ISSN No:-2456-2165

An Automated Hybrid Classification and Prediction Technique (AHCP) for Identifying Brain Tumors from Images

S.Rathnadevi
Research Scholar
PG & Research Dept. of Computer Science
Periyar E.V.R. College (Autonomous)
Tiruchirappalli – 620023

Dr T.N.Ravi
Research Co-ordinator & Asst. Professor
PG & Research Dept. of Computer Science
Periyar E.V.R. College (Autonomous)
Tiruchirappalli - 620023

Abstract: Tumors are abnormal cell growths in the human body. Brain tumours need to be controlled and treated. Detecting brain tumors can be a challenging and especially in its early stages. Detecting tumours non-invasively from medical images is a good option for early detections. There are several hurdles in processing medical images for tumour detection. This paper proposes a novel technique that classifies and predicts tumours in the brain called AHCP. It is an automated technique with an accuracy of 95% in tumour detection.

Keywords:- *Brain Tumors, Image Segmentation, Image Classification, Tumour Prediction, Automated Tumour Detection.*

I. INTRODUCTION

The human brain is a combination of nerve cells. The brain controls human body functions with three major parts namely brain stem, cerebellum and cerebrum. Tumors are cancerous growth of cells that affect the human brain. Brain tumors can be cancerous (malignant) or non-cancerous (benign) Malignant tumours spread the cancerous cells. Brain tumor symptoms depend on the tumour size, location and type. Common signs of tumours include seizure, nausea and vomiting, headache, memory issues, numbness, balance, walking and changes in speech or vision or hearing. Figure 1 depicts a brain tumor.

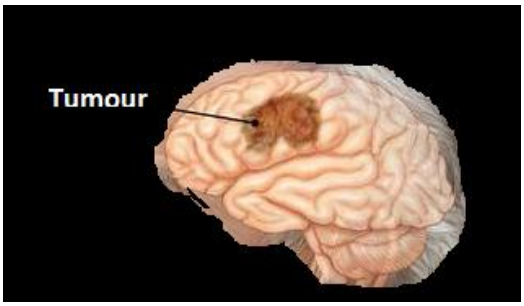

Fig 1:- Tumor in the Brain

Brain tumors can be classified by grades from 1 to 4, four being the most severe. Diagnosis of tumors can be done using neurologic examination or medical imaging. Medical imaging refers to technologies used to view the human body for diagnosis or treatments or monitoring. Information from the area of the body being analyzed can be related to possible diseases for helping treatments of the same. The field of medicinal imaging has gained importance due to mechanizations, non-invasiveness and speed. These technologies have been helpful in advancing image processing, analyses and prediction of diseases. Imaging modalities can generate pictures from different planes, thus allowing specialists to have a close introspect on suspected areas during analysis. It has turned into a generally utilized system for great therapeutic imaging, particularly in mind imaging where delicate tissue contrast and non-obtrusiveness are clear points of interest. X-ray is analyzed by radiologists focused around visual translation of the movies to recognize the vicinity of unusual tissue. Mind pictures have been chosen for the picture reference for this examination on the grounds that the wounds to the cerebrum have a tendency to influence substantial ranges of the human brain part. The classification of information is vital to prune ordinary patients and analyze likelihood of anomalies or tumors in other patients. Thus it is becomes to process and extract information from medical images for identifying tumors in the brain and this paper proposes a novel way of processing images for tumor analysis.

