Breast Cancer Classification Using Extracted Parameters from a Terahertz Dielectric Model of Human Breast Tissue

Bao C. Q. Truong[1] *Student Member, IEEE*, H. D. Tuan[1] *Member, IEEE*, Anthony J. Fitzgerald[2], Vincent P. Wallace[2] *Member, IEEE*, Tuan Nghia Nguyen[1] *Member, IEEE*, and H.T. Nguyen[1] *Senior Member, IEEE*

***Abstract*—Our previous study proposed a dielectric model for human breast tissue and provided initial analysis of classification potential of the eight model parameters and their multiparameter combinations with the support vector machine (SVM). A combination of three model parameters could achieve a leave-one-out cross validation accuracy of 93.2%. However, the SVM approach fails to exploit the combinations of more than three model parameters for classification improvement. Thus, the Bayesian neural network (BNN) method is employed to overcome this problem based on its advantages of handling our small data and high complexity of the multiparamter combinations. The BNN successfully classifies the data using the combinations of four model parameters with an accuracy, estimated by leave-one-out cross validation, of 97.3%. Overall performance assessed by leave-one-out and repeated random-subsampling cross validations for all examined combinations is also remarkably improved by BNN. The results indicate the advance of BNN as compared to SVM in utilising the model parameters for detecting tumour from normal breast tissue.**

Index terms. terahertz (THz), dielectric properties, optimization, support vector machine, neural network, classification

I. Introduction

Biomedical applications of terahertz (THz) radiation has been drawing researchers' attention thanks to recent developments of broadband-pulse generation and detection for the frequency regime. Terahertz imaging has been suggested to be capable of identifying contrast between normal and cancerous tissue in skin, breast, and colon [1]–[3]. Distinct properties of THz radiation make this imaging technique possible for clinical implementation. For instance, with low photon energy THz radiation is non-ionizing and non-destructive as well as its applied power levels in typical terahertz imaging systems comply with safety guidelines [4]. Additionally, the extremely high sensitivity of THz waves to water/highly-hydrated materials gives THz imaging an ultimate advantage for medical applications as most biological tissues have high water content [5]. In fact, this property of THz radiation is considered as a major factor contributing to the observed contrast between normal and diseased tissues in THz images. Other sources of contrast could be differences in content of protein, RNA, DNA, and structure changes in tissue [2], [6].

[1]Centre for Health Technologies, University of Technology Sydney, Ultimo 2007, Australia; Email: cao.q.truong@student.uts.edu.au, tuan.hoang@uts.edu.au, hung.nguyen@uts.edu.au.

[2]School of Physics, University of Western Australia, Crawley 6009, Australia; Email: vincent.wallace@uwa.edu.au.

Contrast features found by previous studies in THz images between normal breast tissue and tumour prove the applicability of the terahertz imaging technique to improve the margin detection of breast cancer in breast conserving surgery (BCS) [2], [7]. Higher optical properties including refractive index and absorption coefficient of breast tumour as compared to healthy breast tissue can result in the aforementioned contrast [8]. As optical properties can be represented by complex dielectric constants at the molecular level, a concrete understanding of dielectric properties of breast tissue provides further insight of the contrast mechanism. Therefore, modeling the dielectric function of breast tissue not only explains the physical characteristics underpinning the contrast features in THz images but also possibly introduces some indicator for diagnosis improvement in BCS.

Our previous study proposed a empirical model describing the dielectric properties of human breast tissue [9]. This modeling was based on analysing the experimental dielectric spectra of fat tissue and highly fat-contained tissues (i.e. breast tissue) in the frequency range from 0.1 to 1.8 THz. Both the non-Debye relaxation responses at the frequencies below 1 THz and the Debye relaxation responses at the higher frequencies of these tissues were taken into consideration to develop the model. The model-fitting procedure using the robust gradient sampling algorithm facilitated optimizing the extraction of the model parameters. We also applied the support vector machine method to classify the extracted values for breast cancer detection and obtained the predictive accuracy up to 93.2% (estimated by leave-one-out cross validation) with a combination of three model parameters [9]. However, the initial results were encouraging but could not fully reflect the classification potential of the model parameters in terms of combining the model parameters. In fact, more parameters incorporated into the classification could not improve the classification accuracy. This motivates us to take further steps towards finding the other applicable methods to improve the classification performance of these parameter combinations.

