



University of Brasilia at Gama – FGA/UnB  
Biomedical Engineering Graduate Program

**ELECTROMYOGRAPHIC  
SIGNAL PROCESSING USING  
MACHINE LEARNING AND ENTROPY**

**LUIZ BARBOSA**

Advisor: Dr. ADSON FERREIRA DA ROCHA



UNIVERSITY OF BRASILIA AT GAMA



**ELECTROMYOGRAPHIC  
SIGNAL PROCESSING USING  
MACHINE LEARNING AND ENTROPY**

**LUIZ BARBOSA**

ADVISOR: ADSON FERREIRA DA ROCHA

MASTER DEGREE THESIS ON  
BIOMEDICAL ENGINEERING

PUBLICATION: 011A/2020

BRASILIA/DF, MARÇO DE 2020

**UNIVERSITY OF BRASILIA AT GAMA**

**GRADUATE PROGRAM**

**ELECTROMYOGRAPHIC  
SIGNAL PROCESSING USING  
MACHINE LEARNING AND ENTROPY**

**LUIZ BARBOSA**

MASTER THESIS SUBMITTED TO THE BIOMEDICAL ENGINEERING GRADUATE PROGRAM, AS A PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER IN BIOMEDICAL ENGINEERING

APPROVED BY:

---

Adson Ferreira da Rocha

(Advisor)

---

Dr. Cristiano Jacques Miosso

(Internal examiner)

---

Dr. Fabiano Araujo Soares

(External examiner)

**CATALOG CARD**

LUCAS BARBOSA, LUIZ JOSÉ

Electromyographic Signal Processing Using Machine Learning and Entropy

[Distrito Federal], 2020.

39p., 210 × 297 mm (FGA/UnB Gama, Mestrado em Engenharia Biomédica, 2020).

Dissertação de Mestrado em Engenharia Biomédica, Faculdade UnB Gama, Programa de Pós-Graduação em Engenharia Biomédica.

- |                        |                               |
|------------------------|-------------------------------|
| 1. Signal Processing   | 2. Biological Signal Analysis |
| 3. Information Entropy | 4. Machine Learning           |
| I. FGA UnB/UnB.        | II. Título (série)            |

**REFERENCE**

LUCAS BARBOSA, LUIZ JOSÉ (2020). Processamento de Sinais Eletromiográficos com o uso de Aprendizagem de Máquina e Entropia. Master thesis in Biomedical Engineering, Publication 011A/2020, Biomedical Engineering Graduate Program, University of Brasilia at Gama, Brasilia, DF, 39p.

**COPYRIGHT**

AUTOR: Luiz Barbosa

TÍTULO: Processamento de Sinais Eletromiográficos com o uso de Aprendizagem de Máquina e Entropia

GRAU: Mestre

ANO: 2020

É concedida à Universidade de Brasília permissão para reproduzir cópias desta dissertação de mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte desta dissertação de mestrado pode ser reproduzida sem a autorização por escrito do autor.

---

[contato@luizbarbosa.net](mailto:contato@luizbarbosa.net)

Brasília, DF – Brasil

*Suppose that we were asked to arrange the following in two categories — distance, mass, electric force, entropy, beauty, melody. I think there are the strongest grounds for placing entropy alongside beauty and melody, and not with the first three.*

Eddington A, The Nature of the Physical World, 1928

## **AGRADECIMENTOS**

Agradeço aos meus colegas de laboratório e trabalho pelas ajudas, conversas e insights. Agradeço também ao meu orientador, Adson, por sua ajuda, conselhos e guias. Principalmente agradeço ao meu co-orientador Alberto, que sem ele não teria conseguido concluir esse projeto.

## RESUMO ESTENDIDO

O sinal eletromiográfico é utilizado em diversas áreas da Medicina e da Biologia, em especial na área de reabilitação muscular. Uma aplicação importante é no controle de próteses robóticas. Atualmente, diversas próteses comerciais de mão utilizam uma malha de controle sequencial, o que torna o movimento da prótese pouco fluido e dependente de sensores externos para execução de movimentos. O presente plano de trabalho propõe o desenvolvimento de um estudo de métodos para o uso do sinal da eletromiografia de superfície (sEMG) para o controle em tempo real de uma prótese de mão. O objetivo é utilizar métodos de extração de características e classificação de padrões em sinais eletromiográficos, assim como o treinamento adaptativo para o reconhecimento de movimentos da mão com vários graus de liberdade, aumentando assim o conforto do usuário e dando naturalidade ao movimento. Por meio desses métodos, foi obtida uma melhoria significativa no processo de reconhecimento de padrões. As propostas e métodos de classificadores foram desenvolvidas e testadas com o uso das bases de dados disponíveis na plataforma *Open Source BioPatRec* [25], a linguagem utilizada para o algoritmo foi o python, com o auxílio das bibliotecas Scikit-learn [28], ScyPy [40] e Tensorflow [1]. Diversos indicadores estatísticos, são aplicados para avaliar o reconhecimento de padrões, *off-line* e *on-line*.

**Palavras-chave:** sinal eletromiográfico de superfície, prótese de mão, reconhecimento de padrões, redes neurais, entropia.

## 1 Introdução

Segundo o IBGE [17], no Brasil, 13,2 milhões de pessoas se declararam portadoras de algum tipo de deficiência motora, sendo que 470 mil foram vítimas de amputações. Uma amputação de mão é uma das lesões mais prejudiciais e pode afetar dramaticamente as capacidades de uma pessoa. Estima-se ainda, segundo o IBGE, que a incidência média anual de amputações seja de 13,9 por 100 mil habitantes. Além disso, segundo dados do Ministério da Saúde [24], o total de nascidos vivos no Brasil no ano de 2015 foi de 3.017.668 e, sabe-se que, de 1 a 2% deles sofrem de alguma anomalia congênita e destes, aproximadamente 10% possuem deformidades dos membros superiores [24, 20, 19].

Houve grandes avanços nas interfaces homem-máquina, na qual os sinais biomédicos, como os sinais mio-elétricos, desempenham um papel fundamental. O controle utilizando o sinal mio-elétrico é uma técnica avançada, subdividida em detecção, processamento, classificação e aplicação de sinais EMG para controlar robôs ou dispositivos de reabilitação humanas. Os sinais mio-elétricos são muito ricos em informações, a partir das quais a intenção de movimento do usuário em forma de contração muscular pode ser detectada, usando eletrodos de superfície. O sinal da eletromiografia de superfície é detectado de forma não invasiva a partir da superfície da pele e pode ser adaptado para força proporcional ou

controle de velocidade em um esquema de controle.

Os sinais de eletromiografia possuem uma propriedade não estacionária, que dificulta a aplicação de reconhecimento de padrões mio-elétricos para o controle de próteses. Na literatura, o reconhecimento de padrões EMG é separado em duas etapas separadas, uma de treinamento e outra de teste. Porém nessa análise não são consideradas as mudanças entre o treinamento e os dados de teste induzidos por mudança de eletrodo, fadiga muscular, mudanças de impedância e fatores psicológicos, que muitas vezes resulta em queda do desempenho [33]. Para solucionar esse problema, vários estudos sobre treinamento adaptativo vem sendo feitos [2, 10, 18, 33, 39, 43], a fim de manter a taxa de certeza da análise do sinal sEMG.

Como os dados do sinal EMG são adquiridos em um curto período de duração, os parâmetros coletados contêm informações limitadas, que não representam todo o período de utilização da prótese. Ampliar o tempo de coleta de dados se torna impraticável, pois adicionaria uma carga muito grande ao usuário. Os sistemas de próteses que utilizam controle de reconhecimento de padrões mio-elétricos não são comercializados por possuírem desempenho insatisfatório [29], justamente por causa das variações dos dados de teste em relação aos do treinamento. Portanto, torna-se necessário desenvolver um sistema de aprendizado que consiga se adaptar e contabilize as mudanças do sinal do EMG. Mais especificamente, para que a prótese se torne o mais natural possível, é necessário que se faça uma reconfiguração gradual do classificador de forma online, levando em consideração as mudanças em tempo real do sinal EMG.

As principais empresas que comercializam próteses, no mundo, são a Touch Bionics, Otto Bock, Steeper e Vicent GmbH. Todas as próteses fabricadas por elas possuem um controle sequencial, sendo que algumas delas possuem sensores para auxiliar a movimentação da prótese ou movimentos já predefinidos [14]. Isso demonstra que o caminho para o controle natural ainda é longo, o que limita a usabilidade da prótese a apenas um membro de apoio. O controle simultâneo visa dar naturalidade ao movimento da prótese, diminuindo o seu desconforto que pode levar ao abandono da mão biônica [6, 23].

## **2 Importância da Informação no Sinal EMG**

A busca e o processamento de informações são um processo cognitivo complexo que exige a identificação extração e organização da informação relevante. Porém, com a evolução da computação em nuvem, técnicas de classificação baseadas em Deep Learning cresceram muito, uma vez que elas veem performando melhor do que técnicas convencionais de Machine Learning. Esse aumento no uso de técnicas de Deep Learning faz com que muitas vezes não seja necessário entender o problema ou delimitar suas bordas.

Essa abordagem gera um problema, como não houve um aprendizado do processo, o resultado fica dependendo da rede neural e ao que ela foi treinada a fazer. Outro problema



é que o conhecimento ganho pode ser usado para outras aplicações.

Em 1948, Claude Shannon publicou um artigo chamado Teoria Matemática da Comunicação [7]. Neste artigo, Shannon descreve como a informação pode ser comunicada por diferentes elementos de um sistema. Shannon mostra como sinais e ruídos se relacionam e como a habilidade de separá-los para extrair informação dos dados é crucial para o processo de comunicação.

A informação é normalmente medida em bits e, um bit de informação permite a escolha de duas alternativas igualmente prováveis. Como exemplo, temos o lançamento de uma moeda, ao se revelar o resultado, é gerado um bit de informação.

A teoria da informação tem os seus próprios conjuntos de termos. Uma mensagem é o uma sequência ordenada de símbolos. Uma mensagem composta por símbolos  $\mathbf{s} = (s_1, \dots, s_n)$  é codificada por uma função  $\mathbf{c} = f(\mathbf{s})$  em uma sequência de códigos  $\mathbf{c} = (c_1, \dots, c_m)$ , aonde o número de símbolos ou códigos não são necessariamente iguais. Esses códigos são transmitidos por um canal de comunicação para produzir saídas  $\mathbf{y} = (y_1, \dots, y_k)$  que são decodificadas para reconstruir a mensagem  $\mathbf{s}$ .

### 3 Introdução à Entropia

Ainda segundo Shannon [37], Para que as definições matemáticas da informação sejam uteis, elas precisam obedecer as seguintes propriedades:

1. **Continuidade:** O total de informação associado à um resultado aumenta ou decrementa continuamente, de acordo com a mudança da probabilidade do resultado.
2. **Simetria:** O total de informação associado com uma sequência de resultados não depende da ordem em que os resultados acontecem.
3. **Valor máximo:** O total de informação com um conjunto de resultados não pode aumentar se esses resultados são igualmente prováveis.
4. **Adição:** A informação associada com um conjunto de resultados é obtida pela adição das informações dos resultados individuais.

#### Surpresa e Entropia

Quanto mais improvável é um resultado, mais surpresa é tida com a sua observação. Então o total de surpresa associado a um resultado aumenta, caso a probabilidade do resultado diminua. Então, para satisfazer a condição de aditividade, Shannon utilizou o logaritmo de  $1/p(x)$ . Isto é conhecido como informação de Shannon de  $x$ , ou seja a informação de Shannon é uma medida da surpresa. A surpresa média de uma variável  $X$ , com uma distribuição  $p(X)$  é chamada entropia de  $p(X)$  e, é representada por  $H(X)$ .

$$H(X) \approx \frac{1}{n} \sum_{i=1}^n \log\left(\frac{1}{p(x_i)}\right) \quad (1)$$

Basicamente, entropia é a medida de incerteza. Quando a incerteza é reduzida, informação é ganha. Portanto, se a incerteza de uma variável  $X$  é resumida por sua entropia  $H(X)$ , se o valor de  $X$  for revelado, o total de informação ganha é, na média, exatamente igual a entropia.

Caso hajam valores consecutivos e relacionados, então eles não fornecem informação independente. Nesse caso a sequência possui menos entropia do que o somatório das entropias individuais calculados com a Equation 1.1. Assim, o teorema de Shannon não se aplica apenas a sequências de elementos independentes, mas também para sequências estruturadas que possuem dependências entre os seus elementos.

## 4 Entropia e o sinal sEMG

O sinal sEMG, assim como todos os sinais naturais, é um sinal diluído, isto é, um sinal que possui muitos dados e pouca informação. Ademais ele é um sinal em que cada sequência de amostragem possui um alto grau dependência entre seus valores.

### Processamento em tempo real

Como uma boa experiência em prótese precisa de respostas rápidas, o processamento em tempo real é essencial. Radhika Menon, et al [31] mostra o impacto do tamanho da janela temporal no erro da classificação do sinal sEMG. Segundo esse estudo com janelas muito pequenas, em média menores do que 200 ms o erro da classificação aumenta, uma vez que o total de informação na janela diminui. Isso dificulta muito o processamento em tempo real que precisa de janelas, em 2011 Peerdeman [29] constatou que a janela de processamento precisa ser menor do que 300 ms ou o atraso se torna inaceitável para o usuário. Instintivamente nota-se a necessidade da diminuição da janela de tempo de processamento para o aumento do conforto do usuário.

Uma das dificuldades na classificação do sinal geradas pelo tamanho reduzido da janela temporal foi identificar exatamente onde o movimento começou. Isso ocorre devido às características estocásticas do sinal, quando as unidades motoras começam a ser recrutadas, o sinal de repouso e movimento é confundido.

Para diferenciar esses dois estados, foi utilizado um autoencoder variacional. Devido a suas características intrínsecas, ele é especializado em separar classes (movimento e repouso) em sua camada latente. Após essa separação, uma percepção simples, foi o suficiente para classificar esses estados.

### Extraindo informação do sinal

Para extrairmos informação da eletromiografia de superfície, foi utilizada uma técnica conhecida como extração de características. O sinal foi dividido em janelas de 50 ms, aonde o sinal sEMG é analisado e as características baseadas em frequência são extraídas, seguindo os principais recursos selecionados e finalmente, a rede neural classifica o movimento.

