IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-5, May 2019 E-ISSN: 2347-2693

Image Segmentation via Genetic Algorithms

Ketna Khanna[1*], Naresh Chauhan[2]

[1,2]Computer Engineering Department, J. C. Bose University of Science and Technology, Faridabad, India

[*]*Corresponding Author: ketnakhanna05@gmail.com, Tel.: +91-9811697195*

DOI: https://doi.org/10.26438/ijcse/v7i5.625630 | Available online at: www.ijcseonline.org

Accepted: 26/May/2019, Published: 31/May/2019

Abstract— Image Segmentation is an immensely important task in Digital Image Processing. It is used in many fields including Medical Imaging, Machine vision, recognition tasks, etc. Many techniques have been proposed to carry out segmentation. Some of them are K Means, Image segmentation using Arithmetic Mean method, Segmentation via Entropy & histogram, Image segmentation using Maximum Between-cluster Variance and so on. This work proposes a novel Genetic Algorithms based image segmentation technique which may produce better results. A new Fitness function based on Entropy has been introduced. In order to check quality of segmented image, a performance evaluation measure is also presented. The proposed technique has been implemented on various images. Two existing approaches K Means and Arithmetic Mean have been thoroughly studied and implemented on same images. The results of proposed technique are then compared by the results of existing approaches using the introduced performance measure Entropy. Entropy measures the image information content. Greater the entropy, more information can be obtained from image. In comparison to the existing techniques, the proposed approach gives encouraging results.

Keywords—Image Segmentation, Genetic Algorithms

I. INTRODUCTION

Image segmentation refers to the splitting up of an image into a set of distinct regions that cover it [1]. The aim of image segmentation is to represent most important areas of an image. It locates the objects and curves or boundaries in images. An image can be partitioned according to attributes like colour, pixel intensity, gray levels or other textural properties [2]. Segmentation is reliable, provided the object's colour is different from that of background in the image. Some major applications of image segmentation include Medical imaging, Content-based image retrieval, Machine vision, object detection and recognition tasks [3]. Therefore, it is important to perform segmentation carefully so that good quality segments can be obtained. The two major techniques of segmentation are Region-based techniques and Edge-based techniques [4].

Many image segmentation techniques have been suggested in the review but they lack analysis of relevant performance measures. One of the commonly used techniques is K Means. Although it is simple but it has lots of drawbacks. The prior knowledge of number of clusters and local optimal solution are its major disadvantages [5]. Also the number of clusters has direct impact on the quality of final segmentation [6]. During Literature review it was found that the use of Soft Computing techniques in image segmentation is limited. This work proposes a Soft Computing technique for region based image segmentation. A performance evaluation measure is also presented to check how good the segmentation is done.

Soft Computing is a combination of various methodologies which are capable of finding the optimal solutions to real world problems [7]. These techniques are used when the conventional techniques fail to provide solution for complex, intractable and ambiguous problems [8]. There are various soft computing methodologies like Neural Networks (NN), Genetic Algorithms (GA) and Ant Colony Optimization (ACO) [9].

In the proposed work, GA is used as it is a meta-heuristic algorithm which imitate the process of natural selection and it is capable of generating high quality optimized solution for complex problems [10]. GA can be used for obtaining global optimal solution thereby avoiding local optima problem of K Means [11]. From review it has been proved that GA may improve the segmentation result of infrared images [12]. GA have also been found effective in the domain of medical image analysis and in segmentation of satellite images and remote images [13, 14].

In the proposed model Entropy is chosen as the performance evaluation standard which measures the image information content [15]. Greater the entropy of a segmented image, more information can be obtained [15]. A Fitness function based on entropy is proposed.

The organization of the paper as follows, Section I contains the introduction, Section II discusses Literature review, section III discusses the background, section IV discusses proposed work, section V presents result and analysis and section VI concludes.

II. RELATED WORK

An extensive literature review has been carried out to find the gaps in the existing literature. A brief review is presented in Table 1.

