Overcoming the Bottleneck in COVID-19 Detection: A Machine-Learning Approach to Improve Accuracy of Electrochemical Impedance Spectroscopy (EIS) Detection Sensitivity

Roya Jafari[1], Zahra Esmaeilbeig[2] and Mohammadreza Nemati[3]*

[1]*Department of Chemistry, Illinois Institute of Technology, USA*

[2]*Department of Electrical & Computer Engineering, University of Illinois at Chicago, USA*

[3]*Department of Electrical Engineering and Computer Science, University of Toledo, USA*

ABSTRACT

In late 2019, coronavirus disease (COVID-19), emerged in Wuhan city, Hubei, China and claimed a large number of lives and affected billions all around the world. Accurate and scalable devices are essential for screening, diagnosis and monitoring of COVID-19 patients. In this Article, we first investigate the Electrochemical Impedance Spectroscopy (EIS) method and its limitations in detecting the positive COVID-19 cases. Afterwards, we demonstrate a machine learning model to determine the optimal parameters for designing an EIS method.

KEYWORDS: COVID-19; Artificial intelligence; Machine learning; Electrochemical impedance spectroscopy

INTRODUCTION

The severe respiratory syndrome coronavirus 2 (SARS-CoV-2) can be transmitted rapidly from human to human. The World Health Organization (WHO) has announced this disease as a pandemic due to the reproduction number 2.2 for SARS-CoV-2 [1]. As far as the implementations for terminating this pandemic is concerned, rapid diagnosis is very important to control the spread of the virus which can help patients to be monitored at the early stages of the infection to get quarantined and treated [2,3].

There are a lot of detection techniques for viruses. The simplest one is laboratory method. This method is based on the epidemiological history of the patient and ancillary exams findings such as CT-scan [4]. This method is applicable only after severe signs of the infection. The most regular method which is used vastly is the quantitative Reverse- Transcription Polymerase Chain Reaction (RT-PCR). In this method RNA probes which have specific sequences are connected to the matched part of the RNA of the virus. After matching RNA with the RNA probe, copies of the virus RNA are made. This method is time consuming since amplification is required [5]. The other used method is Enzyme Immunoassay (EIA) or Enzyme-Linked Immunosorbent Assay (ELISA). ELISA is based on the antibody and enzyme interaction which will give a florescence signal. This serological test can be done only after passing a few days of infection since the antibody made by the immune system is going to be detected [6-8].

There are also bio-sensing methods which can be applied as soon as the patient is infected. These methods can sense either DNA/RNA of the virus or its structural proteins. The most concern for these methods is the low sensitivity due to the high number of intrusive factors. The most rapid and low-cost bio-sensing method is Electrochemical Impedance Spectroscopy (EIS). EIS

uses alternating current (AC) signal to investigate the impedance characteristics of a cell. The AC signal is probed by application of voltage *v* at frequency f to electrodes with spacing d between them to generate the impedance changes ΔZ versus time [9-12].

The schematic diagram of an EIS device is illustrated in Figure 1 that shows the number of intrusive resistances [13]. In this study, our goal is to develop a method for improving EIS sensitivity based on a regression model borrowed from machine learning. Particularly, our model can predict the ΔZ based on given parameter values for *d, f, v* and other relevant parameters. While the parameters of our model are limited to the ones mentioned above, our model can easily be extended to determine other parameters interfering in EIS such as sample handling and time of incubation of the sample.
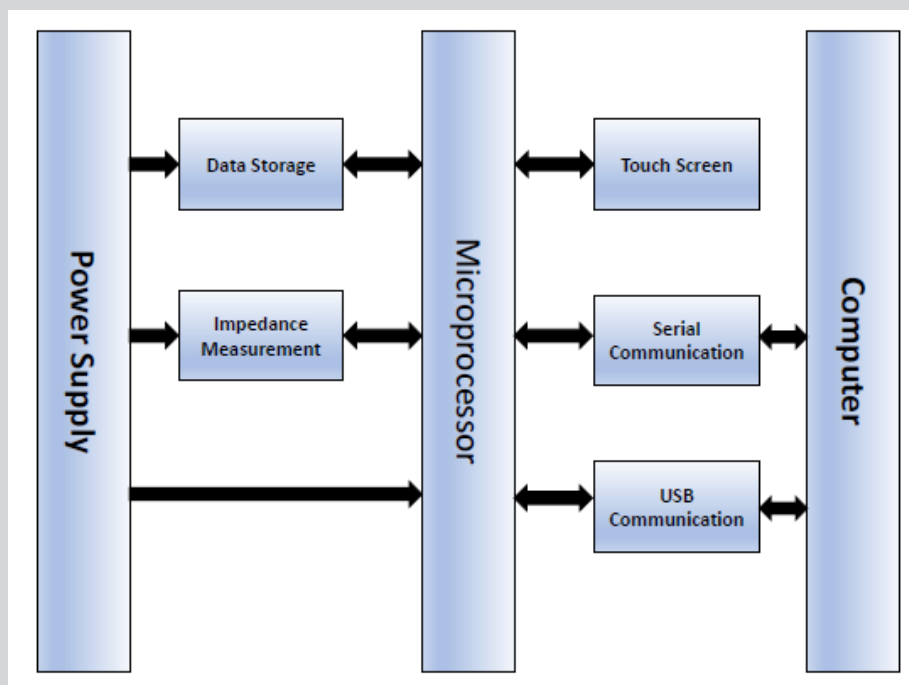

Figure 1: Schematic diagram of the hardware structure of the Impedance detector.

