

# Detection of Leukemia Cancer Cells Using Image Processing Techniques

Moomin Muhiuddeen, M. Sharmila Kumari  
M.Tech. Student, Head Of The Department  
Department of Computer Science and Engineering,  
P. A. College of Engineering, Mangalore, 574153, India

**Abstract:** Leukemia disease can be treated effectively if it is detected early and also by detecting the type of leukemia. This process of observation, which is manually done by a hematologist, may be prone to errors since it is time consuming and tiring to perform such a tedious task. Thus, there is a need for an automated system which performs segmentation of the blood sample image by mainly focusing on automatic recognition and separation of each cell, in order to obtain certain features that will help in detection of the disease. In order to make their work easy and efficient, we have implemented an automated software system that will perform certain image processing techniques such as image enhancement, image segmentation, feature extraction and classification, which will help in early detection of leukemia efficiently and consumes less time. GLCM and GLDM are used as feature extraction techniques. K-means segmentation is used for segmentation. It is done to obtain the area of interest in cells, in this case it is the nucleus.

**IndexTerms**–K-Means, GLCM, GLDM.

## I. INTRODUCTION

The most vital part of any human body is blood as it keeps one alive. It performs various essential limits, for instance, to trade oxygen, carbon dioxide, and mineral thus on to the whole body. In view of the insights it is been understood that blood cancer is the fifth reason for loss of life in human beings.

Leukemia is a cancer in blood, is reparable when it is identified and cared at right time. Its recognition begins with a complete blood count. In the event that there are irregularities in this tally, an investigation was done by pathologists on few of the infected cells to identify whether the cell is cancerous or not and also to find type of leukemia. But there were few drawbacks in this investigation due to many reasons such as time consuming analysis, less accuracy and depending on operators skills. This investigation had an error rate between 30% and 40% in spite of great experience of the pathologists. A channel cytometric test gives an exceptionally precise to arrange leukemia's, yet it is extremely costly and not every one of the healing centers can afford the hardware. The grouping of leukemia sorts encourages the doctors in choosing which analysis suits for a type of cell sort (lymphocytic or myelogenous) infection growth (intense or ceaseless). Paper displays a pre-processing strategy for cancerous cells. K-means algorithm is used for segmentation. Finally the objective of the paper is to generate an element which describes whether the cell is cancerous or not and also identify the type of leukemia, where mainly leukemia are of four types, they are acute myeloid leukemia, acute lymphocytic leukemia, chronic myeloid leukemia and chronic lymphocytic leukemia. It is used in a variety of fields such as pattern recognition and data mining.

