# TUMOR DETECTION USING MORPHOLOGICAL FEATURE EXTRACTION

Nikitha B.V

4th SEM Mtech (Signal Processing)

S.J.C Institute of Technology

Niki.gowda89@gmail.com

Aravind Kumar. N

Prof, Department of ECE

S.J.C Institute of Technology

arvind3861iw2@gmail.com

**Abstract-** Image processing is one of most growing research area these days and now it is very much integrated with the medical and biotechnology field. Image Processing can be used to analyze different medical and MRI images to get the abnormality in the image. This paper proposes an efficient K-means clustering algorithm under Morphological Image Processing (MIP). Medical Image segmentation deals with segmentation of tumor in CT and MR images for improved quality in medical diagnosis. It is an important process and a challenging problem due to noise presence in input images during image analysis. It is needed for applications involving estimation of the boundary of an object, classification of tissue abnormalities, shape analysis, contour detection. Segmentation determines as the process of dividing an image into disjoint homogeneous regions of a medical image. The amount of resources required to describe large set of data is simplified and is selected for tissue segmentation. In our paper, this segmentation is carried out using K-means clustering algorithm for better performance. This enhances the tumor boundaries more and is very fast when compared to many other clustering algorithms. This paper produces the reliable results that are less sensitive to error.

Keywords - Morphological Image Processing (MIP), Image segmentation, K- means, Morphological operations, Fuzzy.

## I. INTRODUCTION

The influence and impact of digital images on modem society is tremendous, and image processing is now a critical component in science and technology. The rapid progress in computerized medical image reconstruction, and the associated developments in analysis methods and computer-aided diagnosis, has propelled medical imaging into one of the most important sub-fields in scientific imaging. Imaging is an essential aspect of medical science to visualize the anatomical structures of the human body. Several new complex medical imaging modalities, such as X-ray, magnetic resonance imaging (MRI), and ultrasound, strongly depend on computer technology to generate or display digital images. With computer techniques, multidimensional digital images of physiological structures can be processed and manipulated to help visualize hidden diagnostic features that are otherwise difficult or impossible to identify using planar imaging methods. Image segmentation may be defined as a technique, which partitions a given image into a finite number of non- overlapping regions with respect to some characteristics, such as gray value distribution, texture. Image segmentation may be

defmed as a technique, which partitions a given image into a finite number of non-overlapping regions with respect to some characteristics, such as gray value distribution, texture Distribution, etc. Segmentation of medical images is required for many medical diagnoses like radiation treatment, planning volume visualization of regions of interest (ROI)defining boundary of brain tumor and intra cerebral brain hemorrhage, etc. Many approaches are based on fuzzy logic means and Neural Networks (NN) distribution, etc. segmentation of medical images is required for many medical diagnoses like radiation treatment, planning volume visualization of regions of interest (ROI) defining boundary of brain tumor and intra cerebral brain hemorrhage, etc. Many approaches are based on fuzzy logic means and Neural Networks (NN). Basically, image segmentation methods can be classified into three categories: edge-based methods, region based methods and pixel-based methods. K-means clustering is a key technique in pixel-based methods. In which pixel-based methods based on K-means clustering are simple and the computational complexity is relatively low compared with other region-based or edge-based methods, the application is more practicable. Furthermore, K-means clustering is suitable for biomedical image segmentation as the number of clusters is usually known for images of particular regions of the human anatomy. It is an unsupervised clustering algorithm that classifies the input data points into multiple classes based on their inherent distance from each other. This paper proposes automatic method to find characteristics of Tumor and Cancer cells using Morphological technique. It is a tool to extract our region of interest among the image. Image compression is a technique that removes the pixel redundancy and compresses the image without any loss of information. PEG is the compression standard normally used. But since we need lossless compression we use JPEG 2000 standard for maximum compressing wit no error.

## II. IMAGE ENHANCEMENT

Image enhancement is one of the major research fields in image processing. In any applications such as medical application, military application, media etc., the image enhancement plays an important role. There are many techniques proposed by different authors in order to remove the noise from the image and produce the clear visual of the image. Also, there are many filters and image smoothing

methods available. All these available techniques are designed for particular kind of noises. Recently, neural networks turn to be a very effective tool to support the image enhancement. Neural network is applied in image enhancement because it provides many advantages over the other techniques. Also, neural network can be suitable for removal of all kinds of noises based on its training data. This paper provides survey about some of the techniques applied for image enhancement. This survey deals with the several existing methods for image enhancement using neural networks.

