

# Tumor Detection in MRI Images using MultiLevel based Segmentation

Poonam Sengar

Department of Computer Science  
L.N.C.T  
Bhopal, India  
[poonamlikerose@gmail.com](mailto:poonamlikerose@gmail.com)

Prof. Alekh Dwivedi

Department of Computer Science  
L.N.C.T  
Bhopal, India

Prof. Vineet Richhariya

H.O.D, Department of Computer  
Science  
L.N.C.T  
Bhopal, India

*Abstract— Brain Tumor detection in MRI images is an important and challenging research since the structure of the brain images is very complicated and detection of tumor in these image is difficult to achieve. Hence various techniques are implemented for the detection of tumor in brain over MRI images so that the detection can be done easily and quickly. The work done in [1] is the comparative analysis of various segmentation techniques for the tumor detection in MRI images such as using clustering and using Genetic algorithm, but the technique implemented here is suitable for only known images and search spaces area is not very large. Here in this paper a new and efficient technique for the segmentation of brain tumor is implemented using single and multiple phase based region based segmentation. The proposed methodology implemented here provides efficient searching of tumor in MRI images as well as provides efficient search area of tumor in images.*

**Keywords—** MRI Images, Segmentation, Genetic Algorithm, Clustering, Tumor, Random Walker.

## I. INTRODUCTION

Brain Tumor mainly consists of cells that may exhibit unstrained growth in the brain. The growth of brain in nature is malignant for the space and invades of brain tissues for the vital body function. Due to this nature of brain tumor is most important for the detection. Since the intensity of MRI images depends on proton Density which can be determined by the relative concentration of water molecules, the rest of the parameters are T1 and T2 and CSF and relaxation. Magnetic Resonance Imaging is a medical imaging technique. Radiologist used it for the visualization of the internal structure of the body. MRI provides rich information about human soft tissues anatomy. MRI helps for diagnosis of the brain tumor. Images obtained by the M Magnetic Resonance Imaging are used for analyzing and studying the behavior of the brain. Image intensity in magnetic Resonance Imaging depends upon four parameters. One is proton density (PD) which is determined by the relative concentration of water molecules. Ones the brain images acquired they are classified as normal and abnormal.

In medical retrieval system can also provide diagnostic support to physicians or radiologists by providing proper display of relevant past cases to assist them in the decision making process. The picture archiving and communications system (PACS) is a tool that works on medical image databases. Instead of storing images it also capable to provide general analysis. It also suggest past relevant cases for better analysis and references. But retrieval of similar images or records is difficult just because of textual information search is available. The classification task begins with extracting appropriate features of the image. It is one of the most important factors in design process of such system. Visual features were categorized into primitive features such as color, shape and texture [2].

Automated image categorization and retrieval system required efficient algorithm solution for diagnostic-level categorization, this will help radiologist to search the radiographic medical images. In content-based image retrieval systems images are categorized or accessed by their features. The image features means image color, its texture, shape and etc. some examples of content based image retrieval system are QBIC, Photo book, Virage, Visual SEEK, Netra. Maximum system uses data mining specially clustering technique to classify images [3].

Clustering is a method that divides data into groups of corresponding objects. Each group is known as cluster. This cluster consists of objects. Objects are like similar to each other and dissimilar as compared to other groups. In preference to data, fewer clusters essentially lose certain excellent details, hence achieve generalization. It represents many objects by merely some clusters, and as a result this models data with its clusters. Clustering can be unsupervised classification of patterns into sets called clusters. Main goal of cluster analysis is to identifying groups of similar objects and, therefore it helps to determine distribution of patterns with interesting correlations in larger datasets. In order that it can be applied in wide research as it arises in various applications. Moreover, in last year's availability of massive transactional and investigational datasets along with the arising requirements for data mining produced highly need of clustering algorithms to facilitate scale and can be useful in miscellaneous domains [4].

