Journal Pre-proof

SmartSPR sensor: Machine learning approaches to create intelligent surface plasmon based sensors

Julio Cartier M. Gomes, Leandro Carlos Souza, Leiva Casemiro Oliveira

PII: S0956-5663(20)30747-8
DOI: https://doi.org/10.1016/j.bios.2020.112760
Reference: BIOS 112760

To appear in: *Biosensors and Bioelectronics*

Received date : 22 May 2020
Revised date : 14 September 2020
Accepted date : 21 October 2020

Please cite this article as: J.C.M. Gomes, L.C. Souza and L.C. Oliveira, SmartSPR sensor: Machine learning approaches to create intelligent surface plasmon based sensors. *Biosensors and Bioelectronics* (2020), doi: https://doi.org/10.1016/j.bios.2020.112760.

SmartSPR Sensor: Machine Learning Approaches to Create Intelligent Surface Plasmon Based Sensors

Julio Cartier M. Gomes[a], Leandro Carlos Souza[b] and Leiva Casemiro Oliveira[a]

[a]Universidade Federal Rural do Semi-Árido, Departament of Computer Science, Rua Francisco Mota Bairro, 572 - Pres. Costa e Silva, Mossoró - RN
[b]Universidade Federal da Paraíba, Informatic Departament, Campus I - Lot. Cidade Universitaria, João Pessoa - PB

ARTICLE INFO

Keywords:
Surface Plasmon Resonance
SPR Intelligent Sensor
Smart Biosensor
Optical Sensor
Machine Learning
Automated Diagnosis Testing

ABSTRACT

Surface plasmon resonance (SPR) based sensors allow to evaluate liquid and gases solutions from real-time measurements of molecular interactions. The reliability of the response generated by a SPR sensor must be guaranteed, especially in substance detection, diagnoses and others routine application, since poorly handled samples, instrumentation noise features or even molecular tampering manipulations can lead to wrong interpretations. This work investigates the use of different machine learning techniques to deal with such issues, and aim to improve and attest the quality of the real-time SPR responses so-called sensorgrams. A new strategy to describe a SPR-sensorgram is show. The results of the proposed approach of using machine learning allows to create the intelligent SPR-sensor which is able to classify sensorgramas and identify the substances presents in it. Also made it possible to order and analyze interest areas of a sensorgrams, standardizing data and supporting eventual official audit. With those embedded intelligence features, the new generation of SPR-intelligent biosensors are qualify to perform automated testing. A properly protocol for Leishmaniasis diagnosis with SPR was used to verify these new feature.

1. Introduction

The surface plasmon resonance (SPR) sensors have been widely used in bio-sensing applications, with the opportunity to obtain a technical solution that does not require markers to make it feasible. The capability to monitor molecular interactions, associating to its structure, in quick responses, has been made SPR sensors a scientifically and attractive tool (Piliarik et al., 2009). The sensing procedures typically occurs by monitoring the behavior changes of the SPR curve parameters, as new analytical substances are admitted. Refractive index changes and resonance position shifts are the indicators for the presence or temporal evolution of absorption layers. The parameters tracking over time can be visualized in a graph called sensorgram.

Sensorgram can be used to graphically detect the parameters of interaction events, such as specific and non-specific binding. Fig. 1 illustrates a sensorgram for the variation of the resonance condition, which indicates by the resonant wavelength (λ_r) over time. When a new substance is admitted, a change in the resonance couple conditions occurs, indicated by the displacement of λ_r. Different phases or areas can be noticed in a sensorgram, which when filtered, identified, and manipulated allows understanding the binding and interaction processes.

The pieces of information contained in the sensorgram are important for the bio-interactions interpretation. To properly manipulate the extracted parameters of a sensorgram, mathematical models, machine learning techniques, and signal descriptors (Amanatiadis et al., 2011) must be applied. Mathematical models, like the Langmuir model, allow to measurement quantitative values for kinetics and affinities

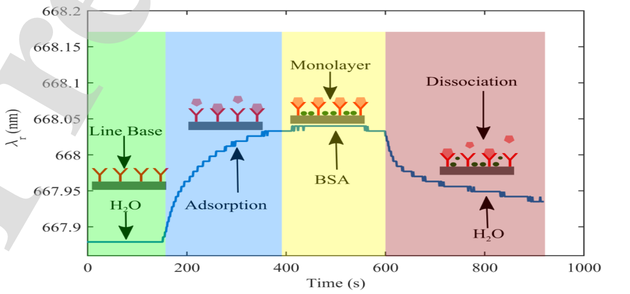

Fig. 1. Sensorgram for real-time monitoring of the resonance position λ_r and indication of its typically phases: Line Base, Adsorption, Monolayer (steady state) and Dissociation.