## II. LITERATURE REVIEW

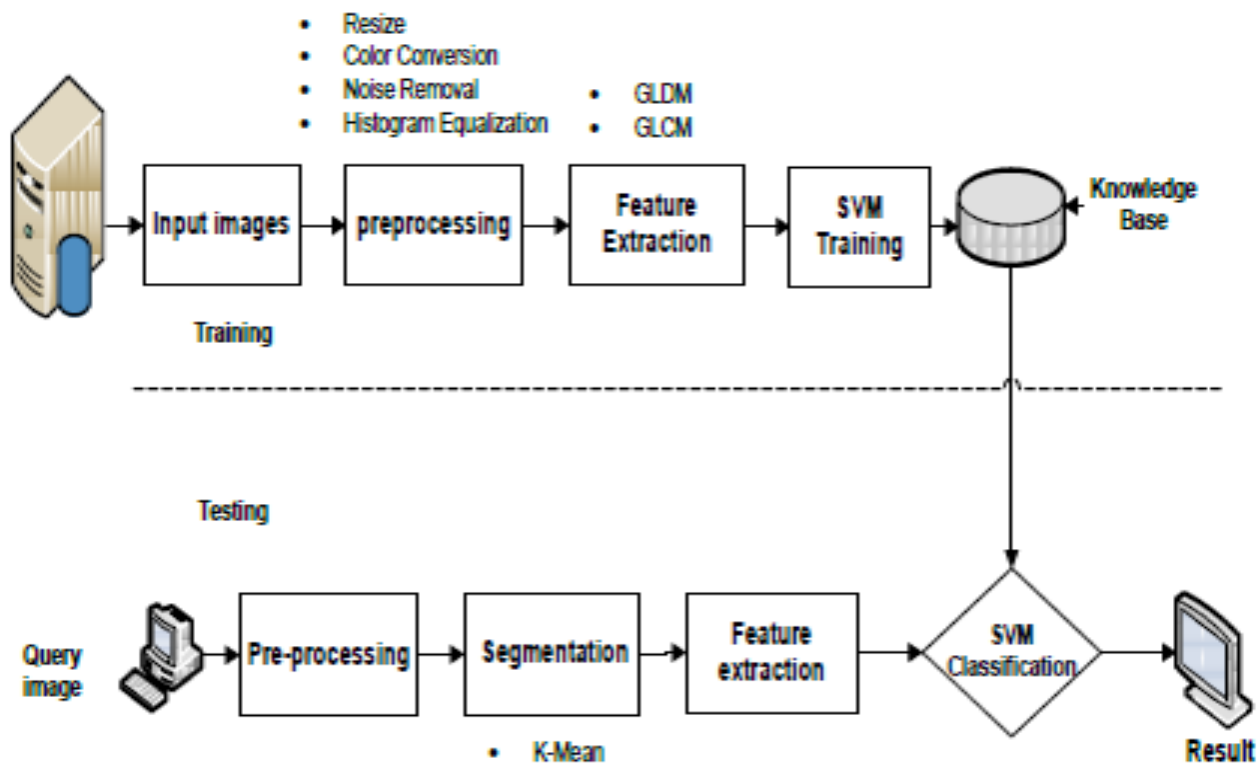
Symptomatic radiography assigns the technological parts of medical pictures and specifically in acquiring therapeutic images. Dr. S Venkatachalam [1] presents the pre-processing strategies for the leukemia injected cells where the final aim is to generate the elements which describe types of leukemia. The undertaken issues contain: the cell segmentation [1] by using the watershed change, determination of distinct cells, and texture quality, statistical and geometrical examination of the cells. Image Processing techniques are commonly used as one of the element under imaging research. These strategies are valuable for representation, enhancement, segmentation and numerous more operations which are helpful for processing medicinal image which perhaps MRI, CT or whatever other images acquired through one of the imaging methodology. One of the advantages of utilizing these methods is to identify any abnormality from the norm in the image of medical application. Some of the applications are, detecting tumor, blocked vessels or broken joints. [2] Vinay Parameshwarappa proposed a strategy for recognizing one such variation from the norm saw in brain image. Utilizing a portion of the traditional picture handling devices and fourier transform.

According to Bhagyashri G. Patil [3] overview, as of late Lung growth cell is gaining the consideration of therapeutic and affected groups under the most recent times in light of its high prevalence unified using hard treatment. Insights from 2008 demonstrate that lung disease, all through world, is the one that assaults the best number of people. Early identification of lung growth is essential for fruitful treatment. There is couple of strategies accessible to identify dangerous cells. Here two techniques for division, for example, thresholding and watershed are utilized to distinguish the disease cell and too discover better approach out of them.

Fundamental mathematical process hypothesis are presented at initially, utilized for identifying the edges and additionally the [4] tumor lung cells MRI and CT images. Since, salt and pepper noise are more pervasive in medical images the routine techniques are not successful in sifting salt and pepper noise. [4] Numerical morphological processes are utilized to identify the edges and the disease cells. Morphological disintegration is a decent channel of salt and pepper commotion. The trial results demonstrate that the proposed calculation is more proficient for restorative picture de-noising, edge location and recognizable proof than the normally utilized layout based on edge identification calculations and morphological edge recognition algorithm.

### III. PROPOSED METHOD

Proposed system has two parts training and testing. Both the parts undergo following steps. Image acquisition is the initial step collecting images of the blood from microscope with proper magnification from any of the hospitals. Second step is image preprocessing where following steps are followed initially color conversion takes place, color image is converted to gray scale. Followed by filtering the image, removal of noise from the image and finally histogram equalization is done to increase the quality of image in terms of contrast. Third step is segmentation using k-mean clustering, where k is considered 3 and nucleus is concentrated for the detection process. Segmentation is followed by feature extraction where features of nucleus are extracted using GLCM and GLDM. In the training part features of pure cancer cell is stored in knowledge base. In the testing part, the cell which needs to be tested is taken as input. And finally SVM classifier with the help of data in the knowledge base is used for classification, where decision is done whether the cell is cancerous or not. If the cell is cancerous then its type is found among acute myeloid leukemia, acute lymphocytic leukemia, chronic myeloid leukemia and chronic lymphocytic leukemia. **Figure 1** describes the architecture of proposed system.



