

## Article

# Machine Learning Techniques for Effective Pathogen Detection Based on Resonant Biosensors

Guoguang Rong , Yankun Xu  and Mohamad Sawan \* 

CenBRAIN Neurotech Center of Excellence, School of Engineering, Westlake University, 600 Dunny Road, Xihu District, Hangzhou 310030, China; rongguoguang@westlake.edu.cn (G.R.); xuyankun@westlake.edu.cn (Y.X.)

\* Correspondence: sawan@westlake.edu.cn

**Abstract:** We describe a machine learning (ML) approach to processing the signals collected from a COVID-19 optical-based detector. Multilayer perceptron (MLP) and support vector machine (SVM) were used to process both the raw data and the feature engineering data, and high performance for the qualitative detection of the SARS-CoV-2 virus with concentration down to 1 TCID<sub>50</sub>/mL was achieved. Valid detection experiments contained 486 negative and 108 positive samples, and control experiments, in which biosensors without antibody functionalization were used to detect SARS-CoV-2, contained 36 negative samples and 732 positive samples. The data distribution patterns of the valid and control detection dataset, based on T-distributed stochastic neighbor embedding (t-SNE), were used to study the distinguishability between positive and negative samples and explain the ML prediction performance. This work demonstrates that ML can be a generalized effective approach to process the signals and the datasets of biosensors dependent on resonant modes as biosensing mechanism.

**Keywords:** machine learning; support vector machine; multilayer perceptron; photonic biosensor; signal processing; Tamm plasmon polariton; localized surface plasmon resonance

## 1. Introduction

The global COVID-19 pandemic has had a huge impact on the world's health and economy [1]. The fast-spreading virus SARS-CoV-2 virus is the main culprit of this phenomenon, and detection of the virus in human populations is crucial for curbing the pandemic [2]. Traditional detection approaches include the nucleic acid amplification test (NAAT) [3] and antigen detection [4] techniques. Currently, the mainstream is quantitative polymerase chain reaction (qPCR) [5], which is a kind of NAAT that has high sensitivity and specificity, but requires a clean environment, bulky and expensive equipment, and trained personnel. Therefore, qPCR is not suitable for onsite, fast turnaround detection or for population-scale screening, which are often required in pandemic control scenarios [6]. To complement qPCR tests, antigen detection based on lateral flow [7] has also been employed in both home use and self-testing. However, antigen detection is limited in detection sensitivity and specificity, hindering its efficacy in fighting a pandemic [8]. There is still a lack of rapid, accurate, and low-cost detection techniques that can be deployed onsite for population-scale epidemic screening and/or surveillance [9], especially for regions with limited resources [10].

Biosensors have been proposed for the detection of SARS-CoV-2 [11]. Biosensor technologies have high sensitivity, good specificity, fast turnaround, ease of operation, low cost, and onsite deployment capability [12,13]. We have previously proposed a photonic biosensor with high sensitivity and specificity for the fast, on-site detection of SARS-CoV-2 [14,15]. The biosensor is based on a nanoporous silicon material fabricated via a CMOS-compatible silicon process and nanophotonic working principles of localized surface plasmon resonance (LSPR) [16] and Tamm plasmon polariton (TPP) [17,18]. The measurement of the biosensor is based on reflection spectroscopy [14].



**Citation:** Rong, G.; Xu, Y.; Sawan, M. Machine Learning Techniques for Effective Pathogen Detection Based on Resonant Biosensors. *Biosensors* **2023**, *13*, 860. <https://doi.org/10.3390/bios13090860>

Received: 3 August 2023

Revised: 24 August 2023

Accepted: 29 August 2023

Published: 31 August 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

We also developed handheld and high-throughput detection systems [19] that can collect the reflection spectrum of biosensors and process the spectral data to determine the detection results efficiently. The high-throughput detection system is suitable for a population-scale screening of infection, and the handheld detection system is for home use or self-tests. The spectral data processing algorithm works by recognizing the characteristic resonant valleys in the reflection spectrum of the biosensor and determines the detection results by judging if there is spectral red shift in the characteristic resonant valleys. This is the often-used and so-called “find peaks” technique, with its name originating from the MATLAB function `findpeaks()`. This technique can also be implemented on field programmable gate arrays (FPGAs) for fast and efficient processing of signals from an array of biosensors [20]. In addition, researchers have proposed the interferogram average over wavelength (IAW) technique to process the signals of optical biosensors that depend on a spectral shift in the characteristic resonant features, which can achieve sensitivity enhancement compared with spectral shift detection [21]. Detection of changes in reflection intensity due to a shift in spectral features in the spectrum has also been used to detect biomolecules in real time [22]. However, both IAW and light-intensity measurement techniques are subject to spectral amplitude fluctuations and thus require highly stable spectroscopy systems, such as a stable light source and high signal-to-noise ratio spectrometers.

