

Smart SPR Sensor: Uso de Aprendizagem de Máquina para Geração de Resposta Inteligente em Sensores Baseado na Ressonância de Plasmon de Superfície

Julio Cartier M. Gomes* Leandro Carlos de Souza*
Miguel Ponce** Leiva Casemiro Oliveira*

* Programa de Pós-Graduação em Ciência da Computação,
Universidade Federal Rural do Semi-Árido e Universidade do Estado
do Rio Grande do Norte, Mossoró, RN, (e-mails:
juliocartier@gmail.com, [leandro.souza, leiva.casemiro]@ufersa.edu.br).
** Cety's University, Mexicali, MX, (e-mail: miguel.ponce@cetys.mx).

Abstract: This work presents a use of machine learning algorithms for data analysis and generation of intelligent response of a sensor based on surface plasmon resonance (SPR). The sensorgrams obtained by the real-time response of an SPR sensor, were used as a database. The sensorgrams were initially described with the discrete cosine transform; and then classified with the k -nearest neighbor (k -NN) algorithm; and the identification of the substances/areas of interest of sensorgram were performed by applying linear regression. The results presented are satisfactory, attesting that the sequence of steps chosen makes it possible to classify the sensorgrams, identify and analyze areas of interest in the sensorgram, therefore creating, a SPR sensor with intelligent response.

Resumo: Este trabalho apresenta o uso de algoritmos de aprendizagem de máquina para análise de dados e geração de resposta inteligente de um sensor baseado na ressonância de plasmons de superfície (SPR). Os sensorgramas obtidos pela resposta em tempo-real de um sensor SPR, foram utilizados como base de dados. Os sensorgramas foram inicialmente descritos com a transformada discreta do cosseno; em seguida classificados com o algoritmo de vizinho mais próximo (k -NN); e a identificação das substâncias/áreas de interesse do sensorgrama foram realizadas com o uso de regressão linear. Os resultados apresentados são satisfatórios, atestando que a sequência de etapas escolhidas possibilita classificar os sensorgramas, identificar e analisar áreas de interesse no sensorgrama, criando assim, um sensor SPR com resposta inteligente.

Keywords: Surface Plasmons Resonance (SPR); Machine Learning; Smart Sensor.

Palavras-chaves: Ressonância de Plasmons de Superfície (RPS); Aprendizagem de Máquina; Sensor Inteligente.

1. INTRODUÇÃO

Os sensores de ressonância de plasmons de superfície (SPR, sigla em inglês para *Surface Plasmons Resonance*) têm sido amplamente utilizados nos últimos anos para projetar e desenvolver aplicações biossensores, devido a oportunidade de se determinar o índice de refração de soluções, detectar a presença de substâncias e monitorar a evolução temporal de camadas de absorção com alta relação custo/benefício Barchiesi (2009).

Os sensores SPR tipicamente exploram as características da estrutura multicamada de diferentes materiais para detectar interações moleculares. Em síntese, uma fina camada metálica é posicionada entre duas camadas dielétrica na qual, em aplicações biossensores, uma das camadas é um meio líquido polar (água) ou não-polar e é comumente denominado de analito.

Sob certas condições, a estrutura multicamadas é iluminada e a luz refletida (sinal) é capturado, fornecendo assim uma assinatura óptica para essa estrutura. O sinal captado passa por processos de filtragem, tendo como objetivo minimizar ruídos e aumentar a faixa do sinal. Em seguida, o sinal pode ser então utilizado para gerar de forma gráfica a ressonância, denominado de curva SPR.

A curva SPR apresenta um valor mínimo de refletividade e esse valor mínimo pode ser alterado a medida que diferentes analitos são inseridos. O monitoramento da alteração dos analitos permite observar as alterações nos parâmetros e comportamento da curva SPR. Tal monitoramento ao longo do tempo pode ser visualizado em um gráfico, denominado de sensorgrama. O sensorgrama apresenta uma evolução temporal nos parâmetros que descrevem a curva Schasfoort (2017).