**Figure1:** Overview of the Proposed Method.

#### 3.1 K-Means Segmentation

The pre-processed K-Means is a least squares apportioning strategy that partition an accumulation of articles into K groups. K-Means method is an unsupervised [5] grouping method that characterizes the information focuses into different classes. The algorithms expect that a vector space shaped by the data features and tries to discover common grouping in them. In the proposed system k is considered as 3. Clustering method works as follows considering large number of data sets, take few of the random center points. We group the data sets to the nearest center points. Find the new centers for these cluster groups, and again the clustering should be done based on the distance to the nearest center. This procedure should be continued till there are no changes in the arrangement of data sets.

The algorithm iterates over following steps:

1. Determine the intensity distribution of the intensities (also called the histogram).
2. Introduce the centroids with k irregular intensities.
3. Iterate the upcoming steps until the cluster a label of the image does not change any longer.
4. Based on the distance of points from centroid intensity, cluster the points. Equation (1)

$$c(i) = \operatorname{argmin}_j \|y^{(i)} - \mu_j\| \quad (1)$$

5. For individual clusters, determine the new centroid.

“k” indicates number of clusters, i is iterations of intensities, and j is iterations of all centroid, and  $\mu_j$  is intensities of centroids .

#### 3.2 GLCM

The “Gray Level Co-Occurrence Matrix” (GLCM) which is applied in feature extraction technique, has turned out to be a well-known factual technique for removing textural highlight from pictures. It is used only for gray-scale image. Basically GLCM calculates number of times pixel i is vertical, horizontal or diagonal to pixel j, where i and j are gray scale intensity values. As per

co-event grid, Haralick characterizes fourteen textural highlights measured from the likelihood network to separate the qualities of surface insights of remote detecting pictures. In our proposed system important features like Angular Second Moment (energy), contrast, auto Correlation, Entropy, variance, dissimilarity, homogeneity, cluster prominence and the Inverse Difference Moment are selected for implementation.

### 3.3 GLDM

Gray Level Difference Method (GLDM) depends on the event of two pixels which have a given total gray level difference and which are isolated by a particular displacement  $\delta$ .

For any displacement,  $\delta = (\Delta p, \Delta q)$ , Let  $S_0(p, q)$  be image intensity function.

Let  $S_0\delta(p, q) = |S_0(p, q) - S_0(p + \Delta p, q + \Delta q)|$  and the estimated probability-density function is defined by

$$f(i|\delta) = \text{prob}(S_0\delta(p, q) = i) \quad (2)$$

Use Equation (2) to calculate probability-density function.

### 3.4 Classifier

SVM is built up as one of the important instruments for machine learning and information mining. SVM is also used in multi-class classification. A special property of SVM is, SVM minimize error in empirical classification and maximize geometric margin. Hence it is called as Maximum Margin Classifiers. This study utilizes SVM to arrange pictures that was a factual order framework suggested by Cortes and Vapnik. [6]

Uses an extensive variety of example acknowledgment issues, picture grouping, money related time arrangement expectation, face recognition, biomedical sign investigation, restorative diagnostics, and information mining utilizes SVM in recent times. SVM Classifier classifies elements between two classes by drawing a hyper-plane. Maximum distance of the nearest element to the hyper-plane is called margin.

In this paper SVM is used as classifier to classify the cell, whether the cell is cancerous or not. The SVM also gives option between four kernel types: (i) Polynomial, (ii) Linear, (iii) Sigmoid, and (iv) Radial Basis Function.

## IV. RESULT AND ANALYSIS

In the proposed method microscopic image is enhanced using pre-processing steps, followed by segmentation using k-mean algorithm. **Figure 3** shows enhanced image using histogram equalization. The proposed method is applied to microscopic images. **Figure 2** shows the microscopic input image.

