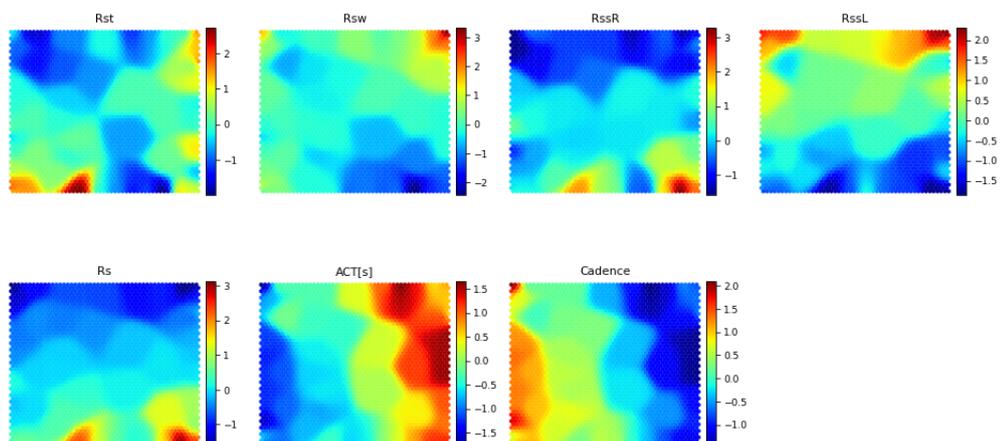
Figura 11 – Heatmap com todas as sete entradas para classificação dos grupos A e C, utilizando os dados de somente um indivíduo



Fonte: Autoria própria.

outro método de avaliação.

Figura 12 – Heatmap SOM + K-MEANS com todas as sete entradas para classificação dos grupos A e C



Fonte: Autoria própria

5 Conclusões e Trabalhos Futuros

Neste trabalho foi descrito o processo de desenvolvimento de um sistema para extrair características e descritores discriminativos de sinais de marcha para identificar um indivíduo de outros. Utilizando-se abordagens híbridas para solucionar o problema, o algoritmo SOM foi hibridizado a um algoritmo de agrupamento, nesse caso: K-Means ou FCM. A escolha dessas ferramentas se deu por causa de (10) e (31), com o objetivo de melhorar os resultados obtidos individualmente com o FCM e o SOM em uma nova abordagem que reunisse o melhor das duas. Durante o andamento da pesquisa percebeu-se que os algoritmos híbridos com o FCM e K-Means apresentaram resultados similares. Ambos os algoritmos de *clustering* apresentam estruturas e comportamentos semelhantes (35) (36). Sendo assim, é recomendado o uso da técnica SOM + K-Means que requer menos recursos computacionais e produz resultados similares.

Os resultados referentes à classificação se mostraram satisfatórios. O algoritmo se mostrou robusto o suficiente para alcançar bons índices de precisão em tarefas de classificação entre grupos com pouca diferença aparente. O algoritmo desmonstrou uma acurácia acima de 80% na classificação dos grupos A e C para os experimentos com sete entradas e acima de 70% na classificação dos grupos E e F. Os resultados dos grupos A e C se assemelham muito aos resultados encontrados em Caldas et al. (2016), em contrapartida os resultados encontrados entre os grupos E e F apresentaram uma melhora significativa. Comprovando que o novo algoritmo é capaz de distinguir indivíduos pertencentes a grupos similares e a grupos significativamente diferentes.

Esta pesquisa também se concentrou em analisar a relevância dos recursos disponíveis como entrada do algoritmo. Os resultados sugerem a importância geral da *Cadência*, especialmente na classificação dos grupos A e C, notou-se que a *Cadência* foi o fator determinante em todos os experimentos que alcançaram índices superiores a 80%. Esse resultado fornece informações úteis aos profissionais de saúde, apoiando e orientando a avaliação clínica. A aplicação desta análise pode contribuir para uma maior eficiência no processo, apoiando a avaliação de distúrbios de marcha.

Como trabalhos futuros, pode-se expandir os experimentos com os métodos relatados para ensaios clínicos, a fim de validar sua aplicabilidade em situações reais. Outra possibilidade, é aplicar as técnicas desenvolvidas neste trabalho para resolver problemas diferentes, como por exemplo, problemas de séries temporais, segmentação de imagens, entre outros.