## III. MORPHOLOGY

Morphological processing is constructed with operations on sets of pixels. Binary morphology uses only set membership and is indifferent to the value, such as gray level or color, of a pixel. Morphological image processing relies on the ordering of pixels in an image and many times is applied to binary and gray scale images. Through processes such as erosion, dilation, opening and closing, binary images can be modified to the user's specifications. Binary images are images whose pixels have only two possible intensity values. They are normally displayed as black and white. Numerically, the two values are often 0 for black, and either 1 or 255 for white. Binary images are often produced by thresholding a gray scale or color image, in order to separate an object in the image from the background. The color of the object (usually white) is referred to as the foreground color. The rest (usually black) is referred to as the background color. However, depending on the image which is to be threshold, this polarity might be inverted, and in which case the object is displayed with 0 and the background is with a non-zero value. Some morphological operators assume a certain polarity of the binary input image so that if we process an image with inverse polarity the operator will have the opposite effect. For example, if we apply a closing operator to a black text on white background, the text will be opened.

### A. Morphological Segmentation

This section details the segmentation of mammograms for identifying the tumor in brain. The proposed approach utilizes mathematical morphology operations for the segmentation. The morphological operations are applied on the gray scale images to segment the abnormal regions [14,15]. Erosion and dilation are the two elementary operations in Mathematical Morphology. An aggregation of these two represents the rest of the operations [11]. The symbols, e, o, and, respectively denote the four fundamental binary morphological operations: dilation, erosion, opening, and closing. A function $f(x, y)$ denotes the image, where $(x, y)$ R2 or Z2 , or simply $f$ , and the function $h(x, y)$ , or $h$ will act as the structuring element. The four operations are defmed as follows:

Dilation:

$$( f \oplus h) (x, y) = \sup \{ x - r, y - s) + h(r, s) \}$$
$$(r, s) \in H$$

Erosion:

$$( f \ominus h) (x, y) = \inf \{ f(x + r, y + s) - b(r, s) \}$$
$$(r, s) \in H$$

Opening:

$$f \circ h = ( f \ominus h) \oplus h$$

Closing:

$$f \cdot h = ( f \oplus h) \ominus h$$

Where sup {} and inf {} denote the supremum and infirmum operation, respectively, and H R2 or Z2 is the support of h(x, y).

## IV. OVERVIEW OF METHODOLOGY

Segmentation is a key step towards derivation of semantics from digital images. The goal of segmentation in general is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. However, since there is no general solution to the image segmentation problem, these methods often have to be combined with domain knowledge in order to effectively solve an image segmentation problem for a particular domain. Some of the general-purpose segmentation methods are region growing, histogram evaluation, graph cut and clustering. Region growing starts with a single pixel (current region) and progresses by recursively examining the adjacent pixels. If they are sufficiently similar, they are added to the current region, otherwise a new region is formed.

Histogram evaluation takes place by computation of a color or intensity based histogram from all of the pixels in the image. The peaks and valleys in the histogram are then used to locate different regions in the image. Graph cut models the image into a weighted undirected graph. Each pixel is a node in the graph, and an edge is formed between every pair of pixels. The weight of an edge is a measure of the similarity between the pixels. The image is partitioned into disjoint sets (segments) by removing the edges connecting the segments. Clustering refers to the process of grouping pixels of an image such that pixels which are in the same group (cluster) are similar among them and are dissimilar to the pixels which belong to the other groups (clusters). Let the feature vectors derived from I clustered data. The generalized algorithm initiates k cluster centroids by randomly selecting k feature vectors from X. Later, the feature vectors are grouped into k clusters using a selected distance measure such as Euclidean j [13]. The next step is to recompute the cluster centroids based on their group members and then regroup the feature vectors according to the new cluster centroids. The clustering procedure stops only when all cluster centroids tend to converge. Similarity is measured by distance and defined by an N dimensional feature space. Feature distance calculation differs from spatial distance calculation. Feature distance calculation is based on features such as color or intensity and texture while spatial distance calculation is based on x, y (width, height) coordinates.

