



UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL
FACULDADE DE COMPUTAÇÃO
CURSO DE GRADUAÇÃO EM ENGENHARIA DE COMPUTAÇÃO

José Pedro Cândido Lopes Peres

**Análise Multivariada de Sinais sEMG para Identificação de Unidades
Motoras e Classificação de Fibras Musculares**

Campo Grande, MS
2025

José Pedro Cândido Lopes Peres

**Análise Multivariada de Sinais sEMG para Identificação de Unidades
Motoras e Classificação de Fibras Musculares**

Trabalho de Conclusão de Curso de Graduação em Engenharia de Computação, da Faculdade de Computação da Universidade Federal de Mato Grosso do Sul (FACOM/UFMS) como requisito para a obtenção do título de Bacharel em Engenharia de Computação.

Orientador: Prof. Dr. Milton Ernesto Romero Romero

Coorientador: Prof. Dra. Vera Regina Fernandes da Silva Marães

Campo Grande, MS

2025

Este trabalho é dedicado aos meus queridos pais,
orientadores, coorientadores e colegas de classe.

RESUMO

A sEMG (Eletromiografia de Superfície - *Surface Electromyography*) é uma técnica não invasiva que capta sinais bioelétricos, permitindo a avaliação da atividade neuromuscular, sendo útil tanto no diagnóstico clínico quanto em aplicações na ciência do esporte. A análise precisa desses sinais apresenta diversos desafios, como interferências eletromagnéticas externas e potenciais elétricos de outros processos fisiológicos, que podem causar sobreposição de sinais. Este estudo foca no processamento de sinais sEMG discretos com o objetivo de estimar, de forma virtual, o número de MUs (Unidades Motoras - *Motor Units*) e fibras musculares envolvidas na atividade muscular (tipo 1 - contração lenta, tipo 2A - contração rápida oxidativa e tipo 2B - contração rápida glicolítica), a partir das correntes iônicas detectadas pelos eletrodos, que atualmente são obtidas, com devida precisão, apenas por meio de métodos invasivos. A metodologia inicia-se com o pré-processamento do sinal, incluindo filtragem Notch, filtragem passa-banda Butterworth de quarta ordem para remoção de ruídos e suavização por média móvel para eliminação de valores atípicos. Em seguida, realiza-se a extração de características, abrangendo variância, desvio padrão, RMS (Valor eficaz - *Root-Mean-Square*), assimetria e curtose, permitindo a avaliação do comportamento estatístico e estocástico do sinal. Para as análises nos domínios do tempo e da frequência, aplica-se a DWT (Transformada Wavelet Discreta - *Discrete Wavelet Transform*). Essa saída alimenta o algoritmo de clusterização K-Means, que segmenta padrões de ativação e identifica as frequências dominantes. As análises foram conduzidas com dados experimentais obtidos em condições controladas e padronizadas, considerando contrações voluntárias de músculos de membros superiores (Protocolo de Membro Superior) e de membros inferiores (Protocolos de Membro Inferior – Vista Anterior e Posterior) de indivíduos, com uma carga de 50 kg em ambos os protocolos. Os resultados experimentais demonstram a eficácia do algoritmo em distinguir níveis de ativação muscular e estimar parâmetros clinicamente relevantes, como as taxas de disparos neurais, além de diferenciar os tipos de fibras musculares envolvidas. Isso fornece uma estimativa virtual das unidades motoras ativadas e dos tipos de fibras. Uma validação clínica adicional é necessária para confirmar a robustez desses achados e respaldar sua aplicação médica. Quando disponíveis, os dados poderão também possibilitar o mapeamento virtual das unidades motoras e fibras para estruturas anatômicas reais. O método proposto é seguro e aplicável em contextos terapêuticos e na prevenção da perda de massa muscular.

Palavras-chave: Eletromiografia de Superfície; Processamento de Sinais; Transformada Wavelet; Unidades Motoras; Clustering.

ABSTRACT

Surface electromyography (sEMG), a non-invasive technique, captures signals that allow for the assessment of neuromuscular activity, proving useful in both patient diagnosis and sports science. Accurate analysis of these signals presents several challenges, including external electromagnetic interference and electrical potentials from other physiological processes, which may cause signal overlap. This study focuses on processing discrete sEMG signals with the objective of estimating a virtual number of Motor Units (MUs) and muscle fibers involved in muscular activity (type 1 - slow-twitch, type 2A - fast-twitch oxidative, and type 2B - fast-twitch glycolytic), based on the ionic currents detected by the electrodes, which are currently obtained, with due precision, only through invasive methods. The methodology begins with signal pre-processing, including Notch filtering, fourth-order Butterworth band-pass filtering for noise removal, and smoothing via a moving average to eliminate outliers. Subsequently, feature extraction is performed, encompassing variance, standard deviation, root mean square (RMS), skewness, and kurtosis, allowing for the assessment of the statistical and stochastic behavior of the signal. For time-domain and frequency-domain analyses, the Discrete Wavelet Transform (DWT) is applied. This output serves as the input for the K-Means clustering algorithm, which segments activation patterns and identifies dominant frequencies. The analyses were conducted using experimental data obtained under controlled and standardized conditions, considering voluntary muscle contractions in upper limb regions (Upper Limb Protocol) and lower limb regions (Lower Limb Protocol – Anterior and Posterior Views) of individuals with a load of 50 kg in both protocols. Experimental results demonstrate the algorithm's effectiveness in distinguishing muscle activation levels and estimating clinically relevant parameters, such as neural firing rates, while differentiating the types of muscle fibers involved. This provides a virtual estimate of both activated motor units and fiber types. Further clinical validation is required to confirm the robustness of these findings and support their medical application. Where data are available, it may also be possible to map virtual motor units and fibers to actual anatomical structures. The proposed method is safe and suitable for therapeutic applications and for the prevention of muscle mass loss.

Keywords: Surface electromyography; signal processing; Wavelet Transform; motor units; clustering.

LISTA DE FIGURAS

Figura 1 – Representação do Trem de Impulsos de Potenciais de Unidades Motoras (MUAPTs) e ruídos de fundo. Fonte: (Alcan; Zinnuroğlu, 2023)	14
Figura 2 – Representação de um sinal de EMG como um somatório de MUAPs. Fonte: (Alcan; Zinnuroğlu, 2023)	15
Figura 3 – Comparação entre os métodos de coleta de sinais de EMG: Needle EMG e Surface EMG. Fonte: (Alcan; Zinnuroğlu, 2023)	16
Figura 4 – Fluxograma das etapas metodológicas. Fonte: Elaborado pelo autor.	26
Figura 5 – Pré-processamento do sinal de sEMG. Fonte: Elaborado pelo autor.	30
Figura 6 – Segmentação dos picos do sinal de sEMG. Fonte: Elaborado pelo autor.	31
Figura 7 – Clusterização K-Means dos rótulos de um sinal muscular processado da região esquerda do membro inferior com Transformada Wavelet. Fonte: Elaborado pelo autor.	31
Figura 8 – Visualização 3D da clusterização K-Means do mesmo sinal processado apresentado na Figura 7. Fonte: Elaborado pelo autor.	32
Figura 9 – Distribuição dos clusters no espaço de componentes principais. Fonte: Elaborado pelo autor.	33

LISTA DE TABELAS

Tabela 1 – Métricas de desempenho médias do pipeline de processamento de sinais sEMG.	34
Tabela 2 – Distribuição percentual aproximada das fibras musculares identificadas nos músculos da região inferior esquerda.	35

LISTA DE ABREVIATURAS E SIGLAS

CSV	Comma-Separated Values
CWT	Transformada Wavelet Contínua - <i>Continuous Wavelet Transform</i>
DWT	Transformada Wavelet Discreta - <i>Discrete Wavelet Transform</i>
EDF	European Data Format
EMG	Eletromiografia - <i>Electromyography</i>
MUAPs	Potenciais de Ação de Unidades Motoras - <i>Motor Unit Action Potentials</i>
MUAPTs	Trem de Impulsos de Potenciais de Unidades Motoras - <i>Motor Unit Action Potentials Train</i>
MUs	Unidades Motoras - <i>Motor Units</i>
PCA	Análise de Componentes Principais - <i>Principal Component Analysis</i>
RMS	Valor eficaz - <i>Root-Mean-Square</i>
sEMG	Eletromiografia de Superfície - <i>Surface Electromyography</i>
SENIAM	Surface Electromyography for the Non-Invasive Assessment of Muscles
VMUAPs	Potenciais de Ação de Unidades Motoras Virtuais - <i>Virtual Motor Unit Action Potentials</i>
WT	Transformada Wavelet - <i>Wavelet Transform</i>
XLSX	Excel Spreadsheet

SUMÁRIO

1	INTRODUÇÃO	10
2	TRABALHOS RELACIONADOS	12
3	FUNDAMENTAÇÃO TEÓRICA	14
3.1	A ELETROMIOGRAFIA	14
3.2	COLETA DE SINAIS DE EMG	15
3.3	DESAFIOS NA ANÁLISE DE SINAIS EMG	16
3.3.1	Problemas Gerais dos Sinais de EMG	17
3.3.2	Problemas Relacionados à Fisiologia Muscular	17
3.3.3	Problemas de Instrumentação e Padronização	18
3.4	ABORDAGENS METODOLÓGICAS PARA A ANÁLISE DE SINAIS EMG	18
3.4.1	Pré-processamento do sinal	18
3.4.2	Extração de características	19
3.4.3	Clustering para identificação de padrões com K-Means	21
3.4.4	Redução de Dimensionalidade com PCA	21
4	JUSTIFICATIVA	23
4.1	MEDICINA E FISIOTERAPIA	23
4.1.1	Diagnóstico mais preciso de doenças musculares	23
4.1.2	Desenvolvimento de novos medicamentos e terapias	23
4.2	CIÊNCIAS DO ESPORTE	24
4.2.1	Otimização do treinamento esportivo	24
4.2.2	Monitoramento da performance	24
5	OBJETIVOS	25
5.1	OBJETIVO GERAL	25
5.2	OBJETIVOS ESPECÍFICOS	25
6	METODOLOGIA	26
6.1	COLETA DOS SINAIS DE SEMG	26
6.2	PRÉ-PROCESSAMENTO DOS SINAIS	27
6.3	EXTRAÇÃO DE CARACTERÍSTICAS E ANÁLISE DE PADRÕES	28
7	RESULTADOS E DISCUSSÕES	30
7.1	RESULTADOS DO PIPELINE DE PROCESSAMENTO DE SINAIS SEMG	30
7.2	ANÁLISE DOS CLUSTERS E IDENTIFICAÇÃO DE FIBRAS MUSCULARES	33
7.3	SÍNTESE DAS MÉTRICAS DE DESEMPENHO DO PIPELINE	34
7.4	DISTRIBUIÇÃO E PRECISÃO NA IDENTIFICAÇÃO DE FIBRAS MUSCULARES DA REGIÃO INFERIOR ESQUERDA	35

8	CONCLUSÕES E PERSPECTIVAS FUTURAS	36
	REFERÊNCIAS	37
	APÊNDICE A – Disponibilidade de arquivos	39
A.1	INSTALAÇÃO E CONFIGURAÇÃO DO AMBIENTE	39
A.2	CÓDIGO FONTE DO PIPELINE DE PROCESSAMENTO DOS SINAIS DE SEMG	39
A.2.1	Pré-processamento dos Sinais	39
A.2.2	Extração de Características e Análise de Padrões . . .	42
A.2.3	Função principal de chamada para análise dos arquivos de sEMG	50

1 INTRODUÇÃO

A sEMG (Eletromiografia de Superfície - *Surface Electromyography*) é uma técnica não invasiva utilizada para registrar sinais bioelétricos da atividade neuromuscular, originados das correntes iônicas a partir do fluxo de íons (principalmente Na, K e Ca), formados pelos potenciais de ação nas fibras musculares, que são controladas pelas MUs (Unidades Motoras - *Motor Units*). Essas correntes iônicas variam conforme o tipo de fibra muscular esquelética envolvida, de contração lenta (Tipo I) ou de contração rápida (IIx/IIba - oxidativa e IIx/IIb - glicolítica), e a intensidade da ativação neuromuscular, refletindo padrões de disparos neuronais e adaptações fisiológicas (Basmajian; De Luca, 1985). Assim, os sinais de sEMG carregam informações essenciais sobre a funcionalidade muscular e as respostas do sistema neuromuscular a estímulos e condições clínicas.

A identificação dos tipos de fibras musculares e a estimativa das taxas de disparo neural são fundamentais para diagnósticos neuromusculares detalhados e para o monitoramento contínuo da saúde muscular (Merletti; Parker, 2004). A diferenciação precisa entre fibras de contração lenta e rápida permite avaliar a funcionalidade muscular, a fadiga e as adaptações decorrentes de treinamento físico. Alterações nesses padrões de ativação estão diretamente relacionadas à perda de massa muscular (*sarcopenia*) e a condições neuromusculares degenerativas, o que evidencia a relevância clínica de análises quantitativas para a prevenção e o tratamento dessas condições (Mani; Nagaraj, 2024).