Referências

- 1 MANNINI, A.; SABATINI, A. M. Gait phase detection and discrimination between walking–jogging activities using hidden markov models applied to foot motion data from a gyroscope. *Gait & posture*, Elsevier, v. 36, n. 4, p. 657–661, 2012. 15
- 2 MURRAY, M. P.; DROUGHT, A. B.; KORY, R. C. Walking patterns of normal men. *Journal of Bone and Joint Surgery-American Volume*, JOURNAL BONE JOINT SURGERY INC 20 PICKERING ST, NEEDHAM, MA 02192, v. 46, n. 2, p. 335–360, 1964. 15
- 3 KATZ, J. N.; DALGAS, M.; STUCKI, G.; KATZ, N. P.; BAYLEY, J.; FOSSEL, A. H.; CHANG, L. C.; LIPSON, S. J. Degenerative lumbar spinal stenosis diagnostic value of the history and physical examination. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, Wiley Online Library, v. 38, n. 9, p. 1236–1241, 1995. 15
- 4 NUTT, J.; MARSDEN, C.; THOMPSON, P. Human walking and higher-level gait disorders, particularly in the elderly. *Neurology*, AAN Enterprises, v. 43, n. 2, p. 268–268, 1993. 15
- 5 KALE, A.; SUNDARESAN, A.; RAJAGOPALAN, A.; CUNTOOR, N. P.; ROY-CHOWDHURY, A. K.; KRUGER, V.; CHELLAPPA, R. Identification of humans using gait. *IEEE Transactions on image processing*, IEEE, v. 13, n. 9, p. 1163–1173, 2004. 15
- 6 SEEL, T.; RAISCH, J.; SCHAUER, T. Imu-based joint angle measurement for gait analysis. *Sensors*, Multidisciplinary Digital Publishing Institute, v. 14, n. 4, p. 6891–6909, 2014. 16
- 7 DEHZANGI, O.; TAHERISADR, M.; CHANGALVALA, R. Imu-based gait recognition using convolutional neural networks and multi-sensor fusion. *Sensors*, Multidisciplinary Digital Publishing Institute, v. 17, n. 12, p. 2735, 2017. 16
- 8 CALDAS, R. R.; ITÚRBIDE, F.; RATIVA, D. Goniometria baseada em sensores inerciais na análise dos movimentos da coluna cervical. In: *Congresso Brasileiro de Engenharia Biomédica*. [S.l.: s.n.], 2014. 16
- 9 CALDAS, R.; MUNDT, M.; POTTHAST, W.; NETO, F. B. de L.; MARKERT, B. A systematic review of gait analysis methods based on inertial sensors and adaptive algorithms. *Gait & posture*, Elsevier, v. 57, p. 204–210, 2017. 16, 20
- 10 CALDAS, R.; HU, Y.; NETO, F. B. de L.; MARKERT, B. Self-organizing maps and fuzzy c-means algorithms on gait analysis based on inertial sensors data. In: SPRINGER. *International Conference on Intelligent Systems Design and Applications*. [S.l.], 2016. p. 197–205. 16, 27, 32, 34, 45
- 11 BASMAJIAN, J. V. *Therapeutic exercise*. [S.l.]: Williams & Wilkins, 1978. 18

- 12 SCHWARTZ, R. P.; HEATH, A. L. The pneumographic method of recording gait. *JBJS*, Citeseer, v. 14, n. 4, p. 783–794, 1932. 18
- 13 JOHANSON, M. Gait laboratory: structure and data gathering. *Human Walking*, Williams and Wilkins Baltimore, MD, p. 201–224, 1994. 18, 20
- 14 NUNES, J. F. *Novel Computational Methodologies for Detailed Analysis and Simulation of Human Motion from Image Sequences*. Tese (Doutorado) — Instituto Politécnico de Viana do Castelo, 2013. 18
- 15 BLANC, Y.; DIMANICO, U. History of the study of skeletal muscle function with emphasis on kinesiological electromyography. *The Open Rehabilitation Journal*, v. 3, p. 84–93, 2010. 18
- 16 STEINDLER, A. Kinesiology of the human body, Charles C. Thomas, Springfield. Illinois USA, p. 646–651, 1955. 18, 25
- 17 VAUGHAN, C. L.; DAVIS, B. L.; JEREMY, C. et al. Dynamics of human gait. Kiboho Publishers, 1999. 19
- 18 VAUGHAN, C.; DAVIS, B.; O'CONNOR, J. The three-dimensional and cyclic nature of gait. *Dynamics of human gait*. Champaign: Human Kinetics Publishers, p. 7–14, 1992. 19
- 19 BRUXEL, Y. Sistema para análise de impacto na marcha humana. 2010. 19
- 20 WHITTLE, M. W. Gait analysis: an introduction. 2007. *Heidi Harrison*, p. 47–100. 19
- 21 SAAD, M.; BATTISTELLA, L. R.; MASIERO, D. Técnicas de análise de marcha. *Acta fisiátrica*, v. 3, n. 2, p. 23–26, 1996. 20
- 22 ABU-FARAJ, Z. O.; HARRIS, G. F.; SMITH, P. A.; HASSANI, S. Human gait and clinical movement analysis. *Wiley Encyclopedia of Electrical and Electronics Engineering*, Wiley Online Library, p. 1–34, 1999. 20
- 23 GAGE, J. R. Gait analysis in cerebral palsy. *Clinics in developmental medicine*, Mac Keith Press, v. 121, 1991. 21
- 24 WOLTRING, H.; HUISKES, R.; LANGE, A. D.; VELDPAUS, F. Finite centroid and helical axis estimation from noisy landmark measurements in the study of human joint kinematics. *Journal of biomechanics*, Elsevier, v. 18, n. 5, p. 379–389, 1985. 21
- 25 KRIESEL, D. A brief introduction on neural networks. Citeseer, 2007. 22
- 26 KOHONEN, T. An introduction to neural computing. *Neural networks*, Elsevier, v. 1, n. 1, p. 3–16, 1988. 22, 24, 27
- 27 HARTIGAN, J. A. Clustering algorithms. Wiley, 1975. 25
- 28 HARTIGAN, J. A.; WONG, M. A. Algorithm as 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, JSTOR, v. 28, n. 1, p. 100–108, 1979. 25
- 29 BOCK, H.-H. Clustering methods: a history of k-means algorithms. In: *Selected contributions in data analysis and classification*. [S.l.]: Springer, 2007. p. 161–172. 27

