

Comparative Analysis Among Feature Selection of sEMG Signal for Hand Gesture Classification by Armband

J. Mendes, M. Freitas, H. Siqueira, A. Lazzaretti, S. Stevan, and S. Pichorim

Abstract—This work presents a comparative study between dimensionality reduction and feature selection to classification problem for six hand gestures by sEMG signal. The classified signals are wrist flexion, wrist extension, wrist flexion for the left, wrist extension to the right, forearm supination, and forearm pronation. An armband with eight channels was used to acquire the signals from 13 subjects (8 male and 5 female). Then, 29 features from time and frequency domain were extracted. Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), and Support Vector Machine (SVM) were used as classifiers. Regarding the dimensionality reduction, Principal Component Analysis and LDA were applied in the signal; for feature selection, the feature combination for wrapper method step wise forward was used. The best scenario with dimensionality reduction was obtained with QDA classifier and 80 attributes from PCA, reaching accuracies of 84%. In the second scenario, with 112 attributes (8 features), a non-linear SVM (with Gaussian kernel) reached accuracies of 91%. Both methods presented similar performances among the accuracies for each class; however, dimensionality reduction approach presented less computational cost whilst has a lower accuracy compared with feature selection approach.

Index Terms—sEMG, Classification, Dimensionality reduction, Feature selection, Armband.

I. INTRODUÇÃO

Reconhecimento e a classificação de gestos manuais possuem diversas aplicações relacionadas ao desenvolvimento de interfaces homem-máquina e ao controle de sistemas robóticos [1]–[3]. Por sua vez, a eletromiografia de superfície (sEMG) tem sido usada e explorada como sinal de entrada em aplicações desta natureza. O sinal de sEMG é formado por meio de fenômenos bioelétricos a partir de variações fisiológicas na membrana das células musculares [4].

Para o processamento destes sinais, são necessárias etapas de aquisição, condicionamento e processamento de sinais, as quais podem ser implementadas tanto em sistemas de *software*, quanto de *hardware*. Estes passos devem ser desenvolvidos de forma que garantam alta acurácia em classificação e reconhecimento [5].

Comumente, a aquisição dos sinais de sEMG é feita utilizando poucos canais, procurando medir os principais músculos responsáveis pelos movimentos em que os gestos que se deseja reconhecer pertencem. Para isso, é necessário conhecer a posição dos eletrodos com certa precisão, nos músculos ou grupos musculares específicos [6], [7]. Em aplicações cotidianas, os usuários podem não possuir o conhecimento necessário para alocar os canais nas regiões específicas para a aquisição.

Para contornar este problema, *armbands* têm sido usadas para auxiliar no processo de reconhecimento de padrões entre grupos musculares, facilitando a sua utilização [8]. Utilizando-se de vários canais, é possível melhorar a resolução espacial, a qual facilita a interpretação dos movimentos relacionados ao sinal de sEMG adquirido. Porém, com o aumento do número de canais, aumenta-se a quantidade de atributos de entrada de um classificador.

Como o sEMG possui uma característica estocástica [9], é necessário realizar operações sobre este sinal para extrair informações (chamadas de características ou *features*, do inglês). Esse processo possui dois objetivos principais: reduzir o tamanho da informação a ser enviada para um classificador e eliminar redundâncias. Assim como a quantidade de canais aumenta a quantidade de atributos de entrada, o aumento de *features* ocasiona o mesmo efeito. Existem *features* que são complementares, outras que apresentam a mesma informação repetidamente e outras que prejudicam o processo de classificação se forem combinadas. Então, surge o problema de como encontrar o melhor conjunto de *features* para otimizar o processo de classificação visando a melhor taxa de acerto, ou seja, é preciso realizar a seleção das melhores *features*.

Dentre as abordagens de seleção de características, dois métodos de destacam: a redução de dimensionalidade e a seleção sistemática de *features* pertinentes ao problema. A primeira utiliza técnicas baseadas em álgebra linear para projetar os dados de um espaço de alta dimensionalidade para outro menor, reduzindo a quantidade de atributos para o classificador [10]. Nessa abordagem, encontram-se técnicas como Análise de Componentes Principais (*Principal*

J. J. A. Mendes Junior, Universidade Tecnológica Federal do Paraná, Curitiba, Paraná, Brasil, (e-mail:josej@alunos.utfpr.edu.br).

M. L. B. Freitas, Universidade Tecnológica Federal do Paraná, Ponta Grossa, Paraná, Brasil, (e-mail:mellabanca@alunos.utfpr.edu.br).

H. Siqueira, Universidade Tecnológica Federal do Paraná, Ponta Grossa, Paraná, Brasil, (e-mail:hugosiqueira@utfpr.edu.br).

A. E. Lazzaretti, Universidade Tecnológica Federal do Paraná, Curitiba, Paraná, Brasil, (e-mail:lazzaretti@utfpr.edu.br).

S. L. Stevan Jr, Universidade Tecnológica Federal do Paraná, Ponta Grossa, Paraná, Brasil, (e-mail:sstevanjr@utfpr.edu.br).

S. F. Pichorim, Universidade Tecnológica Federal do Paraná, Curitiba, Paraná, Brasil, (e-mail:pichorim@utfpr.edu.br).

Component Analysis, PCA) e Análise de Discriminantes Lineares (*Linear Discriminant Analysis*, LDA) no processamento de sinais de sEMG [11].

Por sua vez, a seleção de um conjunto de características consiste em encontrar as *features* redundantes ou irrelevantes para o sistema. Como envolve uma combinação de características, é necessário realizar uma abordagem sistêmica afim de que a redução de características possa apresentar uma vantagem. Técnicas como *wrapper* e metaheurísticas são frequentemente usadas [10].

Nessa perspectiva, este trabalho apresenta uma análise comparativa entre a utilização destas duas abordagens durante a classificação de sinais de sEMG provenientes de gestos manuais. Ao todo, 29 características foram extraídas do sinal de sEMG proveniente de seis gestos manuais adquiridos dos músculos do antebraço por meio de uma *armband* de múltiplos canais, visando fornecer maior quantidade de dados para embasar essa análise. Os classificadores escolhidos foram a LDA, a Análise dos Discriminantes Quadráticos (*Quadratic Discriminant Analysis*, QDA) e Máquinas de Vetores Suporte (*Support Vector Machine*, SVM) devido a sua robustez quanto ao reconhecimento de padrões em aplicações similares [12]–[14]. Como técnicas de redução de dimensionalidade, foram escolhidas a PCA e a LDA; e para a seleção de conjunto de *features*, foi utilizado o algoritmo de *wrapper forward stepwise*.