(Canoa et al., 2015). Identification of other binding parameters in biochemical interactions, such as the association and dissociation phases of the sensorgram, can be performed with the particle swarm optimization algorithm (Yan et al., 2017), using non-linear differential equations (Tiwari et al., 2015) or using finite element simulations (Agrawal et al., 2017). The applicability of these methods is dependent on costly and laborious experiments that sometimes are complex and very specific. Besides, such methods are not useful for sensorgrams classification, identifying the presence of substances, and even to attest the quality of SPR responses.

An approach based on machine learning (ML) algorithms can be used to perform the discovery of the sensorgrams' parameters using data analysis, at the same time that tasks regarding sensorgrams classifying and identification, automated testing, and improvements on the sensor responses can be done without incurring new laborious efforts. Few studies are presented in the literature discussing the application of machine learning to create intelligent SPR sensors. Classification and partitioning methods were used to identify

juliocartier@gmail.com (J.C.M. Gomes); leandro@ci.ufpb.br (L.C. Souza); leiva.casemiro@ufersa.edu.br (L.C. Oliveira)
ORCID(s):

and classify inhibitors (Zhang et al., 2018). However, the usage of the SPR sensor was proposed only to attest to the presence of the inhibitors in the solution. Discussion over linear regression applied on the SPR-curve (not on sensorgram) to calibrate and detect biomarkers related to diseases were present elsewhere (Lee et al., 2014).

Intelligent SPR sensors can be developed to take automatic decisions based on particular and common characteristics of an experimental protocol, expressed by a sensorgram. For this, an intelligent module to deal with the learning tasks must be embedded in the sensor, not affecting the original sensor's design and configuration. The present work aims to use ML algorithms to describe, classify, order, identify, attest quality, and perform an automated test on sensorgrams which express diagnosis protocols, and thus create an intelligent SPR sensor in terms of the sensorgrams responses.

2. Machine Learning Applied to Sensorgram

Fig. 2 shows the two stages of the proposed ML approaches. The preliminary steps are focused on the description, classification, and ordering of sensorgrams. The purpose of this stage is to establish the environment and data to enable inferences. The second stage is related to the smart sensor action, by execution the intelligence task regards to substances identification, quality verification, and automated testing.

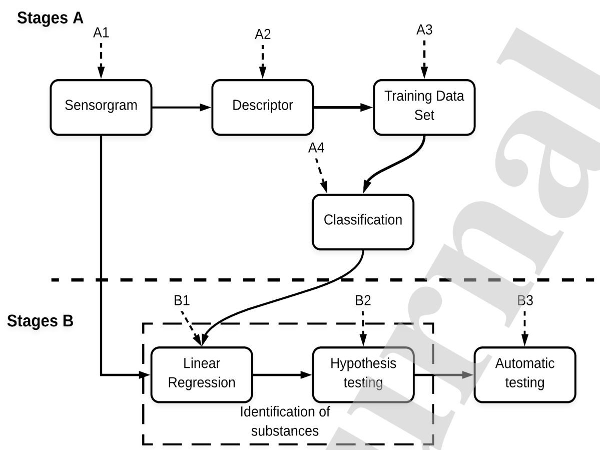

Fig. 2. Block diagram of the propose ML approaches. Describing, classifying, identifying, analyzing and testing are tasks executed based on the sensorgrams responses.

A1 - Sensorgram: Experimental Aspects

For the sensorgrams acquisition, an SPR sensor based on the so-called PPBIO (Polymeric Prism for BIOlogical applications) (Oliveira et al., 2013) was employed. The experimental set-up is shown in Fig. 3 for the sensor operating in the spectral interrogation mode (Oliveira et al., 2019). PPBIO has a trapezoidal structure and reduced dimensions (22 x 10 x 3 mm) allowing up to 10 different sensitive spots. The inclined faces reflect the fan in/fan out light source. Incident light at a fixed angle of 68° and upper surface covered with a thin metal layer (gold), accounts for the low maintenance and robustness. This features also avoids the use of gel/coupling liquid, which causes considerable instrumentation costs. For more details of the SPR PPBIO based sensor, we refer (Oliveira et al., 2013; Moreira et al., 2016; Oliveira et al., 2019). The sensorgrams represent several experimental protocols, and fifteen groups were created. The groups differ in terms of the used substances as well as the quantity of them. The acquired experimental protocols contain three, five, seven, or nine permanent regimes. The permanent regime is defined as a steady-state of a substance in which the value of interest (λ_r) is constant (Patil et al., 2019). Fig. 4 (a-l) shows examples of these groups. It is possible to observe that there are changes in the positions and number of substances, creating sensorgrams groups with 9 (a,e,i,m), 7 (b,f,j,n), 5 (c,g,k,o), and 3 (d,h,l) permanent regimes.