EIS Detection of COVID-19

As all of the connections in the EIS device have their own resistance, the effect of their resistance on ΔZ is a limitation of this technique which also is variable for different devices with different connection parts and materials. In EIS detection, there is significant change in total impedance of the system within just a few seconds after introducing the sample. This is because of the interaction of the protein/RNA of the virus with the electrode which changes the properties of the double layer such as charge distribution, dimension, and shape. This alteration impacts dielectric constant and the conductivity of double layer. Therefore, the overall impedance of the system is changed [14]. The difference of impedance ΔZ in this technique is considered as the detection item i.e., if ΔZ is greater than a threshold, a positive COVID-19 case will be reported.

After binding the protein/DNA with the probes in the electrode surface, the impedance is increased from 0 to 0.25 ω [13]. It should be considered that the change in impedance is dependant on the applied frequency. Dipoles will not react with the overall impedance if the applied signal is more than 1 MHz. In general, the changes in the impedance is more detectable in lower signal frequencies such as 10 KHz. Therefore, we are motivated to develop a method to predict the impedance change Z as a function of frequency *f*.

There are some other factors that limit the detection sensitivity of the EIS technique. The hardware noise, the spacing between interdigitated electrodes, the electrode type, and electrode coating of insulative material are some examples factors that change the sensitivity of the device.

In the next section, we propose our method based on machine learning to process the data of impedance change of all of the previous tested people with inserting all of the parameters such as electrode type, electrode coating, spacing between interdigitated electrodes, voltage applied, etc., that limit the sensitivity of the method. By processing this data, we can adjust the impedance change to a more accurate number.

Proposed Method

The mentioned limitation of the EIS detection for COVID-19 motivated us to develop an algorithm for optimizing the EIS procedure. Recent advances in hardware design have made machine learning algorithm very popular where a sufficient amount of data is available [16,17]. Also, Artificial intelligence and in particular machine learning provide a promising tool that can be extended in pinpointing the problems rising from COVID-19 pandemic [18].

Assume we have features as frequency *f*, Voltage *v* spacing between the electrodes d, time of sample incubation as t, material of the electrodes, and the threshold of Impedance change ΔZ indicating a positive COVID-19 case as τ. We propose a supervised learning model to determine the optimum threshold value τ according to the given features. In Machine learning, such an algorithm to predict a parameter of interest from given experiments is known as regression. Namely, our goal is to predict τ from a set of given parameters *x.*

Deep neural networks (DNN) are one of the most studied approaches due to their high accuracy. In our study, a regression model based on artificial neural networks can be mathematical

demonstrated as

$$\hat{\tau} = F(x, W), \qquad (1)$$

where W are weights of the neural network and x is the vector of features according to which we aim to predict τ. Fitting this model on our dataset i.e. determining the weights of the DNN comprises the following steps:

Data Preparation

In order to prepare a comprehensive dataset for our model, a large amount of data from three categories will be collected. In particular, in the case of detecting the COVID-19 using EIS, one needs to prepare a dataset consisting of not only COVID-19 cases, but also healthy cases as well as hypothetically similar viral cases that resemble similar symptoms as of COVID-19 cases. Such a comprehensive dataset further enables the model to properly determine the optimized parameters for EIS detection. The n combination of features will then go through the EIS detection procedure designed according to pre-recorded feature values x_i = [f, d, t, v, ..], i = 1, . . . , n. The impedance change for each case is recorded as

Next, the collected data is required to be annotated by an expert to determine the threshold τ for the positive cases and preparing the data for the next step which is training the DNN.

Training and Testing the DNN

In order to train our deep neural network, we use the positive COVID-19 cases and minimize the loss function to determining the thresholds τ_i.

$$L(W) = \frac{1}{N} \sum_i (\hat{\tau}_i - \Delta Z_i)^2 \qquad (2)$$

During the test set for any value of the feature vector x, we apply the model to get the optimal value for parameter of interest (threshold τ for impedance change) predicted by the deep neural network [19].

Limitation of our Proposed Method

In the case of the new COVID-19 virus, there is a lack of high quality and well-annotated datasets. On the other hand, there exist a vast amount of information and dataset on other pneumonia-related diseases (e.g. Influenza-A).

The results of our proposed method are awaiting the proper dataset for clinical testing centers and we invite the scientific community in to contribute in preparing a proper dataset for the proposed study in this paper.