Este trabalho destaca essa comparação, pois em muitos sistemas de classificação de sinais de sEMG, uma abordagem é escolhida em detrimento da outra (especialmente a seleção de características) [12], [13], [15], [16], não analisando as vantagens que cada uma possui para aplicações deste gênero, especialmente no desenvolvimento de dispositivos especializados em reconhecimento de gestos. Reforça-se que quando esta comparação é feita, o posicionamento dos eletrodos é fixo [17]. Com o uso de uma *armband* (como a proposta deste trabalho), os canais são distribuídos de forma equidistante ao redor do antebraço.

Este artigo é dividido da seguinte forma: a Seção II apresenta os materiais e métodos envolvidos na comparação; a Seção III apresenta e discute os resultados obtidos, enquanto a Seção IV conclui o presente estudo.

II. MATERIAIS E MÉTODOS

A. Aquisição, Pré-processamento e Extração de Características

A *armband* utilizada neste trabalho possui 8 canais de aquisição, banda de frequência de 10-500 Hz e frequência de amostragem de 2 kHz e uma resolução de 16 bits para a aquisição dos dados [8]. Oito pares de eletrodos de superfície de Prata/Cloreto de Prata (Ag/AgCl) foram alocados na região do antebraço direito (com dois eletrodos por canal) e um eletrodo de referência foi alocado na região do cotovelo.

Para manter um padrão durante as aquisições, a *armband* foi posicionada no meio do antebraço, tendo como referência para o primeiro canal o ventre muscular do músculo flexor do carpo ulnar (*flexor carpi ulnaris*). A Figura 1 apresenta o sistema de aquisição e sua posição na região do antebraço. Os circuitos e o

processo de aquisição estão de acordo com o Comitê de Ética de Pesquisa em Seres Humanos da Universidade Tecnológica Federal do Paraná, registrado sobre CAAE: 89638918.0.0000.5547.



Fig. 1. Ilustração do posicionamento da *armband* de oito canais, instalada no antebraço, contendo a indicação do primeiro canal (CH₁), localizado no músculo do flexor do carpo ulnar.

O protocolo para a aquisição de dados foi feito utilizando um metrônomo com uma frequência de 1,33 Hz (45 batidas por minuto), em que cada voluntário realizou seis movimentos na respectiva sequência: flexão do punho, extensão do punho, flexão do punho para a direita, flexão do punho para a esquerda, supinação do antebraço e pronação do antebraço (apresentados na Figura 2). Estes gestos foram escolhidos pois recrutam grupos musculares similares na sua execução, o que ocorre entre gestos representados nas imagens da Figura 1 (a)-(c), (b)-(d) e (e)-(f), afim de avaliar melhor o sistema de classificação na discriminação de gestos com recrutamentos de classes similares. Além disso, são gestos simples e naturais, os quais indivíduos saudáveis não possuem dificuldade ao realizá-los.

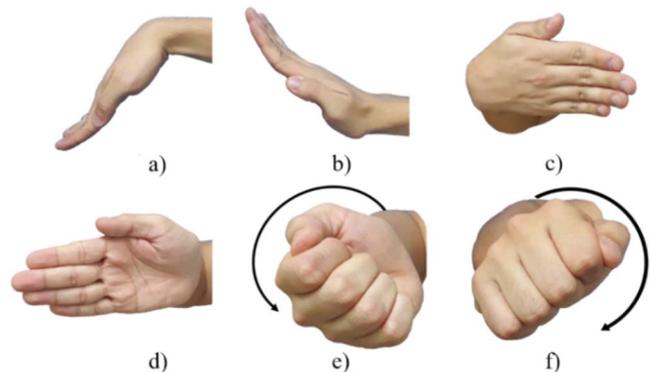


Fig. 2. Gestos desenvolvidos pelos voluntários: a) flexão de punho, b) extensão de punho, c) flexão do punho para a esquerda, d) flexão do punho para a direita, e) supinação do antebraço e f) pronação do antebraço.

Entre cada movimento, o voluntário permanecia em uma posição neutra, relaxando os músculos, deixando-os em repouso. Cada voluntário desenvolveu 50 vezes cada movimento, sendo 300 gestos adquiridos. Como participaram 13 voluntários saudáveis (8 do sexo masculino e 5 do sexo feminino, média de idade de $23,6 \pm 1,5$ anos), foi gerado um *dataset* com 3900 movimentos.

O condicionamento de sinais foi feito por meio do *software* de instrumentação virtual LabVIEW™. Foi usado nesta etapa um filtro passa-banda de quarta ordem, aproximação *Butterworth*, com frequências de corte entre 10 e 450 Hz. Esse foi utilizado para delimitar a faixa de frequências do sinal de

sEMG, atendendo a banda espectral requerida [18]. Também foi implementado um filtro rejeita-faixa (*notch*) de sexta ordem de aproximação *Butterworth*, com frequência central de 60 Hz, sendo este usado para atenuar interferências geradas pela rede de alimentação [18]. As harmônicas de 60 Hz não foram observadas durante o experimento (como 120 Hz e 180 Hz) e um filtro nessa região poderia atenuar frequências de grande energia do sinal de sEMG. Além disso, o amplificador do estágio de entrada do sinal possui alta taxa de rejeição de modo comum, atenuando ruídos nessa natureza.

Os dados foram segmentados por meio da técnica *onset*, a qual detecta o período ativo do sinal de sEMG [19]. Após a detecção automática dos instantes de ativação (usando valor de disparo de 20% do valor de pico do sinal normalizado de cada indivíduo para um período maior de 500 amostras), os dados foram segmentados em janelas de 2000 amostras, o equivalente a 1 segundo. Este processo é apresentado na Figura 3.

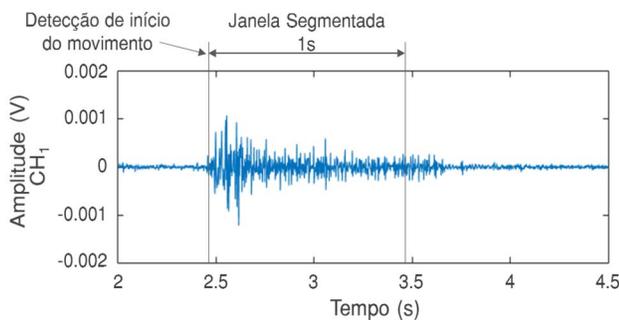


Fig. 3. Exemplo de segmentação e janelamento do sinal de sEMG utilizado. Após a detecção do sinal pelo *onset*, os dados são segmentados em janelas de 2000 amostras, contendo o equivalente a 1 segundo.