- 30 WANG, Z.; CALDAS, R. R. Manual gait events detection and adaptive algorithms on gait analysis based on inertial-sensors data. In: . [S.l.: s.n.], 2018. 28, 34
- 31 CALDAS, R.; RÁTIVA, D.; NETO, F. B. de L. Clustering of self-organizing maps as a means to support gait kinematics analysis and symmetry evaluation. *Medical engineering & physics*, Elsevier, v. 62, p. 46–52, 2018. 28, 45
- 32 TAREK, K. M.; SOFIANE, K.; FAROUK, B. Kohonen maps combined to k-means in a two level strategy for time series clustering application to meteorological and electricity load data. In: *Self-Organizing Maps*. [S.l.]: InTech, 2010. 30
- 33 BEJARANO, N. C.; AMBROSINI, E.; PEDROCCHI, A.; FERRIGNO, G.; MONTICONE, M.; FERRANTE, S. An adaptive real-time algorithm to detect gait events using inertial sensors. In: SPRINGER. *XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013*. [S.l.], 2014. p. 1799–1802. 32
- 34 YANG, M.; ZHENG, H.; WANG, H.; MCCLEAN, S.; HALL, J.; HARRIS, N. A machine learning approach to assessing gait patterns for complex regional pain syndrome. *Medical engineering & physics*, Elsevier, v. 34, n. 6, p. 740–746, 2012. 32
- 35 CEBECI, Z.; YILDIZ, F. Comparison of k-means and fuzzy c-means algorithms on different cluster structures. *AGRÁRINFORMATIKA/JOURNAL OF AGRICULTURAL INFORMATICS*, Hungarian Association of Agricultural Informatics, v. 6, n. 3, p. 13–23, 2015. 45
- 36 GHOSH, S.; DUBEY, S. K. Comparative analysis of k-means and fuzzy c-means algorithms. *International Journal of Advanced Computer Science and Applications*, Citeseer, v. 4, n. 4, 2013. 45

Apêndices

APÊNDICE A – Resultados de Todos os Experimentos Restantes

Este Apêndice contém todos os resultados obtidos que não foram citados no texto de ambos os algoritmos híbridos produzidos (SOM + FCM e SOM + K-Means) para o problema de classificação de indivíduos por faixa etária a partir de dados coletados por sensores inerciais.

Tabela 14 – Resultados do algoritmo SOM + K-Means para classificação dos Grupos A e C com 2 variáveis de entrada

Parâmetros de entrada	Precisão	Recall	F1-Score	Acurácia
Rsw Cadência	0,8571	0,8000	0,8000	0,7916
Rsw Rst	0,4000	0,4000	0,4000	0,4000
Rst Cadência	0,8571	0,8000	0,8000	0,7916
Rsw ACT	0,2916	0,3000	0,3000	0,2929
RssL Cadência	0,8571	0,8000	0,8000	0,7916
Rs Cadência	0,8571	0,8000	0,8000	0,7916
RssR Cadência	0,8571	0,8000	0,8000	0,7916
RssL ACT	0,6000	0,6000	0,6000	0,6000
Rsw RssL	0,5000	0,5000	0,5000	0,4949
RssR Rs	0,5000	0,5000	0,5000	0,4505
Rst RssR	0,6190	0,6000	0,6000	0,5833
Rst ACT	0,7777	0,6000	0,6000	0,5238
Rs ACT	0,5000	0,5000	0,5000	0,4505
Rst RssL	0,3809	0,4000	0,4000	0,3750
Cadência ACT[s]	0,8571	0,8000	0,8000	0,7916
Rst Rs	0,5000	0,5000	0,5000	0,4505
Rsw Rs	0,4000	0,4000	0,4000	0,4000
Rsw RssR	0,2916	0,3000	0,3000	0,2929
RssR RssL	0,5000	0,5000	0,5000	0,4949
RssL Rs	0,5000	0,5000	0,5000	0,4949
RssR ACT[s]	0,5000	0,5000	0,5000	0,4949

Tabela 15 – Resultados do algoritmo SOM + K-Means para classificação dos Grupos E e F com 2 variáveis de entrada

Parâmetros de entrada	Precisão	Recall	F1-Score	Acurácia
Rsw Cadência	0,1875	0,3000	0,3000	0,2307
Rsw Rst	0,7083	0,7000	0,7000	0,7000
Rst Cadência	0,4198	0,4200	0,4200	0,4197
Rsw ACT	0,8125	0,7000	0,7000	0,6703
RssL Cadência	0,2222	0,4000	0,4000	0,2857
Rs Cadência	0,2222	0,4000	0,4000	0,2857
RssR Cadência	0,2222	0,4000	0,4000	0,2857
RssL ACT	0,6190	0,6000	0,6000	0,5833
Rsw RssL	0,2916	0,3000	0,3000	0,2929
RssR Rs	0,7777	0,6000	0,6000	0,5238
Rst RssR	0,6000	0,6000	0,6000	0,6000
Rst ACT	0,5000	0,5000	0,5000	0,4505
Rs ACT	0,5000	0,5000	0,5000	0,4505
Rst RssL	0,7083	0,7777	0,7777	0,6969
Cadência ACT[s]	0,2222	0,4000	0,4000	0,2857
Rst Rs	0,7083	0,7777	0,7777	0,6969
Rsw Rs	0,7083	0,7777	0,7777	0,6969
Rsw RssR	0,7083	0,7777	0,7777	0,6969
RssR RssL	0,7083	0,7777	0,7777	0,6969
RssL Rs	0,7083	0,7777	0,7777	0,6969
RssR ACT[s]	0,7777	0,6000	0,6000	0,5238

Tabela 16 – Resultados do algoritmo SOM + FC-Means para classificação dos Grupos A e C com 2 variáveis de entrada

