Interactive Image Segmentation

Mr. Stafford Michahail, Dr. Bindu A Thomas
Research Scholar, Prof and Head Dept. of ECE
VVIET, Mysore
Stafford03@gmail.com, binduvviet@gmail.com

Abstract— Efficient and effective image segmentation is an important task in computer vision and object recognition. The goal of image segmentation is to simplify and change the representation of an image into something that is more meaningful and easier to analyze. Hence it is typically used to locate objects and boundaries (lines, curves, etc.) in images.

This paper presents an effective image segmentation approach that even provides excellent results for complex images using mean shift and minimal similarity based region merging (MSRM). In the present work, user need to indicate only the approximate location and region of the object and background by using strokes which are called as markers.

A novel maximal similarity based region merging mechanism is proposed to guide the merging process with the help of markers. It achieves these goals by three steps. First merging over-segmented regions according to maximal similarity rule with few strokes as input, second detecting the possible erroneous low contract object boundaries by analyzing image content and third automatically refine those boundary regions using both local and global information.

Keywords– MSRM, Low contrast image boundaries and Image segmentation**.**

I. INTRODUCTION

Image segmentation is an integral part of many applications in image processing which is used to separate the desired section of the input image. In general, the color and textural features in a natural image are very complex so that the fully automatic segmentation of the object from the background is a stiff task. Even though fully automated segmentation techniques have been constantly improving, however to the best of author's knowledge, there is rarely any automated image analysis technique which can be applied autonomously with satisfactory results in general cases. That is why the semi-automated techniques are needed.

A semi-automated segmentation algorithm allows the users to participate in the segmentation procedure and gives some guidance for the definition of the desired contents to be extracted, so we usually call it as an interactive segmentation.

The interactive image segmentation requires the proper user's intervention for segmentation process. Here, user gives the clue for segmenting the image in the process which leads to satisfactory results. Therefore, semiautomatic or interactive segmentation method is proposed, which use human expert knowledge as additional input and makes the segmentation problem more tractable.

The interactive Segmentation methods aims at minimizing the required user interaction time, while maintaining firm user control to guarantee the correctness of the results. In image segmentation, image similarity measure plays an important role. Region merging for object retrieval is an important task in many image processing applications. It has a wide application in area of crime prevention, intellectual properties, medical diagnosis and web searching. The object segmentation results influences the grouping of sub regions. So, accurate object segmentation is possible if both high level and low level priors are combined effectively.

The problem of image segmentation has gained a lot of attention since the early days of computer vision research. Automatic segmentation is a hard problem which requires modeling the problem based on domain knowledge. And even after that, some form of human intervention is required to correct anomalies in the segmentation.

The use of interactive image segmentation process provides solutions to such problems. One such interactive image segmentation proposed here is a novel interactive region merging method based on the initial segmentation of mean shift. In the proposed scheme, the interactive information is introduced as markers, which are input by the users to roughly indicate the position and main features of the object and background. The markers can be the simple strokes. The proposed method will then calculate the similarity of different regions and merge them based on the proposed maximal similarity rule with the help of these markers. The object will then be extracted from the background when the merging process ends.

II. LITERATURE SURVEY

The literature survey explores various image segmentation techniques that are used to partition the image. Some of them are mean shift image segmentation, watershed Image Segmentation, level set, clustering methods, edge detection method, super pixel Image Segmentation and Thresholding. These are classified as low level image segmentation techniques and they usually divide the image into many small regions. Although may have severe over segmentation, these low level segmentation methods provide a good basis for the subsequent high level operations, such as region merging. [4]

A brief description about various image segmentation techniques are as follows:

A. *Mean-shift*

Mean shift technique was first proposed by Fukunaga and Hostetler and later adapted by Chengfor. The purpose of image analysis is more recently extended by Comaneci, Meer and Ramesh to low-level vision problems including segmentation, adaptive smoothing and tracking. The main idea behind mean shift is to treat the points in the d-dimensional feature space as an empirical probability density function where dense region in the feature space corresponds to the local maxima or modes of the underlying distribution. For the data point in the feature space, one performs a gradient ascent procedure on the local estimated density until convergence occurs. It estimates the gradient if the probability density function to detect modes in an interactive fashion. Hence image segmentation that takes colour / intensity similarity as well as local connectivity into account can be obtained by using the algorithm to the combined special range domain. [6]

B. *Water shed image segmentation*

Different watershed lines may be computed in the image processing. The watershed lines can be defined on the nodes, edges, hybrid lines on nodes. Watersheds may also be defined in the continuous domain. There are also many different algorithms to calculate the watersheds. The user can apply different approaches to use the watershed principle for image segmentation.