Clustering is pre-treatment part of other algorithms or kind of independent tool used to achieve data distribution, and can be determine isolated points. Commonly used clustering algorithms are CURE, K-MEANS, DBSCAN, and BIRCH. Every clustering method has respective advantages like: KMEANS is simple & easy to comprehend, DBSCAN is capable to filter noises magnificently, and CURE method is insensitive towards input. Although there is no algorithm that can convince any condition particularly as far as large-scale high dimensional datasets is concerned. Hence it is essential to improve and develop new clustering methods [4].

K-Means is a mostly used algorithm that can deal with small convex datasets preferably. K-Means is a method of cluster analysis that partitioned  $n$  observations into 'k' clusters. Within the cluster each inspection belongs to the cluster with nearest mean. Where  $k$  denotes the number of clusters needed, given that a case is allocated to the cluster wherein its distance to the cluster mean is the insignificant. The accomplishment in the algorithm centers on finding the k-means [5].

## II. LITERATURE SURVEY

Automatic classification of medical X-ray images: hybrid generative-discriminative approach was also proposed by Zare et al [6]. Along with rapid progress in the application of local descriptor in pattern recognition, computer vision and image retrieval, the bag of word (BoW) approach has appeared promising for object classification and image retrieval. The classification task begins with extracting appropriate features of the image. It is one of the most important factors in design process of such system. Moreover, the feature extraction step affects all other subsequent processes. The probabilistic latent semantic analysis (PLSA) has been proposed to learn co-occurrence information between elements in the vector space in an unsupervised manner to disambiguate the BoW representation. PLSA can help to disambiguate visual words because of the ability of the PLSA model to generate a robust, high level representation and low-dimensional image representation since PLSA introduces a latent,

A novel approach was proposed to increase the number of classes with a higher accuracy rate by iterative filtering on the training dataset by Zare et al [5]. Filtering is done according to their classification performance. They presented a novel method to achieve classification of class of Image CLEF 2007 medical database [7]. In this scheme they have four iterations steps. These steps hold different classification models. Within the iteration generation process was performed in two steps. The construction of a model from the entire dataset was the first step. This was used to assess filter high accuracy classes (HAC). This will achieve accuracy of 80 % and using this process they can train 20% of the data set. The classes under HAC were only used to construct the classification model under second step. These steps are also continued to next iterations [5].

A novel feature extraction framework for medical x-ray images classification is proposed by Ghofrani et al [8]. As per this scheme, extract centre symmetric local binary patterns from local part of shape and directional information extracted

from images to achieve a set of capable features after some preprocessing. This method worked in three: preprocessing, feature extraction and classification process. Preprocessing is used to eliminate the effects of noises and also manage grey level variation along with set of capable features. After this feature were extracted local parts of each image. This can be done in three parts. At first, in order to achieve local features, each image is partitioned to 25 sub images. This preserve the information of boundaries of sub images, image partitioning to overlapping sub images is preferable to non-overlapping. Secondly Gabor transform computation was applied before extracting features to achieve more shape and directional information. Finally, in the last stage, CS-CBP features are extracted from filtered images. In last stage after the feature extraction, the images are labeled in to their respective classes using multi-class on-against-one algorithm [8].

Avni et al [9] presented an efficient image categorization and retrieval system for medical image database especially for radiographic images. They presented a patch-based classification system that has demonstrated very strong classification rates while also providing efficiency in the retrieval process. This scheme was composed of a feature extraction phase, a dictionary construction based on the training archive, an image representation phase and a classification phase. In very first step images was represented as collection of small patches. After that sampling techniques were applied. Patches along the border of the image are considered as noise and are ignored. The intensity values within a patch are normalized to have zero mean and unit variance. This provides local contrast enhancement and augments the information within a patch. Patches that have a single intensity value of black are ignored. Based on the representative set on images data dictionary was trained. Each image is represented as set of patches. Using the feature extraction parameters that were learned (PCA, feature weights) and the generated dictionary, each image is represented as a histogram of visual words. In this step images are sampled with a dense grid. As per result obtained they conclude with classification results in a lung pathology detection application [9].

