FUZZY SET THEORETIC APPROACH TO MICROSCOPIC IMAGE EVALUATION OF CHRONIC MYELOGENOUS LEUKEMIA

Abstract

This thesis narrates quantitative characterization and classification of CML using light microscopic images of PBSS. It has been mainly emphasized on the development of integrated methodologies based on fuzzy set theoretic approach for modeling leukocyte and its nucleus.

The system is composed of five basic steps viz. collection of microscopic images of peripheral blood smears from medical college & hospital, preprocessing of microscopic images, segmentation of leukocyte from the background of image, relevant feature extraction from the leukocytes and finally recognition of CML class from the assessment of different crisp and fuzzy classifiers. Microscopic images of peripheral blood smear samples (PBSS) were collected from Midnapur Medical College and Hospital, West Bengal, India. Gray world assumption technique was considered for correcting illumination and midpoint filter was chosen for noise removal on the green channel of grabbed image. In this thesis, two different modifications of discrimination measure for the development of fuzzy divergence measure have been proposed. Firstly the discrimination measure was modified by hedge operator and applied to the development of divergence measure using Shannon's and Yeager's exponential entropy functions. Secondly, the discrimination measure was modified with a proposed function and applied to make of divergence measure using Shannon's exponential entropy function. Both of the proposed measures were applied to segment the leukocyte from background of PBSS and compared with literature by means of quality measures. In addition , a methodology has been proposed for detection of boundary region of nucleus within the cell whereby Chan Vese model of active contour detection method was optimized with three evolutionary algorithms viz. differential evolution, genetic algorithm and nondominated sorting genetic algorithm II. Amongst these, differential evolution provided the best segmentation accuracy in segmenting the nucleus from leukocyte cell.

Features extracted from both the cytoplasm and nucleus of leukocyte were of types entropy, histogram based first order statistical texture, GLCM based texture, gray level run length matrix based texture, fractal dimension, local binary pattern, Hu's invariant moments, Fourier descriptors, wavelet, Gabor wavelets, geometrical shape features. Feature set was statistically analyzed using t-test, kernel density approximation and box plot. t-test was performed separately for 154 features of leukocyte cell of NOM and CML classes as well as leukocyte nucleus of NOM and CML classes. After then, PCA is applied for compression of 160 significant features which have 'p value' < 0.001. The first 18 principal components contributed 98.48% variability of the subset data. It was observed that first principal

component provided 43.69% variability in the data, whereas second one contributed 11.06% of variability. Comparative analysis of different crisp and fuzzy supervised classifiers viz. Bayes, classification & regression tree, probabilistic neural network, back propagation neural network, support vector machines with different kernels and interval type 2 fuzzy set, have been performed for different sub-groups of features. Based on the comparative analysis of classifiers it is found that 97.79%, 97.11%, 96.85% and 97.44% of average sensitivity, specificity, PPV and accuracy respectively have been provided by significant feature set. PCA feature set has achieved 97.24%, 96.61%, 96.16% and 96.85% of average sensitivity, specificity, PPV and accuracy respectively. GT2FS has been applied on significant feature set resulting in average accuracy of 98.78% which is higher than the highest accuracy found in PCA feature set and IT2FS performance. Highest accuracy (98.78%) and highest sensitivity (98.95%), specificity (98.64%), PPV (98.42%) have been achieved by applying general type 2 fuzzy set based classifier among all the classification models.

Keywords: Chronic myelogenous leukemia, Peripheral blood smears, Fuzzy divergence, Differential evolution, Genetic algorithm, Non-dominated Search Genetic Algorithm, Interval Type 2 Fuzzy Set, General Type 2 Fuzzy Set.