Contudo, atualmente a identificação precisa dos tipos de fibras e das taxas de disparo neural é obtida de forma confiável apenas por métodos invasivos do exame, EMG intramuscular, realizado com eletrodos intramusculares, ou a partir biópsias musculares. Procedimentos esses, que são desconfortáveis e apresentam riscos aos pacientes (Basmajian; De Luca, 1985). Interferências eletromagnéticas externas, a sobreposição dos MUAPs (Potenciais de Ação de Unidades Motoras - *Motor Unit Action Potentials*) e a dificuldade de mapear a intensidade das correntes iônicas são desafios significativos neste contexto. Como consequência, os métodos tradicionais não invasivos não conseguem fornecer uma análise precisa e abrangente da atividade neuromuscular, limitando a capacidade de diagnóstico e monitoramento clínico.

Este estudo propõe um método integrativo para o processamento de sinais sEMG, visando a identificação de VMUAPs (Potenciais de Ação de Unidades Motoras Virtuais - *Virtual Motor Unit Action Potentials*) como estimativa estatística dos MUAPs (Potenciais de Ação de Unidades Motoras - *Motor Unit Action Potentials*), com essa aproximação estatística, classificar as fibras musculares a partir da decomposição do sinal de superfície. As análises foram conduzidas com dados experimentais obtidos sob duas condições controladas e protocoladas: repouso e contração com uso de uma carga de 50 kg. Em repouso, espera-se a predominância de fibras do tipo I (contração lenta), enquanto sob esforço intenso observa-se o recrutamento adicional de fibras dos tipos IIa e IIb (contração rápida,

de resistência e de potência). As coletas foram realizadas com o aparelho eletromiógrafo *Miotool* (Miotec) e eletrodos de superfície descartáveis no Laboratório de Fisioterapia da Universidade de Brasília (UnB), considerando contrações voluntárias de músculos de regiões superiores (Protocolo de Membro Superior) e inferiores (Protocolos de Membro Inferior - Vista anterior e posterior) de indivíduos, conforme as recomendações da SENIAM (Surface Electromyography for the Non-Invasive Assessment of Muscles) (Hermens; Freriks; Merletti, 2000). O método proposto utiliza técnicas de pré-processamento com filtragem *Notch*, passa-banda *Butterworth* de 4ª ordem e suavização por média móvel, seguido da extração de características como variância, desvio padrão, RMS (Valor eficaz - *Root-Mean-Square*), assimetria e curtose. Para análises temporais e espectrais, é utilizada a WT (Transformada Wavelet - *Wavelet Transform*), enquanto o algoritmo *K-Means*, descritos no Capítulo 6 e em (Umargono; Suseno; S.K., 2020), segmenta os padrões de ativação e identifica comportamentos dominantes.

Os resultados obtidos, com Inércia 0.75 ± 0.05 e acurácia de aproximadamente $98.5 \pm 2.5\%$, indicam que o algoritmo é capaz de realizar agrupamentos bem definidos e estimar as taxas de disparo neural por meio dos VMUAPs, além de classificar fibras de contração lenta e rápida, segmentando regiões do sinal e identificando frequências de ativação que definem o comportamento e a composição muscular.

A abordagem proposta fornece uma representação quantitativa que amplia a compreensão da relação entre atividade neuromuscular e possíveis irregularidades fisiológicas. No entanto, a validação clínica completa e a generalização do modelo para diferentes populações permanecem desafios futuros.

As seções deste artigo estão organizadas da seguinte forma: o capítulo 2 revisa os trabalhos relacionados, o progresso recente e os desafios no processamento de sinais de sEMG; o capítulo 3 apresenta a justificativa e as aplicações potenciais das análises de sinais de sEMG; o capítulo 5 descreve os objetivos do estudo; no capítulo 6 detalha-se a metodologia proposta, incluindo as etapas de processamento e análise; o capítulo 7 discute os resultados obtidos com dados experimentais; e o capítulo 8 apresenta as conclusões e direções para trabalhos futuros.

2 TRABALHOS RELACIONADOS

A identificação dos tipos de fibras musculares é essencial para o diagnóstico de doenças neuromusculares, como apresentado em (Basmajian; De Luca, 1985). Em (Mani; Nagaraj, 2024) é abordado um grupo de doenças genéticas que enfraquecem progressivamente os músculos responsáveis pelo movimento. Entre as principais, destacam-se a distrofia muscular de Duchenne e a de Becker, que afetam predominantemente os músculos voluntários, embora algumas formas também possam comprometer o coração e outros órgãos involuntários (Viana, 2020).

Em (Mani; Nagaraj, 2024) também discute-se a miosite, que inclui condições como polimiosite e dermatomiosite, caracterizadas por inflamação e degeneração do tecido muscular esquelético. Essas condições podem apresentar manifestações cutâneas e afetar indivíduos de diferentes faixas etárias. Já a atrofia muscular envolve a perda de massa muscular e pode ter causas diversas, incluindo doenças genéticas, desuso prolongado, desnutrição ou envelhecimento (Rodríguez-Tapia et al., 2019). Enquanto as fibras musculares de um músculo saudável apresentam um padrão interferencial típico durante a contração, músculos acometidos por condições miogênicas exibem um padrão interferencial denso e de baixa amplitude, conforme discutido em (Farina; Merletti; Enoka, 2014).

A captura do sinal EMG utilizando dois ou mais eletrodos contribui para amplificar e melhorar a qualidade dos sinais adquiridos, reduzindo o impacto de interferências geradas por outros sinais bioelétricos fisiológicos subjacentes e ruídos, como discutido em (Rodríguez-Tapia et al., 2019). Sendo também apresentadas diretrizes para a captação das correntes iônicas do EMG (Farina; Merletti; Enoka, 2014).

A melhoria de programas de reabilitação adaptados às necessidades específicas de cada indivíduo, resultando em recuperações mais eficientes e melhores desfechos a longo prazo, é abordada em (Basmajian; De Luca, 1985). A caracterização das fibras musculares pode auxiliar fisioterapeutas no desenvolvimento de planos de reabilitação personalizados e eficazes, otimizando o tempo de recuperação e os resultados dos pacientes. Ainda em (Basmajian; De Luca, 1985), discute-se os parâmetros de EMG mais relevantes para nortear a reabilitação musculoesquelética, assim como a variabilidade na resposta dos pacientes a diferentes protocolos de reabilitação, o que pode limitar a efetividade desses planos.

Também se destaca que o entendimento da fisiologia das fibras musculares pode contribuir para o desenvolvimento de novos medicamentos e terapias voltadas ao tratamento de doenças musculoesqueléticas. No entanto, as dificuldades na identificação de alvos terapêuticos, devido à complexidade da resposta muscular a diferentes intervenções terapêuticas e à variabilidade entre pacientes, são discutidas em (Basmajian; De Luca, 1985; Rodríguez-Tapia et al., 2019). A pesquisa sobre a fisiologia das fibras musculares e suas respostas a diferentes tratamentos é essencial para o avanço na compreensão e no tratamento de doenças musculoesqueléticas.

A identificação dos tipos de fibras musculares pode ainda beneficiar atletas, ajudando-os a compreender melhor o funcionamento de seus corpos e a otimizar treinamentos, o que contribui para melhor desempenho e prevenção de lesões, conforme demonstrado em (Mani; Nagaraj, 2024). Limitações na aplicação prática dos dados de EMG em treinamentos esportivos, devido à interpretação dos resultados em contexto dinâmico e à tradução de achados científicos para a prática esportiva cotidiana, são explicadas em (Farina; Merletti; Enoka, 2014). Nesse mesmo trabalho, destaca-se como a EMG pode ser utilizada para monitorar o desempenho de atletas ao longo do tempo, identificando potenciais problemas e ajustando treinamentos conforme necessidades individuais.

3 FUNDAMENTAÇÃO TEÓRICA

3.1 A ELETROMIOGRAFIA

A eletromiografia (EMG) é uma técnica para a análise da atividade neuromuscular, que possibilita a coleta de sinais bioelétricos associados à fisiologia muscular. O sinal é detectado e monitorado ao longo do tempo durante uma contração muscular, sendo formado pela superposição do Trem de Impulsos de Potenciais de Unidades Motoras (MUAPTs), gerado a partir das atividades das Unidades Motoras (MUs) e ruídos de fundo. Como mostra a Figura 1, a qual ilustra o trem de impulsos de potenciais das unidades motoras (MUAPTs) e o ruído de fundo característicos de um sinal de EMG. À esquerda (a), é mostrado o esquema da medula espinal e das fibras musculares, destacando três unidades motoras (MU1, MU2 e MU3), cada uma innervando no nervo motor um conjunto específico de fibras musculares, identificadas por cores distintas. À direita, o gráfico representa a atividade elétrica registrada ao longo do tempo, evidenciando como cada unidade motora contribui para o sinal total, cuja forma final é a soma dessas contribuições mais o ruído de fundo. As características de um sinal EMG dependem de alguns fatores, relacionados aos tamanhos e formatos dos elétrodos utilizados na coleta, a orientação individual de cada elétrodo relativo às fibras musculares, assim como às características fisiológicas e anatômicas da musculatura, que são diretamente vinculadas à idade, condição de saúde e de fadiga do paciente no momento de coleta. Portanto, análises de sinais de eletromiografia fornecem informações que podem ser utilizadas em contextos clínicos e para avaliações neuromusculares.

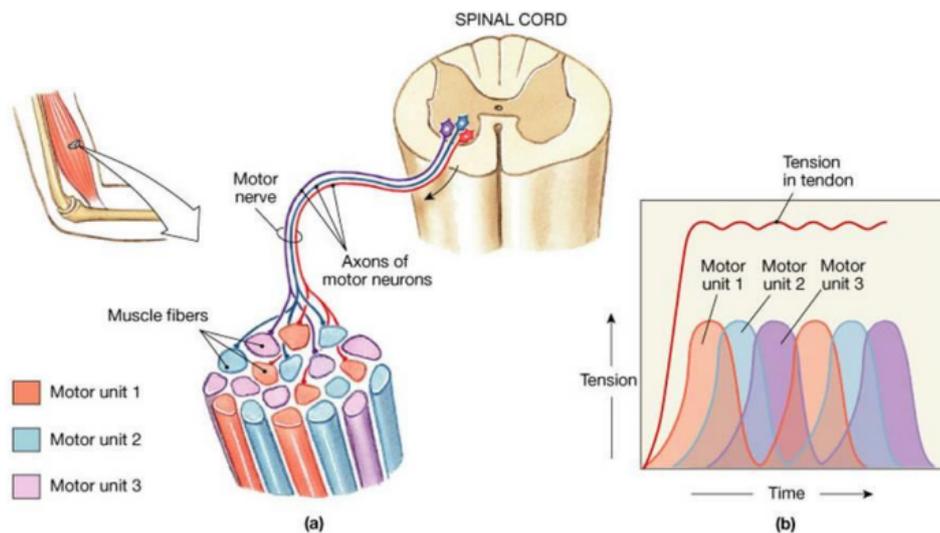


Figura 1 – Representação do Trem de Impulsos de Potenciais de Unidades Motoras (MUAPTs) e ruídos de fundo. Fonte: (Alcan; Zinnuroğlu, 2023)

Os sinais de EMG são adquiridos a partir da utilização de no mínimo dois elétrodos,

devido a uma gravação do sinal se basear em um sistema de medição da diferença de potencial entre duas ou mais regiões musculares, ou seja, um eletrodo ativo, que capta o sinal e o eletrodo de referência. A utilização de dois ou mais eletrodos auxilia por consequência, na amplificação e qualidade dos sinais adquiridos, reduzindo o impacto de interferências geradas por outros sinais bioelétricos fisiológicos subjacentes e ruídos (Alcan; Zinnuroğlu, 2023), permitindo a detecção de pequenas variações elétricas dos MUAPs e a identificação de padrões de ativação muscular, conforme a Figura 2, que mostra a disposição de eletrodos em diferentes regiões, a amplificação do sinal e a soma dos MUAPs.