1. Local minima of the gradient of the image may be chosen as markers. In this case, an over-segmentation is produced and a second step involves region merging.

2. Marker based watershed transformation makes use of specific marker positions which have been either explicitly defined by the user or determined automatically with morphological operators or other ways [2][6].

C. *Thresholding*

Thresholding is an old, simple and popular technique for image segmentation. Image segmentation by thresholding is a simple but powerful approach for segmenting images having light objects on dark background. This technique is based on image space regions i.e. on characteristics of image. This operation converts a multilevel image into a binary image. A threshold is defined and then every pixel in an image is compared with it. If the pixel lies above the threshold it will be marked as foreground, and if it is below the threshold it will be marked as background. The threshold will most often be intensity or colour value. Other forms of thresholding exist where the threshold is allowed to vary across the image. But those techniques are considered to be primitive and will only work for very simple segmentation tasks.

Limitation of thresholding method is that, only two classes are generated, and it cannot be applied to multichannel images and it is sensitive to noise and intensity in homogeneities. [3][6]

D. *Edge detection*

Edge detection technique is finding pixel on the region boundary. This method attempts to resolve image segmentation by detecting the edges or pixels between different regions that have rapid transition in intensity which are extracted and linked to form closed object boundaries. The result is a binary image. Edge detection is a fundamental tool in image processing, machine vision and computer vision, particularly in the areas of feature detection [3][6].

E. *Clustering Methods*

In this case, similarity criteria is defined between pixels, and then similar pixels are grouped together to form clusters. The grouping of pixels into clusters is based on the principle of maximizing the intra class similarity and maximizing the inter class similarity. The quality of a clustering result depends on both the similarity measure used by the method and its implementation [4][6].

F. *Super pixel Image Segmentation*

Many methods for object recognition, segmentation, etc rely on tessellation of an image into \ super pixels". A super pixel is an image patch which is aligned with intensity edges than a rectangular patch. Super pixels can be extracted using any segmentation algorithm. However, most of them produce highly irregular super pixels, with widely varying sizes and shapes. A more regular space tessellation may be desired. The super pixel partitioning problem in an energy minimization frame work is optimized using graph cuts. Energy function explicitly encourages regular super pixels. The Variations of the basic energy allows a trade-off between a less regular tessellation but more accurate boundaries or better efficiency [5].

The current work is based on the Mean Shift algorithm, as it provides lesser over-segmentation compared to other segmentation methods and also provides the individuals with a brilliant building block basis to work from, allowing for important real world analysis of data. It is this concrete mode finding algorithm that not only enhances other forms of segmentation when used in combination, but also produces a more superior form of image segmentation greater than the sum of its parts. This extendibility and compatibility clearly increases the worthiness and opportunities Mean Shift has to offer. However, creating a single segmented image based on Mean Shift may provide us with a good compact description of a segmented image, but not a very flexible one. A natural extension in order to gain such flexibility would be to transform our single level image description into a multi-layered hierarchal image description.

After the low level segmentation using mean shift method we perform high level segmentation using MSRM method which is abbreviated as maximal-similarity based region merging as it is essentially an adaptive region merging method. [2]

G. *Comparison of MSRM Technique with Graph Cut Method*

The comparison of MSRM method with another higher level segmentation method called graph cut method is as shown below: Since the original graph cut segmentation is a pixel based method, for a fair comparison with the proposed region based method. We extended the original pixel based graph cut (denoted by GCP) to a region based graph cut (denoted by GCR), i.e. the nodes in the graph are mean shift segmented regions instead of the original pixels.

Figure 2.1 shows the segmentation results of the three methods on eight test images. The first column shows the mean shift initial segmentation result and the input markers (for the last four images, the image boundary is set as the background marker), the second column shows the results by GCP. The third column shows the results by GCR and the fourth column gives the results by MSRM. We can see that with the same user input markers, the proposed MSRM method achieves the best results, while GCR performs better than GCP.