Foram escolhidas 29 características do sinal de sEMG, sendo estas extraídas no domínio do tempo e no domínio da frequência [13], [20]. A Tabela I apresenta as características utilizadas, separadas por domínio. No total, os dados são fornecidos para as rotinas de classificação como uma matriz de 3900 (300 x 13 movimentos) por 280 (atributos). Vale ressaltar que cada característica possui um valor correspondente aos 8 canais de aquisição de sEMG para cada amostra. Além disso, as características de coeficientes autorregressivos de 4^o ordem (AR4) e coeficientes cepstrais (CEPS) possuem como saída 4 valores cada uma. Cada vetor de características foi normalizado entre os limites de -1 e 1.

B. Técnicas de Classificação e de Redução de Dimensionalidade

Dentre as técnicas de redução de dimensionalidade e projeção dos dados de sEMG, encontram-se a Análises de Componentes Principais (PCA) e a Análise de Discriminantes Lineares (LDA) [21]. A PCA é uma técnica de redução de dimensionalidade não supervisionada. Os dados de entrada são transformados ortogonalmente, buscando, nessa nova projeção, que os eixos não estejam linearmente correlacionados (aos quais, para cada eixo, dá-se o nome de componente principal) [22]. Dessa forma, busca-se a menor quantidade de vetores principais, os quais, por sua vez, devem possuir a maior variância na nova projeção dos dados [16] [17].

TABELA I
CARACTERÍSTICAS EXTRAÍDAS NOS DOMÍNIOS DO TEMPO E FREQUÊNCIA DO SINAL DE sEMG

Domínio	Sigla	Característica
Tempo	WL	Comprimento da Forma de Onda
	SSC	Alteração do gradiente do sinal.
	MAV	Valor Médio Absoluto
	DASDV	Valor da diferença absoluta do desvio padrão
	RMS	Valor da Raiz Quadrada Média
	IEMG	Integral do sinal
	ZC	Cruzamento de zeros
	SampEn	Entropia amostral
	AR4	Coefficientes Autor-regressivos de 4 ordem
	CEPS	Coefficientes Cepstrais
	VAREMG	Variância
	WAMP	Amplitude de Williamson
	MAV1 e MAV2	Valor Médio Absoluto Modificado 1 e 2
	VORDER	Ordem V
	MYOP	Valor de Myopulse
LOGDEC	Detector Logarítmico	
Frequência	TM3, TM4 e TM5	Valor absoluto dos 3 ^o , 4 ^o e 5 ^o momentos
	MNP	Média do Espectro de Potência
	MNF	Frequência Média
	MDF	Frequência Mediana
	FR	Taxa de Frequência
	PKF	Pico de Frequência
	SM1, SM2 e SM3	Momentos Espectrais
	TTP	Potência total do espectro

A Análise de Discriminantes Lineares (LDA) é um método utilizado para reduzir a alta dimensionalidade de um conjunto de dados maximizando a relação entre a variância entre classes e minimizando a variância dos dados da própria classe [24]. Essa técnica interpreta os dados em um espaço n -dimensional e divide as classes em um hiperplano com $n-1$ dimensões, maximizando a separabilidade entre elas [25] e pode ser usada tanto como um redutor de dimensionalidade quanto como classificador. Sua principal vantagem é auxiliar no entendimento da distribuição dos dados, sendo um processo supervisionado.

Encontrar o melhor conjunto de *features* faz parte de um problema de busca, sendo o *wrapper* uma das formas de solucioná-lo. O método *wrapper* escolhido para este trabalho seleciona de forma sequencial o melhor conjunto de características a partir de um conjunto vazio [26]. Este conjunto cresce à medida que a combinação entre os atributos contribui para o processo de classificação, sendo chamado de *forward stepwise*. De forma simplificada, analisa-se a contribuição individual de cada *feature* para o classificador e seleciona um ganhador; sendo feito posteriormente o mesmo processo para a característica vencedora e a combinação desta para todas as demais [27]. Uma das grandes desvantagens desse método é o seu alto custo computacional, além de que o resultado final é uma aproximação da solução ótima [26]. Porém, é uma alternativa quando se tem uma grande quantidade de *features* a serem analisadas. A vantagem principal é que o resultado está diretamente relacionado ao classificador, já que ele precisa ser executado para cada entrada.

C. Sistemas Classificadores

Com relação às técnicas de classificação, LDA também é usada para reconhecimento de padrões, sendo indicada em casos onde as frequências dentro das classes são desiguais. Neste trabalho, para a implementação da LDA, foi assumido que as observações de cada classe são feitas por distribuições gaussianas multivariadas, contendo um vetor de médias específicos para cada classe e uma matriz de covariância comum a todas as classes [18] [22]. Por sua vez, o método de classificação pela Análise de Discriminantes Quadrática (QDA) é uma variação da LDA, na qual as fronteiras de decisão formam superfícies não-lineares, ou seja, quadráticas. Estas são determinadas pela variação do cálculo dos discriminantes de Fischer [24], [29].

As Máquinas de Vetores Suporte (SVMs) são um grupo de algoritmos classificadores. A partir de um conjunto de atributos de entrada, uma SVM prediz o pertencimento ou não destes atributos a uma determinada classe, mapeando os dados em um espaço de maior dimensão que a dos atributos de entrada, procurando por um hiperplano ótimo de separação entre classes [30]. Seu treinamento é baseado em um processo de otimização, o qual deseja-se encontrar o hiperplano que separa as classes contendo a maior margem de separação. Essas separações podem ser feitas de forma linear (SVM Linear), ou usando funções de *Kernel*, como funções de base radial gaussiana (*Radial Basis Function*, RBF) ou polinomiais [30], [31]. Como um problema de otimização, alguns parâmetros devem ser ajustados, como parâmetro de regularização (parâmetro C) e largura das funções de base radial (funções de *Kernel*) [31], sendo estas delimitadas no presente estudo.

D. Metodologia para a Classificação

Todas as rotinas de classificação e redução de dimensionalidade foram implementadas no *software* MATLAB[®]. Foi utilizado um computador com processador Dual Core i5-4210U. Ressalta-se que em todos os testes dos classificadores, os dados passaram pelo processo de validação cruzada pelo método de *k-fold*, sendo $k=10$. Dessa forma, os dados foram separados entre 90% para o conjunto de treinamento e 10% para o conjunto de teste. Para os classificadores do tipo SVM, foram escolhidos 10% do conjunto de treinamento para a validação dos seus parâmetros.

O primeiro modelo baseou-se nas análises dos redutores de dimensionalidades, já que os dados de entrada dos classificadores possuem 280 descritores. Os dados foram enviados para o LDA, QDA e SVM com fronteira de decisão linear (SVM Linear) utilizando abordagem *one-vs-all*. Após analisar o vencedor dos três, decidiu-se aplicar as transformações PCA e LDA para este e verificar as melhores taxas de acerto para os melhores casos.

