A project report on

SATELLITE IMAGE SEGMENTATION USING DIFFERENTIAL EVOLUTION

Submitted in partial fulfilment of the requirement for the degree of

Master of Technology

In

Information Systems

Submitted By:

Ankita Goyal

(2K15/ISY/05)

Under the supervision of

Dr. Anil Singh Parihar

Assistant Professor, Department of Information Technology

Delhi Technological University
Formerly Known as Delhi College of Engineering
ShahbadDaulatpur, Main Bawana Road, New Delhi 110042

Month: June Year: 2017

CERTIFICATE

This is to certify that the project work done on "**Satellite Image Segmentation using Differential Evolution**" submitted to Delhi Technological University, Shahbad Daulatpur, Main Bawana Road, New Delhi by "**Ankita Goyal (2K15/ISY/05)**" in partial fulfillment of the requirement for the award of degree of Master of Technology, is a bonafide work carried out by her under my supervision and guidance. This project work is the original one and has not submitted anywhere else for any other degree.

Dr. Anil Singh Parihar

Assistant Professor, IT Department
Delhi Technological University
(Project Guide)

ACKNOWLEDGEMENT

I express my gratitude to my project guide Dr. Anil Singh Parihar, Assistant Professor, Department of Information Technology, Delhi Technological University, for the valuable support and guidance he provided in making this major project. It is my pleasure to record my sincere thanks to my respected guide for his constructive criticism and insight without which the project would not have shaped as it has. I humbly extend my words of gratitude to other faculty members of this department for providing their valuable help and time whenever it was required.

Ankita Goyal

2K15/ISY/05

M.Tech. (Information Systems)

ABSTRACT

Satellite images are multi-spectral, high resolution images and they have important applications in many fields like agriculture, climate studies, landscape, geology, fishery etc. One of the main applications of satellite images is land cover classification. Satellite image segmentation is a difficult problem as different regions occupy different land cover types. Large regions are occupied by some land cover types like rivers, vegetation, etc. while other land types like roads, bridges, etc. occupy small regions. A fuzzy clustering technique based on differential evolution has been proposed that employs type 2 fuzzy systems for membership representation. The proposed segmentation method is a clustering based segmentation method that uses Xei-Beni index for fitness evaluation in which Euclidean distance measure is used and for membership function type-2 fuzzy system is employed. The proposed algorithm require number of clusters to be specified beforehand. Validity of proposed method is being shown by comparing the value of cluster validity index, silhouette index with well-known algorithms: Fuzzy c-means, variable length genetic algorithm, Type-1 differential evolution and multi-objective genetic algorithm.

TABLE OF CONTENTS

CERTIFICATE…………………………………………………………......ii

ACKNOWLEDGEMENT………………………………………….........iii

ABSTRACT……………………………………………………………....iv

LIST OF FIGURES AND TABLES…………………………………....vii

CHAPTER 1 INTRODUCTION..8

1.1 Image Segmentation..8

1.2 Motivation..9

1.3 Challenges..9

1.4 Objective of Thesis...10

1.5 Outline of Thesis..10

CHAPTER 2 LITERATURE SURVEY...........…………………….....11

2.1 Edge Based Segmentation...11

2.2 Thresholding Based Segmentation...11

2.3 Region Based Segmentation...13

2.4 Clustering Based Segmentation..14

CHAPTER 3 CLUSTERING BASED SEGMENTATION TECHNIQUES..16

3.1 K-Means Clustering Algorithm...16

3.2 Fuzzy C-Means Clustering..17

3.3 Variable Length Genetic Algorithm..18

3.4 Multi-objective Genetic Clustering..20

3.5 Multi-objective Fuzzy Clustering And SVM...21

CHAPTER 4 TYPE-2 FUZZY SYSTEMS...24

4.1 Type-2 Fuzzy System...24

4.2 Type-2 Fuzzy Logic System..24

CHAPTER 5 DIFFERENTIAL EVOLUTION..26

5.1 Introduction..26

5.2 Algorithmic Steps...28

5.2 Advantages Of Differential Evolution..28

CHAPTER 6 PROPOSED APPROACH..29

6.1 The Approach..29

6.2 Flowchart...32

CHAPTER 7 EXPERIMENTAL RESULTS AND ANALYSIS.............33

CHAPTER 8 CONCLUSION...35

LIST OF REFERENCES...36

LIST OF FIGURES AND TABLES

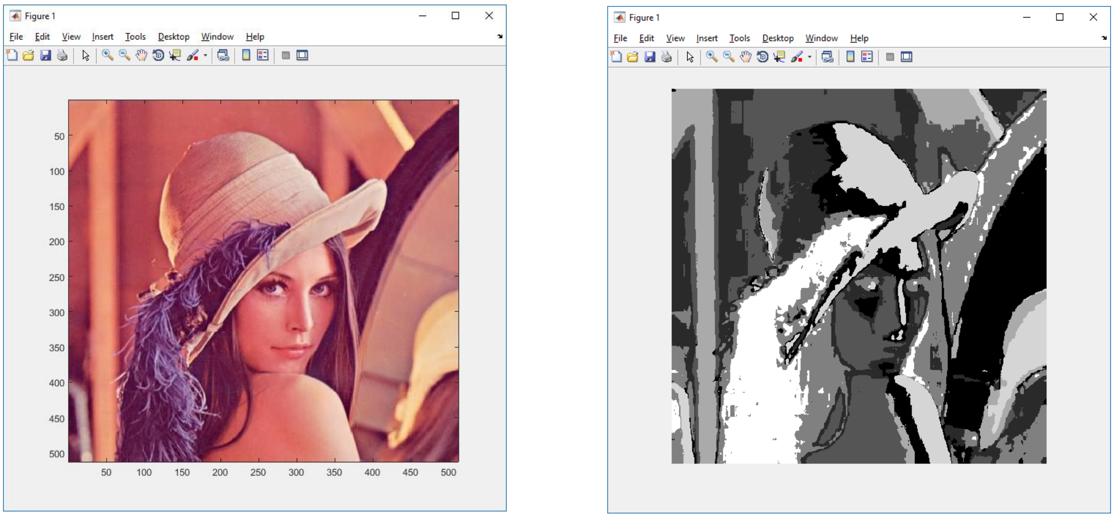
S.NO.	NAME OF FIGURE/TABLE	PAGE NO.
1.1	Image Segmentation Example	8
1.2	Categorization of Image Segmentation Techniques	9
2.1	First Step of Edge Based Segmentation: Edge Detection	11
2.2	Thresholding Based Segmentation: Two Segments	12
1.3	Multi-Level Thresholding Based Segmentation: Three Segments	12
3.1	K-Means Clustering: The Iterative process	16
3.2	SVM Classifier	22
4.1	Type 2 Fuzzy Set	24
4.2	Type-2 Fuzzy Logic System	25
5.1	General Evolutionary Algorithm Procedure	26
6.1	Flowchart of Proposed Algorithm	32
7.1	Experiment Results of Various Algorithms	33
7.2	Comparison of Silhouette Index	37

CHAPTER 1: INTRODUCTION

1.1 Image Segmentation

In Image Segmentation, the image is partitioned into distinct regions that contain pixels with similar attributes like intensity, texture, etc. The main aim is to convert image into something meaningful that is easy to analyze and comprehend. This plays an important part in processing of images and is an initial step in image analysis. Formally, Image Segmentation can be defined as: "Let F be the set of all pixels and P be a homogeneity or similarity predicate defined on group of connected pixels, then segmentation is a partition of F into a set of connected regions S_1, S_2, S_3, ...,Sn such that following conditions are satisfied:"

1. Partitions cover the entire image i.e. $\cup S_j \ = F.$
2. No subset or partition intersect each other i.e. $S_i \cap S_j \ = \emptyset \ for \ i \neq j.$
3. Homogeneity predicate is satisfied by all partitions i.e.$\forall \ S_i, P(S_i) = true.$
4. Homogeneity predicate is not satisfied by union of adjacent partitions i.e. $if \ S_i \ is \ adjacent \ to \ S_j \ and \ i \ \neq j \ then \ \ P(S_i \cup S_j) \ = false.$



Original Image

Segmented Image

Figure 1.1 Image Segmentation Example

Segmentation techniques can be divided into two classes: contextual and non-contextual on the basis of whether they exploit spatial information of an image. If a segmentation method does not use spatial information of image and partitions the image on the basis of attributes like color, intensity, etc. then it is a non-contextual segmentation method. On the other hand, if segmentation method employs spatial information of image, then it is said to be a contextual segmentation method. Segmentation techniques can also be categorized into two categories on basis whether technique is based on detecting some similarity or some discontinuity among the pixels of the image. Categorization is depicted by following figure:

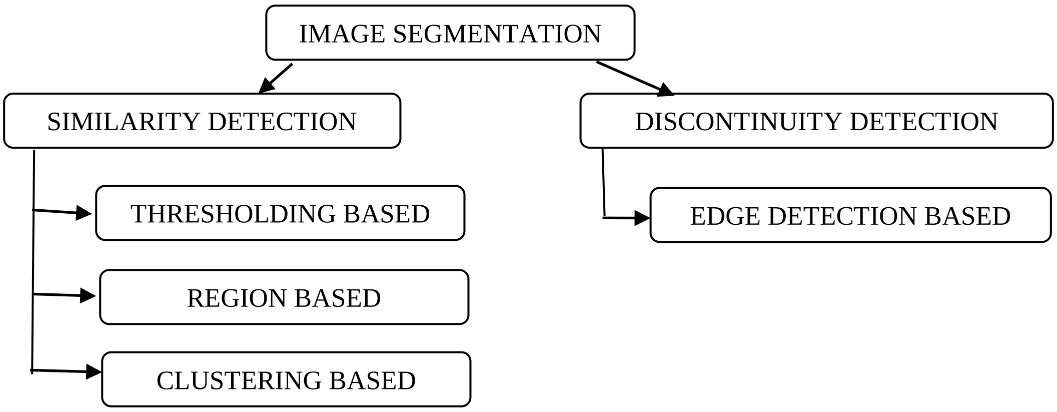


Figure 1.2 Categorization of Image Segmentation Techniques

Image segmentation has applications in many areas like medical imaging, object detection, recognition tasks, machine vision, video surveillance etc.

1.2 Motivation

Satellite images are multi-spectral, high resolution images and they have significant applications in many areas like climate studies, agriculture, landscape, geology, fishery etc. One of the main applications of satellite images is land cover classification. Land cover can be defined as “observed physical cover on the surface of earth”. Information obtained from land cover classification is useful for many organizations like state, city, county government agencies, engineering firms, etc. Information is really useful for many land management applications such as urban planning assessment, vacant land identification, pollution impact studies, etc. Satellite image segmentation aims to partition the image into different land covers. A pixel in a satellite image may not belong to single land cover and thus a lot of imprecision and inaccuracy is associated with satellite images. Due to this uncertainty and large number of dimensional space, fuzzy clustering techniques present an appropriate method for segmentation.

1.3 Challenges

Satellite image segmentation is a difficult problem as different regions occupy different land cover types. Large regions are occupied by some land cover types like rivers, vegetation, etc. while other land types like roads, bridges, etc. occupy small regions. One cannot determine the number of clusters present in the image beforehand. Hughes phenomenon or the curse of dimensionality due to small number of training data available as compared to high number of spectral bands make the problem of segmentation more tough to solve. High correlation between the spectral bands, corruption by atmospheric effects and high computational load pose further challenges to the segmentation task.

