ISSN: 0975-8585

Research Journal of Pharmaceutical, Biological and Chemical Sciences

An Effective Image Segmentation Approach using DSETS and DBSCAN Algorithm.

MD Naresh Karthik, I Kathirvelan, and K Srilatha*.

Department of Electronics and Communication Engineering, Sathyabama University, Chennai, Tamilnadu, India

ABSTRACT

Image segmentation, which can do partition of an image into different regions, plays an important role in computer vision, objects recognition, tracking and image analysis. There are a large number of methods that can extract the required foreground from the background. However, most of these methods are solely based on boundary or regional information which has limited the segmentation result to a large extent. Content based image retrieval is the technique to retrieve similar images from a database that are visually similar to a given query image. It is an active and emerging research field in computer vision. In the proposed system, the Interest points based Histogram of Oriented Gradients (HOG) feature descriptor is used to retrieve the relevant images from the database. To improve the retrieval accuracy of the system the Color Moments along with HOG feature descriptor are used in this system. The Interest points are detected using the Harris-corner detector in order to extract the image features. The classifier is used for object detection. After detect the object the Euclidean is used for matching and indexing the features of the query image with the database images. Detection of moving objects in image is the first relevant step of information and background subtraction is a very popular approach for foreground segmentation. For the classification HOG feature is used.

Keywords: HOG, DSETS, DBSCAN, Segmentation, Histogram

Corresponding author

ISSN: 0975-8585

INTRODUCTION

An image is projection of three dimensional scene into a two dimensional projection plane. An image is two dimensional function that represents a measure of some characteristics such as brightness or color of viewed scene. Mathematical representation can be done by analog images for a continuous range of values representing position and intensity .Digital image is composed of picture elements called pixels. Each pixel is a sample of an original image. More samples typically provide more accurate representations of the original image. The intensity of every pixel is variable. In color image systems, a color is typically represented by three or four component intensities such as red, green, and blue, or cyan, magenta, yellow, and black. Compared to analog images digital image plays an important role in daily application such as Satellite Television, Computer Tomography etc., Digital image processing widely used in fields such as medicine, forensics, remote sensing, communication and automobile.

Segmentation

Image segmentation is a process of partitioning a digital image into multiple segments. The main aim of Image segmentation is domain independent partitioning of an image into a set of disjoint region that are visually different, homogeneous and meaningful with respect to some characteristics. Image segmentation is typically used to locate objects and boundaries (lines, curve) in imagesSiva Kumar. B, K.Srilatha et al [5]. The process of image segmentation is to divide the images into regions. Image segmentation is a major part in digital image processing. Image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain characteristics. There are two properties in image segmentation they are discontinuity and homogeneity. Methods based on discontinuities are called as boundary based methods and methods based on homogeneity are called as region based methods. The main objective of image segmentation is to cluster pixels and salient image regions. There are two segmentation they are local segmentation and global segmentation. The process of segmenting sub-images which are small windows on a whole image is called as Local segmentation. The process of segmenting the whole image is called Global segmentation. The number of pixels used in local segmentation is very much lower than global segmentation. The main objective of image segmentation is to partition an image into a meaningful region with respect to particular application.

Clustering

Clustering is process of grouping a set of objects in the same group are more similar to each other and dissimilar to those in the other group. The requirement that should be satisfied by clustering algorithms are scalability, dealing with different types quality, ability to deal with noise, interpretability and usability. Clustering of data is a method in which large set of data's are grouped into clusters of smaller sets of similar data. A clustering algorithm endeavor to find natural group of components based on some resemblance. The centroid of group of data sets are found by the clustering algorithm Chun-Yan Yu et al [4]. To evaluate cluster membership most algorithm evaluate the distance between cluster centroid and points. The clustering algorithm's output is basically a statistical description of cluster centroids with number of components in each cluster. Some image may not have the appropriate keyword to describe them therefore the image search will be difficult. Relevance feedback is one of the solution to overcome this problem. This technique utilizes the user feedback and hence reduces the possible errors and redundancy. This technique uses a Bayesian classifier. Which also deals with positive and negative feedback.

Existing Method:

- Threshold Segmentation
- PCA Process

Thresholding

The simplest way of image segmentation is called the thresholding method. This process is based on a clip-level (or a threshold value) to turn a gray-scale image into a binary image. The key of this process is to select the threshold value (or values when multiple-levels are selected). Several popular ways are used in industry including the maximum entropy method, Otsu's process (maximum variance), and k-means clustering.