As características extraídas são:

1. SpectralMoment;
2. Sample Entropy;
3. KhushabaSet;
4. Wave Length (frequency);
5. Mean Frequency;
6. Median Frequency.

## 5 Metodologia

A utilização de técnicas de reconhecimento de padrões é de grande importância para o controle de próteses mio-elétricas, trazendo uma melhora nos graus de liberdade e movimentação das mesmas além da capacidade de seu controle sequencial [13]. Tal controle consiste tipicamente na extração de características do sinal e na classificação destas características de dados segmentados no processamento de sinal para comando de um atuador. A qualidade da movimentação dessas próteses é proporcional aos processos de extração das características do sinal mio-elétrico e da classificação dos padrões.

Conforme já mencionado, a maioria dos sistemas de controle empregados em mãos protéticas é o controle sequencial, mas, recentemente, muitas pesquisas estão sendo conduzidas para empregar o controle simultâneo [5, 11, 31, 44]. Uma das principais vantagens do controle simultâneo da prótese é aproximar a movimentação da prótese ao movimento fisiológico natural, porém isso leva a um aumento das características necessárias para se classificar o sinal, além da interferência do sinal de músculos adjacentes nos eletrodos de captura do sEMG.

### **Clusterização, máquina de estados e a classificação do movimento**

O idioma inglês possui uma dívida a Shakespeare. Ele inventou mais de 1700 de palavras, transformando substantivos em verbos, ou verbos em adjetivos, ou conectando palavras nunca antes usadas em conjunto, ou ainda adicionando prefixos e sufixos e criando palavras totalmente originais. Assim como o inglês consiste de sequências de letras não independentes que podem ser eficientemente codificadas como blocos independentes de sub-sequências,

assim também pode ser feito como a maioria dos sinais naturais, como música, imagem, DNA, ou o EMG.

Para a classificação do sinal EMG foi realizado um janelamento no sinal, cada janela com 10ms. Após este janelamento foi realizado as seguintes etapas:

1. Seleção de características: nesta etapa diversas características, tanto no domínio da frequência quanto no domínio do tempo serão extraídas;
2. Redução de dimensionalidade: nesta etapa diversos algoritmos de redução de dimensionalidade foram avaliados para a criação de clusters de sinais de movimentos parecidos. Foram avaliados: NCA (Neighborhood Component Analysis), PCA (Principal Component Analysis), LDA (Linear Discriminant Analysis), Variational Autoencoders;
3. Clusterização do sinal: após avaliação de diversos algoritmos de clusterização, os movimentos foram agrupados em grupos de acordo com a similaridade das características extraídas;
4. Comparação com os possíveis movimentos: após o cluster ser criado, os movimentos são comparados com os possíveis movimentos para determinada posição. Esses movimentos foram extraídos através de um processo chamado de Árvore de Decisão ou Máquina de Estados [23];
5. Classificação do sinal: com a redução das possíveis classes de movimento em intercessão com os clusters criados, diversos classificadores são criados com subconjuntos das classes totais de movimentos.

Cada janela de tempo retorna uma resposta para os possíveis movimentos, o cluster o qual o movimento pertence e a possível classificação. Essas respostas podem ser consideradas letras em uma sentença. Cada janela retorna uma sequência dependente da sequência anterior e aonde cada uma das janelas diminui a entropia do sinal como um todo.

## 6 Resultados e Discussões

Desde que as próteses mio-elétricas foram desenvolvidas, o número de usuários que a rejeitam permaneceu constante [6, ?], O que mostra que não houve avanço significativo na área. A redução da janela de tempo para a classificação do sinal EMG permitirá que o controle da prótese seja realizado de forma mais fluida pelo usuário, aumentando seu conforto ao utilizá-la.

Outro fator muito importante gerado pela diminuição da A janela de tempo diminui a complexidade da avaliação, o que resulta em economia de energia durante o processo de

classificação, o que, por sua vez, aumentaria o tempo de uso da prótese, reduzindo o tempo de recargas que o usuário precisaria fazer.

Além disso, o processamento desenvolvido neste estudo pode ser usado para a classificação de outros sinais naturais que, à medida que são diluídos (muitos dados para pouca informação), são difíceis de classificar.

## 7 Conclusão

Ao fornecer informações a priori para classificação de sinais interativamente, o número de classes possíveis para classificação de sinais diminui bastante. Criar etapas de validação menos complexas também aumentou a precisão, permitindo a redução do tamanho da janela.

As técnicas apresentadas aqui apenas arranham a superfície das aplicações em que a entropia de formações pode e deve ser usada. A ideia principal é que uma sequência de redes simples cuja descrição precise de um pequeno número de bits tenha maior probabilidade de generalizar corretamente do que uma rede mais complexa, porque presumivelmente extraiu a essência dos dados e removeu a redundância. Portanto, fornecer ferramentas que simplifiquem ou forneçam informações de dados é muito importante.

Infelizmente, a coleta de dados do sinal EMG pode ser estressante para o paciente. Portanto, é muito difícil obter grandes bancos de dados de sinais biológicos. Por esse motivo, este estudo foi realizado usando apenas um banco de dados, fornecido junto com a plataforma BioPatRec. Além disso, o banco de dados utilizado possui poucas repetições de movimento, o que dificulta o treinamento com algoritmos de aprendizado de máquina.

No total, essa base possui apenas 17 pacientes e 3 repetições de cada movimento, o que impossibilitou o uso de algumas técnicas de treinamento. Com um banco de dados suficientemente grande e, uma vez criado um classificador redundante, como no caso do estudo, os diversos As etapas da classificação podem ser usadas como um código e, com uma rede neural maior, e os valores das janelas de classificação podem ser previstos usando técnicas como LSTM. A relação de dependências entre as janelas de tempo também pode

ser estudada para se obter um classificador melhor projetado.

## Lista de Referências

- [1] Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, Manjunath Kudlur, Josh Levenberg, Rajat Monga, Sherry Moore, Derek G Murray, Benoit Steiner, Paul Tucker, Vijay Vasudevan, Pete Warden, Martin Wicke, Yuan Yu, Xiaoqiang Zheng, Google Brain, Implementation Osdi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, Manjunath Kudlur, Josh Levenberg, Rajat Monga, Sherry Moore, Derek G Murray, Benoit Steiner, Paul Tucker, Vijay Vasudevan, Pete Warden, Martin Wicke, Yuan Yu, e Xiaoqiang Zheng. TensorFlow : A System for Large-Scale Machine Learning This paper is included in the Proceedings of the TensorFlow : A system for large-scale machine learning. 2016.
- [2] Sebastian Amsuss, Peter M. Goebel, Ning Jiang, Bernhard Graimann, Liliana Paredes, e Dario Farina. Self-correcting pattern recognition system of surface EMG signals for upper limb prosthesis control. *IEEE Transactions on Biomedical Engineering*, 61(4):1167–1176, 2014.
- [3] Daniel A. Bennett e Michael Goldfarb. IMU-Based Wrist Rotation Control of a Transradial Myoelectric Prosthesis. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2017.
- [4] Stephen F. Burrough Brook e Judith A. Patterns of Acceptance and Rejection of Upper Limb Prostheses, 1985.
- [5] ”” C.E. Shannon , vol. 27, pp. , July, October,. A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(April 1928):379–423,623–656, 1948.
- [6] C.P. Diehl e G. Cauwenberghs. Svm incremental learning, adaptation and optimization. *Proceedings of the International Joint Conference on Neural Networks, 2003.*, 4(x):2685–2690.
- [7] Strahinja Dosen, Marko Markovic, Kelef Somer, Bernhard Graimann, e Dario Farina. EMG Biofeedback for online predictive control of grasping force in a myoelectric prosthesis. *Journal of NeuroEngineering and Rehabilitation*, 2015.

- [8] Kevin B. Englehart e Bernard Hudgins. A robust, real-time control scheme for multi-function myoelectric control. *IEEE Trans. Biomed. Eng.*, 50(7):848–854, 2003.
- [9] Purushothaman Geethanjali. Myoelectric control of prosthetic hands: State-of-the-art review, 2016.
- [10] IBGE. No Title, 2010.
- [11] Sarthak Jain, Girish Singhal, Ryan J. Smith, Rahul Kaliki, e Nitish Thakor. Improving long term myoelectric decoding, using an adaptive classifier with label correction. *Proceedings of the IEEE RAS and EMBS International Conference on Biomedical Robotics and Biomechatronics*, páginas 532–537, 2012.
- [12] McFarland. Unilateral upper-limb loss: Satisfaction and prosthetic-device use in veterans and servicemembers from Vietnam and OIF/OEF conflicts, 2010.
- [13] Ministério da Saúde. No Title, 2015.
- [14] Max Ortiz-Catalan, Rickard Brånemark, e Bo Håkansson. BioPatRec: A modular research platform for the control of artificial limbs based on pattern recognition algorithms. *Source Code for Biology and Medicine*, 2013.
- [15] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, e E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
- [16] Bart Peerdeman, Daphne Boere, Heidi Witteveen, Rianne Huis in ‘tVeld, Hermie Hermens, Stefano Stramigioli, Hans Rietman, Peter Veltink, e Sarthak Misra. Myoelectric forearm prostheses: State of the art from a user-centered perspective. *The Journal of Rehabilitation Research and Development*, 48(6):719, 2011.
- [17] IEEE Radhika Menon, Student Member, IEEE, Gaetano Di Caterina, Heba Lakany, Member, IEEE, Lykourgos Petropoulakis, Bernard A. Conway, John J. Soraghan, Senior Member. Study on interaction between temporal and spatial information in classification of EMG signals for myoelectric prostheses.
- [18] Aidan D. Roche, Hubertus Rehbaum, Dario Farina, e Oskar C. Aszmann. Prosthetic Myoelectric Control Strategies: A Clinical Perspective. *Current Surgery Reports*, 2(3):44, 2014.
- [19] 1. Kozin SH. Congenital disorders: classification and diagnosis. *Lippincott Williams & Williams*, páginas 1405–23, 2004.
- [20] 3. Kozin SH. Upper-extremity congenital anomalies. *J Bone Joint Surg Am*, páginas 1564–76, 2003.

- [21] Claude Elwood Shannon. *The Mathematical Theory of Communication by CE Shannon and W. Weaver*. 1949.
- [22] Tatiana Tommasi, Francesco Orabona, Claudio Castellini, e Barbara Caputo. Improving control of dexterous hand prostheses using adaptive learning. *IEEE Transactions on Robotics*, 29(1):207–219, 2013.
- [23] John J. Villarejo, Regina M. Costa, Teodiano Bastos, e Anselmo Frizera. Identification of low level sEMG signals for individual finger prosthesis. In *ISSNIP Biosignals and Biorobotics Conference, BRC*, 2014.
- [24] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, CJ Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, e SciPy 1.0 Contributors. SciPy 1.0—Fundamental Algorithms for Scientific Computing in Python. *arXiv e-prints*, page arXiv:1907.10121, Jul 2019.
- [25] Wu Yi Yang, Sheng Xing Liu, Tai Song Jin, e Xiao Mei Xu. An optimization criterion for generalized marginal Fisher analysis on undersampled problems. *International Journal of Automation and Computing*, 8(2):193–200, 2011.
- [26] Aaron J. Young, Lauren H. Smith, Elliott J. Rouse, e Levi J. Hargrove. Classification of simultaneous movements using surface EMG pattern recognition. *IEEE Transactions on Biomedical Engineering*, 2013.



## ABSTRACT

The electromyographic signal is used in several areas of medicine and biology, especially in the muscle rehabilitation area. Another important application is in the control of robotic prostheses. Currently, several commercial hand prostheses use a sequential control loop, which makes the movement of the prosthesis less fluid and dependent on external sensors to perform movements. The present work plan proposes the development of a study of methods for the use of surface electromyography signal (sEMG) for real time control of a hand prosthesis. The objective is to use methods of feature extraction and pattern classification in electromyographic signals, as well as adaptive training for the recognition of hand movements with varying degrees of freedom, thereby increasing user comfort and giving the movement a natural feel. Through these methods, a significant improvement in the pattern recognition process was obtained. The proposals and methods of classifiers were developed and tested using the databases available on the Open Source BioPatRec [25] platform, the language used for the algorithms was python, with help from the Scikit-learn [28], ScyPy [40] and Tensorflow [1] libraries. Several statistical indicators are applied to assess pattern recognition, offline and online.