A Figura 1 apresenta um esquema das etapas envolvidas para aquisição de dados com um sensor SPR. No uso con-

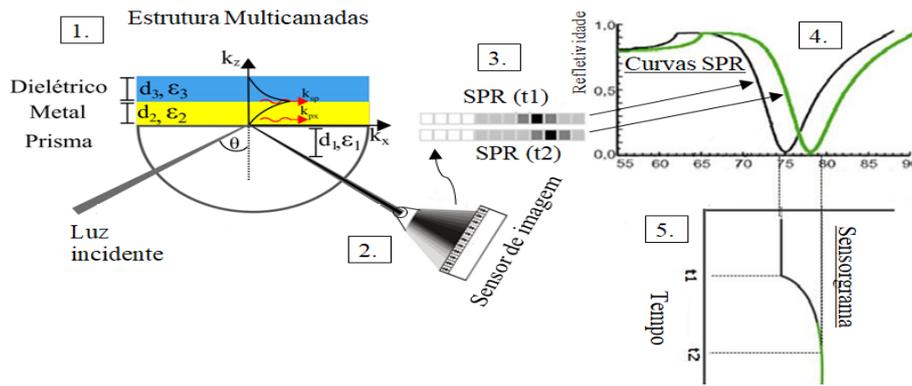


Figura 1. Sensoriamento SPR. 1. Detalhe para uma estrutura com 3 camadas: prisma (substrato óptico), metal e o analito (dielétrico). Situação de ATR, com destaque para o campo evanescente (E) gerado na interface entre os meios e para os vetores de onda da componente horizontal da luz da (k_{px}) e da oscilação do plasmons (k_{sp}). 2. O sinal SPR é captado por um sensor de imagem para daí então ser processado e exibido. 3. Detalhe do mapeamento dos *pixels* da imagem SPR e 4. sua representação gráfica (curva SPR), em que o decaimento na refletividade indica a condição de ressonância. 5. Sensorgrama apresenta a evolução temporal da ressonância. Vetores, ângulo, propriedades ópticas, espessuras e plano de incidência estão indicados na figura.

vencional de um sensor SPR um protocolo experimental é definido com um objetivo específico, e as substâncias são sequencialmente introduzidas no sensor seguindo esse protocolo. Tendo em vista a reprodutibilidade de um sensorgrama SPR para um protocolo experimental bem definido, é possível saber como será a resposta de um sensor SPR, ou seja, qual resposta o sensor deverá fornecer para um determinado ciclo experimental. Desse modo, biossensores SPR inteligentes podem ser desenvolvidos para tomar decisões baseadas nas características particulares e comuns de um protocolo experimental, possibilitando a classificação e a ordenação da resposta.

Nesse sentido, este trabalho tem por objetivo principal utilizar algoritmos de aprendizagem de máquina supervisionada para descrever, classificar e identificar respostas SPR a partir de sensorgramas obtidos com o sensor SPR baseado no prisma de baixo custo PPBIO Oliveira et al. (2016).

2. SENSORGRAMA SPR: CONCEITOS, DEFINIÇÕES E EXPERIMENTOS

A excitação dos plasmons de superfície com sensores SPR desenvolvidos a partir da configuração de Kretschmann Kretschmann and Raether (1968), baseia-se no princípio da reflexão total atenuada (ATR). Esta configuração é a majoritariamente empregada na construção de sensores SPR. Nela, um prisma com superfície metálica é posto em contato com o analito, e iluminado em condições ideais para que ocorra acoplamento óptico entre os fótons e os plasmons de superfície.

A ressonância é obtida variando as condições de acoplamento entre um feixe de luz e os plasmons de superfície. Dentre as formas para se obter tais condições, é possível desenvolver sensores SPR para operar no modo de interrogação angular (AIM - *Angular Interrogation Mode*), onde o feixe de luz incidente é monocromático (comprimento de onda fixo) e variando o ângulo de incidência obtém-se a condição de ressonância Oliveira (2015). No modo AIM a posição de ressonância, denominada de ângulo de

ressonância θ_R , é indicada pelo valor mínimo da curva SPR.

Sob certas condições, a estrutura multicamadas é iluminada e a luz refletida (sinal) é capturada, fornecendo assim uma assinatura óptica para a essa estrutura. O sinal captado passa por processos de filtragem, tendo como objetivo minimizar ruídos e aumentar a faixa do sinal. Em seguida, a imagem pode ser então utilizado para gerar de forma gráfica a ressonância, denominado de curva SPR.

A SPR é um excelente fenômeno para monitorar eventos de adsorção/dissociação entre moléculas. Através das alterações no índice de refração das substâncias em contato com sensor, o valor de θ_R também sofrerá um deslocamento, fornecendo a assinatura óptica para a estrutura. Em síntese, a medida que novas substâncias são inseridas ocorre uma mudança na condição de ressonância, como mencionado anteriormente. As variações de θ_R ao longo do tempo são então apresentados no sensorgrama. Um exemplo de sensorgrama pode ser visto na Figura 2. Na figura é ilustrado o protocolo experimental para as substâncias água (H_2O), solução da proteína BSA (Albumina Sérica Bovina) diluída em água, a concentração de 0,1 mg/ml e hipoclorito (Hypo.) na seguinte ordem: $H_2O \rightarrow BSA \rightarrow H_2O \rightarrow \text{hipoclorito} \rightarrow H_2O$.