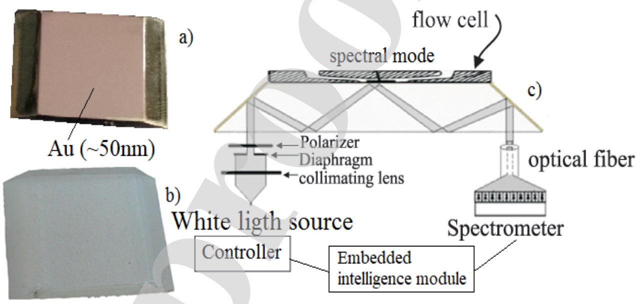

Fig. 3. Sketch of the experimental set-up (c) and photographic image of the PPBIO prism with (a) and without (b) metal layer. The smart SPR sensor is characterized by the embedded intelligence module containing the ML-approaches present here.

The experimental SPR-sensing characteristics and related sensorgrams in the spectral mode were obtained with: plain water (H_2O), high and low concentration of ethanol-water mixture, hypochlorite solution (for non-specific binding) and specific binding of aqueous solutions of BSA, neutral avidin, and other essential substances for protocols of Leishmaniasis (Ferreira et al., 2017) and Dengue fever (Oliveira et al., 2019) detection mediate by SPR.

A2 - Descriptor

To represent a sensorgram, the descriptor so-called TSD (Temporal Sensorgram Descriptor) was created. Since a sensorgram represents one well-defined experimental protocol, the number of substances (permanent regimes) is known. Thus, the TSD algorithm formulated a complex number in which the real part represents the mean value of the permanent regime, and the imaginary part is the length of this regime. Fig. 4 (p) illustrates the application of the TSD descriptor in a sensorgram that corresponds to an experimental protocol with 3 permanent regimes (R1, R2, and R3).

The descriptor TSD was applied in each sensorgram of each group. Fig. 4 (q-s) show the TSD for the 3 sensorgrams