**Keywords:** surface electromyographic signal, hand prosthesis, pattern recognition, neural networks, entropy.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Importance of Information in Electromyography (EMG) Signal . . . . .	2
1.2	Introduction to Entropy . . . . .	3
1.2.1	Surprise and Entropy . . . . .	4
1.3	Entropy and signal Superfice Electromyography (sEMG) . . . . .	4
1.3.1	Real-time processing . . . . .	5
1.3.2	Extracting signal information . . . . .	6
1.4	Objective . . . . .	6
1.4.1	Main Objective . . . . .	6
1.4.2	Specific Objective . . . . .	6
<b>2</b>	<b>State-of-the-Art Overview</b>	<b>8</b>
<b>3</b>	<b>Feature Selection</b>	<b>10</b>
3.1	Introduction . . . . .	10
3.2	Methodology . . . . .	11
3.2.1	BioPatRec Platform . . . . .	11
3.2.2	Features Extraction . . . . .	11
3.2.3	Features Selection . . . . .	12
3.2.4	Neural Network Classifiers . . . . .	13
3.2.5	sEMG Database . . . . .	13
3.2.6	Statistical Evaluation . . . . .	13
3.3	Results and discussion . . . . .	15

3.4	Conclusion . . . . .	16
<b>4</b>	<b>Autoencoder and Anomaly Detection</b>	<b>18</b>
4.1	Introduction . . . . .	18
4.1.1	Model specification . . . . .	19
4.1.2	Anomaly Detection . . . . .	21
4.2	Methodology . . . . .	21
4.2.1	Database . . . . .	21
4.2.2	VAE Algorithm . . . . .	21
4.3	Results . . . . .	22
4.4	Conclusion . . . . .	23
<b>5</b>	<b>Entropy</b>	<b>24</b>
5.1	Introduction . . . . .	24
5.2	Methodology . . . . .	25
5.2.1	The Pipeline . . . . .	25
5.2.2	Database . . . . .	26
5.2.3	Movement Classes . . . . .	27
5.2.4	Feature extraction . . . . .	28
5.2.5	Signal Information . . . . .	28
5.2.6	Classifier Algorithm . . . . .	30
5.3	Results . . . . .	30
5.4	Conclusion . . . . .	32
<b>6</b>	<b>Conclusion</b>	<b>33</b>
	<b>References</b>	<b>34</b>

# List of Tables

2.1	State of the art table of sEMG signal analysis. . . . .	8
3.1	Quantitative indicators obtained with the comparison between old features with feature reduction algorithms. . . . .	15
3.2	Quantitative indicators obtained with the comparison between the new features added to the old features sets with feature reduction algorithms. . . . .	15
5.1	Accuracy and Standard Deviation Comparison . . . . .	32

# List of Figures

1.1	The i-limb line of hands are developed by Touch Bionics, its proposal of easy adaptability and versatility offer great comfort for users. Offering a range of commands that can be changed in a companion application. . .	2
1.2	Discrete noiseless channel. The data from a source is encoded and transmitted through a communication channel. At the other side, a receiver decodes the codeword and recovers the original message. . . . .	3
1.3	The sEMG signal, the vertical axis represents the amplitude and the horizontal axis represents the samples. . . . .	5
2.1	The normal fluxogram used for sEMG prothesis. . . . .	8
2.2	The proposed fluxogram used for in this study for sEMG classification. .	9
3.1	Distribution of the experiment with the old features and the added features.	16
4.1	Flowchart of a variational autoencoder . . . . .	20
4.2	The latent space representation created by the Variational Autoencoder .	22
4.3	An expanded latent space, to represent the difference between the rest and the rest and the movement state. In this image the latent space was dimensioned to be 64 x 64 to facilitate the difference visualization. . . . .	23
5.1	Pipeline representation: the algorithm begins at a starting position; the sEMG signal is measured and pre-processed to help to determine the movement intention; with these two information pieces, the possible next movements are determined as well as the cluster to which the class belongs; finally, the algorithm yields the movement intention. . . . .	26
5.2	Test of different methods of dimensionality reduction for the use in the classificatory algorithm. The methods were tested with a K Nearest Neighbors (KNN) with $k = 3$ . . . . .	29

5.3	Output of the values of the Multi-Layer Perceptron (MLP) neurons of the last layer. The blue color mean the right classification and the red the wrong classification. The left figure is the output of the MLP and the right figure is the result with the possible classes only. . . . .	30
5.4	<b>sMEG! (sMEG!)</b> Clustering the classes by the similarity in the features extracted. After the extraction of twenty-five dimensions, the two more representative dimensions among them were used to generate the plot. A KNN was used to validate the cluster. . . . .	31

## LIST OF NOMENCLATURES AND ABBREVIATIONS

<b>3xCEPS</b>	First Three Cepstral Coefficients
<b>AR</b>	Autoregressive Model
<b>ASM</b>	Absolute value of the Summation of the expth root
<b>ASS</b>	Absolute value of the Summation of the Square root
<b>DWT</b>	Discrete Wavelet Transform
<b>IBGE</b>	Instituto Brasileiro de Geografia e Estatística
<b>EMG</b>	Electromyography
<b>ER</b>	Entropy Representation
<b>GMM</b>	Gaussian Mixture Model
<b>HCA</b>	Hierarchical Agglomerative Clustering
<b>KNN</b>	K Nearest Neighbors
<b>LDA</b>	Linear Discriminant Analysis
<b>LSTM</b>	Long Short Time Memory
<b>MAV</b>	Mean Absolute Value
<b>MICI</b>	Maximal Information Compression Index
<b>MLP</b>	Multi-Layer Perceptron
<b>MSR</b>	Absolute value of the Summation of the The Mean
<b>NCA</b>	Neighborhood Components Analysis
<b>PCA</b>	Principal Component Analysis
<b>PDF</b>	Probability Density Function
<b>RMS</b>	Root Mean Square
<b>sEMG</b>	Surface Electromyography
<b>SM</b>	Spectral Moment
<b>SE</b>	Sample Entropy

**SSC** Signal Slope Change

**UFS** Unsupervised Feature Selection

**VAE** Variational Autoencoder

**WFL** Waveform Length

**WL** Wave Length

**WNN** Wavelet Neuro Network

**ZC** Zero Crossing



# 1 INTRODUCTION

According to Instituto Brasileiro de Geografia e Estatística (IBGE) [17], in Brazil, 13.2 million people declared themselves to have some type of motor deficiency, and 470,000 were victims of amputations. A hand amputation is one of the most harmful injuries and can dramatically affect a person's abilities. It is also estimated, according to IBGE, that the average annual incidence of amputations is 13.9 per 100,000 inhabitants. In addition, according to data from the Ministry of Health [24], the total number of live births in Brazil in 2015 was 3,017,668 and, it is known that from 1 to 2

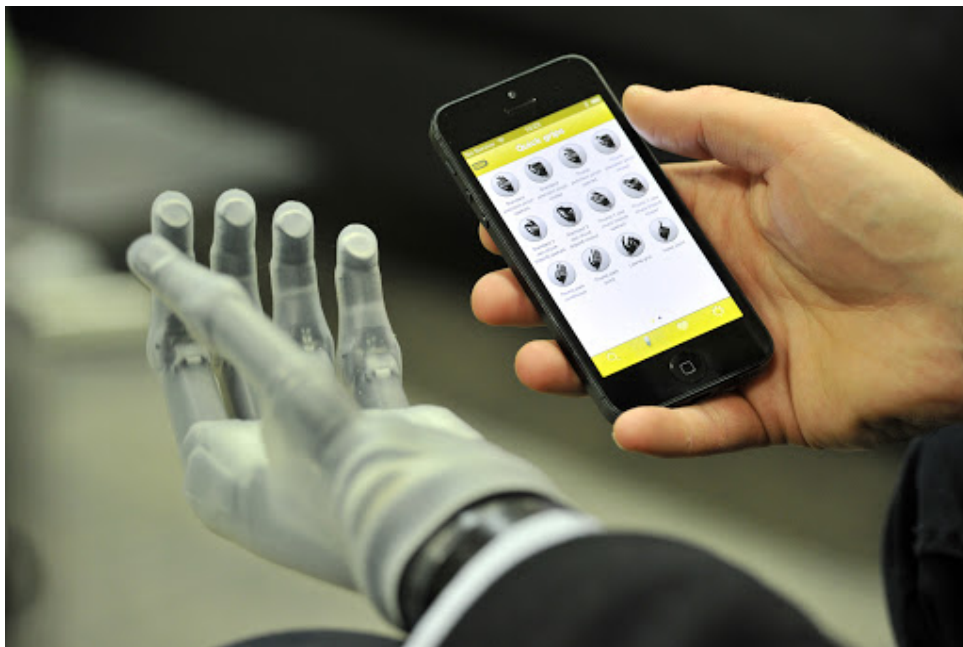
There have been great advances in man-machine interfaces, in which biomedical signals, such as myo-electric signals, play a key role. Control using the mio-electric signal is an advanced technique, subdivided into detection, processing, classification and application of EMG signals to control human robots or rehabilitation devices. The mio-electric signals are very rich in information, from which the user's intention to move in the form of muscle contraction can be detected using surface electrodes. The surface electromyography signal is detected noninvasively from the skin surface and can be adapted for proportional force or speed control in a control scheme.

The signs of electromyography have a non-stationary property, which makes it difficult to apply the recognition of mio-electrical patterns for the control of prostheses. In the literature, emg pattern recognition is separated into two separate steps, one of training and one test. However, the changes between training and electrode change-induced test data, muscle fatigue, impedance changes and psychological factors are not considered, which often results in a drop in performance [33]. To solve this problem, several studies on adaptive training have been done [2, 10, 18, 33, 39, 43], in order to maintain the certainty rate of sEMG signal analysis.

Because EMG signal data are acquired in a short period of duration, the parameters collected contain limited information, which do not represent the entire period of use of the prosthesis. Extending data collection time becomes impractical because it would add a very large load to the user. Prosthesis systems that use mio-electric pattern recognition control are not marketed because they have unsatisfactory performance [29], precisely because of variations in test data in compared to those of training. Therefore, it is necessary

to develop a learning system that can adapt and account for emg signal changes. More specifically, in order for the prosthesis to become as natural as possible, it is necessary to make a gradual reconfiguration of the classifier online, taking into account the real-time changes of the EMG signal.

The main companies that sell prostheses, in the world, are Touch Bionics, Otto Bock, Steeper and Vicent GmbH. All prostheses manufactured by them have a sequential control, some of which have sensors to help drive the prosthesis or movements already predefined [14]. This demonstrates that the path to natural control is still long, which limits the usability of the prosthesis to only one support member. Simultaneous control aims to give naturalness to the movement of the prosthesis, reducing its discomfort that can lead to the abandonment of the bionic hand [6, 23].



**Figure 1.1.** The i-limb line of hands are developed by Touch Bionics, its proposal of easy adaptability and versatility offer great comfort for users. Offering a range of commands that can be changed in a companion application.

## 1.1 IMPORTANCE OF INFORMATION IN EMG SIGNAL

The search and processing of information is a complex cognitive process that requires the identification of extraction and organization of relevant information. However, with the evolution of cloud computing, deep learning-based classification techniques have grown a lot, as they see performing better than conventional Machine Learning techniques. This increase in the use of Deep Learning techniques often makes it often not necessary to understand the problem or delimit its edges.

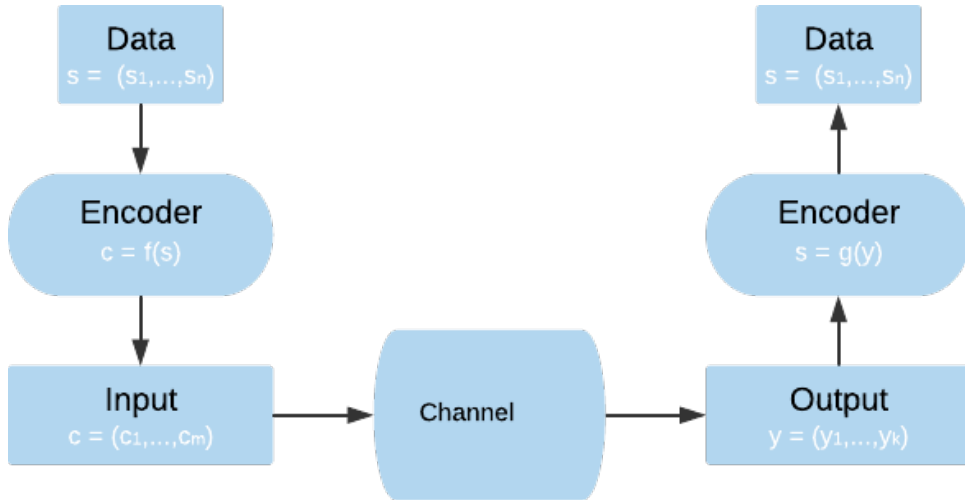
This approach creates a problem, as there was no learning from the process. The

knowledge of how to solve the problem is still lacking and the result is dependent of the neural network, that is trained to do only one task. This dependency creates the need for networks as complex as the first to solve similar problems.

In 1948, Claude Shannon published an article called Mathematical Theory of Communication [7]. In this article, Shannon describes how information can be communicated by different elements of a system. Shannon shows how signs and noises relate and how the ability to separate it to extract information from the data is crucial to the communication process.

Information is usually measured in bits and, a bit of information allows the choice of two equally likely alternatives. As an example, we have the posting of a currency, when revealing the result, a bit of information is generated.

The information theory has its own sets of terms. A message is an orderly sequence of symbols. A message composed of symbols  $\mathbf{s} = (s_1, \dots, s_n)$  is encoded by a function  $\mathbf{c} = f(\mathbf{s})$  in a sequence of codes  $\mathbf{c} = (c_1, \dots, c_m)$ , where the number of symbols or codes are not necessarily the same. These codes are transmitted over a communication channel to produce outputs  $\mathbf{y} = (y_1, \dots, y_k)$  that are decoded to rebuild the message  $\mathbf{s}$ .



**Figure 1.2.** Discrete noiseless channel. The data from a source is encoded and transmitted through a communication channel. At the other side, a receiver decodes the codeword and recovers the original message.

## 1.2 INTRODUCTION TO ENTROPY

Also according to Shannon [37], For mathematical definitions of information to be useful, they must obey the following properties:

1. **Continuity:** The total information associated with a result increases or decrements continuously, according to the change in the probability of the result.

2. **Symmetry:** The total information associated with a sequence of results does not depend on the order in which the results happen.
3. **Maximum value:** The total information with a result set cannot increase if these results are equally likely.
4. **Addition:** The information associated with a result set is obtained by adding the information of the individual results.

### 1.2.1 Surprise and Entropy

The more unlikely a result, the more surprise it is taken with your observation. Then the total surprise associated with a result increases if the probability of the result decreases. Then, to satisfy the condition of additivity, Shannon used the logarithm of  $1/p(x)$ . This is known as Shannon information of  $x$ , i.e. Shannon information is a measure of surprise. The average surprise of a variable  $X$ , with a distribution  $p(X)$  is called entropy of  $p(X)$  and is represented by  $H(X)$ .

$$H(X) \approx \frac{1}{n} \sum_{i=1}^n \log\left(\frac{1}{p(x_i)}\right) \quad (1.1)$$

Basically, entropy is the measure of uncertainty. When uncertainty is reduced, information is gained. Therefore, if the uncertainty of a variable  $X$  is summarized by its entropy  $H(X)$ , if the value of  $X$  is revealed, the total information gains is, on average, exactly equal to entropy.

