Tissue Characterization With an Electrical Spectroscopy SVM Classifier

Shlomi Laufer* and Boris Rubinsky

Abstract—This feasibility study introduces the use of a classifier based on electrical spectroscopy measurements for breast cancer tissue characterization. The classifier is of the support vector machine type, and the vector of data is made of electrical voltage measurements at 12 discrete electrical excitation frequencies over the β dispersion range of the analyzed tissue and at discrete locations selected from information produced by conventional medical imaging. The database was generated through a mathematical simulation model. The performance of the classifier was evaluated through a test of its ability to distinguish between simulations of malignant and benign tissues in the breast. The results demonstrate the feasibility of the concept and illustrate the tissue characterization ability of this classifier.

Index Terms—Classifiers, electrical spectroscopy, mammography, support vector machine (SVM).

I. INTRODUCTION

Breast cancer is the most common cancer among women in the Western world [1], and X-ray mammogram is the standard screening tool for breast cancer. However, while mammograms identify suspicious areas, the character of these areas is frequently inconclusive and a needle biopsy is often performed. There are several types of biopsies, but they are all invasive and expensive, and none are risk-free. Furthermore, in about 60%–85% of the cases, the suspicious tissues are found to be benign [2], [3]. Although imaging techniques, such as sonography, have been shown to reduce the number of unnecessary biopsies [4], [5], the problem is yet to be solved. The classifier concept developed in this study could alleviate the need for needle biopsies for tissue characterization.

Studies have shown that cancerous breast tissue and normal breast tissue have different impedance values [6]–[9], making it reasonable to assume that electrical impedance measurements can be used as an imaging and diagnostic tool. As an imaging system, several electrical impedance tomography (EIT) systems have been developed [10], [11]. The classical approach of EIT is to use the electrical measurements to reconstruct the impedance image of the tissue, from which the tumor location and properties can be determined. However, this method suffers from low resolution [12].

In this paper, we suggest that the mammogram image be used to define the suspicious area and that the usage of spectroscopic electrical measurements be limited to distinguishing between benign and malignant tumors. This makes the problem one of classification, which can be solved using standard classification tools, such as support vector machines (SVMs) [13], [14]. Considering the large number of unnecessary biopsies performed, we believe that reducing this number through

*S. Laufer is with the Research Center for Bioengineering in the Service of Humanity and Society, School of Computer Science and Engineering, Hebrew University of Jerusalem, Jerusalem 91904, Israel (e-mail: shlomi.laufer@mail.huji.ac.il).

B. Rubinsky is with the Research Center for Bioengineering in the Service of Humanity and Society, School of Computer Science and Engineering, Hebrew University of Jerusalem, Jerusalem 91904, Israel (e-mail: rubinsky@cs.huji.ac.il).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.
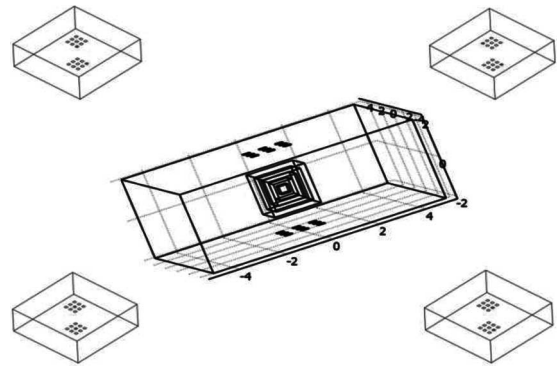
Fig. 1. Three-dimensional model used in this theoretical study. The complete 3-D model with the five square cubes can be seen in the center of the figure. The four pairs of electrodes used are marked in the four smaller images.

the use of this noninvasive, not harmful, and inexpensive method is of practical value.

It should be noted that there are previous publications combining EIT and SVM [15]–[17]. To the best of our knowledge, the novelty of our study is the use of SVM as a two-class classifier, rather than a regression method, and the employment of discrete multifrequency electrical spectroscopy in order to distinguish between benign and malignant breast tumors. This is a first-order feasibility study in which a mathematical simulation model is used to create the database.

