Diagnosis Screening of Ovarian Cancer using DNA-Functionalized Carbon Nanotube Vapor Sensor

Nicholas Kybert[1], Katharine Prokop-Prigge[2], Cynthia M. Otto[3], Lorenzo Ramirez[3], EmmaRose Joffe[3], Janos Tanyi[4], Jody Piltz-Seymour[5], A. T. Charlie Johnson[1,a)], and George Preti[2,6]

University of Pennsylvania, Department of Physics and Astronomy, Research Mentors: A.T. Charlie Johnson and Christopher Kehayias, Aukkarwit Nakprasert, College of Arts and Sciences 2023

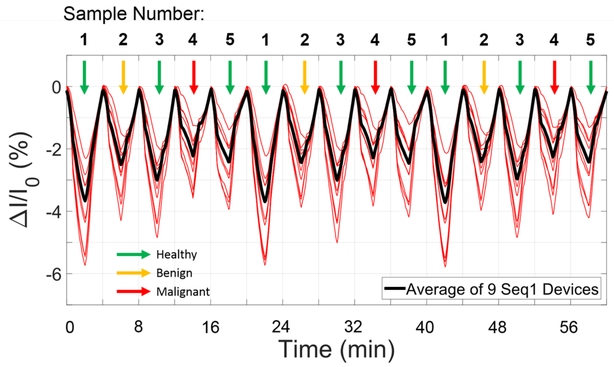
Introduction

- The fatality rate of ovarian cancer varies from 10% if discovered in an early stage up to 80% when discovered in a late stage.
- There is no current approved way to screen this cancer and the early stage symptoms are often confused with other diseases.
- Blood samples from patients with ovarian cancer have different composition of volatile organic compounds(VOCs).
- Carbon nanotubes have sensitive electronic properties, but they do not react with VOCs.
- Functionalized DNA is added to generate the binding sites for detecting the chemicals, creating the sensor.
- The research focuses on the detection of VOCs by measuring the drop in electric current from the sensors after they are exposed to blood samples.
- Hypothesis: Each type of samples(malignant, benign, and healthy) will have different composition in VOCs that can be distinguish and categorize by the sensor.

Methods

- 58 blood sample are used in the experiment with 21 healthy, 16 benign, and 21 malignant samples.
- The sample are connected to the sensor via the gas chamber. They are heated in water bath to extract VOCs. The shift in electric current from the base line is measured. Nitrogen gas is then used to clear all the measurement of the sensor, getting ready for the next one.

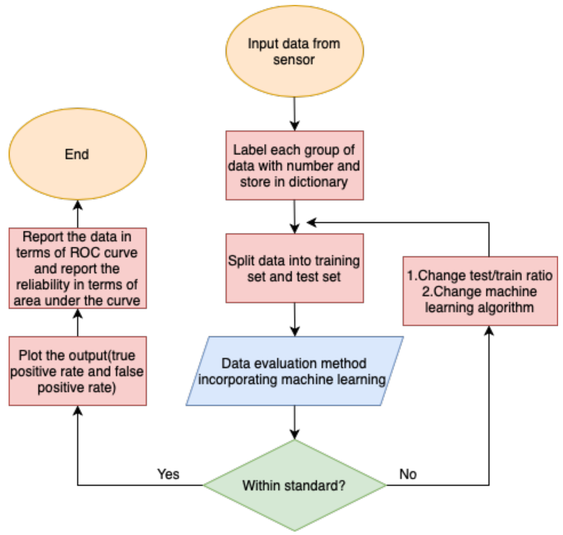
Figure1: Measurement example



- The experiment repeats and the average data are then carried on to the data analysis process to evaluate the accuracy.

- Machine learning algorithms are used to test the accuracy of the raw data. There are four mains algorithms related in this experiment: LDA, SVM, KNN, and Random Forest.
- The most accurate method will be used for presenting the result.

Figure2: Data Analysis process



Results

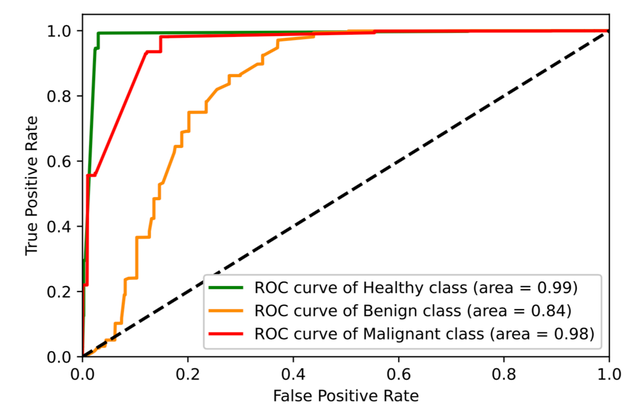
Figure3: Accuracy of each algorithm

Classifier	Overall	Healthy	Benign	Malignant	False Positives	False Negatives
LDA	55/58 (95%)	21/21	15/16	19/21	0	3
SVM	52/58 (90%)	20/21	15/16	17/21	1	5
KNN	53/58 (91%)	21/21	15/16	17/21	0	5
Random Forest	51/58 (88%)	19/21	14/16	18/21	3	4

- Based on Fig.3, LDA algorithm has the highest accuracy of all(95%).
- The number of false negative are consistently high for all algorithm. False negative value appears when the test results incorrectly fails to indicate the presence of the disease when it is present.

- The data analysis for LDA method is repeated for 1000 times and the ROC curve is plotted from the data as shown in Fig.4.
- In comparison to Fig.3, the area under the curve for benign class is lowest(0.88)
  meaning that the classification of this class is the least reliable.
- In contrast, the LDA algorithm consistently correctly categorizes samples from both healthy and malignant groups since the scores are 0.99 and 0.98 respectively

Figure4: ROC curve presentation for LDA method



Conclusions

- ssDNA-CNTs nanosensors are able to discriminate each sample groups clearly from each other
- Although the accuracy for screening ovarian cancer from the sensor is high(above 80%), there are still many concerns related to the result
  - There are still no proof that VOCs from blood sample are unique to ovarian cancer.
  - The samples size is only 58 samples, thus there is still no certainty that the accuracy will still remain high for higher sample size.

Acknowledgements & References

1. Preti,George (2020/03/01). Exploring ovarian cancer screening using a combined sensor approach: A pilot study. *AIP Advances*, 10, 035213-. doi: 10.1063/1.5144532
2. Gelperin, Alan (2005/09/01). DNA-Decorated Carbon Nanotubes for Chemical Sensing. *Nano Letters*, 5, 1774-1778. doi: 10.1021/nl051261f

I am personally thankful to my mentor, Christopher Kehayias for taking care of me through out the whole process