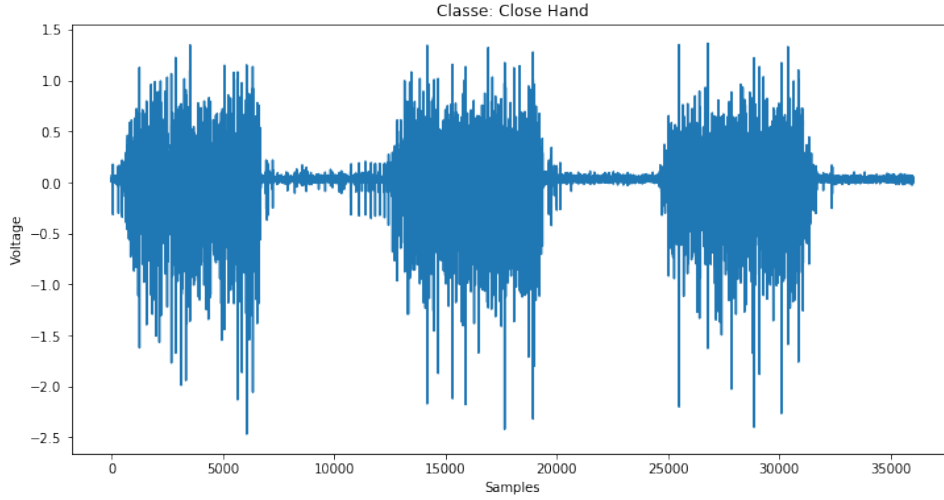
If there are consecutive and related values, then they do not provide independent information. In this case the sequence has less entropy than the sum of the individual entropy calculated with the Equation 1.1. Thus, Shannon's theorem applies not only to independent element sequences, but also to structured sequences that have dependencies between their elements.

## 1.3 ENTROPY AND SIGNAL sEMG

The sEMG signal, like all natural signals, is a diluted signal, that is, a signal that has a lot of data and little information. Moreover it is a sign in which each sampling sequence has a high degree dependence between its values.

The figure 1.3 exemplifies the sEMG signal measured on the forearm for close the hand movement. The movement was repeated three times, for three seconds with a three-second interval between the movements. As we can see from the figure, the three

repetitions led to different patterns in the graph.



**Figure 1.3.** The sEMG signal, the vertical axis represents the amplitude and the horizontal axis represents the samples.

This occurs because in each movement, different motor units are recruited, generating different electrical signals. Other factors that alter the measurement are the displacement of the sensors, sweat and muscle fatigue. One way to mitigate these effects is to acquire information a priori. Where each piece of information extracted facilitates the future classification of the signal.

With the use of a state machine, the possible movements, given a posture for the hand, were mapped and the intersection with the cluster generated a list of possible movements, greatly facilitating the classification of the EMG signal.

### 1.3.1 Real-time processing

Radhika Menon, et al [31] shows the impact of the size of the temporal window on the signal classification error sEMG. According to this study with very small windows, on average less than 200 ms the classification error increases, since the total information in the window decreases. This makes it very difficult to process in real time that needs windows, in 2011 Peerdeman [29] found that the processing window needs to be less than 300 ms or the delay becomes unacceptable to the user. Instinctively, the need for the decrease of the processing time window is in need for increasing user comfort.

A serious problem caused by the decrease in the time window is the difficulty in defining the end of rest and the beginning of movement. Clustering algorithms were used as a way to get more information about the signal. A Neighborhood Components Analysis (NCA) algorithm was used to reduce the dimensionality of the extracted characteristics and a hierarchical clustering algorithm grouped the closest movements.

The solution to this problem once again presented itself with a decrease in the entropy of the signal. An auto-encoder algorithm was used for anomaly analysis clustering the signal at rest and the motion and then, a simple perceptron classified the result.

### **1.3.2 Extracting signal information**

To extract information from surface electromyography, a technique known as characteristic extraction was used. The signal was divided into windows of 50 ms, where the sEMG signal is analyzed and frequency-based characteristics are extracted, following the main selected resources and finally, the neural network classifies movement.

The extracted characteristics are:

1. Spectral Moment (Spectral Moment (SM));
2. Sample Entropy (Sample Entropy (SE));
3. KhushabaSet;
4. Wave Length (WL) Frequency);
5. Mean Frequency;
6. Median Frequency.

## **1.4 OBJECTIVE**

### **1.4.1 Main Objective**

To develop an algorithm for a robotic prosthesis using the latest state of the art technologies art of myoelectric prostheses, from a proportional and simultaneous control, which allow posture recognition of the hand joints.

### **1.4.2 Specific Objective**

When obtaining the maximum information from the myoelectric signal, the dimension of the problem is better understood and with that its classification becomes as efficient as possible.

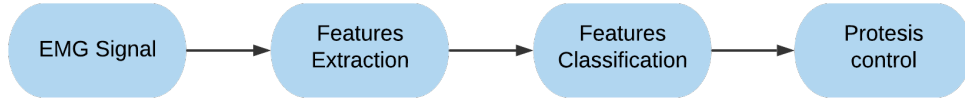
The objective is to use methods of extracting characteristics and classification of patterns in electromyographic signals, as well as adaptive training for the recognition of hand movements with varying degrees of freedom, thus increasing the comfort of the user

and giving naturalness to movement. Through these methods, a significant improvement was achieved in the pattern recognition process.

## 2 STATE-OF-THE-ART OVERVIEW

Pattern recognition for the processing of myo-electrical signals has been the research base for prosthesis control in the last decade [25]. In addition, the use of machines have been disseminated in the EMG signal analysis area, to aid in the selection of characteristics and classification of the same [12, 16, 20, 22]. The project also aims to use adaptive training for that new paradigms can be created and the techniques and methods of analysis of the myo-electric signal can be improved.

The use of pattern recognition techniques is of great importance for the control of myo-electric prostheses, bringing an improvement in the degrees of freedom and their movement beyond the capacity of their sequential control [13]. Such control typically consists of extracting characteristics of the signal and in the classification of these characteristics of segmented data in the processing signal to control an actuator. The quality of the movement of these prostheses is proportional to the processes of extracting the characteristics of the myo-electric signal and classification of standards.



**Figure 2.1.** The normal fluxogram used for sEMG prothesis.

As already mentioned, most control systems employed in prosthetic hands is sequential control, but recently, a lot of research is being conducted to employ simultaneous control [5, 11, 31, 44]. One of the main advantages of simultaneous control of the prosthesis is to approximate the movement of the prosthesis natural physiological, however this leads to an increase in the characteristics necessary to classify the signal, in addition to the interference of the signal from adjacent muscles on the capture of sEMG.

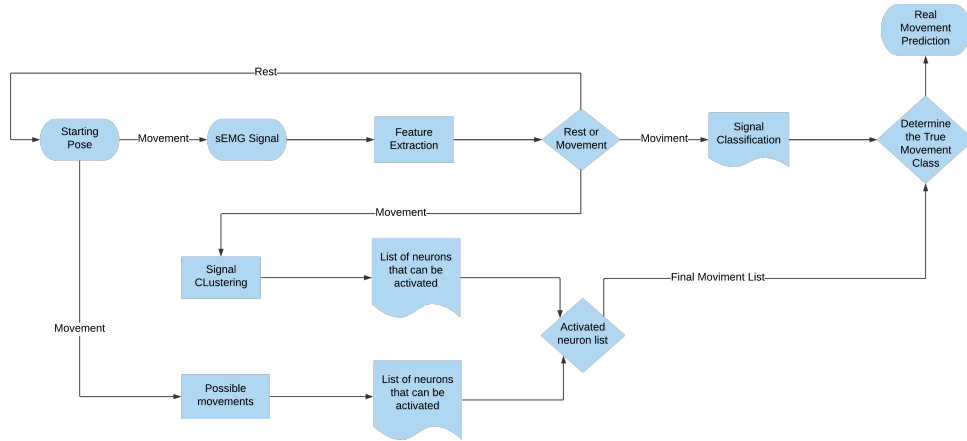
**Table 2.1.** State of the art table of sEMG signal analysis.

Reference	Class	Channel	Features	Classifier
(Huang et al. 2005)	6 classes of movement	4 channels	Gaussian Mixture Model (Gaussian Mixture Model (GMM))	GMM and Majority Vote
(Kheirizadeh et al. 2011)	6 classes of movement	1 channel	Mean Absolute Value (MAV), Signal Slope Change (SSC), Autoregressive Model (AR) and Discrete Wavelet Transform (DWT)	Neuro-Fuzzy
(Young et al. 2013)	3 degrees of freedom	6 channels / 8 channels	MAV, Zero Crossing (ZC), SSC and Wave Length (WL)	Parallel classification strategy
(Benmetti and Goldfarb 2017)	Standardized object relocation and manipulation tasks	2 channels	normalized signal	Result tables
(Yang et al. 2011)	6 classes of movement	4 channels	DWT	Fisher criterion, NWFE (Feature projection)
(Hartmann et al. 2015)	8 classes of movement	6 channels	Root Mean Square (RMS), ZC, SSC, WL and First Three Cepstral Coefficients (3cCEPS)	Linear Discriminant Analysis
(Dean et al. 2016)	6 classes of movement	3 channels	RMS and DWT	Wavelet Neuro Network (Wavelet Neuro Network (WNN))
(Sui, Shah, and Stirling 2016)	two grab and release sequences	7 channels	GMM and Markov	WPD + Random Forest
(Lah et al. 2016)	16 classes of movement	8 channels	Four levels of the Daubechies wavelet transform	Neural Network
(Radhika Menon et al.)	7 classes of movement	128 channels	MAV, SSC, Waveform Length (WFL), ZC	Linear Discriminant Analysis (Linear Discriminant Analysis (LDA))

The purpose of this study takes into account the arrangement of several simple networks to extract information and classify the sEMG signal. The figure 2.2 show the



fluxogram used on this study. An autoencoder analyzes the EMG and classifies the state of rest and movement. If the movement starts, a cluster together with the state machine returns the possible movements that may be being executed. The last step is the classifier that chooses the movements among the possible ones for that state.



**Figure 2.2.** The proposed fluxogram used for in this study for sEMG classification.

## 3 FEATURE SELECTION

### 3.1 INTRODUCTION

Pattern recognition for myoelectric signal processing plays an important role on research for prosthetics [25]. In addition, the application of machine learning techniques has become widespread in the area of surface electromyography (sEMG) signals analysis, to enhance the feature extraction and selection as well the classification of the myoelectric pattern [16, 20, 12, 22]. Furthermore, the use of pattern recognition brings an improvement in the degrees of freedom and movement of the prosthesis beyond the capacity of its sequential control [27].

In this study, the open source BioPatRec platform was used [25, 26, 27]. With a modular and customizable concept, researchers can compare their algorithms easily and efficiently, applying them to control a prosthesis. As advantages, users, by means of this platform, can access the sEMG signals database for both sequential and simultaneous analysis, including quantitative metrics to evaluate the performance of sequential and simultaneous control in a standardized way, as well as to apply methods that the platform provides for feature extraction, feature selection, feature reduction and myoelectric pattern classification.

The objective of this work is to analyze new algorithms for feature extraction and selection methods not provided by BioPatRec platform, four additional feature extraction methods were used: the Levinson-Durbin Recursion, the Absolute value of the Summation of the Expth Root Mean, the Mean value of the Square Root and the Absolute value of the Summation of Square Root [34]. Additionally, an unsupervised method for feature selection (UFS) was used [9].

These additions were made for the improvement of the classification of the myoelectrical signal to provide a better performance for the simultaneous movement of an upper-limb prosthesis, aiming at increasing the user comfort and giving easing for the movement.

## 3.2 METHODOLOGY

For this study, the “6mov8chUFS” (Untargeted Forearm Simultaneous) [26] database was used, which is freely available on the BioPatRec platform [25]. The sEMG signal is analyzed, and the features are extracted, following the main features are selected and finally the neural network classifies the movement.

### 3.2.1 BioPatRec Platform

As previously mentioned, the biomedical signal analysis platform, BioPatRec, was used, using the Multilayer Perceptron Network with backpropagation, already configured on the platform for pattern classification.

### 3.2.2 Features Extraction

In this study were added four time-domain features described below:

- (a) The Absolute value of the Summation of the Square root (Absolute value of the Summation of the Square root (ASS)) [34]: This is the first time-domain feature. For calculation of the ASS, the first step is to first execute a full-wave rectification on the sEMG data, this help in retaining the entire energy content of the signal. Next, the integral of the rectified EMG signal is calculated with respect to the current analysis window, as expressed mathematically in Eq. (1)

$$ASS = \left| \sum_{n=1}^k (x_n)^{\frac{1}{2}} \right| \quad (3.1)$$

where k represents the analysis window,  $x_n$  denote the data within the corresponding analysis window.

- (b) The Mean value of the Square Root (Absolute value of the Summation of the The Mean (MSR)) [34]: This is the second time-do-main feature. It provides an estimated measure of the total amount of activity in the analysis window.

$$MSR = \sum_{n=1}^k (x_n)^{\frac{1}{2}} \quad (3.2)$$

where k represents the analysis window,  $x_n$  denote the data within the corresponding analysis window.

- (c) The Absolute value of the Summation of the expth root (Absolute value of the Summation of the expth root (ASM)) [34] of the data is the third time-domain feature, as shown in Eq. (3). The ASM feature provides a comprehensive insight into the amplitude of the EMG signal since it gives an approximate measure of the power of the signal which also produces a waveform that is easily analyzable. This feature contains information from which the amplitude of the rectified EMG signal could be obtained.

$$ASM = \left| \frac{\sum_{n=1}^k (x_n)^{\text{exp}}}{k} \right| \quad (3.3)$$

$$\text{exp} = \begin{cases} 0.50, & \text{if } (n \leq 0.25 \times k, \text{ if } n \geq 0.75) \\ 0.75 & \text{otherwise} \end{cases}$$

The exp. variable can assume one of two possible values (0.50 or 0.75) based on the characteristic of the EMG signal segment under analysis. The ASM is therefore determined in the following three steps: first the summation of the expth root of all values in a given analysis window is computed; followed by the mean of the resultant values; and lastly, the absolute value of the resultant mean is evaluated.

- (d) The last feature added was the Levinson–Durbin Recursion [45], it is a recursive order-update method to the calculation of linear predictor coefficients, it has applications in filter design, coding, and spectral estimation. This method was used to estimate parameters of the sEMG signal.

