IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-9, Sept. 2018 E-ISSN: 2347-2693

Markov Random Field Based Model In Spatial Domain for Denoising of Microarray Images

Priya Nandihal[1] , Vandana S.Bhat[2*], Jagadeesh D. Pujari[3]

[1,2,3]Dept. of ISE, SDMCET, India

Corresponding Author: talk2priya.nandihal@gmail.com, Tel.: +91-9591814038

Available online at: www.ijcseonline.org

Accepted: 26/Sept/2018, Published: 30/Sept/2018

Abstract— Bioinformatics research is an active area of research that employs DNA microarray technology as a very important tool. Microarray gene expression is acquired through microarray technology in order to monitor the expression of genes under different conditions. Denoising is a major pre-processing step in DNA microarray images. This paper proposes a new spatial denoising technique in spatial domain for DNA microarray image. The method exploits Markov Random Field (MRF) model to reduce the noise in microarray images. Two algorithms developed in this work are Denoising using MRF (DMRF) and Determination of Optimized Values (DOV).Different experimental results and analysis demonstrate the performance of the proposed method with existing methods using various performance metrics.

Keywords— Spatial Filtering, Markov Random Field, Energy function, Non-linear Optimization, Performance Metrics

I. Introduction

The Microarray Technology has actively become a standard platform for functional genomics [1]. As microarray is placement of thousands of cDNAs on a glass microscope slide, it enables profiling of gene expression of tens of thousands of genes. Microarray may soon be capable of providing gene expression analysis of entire human genome in a single experiment. This might open up multiple access to key areas of human health including aging, mental illness, toxicology, drug discovery, disease prognosis and diagnosis. The processing pipeline of microarray images [2] typically involves following steps: gridding and spot fixing, segmentation and intensity extraction. Major part of gridding and spot fixing involves assigning the location of each spot in the image. Segmentation step consists of grouping similar feature pixels (differentiating foreground and background). Intensity extraction involves calculation of red and green foreground intensity pairs and background intensities.

Noise has been an inherent part of microarray image acquisition image acquisition due to its very nature. Other major contribution for the noise comes from complex bio-chemical and optical process involved in the preparation of microarrays [3]. Noise reduction or denoising is considered as major requirement in the pre-processing step as noise might affect subsequent stage of image analysis and finally gene expression profile. Many methods have been proposed for eliminating and reducing noise [4, 5, 6] in a microarray image. Most known methods are transform domain approach and the spatial filtering. In transform approach images are transformed to other domain such as Fourier or wavelet for preprocessing and then transformed back to obtain denoised image. In spatial filtering methods, filters are employed to obtain denoised image. In this paper spatial filtering methods is utilized that employs Markov Random Field (MRF) approach in-order to obtain denoised microarray images.

The detailed description of the paper is organized as follows: Section II focuses on the approach used to denoise microarray image using Markov Random Field. Section III focuses on the results obtained from proposed algorithm. Finally Section IV focuses on the conclusion and discussions of the proposed work.

II. Methodology

Markov Random Field (MRF) is a technique where the probability distribution of a random variable at a node (point) depends on immediate neighbors [7]. MRF exploits the existence of spatial correlations in the present context. In figure 1 i represents row and j represents column. The pixel value is represented by A (i, j). The image intensity values are from 0 to 255 as image intensity values are represented by 8-bits. If a(i, j) be is one of the values between 0-255. Probability is represented as

$$p(A(i,j)=a(i,j)) \qquad 1$$

The conditional probability is given as

$$p(A(i,j)=a(i,j)/A(k,m)=a(k,m)) \text{ for } k \neq i \text{ and } j \neq m. \qquad 2$$

Figure 1 demonstrates the 8 immediate neighbours. Random variables A(i, j) has the ability to form an MRF only if equation (2) is satisfied. This reveals that the value of the pixel is heavily depends on neighbors. This reveals that the value of pixel is heavily depend on neighbors.