II. METHODS

A. Data Collection

1) Geometrical Model: We assume that the data which will be available for the construction of the classifier consists of a mammography image, electrical measurements of currents and voltages from electrodes, and needle biopsy data. The biopsy data are needed only in the construction of the classifier. In this theoretical study, the electrical measurements are replaced by the solution of the field equation obtained from a computer simulation of the problem. The biopsy data correspond to the tissue properties used in the mathematical simulation.

The mathematical model attempts to mimic the information that is available from the X-ray mammography and the biopsy. We assume that the electrical measurements are performed in a configuration similar to the one suggested by Myoung-Hwan *et al.* [10], meaning that the breast tissue has the same geometry during the electrical measurements as during the mammography and that the measurement electrodes are on the mammography plates (see Fig. 1). This particular configuration was chosen since we assume knowledge of the suspicious tissue location from the mammogram image, and this method avoids the need to register the mammogram image. The breast between the mammography plates is modeled as a 3-D box. The rounded edges of the breast are ignored assuming that they are "far enough" from the suspicious tissue.

It is also assumed that the mammogram image provides only rough information of the suspicious tissue's characteristics and dimensions and that it is impossible to extract the exact size and shape of the suspicious tissue. In order to simulate this partial knowledge situation, we assumed that the suspicious tissue has a cube-like geometry, as shown in the center of Fig. 1, with randomly generated variations of the tissue composition.

2) Mathematical Model of Electrical Measurements: Measurements made with electrodes are simulated through the solution of the
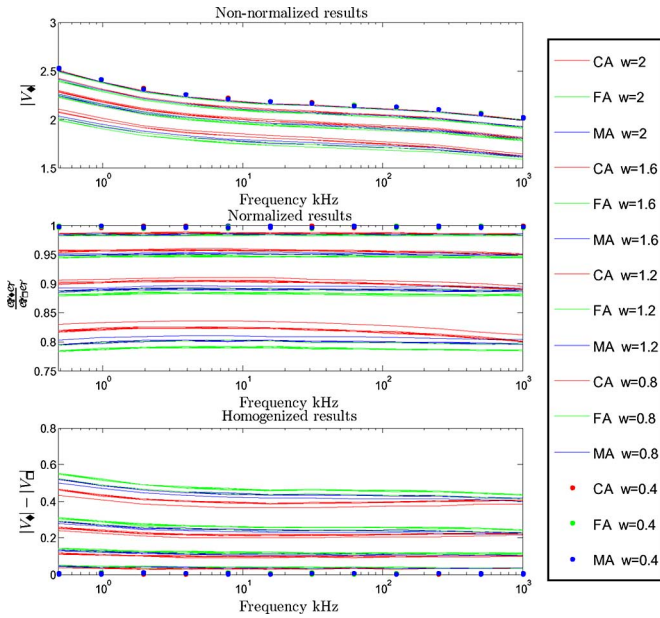
Fig. 2. Typical data used with the classifier: (a) Raw data. (b) Normalized data as in (2). (c) Homogenized data as in (3). Using the normalized preprocessing, it can be seen that the effect of tumor size of 0.4 cm is less than 1%, which may explain the difficulty in classifying small tumors in the presence of noise. w — square cube width.

field equation

$$\nabla \cdot (\sigma \nabla \mathrm{u}) = 0, \qquad x \in \Omega \qquad (1)$$

in the geometry of Fig. 1. It is assumed for boundary conditions that the entire outer surface of the cube is insulated except for two electrodes used for injecting the currents, I and $-I$. The measurement is the calculated voltage on the sites of two measurement electrodes. The voltage measurement electrodes can either be the same electrodes as the current injection electrodes or different ones. For each electrode couple, the voltage is measured for "n" different current injection frequencies.

B. Data Preprocessing

Several different types of data preprocessing were evaluated and found to be useful

$$V_{\mathrm{normalized},j} = V_{\mathrm{tumor},j}/V_{\mathrm{regular},j} \qquad (2)$$

$$V_{\mathrm{homogenized},j} = V_{\mathrm{tumor},j} - V_{\mathrm{regular},j} \qquad (3)$$

$$V_{\mathrm{shift},j} = V_{\mathrm{tumor},j} - V_{\mathrm{mean}} \qquad (4)$$

where $\{V_j \mid j = 1 \ldots n\}$ are a set of measurements (n frequencies) for a specific electrode configuration. V_{tumor} are measurements done near the suspicious tumor. V_{regular} are measurements done far enough from the tumor, or on the other breast. This regular area can be verified using the mammography image. Finally $V_{\mathrm{mean}} = \frac{1}{n} \sum_{j=1}^{n} V_{\mathrm{tumor},j}$.