### 3.2.3 Features Selection

A method based on the Maximal Information Compression Index (Maximal Information Compression Index (MICI)), and the Entropy Representation (Entropy Representation (ER)) was applied for unsupervised feature selection, in order to obtain the best feature sets for the classification [9]. Analyzing through the MICI to obtain combinations of characteristics with lower value or high redundancy. Those redundant features come together with the rejected features, in order to obtain an updated set formed by the features that provide the highest ER value during the combination with non-redundant features.

On the other hand, principal component analysis (Principal Component Analysis (PCA)), an orthogonal linear transformation, that rearrange the components in the inverse order of variance It is used for dimensionality reduction in the BioPatRec platform

as default, as it is a widely used technique. In that study, PCA reduced the 160 features (20 for each channel) to the 64 best features for classification.

Later in chapter 5, new methods of reducing characteristics and dimensionality were tested and evaluated.

### **3.2.4 Neural Network Classifiers**

Despite the existence of a wide variety of different pattern recognition algorithms, the Multi-Layer Perceptron (MLP) as a supervised Artificial Neural Network was chosen because of its inherent capacity of simultaneous classification [25, 27, 3].

An MLP can be used as a logistic regression classifier, where the input is first transformed nonlinearly by a learned transformation. This transformation projects the input data into a space in which it becomes linearly separable. This middle layer is called the hidden layer. A single hidden layer is sufficient to make MLPs a universal approximator. For this study an MLP with 3 hidden layers with eleven neurons was used, the transference function was the softmax function.

### **3.2.5 sEMG Database**

As mentioned before, the 6mov8chUFS database was used, and it is available on the BioPatRec platform. This database is formed by 17 subjects, six classes of individual movements were selected, such as hand opening and closing, flexion and extension of the wrist, and pronosupination of the hand, forming 27 possible movements. The signal was measured as follows: 3 seconds contraction time with 3 seconds for relaxation between each repetition, repetitions of each motion. 8 bipolar electrodes (Disposable Ag/AgCl), 1 cm electrode diameter, 2 cm inter-electrode distance for the bipole. Electrodes were equally spaced around the most proximal third of the forearm.

The signal was extracted using overlapped time windows of 0.2 seconds and time increments of 0.05 seconds.

### **3.2.6 Statistical Evaluation**

The tests were performed on seventeen subjects of the original base "6mov8chUFS". First, the BioPatRec feature selection methods were used with and without PCA. Finally, the unsupervised feature selection algorithm (Unsupervised Feature Selection (UFS)) was used for reducing features in place of the PCA. Thus the tests were completed they were repeated by adding the new characteristic vectors proposed in this study.

The evaluation of the classifier used a cross-validation of 48 trainings with randomized datasets were per subject and for each algorithm, 24 validation sets, and 49 test sets. For this study was used unitary range normalization. The BioPatRec provides statistical tools to evaluate the proposed algorithms on the platform, thus, it has a wide variety of metrics [27] that were used to analyze the results, such as Accuracy (Class Specific), Sensitivity (Recall), PPV (Precision), F1, Specificity (Negative Condition), NPV (Negative Outcome) and Accuracy (Global).

(a) **Accuracy – Class Specific**

$$AccCS = \frac{\text{absolute correct predictions}}{\text{total absolute predictions}} \quad (3.4)$$

(b) **Sensitivity**

$$Sensitivity = \frac{TP_s}{TP_s + FP_s} \quad (3.5)$$

(c) **PPV**

$$PPV = \frac{TP_s}{TP_s + FP_s} \quad (3.6)$$

(d) **F1**

$$F1 = 2 \times \frac{\text{precision} \times \text{sensitivity}}{\text{precision} + \text{sensitivity}} \quad (3.7)$$

(e) **Specificity**

$$Specificity = \frac{TN_s}{TN_s + FP_s} \quad (3.8)$$

(f) **NPV**

$$NPV = \frac{TN_s}{TN_s + FN_s} \quad (3.9)$$

(g) **Accuracy – Global**

$$AccG = \frac{TN_s + TP_s}{TP_s + TN_s + FP_s + FN_s}, \quad (3.10)$$

Were TP means true positive, the correct activation; TN is true negative, the correct inactivation; FP is false positive, the incorrect activation; and FN is false negative, incorrect inactivation.

### 3.3 RESULTS AND DISCUSSION

It has been shown that the individual movements can be successfully predicted offline using pattern recognition algorithms [25, 20, 36, 14] and in this study was demonstrated an increase in terms of classification accuracy was achieved when the new features were used in conjunction with a feature reduction algorithm. Table 1 shows the results obtained using the characteristics present in BioPatRec and the use of PCA or UFS to reduce characteristics.

**Table 3.1.** Quantitative indicators obtained with the comparison between old features with feature reduction algorithms.

Features Selection	Accuracy Class	Sensitivity	PPV	F1	Specificity	NPV	Accuracy Global
None	83.10	85.11	92.30	0.89	99.73	99.43	99.19
PCA	92.74	94.78	94.78	0.95	99.80	99.80	99.61
UFS	90.63	92.29	95.02	0.94	99.81	99.70	99.54

Table 3.2 shows the results obtained using the characteristics present in BioPatRec and those added by this study, in addition to the use of PCA or UFS to reduce the characteristics. As more characteristics are sorted, the MLP begins to diverge, but when selecting the most divergent characteristics, through the methods of selecting characteristics, rating the signal becomes easier. The performance of the proposed and the conventional methods present in the BioPatRec is shown by the gain in the accuracy and deviation shown in table 3.1 and 3.2.

**Table 3.2.** Quantitative indicators obtained with the comparison between the new features added to the old features sets with feature reduction algorithms.

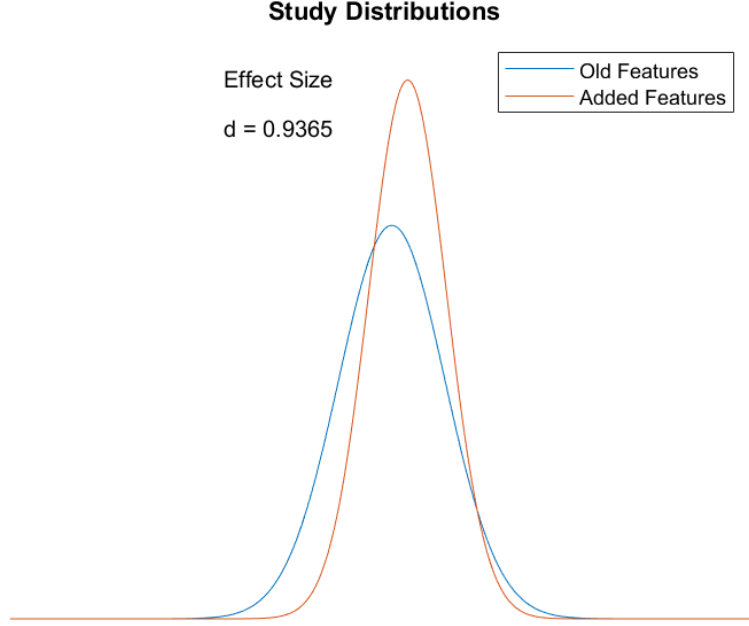
Features Selection	Accuracy Class	Sensitivity	PPV	F1	Specificity	NPV	Accuracy Global
None	82.39	83.60	95.76	0.89	99.86	99.37	99.26
PCA	94.63	96.22	96.95	0.97	99.88	99.85	99.75
UFS	93.50	94.71	96.68	0.96	99.87	99.80	99.68

The below figure shows the distribution curves of the experiment, while the blue curve represents the old features the red one represent the experiment with the added features.

The difference between two means divided by a standard deviation for the data is represented with the "d" letter in the graphic.

According to Cohen and Sawilowsky:

- $d = 0.01 \implies$  *very small effect size*;
- $d = 0.20 \implies$  *small effect size*;
- $d = 0.50 \implies$  *medium effect size*;



**Figure 3.1.** Distribution of the experiment with the old features and the added features.

- $d = 0.80 \implies$  *large effect size*;
- $d = 1.20 \implies$  *very large effect size*;
- $d = 2.00 \implies$  *huge effect size*.

The experiment was offline, and the mean training and validation time was of 34 seconds and the mean testing time was of 0.004 seconds.

The use of BioPatRec allows a fast and accurate simulation of the pattern recognition algorithms, which streamlines the process of development and testing of theories that will be applied in the control of a myoelectric prosthesis. This platform is being used in this study and it is hoped that it can assist in the development of an adaptive learning pattern recognition system for the control of an upper limb electrical prosthesis. In addition, the fact that the BioPatRec platform is modular allows the study to be better divided into stages, such as signal processing, extraction of characteristics, classification and the decision-making system of the prosthesis, making the process agiler.

### 3.4 CONCLUSION

The addition of the new features in conjunction with the selection algorithms improved the characterization of the myoelectric signal, which will facilitate the decision



process for the control of the myoelectric prosthesis. The help of the BioPatRec platform made the work agiler and the statistical metrics helped to evaluate the effectiveness of the algorithms applied in this study. Furthermore, this study showed that simultaneous control can be considered since it improves user comfort. In addition, simultaneous control is required for more natural control of artificial limbs, and pattern recognition has proved to be an excellent means of working with the complexity generated by simultaneous movement.

## 4 AUTOENCODER AND ANOMALY DETECTION

### 4.1 INTRODUCTION

An autoencoder is a neural network that is trained to attempt to copy its input to its output. While copying the input to the output may be useless, a high-dimensional data can be converted to low-dimensional codes by training a multilayer neural network with a small central layer to reconstruct high-dimensional input vectors. Using the inner layer, also called latent space, we have a signal with a reduced size, more easily classified.

The Variational Autoencoder (VAE) is a generative model that estimates the Probability Density Function (Probability Density Function (PDF)) of training data. By training the model to recognize the sEMG signal it will assign a high probability value to a motion class, while the noise will receive a low probability value[42].

The VAE model can also sample examples of the PDF learned, thus generating new examples similar to the original data set. But it is important to emphasize that VAE is not a way of training generative models, but rather that the generative model is a component of VAE, generally being a deeply latent Gaussian model[32].

In the figure 4.1 are the steps to summarize the operation of a VAE. On the left side we have the model definition:

1. Let  $q(z|x)$  be defined as how the signal is encoded into a distribution over the latent space;
2. Let  $z$  be A latent vector sampled from  $q(z|x)$ ,  $z$  will contain the information describing  $x$ . The decode of it is represented as  $p(z|x)$ ;
3.  $z$  is decoded as a signal.

On the right side we have the loss:

1. Reconstruction error: the difference between the output and the input.
2.  $p(z|x)$  should be similar to the prior (multivariate standard Gaussian).

The VAE generating coefficient appears by directly sampling the latent vector from the prior distribution and decoding it into a noisy representation of  $x$ [42].

#### 4.1.1 Model specification

##### 4.1.1.1 Encoder layer

In Bayesian modeling, the distribution of observed variables is governed by latent variables. The latent variables are extracted from a previous density  $p(z)$  and related to the observations through the probability  $p_\theta(x|z)$ . Deep latent gaussian models (DLGMs) are a general class of models where the observed variable is governed by a latent variable hierarchy, and the latent variables at each level of the hierarchy are Gaussian a priori[32].

Normally, in the VAE, a Gaussian distribution is used to sample the latent space.

$$p(z) = N(0, I) \quad (4.1)$$

In this way, each local latent variable is related to its corresponding observation through the likelihood  $p_\theta(x|z)$ , which can be seen as a probabilistic decoder. Using a hidden smaller representation  $z$ , it decodes it into a distribution over observation  $x$ .

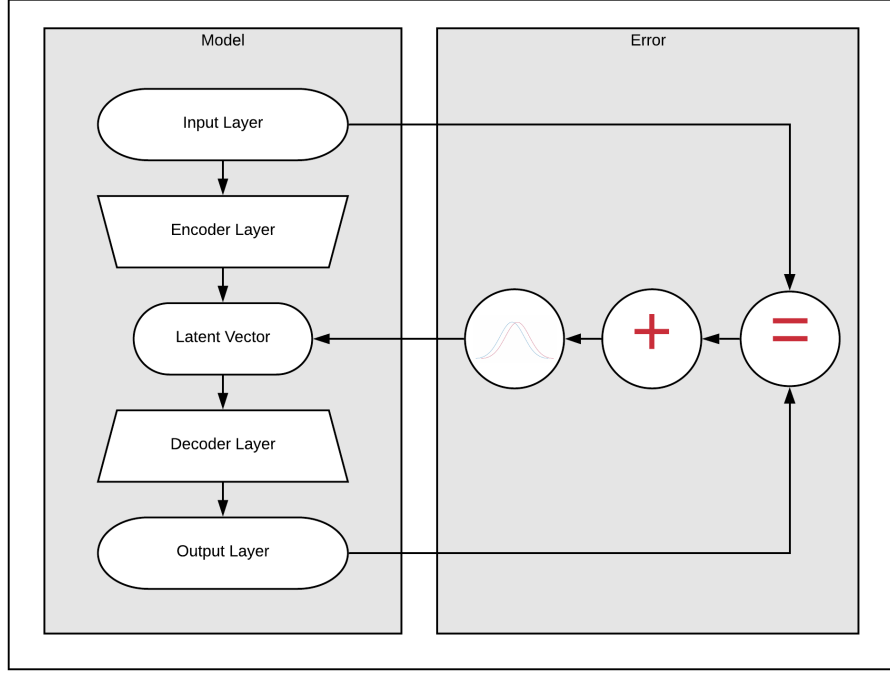
##### 4.1.1.2 Decoder layer

The decoder is another neural network. Its input is the latent vector  $z$ , generates the parameters for the probability distribution of the data and has weights and biases  $\theta$ . The decoder is denoted by  $p_\theta(x|z)$ . The probability distribution is a multivariate Gaussian.

