CHROMOSOME CLASSIFICATION USING SVM

[1]S.Balasubramanian, [2]Dr.P.Marichamy, [3]M.Piramu
[1]Associate Professor, [2]Dean, [3]Assistant Professor
[1] Department of Electronics and Communication Engineering,
[1]PSR Engineering College, Sivakasi, India.

Abstract : In this paper a methodology is proposed for automatic classification of human Chromosomes to improve accuracy and reduce loss. The methodology employs the following steps: preprocessing, training, SVM classifier & Fourier descriptor for segmentation and classification of chromosomes. The chromosomes are classified into normal and dicentric based on the shape. The proposed method is evaluated and compared with the conventional method.

***IndexTerms* - SVM classifier, chromosome, Fourier descriptor, segmentation & classification.**

I. INTRODUCTION

Each cell consists of nucleus; the chromosomes are tightly packaged into thread like structured DNA molecule. Chromosomes consist of DNA molecules tightly twisting around proteins known as histones that sustain its shape. Chromosomes are invisible in the nucleus of the cell. By using microscope also the chromosomes are invisible when the cell is not divided. During cell division the chromosomes are visible under a microscope. Every chromosome has a limitation point known as centromere. The centromere divides the chromosomes into two sections or arms. The representation of "p arm" is known as short arm. The representation of "q arm" is known as long arm. The centromere location in chromosome gives the characteristics shape. If the chromosome is dicentric it will affect the future generation also. According to classify the chromosome into normal or dicentric, it is used to identify the disease easily in human body.

II. PROPOSED SYSTEM

This proposed system is used to classify the chromosome into normal and dicentric by using SVM classifier. The segmentation method used in this system is to improve the accuracy and reduce the loss percentage. Additional features are involved to improve the classifier result. Shape boundary extraction and shape based Fourier feature computation are used. They are validated to compare normal and dicentric chromosomes better than conventional approach.

III. ARCHITECTURE OF PROPOSED SYSTEM

The block diagram of a proposed system is shown in Figure 1. It contains the following steps.

- Pre processing
- Localization and segmentation
- Shape based feature extraction
- SVM classification
- Wi-fi

3.1 Preprocessing:

Pre-processing is a common name for operations with images at the lowest level of abstraction both input and output are intensity images. The aim of pre-processing is an improvement of the image data that suppresses unwilling distortions or enhances some image features important for further processing. The metaphase images are grey-level microscope images shown in Figure 1.1

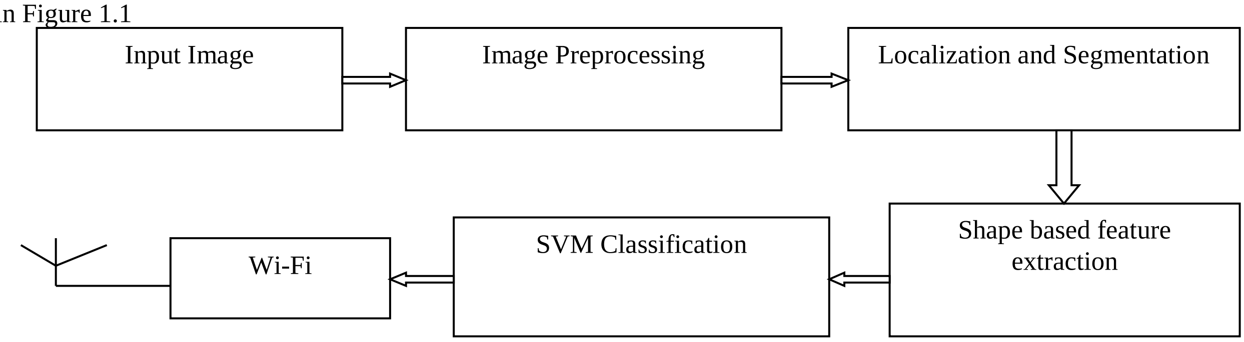

Figure 1.Block Diagram of a Proposed System

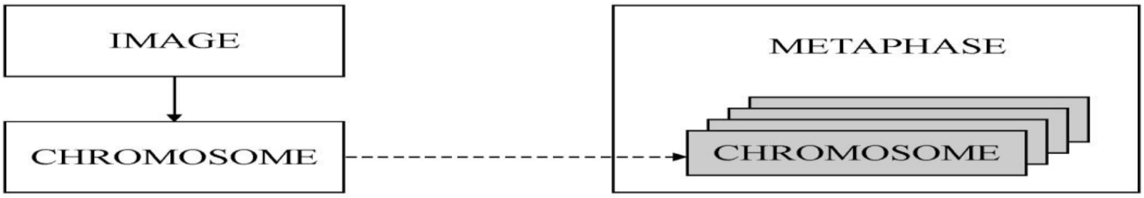

Figure 1.1 Block Diagram of Preprocessing

3.2 Localization and Segmentation

Many biological images comprises of light objects over a constant dark background (especially those obtained using fluorescence microscopy) in such a way that object and background pixels have gray levels grouped into two dominant modes. One obvious way to select a threshold T that separates these

g(x,y) = 1 if f(x,y)>T;
= 0 otherwise;

where g(x,y) is the threshold binary image of f(x,y). We can implement the threshold operation in MATLAB by the following function:

g= im2bw (f,T)

The first argument f gives the input image and the second argument T gives the threshold value.