In year 2010, a learning-based algorithm for automatic medical image annotation based on sparse aggregation of learned local appearance cues was suggested by Tao et al [10]. They adopted a hybrid approach based on robust aggregation of learned local appearance findings, followed by the exemplar-based global appearance filtering. This scheme is used to detect multiple focal anatomical structures within the medical image. It detects multiple focal anatomical structures within the medical image. This is achieved via learning-by-example landmark detection algorithm. It performs simultaneous feature selection and classification at several scales. After that inconsistent findings through a robust sparse spatial configuration (SSC) algorithm were eliminated. This was consistent and reliable local detections will be retained while outliers will be removed. A reasoning module assessing the filtered findings were applied at last i.e., remaining landmarks is used to determine the final content/orientation of

the image. According to classification task, a post-filtering component using the exemplar-based global appearance check for cases with low classification confidence may also be included to reduce false positive (FP) identifications [10].

### III. PROPOSED METHODOLOGY

1. Take an input dataset of disease images.
2. Find the Histogram of the input image.
3. Now apply PSO-SVM classification approach.
4. Apply Single iteration based and multi level based segmentation classified image.
5. Classify the defected portion in the image

The following algorithm is used for the optimization of SVM.

```

1. Initialize max-iterations and number of
   particle and dimensions.
2. for i= 1:no_of_particles
3.   for j= 1:dimensions
4.     particle_position(i,j) = rand*10;
5.     particle_velocity(i,j) = rand*1000;
6.     p_best(i,j) = particle_position(i,j);
7.   end
8. end
9. for count = 1:no_of_particles
10. p_best_fitness(count) = -1000;
11. end
12. for count = 1:max_iterations
13.   for count_x = 1:no_of_particles
14.     x = particle_position(count_x,1);
15.     y = particle_position(count_x,2);
16.     ker = '@linearKernel';
17.     global p1 ;
18.     p1 = x;
19.     C = y;
20.     trnX=X;
21.     trnY=Y;
22.     tstX=X';
23.     tstY=Y';
24.     [nsv,alpha,bias] =
svmTrain(trnX,trnY,C);
25.     actfunc = 0;
26.     predictedY =
svcoutput(trnX,trnY,tstX,ker,alpha,bias,actfunc);
27.     Result = ~abs(predictedY)
28.     Percent = sum(Result)/length(Result)
29.     soln = 1-Percent
30.     if soln~=0
31.       current_fitness(count_x) =
1/abs(soln)+0.0001;
32.     else
33.       current_fitness(count_x)=1000;
34.     end
   end
End
    
```

Support Vector Machine (SVM) is supervised learning approach which operates on the finding of hyperplane which uses an interclass distance or

margin width for the separation of positive and negative samples. For the unequal misclassification cost a coefficient factor of C+ & C- denoted as 'J' is used for the generation of errors can be outweighs both positive and negative examples. Hence the optimization problem of SVM becomes:

$$\text{minimize } \frac{1}{2} \|w\|^2 + C_+ + \sum_{i,y_i=1} \xi_i + C_- \sum_{j,y_j=-1} \xi_j \quad (1)$$

This satisfies the condition,

$$y_k (wx_k + b) \geq 1 - \xi_k, \quad \xi_k \geq 0$$

Parameters	Explanation
$y_i$	Class labels used in the training dataset
W	Normal to the hyperplane
$ b /  w  $	Perpendicular distance from origin to the hyperplane
$  w  $	Euclidean norm of w
C	Regularization parameter used to find the tradeoff between training error and margin width d
$\xi_i$	Slack variable that allows error in classification [8].

Table 4.1. Various Annotation Used

SVM is implemented in linear and non-linear way, the non-linear form or Radial bias kernel are used for the non-linearly separable data with lagrange multiplier

Hence optimization problem becomes:

$$\text{minimize } w(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1,j=1}^l \alpha_i y_i \alpha_j y_j K(x_i, y_j)$$

Where,

$$C \geq \alpha_i \geq 0 \quad \forall_i, \sum_{i=1}^l \alpha_i y_i = 0$$

Due to the chance of non-linearity and error SVM is based on black box models. For the classification of medical diabetes mellitus a final decision is crucial requirement by the end users. Hence Feature Extraction is implemented for the exact working of the SVM.