(i−2, j−2)	(i−2, j−1)	(i−2, j)	(i−2, j+1)	(i−2, j+2)
(i−1, j−2)	(i−1, j−1)	(i−1, j)	(i−1, j+1)	(i−1, j+2)
(i, j−2)	(i, j−1)	(i, j)	(i, j+1)	(i, j+2)
(i+1, j−2)	(i+1, j−1)	(i+1, j)	(i+1, j+1)	(i+1, j+2)
(i+2, j−2)	(i+2, j−1)	(i+2, j)	(i+2, j+1)	(i+2, j+2)

Figure 1: shows immediate neighbours of (i, j) with grey color

The Hamersley-Clifford theorem [8], the probability mass function for MRF is given by Gibbs distribution [9] as shown in equation (3). Quadratic energy function for image analysis is mathematically represented in [10] as in equation (4).

$$p(x(i,j)) = 1/Z * \exp(-E(x(i j))) \qquad 3$$

$$E(x(i,j)) = (x(i,j) - a(i,j))2 + X1 + X2$$
$$X1 = c1* \sum_{(e,f)\in HV(i,j)} \left((x(i,j) - a(e,f))2\right); \; X2 = c2* \sum_{(g,h)\in D(i,j)} \left((x(i,j) - a(g,h))2\right) \qquad 4$$

Here a (i , j) is pixel value at (i, j). HV(i, j) indicates the indices of horizontal and vertical as depicted in fig (2). (e, f) are shaded with grey in the fig (2). D(i, j) indicates the diagonal neighbours (i, j) in fig (3). (g, h) are shaded grey. c_1 and c_2 are optimally chosen for best performance.

Here a (i , j) is pixel value at (i, j). HV(i, j) indicates the indices of horizontal and vertical as depicted in fig (2). (e, f) are shaded with grey in the fig (2). D(i, j) indicates the diagonal neighbours (i, j) in fig (3). (g, h) are shaded grey. c_1 and c_2 are optimally chosen for best performance.

	(i−1, j)	
(i, j−1)	(i, j)	(i, j+1)
	(i+1, j)	

Figure 2. HV(i, j), Horizontal and Vertical neighbors of (i,j)

(i−1, j−1)		(i−1, j+1)
	(i j)	
(i+1, j−1)		(i+1, j+1)

Figure 3.D(i, j), Diagonal neighbors of (i, j)

Denoising using MRF model is achieved through Maximum-a-Posteriori (MAP). The noisy image consists of two components one is the original intensity and the noise component as shown in equation (5). Best estimate of x(i, j) is achieved by utilizing MAP that maximizes the probability of x(i, j) as depicted in the equation (6). Equation (3) and (5) it can observed that maximizing p(x(i, j)) is same as minimizing energy function E(x(i, j)) over x(i, j)

$$a(i,j) = x(i,j) + n(i,j) \qquad 5$$

$$b(i,j) = \arg\max_{X(i,j)}\{p(x(i,j))\} \qquad 6$$

$$b(i,j) = \arg\min_{X(i,j)}\{E(x(i,j))\} \qquad 7$$

Minimizing the energy function can be achieved by differentiating and equating to zero we obtain equation (8).

$$x(i.j) = b(i,j) = (a(i,j) + c1*t1 + c2*t2)/(1 + 4*c1 + 4*c2) \qquad 8$$

t_1 provides with the horizontally and vertically adjacent pixels, t_2 provides diagonally adjacent pixels, b(i, j) in equation (8) for value (i, j) constitute denoised image. t_1 and t_2 are computed as per the equations (9) and (10). The pseudo code for denoising using MRF *(DUMRF)* algorithm is as shown:

$$t1 = a(i-1,j) + a(i,j+1) + a(i,j+1) + a(i+1,j) \qquad 9$$

$$t2 = a(i-1,j-1) + a(i-1,j+1) + a(i+1,j-1) + a(i+1,j+1) \qquad 10$$

Pesudo-code: DUMRF Algorithm
Input: Noisy spot on the microarray image
Output: Denoised spot on the microarray image
STEP 1: read the input image, the size of the noisy spot is (M, N) where M is number of rows and N is number of columns on the image A. Save the pixels a(i, j)
STEP 2: Obtain the suitable values of c_1 and c_2
STEP 3: Implementing the equations (8), (9) and (10) the denoised spot in the image is obtained
STEP 4: Matrix B with the indices b(i, j) is the denoised image
STEP 5: Exit

The denoising performance is determined by the optimal selection of c_1 and $c_2$values as in equation (8). Synthetic noise is added that might be Gaussian, Poisson, salt and pepper with appropriate noise parameters like mean, variance etc. In order to achieve optimal c_1 and c_2, determination of optimum values (DOV) algorithm is employed for minimum denoising error. The obtained c_1 and c_2 values are utilized for rest of the images with unknown noises. Mean square error (MSE) is opted as the metric to measure the performance of the algorithm. MSE is mathematically given by the equation (11).

$$MSE = 1/(M*N)\sum_{i-1}^{M}\sum_{j-1}^{N}(b(i,j) - q(i,j))2 \qquad 11$$

MSE is a non-linear fuction depending on c_1 and c_2. Optimal value of c_1 and $c_2$minimizes the MSE that casts as the non-linear minimization problem. MSE does not possess any constraints for solving c_1 and c_2, this is an unconstrained non-linear optimization problem. Numerous techniques are available [11-13] to solve the above problem. A few significant among them are trust-region [14], Quasi-Newton [15], Nelder-mead [16, 17].