Table 1. Literature Review

S No.	Authors	Proposed Work	Explanation
1	Brzoza A et. al	An unsupervised image segmentation approach based on shortest paths in a graph representation of images has been introduced. This work proposes characterization of pixels in images so as to define the similarity relation between them.	Experiments proves that proposed approach gives superior results to those yielded by other baseline image segmentation methods
2	Akinina A et. al.	Here a method of automatic segmentation as part of the pattern recognition algorithm on satellite images is proposed.	From analysis, the operability of proposed method comes out to be good.
3	Li Y et. al	An improved grab cut algorithm which is the combination of grab cut and graph-based image segmentation is proposed.	From experiments, it has been proved that the proposed algorithm is effective and accurate.
4	Wilhelm T et. al.	In this work an unsupervised bayesian segmentation method that comprises of the edges present in an image as part of the model.	The proposed method produces encouraging results in comparison to supervised methods.
5	Sui H et. al.	In this work, the authors have described a novel stable shape feature-based image registration method via matching the stable region with a set of rotations.	The proposed algorithm is not sensitive to rotation and resolution distortion, through which image registration can be accomplished automatically.
6	Zhang R et. al.	A novel segmentation algorithm based on Markov Random Field and Bayesian theory is proposed. An objective function in image segmentation problem on the basis of optimality criterion of statistical decision and estimation theory is deduced.	The proposed algorithm is proved effective and robust.
7	Meinhold R et. al.	The authors have proposed two improvements in the algorithm for computing piecewise flat embeddings (PFE), so as to reformulate portions of the algorithm to enable various linear algebra operations to be performed in parallel. They have utilized an iterative linear solver so as to quickly solve a linear least-squares problem.	The tests reveal a significant drop in the execution time of proposed algorithm in comparison to other algorithms.
8	Xu W et. al.	Since it is difficult to initialize contour properly in images of complex contents, authors have proposed an ensemble strategy to improve the contour-based segmentation.	The proposed algorithm is proved effective from experiments.
9	Krishna T et. al.	Robust segmentation algorithm on remote sensing images has been proposed and properly implemented.	Proposed approach displays good performance in terms of time, noise and over segmentation in comparison to the traditional segmentation techniques.
10	Ahmadi S et. al.	Authors presented an approach to extract support relation under improvement of segmentation approach.	Proposed approach gives encouraging results in terms of accuracy.

III. BACKGROUND

A. Image Segmentation

Image Segmentation is a process of decomposing an image into sets of regions according to some attributes like pixel intensity or textural properties [26]. Image Segmentation can also be termed as an operation in which regions or features sharing similar characteristics are identified and grouped together [26]. The main aim of image segmentation is to segregate the relevant or meaningful areas or objects present in the image for further inspection. It is an important phase in Image Analysis Process as it differentiates the meaningful part of image from irrelevant parts or from background of image and represents the relevant areas. The two main objectives of image segmentation are as follows-

1. To partition the image into relevant regions. The retrieved regions can be proved helpful in further inspection of the image.
2. To change the representation i.e. pixels of image must be transformed into higher level units so as to either make them more meaningful or its future analysis becomes efficient or both.

Image segmentation can improve pictorial information for human interpretation. It might prove helpful to analyze the image in computer and provide an easy and systematic future analysis and inspection of relevant areas. Accuracy of segmentation can determine the eventual success or failure of computerized analysis procedure. Some Applications of image segmentation are Medical imaging which includes locating tumors, pathologies, diagnosis of anatomical structures; Content-based image retrieval; Machine vision; Object detection and recognition tasks [27]. Two famous existing image segmentation techniques are K Means and Arithmetic mean method.

In K Means, K clusters c_1, c_2--c_k are chosen with respective means m_1, $m_{2..}m_k$ which measures closeness of the data from their assigned clusters [1]. A least-squares clustering procedure inspects all the possible partitions into K clusters and select the one with which the distance can be minimized. Euclidean distance between each data point and chosen cluster centroids. Euclidean Distance can be calculated as-

$$D(c_k) = \sum_{x_i \in c_k} || x_i - c_k ||^2 \quad (1)$$

The data point is assigned to the cluster whose cluster centroid is closest to it.

Arithmetic Mean method uses the average value of random variables which are similar to objects center of mass. So dividing images according to the gray-level center should be the best point of balance [15]. If it is a colored image, then it must be converted first into gray scale image by using rgb2gray inbuilt function.