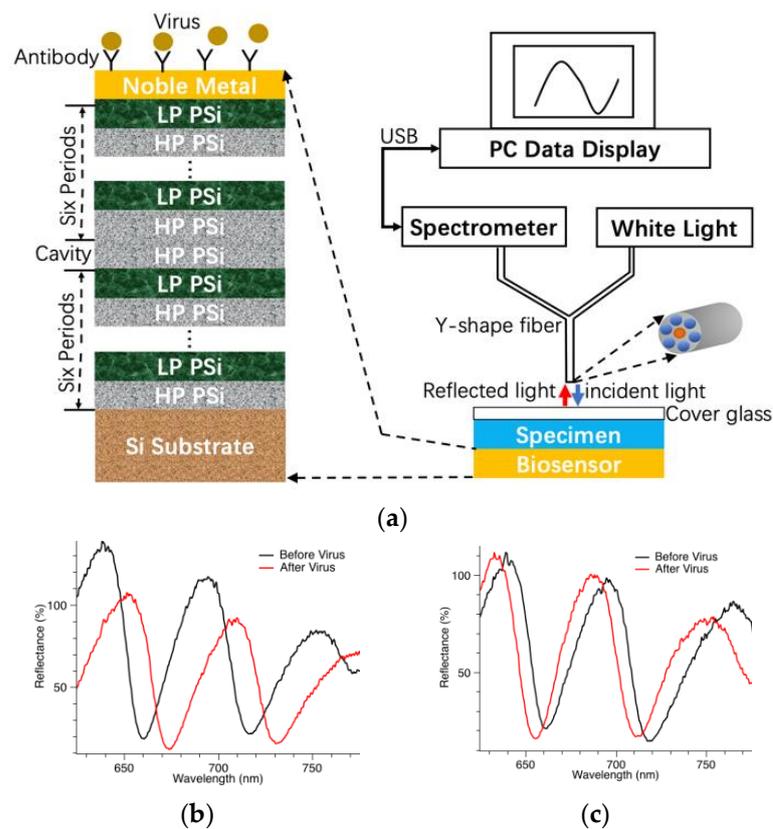
In this work, we demonstrate that it is advantageous to utilize artificial intelligence technology, more specifically machine learning (ML) algorithms, to process the spectral data of the biosensor [23]. Instead of depending on programming, its algorithm is learnt from a big volume of data [24]. Machine learning has been used for computer vision [25], face recognition [26], autonomous driving [27,28], auxiliary decision making [29,30], brain–machine interface [31], cancer diagnosis and assessment [32], and chess game [33]. It includes supervised learning, unsupervised learning, and reinforcement learning [34]. Supervised learning (SL) is an algorithm that learns from massive, labeled datasets and generates prediction models that can work to generate labels for new datasets. SL includes support vector machine (SVM) [35], multilayer perceptron (MLP) [36], linear regression [37,38], linear discriminant analysis [39,40], K-nearest neighbor [41,42], decision tree [43,44], and naïve Bayes [45,46]. In this work, we demonstrate that SVM and MLP can be used for processing of the photonic biosensor signal and dataset. Compared with previously proposed techniques, the ML technique has the following advantages: (1) there is no need to find the appropriate parameters of the algorithm, e.g., the `findpeaks()` function, in a trial-and-error way to guarantee accurate recognition of spectral features; (2) there is no need to discriminate between redshift or blueshift, which can be an extra issue in algorithm design; (3) it is not sensitive to spectral amplitude fluctuations, so the requirements for stable and expensive hardware are relaxed; (4) it is generalizable to all kinds of sensors with salient features in the response signal, which serve as the basis for discriminating between positive and negative responses.

Data visualization approaches can help us to understand the distribution of the dataset and discover the distinguishability of the dataset. T-distributed stochastic neighbor embedding (t-SNE) is a prevalent approach to map high-dimensional data to low-dimensional embedding [47]. In this contribution, we also implemented the t-SNE approach on a SARS-CoV-2 detection dataset to clarify the distinguishability of the biosensor dataset so that a better understanding of the data processing and ML prediction performances could be obtained.

## 2. Materials and Methods

### 2.1. Biosensor Working Principle and Measurement Setup

As shown in Figure 1a, the biosensor is basically a porous silicon microcavity consisting of two Bragg reflectors and one resonant cavity [14,48].



**Figure 1.** Photonic biosensor: (a) structure and its reflection spectroscopy measurement; (b) typical example of redshift showing resonant valleys in reflection spectra; (c) blueshift of resonant valleys in reflection spectra.

One Bragg reflector is six periods of alternating low porosity (LP) and high porosity (HP) porous silicon (PSi) thin films of thickness equal to a quarter resonant wavelength. Noble metal thin film is deposited on top of the porous silicon. Because of the nanoporous structure of the porous silicon material, the conformally deposited noble metal thin film is also porous. When light is incident on the surface of the biosensor, some of its energy is coupled into localized surface plasmon resonance (LSPR) [48] supported by the nanostructures of the noble metal thin film. In addition, some of its energy also couples into the Tamm plasmon polariton (TPP) supported by the interface between the top Bragg reflector and the noble metal thin film [49]. Therefore, the LSPR and TPP are simultaneously excited by the incident light and couple with each other, forming a strong field confinement around the noble metal thin film. If specific antibodies are immobilized beforehand on the surface of the noble metal, they can capture the SARS-CoV-2 virus specifically. Such binding events cause an addition of biomaterials around the noble metal thin film, and the added biomaterial interacts strongly with the coupled LSPR and TPP field. This is the working mechanism of the biosensor for the sensitive detection of the virus. As shown in Figure 1a, in order to measure the signals of the biosensor, reflection spectroscopy is used. A white light source provides the incident light, which passes through the Y-shape fiber and shines vertically onto the biosensor surface. The light reflected from the biosensor surface is collected by the Y-shape fiber and passes into the spectrometer for data analysis. The Y-shape fiber consists of six circumferential fibers guiding incident light, and one central fiber guiding reflected light.

