International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(14s) 2026 pp. 1252-1258
https://doi.org/10.70917/ijcisim-2026-4338
Article

Early Detection of Feminine Disease Using IR Spectroscopy and Graphene Electrodes

Umesh Mhapankar[1], Sonali Sherigar[2]

[1]Agnel Polytechnic, Vashi, Navi Mumbai, Maharashtra, India.
Email: umeshmhapankar@gmail.com
[2] Agnel Polytechnic, Vashi, Navi Mumbai, Maharashtra, India.
Email: sonali.sherigar@gmail.com

Abstract: Artificial intelligence, a technology that empowers computers and machines with human-level intelligence and problem-solving potential, is a game-changer in health care. Its significant ability to improve treatment and epidemiology, diagnosing conditions more rapidly and precisely, is a beacon of hope. Uterine cancer, a leading cause of death among women, underscores the importance of early detection for effective treatment. This paper describes complete non-cutaneous, non-invasive methods that integrate AI with Infrared (IR) spectroscopy and the cutaneous graphene-based electrode. This system is used to test for various female diseases, such as PCOD, fibroids, uterine prolapse, and cancer, using biomarkers, including CA125, a biomarker standardized for uterine cancer, focusing on physical check-ups, medical histology, and biomarker identification. By applying advanced machine learning algorithms such as SVM to analyse IR spectral data, AI significantly enhances the accuracy and speed of classifying biological samples, such as distinguishing between non-cancerous and cancerous tissues, providing a ray of hope for the future of healthcare.
Keywords: IR spectroscopy, biomarker, cutaneous and non-cutaneous, AI and ML

1. Introduction

There were many women aged 15 years and older who suffered from various kinds of gynaecological diseases that might lead to cancer. About 6 of every ten persons were diagnosed with pelvic disorders in the tested patients. Certain female genetic systems are prone to cancer, such as the breast, uterus, ovaries, and endometrium, at the ages of 40 to 60. Diseases such as Polycystic Ovarian Disorders (PCOD), fibroids, uterine prolapse, and cancer are linked with similar symptoms [1]. Out of this, the cancer is fatal if it is not diagnosed early, so there must be a system that detects the cancer in its early stages. This paper utilized and discussed machine Learning (ML) algorithms to detect cancer or pelvic diseases. Before using artificial intelligence (AI) methods, certain parameters from the affected location must be extracted. Researchers use various techniques to identify and detect different cancers. Naturally, scientists use invasive and non-invasive methods. However, these methods have limitations & drawbacks due to multiple aspects. The invasive method requires

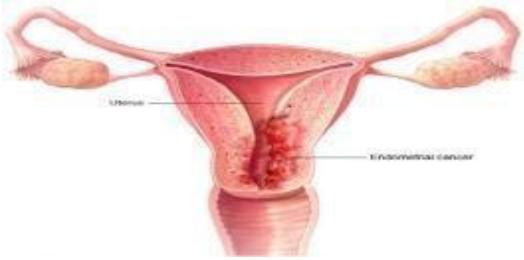

Fig 1. Endometrial/Uterus cancer[11]

Certain semi-invasive methods, like Endoscopy and Colonoscopy, require hospitalization for certain hours. The most prominent method was a biopsy, which involved chemical testing under a microscope after removing the diseased area physically with small surgery or via endoscopy or colonoscopy. To avoid this, this paper introduces the method by which IR sensor-based spectroscopic systems were introduced: non-invasive, non-contact. Various types of literature were available for spectroscopic measurement of cancer. Still, all these methods were semi-invasive, whereas the technique introduced in this paper used an IR pair sensor to detect various types of polyps in the ovarian and uterine areas for further diagnosis or classification, determining whether they were cancer, PCOD, fibroids, or benign. [2]

So, the AI method can handle such problems after understanding the literature. There are various machine learning algorithms were used by scientists such as Linear Discriminate Analyst(LDA), Support Vector Machine (SVM), Random Forest (RF), K-Nearest Neighbour(KNN), Naive Bays Classifier (NB), Relevance Vector Machine (RVM), Deep Neural Network (DNN), Deep Belief Network (DBF), Convolution Neural Network (CNN), Extreme Learning Machine (ELM), Recurrent Neural Network (RNN) Auto encoder(AEC) out of these as per various literature SVM, LDA, RF and DNN were more accurate. So, in

this paper, RF, SVM, and DNN were also used.