Tipicamente as fases de um sensorgrama compreendem a **linha base**: calibração do sensor para uma substância conhecida; **adsorção**: início da injeção da solução de proteína; **monolayer**: regime permanente, saturação no valor do parâmetro de interesse; **dissociação**: remoção das moléculas em suspensão que não aderiram à superfície metálica; **lavagem**: injeção de substância para limpar a superfície e remover as moléculas fixadas na fase de adsorção e **regeneração**: retorno para a condição inicial de calibração Oliveira (2016).

O protocolo experimental representado por um sensorgrama, apresenta padrões e característica do sensor SPR em uso. Tendo em vista que o uso sequencial do sensor por vários ciclos experimentais de adsorção, dissociação, limpeza e regeneração provoca alterações e desgastes na

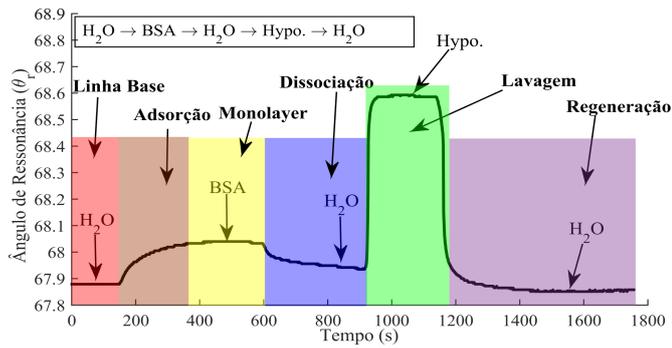


Figura 2. Exemplo de um sensorgrama para o protocolo experimental: $H_2O \rightarrow BSA \rightarrow H_2O \rightarrow Hypo. \rightarrow H_2O$, em um sensor SPR com multicamada composta de vidro BK7 e ouro.

superfície metálica, também é possível verificar por meio do sensorgrama a qualidade do sensor e atestar a reprodutibilidade do mesmo.

2.1 Set-up experimental

Para obtenção dos sensorgramas foi utilizado um sensor SPR baseado no PPBIO (Prisma Polimérico para aplicação BIOLógicas) Moreira et al. (2009). Devido a sua geometria trapezoidal (ver Figura 3a), o PPBIO torna menos dispendiosa a instrumentação de um sensor SPR por evitar partes móveis e/ou rotacionais. As paredes laterais inclinadas do PPBIO são especulares para facilitar a reflexão da luz incidente. Esse prisma possui uma fina camada metálica depositada na base superior, evitando o uso de óleo/gel para acoplamento da superfície metálica. Para mais detalhes sobre o sensor SPR baseado no PPBIO recomenda-se as referências Oliveira et al. (2013); Oliveira et al. (2016).

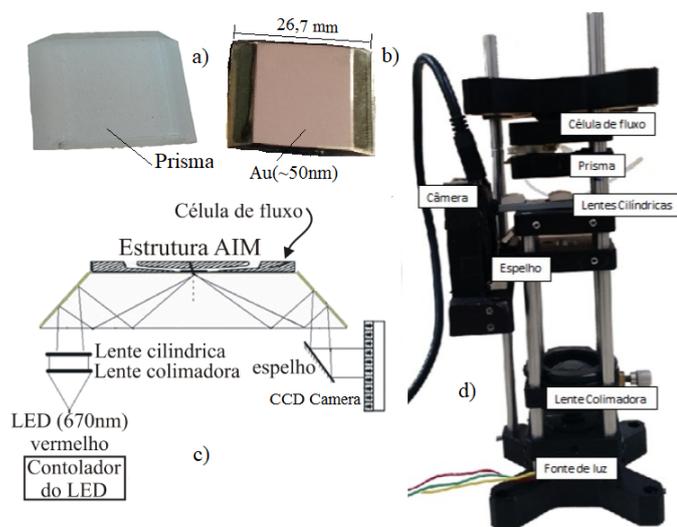


Figura 3. Fotografia do PPBIO a) antes e b) depois da deposição de 50 nm de ouro. c) diagrama de blocos do sensor e d) fotografia do set-up experimental.