It can be seen that GCR will miss some object region sand wrongly label some background regions as object regions

Quantitatively comparing the three methods includes labelling the desired objects in the test images and taking them as ground truth. This is then followed by computation of the true positive rate (TPR) and false positive rate (FPR) for these segmentation results as shown in Table 2.1. The TPR is defined as the ratio of the number of correctly classified object pixels to the number of total object pixels in the ground truth, and the FPR is defined as the ratio of the number of background pixels but classified as object pixels to the number of background pixels in the ground truth. Obviously, the higher the TPR is and the lower the FPR is, the better the method is. Table 2.1 also lists the TPR and FPR results by the three comparison methods on the eight test images in Figure. 2.1. We can see that MSRM has the highest TPR and the lowest FPR simultaneously, which implies that it achieves the best segmentation performance. [5][6]

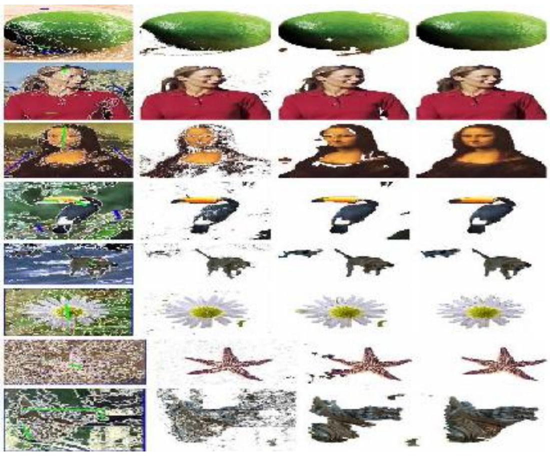

Figure 2.1 Comparison between graph cut and proposed method.

Image	Method	TPR (%)	FPR (%)
Fruit	GC_P	93.14	2.37
	GC_R	96.56	3.37
	MSRM	98.97	0.37
Woman	GC_P	97.58	2.99
	GC_R	96.82	0.73
	MSRM	98.53	0.44
Bird	GC_P	87.49	3.64
	GC_R	90.62	3.55
	MSRM	94.64	0.29
Dogs	GC_P	66.79	0.68
	GC_R	78.99	0.32
	MSRM	92.85	0.11
Mona Lisa	GC_P	54.08	2.02
	GC_R	90.71	2.34
	MSRM	98.85	0.71
Flower	GC_P	95.20	2.09
	GC_R	96.67	2.46
	MSRM	97.59	1.08
Tiger	GC_P	68.50	12.53
	GC_R	79.20	2.42
	MSRM	91.70	0.75
Starfish-1	GC_P	77.50	2.35
	GC_R	87.42	2.66
	MSRM	90.25	0.26

Table 2.1 The TPR and FPR values of different methods on test images

III. DESIGN METHODOLOGY

This section includes the block diagram of interactive image segmentation and required design specifications.

A. *Block Diagram*

The block diagram for interactive image segmentation is as shown in figure 3.1.Figure 3.1 shows the work flow of our approach. The strokes are first input to extract sampling of foreground and background of the source image. After over-segmenting the source image to generate many regions, they are merged into background and foreground using the MSRM rule producing the initial image segmentation. Next, suspicious low-contrast object boundaries are detected. Pixels in those boundary regions are re-classified to decide which class the boundary region belongs to and the region is re-labeled if necessary. After all suspicious boundary regions are processed, the final segmented image is obtained.

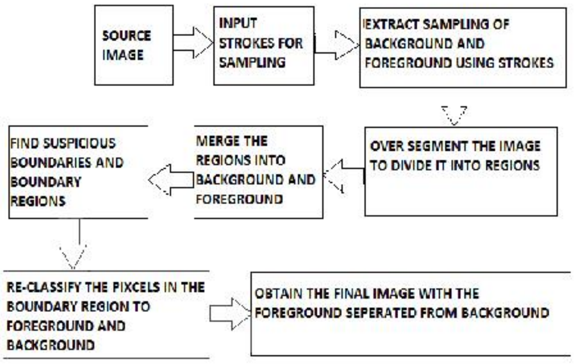

Figure 3.1 Block diagram of interactive image segmentation.

B. *Working*

The detailed descriptions of various steps involved in interactive image segmentation are as follows.

1. Obtaining the source image-This involves selection of the image on which segmentation is required to be performed and giving it as an input to the system in order to obtain the image separated from background.