However, all the methods introduced in this paper are used.

Understanding the anatomy and chemical activity of the pelvic area is essential. Infrared (IR) spectroscopy of blood plasma or serum provides a rapid, reproducible, and relatively non-invasive method for detecting biomolecular changes associated with cancer, making it a promising screening and diagnostic tool. [3] In this study, the spectral range of 600nm to 1200nm was selected to capture key vibrational frequencies linked to biochemical alterations in cancerous tissues. Advanced software, such as artificial intelligence and machine learning algorithms, enhances data interpretation, making IR spectroscopy more practical and accessible across a range of therapeutic applications. [4]

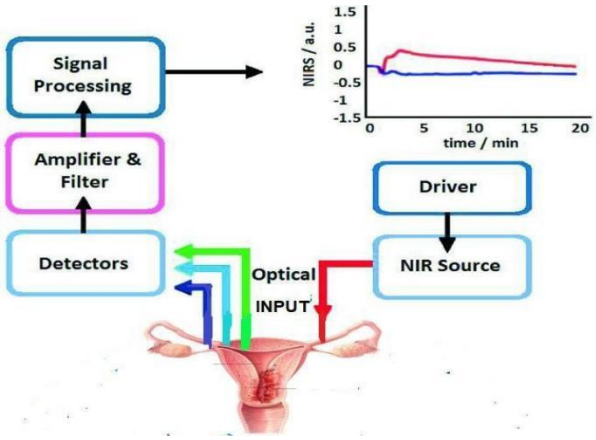

Figure 2: Block Diagram of the developed system

2. METHODOLOGY

A. Signal Acquisition

There were methods for acquiring data from subjects, either contact or non-contact, using electrodes or optical sensors. In this paper, an optical-based sensor method is used. Depending on the uterine conditions, the light absorption coefficient varied. Ten subjects aged 20 to 55 were used for this experiment. Two suffer from PCOS (polycystic ovarian syndrome), and three have fibroid glands.[9]

B. Signal Conditioning

Pelvic signals are very low in frequency and amplitude, and they are complex; the proposed system requires precise Front-End amplification with accurate extraction. Several components are necessary for pelvic amplification. Amplifiers are needed for the relatively weak gastric signal acquired by the cutaneous electrodes (approx. 200500 μV). (Figure 4)

To amplify the differential signal from the uterus region, the system uses an instrumentation amplifier

(INA2128) with variable gain and a high Common Mode Rejection Ratio (CMRR) to amplify Pelvic signals. After amplifying the bio-potential signal, the desired frequency range is filtered to simplify the signal further. This step helps eliminate baseline drift and filter out artifacts caused by respiration and movement. The figure below shows the proposed system's interpretation.

C. Analysis and Classification

After processing the desired signal, further analysis is performed using the Laboratory Virtual Instrument Engineering Workbench (LabVIEW Student Version). Lab-VIEW is a development environment and system-design platform for a visual programming language from National Instruments. LabVIEW.

Using stereo input, the front panel window of LabVIEW, which functions as a graphical user interface, is used for further study. LabVIEW provides a feature for further filtering the signals obtained from various specimens. The final waveform is considered indicative of disorders in the pelvic region.