Devising an appropriate distance calculation method is an important task since it greatly affects final clustering result. Distance calculation method should be able to effectively compromise among multiple parameters. Furthermore, defming similarity parameters based merely on feature space may lead to clustering errors. For instance, two objects with similar color and texture properties but different shapes will be grouped under the same cluster. Errors of this kind could be fatal in content based retrieval system where shape is of substantial importance in production of appropriate search results. Clustering algorithms may be generally classified into four main categories which are hierarchical, overlapping and exclusive [3].Hierarchical clustering algorithms are based on union between two nearest clusters. They start by setting every pixel as a cluster and progress until final result (few desired clusters) reached [3]. Overlapping clustering algorithms are based on fuzzy sets. Each pixel may belong to two or more clusters with different degrees of membership. Final result is produced either in a ranked manner or by selecting an appropriate degree of membership for each pixel. Exclusive clustering algorithms exclusively group pixels, such that if a pixel belongs to a particular cluster then it could not belong to any other cluster. K-means is an instance of exclusive clustering algorithms and is the backbone of this paper's methodology. K means algorithm starts clustering by determining k initial central points, either at random or using some heuristic data. It then groups each image pixel under the central point it is closest to. Next, it calculates new central points by averaging the pixels grouped under each central point. The two former algorithmic steps are repeated alternately until convergence (central points no longer change by averaging). The limitations of K-means clustering are many iterative rounds may be required. This work strives to reduce that limitation.
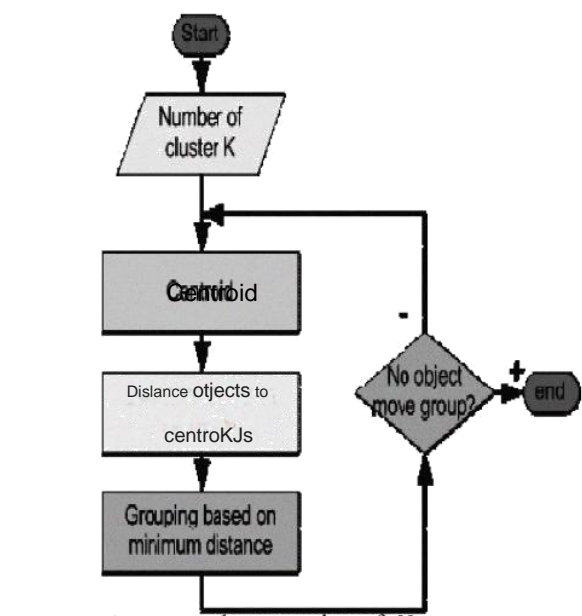


Figure I. Explains the operation of K-means algorithm

V. IMAGE COMPRESSION

Compression techniques are used to reduce the redundant information in the image data in order to facilitate the storage, transmission and distribution of images (e.g. GIF, TIFF, PNG, JPEG). JPEG offers an excellent quality at high and mid bit- rates. However, the quality is unacceptable at low bit-rates (e.g. below0.25 bpp). JPEG cannot provide a superior performance at lossless and lossy compression in a single code stream. the current JPEG standard provides some resynchronization markers, but the quality still degrades when bit-errors are encountered. JPEG was optimized for natural images. Its performance on computer generated images and bilevel (text) images are poor.

A. K-means based segmentation

Currently the amounts of data stored in databases (online and offline) are so huge that create a crucial need for effective and speedy data analysis methods. Cluster analysis is one of the primary data analysis tasks that helps in interpretation and understanding of natural grouping or structure in a dataset.

K means clustering [6,7] is the most widely used and studied method among clustering formulations that are based on minimizing a formal objective function . Modifications to K- means clustering method that makes it faster and more efficient are proposed. The main argument of the proposed modifications is on the reduction of intensive distance computation that takes place at each run (iteration) of K-means algorithm between each data point and all cluster centers. To reduce the intensive distance computation, a simple mechanism by which, at each iteration, the distance between each data point and the cluster nearest to it is computed and recorded in a data structure is suggested. Thus, on the following iteration(s) the distance between each data point and its previous nearest cluster is recomputed.
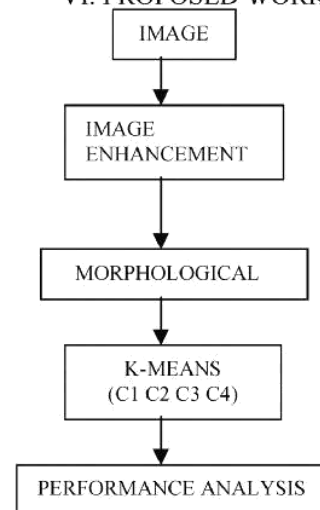