In this paper, we explore the limitation of the support vector machine classifier as well as the applicability of the Bayesian neural networks to improve the classification accuracy using combinations of the model parameters introduced by [9] for breast cancer detection. The Bayesian learning algorithm for neural networks not only enhances the their generalization but also makes the best use of data, thus making it preferable to small data with increasing noisy information and complex-

ity [10]. The parameter combinations are used to train the neural networks for classifying the normal breast tissue and tumour. Classification accuracies are estimated by both leave-one-out cross validation and repeated random-subsampling validation method. The Bayesian neural networks demonstrate a remarkable improvement in applying the model parameters for the breast cancer classification, which is confirmed by the estimated accuracies up to 97.3% with a combination of the four model parameters.

II. Dielectric Model of Human Breast Tissue

A. Applied Model

Since human breast tissue not only has low water content but also possesses inhomogeneous structures of fat cells and proteins, its dielectric spectra in the terahertz range show significant differences from the common spectral response of biological tissues with high water content. Particularly, increases in the real part of the dielectric spectra at frequencies below 1 THz and fairly flat responses over higher frequencies were found in fatty tissues such as adipose and breast tissue [8]. Therefore, the well-known double Debye model, which has been applied for approximating the complex permittivities of the highly hydrated tissues, is not fully capable of dealing with these dielectric responses of breast tissue. To encounter this problem, the multiple Cole-Cole relaxation model was applied for replacing the double Debye model to resolve the low-water-content issue of breast tissue [9]. The authors also considered a non-Debye dielectric relaxation process, which is generalized by the Havriliak-Negami relationship [11], to tackle the dielectric response of breast tissue in the low frequency range. Eventually, [9] proposed the following dielectric model for breast tissue

$$\tilde{\epsilon}(\omega) = \boldsymbol{\epsilon}_\infty + \frac{\omega\boldsymbol{\tau}_1\boldsymbol{\Delta\epsilon}_1 + \boldsymbol{\Delta\epsilon}_2}{1+(\jmath\omega\boldsymbol{\tau}_1)^{\boldsymbol{\alpha}}} + \frac{\boldsymbol{\Delta\epsilon}_3}{1+\jmath\omega\boldsymbol{\tau}_2} + \frac{\boldsymbol{\sigma}}{\jmath\omega}. \quad (1)$$

Here $\boldsymbol{\Delta\epsilon}_1$, $\boldsymbol{\Delta\epsilon}_2$ and the time constant $\boldsymbol{\tau}_1$ describe the dielectric dispersion in the non-Debye relaxation process corresponding to the low frequencies. The dispersive amplitude of the fast relaxation process with the time constant $\boldsymbol{\tau}_2$ is given by $\boldsymbol{\epsilon}_3$. This Debye-like relaxation mode dominates the dielectric response of breast tissue in the higher frequency range. $\boldsymbol{\epsilon}_\infty$ represents the high-frequency limit of the dielectric constant of breast tissue and $\boldsymbol{\sigma}$ reflects the impact of tissue conductivity on dielectric loss.

B. Parameters Extraction

The data used in this study includes the complex permittivities measured with 74 human breast samples, both healthy and cancerous. The samples were taken from the excised specimens of 20 female patients with necessary consents from these patients and the local research committee. All these specimens were preserved in refrigerated and humid environment to maintain their natural moistness. TPIspectra1000 (TeraView Ltd, U.K.), a THz time-domain spectrometer, was used for the measurements which were conducted in transmission mode to collect the transmitted pulses through the samples. Then, the frequency-dependent refractive indices $n(\omega)$ and absorption coefficients $\alpha(\omega)$ of the breast samples were calculated from these time-domain pulses using the method described in [8]. Finally, the measured complex permittivities $\tilde{\epsilon}_m(\omega)$ were easily obtained from these optical properties by the following relationship $\tilde{\epsilon}_m(\omega) = (n(\omega) - \jmath\alpha(\omega))^2$. More specific details of the measurement procedure,experimental equipment, and calculations can be referenced in [8].

To fit the measured data with the applied dielectric model, we minimized the sum of squared error (SSE) between this model and the data. Despite that this optimization problem is highly non-linear and non-convex, it can be effectively solved by the robust gradient sampling algorithm [9], [12]. We applied this method to fit the measured complex permittivities of the breast samples and extracted the respective parameters of the dielectric model (1) as can be seen in Table I.

