Machine Learning Enhanced Optical Fiber Sensor For Detection Of Glucose Low Concentration In Samples Mimicking Tissue

Maria Babińska [1*], Adam Władziński [1], Tomasz Talaśka [2], Małgorzata Szczerska [1]

[1]*Department of Metrology and Optoelectronics, Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, Gdańsk, Poland*
[2]*Department of Decision Systems and Robotics, Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, Gdańsk, Poland*

Received February 14, 2025; accepted March 28, 2025; published March 31, 2025

Abstract— This study presents an optical fiber sensor for detecting low glucose concentrations in a sample mimicking urine. Our research focused on designing a sensor capable of detecting 0.5% glucose concentrations in artificial urine. Algorithms were applied to analyze and accurately classify the data and identify the principal components of the collected data. The Random Forest and XGBoost model achieved the highest accuracy, confirming that frequency domain analysis combined with machine learning can significantly enhance glucose detection accuracy. These findings demonstrate that integrating machine learning with an optical fiber sensor enables the detection of low glucose concentrations.

Interferometric optical techniques are essential for precise chemical analysis, allowing for detecting even slight changes in the composition of biological samples [1–4]. Optical fiber microspheres are an advanced solution in optical spectroscopy and biomedical sensing, known for their high sensitivity to optical changes, such as shifts in the refractive index and light absorption [5–6].

In this study, an optical fiber sensor based on the microsphere has been used. A detailed description of the production procedure of a probe is described in detail in [7].

For the study, an optical fiber microsphere with a diameter of 280 μm was fabricated using a using a fiber fusion splicer, with an image of the microsphere shown in Fig. 1b. A single-mode optical fiber (SMF-28, Thorlabs, USA) was employed and connected to a light source (SLD-1310-18-W, FiberLabs Inc., Fujimi) with a wavelength of 1310 nm (Fig. 1a). The fiber probe, with a spherical structure, was immersed in the prepared artificial urine with varying glucose concentrations. Measurements were conducted for the following glucose concentrations: 0% and 0.5%. A glucose concentration of 0.5% (500 mg/dL) in urine is considered pathologically high [8]. This study represents a preliminary study into both the fiber-optic sensing setup and the machine learning-based analysis.

For each sample, the distribution of interference fringes as a function of wavelength was recorded, and their shifts were analyzed as an indicator of refractive index variations (Fig. 1). For data analysis, a high-resolution optical spectrum analyzer (AQ6374E, YOKOGAWA, Japan) was used, enabling precise monitoring of interference intensity variations under different experimental conditions. The obtained spectra were analyzed for characteristic shifts, and amplitude changes correlated with increasing glucose concentrations. Additionally, Fig. 1a presents a schematic diagram of our constructed measurement system. To evaluate the effectiveness of the proposed solution, artificial urine was used as a controlled test environment, allowing precise regulation of optical parameters and eliminating the variability inherent in actual biological samples [9].

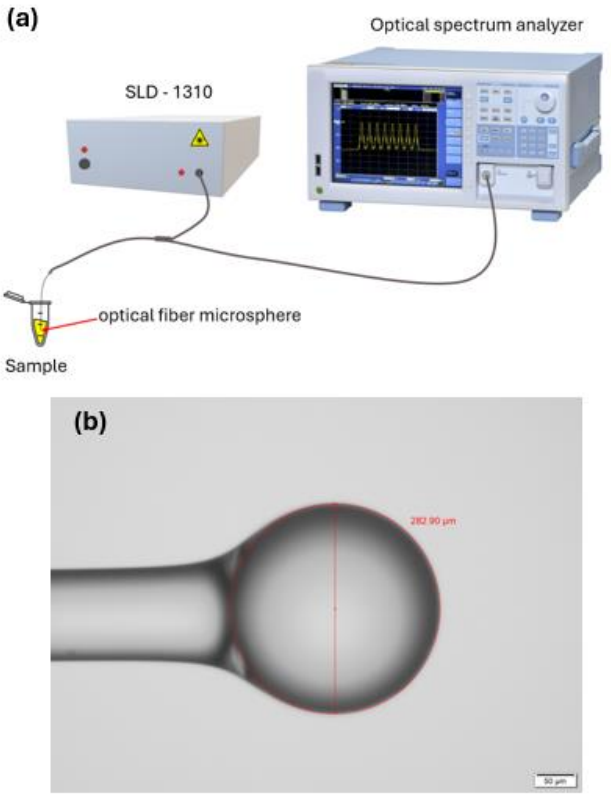

Fig. 1. Schematic of the measurement setup (a). An image of the optical fiber microsphere (b).

* E-mail: marbabi1@pg.edu.pl

This study aims to determine whether the interferometric response recorded in artificial urine accurately reflects the optical changes observed in natural samples, thereby confirming the efficacy of the developed optical fiber microsphere in precisely detecting low glucose concentrations in urine.

The experiments aimed to evaluate whether the developed optical fiber sensor effectively detects low glucose concentrations in urine. Changes in the measured light spectra were analyzed as a function of glucose concentration, focusing on spectra shifts and variations in optical signal intensity [10–11].

The complexity of the spectra and the lack of visually discernible differences make traditional comparative approaches insufficient, as illustrated in Fig. 2, where the spectral data are presented [12]. Although raw values exhibited some variation, this study aimed to validate the feasibility of developing a sensor capable of detecting clinically relevant glucose concentrations. Consequently, measured values were normalized to the range of (0,1) to facilitate interferometric analysis and ensure optimal conditions for subsequent mathematical and statistical processing.

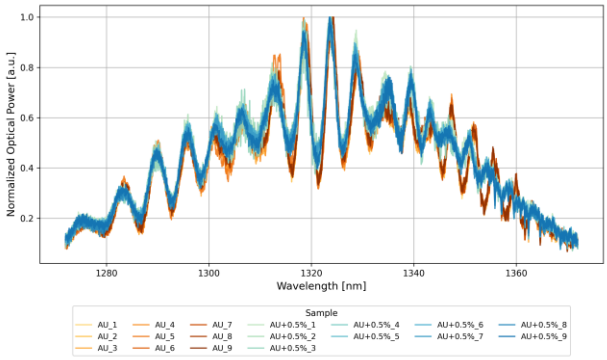

Fig. 2. Spectral data for artificial urine (AU) and artificial urine with 0.5% glucose concentration. Each line corresponds to an individual measurement, labeled using compact notation (e.g., AU_1, AU+0.5%_2) to indicate sample type and repetition number.

The primary motivation for normalization was to reduce variations in spectral intensity introduced by the analyzer. Measured values are influenced by multiple factors, including fluctuations in light source intensity, detector performance, and minor calibration inconsistencies, which should not dominate the analysis [13]. The critical information in interferometric analysis lies in the modulation and shift of the spectral curve with wavelength, not in absolute values. Normalization emphasizes subtle structural differences that may indicate compositional changes [14]. Additionally, normalization stabilizes Fast Fourier Transform (FFT) and Principal Component Analysis (PCA) performance, as these algorithms operate best when data have comparable magnitudes. Scale differences can distort the results by overemphasizing certain frequency components. Normalization also reduces the effect of noise and outliers caused by measurement errors or transient fluctuations. We used min-max normalization (0,1) instead of z-score to ensure consistent scaling across all features, improving both statistical interpretation and machine learning reliability.

We applied frequency-domain analysis to examine spectral differences between samples. The FFT-transformed absorbance spectra were subjected to PCA to reduce dimensionality and extract features relevant to glucose detection. Each FFT index corresponds to a frequency component in this representation, with lower indices reflecting broad, slowly varying features. PCA emphasizes components that explain the greatest variance across all samples, which in our case were located in the low-frequency region. This suggests that glucose-induced changes affect the global spectral shape more than localized high-frequency details. The most influential indices for PC1 were: 5, 12, 7, 8, 15, 6, 11, 9, 4, and 3; for PC2: 8, 4, 9, 6, 3, 7, 5, 12, 15, and 11. The PCA projection of FFT features is shown in Fig. 3.

These results indicate that interferometric differences due to glucose concentration are captured as low-frequency spectral modulations. Random Forest and XGBoost achieved perfect classification accuracy [16], confirming that frequency-domain features are reliable for distinguishing glucose levels in artificial urine, even with limited data.

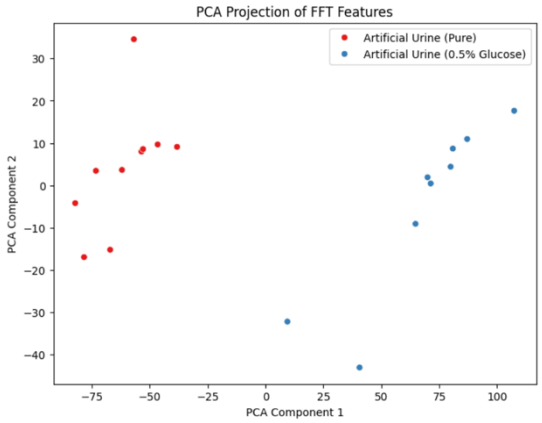

Fig. 3. PCA projection of FFT features.

As shown in the spectra, even after normalization, clear distinctions are not readily observable by the human eye. In such scenarios, mathematical data analysis techniques, such as PCA and other feature extraction methods, become indispensable for detecting subtle signal variations. This is further supported by the classification results presented in Fig. 4, where the accuracy of various models is compared. Notably, Random Forest and XGBoost achieved perfect accuracy (1.0) in distinguishing pure artificial urine from glucose-enriched samples, outperforming SVM.