Parâmetros de entrada	Precisão	Recall	F1-Score	Acurácia
Rsw Cadência	0,8571	0,8000	0,8000	0,7916
Rsw Rst	0,5000	0,5000	0,5000	0,4949
Rst Cadência	0,8571	0,8000	0,8000	0,7916
Rsw ACT	0,2916	0,3000	0,3000	0,2929
RssL Cadência	0,8571	0,8000	0,8000	0,7916
Rs Cadência	0,8571	0,8000	0,8000	0,7916
RssR Cadência	0,8571	0,8000	0,8000	0,7916
RssL ACT	0,6000	0,6000	0,6000	0,6000
Rsw RssL	0,5000	0,5000	0,5000	0,4949
RssR Rs	0,5000	0,5000	0,5000	0,4505
Rst RssR	0,5000	0,5000	0,5000	0,4949
Rst ACT	0,7777	0,6000	0,5238	0,6000
Rs ACT	0,5000	0,5000	0,5000	0,4505
Rst RssL	0,5000	0,5000	0,5000	0,4949
Cadência ACT[s]	0,8571	0,8000	0,8000	0,7916
Rst Rs	0,5000	0,5000	0,5000	0,4505
Rsw Rs	0,4000	0,4000	0,4000	0,4000
Rsw RssR	0,2916	0,3000	0,3000	0,2929
RssR RssL	0,5000	0,5000	0,5000	0,4949
RssL Rs	0,5000	0,5000	0,5000	0,4949
RssR ACT[s]	0,5000	0,5000	0,5000	0,4505

Tabela 17 – Resultados do algoritmo SOM + FC-Means para classificação dos Grupos E e F com 2 variáveis de entrada

Parâmetros de entrada	Precisão	Recall	F1-Score	Acurácia
Rsw Cadência	0,1875	0,3000	0,3000	0,2307
Rsw Rst	0,7083	0,7000	0,7000	0,6969
Rst Cadência	0,2222	0,4000	0,4000	0,2857
Rsw ACT	0,8125	0,7000	0,6703	0,7000
RssL Cadência	0,2222	0,4000	0,4000	0,2857
Rs Cadência	0,2222	0,4000	0,4000	0,2857
RssR Cadência	0,2222	0,4000	0,4000	0,2857
RssL ACT	0,6190	0,6000	0,6000	0,5833
Rsw RssL	0,7083	0,7000	0,7000	0,6969
RssR Rs	0,7777	0,6666	0,6666	0,5238
Rst RssR	0,6000	0,6000	0,6000	0,6000
Rst ACT	0,5000	0,5000	0,5000	0,4505
Rs ACT	0,5000	0,5000	0,5000	0,4505
Rst RssL	0,7083	0,7777	0,7777	0,6969
Cadência ACT[s]	0,2222	0,4000	0,4000	0,2857
Rst Rs	0,7083	0,7000	0,7000	0,6969
Rsw Rs	0,7083	0,7000	0,7000	0,6969
Rsw RssR	0,7777	0,6000	0,5238	0,6000
RssR RssL	0,7083	0,7000	0,7000	0,6969
RssL Rs	0,7083	0,7000	0,7000	0,6969
RssR ACT[s]	0,7777	0,6000	0,6000	0,5238

Tabela 18 – Resultados do algoritmo SOM + K-Means para classificação dos Grupos A e C com 3 variáveis de entrada

Parâmetros de entrada	Precisão	Recall	F1-Score	Acurácia
ACT[s] Cadência Rs	0,8571	0,8000	0,8000	0,7916
Cadência Rst Rsw	0,8571	0,8000	0,8000	0,7916
RssR Rst Rsw	0,5000	0,5000	0,4949	0,5000
ACT[s] Cadência Rsw	0,8571	0,8000	0,8000	0,7916
Rs RssR Rst	0,5000	0,5000	0,4505	0,5000
Cadência Rs RssL	0,8571	0,8000	0,8000	0,7916
ACT[s] RssL RssR	0,5000	0,5000	0,4505	0,5000
RssL Rst Rsw	0,6190	0,6000	0,6000	0,5833
Cadência RssL Rsw	0,8571	0,8000	0,8000	0,7916
Rs RssR Rsw	0,3809	0,4000	0,4000	0,3750
ACT[s] RssR Rst	0,7777	0,6000	0,6000	0,5238
ACT[s] Rs Rst	0,7777	0,6000	0,6000	0,5238
ACT[s] Rs RssL	0,5000	0,5000	0,4949	0,5000
Cadência RssL Rst	0,8571	0,8000	0,8000	0,7916
Cadência Rs Rst	0,8571	0,8000	0,8000	0,7916
Rs RssL Rst	0,6190	0,6000	0,6000	0,5833
ACT[s] Cadência RssL	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência Rst	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência RssR	0,8571	0,8000	0,8000	0,7916
Cadência Rs RssR	0,8571	0,8000	0,8000	0,7916
Cadência Rs Rsw	0,8571	0,8000	0,8000	0,7916
Cadência RssL RssR	0,8571	0,8000	0,8000	0,7916
Cadência RssR Rst	0,8571	0,8000	0,8000	0,7916
Cadência RssR Rsw	0,8571	0,8000	0,8000	0,7916
RssL RssR Rsw	0,6190	0,6000	0,6000	0,5833
RssL RssR Rst	0,5000	0,5000	0,4949	0,5000
Rs Rst Rsw	0,6190	0,6000	0,6000	0,5833
Rs RssL Rsw	0,5000	0,5000	0,4949	0,5000
Rs RssL RssR	0,3809	0,4000	0,4000	0,3750
ACT[s] Rst Rsw	0,5000	0,5000	0,4949	0,5000
ACT[s] RssR Rsw	0,4000	0,4000	0,4000	0,4000
ACT[s] RssL Rsw	0,4000	0,4000	0,4000	0,4000
ACT[s] RssL Rst	0,2222	0,4000	0,4000	0,2857
ACT[s] Rs Rsw	0,4000	0,4000	0,4000	0,4000
ACT[s] Rs RssR	0,5000	0,5000	0,4505	0,5000

