CHAPTER 2

LITERATURE REVIEW

CHAPTER 2

LITERATURE REVIEW

2.1. INTRODUCTION

A literature review has been conducted on the pertinent published topics by eminent research scholars, authors of the book in relevant areas, owners of patents and investigators (Dena, 2015). There are plenty of work has been done on the MIP especially on medical image segmentation and various applications are coming up with the involvement of cutting edge technologies. Meanwhile, there are research labs and IT companies working with the integration of MIP and the next generation technologies. We have carried out the literature review in each area involved in the research to insist on the need of their study. The areas of studies are mentioned as follows.

- Brain tumor and MRI is reviewed in section 2.2.

- Section 2.3 gives the review of segmentation of brain MR images which will enable us to study the segmentation algorithms.

- Different types of segmentation algorithms for brain MR image are studied in section 2.4.

- The study in section 2.5 brings out the importance of Genetic Algorithms in medical image processing. In section 2.5.2, review on Parallel Genetic Algorithms has been emphasized to understand the need for adopting and enhancing it in the research.

- Section 2.6 reviews the implementation of Image Processing in Cloud Computing to support the novel architecture proposed in the research.

2.2. LITERATURE OF BRAIN TUMOR AND MRI

In current times, a brain tumor is one of the most challenging causes for escalation in fatality among the children and adults suffering from this disease. A brain tumor or intracranial neoplasm is a collection of mass accumulated from abnormal cells within the brain. Basically, there are two main types of tumors called malignant and benign tumors (Kleihues P and Ohgaki, 1999). Benign tumors are harmless in nature. Malignant tumors are denoted as cancerous tumors which can be classified as primary and secondary tumors. Primary tumors originate from within the brain itself whereas the secondary tumors are caused by migration of cancerous cells from somewhere else. Secondary tumors are named as brain metastasis tumors.

There has been a study going on in developed countries to understand the death rate of people suffering from brain tumor. Lin et al. (2004) have given in the study that the death rate has increased due to brain tumor over the last thirty years. Normal healthy cells are eaten up by cancerous cells of tumors. These cells spread across brain and instigating swelling in the brain, infection, destruction of the functions of the brain and increasing the pressure within the skull. Since brain tumors may occur in different locations with varying sizes, doctors need more clarity and efficiency to diagnose them clearly.

There are CT scan and MRI scan used to find the problems in the brain when doctors need to proceed with confirmation for further analysis and treatment. As it is well known, the skull is the protector of the brain from any injury. But at the same time, skull prevents healthcare professionals to study the functions of the brain in different health scenario. MRI is capable

of defining the accurate point and position of the tumor and produce images in different planes. When the contrast discrimination is increased in MRI, it can give a still clearer definition of the location of tumor corresponding to key neuroanatomic structures. It is highly essential to treat the patient with a surgical procedure or radiotherapy (Maravilla and Sory, 1986).

MRI has transformed the procedure for diagnosis on the brain. CT and MRI are the great support for the doctors as well as the researchers working on the brain anatomy to have a clear map of brain image without invasive procedures. This revolution of non-invasive procedure of CT and MRI helps doctors to deduce causes of the unusual function of the brain due to different diseases. When there are two different modalities of brain imaging for such as CT (Voros, 2009) and MRI, the question arises why MR images are taken for processing in the research? MRI has its own advantages and overcome the disadvantages of CT to give a better report for analysis (William and Wilbur, 2009).

- First and foremost advantage of MRI is that there is no use of ionizing radiation. So it is highly favoured for children and the patients who need to go through several imaging examinations.

- MRI can give a report of very depth analysis of brain. MRI has the calibre to produce the image of soft tissues with miniature differences in anatomy because it has the sophisticated series of contrast for soft tissues. Because of such sensitiveness, MRI can be highly preferred for abnormalities within the brain itself.

- Physical movement of the patient will have no impact since MRI scanning is not very specific about the imaging plane.

- The risk of instigating possibly lethal allergic reaction by MRI contrast agents is less.

MRI has the advantage of distinguishing between tumor and perifocal derma which is swelling of the brain tissue, whereas CT is not capable of finding the same (Sanghvi, 2009). MRI provisions completely cerebral perfusion imaging which gives a clear portrait of blood flow in the pathway of the vascular network in the brain. So MRI is able to provide a complete radiological evaluation of brain tumors through procedural diagnosis (Sanghvi, 2009).

As it is mentioned earlier, MRI has the calibre to pinpoint the location of tumor more accurately compared to CT. Additionally, MRI better establishes atrophy which is the common biological course of reabsorption and degeneration of tissues, involving the process of programmed cell death called apoptosis, predominantly along the cerebral convexities (Amit and Frederik, 2011).

The study on the brain reveals that the supplier of the brain called the cells in the arteries are so closely bound (Christopher and Derk, 2012). Because of this closely bound nature, even it is difficult for most normal cells in the bloodstream to cross a semipermeable membrane called "blood-brain barrier" which makes a wall between the blood and the cerebrospinal fluid, and constituting a blockade to the cells' passage, units, and bulky molecules.