3. METODOLOGIA PARA INCORPORAÇÃO DE INTELIGÊNCIA NO SENSOR SPR

Neste trabalho, é proposto a utilização da aprendizagem de máquina supervisionada para a análise de um sensorgrama, a fim de descrever, identificar e classificar o protocolo experimental adquirido por um sensor SPR. Na Figura 4 são apresentados as etapas para consecução dessas tarefas de aprendizagem.

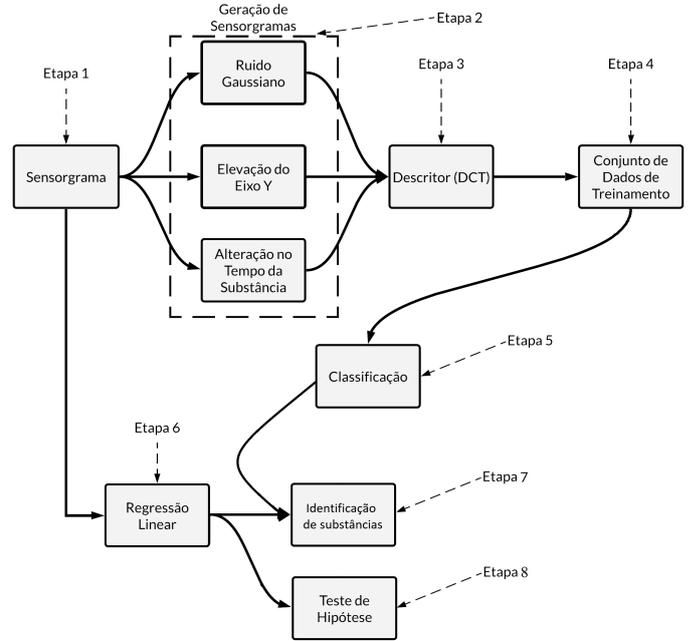


Figura 4. Diagrama de bloco indicando as etapas para realizar as tarefas de criar, descrever, identificar e classificar sensorgramas e suas informações.

A primeira etapa, consistiu em obter grupos de sensorgramas usando o sensor SPR baseado no PPBIO, operando do modo AIM, para prisma de vidro BK7 metalizados com filmes finos de ouro com 50 nm de espessura. Obteve-se 4 grupos cada um representando um protocolo experimental diferente.

A segunda etapa consistiu em criar novas amostras para os sensorgramas já obtidos. Modificações tais como introdução de ruído gaussiano branco, diminuição temporal em regiões de interesse e aumento ou diminuição no valor do eixo y (posição de ressonância) foram adicionadas aos sensorgramas. Essas modificações podem representar o uso de diferentes algoritmos de detecção do valor mínimo da curva SPR, mudança no tempo de inserção de uma substância com vista a economizar material e alteração do PPBIO, camada metálica ou fonte de luz, respectivamente. Já a terceira etapa, foi aplicada a transformada discreta de cosseno (DCT) Zhou and Chen (2009) em cada um dos sensorgramas pertencentes a cada grupo. A quarta etapa consistiu na criação de um conjunto de dados de treinamento com os 16 sensorgramas. Na quinta etapa, aplicou-se o algoritmo de aprendizagem de máquina k -NN Agrawal (2014) para classificação dos sensorgramas.

De posse de sensorgramas classificados, faz-se necessário identificar as regiões que representam as substâncias no sensorgrama, e conseqüentemente, identificar as áreas de

transição entre elas. Esses passos compreendem a sexta etapa. Aplicou-se a regressão linear para estimar as regiões para estimar os regimes permanentes no sensorgrama, observando a mudança no valor do coeficiente da regressão. Em seguida, na sétima etapa, é possível identificar a ordem temporal e a quantidade de substâncias que um sensorgrama apresenta. Por fim, foi aplicado testes de hipótese na resposta do sensor para aferir a qualidade do sensorgrama gerado (etapa 8).

4. RESULTADOS

4.1 Classificação de sensorgramas

Os 4 grupos de sensorgramas obtidos experimentalmente com sensor SPR são apresentados na Figura 5. As substâncias utilizadas foram H₂O, BSA, Hypo. e PBS (*Phosphate Buffered Saline*). Em cada protocolo, tem-se uma sequência temporal diferente para inserção das substâncias.

