IEEE Journal of Translational Engineering in Health and Medicine

Received 15 March 2017; revised 16 August 2017; accepted 15 September 2017. Date of publication 15 November 2017; date of current version 6 December 2017.

Digital Object Identifier 10.1109/JTEHM.2017.2757471

Efficient Cancer Detection Using Multiple Neural Networks

JOHN SHELL, (Member, IEEE), AND WILLIAM D. GREGORY, (Senior Member, IEEE)

Aurora Sinai Medical Center, Novascan Inc., Milwaukee, WI 53233, USA

CORRESPONDING AUTHOR: J. SHELL (jrsjon@gmail.com)

This work was supported in part by Novascan LLC and in part by National Science Foundation under Grant 0944454 and Grant 1058413. This paper has supplementary downloadable material available at http://ieeexplore.ieee.org., provided by the author.

ABSTRACT The inspection of live excised tissue specimens to ascertain malignancy is a challenging task in dermatopathology and generally in histopathology. We introduce a portable desktop prototype device that provides highly accurate neural network classification of malignant and benign tissue. The handheld device collects 47 impedance data samples from 1 Hz to 32 MHz via tetrapolar blackened platinum electrodes. The data analysis was implemented with six different backpropagation neural networks (BNN). A data set consisting of 180 malignant and 180 benign breast tissue data files in an approved IRB study at the Aurora Medical Center, Milwaukee, WI, USA, were utilized as a neural network input. The BNN structure consisted of a multi-tiered consensus approach autonomously selecting four of six neural networks to determine a malignant or benign classification. The BNN analysis was then compared with the histology results with consistent sensitivity of 100% and a specificity of 100%. This implementation successfully relied solely on statistical variation between the benign and malignant impedance data and intricate neural network configuration. This device and BNN implementation provides a novel approach that could be a valuable tool to augment current medical practice assessment of the health of breast, squamous, and basal cell carcinoma and other excised tissue without requisite tissue specimen expertise. It has the potential to provide clinical management personnel with a fast non-invasive accurate assessment of biopsied or sectioned excised tissue in various clinical settings.

INDEX TERMS Biological neural networks, feedforward neural networks, biomedical engineering, bioimpedance, cancer detection.

I. INTRODUCTION

The accurate detection of cancer is a challenging task. According to the Centers of Disease Control and Prevention (CDC), breast cancer is the most prevalent cancer in women and in 2013 accounted for 230,815 women and 2,109 men being diagnosed with breast cancer [1]. We implement a cascaded neural network cancer classification algorithm based solely on utilizing statistical variation and central tendency information between malignant and benign breast tissue impedance data as input to the neural network.

NovaScan Inc. has developed a portable noninvasive device that successfully characterizes benign versus malignant tissue based on an analysis of the frequency sampled impedance data collected [2]. That analysis is based on one of the four parameters in the Cole-Cole model, *fc*, the relaxation or characteristic frequency of the least squares fit to the impedance data [3], [4]. The fitting parameter R_0 (ohms) is low frequency resistance, R_∞ (ohms) represents high frequency resistance, alpha(α) is a unit-less parameter, ranging from 0 to 1 and is a measure of the dispersion of the data, and j is the complex $\sqrt{-1}$. The complex quantity $\boldsymbol{Z}^*$ has a real (resistive) component, Z' and an imaginary (reactive) component Z''. The model is embodied in (1).

$$Z^* = Z' + Z'' = \frac{R_0 - R_\infty}{1 + \left(j\frac{f}{f_c}\right)} + R_\infty \tag{1}$$

Cole and Cole [3], [4] credited with the development of (1) is most noted for his contribution to biophysics, however, his earlier work also included collaborations with [5], [6] that describe the voltage-current relationship of living cell membranes. The effort of [7] expanded this effort to ''obtain simultaneous measurements, in some simple artificial systems, of the electrical properties mentioned and to modify and

extend the kinetic theory of ion motion to cover the observed phenomena to the extent that they may be physical rather than metabolic." A significant physiological characterization related to the Cole model is called the equilibrium or reverse potential. It is the value of transmembrane voltage where the movement of ions and electrical activity exist such that there is no net ion flow across the membrane. The equilibrium potential for any ion can be calculated using the Nernst equation [7] depicted in (2). Nernst [7] attributes this phenomenon as the membrane being in thermodynamic equilibrium where there is no net flux (z) of ions.

$$V_m = \frac{RT}{zF} \ln \frac{[ion]_{outside}}{[ion]_{inside}} \tag{2}$$

V_{m} is the membrane potential (in volts or joules/coulomb), $[\mathrm{ion}]_{\mathrm{outside}}$ is the extracellular ion concentration in (moles/m^3), $[\mathrm{ion}]_{\mathrm{inside}}$ is the intracellular ion concentration in (moles/m^3), R is the ideal gas constant in (joules/Kelvin/mole), T is the temperature in Kelvin, and F is Faraday's constant in (coulombs/mole), z is the number of electrons transferred in the reaction.