Figure 1b,c show the representative reflection spectra of the biosensor. They have characteristic resonant valleys that are in the spectral range of 600–800 nm in wavelength. If there are viruses binding with antibodies on the biosensor surface, the binding events cause a shift in the spectral features to a longer wavelength, which is called “redshift”. For

example, Figure 1b shows such a case where the virus binds with antibodies, redshift occurs, and the detection result is determined to be positive. On the other hand, if there is no virus binding with antibodies on the biosensor surface, there is no shift in the spectral features, i.e., almost overlapping spectra for both before and after binding reaction. In the third case, there could also appear a shift in the resonant features to a shorter wavelength, which is also called “blueshift”. In such cases, the detection result is determined to be negative. Figure 1c shows an example of blueshift. In summary, the principle of the biosensor is based on interactions between biomaterials and photonic energy, and the detection result is determined based on a shift in the spectral features in the optical spectrum collected from the reflection spectroscopy measurement.

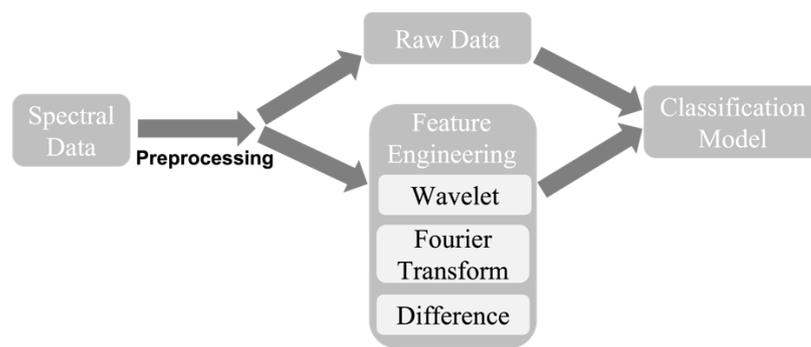
## 2.2. Data Preprocessing

This dataset was obtained from detection experiments of inactivated SARS-CoV-2 in clinical swab specimens, with virus concentrations as low as 1 TCID<sub>50</sub>/mL [14]. Figure 1 shows example spectra of the biosensor for positive and negative detection results. For the positive result, there is spectral redshift; and for the negative result, there is either no spectral shift or there is spectral blueshift. The experimental data were collected via reflection spectroscopy with the corresponding spectra for before and after applying specimens on the biosensor surface. Each spectral data contained 2048 data points representing reflection intensities, with a data-to-data spacing of 0.48 nm in the wavelength range of 200–1200 nm. We usually needed to carry out preprocessing of the spectral data before the data analysis, which included normalization and artifacts removal. Furthermore, normalization was implemented on each of the data samples for the purpose of training convenience. Spectral data of both before and after adding specimens were combined as a single sample, so that the size of the reformed sample was  $2 \times 2048$ , or 4096. Each detection experiment was regarded as a sample for either training or testing purposes. After several outliers were removed to clean the dataset, there were 486 negative samples and 108 positive samples left in total for the classification model training and prediction test.

## 2.3. Feature Engineering

As shown in Figure 2, the input to the model was 4096 data. This required 4096 input neuron nodes, which could be a computational burden. In addition to this raw data approach, the input could also contain features extracted from the data. We propose feature engineering methods comprising three different approaches—wavelet transform, Fourier transform, and spectral difference. For the wavelet domain, we used the wavelet transform with scales of 30 and took the average of each scale, which generated 30 features for each spectral curve. Two curves (before and after virus) generated 60 wavelet-based features. In terms of the Fourier domain, we found that most information appeared in the low-frequency range (<50 Hz), so that we took the average of each 5 Hz in order from 0 to 50 Hz, so that 10 features for each spectral curve and 20 features for spectra pairs were obtained in the Fourier domain. For the spectral difference, we utilized the difference between the spectral data before and after the binding reaction on the biosensor, instead of two separate spectra. There were three features selected from the spectral difference: mean, variance, and sign change rate.

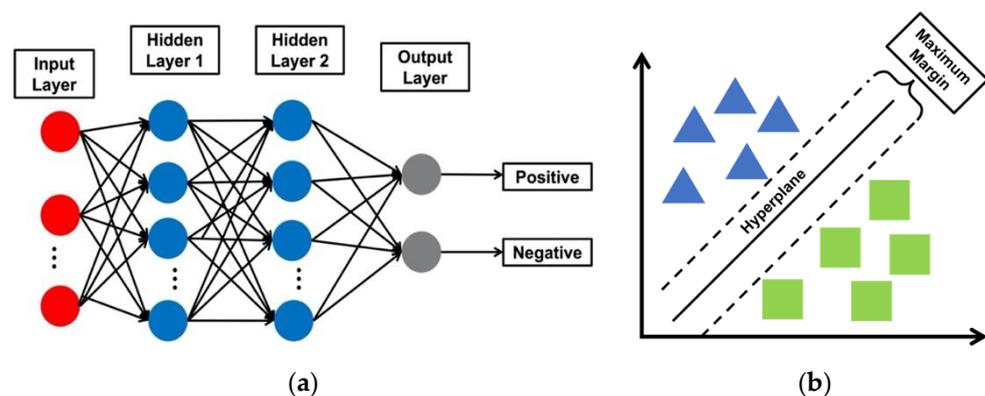
Eventually, for each training sample containing spectral data of before and after the reaction, the wavelet transform and Fourier transform domain features needed the spectra of both before and after the reaction, and spectral difference features only needed the difference between the spectra before and after the reaction. Therefore, there were 83 (60 wavelet domain + 20 Fourier domain + 3 spectral difference) features selected for the classification experiments.