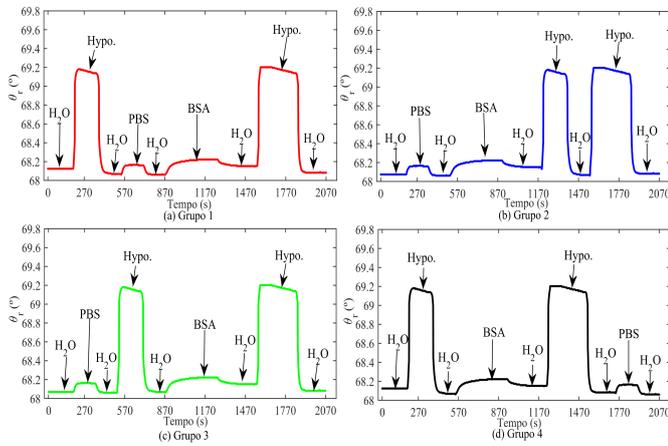


Figura 5. Grupos (a), (b), (c) e (d) de sensorgramas representado 4 protocolos diferentes.

Em seguida, foram criados novos sensorgramas de cada grupo. A Figura 6 ilustra, para o grupo 1, a aplicação de ruído gaussiano branco de 65 dBW, do aumento no valor do y em 0.2° e da mudança no tempo na inserção das substâncias BSA e Hipoclorito. Ao final dessa segunda etapa foram criados 16 sensorgramas, sendo 4 sensorgramas para cada um dos grupos.

Na terceira etapa, foi aplicado a DCT-4 com intuito de descrever os sensorgramas. O uso de um descritor simples faz-se necessário pois a formulação de um modelo matemático que descreva um sensorgrama exige a inclusão de aspectos bioquímicos, processos isotérmicos de absorção e dissociação das substâncias, computados em condições específicas nos moldes do modelo de Langmuir, os quais tornam tal modelagem um trabalho complexo, dependente de experimentos custosos e exigindo uma série de suposições Sousa et al. (2018). A DCT-4 não faz uso de números complexos, representando de forma compacta a informação do sensorgrama. Na Figura 7 tem-se a energia da DCT-4 para o sensorgrama da Figura 5a). Nota-se que os primeiros 8 coeficientes AC da DCT são responsáveis por aproximadamente 80% do sinal.

Com os coeficientes da DCT para os 16 sensorgramas, foi criado o conjunto de dados de treinamento. A Figura 8

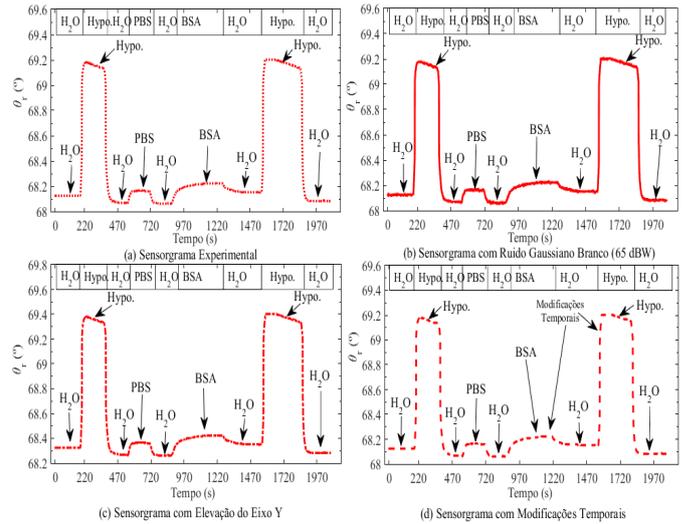


Figura 6. (a) Sensorgrama experimental. (b) aplicação de ruído gaussiano branco de 65 dBW. (c) incremento de 0.2° no eixo y e (d) alteração das regiões de interesse.

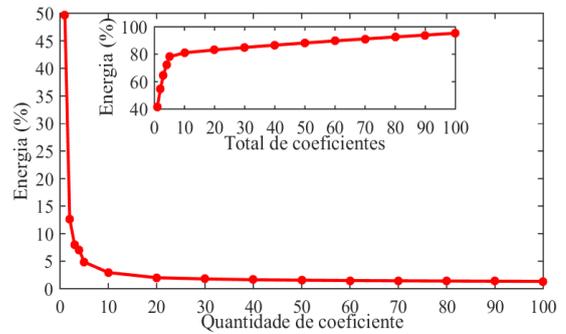


Figura 7. Frequência da energia do sensorgrama utilizando a estrutura multicamada BK7/Au do grupo 1 e a frequência acumulada da energia do sensorgrama.

apresentada os três primeiros coeficientes do nível AC para cada grupo de sensorgrama. Observa-se pelo espalhamento dos dados no gráfico 3D, que os elementos de um grupo estão próximos entre si.

