

# Implication Of Data Mining Techniques Towards Image Segmentation And Classification

<sup>1</sup>Harshita Mishra, <sup>2</sup>Puneet Sharma, <sup>3</sup>Deepak Arora

*Department of Computer Science and Engineering, ASET*

*Amity University, Lucknow, Uttar Pradesh, India*

*mharshu1995@gmail.com, psharma9@lko.amity.edu, darora@lko.amity.edu*

**Abstract-** Image enhancement and image segmentation are highly researched topics of Digital Image Processing (DIP). In image segmentation the image analysis is done in various steps and has important role in digital image processing. The fundamental methods and step required to segment the image for gathering information about image using image segmentation will be studied in this work. The most important method for analyzing image is segmentation. In this research work various methods for segmentation will be studied such as-edge detection method, thresholding method, region growing method, clustering method. Similarity and discontinuity are the most important properties of image segmentation algorithm.

In the present time the highly growing area in the field of data mining and in the computer vision is image classification. The main purpose of the image classification is to remove space between pixels in the image.

**Index Terms** - Image Enhancement, Image Segmentation, image classification, Digital Image Processing, Thresholding, Clustering

## 1. INTRODUCTION

In recent research topics such as digital image processing has various implementations in different fields i.e. medication, remote sensation, production of movies and videos. Many inventions are made in the regarding processing of image in the field of image segmentation.

## 2. IMAGE SEGMENTATION

The first step in compression of image is known as image segmentation. The abilities of image segmentation process are rate of performance, match of the shapes and inter connection between original shape and segmented shape. The method to identify and isolate the area of any digitalized image that refers particular structure is called segmentation. Different features such as color, texture of image are responsible for segmentation to happen.

## 3. IMAGE SEGMENTATION ALGORITHM

Segmentation related to the image this algorithm has been created for the segmentation of digitally processed images. The segmentation of image is on the basis of characteristics that are discontinuity of the image and similarity in the image. The property of the image that is discontinuity of the image depends on the partition of the image and subdivision of the image. This is done after observing upcoming changes in the strength or grey level of the image. In this method our interest mainly focuses on

identification of isolated points, lines and edges. The pixels of the image that are similar are placed together in a clutter. This process is done in the similarity based cluttering of the pixels of the image. The algorithms included in this method are threshold, region based grow and split and merge. The segments of image can be done by the use of image segmentation algorithm. Discontinuity and similarity are properties on which basis image segmentation algorithm is performed. When there is any sudden change in grey level of any image then segmentation preformed is based on discontinuity.

## 4. CLASSIFICATION OF SEGMENTATION TECHNIQUES

Image segmentation is categorized as,

- Segmentation by edge detection
- Segmentation by thresholding
- Segmentation based on region

## 5. SEGMENTATION BY EDGE DETECTION

Edge detection is first stage of the image segmentation. In the process of the edge detection we can divide an image in its objects and related background. In the process edge detection the division of the image is done when there is change in the intensity of the image or change in the pixels of the image. For the detection of the edge in the image segmentation process the two main processes are

used i.e. gray level histogram and image gradient. In the edge detection the operator is of two category i.e. first and second order derivative operators. We get reliable results in second order operator. Canny edge detection comes under second derivative operator.



Fig 1- THE REAL IMAGE

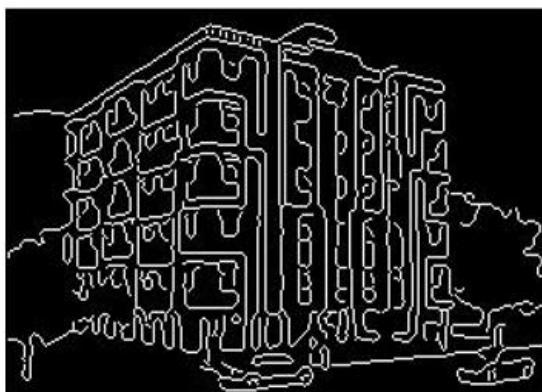


Fig 1.1- IMAGE AFTER APPYING EDGE DETECTION SEGMENTATION

## 6. CANNY EDGE DETECTOR

In canny edge detector the image is firstly captured and then segmentation process is performed. To perform segmentation process firstly image is converted to gray from rgb. The noise present in the image is removed before locating and detecting edges of the image. In canny edge detection we use Gaussian filter that is computed as mask. When image is noise free and after smoothing of the image the next level is to know the strength of the edge by getting gradient of the image. The next required is comparatively correct immensity of gradient of edge strength at each point. Then, the comparatively correct immensity of gradient of the edge durability is known for every point by knowing the gradient in the x direction column and y direction row. After Finding the edge strength, edge direction using the

gradient of x and y directions is found. The tracking of the edge is done by the use of the non-maxima restriction. This process is done in the direction of the edges after that the value of pixel is restricted to zero. The restricted pixel is not judged as edge of the image. The outcome will be a line that will be quite thin in appearance. The outcome of this process will be in the form of a very thin line. The removal of the breaking is done by the use of hysteresis. The breakup of the edge outline is done by breaking process. The requirement of this process is due to the fluctuation in the threshold value from high to low. Thus an image is segmented using edge detection.