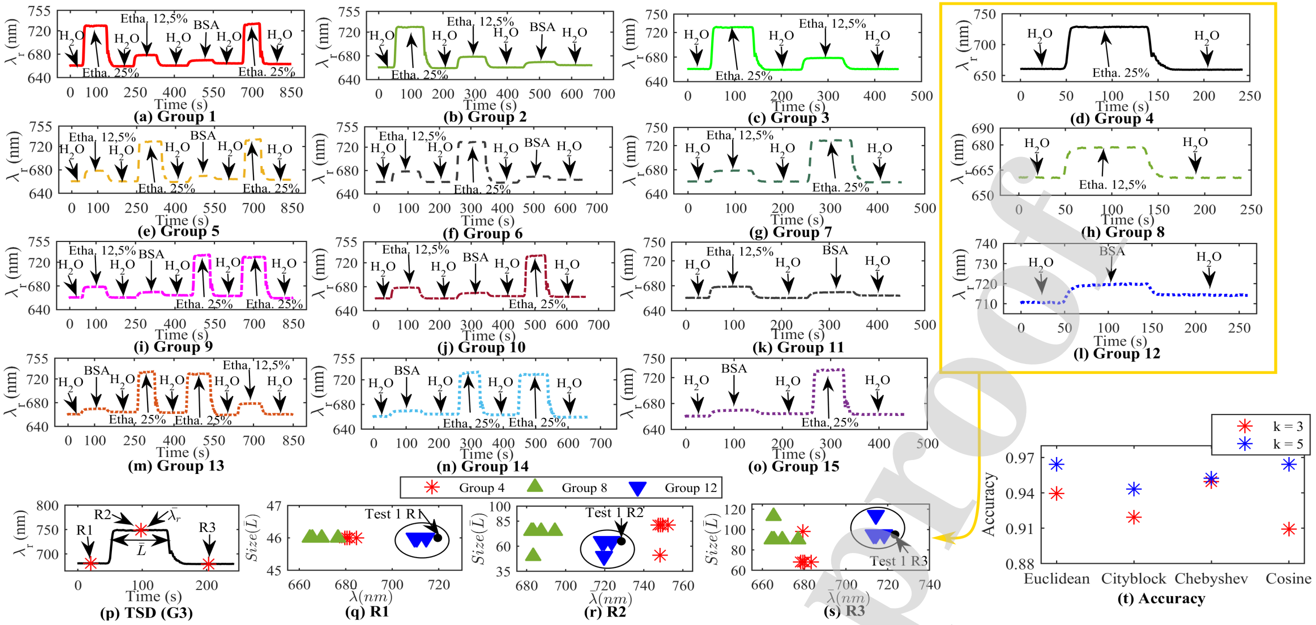

Fig. 4. a-o)Examples of sensorgrams groups. Altogether fifteen different protocols were obtained with a glass BK7 PBBIO and 50 nm gold thin-film. p) Graphical example for TSD of a sensorgram with 3 permanent regimes. Detail of the means and length obtained with TSD. q-s) Example of adding a new sensorgram (black circle) with 3 permanent regimes and TSD labeling it as another element of the group 12. t) Accuracy of the k-NN algorithm with Euclidean, Cityblock, Chebyshev, and Cosine distances.

that are elements representatives of the group with 3 permanent regimes. Each regime is graphically arranged with the mean wavelength value of the regime ($\bar{\lambda}_r$) in the x-axis and its size/length ($\bar{L}$) on the y-axis. It is possible to observe that the sensorgram groups are relatively close to each other but sufficiently distant for their distinction with TSD strategy.

A3 - Training Data Set

New elements for each group of sensorgrams were obtained, applying modifications to the experimental sensorgrams. Theses modifications may represent different algorithms for resonance position detection, changes in the time of the substances admission (aim to material savings), a new multi-layer arrangement (e.g. replace the metal layer), and another noise-pattern intrinsic of the sensor instrumentation. To represents these modifications, the following actions were performed: i) increase or decrease in the resonance position (y-axis value), ii) temporal decrease in the region (steady-state) of interest (x-axis value), iii) introduction of white Gaussian noise, and iv) noise from the microfluidic system (insertion of air bubbles). Such modifications account for the aspects of the reaction kinetics involved in the substances multiplexing.

For training, a dataset was created with the TSD values of all sensorgrams group and its elements, generating 99 sensorgrams. The Fig. 4 (a-l) shows one element for each of the 15 groups of sensorgrams that form the training dataset. The figure also indicates that a new element to be added in the dataset has its TSD extracted and the accuracy metric (Basavaraju et al., 2020) calculated to determine which group this new sensorgram belongs to. With an established training dataset, a target label is assigned to the TSD of each sensorgram group.

A4 - Classification

Finally, to classify sensorgrams the k-NN algorithm was applied in the set of TSD coefficients of all sensorgrams groups. The accuracy of the k-NN algorithm was obtained with the leave-one-out cross-validation (LOO-CV) (Kocaguneli and Menzies, 2013). In the LOO-CV, the dataset consists of $N-1$ training data, and the one left is used for testing. The process is repeated N times, and at each iteration, a different sample is left out for testing. Next, the destination labeling of each sensorgram group is assigned. The distance was used as a classification parameter. The distance describes the elements in space (neighborhood), according to the approximation between them (Deza and Deza, 2016). Euclidean, City block, Chebyshev, and Cosine distances present good performances (Chomboon et al., 2015; Mulak and Talhar, 2015; Alfeilat et al., 2019). The accuracy of the classification was performed for the distances calculated in the neighborhoods $k = 3$ with 99 sensorgrams and $k = 5$ nearest neighbors for 84 sensorgrams, only those which containing five, seven, and nine permanent regimes.

Fig. 6 (t) shows the accuracy values obtained with the LOO-CV classification and validation. Using the TSD descriptor, the accuracy of the k-NN is higher than 0.90, which means that the algorithm response obtained satisfactory performance. The Euclidean distance has the highest mean value for $k = 5$ neighbors, and Chebyshev when used $k = 3$.

