Prediction of genetically-evaluated tumour responses to chemotherapy from breast MRI using machine learning with model selection

Taiguang Yuan1, Ze Jin1,2, Yukiko Tokuda3, Yasuto Naoi4, Noriyuki Tomiyama3, Takashi Obi1,2 and Kenji Suzuki1,2,*

1Department of Information and Communications Engineering, School of Engineering, Tokyo Institute of Technology, Japan
2Laboratory for Future Interdisciplinary Research of Science and Technology, Institute of Innovative Research, Tokyo Institute of Technology, Japan
3Department of Radiology, School of Medicine, Osaka University, Japan
4Department of Breast and Endocrine Surgery, School of Medicine, Osaka University, Japan

*E-mail: suzuki.k.d1@m.titech.ac.jp

Abstract. Genetic tests can provide prognostic information in breast cancer for both diagnosis and treatment planning. However, the cost of a genetic test is still high. In this study, we developed a radiogenomics method to predict genetically-evaluated responses to chemotherapy for breast cancer using our machine-learning technology coupled with model selection. Our proposed method consists of feature extraction, model selection, and prediction by the selected model. In the feature extraction, 318 morphological and texture features were extracted from a tumour region. In the model selection module, there are two major components: (1) selection of imaging biomarkers based on our original sequential forward floating selection (SFFS) feature selection and (2) building of a support vector machine (SVM) classifier including kernel function selection and hyperparameter optimization. The optimized feature set, i.e. imaging biomarkers, coupled with an SVM classifier were chosen by maximizing the area under curve (AUC) of corresponding receiver-operating-characteristic curve (ROC). After the model selection, the optimized SVM classifier operated on the selected imaging biomarkers for prediction. We applied our proposed method to 118 breast MRI studies from 118 patients for predicting genetically-evaluated responses to chemotherapy for breast cancer that evaluated by the genetic test of IRSN-23. We achieved an AUC value of 0.96 using the optimized SVM classifier model coupled with 24 selected imaging biomarkers in predicting the results of IRSN-23 in a five-fold cross-validation procedure.

1. Introduction
Breast cancer is a complex disease caused by multiple factors, a major one is the progressive accumulation of gene mutations, combined with epigenetic dysregulation of critical genes and protein pathways [1]. Recent studies have shown that genetic test has a capability for providing prognostic information on breast cancer. For example, an immune-related 23-gene signature (IRSN-23) which has been recently developed by Naoi, Y., et al., was able to predict the pathological response to
chemotherapy for breast cancer [2]. However, the cost of a genetic test is still high; for example, $3,400 for the first genetic test.

Radiogenomics is an emerging field that employs a machine-learning technique for the selection of imaging biomarkers (effective feature) from medical images to link non-invasive imaging biomarkers with cancer genotypes [3]. Recent radiogenomics studies have demonstrated that magnetic resonance imaging (MRI) features of breast tumours potentially has the capability for predicting molecular classifications of breast cancer subtypes [4] and for predicting the genetically-evaluated recurrence score of breast cancer [5]. Therefore, we hypothesized that MRI features could be potentially helpful in predicting genetically-evaluated responses to chemotherapy for breast cancer.

The purpose of this study was to develop a radiogenomics method to predict genetically-evaluated responses to chemotherapy for breast cancer using machine learning technology with model selection based on our original sequential forward floating selection (SFFS) method coupled with a support vector machine (SVM) classifier that would maximize the area under curve (AUC) of corresponding receiver-operating-characteristic curve (ROC) [6].

2. Materials
Our database contains 118 dynamic-contrast-enhanced (DCE) breast MRI studies from 118 patients, acquired at the Osaka University Hospital. Each DCE breast MRI study consists of three phases (i.e., pre-contrast, early phase, and delayed phase). Each MR image has a matrix size of 256 x 256 pixels with pixel sizes of 0.7-0.9 mm in the sagittal plane, and 1.0-1.5 mm of spacing between slices. “Gold-standard” breast tumour regions were manually determined by an experienced breast radiologist. Figure 1 shows images within a dataset of DCE breast MRI: (a) pre-contrast, (b) early-phase, (c) delayed-phase and (d) manual segmentation of a tumour region. All images were showing the same slice with the largest in-plane diameter of tumour region.

The genetic tests of IRSN-23 were performed on 118 patients, which classified the patients into 45 patients as genomically predicted responders (Gp-R) who are likely to achieve pathological complete response (pCR) and 73 patients as genomically predicted non-responders (Gp-NR) who are unlikely to achieve pCR.

![Figure 1](image)

2. Figure 1. Example images of a dataset of dynamic contrast-enhanced breast MRI: (a) Pre-contrast, (b) Early-phase, (c) Delayed-phase and (d) Manual segmentation of a tumour region.