II. RELATED WORK

Several studies have been reported in image processing in all segments of image processing like filtering, segmentation, feature extraction and classification. A few selected studies are listed in this section. Filtering is the primary part of image processing for enhancing image quality. In filtering, images were de-noised with a median filter and salt and pepper/Poisson noise were removed from images, thus preparing them better for further processing. Order statistics filters which combine median and mean filtering were proposed. They are a simple, yet efficient technique for removing noise from medical images. Anisotropic filter was studied in for removing background noise and preserving image edges. A weighted median filter was also proposed for removing high frequency components along with salt and pepper noise from MRI images in Image segmentation is partitioning an image into regions based on identified or

ISSN No:-2456-2165

required features or characteristics. A Watershed segmentation algorithm for brain tumor segmentation was proposed. The algorithm clustered tumor cells using hierarchical clustering. A color based segmentation using k-means clustering identified brain tumor region significantly from pre-processed MR image. A Histogram based Thresholding segmentation detected brain tumor in medical images. Histograms present intensity values of an image, while Thresholding converts images into a binary image and comparisons result in detecting brain tumors. Region based methods have also been used to find abnormalities in images. Feature extraction techniques extract high-level features needed for performing classifications of targets. Various types of features and extraction techniques were studied. GLCM features were proposed for extracting optimal brain tumor features in MRI image. Several factors like Entropy, Homogeneity, . Energy. Colour, Shape. Color. Intensity and Texture were extracted for efficiency. Classification is the labeling of pixels based on multiple features and a critical step in predicting tumors. Brain Tumor classification based on artificial neural network's feed forward technique was proposed. The work classified subjects as normal or abnormal from medical images. A hybrid approach to classify malignant and benign tumors using fusion of texture and shape futures was proposed in [15]. The proposed technique also used region of interest concept to identify tumors.

III. AHCP METHODOLOGY

The proposed AHCP is a novel classification and prediction mechanism for tumours in the brain using image processing. AHCP uses watershed segmentation steps of Gradient magnitude, Label to RGB conversion, Opening and closing, reconstruction and superimposing, in preprocessing medical images. Segmentation in AHCP is achieved using clustering and GLRLM feature set is used to generate a hybrid feature set. Tumour predictions are based the values of this hybrid feature set. AHCP predicts an image with tumour as Disease or not-disease. Figure 2 depicts the AHCP system architecture.