The first two preprocessing steps can be useful in two ways. First, they help in estimating the influence of the tumor on the measurements with respect to regular tissue, as can be seen in Fig. 2. The second point is related to the fact that the electrical properties of normal tissue vary among different women. In using these data processing steps, we can, to some extent, mask these differences and emphasize the changes caused by the tumor. The third preprocessing step will result in shifting all the measurements around a mean value of zero. By using this shift, we can improve the classifier's generalization for different tumor sizes.
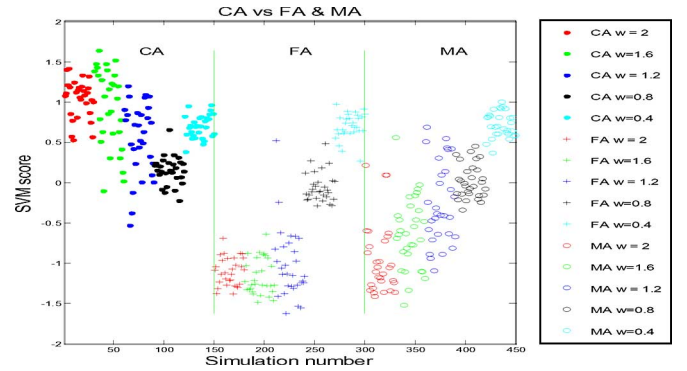

Fig. 3. Results of last classifier in Table I. w — square cube width.

C. SVM Classifier Training

An SVM classifier was used in order to distinguish between malignant and benign breast tumors.

The data are presented as

$$\{(x_1, y_1), \ldots, (x_m, y_m)\}, x \in \mathbb{R}^n, y \in \{-1, 1\}. \qquad (5)$$

In our case, the vector x_i is a vector of length n, and each entry in the vector is a voltage value measured for a different electrical excitation frequency. $y_i = 1$ for a malignant tumor and $y_i = -1$ for a benign tumor. The index $\{i = 1 \ldots m\}$ is one index for each "simulated woman's electrical spectroscopy study."

D. Measurement Noise

The robustness of the classifier to measurement noise was also checked. Noise was added in the following manner [18]:

$$(V_0)^* = V_0 (1 + A_\nu) \qquad (6)$$

where V_0 is a vector of voltage measurements, ν is a vector of normally distributed random numbers, with mean 0 and variance 1 (same dimensions as V_0), and A is the noise level.

E. Electrode Combination

The basic configuration used only one pair of electrodes. A more complex classifier can use multiple electrode configurations. These multiple measurements were used in two ways.

Majority: Train a separate classifier for each configuration, then give each classifier a binary score (1 for malignant and -1 for benign), and finally sum all the classifiers. (7)
Summation: Sum all the measurements and then train one classifier. (8)

III. RESULTS

A. Model Description

A 3-D mesh ($\sim$40 000 elements) was created using Comsol Multiphysics Version 3.3. The solution of (1) was obtained using the forward solver of EIDORS [19]. The size of the box was $10 \times 10 \times 4$ (length $\times$ width $\times$ height). In total, we used 18 electrodes (3×3 on each side), of which four pairs were sufficient for this study (Fig. 1). Five different-sized square cubes were used to represent the suspicious tissue, with widths of 0.4, 0.8, 1.2, 1.6, and 2.0, respectively. In a study conducted by Jossinet [8], [9], the electrical impedance values of breast tissue were measured for 12 frequencies in the β dispersion range. The

TABLE I
RESULTS OF CLASSIFIERS WITH NO SUSPICIOUS AREA SIZE

Noise Level A	Number of electrodes used	Electrode combinations	sensitivity	specificity	AUC	90AUC[20]
0	1	-	100%	100%	1±0	1±0
1%	1	-	95%	38%	0.8±0.02	0.38±0.04
1%	16	Majority (7)	95%	70%	0.89±0.01	0.61±0.03
1%	16	Summation (8)	95%	63.3%	0.87±0.01	0.63±0.03

