Chapter 40
Mammogram Mining Using Genetic Ant-Miner

K. Thangavel
Periyar University, India

R. Roselin
Sri Sarada College for Women (Autonomous), India

ABSTRACT

Image mining deals with the extraction of implicit knowledge, image data relationship, or other patterns not explicitly stored in the images. It is an extension of data mining to image domain and an interdisciplinary endeavour. This chapter focuses on mammogram classification using genetic Ant-Miner. The key idea is to generate classifier for classifying mammograms as normal or abnormal using the proposed Genetic Ant-Miner algorithm. The Genetic Algorithm has been employed to optimize some of the ant parameters. A comparative analysis is performed in order to achieve the efficiency of the proposed algorithm. Further, the experimental results reveals that the improvement of the proposed Genetic Ant-Miner in the domain of Biomedical image Analysis.

INTRODUCTION

A mammogram is an X- ray of the breast. Mammography is a specific type of imaging that uses low-dose x-ray system to examine breasts. A mammography examination called a mammogram is used to aid in the early detection and diagnosis of breast diseases.

Breast cancer is a cancer that starts in the breast, usually in the inner lining of the milk ducts or lobules. It occurs in both men and women, although male breast cancer is rare.

DOI: 10.4018/978-1-4666-2455-9.ch040

A lot of research has been focused on the development of algorithms for the automated classification of abnormal mammograms. These algorithms are based either on morphology and distribution features of Micro Classifications (MCs) extracted by radiologists or on computer-extracted image features (Baker, Kornguth, Iglehart & Floyd, 1996). There are two methods to extract image features. The first category accounts for morphology/shape features of individual MCs or of MC clusters, while the second category corresponds to texture features extracted from Regions of Interest (ROI) containing the MCs.

The performance of the Computer Aided Diagnosis schemes is differentiated with respect to the features investigated, the classifiers used and the data sets analyzed. The success of the morphological features-based schemes strongly depends on the robustness of the MCs segmentation algorithms (Veldkamp & Karssemeijer, 1996). Specifically, in the case of dense breast parenchyma abutting the MCs, classification is a challenging task due to difficulty in the segmentation process. The texture analysis approach seems to overcome this limitation as no segmentation stage is required. The rationale of using texture features is based on capturing changes in the texture of the tissue surrounding MCs.

Mining information and knowledge from large data-base has been recognized by many researchers as a key research topic in database system and machine learning. One of the data mining tasks gaining significant attention is the classification rules extraction from databases. The goal of this task is to assign each case (object, record, and instance) to one class, out of a set of predefined classes, based on the values of some attributes for the case. There are different classification algorithms used to extract relevant relationship in the data as decision trees which performing a successive partitioning of cases until all subsets belong to single class (Quinlan, 1987). Medical images have been classified by Osmar R. Zaiane, M.L Antonie and A. Coman (2002) using Association Rule based classifiers. Shuyan Wang, Mingquan Zhou and Guohua Geng (2005) proposed medical image classifier based on decision tree algorithm. The experimental results show the classification accuracy is around 69% to 80%. Walid Erray, and Hakim Hacid (2006) proposed a method that was able to take into account the costs in the automatic learning process using decision trees and got promising results.

Metaheuristics are generally applied to problems for which there is no satisfactory problem-specific algorithm or heuristic; or when it is not practical to implement such a method. Most commonly used metaheuristics are targeted to combinatorial optimization problems. The Ant-Miner algorithm is proposed by R. S. Parpinelli, H. S. Lopesand, A. Freitas (2002) applies an ant colony optimization heuristic to the classification task of data mining to discover an ordered list of classification rules. In a colony of social insects, such as ants, bees, wasps and termites, each insect usually performs its own tasks independently from other members of the colony. However, the tasks performed by different insects are related to each other in such a way that the colony, as a whole, is capable of solving complex problems through cooperation. Ant-Miner is interested in a particular behavior of real ants, namely the fact that they are capable of finding the shortest path between a food source and the nest without the use of visual information. As ant move, a certain amount of pheromone is dropped on the ground, marking the path with a trail of this substance. The more ants follow a given trail, the more attractive this trail becomes to be followed by other ants. This process can be described as a loop of positive feedback, in which the probability that an ant chooses a path is proportional to the number of ants that have already passed by that path.

This chapter attempts to apply Ant-Miner in mammogram processing and the Ant-Miner Parameters are optimized by using Genetic Algorithms.

METHODS AND MATERIALS

The Mammographic Image Analysis Society (MIAS) Database is used to perform the analysis of the efficiency of the proposed algorithm. The data is available at http://peipa.essex.ac.uk. The database contains 322 mammograms including normal, mass, and micro calcification cases. It indicates different classes of abnormalities such as calcification, well-defined circumscribed masses, speculated masses, ill-defines masses, architectural distortion, asymmetry and normal.