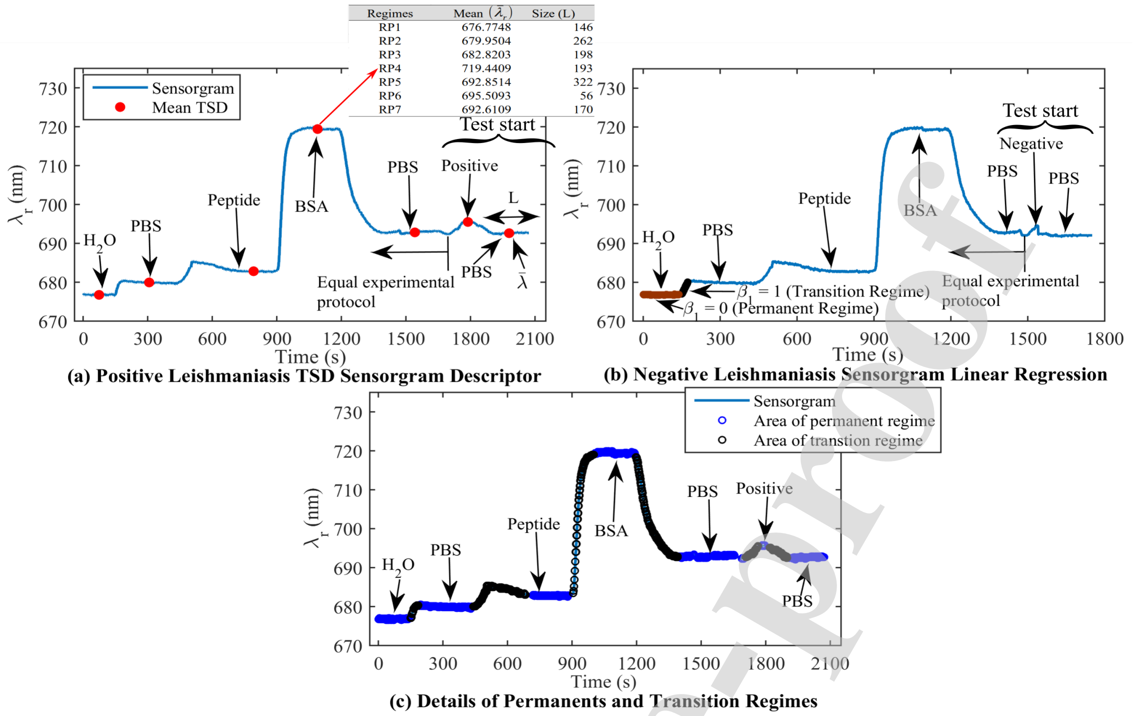

Fig. 5. Proof-of-concept sensorgrams for the Leishmaniasis detection. The experimental protocol recipe is: $H_2O \rightarrow PBS \rightarrow PeptideLC2 \rightarrow BSA \rightarrow$. Exemples of (a) positive and (b) negative samples tests. Temporal and vertical difference does not affect classification. Detail of TSD applied to the Leishmaniasis experimental protocol for positive test. The $\bar{\lambda}_r$ of each substance in the recipe is indicated by the red dot. Estimated values of the linear regression parameter (β_1) for the permanent (brown dots) and transition (black dots) regime, indicated in negative test. (c) Identification of permanent which characterizing the substances identify and the complementary transition regimes.

3. Stage B - Smart Sensor Action

Once the dataset has been classified, it is possible to use the smart SPR sensor to identify sensorgrams and substances, attests to the quality of the sensor response, and on sensorgrams which express diagnosis protocols perform an automated test. The experimental proof-of-concept protocol for Leishmaniasis' diagnosis, presented in (Ferreira et al., 2017), will be used in stage B.

Fig. 5 shows the Leishmaniasis experimental protocol for detection mediate by SPR. Indication of positive and negative test examples. Once the protocol recipe has been established until the 1800s in positive sensorgram (and until 1500s negative sensorgram), human samples diluted 1:1000 in PBS are tested. The recipe is composed of seven permanent regimes. The LC2 antigen was used, a highly specific and sensitive synthetic peptide, derived from a Leishmania chagasi mimotype. For more details see (Ferreira et al., 2017).

Identification and Analysis of Substances

The intelligent SPR sensor classifies new sensorgrams as its permanent regimes are being formed. For Leishmaniasis protocol, the TSD strategy creates a training and testing dataset for sensorgrams with 3, 5, and 7 permanent regimes, obtaining 42 sensorgrams in total. Fig. 5 (a) shows the TSD values calculated for the Leishmaniasis sensorgram, composed of 7 permanent regimes. The k-NN algorithm with k = 3 and Euclidean distance is used for recognition. With the LOO-CV cross-validation process, the accuracy was 0.92.

B1 - Linear Regression

After the sensorgram has been classified, the regions that represent substances, and consequently the transition areas between them, are segmented by linear regression. The regression follows the Eq. 1. Ideally, permanent regimes exhibit a straight line, and the slope coefficients β_1 is zero. The ML approach calculates the coefficients β_0 and β_1, in order to minimize the error (ϵ_i) of the sum of the squares of the distances (x) from each point to the regression; β_0 is the intersection variable, β_1 the steady-state slope and x_i is the real-time data from sensorgram.

$$y_i = \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij} + \epsilon_i, \qquad i = 1, 2, \cdots, n \tag{1}$$

Therefore, the transition regions are reveled when $\beta_1 \neq 0$. The regression coefficient β_1 is tracked within a predefined time window, and its behavior indicates if there is a substance (permanent regime) or a transition regime, as depicted in Fig. 5 (c).