The mathematical expectation of a gray image can be determined by the following equation:

$$\mu_{threshold} = \sum_{n=1}^{N} L_n P(L_n) \quad (2)$$

Where

L_n= nth gray-level of image

$P(L_n)$= Probability of L_n

Threshold is used for dividing the image into distinct regions

$$\text{Value} = \begin{cases} 1, & if\,(img[i] \geq threshold) \\ 0, & otherwise \end{cases}$$

B. Genetic Algorithms

GA is a meta-heuristic algorithm which imitate the process of natural selection [10]. It is an evolutionary algorithm that is capable of generating high quality optimized solution for complex problems. When the classical techniques fail to give good enough solutions to intractable problems, GA can be used. GA is known for finding near optimal solution in very large search space based on the fitness function. The process of GA starts by randomly initialization of the start pool. Fitness values of chromosomes are calculated using fitness function and the fittest chromosomes i.e. chromosomes having maximum fitness values are selected using Roulette Wheel Selection method. Selected chromosomes are used to generate children for the next generation. As the algorithm proceeds, crossover and mutation are applied to the selected chromosomes and offsprings are produced which inherit attributes of parents. These are then added to the current population and this forms new generation. If parents have better fitness, their offspring will be better than parents and have a better chance at surviving. This is an iterative process which halts when stop criteria is satisfied and last generation includes the fittest individuals. Over successive generations, the population evolves toward an optimal solution [28].

A GA consists of four major components; chromosomes, Fitness function, Selection method and genetic operators (crossover and mutation operators). Chromosomes are also referred as candidate solutions of the given problem. Collection of chromosomes is termed as population. The first generated population is called a start pool. This pool can be generated either randomly or manually by using some techniques like Greedy Algorithm [29]. A fitness function is one which predicts the quality of a solution [29]. It contains whole information of problem. The selection function is used to determine the chromosomes that will take part in the execution process of a GA. One of commonly used function is Roulette Wheel Selection i.e. individual having highest fitness value will be given more priority than others [30]. Crossover operator shares the information between two parents i.e. chromosomes and generating new offspring (children) from them [31].

There are various conditions which are used to depict the end of the process of genetic algorithm. Most common terminating criteria are when there no further improvements in subsequent generations or when maximum number of specified generations is reached [29].

IV. METHODOLOGY

This work proposes a new fitness function which may give superior results than existing ones. An Entropy based performance measure is also introduced. Fitness function depends on entropy. First an image of dimension $(n * m)$ is selected. If selected image is colored, then it is converted into gray-scale image else this step is skipped. The gray-scale image is then reshaped and transformed into one-Dimensional array of $1 * (n * m)$. Start population is initialized randomly where chromosomes are the combination of 1s and 0s. Size of each candidate chromosome is $(n * m)$. Entropy of each candidate chromosome can be calculated as-

$$\text{Entropy}\ (\lambda) = \sum_{n=1}^{N} [-P(L_n) * \ln P(L_n)] \quad (3)$$

Where

L_n = n^{th} gray level of image

$P(L_n)$ = Probability of L_n

Fitness Function is based on entropy of chromosomes. It is computed as-

$$\text{Fitness Function} = \frac{1}{1+e^{-\lambda f}} \quad (4)$$

Where
f = Entropy of segmented image
λ = Constant value proportional to entropy (f)

The Fitness values of candidate chromosomes are calculated. A counter Generation_no is used to keep track of count of generations. Generation_no is initialized to 1. A number of maximum generations is specified prior to the start of execution. This specified maximum number of generation will serve as a stop criteria.
Following process is repeated till stop conditions are not satisfied.

1. Fittest chromosomes i.e. chromosomes with highest fitness values are picked using roulette Wheel Selection method.
2. Selected chromosomes undergo crossover and mutation.
3. Modified offsprings are inserted into population
4. If stop criteria is not satisfied then go to step 1. Else the fittest chromosomes from the last generation are returned as solution.

The process will halt if any of the following stop criteria is satisfied-

a) When there is no further improvement in population i.e. the chromosomes present in the i^{th} generation are same as the ones present in $(i+1)^{th}$ generation.
b) When the specified number of generations is reached.

The chromosomes present in the last generations are fittest i.e. they have highest fitness values. Maximum fitness value indicates higher entropy. Greater the entropy of an image, more information can be retrieved from it. The selection of chromosomes from last generation is done randomly. From the chosen chromosome, segmentation is performed.

A. Algorithm of proposed Work

The Algorithm for proposed work is as follows:

1. Select n*m image to be segmented. If it is a colored image, convert it into gray scale.
2. Reshape it into a 1D array of dimension (n*m,1).
3. Initialize the initial population by generating chromosomes randomly. Chromosomes are the strings of alleles of length (n*m) where the value of each allele is either 0 or 1.
4. Fitness Function is computed in such a way that maximum entropy can be attained. It is defined as –

$$\text{Fitness Function} = \frac{1}{1+e^{-\lambda f}}$$

Where
f = Entropy of segmented image
λ = Constant value proportional to entropy (f)

5. Evaluate fitness values of candidate chromosomes of start pool. Counter Generation_no represents the current version of generation. Initialize Generation_no to 1.