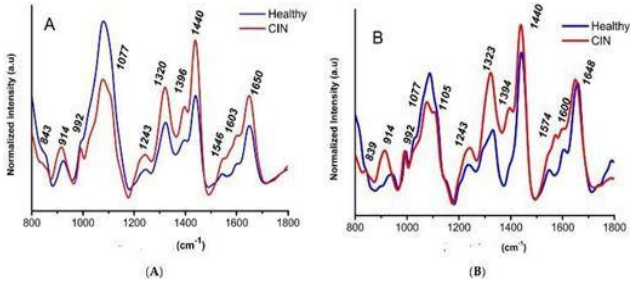

Figure 3: Raman shift within various spectra in the endodermis area

Support Vector Machines (SVMs) began gaining prominence in cancer classification with the advent of high-throughput gene expression microarray technologies in the early 2000s. One of the early applications was reported by Wong et al. [9], who used a linear SVM to differentiate between two categories of uterine cancer based on gene expression profiles. Their study utilized data from 38 patients as the training cohort. In addition to SVM, a basic supervised learning approach, “weighted voting,” was employed to discriminate between two clinically recognized conditions—polycystic ovarian disorder (PCOD) and uterine fibroids. The gene expression data were generated using an Affymetrix HG-U6800 microarray.

A total of 7,129 gene expression probes were analysed, with each gene assigned a weight reflecting its contribution to distinguishing between the two diagnostic categories. The trained SVM model was subsequently validated using an independent dataset comprising 34 patient samples. The findings highlighted the effectiveness of SVMs in handling datasets with very high dimensionality (thousands of gene features)

and relatively small sample sizes. Later work [10] further enhanced the weighted voting approach integrated with SVM, reducing the misclassification rate from 6% (2 incorrect predictions out of 34 cases) to 0%.

chips.

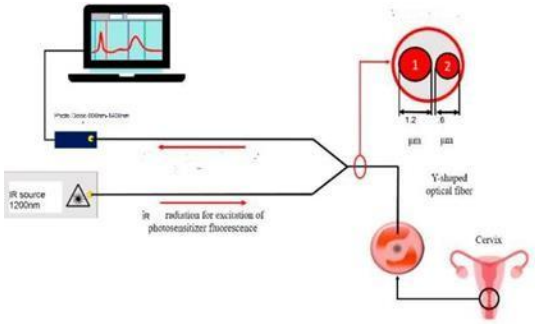

Figure 4: Schematic of the signal acquisition method

It is important to note that this initial investigation did not incorporate any prior feature selection before model construction. In a subsequent study, an SVM was implemented for colon cancer tissue classification after performing gene selection. Their dataset included 40 uterine tumor samples and 22 normal ovarian tissue specimens. They

introduced a feature-ranking method, the naive Bayes relevance (NBR) score, which estimates the likelihood of a class given a feature value under a Gaussian assumption. Using this ranking strategy, SVM models were trained on the top 50–200 samples. These reduced gene subsets achieved comparable or improved generalization performance relative to the full set of 1,988 samples when distinguishing tumor from non-tumor samples. Furthermore, SVM performance with varying numbers of selected genes was compared to that of a naive Bayes classifier trained on the same subsets [12], with SVM consistently demonstrating superior results.

Another investigation [13] employed linear SVMs combined with feature selection across three different cancer datasets. One dataset comprised 31 tissue samples, including malignant ovarian tissues and normal specimens (both ovarian and non-ovarian) [14], [15]. Gene selection was performed using the signal-to-noise ratio metric [16], which is closely related to the Fisher criterion used in Fisher's linear discriminant analysis. Across multiple datasets, SVM achieved strong classification performance, although certain perceptron-based algorithms showed comparable outcomes in some cases.

Segal et al. [17] introduced a genome-based SVM framework for classifying clear cell sarcoma, a soft-tissue malignancy with characteristics similar to those of melanoma. Initially, 256 genes were identified using Student's t-test. A linear SVM was then trained to differentiate melanoma from soft-tissue sarcoma using

these selected genes. In a leave-one-out cross-validation setting, the classifier correctly predicted 75 out of 76 samples. In another related study, the SVM methodology was also applied to address the complex histopathological patterns observed in adult soft-tissue sarcomas.