III. Classification

A. Support Vector Machines

The support vector machine (SVM) has been emerging as one of the most popular learning algorithms for pattern recognition [13]. This method aims to search for a hyperplane separating two data classes in a multi-dimensional space. The optimal hyperplane should create maximal gaps between itself and support vectors falling on two sides of this plane. For this study, we implemented the SVM classification with the toolbox of [14] in the Matlab environment. Based on a number of trial simulations, the kernel function, the Gaussian radial basic function (RBF), was applied due to its best classification performance for the data. However, this kernel function requires adjusting its parameter γ to a suitable value in order to optimize the classification performance. Apart from that, the cost C, which controls the trade-off between the complexity of learning model and training errors, also needs to be selected by users. Therefore, we applied the grid-search to simultaneously find the optimal (C, γ) that can provide the best classification performance.

B. Bayesian Neural Network

A traditional learning method such as SVM and regular neural networks commonly encounters crucial issues of generalization which is defined by how well an obtained prediction model can detect new cases excluding from training data. The generalization loss leads to either underfitting or overfitting the data structure. The problem can be solved by determining appropriate complexity of the prediction model through globally searching its design parameters. This approach is very intensive and requires using a part of data for validation of the parameter search, thus not being able to optimize the use of data source [15].

Bayesian neural networks (BNN) have been seen as a practical and powerful tool to improve the generalization and performance of neural networks since they were introduced by [16]. The Bayesian framework applied in this method allows the learning process to overcome the aforementioned challenge. Particularly, based on the Bayes' theorem a probability distribution of network parameters is obtainable in the Bayesian learning. By that it means that uncertainty and noisy

Table I
THE GROUP-AVERAGE VALUES OF THE MODEL PARAMETERS IN (1) OBTAINED BY FITTING THE 74 BREAST SAMPLES FROM [8].

Group	$\boldsymbol{\epsilon}_\infty$	$\boldsymbol{\Delta\epsilon}_1$	$\boldsymbol{\Delta\epsilon}_2$	$\boldsymbol{\Delta\epsilon}_3$	$\boldsymbol{\sigma}$	$\boldsymbol{\tau}_1(ps)$	$\boldsymbol{\tau}_2(ps)$	$\boldsymbol{\alpha}$
Normal	2.61 ± 0.10	21.75 ± 9.50	-1.84 ± 0.25	0.99 ± 0.11	2.89 ± 0.36	2.84 ± 0.45	0.13 ± 0.01	1.80 ± 0.13
Tumour	3.15 ± 0.07	545.6 ± 500.3	2.82 ± 0.22	1.34 ± 0.07	7.89 ± 0.36	4.67 ± 2.17	0.10 ± 0.01	1.90 ± 0.14

Table II
THE ESTIMATED ACCURACIES (%) BY LOO-CV AND RRS FOR APPLYING THE DOUBLE DEBYE PARAMETERS WITH THE SVM TO CLASSIFY THE HEALTHY BREAST TISSUE AND BREAST TUMOUR.

Parameter Combinations	Kernel Parameters C	Kernel Parameters γ	Leave-One-Out Accuracy(%)	Repeated Random Subsampling Accuracy(%)	Repeated Random Subsampling Sensitivity(%)	Repeated Random Subsampling Specificity(%)
C1. $\boldsymbol{\sigma}$	1	0.25	86.5	85.0 ± 8.1	84.1 ± 12.3	85.8 ± 12.1
C2. $(\boldsymbol{\epsilon}_\infty, \boldsymbol{\sigma})$	512	0.125	91.9	86.6 ± 8.9	85.2 ± 14.3	87.8 ± 12.4
C3. $(\boldsymbol{\sigma}, \boldsymbol{\tau}_1)$	16	0.125	91.9	88.3 ± 8.5	86.0 ± 13.6	90.4 ± 10.9
C4. $(\boldsymbol{\epsilon}_\infty, \boldsymbol{\Delta\epsilon}_1, \boldsymbol{\sigma})$	4	0.0625	87.8	81.0 ± 9.5	74.6 ± 19.0	86.6 ± 13.4
C5. $(\boldsymbol{\epsilon}_\infty, \boldsymbol{\Delta\epsilon}_2, \boldsymbol{\sigma})$	1	0.03125	85.1	85.6 ± 8.5	85.3 ± 12.5	85.9 ± 12.4
C6. $(\boldsymbol{\epsilon}_\infty, \boldsymbol{\Delta\epsilon}_3, \boldsymbol{\sigma})$	256	0.25	93.2	87.6 ± 8.8	85.4 ± 14.4	89.6 ± 11.3
C7. $(\boldsymbol{\epsilon}_\infty, \boldsymbol{\sigma}, \boldsymbol{\tau}_2)$	512	0.125	90.5	86.4 ± 8.6	85.6 ± 13.9	87.1 ± 12.5
C8. $(\boldsymbol{\epsilon}_\infty, \boldsymbol{\sigma}, \boldsymbol{\alpha})$	512	0.125	89.2	84.2 ± 9.6	82.6 ± 15.0	85.6 ± 12.9
C9. $(\boldsymbol{\epsilon}_\infty, \boldsymbol{\Delta\epsilon}_2, \boldsymbol{\Delta\epsilon}_3, \boldsymbol{\sigma})$	1	0.03125	85.1	85.3 ± 8.3	84.2 ± 12.5	86.3 ± 12.1
C10. $(\boldsymbol{\epsilon}_\infty, \boldsymbol{\sigma}, \boldsymbol{\tau}_2, \boldsymbol{\alpha})$	512	0.125	89.2	84.6 ± 9.6	83.1 ± 15.3	86.0 ± 12.8