A segunda análise consistiu em escolher o melhor grupo de características do sinal de sEMG utilizando o método *wrapper* pelo *forward stepwise*. Para esse caso, optou-se por utilizar os métodos de LDA e QDA devido ao seu baixo custo computacional e alta taxa de acerto com robustez em relação a trabalhos similares [32]. A rotina do *wrapper* foi repetida 20 vezes e foram analisadas as acurácias para os dois classificadores, tomando os maiores valores de acurácia com a menor quantidade de atributos. Definida a quantidade de atributos ótima, foi analisada a frequência de ocorrência destes

dados em um histograma contendo as n características para todas as 20 repetições. Como critério de seleção, analisaram-se as características que se repetem com uma frequência maior ou igual a 50% de ocorrência. Tendo o grupo estabelecido, foram analisados os desempenhos dos classificadores LDA, QDA e SVM com decisão linear e SVM não-linear com função de *kernel* RBF. Os parâmetros variados para a validação das SVMs são apresentados na Tabela II.

TABELA II
PARÂMETROS AVALIADOS PARA OS CLASSIFICADORES DO TIPO SVM

Param SVM	Valores								
C	0,01	0,05	0,1	0,5	1	5	10	20	100
Kernel (apenas para RBF)	0,01	0,1	0,5	1	5	10	50	100	

III. RESULTADOS E DISCUSSÃO

Com relação aos dados obtidos dos voluntários, a Figura 4 apresenta um exemplo de uma sequência de aquisição para os oito canais. Os seis gestos são apresentados na sequência de (a) a (f) da Figura 2. Por meio da Figura 4, pode-se perceber a distinção entre os grupos musculares adquiridos pela *armband* para cada gesto, além da característica do sinal de sEMG (baixas amplitudes e natureza estocástica).

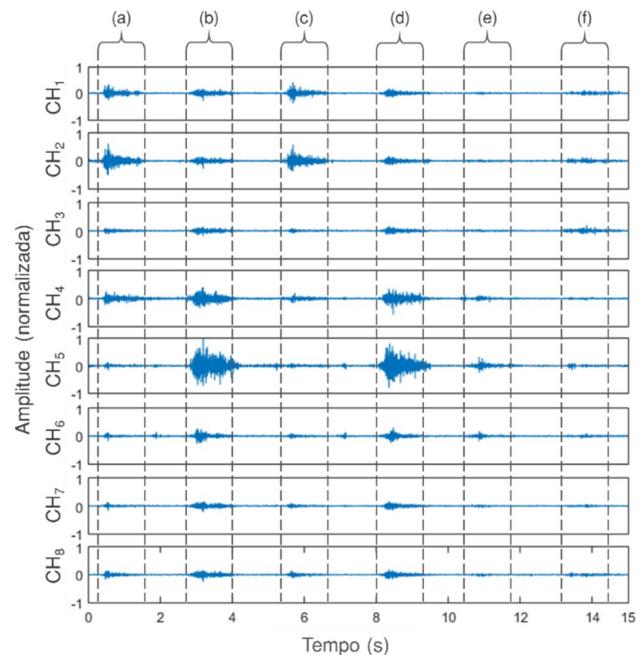


Fig. 4. Exemplo de sinal adquirido de um voluntário para os oito canais de aquisição. Nessa sequência, são apresentados os respectivos movimentos nas janelas de 1,33 segundos: a) flexão de punho, b) extensão de punho, c) flexão do punho para a esquerda, d) flexão do punho para a direita, e) supinação da mão e f) pronação da mão.

A. Abordagem por Redução de Dimensionalidade

A Figura 5 apresenta os resultados obtidos para as técnicas de redução de dimensionalidade. Nota-se na Figura 5a que dentre os três primeiros classificadores (sem a utilização de métodos de redução de dimensionalidade), o SVM linear

possuiu a menor das taxas de acerto que as demais, variando em torno de $55\pm 3\%$. Esse foi o melhor valor encontrado utilizando as faixas de variação de C , que nesse caso, situou-se em 10. Além de apresentar a menor acurácia, o SVM apresentou o maior tempo de treinamento. Por sua vez, os classificadores QDA e LDA chegaram a taxas de acerto próximas para a apresentação de todos os 280 atributos de entrada, com uma acurácia pouco maior para o QDA ($83\pm 1\%$).

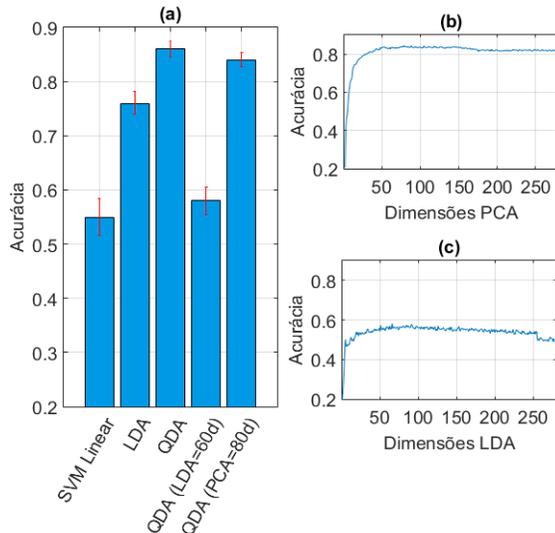


Fig. 5. Resultados obtidos pelos métodos de redução de dimensionalidade: a) refere-se as taxas de acerto dos classificadores, b) as acurácias do classificador QDA com o incremento do vetor de coeficientes da PCA e c) as acurácias do classificador QDA com o incremento do vetor de coeficientes da LDA.

Com o desempenho ligeiramente maior que o classificador LDA, o classificador QDA foi submetido à análise de redução de dimensionalidade pela PCA e LDA. As Figuras 5b e 5c apresentam as acurácias obtidas do classificador QDA em função do aumento das dimensões após a aplicação da PCA e da QDA. Neste caso, percebe-se que a técnica PCA possui maior taxa de acerto (acima de 0,8 para 25 componentes ou mais) que a LDA (próximas a 0,6 acima das 25 dimensões).

Ambas as técnicas apresentaram um platô acima de um certo número de componentes e dimensões, o qual decresce após o aumento das dimensões. Visando reduzir a quantidade de atributos, as dimensões escolhidas foram formadas pela combinação entre a maior taxa de acerto antes que houvesse o decréscimo da sua acurácia. Sendo assim, percebeu-se que os valores que satisfaziam essas condições representavam 60 dimensões para a LDA e 80 componentes para a QDA. Estes dados são apresentados na Figura 5a comparativamente com os classificadores sem a redução de dimensionalidade. A maior taxa de acerto pertenceu à redução por PCA, com $84\pm 1\%$ em comparação a LDA, com $58\pm 2\%$.