1.4 Objective of Thesis

The objective of the thesis is to propose a segmentation method for segmenting satellite images that performs efficiently taking into consideration various aspects of satellite images like hughes phenomenon, high correlation between spectral bands, etc. The proposed segmentation method is a clustering based segmentation method that partitions image into different regions which represents land cover map. It uses type 2 fuzzy systems and differential evolution for producing accurate and precise segments. The proposed approach algorithm require number of clusters to be specified beforehand. Validity of proposed method is being shown by comparing the value of cluster validity indexes, silhouette index with well-known algorithms: Fuzzy c-means, variable length genetic algorithm, Type-1 differential evolution and multi-objective genetic algorithm.

1.5 Outline of Thesis

The remainder of this thesis is organized as follows: in chapter 2 a literature survey of various segmentation techniques is presented. In chapter 3, different clustering based segmentation techniques for satellite image segmentation are being discussed. Chapter 4 gives an introduction to type-2 fuzzy systems. In chapter 5, differential evolution algorithm is being described. Chapter 6 defines the newly proposed approach for satellite image segmentation using DE and type-2 fuzzy system. In chapter 7, experimental results are provided and analyzed to show the validity of proposed method. Finally, chapter 8 gives the summary and conclusions.

CHAPTER 2: LITERATURE SURVEY

2.1 Edge Based Segmentation

In this, edges of image are source of information for performing segmentation of the image. Post processing of result obtained after edge detection is required to produce segmentation result. Main algorithmic steps are:

1. Compute image containing all the credible edges of the image to be segmented.
2. Process the above image so that it contains only closed object boundaries.
3. Fill the object boundaries to get the segmentation result

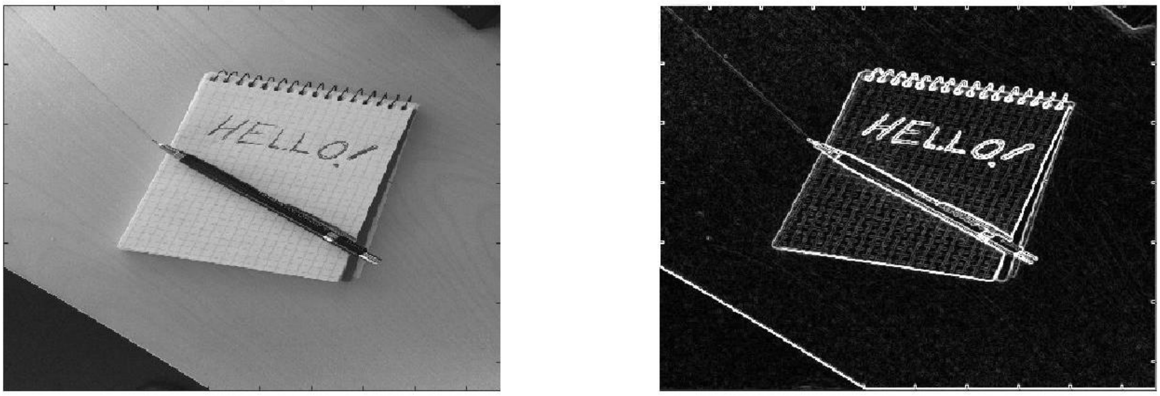


Figure 2.1 First Step of Edge Based Segmentation: Edge Detection

Many edge based segmentation techniques have been proposed by the researchers. In [6], a CNN based edge detection algorithm has been proposed for satellite images. CNN is a connected parallel information processing system. They are similar to neural networks but communication is allowed to take place in between neighboring units. They can efficiently solve real world problem due its parallel processing power. [7] presents a method for ‘multi-scale segmentation’ of satellite images using information obtained from edges. For edge detection, canny edge detector is employed. The proposed method uses half partition structure and basically comprises of three steps: detection of edges, separated pixel merging and calculation of significant features.

2.2 Thresholding Based Segmentation

In this, process of thresholding is performed to segment the image into distinct segments. A ‘threshold value’ is chosen and pixels having intensity less than threshold are turned to ‘0’ and pixels having value greater than threshold are turned to’1’.Thresholding partitions image into 2 segments : pixel having value ‘0’ represents one segment and pixel having value ‘1’ represents other. Thus, it converts a grayscale image into binary image. It is useful for those images that contain light background with dark objects or vice-versa.

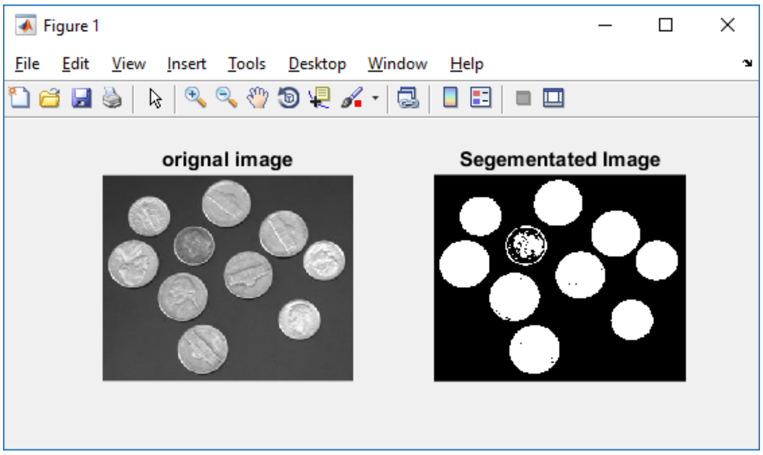


Figure 2.2 Thresholding Based Segmentation: Two Segments

When more than two segments are required multiple threshold values are taken and same procedure is followed as in case of single threshold.

$$g(v) = \left\{ \begin{array}{ccc} 0 & if & v \le t_1 \\ 1 & if & t_1 \le v \le t_2 \\ 2 & if & t_2 \le v \le t_3 \\ \vdots & \vdots & \vdots \\ n-1 & if & t_{n-1} \le v \le t_n \\ n & if & t_n \le v \end{array} \right\} \qquad (1)$$

The above equation will partition the input image into ‘n+1’ segments.

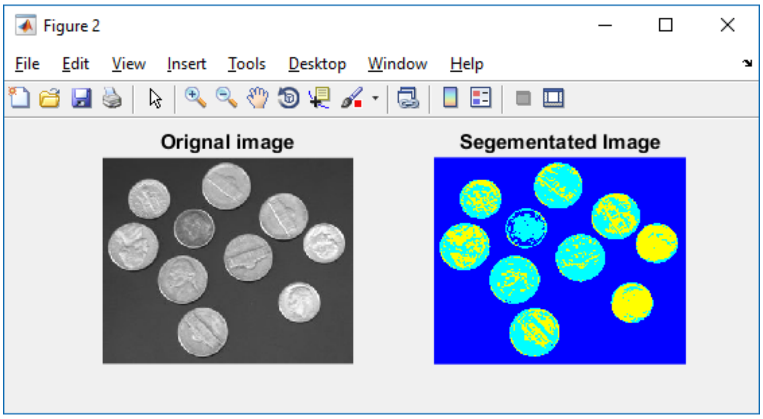


Figure 2.3 Multi-Level Thresholding Based Segmentation: Three Segments

In literature, many thresholding based segmentation techniques have been proposed for segmenting satellite image. In [8] satellite image segmentation is done by performing multi-level thresholding using maximum Rényi entropy (MRE).Firstly, a segmentation map is produced MRE based multi-level thresholding. Then segmentation map and spatial feature of image are used to create kernel that is used to train SVM classifier. SVM produces the final segmentation result. [9] proposed an approach for segmentation using multi-level thresholding based on Chaotic Darwinian Particle Swarm Optimization. In [10], segmentation technique based on cuckoo search algorithm and energy curve has been proposed. In this paper, spatial information needed for thresholding is extracted from energy curve of the image. The energy curve concept has been deployed in the objective function namely, Kapur's entropy. Further, Cuckoo Search is utilized to determine the optimal threshold values on the energy curve by optimizing the objective criterion.

2.3 Region Based Segmentation

Algorithms of region based segmentation is categorized into two categories merging and splitting. The key algorithmic steps region based segmentation using merging are:

1. Obtain initial segmentation i.e. over segmentation of the image to be segmented.
2. Adjacent segments that are similar in one or other aspect are merged to create single segment.
3. Repeat step 2 until no more segments can be merged.

The main step in segmentation using merging is the 'selection of similarity criterion' that will be used to determine whether two segments should be merged or should not be merged. The basic algorithmic steps in segmentation using splitting are:

1. Obtain initial segmentation i.e. under segmentation of the image to be segmented.
2. Split the segment that is not similar in one or other aspect to form multiple segments.
3. Repeat step 2 until no more segments can be split.

Similarly, the main step in segmentation using splitting is the 'selection of dis-similarity criterion' that will be used to determine whether two segments should be split or should not be split.

In literature, many region based segmentation approaches have been proposed for segmentation of satellite. In [11], a segmentation technique using 'active contours model' and 'level set evolution' is proposed. Statistical information present inside and outside of the contour is exploited to formulate region based sign pressure function that will be used to guide the direction of evolution. [12] proposes an enhanced version of Seeded region growing technique for satellite image segmentation. Seeded region growing method require automatic selection of seeds. The location information i.e. latitude, longitude and feature metadata is used as seed, which is superimposed on the satellite image. These seed pixels generate training sites for proposed method. Then supervised classification can be performed. In [13], the proposed method uses the Tree of Shapes representation for region-based classification of remote sensing images.

2.4 Clustering Based Segmentation

Clustering is an unsupervised learning algorithm whose main goal is to organize objects into groups that are similar in one way or another. Different clustering techniques are used in image segmentation. These algorithms are useful for segmenting data sets that have high dimensional space and thus these techniques are suitable for segmenting satellite images that contain several spectral bands.