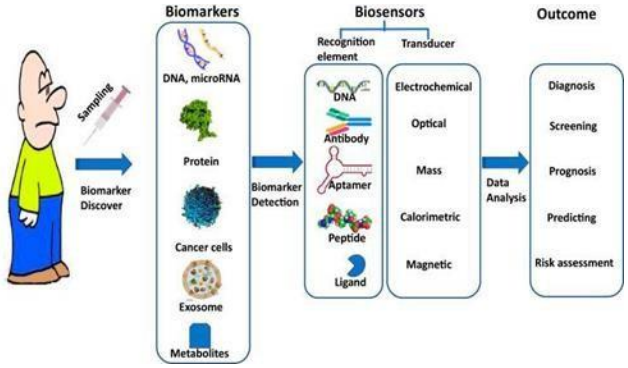

Figure 5: Complete Block diagram of the System [21]

3. Empirical results

The proposed system helps obtain results for various specimen observations. For a normal specimen, the signal cycle repeats periodically. The cyclic nature of the signal, repeating over time, reflects the blood flow pattern in that region. The blood flow pattern resulting in the cyclic signal may vary from specimen to specimen, but the magnitude will remain approximately constant in normal healthy specimens.

Figure 6: different subject amplitude variations in various locations of the uterus

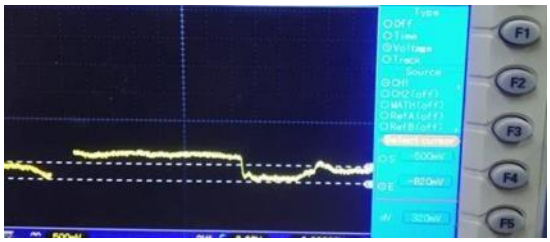

Figure 7: Pelvic region signal during normal days for specimen 1.

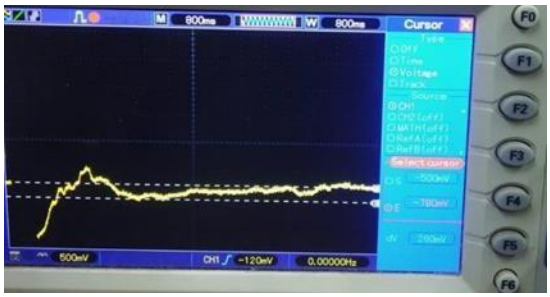

Figure 8: Pelvic region signal during normal days for specimen 3

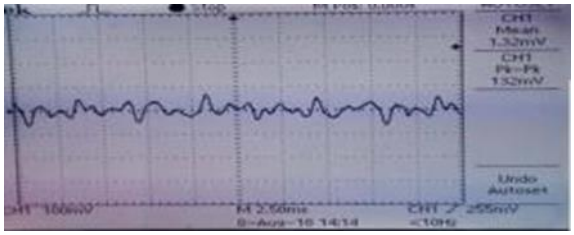

Figure 9: Pelvic region signal during 1st day of the menstrual cycle for specimen 5.

Whereas, in the case of the specimen with a disorder such as PCOS, the rhythm of the waveform observed would be different [19]. Similarly, during a female's menstrual cycle, blood flow in the pelvic region would be more intense than on normal days. As a result, the signal obtained during this period would be of high magnitude. The classifier that is proposed and compared is SVM with other methods, whose accuracy is slightly above, as shown in Table 3

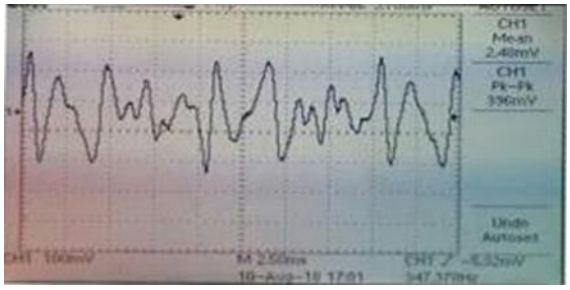

Figure 10: Pelvic region signal during the 3rd day of the menstrual cycle for specimen 5.