Nota-se que a LDAr (como redução de dimensionalidade) não apresentou vantagens para esses conjuntos de gestos manuais. Em compensação, a PCA demonstrou uma ligeira melhora na acurácia com uma diminuição significativa da dimensionalidade (80 componentes em comparação aos 280). A soma destes componentes representou cerca de 99,5% da

variância dos dados. Estes resultados estão sintetizados na Tabela III.

TABELA III
RESULTADO OBTIDO PELOS CLASSIFICADORES PELA ANÁLISE DE REDUÇÃO DE DIMENSIONALIDADE

Classificador	Acurácia
SVM Linear (C=10)	55,3±3,4%
LDA	81,2±2,1%
QDA	83,5±1,5%
QDA (LDAr 60 dim)	58,2±2,6%
QDA (PCA 80 dim)	84,4±1,3%

B. Abordagem por Grupo de Características

Com relação a segunda abordagem (a escolha do melhor grupo de características), a Figura 6 apresenta o resultado obtido pelo método *wrapper* para os métodos de classificação QDA e LDA. Nota-se uma diferença de comportamento para estas duas curvas. As acurácias para o classificador QDA possuem valor de pico (próximo a 0,89) e começam a decrescer, atingindo valores próximos aos obtidos com a análise anterior (sem a aplicação das técnicas de redução de dimensionalidade).

Por outro lado, os resultados obtidos para o classificador LDA apresentaram um crescimento na acurácia com o aumento da quantidade de características (*features*), possuindo pequenas variações a partir de 20 características (estabilizando em 0,82). Logo, para esse problema, o classificador LDA necessita de outras características de sEMG para poder aumentar a sua taxa de acerto, as quais podem ser complexas a ponto de aumentar o tempo de processamento. Além disso, foi realizado um teste de correlação pareado (Wilcoxon) para cada *feature*, excluíram-se as que possuíam maior correlação (mantendo-se apenas WAMP, VORDER, ZC, SampEn e TTP) e mesmo assim, obteve-se uma acurácia de 0,62. Dessa forma, percebe-se que a utilização de *features* correlatas auxilia no processo de classificação da LDA.

Analisando os dados visualizados pelo classificador QDA da Figura 6a, percebeu-se que entre 6 e 10 características não há diferença significativa para os valores encontrados.

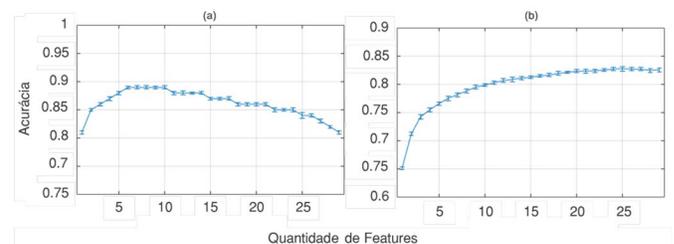


Fig. 6. Média das acurácias e seus respectivos desvios padrões obtidos pelos classificadores (a) QDA e (b) LDA pelo aumento da quantidade de características (*features*) utilizando o método *wrapper*.

Sendo assim, escolheu-se o valor das 10 primeiras características obtidas pelo método *wrapper* recolhido das 20 iterações. Com estes dados, foi formado o histograma presente da Figura 7. Logo, as características que satisfazem a condição de se repetirem com uma frequência de 50% ou mais foram: “WL”, “SSC”, “DASDV”, “ZC”, “AR4”, “CEPS”, “WAMP” e

“MYOP”. Assim, 8 características foram usadas dentre as 29 extraídas previamente, possuindo ao total, 112 atributos.

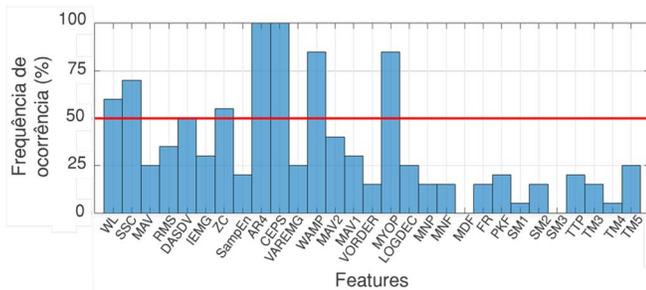


Fig. 7. Histograma de frequências para as 10 primeiras características resultantes do método wrapper.

A próxima etapa foi analisar o desempenho dos classificadores para esse conjunto reduzido de características. Porém, antes da comparação, é necessário verificar os melhores parâmetros para SVM. Para o caso Linear, o único parâmetro ajustável é o C , enquanto que para o SVM RBF, além deste, analisou-se também a abertura da gaussiana ($Kernel$). Com relação ao SVM Linear, o melhor resultado para o parâmetro C foi observado para este tendo o valor igual a 1. Com o decréscimo e aumento deste parâmetro (verificado de forma manual), percebe-se que não havia ganhos para este tipo de classificador para essa base de dados.

Com relação a SVM RBF, verificou-se a mesma tendência, sendo apresentada na Figura 8. Analisando a Figura 8, percebe-se que tanto o parâmetro C quanto a função de $Kernel$ gaussianas influenciam nos valores da acurácia. Nota-se que a região com maiores classificações está situada com os valores de $Kernel$ s entre 0,5 e 10 da Tabela II. Dentre eles, o valor de $Kernel$ igual a 1 possui as maiores taxas de acerto. Percebe-se que seu aumento reflete em uma maior taxa de acerto, já que há uma diminuição do tamanho das margens de decisão do SVM. Valores de C acima de 1 (Tabela II) apresentaram taxas de acerto com variações não significativas. Por isso, o parâmetro foi escolhido de forma que obtivesse maior acurácia com melhor desvio padrão. Nesse caso, a configuração da SVM RBF com o parâmetro igual a 5 foi escolhida.

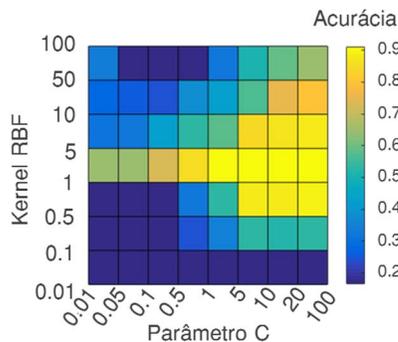


Fig. 8. Acurácias para a SVM RBF em função dos parâmetros C e da função de $Kernel$ RBF.