Tabela 19 – Resultados do algoritmo SOM + K-Means para classificação dos Grupos E e F com 3 variáveis de entrada

Parâmetros de entrada	Precisão	Recall	F1-Score	Acurácia
ACT[s] Cadência Rs	0,2222	0,4000	0,2857	0,4000
Cadência Rst Rsw	0,2222	0,4000	0,2857	0,4000
RssR Rst Rsw	0,7777	0,6000	0,6000	0,5238
ACT[s] Cadência Rsw	0,2222	0,4000	0,2857	0,4000
Rs RssR Rst	0,4000	0,4000	0,4000	0,4000
Cadência Rs RssL	0,6190	0,6000	0,6000	0,5833
ACT[s] RssL RssR	0,6000	0,6000	0,6000	0,6000
RssL Rst Rsw	0,7083	0,7000	0,7000	0,6969
Cadência RssL Rsw	0,2222	0,4000	0,2857	0,4000
Rs RssR Rsw	0,7777	0,6000	0,6000	0,5238
ACT[s] RssR Rst	0,6000	0,6000	0,6000	0,6000
ACT[s] Rs Rst	0,2916	0,3000	0,3000	0,2929
ACT[s] Rs RssL	0,7083	0,7000	0,7000	0,6969
Cadência RssL Rst	0,2222	0,4000	0,2857	0,4000
Cadência Rs Rst	0,2222	0,4000	0,2857	0,4000
Rs RssL Rst	0,7083	0,7000	0,7000	0,6969
ACT[s] Cadência RssL	0,2222	0,4000	0,2857	0,4000
ACT[s] Cadência Rst	0,2222	0,4000	0,2857	0,4000
ACT[s] Cadência RssR	0,2222	0,4000	0,2857	0,4000
Cadência Rs RssR	0,2222	0,4000	0,2857	0,4000
Cadência Rs Rsw	0,2222	0,4000	0,2857	0,4000
Cadência RssL RssR	0,2222	0,4000	0,2857	0,4000
Cadência RssR Rst	0,2222	0,4000	0,2857	0,4000
Cadência RssR Rsw	0,2222	0,4000	0,2857	0,4000
RssL RssR Rsw	0,7083	0,7000	0,7000	0,6969
RssL RssR Rst	0,7083	0,7000	0,7000	0,6969
Rs Rst Rsw	0,6000	0,6000	0,6000	0,6000
Rs RssL Rsw	0,2916	0,3000	0,3000	0,2929
Rs RssL RssR	0,6000	0,6000	0,6000	0,6000
ACT[s] Rst Rsw	0,7083	0,7000	0,7000	0,6969
ACT[s] RssR Rsw	0,7777	0,6000	0,6000	0,5238
ACT[s] RssL Rsw	0,6000	0,6000	0,6000	0,6000
ACT[s] RssL Rst	0,7083	0,7000	0,7000	0,6969
ACT[s] Rs Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Rs RssR	0,7777	0,6000	0,6000	0,5238

Tabela 20 – Resultados do algoritmo SOM + FC-Means para classificação dos Grupos A e C com 3 variáveis de entrada

Parâmetros de entrada	Precisão	Recall	F1-Score	Acurácia
ACT[s] Cadência Rs	0,8571	0,8000	0,8000	0,7916
Cadência Rst Rsw	0,8571	0,8000	0,8000	0,7916
RssR Rst Rsw	0,5000	0,5000	0,4949	0,5000
ACT[s] Cadência Rsw	0,8571	0,8000	0,8000	0,7916
Rs RssR Rst	0,5000	0,5000	0,4505	0,5000
Cadência Rs RssL	0,8571	0,8000	0,8000	0,7916
ACT[s] RssL RssR	0,5000	0,5000	0,4505	0,5000
RssL Rst Rsw	0,6190	0,6000	0,6000	0,5833
Cadência RssL Rsw	0,8571	0,8000	0,8000	0,7916
Rs RssR Rsw	0,3809	0,4000	0,3750	0,4000
ACT[s] RssR Rst	0,8125	0,7777	0,7777	0,6703
ACT[s] Rs Rst	0,7777	0,6000	0,6000	0,5238
ACT[s] Rs RssL	0,5000	0,5000	0,4949	0,5000
Cadência RssL Rst	0,8571	0,8000	0,8000	0,7916
Cadência Rs Rst	0,8571	0,8000	0,8000	0,7916
Rs RssL Rst	0,6190	0,6000	0,6000	0,5833
ACT[s] Cadência RssL	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência Rst	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência RssR	0,8571	0,8000	0,8000	0,7916
Cadência Rs RssR	0,8571	0,8000	0,8000	0,7916
Cadência Rs Rsw	0,8571	0,8000	0,8000	0,7916
Cadência RssL RssR	0,8571	0,8000	0,8000	0,7916
Cadência RssR Rst	0,8571	0,8000	0,8000	0,7916
Cadência RssR Rsw	0,8571	0,8000	0,8000	0,7916
RssL RssR Rsw	0,6190	0,6000	0,6000	0,5833
RssL RssR Rst	0,5000	0,5000	0,4949	0,5000
Rs Rst Rsw	0,6190	0,6000	0,6000	0,5833
Rs RssL Rsw	0,5000	0,5000	0,4949	0,5000
Rs RssL RssR	0,3809	0,4000	0,3750	0,4000
ACT[s] Rst Rsw	0,5000	0,5000	0,4949	0,5000
ACT[s] RssR Rsw	0,4000	0,4000	0,4000	0,4000
ACT[s] RssL Rsw	0,4000	0,4000	0,4000	0,4000
ACT[s] RssL Rst	0,2222	0,4000	0,4000	0,2857
ACT[s] Rs Rsw	0,4000	0,4000	0,4000	0,4000
ACT[s] Rs RssR	0,5000	0,5000	0,4505	0,5000