Table 2.1 Advantages and disadvantages of MRI.

Advantages	Disadvantages
1. Defines well the edema of parenchyma which is the excess accumulation of fluid in the intra or extracellular spaces of the brain.	1. Calcification and bone erosions cannot be found in a much accurate manner.
2. Edema and compression effects can be found with the maximum accuracy.	2. Intraoperative assessment is not possible.
3. It gives better results in detecting mass effects and atrophy.	3. Spatial fidelity is lower. Trying to achieve high spatial fidelity.
4. The neuroanatomical definition is high to facilitate the tissue distinction clearer.	4. It is a time taking procedure.
5. It can detect accurately the tumor vascularity which is the development of new micro-vessels in the dermis at the base of an invasive tumor.	

The emerging advancements in MRI will definitely provide a novel and precise insights in the area of neuro-oncology to plan well for the treatment and manage the patients suffering from brain tumors.

2.3. SEGMENTATION OF BRAIN MR IMAGES

One of the most important objectives of the brain tumor image processing is to distinguish between the normal region and the region where tumor has spread. This process can be termed as segmentation of the image (Gonzalez

and Woods, 1992). There is an incessant and swift pace research going on to construct an efficient fully automated segmentation of brain tissues from magnetic resonance images (MRI) in neurological pathology.

Following are the different tissue types intended for classification by the segmentation of MR images (Kapur et al., 1996).

- Gray Matter (GM),

- White Matter (WM) and

- Cerebrospinal Fluid (CSF).

In the cases of brain tumors with the aggressive nature of spreading, it is a difficult task to determine the neoplastic borders to fix the locality of the tumor. It may lead to negative impact on the accuracy of diagnosis at the macroscopic level to decide the treatment and patient management. So further examination may also be required after segmentation. The size or shape of the tumor, metabolic-related statistics and the deviations in molecules after the therapy may give better insight during supplementary analysis.

Gaensler (1998), could use segmentation to study how the edema of white matter is the reason for the aggressiveness of the tumor. The author realized the use of segmentation by computational-based procedures in such scenario like the diversity of tissue classification and edema since patient management is a crucial task in this situation.

From the study of Kaye et al. (2001), we could understand the complication of heterogeneity-related parameters of brain tumors. The complexity due to the strong bond among these parameters can make the distinction of a low-grade glioma from a malignant tumor very difficult.

28

Levner et al. (2009) have extensively used segmentation technique to analyse the Glioblastoma Multiforme after applying the treatment of chemotherapy. Glioblastoma Multiforme is an aggressive primary brain tumor where processing technique needs extremely high efficiency. Otherwise the results out of processing may be in vain.

Rees et al. (2009) have highlighted in their work that image analysis through segmentation is helpful to understand the progression rate of lower grade gliomas develops in the glial cells of the brain. It enables evaluation of the risk of conversion of this type of tumor state to a malignant one.

In reality, medical images processing with the support of computational techniques is complicated because of the dissimilar and diffuse of tissues and fluids with different patterns in the cerebral with the intention of constructing a diagnosis of disease. Meantime the process of segmentation needs to be highly efficient to deal with the limitations and artifacts of various imaging modalities.

Since the research concentrates on MR images, the following problems should be addressed before segmentation procedure begins with.

- Artifacts are invisible to the human eye, because of a non-uniform intensity pattern observed over an image for a certain tissue in the brain. But it can be made clear only by an automated image processing, which creates a complicated environment for segmentation.

- Bankman (2009) has highlighted the downside impacts Partial Volume Effect (PVE). PVE occurs in MR images because of a pixel that contains more than one tissue, and there is no MR scanner with an adequate resolution to address this issue. This is the reason

different tissue types reflects same pixel intensity at the same time. Due to the PVE, the boundary between the tissues are not clear and it ends with misclassification. There are algorithms evolving to deal with this problem, since PVE has negative impact on the results.

- Intensity inhomogeneity is a usual problem in MR images which are caused by magnetic settings, patients' position, and other factors.

We can understand from the literature that various computational algorithms and procedures are available to segment tumors from the brain. Researchers are working on pre-segment and post-segment procedures also to give most accurate results. The images produced in different contrasts help to identify the tumors more precisely. But at the same time, it adds more complication to choose the image segmentation algorithms.

We are sure from the above segmentation study that it is highly important to understand and choose the different image segmentation algorithms to get the best accurate results in various conditions.

2.4. BRAIN MR IMAGE SEGMENTATION ALGORITHMS

There are various segmentation algorithms proposed and each of them has its advantages and disadvantages. One of them is intensity-based procedures and it is capable of addressing the most significant barriers listed below (Yongyue, Michael and Stephen, 2001).

a) The shape of the hippocampus which is the extended ridges on the floor of each lateral ventricle of the brain can be constantly deformed onto a sphere. It requires more prediction efficiency before starting any treatment based on the image analysis done.

b) When the midline connections between the left and right hemisphere are artificially closed, these two surfaces have the topology of a sphere.

c) Differentiating opposite banks of a gray matter is hard because of the influence of partial volume effect, subject motion, uncertainty in defining borders of tissues, the heterogeneous behavior of brain tumors and artifacts arise from MRI, etc.