The model in (2) is valid when expressing one specific ion, however since various ion channels and transport mechanisms provide varying permeability to distinct ions such as Na^+, K^+, Ca^{2+}, and Cl^-, a voltage differential exists between the intracellular cytoplasm and extracellular space. This potential difference, *Vm*, is the membrane potential [8]. A cell is considered depolarized when *Vm* is less negative and depolarized due to conductance charges, or 1/***Z*** **'**in relationship to the Cole model in (1). A model that encompasses varying ionic conductance and takes into account multiple ions is the model equation in (3) developed by [9] and [10].

$$V_m = \frac{RT}{F} ln \frac{P_{Na^+}[Na^+]_o + P_{K^+}[Na^+]_o + P_{Cl^-}[Cl^-]_o}{P_{Na^+}[Na^+]_i + P_{K^+}[Na^+]_i + P_{Cl^-}[Cl^-]_i} \tag{3}$$

It is well established that cancer cells possess bioelectrical properties distinctly different than normal cells and electrophysiological analysis in cancer cell types report a depolarized *Vm* that contributes to cellular proliferation and migration. Recent research suggests that *Vm* has a functional role in cancer migration [8].

Figure 1 displays a list of *Vm* values gathered by and provided courtesy of the authors in [8].

We can now calculate the Nernst equation for generally accepted values of transmembrane ionic K^+, where $K^+_{\mathrm{inside}} = 100$ (mM) and $K^+_{\mathrm{outside}} = 10$ (mM) to calculate *Vm,* which is approximately -70mV. If we now vary the range of K^+ inside and outside the cell, we can depict depolarized *Vm* and its relationship to malignancy and the correlation to the Cole relaxation frequency, *fc*, Fig 2.

The data used in this work were collected in an IRB approved study at two different Aurora Medical Center facilities in Milwaukee, WI. The database consisted of 232 histologically confirmed malignant and 191 benign excised breast tissue patient impedance files consisting of 47 frequency samples in the range, 1 Hz – 32 MHz. Statistics reported

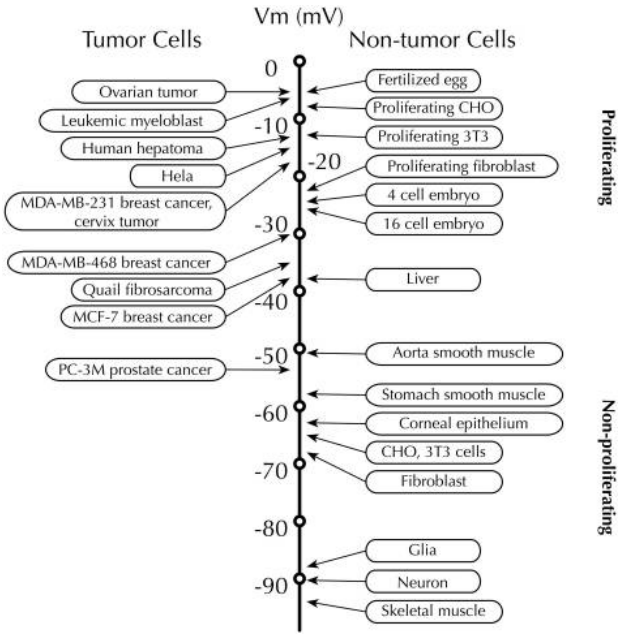

FIGURE 1. **Vm (mV) for various tissue types provided courtesy of [8].**

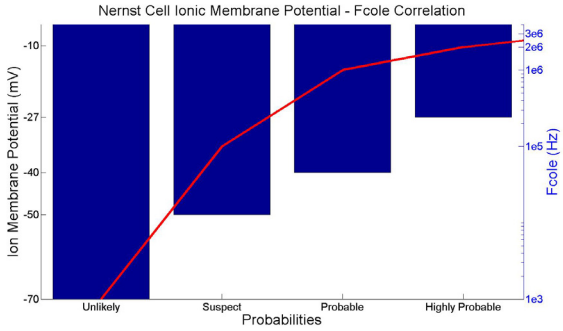

FIGURE 2. **Nernst calculation for K^+ depicting depolarized *Vm* and the resulting Fcole, *fc*, probabilities.**

in a publication of the results from that study [2] indicated sensitivity of 100% and specificity of 85%. Erroneous false positive (FP) or false negative (FN) results can lead to over diagnosis or under diagnoses respectively. Screening mammograms can account for about a 20% false negative rate [11]. The analysis reported in [2] reduced the FN rate compared with screening mammography to 0%; however, the goal of this endeavor is to further reduce the FP rate of 15%.

II. METHODS AND PROCEDURES

The raw impedance based files utilized to implement the multiple neural network classification schemes were collected from a previous study database [2]. The device developed by Novascan used to collect impedance data from excised tissue samples were: a symmetric planar four electrode sensor array consisting of silver/silver chloride (Ag/AgCl) or blackened platinum (BPt) with each electrode measuring 5 mm long by 0.5 mm wide separated by 0.5 mm spacing. The outer electrodes serve to apply current while the inner electrodes sense the resulting specimen voltage. Applying Ohm's Law,

IEEE Journal of Translational Engineering in Health and Medicine

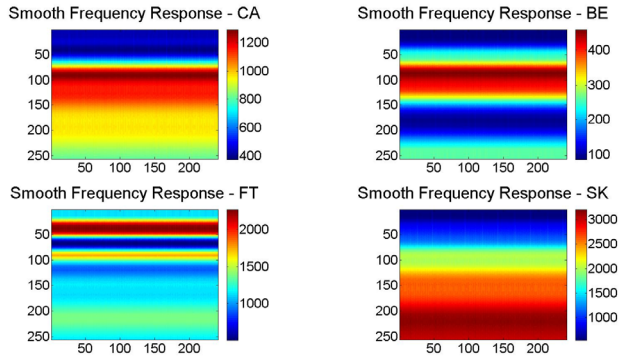

FIGURE 3. Smooth Frequency Response windowed for 256x256 Mel filter bank interpolation for 4 Tissue types displaying sub-band power variability.