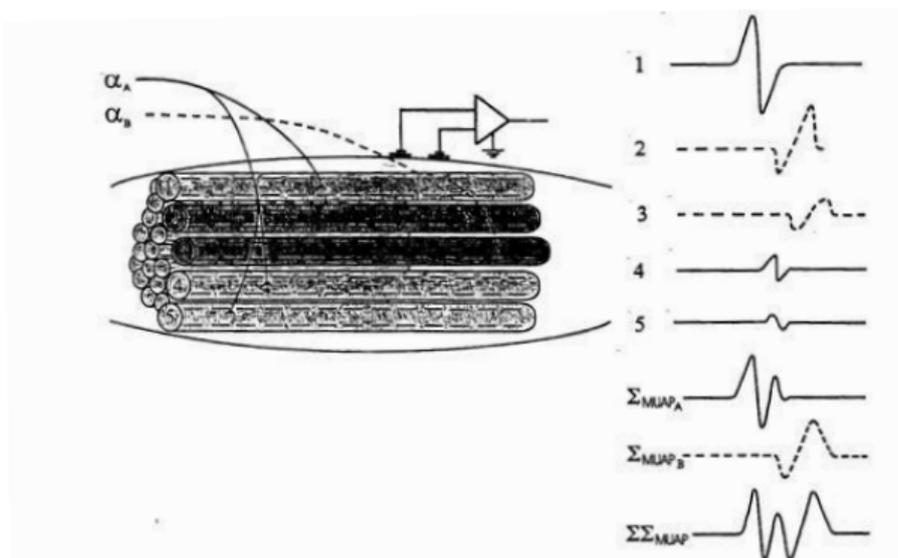


Figura 2 – Representação de um sinal de EMG como um somatório de MUAPs. Fonte: (Alcan; Zinnuroğlu, 2023)

3.2 COLETA DE SINAIS DE EMG

Duas formas principais de coleta dos sinais são comumente utilizadas, conforme representado na Figura 3: a forma invasiva (3-b), também conhecida como intramuscular ou *Needle EMG*, que se baseia na inserção de uma agulha concêntrica diretamente no músculo para registrar a atividade elétrica de uma área específica, e a forma não invasiva, ou *Surface EMG* (sEMG), que consiste na disposição de eletrodos de superfície na pele, sobre os músculos alvo (3-a). O *Needle EMG* oferece uma maior precisão na medição da atividade elétrica muscular, sendo capaz de captar sinais de uma região mais restrita e profunda do músculo, o que o torna particularmente útil para a avaliação de distúrbios neuromusculares e para estudos mais detalhados de áreas específicas de ativação muscular. Contudo, esse método invasivo envolve riscos como infecção, dor e um tempo de recuperação relativamente longo, o que restringe sua aplicação a ambientes controlados, como clínicas e centros de reabilitação. Por outro lado, o sEMG apresenta uma abordagem mais simples e segura, sem a necessidade de inserção de agulhas, mas em compensação, oferece uma resolução espacial menor e é mais suscetível a ruídos e interferências, o que pode

comprometer a precisão dos dados obtidos, especialmente em músculos profundos ou de difícil acesso.

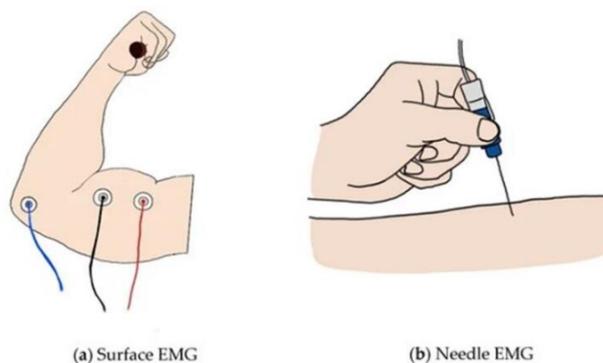


Figura 3 – Comparação entre os métodos de coleta de sinais de EMG: Needle EMG e Surface EMG. Fonte: (Alcan; Zinnuroğlu, 2023)

As extrações de dados dos sinais de EMG se baseiam em um processo de decomposição, nos quais cada MUAPT é extraído do sinal, de forma que sejam identificadas cada uma das MUs responsáveis por esse trem de impulsos. A partir disso, é possível estimar um padrão de disparos e o *template* de ativação das contribuições individuais de cada Potencial de Ação de Unidade Motora (MUAP), conforme (Alcan; Zinnuroğlu, 2023). Devido aos avanços tecnológicos, os métodos de análise, processamento de sinais e, em especial, o reconhecimento de padrões, levaram ao desenvolvimento de novas técnicas para a extração de informações dos sinais de EMG. A partir disso, diversos procedimentos têm se beneficiado dessas inovações, principalmente no diagnóstico de condições musculares e a personalização de estratégias de reabilitação especialmente para grupos-alvo específicos, como idosos que enfrentam a perda progressiva de massa muscular ou no setor esportivo, no qual o conhecimento da composição muscular e dos padrões de recrutamento de unidades motoras é crucial para otimizar o desempenho atlético e prevenir lesões (Alcan; Zinnuroğlu, 2023). Isso possibilita o desenvolvimento de programas de treinamento, otimizando o desempenho atlético e reduzindo o risco de lesões. Devido aos impactos positivos, a eletromiografia tem sido utilizada amplamente como uma das principais ferramentas para a realização de exames clínicos e estudos no diagnóstico de distúrbios neuromusculares, principalmente por meio da decomposição do sinal de EMG (Alcan; Zinnuroğlu, 2023).

3.3 DESAFIOS NA ANÁLISE DE SINAIS EMG

A análise de sinais de eletromiografia (EMG) enfrenta desafios que afetam a precisão e a aplicabilidade dos dados obtidos. Esses desafios abrangem desde a natureza do sinal até questões relacionadas à instrumentação e à fisiologia muscular. Esta seção discute os principais problemas enfrentados na aquisição, processamento e interpretação de sinais EMG.

3.3.1 Problemas Gerais dos Sinais de EMG

Os sinais de EMG são caracterizados por sua complexidade, devido à sobreposição de múltiplas fontes de ativação neuromuscular e à presença de ruídos. Alguns dos principais desafios incluem:

- **Baixa amplitude do sinal:** Os sinais de EMG são tipicamente de baixa amplitude, frequentemente na ordem de microvolts ($1\mu V$), o que torna a detecção e análise mais desafiadoras, especialmente em ambientes ruidosos (Mandrile et al., 2003);
- **Não estacionariedade:** Os sinais de EMG apresentam características não estacionárias, com variações nas propriedades estatísticas (como as médias e variâncias) ao longo do tempo que dificultam a análise e interpretação de tendências (Mandrile et al., 2003);
- **Baixa relação sinal-ruído (SNR):** O sinal de EMG é frequentemente contaminado por ruídos ambientais, como interferências eletromagnéticas, artefatos de movimento do corpo e interferências de outros sinais bioelétricos, como os do coração (ECG - Eletrocardiograma) e da atividade cerebral (EEG - Eletroencefalograma) (Merletti et al., 2010);
- **Dificuldade na decomposição do sinal:** A separação dos MUAPs individuais é um problema desafiador, devido à sobreposição de atividades musculares (Rodríguez-Tapia et al., 2019)(Merletti et al., 2010) e à variabilidade na morfologia dos sinais;
- **Influência de fatores externos:** O posicionamento dos eletrodos, a variabilidade anatômica e as condições fisiológicas do sujeito impactam significativamente a qualidade do sinal adquirido (Mandrile et al., 2003).

3.3.2 Problemas Relacionados à Fisiologia Muscular

A fisiologia do músculo também impõe desafios para a análise do EMG:

- **Heterogeneidade das unidades motoras:** Os músculos são compostos por diferentes tipos de fibras musculares controladas por unidades motoras, sendo essas classificadas em tipos I (fibras de contração lenta), IIa - oxidativas (fibras de contração rápida e resistência) e IIb - glicolíticas (fibras de contração rápida e potência), cada uma com propriedades distintas de ativação e recrutamento, o que dificulta a identificação de padrões específicos;
- **Fadiga muscular:** A fadiga pode alterar a morfologia do sinal EMG, tornando a análise comparativa mais complexa;

- **Atividade de tecidos adjacentes:** A interferência de sinais residuais provenientes de outros tecidos musculares pode comprometer a especificidade da análise.

3.3.3 Problemas de Instrumentação e Padronização

A instrumentação utilizada na aquisição do EMG desempenha um papel crítico na qualidade do sinal. Contudo, a falta de padronização nos sistemas de medição e processamento dificulta a reprodutibilidade dos estudos. Algumas das principais questões incluem:

- **Falta de padrões universalmente aceitos:** Embora existam diretrizes, como as estabelecidas pelo “Standards for EMG Instrumentation”, a variabilidade nos protocolos de aquisição e análise ainda persiste;
- **Diferenças entre equipamentos:** Sistemas de diferentes fabricantes podem apresentar variações nos filtros, na calibração e na amplificação, impactando a comparação entre estudos;
- **Necessidade de filtragem eficiente:** A remoção de artefatos e ruídos exige filtros adaptativos robustos para garantir um sinal mais limpo e confiável para análise (Mandrile et al., 2003).

A resolução desses problemas exige o desenvolvimento de métodos para a filtragem e análise dos sinais, além da adoção de padrões rigorosos para garantir a confiabilidade dos dados obtidos com validade clínica. A implementação de diretrizes, como as propostas pelo “Standards for EMG Instrumentation”, representa um avanço significativo na padronização dos procedimentos de aquisição e análise, contribuindo para a reprodutibilidade e precisão das medições de EMG. Mas para análises mais específicas, como a identificação de padrões de ativação muscular em tarefas motoras complexas, ainda há uma grande necessidade de padronização.

3.4 ABORDAGENS METODOLÓGICAS PARA A ANÁLISE DE SINAIS EMG

Para lidar com as limitações inerentes ao processamento de sinais mioelétricos de superfície (sEMG), este trabalho adota um método estruturado de análise de sinais, em etapas de pré-processamento, extração de características e clusterização, apoiando-se em técnicas implementadas na linguagem de programação Python e bibliotecas como NumPy, SciPy e scikit-learn, conforme descritas na Seção 6.

3.4.1 Pré-processamento do sinal

O estágio inicial de pré-processamento envolve a remoção de artefatos e interferências. Primeiramente, aplica-se uma filtragem *Notch* para eliminar ruídos de linha de energia

em frequências específicas (encontrados no Brasil na faixa de 60 Hz), que é representado pela seguinte equação:

$$H(f) = \begin{cases} 0, & \text{se } f = 60 \text{ Hz} \\ 1, & \text{caso contrário} \end{cases} \quad (1)$$

No qual $H(f)$ é a resposta em frequência do filtro, atenuando a frequência de 60 Hz para mitigar interferências eletromagnéticas provenientes da rede elétrica.

Após isso, é aplicada a filtragem *bandstop* Butterworth de 4ª ordem. O filtro é ajustado para garantir atenuação mínima, baseada na frequência de corte, sendo descrito na equação:

$$H(f) = \begin{cases} 0, & \text{se } f \in [f_1, f_2] \\ 1, & \text{caso contrário} \end{cases} \quad (2)$$

onde f_1 e f_2 são as frequências de corte inferior e superior, respectivamente, definidas para eliminar ruídos específicos.

Em seguida, aplica-se um filtro *Butterworth* passa-banda de 4ª ordem, restringindo o espectro a 20–240 Hz, que é descrito na literatura como parte do espectro de interesse (Wang; Tang; Bronlund, 2013). Para reduzir ruídos residuais de alta frequência, o sinal é suavizado por média móvel com janela deslizante de comprimento N para garantir uma melhor representação do sinal, dada por:

$$y[n] = \frac{1}{N} \sum_{k=0}^{N-1} x[n-k]$$

onde $x[n]$ é o sinal filtrado e $y[n]$ é o sinal suavizado, obtido através do cálculo da média dos últimos N pontos do sinal, deslocando-se um ponto por vez, de 0 a $N - 1$. A escolha do tamanho da janela N é crucial, pois janelas muito curtas podem resultar em perda de informações, enquanto janelas muito longas podem suavizar excessivamente o sinal, dificultando a identificação de características.

3.4.2 Extração de características

Após o pré-processamento, são extraídas métricas estatísticas para caracterizar a atividade muscular:

- **Média** (μ):

$$\mu = \frac{1}{N} \sum_{n=1}^N x[n]$$

Onde N é o número total de amostras e $x[n]$ representa o valor do sinal na amostra n .

- **Variância** (σ^2):

$$\sigma^2 = \frac{1}{N-1} \sum_{n=1}^N (x[n] - \mu)^2$$

No qual μ é a média do sinal, $x[n]$ é o valor do sinal na amostra n , e N é o número total de amostras. A variância mede a dispersão dos valores em relação à média, indicando a variabilidade do sinal.

- **Desvio padrão** (σ):

$$\sigma = \sqrt{\sigma^2}$$

Calculado como a raiz quadrada da variância. Fornece a medida da dispersão dos valores em relação à média.

- **RMS (Valor eficaz - *Root-Mean-Square*) em janelas móveis:**

$$\text{RMS} = \sqrt{\frac{1}{N} \sum_{n=1}^N x[n]^2}$$

Onde N é o número de amostras na janela e $x[n]$ é o valor do sinal na amostra n . O cálculo é realizado a partir da média dos quadrados dos valores do sinal, seguido pela raiz quadrada do resultado. O RMS é uma medida da magnitude do sinal, sendo especialmente útil para sinais de EMG, pois reflete a energia total do sinal ao longo do tempo.

- **Assimetria (Skewness):**

$$\text{Skewness} = \frac{1}{N} \sum_{n=1}^N \left(\frac{x[n] - \mu}{\sigma} \right)^3$$

Obtida a partir da média e do desvio padrão, a assimetria é uma medida que avalia a simetria do sinal em relação à média, onde μ é a média e σ o desvio padrão.

- **Curtose (Kurtosis):**

$$\text{Kurtosis} = \frac{1}{N} \sum_{n=1}^N \left(\frac{x[n] - \mu}{\sigma} \right)^4$$

A curtose é uma medida que avalia o “formato” da distribuição dos valores do sinal, indicando o quão “pontuda” ou “achatada” ela é. Valores altos de curtose indicam que o sinal tem picos mais agudos e valores extremos mais frequentes, enquanto valores baixos indicam uma distribuição mais achatada e com menos picos, sugerindo uma atividade muscular mais estável. É calculada a partir da média e do desvio padrão, onde $x[n]$ é a amostra n atual, μ é a média e σ o desvio padrão.