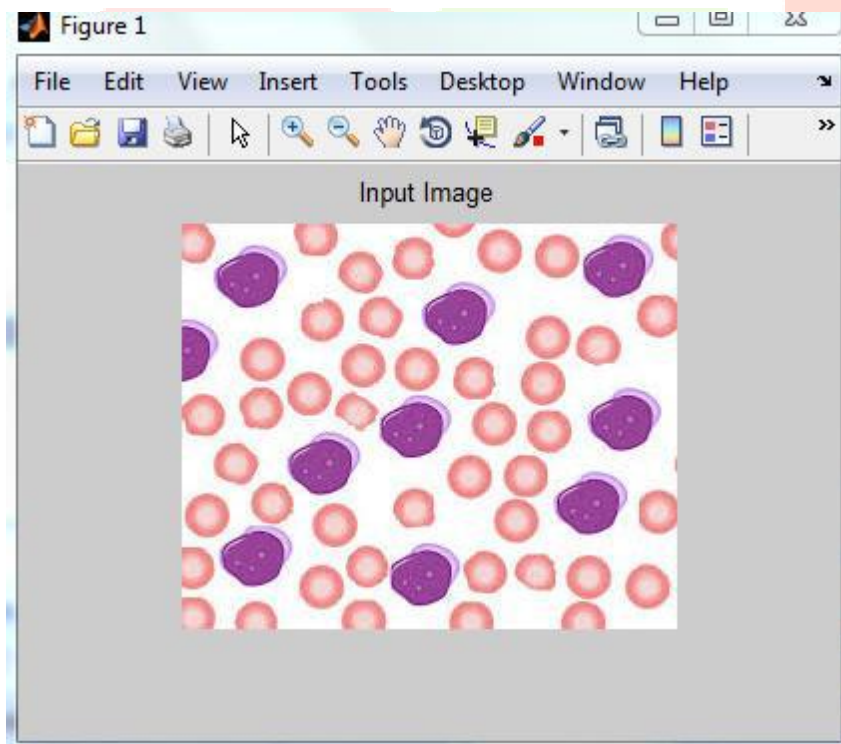


Figure2:Input Image

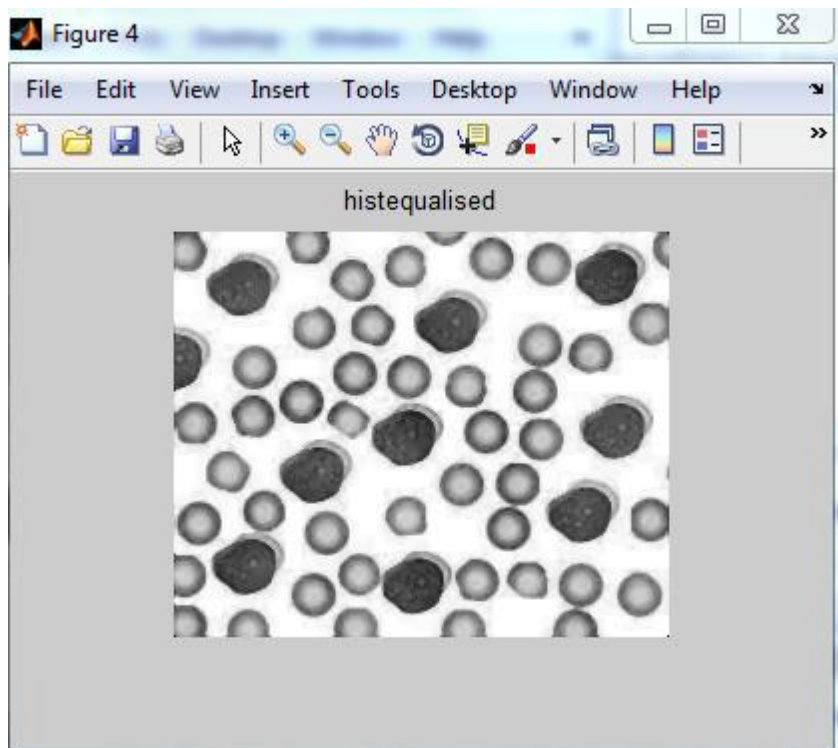


Figure3: Enhanced image

For K-mean segmentation we have selected 3 classes which is, “k” is taken 3. **Figure 4**, **Figure 5** and **Figure 6** shows k-means segmentation outputs.

