Chapter 2

Literature Survey

Sickle cell anaemia is a well understood medical condition. There is no widely available standard cure. However there are certain treatments to prevent the symptoms and thus avoid complications resulting from it. These treatments require early detection as mentioned above. However there isn't any automated test for sickle cell anaemia. This presumably stems from insufficiency of literature or studies involving automation of diagnostic process. Correspondingly, literature involving automated detection of sickle celled disease remains considerably scarce.

One of such recherché studies published on Science Direct is 'Detection of Abnormal Findings in Human RBC in Diagnosing Sickle Cell Anaemia Using Image Processing' authored by Pranati Rakshita and Kriti Bhowmikb [3]. This study treads along parallel ideologies as our project. The proposed methodology in it involves preprocessing, edge detection and region selection.

Preprocessing requires initially converting the blood smear image into a binary form. Then an adaptive filtering method is used to eliminate unwanted noise present in it. The filter of choice used in this paper was Weiner's filer. The next step is edge detection which demarcates boundaries for the red blood cells. The study approves the use of any of the following edge detectors: Sobel Operator, Robert's Operator, Canny Operator, LoG Operator, Zerocross Operator and Prewitt Operator [4]. The third step is region selection, wherein we measure properties of connection image components which satisfy certain predefined conditions. It is used to compute the shape measurements like the centroid, area, bounding box convexHull, convexArea, perimeter etc. An extensive use of properties area and perimeter is done. Area is a scalar value which represents the actual number of pixels in the region and perimeter is used to calculate the distance around the boundary of the region. In MATLAB, these measurements can be computed using the inbuilt function 'RegionProps'.

Using the measurements obtained, the study proposes a metric to determine the circularity of objects (which are the RBCs) in the image. This metric is defined as: (4*pi*area)/perimeter^2 [1]. The metric ranges from values of 0 to 1; 1 for a perfectly 2D circle, 0.785 for a 2D square and so on goes decreasing as the shape resembles less of a circle. So a typical RBC would have a metric higher than 0.82 whereas a sickle shaped RBC would have a much lower metric of about 0.4 – 0.5. This we obtain a clear distinction between different shapes of RBCs present.

The final step involves classification of the input image as anaemic or not. In a study by NIIT Rourkela, they use clustering algorithms for segmentation of images [5]. One such algorithm is K-means clustering. It is an algorithm based on finding data clusters in a data set such that a cost function (or an objection function) of dissimilarity (or distance) measure is minimized. K-means uses an iterative algorithm that minimizes the sum of distances from each object to its cluster centroid, over all clusters. This algorithm moves objects between clusters until the sum cannot be decreased further. The result is a set of clusters that are as compact and well-separated as possible.

Another approach suggested here is the use of Fuzzy C-means clustering. It is one of the commonly used methods for image segmentation and its success is mainly due to the introduction of fuzziness for the belongingness of each image pixels. Compared with crisp or hard segmentation methods, FCM is able to retain more information from the original image, with the only disadvantage being increased sensitivity to noise and other imaging artifacts. Fuzzy c-means (FCM) is a data clustering technique in which a dataset is grouped into ‘n’ clusters with every data point in the dataset belonging to every cluster to a certain degree. For example, a certain data point that lies close to the center of a cluster will have a high degree of belonging or membership to that cluster and another data point that lies far away from the center of a cluster will have a low degree of belonging or membership to that cluster.