## 7. SEGMENTATION BY THRESHOLDING

The simplest process to segment the image is done on the basis of the intensity level known as threshold segmentation. The implementation of threshold can be done at global level and local level. In global threshold difference between object and pixels of background by making comparisons between threshold that are selected and binary partition is used for segmentation of the image. Local threshold technique is also known as adaptive threshold technique. In local threshold technique the values of threshold are different all over the image that depends on the local characteristic of the segmented image. In the histogram threshold the process of segment is done of the image. In threshold segmentation there is requirement of the various preprocessing and post processing techniques. At present times various threshold segmentation algorithm are mean algorithm, P-tile algorithm, histogram algorithm, maximization of edge algorithm and visualization algorithm.



Fig 2-THE REAL IMAGE



Fig 2.1- IMAGE AFTER APPYING THRESHOLD SEGMENTATION

## 8. SEGMENTATION BY REGION BASED

In the segmentation process by use of the region growing the pixels are grouped together in the same group. In region based segmentation algorithm the Thresholding technique is included. The part of the image on which segmentation process is applied, that part is covered or closed part of the image. The region based image segmentation process is also known as similarity based image segmentation process. There would not be any kind of problem of the gaps in the region based image segmentation due to lost edge pixels. All boundaries of the image in this segmentation process of image. When all the changes in the image are observed such as change in color and change in texture or form of the image all edge flow are transformed into vector form. This transformation helps in detection of edge for upcoming segmentation process of the image.



Fig 3.1- IMAGE AFTER APPYING REGION BASED SEGMENTATION

## 9. DATA MINING TECHNIQUES-

The discovery of the invisible information from large amount data i.e. database is the most important aim of data mining. Descriptive and predictive data mining is two classifications for the data mining. Descriptive data mining is used to produce different patterns which help in knowing important connection between different information. In the predictive data mining particular feature that are identified are valued based on the related features. We require different target to be achieved to get all required aims of the data mining. Different targets are-

- Association Rule
- Clustering
- Classification

### 9.1 Association Rule

Statements such as if then statement are includes in the association rule. This is helpful to specific connection among different data items in the huge amount of the data sets related to the different database. There are various applications of the association rule.

### 9.2 Clustering

The grouping of the objects is clustering. The clustering process is done on the basis of the data extracted that explains different objects and connection between them. The concept of the clustering is to place same objects in same group and different object are grouped together.

### 9.3 Classification

The technique of the data mining that is used to identify different groups of the data items is called classification. In this technique data set are trained are set as input in the process of the classification and their label is well known. The formation of the frames in the classification is done by the help of the well-known dataset. This is known as supervised learning. The well known methods of the classification are-



Fig 3- THE REAL IMAGE

- Neural network
- Decision tree
- Naïve bayes classification
- Nearest neighbor
- Support vector machine

For getting data from the image, image segmentation is used in the classification.

## **10. IMAGE CLASSIFICATION**

The process of obtaining data from image is known as image classification. The thematic map is created from the rectangular part of the image obtained through image segmentation. The two main types of the image segmentation are supervised image classification and unsupervised image classification. The two main characteristics for the image classification process are:

- Classification using Learning
- Classification using feature sets

### **10.1 Classification using Learning**

#### **a) Supervised Image Classification-**

The classification of the image with the selection of the pixels of the image that represent to the class using training sites is called supervised image classification. The training sites are used to form signatures. The signatures that are formed are used for the classification of the image.

#### **b) Unsupervised Image classification-**

The process of image classification in which sample of classes is not required for classification purpose is called unsupervised image classification. The classification is based on the software analysis of the pixels of the image.

#### **10.2- Classification using feature set-**

Different attribute of the information are called feature set. The use of the feature is to set elements to the different classes. In the case of the satellite sensing the features used are different measurement done by the help of the sensor with respect to the various wavelengths in the electromagnetic spectrum that are visible light, infrared rays, microwave rays etc.

## **11. VARIOUS EXAMPLES OF THE IMAGE CLASSIFICATION-**

- **Case of multiple class-** Different characters is recognized of the scanned image or text by the use of the bitmap.
- **Case of two classes:** The differentiation between text and graphics from the scanned image.

## **12. ANALYSIS AND DISCUSSION**

In this paper we have mentioned different algorithm for segmentation of image and the classification of the image. The results given by the edge detection method of the segmentation of the image is better than other algorithms for images that have excellent characteristics as in flowers, rivers and image captured by satellite. Segmentation of image such as threshold algorithm is useful for images that have very less characteristics as in images of face and fruit. The partition of the image is done by single seeded region growing depending on the varying characteristics within the image. In the segmentation of image on the basis of clustering depends on the color variation in the image. The outcome of the segmentation process is more prefects as compared to the result of the clustering algorithm and edge detection algorithm with respect to the image, region growing and thresholding. With the help of the image classification process the extraction of the features is done. For better result of the image classification we can use data mining algorithms.