Many clustering based approaches have been proposed for segmentation of satellite images. Fuzzy C-Means (FCM) clustering [14] uses the principles of fuzzy sets to evolve a partition matrix U while minimizing the measure $J_m = \sum_{k=1}^{K} \sum_{i=1}^{n} u_{ki}{}^{m} D^2(v_k,x_i)$ that shows the compactness of clusters. However, FCM has two major disadvantages: it requires the number of clusters to be specified beforehand and it may get stuck to suboptimal solutions based on the initial configuration of the system. To overcome the above mentioned limitations of FCM, a new approach using the search capability of genetic algorithms to automatically evolve optimum number of partitions of data such that some measure of goodness of the partitions is optimized is used. Genetic Algorithms are randomized search and optimization algorithm that mimics biological evolution process that fittest will survive in the generations to come. VGA proposed in [17] automatically evolves the optimum number of clusters for given image or dataset and produces the fuzzy partitioning of the data. Each chromosome encodes a partitioning of the data that has variable number of clusters and the fitness of the chromosome is computed as a function of cluster validity index, “Xie-Beni”[15].This technique use a single cluster validity measure for computing fitness of chromosome in the population. However, a single cluster validity measure may not be able to produce optimum results for different kinds of data sets with different characteristics. To handle this requirement, in [18] fuzzy partitioning is posed as multi-objective optimization (MOO) problem, where search is performed over a number of conflicting objective functions. The Xie-Beni (XB) index and the fuzzy C-means (FCM) measure (J_m) are being optimized in MOO. In Multi-

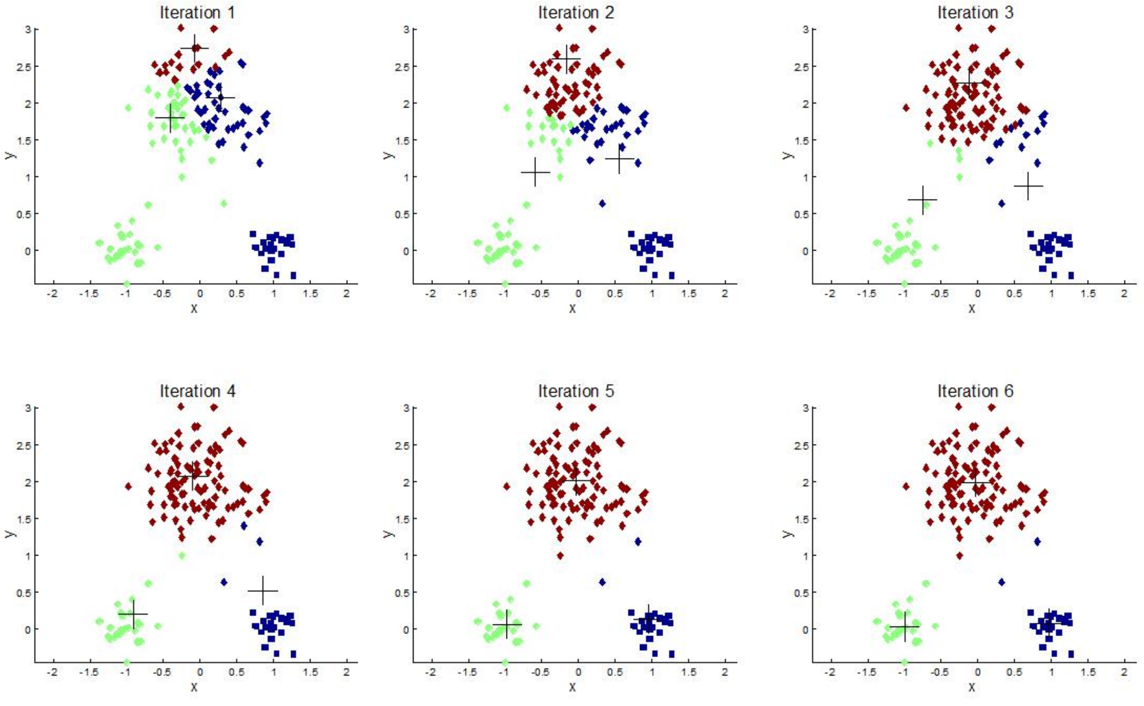
Objective Optimization, the final solution set contains a number of non-dominated Pareto-optimal solutions and it is very difficult to obtain final solution from them. Thus, in [20] SVM is used as a solution to the above mentioned problem. In this approach, a set of non-dominated solutions is obtained by executing NSGA-II [19] on data set. A fuzzy voting is performed on solution set to get high confidence points which becomes training data set where input label is taken from original data set and output label is the cluster center obtained in solution set for the SVM classifier. Remaining points are classified using SVM classifier. [4] proposes a segmentation technique using FCM that uses type-2 fuzzy sets for representing uncertainty.

3.1 K-Means Clustering Algorithm

K-means is an unsupervised learning technique that is used to organize data into clusters where objects similar in some respect belong to the same cluster and objects dissimilar in the same respect belong to different clusters. It is an iterative procedure in which following objective function is being minimized:

$$J = \sum_{j=1}^{K} \sum_{i=1}^{N_j} ||x_i - c_j||^2 \tag{2}$$

where $||x_i - c_j||$ represents the distance between data point $\boldsymbol{x_i}$belonging to cluster j and cluster center $\boldsymbol{c_j}$, N$_j$ represents number of data set points belonging to a specific cluster j and K represents total number of clusters.



Iteration 1

Iteration 2

Iteration 3

Iteration 4

Iteration 5

Iteration 6

Figure 3.1 K-Means Clustering: The Iterative process

Algorithmic steps for K-means are:

1. First select randomly K cluster centers.
2. The distance between each data point and randomly selected centers of cluster is calculated. This distance can be Euclidean distance, manhattans distance, etc.
3. Each data set point is assigned cluster on the basis of distance calculated in previous step. If the distance between data point and specific cluster center is minimum, then that data point is assigned to that particular cluster.

4. New cluster centers are calculated using following equation:

$$c_i = \frac{1}{N_i} \sum_{j=1}^{N_i} x_j, \ for\ 1 \le i \le K \tag{3}$$

5. Repeat steps 2, 3 and 4 till no change in centers of clusters takes place.

3.1.1 Advantages of K-means Clustering Algorithm

1. It is simple to use and easy to understand.
2. It gives good result in case data set is well separated.

3.1.2 Disadvantages of K-means Clustering Algorithm

1. It needs the number of partitions or clusters as one of the inputs.
2. Result depends upon the initial selection of cluster centres.
3. It is unable to partition non-linear data into clusters.

3.2 Fuzzy C-Means Clustering

Fuzzy C-means [14] is a clustering technique in which each point of data set is allowed to belong to more than one cluster, unlike K-means algorithm. It tries to minimize the J_m measure which shows the compactness of clusters. Thus, smaller the value of J_m, better are the clusters obtained. The “J_m" is defined as:

$$J_m = \sum_{k=1}^{K} \sum_{i=1}^{n} {u_{ik}}^m D^2(v_k, x_i) \tag{4}$$

where"u_{ik} " is the degree of membership of point $x_i$to cluster center v_k,"$D(v_k,x_i)$” represents distance between x_i and v_k.and m is the degree of membership weight, m ϵ [0, ∞). Algorithmic steps for FCM are:

1. Membership weight m, total number of clusters, K and $C^{(0)}$ as the initial set of cluster centroids. The value of threshold, ‘ε’ in order to stop iteration is chosen. The membership value $U^{(0)}$ is calculated by following equation with the value of $C^{(0)}$:

$$u_{ki} = \frac{1}{\sum_{j=1}^{K} \frac{D(v_k, x_i)}{D(v_j, x_i)}^{\frac{2}{m-1}}} \ , for\ 1 \le k \le K;\ 1 \le i \le n \tag{5}$$

2. $C^{(q+1)}$ is updated using following equation:

$$c_i = \frac{\sum_{k=1}^{n} {u_{ik}}^m x_k}{\sum_{k=1}^{n} {u_{ik}}^m} \ , for\ 1 \le i \le K \tag{6}$$

3. $U^{(q+1)}$ is calculated according to $C^{(q+1)}$ using equation (5).

4. Values of $U^{(q+1)}$ are compared with $U^{(q)}$. If $\| U^{(q+1)} - U^{(q)} \| \leq \epsilon$, stop iteration. Otherwise, go to 2.

3.2.1 Advantages of FCM

1. It allows each data set point to belong to more than one cluster and assigns it to one of the clusters on the basis of membership function which results in good clustering.
2. It gives good result in case of overlapping data set.

3.2.2 Disadvantages of FCM

1. It needs the number of partitions or clusters as one of the inputs.
2. It often provide suboptimal solutions based on the initial parameter settings of the system.

3.3 Variable Length Genetic Algorithm

Genetic Algorithms are randomized search and optimization algorithm that mimics biological evolution process that fittest will survive in the generations to come. VGA [17] automatically evolves the optimum number of clusters for given image or dataset and produces the fuzzy partitioning of the data. Each chromosome depicts a partitioning of the data set which consists of variable number of clusters. The fitness of each chromosome is calculated as a function of “Xie-Beni (XB)” index, a cluster validity index. This index is optimized for obtaining the best clusters.

3.3.1 String Representation and Population Initialization

In this approach, centers of clusters are depicted in the chromosomes as real numbers. If dimensional space is ‘D’ and ‘c’ clusters are to be encoded then the length of chromosome will be D * c.

3.3.2 Fitness Computation

The fitness of a chromosome implies “the degree of goodness of the solution it represents”. The ‘Xie-Beni (XB)’ index has been used for fitness calculation. The XB index is defined as “a function of the ratio of the total variation $'\sigma'$ to the minimum separation ‘sep’ of the clusters.”

$$XB(U, V;X) = \frac{\sigma(U,V;X)}{n \times sep(V)} = \frac{\sum_{k=1}^{K}\sum_{i=1}^{n} {u_{ki}}^2 D^2(v_k,x_i)}{n \times min_{k\neq l} D^2(v_k,v_l)} \tag{7}$$

where"u_{ki}" represents the degree of membership of data set point x_i to cluster center v_k,"$D(v_k,x_i)$" represents the distance between data set point x_i and center of cluster v_k and n represents the size of data set.

The fuzzy membership values u_{ki}, the degree of membership of data set point x_i to cluster center v_k are computed as:

$$u_{ki} = \frac{1}{\sum_{j=1}^{K} \frac{D(v_k,x_i)}{D(v_j,x_i)}^{\frac{2}{m-1}}} \text{ ,for } 1\leq k\leq K;\ 1\leq i\leq n \tag{8}$$

The new cluster centres are obtained using the following equation:

$$v_i = \frac{\sum_{k=1}^{n} {u_{ik}}^m x_k}{\sum_{k=1}^{n} {u_{ik}}^m} \text{ , for } 1\leq i\leq K \tag{9}$$

Here, K represents the total number of cluster centers encoded in a chromosome.

3.3.3 Genetic Operators

3.3.3.1 Selection

Conventional fitness proportional selection algorithm is used for selection.

3.3.3.2 Crossover

The centres of clusters represented by the chromosome are not divisible. This implies that the points of crossover will lie in between two centres of clusters. Two parent chromosomes P_1 and P_2 are selected for process of crossover. Let P_1 represents partitioning containing M_1 clusters and P_2 represent partitioning containing M_2 clusters.

μ_1, point of crossover in P_1, is generated as μ_1 = rand() mod M_1 .

μ_2, the point of crossover in P_2 is given as

$$\begin{aligned} \mu_2 &= LB(\mu_2) + rand()mod(UB(\mu_2) - LB(\mu_2)) \quad if\ UB(\mu_2) >= LB(\mu_2) \\ &= 0 \qquad\qquad\qquad\qquad\qquad\qquad else \end{aligned} \tag{10}$$

where, $LB(\mu_2) = min[2,max[0, 2 - (M_{1-}\ \mu_1)]]$

$UB(\mu_2) = [M_2 , max[0,2 - \mu_1)]]$

3.3.3.3 Mutation

Every index position of a chromosome is mutated with a probability μ_m and thus chromosome gets mutated thoroughly. A random number ∂ in the range [0, 1] is generated. Values are changed according to following equation:

$$\begin{aligned} v_{new} &= (1 \pm 2 * \partial) * v, && when\ v = 0, \\ &= \pm 2 * \partial, && when\ v = 0. \end{aligned} \tag{11}$$

3.3.4 Termination

The algorithm is executed for a predefined number of generations. The best string or the chromosome with the maximum fitness in the last generation provides the solution to the clustering problem.

3.4 Multi-objective Genetic Clustering

The Multi-objective optimization problem [18] can be defined as: “Find the vector $x^* = [x^*_1, x^*_2, \dots, x^*_n]^T$ of decision variables that satisfies a number of equality and inequality constraints and optimizes the vector function $f(x) = [f_1(x), f_2(x), \dots, f_r(x)]^{T}$”. The constraints define the feasible region F which contains all the admissible solutions. The vector x^* denotes an optimal solution in F. x^* is Pareto optimal if there is no feasible vector x in which one objective function cannot be decreased without increasing other objective function.NSGA-II [19] is used for obtaining set of non-dominated solutions.