   Generation_no = 1
6. Repeat
   a) Select fittest chromosomes (with maximum fitness value) using Roulette Wheel Selection.
   b) Apply crossover and mutation to the chosen chromosomes.
   c) Insert the generated offsprings into the population and increment Generation_no.

      Generation_no+= Generation_no

      Until stop criteria is not satisfied.
7. Random selection of fittest chromosomes from the last generation.

B. Flow Chart of proposed Work

The flow chart of proposed algorithm is as follows:

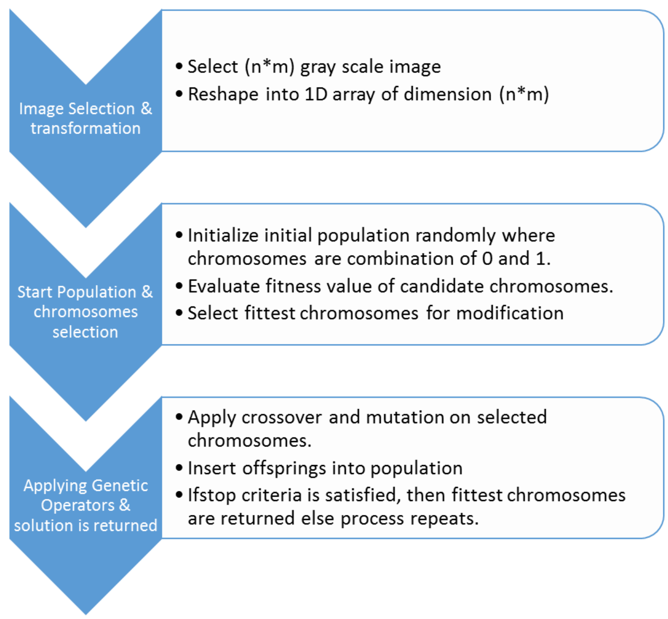
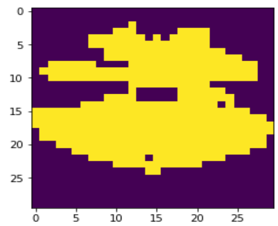


Figure 1. FlowChart of proposed technique

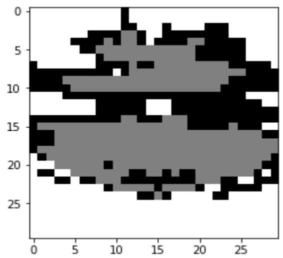
V. Results and Discussion

Along with the implementation of proposed technique, two existing image segmentation approaches K means and arithmetic Mean are also implemented on five images. Implementation is performed in Python with the help of Anaconda. The results of proposed approach are then compared to that of existing techniques on the basis of introduced performance measure Entropy. For a given beach image to be segmented, the result of proposed technique and existing approaches are as follows:

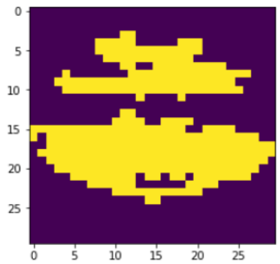
a: Original Image



b: Segmented Image via AM



c: Segmented Image via K Means



d: Segmented Image via Proposed technique

Figure 2. Results

Entropy for the above images are as follows:

Entropy of segmented image via Arithmetic Mean method- 0.6809
Entropy of segmented image via K Means- 0.4149
Entropy of segmented image via proposed technique- 0.693137

From the results, it is proved that the proposed approach is better in comparison to Arithmetic Mean method and K Means algorithm.

Table 2 displays the entropy values of proposed technique and existing approach for five distinct images.

Table 2. Observation Table

S No.	Image	Entropy via K Means	Entropy via Arithmetic Mean	Entropy via Proposed approach
1.	**Beach**	0.4149	0.6809	0.693137
2.	**Home**	0.6612	0.6931	0.724011
3.	**Mountain**	0.3713	0.6118	0.682433
4.	**Room**	0.4846	0.6640	0.691069
5.	**Sunset**	0.6756	0.6913	0.703137

A comparison chart between K Means, Arithmetic Mean and the proposed technique on the basis of Entropy of segmented image is displayed in Figure 3.