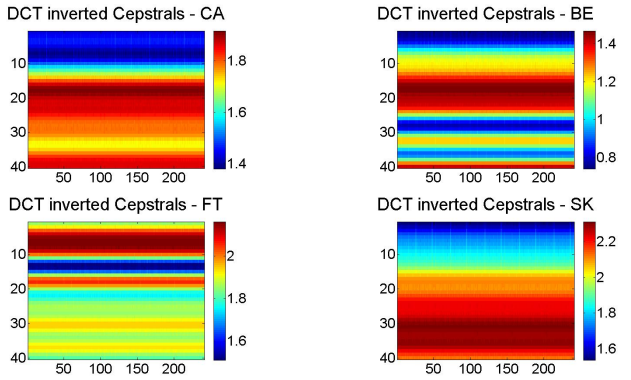

FIGURE 4. Direct Cosine Transfrom of inverted cepstrals windowed for 256x256 Mel filter bank interpolation for 4 Tissue types displaying sub-band power variability.

impedance (Z^*) = voltage (V) / current (I) provides the measured complex impedance based on a fixed complex applied alternating current voltage and known current. The device and electrode circuitry is optically connected to a laptop computer which stores the real (Z') and imaginary (Z'') components of Z^* from 1 Hz – 32 MHz. The laptop also provides a graphical user interface to monitor the Z^* components and view the final Cole analysis result. A local regression nonparametric smoothing method, Loess [13]–[15], is applied to the data before the smoothed data is then fit to the Cole-Cole function where the Cole relaxation constant, *fc* is extracted. This parameter is then used to designate or classify data files cancer (CA) or benign (BE). A total of 360 files comprised on 180 CA and 180 BE files confirmed through pathology were randomly chosen from the database for inclusion as input to the neural networks. The dataset was then apportioned as follows; 70% of each CA and BE set (126 each) were set aside for training, 15% of each set (26 each) reserved for a validation set and the remaining 15% (26 each) was used as test sets that would not be introduced to the networks until the final stage.

The primary focus of this study is to distinguish cancer (inclusive of all types mentioned above) from benign normal tissue. The study did include tissue samples of fatty tissue and skin which contributed to confirmation that impedance

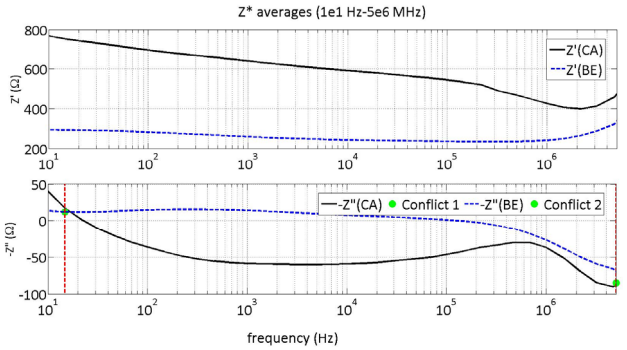

FIGURE 5. 47-sample dataset averages of 180 CA and 180 BE impedance files from 10 Hz – 5 MHz displaying potential conflicts in data differentiation.

data agreed with established and well documented impedance values.

One consideration to extract features for the neural network examined spectral features to discern malignant from normal tissue. Fig. 3 and Fig. 4 utilize analysis developed in [16] show the viability and feasibility of Mel-frequency cepstral coefficients as neural network input, however, the statistical variability analysis explained later formed a solid basis for neural network input.

An important aspect in the design of a neural network is judicious scrutiny of the datasets that will be used as input to the network. Fig 5 depicts the 47 sample raw impedance datasets averaged over each frequency in the 1 Hz – 32 MHz range. It can be observed that data in the imaginary part of Z^* for CA and BE overlap at approximately 15 Hz and data at the higher frequencies around 5 MHz could also induce or contribute to error between the two datasets.

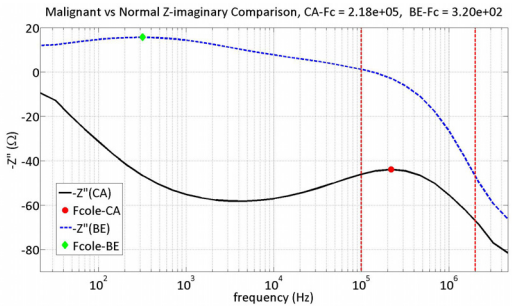

FIGURE 6. 33-sample dataset averages of Z″ for CA and BE from 10 Hz – 4 MHz displaying the Cole frequency, fc and the frequency range of interest.

