CELL-GRAPH MINING FOR BREAST & BONE TISSUE MODELING AND CLASSIFICATION

By

Cemal Çağatay Bilgin

A Thesis Submitted to the Graduate

Faculty of Rensselaer Polytechnic Institute

in Partial Fulfillment of the

Requirements for the Degree of

MASTER OF SCIENCE

Major Subject: COMPUTER SCIENCE

Approved:

_____

Bülent Yener
Thesis Adviser

Rensselaer Polytechnic Institute
Troy, New York

November 2007
(For Graduation December 2007)

ABSTRACT

The most reliable way to diagnose cancer in the current practice of medicine is through pathological examination of a biopsy which has a certain level of subjectivity. To reduce this subjectivity and have a mathematical model for diagnosing cancer tissues we consider the problem of automated cancer diagnosis tools in the context of breast and brain tissues. In this work we present graph theoretical techniques that identify and compute quantitative metrics for tissue characterization and classification. We segment the digital images of histopatological tissue samples of breast and bone using k-means algorithm. For each segmented image we generate different cell-graphs using positional coordinates of cells and surrounding matrix components. These cell-graphs have 500-2000 cells(nodes) with 1000-10000 links depending on the tissue and the type of cell-graph being used. Having generated the graphs, we calculate a set of global metrics from cell-graphs and use them as the feature set for learning. We compare our techniques in different scenarios, and show that hierarchical cell graphs, gives 81.8% accuracy whereas with Learning based on intensity values of images, Delaunay triangulation of the cells, the previous technique we proposed for brain tissue images we obtain 61.0%, 54.1% and 75.9% accuracies. The ECM-aware cell-graph approach is tesed on the bone tissue samples and gives 88.7% accuracy, simple-cell graphs perform 81.4% and Delaunay triangulation method obtains 71.8 learning ratio.