Tabela 21 – Resultados do algoritmo SOM + FC-Means para classificação dos Grupos E e F com 3 variáveis de entrada

Parâmetros de entrada	Precisão	Recall	F1-Score	Acurácia
ACT[s] Cadência Rs	0,2222	0,4000	0,4000	0,2857
Cadência Rst Rsw	0,2222	0,4000	0,4000	0,2857
RssR Rst Rsw	0,7777	0,6000	0,6000	0,5238
ACT[s] Cadência Rsw	0,1875	0,3000	0,3000	0,2307
Rs RssR Rst	0,4000	0,4000	0,4000	0,4000
Cadência Rs RssL	0,5000	0,5000	0,4505	0,5000
ACT[s] RssL RssR	0,6000	0,6000	0,6000	0,6000
RssL Rst Rsw	0,7083	0,7000	0,7000	0,6969
Cadência RssL Rsw	0,2222	0,4000	0,2857	0,4000
Rs RssR Rsw	0,7777	0,6000	0,6000	0,5238
ACT[s] RssR Rst	0,6000	0,6000	0,6000	0,6000
ACT[s] Rs Rst	0,2916	0,3000	0,3000	0,2929
ACT[s] Rs RssL	0,7083	0,7000	0,7000	0,6969
Cadência RssL Rst	0,2222	0,4000	0,2857	0,4000
Cadência Rs Rst	0,2222	0,4000	0,2857	0,4000
Rs RssL Rst	0,7083	0,7000	0,7000	0,6969
ACT[s] Cadência RssL	0,2222	0,4000	0,2857	0,4000
ACT[s] Cadência Rst	0,2222	0,4000	0,2857	0,4000
ACT[s] Cadência RssR	0,2222	0,4000	0,2857	0,4000
Cadência Rs RssR	0,2222	0,4000	0,2857	0,4000
Cadência Rs Rsw	0,2222	0,4000	0,2857	0,4000
Cadência RssL RssR	0,2222	0,4000	0,2857	0,4000
Cadência RssR Rst	0,2222	0,4000	0,2857	0,4000
Cadência RssR Rsw	0,2222	0,4000	0,2857	0,4000
RssL RssR Rsw	0,6000	0,6000	0,6000	0,6000
RssL RssR Rst	0,7083	0,7000	0,7000	0,6969
Rs Rst Rsw	0,6000	0,6000	0,6000	0,6000
Rs RssL Rsw	0,2916	0,3000	0,3000	0,2929
Rs RssL RssR	0,6000	0,6000	0,6000	0,6000
ACT[s] Rst Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] RssR Rsw	0,7777	0,6000	0,6000	0,5238
ACT[s] RssL Rsw	0,6000	0,6000	0,6000	0,6000
ACT[s] RssL Rst	0,7083	0,7000	0,7000	0,6969
ACT[s] Rs Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Rs RssR	0,5000	0,5000	0,4505	0,5000

Tabela 22 – Resultados do algoritmo SOM + K-Means para classificação dos Grupos A e C com 4 variáveis de entrada

Parâmetros de entrada	Precisão	Recall	F1-Score	Acurácia
ACT[s] Cadência Rs RssL	0,8571	0,8000	0,8000	0,7916
RssL RssR Rst Rsw	0,6190	0,6000	0,6000	0,5833
ACT[s] Cadência RssL Rsw	0,8571	0,8000	0,8000	0,7916
Cadência RssL Rst Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] RssL RssR Rst	0,5000	0,5000	0,5000	0,4505
Cadência Rs RssL Rst	0,8571	0,8000	0,8000	0,7916
ACT[s] Rs RssL Rst	0,5000	0,5000	0,5000	0,4949
Cadência Rs Rst Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência Rst Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Rs RssL Rsw	0,5000	0,5000	0,5000	0,4949
Rs RssL RssR Rst	0,5000	0,5000	0,5000	0,4505
Cadência Rs RssR Rst	0,8571	0,8000	0,8000	0,7916
ACT[s] RssL Rst Rsw	0,5000	0,5000	0,5000	0,4505
ACT[s] Cadência Rs Rst	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência RssL RssR	0,8571	0,8000	0,8000	0,7916
Rs RssL Rst Rsw	0,5212	0,5200	0,5200	0,5129
ACT[s] Rs RssR Rsw	0,3809	0,4000	0,4000	0,3750
ACT[s] Cadência Rs RssR	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência Rs Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência RssL Rst	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência RssR Rst	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência RssR Rsw	0,8571	0,8000	0,8000	0,7916
Cadência Rs RssL RssR	0,8571	0,8000	0,8000	0,7916
Cadência Rs RssL Rsw	0,8571	0,8000	0,8000	0,7916
Cadência Rs RssR Rsw	0,8571	0,8000	0,8000	0,7916
Cadência RssL RssR Rst	0,8571	0,8000	0,8000	0,7916
Cadência RssL RssR Rsw	0,8571	0,8000	0,8000	0,7916
Cadência RssR Rst Rsw	0,8571	0,8000	0,8000	0,7916
Rs RssR Rst Rsw	0,3809	0,4000	0,4000	0,3750
Rs RssL RssR Rsw	0,3809	0,4000	0,4000	0,3750
ACT[s] RssR Rst Rsw	0,4000	0,4000	0,4000	0,4000
ACT[s] RssL RssR Rsw	0,5000	0,5000	0,5000	0,4949
ACT[s] Rs Rst Rsw	0,2916	0,3000	0,3000	0,2929
ACT[s] Rs RssL Rsw	0,5000	0,5000	0,5000	0,4949
ACT[s] Rs RssL Rst	0,5000	0,5000	0,5000	0,4949
ACT[s] Rs RssL RssR	0,5000	0,5000	0,5000	0,4505