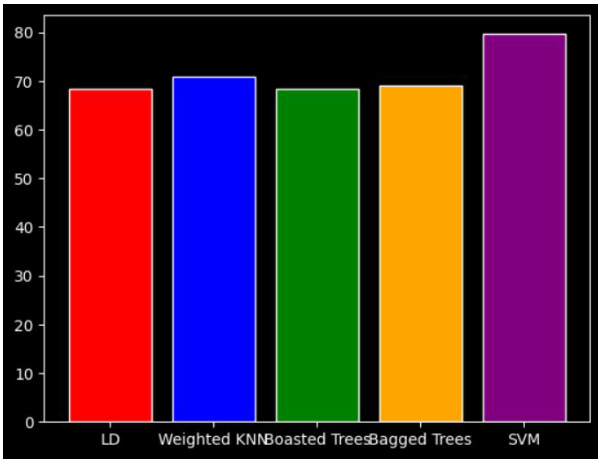

Figure 11: SVM achieves the highest Accuracy in comparison with other classification algorithms.

Table 1: Observation of 3 specimens

SPECIMEN 1		SPECIMEN 2		SPECIMEN 3	
DAY	Voltage	DAY	Voltage	DAY	Voltage
2nd	4V	2nd	3V	22nd	150mV
3rd	6V	3rd	2V	23rd	40mV
4th	960mV	4th	680mV	24th	80mV
5th	170mV	5th	220mV	25th	150mV
8th	170mV	8th	260mV	28th	210mV
9th	170mV	9th	170mV	29th	1V
10th	400mV	10th	500mV	1st	300mV
12th	150mV	12th	291mV	3rd	300mV
13th	200mv	13th	270mV	4th	2V

4. Conclusion

This paper proposes a simple, non-invasive, non-contact method to identify uterine or cervical cancer early. At the same time, machine learning algorithms are used. SVM is the simplest and most used algorithm. The comparison in Figure 11 concluded that accuracy is much better than that of other algorithms and Methods.

References

1. Alvarez WC. “The electrogastrogram and what it shows”. JAMA 1922; 78:1116–1118.
2. Gajjar K, Trevisan J, Owens G, Keating PJ, Wood NJ, Stringfellow HF, Martin-Hirsch PL, Martin FL. Fourier-transform infrared spectroscopy coupled with a classification machine for the analysis of blood plasma or serum: a novel diagnostic approach for ovarian cancer. Analyst. 2013 Jul 21;138(14):3917-
3. 26. doi:10.1039/c3an36654e. Epub 2013 Jan 17. PMID:
4. 23325355.Online: Better Health Channel.
5. Backhaus, J.; Mueller, R.; Formanski, N.; Szlama, N.; Meerpohl, H.-G.; Eidt, M.; Bugert, P. Diagnosis of Breast Cancer with Infrared
6. Spectroscopy from Serum Samples. Vib. Spectrosc. 2010, 52, 173–177. [CrossRef] V. Balan, C.-T.
7. E. Kontsek et al., "Mid-Infrared Imaging Can Characterize and Separate Cancer Cell Lines,"
8. *Pathol. Oncol. Res.*, vol. 26, pp. 2401–2407, 2020.
9. Huang S, Cai N, Pacheco PP, Narrandes S, Wang Y, Xu W. Applications of Support Vector Machine (SVM) Learning in Cancer Genomics. Cancer Genomics Proteomics. 2018 Jan-Feb;15(1):41-51. doi:
10. 10.21873/cgp.20063. PMID: 29275361; PMCID: PMC5822181.
11. M. A. Cohen Ford et al., "Infrared Spectroscopy of Normal and Abnormal Cervical Smears: Evaluation by Principal Component Analysis,"
12. *Gynecol. Oncol.*, vol. 66, pp. 59–65, 1997.
13. Seelenmeyer C, Wegehingel S, Lechner J, Nickel W. The cancer antigen CA 125 represents a novel counter-receptor for galectin-1. J Cell Sci. 2003;116(Pt 7):1305–1318. [PubMed] [Google
14. Scholar]
15. 8 D. R. Whelan et al., "Monitoring the Reversible B to A-like Transition of DNA in Eukaryotic Cells Using Fourier Transform Infrared Spectroscopy,"
16. *Nucleic Acids Res.*, vol. 39, pp. 5439–5448, 2011.
17. P. T. Wong et al., "Infrared Spectroscopy of Exfoliated Human Cervical Cells: Evidence of Extensive Structural Changes during Carcinogenesis,"
18. *Proc. Natl. Acad. Sci. USA*, vol. 88, pp. 10988–10992, 1991.
19. Vapnik V and Mukherjee S: Support
20. VectorMachine for Multivariate Density Estimation. In: Advances in Neural Information
21. Processing Systems. Leen T, Solla S, and Muller KR (eds.). Cambridge, MA, MIT Press, pp. 659-665, 2000.
22. Moler E, Chow M, and Mian I: Analysis of molecular profile data using generative and discriminative methods. Physiol Genomics 4(2): 109-
23. 126, 2000.