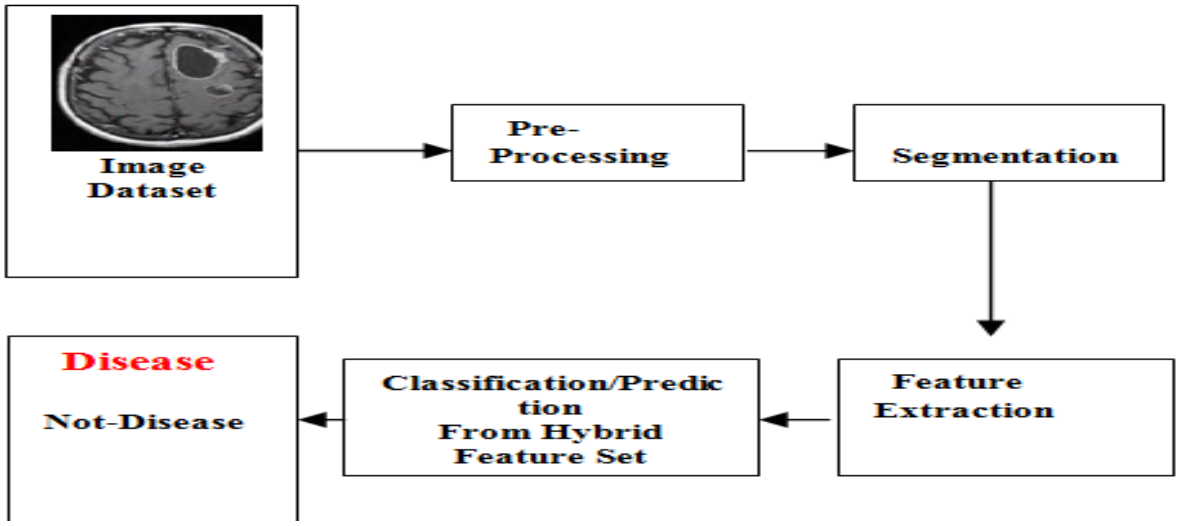

Fig 2:- AHCP system architecture

A. AHCP Image Pre-Processing

Medical images reconstructed from multiple channels of analogous devices may have impurities. Preprocessing is the primary step of AHCP, where Images from an existing image dataset or images collected from patients can be processed by AHCP for enhancements. Many preprocessing techniques are used in AHCP to remove unwanted contents like noise, artifacts and background in images. Its noise reduction procedure improves the quality of images, after removal of artifacts. The steps followed by AHCP are Gradient magnitude, Label to RGB conversion, Opening by Erosion, closing by reconstruction (opening with a closing can remove dark spots and stem marks), Right Ground Marker, Black and White and Background Marker. AHCP pre-processing is exhaustive and results in an enhanced quality image with its edges detected and removed and thus obtaining a clearer picture of the objects. Figures 3 and 4 depict the outputs of AHCP preprocessing steps, while Figure 5 depicts the final output of AHCP pre-processing, where the objects are marked.