**Figure 2.** Simplified block diagram of the data-processing procedure. Raw data and feature engineering methods were used in the experiments.

#### 2.4. Classification Models

All the samples were randomly shuffled and separated as 70% for training and 30% for testing. This allocation ratio is a practical standard for benchmark performance. Multilayer perceptron (MLP) and support vector machine (SVM) models were used since they are usually considered as efficient ML models capable of achieving baseline performance. As shown in Figure 3, in terms of the MLP model, two hidden layers with 100 and 50 neuron nodes with a sigmoid activation function were implemented, with the optimizer as a stochastic gradient descent solver, and the learning rate and epoch set as 0.1 and 30, respectively. The number of layers and the number of neurons in each layer were optimized through N ablation study, in which different numbers of layers and different pairs of the number of neurons were tested. Finally, we found that the two layers with 100 and 50 neurons were expected to be the best in final performance and learnable parameters. Decreasing the number of neurons decreased the prediction accuracy slightly. As for the SVM model, we set the gamma parameter of the radial basis kernel function as 1.



**Figure 3.** (a) Simplified diagram for illustration of MLP architecture; (b) SVM explanation.

#### 2.5. Control Experiments

For the control experiments, we detected SARS-CoV-2 specimens with photonic biosensors, which did not have specific antibodies immobilized on the biosensor surface beforehand. There were a total of 732 data samples for detecting SARS-CoV-2 virus specimens of various concentrations, and 36 data samples for detecting specimens containing no SARS-CoV-2 viruses. This new dataset was processed using the already trained SVM and MLP models, as shown in Figure 3.

#### 2.6. Dataset Distinguishability Analysis

Nowadays, data visualization approaches can help us understand the distribution of a dataset and intuitively investigate whether a dataset is distinguishable or not. T-distributed stochastic neighbor embedding (t-SNE) is a tool to visualize high-dimensional data. It

converts similarities between data points to joint probabilities and tries to minimize the Kullback–Leibler divergence [50] between the joint probabilities of the low-dimensional embedding and the high-dimensional data.

We implemented the t-SNE tool on the specimen detection datasets to interpret the distinguishability of the datasets. The data distribution patterns could help interpret the performance of the models on the dataset. Both the raw dataset and the features extracted from the raw data were considered in terms of their distinguishability. We also investigated whether the extracted features had distributions different from those of the raw data.

### 3. Results and Discussion

In terms of the experiments, we used SVM and MLP models to test the raw data processing and feature engineering method. Two performance metrics were considered in the experiments, sensitivity (SEN) and specificity (SPE), which are defined as

$$\text{SEN} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (1)$$

$$\text{SPE} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (2)$$

where TP, FN, TN, and FP stand for true positive, false negative, true negative, and false positive, respectively.

Table 1 shows the performance of the ML model predictions. We can see from the last two rows that perfect performance was achieved for both the raw data and the feature engineering methods, combined with either the SVM or MLP model. The fourth and fifth rows in Table 1 show the performance of the models in processing the control experiment dataset. The performance was very poor, and this was due to the fact that the biosensors had not been functionalized with specific antibodies and, thus, could not detect the SARS-CoV-2 virus effectively. The low *p*-values for both the control detection and valid detection datasets demonstrate the reliability of the classification of the two types of sample datasets.

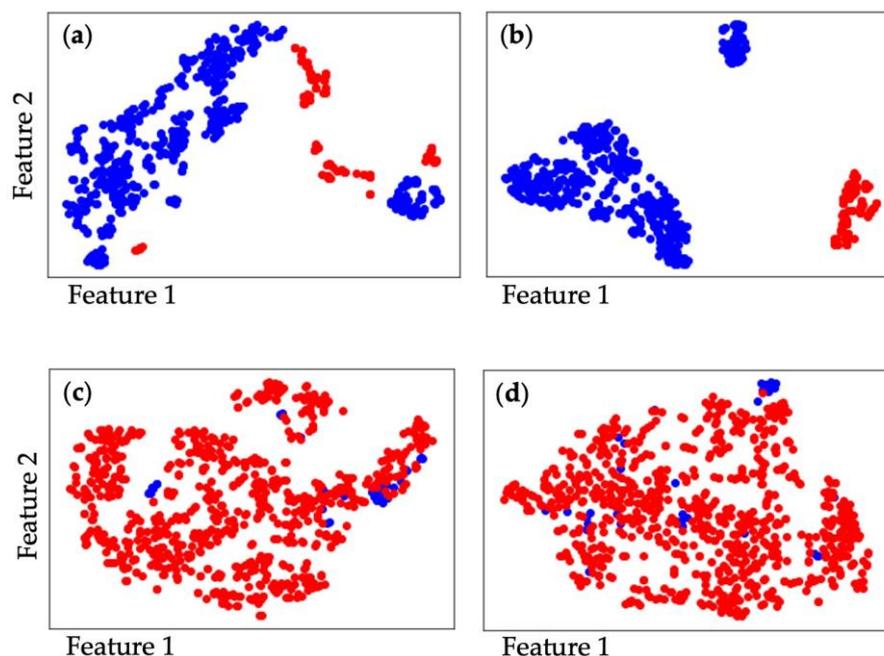
**Table 1.** Performance of raw data and feature engineering processing methods with two machine learning models.

