Chapter 31
Image Based Classification Platform:
Application to Breast Cancer Diagnosis

Paulo J. S. Gonçalves
Polytechnic Institute of Castelo Branco, Portugal & Technical University of Lisbon, Portugal

Rui J. Almeida
Erasmus University Rotterdam, The Netherlands

João R. Caldas Pinto
Technical University of Lisbon, Portugal

Susana M. Vieira
Technical University of Lisbon, Portugal

João M. C. Sousa
Technical University of Lisbon, Portugal

ABSTRACT

The high number of exams that is done in healthcare institutions increases the medical doctors' workload, leading to poor working conditions and the increase of wrong diagnoses. As consequence, an automatic system that can help medical doctors in diagnostic tasks is of major interest to any healthcare institution. The chapter proposes an Image Based Classification Platform suitable to help Medical Doctors diagnosing breast cancer, based on mammograms, i.e., to detect if a tumor is present in the image. The platform is twofold, i.e., in the first part the image descriptors are extracted from the image using image-processing algorithms. The obtained descriptors are used in the second part. The second part is related to classification, where computational intelligence methods are used to classify a given image, based on the descriptors obtained in the first phase. Texture analysis based on co-occurrence matrices are applied to obtain the descriptors from the MIAS database of mammograms. From these descriptors, fuzzy models, neural networks, and support vector machines are successfully used to classify the mammograms and obtain a diagnosis.

DOI: 10.4018/978-1-4666-3990-4.ch031

1. INTRODUCTION

This chapter presents a software tool to help medical doctors in diagnostic tasks, using medical images. This software tool is divided in two complementary parts, which can work together or separately; the first is dedicated to image processing and the second to classification tasks. In this chapter, both parts work together to perform classification tasks based on information obtained from images.

The first part of the system gives the image descriptors of the captured image. The algorithms implemented in the software tool are the wavelet decomposition for texture and color (Alexandre & Pinto, 2006), gray-level co-occurrence matrix (Haralick, 1973), quadtree decomposition (Ramalho, Caldas Pinto and Marcolino, 2000) and Markov random fields (Yu & Cheng, 2003). These algorithms give a large number of features that are used in the second part of the system. The descriptors to be used can be chosen by the user. The second part of the system uses the image descriptors obtained in the image processing part of the system. The computational intelligence systems implemented in the platform are based in Fuzzy Models (Takagi & Sugeno, 1985; Zadeh, 1973; Mamdani, 1977), Neural Networks (Zhang, 2000) and Support Vector Machines (Vapnick, 1998). The software tool described in this chapter is applied to Breast Cancer diagnosis, based in mammograms. The tool classifies the mammograms in two classes: with tumor and without tumor. From the diagnostic data, mammograms, the software tool can propose to the medical doctor a possible assessment.

Breast Cancer is a major cause of death amongst women in the world (Ferlay, 2010). In the world, 22.9% of Cancer diagnoses are Breast Cancer, which causes 13.7% of Cancer deaths amongst women (Ferlay, 2010). In the European Union (EU) this reality is been tackled to increase the survival rates of women with Breast Cancer diagnose, by governments. In the EU, the overall spending of Breast Cancer care amounts to about 0.5-0.6% of total health care expenditure (Organisation for Economic Co-operation and Development (OECD), 2010). This effort has been rewarded with the decrease or stabilization of the death rates in the EU, despite the increase of incidence rates of Breast Cancer over the past decade (OECD, 2010). These facts show the increase of the survival rates due to earlier diagnosis and/or better treatments (OECD, 2010).

According to the OECD Health Data (OECD, 2010) the survival rate is 79.4%, ranging from 61.6% in Poland to 88.3% in Iceland, with only 51.4% of the women (with 50 to 69 years) screened in the OECD countries, ranging from 10.7% in Turkey to 84.4% in Finland. This last fact, i.e., the percentage of screened women, is of great importance to detect Breast Cancer in the early stages, which increase can contribute to diminish the mortality rates. The governments screening programs based on mammograms have been discussed by (Cady & Chung, 2005) concluding that its benefits overtake its drawbacks, suggesting that its usage should be expanded.

The effort to expand mammograms screening programs to the world entire population will induce the problem of interpreting an extremely high amount of data. The number of mammographic equipment must increase and most importantly is the high number of experts (radiologists) needed, which have to examine the mammograms. Another issue is the lack of radiologists in remote and under populated areas around the world. The level of expertise of the current radiologists must be maintained, or preferably be increased. The facts mentioned above justify the efforts on developing automatic systems to help medical doctors (radiologists) in their diagnosis. Such systems allow better result diagnosis and/or lower the volume of work of radiologists, when used has a second opinion or a first opinion to discard evident diagnostic cases, respectively. These systems are called in the literature Computer Aided Diagnosis or Detection (CAD) Systems (Mohanty, Champati,