The decision to exclude lower frequency data was motivated by several works expounding the advantage of eliminating adverse electrode polarization effects that is particularly noticeable below 1 kHz [17]–[19]. It was then necessary to consider excluding data above 3.2 MHz, Fig. 6 displays the result of the modified 1e3 kHz – 3.2 MHz datasets. The dashed vertical lines in Fig. 6 demarcate the frequency range where the probability that the highest incidence of CA is expected based on our experimentation in [2].

We consider 33 data samples in the frequency range 1 kHz to 3.2 MHz excluding potential error in the classification

paradigm to address adverse polarization effects particularly noticeable below 1 kilohertz (kHz). Fig 6 includes the Cole-Cole parameter, *fc*, for the averaged and smoothed Z''-CA and Z''-BE.

The vertical lines in Fig. 6 demarcate the frequency range where the probability that the highest incidence of CA is expected based on our observations in [2].

To confirm these observations a statistical analysis was performed. The first approach was a two-sided two-sample Kolmogorov-Smirnov test for $\boldsymbol{Z}^*$ of the CA and BE averaged 47 and 33 sample datasets to determine the probability they are from the same distribution. The significantly small p-value for both datasets were an indication to reject the null hypothesis that the 2 datasets are from the same distribution and led to acceptance of the alternative hypothesis that the two datasets, CA and BE, are different. Another tool utilized to provide insight into the 47 and 33 sample datasets is the Kruskal-Wallis test, a nonparametric version of one-way ANOVA, and an extension of the Wilcoxon rank sum test. The test makes a rank comparison of the medians of the groups of data in question to determine if the samples come from the same dataset. A p-value of 0.999 for the Kruskal-Wallis Z' comparison and 0.987 for the Z'' test indicate significant differences in both sets. An examination of one final comparison is a balanced one way ANOVA to ascertain some insight regarding the mean of the CA and BE datasets. Fig. 6 displays the modified CA and BE datasets.

Results of the statistical analysis displayed in Figs. 7, 8, and 9 for the modified datasets indicate that the CA and BE datasets display a significant amount of statistical variation that can be utilized as features that will contribute to a robust input data set to the neural networks. The results of the statistical analysis are summarized in Table 1. The high amount of outliers in the original 47 sample dataset can have deleterious effects on statistical analysis that could contribute an increased amount of error in the variance and bias and subsequently might adversely affect the regression algorithms utilized in the neural networks [20].

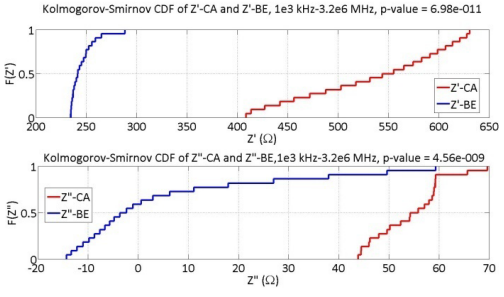

FIGURE 7. Kolmogorov-Smirnov nonparametric test of the comparison of the probability distribution for the 47 sample and 33 sample datasets.

Cross correlation and the results based on a 95% confidence interval in Table 2 was also performed to confirm that there was a significant difference in the 47 sample and

TABLE 1. Group statistics for ca and be datasets.

	Kruskal Wallis p-value	KW Result, Group medians equal	Kolmogorov Smirnov p-value	KS Result, Group medians equal	One Way ANOVA p-value	ANOVA Result, Group means equal	Intra-Group Correlation p-value	Correlated p-value < 0.05
Z' CA v Z' BE-47	0.9970	NO	0.0000	NO	0.0000	NO	0.0300	YES
Z''CA v Z''BE-47	0.0329	YES	0.0000	NO	0.5819	YES	0.0000	YES
Z' CA v Z' BE-33	0.9873	NO	0.0000	NO	0.0000	NO	0.4700	NO
Z''CA v Z''BE-33	0.9986	NO	0.0000	NO	0.0000	NO	0.1100	NO
	Z'CA-47	Z'CA-33	Z'BE-47	Z'BE-33	Z''CA-47	Z''CA-33	Z''BE-47	Z''BE-33
Mean	621.38	669.50	273.42	268.47	-36.14	31.14	-22.27	-12.47
VAR	30714.46	3713.57	2640.77	378.52	29731.22	1188.02	5856.32	9.27
STD	175.26	60.94	51.39	19.46	172.43	34.47	76.53	3.04
ANOVA OUTLIERS	3	0	4	1	8	2	13	0

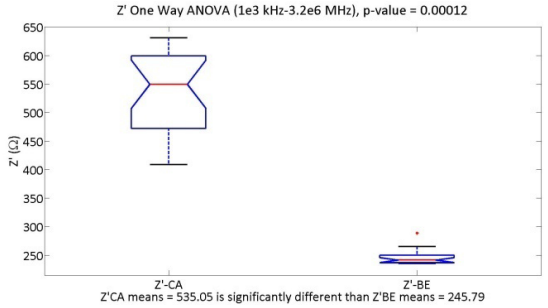

FIGURE 8. One way ANOVA for the 33 sample CA and BE datasets with a p-value of 0.00012 indicating significant difference in the datasets.