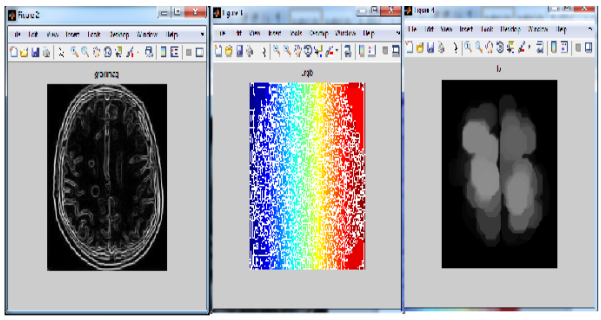

Fig 3:- AHCP Pre-process Output – 1

ISSN No:-2456-2165

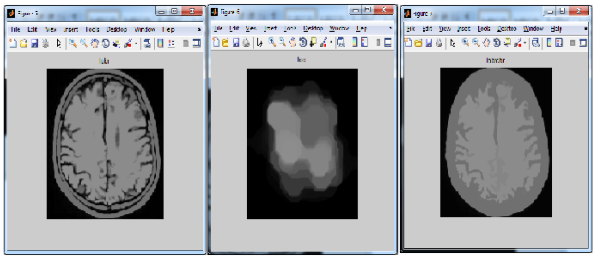

Fig 4:- AHCP Pre-process Output – 2

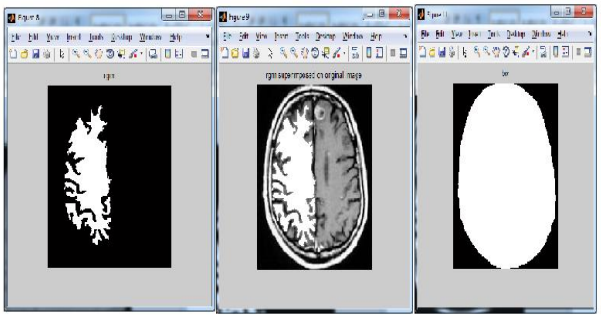

Fig 5:- AHCP Pre-process Output – 3

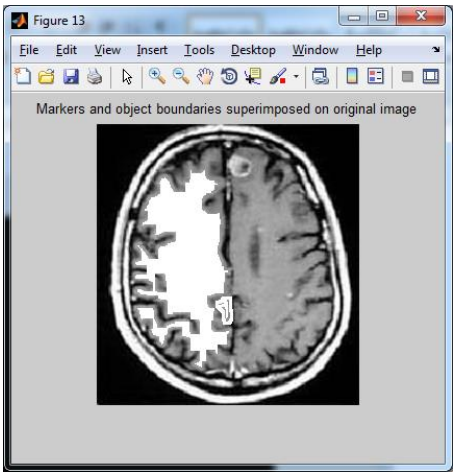

Fig 6:- AHCP Pre-Process Final Output

B. AHCP Segmentation

Segmentation of medical images plays an important for extracting required areas from images. Tumor segmentation from images is a time consuming manual task and is generally performed by medical experts. Various experiments with published benchmarks are required in segmentation [16]. AHCP addresses the challenge of disease identification in segmentation. AHCP uses clustering technique to segment brain images. Clustering is a process of partitioning or grouping a given sector unlabeled pattern into a number of clusters. Similar patterns are assigned to a group, which is considered as a cluster. K-Means is an unsupervised clustering algorithm that classifies the input data points into multiple classes based on their intrinsic distance from each other. The algorithm assumes that the data features form a vector space and tries to find natural clustering in them. The area of the tumor extracted by Ground Truth image is approximately the same as that of the area extracted using AHCP. Figure 7 depicts the output of AHCP Segmentation.

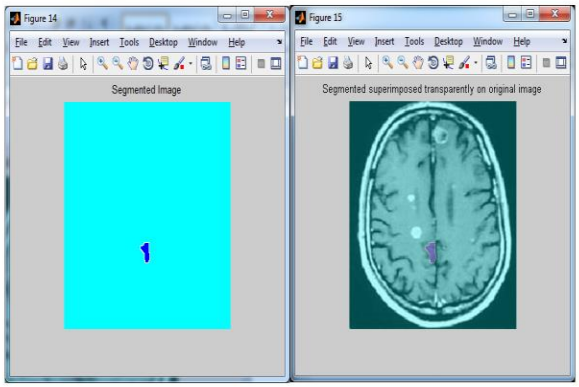

Fig 7:- AHCP Segmentation Output

C. AHCP Feature extraction

The Texture characteristic of an image can be used to describe local spatial variations in image brightness and related to image properties. Numerical manipulation of digitized images helps in quantitative measurements of these factors. The spatial dependence of gray-level values contribute to texture in an image which is a two-dimensional matrix. Structural approaches seek to understand the hierarchal structure of the image. Transform approaches generally perform modifications to the image for obtaining a new response image for analysis. For efficient classifications and predictions features need to be extracted to the diagnostic system, so that it can easily discriminate between the normal and the abnormal parts. AHCP feature extraction is dependent on GLCM and GLRLM features. Grey-Level Run-Length Matrix (GRLM) is a matrix from which the texture features can be extracted for texture analysis [21] It is a way of searching the image, always across a given direction, for runs of pixels having the same gray level value. Run length is the number of adjacent pixels that have the same grey intensity in a particular direction. Gray-level run-length matrix (GLCM) is a two-dimensional matrix where each element is the number of elements j with the intensity i, in the direction θ. Thus, given a direction, the run -length matrix measures for each allowed gray level value how many times there are runs of, for example, 2 consecutive pixels with the same value. Next it does the same for 3 consecutive pixels, then for 4, 5 and so on. Note that many different run-length matrices may be computed for a single image, one for each chosen direction. The GLRLM is based on computing the number of gray level runs of various lengths . A gray level run is a set of consecutive and collinear pixel points having the same gray level value. The length of the run is the number of pixel points in the run. The gray level run-length matrix is depicted as

$$R(\ ) = (g(i, j) \mid\ ),\ 0 \le i \le Ng,\ 0 \le j \le Rmax,$$

where Ng is the maximum gray level and Rmax is the maximum length. Figure 5 shows the sub image with 4 gray

levels for constructing the GLRLM. Figure 8 depicts an image matrix with its GLRLM matrix.

1	2	3	4
1	3	4	4
3	2	2	2
4	1	4	1

Image Matrix

Gray Level	Run Length(j)			
	1	2	3	4
1	4	0	0	0
2	1	0	1	0
3	3	0	0	0
4	3	1	0	0

GLRLM Matrix of the Image

Fig 8:- Image and GLRLM Matrix

Seven texture features can be extracted from the GLRLM. These features uses grey level of pixel in sequence and is intended to distinguish the texture that has the same value of SRE and LRE but have differences in the distribution of gray levels. Once features sets are constructed using Histogram features, GLCM, GRLM, and their combination. Table 1 lists the GLRLM features extracted from images and Table 2 lists the GLCM features.