2. Giving user input in terms of strokes for sampling and extracting the foreground and background- This involves user giving some interactive information as inputs in terms of green and blue markers, and extracting the foreground and background. The green marker represents the object while the blue markers represent the background and extracting the foreground and background.

3. Over segmenting the image to divide it into region-. This involves the mean shift segmentation which results in over segmentation for both the target object and background. In this case, the user is required to implicitly specify the regions which are located in the border of the image as background markers by drawing the object markers (green strokes) in the image.

4. Merging the regions into foreground and background- The user first marks the foreground and background regions with short strokes. The background regions are then merged using the MSRM rule. However, instead of using computationally expensive colour histograms, the mean colour of each region are used. Then the initial labeling of each region, either foreground (marked as 1) or background (marked as 0) is generated. Our initial over-segmentation outputs regions that are bigger than 20 pixels. It is assumed that the strokes provided by the user are sparse but pick the visually distinct regions. Given these two conditions, one round of boundary refinements was enough in our experiments.

5. Detecting suspicious low-contrast object boundary regions-This involves finding candidate regions for more careful analysis. Here, it is required to consider regions at the boundary between the foreground and the background, and also the ones whose colours are similar but whose labels differ. The boundary regions U_{bd} are defined as the regions that have atleast one B_j neighboring region with a different initial labeling. For example, if a foreground region A_i has a neighbor B_j that is marked as background, then A_i and B_j are boundary regions. It also required that such regions share a boundary that is at least 4 pixels long. For each boundary region, the mean colour μ_{bd} is calculated, and the neighbor of opposite label with the most similar mean colour is found. Finally all the boundary regions are sorted according to their minimal color differences $d_{bd}^{i,j}$, and then the regions with $d_{bd}^{i,j} < d_{thresh}$ are selected as the suspicious low-contrast object boundary regions to be refined. The threshold d_{thresh} is simply the median colour difference over all boundary region pairs such that one region is foreground and the other one is background.

6. Refine suspicious boundary regions and obtaining the final image- After all suspicious low-contrast object boundary regions are detected they are analyzed and possibly reclassified. Here, it is assumed that the initial segmentation using the mean-shift algorithm includes the correct region boundary. Using the local and global information of the pixels inside each region, each pixel is classified to be foreground or background. Then the number of foreground and background pixels is counted inside each region. If one region has more foreground pixels, it is classified as a foreground region, otherwise as background. The final image is then obtained by separating foreground from background.

C. *Mean shift Image Segmentation Algorithm*

1. Given $\mathbf{s_i}$, form $\mathbf{x_i} = (\mathbf{s_i}, \mathbf{F}(\mathbf{s_i}))^{\mathbf{T}}$ set j=1 and $\mathbf{Y_{i,1}} \leftarrow \mathbf{x_i}$.
2. Form a new set $\mathbf{s} = \{\mathbf{x_k}$, k=1, 2, 3…….n} from the image centred at $\mathbf{s_i}$ and within domain bandwidth $\mathbf{h_d}$.Compute the new center:

$$\mathbf{y_{i,j+1}} = \frac{\sum_{k=1}^{n} \mathbf{x_k}\, \mathbf{g}\left(\left\|\frac{\mathbf{y^s_{i,j}} - \mathbf{s_k}}{\mathbf{h_d}}\right\|^2\right) \mathbf{g}\left(\left\|\frac{\mathbf{y^r_{i,j}} - \mathbf{F}(\mathbf{s_k})}{\mathbf{h_r}}\right\|^2\right)}{\sum_{k=1}^{m} \mathbf{g}\left(\left\|\frac{\mathbf{y^s_{i,j}} - \mathbf{s_k}}{\mathbf{h_d}}\right\|^2\right) \mathbf{g}\left(\left\|\frac{\mathbf{y^r_{i,j}} - \mathbf{F}(\mathbf{s_k})}{\mathbf{h_r}}\right\|^2\right)}$$

3. Compute the mean shift $\mathbf{m_G}(\mathbf{y_{i,j}}) = \mathbf{y_{i,j+1}} - (\mathbf{y_{i,j}})$.
4. If $\|\mathbf{m_G}(\mathbf{y_{i,j}})\| < \boldsymbol{\epsilon}$, (e.g. $\boldsymbol{\epsilon}$ = 0.01) go to step 5. Otherwise: j ← j+1 and repeat Steps 2-3.
5. Set the converged value, assign filtered value $\mathbf{y_{i,c}} \leftarrow \mathbf{y_{i,j+1}}$. Assign filtered value as converged range vector $\mathbf{F}(\mathbf{s_i}) \leftarrow \mathbf{y^r_{i,c}}$.
6. Repeat Steps 1-5 until all pixels in image converges.