In computer vision and image processing, Otsu's method, named after Nobuyuki Otsu, is used to automatically perform clustering-based image thresholding, or, the reduction of a gray level image to a binary image. The algorithm assumes that the image contains two classes of pixels following bi-modal histogram (fore ground pixels and back ground pixels), it then calculates the optimum threshold separating the two classes so that their combined spread (intra-class variance) is minimal, or equivalently (because the sum of pairwise squared distances is constant), so that their inter-class variance is maximal. Consequently, Otsu's method is roughly a one-dimensional, discrete analog of Fisher's Discriminant Analysis. Otsu's method is also directly related to the Jenks optimization method. One way to find boundaries of objects is to detect discontinuities in intensity values at the edge of a region. These discontinuities can be found by calculating the first and/or second order derivatives of an image.

Smoothing is used to eliminate edge effects, weight contribution of neighborhood pixels according to their closeness to the center. Median filter is involved in filtering process. A median filter operates over a window. We consider, an important nonlinear filter, the median filter, which has its own implementation in MATLAB. MATLAB also has a function to generate noise in an image.

Median filtering follows this basic prescription. The median filter is normally used to reduce noise in an image, somewhat like the mean filter. However, it often does a better job than the mean filter of preserving useful detail in the image. This class of filter belongs to the class of edge preserving smoothing filters which are non-linear filters. This means that for two images *A(x)* and *B(x)*: Median [A(x) + B(x)] ≠ Median [A(x)] + median [B(x)]

3.3 Shape based feature extraction

Shape has been shown to be an important component in the visual scene understanding by humans. In the extracted chromosomes the distinguishing feature in the dicentric chromosomes are two additional constrictions along their body length. Further, the chromosomes exhibit variability in size and shapes even amongst normal chromosomes. Another observation is that the chromosomes are manifested differently oriented. Shape feature methods are widely used over and above segmentation in image analysis [7]. The shape features for our use had to be robust to scale, rotation and boundary start point. Various methods of shape representation and feature extraction are found in literature such as those based on geometry [8, 9].

Fourier based shape descriptor was adopted as many desirable invariance properties are achievable in the Fourier space representation [10, 11].

3.4 SVM for Classification

SVM is a useful technique for data classification. Even though it's considered that Neural Networks are easier to use than this, however, sometimes unsatisfactory results are obtained. A classification task usually involves with training and testing data which consist of some data instances. Each instance in the training set contains one target values and several attributes. The goal of SVM is to produce a model which predicts target value of data instances in the testing set which are given only the attributes. Classification in SVM is an example of Supervised Learning. Known labels help indicate whether the system is performing in a right way or not. This information points to a desired response, validating the accuracy of the system, or be used to help the system learn to act correctly. A step in SVM classification involves identification as which are intimately connected to the known classes. This is called feature selection or feature extraction. Feature selection and SVM classification together have a use even when prediction of unknown samples is not necessary. They can be used to identify key sets which are involved in whatever processes distinguish the classes.

IV. PROPOSED SYSTEM RESULTS

The result of proposed system is shown in the following figures.