ISSN: 0975-8585

Recently, methods have been developed in thresholding computed tomography (CT) images. The key idea is that, unlike Otsu's method, the thresholds are derived from the radiographs instead of the (reconstructed) image.

Principal Component analysis

PCA is a mathematical procedure that uses aquadratical transformation to convert a set of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. The number of principal components is less than to the number of original variables. This transformation is proposed in such a way that the first PCA has the largest possible variance and each succeeding component in turn has the highest variance under the constraint that it be quadratical to the preceding components. PCA are guaranteed to be independent only if the data are set is jointly normally distributed. PCA is sensitive to the relative scaling of the original variables. Depending on the field of application, it is also named the discrete Karhunen–Loève transform (KLT), the Hotelling transform or proper quadratical decomposition.

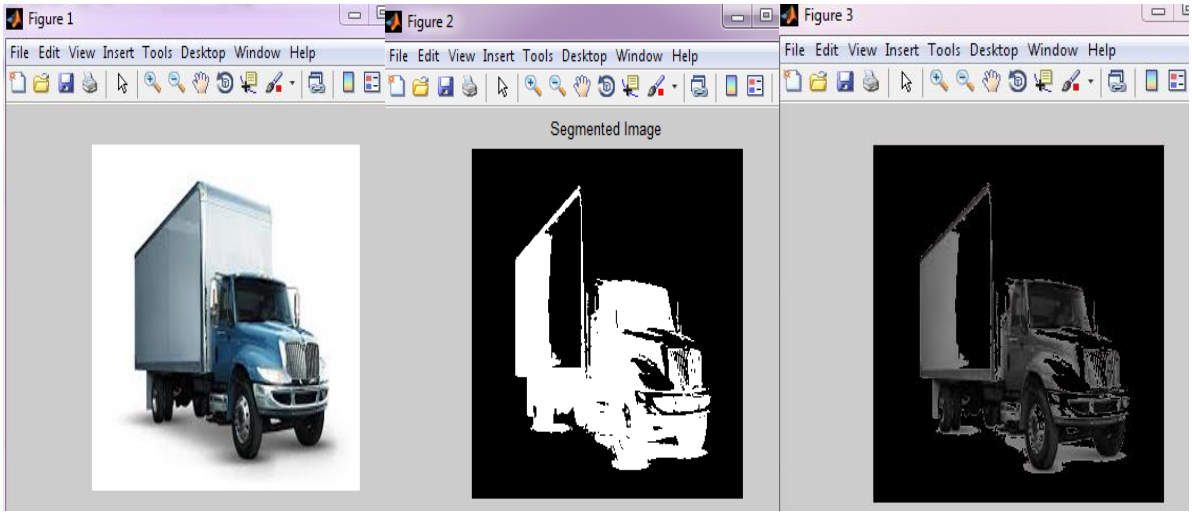

Fig 1: Input image **Fig 2: PCA Process** **Fig 3: Threshold Segmentation**

Proposed Method

Dominant sets

Dominant sets or Dsets which shows the graph theoretic concept of cluster and seeks to extract clusters sequentially. Dominant set is a sun set of data satisfying the constraints of high internal similarity and low external similarity required for a cluster, and therefore can be regarded as a cluster. The objective is to detect the dominant set and remove the data inside from the input data. Repeating process in the remaining data until a stopping criterion is satisfied we will get a partition of input data where each part corresponds to a cluster and number of clusters is determined automatically. The clustering process can be accomplished in a game theoretic framework to obtain soft clustering results and the numbers of clusters is also determined automatically. Dominant sets group similar data's only and the noise will be left unclustered. It uses pairwise similarity matrix of data as input and has no requirement as to the representation form of data, and works well with the asymmetric and negative similarity Melissa.S, K.Srilatha, et al [10]. DSets clustering has found successful applications in such various domains as image segmentation, object detection, bioinformatics , human activity analysis and object classification and some other applications.

DBSCAN

DBSCAN is one of the most commonly used density based clustering algorithm. DBSCAN is able to generate clusters of arbitrary shapes efficiently. DBSCAN uses neighborhood radius Eps and the minimum number of points in the neighborhood MinPts to classify the points into the core points in cluster, border points in clusters and outliers. One point is called a core point ifit have at least MinPts points in its Eps-

neighborhood .If one non-core point is reachable from a core point, it is aborder point. Otherwise, the non-core point is regarded as an outlier. DBSCAN accomplishes the clustering process by extracting cluster sequentially. Starting from an arbitrary point *p*unvisited, DBSCAN retrieves all points in the Eps-neighborhood of *p*. If *p* is a core point, all the points in its Eps-neighborhood are regarded as in a cluster. For each point in the cluster, we find all the density-reachable points and add them into the cluster. Repeating this process until no new density-reachable points are found, we obtain a finalcluster. If *p* is not a core point, we move to the next pointJianHou et al [1].

