Comparison of 1st and higher order structural diagrams as graph based features for breast histopathology whole slide images

CAD techniques are being used for medical image analysis over a decade with increasingly more complex and more accurate results. The use of CAD started with the cytopathology images about 15 years ago [1], and was relatively successful due to the simple and easily characterizable nature of the cytopathology images which usually do not contain much of unrelated tissue and large diverse clusters, since inside a cell structure the arrangement and existence of different parts are more deterministic. However, the histopathology images should be used for problems such as cancer diagnosis and gradation do not conform to these standards for several reasons: More complex structure with a lot of tissue and ductal regions, imperfect Hematoxylin-Eosin staining that should be dealt with, larger image sizes (around 600 GB's per image with 80K x 90K pixels), random geometrical arrangement of nuclei and so on. My research focuses on the extraction of ROI's and classifying them with proper graph features so that the whole-slide image can be scanned faster and accurately with the aid of computers and can help doctors save time and increase the accuracy of diagnoses.

The general trend in CV for any branch usually follows the same stages: First, color based data (histograms) are analyzed, then textural analysis (Gabor features, co-occurrence matrices etc.), and finally structural (graph based) analysis is employed. Currently, the CAD usually relies on textural analysis, and some graph-based analysis. This project will focus on determining the best graph features (via 1st and higher order Voronoi and/or Delaunay diagrams) to extract (average/std. dev. of area/ perimeter/chord, average side length/area of triangles, and so on), so that the accuracy in determining if a whole slide image consists of any cancerous cell clustering, and if so, what type of cancer that is by the appropriate clustering and matching of the features by general classification techniques. The results will be compared with the previous research efforts (which claim very high accuracy rates yet may be biased due to the "appropriate" selection of easily discernible validation data [2]) and presented as a useful resource as a later reference for my research and possible for other related research.

The main focus of this research is determining if a higher order Voronoi diagram based classification can yield better results in terms of detection accuracy. The general approaches in CV field tend to stop at the first order Voronoi diagrams and the higher order diagrams are not explored enough to make a conclusive remark whether the added complexity and cost is a good trade-off that can improve the performance or it is negligible.

References