For effective substance identification, the time window must be adjusted so that the present noise-level instrumentation does not characterize the occurrence of transition between substances (Kavitha et al., 2016). Fig. 5 (c) shows the time location and resonance position of the substances present in the Leishmaniasis positive test. The identified substance is expressed by the condition of stable resonance when a monolayer solution is in contact with the sensor.

The transition patterns are expressed when resonance condition still not stable when there are a fraction of two or more substances, and the resonance position λ_r is ascending or descending. In Fig. 5 (c) also depicted the complementary transition areas. Knowing the sensorgram and its regions, the intelligent SPR sensor discovers the temporal location and resonance position of both regimes. Added to the fact that the sensorgram has already been classified, the substances of the sensorgram are then identified.

B2 - Hypothesis Testing

To attest to the quality of the sensorgrams responses, a hypothesis test about the calculated regression coefficient is performed. The test verifies whether the regression coefficient, at the established time window, is contained in the acceptable confidence interval (CI). The CI certifies the quality level of the SPR-sensor response. Thus, it was tested whether the sensorgram and its regimes (substances) reflect with 99.9% confidence, the Leishmaniasis experimental protocol. The confidence level defines the upper and lower limits of the CI.

Fig. 6 (a) depicted the CI and the hypothesis test, which is performed on the Leishmaniasis positive sensorgram. When the hypothesis test (H_0) for the regression slope (β_1) is zero ($H_0 \Rightarrow \beta_1 = 0$, True) it means that the value of β_1 and its test are within in the established CI limits. Thereby, it is possible to confirm that the permanent regime has been correctly identified and refers to a substance. In the insert, highlight that in the interval between 870s to 1040s it was not possible to confirm with 99.9% confidence that it is a substance since the hypothesis test is out of the CI. Therefore, this time region refers to a transition between substances.

With the hypothesis test approach, it is possible to verify the behavior of an SPR sensor with an embedded intelligence module, and if the standardized-responses regulated by government agencies are been following.

B3 - Results for the Automated Testing

The embedded intelligence module also performs the task for automated testing on sensorgrams which express diagnosis protocols. With the previous learning of the positive sensorgram for the Leishmaniasis test (Fig. 5 a), its patterns as a group become identifiable, and differentiate it from a negative test becomes possible. An eventual new Leishmaniasis test is compared with the positive sensorgram previously-stored, through non-parametric bootstrap test (Souza et al., 2020) and counter-test with the parametric Student's t-test. Typically, the Student's t-test is performed with unknown differences between means (ϵ_i) and variations, and the non-parametric bootstrap computed with 2000 replications. By using statistical tests, the smart SPR sensor infer if the sample has a positive or negative response to Leishmaniasis, and thus performs an automated testing.

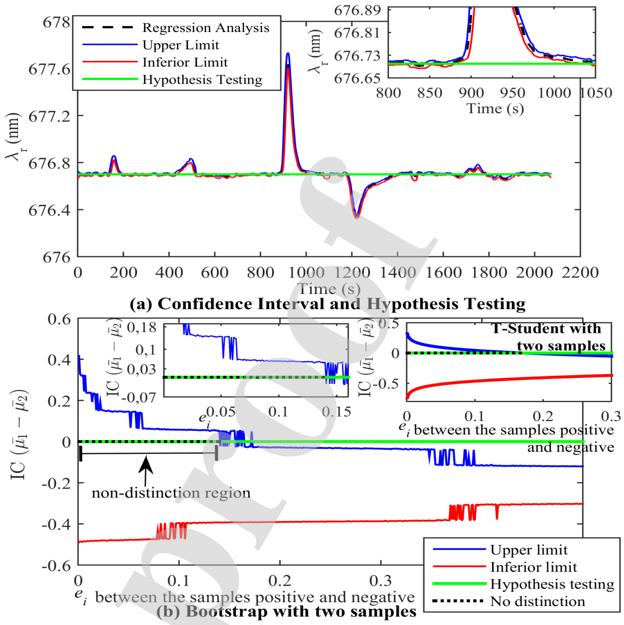

Fig. 6. (a) Hypothesis test to verify with 99.9% confidence the presence of a certain substance in the experimental protocol. The confidence level defines the upper and lower limits of the confidence interval, and the hypothesis test verifies if the regression coefficient is statistically zero, i.e., if the initial hypothesis is true in the established time-window. Inset shown that in the range 807s to 1040s it was not possible to confirm with 99.9% confidence the presence of a substance, indicating a transition regime. (b) Confidence interval for two sensorgram samples of the Leishmaniasis experimental protocol: one positive (database) and one negative (new test), as depicted in Fig. 5. Details for the non-distinction ϵ_i interval. Inset for the Student's t-test. For both statistical tests the hypothesis test was equivalent.