Histogram of Oriented Gradients (HOG) features

HOG is a standard image feature used, among others, in object detection and deform object detection. It perish the image into square cells of a given size compute a histogram of oriented gradient in each cell, and then renormalizes the cells by looking into adjacent blocks.

Preprocessing

Firstly, video frames captured from a camera are input to the background Preprocessing stages are used for filtration and to change the raw input video to a processable format.

HOG is a feature descriptor used to detect objects in image processing. The HOG descriptor technique counts occurrences of gradient allocation in localized portions of an image - detection window, or region of interestM. Pavan et al [12].

Block Diagram:

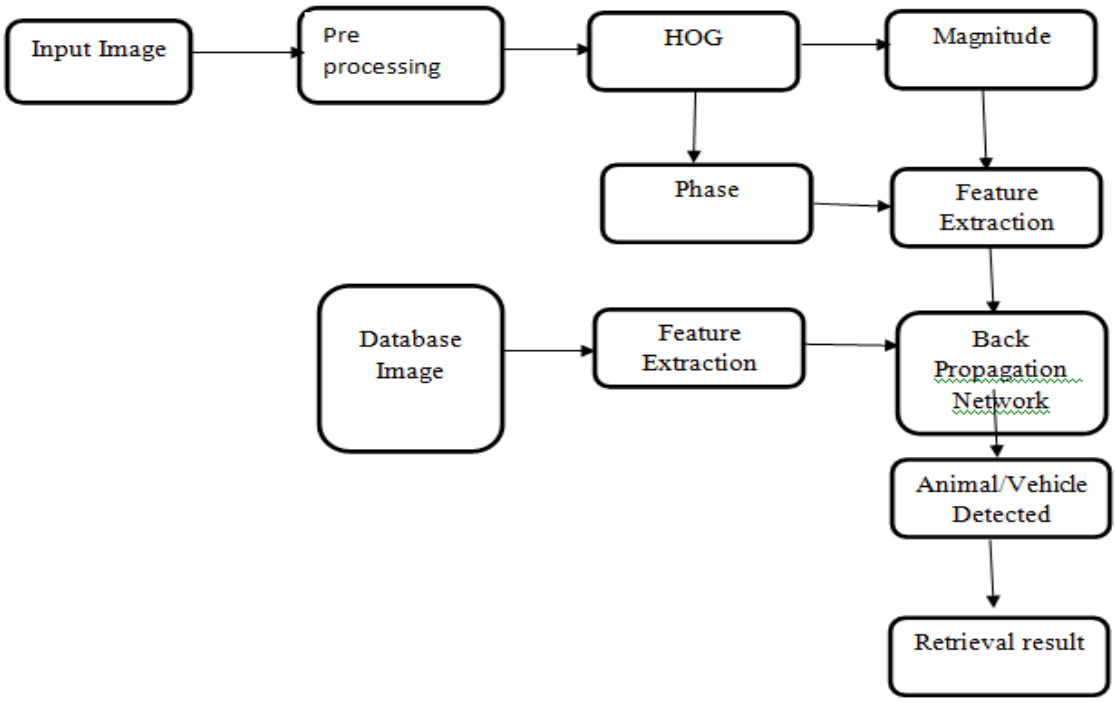

Fig 4: Block Diagram of Proposed method

Implementation of the HOG descriptor algorithm is as follows:

1. Divide the image into small combined regions called cells, and for each cell compute the histogram of gradient directions or edge orientations for the pixels within the cell.
2. Discrete cell into angular bins according to the gradient process.

3. Each cell's pixel contributes larger gradient to its corresponding angular bin.
4. Groups of adjacent cells are considered as free regions called blocks. The grouping of cells into a block is the basis for grouping and normalization of histograms.
5. Normalized group of histograms represents the block histogram. The set of these block histograms represents the descriptor.

Back propagation Neural Network

Artificial neural networks consider classification as one of the most dynamic research and application areas. The disadvantage in using ANN is to find the most appropriate grouping of training and transfer function for classifying the data sets with growing number of features and classified sets. The different combinations of functions and its effect while using ANN as a classifier are studied and the correctness of these functions are analyzed for various kinds of datasets Vinod Kumar Dehariya et al [3].