- **DWT (Transformada Wavelet Discreta - *Discrete Wavelet Transform*)**

$$\text{DWT}(a, b) = \int_{-\infty}^{\infty} x(t) \psi^* \left(\frac{t-b}{a} \right) dt$$

A DWT é uma técnica de análise tempo-frequência que decompõe o sinal $x(t)$ em diferentes escalas a e posições b , utilizando funções chamadas *wavelets-mãe* (ψ). Enquanto a

Transformada de Fourier fornece apenas a decomposição em frequência, a DWT preserva também a localização temporal de cada componente de frequência, permitindo identificar variações rápidas (alta frequência) e tendências mais lentas (baixa frequência). Diferente da Transformada Wavelet Contínua (CWT), a DWT utiliza uma base discreta de wavelets, o que a torna mais eficiente computacionalmente e adequada para sinais discretos, como os de EMG (Debnath, 2002).

No contexto dos sinais sEMG, essa propriedade é fundamental para detectar eventos transitórios, como ativações musculares curtas e mudanças nos padrões de disparo das MUs. Neste trabalho, foi utilizada a wavelet Daubechies-4, por seu bom desempenho na análise de sinais biomédicos com transições abruptas.

3.4.3 Clustering para identificação de padrões com K-Means

- K-Means

A segmentação é realizada via *K-Means*, minimizando a soma das distâncias quadráticas intra-cluster:

$$\mathbf{C} \arg \min_C \sum_{i=1}^k \sum_{x_j \in C_i} \|x_j - \mu_i\|^2$$

O algoritmo *K-Means* é um método de aprendizado não supervisionado que agrupa os dados em k clusters, minimizando a soma das distâncias quadráticas entre cada ponto x_j e o centróide μ_i do seu respectivo cluster (Umargono; Suseno; S.K., 2020).

No processamento de sEMG, cada vetor x_j representa um conjunto de características extraídas do sinal (por exemplo: variância, RMS, curtose, coeficientes de wavelet). O objetivo é separar automaticamente padrões mioelétricos que correspondam a diferentes tipos de fibras musculares (Tipo I, IIa e IIb).

- Avaliação da Homogeneidade intra-cluster

A homogeneidade dos grupos é medida pela variância intra-cluster:

$$\sigma_{\text{intra}}^2 = \frac{1}{|C_i|} \sum_{x_j \in C_i} \|x_j - \mu_i\|^2$$

Assim, os centróides podem ser revertidos ao espaço original via inversão de normalização e PCA.

3.4.4 Redução de Dimensionalidade com PCA

A Análise PCA (Análise de Componentes Principais - *Principal Component Analysis*) é uma técnica estatística multivariada utilizada para reduzir a dimensionalidade dos dados mantendo a maior parte da variabilidade original. Seu objetivo é encontrar novas direções — chamadas de componentes principais — que sejam combinações lineares das variáveis originais e que expliquem a maior parte da variância dos dados.

Matematicamente, o PCA resolve o problema de maximizar a variância projetada:

$$\mathbf{w}_1 = \arg \max_{\mathbf{w}} \text{Var}(\mathbf{w}^T \mathbf{X}) \quad \text{com} \quad \|\mathbf{w}\| = 1$$

onde \mathbf{X} representa a matriz de dados centrados na média, e \mathbf{w}_1 é o vetor de pesos da primeira componente principal.

Os vetores de pesos são obtidos como autovetores da matriz de covariância:

$$\mathbf{S} = \frac{1}{n-1} \mathbf{X}^T \mathbf{X} \quad \text{com} \quad \mathbf{S} \mathbf{w} = \lambda \mathbf{w}.$$

onde \mathbf{X} é a matriz de dados centrados na média, \mathbf{S} é a matriz de covariância e \mathbf{w} é o vetor de pesos (autovetor).

No processamento de sinais sEMG, o PCA é utilizado para:

- Visualizar os clusters obtidos pelo K-Means em 2 ou 3 dimensões.
- Reduzir ruído e manter apenas a informação mais relevante para classificação das fibras.
- Reverter normalizações, reconstruindo os centróides no espaço original para interpretação fisiológica.

A interpretação é simples: pontos próximos no espaço das componentes principais tendem a representar ativações musculares com padrões semelhantes de frequência e amplitude.

4 JUSTIFICATIVA

Os dispositivos de coleta e análise de sinais de sEMG atuais ainda não permitem a identificação de características intrínsecas da fisiologia neuromuscular, especialmente no que se refere à caracterização de comportamentos individuais das MUs e suas fibras musculares de controle, limitando a aplicação dessa tecnologia em contextos que exigem uma análise detalhada da atividade muscular (Rodríguez-Tapia et al., 2019). Ao fornecer informações detalhadas sobre a atividade muscular, as análises de sEMG possibilitam a criação de programas de reabilitação adaptados às necessidades específicas de cada indivíduo, resultando em recuperações mais eficientes e melhores resultados a longo prazo (Basmajian; De Luca, 1985), como descrito nas seções a seguir.

4.1 MEDICINA E FISIOTERAPIA

4.1.1 Diagnóstico mais preciso de doenças musculares

A identificação dos tipos de fibras musculares é crucial para o diagnóstico de doenças neuromusculares (Basmajian; De Luca, 1985). A distrofia muscular, por exemplo, é um grupo de doenças genéticas que enfraquecem progressivamente os músculos que controlam o movimento. Entre os principais tipos estão a distrofia muscular de Duchenne e Becker, que afetam principalmente os músculos voluntários, mas algumas formas também podem atingir o coração e outros órgãos involuntários. A miosite, que inclui condições como a polimiosite e dermatomiosite, é caracterizada pela inflamação e degeneração do tecido muscular esquelético, podendo apresentar manifestações cutâneas e impactar pessoas de todas as idades. A atrofia muscular, por sua vez, envolve a perda de massa muscular, podendo ser causada por uma variedade de fatores, incluindo doenças genéticas, desuso prolongado, desnutrição ou envelhecimento (Hermens; Freriks; Merletti, 2000). As fibras musculares de um músculo saudável exibem um padrão interferencial típico durante a contração, enquanto um músculo afetado por uma condição miogênica apresenta um padrão de interferência denso e de baixa amplitude (Meekins; So; Quan, 2008).

4.1.2 Desenvolvimento de novos medicamentos e terapias

Um melhor entendimento da fisiologia das fibras musculares pode levar ao desenvolvimento de novos medicamentos e terapias para o tratamento de doenças musculoesqueléticas (Hermens; Freriks; Merletti, 2000). Considerando que a complexidade da resposta muscular a diferentes intervenções terapêuticas e a variabilidade entre os pacientes podem dificultar a identificação de alvos terapêuticos (Merletti; Parker, 2004), a análise de dados minuciosos de sEMG pode fornecer informações complementares sobre a atividade muscular e a resposta ao tratamento, permitindo uma abordagem mais personalizada e eficaz em diferentes contextos.

4.2 CIÊNCIAS DO ESPORTE

4.2.1 Otimização do treinamento esportivo

A identificação das fibras musculares pode ajudar os atletas a compreender melhor seu próprio corpo e otimizar seu treinamento, levando a um melhor desempenho e prevenção de lesões (Rodríguez-Tapia et al., 2019). Contudo, a aplicação prática dos dados de sEMG no treinamento esportivo enfrenta desafios, como a interpretação dos resultados em um contexto dinâmico e a tradução dos achados científicos para a prática esportiva cotidiana (Garcia; Martinez; Rodriguez, 2020).

4.2.2 Monitoramento da performance

A sEMG pode ser utilizada para monitorar a performance dos atletas ao longo do tempo, identificando possíveis problemas e ajustando o treinamento de acordo com as necessidades individuais (Barbosa et al., 2019). No entanto, a interpretação dos dados de sEMG em um contexto esportivo requer uma compreensão profunda da biomecânica do movimento humano e das demandas específicas de cada modalidade esportiva (Hermens; Freriks; Merletti, 2000).

5 OBJETIVOS

A análise de sinais de EMG apresenta desafios significativos. A complexidade dos sinais e a variabilidade dos padrões de ativação muscular dificultam a interpretação dos dados e a extração de informações relevantes para a avaliação da função neuromuscular. Este capítulo apresenta as metas deste estudo, focando no desenvolvimento de métodos para processamento, filtragem e análise dos sinais de sEMG.

5.1 OBJETIVO GERAL

O objetivo geral deste trabalho é realizar uma análise aprofundada da atividade neuromuscular utilizando sinais de eletromiografia, com ênfase na diferenciação de tipos de fibras musculares e na estimativa das taxas de disparo neural das MUs, possibilitando a identificação de padrões de ativação muscular e a avaliação da coordenação e do recrutamento dessas fibras durante contrações voluntárias.

5.2 OBJETIVOS ESPECÍFICOS

Os objetivos específicos deste trabalho são:

- **Identificação das atividades neuromusculares no decorrer de todo o sinal:** Analisar e segmentar os sinais de sEMG para identificar as atividades neuromusculares ao longo do tempo, utilizando técnicas de processamento de sinais que permitam a decomposição e a extração de informações relevantes das Unidades Motoras.
- **Estimativa das taxas de disparos neurais:** Estimar as taxas de disparo das Unidades Motoras, considerando a atividade elétrica registrada pela sEMG, para avaliar a coordenação e o recrutamento muscular durante contrações voluntárias e involuntárias.
- **Diferenciação dos tipos de fibras musculares:** Aplicar métodos de análise multivariada com base nos sinais sEMG para identificar e diferenciar os tipos de fibras musculares presentes no sinal, com o objetivo de obter estimativas das características fisiológicas do músculo.

Esses objetivos visam contribuir para o entendimento da função neuromuscular e para o desenvolvimento de técnicas de análise de sinais que possam ser aplicadas em contextos clínicos e de pesquisa, melhorando a capacidade de diagnóstico e monitoramento de condições neuromusculares. No capítulo 6, serão apresentados os métodos utilizados para alcançar esses objetivos, incluindo as técnicas de processamento de sinais e análise estatística aplicadas aos dados de sEMG.

6 METODOLOGIA

Para abordar esse problema, este estudo propõe uma abordagem integrativa baseada no processamento de sinais discretos e em métodos de aprendizado de máquina, com o objetivo de explorar padrões característicos dos sinais da sEMG e, assim, realizar a identificação das diferentes tipos de fibras musculares a partir dos dados obtidos pelo procedimento não invasivo.

Este capítulo apresenta as etapas metodológicas principais realizadas no processamento e análise dos sinais conforme o fluxograma da Figura 4, que ilustra o fluxo de trabalho desde a coleta dos sinais até a análise final, seguindo a ordem: Coleta dos Sinais, Pré-processamento, Análise e Extração de Características, Seleção e Segmentação baseada nos padrões de ativação dos picos do sinal, Clusterização das informações extraídas, Validação e Redução de Dimensionalidade.

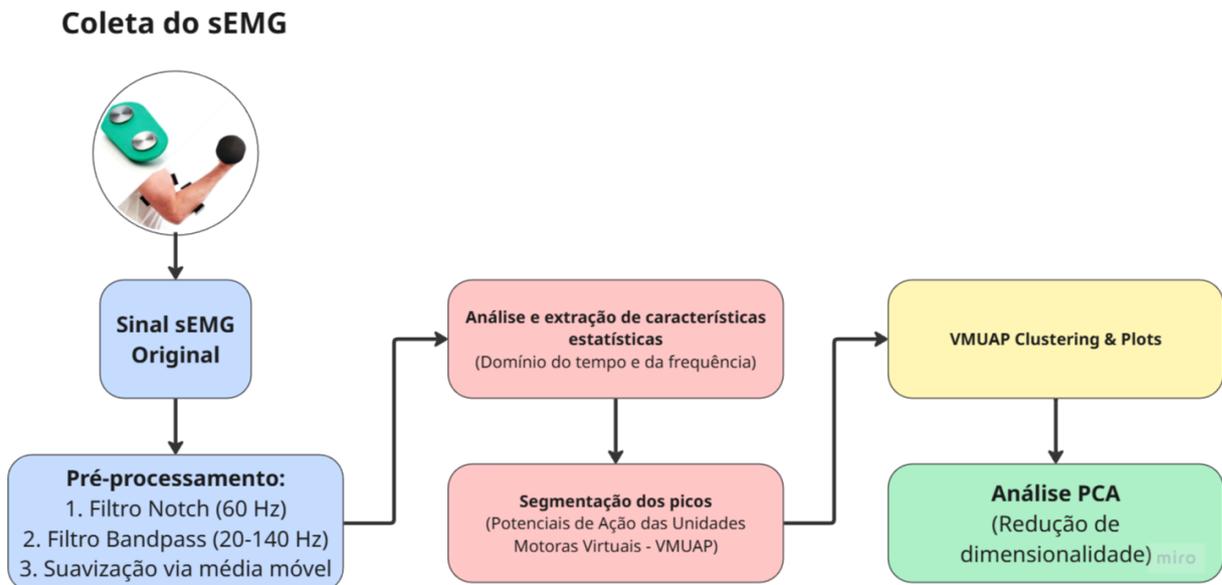


Figura 4 – Fluxograma das etapas metodológicas. Fonte: Elaborado pelo autor.