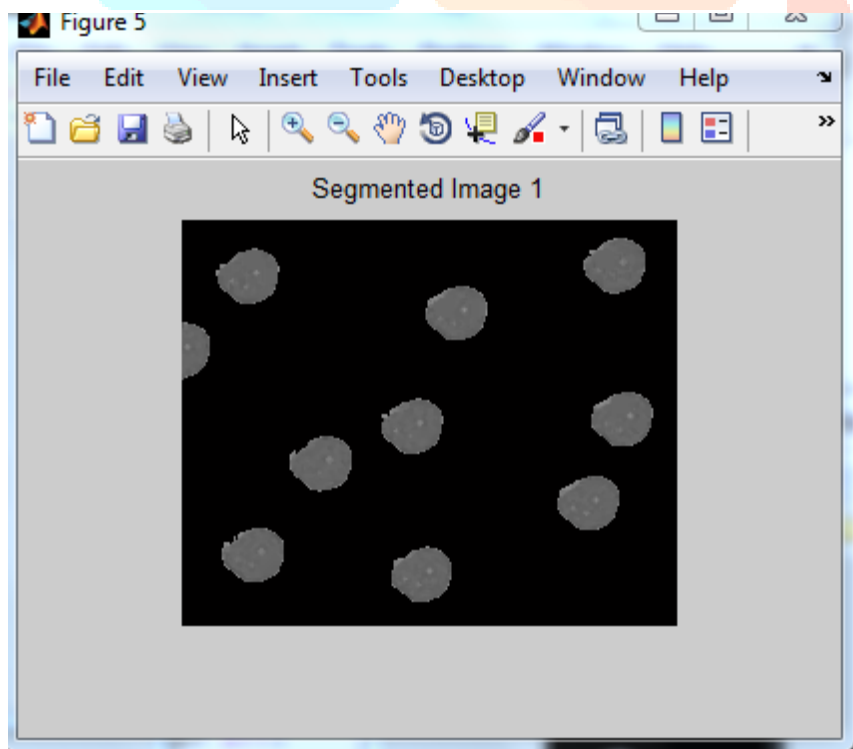


Figure4: Segmented Output of K-mean

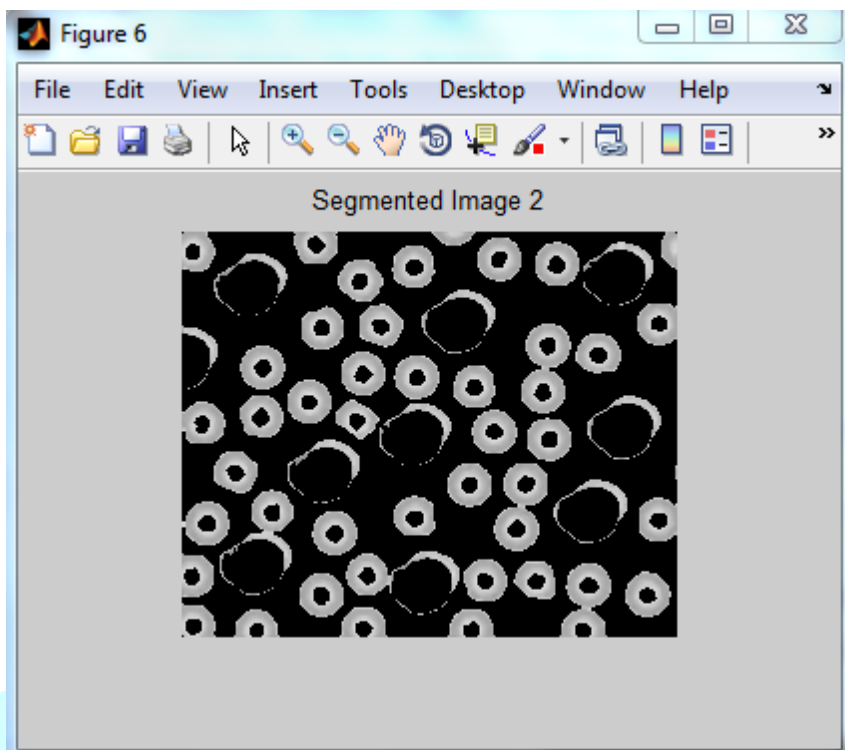


Figure5: Segmented Output of K-mean

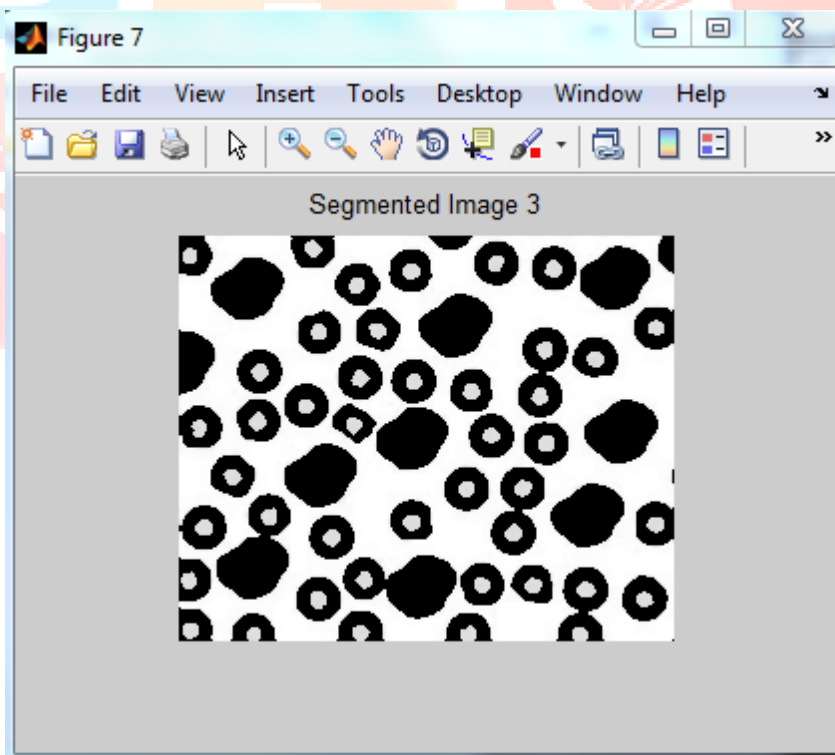


Figure6: Segmentation Output of K-mean

After obtaining segmentation, **Figure 4** is taken for feature extraction as it has complete nucleus which is the region of interest. Feature extraction is done through GLDM and GLCM. Finally SVM classifier is used to classify input microscopic image as cancerous or not.



## V. CONCLUSION

The main aim of this paper is cell segmentation followed by feature extraction to detect cancer cells. Features such as Angular Second Moment (energy), contrast, auto Correlation, Entropy, variance, dissimilarity, homogeneity, cluster prominence and the Inverse Difference Moment etc. are considered for accurate precision of identification. The results show that the k means method is applied for best segmentation performance. Likewise, the completely divided core can be better accomplished by utilizing Matlab based machine learning calculations since it is less sensitive to variations in the input image.

## VI. ACKNOWLEDGMENT

We would like to thank all my friends and staffs of P.A College of Engineering, Mangalore for their feedback and support.