It is understood from the literature that the intensity-based procedures are adopted excessively for the segmentation of the medical image. The facts hidden in the intensified voxels i.e. pixels defined in 3D space, are used by these procedures for processing. These intensity-based procedures partition the MR images into a definite number of chunks and each chunk is labeled to portray its association within one of the image chunks partitioned. There are various conventional procedures, classification and clustering procedures and deformable model procedures are being equipped for the brain MR image processing. Largely the conventional procedures are much used in the segmentation of 2-D images for analysis.

From the of Liu et al., 2014, it is understood that most of the brain tumor segmentation algorithms are based on classification and clustering procedures and they are listed below.

- Fuzzy C-Means (FCM)

- K-means

- Markov Random Fields (MRF)

- Bayes

- Artificial Neural Networks (ANN)

- Support Vector Machines (SVM)

- Atlas-based, etc.

However, implementation of conventional methods, as well as the clustering and classification algorithms by different authors for brain MRI processing, are reviewed in the following.

2.4.1. Threshold-Based Procedures

Image histograms have made the foundation for various image processing procedures. The histogram is a discrete function of a digital image and the function enumerates the number of pixels that falls in a certain series of the gray scale. It is a simple procedure and not so difficult to define the best threshold based on the histogram to reach the segmented output.

Otsu (1979) has proposed an efficient methodology which derives a discriminant function and ultimately it gets maximized between class variance. The implementation of this proposed methodology is easy. But the algorithm is not able to give better accurate results with the complex behaviour of a brain tumor.

The statistical properties of a neighbourhood pixels in an image are confined characteristics. They are identified across the image and they are the deciding factors of threshold (Suzuki H. and Toriwaki, 1991). There are various algorithms proposed by different authors to choose the threshold automatically. The algorithms run by the calculated threshold value and separates the intensities of an image to render the optimum result with the nearly uniform region.

Even though algorithms are adopted for correcting bias field effect which is one of the drawbacks evolved with brain MR image before segmentation,

thresholding algorithms still could not work well with the healthy as well as unhealthy overlapping tissues.

Gibbs et al. (1996) have mentioned in their work that the efficiency of using threshold may be improved with global values. Firstly, selection of a global threshold is done manually from trials of an enhanced brain tumor and its surrounding tissues. In the second step, segmentation is refined by combining region-based methods and edge information.

It is evident from the analysis done by Choi et al. (2000) that when the application of single threshold value does not yield better result in segmentation of the image in a global context, the local threshold values for each region with PVE can be considered to segment the image.

Several existing threshold techniques were used by Bhattacharyya and Tai-hoon (2011) to segment brain tumor from MRI. Each of the technique gave dissimilar results in each image. As a result, they have analysed different threshold techniques and edge detection methods, to acquire a fitting result on brain tumor images. It is evident that each threshold-based algorithm suffers from different identified drawbacks.

2.4.2. Region Based Procedures

Region Growing methods are good enough to segment the regions of similar characteristics with optimum accuracy (Johnston and Atkins, 1996).

A study by Zadeh, Windham and Peck (1996) revealed that the region growing method for brain image segmentation is an efficient method because of its simplicity as well as lesser amount of computation time compared to non-region-based algorithms. But the major problem with the

region growing method is PVE which reduces the accuracy of MR brain image segmentation.

In region-based segmentation procedures, pixels of homogenous properties are brought together to form disjoint regions. The properties of pixels to form a region are defined in advance with likeness criterion (Mitchell et al., 1997).

In the "Region Growing" algorithm, segmentation starts with as a minimum of one seed which satisfies the properties of interest. Based on this seed, the neighbourhood seeds of similar characteristics are chosen and the region is formed. The seed to start with segmentation is chosen either manually or by automated procedure (Mark and Alberto, 2002). There are different features like the range of pixel values, contrast, correlation, homogeneity, entropy, energy, shape, texture, intensity, etc. used to identify the similarity in the image. The iteration of the algorithm continues till there is no more pixel for adding to the region.

2.4.3. Clustering and Classification Methods for Segmentation

Unsupervised clustering method has high reproducibility since its results are mainly based on the information of image data itself. The method requires little or no assumption of the model, and the distribution of the image data. The algorithms for unsupervised classification, or clustering of voxels for image segmentation group a dataset according to their similarity. The data set is composed of p-dimensional feature vectors that represent voxel intensities. These intensities are established from T1- or T2-weighted MR images or from local texture measures. Dunn (1973) proposed a very popular unsupervised algorithm namely fuzzy c-means (FCM) clustering algorithm to segment the medical images.