## **13. CONCLUSION**

Segmentation of image is most favorable for coming future and is globally used segmentation process and highly researched topic at present time. Many issues are related to segmentation of image i.e. similarity in images, various image characteristics, and appearance of image. The paper include various algorithm related to the segmentation of image and discussed briefly in this paper. All important types of segmentation of image are mentioned in this paper. The mentioned types of segmentation of image are applicable to digital processing of the image. Image classification is the most useful application for the separation of the image with the text. The process of the image classification has wide application in the different fields such as medical sciences especially in the heart diseases.

## **REFERENCE**

- [1] Pal, N. R., & Pal, S. K. (1993). A review on image segmentation techniques. *Pattern recognition*, 26(9), 1277-1294.
- [2] Seerha, G. K., & Kaur, R. (2013). Review on recent image segmentation techniques. *International Journal on Computer Science and Engineering*, 5(2), 109.
- [3] Al-Amri, S. S., & Kalyankar, N. V. (2010). Image segmentation by using threshold techniques. *arXiv preprint arXiv:1005.4020*.
- [4] Yogamangalam, R., & Karthikeyan, B. (2013). Segmentation techniques comparison in image

- processing. International Journal of Engineering and Technology (IJET), 5(1), 307-313.
- [5] Khan, A. M., & Ravi, S. (2013). Image segmentation methods: A comparative study.
- [6] Zhang, Y. J. (2001). A review of recent evaluation methods for image segmentation. In Proceedings of the Sixth International Symposium on Signal Processing and its Applications (Cat. No. 01EX467) (Vol. 1, pp. 148-151). IEEE.
- [7] Khan, W. (2013). Image segmentation techniques: A survey. Journal of Image and Graphics, 1(4), 166-170.
- [8] Peli, T., & Malah, D. (1982). A study of edge detection algorithms. Computer graphics and image processing, 20(1), 1-21.
- [9] Ding, L., & Goshtasby, A. (2001). On the Canny edge detector. Pattern Recognition, 34(3), 721-725.
- [10] Singh, T. R., Roy, S., Singh, O. I., Sinam, T., & Singh, K. (2012). A new local adaptive thresholding technique in binarization. arXiv preprint arXiv:1201.5227.
- [11] Mancas, M., Gosselin, B., & Macq, B. (2005, March). Segmentation using a region-growing thresholding. In Image Processing: Algorithms and Systems IV (Vol. 5672, pp. 388-399). International Society for Optics and Photonics.
- [12] Chandhok, C., Chaturvedi, S., & Khurshid, A. A. (2012). An approach to image segmentation using K-means clustering algorithm. International Journal of Information Technology (IJIT), 1(1), 11-17.
- [13] Jensen, J. R., & Lulla, K. (1987). Introductory digital image processing: a remote sensing perspective.
- [14] Boyat, A. K., & Joshi, B. K. (2015). A review paper: Noise models in digital image processing. arXiv preprint arXiv:1505.03489.
- [15] Jensen, J. R., & Lulla, K. (1987). Introductory digital image processing: a remote sensing perspective.
- [16] Ngai, E. W., Xiu, L., & Chau, D. C. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. Expert systems with applications, 36(2), 2592-2602.
- [17] Krishna, K., & Murty, N. M. (1999). Genetic K-means algorithm. IEEE Transactions on Systems Man And Cybernetics-Part B: Cybernetics, 29(3), 433-439.
- [18] Berkhin, P. (2006). A survey of clustering data mining techniques. In Grouping multidimensional data (pp. 25-71). Springer, Berlin, Heidelberg.
- [19] Bishop, C. M. (2006). Pattern recognition and machine learning. Springer.
- [20] Kohavi, R., & Quinlan, J. R. (2002, January). Data mining tasks and methods: Classification: decision-tree discovery. In Handbook of data mining and knowledge discovery (pp. 267-276). Oxford University Press, Inc.
- [21] Rish, I. (2001, August). An empirical study of the naive Bayes classifier. In IJCAI 2001 workshop on empirical methods in artificial intelligence (Vol. 3, No. 22, pp. 41-46).
- [22] Agrawal, R., & Srikant, R. (1994, September). Fast algorithms for mining association rules. In Proc. 20th int. conf. very large data bases, VLDB (Vol. 1215, pp. 487-499).
- [23] Lee, J. S. (1980). Digital image enhancement and noise filtering by use of local statistics. IEEE Transactions on Pattern Analysis & Machine Intelligence, (2), 165-168.
- [24] Agnelli, D., Bollini, A., & Lombardi, L. (2002). Image classification: an evolutionary approach. Pattern Recognition Letters, 23(1-3), 303-309.
- [25] Conci, A., & Castro, E. M. M. (2002). Image mining by content. Expert Systems with Applications, 23(4), 377-383.