3.4.1 String Representation and Population Initialization

In this approach, cluster centers are encoded in the chromosomes as real numbers. If dimensional space is ‘D’ and ‘c’ clusters are to be encoded then the length of chromosome will be D * c.

3.4.2 Computing the Objectives

The Xie-Beni (XB) index and J_m measure are taken as the two objectives which need to be simultaneously optimized.

$$XB(U, V; X) = \frac{\sigma(U, V; X)}{n \times sep(V)} = \frac{\sum_{k=1}^{K} \sum_{i=1}^{n} {u_{ki}}^2 D^2(v_k, x_i)}{n \times min_{k \neq l} D^2(v_k, v_l)} \tag{12}$$

$$J_m = \sum_{k=1}^{K} \sum_{i=1}^{n} {u_{ik}}^m D^2(v_k, x_i) \tag{13}$$

The fuzzy membership values u_{ki} are computed using equation (8).The new cluster centres are computed using equation (9).

3.4.3 Genetic Operators

Crowded binary tournament selection is done and then conventional crossover and mutation are performed. The results of the last generation provide the different solutions to the problem.

3.5 Multi-objective Fuzzy Clustering and SVM

In Multi-Objective Optimization, the final set of solutions provide a number of non-dominated Pareto-optimal solutions and it is a tedious process to obtain final optimal result from them. Thus, SVM is used as an answer to the above mentioned difficulty. In this approach [20], a set of non-dominated solutions is obtained by executing 'NSGA-II' on given data set. A fuzzy voting is performed on solution set to get high confidence points which becomes training data set where input label is taken from original data set and output label is the cluster center obtained in solution set for the SVM classifier. Remaining points of the dataset are classified using SVM classifier . NSGA-II [19] is used for obtaining initial set of non-dominated solutions.

3.5.1 Algorithmic Steps

1. Obtain a set S of non-dominated solutions by performing NSGA-II based clustering.
2. The fuzzy membership matrix U is computed for every result in the solution set.
3. Membership matrices obtained in step 2 are reorganised such that cluster 'k' in one solution string corresponds to cluster 'k' in other strings also.
4. The data set points having maximum membership degree to some cluster k more than user defined membership threshold 'α' for minimum β*N number of solutions are chosen. These high confidence points forms the training set for SVM. The output class label of these points is cluster 'k'.
5. The SVM classifier is trained using above obtained training set.
6. The class labels or the clusters for the remaining data points are obtained using the trained SVM classifier.

7. Combine the result obtained in step 4 and step 6 to get final clusters for complete data set.

3.5.2 NSGA-II Based Multi-objective Clustering

The algorithm uses chromosomes that depict the coordinates of the centres of clusters as real numbers. If dimensional space is ‘D’ and ‘c’ clusters are to be encoded then the length of chromosome will be D * c. XB and FCM measure (J_m) are two objective functions that are being optimized simultaneously. The fuzzy membership values u_{ki} are computed using equation (8). The cluster centre v_k, $1 \leq k \leq K$ are updated using equation (9).The membership values are then computed again using modified cluster centres. Crowded binary tournament selection is done and then conventional crossover and mutation are performed. The last generation results give the different solutions to the problem.

3.5.3 SVM Classifier

SVM classifier constructs a hyperplane that separates the two classes in‘d’ dimensional space data set as maximally possible. Two hyper-planes, parallel to hyper-plane separating the two classes, are constructed that are pushed maximally towards their respective classes. When the distance between the two parallel planes is maximum, then the hyper-plane or classifier is said to provide a good separation. Greater the distance between the two parallel planes, better is the generalization error of the classifier. Basically, the SVM classifier has been designed for separating problems containing only two classes. It is extended for multiclass i.e. more than two class problems by constructing a multiple two-class SVMs.

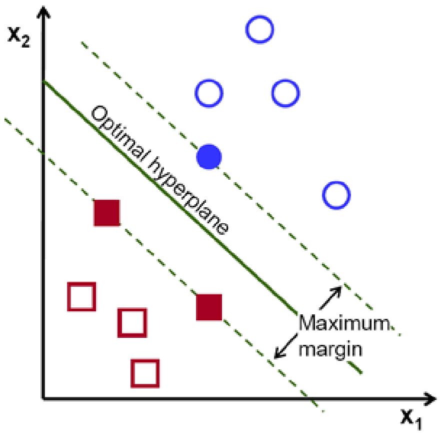


Figure 3.2 SVM Classifier

3.5.3.1 Extending SVM Classifier

Two approaches for extending SVM binary classifier:

3.5.3.1.1 One-vs-All Classification (OVA)

In this approach, N two-class or binary classifiers are built for distinguishing N classes. Let the points with positive value represent the points that belong to class ‘i', and the points with negative value represent the points not belonging to class ‘i' respectively for a classifier ‘i'. Here, f_i reprsents the i^{th} classifier. Classification is performed using following equation

$$f(x) = arg\overset{max}{i} f_i(x) \qquad (14)$$

3.5.3.1.2 All-vs-All Classification (AVA)

In this approach, N (N −1) two-class classifiers are built for distinguishing N classes, one classifier is used to distinguish two classes i and j. Here, f_{ij} represents the classifier two classes i and j where class i are selected as positive value points and class j are selected as negative value points. Classification is performed using following equation

$$f(x) = arg\overset{max}{i} \left(\sum_j f_{ij}(x)\right) \qquad (15)$$

CHAPTER 4: TYPE-2 FUZZY LOGIC SYSTEMS

4.1 Type-2 Fuzzy Set

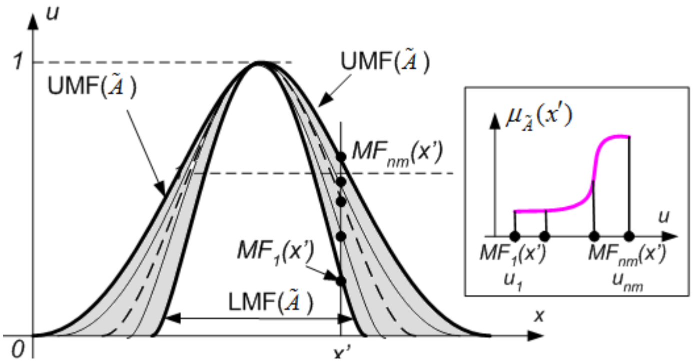
A type 2 fuzzy set is ‘a fuzzy set which has membership function as type 1fuzzy set on [0, 1]’. These sets generalise type 1 fuzzy sets . 'Uncertainty was not present in the membership function in type 1 fuzzy sets, they had a crisp value. This problem can be overcome by use of type 2 fuzzy sets as membership function in case of type 2 fuzzy sets is fuzzy. “A type-2 fuzzy set A, or $\tilde{A}$, is characterized by its membership function μ(x, u), where x ∈ X and u ∈ J_x ⊆ [0, 1]”, where X represents the universe of discourse. It can be expressed as follows:

$$\tilde{A} = \{((x,u),\ \mu(x,\ u)))I\ \forall\, x \in X,\ \forall\, u \in J_x \subseteq [0,1]\} \qquad (16)$$

where 0 ≤ μ(x, u) ≤ 1. $\tilde{A}$ can also be written as following:

$$\tilde{A} = \int_{x \in X} \int_{u \in Jx} \frac{\mu(x,u)}{x,u} \qquad (17)$$

Here, J_x⊆ [0, 1] represents the primary membership of x. μ(x, u) is a type-1 fuzzy set that provides the possibilities for the primary membership. Uncertainty in the primary membership function is the union of all membership functions. Shaded region in below figure is known as the ‘footprint of uncertainty (FOU)’ which depicts uncertainty. If μ(x, u) =1 ∀u ∈ Jx⊆ [0,1] then it is an interval type -2 fuzzy set. Two type-1 fuzzy sets called as UMF (upper membership function) and LMF (lower membership function) bound FOU.



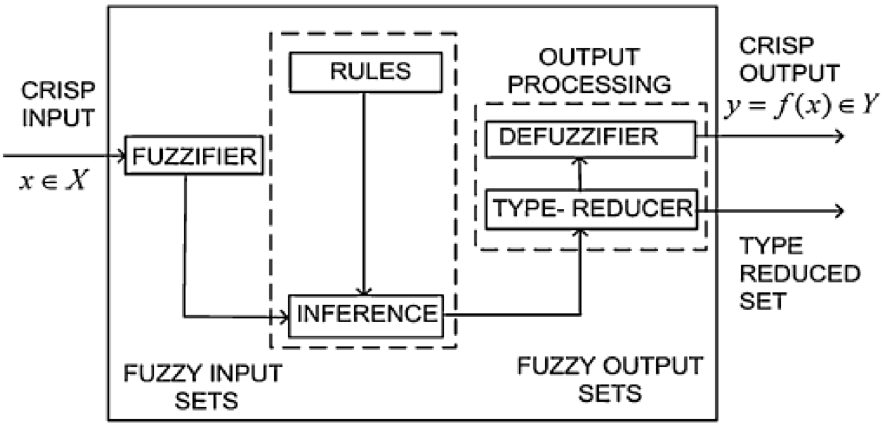
4.1 Type 2 Fuzzy Set

4.2 Type-2 Fuzzy Logic System

A fuzzy reasoning system is said to be type-2 fuzzy logic system if type-2 fuzzy sets are employed for fuzzy reasoning. A type -1 fuzzy logic system could not handle uncertainty related with rule base as membership function of type-1 fuzzy set has a crisp values.

Type-2 Fuzzy logic systems are helpful in the cases where exact membership values cannot be determined. A type-2 FLS comprises of following components:

1. Fuzzifier
2. Rule base
3. Fuzzy inference engine
4. An output processor which further consist of
    - Type-reducer
    - Defuzzifier



4.2 Type 2 Fuzzy Logic System

4.2.1 Fuzzifier

Fuzzification is the first step in the fuzzy inferencing process and this is performed by fuzzifier. Here, domain transformation takes place and crisp inputs are transformed into type- 2 fuzzy inputs.

4.2.2 Rules and Fuzzy Inference Engine

Here, the antecedents and the consequents of IF-THEN rules consist of type-2 fuzzy sets. The fuzzy inference engine using rule base, converts input fuzzy sets to output fuzzy sets. Many operations like union, intersection are required in this process.

4.2.3 Output Processor

The output processor takes fuzzy input from inference engine and converts to crisp outputs. Type-2 fuzzy set are converted to type-1 fuzzy set using type reducer and these converted output of type reducer is fed to defuzzifier. Defuzzifier converts type-1 fuzzy sets obtained from type reducer to crisp outputs.

CHAPTER 5: DIFFERENTIAL EVOLUTION

5.1 Introduction

Differential Evolution algorithm is an evolutionary technique that is used to solve optimisation problems over continuous domain. It is a population based algorithm and uses real numbers for representing decision variables. Differential Evolution uses three stochastic operators: mutation, recombination and selection for directing search in right direction towards the optimal solution. The basic difference between Differential Evolution and Genetic Algorithms is that Genetic Algorithm use crossover as basic search mechanism whereas DE uses mutation as basic search mechanism for solving optimisation problems.