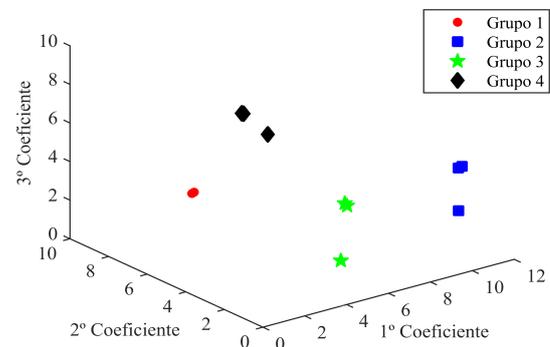


Figura 8. Disposição 3D dos coeficientes de DCT de cada grupo de sensorgrama.

Atribuído o dado de destino para cada sensorgrama, esses dados foram inseridos no algoritmo de aprendizagem de máquina *k*-NN. No treinamento, ajustou-se o algoritmo

para utilizar a distância euclidiana e computá-la na vizinhança de 3 vizinhos mais próximos.

Com os parâmetros ajustados, a quinta etapa consistiu em classificar sensorgramas. Para o teste de classificação, foram criados dois novos sensorgramas, denominados de Teste 1 e Teste 2 (do Grupo 1 e do Grupo 2), aplicando para isso as modificações. Para o Teste 1 foi aumentado o valor do eixo y em 0.2° e diminuído o tempo da inserção das substâncias Hipoclorito e BSA. O Teste 2 foi diminuído o tempo da inserção da substância Hipoclorito. Os novos sensorgramas encontram-se na Figura 9 e sua posição no gráfico 3D dos coeficientes de DCT estão indicadas na Figura 10.

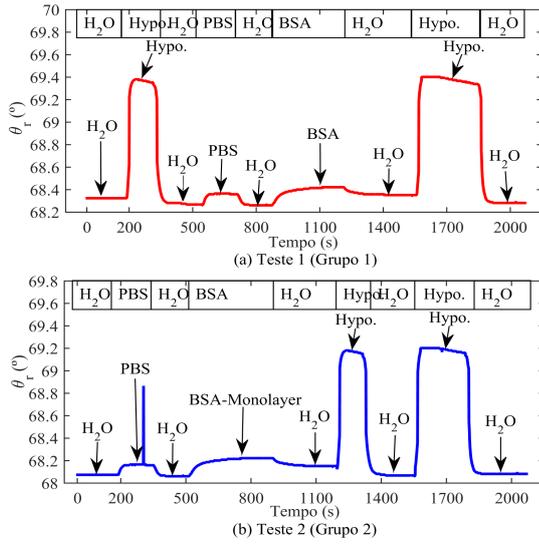


Figura 9. Sensorgramas de teste para classificação utilizando o algoritmo k -NN. (a) Sensorgrama do Grupo 1 com alteração no valor da posição de ressonância (b) sensorgrama do Grupo 2, com alteração da inserção da substância Hipoclorito e presença de bolha de ar na inserção do PBS.

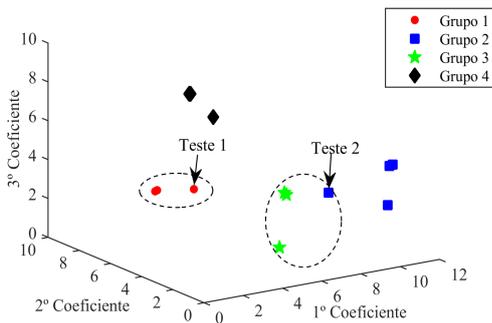


Figura 10. Novos exemplos inseridos para classificação utilizando o algoritmo k -NN. O sensorgrama Teste1 foi classificado corretamente, enquanto o Teste2 for gerado na fronteira entre dois grupos, gerando uma classificação errada.

O classificador então foi usado para verificar em qual grupo cada um desses sensorgramas pertencia. Com os testes de exemplo, o algoritmo classificou o sensorgrama Teste 1 como sendo do Grupo 1 de forma satisfatória e

o sensorgrama Teste 2 com sendo do Grupo 3 de forma insatisfatória. O classificador apresentou boas taxas de acertos na classificação de sensorgramas que apresentam 9 (nove) regimes permanentes, isto é, sensorgramas que representam protocolo experimental composto de ciclos de nove substâncias. Contudo, a presença de bolhas de ar na mediação, caracterizada pelo pico em 350 s na Figura 9(b), não são foram modeladas levando a falha de classificação.