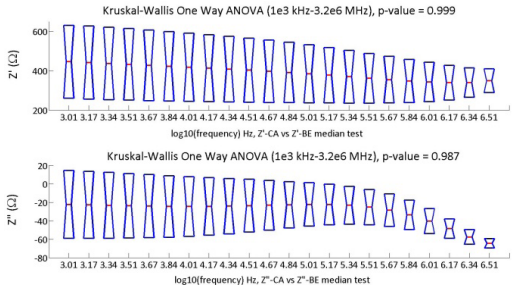

FIGURE 9. Kruskal-Wallis nonparametric test displaying a rank comparison of the medians between the 33 samples BE and CA datasets.

TABLE 2. Correlation statistics for ca and be.

	Number of Samples	Correlation	P Value	Significant
Z' - CA v Z' - BE	47	0.3176	0.0296	YES
Z'' - CA v Z'' - BE	47	0.6625	0.0000	YES
Z' - CA v Z' - BE	33	0.1642	0.4654	NO
Z'' - CA v Z'' - BE	33	0.3485	0.1119	NO

33 sample datasets. A Receiver Operating Characteristic (ROC) curve was also utilized to discern the dataset integrity and is displayed in Fig 10.

III. NEURAL NETWORK BACKGROUND

The use of artificial neural networks (ANN) as computational implementations of biological neural networks has been utilized to solve various machine learning tasks such as speech recognition [16]. The supervised learning algorithm of a binary classifier utilizing a single layer or perceptron can be attributed to [21]. Problems defined by equation 4 can be

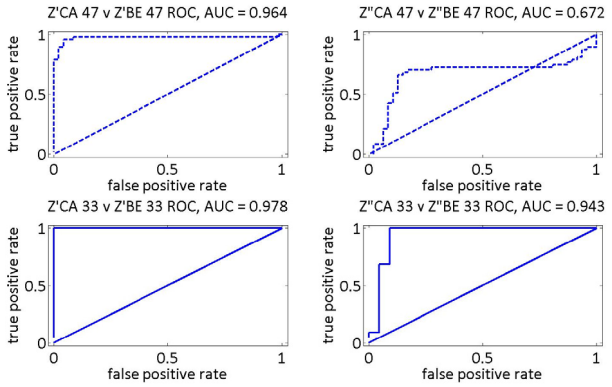

FIGURE 10. 33-sample dataset averages of the imaginary CA and BE impedance files from 10 Hz – 5 MHz displaying Fcole (*fc*) and the frequency range of interest.

addressed by a two dimensional perceptron.

$$y = w_n x_n + w_{n-1} x_{n-1} + \ldots + w_1 x_1 + w_0 \tag{4}$$

This can be expressed as:

$$y = w_1 x_1 + w_0 = mx + b \tag{5}$$

Where w_n indicates the weight vector, x_n the input vector, and w_o the bias, the equation of a line or hyperplane and is the limiting factor that determines accurate perceptron output.

Machine learning problems that involve vectors that are not linearly separable may require one or more additional hidden layers for increased computational capability are referred to as multilayered perceptrons (MLP) [15], [21]–[26].

The six backpropagation MLPs chosen for this implementation are:

1) Gradient descent with momentum and adaptive learning rate backpropagation [27].
2) Levenberg-Marquardt backpropagation [28].
3) Resilient backpropagation [29].
4) Scaled conjugate gradient backpropagation [30].
5) Conjugate gradient backpropagation with Powell-Beale restarts [31], [32].
6) Conjugate gradient backpropagation with Fletcher-Reeves updates [33].

An important facet of neural network design is the development of a network that generalizes well to avoid overfitting. K-fold cross validation is sometimes used to increase the likelihood that the random partition of data for training, validation, and test sets will provide better classification results providing a scheme where all the data is seen by the network in the test set once and all the data is trained at least k-1 times. We observed however in the statistical analysis that standard error in the CA and BE datasets was minimal and should provide robust and reliable input to the neural networks. One can achieve a similar result to cross validation by randomly dividing the data into training and test sets k different times.

Three of the neural network architectures, (4-6) employ conjugate gradient methods and are similar in their optimization approach to reset the search direction to the negative

TABLE 3. Correlation statistics for ca and be.

Network Adjustable Parameters					
	EPOCHS	Perf Goal	Max Validation Fails	Min Perf Gradient	Linesearch routines (10)
GD w/M (10)	500	1e-5	4	1e-5	
LM (7)	40	1e-6	4	1e-6	
RP (9)	300	1e-5	4	1e-5	
SCG (7)	300	1e-6	4	1e-6	Trust Region*
CGB (5)	500	1e-5	4	1e-5	Charalambous
CGF (5)	500	1e-6	4	1e-6	Charalambous

of the gradient but differ in how the reset point is determined. [30]–[35]. Network architectures (1-3) vary in how the weights and bias are altered to achieve the same end result of minimization. All of the networks provide a common parameter to stop training early that helps avoid over fitting by monitoring validation data performance. Early stopping of training occurs when error on the validation data increases. Table 3 lists a summary of some of the parameters that can be adjusted or initialized at the initial onset of training. The values in parenthesis are the actual number of parameters that can be changed.