Method	Raw Data						Feature Engineering					
	SVM			MLP			SVM			MLP		
Parameter	SEN	SPE	AUC	SEN	SPE	AUC	SEN	SPE	AUC	SEN	SPE	AUC
Performance on Control Detection Data	100%	46%	60.1%	86%	46%	63.4%	29%	33%	38.8%	27%	32%	38.1%
<i>p</i> -Value for Control Detection Data							<0.05					
Performance on Valid Detection Data							100%					
<i>p</i> -Value for Valid Detection Data							<0.05					

SVM: support vector machine; MLP: multilayer perceptron; SEN: sensitivity; SPE: specificity; AUC: area under ROC curve; ROC: receiver operating characteristic.

Figure 4a shows the data distribution of the raw datasets in 2D space with the t-SNE data visualization approach. We can see that the positive and negative samples from the dataset of the valid detection experiments are clustered without any overlapping. Thus, the valid experimental dataset is distinguishable. Figure 4b shows the data distribution of the features extracted from the dataset in Figure 4a. The extracted features changed the data distribution while maintaining distinguishability because the samples were separated into different clusters. Figure 4c shows the data distribution of the datasets obtained from the control experiments wherein biosensors were not functionalized with specific antibodies. Negative samples overlap the positive samples, and the dataset is indistinguishable

according to the visualization results. Figure 4d shows the data distribution of the features extracted from the dataset in Figure 4c. The distribution of the features' dataset is still mixed up, so that feature engineering cannot help the dataset to be classified effectively. These dataset distribution results can serve to interpret the performance comparisons demonstrated in Table 1.



**Figure 4.** The t-SNE data visualization results of experimental SARS-CoV-2 detection dataset; red and blue represent the positive and negative samples, respectively. (a) Raw dataset of valid detection experiment; (b) feature engineering dataset of valid detection experiment; (c) raw dataset of the control detection experiment; (d) feature engineering dataset of control detection experiment.

Table 2 demonstrates the advantages of the ML data processing technique compared with other techniques. The general advantages of ML are valid, in addition to the eased hardware requirement.

**Table 2.** Comparison of machine learning techniques with other signal processing techniques.

Technique	Factor	Need Data Filtering and Denoising	Need to Take Care of Shift Direction	Need Stable Light Source and Low Noise Spectroscopy System	Needed Researcher Work
Find peaks and calculate spectral shift		Yes	Yes	No	Algorithm design and test
Interferogram average over wavelength		Yes	No	Yes	Algorithm design and test
Intensity interrogation		Yes	No	Yes	Algorithm design and test
Machine learning		Yes	No	No	Model training from data

To verify the efficacy of the ML data processing technique for biosensors, detection experiments of inactivated SARS-CoV-2 at the vaccination sites of the Hangzhou Center for Disease Control and Prevention (CDC) were carried out, and the detection results were compared with the gold standard: reverse-transcription qPCR technique. The environmental specimens were collected from various locations at different vaccination sites, delivered to Hangzhou CDC within 4 h, and were simultaneously analyzed using both techniques. Table 3 shows that the biosensors, together with the ML data processing, generated detection results that were consistent with the qPCR results. Note that qPCR

provides semiquantitative results dependent on the Ct value [5], while the ML processing of biosensor data only provides qualitative results. This comparative study demonstrates that the ML technique is an effective tool for biosensor signal and data processing,

**Table 3.** Comparison of detection results of inactivated SARS-CoV-2 at vaccination sites of Hangzhou CDC using both qPCR technique and biosensor with ML technique.

Specimen Collection Location		qPCR Result	Biosensor with ML Result
Vaccination Site 1	Operation Desktop	Weak positive	Positive
Vaccination Site 1	Vaccination Station	Strong positive	Positive
Vaccination Site 2	Operation Desktop	Weak positive	Positive
Vaccination Site 2	Vaccination Station	Weak positive	Positive
Vaccination Site 2	Ventilation Plate	Strong positive	Positive
Vaccination Site 2	Inoculation Table Handle	Weak positive	Positive
Vaccination Site 4	Keyboard and Mouse	Negative	Negative
Vaccination Site 5	Pen and White Board	Strong positive	Positive
Vaccination Site 55	Inoculation Table Handle	Negative	Negative
No. 4 and No. 5 Inoculation Desk Room	Door Handle and Switch	Negative	Negative
Other	Hemostatic Swab	Weak positive	Positive
Other	Cleaner's Hand	Negative	Negative

#### 4. Conclusions

In this work, machine learning techniques were used to process the signals and datasets of photonic biosensors. Both SVM and MLP were used to process raw data and future engineering data, and perfect results were obtained that distinguished between negative and positive detections. Control experiments were also carried out, wherein biosensors not functionalized with specific antibodies were used to detect SARS-CoV-2 virus. Both the SVM and MLP models trained with valid experimental data could not distinguish between the negative and positive detections in the control experiments. To demonstrate the distinguishability of the raw data and the future engineering data for both valid experiments and control experiments, we implemented a t-SNE data visualization approach. The results showed that the valid experimental dataset was distinguishable, and the control experimental dataset was indistinguishable according to both the raw data and the feature engineering methods. The results were consistent with the data processing performance of machine learning techniques achieved for the valid experimental dataset and the control experimental dataset. Future research will focus on ML techniques for the determination of quantitative detection results so that the quantity of target biospecies in specimens can be obtained. ML can be a powerful tool in processing the signals and datasets of biosensors for which there are salient features in the response signals of such biosensors. This includes optical, electrochemical, thermal, and mechanical biosensors.