$\gamma_{1,2}$	Random numbers on the interval [0,1] applied to ith particle.
----------------	---

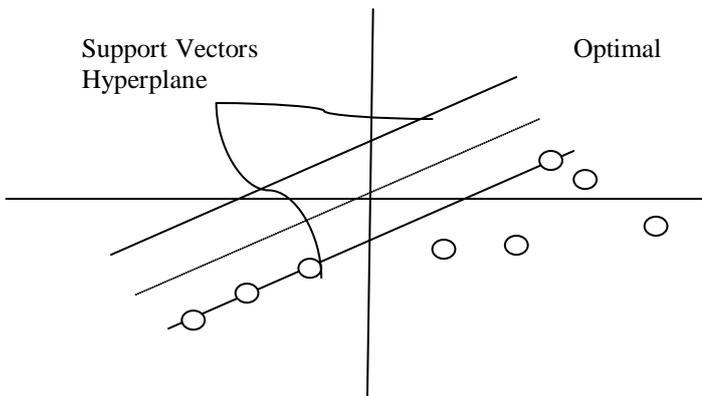


Figure 4.1 Basic Architecture of Linear SVM

Particle Swarm Optimization (PSO) is an efficient Intelligent based optimization technique developed by Eberhart and Kennedy [15], [16] in 1995. Particle Swarm Optimization (PSO) is easier to implement and it is easy the parameters of PSO. Particle Swarm Optimization (PSO) is also used for maintaining the variety of swarm [17].

The Basic form of Particle Swarm Optimization (PSO) consists of the moving velocity of the form:

$$V_i(k+1) = V_i(k) + \gamma_{1i}(P_i - X_i(k)) + \gamma_{2i}(G - X_i(k))$$

And accordingly its position is given as:

$$X_i(k+1) = X_i(k) + V_i(k+1)$$

Where,

Table 4.2. Basic Parameter or Notations of PSO

Parameter	Summary
I	Particle Index
K	Discrete time index
V	Velocity of the ith particle
X	Position of ith particle
P	Best position found by ith particle
G	Best position found by swarm

### Pseudo Code for Image Classification using PSO based SVM

Start with the Initialization of Population

While! ( Ngen || Sc)

For p=1 :Np

If fitness Xp > fitness pbestp

Update pbestp = Xp

For

If fitness Xk > gbest

Update gbest = Xk

Next K

For each dimension d

$$v_{pd}^{new} = w * v_{pd}^{old} + c_1 * rand_1 * (pbest_{pd} - x_{pd}^{old}) * (gbest_d - x_{pd}^{old})$$

$$v_{pd} \in$$

(8)

$$v_{pd} = \max(\min(V_{max}, v_{pd}), V_i)$$

(9) (5)

$$x_{pd}$$

(6)(10)

Next d

Next p

Next generation till stop

The particles are first encoding into a bit string  $S=F1F2\dots Fn$ ,  $n=1,2\dots m$  and the bit {1} represents for the selected feature from the dataset and the bit string {0} is the non-selected feature from the dataset. The evaluation parameters can be computed using SVM. Let us suppose in the dataset the available feature set is 10 then set {F1F2F3.....F10} is then analyzed using PSO and selection of any number of features say 5 a dimensional evaluation of these 5 features is computed using SVM. Each particle in PSO is renewed using adaptive computation of SVM, hence on the basis of which pbest is chosen. Now for the final feature selection each of the particle is then updated according to operation.

$$v_{pd}^{new} = w * v_{pd}^{old} + c_1 * rand_1 * (pbest_{pd} - x_{pd}^{old}) + c_2 * rand_2 * (gbest_d - x_{pd}^{old})$$

$V_{max}$	(11)	factor 2 Maximum Velocity
-----------	------	------------------------------

$$S(v_{pd}^{new}) = \frac{1}{1 + e^{-v_{pd}^{new}}}$$

$$(rand < S(v_{pd}^{new}))$$

$$x_{pd}^{new} = 1 \text{ else}$$

(13)

Table 4.3. Various Notations used in Pseudo Code

Parameter	Summary
Ngen	Number of generations or iterations
Sc	Stopping Criteria
Np	Number of particles
Xp	Current position of pheromone
Pbestp	Pheromone with best fitness
Xk	Current particle position
Gbest	Best fitness value
K	Current particle number
$v_{pd}^{new}$	Updated particle velocity
$v_{pd}^{old}$	Current particle velocity
rand1	Random number 1
rand2	Random number 2
a1	Acceleration factor 1
a2	Acceleration

(12) IV. RESULT ANALYSIS

The table shown below is the analysis and comparison of the existing methodology and the proposed methodology. The analysis is done on three images in which the accuracy of the proposed methodology is better as compared to the existing methodology.