"A" is the noise parameter taken from (6). Electrode combination definition as in (7) and (8). ± standard deviation. The training data subset included 70 of the 100 simulations for each combination of suspicious tissue type and cube size (70 $_{simulations}$ ×3 $_{tissue\ type}$ ×5 $_{cube\ size}$, a total of 1050 simulations), and the test data subset included the other 30 (30 $_{simulations}$ ×3 $_{tissue\ type}$ ×5 $_{cube\ size}$, a total of 450 simulations).

TABLE II
RESULTS OF CLASSIFIERS WITH SUSPECT AREA SIZE

	A = 1%				A =2%				A=5%			
Width	Sens.	Spec.	AUC	90AUC	Sens.	Spec.	AUC	90AUC	Sens.	Spec.	AUC	90AUC
2	100%	100%	1	1	100%	100%	1		96.7%	98.3%	0.99	0.95
1.6	100%	100%	1	1	96.7%	100%	0.99	0.99	96.7%	85%	0.97	0.81
1.2	96.7%	96.7%	0.99	0.97	96.7%	85%	0.97	0.85	96.7%	56.6%	0.84	0.55
0.8	96.7%	63%	0.94	0.68	96.7%	21.7%	0.78	0.16	96.7%	8.3%	0.58	0.09
0.4	96.7%	5%	0.59	0.24	100%	0%	0.5	-	100%	6.5%	0.58	0.1

"A" is the noise level from (6). Sens.: sensitivity. Spec.: specificity. All 16 electrode combinations were used as in (8). The training data subset included 70 simulations for each tissue type (70 $_{simulations}$ ×3 $_{tissue\ type}$, a total of 210 simulations), and the test data subset included the remaining 30 for each tissue type (30 $_{simulations}$ ×3 $_{tissue\ type}$, a total of 90 simulations).

mean values from that study were used in our simulations. For the normal tissue, we used the properties of adipose tissue (AT: normal breast tissue). We differentiated between the following pathological tissues: carcinoma (CA: malignant tumors) versus mastopathy (MA: a general term covering various benign breast diseases) and fibroadenoma (FA: benign tumors of the breast). In order to simulate the variations in tissue properties and their uncertainty associated with imaging and biopsy (see Section A.1), we randomly chose 75%–90% of the cube and set its connectivity to one of the three pathological tissues. The rest of the box was left with AT.

B. Data Collection

All the data for training and testing the classifiers were obtained from the computer simulations. For each different pathological tissue type (three types) and for each different suspicious tissue cube size (five sizes), we ran 100 different simulations (a total of three tissue types × five cube sizes ×100 random selections). Thus, a total of 1500 simulations were completed. In each simulation, we randomly chose 75%–90% of the cube to be the pathological tissue and the rest to be adipose tissue. This means that each simulation had a unique impedance distribution. Each simulation included 16 different combinations of electrode current and voltage pairs and was simulated for the 12 different frequencies. In relation to real life classifiers, each model represented one woman for whom the mammography and biopsy were known and on whom electrical measurements were done.

C. Classifiers

Several classifiers were trained in order to examine the different parameters discussed in the Section II. In the first group of classifiers, the suspicious area size was ignored. The different classifiers in this group are presented in Table I, and the SVM scores for one of them in Fig. 3. In the second group, the information on the suspicious area size was used, and five different classifiers were trained, one for each size. The results for several different noise levels are provided in Table II.

Since the main aim of this method is to reduce the number of unnecessary biopsies with a minimal number of subsequent malignancy misses, we set the minimal sensitivity (the percent of malignant tumors

classified as such) at 95%. The specificity (the number of benign tumors classified as such) can be seen as the percent of unnecessary biopsies avoided. Similar assumptions have been made in other studies [2], [20]. All the classifiers in this study were trained and tested using the program svmLight [21].