Por fim, a Tabela IV apresenta os resultados obtidos para todos os classificadores nessa abordagem. Pode-se perceber que a SVM não-linear com função RBF possui a maior taxa de

acerto para esse conjunto de dados, com acurácia de $91\pm 1\%$.

TABELA IV
RESULTADO OBTIDO PELOS CLASSIFICADORES PELA ANÁLISE DE REDUÇÃO DE GRUPO DE CARACTERÍSTICAS

Classificador	Acurácia
SVM Linear ($C=1$)	$53,2\pm 5,1\%$
SVM RBF ($C=5$, $Kernel=1$)	$90,9\pm 1,3\%$
QDA	$85,4\pm 0,3\%$
LDA	$78,4\pm 0,2\%$

C. Comparação entre Redução de Dimensionalidade e Redução de Grupo de Características

Como forma de comparar o desempenho dos dois métodos entre os movimentos classificados, as matrizes de confusão dos seis gestos em cada caso são apresentadas na Figura 9. Além disso, as Tabelas V e VI apresentam os dados referentes as métricas da matriz de confusão, sendo analisado a sua acurácia, sensibilidade, especificidade e eficiência.

Na Figura 9, a primeira semelhança entre as matrizes concentra-se nas regiões em que ocorrem os acertos e erros, os quais são, em sua maioria, nos movimentos que recrutam grupos musculares semelhantes. Por exemplo, o gesto ‘a’ é erroneamente classificado com o ‘c’ devido aos grupos musculares próximos na sua ativação. Igualmente, isso ocorre entre os gestos ‘b’ e ‘d’ e com os gestos ‘e’ e ‘f’.

		(a) Obtido						(b) Obtido					
		a	b	c	d	e	f	a	b	c	d	e	f
Desejado	a	98,3		1,7				96,6		3,4			
	b		75,9		22,8	1,3			92,2		7,8		
	c	10,4		89,6				9,5		85,1	4,1	1,4	
	d		6		94				9,5		90,5		
	e	1,3		3,9		80,3	14,5					95,4	4,6
	f		3,3	1,7		5	90					3,1	95,4

Fig. 9. Matrizes de Confusão para os seis gestos para os melhores pelas abordagens de (a) redução de dimensionalidade (QDA com PCA e 80 dimensões) e de (b) redução de características (SVM RBF com 8 características). Todos os valores estão em porcentagem (%).

Para os dados referentes à classificação com redução de dimensionalidade (Figura 9a), os gestos com menores taxas de acerto foram os gestos ‘b’ e ‘e’, os quais foram mais classificados erroneamente com os gestos ‘f’ e ‘d’, respectivamente. Isso reflete na sensibilidade do classificador, o qual mostra a dificuldade do mesmo em reconhecer corretamente os gestos quando estes são verdadeiros, afetando a eficiência para essas classes (o qual pode ser visto na Tabela V). Todos os demais possuíram taxas acima de 89% e altos valores de eficiência.

TABELA V
MÉTRICAS DA MATRIZ DE CONFUSÃO DOS RESULTADOS OBTIDOS PELO CLASSIFICADOR PELO MÉTODO DE REDUÇÃO DE DIMENSIONALIDADE

Parâmetro (%)	Classes					
	a	b	c	d	e	f
Acurácia	97,4	93,4	96,6	94,2	94,7	94,7
Sensibilidade	98,2	75,9	89,6	94	80,3	90
Especificidade	97,2	98,2	98,2	94,1	98,6	96,3
Eficiência	97,7	87,05	93,9	94,05	89,45	93,15

Sobre os dados advindos pelo método de reduzir o grupo de características, a matriz de confusão da Figura 9b, o único gesto com taxa de classificação abaixo de 90% foi o gesto ‘c’, confundido em sua maioria com o gesto ‘a’. Os demais gestos apresentaram altos valores de acurácia e eficiência, sempre acima de 90% (conforme a Tabela VI).

TABELA VI
MÉTRICAS DA MATRIZ DE CONFUSÃO DOS RESULTADOS OBTIDOS PELO CLASSIFICADOR PELO MÉTODO DE REDUÇÃO DE GRUPO DE CARACTERÍSTICAS

Parâmetro (%)	Classes					
	a	b	c	d	e	f
Acurácia	97,3	97	96,5	96,3	98,4	98,4
Sensibilidade	96,6	92	85,1	90,5	95,4	95,4
Especificidade	97,4	98	99,3	97,4	99	99
Eficiência	97	95	92,2	93,95	97,2	97,2

Pode-se notar que entre as duas abordagens, o classificador com SVM RBF com o grupo reduzido de características obteve maior separabilidade entre as classes. Entretanto, o classificador demandou alto custo computacional de treinamento entre os demais classificadores testados, além de utilizar 114 atributos no vetor de entrada. Por outro lado, combinar a PCA com 80 componentes e o classificador QDA, obteve-se um desempenho pouco inferior com menor tempo de treinamento que a SVM RBF e uma menor quantidade de atributos no vetor de entrada.

A última avaliação dos métodos foi feita por meio do tempo de execução dos algoritmos, visando analisar o seu custo computacional. Estes dados são apresentados na Tabela VII, divididos entre o tempo de treinamento dos modelos e o tempo de teste para cada amostra. A primeira grande observação se deve a disparidade de tempo entre o método de seleção de *wrapper* em relação com as técnicas de redução de dimensionalidade. Mesmo sendo a técnica que anteriormente apresentou maior acurácia, seu custo computacional é elevado, porém, após execução, não há mais preocupação com a escolha das *features*. Além disso, na etapa de teste, não é necessário realizar cálculos adicionais, como transformações, como é o caso da PCA.

TABELA VII
MEDIDAS DE TEMPO DE EXECUÇÃO ENTRE CADA MÉTODO AVALIADO

Método	Treinamento (s)	Execução por amostra (ms)
Wrapper	410	-
PCA	0,3071	0,168
LDAr	0,48	8
LDA	0,048	0,0015
QDA	0,0945	0,0052
SVM Linear	445	0,01
SVM RBF	649,5	50

Com relação aos classificadores, as SVMs apresentaram maior tempo de processamento durante o treinamento e teste. Contudo, ao embarcar o classificador em um microcontrolador, leva-se em consideração o tempo de execução por amostra. Sendo assim, o tempo de processamento para a SVM RBF, que possuiu as maiores taxas de acerto, é da ordem de 50 ms. Dessa forma, considerando o desenvolvimento de todas as etapas em

uma aplicação embarcada (com a escolha das características já realizadas) a seleção de *features* pode ser implementada sem grandes atrasos. Porém, se o projetista desejar tempos menores de processamento, a abordagem por redução de características apresenta menor custo computacional ao passo que há uma diminuição de acurácia para alguns gestos.