Fig. 6 (b) shows the confidence interval (CI) of the bootstrap test as a function of the error for the two sensorgram samples: one positive (database) and the other negative (new test). The hypothesis test means that the new test (new sensorgram) is statically equal to the positive sensorgram ($H_0 \Rightarrow IC < \epsilon_i$, True). In this case, there is no distinction between them, and thus, the tested sample is positive. For the sensorgrams expressed in Fig. 4, the non-distinction with 99.9% confidence occurs for $\epsilon_i <= 0.14$. The hypothesis test was equivalent for both statistical tests, as shown in the figure inset.

As the error increases, it is possible to observe the distinction between the tests. For the example case, values of $\epsilon_i > 0.14$ mean that the new test (new sensorgram) is statically different from the positive sensorgram, and the smart sensor affirms that this tested sample is negative.

4. Conclusion

Approaches for the uses of machine learning algorithms to describe, classify, analyze, identify, and automated diagnostic tests to create smart SPR sensors have been reported. The sensorgrams responses were acquired with the SPR sensors based on the PPBIO solution. To describe and classify the sensorgrams, the new descriptor so-called TSD was used, combined with the k-NN machine learning algorithm, to classify and recognize sensorgrams. The TSD and k-NN approach reveled a satisfactory results for distinct neighborhoods ($k = 3$ and $k = 5$) and different distance metrics. It is worth mentioning that the created new elements based on modifications in the experimental sensorgrams (temporal and vertical difference, noise, and changes in the resonance position), does not affect the classification and the robustness of the applied approaches.

To identify substances, attest the quality, and perform automated diagnostics, the ML-approach uses hypothesis testing with confidence intervals. For this, the linear regression segments satisfactorily the permanent and the transitions regimes of the sensorgram, enable the smart sensor to attests its response quality. With the substance identification, the smart sensor also predicting anomalies in the sensor behavior. For diagnostics represented by sensorgrams, the intelligent sensor automates the diagnostic tests. Taking as an example the positive Leishmaniasis' sensorgram test, a new test is compared with the positive one, and statistical verification is performed between the samples to verify if there is non-distinction.

The automation of the described intelligence tasks was embedded in the so-called *intelligence module*. Following the proposed ML-approaches, it is possible to assess the quality of SPR sensors and attest to the conditions under which it will operate with an established confidence level. All the described procedures guarantee the reliability of the responses provided by the smart SPR sensors, and thus eventual audits of regulatory agencies can be achieved satisfactorily.

Acknowledgment

The authors would like to thank CAPES and UFERSA for the financial support and research grants.

References

Agrawal, A., Singh, A., Yazdi, S., Singh, A., Ong, G.K., Bustillo, K., Johns, R.W., Ringe, E., Milliron, D.J., 2017. Resonant coupling between molecular vibrations and localized surface plasmon resonance of faceted metal oxide nanocrystals. Nano Letters 17, 2611–2620. pMID: 28337921.

Alfeilat, H.A.A., Hassanat, A.B., Lasassmeh, O., Tarawneh, A.S., Alhasanat, M.B., Eyal Salman, H.S., Prasath, V.S., 2019. Effects of distance measure choice on k-nearest neighbor classifier performance: A review. Big Data 7, 221–248. pMID: 31411491.

Amanatiadis, A., Kaburlasos, V., Gasteratos, A., Papadakis, S., 2011. Evaluation of shape descriptors for shape-based image retrieval. IET Image Processing 5, 493–499.

Basavaraju, A., Du, J., Zhou, F., Ji, J., 2020. A machine learning approach to road surface anomaly assessment using smartphone sensors. IEEE Sensors Journal 20, 2635–2647.

Canoa, P., Simón-Vázquez, R., Popplewell, J., González-Fernández, ., 2015. A quantitative binding study of fibrinogen and human serum albumin to metal oxide nanoparticles by surface plasmon resonance. Biosensors and Bioelectronics 74, 376 – 383.

Chomboon, K., Chujai, P., Teerarassammee, P., Kerdprasop, K., Kerdprasop, N., 2015. An empirical study of distance metrics for k-nearest neighbor algorithm. Proceedings of the 3rd International Conference on Industrial Application Engineering , 280–285.