The real world problems which are represented by multidimensional datasets are taken from medical background. The clustering of these data sets are significant. The data set is divided into training set and testing set and it has no usage in the training process. The results are produced with these datasets and it is used for testing. The training set is taken from 2/3rd of the dataset. This is made through the accuracy achieved through testing against these data sets. Then the network is simulated with the same data. The back propagation algorithm trains the neural network Long Chen et al [8]. Gradient descent method was used to decrease the mean squared error between network output and the actual error rate. The following parameters are considered to measure the efficiency of the network J. Hou et al [14].

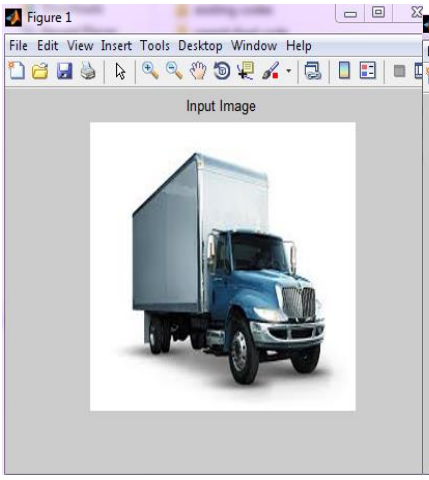

Fig 5: Input image

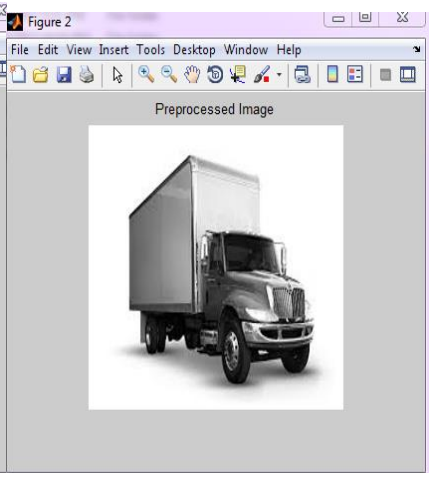

Fig 6: Preprocessed image

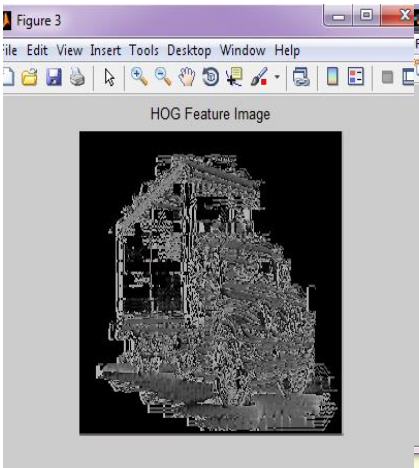

Fig 7: HOG feature image

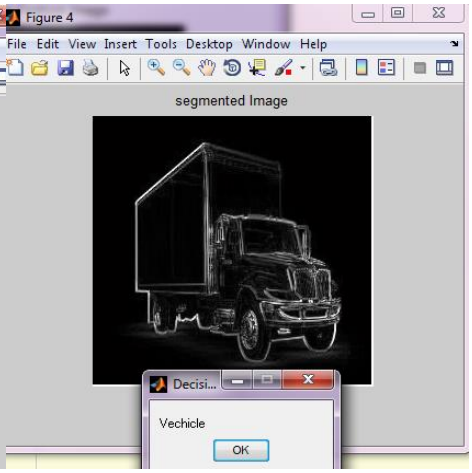

Fig 8: Segmented image

ISSN: 0975-8585

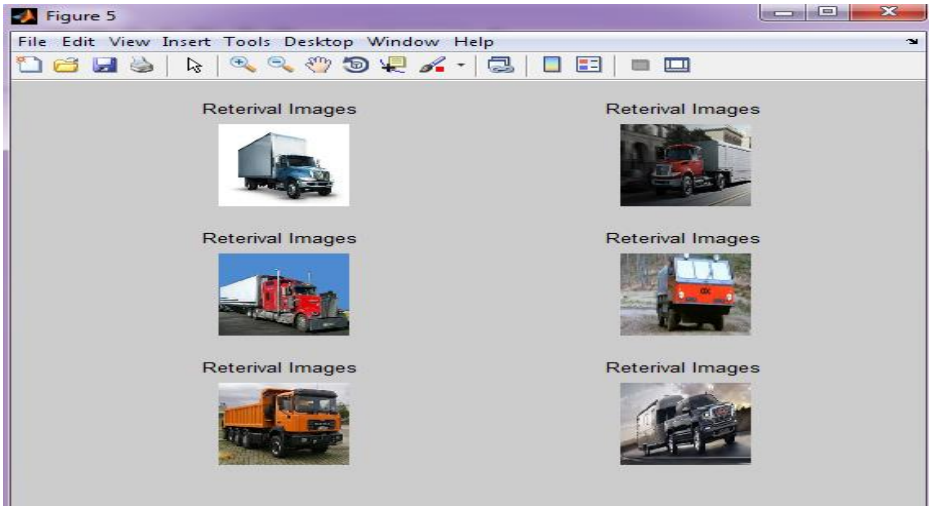

Fig 9: Retrived images

CONCLUSION

In this paper we present the Dsets and DBSCAN as a parameter free clustering algorithms based on the combination of Histogram of Gradient (HOG) and Back Propagation network which is a classifier. We firstly investigate the histogram of gradient before this the input image is preprocessed. In histogram of gradient process we are finding the magnitude and phase of the image and with help of back propagation classifier we can find the images which are similar to the input images.

In this method that the images are retrieved with the help of histogram of gradient and back propagation network. Therefore DSets-DBSCAN algorithm is a density based clustering approach and the minimum density inside a dominant set is used to determine the parameters used in DBSCAN-based cluster extension.

REFERENCES

[1] JianHou, HuijunGao, Xuelong Li, "Dsets – DBSCAN: A parameter free clustering algorithm", IEEE transaction on image processing, vol. 25, No. 7, july 2016.
[2] James C Bezdek, Robert Ehrlich, William Full, "FCM : The Fuzzy C means clustering algorithm", IEEE transaction on image processing, vol.10, No.2-3, pp.191-201,1984.