Com relação trabalhos similares, reforça-se que a comparação entre métodos de redução de dimensionalidade e de seleção de características em sEMG é escassa devido a escolha de uma abordagem em detrimento a outra. Porém, comparando com o conjunto de gestos selecionados, este trabalho apresenta uma otimização a [32], que utiliza o mesmo conjunto de gestos e abordagem de *armband* de múltiplos canais, mantendo a mesma acurácia com menor quantidade de atributos ao classificador.

IV. CONCLUSÃO

Este trabalho apresentou a comparação de métodos de seleção de atributos na classificação de seis gestos manuais por meio de sEMG. Ao total, 29 características foram extraídas do sinal e as duas abordagens foram utilizadas para diminuir a quantidade de entradas em um classificador: redução de dimensionalidade e seleção do conjunto de características. Mesmo escolhendo-se gestos com alta similaridade entre grupos musculares, os classificadores obtiveram êxito em suas taxas de acerto, com acurácias acima de 80%.

Pela abordagem de redução de dimensionalidade, os melhores resultados foram obtidos por meio do classificador QDA, sendo a entrada deste um vetor de 80 atributos resultantes dos 80 primeiros componentes de uma PCA, com taxas de acerto média de $84,4 \pm 1,3\%$. Os resultados obtidos pela redução do conjunto de características apontaram a SVM com função de RBF ($C=5$, $Kernel=1$) como o melhor resultado, com taxas de acerto de $90,9 \pm 1,3\%$. Porém, essa última abordagem possui um vetor de entrada com 112 atributos. As classes obtiveram taxas de acerto acima de 75%, não havendo sobreposição entre classes para ambas as técnicas. Cada abordagem possuiu diferentes gestos com menor taxa de classificação (gesto ‘b’ para a primeira e gesto ‘c’ para a segunda). Por fim, pode-se analisar que mesmo com uma menor taxa de acerto, menor quantidade de atributos de entrada e menor tempo de execução, a abordagem com redução de dimensionalidade obteve resultados próximos à análise de redução de características (diferença cerca de 6%) pelo *wrapper*.

Trabalhos futuros estão em analisar a relação entre diferentes classificadores (como Redes Neurais Artificiais e SVMs com *Kernels* polinomiais), outros métodos de redução de dimensionalidade, verificando o seu impacto, uso de algoritmo genético para otimização dos parâmetros de SVM e das características e a análise taxa de acerto do sistema por usuário para classificação *online* de gestos.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

REFERÊNCIAS

- [1] E. Akan, H. Tora, and B. Uslu, "Hand gesture classification using inertial based sensors via a neural network," 2017, pp. 140–143.
- [2] A. D. de Sarro Junior, J. J. A. Mendes Junior, and S. H. Frantz, "Controle de um Braço Robótico através de Eletromiografia," trabalho de Conclusão de Curso (Tecnologia em Automação Industrial), Universidade Tecnológica Federal do Paraná, Ponta Grossa, Paraná, 2014.
- [3] D. R. Freer, J. Liu, and G. Yang, "Optimization of EMG movement recognition for use in an upper limb wearable robot," in *2017 IEEE 14th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, 2017, pp. 202–205.
- [4] M. Z. Jamal, "Signal Acquisition Using Surface EMG and Circuit Design Considerations for Robotic Prosthesis," in *Computational Intelligence in Electromyography Analysis - A Perspective on Current Applications and Future Challenges*, InTech, 2012.
- [5] J. J. A. M. Junior, M. B. Pires, S. Okida, and S. L. Stevan, "Robotic Arm Activation using Surface Electromyography with LABVIEW," *IEEE Lat. Am. Trans.*, vol. 14, no. 8, pp. 3597–3605, Aug. 2016.
- [6] A. Saikia, S. Mazumdar, N. Sahai, S. Paul, and D. Bhatia, "Comparative study and feature extraction of the muscle activity patterns in healthy subjects," 2016, pp. 147–151.
- [7] S. Mazumdar, A. Saikia, N. Sahai, S. Paul, and D. Bhatia, "Determination of significant muscle in movement of upper limb using maximum voluntary contraction of EMG signal," in *2017 4th International Conference on Signal Processing and Integrated Networks (SPIN)*, 2017, pp. 96–99.
- [8] J. J. A. Mendes Junior, M. B. Pires, M. E. M. Vieira, S. Okida, and S. L. Stevan Jr., "Desenvolvimento de armband com fusão de sEMG e giroscópio para identificação de grupos musculares do braço," in *Anais do XXV Congresso Brasileiro de Engenharia Biomédica*, Foz do Iguaçu, Brasil, 2016, pp. 528–531.
- [9] K. Najarian and R. Splinter, *Biomedical signal and image processing*, 2nd ed. Florida, US: CRC Press, 2012.
- [10] P.-N. Tan, M. Steinbach, and V. Kumar, *Introdução ao datamining: mineração de dados*. Ciência Moderna, 2009.
- [11] R. H. Chowdhury, M. B. I. Reaz, M. A. B. M. Ali, A. A. A. Bakar, K. Chellappan, and T. G. Chang, "Surface Electromyography Signal Processing and Classification Techniques," *Sensors*, vol. 13, no. 9, pp. 12431–12466, Sep. 2013.
- [12] A. Phinyomark, R. N. Khushaba, and E. Scheme, "Feature Extraction and Selection for Myoelectric Control Based on Wearable EMG Sensors," *Sensors*, vol. 18, no. 5, p. 1615, May 2018.
- [13] A. Phinyomark, P. Phukpattaranont, and C. Limsakul, "Feature reduction and selection for EMG signal classification," *Expert Syst. Appl.*, vol. 39, no. 8, pp. 7420–7431, Jun. 2012.
- [14] K. S. Kim, H. H. Choi, C. S. Moon, and C. W. Mun, "Comparison of k-nearest neighbor, quadratic discriminant and linear discriminant analysis in classification of electromyogram signals based on the wrist-motion directions," *Curr. Appl. Phys.*, vol. 11, no. 3, pp. 740–745, May 2011.
- [15] A. Phinyomark, H. Hu, P. Phukpattaranont, and C. Limsakul, "Application of Linear Discriminant Analysis in Dimensionality Reduction for Hand Motion Classification," *Meas. Sci. Rev.*, vol. 12, no. 3, pp. 82–89, Jan. 2012.
- [16] Z. Yan and Z. Liu, "The study on feature selection strategy in EMG signal recognition," *2013 ICME Int. Conf. Complex Med. Eng.*, pp. 711–716, 2013.
- [17] J. Liu, "Feature dimensionality reduction for myoelectric pattern recognition: A comparison study of feature selection and feature projection methods," *Med. Eng. Phys.*, vol. 36, no. 12, pp. 1716–1720, Dec. 2014.
- [18] S. L. Stevan Jr, J. J. A. M. Jr, F. C. Janzen, E. A. Jr, and S. Okida, "Circuit Design for Surface Electromyography Data Acquisition," *IEEE Lat. Am. Trans.*, vol. 13, no. 10, pp. 3193–3200, Oct. 2015.
- [19] G. Kamen and D. Gabriel, *Fundamentos da Eletromiografia*. São Paulo: Phorte Editora, 2015.
- [20] N. Nazmi, M. A. Abdul Rahman, S.-I. Yamamoto, S. A. Ahmad, H. Zamzuri, and S. A. Mazlan, "A Review of Classification Techniques of EMG Signals during Isotonic and Isometric Contractions," *Sensors*, vol. 16, no. 8, p. 1304, Aug. 2016.
- [21] M. Hakonen, H. Piitulainen, and A. Visala, "Current state of digital signal processing in myoelectric interfaces and related applications," *Biomed. Signal Process. Control*, vol. 18, pp. 334–359, Apr. 2015.
- [22] P. Geethanjali, "Comparative study of PCA in classification of multichannel EMG signals," *Australas. Phys. Eng. Sci. Med.*, vol. 38, no. 2, pp. 331–343, Jun. 2015.
- [23] J. V. Stone, "Principal Component Analysis and Factor Analysis," in *Independent Component Analysis: A Tutorial Introduction*, MITP, 2004.
- [24] A. Tharwat, "Linear vs. quadratic discriminant analysis classifier: a tutorial," *Int. J. Appl. Pattern Recognit.*, vol. 3, no. 2, pp. 145–180, Jan. 2016.
- [25] S. Theodoridis, A. Pikrakis, K. Koutroumbas, and D. Cavouras, *Introduction to Pattern Recognition: A Matlab Approach*, 1 edition. Burlington, MA: Academic Press, 2010.
- [26] N. E. Aboudi and L. Benhlila, "Review on wrapper feature selection approaches," in *2016 International Conference on Engineering MIS (ICEMIS)*, 2016, pp. 1–5.
- [27] R. Panthong and A. Srivihok, "Wrapper Feature Subset Selection for Dimension Reduction Based on Ensemble Learning Algorithm," *Procedia Comput. Sci.*, vol. 72, pp. 162–169, Jan. 2015.
- [28] S. Balakrishnama and A. Ganapathiraju, "Linear discriminant analysis - a brief tutorial." Institute for Signal and Information Processing, Mississippi State University, 1998.
- [29] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning: with Applications in R*, 1st ed. 2013, Corr. 7th printing 2017 edition. New York: Springer, 2017.
- [30] D. J. Hand, H. Mannila, P. in the D. of S. D. J. Hand, and P. Smyth, *Principles of Data Mining*. MIT Press, 2001.
- [31] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*. Elsevier, 2011.
- [32] M. L. B. Freitas, J. J. A. Mendes Junior, D. P. Campos, and S. L. Stevan Jr., "Hand gestures classification using multichannel sEMG armband," in *Anais do XXVI Congresso Brasileiro de Engenharia Biomédica*, Armação dos Búzios, RJ, Brasil, 2018.