Pesudo-code: Algorithm (DOV)
Input: Noiseless image Q, with synthetic noise matrix G
Output: Optimized values of c_1 and c_2
STEP 1: opt suitable synthetic noise matrix G
STEP 2: Obtain noisy image A by adding Q and G
STEP 3: Declare vector c as, c=[c_1 and c_2].
STEP 4: Initialize vector c_0=[0.0, 0.0].
STEP 5: Create a function get_mse(c, Q, A) with equations (11) and (8)
STEP 6: Execute function get_mse(c, Q, A).
STEP 7: Output is ready in vector c=c_1 and c_2
STEP 8: Exit

Three main experiments are carried out in the paper. The first experiment provides the optimal values of c_1 and c_2. In the second experiment, the obtained values are tested on the other spot and third experiment carries out the comparison with other methods.
In the first experiment a sample microarray spot of size 20×20 is obtained and converted to gray scale image Q. Gaussian noise G is added with 0 mean, 0.01 variance to get A. The noisy image image A along with it noiseless image Q and denoised image are shown in figure 4. The second experiment tests on Gaussian, salt and pepper, poisson and speckle noise are considered taking the values of c_1 and c_2 from experiment 1. In the third experiment the algorithm is tested with noise levels ranging from 0.001 to 0.010 and compared with averaging filter [18], sure shrink wavelet filter [19], soft-thresholding wavelet filter [20] and SUSAN filters [21].

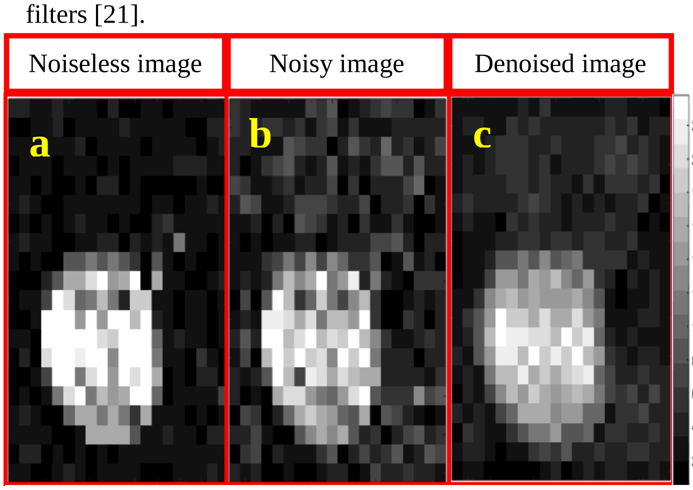

Figure 4: (a) noiseless image, (b) contains synthetically added noise component, (c)denoised image

III. Results And Discussion

In the first experiment carried out the values of c_1 and c_2 obtained were $\mathbf{c_1 = 0.1328}$ and $\mathbf{c_2 = 0.0387}$ at iteration 7. In figure 4 it can be observed that the figure 4(a) is the ground truth i.e, image with no noise, figure 4 (b) is the image that has the synthetically added noise, figure 4 (c) is the obtained denoise image with the proposed algorithm in the work. The converging values of c_1 and $c_2$along with the corresponding MSEare shown in table 1. If we observe carefully it can be noted that the values of $c_{1,}c_2$ and MSE do not vary after the 6^{th} iteration. The algorithm was tested till the 9^{th} iteration in order to confirm was it the local minima or global minima. Turn out that the values do not change even after 9^{th} iteration. Hence the mid iteration 7^{th} iteration values were chosen as the optimal values.