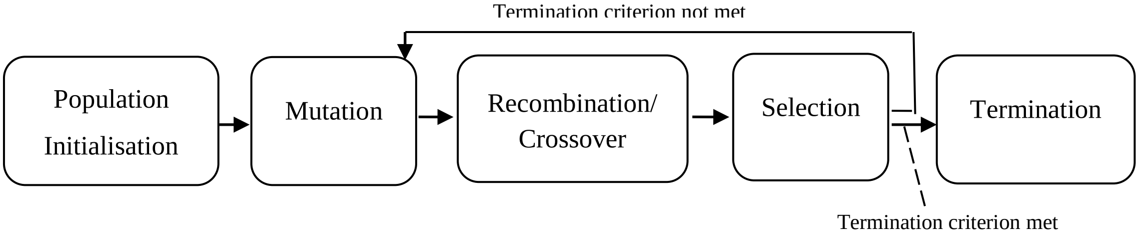

Termination criterion not met
Population Initialisation
Mutation
Recombination/ Crossover
Selection
Termination
Termination criterion met

Fig 5.1 General Evolutionary Algorithm Procedure

5.1.1 Population Initialization

In the case no information is available regarding the problem then initially, population is chosen randomly. For 'D' dimensional space, parameter vector can be expressed as follows:

$$x_{i,g} = \left[x_{1,i,G}, x_{2,i,G} \dots \dots .., x_{D,i,G}\right] \quad i = 1,2,3 \dots N \tag{18}$$

Here, g represents the number of generation and N represents size of population. Upper x_j^U and lower x_j^L bounds of the parameters x_j are chosen and then parameters are assigned values in interval [x_j^L , x_j^U].

5.1.2 Mutation

This step enhances the search space. In this, each parameter vector is mutated and a donor vector is created for each parameter vector. For each parameter vector $\boldsymbol{x_{i,g}}$, three parameter vectors $\boldsymbol{x_{i1,g}}, \boldsymbol{x_{i2,g}}$ and $\boldsymbol{x_{i3,g}}$ from population are selected in such a way that the indexes i, i1, i2 and i3 are different. The donor vector is created using following equation:

$$v_{i,G+1} = x_{i1,G} + F\left(x_{i2,G} - x_{i3,G}\right) \qquad (19)$$

Here F is mutation factor and it can be any value in interval [0, 2].

Other mutation operators that are used commonly are defined as follows:

1. $v_{i,G+1} = x_{best} + F\left(x_{i1,G} - x_{i2,G}\right)$
2. $v_{i,G+1} = x_{best} + F\left(x_{i1,G} - x_{i2,G}\right) + F\left(x_{i3,G} - x_{i4,G}\right)$
3. $v_{i,G+1} = x_{i1,G} + F\left(x_{i2,G} - x_{i3,G}\right) + F\left(x_{i4,G} - x_{i5,G}\right)$
4. $v_{i,G+1} = x_{i,G} + F\left(x_{bext} - x_{i,G}\right) + F\left(x_{i2,G} - x_{i3,G}\right)$
5. $v_{i,G+1} = x_{i1,G} + F\left(x_{bext} - x_{i1,G}\right) + F\left(x_{i2,G} - x_{i3,G}\right)$

Here, $\boldsymbol{x_{i,G}}$ represents parameter vector of current generation for which donor vector is being created. $\boldsymbol{x_{best}}$ represents parameter vector with maximum fitness. $\boldsymbol{x_{i1,G}}$, $\boldsymbol{x_{i2,G}}$, $\boldsymbol{x_{i3,G}}$ $\boldsymbol{x_{i4,G}}$ and $\boldsymbol{x_{i5,G}}$ represent randomly chosen parameter vectors from current generation such that condition i1 ≠ i2 ≠ i3 ≠ i4 ≠ i5 ≠ i ≠ best is satisfied.

Some mutation operators like 3 and operator defined in (18) are explorative whereas mutation operators like 2 and 4 are more exploitative.

5.1.3 Recombination

In this step, for each parameter vector, a trial vector is constructed using donor vector obtained in mutation step and original parameter vector. Elements of donor vector become a part of trial vector with probability CR. Here, CR denotes the probability of crossover. If a random value generated $rand_{j,i}$ is greater than CR then trial vector contains elements from donor vector else not.

$$\boldsymbol{u_{j,i,G+1}} = \begin{Bmatrix} \boldsymbol{v_{j,i,G+1}} & \boldsymbol{if} & \boldsymbol{rand_{j,i}} \leq \boldsymbol{CR\ or\ j = I_{rand}} \\ \boldsymbol{x_{j,i,G+1}} & \boldsymbol{if} & \boldsymbol{rand_{j,i}} > \boldsymbol{CR\ and\ j \neq I_{rand}} \end{Bmatrix}$$

$$for\ i = 1,2,3 \ldots N\ and\ j = 1,2,3 \ldots D \qquad (20)$$

Here, $rand_{j,i}$ is any value in interval [0, 1] and I_{rand} is a random integer from [1, 2… D]. The purpose of I_{rand} is to ensure that $u_{i,G+1}$ is never equal to $x_{i,G+1}$.

5.1.4 Selection

In this step fitness of trial vector obtained in previous step is compared with that of original parameter vector. If fitness value of trial vector is greater than fitness of original parameter vector then it is selected against original parameter vector for the next generation else vice-versa.

$$\boldsymbol{x_{i,g+1}} = \begin{Bmatrix} \boldsymbol{u_{i,G+1}} & \boldsymbol{if\ f(u_{i,G+1}) \leq\ f(x_{i,G})} \\ \boldsymbol{x_{i,G}} & \boldsymbol{otherwise} \end{Bmatrix} for\ 1 \leq i\ \leq N \tag{21}$$

5.1.5 Termination

Mutation, recombination and selection process are performed for a predefined number of iterations or until a particular termination criterion is not met. The parameter vector with the maximum fitness in the last iteration provides the solution to the problem.

5.2 Algorithmic Steps

1 Initialize population size N and termination criterion.

2 Randomly initialize parameter vector values x_j in interval [x_j^L , x_j^U].

3 Evaluate the fitness of parameter vectors.

4 Perform mutation using equation (18). Other mutation operators can be used.

5 For each parameter vector, create trial vector using equation (20).

6 Compare fitness of original parameter vector and trial vector created to find which vector will enter next generation.

7 Check termination criterion is met. If yes, output parameter vector with maximum fitness value. If not, go to step 3.

5.3 Advantages of Differential Evolution

1 DE has simple structure and is easy to implement.

2 Differential Evolution Enhances the capacity of local search

3 DE provides speedy results due to fast convergence.

4 Few parameter settings are required as compared to other algorithms.

5 DE is robust and is able to find true global minimum of search space without depending upon values of the initial parameters.

6 It can handle complex high dimensional optimization problems.

CHAPTER 6: PROPOSED APPROACH

6.1 The Approach

A fuzzy clustering technique based on differential evolution based has been proposed for segmentation. The proposed approach is a clustering based segmentation method that uses Xei-Beni index for fitness evaluation, in which Euclidean distance measure is used and for computing membership values, type-2 fuzzy system is employed. The basic steps of approach are described in detail.

6.1.1 Vector Representation and Population Initialization

In this approach, centers of clusters are depiced in the vectors as real numbers. If dimensional space is 'D' and 'c' clusters are to be encoded then the length of vector will be D * c. Initially, 'c' cluster centers of the vector are chosen randomly.

$$x_{j,i,G} = x^{L}_{j,i,G+1} + rand(0,1) * \{ x^{U}_{j,i,G+1} - x^{L}_{j,i,G+1} \}$$
$$for\ i = 1,2,3 \ldots N\ and\ j = 1,2,3 \ldots D \qquad (22)$$

Here, $\boldsymbol{x^{L}_{j,i,G+1}}$ represents the lower bound, $\boldsymbol{x^{U}_{j,i,G+1}}$ represents the upper bound, N is size of population and D is dimensional space. For example, D=3 and c=4, i.e. number of dimensions is 3 and number of clusters is 4, then vector is represented as follows, (2.9, 7.2, 5.2, 5.6, 9.5, 3.4, 2.0, 4.6, 6.3, 5.9, 4.9, 8.5). Here, clusters encoded are (2.9, 7.2, 5.2), (5.6, 9.5, 3.4), (2.0, 4.6, 6.3) and (5.9, 4.9, 8.5) respectively.

6.1.2 Fitness Computation

The method uses Xei-Beni (XB) index for fitness evaluation. The XB index is defined as "a function of the ratio of the total variation 'σ' to the minimum separation 'sep' of the clusters."

$$XB(U, V;X) = \frac{\sigma(U, V;X)}{n \times sep(V)} = \frac{\sum_{k=1}^{K}\sum_{i=1}^{n} {u_{ki}}^{2} D^{2}(v_k, x_i)}{n \times min_{k \neq l} D^{2}(v_k, v_l)} \qquad (23)$$

where"u_{ki} " is the degree of membership of point x_i to cluster center v_k,"$D(v_k,x_i)$" is the distance between x_i and v_k and n is the total number of data points.

For computing fuzzy membership values u_{ki} which is the degree of membership of point x_i to cluster center v_k , type 2 fuzzy sets are employed. Two fuzzifiers $m_1 = 2$ and $m_2 = 5$

are used to handle uncertainty. Upper ($\bar{\boldsymbol{u}}_{\boldsymbol{ki}}$) and lower ($\underline{\boldsymbol{u}}_{\boldsymbol{ki}}$) membership values are calculated as follows:

$$\bar{\boldsymbol{u}}_{\boldsymbol{ki}} = \left\{ \begin{array}{ll} \dfrac{\mathbf{1}}{\sum_{j=1}^{K} \left(\frac{D(v_k,x_i)}{D(v_j,x_i)}\right)^{\frac{2}{m_1-1}}} & \boldsymbol{if} \; \dfrac{\mathbf{1}}{\sum_{j=1}^{K} \frac{D(v_k,x_i)}{D(v_j,x_i)}} < \dfrac{\mathbf{1}}{\boldsymbol{K}} \\ \dfrac{\mathbf{1}}{\sum_{j=1}^{K} \left(\frac{D(v_k,x_i)}{D(v_j,x_i)}\right)^{\frac{2}{m_2-1}}} & \boldsymbol{otherwise} \end{array} \right\} \qquad (24)$$

$$\underline{\boldsymbol{u}}_{\boldsymbol{ki}} = \left\{ \begin{array}{ll} \dfrac{\mathbf{1}}{\sum_{j=1}^{K} \left(\frac{D(v_k,x_i)}{D(v_j,x_i)}\right)^{\frac{2}{m_1-1}}} & \boldsymbol{if} \; \dfrac{\mathbf{1}}{\sum_{j=1}^{K} \frac{D(v_k,x_i)}{D(v_j,x_i)}} \geq \dfrac{\mathbf{1}}{\boldsymbol{K}} \\ \dfrac{\mathbf{1}}{\sum_{j=1}^{K} \left(\frac{D(v_k,x_i)}{D(v_j,x_i)}\right)^{\frac{2}{m_2-1}}} & \boldsymbol{otherwise} \end{array} \right\} \qquad (25)$$