Table III
THE ESTIMATED ACCURACIES (%) BY LOO-CV AND RRS FOR APPLYING THE DOUBLE DEBYE PARAMETERS WITH THE BNN TO CLASSIFY THE HEALTHY BREAST TISSUE AND BREAST TUMOUR.

Parameter Combinations	Leave-One-Out Accuracy(%)	Repeated Random Subsampling Accuracy(%)	Repeated Random Subsampling Sensitivity(%)	Repeated Random Subsampling Specificity(%)
C1. $\boldsymbol{\sigma}$	86.5	86.5 ± 8.7	86.0 ± 11.4	88.8 ± 9.6
C2. $(\boldsymbol{\epsilon}_\infty, \boldsymbol{\sigma})$	91.9	92.9 ± 6.6	93.2 ± 9.0	94.4 ± 7.8
C3. $(\boldsymbol{\sigma}, \boldsymbol{\tau}_1)$	94.6	88.6 ± 7.4	91.3 ± 9.6	88.5 ± 9.6
C4. $(\boldsymbol{\epsilon}_\infty, \boldsymbol{\Delta\epsilon}_1, \boldsymbol{\sigma})$	94.6	93.1 ± 6.5	93.7 ± 8.7	94.2 ± 7.5
C5. $(\boldsymbol{\epsilon}_\infty, \boldsymbol{\Delta\epsilon}_2, \boldsymbol{\sigma})$	96.0	92.3 ± 6.5	94.0 ± 8.2	92.2 ± 8.3
C6. $(\boldsymbol{\epsilon}_\infty, \boldsymbol{\Delta\epsilon}_3, \boldsymbol{\sigma})$	94.6	93.0 ± 5.8	94.0 ± 8.5	93.9 ± 7.6
C7. $(\boldsymbol{\epsilon}_\infty, \boldsymbol{\sigma}, \boldsymbol{\tau}_2)$	93.2	93.4 ± 6.5	93.2 ± 8.7	95.0 ± 7.4
C8. $(\boldsymbol{\epsilon}_\infty, \boldsymbol{\sigma}, \boldsymbol{\alpha})$	96.0	92.2 ± 6.6	92.8 ± 9.2	93.5 ± 8.1
C9. $(\boldsymbol{\epsilon}_\infty, \boldsymbol{\Delta\epsilon}_2, \boldsymbol{\Delta\epsilon}_3, \boldsymbol{\sigma})$	97.3	92.4 ± 6.4	93.3 ± 9.0	93.4 ± 7.7
C10. $(\boldsymbol{\epsilon}_\infty, \boldsymbol{\sigma}, \boldsymbol{\tau}_2, \boldsymbol{\alpha})$	97.3	93.6 ± 6.4	94.3 ± 8.3	94.4 ± 7.6

information of data can be taken into consideration to improve the prediction performance. In addition, the learning process using the Bayesian regulation facilitates automatic adjustment of network hyper-parameters, which are regulation constants controlling the complexity of the prediction model, to the most appropriate values. This allows the elimination of using the validation set of data, thus maximizing data resource for training. As a result, BNN is of great interest to handling our small data. Besides, by viewing our multiparameter problem in this paper from the advantages of BNN, we can find it a probable solution to dealing with the increasing complexity of the prediction model when more model parameters are incorporated into the classification. Indeed, this complexity issue is directly concerned with adjusting more regulation constants, which is considered as an important advantage of the Bayesian approach [16].