An indication of neural network performance can include a linear regression to compare network output versus input for each phase of network performance as depicted in Fig 11 of the Levenberg-Marquadt network. The correlation coefficient (R) based on the regression fit in this iteration indicated good performance in the training and validation phases (~0.99 for each) and acceptable overall performance (~0.99). This measure indicates that this network should reliably generalize test data newly introduced to the network.

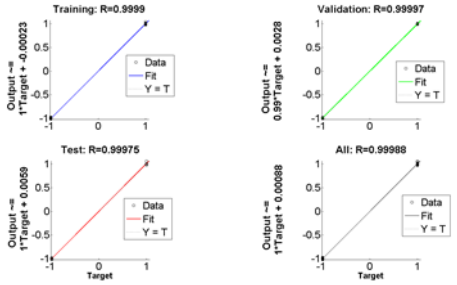

FIGURE 11. Linear regression fits with correlation coefficients for initial Levenberg-Marquadt neural network.

The process to iteratively and exhaustively test network performance after adjusting single parameters or experimentation with additional hidden layers led to reliable and robust final configurations for the six neural networks. Table 3 summarizes individual adjustable network parameters generally common to each network with some algorithm specific parameters excluded.

IV. RESULTS

Six feedforward backpropagation neural networks were configured for this implementation using MathWorks ©Matlab software [34] augmented with modified regression algorithms provided by the author in [16]. Individual network architecture of each network is displayed in Fig. 12. Network

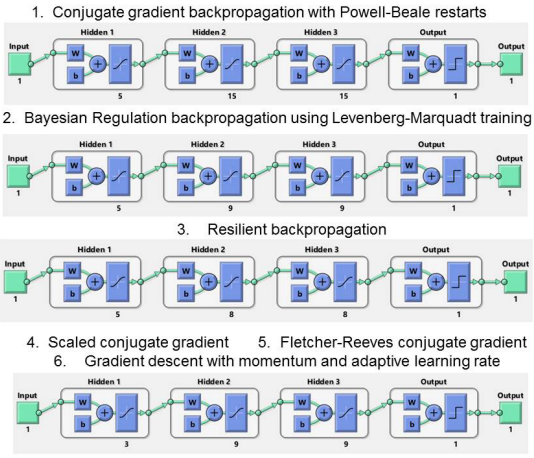

FIGURE 12. Neural network architectures depicting the number of hidden layers and neurons in each hidden layer for the six systems implemented.

TABLE 4. Simulation results for test data.

Optimized Simulated Test Data Neural Network Sensitivity						
	CG POWELL-BEALE	LEVENBERG-MARQUARDT	SCALED CG	RESILIENT BP	CG FLETCHER-REEVES	VAR LR w/MOMENTUM
TP	27	27	27	27	27	27
FP	0	0	0	0	0	0
TN	27	27	27	27	27	27
FN	0	0	0	0	0	0
Sensitivity (%)	100.00	100.00	100.00	100.00	100.00	100.00
Specificity (%)	100.00	100.00	100.00	100.00	100.00	100.00
PPV (%)	100.00	100.00	100.00	100.00	100.00	100.00
NPV (%)	100.00	100.00	100.00	100.00	100.00	100.00

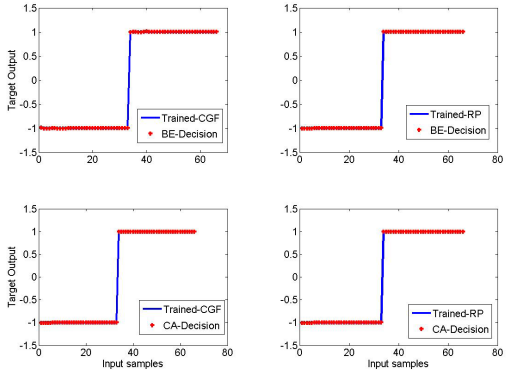

FIGURE 13. Simulation of CGF and RP neural network output response to correct classifications for BE and CA.

implementation is discussed in [15]. The final simulated test data results are summarized in Table 4.

The validation stage provided opportunity to adjust early stopping parameters of the neural networks such as regularization with retesting network performance. The test set led to robust results listed in Table 4.

Simulation results of the test data unseen by the networks until the final stage confirm that the networks generalized to data not seen previously by the networks. The possibility that sample files from the CA and BE datasets could be correlated providing similar output were addressed with software decision analysis based on network output. The simulation in Fig.13 depicts CGF and RP network response to random CA and BE test files. The output responses for the four neural networks not shown were similar.

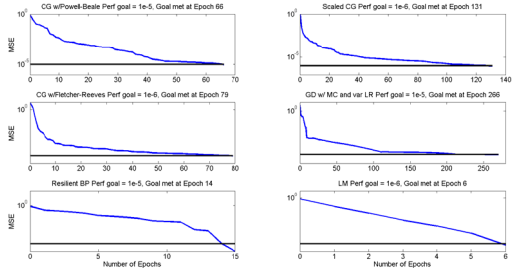

FIGURE 14. Network performance based on mean squared error.