José Jair Alves Mendes Junior possui graduação em Tecnologia em Automação Industrial pela Universidade Tecnológica Federal do Paraná (UTFPR) – Campus Ponta Grossa (2015). Possui Mestrado em Engenharia Elétrica pela mesma instituição (2017). Atualmente, é aluno de doutorado do Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial (CPGEI) da UTFPR – Campus Curitiba.



Melissa La Banca Freitas possui graduação em Engenharia Eletrônica pela Universidade Tecnológica Federal do Paraná (UTFPR) – Campus Ponta Grossa (2019). Atualmente, é aluna de mestrado do Programa de Pós-Graduação em Engenharia Elétrica (PPGEE) da UTFPR – Campus Ponta Grossa.



Hugo Valadares Siqueira possui graduação em Engenharia Elétrica pela Universidade Estadual Paulista (UNESP) - Câmpus Ilha Solteira (2006), mestrado (2009), doutorado (2013) e pós-doutorado (2014) pela Universidade Estadual de Campinas (UNICAMP). Concluiu em 2017 outro estágio de pós-doutoramento no Instituto Politécnico da Universidade de Pernambuco (Poli-UPE) e Illinois State University - EUA.

Atualmente é professor do Departamento Acadêmico de Engenharia Eletrônica da Universidade Tecnológica Federal do Paraná (UTFPR), Campus Ponta Grossa.



André Eugênio Lazzaretti recebeu o grau de Engenheiro Eletricista e os títulos de Mestre e Doutorado em Ciências na Universidade Tecnológica Federal do Paraná, Curitiba, Brasil, em 2007, 2010 e 2015 respectivamente, todos em Engenharia Elétrica. Atualmente é professor adjunto na Universidade Tecnológica Federal do Paraná, Curitiba.

Sua pesquisa tem o foco em aprendizado de máquina, instrumentação e processamento digital de sinais.



Sergio Luiz Stevan Junior é graduado em Engenharia Elétrica pela Universidade Federal do Paraná (UFPR) em 1999. Possui mestrado em Ciências (área de Telemática) pela CEFETPR em 2002. Possui Ph.D. Engenheiro Eletrônico pela Universidade de Aveiro - em 2011. Atualmente é professor adjunto da Universidade Tecnológica Federal do

Paraná (UTFPR), Campus Ponta Grossa, onde participa como professor permanente do Programa de Pós-Graduação em Engenharia Elétrica. Participa como professor permanente do Programa de Pós-Graduação em Computação Aplicada da Universidade Estadual de Ponta Grossa (UEPG). Atualmente é editor-chefe do Journal of Applied Instrumentation and Control.



Sérgio Francisco Pichorim nasceu em São José dos Pinhais, Brasil, em 23 de abril de 1965. Recebeu o grau de Engenheiro Eletricista e os títulos de Mestre e Doutorado em Ciências no Centro Federal de Educação Tecnológica, Curitiba, Brasil, em 1990, 1995 e 2003 respectivamente, todos em Engenharia Biomédica. É

professor da Universidade Tecnológica Federal do Paraná, Curitiba, desde 1998. Sua pesquisa inclui o desenvolvimento de instrumentação para biotelemetria, processamento de sinais e sistemas biomédicos.