Image ID	GL Non-uniformity	RL Non-uniformity	Short Runs	Run %	Long Runs
Img1	0.015525	0.01138	0.004645	0.00049	0.001679
Img2	0.022875	0.02392	0.014855	0.00251	0.001179
Img3	0.021625	0.01558	0.009845	0.00229	0.000721
Img4	0.004625	0.01632	0.012845	0.00499	0.000379
Img5	0.045525	0.00442	0.001655	0.00181	0.001679
Img6	0.006975	0.02752	0.001955	0.00011	0.001079
Img7	0.033325	0.02192	0.015945	0.00389	0.000779
Img8	0.043675	0.04138	0.005055	0.00171	0.002121
Img9	0.014725	0.01508	0.000745	0.00021	0
Img10	0.020075	0.01932	0.032355	0.00971	0.002892
Img11	0.005525	0.01238	0.006855	0.00311	0.000179
Img12	0.021625	0.01558	0.009845	0.00229	0.000721
Img13	0.026175	0.00098	0.015055	0.00331	0.001021
Img14	0.021625	0.02622	0.001545	0.00049	0.000179
Img15	0.002875	0.04362	0.013545	0.00459	0
Img16	0.004925	0.00982	0.022745	0.00669	0.001021
Img17	0.043675	0.04138	0.005055	0.00171	0.002121
Img18	0.025675	0.00752	0.001155	0.00049	0.001821
Img19	0.032725	0.02118	0.001055	0.00121	0.001079
Img20	0.029775	0.02568	0.006655	0.00081	0.001721

Table 1:- Difference Table Glrlm Features Set

Image ID	Energy	Entropy	Contrast	IDM	DM
Img1	0.571835	0.906705	0.577235	0.03783	0.0775
Img2	0.034765	0.220705	0.048035	0.01823	0.197
Img3	0.369565	2.180205	0.170635	0.02497	0.1674
Img4	0.344665	1.903905	0.014935	0.00777	0.056
Img5	0.419865	2.365705	0.460365	0.03733	0.3484
Img6	0.159435	0.380305	0.800735	0.01163	0.1118
Img7	0.393265	2.853605	0.501265	0.00243	0.3662
Img8	0.572835	4.862795	0.737165	0.03947	0.6232
Img9	0.078835	0.692605	1.579835	0.01347	0.168
Img10	0.488535	2.792895	0.421635	0.19703	0.0473
Img11	0.240135	0.497305	0.159035	0.01137	0.0574
Img12	0.369565	2.180205	0.170635	0.02497	0.1674
Img13	0.206035	5.006495	0.254335	0.02967	0.3193
Img14	0.443265	2.580805	0.506265	0.01137	0.2888
Img15	0.288865	1.655405	1.111765	0.01277	0.1913
Img16	0.313365	1.653905	0.037635	0.02897	0.0688
Img17	0.572835	4.862795	0.737165	0.03947	0.6232
Img18	0.126535	2.937395	1.155735	0.03667	0.3339
Img19	0.080565	1.945405	0.022465	0.01243	0.3771
Img20	0.040735	1.554395	1.313965	0.03597	0.3022

Table 2:- Difference Table Glcm Feature Set

D. AHCP Classification and Prediction

There are many possible techniques for classification of data. Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) , Support Vector Machines (SVM). Commonly used techniques classify and apply dimensionality reduction for predictions. In the case of AHCP, segmentation and feature set generation becomes the most important base for predictions. Based on the GLCM and GLRLM feature set, a hybrid dataset feature set is created and is listed in Table 3.