Tabela 23 – Resultados do algoritmo SOM + K-Means para classificação dos Grupos E e F com 4 variáveis de entrada

Parâmetros de entrada	Precisão	Recall	F1-Score	Acurácia
ACT[s] Cadência Rs RssL	0,2222	0,4000	0,4000	0,2857
RssL RssR Rst Rsw	0,7083	0,7777	0,7777	0,6969
ACT[s] Cadência RssL Rsw	0,2222	0,4000	0,4000	0,2857
Cadência RssL Rst Rsw	0,2222	0,4000	0,4000	0,2857
ACT[s] RssL RssR Rst	0,7083	0,7000	0,7000	0,7000
Cadência Rs RssL Rst	0,5000	0,5000	0,5000	0,4505
ACT[s] Rs RssL Rst	0,6000	0,6000	0,6000	0,6000
Cadência Rs Rst Rsw	0,2916	0,3000	0,3000	0,2929
ACT[s] Cadência Rst Rsw	0,1875	0,3000	0,3000	0,2307
ACT[s] Rs RssL Rsw	0,2916	0,3000	0,3000	0,2929
Rs RssL RssR Rst	0,4000	0,4000	0,4000	0,4000
Cadência Rs RssR Rst	0,2222	0,4000	0,4000	0,2857
ACT[s] RssL Rst Rsw	0,2916	0,3000	0,3000	0,2929
ACT[s] Cadência Rs Rst	0,2222	0,4000	0,4000	0,2857
ACT[s] Cadência RssL RssR	0,5000	0,5000	0,5000	0,4505
Rs RssL Rst Rsw	0,7083	0,7000	0,7000	0,7000
ACT[s] Rs RssR Rsw	0,7777	0,6000	0,6000	0,5238
ACT[s] Cadência Rs RssR	0,2916	0,3000	0,3000	0,2929
ACT[s] Cadência Rs Rsw	0,2916	0,3000	0,3000	0,2929
ACT[s] Cadência RssL Rst	0,5603	0,5600	0,5600	0,5592
ACT[s] Cadência RssR Rst	0,5603	0,5600	0,5600	0,5592
ACT[s] Cadência RssR Rsw	0,1875	0,3000	0,3000	0,2307
Cadência Rs RssL RssR	0,6000	0,6000	0,6000	0,6000
Cadência Rs RssL Rsw	0,7083	0,7000	0,7000	0,7000
Cadência Rs RssR Rsw	0,7777	0,6000	0,6000	0,5238
Cadência RssL RssR Rst	0,2222	0,4000	0,4000	0,2857
Cadência RssL RssR Rsw	0,2222	0,4000	0,4000	0,2857
Cadência RssR Rst Rsw	0,1875	0,3000	0,3000	0,2307
Rs RssR Rst Rsw	0,7777	0,6000	0,6000	0,5238
Rs RssL RssR Rsw	0,7083	0,7000	0,7000	0,7000
ACT[s] RssR Rst Rsw	0,6190	0,6000	0,6000	0,5833
ACT[s] RssL RssR Rsw	0,7083	0,7000	0,7000	0,6969
ACT[s] Rs Rst Rsw	0,8125	0,7000	0,7000	0,6703
ACT[s] Rs RssL Rsw	0,2916	0,3000	0,3000	0,2929
ACT[s] Rs RssL Rst	0,6000	0,6000	0,6000	0,6000
ACT[s] Rs RssL RssR	0,6000	0,6000	0,6000	0,6000

Tabela 24 – Resultados do algoritmo SOM + FC-Means para classificação dos Grupos A e C com 4 variáveis de entrada