24. Furey TS, Cristianini N, Duffy N, Bednarski DW, Schummer M, and Haussler D: Support vector machine classification and validation of cancer tissue samples using microarray expressiondata. Bioinformatics 16(10): 906-914, 2000.
25. Segal NH, Pavlidis P, Noble WS, Antonescu CR, Viale A, Wesley UV, Busam K, Gallardo H,
26. .
27. DeSantis D, and Brennan MF: Classification of clear cell sarcoma as melanoma of soft parts by genomic profiling. J Clin Oncol 21: 1775-1781, 2003.
28. Segal NH, Pavlidis P, Antonescu CR, Maki RG, Noble WS, DeSantis D, Woodruff JM, Lewis JJ, Brennan MF, and Houghton AN: Classification and subtype prediction of adult soft tissue sarcoma by functional genomics. Am J Pathol 163(2): 691-700, 2003.
29. Gramfort et al., “MNE software for processing MEG and EEG data”, NeuroImage Volume 86, 1 February 2014
30. A. C. Rodriguez, Z. Blanchard, K. A. Maurer, and J. Gertz, "Estrogen Signaling in Endometrial Cancer: A Key Oncogenic Pathway with Several Open Questions," *Horm. Cancer*, vol. 10, pp. 51–63, 2019.
31. Searle, A. and Kirkup, L. A direct comparison of wet, dry, and insulating bioelectric recording electrodes. Physiological Measurement, 21(2),
32. pp.271-283.2000
33. \18. Góngora Alonso S, Marques G, Agarwal D, De la Torre Díez I, Franco-Martín M. Comparison of Machine Learning Algorithms in the Prediction of Hospitalized Patients with Schizophrenia. Sensors (Basel). 2022 Mar 25;22(7):2517. doi: 10.3390/s22072517. PMID: 35408133; PMCID: PMC9003328.
34. K. K. Zorn et al., "Gene Expression Profiles of Serous, Endometrioid, and Clear Cell Subtypes of Ovarian and Endometrial Cancer," *Clin. Cancer Res.*, vol. 11, pp. 6422–6430, 2005.
35. S. G. Silverberg, "Problems in the Differential Diagnosis of Endometrial Hyperplasia and Carcinoma,"
36. *Mod. Pathol.*, vol. 13, pp. 309–327, 2000.
37. Sengupta Cancer Detection: A Review. Biomolecules. 2023 Jun 22;13(7):1024. doi: 10.3390/biom13071024. PMID: 37509060; PMCID:
38. PMC10377131J, Hussain CM. CNT and Graphene-Based Transistor Biosensors for