3. Methods
Our proposed method consists of feature extraction, model selection and prediction by the selected model. Figure 2 illustrates the block diagram of our proposed method. In the feature extraction module, 318 image features including morphological, grey-level-based, and texture features, were extracted from each segmented tumour region on DCE breast MRI in three phases (i.e., pre-contrast, early phase, and delayed phase). In the model selection module, there are two major components: (1) selection of imaging biomarkers based on our original sequential forward floating selection (SFFS) feature selection and (2) building of an SVM classifier including kernel function selection and hyperparameter optimization. An optimized feature set, i.e. imaging biomarkers, coupled with an SVM classifier were chosen by maximizing the AUC. After the model selection, the optimized SVM classifier coupled with the selected imaging biomarkers was applied to 118 DCE breast MRI studies for prediction of results of IRSN-23 as a Gp-R or a Gp-NR.
3.1. Feature extraction

Feature extraction is one of the most important steps in our prediction system. Given the manually segmented breast tumour regions, 106 two-dimensional (2D) and three-dimensional (3D) features including 42 morphological and grey-level-based features, 36 margin and peak-related features [7], and 28 texture features [8] were extracted from each segmented tumour region per phase on DCE breast MRI to form an initial feature set of 318 features in total. For each tumour, 2D features were computed in a slice with the largest area of a segmented tumour region; and 3D features were extracted in the overall segmented volume. Shape information, such as radial and tangential gradient indices were calculated as morphological features. Grey-level information characterized tumour intensity information. These features aimed at serving as the initial feature set for the selection of imaging biomarkers for the prediction of genetic test results. All feature values were normalized to z-scores.

3.2. Model selection

3.2.1. Imaging biomarkers selection

Not all of the extracted features would be useful for the prediction. To select the imaging biomarkers, we used our original maximal AUC SFFS feature selection method coupled with an SVM classifier to remove ineffective features.

Table 1 shows the main procedures of our SFFS feature selection. The selected feature set starts with an empty set $F_0$. Then in the optimization procedure of the feature subset, add one feature $\{x\}$ at a time while monitoring the change of AUC, which is calculated by an SVM classifier, until a maximal AUC value was achieved. This is given in Eq. (1) where the criterion $J(F_k + \{x\})$ is the AUC value of the SVM classification with the selected feature set $(F_k + \{x\})$. Thus, Eq. (1) guarantees that the selected feature set would achieve the maximal AUC value with the combination of the existing features in the subset. This step, however, is only considering adding features without removing any existing ones. It might be possible to improve the performance (the AUC value) by removing some features from the selected subset. A removing procedure was implemented by Eq. (2). One feature was removed from the selected feature set at a time if the remaining feature subset performs better than the previous maximal AUC value. The procedure continued until all features in the subset were tested. The feature subset with the maximal AUC value would be selected as the final output of the procedure.
Table 1. Procedure of Max-AUC SFFS feature selection.

| Initialization: |
|------------------|
| Full feature set from MR images \( \{X\} \), selected feature set at step 0 \( F_{0\omega}=\emptyset \), predefined feature number \( l=318 \), \( k=0 \). |

\[
\text{while } k<l \\
\quad x^+ = \arg \max_{x \in k-1} f_k + \{x\} \\
\quad F_{k+1} = F_k + \{x^+\} \\
\text{if } k>2 \\
\quad x^- = \arg \max_{x \in k-1} f_k - \{x\} \\
\text{while } J(F_k-\{x\}) > \max J(F_k) \text{ and } k>2 \\
\quad F_{k-1} = F_k - \{x^-\} \\
\text{if } k>2 \\
\quad x^- = \arg \max_{x \in k-1} f_k - \{x\} \\
\text{end} \quad \text{end} \quad \text{end} \quad \text{end}
\]

Output: Selected feature set \( Y \).

3.2.2. Optimization of the SVM classifier

We used an SVM classifier as the classifier in our prediction system for its excellent performance and robustness to outliers. However, hyperparameters selection, e.g., an optimal kernel function with suitable parameters, is important to the robust classification performance of an SVM classifier. The optimal hyperparameter is defined as the one that can maximize the between-class separability and minimize the within-class separability.