## REFERENCES

- [1] S.Jagadeesh , Dr.E.Nagabhooshanam , Dr.S.Venkatchalam, "Image Processing Based Approach to Cancer Cell Prediction in Blood Sample", International Journal of Technology and Engineering Sciences, Vol.1 (1), ISSN: 2320-8007.
- [2] Vinay Parameshwarappa, Nandish S., "A Segmented Morphological Approach to Detect Tumour in Brain Images", International Journal of Advanced Research in Computer Science and Software Engineering, ISSN 2277 128X, Volume (4), Issue (1), January 2014.
- [3] Bhagyashri G. Patil, Prof. Sanjeev N. Jain, "Cancer Cells Detection Using Digital Image Processing Methods", International Journal of Latest Trends in Engineering and Technology (IJLTET), vol.3 ISSN: 2278-621X, 2014.
- [4] Sivappriya T , Muthukumaran K, "Cancer Cell Detection Using Mathematical Morphology", International Journal of Innovative Research in Computer and Communication Engineering, Vol.2, Special Issue 1, March 2014.
- [5] Ms.Chinki Chandhok, Mrs.Soni Chaturvedi, Dr.A.A Khurshid, "An Approach to Image Segmentation using K-means Clustering Algorithm", International Journal of Information Technology (IJIT), Volume – 1, Issue – 1, August 2012.
- [6] Richa Dandotiya, Shailendra Singh, and Shailesh Jalori, "A survey of image classification techniques for flood monitoring system" , International Conference on Emerging Trends in Computer and Image Processing (ICETCIP'2014) Dec. 15-16, 2014 Pattaya (Thailand)