Todo o fluxo foi implementado com bibliotecas como *Pandas*, *Scipy*, *Numpy*, *Scikit-Learn* e *Matplotlib*, sendo executado em ambiente Colab Notebook. O código se encontra disponível na seção A.2 do Apêndice A deste trabalho.

6.1 COLETA DOS SINAIS DE SEMG

Os sinais de sEMG utilizados neste estudo foram coletados por professoras da Universidade de Brasília (UnB), seguindo um protocolo rigoroso para garantir a qualidade e a reprodutibilidade dos dados. A coleta foi realizada utilizando um sistema de eletromiografia não invasivo, com eletrodos posicionados em grupos musculares alvo conforme as diretrizes do SENIAM (Surface Electromyography for the Non-Invasive Assessment of

Muscles), que padroniza a localização dos eletrodos para diferentes grupos musculares. O processo de coleta seguiu três etapas principais:

1. **Preparação da pele:** Antes da aplicação dos eletrodos, foi realizada a tricotomia (remoção dos pelos) e a higienização do local, a fim de minimizar a impedância do contato eletrodo-pele e melhorar a qualidade do sinal adquirido.
2. **Posicionamento dos eletrodos:** O paciente foi posicionado de acordo com a área de coleta, com base nos posicionamentos estabelecidos pelo SENIAM. A região de maior contração motora foi identificada visualmente e utilizada como referência para a fixação dos eletrodos.
3. **Execução dos protocolos de testagem:** Foram seguidos três protocolos distintos para aquisição dos sinais sEMG, conforme descrito abaixo:
 - **Protocolo de membro superior:** O paciente foi posicionado em sedestação (sentado), com as mãos em posição anatômica (viradas para cima). Foi aplicada uma resistência de 50 kg por meio de uma faixa elástica posicionada a 15 cm do braço. O voluntário realizou três contrações máximas de elevação dos braços, cada uma com duração de 5 segundos, com descanso de 5 segundos entre as repetições. O estímulo verbal "*Força*" foi utilizado ritmadamente a cada 1 segundo para padronizar o esforço.
 - **Protocolo de membro inferior – vista anterior:** O paciente foi posicionado sentado, com os joelhos flexionados a 90 graus e os pés apoiados em uma escada. A resistência de 50 kg foi aplicada da mesma forma que no protocolo anterior. O voluntário realizou três contrações máximas de extensão dos joelhos, seguindo o mesmo padrão de tempo e estímulo verbal.
 - **Protocolo de membro inferior – vista posterior:** O paciente foi posicionado em decúbito ventral (barriga para baixo), com as pernas estendidas. A resistência foi aplicada da mesma forma, e o voluntário realizou três contrações máximas de flexão do joelho, também seguindo o mesmo padrão de tempo e estímulo verbal.

6.2 PRÉ-PROCESSAMENTO DOS SINAIS

Os sinais coletados foram armazenados automaticamente pelo software do dispositivo de aquisição em arquivos no formato EDF (European Data Format), um padrão amplamente utilizado para dados biomédicos, mas que apresenta limitações para manipulação e análise computacional direta. Para facilitar o processamento posterior, os arquivos EDF foram convertidos para formatos mais acessíveis, como CSV (Comma-Separated

Values) e XLSX (Excel Spreadsheet), permitindo uma análise mais eficiente e compatível com ferramentas estatísticas e de aprendizado de máquina.

Após a conversão dos dados, foi realizado o pré-processamento para melhorar a relação sinal-ruído, removendo artefatos e ruídos indesejados, e preparar os sinais para as análises subsequentes. As principais etapas de pré-processamento incluem:

- **Filtro Notch (60 Hz):** Remoção da interferência proveniente da rede elétrica por meio da qual o aparelho de coletas é alimentado. Esse filtro é essencial para eliminar ruídos de 60 Hz, comuns em ambientes urbanos, que podem distorcer a análise dos sinais de sEMG (Merletti et al., 2010).
- **Filtro Passa-Banda Butterworth de 4ª Ordem:** Este filtro permite o isolamento das frequências associadas à atividade neuromuscular. Conforme descrito em (Merletti et al., 2010), a faixa de 20-240 Hz é considerada a mais apropriada para capturar a atividade das diferentes fibras musculares, tanto de contração lenta quanto rápida.
- **Suavização do sinal:** Aplicação do cálculo da média móvel com uma janela de 50 amostras para suavizar o sinal e reduzir flutuações rápidas e ruídos de alta frequência, proporcionando um sinal mais estável para análise. Essa quantidade de amostras foi escolhida com base na frequência de amostragem de aproximadamente 2000 Hz da captura original, garantindo uma resposta adequada às variações do sinal.

6.3 EXTRAÇÃO DE CARACTERÍSTICAS E ANÁLISE DE PADRÕES

Após o pré-processamento, foram extraídas características relevantes do sinal para avaliar padrões de ativação muscular e dinâmica das unidades motoras:

- **Análise no Domínio do Tempo e da Frequência:** Extração de características estatísticas (média, variância, desvio padrão, RMS, skewness e curtose), bem como análise espectral baseada na Transformada Wavelet para avaliação da dinâmica dos VMUAPs, utilizando-a também para capturar as variações de frequência ao longo do tempo, permitindo uma análise mais detalhada das características dos sinais de sEMG.
- **Detecção de Potenciais de Ação das Unidades Motoras Virtuais (VMUAPs):** Identificação de VMUAPs por meio de detecção adaptativa de picos, utilizando um limiar dinâmico baseado na média e desvio padrão do sinal filtrado. Adicionalmente, a segmentação dos VMUAPs foi refinada através da correspondência de templates, baseados nas características estatísticas predominantes na região do sinal, para garantir a consistência das formas de onda.

- **Clusterização de VMUAPs:** Devido às características estocasticamente variáveis e não estacionárias dos MUAPs (Rodríguez-Tapia et al., 2019), a clusterização foi realizada para agrupar os sinais em classes distintas com base nas características extraídas. O algoritmo de aprendizado não supervisionado K-Means foi utilizado para identificar padrões comuns e diferenciar as estimativas de fibras musculares de contração lenta (Tipo I) e de contração rápida (IIx/IIba e IIb) a partir dos dados processados. A escolha do número de clusters foi baseada na análise do método do cotovelo (*Elbow Method*) (Umargono; Suseno; S.K., 2020), que avalia a variação intra-cluster em relação ao número de clusters especificado, permitindo identificar o ponto onde a adição de mais clusters não resulta em uma redução significativa da variação intra-cluster.
- **Visualização 3D dos Clusters:** Utilização de gráficos tridimensionais para representar os clusters formados, onde cada ponto representa um rótulo de cluster. Essa visualização permite uma análise mais intuitiva das relações entre os clusters e facilita a identificação de padrões complexos nos dados e suas interações.
- **Análise de Componentes Principais (PCA) dos Clusters:** Aplicada para reduzir a dimensionalidade dos dados e facilitar a visualização dos clusters formados, mantendo as características mais relevantes do sinal. A PCA ajuda a identificar as direções de maior variância nos dados conforme descrito no Capítulo 3, permitindo uma melhor interpretação dos padrões de ativação muscular.

7 RESULTADOS E DISCUSSÕES

Com a aplicação das etapas de pré-processamento e segmentação dos sinais de sEMG, é possível analisar a atividade muscular de forma mais precisa, conforme demonstrado nas seções anteriores. Nesta seção, apresentamos os resultados obtidos com a aplicação do algoritmo K-Means para clusterização dos rótulos dos picos dos sinais processados, utilizando a Transformada Wavelet como principal ferramenta de análise (Debnath, 2002).

7.1 RESULTADOS DO PIPELINE DE PROCESSAMENTO DE SINAIS SEMG

O pré-processamento dos sinais de sEMG, conforme descrito na Seção 6.2, tem como resultado a remoção de ruídos e artefatos, permitindo uma análise mais clara da atividade muscular, conforme a Figura 5, que destaca as etapas: Sinal após aplicação do filtro Notch (60 Hz), Sinal após o filtro passa-banda (20-240 Hz), suavização do sinal por média móvel e por fim, após o cálculo do RMS.

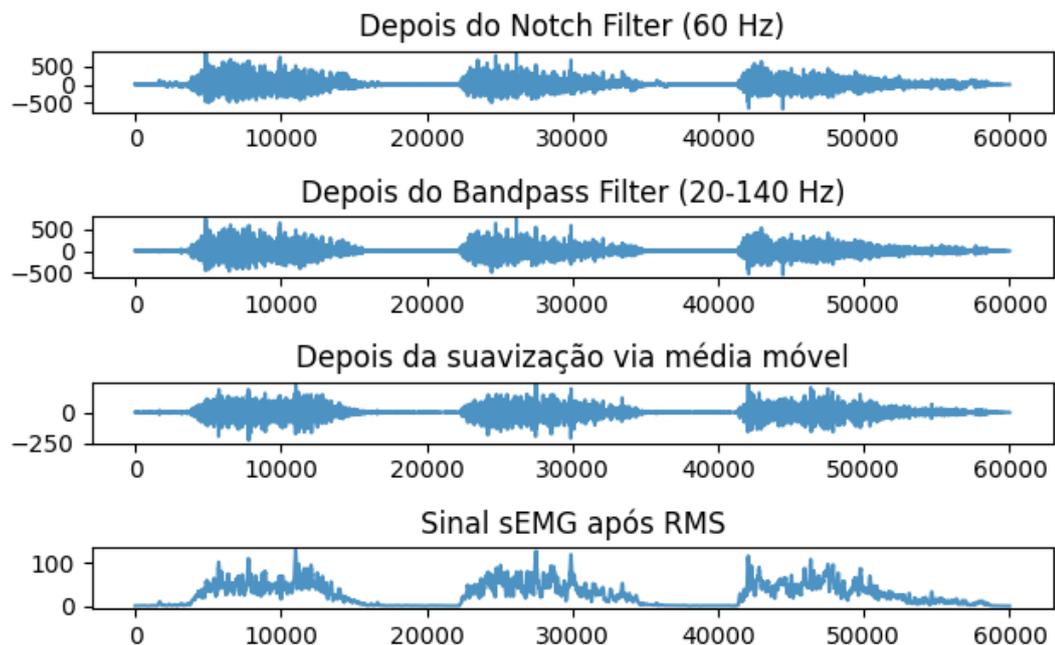


Figura 5 – Pré-processamento do sinal de sEMG. Fonte: Elaborado pelo autor.

A segmentação dos picos do sinal processado é realizada com base na Transformada Wavelet, conforme descrito na Seção 6.3 e em (Debnath, 2002). A Figura 6 ilustra o processo de segmentação dos picos do sinal de sEMG, destacando os picos identificados, em vermelho, após a aplicação da Transformada Wavelet, que permite uma análise mais nítida da atividade muscular.

Com os picos segmentados, os rótulos são atribuídos baseando nas características dos picos, como altura, largura e distância entre os picos. Esses rótulos são então utilizados

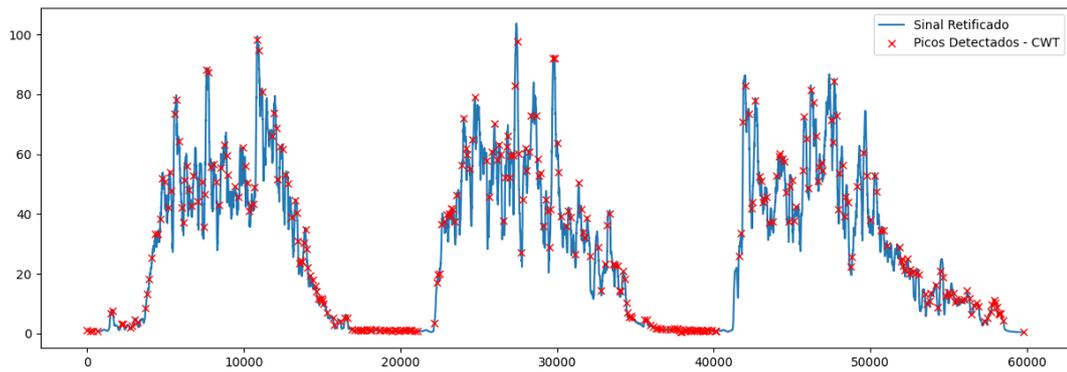


Figura 6 – Segmentação dos picos do sinal de sEMG. Fonte: Elaborado pelo autor.

para a clusterização dos dados, conforme descrito na Seção 6.3. Posteriormente, a Figura 7 apresenta o gráfico de clusterização dos rótulos do algoritmo K-Means aplicado a um sinal processado da região inferior esquerda de um membro inferior, detectado com o auxílio da Transformada Wavelet. Cada ponto colorido representa um pico identificado no sinal, e as cores indicam os diferentes clusters formados pelo K-Means. A análise visual dos clusters permite observar a separação clara entre os diferentes tipos de fibras musculares, sugerindo que o método é eficaz na segmentação dos dados.