Because each data point has two membership values: lower and upper, therefore cluster center is represented by an interval between v^L and v^R. v^L and v^R are computed using KM algorithm and then cluster center encoded in vector is computed using following equation:

$$v_i = \frac{v_i^L + v_i^R}{2} \;, \; for\ 1 \leq i \leq K \qquad (26)$$

Membership values u_{ki} are updated using following equation:

$$u_{ki} = \frac{u_{ki}^L + u_{ki}^R}{2} \;, \; for\ 1 \leq i \leq n\ and\ 1 \leq k \leq K \qquad (27)$$

u_{ki}^L and u_{ki}^R are computed using following equations:

$$u_{ki}^L = \frac{\sum_{l=1}^{D} u_{kl}(x_i)}{D}$$

$$\text{where } \boldsymbol{u}_{\boldsymbol{kl}}(\boldsymbol{x}_{\boldsymbol{i}}) = \left\{ \begin{array}{ll} \bar{\boldsymbol{u}}_{\boldsymbol{ki}} & \boldsymbol{if} \; \bar{\boldsymbol{u}}_{\boldsymbol{ki}} \; \boldsymbol{is\ used\ for} \; \boldsymbol{v}_{\boldsymbol{i}}^{\boldsymbol{L}} \\ \underline{\boldsymbol{u}}_{\boldsymbol{ki}} & \boldsymbol{otherwise} \end{array} \right\} \qquad (28)$$

$$u_{ki}^R = \frac{\sum_{l=1}^{D} u_{kl}(x_i)}{D}$$

$$\text{where } \boldsymbol{u}_{\boldsymbol{kl}}(\boldsymbol{x}_{\boldsymbol{i}}) = \left\{ \begin{array}{ll} \bar{\boldsymbol{u}}_{\boldsymbol{ki}} & \boldsymbol{if} \; \bar{\boldsymbol{u}}_{\boldsymbol{ki}} \; \boldsymbol{is\ used\ for} \; \boldsymbol{v}_{\boldsymbol{i}}^{\boldsymbol{R}} \\ \underline{\boldsymbol{u}}_{\boldsymbol{ki}} & \boldsymbol{otherwise} \end{array} \right\} \qquad (29)$$

Here, D denotes number of dimensions of data set, k represents the number of clusters in vector and x_i represents a data set point.

6.1.3 Mutation

In this, each parameter vector is mutated and a donor vector is created for each parameter vector. Random value between 0 and 1 decides which mutation operator will be applied.

$$\boldsymbol{v_{i,l,G+1}} = \left\{ \begin{array}{cc} \boldsymbol{x_{GBest,l,G} + \alpha\,(x_{LBest,l,G} - x_{i1,l,G})} & \boldsymbol{if\ rand(0,1) < \alpha} \\ \boldsymbol{x_{i1,l,G} + F(x_{i2l,,G} - x_{i3,l,G})} & \boldsymbol{otherwise} \end{array} \right\} \quad (30)$$

Here, i varies from 1 to N where N is the size of ppulation and l denotes the number of elements in the parameter vector i.e. l = c*D. $\boldsymbol{x_{GBest}}$ and $\boldsymbol{x_{LBest}}$ represents global and local best i.e. vector with maximum fitness value of current generation. Parameter vectors $\boldsymbol{x_{i1,g}}, \boldsymbol{x_{i2,g}}$ and $\boldsymbol{x_{i3,g}}$ are selected from population in such way that the indexes i, i1, i2 and i3 are different. $\boldsymbol{\alpha}$ is defined by following equation:

$$\alpha = \frac{1}{1+ e^{\left(^{-1}/_{generation}\right)}} \quad (31)$$

6.1.4 Recombination

In this step, for each parameter vector, using donor vector obtained in mutation step and original parameter vector a trial vector is created. Elements of donor vector enter into trial vector with probability CR. If a random value generated $rand_{j,i}$ is greater than CR then trial vector contains elements from donor vector else not.

$$\boldsymbol{u_{i,j,G+1}} = \left\{ \begin{array}{lll} \boldsymbol{v_{j,i,G+1}} & \boldsymbol{if} & \boldsymbol{rand_{j,i} \leq CR\ or\ j = I_{rand}} \\ \boldsymbol{x_{j,i,G+1}} & \boldsymbol{if} & \boldsymbol{rand_{j,i} > CR\ and\ j \neq I_{rand}} \end{array} \right\}$$

$$for\ i = 1,2,3 \ldots N\ and\ j = 1,2,3 \ldots c * D \quad (32)$$

Here, $rand_{j,i}$ is any value in interval [0, 1] and I_{rand} is a random integer from [1, 2… c* D].The purpose of I_{rand} is to ensure that $u_{i,G+1}$ is never equal to $x_{i,G+1}$.

6.1.5 Selection

In this step fitness of trial vector obtained in previous step is compared with that of original parameter vector. If fitness value of trial vector is greater than fitness value of original parameter vector then it is selected against original parameter vector for the next generation else vice-versa.

6.1.6 Termination

Mutation, recombination and selection process are performed for a predefined number of iterations or until a particular termination criterion is not met. The parameter vector with the maximum fitness of the last iteration gives the solution to the problem.

6.2 Flowchart

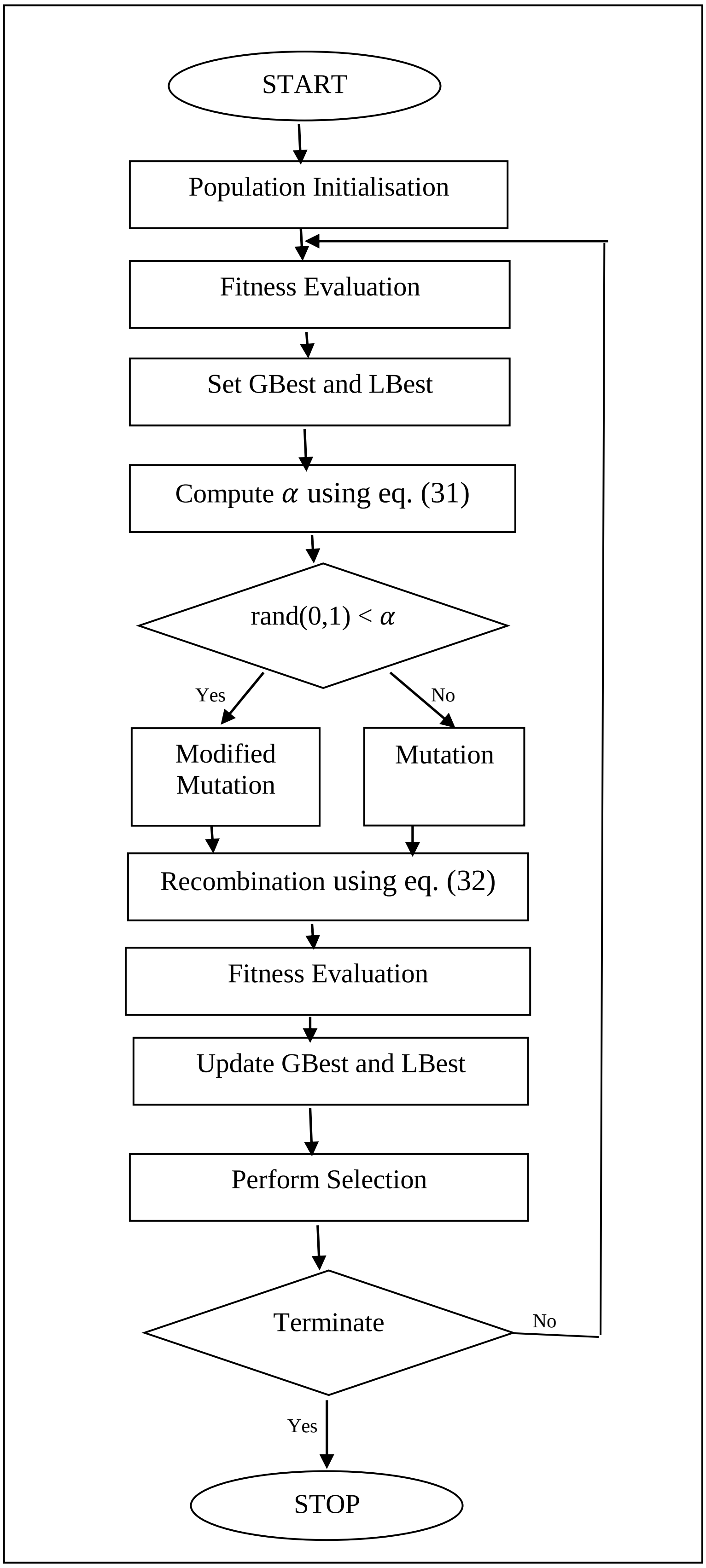
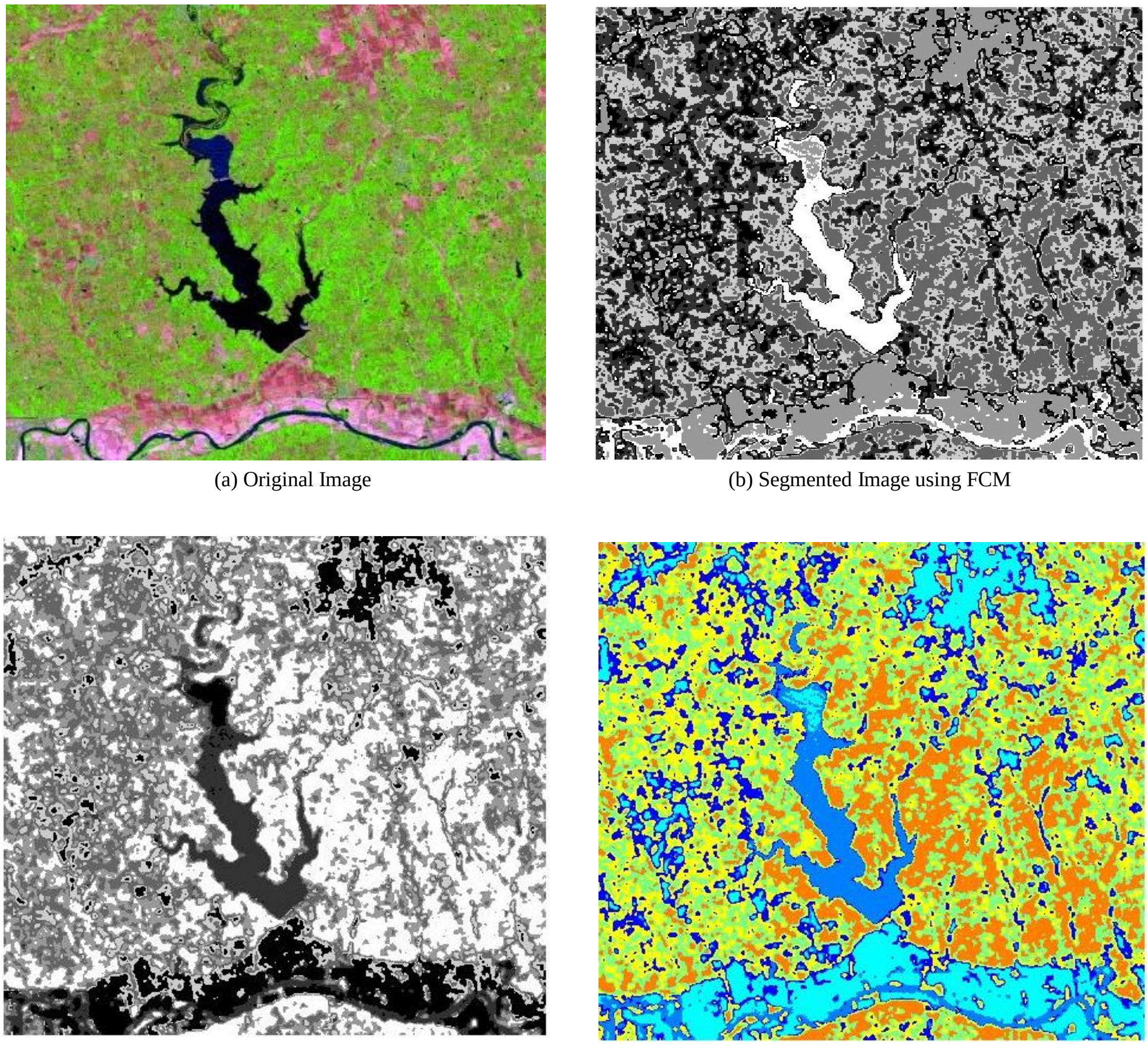