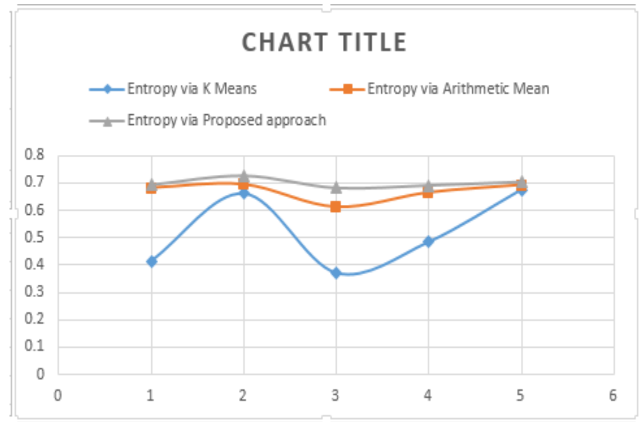


Figure 3. Comparison Chart

Entropy of the proposed technique for test images is greater than that of the existing approaches. From the above results it is proved that the proposed approach can produce better results as compared to K means and Arithmetic Mean method.

VI. CONCLUSIONS AND FUTURE SCOPE

Image segmentation breaks down the image into meaningful regions. It is an important phase in DIP as it segregates the relevant objects from the entire image. Detection of meaningful or required objects from image can prove helpful in its future inspection. Arithmetic Mean method for image segmentation uses mean of the pixel intensities for the purpose of Image Segmentation. K Means is another segmentation method which is known for its simplicity but it has many drawbacks. K Means has a tendency to stuck in the local optimal. GA is used in this work as it is capable of finding unique optimal solutions for complex real world problems unlike K means. The number of iterations required in the process of GA are much less as compared to that in K Means. GA can be used even if there is no prior knowledge of number of clusters in which the image is to be classified. Through the process of GA, the initial population is derived towards better solution. This work proposed a novel Genetic Algorithms based technique for image segmentation where a Fitness function is introduced with which maximum entropy can be achieved. K Means, Arithmetic Mean and proposed technique have been applied on five images. Results show that the proposed approach gives promising results.

The extension to this work will see the effect of variation in crossover rate, keeping other parameters constant. Mutation rate, Selection function etc. would also be changed to find the impact of these parameters. Researches have also suggested new variants of GA like diploid GA. It is intended to use these variants in future work.

REFERENCES

[1] Shapiro and Stockman, "*Computer vision",* Pearson publication, 1st edition, 2001.

[2] P. Umorya1 , R. Singh, *"A Comparative Based Review on Image Segmentation of Medical Image and its Technique"*, International Journal of Computer Sciences and Engineering, Vol.5, Issue.2, pp.71-76, 2017.

[3] Pham, Dzung L.; Xu, Chenyang; Prince, Jerry L, *"Current Methods in Medical Image Segmentation"*, Annual Review of Biomedical Engineering, 2000, pp: 315–337.

[4] Mihai Ivanovici Ivanovici, Noël Richard and Dietrich Paulus, *"Advanced Color Image Processing and Analysis"*, Springer, 2012, pp 219-277.

[5] Mathew, J., & Vijayakumar, R, *"Scalable parallel clustering approach for large data using parallel K means and firefly algorithms"*, International Conference on High Performance Computing and Applications (ICHPCA) (pp. 1–8). IEEE, 2014.

[6] Kuo, R. J., Syu, Y. J., Chen, Z.-Y., & Tien, F., *"Integration of particle swarm optimization and genetic algorithm for dynamic clustering"* Information Sciences, 2012, pp.124–140.

[7] Padhy, N.P. and Simon, S.P., *"Soft Computing: With Matlab Programming"*, Oxford, 2015, ISBN: 9780199455423, 0199455422. Edition: 1st Edition.

[8] Ketna Khanna, Naresh Chauhan and Vedpal, *"Genetic Algorithms based Black Box Testing"*, International Journal of Emerging Technologies and Innovative Research, 2018, pp. 537- 544.

[9] Dorigo, M. and Krzysztof Socha, *"Approximation Algorithms and Metaheuristics"*, IRIDIA Institut de Recherches Interdisciplinaires et de D\u00b4eveloppements en Intelligence Arti\ufb01cielle publisher, 2006, pp. 1-25.

[10] Goldberg, D., *"Genetic Algorithms in search, Optimization and Machine Learning"*, Addison-Wesley publishing co., 1989, pp.1-21.