The loss function of VAE is the negative log-likelihood with a regularizer. Since it is not possible to generalize the global representation shared by all data points, we can decompose the loss function into terms that depend only on a single data point  $l_i$ . The parameters are typically the weights and biases of the neural networks represented as  $\theta$  and  $\phi$ . The total loss will then be represented by the sum of all losses  $l_i$  for all data points. The loss function  $l_i$  for the data point  $x_i$  is:

$$l_i(\theta, \phi) = \mathbf{E}_{z \sim q_\theta(z|x_i)}[\log(p_\phi(x_i|z))] + \mathbf{KL}(q_\theta(z|x_i)||p(z)) \quad (4.2)$$

The first term represent the reconstruction loss, and has the form of a negative log-likelihood of the i-th data point. The expectation is taken with respect to the encoder's distribution over the representations. This term encourages the decoder to learn how to reconstruct the data. If the decoder's output does not have similarity with the data,



**Figure 4.1.** Flowchart of a variational autoencoder

statistically the decoder parameterizes a likelihood distribution that does not place much probability mass on the true data. Poor reconstruction will incur a large cost in this loss function.

The second term is a regularizer. This is the Kullback-Leibler divergence between the encoder’s distribution  $q_{\theta}(z|x)$  and  $p(z)$ . This divergence measures how much information is lost when using  $q$  to represent  $p$ , in other words, the divergence of the approximate from the true posterior.

#### 4.1.1.3 Inference Network

An inference network is a flexible construction for parameterizing approximating distributions during inference[8], and it is used on VAE[32, 42] to infer the optimal values of the latent variables given observed data, or to calculate the posterior  $p(z|x)$ . By modeling the true distribution  $P(z—X)$  using simpler distribution that is easy to evaluate, e.g. Gaussian, and minimize the difference between those two distribution using KL divergence metric, as in 4.2, which tells us how difference it is  $p$  and  $q$ .

### 4.1.2 Anomaly Detection

The use of auto-ponder for anomaly detection is already common in the area of artificial intelligence [41, 47, 4]. In this study, a variational autoencoder was used to detect the onset of movement, which is otherwise confused with the resting state.

The test of the usability of the auto-encoder was done for the classification of the 26 classes of movement of the database, it was later retrained to detect the change of the resting state.

## 4.2 METHODOLOGY

### 4.2.1 Database

In this work the "6mov8chFUS" database made available with the BioPatRec platform[25] was used. The "6mov8chUFS" database consists of 17 patients, with six individual movement classes selected, such as opening and closing the hand, flexion and extension of the wrist, and prono-supination of the hand, forming 27 possible movements.

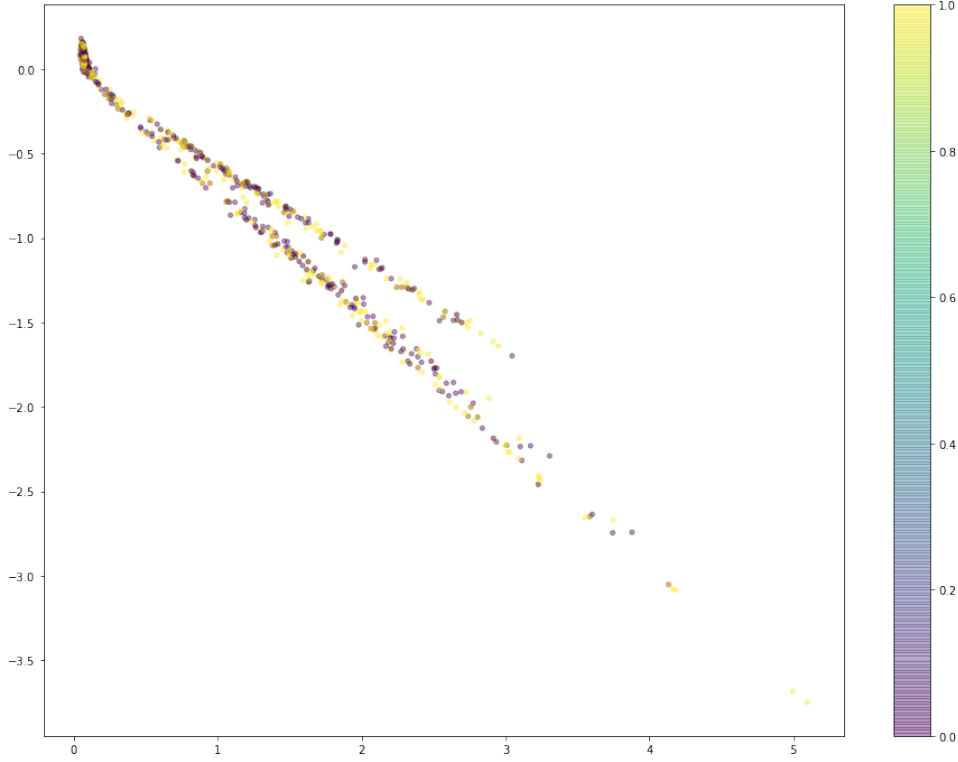
The signal was measured as follows: 3 seconds of contraction time with 3 seconds for relaxation between each repetition, repetitions of each movement. 8 bipolar electrodes (disposable Ag / AgCl), 1 cm of electrode diameter, 2 cm of inter-electrode distance for the bipolar. Electrodes were equally spaced around the proximal third of the forearm.

### 4.2.2 VAE Algorithm

A simple Variational Autoencoder was used for the anomaly detection phase. The encoder stage was composed of three dense layers. The first one has 32 neurons, the second 16 neurons and the third 8 neurons. The sampling inference, which characterizes the variational autoencoder, is made on top of the last layer of the encoder step, which has 8 neurons. All layers in this phase have the Rectified Linear Unit (ReLU) as the activation function.

The decoding stage is the reverse of the encoding. It starts with 8 neurons in the first layer, 16 neurons in the second and 32 neurons in the last. The first two layers have ReLU as the activation function and the third layer, the sigmoid as the activation function.

The figure 4.2 is the dimensionless values of the autoencoder trained with two neurons in the latent layer. The figure represents the representation of the anomalies found when the patient starts to move. With that output of the latent space, a simple perceptron



**Figure 4.2.** The latent space representation created by the Variational Autoencoder

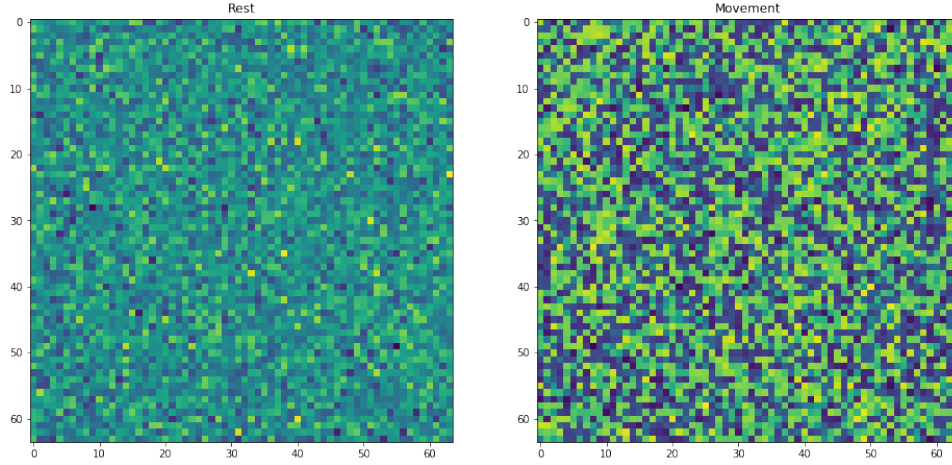
network is used to detect whether the patient is moving or resting.

### 4.3 RESULTS

In this experiment, a 0.01s window (10 samples) was used. With this window size, normal classification methods cannot accurately classify the movements classes or even differentiate the resting state of movement. Through the use of auto-encoder as an anomaly detector, it was possible to determine the moment when the movement really started.

The figure 4.3 is the representation of the difference between the rest movement state in the latent space of the variational autoencoder. For the construction of the image the latent space was dimensioned with 4096 neurons. The output of the neurons has been scaled to a 64 by 64 square shape and the image was made.

It is worth remembering that, although the detection of the anomaly only needed eight neurons to be able to perform successfully, the representation of the latent space needed to be enlarged so that the difference was visible to our eyes.



**Figure 4.3.** A expanded latent space, to represent the difference between the rest and the rest and the movement state. In this image the latent space was dimensioned to be 64 x 64 to facilitate the difference visualization.

#### 4.4 CONCLUSION

The results obtained demonstrate the feasibility of the variational autoencoder as an anomaly detector for the sEMG signal. Furthermore, by separating the resting state from the other classes of movement, the entropy of the system is reduced, facilitating the future classification of the other classes of movement.

By using VAE, which is specialized in separating classes in their latent space, movement detection has been simplified and its computational cost has been greatly reduced. Although the training was done only using the offline signal, simulating the acquisition of the signal online proved to be extremely efficient.

In this study, the auto encoder was used only as an anomaly detection instrument, but its use is not limited to that. As shown in figure 4.3, the autoencoder can be used to increase the separation between classes just by increasing the size of the latent space. This technique can be used in further studies to improve the system accuracy.

## 5 ENTROPY

### 5.1 INTRODUCTION

The EMG signal varies according to the posture, position, and force applied by the person that is performing a muscle contraction. Therefore, to have perfect control of a prosthesis, extensive training with a myoelectric unit is required. However, as these prostheses have a high cost investment, it is often not possible to offer longterm training for their use.

One of the difficulties regarding EMG signal processing for use in prosthesis control is the need for real-time processing. This creates the need for ever-smaller time windows that, in turn, have less signal information, which limits the amount of information that can be extracted from them.

In the last few years, deep learning techniques have become increasingly used, but the more complex the applications are, the more complex the neural networks become. However, the more parameters the network has, the higher the chance of having overfitting, which causes considerable deterioration to the network's generalization capacity. Otherwise, if the neural network has few parameters, it will probably not be able to represent the data accurately. Notably, the best way to achieve generalization is to seek a balance between training error and network complexity [21, 46].

In the tasks of biosignal classification, there are frequently too few biomedical signal samples to allow for the achievement of good results with deep learning [35]. Another problem associated with deep learning is the cost of training processing, which demands specific hardware and high computational cost due to the complexity of the current neural networks architectures [46]. This thesis suggests extracting maximum signal information before signal classification as a method to reduce the system complexity and create redundancy for the classification and thereby managing to decrease the time window for the real-time processing of the signal. One of the steps of the algorithm proposed includes the development of a pipeline that extracts a priori information from the EMG signal by computing the current hand and forearm postures and the similarity of the EMG signals from the forearm.



As mentioned before, one way to ensure entropy reduction (information gain) is to obtain a priori information [37]. The starting position of the hand and forearm provide a great deal of information to the classifier as it restricts possible movement classes. For this purpose, a state machine [3] that counts the possible movements from the initial classes was created. This state machine reduced by just over three times the number of classes to classify, improving the classification of the system.

A second method used in this work for entropy reduction, was the classification of signals by similarity. Therefore, a Hierarchical Agglomerative clustering (Hierarchical Agglomerative Clustering (HCA)) technique was selected. In this technique, each data point is considered an individual cluster. At each iteration, using distance, a similar pair of clusters are merged as they move up the hierarchy, until there is a formation of one cluster or K clusters.

The proposed methodology led to a substantial decrease in the size of the temporal window used for sEMG signal processing. Besides, there was also an improvement in the generalization and processing speed due to the simplification of the model used for classification.

This chapter is organized as follows: Section II shows a detailed explanation of the algorithms used. The results are listed in Section III, which also presents a detailed analysis of the experiment. Finally, there is a conclusion, where the results are summarized.

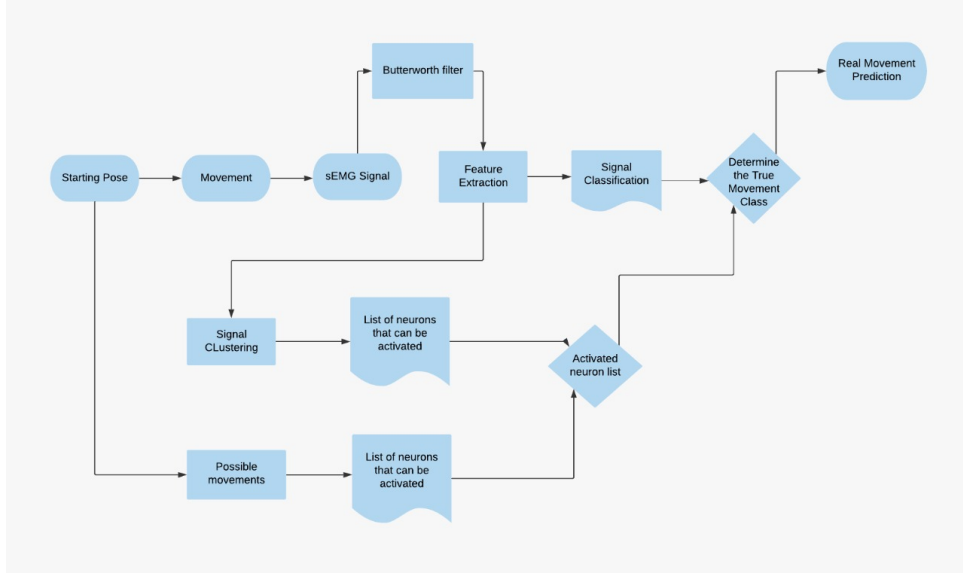
## **5.2 METHODOLOGY**

### **5.2.1 The Pipeline**

Initially, there is an estimation of the original positions of the hand and forearm. The possible movements are checked by analyzing the state machine; these two steps lead to a list of values associated with the neurons that are activated in the final layer of the classifier.

The second step is processing the signal with the Butterworth filter, followed by the feature extraction and standardization of the data. Furthermore, a reduced space transform created by the NCA algorithm is applied, and the signal is ready for the classification and clustering algorithms. This transformation reduces the vector of characteristics from 36 to 25 dimensions.

The third step for signal classification is to find the group to which the analyzed window belongs. The HCA clustering algorithm will provide a new list of activated neurons in the classifier. In the HCA algorithm, each observation starts in its cluster, and the algorithm merges pairs of clusters when one moves up the hierarchy. The intersection



**Figure 5.1.** Pipeline representation: the algorithm begins at a starting position; the sEMG signal is measured and pre-processed to help to determine the movement intention; with these two information pieces, the possible next movements are determined as well as the cluster to which the class belongs; finally, the algorithm yields the movement intention.

between the state machine-generated list and the cluster generates the final list of neurons that can be activated.