**Author Contributions:** Conceptualization, G.R. and M.S.; methodology, G.R. and Y.X.; software, Y.X.; writing—original draft preparation, G.R.; writing—review and editing, M.S. and Y.X.; supervision, M.S.; funding acquisition, M.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Leading Innovative and Entrepreneur Team Introduction Program of Zhejiang grant number 2020R01005, Westlake University grant number 10318A992001, Tencent Foundation grant number XHTX202003001, and Zhejiang Key R&D Program grant number 2021C03002. The APC was funded by the Leading Innovative and Entrepreneur Team Introduction Program of Zhejiang.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Tao, Y.; Yang, C.; Wang, T.; Coltey, E.; Jin, Y.; Liu, Y.; Jiang, R.; Fan, Z.; Song, X.; Shibasaki, R.; et al. A Survey on Data-driven COVID-19 and Future Pandemic Management. *ACM Comput. Surv.* **2022**, *55*, 1–36. [[CrossRef](#)]
2. Tenali, N.; Babu, G.R.M. A Systematic Literature Review and Future Perspectives for Handling Big Data Analytics in COVID-19 Diagnosis. *New Gener. Comput.* **2023**, *41*, 243–280. [[CrossRef](#)] [[PubMed](#)]
3. Duncan, D.B.; Mackett, K.; Ali, M.U.; Yamamura, D.; Balion, C. Performance of saliva compared with nasopharyngeal swab for diagnosis of COVID-19 by NAAT in cross-sectional studies: Systematic review and meta-analysis. *Clin. Biochem.* **2023**, *117*, 84–93. [[CrossRef](#)]
4. Tng, D.J.; Yin, B.C.; Cao, J.; Ko, K.K.; Goh, K.C.; Chua, D.X.; Zhang, Y.; Chua, M.L.; Low, J.G.; Ooi, E.E.; et al. Amplified parallel antigen rapid test for point-of-care salivary detection of SARS-CoV-2 with improved sensitivity. *Microchim. Acta* **2022**, *189*, 14. [[CrossRef](#)] [[PubMed](#)]
5. Dutta, D.; Naiyer, S.; Mansuri, S.; Soni, N.; Singh, V.; Bhat, K.H.; Singh, N.; Arora, G.; Mansuri, M.S. COVID-19 Diagnosis: A Comprehensive Review of the RT-qPCR Method for Detection of SARS-CoV-2. *Diagnostics* **2022**, *12*, 1503. [[CrossRef](#)]
6. Larremore, D.B.; Wilder, B.; Lester, E.; Shehata, S.; Burke, J.M.; Hay, J.A.; Tambe, M.; Mina, M.J.; Parker, R. Test sensitivity is secondary to frequency and turnaround time for COVID-19 screening. *Sci. Adv.* **2021**, *7*, eabd5393. [[CrossRef](#)]
7. He, J.; Zhu, S.; Zhou, J.; Jiang, W.; Yin, L.; Su, L.; Zhang, X.; Chen, Q.; Li, X. Rapid detection of SARS-CoV-2: The gradual boom of lateral flow immunoassay. *Front. Bioeng. Biotechnol.* **2023**, *10*, 1090281. [[CrossRef](#)]
8. Al-Hashimi, O.T.; AL-Ansari, W.I.; Abbas, S.A.; Jumaa, D.S.; Hammad, S.A.; Hammoudi, F.A.; Allawi, A.A. The sensitivity and specificity of COVID-19 rapid anti-gene test in comparison to RT-PCR test as a gold standard test. *J. Clin. Lab. Anal.* **2023**, *37*, e24844. [[CrossRef](#)]
9. Liu, K.S.; Mao, X.D.; Ni, W.; Li, T.P. Laboratory detection of SARS-CoV-2: A review of the current literature and future perspectives. *Heliyon* **2022**, *8*, e10858. [[CrossRef](#)]
