Support Vector Machines to Aid Breast Cancer Diagnosis Using a Microwave Radar Prototype

Abstract

In this study, we present a method to improve breast cancer diagnosis using a microwave radar prototype. The system is used to collect electromagnetic signals from breast phantoms and classify them using a Support Vector Machine (SVM). The proposed work is an extension of our previous work and aims to improve the accuracy of breast cancer diagnosis.

1. Introduction

Previous studies have addressed the use of classifiers to assist breast cancer diagnosis. Several algorithms have been used for this purpose, including Support Vector Machines (SVM), Quadratic Discriminant Analysis (QDA), and Linear Discriminant Analysis (LDA). In this study, we extend our previous work by using a different classifier, namely SVM.

2. Materials

2.1 Microwave Imaging System Description

The microwave imaging system used in this study is a prototype developed at the University of Bristol. It consists of a multiport VNA and a conformal array of 60 wide-slot antennas. The system operates in the 3-8 GHz frequency range and has a loss matching ceramic shell.

2.2 Phantoms

A total of 36 breast phantoms were tested in this system. Each phantom contains different tumour profiles and is imaged as a whole. The model considered in this study contains a thickness varying between 1 and 3 mm.

3. Methodology

In this study, we have used a SVM classifier to classify the electromagnetic signals recorded with a microwave radar prototype. The classifier will give an indication on whether an area in the breast is healthy or has any tumours.

4. Results

The proposed work has been evaluated using synthetic breast phantoms. 779 of these phantoms were completed with a chest wall TMM, and 1,925,455 measurements of "hits" and "misses" on 36 SFPs were recorded. The model correlates to our "probabilistic map" so "misses" = 1 - "hit". Hence creating a "probabilistic map" will help avoid false positives and detect tumours that were not imaged as false negatives.

Conclusions

The proposed work has shown promising results in improving breast cancer diagnosis using a microwave radar prototype. Further research is needed to validate the findings and improve the accuracy of the system.

References


data of each focal point and a SVM classifier was used to distinguish “hits” and “misses”. The dataset was then divided into training and testing sets using cross-validation. A k-fold cross-validation with k=10 was implemented and the calculated metrics correspond to the sum of the classification results for each fold. The classification performance was evaluated by metrics such as accuracy, sensitivity, specificity, positive predictive value and negative predictive value.

SVM classifier maps the input vectors (i.e. the 24 features for each focal point) into high-dimensional feature spaces according to the chosen kernel, creating a hyperplane that separates the data into classes. The kernel considered in this study was the Radial Basis Function (RBF) which has a scaling factor \( \gamma \) which needs to be optimised for each dataset. There is also another important parameter of SVM, usually represented with \( C \), which is a penalty parameter of the error term. Both \( \gamma \) and \( C \) need to be optimised to guarantee good classification results.

We performed a parameter grid-search [8] to optimise the parameters \( C \) and \( \gamma \) for the present database. We compared the performance for \( \gamma = 4 \) and \( C = 2 \).

For each focal point, the SVM classifier returns a classification value which we normalized from 0 to 1. We have both binary and regression results. For binary results, we apply a threshold of 0.5, and so any classification value below 0.5 becomes “0” (which represents “miss” or a healthy point), or else it becomes “1” (which represents “hit” or a tumour point). The regression results have all values ranging between 0 and 1.

4. Results

Figure 1. An example of a “hit” and a “miss” synthetic focal point in the synthetic breast model. Orange and yellow pixels represent true “miss” and “hit” SFPs, respectively.

Figure 2. Reconstructed profile of the synthetic breast model using the SVM classifier in the binary mode. Orange and yellow pixels represent the classified “miss” and “hit” SFPs, respectively.

Figure 3. Reconstructed profile of the synthetic breast model using the SVM classifier in the regression mode. Orange and yellow pixels represent the classified “miss” and “hit” SFPs, respectively.

5. Conclusions

In general, the approach with SVM classifier presented in this paper outperforms the regression classification results with LDA, presented in [6], and will be considered in future related work.

6. Acknowledgements

This paper is supported by Fundação para a Ciência e a Tecnologia – FCT / MEC (PIDDAC) under the Strategic Programme UID/BIO/00645/2013. This paper has been developed in the framework of COST Action TD1301, MiMed.

7. References