For selecting the optimal hyperparameters, we applied our SFFS feature selection method to choose an optimal set of feature vectors for each hyperparameter of the SVM classifier. Given a set of \( N \) training samples \( \{(x_i,y_i)\}_{i=1}^N \) where \( x_i \) is the feature vectors of initial feature set (318 features in total) with \( x_i \in \Re^d \) and \( y_i \) represents the genetic test results (e.g., Gp-R or Gp-NR) with a binary range of \( y_i \in \{-1, 1\} \). The decision function for the SVM classifier is defined as:

\[
f(x) = \sum_{i=1}^N \alpha_i y_i K(x_i,x) + \alpha_0
\]

where \( \alpha_i \geq 0 \) are Lagrange multipliers optimized through quadratic programming. \( K(x_i,x) \) is a symmetric non-negative inner-product kernel, where \( (x_i,x_j) \) is the feature vectors of the feature subset in our SFFS feature selection procedure. In this paper, we only focused on the linear kernel function and the radial basis function (RBF) kernel function.

Finally, the hyperparameters with the selected optimal feature sets that reached the maximum AUC value was selected for the prediction system.

4. Results

4.1. Model selection

4.1.1. Optimization of the SVM classifier

A five-fold cross-validation procedure was used to select the optimal hyperparameters of the SVM classifier including the optimal kernel function and other hyperparameters. In this study, we
experimented with the linear kernel function and the RBF kernel function with four different gamma values. For each hyperparameter, we applied our SFFS feature selection method to choose an optimal set of feature vectors. The AUC values were obtained for the SVM classifier with the optimal feature set in a five-fold cross-validation procedure. Table 2 presents the AUC values indicating the performance of the SVM classifier with different kernels and hyperparameters for our breast MRI database. Finally, the RBF kernel function with a gamma value of 0.5 was selected for the following experiments which reached the maximum AUC value.

Table 2. AUC values for SVMs with different kernels and hyperparameters for predicting results of IRSN23.

| Model Parameters | Linear | RBF | Γ=0.5 | Γ=1 | Γ=2 | Γ=4 |
|------------------|--------|-----|-------|-----|-----|-----|
| AUC values       | 0.73   | 0.93| 0.89  | 0.87| 0.85|

4.1.2. Selection of the imaging biomarkers

We applied our maximal AUC SFFS feature selection coupled with the optimized SVM classifier for predicting the genetic test results of IRSN-23 in a five-fold cross-validation procedure. With the optimized SVM classifier and feature selection, 13 imaging biomarkers, e.g., Texture contrast, Histogram contrast, Tangential gradient index, Inverse difference moment, etc. have been selected with the maximum AUC value.

4.2. Performance evaluation of the selected model

We applied the selected model for predicting the genetic test results of IRSN-23 in a five-fold cross-validation procedure. To evaluate the performance of the selected model, a binormal model was used to estimate the AUC value from the outputs of the optimized SVM classifier [9]. Figure 3 plots the ROC curve for the selected model. The selected model achieved an AUC value of 0.93 in predicting the genetic test results of IRSN-23.

![Figure 3. The binormal ROC curve of the selected model for predicting the results of IRSN-23.](image)

5. Conclusions

In this study, we developed a radiogenomics method to predict genetically-evaluated responses to chemotherapy for breast cancer using our machine-learning technology coupled with model selection. The proposed method discovered the optimal model with the most effective features as imaging biomarkers that maximized the AUC value of the ROC curve. Our results indicated that imaging biomarkers would be alternatives to genetic tests in the evaluation of responses to chemotherapy for
breast cancer for treatment planning. Note that genetic tests are still very useful in the diagnosis and prediction of risks in the development of breast cancer.

Acknowledgements
This work was supported by JST-Mirai Program Grant. The authors are grateful to the members of Suzuki Lab for their valuable discussions.

References
[1] O. I. Olopade, T. A. Grushko, R. Nanda, and D. Huo, Clinical Cancer Research, 14, 7988–99 (2008)
[2] Sota, Y., Y. Naoi, R. Tsunashima, N. Kagara, K. Shimazu, N. Maruyama, A. Shimomura, et al., Annals of Oncology, 25, 100-06 (2014)
[3] J. P. O’Connor, E. O. Aboagye, J. E. Adams, et al., Nature Reviews Clinical Oncology 14, 169-86 (2017)
[4] Elizabeth J. Sutton, et al., Journal of Magnetic Resonance Imaging, 44, 122-29 (2016).
[5] Hui Li, et al., Radiology, 281, 382-391 (2016).
[6] J.-W. Xu and K. Suzuki, IEEE Journal of Biomedical and Health Informatics, 18, 585–93 (2014)
[7] G. M. T. Brake, N. Karssemeijer, and J. H. C. L. Hendriks, Physics in Medicine and Biology, 45, 2843-57 (2000)
[8] R. M. Haralick, K. Shanmugam, I. Dinstein, IEEE Transactions on Systems, Man, and Cybernetics, SMC-3, 610–21 (1973)
[9] E. Metz, B. A. Herman, J.-H. Shen, Statistics in Medicine, 17, 1033–53 (1998)