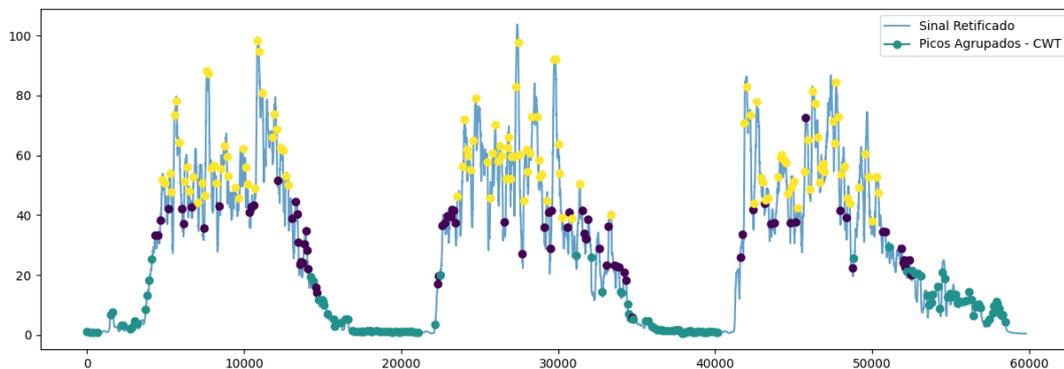


Figura 7 – Clusterização K-Means dos rótulos de um sinal muscular processado da região esquerda do membro inferior com Transformada Wavelet. Fonte: Elaborado pelo autor.

A Figura 8 mostra o mesmo sinal da Figura 7, porém com a visualização em 3D dos clusters formados pelo K-Means, onde cada ponto representa um rótulo de cluster. É possível visualizar o comportamento tridimensional dos picos de um dos sinais processados, sendo o eixo X a altura dos picos, o eixo Y a largura dos picos e o eixo Z a distância entre os picos, proporcionando uma perspectiva mais abrangente da segmentação dos dados.

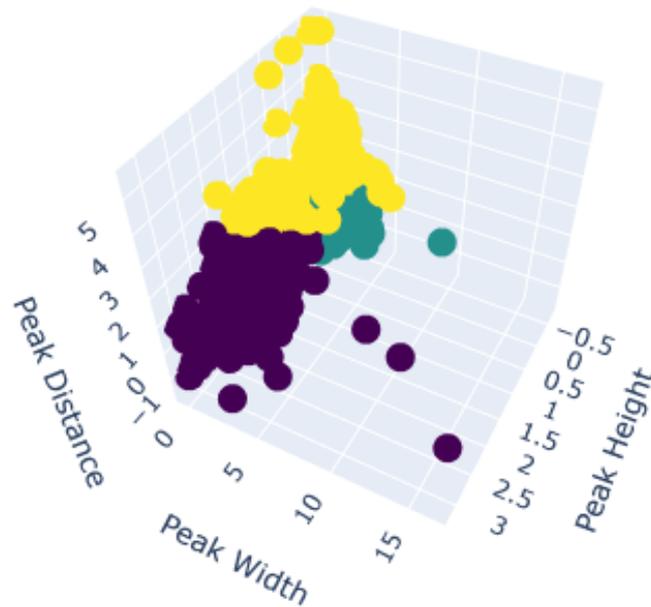


Figura 8 – Visualização 3D da clusterização K-Means do mesmo sinal processado apresentado na Figura 7. Fonte: Elaborado pelo autor.

A visualização tridimensional do comportamento dos picos do sinal processado, sendo o eixo X a altura dos picos, o eixo Y a largura dos picos e o eixo Z a distância entre os picos, proporciona uma perspectiva mais abrangente da segmentação dos dados. Essa abordagem tridimensional não apenas facilita a identificação de padrões complexos, mas também permite uma análise mais intuitiva das relações entre os clusters. A visualização em 3D é especialmente útil quando se trabalha com dados de alta dimensionalidade, pois ajuda a evitar a sobreposição de informações e a destacar as características distintivas de cada cluster.

As Figuras 7 e 8 ilustram a segmentação e visualização das diferentes fibras musculares identificadas por meio da clusterização K-Means. Os resultados demonstram experimentalmente a viabilidade da abordagem proposta, evidenciando a diferenciação das atividades musculares em repouso e sob esforço.

A análise permite estimar o número virtual de Unidades Motoras e das diferentes fibras musculares do tipo I, IIa e IIb. Considera-se, para efeito de interpretação, que nos primeiros 5 segundos apenas fibras do tipo I, responsáveis pela postura, estão ativas. Esse nível inicial define uma intensidade de referência para as fibras do tipo I, servindo como

base para o processamento adicional. Após os 5 segundos, as fibras do IIx/IIba começam a ser recrutadas de forma incremental até que, sob esforço máximo, as fibras do IIx/IIbb também são acionadas em menor proporção. O processo estatístico baseado no K-Means deve discriminar principalmente as fibras do IIx/IIba e IIb, já que há uma referência prévia para as fibras do tipo I.

7.2 ANÁLISE DOS CLUSTERS E IDENTIFICAÇÃO DE FIBRAS MUSCULARES

A partir da análise dos clusters formados, aplicando-se o PCA (Análise de Componentes Principais - *Principal Component Analysis*) com dimensionalidade reduzida para duas dimensões, é possível observar que os clusters correspondem a diferentes tipos de fibras musculares, conforme esperado. A Figura 9 ilustra a distribuição dos clusters no espaço de componentes principais, evidenciando a separação entre as fibras do tipo I, IIa e IIx/IIb. O que também é observado é que os clusters formados estão bem definidos, com uma clara distinção entre as diferentes fibras musculares, o que valida a eficácia do método de clusterização aplicado.

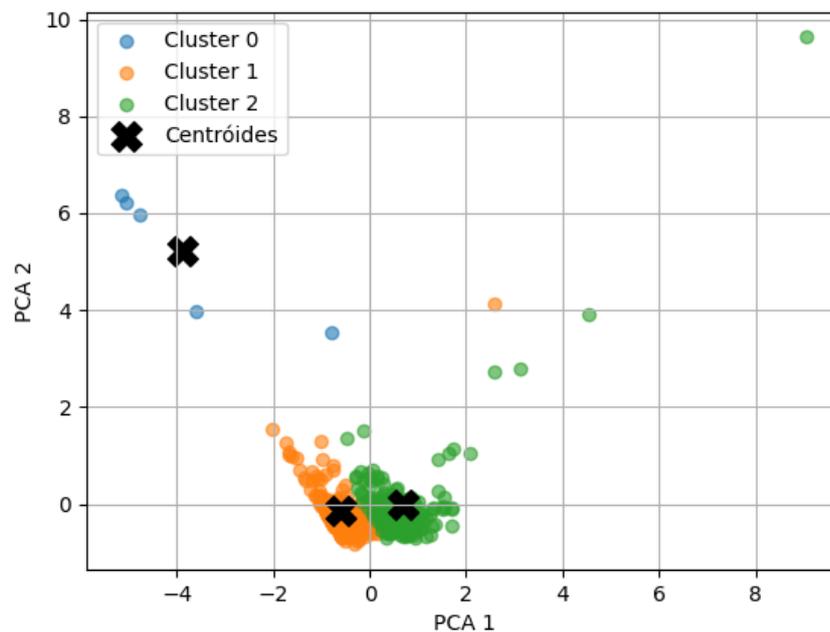


Figura 9 – Distribuição dos clusters no espaço de componentes principais. Fonte: Elaborado pelo autor.

7.3 SÍNTESE DAS MÉTRICAS DE DESEMPENHO DO PIPELINE

Para avaliar a eficácia do pipeline de processamento proposto, foram utilizadas métricas tanto de classificação quanto de agrupamento, considerando todos os sinais amostrados entre as diferentes regiões musculares. As métricas adotadas foram: Acurácia, Silhouette Score, Inertia e Calinski-Harabasz Index. A Tabela 1 apresenta os valores médios obtidos.

Tabela 1 – Métricas de desempenho médias do pipeline de processamento de sinais sEMG.

Métrica	Pré-DWT	DWT - Integrada
Acurácia (%)	75.5	98.5
Inertia	0.30	0.75
Silhouette Score	0.15	0.65
Calinski-Harabasz Index	270.5	875

- **Acurácia:** representa a proporção de classificações corretas em relação ao total de predições baseando-se nas correspondências de ativação em períodos de repouso e contração além das faixas de frequência dispostas na literatura para cada tipo de fibra muscular (Tipo I - 60-100 Hz, Tipo IIa - 30-60 Hz, Tipo IIX/IIB - 0-30 Hz)(Rodríguez-Tapia et al., 2019), sendo adequada para avaliar a capacidade discriminativa do pipeline. Os resultados indicam um ganho expressivo de desempenho após a integração da DWT, com acurácia média passando de 75.5% para 98.5%.

No contexto de agrupamento, o **Silhouette Score** varia entre -1 e 1, e mede o quão bem cada amostra está posicionada dentro do seu cluster, considerando a separação em relação aos clusters vizinhos. Valores acima de 0.5, como observado após a integração da DWT, indicam agrupamentos bem definidos.

A métrica **Inertia** corresponde à soma das distâncias quadradas internas aos clusters. Embora valores menores indiquem maior coesão intra-cluster, a interpretação isolada dessa métrica pode ser limitada; por isso, seu uso é mais eficaz em comparação relativa entre diferentes configurações.

Já o **Índice de Calinski-Harabasz** avalia a razão entre a dispersão entre clusters e a dispersão interna. Quanto maior o valor, melhor a separação entre os grupos. O aumento significativo desse índice após a aplicação da DWT reforça a eficiência do pipeline em produzir agrupamentos mais coesos e distintos.

Em conjunto, os resultados demonstram que o pipeline integrado com a transformada wavelet promove melhorias substanciais tanto na classificação quanto na estrutura dos agrupamentos extraídos dos sinais sEMG.

7.4 DISTRIBUIÇÃO E PRECISÃO NA IDENTIFICAÇÃO DE FIBRAS MUSCULARES DA REGIÃO INFERIOR ESQUERDA

Os resultados consolidados para os músculos da região inferior esquerda indicam uma distribuição estimada de **56% de fibras do tipo I** (classe 0), **38% de fibras do IIx/IIba** (classe 1) e **6% de fibras do IIx/IIbb** (classe 2), considerando o total de 1536 MUAPs analisados, nos arquivos dessa classe.

As métricas de desempenho foram elevadas em todos os agrupamentos após a integração da DWT, como o Silhouette Score que teve uma melhoria percentual de 30% e o Calinski-Harabasz que aumentou em 50%, reforçando a eficácia do pipeline proposto na identificação dos diferentes tipos de fibras musculares, mesmo em músculos de grande complexidade anatômica e funcional.

Tabela 2 – Distribuição percentual aproximada das fibras musculares identificadas nos músculos da região inferior esquerda.

Tipo de Fibras	Proporção (%)
Tipo I (Classe 0)	56%
IIx/IIba (Classe 1)	38%
IIx/IIbb (Classe 2)	6%

Obtendo-se uma média de acurácia para as amostras disponíveis nesta classe, para os tipos de fibras musculares, em cerca de 98.5% de acerto e Inércia 0.75 ± 0.05 . Esses resultados demonstram a eficácia do método proposto para a identificação e diferenciação de características presentes dos sinais de sEMG, possibilitando a análise detalhada da atividade muscular e a identificação das diferentes fibras musculares. A metodologia aplicada, combinando o pré-processamento dos sinais, a segmentação por Transformada Wavelet e a clusterização K-Means, mostra-se promissora para futuras pesquisas na área de eletromiografia e suas aplicações em reabilitação, ergonomia e treinamento esportivo.

8 CONCLUSÕES E PERSPECTIVAS FUTURAS

A sEMG (Eletromiografia de Superfície - *Surface Electromyography*) mostrou-se uma ferramenta promissora para a avaliação não invasiva da atividade neuromuscular. Este trabalho demonstrou a viabilidade de estimar virtualmente o número de Unidades Motoras e a composição de fibras musculares a partir do processamento de sinais discretos de sEMG. A abordagem proposta integrou etapas de filtragem, extração de características, análise tempo-frequência e clusterização, permitindo discriminar níveis de ativação e diferenciar tipos de fibras musculares em distintas condições de esforço.

Os resultados obtidos indicam potencial para aplicação clínica em diagnósticos, monitoramento de reabilitação e prevenção de condições associadas à perda de massa muscular, devido aos índices de separação dos padrões de ativação muscular ao longo do sinal. No entanto, são necessários estudos adicionais para validar a metodologia em populações diversas e correlacionar os parâmetros estimados com medidas invasivas de referência. A validação clínica completa e a generalização do modelo para diferentes populações permanecem desafios futuros. A integração com dados obtidos por métodos invasivos pode ampliar a robustez dos resultados, viabilizando aplicações em diagnósticos, monitoramento em reabilitação e prevenção de condições associadas à perda de massa muscular.