Table 1. Converging values of c1, c2 and MSE

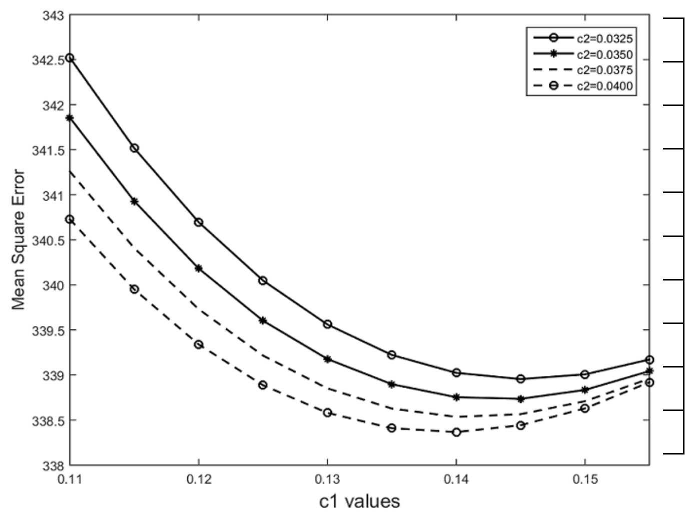

Figure 5. Variation of MSE with c_1 and c_2

The detailed relation with MSE, c_1 and c_2 is clearly shown in figure 5. As $c_2$goes on increasing, at higher value of c_1 MSE tries to attain a steady value. In the second experiment the values of c_1 and c_2 were plugged in for 4 kinds of noise as mentioned in the methods section the corresponding values. The corresponding percentage reduction of noise for each noise type is shown in table 2. The higher value of percentage reduction is reported in the Gaussian noise, so the developed algorithm works good for the Gaussian kind of noise. In the third experiment the developed algorithm is compared with standard algorithm as mentioned in the methods section. The values projected in the table 3 to 8 as in (supporting document) are plotted in the figure 6 to 11. The MSE and Mean Absolute Error (MAE) values obtained from developed algorithm and the standard algorithms are reported in the figure 6 and 10. Lower values of MSE and MAE indicate that the proposed algorithm perform well than the existing standard algorithms that are compared.

Table 2. Percentage Reduction for different types of noises

Type of Noise	Gaussian	Salt & pepper	Poisson Noise	Speckle Noise
Per_red	40.85	32.64	16.69	26.48