4.1 DICENTRIC IMAGE OUTPUT

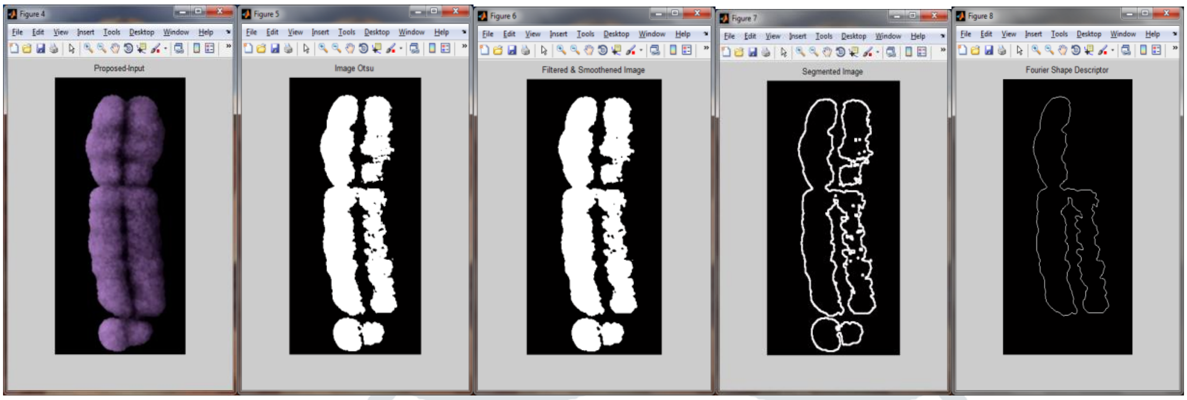

Figure. 2. Dicentric images

4.2. NORMAL IMAGE OUTPUT

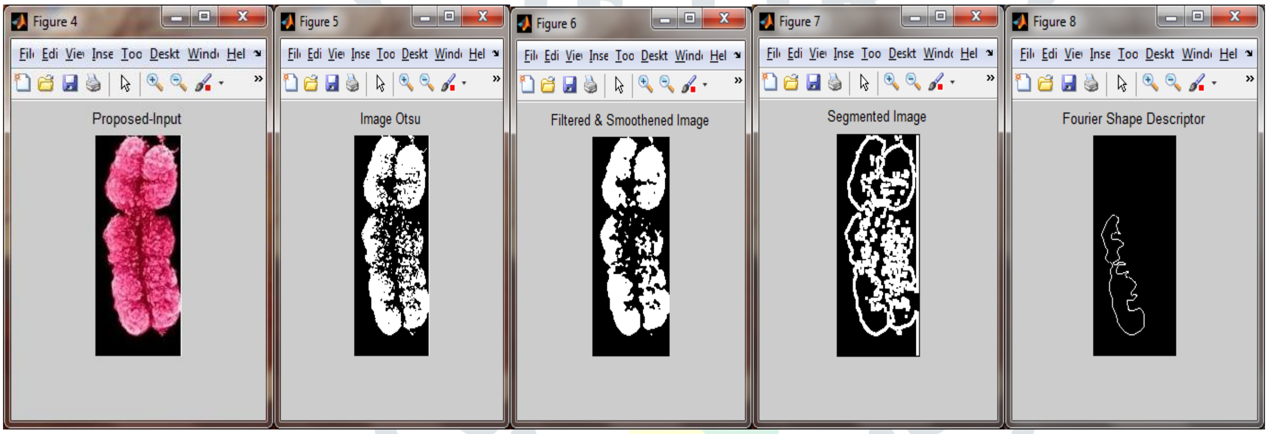

Figure 3. Normal images

V. RESULT ANALYSIS

5.1 Previous Work Result

TYPE	NORMAL	DICENTRIC
Accuracy	99.0383	94.6645
Loss	0.9617	5.3355

5.2 Proposed Work Result

TYPE	NORMAL	DICENTRIC
Accuracy	99.0624	94.95
Loss	0.9576	5.4540

VI. CONCLUSION AND FUTURE ENHANCEMENT

By using this proposed system, we can easily classify the normal and dicentric chromosome. We can improve the accuracy level and reduce the loss percentage by using Segmentation and SVM classification method. It will be useful for easily identify diseases.

The proposed method presents a framework for incorporating additional features into the Laplacian based thickness measurement process. This framework may be used to integrate intensity information into the thickness measurement process in the research. However, the possibility of incorporating other relevant features into thickness measurement processes in different applications

VII. REFERENCES

[1]. A. S. Arachchige, J. Samarabandu, J. H. M. Knoll, and P. K. Rogan "Intensity Integrated Laplacian-Based Thickness Measurement for Detecting Human Metaphase Chromosome Centromere Location", IEEE Trans. on Biomedical Imag, vol. 60, no. 7, Jul 2013.

[2]. W. Yan, D. Li, "Segmentation Algorithms of Chromosome Images", 3rd International Conference on Computer Science and Network Technology, 2016.

[3]. Christoph Dalitz, Christian Brandt, Steffen Goebbels, David Kolanus, "Fourier Descriptors for Broken Shapes", EURASIP Journ. On Advances in Signal Proc, 2013.

[4]. Christoforos Markou, Christos Maramis, Anastasios Delopoulos, "Automatic Chromosome Classification using Support Vector Machines".

[5]. Shervin Minaee, Mehran Fotouhi, Babak Hossein Khalaj, "A Geometric Approach to Fully Automatic Chromosome Separation", Signal Processing in Medicine and Biology Symposium (SPMB), 2014 IEEE

[6]. J. Piper and E. Granum, "On fully automatic feature measurement for banded chromosome classification," *Cytometry*, vol. 10, pp. 242–255,1989.

[7]. M. Moradi and S. K. Saterahdan, "New features for automatic classification of human chromosomes: A feasibility study," *Pattern Recognit. Lett.*,no. 27, pp. 19–28, 2006.

[8]. X. Wang, B. Zheng, S. Li, J. J. Mulvihill, and H. Liu, "A rule-based computer scheme for centromere identification and polarity assignment of metaphase chromosomes," *Comput. Methods Programs Bio Med.*, vol.89, pp. 33–42, 2008.

[9]. A. S. Arachchige, J. Samarabandu, J. Knoll, W. Khan, and P. Rogan, "An image processing algorithm for accurate extraction of the centerline from human metaphase chromosomes," in *Proc. Int. Conf. Imag. Process.*, Sep.2010, pp. 3613–3616.

[10].R. M. Mohammad, "Accurate localization of chromosome centromere based on concave points," *J.Med. Signals Sens.*, vol. 2, no. 2, 2012.

[11]. A. S. Arachchige, J. Samarabandu, J. Knoll, W. Khan, and P. Rogan, "An accurate image processing algorithm for detecting fish probe locations relative to chromosome landmarks on DAPI stained metaphase chromosome images," in *Proc. 7th Can. Conf. Comput. Robot Vis.*, May 2010.