FCM algorithm gives privilege to the user to insist on the number of clusters to be designed. According to Yang (1993), this is a significant feature for the application of FCM algorithms in image segmentation where the accuracy can be improved. The restriction such as the sum of the degree of membership must be equal to one in FCM has been overcome by the Possibilistic Fuzzy c-Means Clustering Algorithm (PFCM) proposed by Krishnapuram and Keller (1993). In PFCM, the degree of the member is assigned between 0 and 1 depends on their closeness to the centre of the clusters.

The algorithms of FCM type have contributed more in the field of MRI brain tumor segmentation research. Philips II et al. (1995) used FCM algorithm to achieve multispectral segmentation in T1-, T2- weighted and proton density images of Glioblastoma Multiforme. They could get neuroanatomic and neuropathologic tissue contrast information with this algorithm which are clinically substantial from raw MR image data. The algorithm improved accuracy significantly in the final tumor segmentation.

Bezdek et al. (1999) gave the improved version of FCM algorithm. The fuzzy-c-means algorithm recommends a fuzzy logic in partitioning clusters where each pixel holds the degree of membership between 0 and 1. In this manner, FCM could overcome the weakness of k-means algorithm where each pixel has rigid membership in the set of clusters.

A knowledge-based fuzzy clustering method was proposed by Fletcher-heath et al. (2001) for multispectral segmentation of brain tumors. The algorithm could extract all the six non-enhancing brain tumors from 31 images of four patients. It could efficiently separate the tumors region from

physically connected CSF regions. At the same time, the efficiency was affected by the drawbacks of FCM algorithms.

Following are the advantages of PFCM (Nikhil et al., 2005):

1) PFCM has eliminated the noise sensitivity deficiency problem of FCM algorithm has more adverse influence in the results.

2) PFCM supports to get rid of the coincident clusters problem of Possibilistic C-Means algorithm.

3) PFCM provides an enhancement to FPCM by eliminating the row sum constraints must be equal to one in FPCM.

PFCM algorithm will work as PCM algorithm when the membership degree value is very high. Because of this high degree value, the influence of memberships on the prototypes will be condensed.

Hassan et al. (2009) have proposed a technique for segmenting the brain tumors in 3D magnetic resonance images. The authors used Fuzzy Possibilistic c-Means for classification of tumors. Histogram based FPCM was used to detect and label the tumors in their procedure. The authors have combined contour-based and region-based methods for the segmentation of brain tumors in 3D MR images. The proposed algorithm worked well with five fully enhanced tumors. But the algorithm could not assess spatial relations to other structures around the tumor and the number of images tested is also less.

Chandra et al. (2009) have proposed a Particle Swarm Optimization (PSO) based clustering algorithm. The proposed algorithm identified the centroids of a number of clusters, where each cluster is a group of the similar brain tumor patterns found from MR Images. The results obtained for three

performance measures were compared with those acquired from Support Vector Machine (SVM). The performance analysis has shown that the qualitative results of proposed model are analogous with those obtained from SVM. Furthermore, the different values of PSO control parameters have been selected in order to acquire better results from the algorithm.

Another procedure for segmentation using neural networks technique was used by Badran et al. (2010) to categorize the tumor region precisely in the brain MR images. They adopted the steps of procedure such as preprocessing, image segmentation, feature extraction and image classification to recognise whether the tumor is a benign or malignant one. The Region of Interest (ROI) method was used at the end to locate the tumor area.

Qurat-ul et al. (2010) have proposed a method for brain tumor segmentation and the extraction of malignant tumor region using FCM. The proposed method was built on fuzzy entropy that detects ideal and dynamic threshold in line with the clusters formed by FCM. The authors claim that they could attain 99% accurateness for diagnosis. The proposed algorithm needs multiple steps to get the desired result. For the meantime, there was no clarification on the efficiency of the algorithm with the border pixels and overlapping tissues.

Koley and Majumder (2011) have presented a cohesion based self-merging (CSM) algorithm for the segmentation of brain MRI in order to find the exact region of a brain tumor. CSM (Cohesion Self-Merging) based partitional K-means clustering algorithm has reasonable outcome when matched to other merging procedures. Even though their procedure was much simpler, the tedious computation has its own disadvantage and

cannot even differentiate two cases of overlapping clusters when average complete distances of these two clusters are very nearer to each other.

The sum of degrees of each member in the cluster from fuzzy-c-means must be equal to 1. It is considered as a drawback because of the following reasons (Binsy and Madhu, 2012).

1. In the outlier points, the degree of membership is high and due to which the algorithm has trouble in dealing with outlier points.

2. Depends on the degree of membership only the membership of the pixel in the cluster will be decided. It may cause an adverse impact on the results.

3. The study says that the FCM has hitches in dealing with high-dimensional data sets and a huge amount of prototypes.

4. Even though the initialization of clusters is an advantage, there is a chance for FCM getting easily stuck in local optima.

All the above reviewed segmentation algorithms have their own advantages and disadvantages in different circumstances. Since the medical images especially the brain images are considered as a huge search space due to various visible and invisible parameters, there is still hope to use better algorithms such as Genetic Algorithms to make the search more and more efficient.