Figure.2. Block Diagram of the proposed method

Figure.2 explains the flowchart of the proposed method. In the proposed method, we combine segmentation and K-means clustering. A brain Image consists of four regions i.e. gray matter (GM), white matter (WM), cerebra spinal fluid (CSF) and background. These regions can be considered as four different classes. Therefore, an input image needs to be divided into these four classes. In order to avoid the chances of mis- classification, the outer elliptical shaped object should be removed. After the enhancement of image morphological process is carried out to extract the required region. Next by implementing K-means with clusters exact result is produced.

Figure.3 shows an image of the brain with skull seen as an outer elliptical ring. In figureA, this elliptical ring is removed and we are left with only soft tissues. This is done by employing the following morphological function, i.e. erosion and dilation.
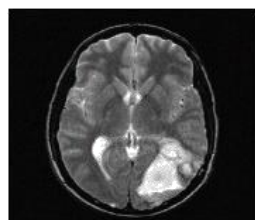


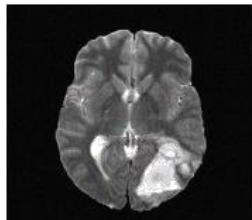Figure. 3 Image with
Outer Ring (Skull)
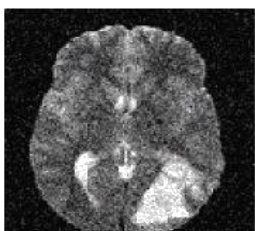


Figure.4 Removing Skull
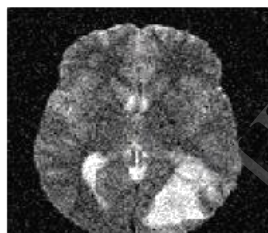Tissues



Figure.S Noisy image



Figure.6 Enhanced image

In Figure.5, the noisy image is shown, which is enhanced by removing noise in the image by a suitable filter. The enhanced image is shown in figure .6. Due to unsupervised nature of the approach, the proposed system is efficient and is less error sensitive. It can be deduced from the results that unsupervised segmentation methods are better than the supervised segmentation methods. Because for using supervised segmentation method a lot of preprocessing is needed. More importantly, the supervised segmentation method requires considerable amount of training and testing data which comparatively complicates the process. Whereas, this study can be applied to the minimal amount of data with reliable results. However, it may be noted that, the use of K-Means clustering method is fairly simple when compared with frequently used fuzzy clustering methods. The Figure.7 shows the segmentation using the proposed work. In this system brain tumors have been segmented with the help of K-means algorithm. The execution time for K-means clustering was less compared to the other clustering methods. Regarding the
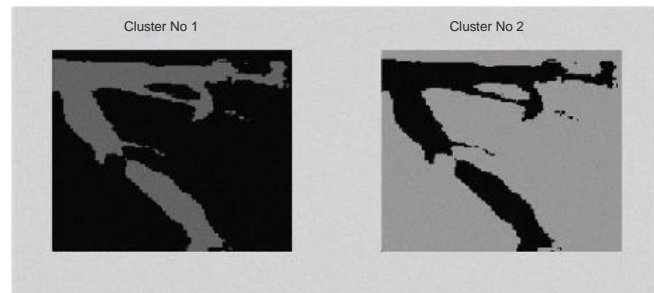


Figure.7 Segmentation using improved K- Means (Proposed work)

number of tumor pixels, K-means clustering gave a better result than the other methods. The clustering algorithm was tested with a database of 100 MRI brain images. K-means clustering achieved about 95% result.

JPEG2000 is a new compression standard for still images intended to overcome the shortcomings of the existing JPEG standard. JPEG2000 makes use of the wavelet and sub-band technologies. Some of the markets targeted by the JPEG2000 standard are Internet, printing, digital photography, remote sensing, mobile, digital libraries and E-commerce. the standard provides lossy compression with a superior performance at low bit-rates. It also provides lossless compression with progressive decoding. Applications such as digital libraries/databases and medical imagery can benefit from this feature. the standard incorporate a set of error resilient tools to make the bit-stream more robust to transmission errors. In this mode, regions of interest (ROI's) can be defined. These ROI's can be encoded and transmitted with better quality than the rest of the image .