10. Chong, Y.P.; Choy, K.W.; Doerig, C.; Lim, C.X. SARS-CoV-2 Testing Strategies in the Diagnosis and Management of COVID-19 Patients in Low-Income Countries: A Scoping Review. *Mol. Diagn. Ther.* **2023**, *27*, 303–320. [[CrossRef](#)]
11. El-Sherif, D.M.; Abouzid, M.; Gaballah, M.S.; Ahmed, A.A.; Adeel, M.; Sheta, S.M. New approach in SARS-CoV-2 surveillance using biosensor technology: A review. *Environ. Sci. Pollut. Res.* **2022**, *29*, 1677–1695. [[CrossRef](#)] [[PubMed](#)]
12. Abid, S.A.; Muneer, A.A.; Al-Kadmy, I.M.; Sattar, A.A.; Beshbishy, A.M.; Batiha, G.E.-S.; Hetta, H.F. Biosensors as a future diagnostic approach for COVID-19. *Life Sci.* **2021**, *273*, 119117. [[CrossRef](#)]
13. Wei, H.; Zhang, C.; Du, X.; Zhang, Z. Research progress of biosensors for detection of SARS-CoV-2 variants based on ACE2. *Talanta* **2023**, *251*, 123813. [[CrossRef](#)]
14. Rong, G.; Zheng, Y.; Li, X.; Guo, M.; Su, Y.; Bian, S.; Dang, B.; Chen, Y.; Zhang, Y.; Shen, L.; et al. A high-throughput fully automatic biosensing platform for efficient COVID-19 detection. *Biosens. Bioelectron.* **2023**, *220*, 114861. [[CrossRef](#)]
15. Rong, G.; Zheng, Y.; Yang, X.; Bao, K.; Xia, F.; Ren, H.; Bian, S.; Li, L.; Zhu, B.; Sawan, M. A Closed-Loop Approach to Fight Coronavirus: Early Detection and Subsequent Treatment. *Biosensors* **2022**, *12*, 900. [[CrossRef](#)] [[PubMed](#)]
16. Takemura, K. Surface Plasmon Resonance (SPR)- and Localized SPR (LSPR)-Based Virus Sensing Systems: Optical Vibration of Nano- and Micro-Metallic Materials for the Development of Next-Generation Virus Detection Technology. *Biosensors* **2021**, *11*, 250. [[CrossRef](#)] [[PubMed](#)]
17. Zhu, Y.; Hu, W.; Fang, Y. Direct excitation of the Tamm plasmon-polaritons on a dielectric Bragg reflector coated with a metal film. *Opto-Electron. Rev.* **2013**, *21*, 338–343. [[CrossRef](#)]
18. Liu, C.; Kong, M.; Li, B. Tamm plasmon-polariton with negative group velocity induced by a negative index meta-material capping layer at metal-Bragg reflector interface. *Opt. Express* **2014**, *22*, 11376–11383. [[CrossRef](#)]
19. Zheng, Y.; Bian, S.; Sun, J.; Wen, L.; Rong, G.; Sawan, M. Label-Free LSPR-Vertical Microcavity Biosensor for On-Site SARS-CoV-2 Detection. *Biosensors* **2022**, *12*, 151. [[CrossRef](#)]
20. Cao, Y.; Zhu, Y.; Rong, G.; Lin, Z.; Wang, G.; Gu, Z.; Sawan, M. Efficient Optical Pattern Detection for Microcavity Sensors Based Lab-on-a-Chip. *IEEE Sens. J.* **2012**, *12*, 2121–2128. [[CrossRef](#)]
21. Mariani, S.; Pino, L.; Strambini, L.M.; Tedeschi, L.; Barillaro, G. 10 000-Fold Improvement in Protein Detection Using Nanostructured Porous Silicon Interferometric Aptasensors. *ACS Sens.* **2016**, *1*, 1471–1479. [[CrossRef](#)]
22. Wu, C.; Rong, G.; Xu, J.; Pan, S.; Zhu, Y. Physical analysis of the response properties of porous silicon microcavity biosensor. *Phys. E Low-Dimens. Syst. Nanostructures* **2012**, *44*, 1787–1791. [[CrossRef](#)]
23. Kotsiantis, S.B.; Zaharakis, I.D.; Pintelas, P.E. Machine learning: A review of classification and combining techniques. *Artif. Intell. Rev.* **2006**, *26*, 159–190. [[CrossRef](#)]