IV. DISCUSSION

This feasibility study introduced the use of a classifier based on multifrequency electrical spectroscopy measurements for breast cancer tissue characterization. The results of the study show that the usage of more than one electrode can, in fact, improve the classification of benign and malignant breast tumors. It also appears that knowledge of the estimated tumor size can improve the classifier's capabilities. This size estimation can be provided by the radiologist or by a CAD program. In clinical examinations, more information can be extracted, such as different classifiers for microcalcifications and for suspicious masses. Further clinical work can also investigate the option of adding electrical data to existing CADx systems in order to improve their present classification capabilities and to facilitate the construction of one multimodality classifier. In the presence of noise, the electrical measurements do not appear to be sufficient for the classification of small tumors (0.4 cm and 0.8 cm for 2% noise and above).

V. CONCLUSION

This study was conducted to evaluate the ability of a classifier using noninvasive electrical spectroscopy measurements to distinguish between malignant and benign tumors in a suspicious area identified on mammography. The results demonstrate the feasibility of this tissue characterization technique. Obviously, much more research is needed to optimize the design of the classifier for this use and to build a clinical database of information. However, the technique should have the potential to improve the characterization abilities of conventional imaging with relatively simple electrical measurements. Furthermore, it may provide a convenient alternative to needle biopsies.

REFERENCES

[1] American Computer Society, *Cancer Facts and Figures 2007.*   Atlanta, GA: American Cancer Society, 2007.

[2] A. O. Bilska-Wolak and C. E. Floyd, Jr., "Development and evaluation of a case-based reasoning classifier for prediction of breast biopsy outcome with BI-RADS lexicon," *Med. Phys.*, vol. 29, pp. 2090–2100, Sep. 2002.

[3] E. S. Burnside, E. A. Sickles, R. E. Sohlich, and K. E. Dee, "Differential value of comparison with previous examinations in diagnostic versus screening mammography," *Amer. J. Roentgenol.*, vol. 179, pp. 1173–1177, Nov. 2002.

[4] A. T. Stavros, D. Thickman, C. L. Rapp, M. A. Dennis, S. H. Parker, and G. A. Sisney, "Solid breast nodules: Use of sonography to distinguish between benign and malignant lesions," *Radiology*, vol. 196, pp. 123–134, Jul. 1995.

[5] M. A. Dennis, S. H. Parker, A. J. Klaus, A. T. Stavros, T. I. Kaske, and S. B. Clark, "Breast biopsy avoidance: The value of normal mammograms and normal sonograms in the setting of a palpable lump," *Radiology*, vol. 219, pp. 186–191, Apr. 2001.

[6] D. Miklavi, N. Pavšelj, and F. X. Hart, "Electric properties of tissues," in *Wiley Encyclopedia of Biomedical Engineering.*   Hoboken, NJ: Wiley, 2006.

[7] A. J. Surowiec, S. S. Stuchly, J. R. Barr, and A. A. S. A. Swarup, "Dielectric properties of breast carcinoma and the surrounding tissues," *IEEE Trans. Biomed. Eng.*, vol. 35, no. 4, pp. 257–263, Apr. 1988.

[8] J. Jossinet, "Variability of impedivity in normal and pathological breast tissue," *Med. Biol. Eng. Comput.*, vol. 34, pp. 346–350, Sep. 1996.

[9] J. Jossinet and M. Schmitt, "A review of parameters for the bioelectrical characterization of breast tissue," *Ann. N.Y. Acad. Sci.*, vol. 873, pp. 30–41, Apr. 1999.

[10] C. Myoung Hwan, K. Tzu-Jen, D. Isaacson, G. J. A. S. G. J. Saulnier, and J. C. A. N. J. C. Newell, "A reconstruction algorithm for breast cancer imaging with electrical impedance tomography in mammography geometry," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 4, pp. 700–710, Apr. 2007.

[11] H. Ryan, H. Alex, and D. P. Keith, "Design and implementation of a high frequency electrical impedance tomography system," *Physiol. Meas.*, vol. 25, pp. 379–390, 2004.

[12] A. Adler and W. R. B. Lionheart, "Information content of EIT measurements," in *Proc. 13th Int. Conf. Elect. Bioimpedance and 8th Conf. Elect. Impedance Tomogr.*, 2007, vol. 17, pp. 360–363.

[13] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Min. Knowl. Discov.*, vol. 2, pp. 121–167, 1998.