ISSN No:-2456-2165

Img ID	GL Non-uniformity	RL Non-uniformity	Short Runs	Run %	Long Runs	Energy	Entropy	Contrast	IDM	DM	Mean Diff
Img1	0.015525	0.01138	0.004645	0.00049	0.001679	0.571835	0.906705	0.577235	0.03783	0.0775	0.220482
Img2	0.022875	0.02392	0.014855	0.00251	0.001179	0.034765	0.220705	0.048035	0.01823	0.197	0.058407
Img3	0.021625	0.01558	0.009845	0.00229	0.000721	0.369565	2.180205	0.170635	0.02497	0.1674	0.296284
Img4	0.004625	0.01632	0.012845	0.00499	0.000379	0.344665	1.903905	0.014935	0.00777	0.056	0.236643
Img5	0.045525	0.00442	0.001655	0.00181	0.001679	0.419865	2.365705	0.460365	0.03733	0.3484	0.368675
Img6	0.006975	0.02752	0.001955	0.00011	0.001079	0.159435	0.380305	0.800735	0.01163	0.1118	0.150154
Img7	0.033325	0.02192	0.015945	0.00389	0.000779	0.393265	2.853605	0.501265	0.00243	0.3662	0.419262
Img8	0.043675	0.04138	0.005055	0.00171	0.002121	0.572835	4.862795	0.737165	0.03947	0.6232	0.692941
Img9	0.014725	0.01508	0.000745	0.00021	0	0.078835	0.692605	1.579835	0.01347	0.168	0.256351
Img10	0.020075	0.01932	0.032355	0.00971	0.002892	0.488535	2.792895	0.421635	0.19703	0.0473	0.403175
Img11	0.005525	0.01238	0.006855	0.00311	0.000179	0.240135	0.497305	0.159035	0.01137	0.0574	0.099329
Img12	0.021625	0.01558	0.009845	0.00229	0.000721	0.369565	2.180205	0.170635	0.02497	0.1674	0.296284
Img13	0.026175	0.00098	0.015055	0.00331	0.001021	0.206035	5.006495	0.254335	0.02967	0.3193	0.586238
Img14	0.021625	0.02622	0.001545	0.00049	0.000179	0.443265	2.580805	0.506265	0.01137	0.2888	0.388056
Img15	0.002875	0.04362	0.013545	0.00459	0	0.288865	1.655405	1.111765	0.01277	0.1913	0.332474
Img16	0.004925	0.00982	0.022745	0.00669	0.001021	0.313365	1.653905	0.037635	0.02897	0.0688	0.214788
Img17	0.043675	0.04138	0.005055	0.00171	0.002121	0.572835	4.862795	0.737165	0.03947	0.6232	0.692941
Img18	0.025675	0.00752	0.001155	0.00049	0.001821	0.126535	2.937395	1.155735	0.03667	0.3339	0.46269
Img19	0.032725	0.02118	0.001055	0.00121	0.001079	0.080565	1.945405	0.022465	0.01243	0.3771	0.249521
Img20	0.029775	0.02568	0.006655	0.00081	0.001721	0.040735	1.554395	1.313965	0.03597	0.3022	0.331191

Table 3:- Hybrid Feature Set

ISSN No:-2456-2165

The predictions of brain tumor are based on the hybrid feature set values. In AHCP analysis a difference of all the features is 0.25. and hence 0.25 is set as the prediction factor. The last column of the hybrid feature set depicts the mean difference based on which brain tumour is identified. In this phase a series of initial processing procedure. In this case >0.25 is a disease <0.25 is not a disease. The final results of AHCP is listed in Table 4, while Figure 9 depicts the screen shot of AHCP prediction.