2.5 GENETIC ALGORITHMS

Genetic algorithms (GAs) are a novel approach to enhance the efficient search. GAs are constructed on the philosophy of natural selection process invented by Charles Darwin. Natural selection of fittest individuals leads to reach the optimum solution for a given problem. Optimum solution can

be achieved by natural exchange of genetic material between parents and the offspring which is produced from parent genes. The offspring are tested for fitness and the fittest individuals are only endorsed to breed for further production.

Holland (1975) has pioneered in proposing GA with comprehensive optimization method for any kind of complex and huge search space based on the idea of randomized search. Some of the characteristics of iterative procedure were incorporated to reach the desired solution space. Genetic Algorithm has high caliber for tedious search problems without being trapped in a local minimum or maximum values. The operators in GA such as selection, mutation, and crossover are updated repeatedly in a population of individuals defined by some genes, to elucidate the problem. Each individual is assessed by a fitness function which decides the population evolution in order to optimize it.

Several authors have adopted GAs to solve a vast diversity of simple as well as complex problems (Bethke, 1980). Each author has given a distinct methodology with different parameters like fitness function, mutation, and selection schemes. For the reason that the accomplishment of optimization intensely hinges on the parameters selected.

In an image, one genetic material called chromosome may have a part of an image or the whole image incorporating few parameters or the complete range of parameters. Mutation among the selected children chromosomes based on the desired fitness function occurs through crossover in different manners. Migration of individual chromosomes among the various clusters depends on the different strategies involved in the algorithm. Goldberg,

Deb, and Clark (1992) have given a study on GAs that population size plays a vital role in determining the solution of the best quality in serial GAs.

Because of a definite characteristic of unprejudiced optimization of GA, it is used as a guaranteed method of image processing and it constantly acquires admiration in the same field. Since GAs were recognized as a most efficient optimization method when a solution space is large enough, it has been mostly used in the applications in image processing (Kharrat et al., 2010) and other different areas. GAs can be used to get the best solution when the search space is massive, and to reach the solution with reduced execution time. Execution time for processing multi-structural images is one of the most important factors to evaluate the efficiency of an algorithm.

In image processing, GA is used for image processing even from the elementary image contrast to composite filters and deformable models constraints for the better results. GA is used for detecting edges to conclude difficult decisions or classifying identified features during image processing. GA enables the search process in an efficient manner without being stopped in local extremes. Because getting trapped in local extremes will definitely lead to an undesirable solution in image processing.

Genetic Algorithms have been applied to segmentation in image processing predominantly. Since GAs are working with the idea of finding the optimal regions of segmentation with the repeated evolution of populations of image pixels based on the selected criteria, GA has been tailored to reach the regions of interest. Since MIP demands more accuracy and precision, Certainly, GA can work well in the brain MRIs where fitness evaluation is numerically complicated. GA for segmentation has the major advantage of

finding the perfect number of areas of interest from the segmented outcome.

2.5.1 Genetic Algorithms for brain MRI Segmentation

Bhanu and Lee (1995) have used adaptive image segmentation in GA by considering different image acquisition environments to optimize the variables in a segmentation method. They have adopted weighted sum approach where they considered five quality measures into a single, scalar measure of segmentation quality where each quality measure takes the same weight. It is very effective in compensating for the changes observed in the images. The stopping criteria for further genetic evolution are one of the three tests where average fitness function is considered. Migration rate is not considered for covering the complete surface of the image.

Gulsen, Smith and Tate (1995) have combined genetic algorithm and curve fitting to establish a periodical evolution method to reduce hasty arrival to local minima when some terms control others in measure. By reducing the hasty arrival of the search space, GA could give better performance. But there is a probability of weak in solution space due to the migration operator adopted.

There is an application of Genetic Algorithm in image processing by Delignon, Marzouki, and Pieczynski (1997). They used the adaptive versions of SEM, EM, and ICE in the case of "pixel by pixel", segmentation. The authors have mentioned that there is an improvement in efficiency of image processing with genetic algorithm.

Lucia (1998) has made an in-depth analysis of snakes' algorithm parameters optimization by using genetic algorithms. They have used a framework in computer vision called active contour model, also known as

snakes, for medical image segmentation and proposed an energy minimization method evolving from Genetic Algorithms. They address the limitations in the application of snakes to extract a region of interest by exploring the supremacy of GA to succeed these restrictions such as initialization, the existence of multiple minima, and the selection of elasticity parameters.

Yoshimura and Oe (1999) have used GA for segmentation where they combined GA and Kohonen's Self-Organizing Map (SOM) to cluster the textured images.

Cagnoni et al. (1999) brought out the use of GA in elastic contour model to segment medical images. They use GA to optimize energy function derived from snakes' algorithm, thus driving contour into an optimal position. They adopt GA to optimize parameters of the interpolator and coefficients of the edge detector, and consequently on the efficiency of their coupling in the modules of the contour-tracking model such as a non-linear edge detector and an interpolator based on an elastic-contour model.

Singh et al. (2005) have used Fuzzy adaptive radial basis function algorithm for MR image segmentation. It could preserve the sharpness of the images. But it is able to do only one task related to fusion.