Images	Existing Work	Proposed Work
1	0.89	0.96
2	0.92	0.963
3	0.91	0.97

Table 1. Comparison of Accuracy

The table shown below is the analysis and comparison of the existing methodology and the proposed methodology. The analysis is done on three images in which the Elapsed Time of the proposed methodology is better as compared to the existing methodology.

Images	Existing Work	Proposed Work
1	2.56	1.265
2	3.67	1.428
3	3.41	1.732

Table 2. Comparison of Elapsed Time in sec

The figure shown below is the analysis and comparison of the existing methodology and the proposed methodology. The analysis is done on three images in which the accuracy of the proposed methodology is better as compared to the existing methodology. The two methodologies implemented here for the classification of Disease in MRI Images using Support vector machine and the optimization of Support vector machine using Particle Swarm Optimization is done here and the experimental results are performed on various MRI images on the existing and the proposed methodology. The proposed

methodology implemented here provides better classification of disease in MRI images.

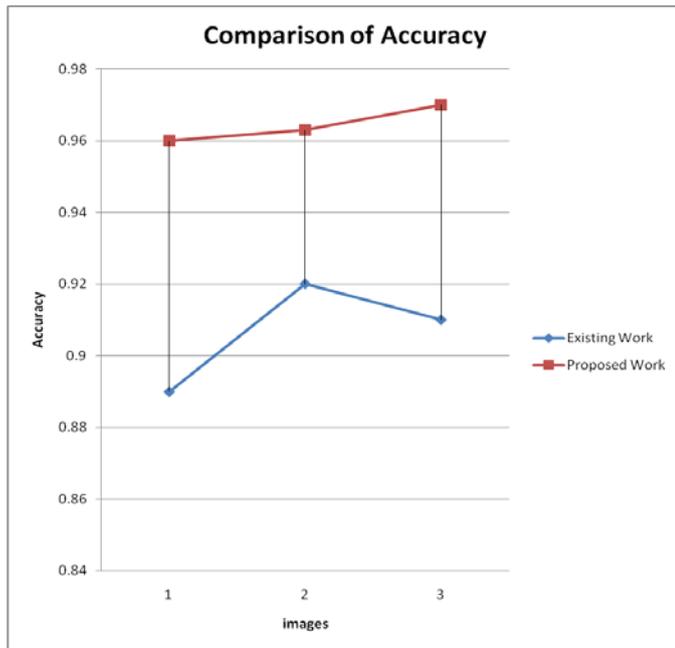


Figure 3. Comparison of Accuracy of the existing and proposed methodology

The figure shown below is the analysis and comparison of the existing methodology and the proposed methodology. The analysis is done on three images in which the Elapsed Time of the proposed methodology is better as compared to the existing methodology. The two methodologies implemented here for the classification of Disease in MRI Images using Support vector machine and the optimization of Support vector machine using Particle Swarm Optimization is done here and the experimental results are performed on various MRI images on the existing and the proposed methodology. The proposed methodology implemented here provides better classification of disease in MRI images.