## VII. RESULTS & DISCUSSION

The tumor affected cells are found out by applying K-means Clustering algorithm using Matlab Simulator. Same process can be applied for brain cell without tumor. The figure.8. shows the brain Image with tumor cells. Brain tumor

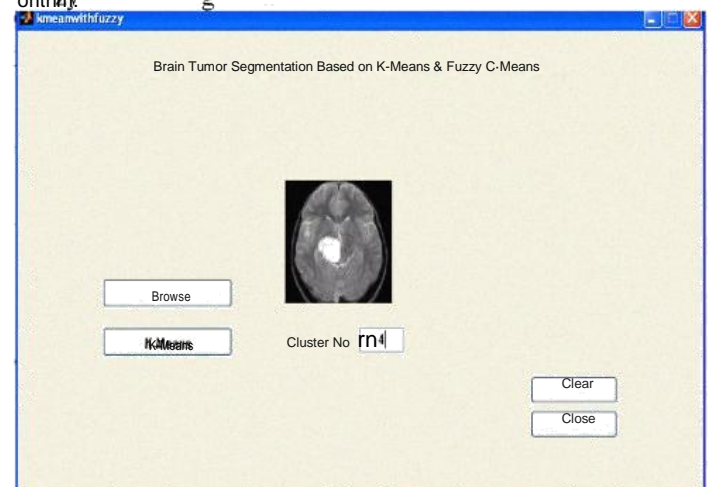segmentation is takes place after applying the K-means & Fuz c-means al orithm.
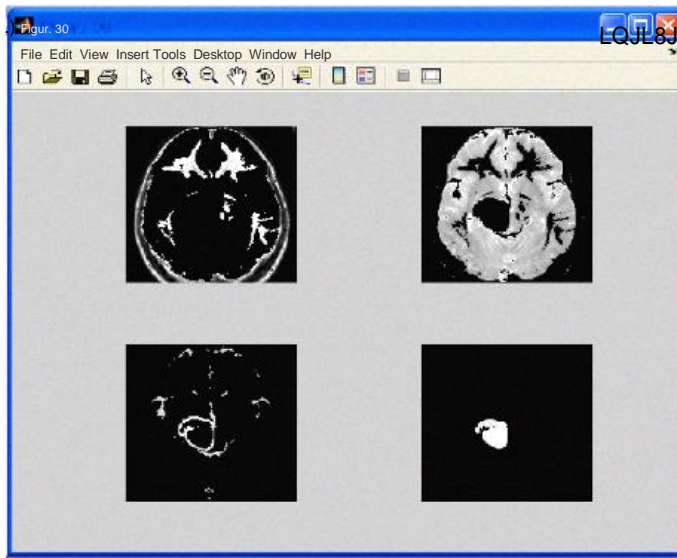


Figure.S. Brain Image with tumor

Figure.9. Clustering of Brain Image

Thus the tumor cells are identified in the 4th cluster as shown in Figure.9.

## VIII. CONCLUSION

Thus in our proposed method, we combine segmentation and K-means clustering. It can be deduced from the results that unsupervised segmentation methods are better than the supervised segmentation methods. Because for using supervised segmentation method a lot of preprocessing is needed. More importantly, the supervised segmentation method requires considerable amount of training and testing data which comparatively complicates the process. Whereas, this study can be applied to the minimal amount of data with reliable results. However, it may be noted that, the use of K- Means clustering method is fairly simple when compared with frequently used fuzzy clustering methods. Also in proposed work JPEG2000 is a new compression standard for still images intended to overcome the shortcomings of the existing JPEG standard. It also provides lossless compression with progressive decoding. Applications such as digital libraries/databases and medical imagery can benefit from this feature. The standard incorporates a set of error resilient tools to make the bit-stream more robust to transmission errors. In this mode, regions of interest (ROJ's) can be defined. These ROJ's can be encoded and transmitted with better quality than the rest of the image .A K-means based segmentation process to detect Brain tumor and cancer cells was implemented. Less execution time is achieved when compared with other clustering methods due to less number of iterations. Maximum lossless compression is also achieved using new standard and transformations. The proposed system is efficient and is less error sensitive.