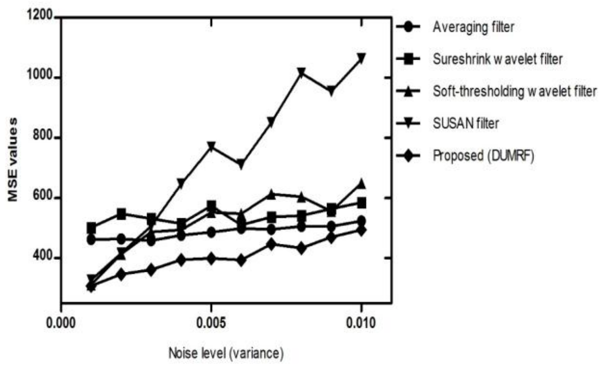

Figure 6. Variation of Mean Square Error with noise level

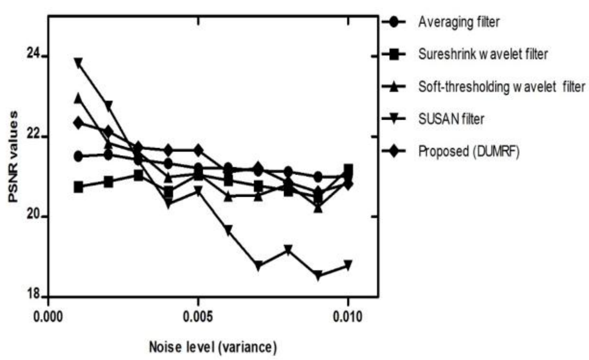

Figure 7. Variation of Peak signal-to-noise ratio with noise level

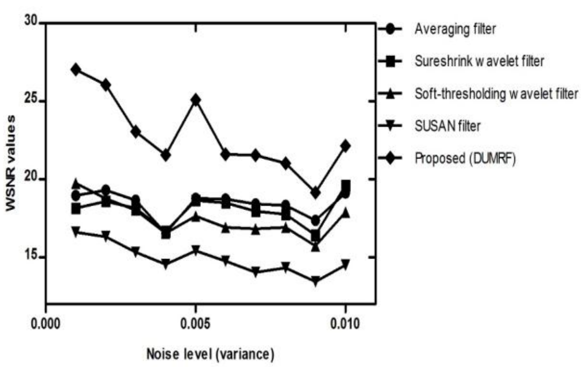

Figure 8. Variation of Weighed signal-to-noiseratiowith noise level

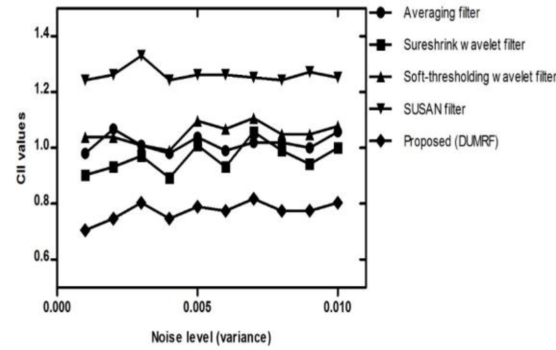

Figure 9. Variation of Contrast improvement index with noise level

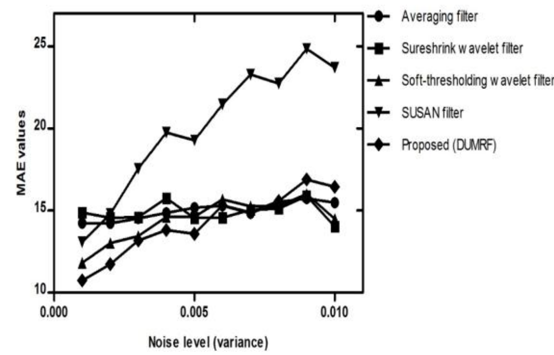

Figure 10. Variation of Mean Absolute Error with noise level

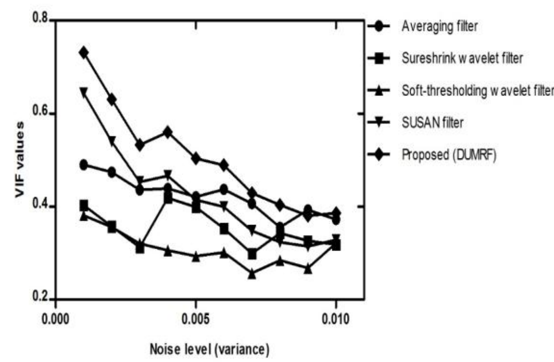

Figure 11. Variation of Visual information fidelity with noise level

Higher values of Peak signal-to-noise (PSNR), Weighed signal-to-noise (WSNR), Contrast Improvement index (CII)and Visual Information Fidelity (VIF) indicate better performance of the algorithm. Figure 7 and 8 shows that the higher values of PSNR and WSNR obtained from proposed algorithm. Figure 10 shows the lower of contrast of proposed algorithm from rest of the algorithm compared. The VIF value in figure 11 is higher in case of the proposed algorithm, that shows that the algorithm is better for denoising.

IV. CONCLUSION AND FUTURE SCOPE

The proposed *DUMRF* is one of the denoising algorithms specifically tailored for denosing of microarray images. As discussed previously denosing is one of the main step in the preprocessing of the microarray image. The *DUMRF* algorithm mainly depends on the optimal selection of c_1 and c_2. The optimal selection of values was obtained using *DOV* algorithm. The DOV algorithm uses non-linear optimization technique to get relevant parameters for DUMRF algorithm. Experimental results and analysis shows proposed method produces promising results compared to other standard denoising method. In the figure 6, it is observed that the MSE increases as the variance increases. The developed algorithm can be applied for the noise models that have lower variance value.

In figure 9, the lower values of CII might be due to the loss of contrast during the processing of the images. That intern directs to the initial values of $c_1$and c_2.One of the future work will include more optimization for obtaining the initial values of $c_1$and c_2. The same trend of cures is seen in figure 10 in case of MSE in figure 6.The algorithm works more optimally for the Gaussian noise models indicated by the second experiment. The work has been demonstrated on a single spot of microarray. The same can be tested on the large micro array images. In conclusion, the DUMRF algorithm is an efficient pre-processing method in microarray image analysis for accurate gene expression profiling.

References

[1] Y. F. Leung and D.Cavalieri, "Fundamentals of cdna microarray data analysis, *"Trends Genet.*, vol. 19, no. 11, pp. 649–659, 2003.

[2] Y.H. Yang., M.I. Buckley, S.Dudoit, and T.P. Speed, "Comparison of methods for image analysis on cDNA microarray data," in *J.Computational and Graphical Stat*, vol. 11, no. 1, pp. 108-136, 2002.

[3] Mastrogianni Aikaterini, Dermatas Evangelos and Bezerianos Anastasios, "Robust pre-processing and noise reduction in microarray images", Proceedings of the fifth International Conference: biomedical engineering, pp.360-364, 2007.