Parâmetros de entrada	Precisão	Recall	F1-Score	Acurácia
ACT[s] Cadência Rs RssL	0,8571	0,8000	0,8000	0,7916
RssL RssR Rst Rsw	0,6190	0,6000	0,6000	0,5833
ACT[s] Cadência RssL Rsw	0,8571	0,8000	0,8000	0,7916
Cadência RssL Rst Rsw	0,8652	0,8600	0,8600	0,8594
ACT[s] RssL RssR Rst	0,5000	0,5000	0,5000	0,4505
Cadência Rs RssL Rst	0,8571	0,8000	0,8000	0,7916
ACT[s] Rs RssL Rst	0,5000	0,5000	0,5000	0,4949
Cadência Rs Rst Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência Rst Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Rs RssL Rsw	0,5000	0,5000	0,5000	0,4949
Rs RssL RssR Rst	0,5000	0,5000	0,5000	0,4505
Cadência Rs RssR Rst	0,8571	0,8000	0,8000	0,7916
ACT[s] RssL Rst Rsw	0,5000	0,5000	0,5000	0,4505
ACT[s] Cadência Rs Rst	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência RssL RssR	0,8571	0,8000	0,8000	0,7916
Rs RssL Rst Rsw	0,6190	0,6000	0,6000	0,5833
ACT[s] Rs RssR Rsw	0,3809	0,4000	0,4000	0,3750
ACT[s] Cadência Rs RssR	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência Rs Rsw	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência RssL Rst	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência RssR Rst	0,8571	0,8000	0,8000	0,7916
ACT[s] Cadência RssR Rsw	0,8571	0,8000	0,8000	0,7916
Cadência Rs RssL RssR	0,8571	0,8000	0,8000	0,7916
Cadência Rs RssL Rsw	0,8571	0,8000	0,8000	0,7916
Cadência Rs RssR Rsw	0,8571	0,8000	0,8000	0,7916
Cadência RssL RssR Rst	0,8571	0,8000	0,8000	0,7916
Cadência RssL RssR Rsw	0,8571	0,8000	0,8000	0,7916
Cadência RssR Rst Rsw	0,8571	0,8000	0,8000	0,7916
Rs RssR Rst Rsw	0,3809	0,4000	0,4000	0,3750
Rs RssL RssR Rsw	0,3809	0,4000	0,4000	0,3750
ACT[s] RssR Rst Rsw	0,4000	0,4000	0,4000	0,4000
ACT[s] RssL RssR Rsw	0,5000	0,5000	0,5000	0,4949
ACT[s] Rs Rst Rsw	0,2916	0,3000	0,3000	0,3600
ACT[s] Rs RssL Rsw	0,5000	0,5000	0,5000	0,4949
ACT[s] Rs RssL Rst	0,5000	0,5000	0,5000	0,4949
ACT[s] Rs RssL RssR	0,5000	0,5000	0,5000	0,4505

Tabela 25 – Resultados do algoritmo SOM + FC-Means para classificação dos Grupos E e F com 4 variáveis de entrada

Parâmetros de entrada	Precisão	Recall	F1-Score	Acurácia
ACT[s] Cadência Rs RssL	0,2222	0,4000	0,4000	0,2857
RssL RssR Rst Rsw	0,7083	0,7000	0,7000	0,6969
ACT[s] Cadência RssL Rsw	0,2222	0,4000	0,4000	0,2857
Cadência RssL Rst Rsw	0,2222	0,4000	0,4000	0,2857
ACT[s] RssL RssR Rst	0,7083	0,7000	0,7000	0,6969
Cadência Rs RssL Rst	0,5000	0,5000	0,5000	0,4505
ACT[s] Rs RssL Rst	0,6000	0,6000	0,6000	0,6000
Cadência Rs Rst Rsw	0,2916	0,3000	0,2929	0,3000
ACT[s] Cadência Rst Rsw	0,1875	0,3000	0,3000	0,2307
ACT[s] Rs RssL Rsw	0,2916	0,3000	0,3000	0,2929
Rs RssL RssR Rst	0,4000	0,4000	0,4000	0,4000
Cadência Rs RssR Rst	0,5000	0,5000	0,5000	0,4949
ACT[s] RssL Rst Rsw	0,2916	0,3000	0,3000	0,2929
ACT[s] Cadência Rs Rst	0,2222	0,4000	0,4000	0,2857
ACT[s] Cadência RssL RssR	0,5000	0,5000	0,5000	0,4545
Rs RssL Rst Rsw	0,7083	0,7000	0,7000	0,6969
ACT[s] Rs RssR Rsw	0,7777	0,6000	0,6000	0,5238
ACT[s] Cadência Rs RssR	0,3809	0,4000	0,4000	0,3750
ACT[s] Cadência Rs Rsw	0,1875	0,3000	0,3000	0,2307
ACT[s] Cadência RssL Rst	0,2222	0,4000	0,4000	0,2857
ACT[s] Cadência RssR Rst	0,2222	0,4000	0,4000	0,2857
ACT[s] Cadência RssR Rsw	0,1875	0,3000	0,3000	0,2307
Cadência Rs RssL RssR	0,6000	0,6000	0,6000	0,6000
Cadência Rs RssL Rsw	0,7083	0,7000	0,7000	0,6969
Cadência Rs RssR Rsw	0,7777	0,6000	0,6000	0,5238
Cadência RssL RssR Rst	0,5000	0,5000	0,5000	0,4505
Cadência RssL RssR Rsw	0,2222	0,4000	0,4000	0,2857
Cadência RssR Rst Rsw	0,1875	0,3000	0,3000	0,2307
Rs RssR Rst Rsw	0,7777	0,6000	0,6000	0,5238
Rs RssL RssR Rsw	0,7083	0,7000	0,7000	0,6969
ACT[s] RssR Rst Rsw	0,6190	0,6000	0,6000	0,5833
ACT[s] RssL RssR Rsw	0,7777	0,6000	0,6000	0,5238
ACT[s] Rs Rst Rsw	0,7777	0,6000	0,6000	0,5238
ACT[s] Rs RssL Rsw	0,2916	0,3000	0,3000	0,2929
ACT[s] Rs RssL Rst	0,6000	0,6000	0,6000	0,6000
ACT[s] Rs RssL RssR	0,6000	0,6000	0,6000	0,6000