The last step is the signal classification by the MLP and the multiplication of the result by the list generated in the previous step. The image below shows the value of neurons before and after applying the list of neurons to be used.

### 5.2.2 Database

The "6mov8chUFS" database [26] consists of 17 patients, with six individual movement classes selected, such as opening and closing the hand, flexion and extension of the wrist, and pronosupination of the hand, forming 27 possible movements. The signal was measured as follows: 3 seconds of contraction time with 3 seconds for relaxation between each repetition, repetitions of each movement. 8 bipolar electrodes (disposable Ag / AgCl), 1 cm of electrode diameter, 2 cm of inter-electrode distance for the bipolar Electrodes were equally spaced around the proximal third of the forearm.

The algorithm was tested with the "6mov8chFUS" database, made available by BioPatRec platform [25]. Also, with the help of the BioPatRec platform, an analysis was made using seventeen characteristics. With the use of principal component analysis (PCA) the four best characteristics for the execution of the algorithm were selected. The number of features was chosen to take into account the speed of the information

processing, which needed to be as fast as possible to create the most natural movement possible.

### **5.2.3 Movement Classes**

The movement classes used in the "6mov8chUFS" database are listed as follows:

1. Open Hand
2. Close Hand
3. Flex Hand
4. Extend Hand
5. Pronation
6. Supination
7. Open Hand + Flex Hand
8. Close Hand + Flex Hand
9. Open Hand + Extend Hand
10. Close Hand + Extend Hand
11. Open Hand + Pronation
12. Close Hand + Pronation
13. Open Hand + Supination
14. Close Hand + Supination
15. Flex Hand + Pronation
16. Extend Hand + Pronation
17. Flex Hand + Supination
18. Extend Hand + Supination
19. Open Hand + Flex Hand + Pronation
20. Close Hand + Flex Hand + Pronation
21. Open Hand + Flex Hand + Supination

- 22. Close Hand + Flex Hand + Supination
- 23. Open Hand + Extend Hand + Pronation
- 24. Close Hand + Extend Hand + Pronation
- 25. Open Hand + Extend Hand + Supination
- 26. Close Hand + Extend Hand + Supination

#### 5.2.4 Feature extraction

In order to reduce signal noise, the first step in extracting features was to use a sixth-order bandpass Butterworth filter, 80-450Hz. Additionally, a time window for sampling the signal is selected. In this study, the sampling frequency of the sEMG was 2000 Hz, with a 0.01s window (i.e., ten samples per window).

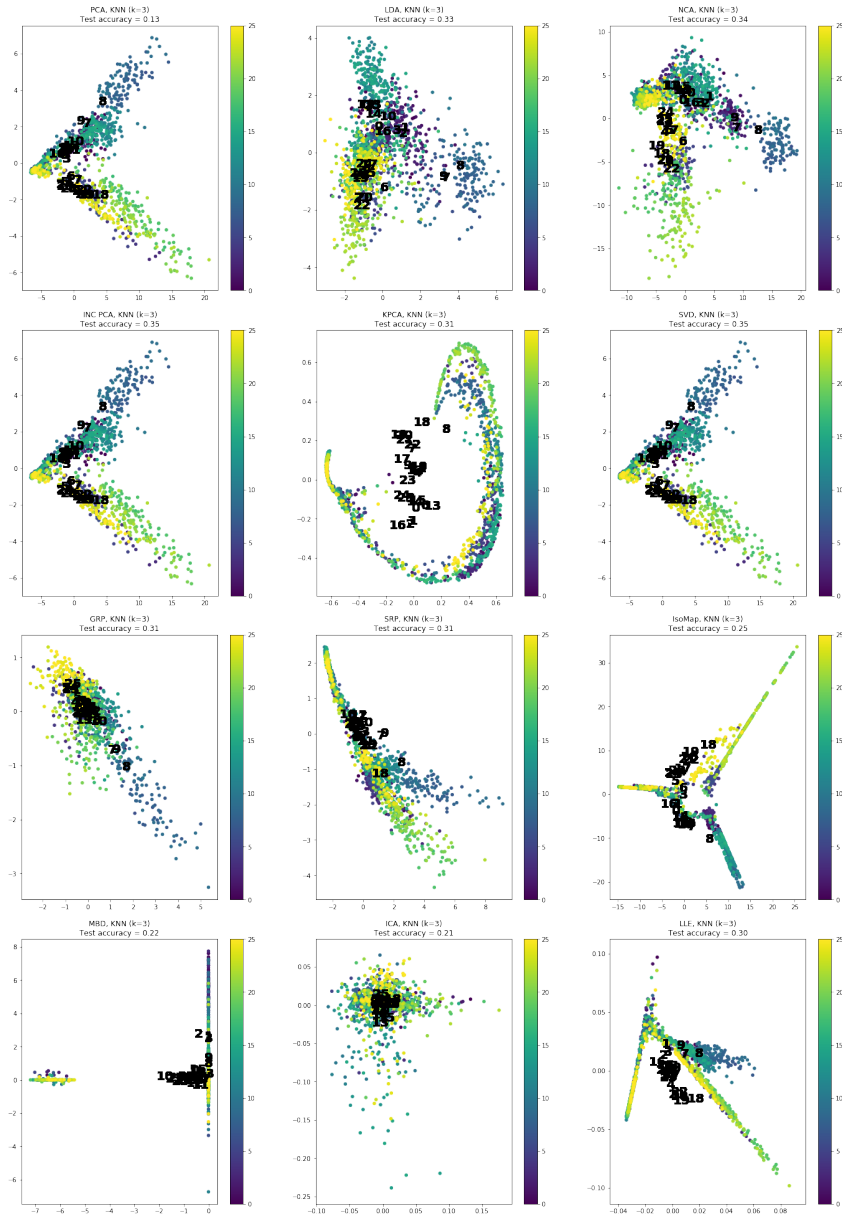
This stage was subdivided into two steps:

1. **Selection of characteristics:** to extract information from the signal four frequency domain features were used [19]:
  - Spectral Moment;
  - Waveform Length (acumulative changes in the length);
  - Mean;
  - Median;
2. **Dimensional reduction:** (NCA) [30] allowed for the dimensional reduction by helping to select the most significant features of the signal. NCA is a supervised learning algorithm for distance metric learning. It learns a linear transformation (of input data) that maximizes, in the transformed space, the average leave-one-out classification performance.

The figure 5.2 shows the study made to select the best dimensionality reduction algorithm. The methods were tested with a KNN with  $k = 3$ , the method was select base on the accuracy of the knn. The plots represent the features on the dimensionality reduction algorithm space, therefore they don't have dimensions.

#### 5.2.5 Signal Information

To extract the maximum information of the signal the processes were divided into two steps:

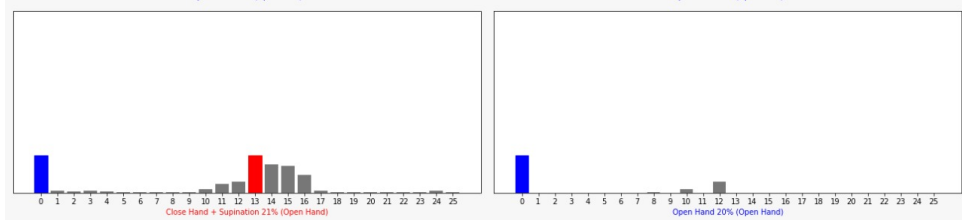


**Figure 5.2.** Test of different methods of dimensionality reduction for the use in the classificatory algorithm. The methods were tested with a KNN with  $k=3$ .

1. Signal Clustering: Agglomerative Hierarchical Clustering (HCA) clusterized the signal into three groups according to the similarity of the features. The HAC algorithm recursively merges the pair of clusters that minimally increases a given linkage distance [38, 15];
2. Comparison with possible movements: after the creation of the cluster, the algorithm compares movements classes with the possible movements for a given position, and the classes are extracted through a process called Decision Tree or State Machine [9].

### 5.2.6 Classifier Algorithm

A simple multi-layer perceptron (MLP), with three layers, was used to classify the signal. The first layer was composed of 25 neurons, the middle layer by 52 neurons and the last by 26 neurons. In the first two layers, the linear rectifier was used as activation function. The last layer had a softmax function helping in the classification process. A dropout function (20%), placed between the MLP layers, reduce the chance of over-fitting. The MLP was chosen because of its inherent capacity of simultaneous classification [37].



**Figure 5.3.** Output of the values of the MLP neurons of the last layer. The blue color mean the right classification and the red the wrong classification. The left figure is the output of the MLP and the right figure is the result with the possible classes only.

## 5.3 RESULTS

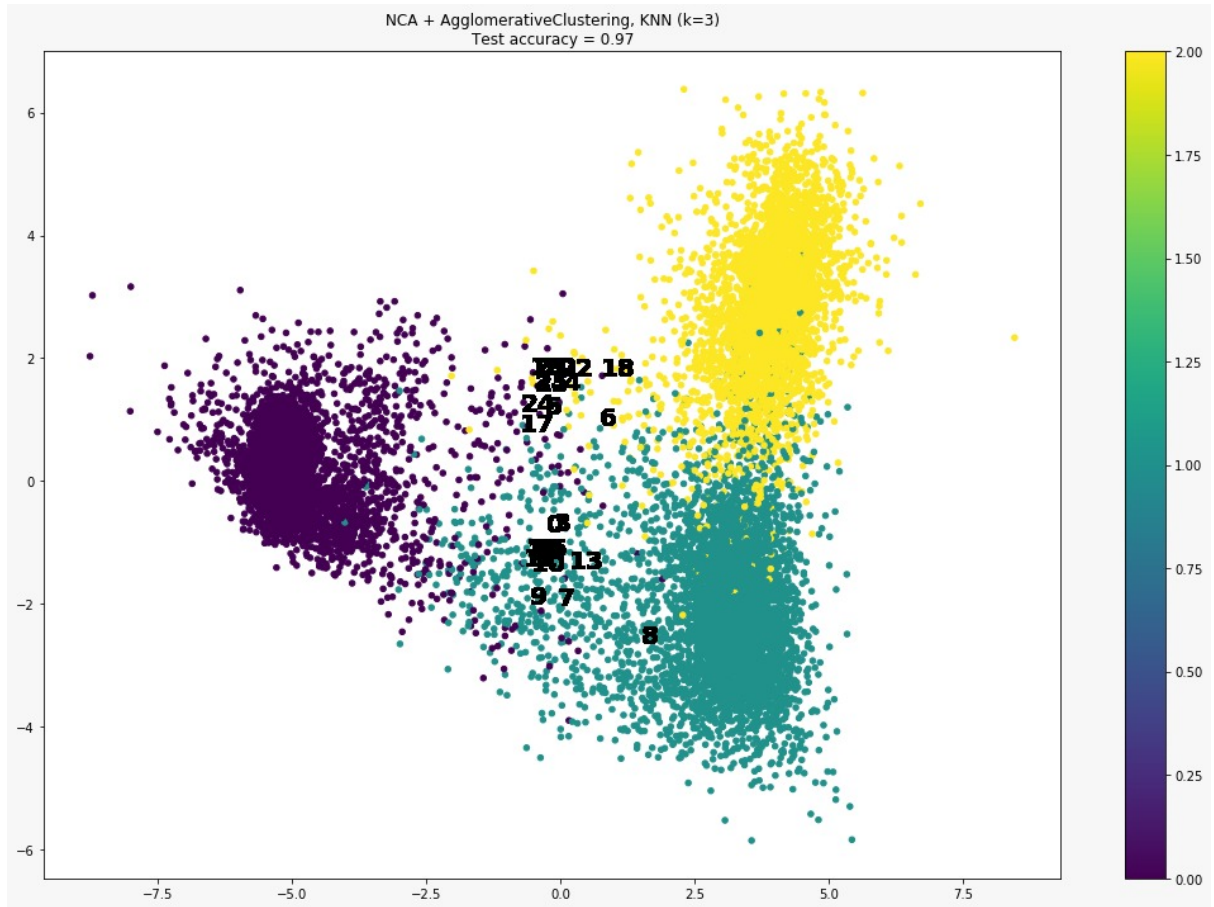
A recent study [31] shows the impact of the temporal window size on the EMG signal classification error. According to this study, with very small windows, (on average less than 200 ms), the classification error increases as the total information in the window decreases. This feature makes real-time processing very difficult, since it requires small temporal windows. Moreover, in 2011 Peerdeman [29] found that the processing window needs to have less than 300 ms, or the delay becomes unacceptable to the user.

The proposals studied in this article aim to reduce this limit from 200 to 300 mil-

liseconds by creating mechanisms that allow for using smaller windows. Through these smalltime windows, it is possible to perform real-time processing.

Because of the small size (10ms) of the window, the MLP did not achieve a good signal classification, but the selection of which neurons are active for the classification, significantly increased the accuracy. Figure 2 shows the difference in the classification using only the possible classes.

For the clustering algorithm, three data clusters were created and evaluated using a KNN ( $k = 3$ ). This KNN achieved 97% accuracy by classifying the groups created, as seen in figure 5.4.



**Figure 5.4. sMEG!** Clustering the classes by the similarity in the features extracted. After the extraction of twenty-five dimensions, the two more representative dimensions among them were used to generate the plot. A KNN was used to validate the cluster.

Table 5.1 shows the accuracy and standard deviation of MLP used in this article. The left side shows the results for the full MLP, and the right column shows the results after the application of the pipeline shown in figure 5.1.

Although the standard deviation remained relatively unchanged, the improvement in accuracy was immense. This achievement can be reached by excluding very close

**Table 5.1.** Accuracy and Standard Deviation Comparison

	Ful MLP Output	MLP with Selectd Neurouns
Acc	0.382	0.913
Std	0.013	0.011

classes, such as closing and flexing the hand or opening and flexing the hand, where misclassifications generally occur.

## 5.4 CONCLUSION

In this thesis, two methods were used to obtain a priori information and thus reduce signal entropy before classification. New methods can bring even more significant improvement to the system. By providing a priori information for signal classification interactively, the number of possible classes for signal classification dramatically decreases. Creating less complex validation steps also increased accuracy while allowing window size reduction. We believe that the techniques presented here only scratch the surface of the applications where information entropy can and should be used.