D. *MSRM Algorithm*

Input: the initial mean shift segmentation result. Output: the final segmentation map.

While there is region merging in the last loop

Stage 1. Merging non-marker regions in N, with marker background regions in M_B Input: the initial segmentation result or the merging result of the second stage.

(1-1) For each region $B \in M_B$, form these to fits adjacent regions $S_B = \{A_i\}_{i=1,2,3...r}$. .

(1-2) For each A_i and A_i / □ M_B, form its set of adjacent regions $S_{Ai} = \{S_j^{Ai}\}_{j=1,2,3...k}$ .There is B □ S_{Ai} .

(1-3) Calculate ρ (A_i, S_j^{Ai}). If ρ $(A_i, B) = max_{j=1,2,3..k}$ = ρ (A_i, S_j^{Ai}), then $B = B \sqcup A_i$. Otherwise, B and will A_i not merge.

(1-4) Update M_B and N accordingly.

(1-5) If the regions in M_B will not find new merging regions, the first stage ends. Otherwise, go to back to (1-1). .

Stage 2. Merging non-marker regions in N adaptively
Input: the merging result of the first stage.

(2-1) for each region P □ N, form the set of its adjacent regions $S_p = .\{H_i\}_{i=1,2,3...p}$

(2-2) for each H_i that $H_i / \in M_B$ and $/ H_i$ □ M_o, form its set of adjacent regions $S_{Hi} = \{S_j^{Hi}\}_{j=1,2,3...k}$. There is P □ S_{Hi}

(2-3) Calculate ρ (H_i , S_j^{Hi}). If ρ(P, H_i)= $max_{j=1,2,3..k}$ =1,2,...,k ρ(H_i , S_j^{Hi}), then $P = P \sqcup H_i$. Otherwise, P and H_i will not merge.

(2-4) Update N.

(2-5) If the regions in N will not find new merging region, the second stage stops. Otherwise, go back to (2-1).

END

E. *System Requirements*

1. Windows XP (Service Pack 2 or 3) or Windows 2003 (Service Pack 2 or R2).
2. Intel Pentium 4 processor or above.
3. 512 MB RAM (at least 1024MB RAM recommended).
4. 600 MB disk space.
5. 16-, 24-, or 32-bit OpenGL capable graphics adapter.
6. CD-ROM or DVD drive (for installation).
7. E-mail (required), internet access (recommended) for product activation.

IV. Experimental Result

Figure 5.1(b) and Figure 5.1(c) gives the experimental results of mean shift and MSRM segmentation techniques when applied to Figure 5.1(a).

Figure 5.1(a): Original image.

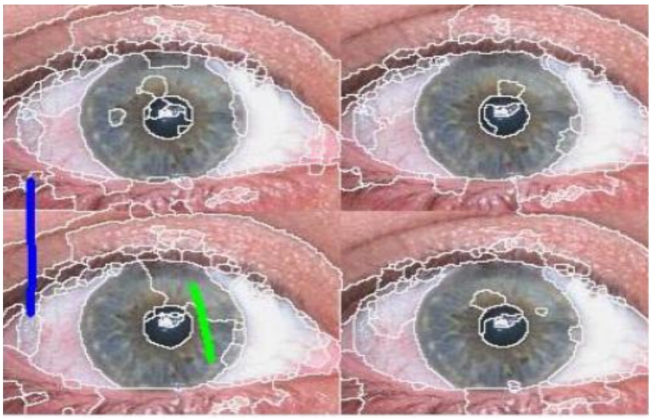

Figure 5.1(b) Mean shift segmented image.

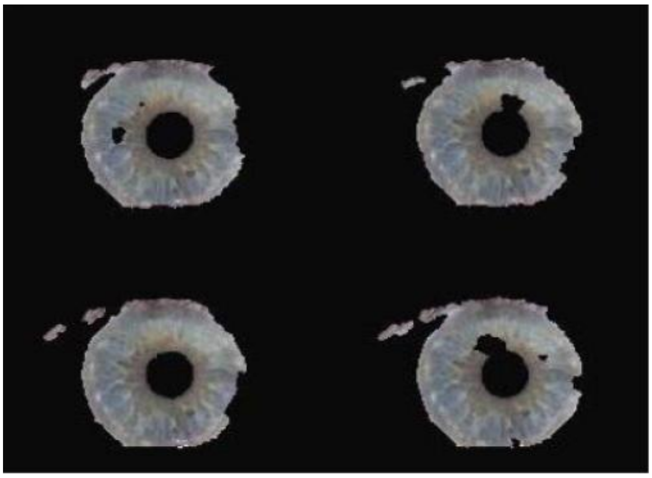

Figure 5.1(c) Result of MSRM technique.