Como perspectivas para continuidades futuras, destaca-se a aplicação da abordagem proposta a outros sinais biomédicos com comportamento não linear e não estacionário, como séries de intervalos RR da variabilidade da frequência cardíaca, além de sinais acústicos e vibroacústicos. Investigações futuras podem também explorar melhorias na combinação dessa metodologia com algoritmos de aprendizado de máquina supervisionado, a fim de otimizar a classificação automática de padrões neuromusculares e ampliar as possibilidades de uso em sistemas portáteis de monitoramento.

REFERÊNCIAS

ALCAN, V.; ZINNUROĞLU, M. Current developments in surface electromyography. **Turkish Journal of Medical Sciences**, v. 53, n. 5, p. 1019–1031, 2023.

BASMAJIAN, J. V.; DE LUCA, C. J. **Muscles alive: their functions revealed by electromyography**. [S.l.]: Williams & Wilkins, 1985.

DEBNATH, L. **Wavelet Transform and Their Applications**. Boston: Birkhauser Verlag, 2002.

FARINA, D.; MERLETTI, R.; ENOKA, R. M. Surface EMG crosstalk evaluated from experimental recordings and simulated signals. **Critical reviews in biomedical engineering**, v. 42, n. 4-5, p. 1–20, 2014.

GARCIA, B.; MARTINEZ, C.; RODRIGUEZ, D. Application of Electromyography in Sports Science: Current Trends and Future Directions. **Sports Medicine**, v. 48, n. 3, p. 211–225, 2020.

HERMENS, H. J.; FRERIKS, B.; MERLETTI, R. Development of recommendations for EMG sensors and sensor placement procedures. **Journal of Electromyography and Kinesiology**, v. 10, n. 5, p. 361–374, 2000.

MANDRILE, F.; FARINA, D.; POZZO, M.; MERLETTI, R. Stimulation artifact in surface EMG signal: effect of the stimulation waveform, detection system, and current amplitude using hybrid stimulation technique. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, v. 11, n. 4, 2003.

MANI, R.K.; NAGARAJ, B. Diagnosis and classification of neuromuscular disorders using Bi-LSTM optimized with grey Wolf optimizer for EMG signals. **Scientific Reports**, v. 19274, n. 15, 2024.

MEEKINS, G. D.; SO, Y.; QUAN, D. American Association of Neuromuscular & Electrodiagnostic Medicine evidenced-based review: use of surface electromyography in the diagnosis and study of neuromuscular disorders. **Muscle Nerve**, v. 38, n. 4, p. 1219–1224, 2008.

MERLETTI, R.; BOTTER, A.; CESCO, C.; MINETTO, M. A.; VIEIRA, T. M. Advances in surface EMG: recent progress in clinical research applications. **Critical Reviews in Biomedical Engineering**, v. 38, n. 4, p. 347–379, 2010.

MERLETTI, R.; PARKER, P. A. **Electromyography: physiology, engineering, and noninvasive applications**. [S.l.]: John Wiley & Sons, 2004.

RODRÍGUEZ-TAPIA, B.; SOTO, I.; MARTÍNEZ, D. M.; CANDOLFI ARBALLO, N. Myoelectric Interfaces and Related Applications: Current State of EMG Signal Processing—A Systematic Review. **IEEE Access**, v. 8, p. 123456–123467, 2019.

UMARGONO, E.; SUSENO, J. E.; S.K., V. G. K-Means Clustering Optimization Using the Elbow Method and Early Centroid Determination Based on Mean and Median Formula. **International Journal of Advanced Computer Science and Applications**, v. 11, n. 1, p. 1–6, 2020.

VIANA, R. G. de O. **Classificação de Hipertensos Utilizando o Sinal da Variabilidade da Frequência Cardíaca (VFC) com Auxílio de Inteligência Artificial**. 2020. Diss. (Mestrado) – Universidade Federal de Mato Grosso do Sul, Campo Grande.

WANG, J.; TANG, L.; BRONLUND, J.E. Surface EMG Signal Amplification and Filtering. **International Journal of Computer Applications**, v. 82, n. 1, 2013.

APÊNDICE A – Disponibilidade de arquivos

A.1 INSTALAÇÃO E CONFIGURAÇÃO DO AMBIENTE

O ambiente de desenvolvimento utilizado neste trabalho foi configurado utilizando o *Anaconda*, que permite a criação de ambientes virtuais e a gestão de pacotes de forma eficiente.

A partir disso, é necessária a instalação dos pacotes utilizados no desenvolvimento do projeto. Os principais pacotes utilizados incluem:

- `numpy`: para manipulação de arrays e operações matemáticas.
- `scipy`: para processamento de sinais e análise de dados.
- `matplotlib`: para visualização de dados.
- `scikit-learn`: para implementação de algoritmos de aprendizado de máquina.
- `pywavelets`: para a Transformada Wavelet.

O seguinte comando pode ser utilizado para instalar os pacotes necessários no ambiente Anaconda:

```
1 !pip install pandas numpy matplotlib mplcursors scipy
   scikit-learn pywavelets padasip EMD-signal PyWavelets
   plotly fastdtw
```

A.2 CÓDIGO FONTE DO PIPELINE DE PROCESSAMENTO DOS SINAIS DE SEMG

O código fonte do pipeline de processamento dos sinais de sEMG foi dividido nas etapas de pré-processamento, extração de características e análise de padrões. As subseções a seguir detalham o código utilizado em cada etapa, com o objetivo de facilitar a compreensão e a replicação dos procedimentos realizados neste estudo.

A.2.1 Pré-processamento dos Sinais

```
1 from scipy.signal import butter, sosfilt, sosfreqz
2
3 def notch_filter(data, notch_freqs, sampling_rate,
4     quality_factor=30):
5     """
6     Aplica filtros Notch para remover interferencias em frequencias
7     especificas (por exemplo, 60 Hz).
8
9     Parameters:
```

```
8 data (array): Sinal EMG bruto.
9 notch_freqs (list): Lista de frequencias de notch (por exemplo,
    [50, 60] para remover interferencias de 60 Hz).
10 sampling_rate (float): Taxa de amostragem em Hz.
11 quality_factor (float): Fator de qualidade para controle da
    largura da banda do notch.
12
13 Returns:
14 array: Sinal filtrado com notch.
15 """
16 filtered_data = data.copy() # Cria uma copia do sinal para
    aplicar os filtros sequenciais
17
18 # Aplica um filtro Notch para cada frequencia em notch_freqs
19 for notch_freq in notch_freqs:
20     nyquist = 0.5 * sampling_rate
21     normal_cutoff = notch_freq / nyquist
22     # Utilizando a função butter para gerar filtro bandstop (Notch)
23     sos = butter(4, [normal_cutoff - 1/(2*nyquist), normal_cutoff +
        1/(2*nyquist)], btype='bandstop', output='sos')
24
25     # Filtragem do sinal
26     filtered_data = sosfilt(sos, filtered_data)
27
28     # Validação da atenuação de -60 dB na frequência de notch_freq
29     w, h = sosfreqz(sos, worN=2000)
30     freq = w * nyquist / np.pi
31     magnitude = 20 * np.log10(abs(h))
32
33     # Exibe o gráfico de resposta de frequência para cada filtro
34     plt.figure(figsize=(10, 6))
35     plt.plot(freq, magnitude)
36     plt.xlim(0, sampling_rate / 2)
37     plt.title(f'Resposta de Frequência do Filtro Notch (60 Hz)')
38     plt.xlabel('Frequência (Hz)')
39     plt.ylabel('Magnitude (dB)')
40     plt.grid(True)
41     plt.axvline(notch_freq, color='r', linestyle='--',
        label=f'{notch_freq} Hz')
42     plt.axhline(-60, color='g', linestyle='--', label='Atenuação -60
        dB')
43     plt.legend()
```

```
44 plt.show()
45
46 return filtered_data
47
48
49 def bandpass_filter(data, lowcut, highcut, sampling_rate,
50                     order=4):
51     nyquist = 0.5 * sampling_rate
52     low = lowcut / nyquist
53     high = highcut / nyquist
54     b, a = butter(order, [low, high], btype='band')
55     filtered_data = filtfilt(b, a, data)
56     return filtered_data
57
58 def smooth_signal(data, window_len=25):
59     window = np.ones(window_len) / window_len
60     smoothed_data = np.convolve(data, window, mode='same')
61     return smoothed_data
62
63 def compute_rms(signal, sampling_rate, window_ms=100):
64     """
65     Calcula o RMS usando janelas adaptativas com base na taxa de
66     amostragem.
67
68     Parameters:
69     signal (array): Sinal EMG a ser retificado.
70     sampling_rate (float): Taxa de amostragem em Hz.
71     window_ms (int): Tamanho da janela em milissegundos.
72
73     Returns:
74     array: Sinal RMS calculado em janelas móveis.
75     """
76     window_samples = int((window_ms / 1000) * sampling_rate) #
77         Converte ms para número de amostras
78
79     if window_samples < 1:
80         window_samples = 1 # Garante que haja pelo menos 1 amostra por
81             janela
82
83     rms_signal = np.sqrt(np.convolve(signal**2,
84                                     np.ones(window_samples) / window_samples, mode='valid'))
```

```
81 return rms_signal
```

