DETECTION OF BREAST CANCER USING DISCRETE THERMAL DATA RECORDED USING THE FIRST WARNING SENSOR SYSTEM

NANYANG TECHNOLOGICAL UNIVERSITY

FINAL RESEARCH REPORT

By:

DR. EDDIE NG YIN-KWEE
(Associate Professor)

DR. VINITHA SREE SUBBHURAAM
(Research Associate)

School of Mechanical and Aerospace Engineering
Nanyang Technological University
Nanyang Avenue
Singapore 639798

2011

(Confidential excerpt provided by FWS exclusively for viewer)

CHAPTER 8: CONCLUSIONS AND FUTURE DIRECTIONS

8.1 Conclusions

Handling the complexity of cancer diagnosis, treatment, and prognosis would not have been possible without the availability of the currently improving imaging systems. Today's latest imaging systems play a major role in the primary cancer screening, in the diagnosis and characterization of lesions, staging and restaging, treatment selection and treatment progress monitoring, and in the determination of the cancer recurrence. Moreover, adjunct modalities, which provide added confidence to the diagnosis made by primary modalities such as mammography, are finding their way into the market these days. One such adjunct modality is the First Warning System™. It uses discrete temperature data recorded from the breast surface to detect the presence of cancer. In this report, several data mining techniques that were used to analyze the FWS dataset were presented. The following are the key conclusions made in this study.

1. Earlier studies using FWS dataset employed the time independent extrapolation technique. This technique does not utilize the most important feature of the FWS modality, namely, the dynamic monitoring of the temperature. Such a feature is unique to FWS, and hence, methods are needed to utilize the complete time dependent information to predict the class – benign and malignant.

2. Moreover, these studies extrapolated the sample size to train the classifiers efficiently. The higher accuracies obtained in these studies can be attributed to the random selection of temperature sets for extrapolation. Since the selection procedure is random, there would not be any control of which samples to select and which to omit. Even if there was such a choice, such selected datasets will not be a true representation of the original dataset. For submission to FDA, results based on the original data would improve the chances of the device getting accepted into the medical community in a lesser time.

3. Therefore, in this work, the potential of data mining tools and techniques were utilized to extract class-differentiating information and features from the multidimensional time series dataset. Several algorithms (feature subset and classifier combinations) were developed and evaluated. Among these algorithms, the sensor-wise derived features used in the {intentionally not disclosed} resulted in the highest accuracy of 74.38%.

4. The 48-hour recording duration implies that the patient has to wear the sensors for 48 hours. This may sometimes be a discomfort to the patients, and there are more chances of encountering wrong readings due to loose sensors, movement artifacts, and temperature changes in the surrounding environment. To address this issue, two hours of data recorded during the 6 pm to 8 pm time period were selected, and evaluated the algorithms. A key observation was that both the 48-hrs dataset and the 2-hrs reduced time datasets present accuracies that are not drastically dissimilar (74.38% vs. 67.98%, respectively). This is an indication that the FWS data recording time could be reduced from 48 hours to two hours with confidence.

5. Several resampling techniques were studied to improve the classification accuracy. It was found that the classifiers {intentionally not disclosed} can be expected to give higher accuracies (range of 85% to more than 90%), if the future new temperature sets closely resemble the current temperature sets.