After preparing the dataset and designing the DNN model proposed in this paper, there exist one essential challenge that one might face in all other aforementioned stages for training and testing the DNN: In machine learning, the testing and training stages are closely related and a successful training might not result in a good performance in testing stage. For instance, if the neural network has a parameter W with high dimension to be optimized and the training set has a limited size, the model is susceptible to "generalization error" also known as "over-fitting", leading to a small training error and a large test error. In such a case, one needs to either provide more data via either collecting new samples or using advanced techniques used by machine learning engineers such as data augmentation techniques, or to change the underlying model entirely.

CONCLUSION

Electrochemical Impedance Spectroscopy (EIS) is a well-established and known procedure for detecting different types of biomolecules such as proteins and nucleotides. However, it has lower sensitivity in comparison to the traditional methods such as PCR. Thus, it is newly introduced for COVID-19 detection. As a result, the improvement of this method is highly promising for the rapid and low-cost detection of the COVID-19. In this paper, we proposed a method using machine learning algorithms to improve the EIS technique based on a dataset of previously tested cases.

REFERENCES

1. Bergna LA, Acciarri C, Galli M, Zehender G (2020) Early phylogenetic estimate of the effective reproduction number of sars-cov-2. Journal of Medical Virology 92(6): 675-679.

2. Bikdeli B, Madhavan MV, Gupta A, Jimenez D, Burton JR, et al. (2020) Pharmacological agents targeting thromboinflammation in covid-19: review and implications for future research. Thrombosis and haemostasis 120(7): 1004-1024.

3. Shang J, Ye G, Shi K, Wan Y, Luo C, et al. (2020) Structural basis of receptor recognition by sars-cov-2. Nature 581 (2020): 221-224.

4. Ai T, Yang Z, Hou H, Zhan C, Chen C, et al. (2020) Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (covid-19) in china: a report of 1014 cases. Radiology 296(2): E32-E40.

5. Fouchier RA, Bestebroer TM, Herfst S, Van Der Kemp L, Rimmelzwaan GF, et al. (2000) Detection of influenza a virus from different species by pcr amplification of conserved sequences in the matrix gene. J Clin Microbiol 38(11): 4096-4101.

6. Adams ER, Anand R, Andersson MI, Auckland K, Baillie JK, et al. (2020) Evaluation of antibody testing for sars-cov-2 using ELISA and lateral flow immunoassays. MedRxiv.

7. Liu W, Liu L, Kou G, Zheng Y, Ding Y, et al. (2020) Evaluation of nucleocapsid and spike protein-based enzyme-linked immunosorbent assays for detecting antibodies against sars-cov-2. Journal of Clinical Microbiology 58(6): e00461-e00420.

8. Udugama P, Kadhiresan HN, Kozlowski A, Malekjahani M, Osborne V, et al. (2020) Diagnosing covid-19: the disease and tools for detection, ACS nano 14(4): 3822-3835.

9. Fu Y, Callaway Z, Lum J, Wang R, Lin J, et al. (2014) Exploiting enzyme catalysis in ultra-low ion strength media for impedance biosensing of avian influenza virus using a bare interdigitated electrode, Analytical chemistry 86(4): 1965-1971.

10. Singh KV, Whited AM, Ragineni Y, Barrett TW, King J, et al. (2010) 3d nanogap interdigitated electrode array biosensors. Analytical and Bioanalytical Chemistry 397(4): 1493-1502.

11. Rahimi R, Mehrehjedy A, Zargari S (2017) Bivo4/mn3o4 a novel p–n heterojunction photocatalyst functionalized with metalloporphyrins: Synthesis, charge transfer mechanism, and enhanced visible-light photocatalysis for degradation of dye pollutant. Environmental Progress & Sustainable Energy 36: 1439-1448.

12. Bonanni AH, Loo M, Pumera (2012) Graphene for impedimetric biosensing, TrAC Trends in Analytical Chemistry 37: 12-21.

13. Wang X, Zhao Z, Wang Y, Lin J (2020) A portable impedance detector of interdigitated array microelectrode for rapid detection of avian influenza virus, In: International Conference on Computer and Computing Technologies in Agriculture, Springer, pp. 247-256.

14. Hnaien M, Diouani MF, Helali S, Hafaid I, Hassen WM (2008) Immobilization of specific antibody on sam functionalized gold electrode for rabies virus detection by electrochemical impedance spectroscopy. Biochemical Engineering Journal 39(3): 443-449.

15. Esmaeilbeig Z, Ghaemmaghami S (2018) Compressed video watermarking for authentication and reconstruction of the audio part, In: 2018 15th International ISC (Iranian Society of Cryptology) conference

on information security and cryptology (ISCISC), IEEE, pp. 1-6.

16. Ajirlou F, Partin-Vaisband I (2020) A machine learning pipeline stage for adaptive frequency adjustment. arXiv preprint arXiv:2007.01820.

17. Nemati M, Ansary J, Nemati N (2020) Machine-learning approaches in covid-19 survival analysis and discharge-time likelihood prediction using clinical data. Patterns 1(5): 100074.

18. Dana PA, Esmaeilbeig Z, Sadeghi MR (2020) Reliability enhancement and packet loss recovery of any steganographic method in voice over IP. Wireless Networks.