Hai-Yan and Jiu-Lun (2008) proposed a technique of Fuzzy Partition Entropy of 2D Histogram and Genetic Algorithm. QGA is selected for an optimal combination of parameters. But it is practically complicated.

Fitness function also takes a major role in GA. Lai and Chang (2009) have proposed an idea that fitness function as an evaluation measure since the process forwards in a hierarchical manner. It is not mentioned in the study of the fitness function adopted in earlier works to compute its consistency.

Kharrat et al. (2010) used hybrid techniques are implemented on a real human brain dataset. The classification accuracy varies from 96.36±1.23 to 97.59±1.2 %, which needs further improvement for analysis.

Kumar, Mehta, and Bhilai (2011) have proposed a technique of Texture Based Tumor Detection and Automatic Segmentation. They have used Seeded Region Growing algorithm. It is efficient for segmentation of brain tumor in MRI, in which it is possible to determine abnormality is present in the image or not. But it takes more time to run. The fine segmentation result was generated by fuzzy C-means algorithm. They adopted GA on the images acquired after SOM to group small regions in a feature space for further analysis.

Probabilistic Neural Network Techniques have been used by Dahab et al. (2012) to detect tumor present in the brain with the idea of PNN (probabilistic neural network) a method is based on LVQ (Learning Vector Quantization).

Another method using GA and SVM (Support Vector Machine) has been addressed by Jafari and Shafaghi (2012) to detect brain tumor. Firstly, they extract a set of features using GA and then SVM works on them to classify.

Nagori, Mutkule, and Sonarkar (2013) have suggested a new approach by comparing various algorithms to achieve correctness in the detection of brain tumor. They applied Genetic Algorithm to narrow down the search in the population with the equal fitness values and then detecting the tumor present in the brain. The authors could not explain the impact of local minima in efficiency.

Since the brain MR image data is huge in size and the number of images are increasing every day, processing time and memory utilization is a challenge before the algorithm developers. As the study shows that GAs are strong enough and efficient with large search space, it would be highly beneficial if GAs are made as Parallel GAs.

From all of the above studies, it has been understood that the accuracy can be increased by addressing the following three main problems found in using GAs especially in brain MR image processing.

1. The size and the number of demes.

2. Topology chose to interconnect the demes.

3. Migration rate to decide the number of demes can migrate among the populations at any given time.

2.5.2 Parallel Genetic Algorithm for Image Segmentation

Following study of different works from various authors exhibits that how Parallel Genetic Algorithms (PGAs) are efficient in search techniques and to overcome the issues arose in GAs.

There are two main types of PGAs namely coarse grain and fine grain PGAs. Coarse grain PGAs are most popular ones in the category (Grefenstette, 1984).

Local optimizer was used by Petty, Schomisch, and Born (1991) in the algorithm to make the PGA more efficient. But it is not obvious that the results attained are due to the distributed population or the local optimizer. Still it is realized that the program they designed was efficient and influential.

44

It is an interesting characteristic of PGA that they enable parallelism for sequential GAs along with they exhibit a sort of advanced competence and effectiveness in reaching optimum solution because of their structured population and parallel execution (Pettey, Leuze and Grefenstette, 1987).

PGAs are well required by the applications in the fields of artificial intelligence, numeric and combinatorial optimization, business, engineering, etc. because of their robust nature in dealing with highly complex problems to attain optimum results. Following are some of the characteristics of a PGA gives advantages (Enrique and José, 1999):

- Reliability in using depictions of the problem parameters instead of parameters themselves

- Robustness

- Flexibility in customization for a new problem

- Capable of multiple efficient solutions for a given problem.

- Faster.

- Less prone to finding only sub-optimal solutions.

- Accomplished by working with other search techniques in parallel.

PGAs have the advantage of making the procedures on the demes independent from each other because of parallelism. At the Same time, PGAs are more efficient as a result of the division of the whole population as subpopulations to localize competitive selection among the subpopulations. The literature proves that PGAs have the additional features such as higher efficiency (Tanese, 1989) capable of maintenance

in the wider range (Lin, Punch and Goodman, 1994) and making memory and CPU freer.

PGAs are not only sped up the execution but also give excellent numerical performance even when these algorithms are executed on a sole processor (Spiessens and Manderick, 1991). Optimization is achieved by the use of structured population which are either formed as a set of islands or a diffusion grid (Gordon and Whitley, 1993) because of the numerical performance of PGAs. Therefore, it has been proven that even without parallel machines, PGAs can contribute better results as a result of structured-population models (Lozano, 1996). It does not mean that parallel machines have no influence in the result optimization. In the meantime, it is experienced in the studies that hardware parallelization is a supplement to accelerate the process, and it can be achieved through several ways on a given structured-population GA.

Manderick and Spiessens (1989) have unveiled excellent results demonstrated in their work of implementing a parallel genetic algorithm for surface model fitting in 3D space. Still there is a need to address the issue of execution time in different parameters like number of nodes and the topology used.