4.2 Identificação e análise em sensorgramas

Por meio da análise de regressão, o sensor estima as regiões que representam as substâncias analisadas, e consequentemente, identifica as áreas de transição entre elas. A região de transição entre substâncias é expressa por:

$$S(a, b) = \sum_{i=1}^{L_t} (\epsilon_i)^2 = \sum_{i=1}^{L_t} (y_i - a - bx_1)^2 \quad (1)$$

em que a janela de tempo L_t define os pontos usados para calcular o coeficiente de inclinação dos pontos. A janela deve ser ajustada de modo que ruídos presentes na instrumentação do sensor não caracterizem a ocorrência de transição entre substâncias.

Na Figura 11, tem-se a identificação de padrões no protocolo experimental. Os padrões identificados são expressos pela condição de ressonância estável quando se tem fração de uma única substância em contato com sensor, e as regiões de transição entre as substâncias.

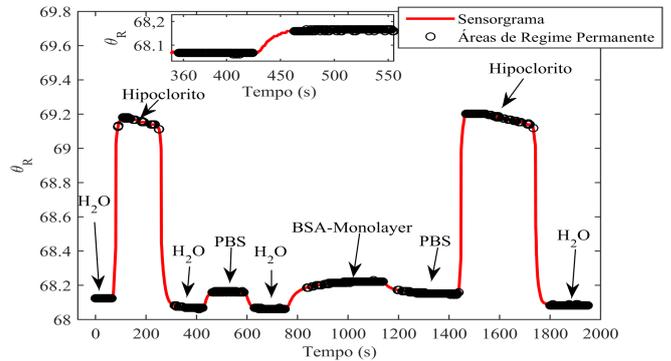


Figura 11. Identificação de padrão na resposta do sensor SPR. Identificação dos regimes permanente, caracterizando as regiões que identificam as substâncias analisadas.

Para verificar se os dados estão distribuídos na forma normal no modelo de regressão do regime permanente foi aplicado a análise dos resíduos. Na Figura 12 (a) é possível observar a análise do resíduo do regime permanente, em que sua distribuição está na forma normal, diante que segue a condição de ressonância estável. Já a Figura 12 (b) em que é um regime de transição, os dados não estão na forma normal, não seguindo a condição de ressonância estável.

Com o uso da regressão linear é possível proceder com testes de hipótese, a fim de atestar a confiabilidade dos dados obtidos com o sensor. O teste consiste em verificar a inclinação da reta gerada na fase de regressão. Nas áreas de regime permanente, a inclinação deve ser próxima de zero. O nível de confiança define os limites superior e inferior do

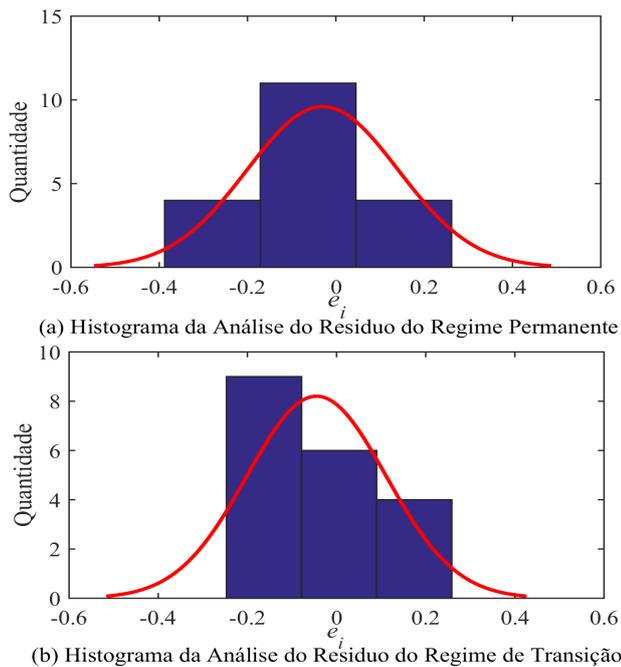


Figura 12. Histogramas da análise do resíduo do modelo de regressão. (a) Histograma do regime permanente. (b) Histograma do regime de transição.

intervalo de confiança e o teste de hipótese verifica se na janela temporal estabelecida o coeficiente de regressão está contido nesses limites. Assim, foi testado se o sensorgrama reflete o protocolo experimental com 99.9% de confiança.