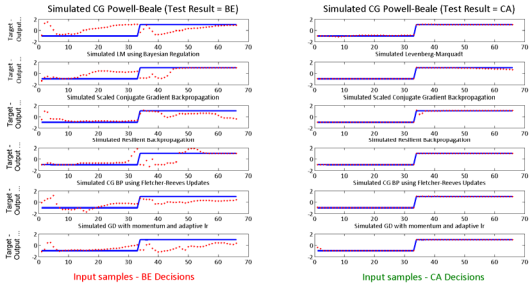

FIGURE 15. Network performance based on mean squared error.

Fig. 14 depicts trained network performance providing insight and indication of how fast some of the regression algorithms converge and how well the networks might generalize based on the slope of the mean squared error (mse).

Fig. 15 displays the network output for one test file. Two network outputs are provided that compare the input data under test to an expected BE result and an expected CA result such that the network output is not absolute but relative to the ensuing comparative statistical analysis of each output response.

V. CONCLUSIONS

The incentive to improve cancer detection accuracy utilizing neural networks appears plausible based on the positive neural network implementation, Frequency dependent variation in impedance data appears to be a viable discerning feature for cancer classification. Data feature selection at the onset contributed to a successful network utilization. Systematic optimization of each neural network that differ in their approach to optimization contributed to highly predictive networks. The overall majority rule analysis calling for 4 of the 6 networks to agree seems prudent to ensure a high level of sensitivity and specificity. This algorithm can potentially be applied in real time immediately after an impedance based tissue scan and can be a useful tool particularly in environments lacking diagnostic apparatus or the requisite skilled practitioners needed to operate that equipment. The NovaScan device is currently being tested on excised squamous and basal cell carcinoma tissue. Successful completion of this feasibility study will lead to applying the technology directly to suspicious lesions on patients. The positive results thus far with breast tissue (N = 423) and limited skin tissue (N = 50) provide strong evidence that the self contained

IEEE Journal of Translational Engineering in Health and Medicine

device consisting of a laptop computer, handheld Probe, and all requisite hardware and software providing an automated analysis is user friendly and highly portable and has the potential to be a beneficial tool in cancer detection. The definitive high level of classification accuracy, sensitivity, and specificity indicated by the results has the potential to dramatically increase and enhance clinical management flexibility by providing an effective patient tissue assessment tool in clinical or non-clinical environments. This device and software could help alleviate the challenges associated with tissue specimen interpretation and ease workflow integration by providing reliable, accurate, and consistent results. Specifically the technology could provide accurate prognostic value as a precursor to histologic or cytologic evaluation of biopsied tissue assisting in pathologic time management. It also has the potential to be utilized intraoperatively providing an oncologist or surgeon with valuable tissue assessment information.

REFERENCES

[1] U.S. Cancer Statistics Working Group. (2016). "United states cancer statistics: 1999-2013 Incidence and mortality Web-based report." Dept. Health Hum. Services, Centers Disease Control Prevention, Nat. Cancer Inst., Atlanta, GA, USA, Tech. Rep., [Online]. Available: http://www.cdc.gov/uscs

[2] W. D. Gregory, J. J. Marx, C. W. Gregory, W. M. Mikkelson, J. A. Tjoe, and J. Shell, "The Cole relaxation frequency as a parameter to identify cancer in breast tissue," *Med. Phys.*, vol. 39, no. 2, pp. 4167–4174, 2012.

[3] K. S. Cole and R. H. Cole, "Dispersion and absorption in dielectrics I. Alternating current characteristics," *J. Chem. Phys.*, vol. 9, no. 4, pp. 341–352, 1941.

[4] K. S. Cole and R. H. Cole, "Dispersion and absorption in dielectrics II. Direct current characteristics," *J. Chem. Phys.*, vol. 10, no. 2, pp. 98–105, 1942.

[5] K. S. Cole and A. L. Hodgkin, "Electric impedance of *Nitella* during activity," *J. Gen. Physiol.*, vol. 22, no. 1, pp. 37–64, 1939.

[6] K. S. Cole and H. J. Curtis, "Electric impedance of the squid giant axon during activity," *J. Gen. Physiol.*, vol. 22, no. 5, pp. 649–670, May 1939.

[7] W. Nernst, "The electromotoric effectiveness of ions," Habilitation thesis, 1889.

[8] W. Yang and W. J. Brackenbury, "Membrane potential and cancer progression," *Front. Physiol.*, vol. 4, p. 185, Jul. 2013.

[9] D. E. Goldman, "Potential, impedance, and rectification in membranes," *J. Gen. Physiol.*, vol. 27, no. 1, pp. 37–60, 1943.

[10] A. L. Hodgkin and B. Katz, "The effect of sodium ions on the electrical activity of the giant axon of the squid," *J. Physiol.*, vol. 108, no. 1, pp. 37–77, 1949.

[11] National Cancer Institute at the National Institutes of Health. (Mar. 25, 2015). *Mammograms*. [Online].Available: http://www.cancer.gov/types/breast/mammograms-fact-sheet

[12] W. S. Cleveland, "Robust locally weighted regression and smoothing scatterplots," *J. Amer. Stat. Assoc.*, vol. 74, no. 368, pp. 829–836, 1979.