[3] Vinod Kumar Dehariya, Shailandra Kumar Srivastva, R.c.Jain, "Clustering of image data sets using K means and Fuzzy c means algorithm", IEEE transaction on 2010 International Conference on Computational Intelligence and Communication Networks.
[4] Chun-Yan Yu, Ying Li, Ai-lian Liu, Jing-hong Liu, "A novel modified kernal fuzzy c means clustering algorithm on image segmentation", IEEE International Conference on the 14th IEEE International Conference on Computational Science and Engineering.
[5] Siva Kumar, B., Srilatha, K, "A novel method to segment blood vessels and optic disc in the fundus retinal images", Research Journal of Pharmaceutical, Biological and Chemical Sciences, VOL 7/ISSUE 3/MAY 2016/PP 365-373.
[6] X. Zhu, C. C. Loy, and S. Gong, "Constructing robust affinity graphsfor spectral clustering," in Proc. IEEE Int. Conf. Comput. Vis. PatternRecognit., Jun. 2014, pp. 1450–145.
[7] B. J. Frey and D. Dueck, "Clustering by passing messages between data points," Science, vol. 315, no. 5814, pp. 972–976, Feb. 2007.
[8] Long Chen, Jing Zou and C. L. Philip Chen, "Kernal Non-local shadowed C means for image segmentation", 2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE) July 6-11, 2014.
[9] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in proc.Int.Conf. Knowl Discovery Data Mining, 1996, pp. 226–231.
[10] Melissa.S, Srilatha.K, " A novel approach for pigmented epidermis layer segmentation and

classification" International Journal of Pharmacy and Technology, VOL 8/ISSUE 1/MAR 2016/PP 10449-10458

[11] X. Wang, Y. Tang, S. Masnou, and L. Chen, "A global/local affinitygraph for image segmentation," IEEE Trans. Image Process., vol. 24,no. 4, pp. 1399–1411, Apr. 2015.

[12] M. Pavan and M. Pelillo, "Dominant sets and pairwise clustering," IEEETrans. Pattern Anal. Mach. Intell., vol. 29, no. 1, pp. 167–172, Jan. 2007.

[13] J. Hou, Q. Xia, and N.-M. Qi, "Experimental study on dominant setsclustering," IET Comput.Vis., vol. 9, no. 2, pp. 208–215, 2015.

[14] J. Hou and M. Pelillo, "A simple feature combination method basedon dominant sets," Pattern Recognition., vol. 46, no. 11, pp. 3129–3139,2013.