Deza, M.M., Deza, E., 2016. Encyclopedia of Distances. 4 ed., Springer.

Ferreira, E., Lima, J., Alves-Balvedi, R.P., Bonan, P.R., Medeiros, E., Goulart, L., Lima, A., Neff, H., Oliveira, L., Castellano, L., Araújo, A., Moreira, C., 2017. Leishmania spp. detection using a surface plasmon resonance biosensor. Proceedings 1, 1 – 5.

Kavitha, S., Varuna, S., Ramya, R., 2016. A comparative analysis on linear regression and support vector regression , 1–5.

Kocaguneli, E., Menzies, T., 2013. Software effort models should be assessed via leave-one-out validation. Journal of Systems and Software 86, 1879 – 1890.

Lee, Y.K., Lee, K., Kim, W.M., Sohn, Y., 2014. Detection of amyloid-β42 using a waveguide-coupled bimetallic surface plasmon resonance sensor chip in the intensity measurement mode. Plos One 9, 1–7.

Moreira, C.d.S., Oliveira, L.C., Fischer, R., Medeiros, E.S., Lima, A.M.N., Neff, H., 2016. Polymer-based surface plasmon resonance biochip: construction and experimental aspects. Research on Biomedical Engineering 32, 92 – 103.

Mulak, P., Talhar, N.R., 2015. Analysis of distance measures using k-nearest neighbor algorithm on kdd dataset. International Journal of Science and Research , 2101–2104.

Oliveira, L.C., Lima, A.M.N., Thirstrup, C., Neff, H.F., 2019. A Materials Guide to Design, Characterization, Optimization, and Usage. volume 70. 2 ed., Springer International Publishing.

Oliveira, L.C., d. S. Moreira, C., Thirstrup, C., Melcher, E.U.K., Lima, A.M.N., Neff, H., 2013. A surface plasmon resonance biochip that operates both in the angular and wavelength interrogation modes. IEEE Transactions on Instrumentation and Measurement 62, 1223–1232.

Patil, P.O., Pandey, G.R., Patil, A.G., Borse, V.B., Deshmukh, P.K., Patil, D.R., Tade, R.S., Nangare, S.N., Khan, Z.G., Patil, A.M., More, M.P., Veerapandian, M., Bari, S.B., 2019. Graphene-based nanocomposites for sensitivity enhancement of surface plasmon resonance sensor for biological and chemical sensing: A review. Biosensors and Bioelectronics 139, 1–22.

Piliarik, M., Vaisocherová, H., Homola, J., 2009. Surface plasmon resonance biosensing 503, 65–88.

Souza, L.C.d., Souza, R.M.C.d.R., do Amaral, G.J.A., 2020. Dynamic clustering of interval data based on hybrid l_q distance. Knowledge and Information Systems 62, 687–718.

Tiwari, P.B., Üren, A., He, J., Darici, Y., Wang, X., 2015. Note: Model identification and analysis of bivalent analyte surface plasmon resonance data. Review of Scientific Instruments 86, 105–108.

Yan, Y., Chen, L., Tay, P.C., Tanaka, M.L., 2017. Molecule binding parameter estimation from surface plasmon resonance measurements 1, 1 – 9.

Zhang, H., Liu, W., Liu, Z., Ju, Y., Xu, M., Zhang, Y., Wu, X., Gu, Q., Wang, Z., Xu, J., 2018. Discovery of indoleamine 2,3-dioxygenase inhibitors using machine learning based virtual screening. Med. Chem. Commun. 9, 937–945.

Highlights

- Automated diagnosis mediated by SPR sensorgrams real-time responses
- High level strategy for processing information to create Smart SPR sensors.
- Characterization of the machine learning approaches to describe, classify, analyze, identify, and automated diagnostic tests using sensorgrams responses.
- Assurance of the response reliability provided by Smart SPR sensors, and satisfactorily achievement of regulatory agencies audit proceeds.
- New strategy to describe sensorgrams, the so-called TSD (Temporal Sensogram Descriptor)
- Possibility to assess the quality of SPR sensors and attest to the conditions under which it will operate with an established confidence level.

CRediT author statement

- **Julio Cartier Maia Gomes**

  Software development, Investigation, data visualization, validation, draft preparation.

- **Leandro Carlos Souza**

  Development, reviewing and editing, Machine Learning specialist.

- **Leiva Casemiro Oliveira**

  Supervision, conceptualization, development, draft preparation, writing- reviewing and editing, SPR specialist.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: