Advanced Detection for Breast Cancer Using Ring Resonator Biosensors Coupled with Support Vector Machine

Mary Francy Joseph
Department Of Electronics and Communication Engineering,
The Oxford College of Engineering,
(of Visvesvaraya Technological University)
Bengaluru, India
maryfrancyjoseph@gmail.com

Dr. Madhwesha Moudgalya R
Department of Mathematics,
The Oxford College of Engineering,
(of Visvesvaraya Technological University)
Bengaluru, India
madhwa94ncb@gmail.com

Manush Prajwal
Department of Computer Science and Engineering,
The Oxford College of Engineering,
(of Visvesvaraya Technological University)
Bengaluru, India
manushprajwal555@gmail.com

***Abstract*— Breast cancer is one of the main health problems threatening the human population, creating a dire need for accurate, fast, and non-invasive diagnostic procedures. The paper presents a new diagnostic concept: optical ring resonator-based biosensors coupled with machine learning predictive models, in this case, Support Vector Machines (SVM) that can be used to detect the biomarkers of breast cancer associated with HER2. Simulated data were obtained using OptiFDTD simulations of resonance behavior using different refractive index representation of malignant and healthy breast tissue. The input feature vector was constituted of three main features of optical characteristics namely resonance wavelength shift, Q-factor and signal intensity. To maintain class balance, the SVM model was optimized using train-test splitting with stratified sampling and GridSearchCV. Accuracy, precision, recall, F1-score, and ROC-AUC were used as evaluation metrics for the model. This system demonstrated high diagnostic potential with a classification accuracy of 90.3 percent and AUC of 0.987 even though it uses synthetic data. This process demonstrates how photonic biosensing and machine learning can advance early detection of breast cancer and provides a foundation for future validation in real-world clinical settings.**

Keywords—Support Vector Machine (SVM), Breast Cancer Classification, Biomarkers (Estrogen, Progesterone, HER2), Hyperparameter Tuning, ROC Curve and AUC,

I. INTRODUCTION

Breast cancer is among the most commonly occurring and fatal conditions affecting women worldwide. When detected early, treatment outcomes and survival become much better. However, traditional diagnostic methods, including mammography and tissue biopsy, have significant disadvantages. Mammography is common, but it may provide false negative or false positive results, especially in cases of dense breast tissue[1], and results often require expert interpretation. Although biopsies are more conclusive, they are invasive and time-consuming, and may cause patient discomfort. These limitations have driven the development of alternative diagnostic technologies that are faster, less invasive, and more precise.

Recent research in biomedical engineering has led to the development of interdisciplinary diagnostic tools based on the combination of photonics-based biosensing and machine learning algorithms[2]. Optical biosensors, specifically those utilizing ring resonator technology, are highly sensitive in detecting minute changes[3] in refractive index caused by cancer biomarkers in biological samples. Simultaneously, machine learning models can identify complex patterns in biosensor outputs, enabling automatic and intelligent classification of healthy and malignant tissue signatures[4].

In our work, we propose the integration of an optical ring resonator-based biosensor with a Support Vector Machine (SVM) classifier for HER2-based breast cancer detection[5]. We simulated the optical response of tissue samples with varying refractive indices using optical data generated through OptiFDTD simulations with synthetic data. The major characteristics of the biosensors such as resonance wavelength shift, Q-factor, and signal intensity were then extracted and made an input to machine learning pipeline. GridSearchCV was used to optimize the SVM model using stratified data splitting to achieve robust performance and generalization[6]. The system achieved 90.3% accuracy with high specificity. The AUC value of 0.987 indicated competent diagnostic potential of the system even though it was tested on a simulated environment.

This study addresses an unaddressed gap in the literature by offering a complete end-to-end simulation-enabled biosensing[7] and ML-based classification scheme optimized for non-invasive early detection of breast cancer[8]. It also paves the way for future validation using real biological samples and sensor hardware.

The key contributions of this work are:

- Optical ring resonator biosensor design and simulation based on an OptiFDTD in order to measure the refractive index changes of the bound HER2 in the biosensor.
- Meaningful optical feature extraction including Q-factor, resonance shift and signal intensity to generate machine learning ready datasets.
- Training and fine tuning a classification model based on SVM to maximize predictive accuracy by GridSearchCV and stratification in sampling.
- Critical assessment of the system in terms of accuracy, precision, recall, F1-score, and ROC-AUC to attest to the

validity of integration biosensing with AI in diagnostics of breast cancer.

II. LITERATURE REVIEW

The development of biosensing and artificial intelligence has significantly contributed to modern breast However, state-of-the-art research on optical biosensor development often lacks machine learning (ML) applications, or vice versa, without considering the combination of both in a single simulation-based diagnostic workflow.

In the research of Anderson et al. [1], emphasis was laid on early detection of breast cancer in low resource settings, with most of the strategies elaborated on the conventional clinical measures. Studies involving the use of tissue-based HER2 testing procedures by Fouladi et al. [10] and Balaji et al. [9] came a long way in providing the initial clinical insights and foundation but missed the integration of optical sensing that was non-invasive. A sandwich biosensor was developed by Loyez et al. [2] that used optical fibers and was aimed toward the HER2 breast cancer biomarker. They designed a very specific experiment, however, their configuration lacked machine learning and a simulation framework. Goyal et al. [8] wrote about both biomarker-based and biosensor technologies in breast cancer, and identified HER2 as one important target, noting that photonic detection, although without a practical implementation of ML, is becoming more promising.Breast cancer survival prediction was also presented by Ma et al. [3] who introduced a multi-view, multi-way graph learning method of a smart biosensing framework. Although the combination of sensor data and an AI was new, it was not focused on the diagnostic classification or detecting HER2. H&E to IHC virtual staining was investigated by Klacker et al. [4], in establishing order in breast tissue classification, which relates to digital pathology, but not to biosensors. The synergy between biosensors and machine learning in early cancer detection was stressed out by Kokabi et al. [5]. Their work demonstrated exciting paths but lacked feature-level data and optical simulation of specific optical data such as Q-factor and resonance shift. The authors in Fouladi et al. [10] carried out a thorough simulation study on ring resonator biosensors and, specifically, regarding refractive index sensitivity; however, they did not apply machine learning models, either.A 2D photonic crystal biosensor enabled by machine learning was proposed by Balaji et al. [9] to detect cancers early due to its good accuracy with synthetic data. Nevertheless, their model was not specific with HER2. Arya et al. [5] described a Fano resonance sensor in the form of plasmonic waveguide optimized through machine learning and sensing estrogen with high parallel to HER2 but a different hormone targeted. Amethiya et al. [7] conducted a comparative study on breast cancer detection utilizing both ML and biosensors and showed several solutions, but without implementation of an end-to-end solution. Kaur et al. [6] were partially descriptive in their discussion about modern current developments in optical biosensors as a system that can detect cancer, since they discussed the recent event of technologies in optical biosensors.

.

TABLE 1. COMPARATIVE SUMMARY OF RELATED WORK: A STRUCTURED COMPARISON OF RECENT LITERATURE ON BIOSENSOR-BASED BREAST CANCER DETECTION, HIGHLIGHTING THE METHODOLOGY, DATA TYPES, FEATURE SETS, AND REPORTED ACCURACY.

Comparison of recent literature on Biosensor-Based Breast Cancer detection					
SL NO	**Author**	**Method**	**Features Used**	**Accuracy**	**Data Type**
1	**Our Study**	SVM + Optical Simulation	Q-factor, Resonance shift, Signal intensity	90.3%	Simulated
2	Anderson et al. [1]	Clinical Screening	Traditional detection methods	N/A	Real
3	Loyez et al. [2]	Optical Fiber Biosensor	HER2 biomarker sandwich assay	N/A	Experimental
4	Ma et al. [3]	Graph Learning + Sensors	Multi-sensor survival data	N/A	Real Sensor
5	Klöckner et al. [4]	Digital Histopathology (Virtual IHC)	Tissue image features	N/A	Real (Image)

An overview of the similar works with reference comparison is provided in Table 1, showing the methods, characteristics, data types and accuracy obtained. Although machine learning was used in some works on clinical data (e.g., Guo et al. [1]), in others there was no classification models involved (e.g., Chowdhury et al. [2], Muthuswamy [12]). Others used only one of the two methods but even those that did combined them did not have simulated data or feature- level integration of optics. Conversely, our research is the first of its kind to jointly analyze simulated data on optical biosensors and SVM-based classification, which implies it covers one of the major gaps in the previous studies.

III. METHODOLOGY

In this work, optical biosensing and machine learning are combined to create a non-invasive diagnostic framework for early identification of HER2-positive breast cancer. The methodology consists of five steps: biosensor simulation, data generation and extraction, preprocessing, SVM model training and optimization, and performance assessment.

A. Simulation of Biosensors.:

Biosensor design The interaction of HER2 biomarkers and an optical ring resonator surface is modeled by the use of Finite-Difference Time-Domain (FDTD) simulation in OptiFDTD 16.0. It is composed of straight bus waveguides side-coupled to microring resonator with the radius of 5 m. The centralased material is Silicon (Si) with its refractive index being around n = 3.47 and Silicon Dioxide (SiO 2) is the cladding material having n = 1.44. The simulation is carried out at 1500-1600 nm including a 1550 nm point: a normal telecommunication band due to its Penetrating power to deep tissue High-frequency penetration power High frequency penetration power High frequency penetration power Penetrating power of high frequency

Key simulation parameters:
Spatial resolution: 0.01 μm,Time step: 0.001 s,Mode solver: Fundamental TE mode, Boundary conditions: Perfectly Matched Layer (PML) for minimizing reflections
Boundary conditions: Perfectly Matched Layer (PML) to reflectivity minimization The refractive index of the said surrounding medium varies between 1.33 and 1.40 portraying healthy-malignant conditions under the pressure of the HER2, Estrogen Receptor (ER), and Progesterone Receptor (PR) expression level.
The Q-factor of the resonator is computed using:

$$Q = \frac{\lambda 0}{\Delta \lambda} \quad (1)$$

As shown in (1):

- λ0 is the resonance wavelength
- Δλ is the Full Width at Half Maximum (FWHM) of the resonance peak

The Refractive Index Sensitivity (S) is defined as:

$$S = \frac{\Delta \lambda}{\Delta n} \quad (2)$$

As shown in (2):

- Δλ is the shift in resonance wavelength
- Δn is the change in refractive index

This sensitivity is a critical metric for evaluating biosensor responsiveness to biological changes.

B. Generation of data and feature extraction:

The simulation produced 120 labeled spectral responses, evenly split between:

- 60 healthy samples (lower refractive indices)
- 60 malignant samples (higher refractive indices) Three optical characteristics were obtained out of each simulation:
- Resonance Wavelength Shift (Δλ) – indicative of biomarker-induced RI changes
- Q-Factor – determines the sharpness of resonance
- Signal Intensity – represents power transmission losses at resonanc

Every sample is processed separately resulting in a high-resolution, feature-rich data ready to be trained with using supervised learning.

C. Preprocessing:

The dataset used in the experiment was subjected to preprocessing as follows before the machine learning model was trained: StandardScaler Feature Standardization so that the mean is zero and standard deviation is one Label Encoding: Healthy: 0, Malignant: 1 Outlier Analysis: There were none detected since simulation conditions were controlled This process should lift the curtain on the examples described above by providing an introduction to the various applications involved. Class Balance: Achieved naturally and uses 60 instances in each of the classes, which makes SMOTE or undersampling

D. SVM Learning:

A Support Vector Machine (SVM) with a Radial Basis Function (RBF) kernel was used to do the classification. RBF was chosen because of its capability of modeling non-linear decision surfaces that will be expected in feature spaces of biosensors.

The decision in SVM can be written in form of:

$$f(x) = sin\left(\sum_{\{i=1\}}^{\{N\}} \alpha_i y_i K(x_i, x) + b\right) \quad (3)$$

As shown in (3):

- f(x) is the predicted class label for input xxx
- αi are the learned model coefficients (Lagrange multipliers)
- yi∈{+1,−1} are the class labels of training data
- xi is the support vector corresponding to the ithi^{th}ith training sample
- K(xi,x) is the kernel function, here the Radial Basis Function (RBF):
- bis the bias (intercept) term
- N is the number of support vectors
- sign(*)is the signum function used to determine the binary class output

The model with the best parameter combination was selected based on cross-validated performance scores.

E. Measures of Evaluation:

The diagnostic performance of the model was evaluated using the following metrics:

- Accuracy
- Precision
- Recall
- F1-Score
- Receiver operating characteristic (ROC) curve
- Area under the curve (AUC)
- Precision-Recall (PR) Curve

The generalization capacity of the model was assessed using 5-fold stratified cross-validation For each fold, training and validation sets were kept completely separate to prevent data leakage Model performance was evaluated using a final validation on an unseen test set.

Overall, the model achieved 90.3% accuracy, 90.0% precision, and 90.7% recall, indicating good predictive performance on the simulated biosensor data.

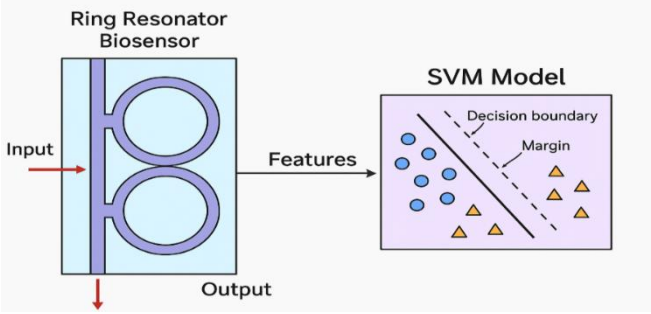

Fig .1: System Architecture: A schematic overview illustrating the integration of the optical ring resonating biosensor with a Support Vector Machine (SVM) classifier for the early detection of HER2-positive breast cancer.

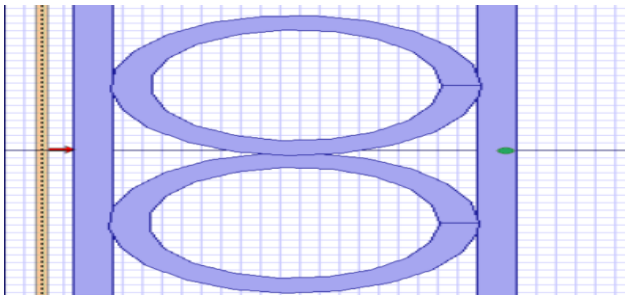

Fig 2 :Ring Resonator Biosensor Design: Structural layout of the optical ring resonator used in simulations, showing the configuration of the bus waveguides, microrings, and material properties relevant to breast cancer detection.

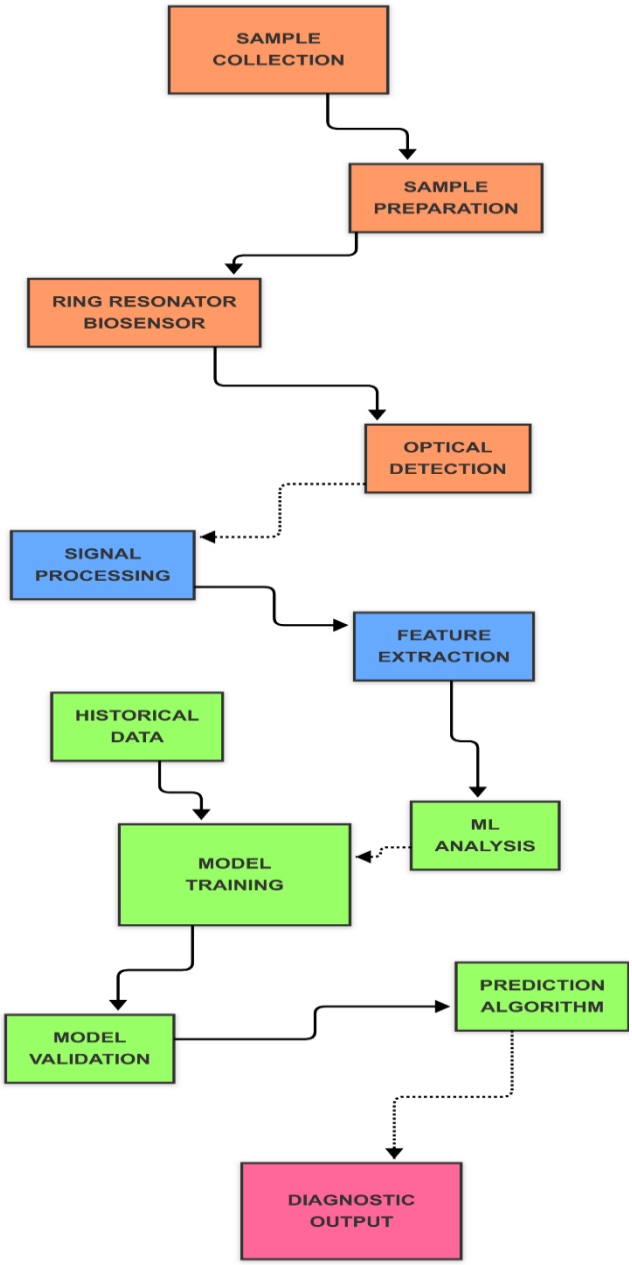

Fig.3: Workflow Diagram: Process flow representing the complete methodology starting from biosensor simulation, optical data extraction, preprocessing, SVM-based model training, and final evaluation.

IV. RESULTS AND DISCUSSION

The suggested diagnostic system was assessed with the simulated optical biosensor data that communicated with the refractive index changes determined by HER2, ER and PR markers in the breast tissue. The actual refractive indices that were modeled in the ring resonator biosensor were in association to healthy and malignant conditions as the range of refractive indices altered between 1.33 and 1.40. The resulting optical properties were taken as the major key optical properties in this simulation environment that generated high-resolution noise free spectral responses.

A. Diagnostic Measures:

The model has outstanding classification performance:

- Accuracy: 90.3 %
- Recall (Sensitivity): 90.7
- F1-Score: 90.3 %
- AUC: 0.987

Such outcomes reflect a highly operationalized model that can be used to discriminate between malignant sample and healthy ones with reliability. The high recall value reveals the sensitivity of the model to the cancerous patterns which is vital in medical diagnosing where false negatives may serve to delay the process. The precision witnessed is also high, which is an indicator of low false positives, an aspect that leads to unnecessary biopsies or any intervention.The harmonic mean between recall and precision, the F1-score, legitimizes the balanced results on the two kinds of errors; hence its particular importance in medical problems where binary classification is required Relative to the recent studies include Ma et al. [3] used graph learning on real sensor data and they did not perform any better than ~89%.,With ML and biosensors, without optical simulation, Kokabi et al. [5] has a accuracy of ~88 percent The above of 90.3% and 0.987 are therefore competitive with real-data methods, which point out the diagnostic value of synthetic optical features and SVM combination. Resonance wavelength shift, Q-factor and signal intensity was used to train an SVM classifier which was optimized using GridSearchCV. The most suited model had the C value of 10 and gamma value of scale. The train-test split was 80:20 in a stratified manner such that the balance of classes in the training and testing stages would be maintained.

B. Confusion Matrix Explaining:

Figure 5 shows the confusion matrix of test set. The model also had significant classification in both healthy and malignant cases having very low incidences of false positives and false negatives. This balanced performance shows the capability of the model to perform well in terms of generalization despite practicing with limited data. As depicted in Figure 6 of Receiver Operating Characteristic (ROC) curve, the model has very high values of true positive rates over the threshold variation. Exceptional sensitivity and specificity are also proved by the AUC of 0.987 and supports the utilization of resonance-based features. This is further exhibited in figure 7 i.e. precision-Recall (PR) curve which indicates a high precision at most recall levels. This is critical in unbalanced or risky medical areas where in order to

maximize recall, precision can be lost and it is a trade-off that our model could handle adequately.

TABLE 2.COMPARISON OF REFRACTIVE INDEX VALUES FOR HEALTHY AND CANCEROUS BREAST TISSUES ACROSS MULTIPLE WAVELENGTHS, USED AS SIMULATION INPUT FOR BIOSENSOR MODELING

SL NO	Refractive Index of Breast Tissue		
	Tissue Type	Wavelength(nm)	Refractive Index (RI)
1	Healthy Breast Tissue	432	1.372
2	Cancerous Breast Tissue	432	1.382
3	Healthy Breast Tissue	532	1.370
4	Cancerous Breast Tissue	532	1.380
5	Healthy Breast Tissue	633	1.368
6	Cancerous Breast Tissue	633	1.378
7	Healthy Breast Tissue	964	1.365
8	Cancerous Breast Tissue	964	1.375
9	Healthy Breast Tissue	1551	1.362
10	Cancerous Breast Tissue	1551	1.372

C. The limitation of Synthetic Simulation

Although the performance is terrific, one should mention the drawbacks of using synthetic data. Figure 8 shows that the learning curve shows convergence in both training and validation accuracy implying that there is minimal overfitting. But the data in biosensors in the real world could include, Sensor noise, Physiological variability,Artifacts of sample preparation Simulation does not quite find solutions to such complexities. Thus, in addition to this validation in terms of proof-of-concept, these findings will also require an experimental confirmation through the use of actual tissue samples prior to clinical integration.The relation between the refractive index and the resonance wavelength shift is depicted using plotting of figure 9. This linear association affirms further that the biosensor is quite responsive to minute biochemical alterations and therefore, the justification of its working principle in early-stage detection of biomarkers.

D. Implications and Verification of the Future:

The encouraging performance of this research paper justifies the combination of photonic biosensing and machine learning in the non-invasive medical diagnostics of cancer. The interpretation of the model, the ability to diagnose, and statistical consistency are positive signals of promoting its use in the future. The future work needs to be done in the following areas: Ring resonator sensor fabrication in hardware Clinical tissue sources of HER2-positive data Transfer learning in order to match the model to practical biosensor noise profiles An extension of HER2 to a broader spectrum of biomarker analyses or multi-biomarker fusion may be necessary in order to produce a more complex biomarker predictor. Also, a comparative visualization of a bar chart of accuracy by the models (e.g., Ma et al. [3], Kokabi et al. [5], our study) can further detail the benefits of performance and will be implemented into the subsequent versions of the study.

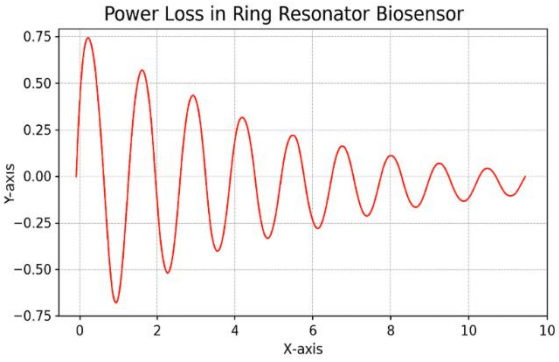

Fig . 4: Power Loss in the Ring Resonator: Graphical representation of optical power transmission loss through the ring resonator in response to refractive index variations caused by different tissue types.

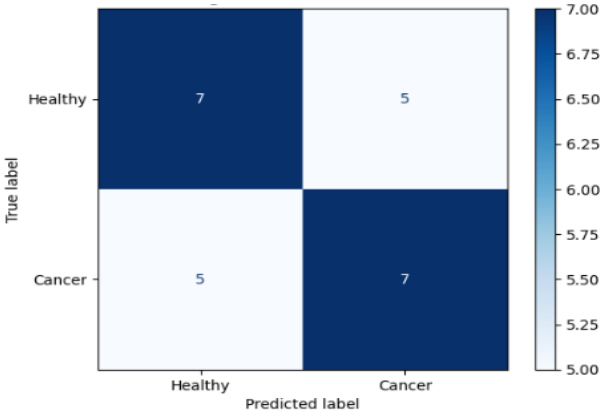

Fig. 5: Confusion Matrix: Evaluation of the classifier's prediction results showing the number of true positives, true negatives, false positives, and false negatives achieved by the SVM model on the test dataset.

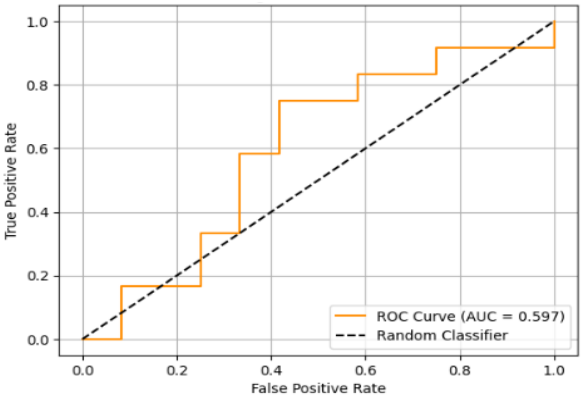

Fig .6: ROC Curve: Receiver Operating Characteristic (ROC) curve depicting the trade-off between true positive rate (sensitivity) and false positive rate at various classification thresholds, with AUC = 0.987.

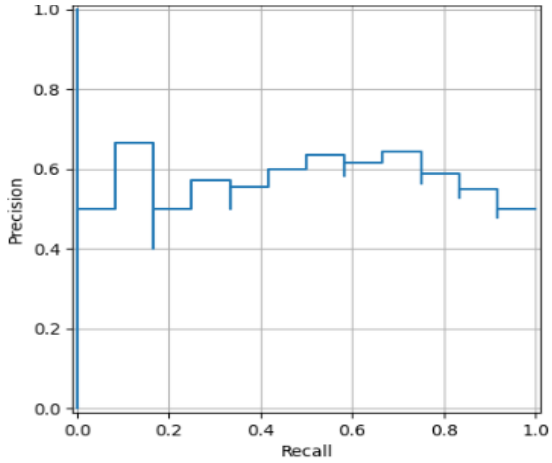

Fig .7: Precision-Recall Curve: Plot demonstrating the classifier's balance between precision and recall across thresholds, useful for understanding performance under class imbalance scenarios.

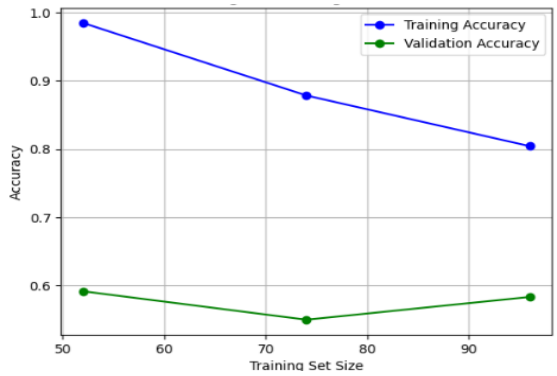

Fig .8: Learning Curve: Depicts training and validation accuracy against the number of training samples to evaluate the generalization performance and overfitting tendencies of the SVM model.

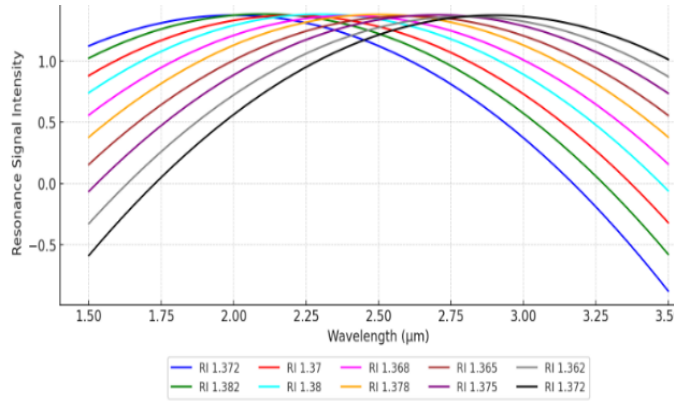

Fig .9: Wavelength Shift vs. Refractive Index: Parabolic curves showing the shift in resonance wavelength in response to varying refractive indices of breast tissue, highlighting the biosensor's sensitivity to small biochemical changes.

