Multi-source automatic breast cancer diagnosis

Inês Domingues
INESC Porto
FEUP, Portugal
inesdomingues@gmail.com

Jaime S. Cardoso
INESC Porto
FEUP, Portugal
jaime.cardoso@inescporto.pt

Maria J. Cardoso
FMUP, Portugal
INESC Porto
mjcard@med.up.pt

Abstract

Breast cancer remains a major health problem, with x-ray mammography being the commonest cancer detection technique. However, in several scenarios, including high-risk population, radiologists turn to multiple exams, like ultrasonography and resonance imaging to have a more reliable assessment. In this paper, a research work under-development is described and some preliminary results are shown.

1. Introduction

According to the American Cancer Society, approximately forty thousand women are expected to be diagnosed with breast cancer in 2009. Since the cause of the disease remains unknown, early detection and diagnosis is the key for breast cancer control, and it can increase the success of treatment, save lives and reduce cost. X-ray mammography screenings are the most common breast cancer detection method. However, x-ray has well-recognized limitations, and recently, other imaging including ultrasound (US) and magnetic resonance imaging (MRI) have been used as adjunctive screening tools. In order to eliminate the operator dependency and improve the diagnostic accuracy, computer-aided diagnosis (CAD) systems are a valuable and beneficial means for breast cancer detection and classification. The present document describes a research work under-development for the automatic interpretation of screening images in the breast cancer field.

2. State-of-the-art

Breast image registration is a difficult task due to the non-rigid propriety of the breast, breast motion, and breast compression during some image acquisition techniques. Registration methods can be divided into intra and inter modality. While the first ones deal with problems like cancer evolution in a patient, the second deal with the fusion of information from different types of images.

The intramodality may be further classified accordingly to the technique used. Feature based registration techniques aim at finding corresponding points in pairs of images. These points can be manually or automatically marked. It has been stated that intensity registration by mutual information outperforms other methods. The intermodality methods focus on the combination of the advantages of different acquisition methods. X-ray is the most used screening breast cancer detection method. Moreover, in some cases, other methods as US and MRI are also used in order to have a reliable diagnostic. MRI is usually acquired in women with BRCA gene mutations, women with family history of breast cancer, and women with personal history of other types of cancer. US, as it does not involve ionizing radiation, may be the only viable modality in pregnant and lactating women. Its main disadvantage is being user dependent, making x-ray mammography a more suitable imaging method for automated analysis.

In spite of the vast amount of work in automated breast image processing, there is no satisfactory way to register the images from different modalities. Moreover, even for x-ray mammograms, the research community has mostly focused in the detection of lesions in the image, rather than in the automation of the complete decision in the normalized BI-RADS scale. In the other modalities, the research on automated processing is even more insufficient.

3. Current Work

We are working on new techniques to combine information from different types of sources, namely x-ray, US and MRI, for the detection of breast cancer, creating a multi-modality breast cancer detection system. Image processing and automated learning techniques are being applied to the problem of image registration from different modalities and detection and diagnosis of suspicious lesions in the multiple views of the breast. The extraction of relational structure in the image is receiving special attention since, by capturing

not only the statistical description of the objects in the image but also the relations between them, the identification is facilitated.

We intend to achieve results that exceed the current state-of-the-art in multi-modality breast cancer detection systems. The correct fusion of information from x-ray, US and MRI for an improved CAD system of the breast is the primary goal of the project.

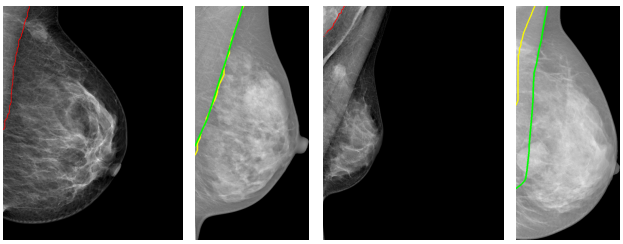
The work plan can be divided into a theoretical and an implementation part. Firstly, the medical knowledge will be used together with a mathematical analysis of the problem in order to achieve a deep understanding on both breast cancer and the current image processing and analysis methodologies. The second part consists on the implementation and performance evaluation of the algorithms.

The following constitute the main scientific and technical objectives of the project:
- to create a multi-modal database of breast images, manually annotated and stored with the corresponding medical reports. This database is to be made publicly available for the benefit of the research community;
- to study and acquire a deep understanding of the image processing and automatic learning techniques in use in the breast image field, and their critical assessment;
- to become proficient in image registration and fusion techniques and their application in the image field;
- to conduct research on new representations and formulations for image understanding in the breast cancer field;
- to conduct research on machine learning algorithms for structured representations;
- to develop tools that implement the aforementioned methodologies, at least at a prototype phase;
- to integrate the new techniques in medical decision support systems currently being developed in the host institution.

For the image registration, we are considering two classes of fusion schemes, namely early fusion and late fusion. The former fuses modalities in feature space, the latter fuses modalities in semantic space. However, since the data originates from different physical devices taken under very different conditions (due to breast motion, the non-rigidity of the breast and the breast compression with some of the acquisition techniques), we anticipate the advantage of late fusion techniques, where some form of analysis and object identification is important for a proper image registration.

4. Pectoral muscle detection

Some of the registration methods try to find a function that maps correspondent landmarks on the images. These landmarks need to be previously determined, in a manual or automatic way. Having this in mind, our first work consisted in the automatic detection of a structure that can be used for registration. The selected structure was the pectoral muscle, and the work developed was lead to two publications [1, 2]. Fig. 1 shows some results. Both methods use the short path algorithm. But, while in [1] the short path is calculated on a polar coordinate transformed image, in [2] the pectoral muscle contour endpoints were estimated by Support Vector Regression. The polar coordinates algorithm reached a normalized average hausdorff distance of 0.14, and the regression algorithm of 0.12 for the database collected at Hospital S. João, Porto, Portugal.



(a) img 05 (b) img 18 (c) img 39 (d) img 26

Figure 1. Selected successful (two figures on the left) and poor (two figures on the right) results. (a) and (c) were obtained with the polar coordinates algorithm and (b) and (d) with the regression algorithm.

5. Conclusion

The impact of the research conducted in the described project is reflected in its ability to improve the quality of breast cancer detection, speeding up the time to output a diagnostic with the correspondent beneficial implications in treatment possibilities and psychological patients well being. The deployment of a multi-modality annotated database allows not only other researchers to study the same problem but also the comparison of results. A second contribution is in the field of learning in computer vision, namely in problems pertaining to the semantic understanding of visual data. The project will allow researchers to move ahead in the application of computational intelligence ideas to real-life problems.

References

[1] J. S. Cardoso, I. Domingues, I. Amaral, I. Moreira, P. Passarinho, J. S. Comba, R. Correia, and M. J. Cardoso. Pectoral muscle detection in mammograms based on polar coordinates and the shortest path. In *EMBC*, 2010.

[2] I. Domingues, J. S. Cardoso, I. Amaral, I. Moreira, P. Passarinho, J. S. Comba, R. Correia, and M. J. Cardoso. Pectoral muscle detection in mammograms based on the shortest path with endpoints learnt by SVMs. In *EMBC*, 2010.