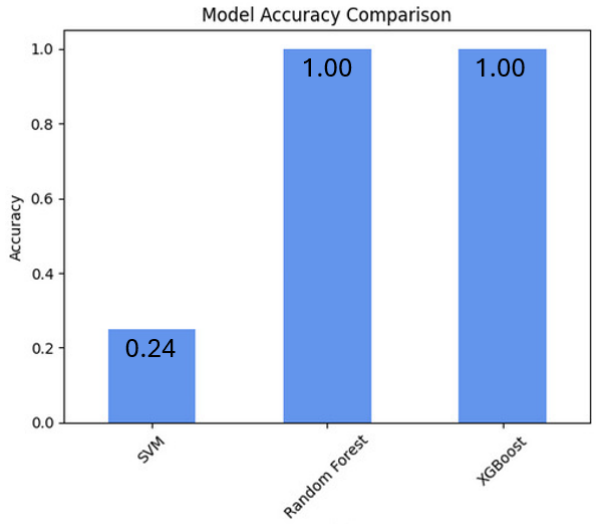

Fig. 4. Comparison of model accuracy for classification algorithms (SVM, Random Forest, and XGBoost) used to distinguish pure artificial urine from glucose-enriched samples. XGBoost and Random Forest have the highest accuracy, up to 1. This high accuracy indicates that these methods can effectively capture patterns in the data, allowing for a better understanding and interpretation of the underlying differences between sample types.

The conducted analysis confirmed that the most critical differences between the samples are embedded in the low-frequency FFT components - a characteristic signature of interferometric effects.

These results strongly suggest that integrating frequency-domain analytical techniques with dimensionality reduction methods is essential for effective interferometric analysis in optical biosensors. Specifically, PCA enabled the identification of signal components that best differentiate the samples, highlighting its significance for further research on optical detection methods for biomarkers in biological fluids.

The high sensitivity of the proposed optical fiber sensor suggests significant potential for detecting low glucose concentrations, which could be beneficial for non-invasive monitoring of clinically relevant glucose levels. Future research will focus on further refining the sensor's detection capabilities and optimizing its application for real-time biomedical diagnostics.

This research was supported by the Ministry of Education and Science under project NdS-II/SP/0438/2024/01 at Gdansk Medical University, by the DS programs of the Faculty of Electronics, Telecommunications and Informatics at Gdansk University of Technology, by grant 14/1/2024/IDUB/III.4.1/Tc and 7/1/2024/IDUB/III.4c/Tc under the TECHNETIUM Talent Management Grants, and by COST Action [CA21159].

References

[1] S. Garcia-Galan, J.A. Cabrera-Rodriguez, F.J. Maldonado-Carrascosa, N. Ruiz-Reyes, M. Szczerska, P. Vera-Candeas, F.D. Gonzalez-Martinez, F.J. Canadas-Quesada, A.J. Cruz-Lendinez, Biomedical Signal Processing and Control **102**, 107255 (2025).
[2] D.D. Nolte, Optical Interferometry for Biology and Medicine (Springer 2012).
[3] X. Wang, O.S. Wolfbeis, Anal Chem. **92**(1), 397 (2020).
[4] A. Sękowska, D. Majchrowicz, A. Sabisz, M. Ficek, B. Bułło-Piontecka, M. Kosowska, L. Jing, R. Bogdanowicz, M. Szczerska, Scient. Rep. **10**(1), 6446 (2020).
[5] S. Soria, S. Berneschi, M. Brenci, F. Cosi, G. N. Conti, S. Pelli, G.C. Righini, Sensors **11**(1), 785 (2011).
[6] H. Aouani, F. Deiss, J. Wenger, P. Ferrand, N. Sojic, H. Rigneault, Opt. Expr. **17**(21), 19085 (2009).
[7] P. Listewnik, M. Bechelany, M. Szczerska, Smart Mat. Struct. **30**(10), 10LT01 (2021).
[8] S.L. Cowart, M.E. Stachura, Jama **264**(21), 2808 (1990).
[9] N. Sarigul, F. Korkmaz, İ. Kurultak, Scient. Rep. **9**, 20159 (2019).
[10] X. Li, Y. Shao, Y. Yu, Y. Zhang, S. Wei, Sensors **16**(6), 794 (2016).
[11] E. Ujah, M. Lai, G. Slaughter, Scient. Rep. **13**, 4495 (2023).
[12] M. Grochowski, M. Wąsowicz, A. Mikołajczyk, M. Ficek, M. Kulka, M.S. Wróbel, M. Jędrzejewska-Szczerska, Metrology Measur. Syst. **26**(1), 81 (2019).
[13] I. Feder, M. Wróbel, H. Duadi, M. Jędrzejewska-Szczerska, D. Fixler, Biomedical Opt. Expr. **7**(11), 4695 (2016).
[14] P. Kluczynski, J. Gustafsson, A.M. Lindberg, O. Axner, Science Direct **56**(8), 1277 (2001).
[15] M. Hirsch, D. Majchrowicz, P. Wierzba, M. Weber, M. Bechelany, M. Jędrzejewska-Szczerska, Sensors **17**(2), 261 (2017).
[16] S. Shi, L. Ye, H. Xiao, T. Luo, 16th ICME International Conference on Complex Medical Engineering (CME), pp. 349–355 (2022).