24. Kliegr, T.; Bahník, Š.; Fűrnkranz, J. A review of possible effects of cognitive biases on interpretation of rule-based machine learning models. *Artif. Intell.* **2021**, *295*, 103458. [[CrossRef](#)]
25. Metri-Ojeda, J.; Solana-Lavalle, G.; Rosas-Romero, R.; Palou, E.; Rodrigues, M.R.; Baigts-Allende, D. Rapid screening of mayonnaise quality using computer vision and machine learning. *J. Food Meas. Charact.* **2023**, *17*, 2792–2804. [[CrossRef](#)]
26. Mughaid, A.; Obeidat, I.; AlZu'bi, S.; Abu Elsouid, E.; Alnajjar, A.; Alsoud, A.R.; Abualigah, L. A novel machine learning and face recognition technique for fake accounts detection system on cyber social networks. *Multimed. Tools Appl.* **2023**, *82*, 26353–26378. [[CrossRef](#)]
27. Xu, Y.; Lin, J.; Gao, H.; Li, R.; Jiang, Z.; Yin, Y.; Wu, Y. Machine Learning-Driven APPs Recommendation for Energy Optimization in Green Communication and Networking for Connected and Autonomous Vehicles. *IEEE Trans. Green Commun. Netw.* **2022**, *6*, 1543–1552. [[CrossRef](#)]
28. Balkus, S.V.; Wang, H.; Cornet, B.D.; Mahabal, C.; Ngo, H.; Fang, H. A Survey of Collaborative Machine Learning Using 5G Vehicular Communications. *IEEE Commun. Surv. Tutor.* **2022**, *24*, 1280–1303. [[CrossRef](#)]
29. Dai, Z.-H.; Wang, R.-H.; Guan, J.-H. Auxiliary Decision-Making System for Steel Plate Cold Straightening Based on Multi-Machine Learning Competition Strategies. *Appl. Sci.* **2022**, *12*, 11473. [[CrossRef](#)]
30. Hoque, M.N.; Mueller, K. Outcome-Explorer: A Causality Guided Interactive Visual Interface for Interpretable Algorithmic Decision Making. *Ieee Trans. Vis. Comput. Graph.* **2022**, *28*, 4728–4740. [[CrossRef](#)]
31. Fidencio, A.X.; Klaes, C.; Iossifidis, I. Error-Related Potentials in Reinforcement Learning-Based Brain-Machine Interfaces. *Front. Hum. Neurosci.* **2022**, *16*, 806517. [[CrossRef](#)] [[PubMed](#)]
32. Tabari, A.; Chan, S.M.; Omar, O.M.F.; Iqbal, S.I.; Gee, M.S.; Daye, D. Role of Machine Learning in Precision Oncology: Applications in Gastrointestinal Cancers. *Cancers* **2022**, *15*, 63. [[CrossRef](#)] [[PubMed](#)]
33. Brown, J.A.; Cuzzocrea, A.; Kresta, M.; Kristjanson, K.D.; Leung, C.K.; Tebinka, T.W. A Machine Learning System for Supporting Advanced Knowledge Discovery from Chess Game Data. In Proceedings of the 2017 16th Ieee International Conference on Machine Learning and Applications (Icmla), Cancun, Mexico, 18–21 December 2017; pp. 649–654.
34. Hu, J.; Sasakawa, T.; Hirasawa, K.; Zheng, H. A hierarchical learning system incorporating with supervised, unsupervised and reinforcement learning. In *Advances in Neural Networks—Isnn 2007, Pt 1, Proceedings*; Springer: Berlin/Heidelberg, Germany, 2007; Volume 4491, p. 403.
35. Kumar, B.; Vyas, O.P.; Vyas, R. A comprehensive review on the variants of support vector machines. *Mod. Phys. Lett. B* **2019**, *33*, 1950303. [[CrossRef](#)]
36. Champati, B.B.; Padhiari, B.M.; Ray, A.; Halder, T.; Jena, S.; Sahoo, A.; Kar, B.; Kamila, P.K.; Panda, P.C.; Ghosh, B.; et al. Application of a Multilayer Perceptron Artificial Neural Network for the Prediction and Optimization of the Andrographolide Content in *Andrographis paniculata*. *Molecules* **2022**, *27*, 2765. [[CrossRef](#)] [[PubMed](#)]
37. Fang, X.; Ghosh, M. High-dimensional properties for empirical priors in linear regression with unknown error variance. *Stat. Pap.* **2023**. [[CrossRef](#)]
38. Laria, J.C.; Clemmensen, L.H.; Ersbøll, B.K.; Delgado-Gómez, D. A Generalized Linear Joint Trained Framework for Semi-Supervised Learning of Sparse Features. *Mathematics* **2022**, *10*, 3001. [[CrossRef](#)]
39. Li, M.; Yuan, B. 2D-LDA: A statistical linear discriminant analysis for image matrix. *Pattern Recognit. Lett.* **2005**, *26*, 527–532. [[CrossRef](#)]
40. Graf, R.; Zeldovich, M.; Friedrich, S. Comparing linear discriminant analysis and supervised learning algorithms for binary classification—A method comparison study. *Biom. J.* **2022**. [[CrossRef](#)]
41. Valero-Mas, J.J.; Gallego, A.J.; Alonso-Jiménez, P.; Serra, X. Multilabel Prototype Generation for data reduction in K-Nearest Neighbour classification. *Pattern Recognit.* **2023**, *135*, 109190. [[CrossRef](#)]
42. Cannings, T.I.; Berrett, T.B.; Samworth, R.J. Local nearest neighbor classification with applications to semi-supervised learning. *Ann. Stat.* **2020**, *48*, 1789–1814. [[CrossRef](#)]
43. Mohanty, S.K.; Swetapadma, A.; Nayak, P.K.; Malik, O.P. Decision tree approach for fault detection in a TCSC compensated line during power swing. *Int. J. Electr. Power Energy Syst.* **2023**, *146*, 108758. [[CrossRef](#)]
44. Borchert, S.; Mathilakathu, A.; Nath, A.; Wessolly, M.; Mairinger, E.; Kreidt, D.; Steinborn, J.; Walter, R.F.H.; Christoph, D.C.; Kollmeier, J.; et al. Cancer-Associated Fibroblasts Influence Survival in Pleural Mesothelioma: Digital Gene Expression Analysis and Supervised Machine Learning Model. *Int. J. Mol. Sci.* **2023**, *24*, 12426. [[CrossRef](#)] [[PubMed](#)]
45. Kim, T.; Lee, J.-S. Maximizing AUC to learn weighted naive Bayes for imbalanced data classification. *Expert Syst. Appl.* **2023**, *217*, 119564. [[CrossRef](#)]
46. Askari, A.; D'aspremont, A.; El Ghaoui, L. Naive Feature Selection: A Nearly Tight Convex Relaxation for Sparse Naive Bayes. *Math. Oper. Res.* **2023**. [[CrossRef](#)]
47. Menialov, I.; Krivtsov, S.; Chumachenko, T. *Dimensionality Reduction of Diabetes Mellitus Patient Data Using the T-Distributed Stochastic Neighbor Embedding*; Springer: Cham, Switzerland, 2022; pp. 86–95. [[CrossRef](#)]
48. Wu, B.; Rong, G.; Zhao, J.; Zhang, S.; Zhu, Y.; He, B. A Nanoscale Porous Silicon Microcavity Biosensor for Novel Label-Free Tuberculosis Antigen–Antibody Detection. *Nano* **2012**, *7*, 1250049. [[CrossRef](#)]

49. Kaliteevski, M.; Iorsh, I.; Brand, S.; Abram, R.A.; Chamberlain, J.M.; Kavokin, A.V.; Shelykh, I.A. Tamm plasmon-polaritons: Possible electromagnetic states at the interface of a metal and a dielectric Bragg mirror. *Phys. Rev. B* **2007**, *76*, 165415. [[CrossRef](#)]
50. van Erven, T.; Harremoës, P. Rényi Divergence and Kullback-Leibler Divergence. *Ieee Trans. Inf. Theory* **2014**, *60*, 3797–3820. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.