Mejía-Olvera and Cantú-Paz (1994) have shown in the experiment that parallel genetic algorithm can be a better optimizer than the classical GA. PGAs are made to coincide with the natural process of evolution by incorporating migration operation. In this scenario, PGAs produce numerous isolated subpopulations in parallel and the subpopulations from time to time exchange their best individuals based on the predetermined migration technique.

As it was mentioned, migration of individuals among the subpopulations plays a major role in the optimization of PGA. In this context, Marin, Trelles-Salazar, and Sandoval (1994) proposed a reformed migration technique. In the given architecture, master-salve processors are used where GA on each subpopulation is run on the slave processor and the best partial demes will be sent to master process every so often. Now it is the job of the master process to choose the best-fit individual received from different slave process at a point of time and broadcast to the slaves. The authors have used six nodes on the network and could experience that there is a significant increase in the speed with the smaller number of nodes. The algorithm did not forecast the efficiency when the dataset increases extremely.

Yong, Tianzi and Evans (2002) proposed an active model using a parallel genetic algorithm to segment the lateral ventricles from brain MR images. They have defined an objective function to improve the performance of PGA. They have randomly chosen the number of subpopulations and the subpopulation from which the best fit individuals can migrate to others. The disadvantages of the island model and the stepping stone model are overcome in their proposed model by reducing the communication and make the parallel genetic algorithm behavior more close to nature (Chipperfield and Fleming, 1996).

Jafari and Shafaghi (2012) have used genetic algorithm for feature selection after the segmentation process. They have used PCA also for feature reduction. Since the authors have worked more on the classification of normal and abnormal brains, the power of genetic algorithm could not be realized much in terms of efficiency.

Li and Huang (2012) proposed a Distributed Parallel Genetic Algorithm oriented adaptive migration strategy to improve the convergent speed without affecting the efficiency of the result by considering the benefits of parallel computer architecture and using PGA. There is a need for considering the parameter like communication topology which has greater influence in execution time.

Narayana, Krishna, and Judy (2014) have used optimum population size, mutation rate and selection strategy which is parallelized with MapReduce architecture for finding the optimal conformation of a protein using the two-dimensional square HP model. They could find the result that GA converges is faster than the traditional to the optimum state.

The following table elucidates that the adaptation of PGAs is still an open problem and it gives continuous challenges to the researchers to work with different parameters and operators.

Table 2.2 Characteristics of Parallel / Distributed Genetic Algorithms
(Gabriel and Enrique, 2011)

Algorithm	Author	Year	Main Characteristics
PGA	Pettey, Leuze and Grefenstette	1987	Generational islands on an Intel iPSC hypercube (8 CPUs). Migrate the best. Dynamic Top.
dGA	Tanese	1989	Distributed populations. Good results with 20% of population migrants every 20 generations
GENITOR II	Whitley and Starkweather	1990	Steady-State islands with ranked selection and reduced surrogate crossover
PGA	Mühlenbein	1991	Sub-populations in a circular ladder-like 2-D topology. Migrate the best, local hill-climbing.
SGA-cube	Erickson, Smith and Goldberg	1991	Made for nCUBE2. This is the parallel extension of the well-known simple GA of Goldberg.
PARAGEN ESIS	Michael	1992	Made for the CM-200. This places one individual in every CPU.

PeGAsuS	Ribeiro, Alippi and Treleaven	1993	Targeted for MIMD machines and written in a very high and flexible description language.
GAMAS	Potts, Giddens and Yadav	1994	Uses 4 very heterogeneous species (islands) and quite specialized migrations and genotypes.
iiGA	Lin, Punch and Goodman	1994	Injection island GA with hierarchical heterogeneous nodes and asynchronous migrations.
SP1-GA	Levine	1994	128 steady-state islands on an IBM SP1 machine of 128 nodes. 2-D toroidal mesh. mr=1
DGENESIS	Mejía-Olvera and Cantú-Paz	1994	Free topology, flexible migration, and policies for selection. Implemented with sockets (UDP)
GALOPPS	Goodman	1996	Very flexible. Implemented with PVM and comprising a large number of operators.
GDGA	Herrera and Lozano	1996	Synchronous. Simulated on one processor. Generational. Uses Fuzzy crossover and FP genes.

CoPDEB	Adamidis and Petridis	1996	Every island uses its own probabilities for mutation, crossover, and specialized operators.
MARS	Talbi et al.	1999	Parallel environment with fault tolerance
GDGA	Herrera and Lozano	2000	Coarse Grain. Hypercube topology.
MALLBA	Alba et al.	2004	Efficient general framework.
ParadisEO	Cahon et al.	2004	A general framework for Parallel Algorithms.
JGAP	Helseth et al.	2008	Framework for Genetic Algorithm and Genetic Programming.
apcGA	Luque et al.	2009	Fine Grain PGA.
Assembling DNA Fragments		2011	GA in Bioinformatics. Independent runs.
Dense pixel matching algorithm.	Santos-Paulino, Nebel and Florez-Revuelta	2014	Optimizing a different fitness function for each scanline.
Alternate search drivers	Krzysztof and Mikołaj	2015	Search Drivers for Detection of Retinal Blood Vessels.