[14] B. Schölkopf and A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond.*   Cambridge, MA: MIT Press, 2002.

[15] Y. Wu, Y. Li, L. Guo, W. Yan, X. Shen, and K. Fu, "SVM for solving forward problems of EIT," in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, 2005, vol. 2, pp. 1559–1562.

[16] Y. Wu, L. Guo, G. Dong, Q. Wu, X. Shen, G. Xu, and W. Yan, "Tissue conductivity estimation in two-dimension head model based on support vector machine," *Conf. Proc. IEEE Eng. Med. Biol. Soc.,* 2006, vol. 1, pp. 1130–1133.

[17] G. Goldswain and J. Tapson, "Kernel ridge regression for volume fraction prediction in electrical impedance tomography," *Meas. Sci. Technol.*, vol. 17, pp. 2711–2720, 2006.

[18] J. F. Edd and B. Rubinsky, "Detecting cryoablation with EIT and the benefit of including ice front imaging data," *Physiol. Meas.*, vol. 27, pp. S175–S185, May 2006.

[19] A. Adler and W. R. Lionheart, "Uses and abuses of EIDORS: An extensible software base for EIT," *Physiol. Meas.*, vol. 27, pp. S25–S42, May 2006.

[20] Y. Jiang, C. E. Metz, and R. M. Nishikawa, "A receiver operating characteristic partial area index for highly sensitive diagnostic tests," *Radiology*, vol. 201, pp. 745–750, Dec. 1996.

[21] T. Joachims, "Making large-scale SVM learning practical," in *Advances in Kernel Methods—Support Vector Learning*, B. Schölkopf, C. Burges, and A. Smola, Eds.   Cambridge, MA: MIT Press, 1999.

Metabolic Prosthesis for Oxygenation of Ischemic Tissue

Elias Greenbaum*, *Member, IEEE*,
Mark S. Humayun, *Member, IEEE*, Charlene A. Sanders, Dan Close,
Hugh M. O'Neill, and Barbara R. Evans

Abstract—This communication discloses new ideas and preliminary results on the development of a metabolic prosthesis for local oxygenation of ischemic tissue under physiologically neutral conditions. We report for the first time selective electrolysis of physiological saline by repetitively pulsed, charge-limited electrolysis for the production of oxygen and suppression of free chlorine. Using 800-μA amplitude current pulses and <200 μs pulse duration, we demonstrate prompt oxygen production and *delayed* chlorine production at the surface of a fused 0.85-mm diameter spherical platinum electrode. The data, interpreted in terms of the ionic structure of the electric double layer, suggest a strategy for *in situ* production of metabolic oxygen via a new class of "smart" prosthetic implants for ischemic disease such as diabetic retinopathy. We also present data indicating that collateral pH drift, if any, can be held constant using a feedback-controlled three-electrode electrolysis system that chooses an anode and cathode pair based on pH data provided by a local sensor.

Index Terms—Debye length, ischemic tissue, metabolic prosthesis, oxygenation, pH clamp.

I. INTRODUCTION

Physiological saline (and all physiological fluids) contains dissolved sodium (Na^+) and chloride (Cl^-) ions. When a dc electric current is passed through saline, oxygen and chlorine gases are produced at the anode via oxidation of water and Cl^- and hydrogen is primarily produced at the cathode. The anodic reactions are
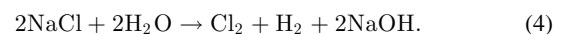
$$2H_2O \rightarrow O_2 + 4H^+ + 4e^- \tag{1}$$

and

$$2Cl^- \rightarrow Cl_2 + 2e^-. \tag{2}$$
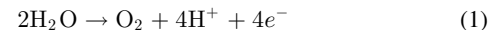
The primary cathodic reaction is

$$2H_2O + 2e^- \rightarrow H_2 + 2OH^-. \tag{3}$$

When the objective of the electrochemistry is to produce chlorine and sodium hydroxide, as in the chloralkali process, the net reaction of interest is

$$2NaCl + 2H_2O \rightarrow Cl_2 + H_2 + 2NaOH. \tag{4}$$

Oxygen evolution is always present. Unlike the present application, the goal of the chloralkali process is to maximize chlorine production