Figure 6.2 Flowchart of Proposed Algorithm

CHAPTER 7: EXPERIMENTAL RESULTS AND ANALYSIS

The parameters of the proposed algorithm (type-2 DE) are kept as follows: size population size is kept 30, crossover probability =0.8 and mutation factor F is also 0.8. Values of two fuzzifiers are taken as m_1=2 and m_2=5 respectively. The algorithm is run for a maximum 50 iterations. The parameters of the MOGA algorithm and variable genetic algorithm are kept as follows: population size is kept 20, crossover probability = 0.8 and mutation probability=0.01 and the weighting coefficient m=2.0. The algorithm is run for a maximum 50 iterations. The values of both α and β are 0.5. The SVM parameter C = 100 and the RBF kernel parameter γ is kept 0.1. FCM is run for maximum 100 iterations. For type 1 DE, population size is kept 30, crossover probability = 0.8 and mutation factor F is equal to 0.8. Experiment is conducted on various images and results obtained are as follows:

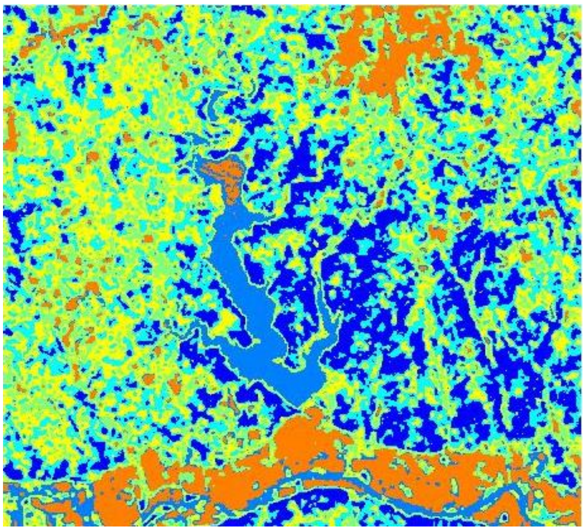


(a) Original Image

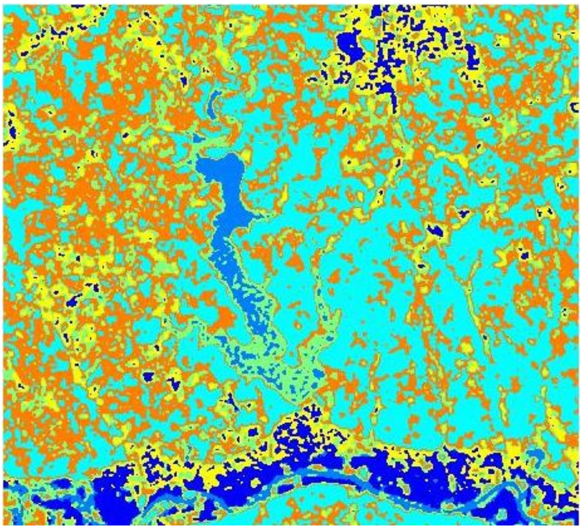
(b) Segmented Image using FCM

(c) Segmented Image using VGA

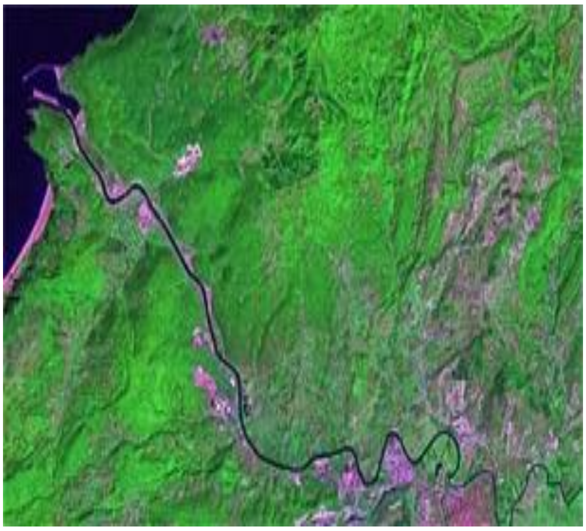
(d) Segmented Image using MOGA



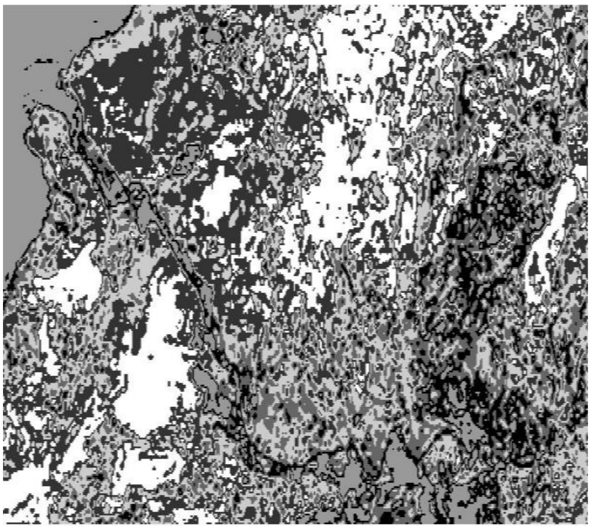
(e) Segmented Image using Type-1 DE



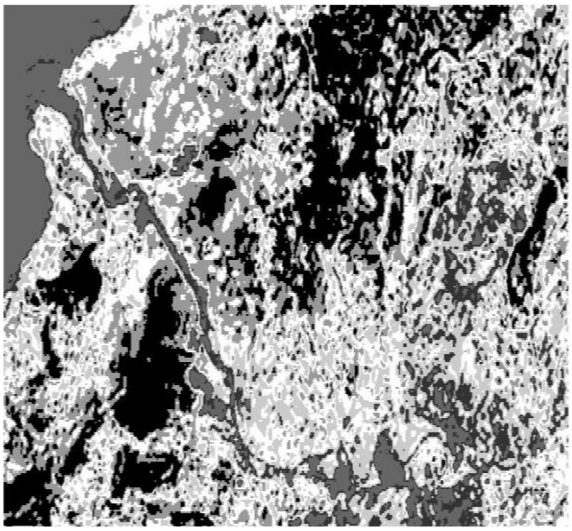
(f) Segmented Image using Type-2 DE



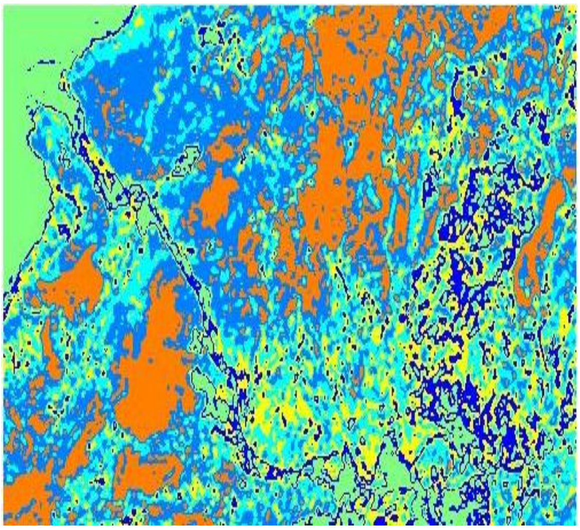
(g) Original Image



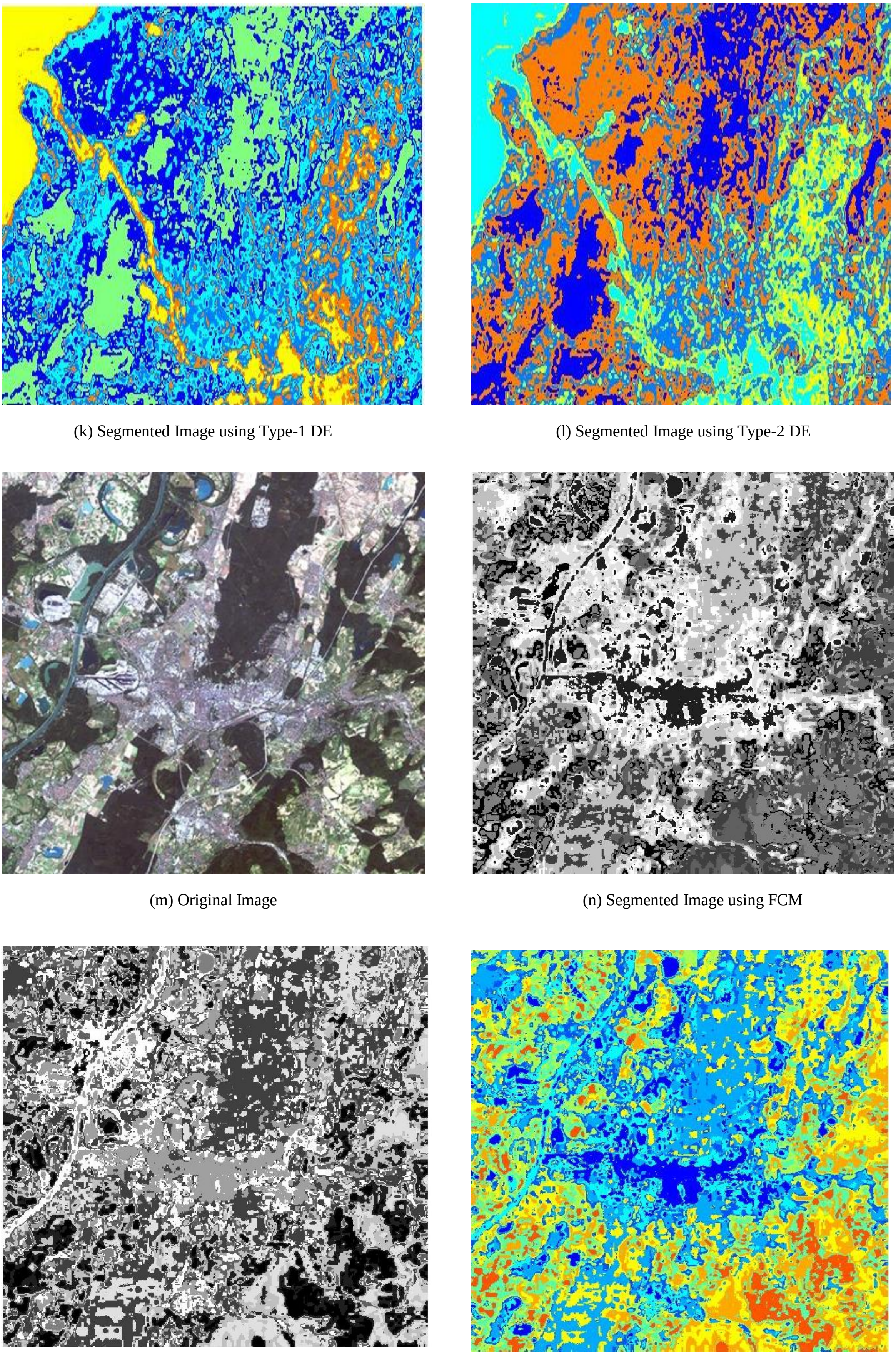
(h) Segmented Image using FCM



(i) Segmented Image using VGA



(j) Segmented Image using MOGA



(k) Segmented Image using Type-1 DE

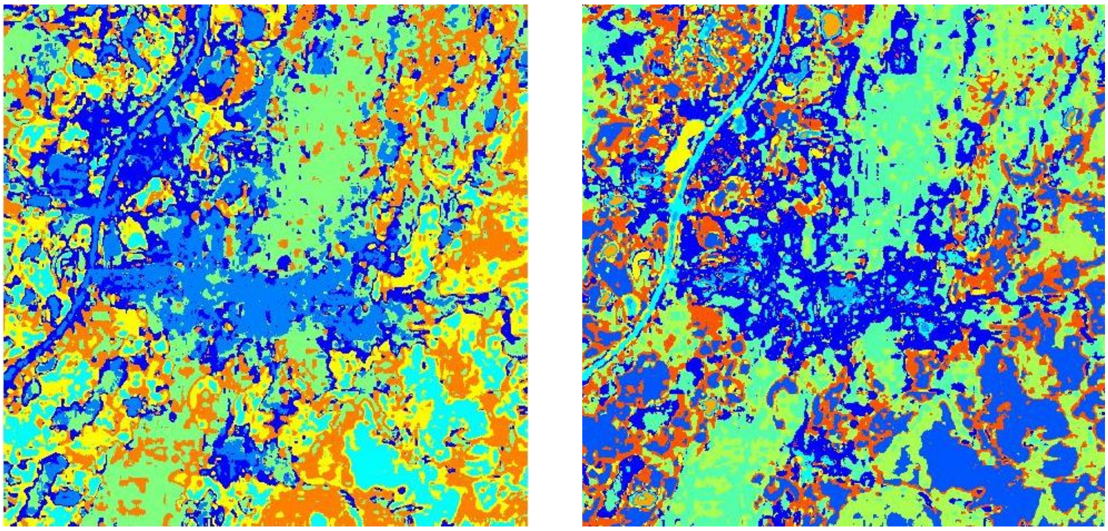
(l) Segmented Image using Type-2 DE

(m) Original Image

(n) Segmented Image using FCM

(o) Segmented Image using VGA

(p) Segmented Image using MOGA



(q) Segmented Image using Type-1 DE

(r) Segmented Image using Type-2 DE

Figure 7.1 Experiment Results of FCM and DE

For demonstration of the performance of the proposed scheme quantitatively, silhouette index of images classified using various algorithms are compared.

Silhouette index

Silhouette Index is a cluster validity index that measures how efficiently clustering has been performed.The silhouette value measures similarity of an object to its own cluster that is also known as cohesion as compared to similarity to the other clusters. Minimum value for silhouette index is -1 and maximum is 1.

A high value indicates high similarity of the data object to its own cluster and low or poor similarity to the other clusters. If maximum number of data points have a high value, then the clustering ouput is considerred apt. If maximum number of data points have a low or negative value, then the the clustering ouput is not considerred to be apt.Clustering output is said to contain eihter too large the number of clusters or too low number of clusters. Euclidean distance or the Manhattan distance or any other distance metric can be used for calculting Silhoutte Index.

Let the cluster number is k. For each data point i, a(i) represents the average dissimilarity of data point i with all other data points in its own cluster and b(i) represents the lowest average dissimilarity of data set point i to any other cluster. Silhouette Index is defined as

$$S(i)=\begin{cases} 1-a(i)/b(i), & if\ a(i)<b(i) \\ 0, & if\ a(i)=b(i) \\ b(i)/a(i)-1, & if\ a(i)>b(i) \end{cases} \tag{33}$$

The average s(i) of entire data set of a cluster is an index of how data in a cluster is grouped tightly or not i.e. it measures cohesion of a cluster. Thus the average s(i) over all data of the dataset measures whether the clustering is appropriate or not.

Table 7.2 Comparison of Silhouette Index

Image	**No. of Clusters**	**FCM**	**VGA**	**MOGA**	**Type-1 DE**	**Type-2 DE**
Perry	6	0.49	0.51	0.58	0.70	0.72
Landsat	6	0.41	0.49	0.54	0.69	0.75
Kalu-Sarai	9	0.33	0.39	0.47	0.57	0.66
Average Value		0.41	0.46	0.53	0.65	0.71

Average value of Silhouette index for proposed algorithm, Type-2 DE is more than that for other algorithms which confirms superiority of proposed algorithm over existing algorithms.

CHAPTER 8: CONCLUSION

The problem of satellite image segmentation has been modelled as fuzzy clustering that is solved by optimization of XB index, a cluster validity measure. A fuzzy clustering technique based on differential evolution based has been proposed for segmentation and the uncertainty involved has been represented using type-2 fuzzy sets. Results obtained on different satellite images and real life date sets indicate that proposed algorithm consistently performs better than many existing algorithms. Good performance of proposed method for such large image data sets shows that it can be employed in other data mining application.

LIST OF REFERENCES