C. Accuracy Estimation

Both leave-one-out cross validation (LOOCV) and repeated random-subsampling validation (RRS) are applied to validate the classification accuracy. LOOCV has been among the most popular methods to estimate the accuracy of a classifier [17]. With consecutively holding out only one point and using the whole remaining set of the data for training, this cross validation provides an unbiased and high-variance estimation of accuracy. Thus, the accuracy prediction with LOOCV could be too optimistic. The RRS with a significant proportion of data left out for testing does not make the best use of data for training but offers a better balance between bias and variance when estimating the classification accuracy. Combing the two validation methods is necessary for more accurately justifying the classification performance. For this study, we chose to use 80% of the data (59 samples) for training and 20% (15 samples) for testing in the RRS with 1000 repetitions of the training-testing process for each classifier.

IV. RESULTS AND DISCUSSION

In this study, ten classification combinations of the model parameters are chosen to investigate and annotated by C1 to C10 respectively as can be seen in both Table II and III. In fact, based on the statistical analysis in [9], we could form a variety of potential combinations for the classification. However, by analysing the LOOCV and RRS accuracy simultaneously with either the SVM or BNN a number of the combinations were

filtered, and hence, only the best ten combinations were selected to present. They are not only able to achieve the highest classification accuracies but also optimal in terms of low dimension and complexity.

Table II shows the LOO-CVs and the average classification accuracies with their standard deviations estimated by RRS for using SVM classifiers with combinations of the model parameters in (1). The optimal kernel parameters including C and γ for each set of the model parameters are accountable for the best LOOCV of the combination. Under the impact of the smaller training set, C6 obtains the highest LOOCV (93.2%) but a far lower RRS accuracy (87.6%). Despite that C1 only contains one model parameter $\boldsymbol{\sigma}$, its classification performance is still better than the higher-dimension combinations such as C5 and C9. In fact, $\boldsymbol{\sigma}$ has been considered as the most potential parameter of the model (1) for breast cancer classification [9]. C3 should be the most suitable for the SVM method thanks to its high and stable accuracies predicted by LOOCV (91.9%) and RRS (88.3%). However, the combinations with more model parameters such as C4-C10 do not improve or even weaken the classification performance using SVM. As mentioned earlier, this remains the challenge of applying the SVM approach for the data, which motivates further investigation into the applicability of BNN.

According to Table III, the problem of SVM indeed can be overcome by BNN structured by 10 hidden nodes. To be more specific, the overall classification performance of the combinations is improved whenever an extra model parameter is added to the classification. Accordingly, the best accuracy obtained with the four-parameter combinations including C9 and C10 is 97.3% in LOOCV and 93.6% in RRS. The highest LOOCVs of the three- and two-parameter combinations including C2-C8 are 96.0% and 94.6% respectively. However, although the average accuracies of C2-C10 in RRS achieve very high values from 92.2%, the impact of increasing the number of input parameters on the classification is not significant. The similarity in C1,C2,C4 between LOOCVs and RRS accuracies suggests BNN can learn the data structure of these combinations very well regardless of the smaller data set for training in RRS. Conversely, the impressive LOOCVs of the rest are unachievable in RRS due to the shortage of training data. By and large, classifying the breast tumour using the combinations of the model parameters in (1) with BNN offers better overall performance than that with SVM.

V. Conclusion

The SVM method is limited in terms of efficiently learning the data structure of the combinations of the model parameters in (1) for classification. Therefore, we revised the problem and successfully applied the BNN classifier to improve performance of the combinations. Particularly, ten parameter combinations C1-C10 were introduced for investigation with both SVM and BNN. Using the BNN, the best LOOCV is enhanced to 97.3% with the four-parameter combinations as compared to 93.2% with the SVM. The advance of BNN in classifying the data is also expressed over the estimated accuracies in RRS. The average accuracies vary between about $92.17 - 93.57\%$ corresponding the different combinations, which are also by far higher than 88.3% (RRS) with the SVM. Apart from that, the classification accuracies predicted by cross validation methods in this study may be statistically insufficient for making a confirmation of the true classification accuracies in practice due to the used small data. However, our encouraging results should be basic to future studies which examines larger data such as THz images of breast tumour in order to improve the cancer-margin detection in BCS. Further developments of this study will also include selection of the best parameter combinations and improvement of classification methodologies.