A.2.2 Extração de Características e Análise de Padrões

```
1 # Discrete Wavelet Detailing Coefficients
2 def reconstruct_single_band(coeffs, wavelet, level_to_keep):
3     """
4     Reconstrói o sinal apenas para o coeficiente de detalhe
5     level_to_keep,
6     mantendo o coeficiente de aproximação, substituindo os demais
7     detalhes por zeros.
8
9     Args:
10    coeffs: lista original de coeficientes da DWT.
11    wavelet: nome da wavelet usada.
12    level_to_keep: índice do coeficiente de detalhe a manter (1 a N).
13
14    Retorna:
15    Sinal reconstruído para aquele nível.
16    """
17    recon_coeffs = []
18    # coeficiente de aproximação mantido
19    recon_coeffs.append(coeffs[0])
20
21    # Para os coeficientes de detalhe:
22    for i in range(1, len(coeffs)):
23        if i == level_to_keep:
24            # mantém o coeficiente desejado
25            recon_coeffs.append(coeffs[i])
26        else:
27            # substitui por zeros do mesmo tamanho do coeficiente original
28            recon_coeffs.append(np.zeros_like(coeffs[i]))
29
30    # Reconstrução do sinal
31    return pywt.waverec(recon_coeffs, wavelet)
32
33 # Wavelet a ser utilizada (Daubechies-4)
34 wavelet = 'db4'
35
36 # Aplicando Wavelet Transform Discreta (DWT) - Decomposição
37 Multinível:
```

```
35 max_levels = pywt.dwt_max_level(len(rectified_data),
    pywt.Wavelet(wavelet).dec_len)
36 levels = 4
37 levels = min(levels, max_levels)
38 coeffs = pywt.wavedec(rectified_data, wavelet, level=levels)
39
40 # Validando DWT
41 if coeffs is None or len(coeffs) == 0:
42     print("Erro ao aplicar a Discrete Wavelet Transform (DWT).
    Verifique os dados de entrada.")
43     return None
44 else:
45     print("DWT aplicada com sucesso: ", len(coeffs))
46
47 # Plot dos coeficientes
48 plt.figure(figsize=(15, 10))
49 for i in range(1, levels + 2):
50     plt.subplot(levels + 1, 1, i)
51     plt.plot(coeffs[i-1])
52     plt.title(f'Detail Coefficients (Level {i-1})')
53     plt.xlabel('Samples (Index)')
54     plt.ylabel('Amplitude ( $\mu$ V)')
55     plt.tight_layout()
56     plt.savefig(f'{output_dir}/dwt_emg.png')
57     plt.show()
58
59 # Energia por banda (opcional)
60 energies = [np.sum(c**2) for c in coeffs[1:]]
61 bands = [f"D{levels - i}" for i in range(len(coeffs)-1)]
62
63 plt.figure(figsize=(8, 5))
64 plt.bar(bands, energies)
65 plt.title("Energia por Banda da DWT")
66 plt.xlabel("Banda (Detalhe)")
67 plt.ylabel("Energia")
68 plt.tight_layout()
69 plt.savefig(f'{output_dir}/dwt_energy_emg.png')
70 plt.show()
71
72 # Frequências associadas aos níveis da DWT:
73 nyquist_freq = sampling_rate / 2
74
```

```
75 # Reconstrução usando a função segura:
76 bands_signals = {}
77 bands_signals['Type I'] = reconstruct_single_band(coeffs,
78     wavelet, 4) # detalhe nível 4 (baixas freq)
79 bands_signals['Type IIa'] = reconstruct_single_band(coeffs,
80     wavelet, 3) # detalhe nível 3 (médias freq)
81 bands_signals['Type IIx'] = reconstruct_single_band(coeffs,
82     wavelet, 2) # detalhe nível 2 (altas freq)
83
84 # Ajustar tamanho para coincidir com o original
85 for k in bands_signals:
86     bands_signals[k] = bands_signals[k][:len(rectified_data)]
87
88 # Etapa 5: Refinamento da Detecção de Picos
89 # Ajustando os parâmetros dinamicamente
90
91 altura_minima = np.mean(rectified_data) + np.std(rectified_data)
92     # Média + Desvio padrão como referência
93 proeminencia_minima = np.std(rectified_data) * 0.5 # Metade do
94     desvio padrão como referência
95
96 picos, propriedades = find_peaks(rectified_data,
97     height=altura_minima, prominence=proeminencia_minima,
98     distance=sampling_rate * 0.005)
99
100 # Exibir os picos detectados
101 plt.figure(figsize=(15, 5))
102 plt.plot(rectified_data, label="Sinal Retificado")
103 plt.plot(picos, rectified_data[picos], "rx", label="Picos
104     Detectados")
105 plt.legend()
106 plt.title("Detecção de Picos no Sinal EMG")
107 plt.savefig(f'{output_dir}/picos_detectados-percentile-proeminence.png')
108 plt.show()
109
110 # Filtros Adaptativos para Melhorar a Qualidade do Sinal
111 wavelet = 'db4'
112 coeffs = pywt.wavedec(rectified_data, wavelet, level=4)
113 threshold = np.median(np.abs(coeffs[-1])) / 0.6745 #
114     Thresholding baseado na estatística do ruído
115
116 # Aplicação de Thresholding nos coeficientes de detalhe
```

```
108 coeffs_thresh = [pywt.threshold(c, threshold, mode='soft') for c
    in coeffs]
109
110 # Reconstrução do sinal
111 filtered_signal = pywt.waverec(coeffs_thresh, wavelet)
112
113 plt.figure(figsize=(15, 5))
114 plt.plot(rectified_data, label="Sinal Original")
115 plt.plot(filtered_signal, label="Sinal Filtrado", alpha=0.7)
116 plt.legend()
117 plt.title("Filtro Wavelet para Suavização de Ruído")
118 plt.show()
119
120 # Estimativa da Taxa de Disparo das Unidades Motoras
121
122 isi = np.diff(picos) / sampling_rate # Intervalo entre picos em
    segundos
123 taxa_disparo = 1 / isi # Firing Rate em Hz
124
125 plt.figure(figsize=(20, 5))
126 plt.plot(taxa_disparo, marker='o', linestyle='-', label="Firing
    Rate")
127 plt.xlabel("Número do disparo")
128 plt.ylabel("Firing Rate (Hz)")
129 plt.title("Estimativa da Taxa de Disparo das Unidades Motoras")
130 plt.legend()
131 plt.savefig(f'{output_dir}/estimativa_de_disparos.png')
132 plt.show()
133
134 #
    -----
135 # Aplicando o K-Means para agrupamento dos picos (MUAPs) com
    base nas propriedades obtidas com DWT
136 #
    -----
137
138 # Escolha do sinal da DWT para análise
139 # Aqui você pode usar um sinal reconstruído a partir de um nível
    específico ou de uma banda de interesse
140 # Exemplo: banda já separada no passo de decomposição
141 signal_dwt = bands_signals['Type I'] # Pode trocar para 'Type
    IIa' ou 'Type IIx'
```

```
142
143 picos_dwt, propriedades_dwt = find_peaks(signal_dwt,
144     height=0.05, distance=20, width=5)
145 # Extração das propriedades dos picos
146 altura_dwt = propriedades_dwt["peak_heights"]
147 largura_dwt = propriedades_dwt["widths"]
148 distancia_dwt = np.diff(picos_dwt, prepend=picos_dwt[0])
149
150 #
151     -----
152 # Normalização dos dados
153 X_dwt = np.array([altura_dwt, largura_dwt, distancia_dwt]).T
154 scaler_dwt = StandardScaler()
155 X_norm_dwt = scaler_dwt.fit_transform(X_dwt)
156
157 # Ajusta automaticamente número de componentes do PCA
158 n_features = X_norm_dwt.shape[1]
159 n_components = min(3, n_features)
160 #
161     -----
162 # Aplicando PCA para redução de dimensionalidade
163 pca_dwt = PCA(n_components=n_components)
164 X_pca_dwt = pca_dwt.fit_transform(X_norm_dwt)
165
166 print("Aplicando o K-Means para agrupamento dos picos (MUAPs)
167     com base nas propriedades obtidas com DWT")
168 print("Variância Explicada (DWT):",
169     pca_dwt.explained_variance_ratio_)
170 print("Variância Explicada Acumulada (DWT):",
171     np.cumsum(pca_dwt.explained_variance_ratio_))
172
173 #
174     -----
175 # Aplicando o K-Means
176 kmeans_dwt = KMeans(n_clusters=3, random_state=42)
177 kmeans_dwt.fit(X_pca_dwt)
178
179 #
180     -----
181 # Avaliação dos clusters
```

```
176 print("Centroids (DWT):", kmeans_dwt.cluster_centers_)
177 print("Labels (DWT):", kmeans_dwt.labels_)
178 print("Inertia (DWT):", kmeans_dwt.inertia_)
179 print("Silhouette Score (DWT):", silhouette_score(X_norm_dwt,
    kmeans_dwt.labels_))
180 print("Calinski-Harabasz Score (DWT):",
    calinski_harabasz_score(X_norm_dwt, kmeans_dwt.labels_))
181
182 #
    -----
183 # Visualização dos agrupamentos - Matplotlib (3D)
184 fig = plt.figure(figsize=(15, 5))
185 ax = fig.add_subplot(111, projection='3d')
186 ax.scatter(X_norm_dwt[:, 0], X_norm_dwt[:, 1], X_norm_dwt[:, 2],
187 c=kmeans_dwt.labels_, cmap='viridis')
188 ax.set_xlabel("Peak Height (Normalized)")
189 ax.set_ylabel("Peak Width (Normalized)")
190 ax.set_zlabel("Peak Distance (Normalized)")
191 ax.set_title("Estimativa de Agrupamento de Picos (MUAPs) - DWT")
192 plt.savefig(f'{output_dir}/MUAP_clustering_estimative_dwt.png')
193 plt.show()
194
195 #
    -----
196 # Salvar resultados em CSV
197 df_dwt = pd.DataFrame({
198     'Peak Height': altura_dwt,
199     'Peak Width': largura_dwt,
200     'Peak Distance': distancia_dwt,
201     'Cluster': kmeans_dwt.labels_
202 })
203 df_dwt.to_csv(f'{output_dir}/peak_clusters_dwt.csv',
    index=False, encoding='utf-8')
204
205 #
    -----
206 # Criando DataFrame para visualização com Plotly
207 df_plot_dwt = pd.DataFrame({
208     'Peak Height': X_norm_dwt[:, 0],
209     'Peak Width': X_norm_dwt[:, 1],
210     'Peak Distance': X_norm_dwt[:, 2],
211     'Cluster': kmeans_dwt.labels_
```

```
212 })
213
214 fig = px.scatter_3d(
215 df_plot_dwt, x='Peak Height', y='Peak Width', z='Peak Distance',
216 color='Cluster', title="Estimativa de Agrupamento de Picos
      (MUAPs) - DWT",
217 color_continuous_scale='viridis'
218 )
219 fig.show()
220
221 #
      -----
222 # Plotando picos no sinal com cores por cluster
223 viridis = plt.cm.get_cmap('viridis',
      len(np.unique(kmeans_dwt.labels_)))
224
225 plt.figure(figsize=(15, 5))
226 plt.plot(signal_dwt, label="Sinal DWT", alpha=0.7)
227 for i in range(len(picos_dwt)):
228 plt.plot(
229 picos_dwt[i], signal_dwt[picos_dwt[i]], marker='o',
230 color=viridis(kmeans_dwt.labels_[i] /
      len(np.unique(kmeans_dwt.labels_))),
231 markersize=6
232 )
233 plt.legend(["Sinal DWT", "Picos Agrupados - DWT"])
234 plt.title("Clusters de Picos no Sinal EMG - DWT")
235 plt.savefig(f'{output_dir}/peaks_clustering_dwt.png')
236 plt.show()
237
238 #
      -----
239 # Cálculo de métricas adicionais
240 resultados_contribuicao_dwt = calcular_contribuicao_por_label(
241 kmeans_dwt.labels_, kmeans_dwt.cluster_centers_, X_norm_dwt,
      pca_dwt, scaler_dwt
242 )
243
244 print("Distribuição de labels e centróides (DWT):")
245 print(resultados_contribuicao_dwt["contribuicao_por_label"])
246 print("Centróides dos clusters (em dados originais - DWT):")
247 print(resultados_contribuicao_dwt["centroides"])
```

```
248
249 #
    -----
250 # Variância intra-cluster
251 variancias_dwt = calcular_variancia_intra_cluster(
252 kmeans_dwt.labels_, X_norm_dwt, kmeans_dwt.cluster_centers_
253 )
254 print("Dispersão (variância intra-cluster - DWT):")
255 print(variancias_dwt)
256
257 #
    -----
258 # Treinamento de Rede Neural (MLP)
259 modelo_dwt, resultados_nn_dwt = treinar_rede_neural(X_norm_dwt,
    kmeans_dwt.labels_)
260 print("Relatório de Classificação (DWT):")
261 print(resultados_nn_dwt["relatorio_classificacao"])
262 print("Matriz de Confusão (DWT):")
263 print(resultados_nn_dwt["matriz_confusao"])
264
265 #
    -----
266 # Visualização final dos clusters no PCA
267 visualizar_centroides_clusters(
268 X_pca_dwt, kmeans_dwt.labels_, kmeans_dwt.cluster_centers_,
269 np.unique(kmeans_dwt.labels_), output_dir=output_dir
270 )
271
272 # Dynamic Time Warping (DTW) para comparar as formas de onda dos
    picos (MUAPs) em cada cluster
273 # Função para calcular matriz de distâncias DTW entre as formas
    de onda dos MUAPs
274 def compute_dtw_matrix(waveforms):
275     n = len(waveforms)
276     dtw_matrix = np.zeros((n, n))
277
278     for i, wf in enumerate(waveforms):
279         print(f"Waveform {i}: Shape {np.array(wf).shape}")
280
281         for j in range(i):
282             for j in range(i + 1, n): # Calculamos apenas a metade superior
                (simetria)
```

```
283 # distance, _ = fastdtw(np.array(waveforms[i]).flatten(),
    np.array(waveforms[j]).flatten(), dist=euclidean)
284 distance, _ = fastdtw(np.squeeze(waveforms[i]),
    np.squeeze(waveforms[j]), dist=euclidean)
285 dtw_matrix[i, j] = distance
286 dtw_matrix[j, i] = distance # Matriz simétrica
287
288 return dtw_matrix
289
290 print("Pipeline de processamento do sinal EMG com Direct R
    concluído com sucesso.")
291
292 return rectified_data
```

A.2.3 Função principal de chamada para análise dos arquivos de sEMG

```
1 # Diretório raiz do dispositivo
2 root_directory = 'Análise dos sinais - Eletromiografia/EMG/'
3
4 # Sinais a processar - Alterar conforme a quantidade de sinais
  por arquivo
5 signals = ['Signal 1 ( $\mu$ V)', 'Signal 2 ( $\mu$ V)', 'Signal 3 ( $\mu$ V)',
    'Signal 4 ( $\mu$ V)']
6
7 # Função auxiliar para verificar se há csv_files em uma pasta
8 def find_csv_folders(base_folder):
9     csv_folders = []
10    for dirpath, dirnames, filenames in os.walk(base_folder):
11        if os.path.basename(dirpath) == 'csv_files':
12            csv_folders.append(dirpath)
13    return csv_folders
14
15 # Busca todas as pastas csv_files em todos os níveis dentro de
    root_directory
16 csv_directories = find_csv_folders(root_directory)
17
18 print(f"Encontradas {len(csv_directories)} pastas 'csv_files'")
19
20 # Processa cada pasta csv_files encontrada
21 for csv_directory in csv_directories:
22    print(f"\nProcessando pasta: {csv_directory}")
23
```

```
24 # Define pasta base (pasta mãe de csv_files)
25 parent_folder = os.path.dirname(csv_directory)
26
27 try:
28 dataframes_list, selected_files = load_csv_files(csv_directory)
29 except Exception as e:
30 print(f"Erro ao carregar CSVs em {csv_directory}: {e}")
31 continue
32
33 for i, dataframe in enumerate(dataframes_list):
34 selected_file = selected_files[i]
35 patient_file = os.path.basename(selected_file)
36
37 # Calcula taxa de amostragem
38 sampling_rate = 1 / np.mean(np.diff(dataframe['Time
    (Seconds)'].to_numpy()))
39 print(f"Sampling rate: {sampling_rate:.2f} Hz")
40
41 for signal in signals:
42 print(f"Processando arquivo: {patient_file} - {signal}")
43 emg_signal = dataframe[signal].to_numpy()
44
45 # Define diretório de saída relativo à pasta onde achou o
    csv_files
46 output_dir = os.path.join(
47 parent_folder,
48 'output_without_mass_processing',
49 patient_file,
50 signal
51 )
52
53 # Verifica se output já existe e não está vazio
54 if os.path.exists(output_dir) and os.listdir(output_dir):
55 print(f"Output já existe. Pulando: {output_dir}")
56 continue
57
58 os.makedirs(output_dir, exist_ok=True)
59
60 processed_signal = process_emg_pipeline(
61 emg_signal,
62 sampling_rate=sampling_rate,
63 output_dir=output_dir
```

```
64 )  
65  
66 print("\n TODAS AS ANÁLISES CONCLUÍDAS")
```