[4] B.Smolka, R.Lukac, K.N.Plataniotis, "Fast noise reduction in cDNA microarray images", 23rd Biennial Symposium on Communications, pp. 348-351, 2006.

[5] Sachin singh et. Al,"Performance Evaluation of Different Adaptive Filters ECG Signal Processing", (IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 05, 2010, 1880-1883.

[6] Zinat Afrose, "Relaxed Median Filter: A Better Noise Removal Filter for Compound Images", International Journal on Computer Science and Engineering (IJCSE), ISSN: 0975-3397 Vol. 4 No. 07 July 2012.

[7] Ross Kindermann and J. Laurie Snell. "Markov Random Fields and Applications," (Contemporary Mathematics v.1) American mathematical society, Providence, Rhode Island.

[8] S. Cheung, Proof of Hammersley–Clifford Theorem, Technical Report, Technical Report, 2008.

[9] Bouman, C.A.& Shapiro, M. (1994) A Multiscale Random Field Model for Bayesian Image Segmentation.IEEE Trans. Image Proc,pp.162-167.

[10] A.Bouman, "Markov random fields and stochastic image models," in IEEE Int. Conf. Image Processing Tutorial, Oct. 1995.

[11] Rafael C. Gonzalez and Richard E.Woods, "Digital Image Processing", Third Edition, PHI Learning Private Ltd., 2008.

[12] Swann, W.H.ScienceDirect, "A Survey on Nonlinear Optimization Techniques", FEBS Letters, vol 2, supplement 1, pp. 39-55, 1969.

[13] A.Forsgren, P. E. Gill, and M. H. Wright, "Interior methods for nonlinear optimization", SIAM Review 44,pp. 525–597,2002

[14] Coleman, T.F. and Y. Li, "An Interior, Trust Region Approach for Nonlinear Minimization Subject to Bounds", SIAM, Journal on Optimization, Vol. 6, pp. 418-445, 1996.

[15] Shanno, D.F., "Conditioning of Quasi-Newton Methods for Function Minimization," Mathematics of Computing, Vol. 24, pp. 647-656, 1970.

[16] Nelder, J. A., and Mead, R., "A simplex method for function minimization," The Computer Journal, vol. 7, pp. 308-313, 1965.

[17] Lagarias, J. C., Reeds, J. A., Wright, M. H., and Wright, P. E., "Convergence properties of the Nelder-Mead simplex method in low dimensions," SIAM Journal on Optimization, vol. 9, pp. 112-147, 1998.

[18] Rafael C. Gonzalez and Richard E. Woods, "DigitalImage Processing", Third Edition, PHI Learning PrivateLtd., 2008.

[19] David L.Donoho and Iain M. Johnstone. "Adapting to unknown smoothness via waveletshrinkage", Journal of the American Statistical Association, pages 1200–1224, 1995.

[20] Donoho, D.L. "De-noising by soft-thresholding," IEEE Trans. on Inf. Theory, 41, 3, pp. 613–627. 1995.

[21] S.M. Smith, J.M. Brady, SUSAN—a new approach to low level image processing, International Journal of Computer Vision 23, pp. 45–78, 1997.

Authors Profile

Mrs.Priya Nandihal is currently pursuing Ph.D in the department of Information Science at SDM college of Engineering and Technology ,Dharwad. She has published 04 papers at international confercences and 02 papers at international jounals. Her main research work focuses on image processing,Cryptography and Artifical Intelligence. She has 8 years of teaching experience and 3 years of Research Experience.She is a life member of ISTE, CSI.

Mrs Vandana S Bhat is an Assistant Professor in the department of Information Science and Engineering at SDM College of Engineering and Technology, Dharwad, Karnataka, INDIA. She received her Ph.D. from Visvesvaraya Technological University, Belagavi, Karnataka, India. She has published 17 papers at International Journal and 03 at conferences. Her main research work focuses on Data Mining,image processing and Artifical intelligence.

Dr. Jagadeesh D. Pujari is a Professor and Head in the department of Information Science and Engineering at SDM College of Engineering and Technology, Dharwad. He received his Ph.D. from GULBARGA UNIVERSITY, Gulbarga, Karnataka, INDIA. He is guiding 06 Ph.D. students at Visvesvaraya Technological University. Two candidates have been awarded Ph.D. under his guidance from VTU Belgaum. He has published 55 papers at International Journal and conferences and 03 at National conferences. He is a life member of ISTE, CSI.