Img ID	Previous Result	Mean Diff of hybrid features	Prediction Result
Img1	No Disease	0.220482	No Disease
Img2	No Disease	0.058407	No Disease
Img3	No Disease	0.296284	Disease
Img4	No Disease	0.236643	No Disease
Img5	Disease	0.368675	Disease
Img6	No Disease	0.150154	No Disease
Img7	Disease	0.419262	Disease
Img8	Disease	0.692941	Disease
Img9	Disease	0.256351	Disease
Img10	Disease	0.403175	Disease
Img11	No Disease	0.099329	No Disease
Img12	Disease	0.296284	Disease
Img13	Disease	0.586238	Disease
Img14	Disease	0.388056	Disease
Img15	Disease	0.332474	Disease
Img16	No Disease	0.214788	No Disease
Img17	Disease	0.692941	Disease
Img18	Disease	0.46269	Disease
Img19	No Disease	0.249521	No Disease
Img20	Disease	0.331191	Disease

Table 4:- AHCP Prediction Output Table

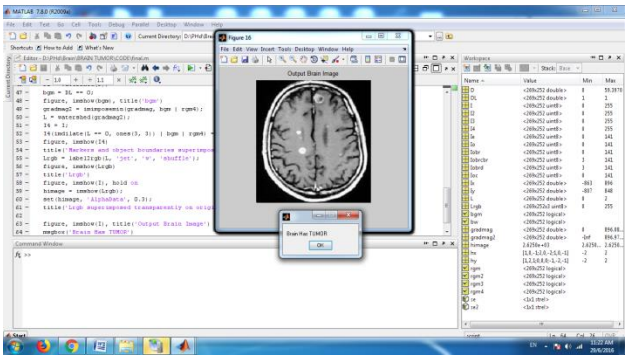

Fig 9:- AHCP prediction Screen Output

IV. PERFORMANCE ANALYSIS

The performance of AHCP was assessed in terms of True Positives (TP) for properly classified positive cases, True Negative (TN) for properly classified negative cases, False Positives (FP) for imperfectly classified negative cases and False Negative (FN) for imperfectly classified positive cases.

True Positive = Disease Predict as a Disease = 12
True Negative = No disease predict as no disease = 7
False Positive = No disease predict as a disease =1
False negative = disease predict as no disease = 0
Accuracy = (tp+tn)/ (tp+fp+tn+fn)
= (12+7)/20
= 0.95

Thus it is evident that AHCP with 95% accuracy is better and can predict tumors in images accurately as early as possible which helps for proper medication.

V. CONCLUSION

Texture analysis is a potentially valuable area in image processing and tumour detection from images. The proposed AHCP system implements a novel classification and prediction mechanism for efficiently analyzing brain tumor images. AHCP involves a separate direction in all stages starting from pre-processing. AHCP is significantly efficient for prediction of tumours in human brain image. The demonstrated performance of AHCP shows its advantages of accuracy, robustness and non-invasiveness. It can be concluded that AHCP is viable and implementable for tumour detection in medical images.

REFERENCES

[1]. T. Kesavamurthy, S. SubhaRani. 2006. “Pattern Classification using imaging techniques for Infarct and Hemorrhage Identification in the Human Brain”, Calicut Medical Journal.
[2]. D. Selvathi, R.S. Ram Prak ash, Dr. S.Thamarai Selvi. 2007. Performance Evaluation of Kernel Based Techniques for Brain MRI Data Classification, International Conference on Computational Intelligence and Multimedia Applications.