References

[1] V. Wallace, A. Fitzgerald, S. Shankar, N. Flanagan, R. Pye, J. Cluff, and D. Arnone, "Terahertz pulsed imaging of basal cell carcinoma ex vivo and in vivo," *British Journal of Dermatology*, vol. 151, no. 2, pp. 424–432, 2004.

[2] A. J. Fitzgerald, V. P. Wallace, M. Jimenez-Linan, L. Bobrow, R. J. Pye, A. D. Purushotham, and D. D. Arnone, "Terahertz pulsed imaging of human breast tumors," *Radiology*, vol. 239, no. 2, pp. 533–540, 2006.

[3] C. B. Reid, A. Fitzgerald, G. Reese, R. Goldin, P. Tekkis, P. S. OKelly, E. Pickwell-MacPherson, A. P. Gibson, and V. P. Wallace, "Terahertz pulsed imaging of freshly excised human colonic tissues," *Physics in Medicine and Biology*, vol. 56, no. 14, p. 4333, 2011.

[4] E. Berry, G. C. Walker, A. J. Fitzgerald, N. N. Zinovev, M. Chamberlain, S. W. Smye, R. E. Miles, and M. A. Smith, "Do in vivo terahertz imaging systems comply with safety guidelines?," *Journal of Laser Applications*, vol. 15, no. 3, pp. 192–198, 2003.

[5] V. P. Wallace, P. F. Taday, A. J. Fitzgerald, R. M. Woodward, J. Cluff, R. J. Pye, and D. D. Arnone, "Terahertz pulsed imaging and spectroscopy for biomedical and pharmaceutical applications," *Faraday Discussions*, vol. 126, pp. 255–263, 2004.

[6] S. Sy, S. Huang, Y.-X. J. Wang, J. Yu, A. T. Ahuja, Y. ting Zhang, and E. Pickwell-MacPherson, "Terahertz spectroscopy of liver cirrhosis: investigating the origin of contrast," *Physics in Medicine and Biology*, vol. 55, no. 24, p. 7587, 2010.

[7] A. J. Fitzgerald, S. Pinder, A. D. Purushotham, P. OKelly, P. C. Ashworth, and V. P. Wallace, "Classification of terahertz-pulsed imaging data from excised breast tissue," *Journal of Biomedical Optics*, vol. 17, no. 1, pp. 016005–1–016005–10, 2012.

[8] P. C. Ashworth, E. Pickwell-MacPherson, E. Provenzano, S. E. Pinder, A. D. Purushotham, M. Pepper, and V. P. Wallace, "Terahertz pulsed spectroscopy of freshly excised human breast cancer," *Opt. Express*, vol. 17, pp. 12444–12454, Jul 2009.

[9] B. C. Q. Truong, H. D. Tuan, A. J. Fitzgerald, V. P. Wallace, and H. T. Nguyen, "A dielectric model of human breast tissue in terahertz regime," *IEEE Transactions on Biomedical Engineering*, vol. 62, pp. 699–707, Feb 2015.

[10] C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford: Clarendon Press, 1995.

[11] S. Havriliak and S. Negami, "A complex plane representation of dielectric and mechanical relaxation processes in some polymers," *Polymer*, vol. 8, no. 0, pp. 161 – 210, 1967.

[12] J. Burke, A. Lewis, and M. Overton, "A robust gradient sampling algorithm for nonsmooth, nonconvex optimization," *SIAM Journal on Optimization*, vol. 15, no. 3, pp. 751–779, 2005.

[13] C. J. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, pp. 121–167, Jun 1998.

[14] C. C. Chang and C. J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

[15] M. T. Hagan, H. B. Demuth, M. H. Beale, and O. D. Jes, *Neural Network Design*. Martin Hagan, 2014.

[16] D. J. MacKay, "Bayesian neural networks and density networks," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 354, no. 1, pp. 73 – 80, 1995.

[17] L. D. Fisher and G. van Belle, *Biostatistics: A Methodology for the Health Sciences*. John Wiley & Sons, 1993.