## REFERENCES

[I]. M. Masroor Ahmed and Dzulkifli Bin Mohamad "Segmentation of Brain MR Images for Tumor Extraction by Combining Kmeans Clustering and Perona-Malik Anisotropic Diffusion Model" International Journal of Image Processing, Volume (2) : Issue(I ),2010.

[2] Arash Azim Zadeh Irani and Bahari Belaton "A K-means Based Generic Segmentation System" 2009 Sixth International Conference o n C o m p u ter Grap h i c s , I m a g i n g a n d V i s u a l i z a t i o n

[3] Ming-Ni Wu, Chia-Chen Lin and Chin-Chen ChangI3 "Brain Tumor Detection Using Color-Based K-Means Clustering Segmentation" International Conference on Computer Graphics,2009

[4] S.Sathish Kumar, M.Moorthi, M.Madhu3and Dr.R.Amutha "An Improved method of Segmentation using Fuzzy-Neuro Logic" Second International Conference on Computer Research and development,2009

[5] S. Murugavalli and V. Rajamani" An improved Implementation of Brain tumor Detection Using Segmentation Based on Neuro Fuzzy Technique", Journal of Computer Science 3 ,2007 october.

[6] Kharrat, A. Benamrane, N. Ben Messaoud, M. Abid, Comput. & Embedded Syst. Lab. (CES), Nat. Eng. Sch. of Sf ax, Sfax, Tunisia, Nov. 2009 .

[8] M. Schmidt, I. Levner, R. Greiner, A. Murtha, and A. Bistritz. segmenting Brain Tumors using Alignment-Based Features". International Conference on Machine Learning and Applications,

Los Angeles, December 2005.

[9] S.R. Kannan , Jan. 2005 "Segmentation of MRI using new Unsupervised Fuzzy C means algorithm" ICGST-GVIP Journal, Volume 5, Issue2.

[10] S ..Saheb Baha and Dr.K.Satya Prasad, "Morphological image Processing in Bio- Medical Application", Proceedings of PCEA-IFTOMM- International conference - PICA- 2006, held on 11th to 14th of July at Nagapur and NCBME-2006 - National conference on Bio-Medical Engineering held on 28th to 29th of March at Mumbai.

[II] S.Saheb Baha and Dr. K.Satya Prasad,"Automatic detection of Hard Exudates in Diabetic Retinopathy using Morphological segmentation and Fuzzy Logic" in [JCSNS International Journal of Computer Science and Network Security, VOL.8 No.12, December 2008.

[12] Juraj Horvath, 2006 "Image Segmentation using Fuzzy C-means" SAMI 2006.

[13] Lei jiang, Wenhui Yang , "A modified Fuzzy C-means algorithm for Magnetic Resonance Images" Proc. VIIth Digital Image Computing: Techniques and Applications, 2003.

[14 ] DIGTAL IMAGE PROCESSING,- Rafael C. Gonzalez, -Richard E. Woods, ADDISON-WESLEY, An imprint of Pearson Education, 1st Edition.

[15] J. Goldberger, S. Gordon, H. Greenspan, "Unsupervised Image-Set Clustering Using an Information Theoretic Framework", IEEE Transactions on Image Processing, Vol. IS, No. 2, February 2006,

[16] Fuzzy Logic with Engineering Applications, International edition 1997,-Timothy J. Ross , McGRAW-HILL International Edition.

[17] Maj6s, M. Julia-Sape, J. Alonso, M. Serrallonga, C. Aguilera, 1. Acebes, C. Arus and 1. GilL "Brain Tumor Classification by Proton MR Spectroscopy: Comparison of Diagnostic Accuracy at Short and Long TE". American Journal of Neuroradiology 25:1696-1704, November- December 2004.

[18] Haralick, R.M. and L.G. Shapiro, 1985. "Survey ,image segmentation techniques", Computer Vision, Graphics Image Process, 29: 100-132[9]

[19] Christophe Ambroise,Fouad Badrean,Sylvie Thira,Genive Seze, Hierarchical Clustering of SOM for Cloud Classification, neuro Computing, Volume 30, January 2000, Pages 47-52

[20] Y. Zhang, "A survey on evaluation methods for image segmentation," Pattern Recognition, vol. 29, pp. 1335-1346, 1996.