When researchers are working with an accuracy of segmentation of medical images particularly in brain images, there is another challenge as it was mentioned earlier, there is a need to look up the huge volume of brain images to be processed to get the knowledge. Even though there are not

many research articles especially on brain MR image processing on a large scale in Hadoop, we could find some articles which can support large scale image processing in Hadoop using MapReduce.

2.6   IMAGE PROCESSING IN CLOUD COMPUTING

Currently, a large amount of information is generated by several image modalities and computational techniques are required to support decisions in tasks related to diagnosis, surgical planning, and evaluation of treatments. Due to the availability of various services in Cloud, healthcare applications are able to reap the benefits such as hospitals scale with easiness, better various resources management, and offering access, viewing, and sharing of medical images across hospitals and laboratories.

Medical Imaging Clouds is a suitable example which offers remote services of medical imaging information technology (IT) over a network.

It may be thought as a difficult process of transporting medical images from one place to another because of their large size varies from 2 to 500 MB even when compressed. Indeed, it needs high bandwidth to transmit such a huge data. But in reality, there is a fast progress in making new devices for imaging like higher slice count scanners and they produce higher resolution images of huge size.

Cloud services can support healthcare industry by Storing, archiving, sharing and accessing images in the cloud. In turn, it is easy to manage data more efficiently and profitably (Medical Imaging in the Cloud, 2012). The cloud enables hospitals to:

- Deal with huge bandwidth images efficiently.

- Use non-proprietary, standards-based, vendor-neutral architecture.

- Expand or contract storage capacity easily as needed.

- Manage authentication, encryption and security protocols.

- Conduct efficient system-wide application upgrades.

- Extend the life of existing infrastructure/investments.

Image visualization and analysis by a radiologist take first place in medical image processing with respect to the PACS concept. Cloud can support PACS by enabling the administration and help PACS to run more competently. To avoid the legal issues of privacy, proprietary data can be converted to vendor-neutral imaging data with software. Cloud can work as the thread that bonds all PACS organised by combining storage from dissimilar PACS, evading migration for new PACS and upgrades, and providing for backup and business continuity.

The methodology of knowing the growth of the tumor in a human brain has leveraged cloud computing architecture and its power to remotely compute the processing of MR image. The entire application designed for the above computational purpose is available to users through cloud network called as Software as a Service (SaaS). Using the architectural understanding of Software as a Service (SaaS) which is one of the cloud computing architecture it is attempted to develop a fast and efficient technique for processing of MR images.

2.6.1 Hadoop / Spark applications for Cloud

Wang et al. (2010) revealed that the proposed high throughput analysis platform significantly enhances the reliability and speed for biomarker research in their paper. There is no indication that the loss of cost in terms of memory and complexity which we need to pay for ultra-fast processing.

Powell et al. (2010) proposed a Scalable Image Processing Framework. The SIPF has the ability to perform a variety of scalable image operations such as cropping, rotation, scaling (bilinear and nearest neighbour interpolation) edge detection, sharpening, convolution (filters) brightness, contrast, and Gaussian blurring. In the given framework huge image obtained from the Mars Exploration Rovers is taken as a small portion of the memory at a time for processing. The paper concentrates more on memory usage with the help of high computing platforms. But it does not reveal how the speed of computing will be affected due to the division of images as small parts and the efficiency of the processing will be improved in such a circumstance.

Potisepp (2013) has brought image processing in his research where he processed with over 250 GB of large-scale image data using Hadoop's MapReduce. The author has used non-iterative algorithms for regular image processing where the time has drastically come down.

Demir and Sayar (2014) have done a case study regarding the Hadoop's efficiency of massive image processing with face recognition applications and suggest the way to optimize Hadoop. They have worked with a huge volume of small size images and optimized Hadoop. Chan et al. demonstrated the application of Graphical Processing Unit (GPU) with the parallel Genetic Algorithm (GA) for analyzing the effective connectivity using fMRI data. The study shows how the processing time has extremely reduced from 30 hrs to 20 hrs. This gives an evidence that there is a high probability while using PGA with current cutting edge software.

There is a scalable framework using Apache Spark and the Resilient Distributed Datasets developed by ETH Zurich (2014) to enable parallel,

distributed, real-time image processing and quantitative analysis. The framework works as an application tool in the cloud with multiple clusters to do filtering, segmentation and shape analysis qualifying data search. The application tests hypothetically over millions of structures with the time frame of an experiment.

Merv and Nick (2015) have given an idea that the plans for deployment of Hadoop in the cloud nearly equal those for on premises operation and an additional 29% intend continued use or new deployment in both places, challenging assumptions new adopters may have about mostly on premises adoption.

Meanwhile, there are MIP applications evolving with the implementation of cloud computing to ease the process by doing it online instantaneously. This can help the doctors, hospitals as well as the patients to get the maximum benefit with low cost and high speed.