[13] W. S. Cleveland and S. J. Devlin, "Locally weighted regression: An approach to regression analysis by local fitting," *J. Amer. Stat. Assoc.*, vol. 83, no. 403, pp. 596–610, 1988.

[14] J. Shell and W. D. Gregory, "A cancer detection device utilizing multi-tiered neural networks for improved classification," in *Proc. 3rd IEEE-NIH Spec. Topics Conf. Healthcare Innov. Point Care Technol.*, Nov. 2016, pp. 92–96.

[15] J. R. Shell, "Robust speaker detection using neural networks," in *Intelligent Engineering Systems through Artificial Neural Networks*, vol. 16, C. H. Dagli, A. L. Buczak, D. L. Enke, M. Embrechts, and O. Ersoy, Eds. New York, NY, USA: American Society of Mechanical Engineers, 2006, pp. 523–526.

[16] H. P. Schwan, "Electrode polarization impedance and measurements in biological materials," *Ann. New York Acad. Sci.*, vol. 148, pp. 191–209, Feb. 1968.

[17] H. P. Schwan, "Linear and nonlinear electrode polarization and biological materials," *Ann. Biomed. Eng.*, vol. 20, no. 3, pp. 269–288, 1992.

[18] H. P. Schwan, *Determination of Biological Impedances Physical Techniques in Biological Research*, vol. 6. New York, NY, USA: Academic, 1963, pp. 323–407.

[19] R. V. Hogg and J. Ledolter, *Engineering Statistics*. New York, NY, USA: MacMillan, 1987.

[20] F. Rosenblatt, "The perceptron—A perceiving and recognizing automaton," Cornell Aeronautical Lab., Buffalo, NY, USA, Tech. Rep. 85-460-1, 1957.

[21] Y. Freund and R. E. Schapire, "Large margin classification using the perceptron algorithm," *Mach. Learn.*, vol. 37, no. 3, pp. 277–296, 1999.

[22] F. Rosenblatt, *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. Washington, DC, USA: Spartan Books, 1961.

[23] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation," in *Parallel Distributed Processing: Explorations in the Microstructure Of Cognition* (Foundations), vol. 1, D. E. Rumelhart and J. L. McClelland, Eds. Cambridge, MA, USA: MIT Press, 1986.

[24] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Math. Control, Signals Syst.*, vol. 2, no. 4, pp. 303–314, 1989.

[25] S. S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 1998.

[26] M. T. Hagan, H. B. Demuth, and M. H. Beale, *Neural Network Design*. Boston, MA, USA: PWS, 1996.

[27] M. T. Hagan and M. B. Menhaj, "Training feedforward networks with the Marquardt algorithm," *IEEE Trans. Neural Netw.*, vol. 5, no. 6, pp. 989–993, Nov. 1994.

[28] M. Riedmiller and H. Braun, "A direct adaptive method for faster backpropagation learning: The RPROP algorithm," in *Proc. IEEE Int. Conf. Neural Netw.*, Mar./Apr. 1993, pp. 586–591.

[29] M. F. Møller, "A scaled conjugate gradient algorithm for fast supervised learning," *Neural Netw.*, vol. 6, no. 4, pp. 525–533, Nov. 1993.

[30] M. J. D. Powell, "Restart procedures for the conjugate gradient method," *Math. Program.*, vol. 12, no. 1, pp. 241–254, 1977.

[31] E. M. L. Beale, "A derivation of conjugate gradients," in *Numerical Methods for Non-linear Optimization*, F. A. Lootsma, Ed. London, U.K.: Academic, 1972.

[32] R. Fletcher and C. M. Reeves, "Function minimization by conjugate gradients," *Comput. J.*, vol. 7, no. 2, pp. 149–154, 1964.

[33] *MATLAB and Statistics Toolbox Release 2013b*, The MathWorks Inc., Natick, MA, USA, 2013.

JOHN SHELL received the Ph.D. degree in electrical and computer engineering from Southern Illinois University Carbondale in 2008. He is currently a Senior Engineer at NovaScan Inc., Milwaukee, WI, USA. His research interests include biomedical engineering, software and analytical analysis, and algorithm development.

WILLIAM D. GREGORY received the B.S. degree in physics from Georgetown University in 1961, and the Ph.D. degree from MIT in 1966. He is a registered Professional Engineer in WV and WI and a licensed Patent Agent in U.S. His commercial experience includes the Westinghouse Corporation Commercial Nuclear Division, Raytheon Research Division, BDM Corp, U.S. Navy Mine Defense Center, and Wandel und Goltermann, GmbH. His academic experience was as a Professor or the Dean at Georgetown University, Clarkson University, Gannon University, West Virginia University, and most recently at the University of Wisconsin-Milwaukee, where he is a Professor Emeritus of health sciences and electrical engineering. He is currently a Board Chair and the Chief Science Officer of NovaScan Inc., Milwaukee, WI, USA. He was awarded an Alexander von Humboldt Senior Scientist Fellowship in 1978 to Universität Tübingen. He is an inventor on over 100 U.S. and foreign patents, has published 150 peer-reviewed papers and co-authored one textbook (*The Science and Technology of Superconductivity*). His current research interests are in the use of electrical property measurements in biology and medicine.