[11] S. Josephine," *Brain Tumor MRI Image Detection and Segmentation Using Genetic Algorithm*", International Journal of Computer Sciences and Engineering, Vol.6, Issue.2, pp.358-364, 2018.

[12] Ya Wang, *"Improved OTSU and Adaptive Genetic Algorithm for Infrared Image Segmentation"*, The 30th Chinese Control and Decision Conference, 2018, pp. 5644-5648.

[13] Pathak S, *"Optimized Noisy Image Segmentation Using Genetic Algorithm"*, International Conference on Intelligent Computing and Control Systems ICICCS, 15-16 June 2017, Madurai, India, pp. 1311-1316.

[14] Maulik U, *"Medical Image Segmentation Using Genetic Algorithms"*, IEEE Transactions on Information Technology in Biomedicine (Volume: 13 , Issue: 2 , March 2009), 2009, pp. 166-173.

[15] Yide Ma, Kun Zhan and Zhaobin Wang,*"Applications of Pulse-Coupled Neural Networks"*, Higher Education Press, Beijing and Springer-Verlag Berlin Heidelberg, 2010, pp.1-9.

[16] Brzoza A et. al, *"An approach to image segmentation based on shortest paths in graphs"*, IEEE, 2017.

[17] Akinina A et. al., *"Multiscale Image Segmentation using Normalized Cuts in Image Recognition on Satellite Images"*, IEEE, 2018.

[18] Li Y et. Al, *"Grab Cut Image Segmentation Based on Image Region"*, IEEE, 2018.

[19] Wilhelm T et. al., *"Boundary aware image segmentation with unsupervised mixture models"*, IEEE, 2017.

[20] Sui H et. al., *"Automatic Image Registration Based on Shape Features and Multi-Scale Image Segmentation"*, IEEE, 2017.

[21] Zhang R et. al., *"The Development of Markov Random Field Theory and Applications on Image Segmentation Algorithm"*, IEEE, 2016.

[22] Meinhold R et. al., *"Efficiently Computing Piecewise Flat Embeddings for Data Clustering and Image Segmentation"*, IEEE, 2016.

[23] Xu W et. al., *"Ensemble of active contour based image segmentation"*, IEEE, 2017.

[24] Krishna T et. al., *"Three Phase Segmentation Algorithm for High Resolution Satellite Images"*, IEEE, 2016.

[25] Ahmadi S et. al., *"Enhance Support Relation Extraction Accuracy Using Improvement Of Segmentation In RGB-D Images"*, IEEE, 2017.

[26] H. T. T. Binh, M. D. Loi and N. T. Thuy, *"Improving Image Segmentation Using Genetic Algorithm"*, 11th International Conference on Machine Learning and Applications, Boca Raton, pp.18-23, 2012

[27] Gao H., *"The applications of image segmentation techniques in medical CT images"*, IEEE, 2011.

[28] Dhiraj Kumar, M. Tech Thesis, National Institute of Technology, Rourkela, 2009.

[29] Last, M. et. al., *"Effective black box testing via genetic algorithms"*, Springer-Verlag Berlin Heidelberg, 2006.

[30] Malhotra, R. et. al., *"Genetic Algorithms: Concepts, Design for Optimization of Process Controllers"*, 2011, Vol. 4, No. 2.

[31] Stender, J., *"Introduction to genetic algorithms"*, IEEE, 2004.

[32] Fischer, M. and Tonjes, R., *"Generating test data for black box testing using Genetic Algorithm"*, Conference: Emerging Technologies & Factory Automation (ETFA). IEEE 17th Conference, 2012.

[33] Hatem, M., *"Solving large problems with heuristic search: general-purpose parallel external-memory search"*, Journal of Artificial Intelligence Research 62, 2018.

Authors Profile

Ms. Ketna Khanna is currently pursuing Masters of Technology from J.C. Bose University of Science and Technology, YMCA, Faridabad. She did her Bachelors of Technology from Maharashi Dayanand University, Rohtak in 2017. Her main research area focuses on Data Mining and Software Testing.

Dr. Naresh Chauhan is PHD in Computer Engineering, Master of Technology in Information Technology and Bachelor of Technology in Computer Engineering. He is currently working as a Professor in Computer Engineering Department in J. C. Bose University of Science and Technology, YMCA, Faridabad. His research areas include Web Crawlers, Software Testing and Agile Software Development. He has around 150 research papers. He has 26 years of teaching experience and 5 years of Industrial experience.