[3]. Mohd FauziBin Othman, Noramalina Bt Abdullah, Nurul Fazrena Bt Kamal. 2010. MRI Brain Classification Using Support vector machine, Cairo University, Malaysia.
[4]. D. Bhattacharyya, Tai-hoon Kim. 2011. Brain Tumor Detection Using MRI Image Analysis, Communications in Computer and Information Science, Vol: 151, pp: 307-314.
[5]. E.F. Badran, E.G. Mahmoud, N. Hamdy. 2010. An algorithm for detecting brain tumors in MRI images, Proceedings of the International Conference on Computer Engineering and Systems (ICCES), Cairo, pp: 368-373.
[6]. S.Koley, A. Majumder. 2011. Brain MRI segmentation for tumor detection using cohesion based self merging algorithm, Proceedings of the IEEE 3rd International Conference on Communication Software and Networks (ICCSN), Xi'an, pp: 781-785.
[7]. D. M. Joshi, N. K. Rana, V. M. Misra. 2010. Classification of brain cancer using artificial neural network, Proceedings of the 2010 2nd International Conference on Electronic Computer Technology, 112-116, Kuala Lumpur, Malaysia, May.
[8]. P. Georgiadis, D. Cavouras, I. Kalatzis, A. Daskalakis, G. C. Kagadis, K. Sifaki, M. Malamas, G. Nikiforidis,E. Solomou. 2008. Improving brain tumor characterization on MRI by probabilistic neural networks and non-linear transformation of textural features, Computer Methods and Programs in Biomedicine, 89(1): 24-32.
[9]. E. I. Zacharaki, S. Wang, S. Chawla, D. Soo Yoo, R. Wolf, E. R. Melhem, C. Davatzikos. 2009. Classification of brain tumour type and grade using MRI texture and shape in a machine learning scheme, Magnetic Resonance in Medicine, 62(6): 1609-1618.
[10]. P. Narendran, V.K. Narendira Kumar, K. Somasundaram. 2012. 3D Brain Tumors and Internal Brain Structures Segmentation in MR Images, I.J. Image, Graphics and Signal Processing, 1, 35-43.
[11]. Dou, W., Ruan, S., Chen, Y., Bloyet, D., and Constans, J. M. 2007. A framework of fuzzy information fusion for segmentation of brain tumor tissues on MR images, Image and Vision Computing, 25: 164-171.
[12]. Y. Boykov and V. Kolmogorov. 2004. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision, IEEE Transactions on Pattern Analysis and Machine Intelligence, pages 1124-1137.
[13]. A. Lefohn, J. Cates, and R. Whitaker. 2003. Interactive, gpu-based level sets for 3d segmentation, MICCAI, pp. 564-572.
[14]. Kailash, D.Kharat,and PradyumnaKulkarni,"Brain Tumor Classification Using Neural Network Based Methods", International Journal of Computer Science and Informatics, Vol.1, pp. 2231 –5292, 2012.
[15]. Neelam Marshkole, Bikesh Kumar Singh,and Thoke, A.S., "Texture and Shape based Classification of Brain Tumors using Linear Vector Quantization" International Journal of Computer Applications,Vol.30, pp.0975 – 8887, 2011.
[16]. B.Sathya and R. Manavalan ,"Image Segmentation by Clustering Methods: Performance Analysis", IJCA vol 29-No.11,September 2011.
[17]. Siddheswar Ray and Rose H. Turi,"Determination of Number of Clusters in K-Means Clustering andApplication in Colour Im- age Segmentation",School of Computer Science and Software Engineering,Monash University, Wellington Road, Clayton, Vic-toria, 3168, Australia.
[18]. Prof.A.S.Bhide1, Priyanka Patil2,and Shraddha Dhande3," Brain Segmentation using Fuzzy C means clustering to detect tumor Region.", , 1Electronics and Communication Engineering, North Maharashtra University, Jalgaon, India., 2Electronics and Communication Engineering North Maharashtra University, Jalgaon, India ,3Electronics and Communication Engineering, Vishwakarma Institute of Technology, Pune, India, ISSN: 2277 – 9043 International Journal of Advanced Research in Computer Science and Electronics Engineering Volume 1, Issue 2, April 2012.
[19]. Mohamed, S.S. and Salama M.M. (2005) 'Computer Aided diagnosis for Prostate cancer using Support Vector Machine' Publication: Proc., medical imaging conference, California, SPIE Vol. 5744, pp. 899-907.
[20]. Haralick, R.M. (1979) 'Statistical and Structural Approaches to Texture', Proceedings of the IEEE, Vol. 67, No. 5, pp. 786-804.
[21]. K.Thangavel, M.Karnan, R.Sivakumar and A. Kaja Mohideen " Ant Colony System for Segmentation and Classification of Microcalcification in Mammograms" AIML Journal, Volume (5), Issue (3), pp.29-40., September, 2005.