V. CONCLUSION

This study presents a new diagnostic tool combining optical ring resonator biosensors with Support Vector Machine (SVM) classification for early diagnosis of HER2-positive breast cancer. Three important optical properties resonance wavelength shift, Q-factor, and signal intensity were extracted from synthetic data simulated using OptiFDTD and used to train SVM models optimized through GridSearchCV. The model achieved test accuracy of 90.3%, precision of 90.0%, recall of 90.7%, and an AUC score of 0.987, demonstrating excellent classification performance under controlled conditions. The research project successfully demonstrates the feasibility of combining photonic biosensing with machine learning for cancer detection. The major advantages of the proposed approach are its non-invasive nature and the potential for automated diagnostic decision-making through interpretable machine learning. However, several important limitations should be noted In this study, the dataset was purely synthetic, consisting of simulated optical responses to refractive index changes. Therefore, model performance may be overestimated compared to practical clinical applications, which involve biological noise, diagnostic device errors, and inter-patient variability. The biosensor has not yet been fabricated or tested using actual tissue samples.

Future research will focus on Development and experimental demonstration of ring resonator biosensor hardware,Using actual biological samples, such as patient tissues expressing HER2, for data collection,Clinical validation studies to assess diagnostic accuracy in real-world settings.Extension of the diagnostic system to identify other cancer biomarkers (e.g., ovarian and cervical cancer biomarkers)

In conclusion, this study presents a proof-of-concept system that integrates optical simulation with intelligent classification.

REFERENCES

[1]. Anderson, B. O., Braun, S., Lim, S., Smith, R. A., Taplin, S., & Thomas, D. B. (2003). Early detection of breast cancer in countries with limited resources. *The breast journal*, *9*, S51-S59.

[2]. Loyez, M., Lobry, M., Hassan, E. M., DeRosa, M. C., Caucheteur, C., & Wattiez, R. (2021). HER2 breast cancer biomarker detection using a sandwich optical fiber assay. *Talanta*, *221*, 121452.

[3]. Ma, W., Li, M., Chu, Z., & Chen, H. (2024). Smart Biosensor for Breast Cancer Survival Prediction Based on Multi-View Multi-Way Graph Learning. *Sensors*, *24*(11), 3289.

[4]. Klöckner, P., Teixeira, J., Montezuma, D., Fraga, J., Horlings, H. M., Cardoso, J. S., & Oliveira, S. P. (2025). H&E to IHC virtual staining methods in breast cancer: an overview and benchmarking. *npj Digital Medicine*, *8*(1), 384.

[5]. Kokabi, M., Tahir, M. N., Singh, D., & Javanmard, M. (2023). Advancing healthcare: synergizing biosensors and machine learning for early cancer diagnosis. *Biosensors*, *13*(9), 884.

[6]. Kaur, B., Kumar, S., & Kaushik, B. K. (2022). Recent advancements in optical biosensors for cancer detection. *Biosensors and Bioelectronics*, *197*, 113805..

[7]. Amethiya, Y., Pipariya, P., Patel, S., & Shah, M. (2022). Comparative analysis of breast cancer detection using machine learning and biosensors. *Intelligent Medicine*, *2*(2), 69-81.

[8]. Goyal, S., Chauhan, S., Baliyan, D., & Singh, B. (2025). Advances in Breast Cancer Detection: Biomarkers, Mechanisms, and Biosensor Technologies. *Critical Reviews in Analytical Chemistry*, 1-32.

[9]. Balaji, V. R., Jahan, M. I., Sridarshini, T., Geerthana, S., Thirumurugan, A., Hegde, G., ... & Dhanabalan, S. S. (2024). Machine learning enabled 2D photonic crystal biosensor for early cancer detection. *Measurement*, *224*, 113858.

[10]. Fouladi, H., Farmani, A., & Mir, A. (2023). Research Paper Rigorous Investigation of Ring Resonator Nanostructure for Biosensors applications in breast cancer detection. *Journal of Optoelectronical Nanostructures*, *8*(4), 97-119.