A Figura 13 apresenta o teste de hipótese para o coeficiente da regressão linear. Quando o teste da hipótese de que a inclinação da reta é nula encontra-se dentro dos limites estabelecidos, é possível confirmar que é uma substância. Em destaque, tem-se o intervalo em que foi possível confirmar com 99.9% de certeza que essa região do sensorgrama representa uma substância.

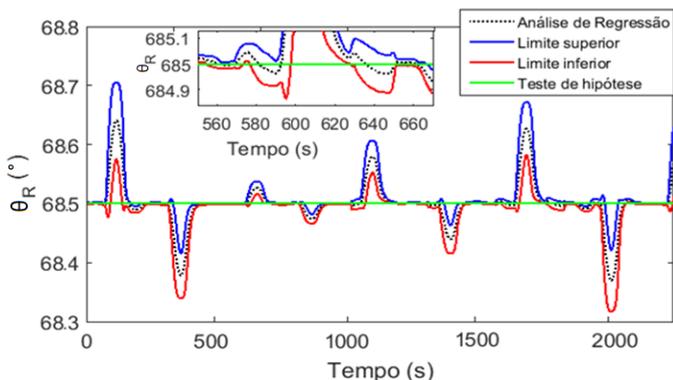


Figura 13. Teste de hipótese para verificar confiabilidade da presença de uma determinada substância em um protocolo experimental. Em destaque tem-se o intervalo em que foi possível confirmar com 99.9% de certeza que é uma substância.

5. CONCLUSÃO

Foi apresentado a utilização de algoritmos de aprendizagem de máquina com objetivo de descrever, analisar e

classificar o sensorgrama de um determinado protocolo experimental. Os algoritmos apresentaram resultados satisfatórios, permitindo classificar sensorgramas e identificar padrões na resposta SPR.

Com o uso apropriado de algoritmos inteligentes é possível aferir a qualidade de um sensor SPR e atestar em que condições de operação este funcionará com determinado nível de confiabilidade. Assim, o sensor inteligente garante a confiabilidade nas respostas fornecidas, condição essencial para auditoria por órgãos fiscalizadores

Ademais, a análise sobre o uso de técnicas de inteligência em sensores SPR possibilita prever anomalias no comportamento do sensor, garantido melhores resultados nas medições.

AGRADECIMENTOS

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

REFERÊNCIAS

- Agrawal, R. (2014). K-nearest neighbor for uncertain data. *Int. J. of Comp. Appl.*, 105(11).
- Barchiesi, D. (2009). Adaptive non-uniform, hyper-elitist evolutionary method for the optimization of plasmonic biosensors. In *Int. Conf. on Comp. Ind. Eng.*, 542–547.
- Kretschmann, E. and Raether, H. (1968). Radiative decay of nonradiative surface plasmons excited by light. *Z. Naturforsch. A*, 23, 2135.
- Moreira, C., Neto, A., Lima, A., Thirstrup, C., and Neff, H. (2009). Exchangeable low cost polymer biosensor chip for surface plasmon resonance spectroscopy. *Procedia Chemistry*, 1, 1479–1482.
- Oliveira, L.C. (2016). *Construção e Caracterização de Sensores SPR: influências da camada metálica e do substrato dielétrico*. Tese de doutorado, Universidade Federal de Campina Grande.
- Oliveira, L., Moreira, C., Lima, A., and Neff, H. (2016). A prism-based polymeric surface plasmon resonance biochip for angular and spectral modes. *Procedia Engineering*, 168, 1350 – 1353.
- Oliveira, L., Morerira, C., Thirstrup, C., Melcher, E., Lima, A., and Neff, H. (2013). A surface plasmon resonance biochip that operates both in the angular and wavelength interrogation modes. *IEEE TIM*, 62(5), 1223–1232.
- Oliveira, L.e.a. (2015). *Surface Plasmon Resonance Sensors - A Materials Guide to Design and Optimization*. SpringerBriefs in Physics. Springer, 1st edition.
- Schasfoort, R.B.M. (ed.) (2017). *Handbook of Surface Plasmon Resonance*. The Royal Society of Chemistry, 2 edition.
- Sousa, T., Oliveira, L., Neff, H., Laborde, H., and Lima, A. (2018). Numerical tool for estimating the dielectric constant, the thickness, and the coverage of immobilized inhomogeneous protein films on gold in aqueous solution. *Appl. Opt.*, 57(24), 6866–6875.
- Zhou, J. and Chen, P. (2009). Generalized discrete cosine transform. 449–452. Pacific-Asia Conf. on Cir., Comm. and Sys.