The main idea of this study was to create a network that is simple and, through a small number of bits, can generalize the data better than a more complex network. Therefore, it is imperative to provide tools that simplify or provide data information. Besides, a simple neural network will have faster processing time and use less energy, being cheaper to train and more efficient to apply. Thus, providing tools that simplify or provide data information is critical. Furthermore, more studies to account for the trade-off between the number of steps and the processing time of the pipeline is needed.



## 6 CONCLUSION

This study was done with the intention of raising a greater amount of information about the EMG signal. This survey serves to build a better definition of the problem. Knowing how the signal behaves it is possible to create simpler and more efficient solution. For this purpose, a processing protocol of feature extraction, clustering and classification of the myo-electric signal were followed.

As a good prosthesis experience needs quick responses, real-time processing is critical. The number of features and their choice was such as to guarantee faster signal processing. Four characteristics in the frequency domain were selected:

- Spectral Moment;
- Waveform Length (cumulative changes in the length);
- Mean;
- Median;

One of the difficulties in classifying the signal was knowing exactly where the movement started. This happens because of the stochastic characteristics of the signal, when motor units start to be recruited, the signal of rest and movement are confused. To differentiate these two states, a variational self-hiding was used. Due to its intrinsic characteristics, it is specialized in separating classes (movement and rest) in its latent layer. With a simple perceptron it was possible to ascertain the change in the user's states and activate the prosthesis movement classifier.

The last part of this study was the development of a classifier. For this, the concept of information entropy was used. The first part of the classifier was the construction of a state machine that maps the possible combinations of movement given an initial position of the prosthesis. Another step was the use of clusters to catalog nearby movements, the intersection of these two steps generates a list of possible movements. Because of the extremely small time window, without using these two steps, classification would be impossible.

Since the myo-electric prostheses were developed, the number of users who reject it has remained constant [6], Which shows that there has been no significant advance in the area. The reduction of the time window for the classification of the myo-electric signal will allow the control of the prosthesis to be performed in a more fluid way by the user, increasing his comfort when using it.

Another very important factor generated by the decrease in the time window is the decrease in the complexity of the evaluation, which results in energy savings during the classification process, which, in turn, would increase the time of using the prosthesis, reducing the amount of time of refills that the user would need to do. In addition, the processing developed in this study can be used for the classification of other natural signals which, as they are diluted (too much information for little information), are difficult to classify.

The collection of EMG signal data can be stressful for the patient, so it is very difficult to get large databases of biological signals. Because of this, this study was done using only one database, provided along with the BioPatRec platform. In addition, the database used has few repetitions of movement, which makes training with machine learning algorithms difficult. In total, this base has only 17 patients and 3 repetitions of each movement, which made it impossible to use some training techniques.

With a sufficiently large database and, once a redundant classifier has been created, as in the case of the study, the various steps of the classification can be used as a code and, with a larger neural network and the values of the classification windows can be predicted using techniques such as Long Short Time Memory (LSTM).

One of the main future works would be the implementation of the system and tests with patients with live acquisition of the sEMG. The correlations between the time windows can also be studied in order to have a better designed classifier. New methods of reducing entropy before the classifier must be tested since tools that simplify or provide data information is critical.

Despite the few samples of the data set used, the methodology created in this thesis proved to be efficient for the classification of the sEMG signal. However, further studies must be carried out to find new ways to decrease the entropy of the signal before its final classification.

# References

- [1] Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, Manjunath Kudlur, Josh Levenberg, Rajat Monga, Sherry Moore, Derek G Murray, Benoit Steiner, Paul Tucker, Vijay Vasudevan, Pete Warden, Martin Wicke, Yuan Yu, Xiaoqiang Zheng, Google Brain, Implementation Osdi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, Manjunath Kudlur, Josh Levenberg, Rajat Monga, Sherry Moore, Derek G Murray, Benoit Steiner, Paul Tucker, Vijay Vasudevan, Pete Warden, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow : A System for Large-Scale Machine Learning This paper is included in the Proceedings of the TensorFlow : A system for large-scale machine learning. 2016.
- [2] Sebastian Amsuss, Peter M. Goebel, Ning Jiang, Bernhard Graimann, Liliana Paredes, and Dario Farina. Self-correcting pattern recognition system of surface EMG signals for upper limb prosthesis control. *IEEE Transactions on Biomedical Engineering*, 61(4):1167–1176, 2014.
- [3] Mohammadreza Asghari Oskoei and Huosheng Hu. Myoelectric control systems-A survey, 2007.
- [4] Caglar Aytekin, Xingyang Ni, Francesco Cricri, and Emre Aksu. Clustering and Unsupervised Anomaly Detection with l2 Normalized Deep Auto-Encoder Representations. *Proceedings of the International Joint Conference on Neural Networks*, 2018-July, 2018.
- [5] Daniel A. Bennett and Michael Goldfarb. IMU-Based Wrist Rotation Control of a Transradial Myoelectric Prosthesis. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2017.
- [6] Stephen F. Burrough Brook and Judith A. Patterns of Acceptance and Rejection of Upper Limb Prostheses, 1985.

- [7] ”” C.E. Shannon , vol. 27, pp. , July, October,. A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(April 1928):379–423,623–656, 1948.
- [8] Peter Dayan. Recurrent sampling models for the Helmholtz machine. *Neural Computation*, 11(3):653–677, 1999.
- [9] Denis Delisle-Rodriguez, Ana Villa-Parra, Teodiano Bastos-Filho, Alberto López-Delis, Anselmo Frizera-Neto, Sridhar Krishnan, and Eduardo Rocon. Adaptive Spatial Filter Based on Similarity Indices to Preserve the Neural Information on EEG Signals during On-Line Processing. *Sensors*, 17(12):2725, 2017.
- [10] C.P. Diehl and G. Cauwenberghs. Svm incremental learning, adaptation and optimization. *Proceedings of the International Joint Conference on Neural Networks*, 2003., 4(x):2685–2690.
- [11] Strahinja Dosen, Marko Markovic, Kelef Somer, Bernhard Graimann, and Dario Farina. EMG Biofeedback for online predictive control of grasping force in a myoelectric prosthesis. *Journal of NeuroEngineering and Rehabilitation*, 2015.
- [12] Feng Duan, Lili Dai, Wennan Chang, Zengqiang Chen, Chi Zhu, and Wei Li. SEMG-Based Identification of Hand Motion Commands Using Wavelet Neural Network Combined with Discrete Wavelet Transform. *IEEE Transactions on Industrial Electronics*, 2016.
- [13] Kevin B. Englehart and Bernard Hudgins. A robust, real-time control scheme for multifunction myoelectric control. *IEEE Trans. Biomed. Eng.*, 50(7):848–854, 2003.
- [14] Purushothaman Geethanjali. Myoelectric control of prosthetic hands: State-of-the-art review, 2016.
- [15] Yuri Gloumakov, Adam Spiers, and Aaron Dollar. *A Clustering Approach to Categorizing 7 Degree-Of-Freedom Arm Motions During Activities of Daily Living*. 2019.
- [16] Yonghong Huang, Kevin B. Englehart, Bernard Hudgins, and Adrian D.C. Chan. A Gaussian mixture model based classification scheme for myoelectric control of powered upper limb prostheses. *IEEE Transactions on Biomedical Engineering*, 2005.
- [17] IBGE. No Title, 2010.
- [18] Sarthak Jain, Girish Singhal, Ryan J. Smith, Rahul Kaliki, and Nitish Thakor. Improving long term myoelectric decoding, using an adaptive classifier with label correction. *Proceedings of the IEEE RAS and EMBS International Conference on Biomedical Robotics and Biomechatronics*, pages 532–537, 2012.

- [19] Luiz José, Lucas Barbosa, Paulo Roberto, Fernandes De Oliveira, Philippe Dias Araujo, Adson Ferreira, Denis Delisle-rodriguez, John Jairo, Villarejo Mayor, Suélia Rodriguez, Fleury Rosa, and Alberto López-delis. Simultaneous Myoelectric Pattern Recognition using BioPatRec Platform for Hand Prosthesis. *XXVI Brazilian Congress on Biomedical Engineering*, page 869, 2018.
- [20] Mahdi Khezri and Mehran Jahed. A neuro-fuzzy inference system for sEMG-based identification of hand motion commands. *IEEE Transactions on Industrial Electronics*, 2011.
- [21] Yann LeCun, John S Denker, and Sara A. Solla. Optimal Brain Damage (Pruning). *Advances in neural information processing systems*, pages 598–605, 1990.
- [22] Guan-Chun Luh, Yi-Hsiang Ma, Chien-Jung Yen, and Heng-An Lin. Muscle-Gesture Robot Hand Control Based On Semg Signals With Wavelet Transform Features And Neural Network Classifier. 2016.
- [23] McFarland. Unilateral upper-limb loss: Satisfaction and prosthetic-device use in veterans and servicemembers from Vietnam and OIF/OEF conflicts, 2010.
- [24] Ministério da Saúde. No Title, 2015.
- [25] Max Ortiz-Catalan, Rickard Brånemark, and Bo Håkansson. BioPatRec: A modular research platform for the control of artificial limbs based on pattern recognition algorithms. *Source Code for Biology and Medicine*, 2013.
- [26] Max Ortiz-catalan, Bo Håkansson, and Rickard Brånemark. Real-Time and Simultaneous Control of Artificial Limbs Based on Pattern Recognition Algorithms. 22(4):756–764, 2014.
- [27] Max Ortiz-Catalan, Faezeh Rouhani, Rickard Branemark, and Bo Hakansson. Offline accuracy: A potentially misleading metric in myoelectric pattern recognition for prosthetic control. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2015-Novem:1140–1143, 2015.
- [28] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
- [29] Bart Peerdeman, Daphne Boere, Heidi Witteveen, Rianne Huis in ‘tVeld, Hermie Hermens, Stefano Stramigioli, Hans Rietman, Peter Veltink, and Sarthak Misra. Myoelectric forearm prostheses: State of the art from a user-centered perspective. *The Journal of Rehabilitation Research and Development*, 48(6):719, 2011.

- [30] Angkoon Phinyomark, Franck Quaine, Sylvie Charbonnier, Christine Serviere, Franck Tarpin-Bernard, and Yann Laurillau. EMG feature evaluation for improving myoelectric pattern recognition robustness. *Expert Systems with Applications*, 40(12), 2013.
- [31] IEEE Radhika Menon, Student Member, IEEE, Gaetano Di Caterina, Heba Lakany, Member, IEEE, Lykourgos Petropoulakis, Bernard A. Conway, John J. Soraghan, Senior Member. Study on interaction between temporal and spatial information in classification of EMG signals for myoelectric prostheses.
- [32] Danilo Jimenez Rezende, Shakir Mohamed, and Daan Wierstra. Stochastic Back-propagation and Approximate Inference in Deep Generative Models. 2014.
- [33] Aidan D. Roche, Hubertus Rehbaum, Dario Farina, and Oskar C. Aszmann. Prosthetic Myoelectric Control Strategies: A Clinical Perspective. *Current Surgery Reports*, 2(3):44, 2014.
- [34] Oluwarotimi Williams Samuel, Hui Zhou, Xiangxin Li, Hui Wang, Haoshi Zhang, Arun Kumar Sangaiah, and Guanglin Li. Pattern recognition of electromyography signals based on novel time domain features for amputees’ limb motion classification. *Computers and Electrical Engineering*, (April), 2017.
- [35] J. Raudys & Anil K. Jain Sarunas. Small Sample Size Effects in Statistical Pattern Recognition: Recommendations for Practitioners, 1991.
- [36] Jonathon W. Sensinger, Blair A. Lock, and Todd A. Kuiken. Adaptive pattern recognition of myoelectric signals: Exploration of conceptual framework and practical algorithms. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 17(3):270–278, 2009.
- [37] Claude Elwood Shannon. *The Mathematical Theory of Communication by CE Shannon and W. Weaver*. 1949.
- [38] Matija Štrbac, Milica Isaković, Minja Belić, Igor Popović, Igor Simanić, Dario Farina, Thierry Keller, and Strahinja Došen. Short-and long-term learning of feedforward control of a myoelectric prosthesis with sensory feedback by amputees. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2017.
- [39] Tatiana Tommasi, Francesco Orabona, Claudio Castellini, and Barbara Caputo. Improving control of dexterous hand prostheses using adaptive learning. *IEEE Transactions on Robotics*, 29(1):207–219, 2013.
- [40] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan

- Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, CJ Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0—Fundamental Algorithms for Scientific Computing in Python. *arXiv e-prints*, page arXiv:1907.10121, Jul 2019.
- [41] L. N. Vostrova, T. A. Voronina, T. L. Karaseva, S. A. Gernega, E. I. Ivanov, A. M. Kirichenko, and M. Yu Totrova. Synthesis and anticonvulsant activity of benzimidazo[1, 2-c]quinazolines. *Pharmaceutical Chemistry Journal*, 20(6):404–406, 1986.
- [42] Max Welling. Auto-Encoding Variational Bayes arXiv : 1312 . 6114v10 [ stat . ML ] 1 May 2014. (ML):1–14.
- [43] Wu Yi Yang, Sheng Xing Liu, Tai Song Jin, and Xiao Mei Xu. An optimization criterion for generalized marginal Fisher analysis on undersampled problems. *International Journal of Automation and Computing*, 8(2):193–200, 2011.
- [44] Aaron J. Young, Lauren H. Smith, Elliott J. Rouse, and Levi J. Hargrove. Classification of simultaneous movements using surface EMG pattern recognition. *IEEE Transactions on Biomedical Engineering*, 2013.
- [45] Mahdi Zeinali and Masoud Shafiee. A new Levinson–Durbin based 2-D AR model parameter estimation method. *Multidimensional Systems and Signal Processing*, 20(1):35–47, 2014.
- [46] Chiyuan Zhang, Benjamin Recht, Samy Bengio, Moritz Hardt, and Oriol Vinyals. Understanding deep learning requires rethinking generalization. *5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings*, 2019.
- [47] Bo Zong, Qi Song, Martin Renqiang Min, Wei Cheng, Cristian Lumezanu, Daeki Cho, and Haifeng Chen. Deep autoencoding Gaussian mixture model for unsupervised anomaly detection. *6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings*, pages 1–19, 2018.