V. CONCLUSION

Image segmentation has a promising future as the universal segmentation algorithm and has become the focus of contemporary research. In spite of several decades of research upto now to the knowledge of authors, there is no universally accepted method for image segmentation. As the result of image segmentation is affected by lots of factors, such as: homogeneity of images, spatial characteristics of the image continuity, texture, image content. The image is initially segmented by mean shift segmentation as it reduces the over segmentation .The users only need to roughly indicate the main features of the object and background by using some strokes, which are called markers. Since the object regions will have high similarity to the marked object regions and so do the background regions, a novel maximal similarity based region merging mechanism may be used to extract the object. This scheme is simple yet powerful and it is image content adaptive.

With the similarity based merging rule, a two stage iterative merging algorithm is presented to gradually label each non-marker region as either object or background. This method provides a general region merging framework.

REFERENCES

[1] “Banudevareddy.B, Prasanna Kumari P, (2012, sep-oct), ”Similarity based region merging interactive image segmentation”,*International journal of information system and computer sciences, 01(01).*

[2] NehaVerma, Deepak Sharma, (2013, Jun), “Region merging based image segmentation using maximal similarity mechanism”, *International journal of engineering research and development* , 07(07).

[3] NehaVerma,DeepakSharma,(2013,julaug),“Advanced maximal similarity based region merging by user interactions”,*International journal of engineering research and application ,03(04).*

[4] Jifeng Ning, Lei Zhang, David Zhang, Chenjke Wu (2010, Jun), “Interactive image segmentation by maximal similarity based region merging”, *journal of pattern recognition, 04(03).*

[5] Garima Singh Rawath, Joy Bhattacharjee, Roopali Soni,(2013,May), “Improved technique for object detection using similarity based region merging”, *International journal of research in JRCCJ computer and communication technology* , 02(05).

[6] G. Sumana, G. Anjan Babu (2013) “Efficient interactive image segmentation by maximal similarity based region merging for artificial neural networks for MRI medical images”, *international journal of engineering sciences research-IJESR ,04(01).*

[7] Fei Wang, Xin Wang, Tao Li. “Efficient label propagation for interactive image segmentation”. Sixth conference on machine learning and application.

[8] Ding ding Li, Yingen Xiong, Linda Shapiro, Kari Pulli “Robust interactive image segmentation with automatic boundary refinement”, international conference on image processing.

[9] Vikram M Karkade (2013, jun) “Overview on approaches of image segmentation with its algorithm and applications”, *International journals of engineering and computer sciences,02(06).*

[10] K. Vani Sree, A Vanaja (2013, Sep), “A novel marker based interactive image segmentation method”, *International journal of computational engineering and research,03(09).*

AUTHOR’S BIOGRAPHY

Stafford Michahial has received Master’s degree in Digital electronics and Communication from Visvesvaraya Technological University in the year 2012, Pursing Ph. D degree in the area of Image

Processing in VTU. My area of interest includes Neural Network, Data Mining, Image Processing, Pattern Recognition, Robotics electronics. He has published 17 papers in international and national journal. Successfully guided 10 UG projects.

Dr.Bindu A. Thomas received the Ph. D degree in Electronics from Mysore University, 2010, M.Tech degree in Industrial Electronics from Visvesvaraya Technological University with honors in 2003 and Bachelors degree in Electronics and Communications from Kuvempu University with Distinction in the year 1996. She is currently serving as the Professor and Head of the department in Electronics and Communications Engg., VidyaVikas Institute of Engineering and Technology, Mysore, Karnataka, India, where she is involved in Guiding research scholars and teaching UG students, and her major area of research include Image Processing, Pattern Recognition, Embedded Systems, Computational Intelligence applications in Robotics. she has Authored 60 research publications including international journals. Successfully guided 50 UG Projects co guide for 9 PG Projects