[1] Ujjwal Maulik and Indrajit Saha, “Modified differential evolution based fuzzy clustering for pixel classification in remote sensing imagery”, Science Direct Journal of Pattern Recognition, Vol. 42, Issue 9, Pages 2135–2149, September 2009.

[2] Ujjwal Maulik and Indrajit Saha, “Automatic Fuzzy Clustering Using Modified Differential Evolution for Image Classification”, IEEE Transactions on Geoscience and Remote Sensing, Vol. 48, No. 9, September 2010.

[3] Dzung Dinh Nguyen, Long Thanh Ngo and Long The Pham, “GMKIT2-FCM: A Genetic-based Improved Multiple Kernel Interval Type-2 Fuzzy C-Means Clustering”, IEEE International Conference on Cybernetics, Pages 104-109, October 2013.

[4] Cheul Hwang and Frank Chung-Hoon Rhee, “Uncertain Fuzzy Clustering: Interval Type-2 Fuzzy Approach to C-Means”, IEEE Transactions on Fuzzy Systems, Vol. 15, No. 1, February 2007.

[5] Nilesh N. Karnik and Jerry M. Mendel, “Centroid of a type-2 fuzzy set”, Science Direct Journal of Information Sciences, Vol. 132, Issues 1–4, Pages 195–220, February 2001.

[6] Guo-bao Xu, Gui-yan Zhao, Lu yin, Yi-xin Yin and Yu-li Shen, “A CNN-based Edge Detection Algorithm for Remote Sensing Image”, IEEE Chinese Control and Decision Conference, Pages 2558- 2561, August 2008.

[7] Jianyu Chen, Jonathan Li, Delu Pan, Qiankun Zhu, and Zhihua Mao, “Edge-Guided Multiscale Segmentation of Satellite Multispectral Imagery”, IEEE Transactions on Geoscience and Remote Sensing, Vol. 50, No. 11, November 2012.

[8] Soham Sarkar, Swagatam Das and Sheli Sinha Chaudhuri, “Hyper-spectral image segmentation using Rényi entropy based multi-level thresholding aided with differential evolution”, Science Direct Journal of Expert Systems with Applications, Vol. 50, Pages 120-129, May 2016.

[9] Shilpa Suresh and Shyam Lal,” Multilevel thresholding based on Chaotic Darwinian Particle Swarm Optimization for segmentation of satellite images”, Science Direct Journal of Applied Soft Computing, Vol. 55, Pages 503–522, June 2017.

[10] S. Parea, A. Kumar, V. Bajaj and G.K. Singh, “A multilevel color image segmentation technique based on cuckoo search algorithm and energy curve”, Science Direct Journal of Applied Soft Computing, Vol. 47, Pages 76–102, October 2016.

[11] G.Raghotham Reddy, K.Ramudu, A.Sriniva and Dr. R. Rameshwar Rao, “Region Based Segmentation of Satellite and Medical Imagery with Level set evolution”, Recent Advances in Intelligent Computational Systems (RAICS) IEEE Conference, Pages 642-646, September 2011.

[12] Astha Baxi, Manoj Pandya, M .H. Kalubarme and M .B. Potdar, “Supervised Classification of Satellite Imagery using Enhanced Seeded Region Growing Technique”, Nirma University International Conference on Engineering (NUiCONE), December 2011.

[13] Gabriele Cavallaro, Mauro Dalla Mura, Edwin Carlinet ,Thierry G´eraud, Nicola Falco and Jon Atli Benediktsson,”Region-based classification of remote sensing images with the morphological tree of shapes” Geoscience and Remote Sensing Symposium (IGARSS), IEEE International Conference, Pages 5087-5090, July 2016.

[14] J. C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms. New York, NY, USA: Plenum, 1981.

[15] X. L. Xie and G. Beni, “A validity measure for fuzzy clustering,” IEEE Trans. Pattern Anal. Machine Intelligence, Vol. 13, pp. 841–847, Aug. 1991.

[16] U. Maulik and S. Bandyopadhyay, “Performance evaluation of some clustering algorithms and validity indices,” IEEE Trans. Pattern Anal. Mach. Intelligence, Vol. 24 No. 12, pp. 1650–1654, Dec. 2002.

[17] U. Maulik and S. Bandyopadhyay, “Fuzzy partitioning using a real coded variable-length genetic algorithm for pixel classification,” IEEE Transactions on Geoscience and Remote Sensing, vol. 41, no. 5, pp.1075– 1081, May 2003.

[18] Ujjwal Maulik,Sanghamitra Bandyopadhyay & Anirban Mukhopadhyay, “ Multiobjective Genetic Clustering for Pixel Classification in Remote Sensing Imagery”, IEEE Transactions on Geoscience & Remote Sensing, Vol. 45,No. 5,pp. 1650-1654,May 2007.

[19] K. Deb, A. Pratap, S. Agrawal, and T. Meyarivan, “A fast and elitist multi-objective genetic algorithm: NSGA-II,” IEEE Trans. Evolutionary Computing, Vol. 6, No. 2, pp. 182–197, Apr. 2002.

[20] Ujjwal Maulik and Sanghamitra Bandyopadhyay, “Unsupervised Pixel Classification in Satellite Imagery Using Multi-objective Fuzzy Clustering Combined With SVM Classifier” IEEE Transactions on Geoscience and Remote Sensing Vol. 47, No. 4, pp. 1132-1138, April 2009.

[21] Chen Yang, Lorenzo Bruzzone, Haishi Zhao, Yanchun Liang and Renchu Guan, "Decorrelation–Separability Based Affinity Propagation for Semi-supervised Clustering of Hyperspectral Images" IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Vol.9, No. 2, pp. 568-582, Aug. 2016.

[22] Anirban Mukhopadhyay, Ujjwal Maulik, Sanghamitra Bandyopadhyay and Carlos Artemio Coello , "A Survey of Multiobjective Evolutionary Algorithms for Data Mining: Part I" IEEE Transactions on Evolutionary Computation, Vol. 18, No. 1, pp. 4-19 ,Feb. 2014.