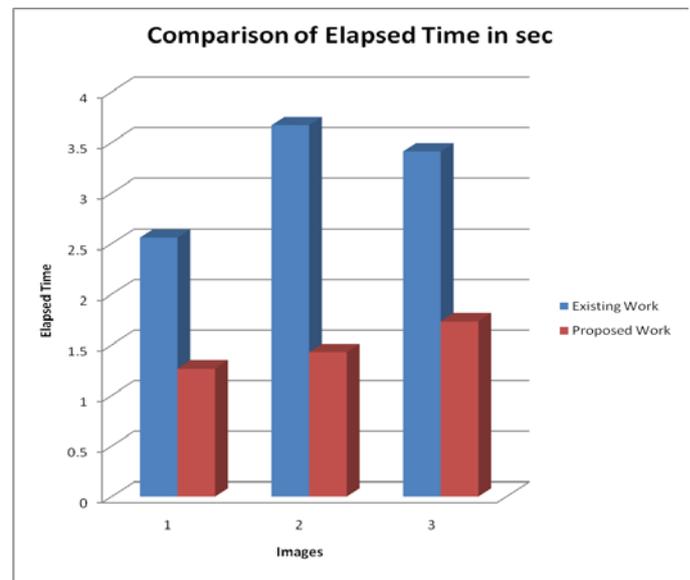


Figure 2. Comparison of Elapsed Time of the existing and proposed methodology

### V. CONCLUSION

The proposed methodology implemented here for the detection of brain tumor in MRI images is efficient in terms of searching of tumor in MRI images as well as it also provides high tumor area in the image. The experiment is performed on various MRI images and compared with some of the existing techniques implemented for detection of brain tumor in MRI images. The proposed methodology outperforms well as compared to other existing techniques.

### References

- [1] Kailash Sinha, G.R. Sinha, "Efficient Segmentation Methods for Tumor Detection in MRI Images", *2014 IEEE Student's Conference on Electrical, Electronics and Computer Science*, 2014.
- [2] Zare, Mohammad Reza, Ahmed Mueen, Mohammad Awedh, and Woo Chaw Seng. *Automatic classification of medical X-ray images: hybrid generative-discriminative approach*. IET Image Processing, vol. 7, no. 5, pp. 523-532, 2013.
- [3] Selvi, S. Malar, and C. Kavitha. *Radiographic medical image retrieval system for both organ and pathology level using bag of visual words*, International Journal of Engineering Sciences & Emerging Technologies, ISSN: 22316604, vol. 6, no. 4, pp. 410 – 416, 2014.
- [4] Ji Dan, Qiu Jianlin, Gu Xiang, Chen Li, and He Peng. *A Synthesized Data Mining Algorithm Based on Clustering and Decision Tree*, IEEE 10th International Conference on Computer and Information Technology (CIT), pp. 2722 – 2728, 2010.
- [5] Zare, Mohammad Reza, Ahmed Mueen, and Woo Chaw Seng. "Automatic classification of medical X-ray images

- using a bag of visual words." IET Computer Vision, vol. 7, no. 2, pp. 105-114, 2013.
- [6] Zare, Mohammad Reza, Ahmed Mueen, Mohammad Awedh, and Woo Chaw Seng. "Automatic classification of medical X-ray images: hybrid generative-discriminative approach." IET Image Processing, vol. 7, no. 5, pp. 523-532, 2013.
- [7] Müller, Henning, Thomas Deselaers, Thomas Deserno, Paul Clough, Eugene Kim, and William Hersh. "Overview of the ImageCLEFmed 2006 medical retrieval and medical annotation tasks." In Evaluation of Multilingual and Multi-modal Information Retrieval, pp. 595-608, Springer Berlin Heidelberg, 2007.
- [8] Ghofrani, Fatemeh, Mohammad Sadegh Helfroush, Habibollah Danyali, and Kamran Kazemi. "Medical X-ray image classification using Gabor-based CS-local binary patterns." In Proceedings International Conference on Electronics, Biomedical Engineering and its Applications, pp. 7-8, 2012.
- [9] Avni, Uri, Hayit Greenspan, Eli Konen, Michal Sharon, and Jacob Goldberger. "X-ray categorization and retrieval on the organ and pathology level, using patch-based visual words." IEEE Transactions on Medical Imaging, vol. 30, no. 3, pp. 733-746, 2011.
- [10] Tao, Yimo, Zhigang Peng, Bing Jian, Jianhua Xuan, Arun Krishnan, and Xiang Sean Zhou. "Robust learning-based annotation of medical radiographs." In Medical Content-Based Retrieval for Clinical Decision Support, pp. 77-88, Springer Berlin Heidelberg, 2010.

IJOART