Technical Journal, University of Engineering and Technology (UET) Taxila, Pakistan Vol. 24 No. 1-2019
ISSN:1813-1786 (Print) 2313-7770 (Online)

Brain Tumor Detection and Classification Using Geometrical Shapes as Texture Descriptors

M.Z.Zafar[1], S.M.Adnan[2], W.Ahmad[3], J.Ikram[4], J.Rashid[5]

[1,2,3,5] *Computer Science Department, U.E.T. Taxila, Pakistan*
[4] *School of Computer Sciences, Beijing Institute of Graphics Communication, China*
[1] zeeshanuet27@gmail.com

Abstract- The existence of abnormal cells in the brain causes a brain tumor. There are two kinds of the tumor; Low grade or slow-growing tumor and High-grade or fast-growing tumor. Patient's successful treatment depends on the accuracy of tumor detection. Therefore, an automatic system with improved accuracy for tumor detection and classification is required. The proposed method consists of three phases to determine the presence of a brain tumor. Red, Green, and Blue (RGB) input image is converted into gray level image and skull is stripped using image masking in the preprocessing phase. In the second phase, the features are extracted using geometrical descriptors. These geometrical descriptors comprise of three geometrical shapes eclipse, parabola, and hyperbola. The performance of these geometrical shape descriptors is evaluated using Local Ternary Pattern (LTPs) and Local Quinary Patterns (LQPs). Support Vector Machine (SVM) and k-Nearest Neighbors (KNN) classifiers are used for classifying MR Images into the healthy and unhealthy brain. Experiments are performed on the Kaggle brain MR Imaging dataset and results are compared with existing techniques. Our experimental results show that the parabola descriptor achieved 97.5% as compared to other geometrical shape descriptors like eclipse and hyperbola.

Keywords- Brain tumor, MRI, Tumor detection, Tumor classification, Geometrical texture descriptors, SVM.

I. INTRODUCTION

The brain is the major controlling part of the human body. It is a complex massive network comprises of 50-100 billion neurons [1]. The unnecessary mass or abnormal growth of brain cells is known as a brain tumor or in other words, the rise of intracranial pressure due to the presence of intracranial lesion within the skull is called a brain tumor —[2]. Two major classifications of the tumor are Benign and Malignant. The slow growth of abnormal cells lies in the category of the benign tumor while the fast-growing abnormal cells lie in the category of a malignant tumor. A benign tumor is noncancerous in nature which meant that they do not affect the nearby healthy brain tissues while the malignant tumor is cancerous in nature. The malignant tumor may cause immediate death if remained uncured [3]. For tumor detection, doctors either use a Computerized Tomography (CT) scan or MRI. The radio waves and magnetism properties are used in MRI to generate an accurate internal image of the brain. Neurosurgeons prefer MRI because it is capable of capturing even the smallest abnormalities in the brain.

The tumor is curable if it is detected at an early stage. Therefore, it is important for doctors to have an automatic system which could accurately detect the location, size and type of tumor from MR Images. [4] Researchers used geometrical descriptors on pap smear dataset for the detection of abnormal smear cells and on face dataset for classifying pain states. In this paper, we use geometrical descriptors Eclipse, parabola and hyperbola shapes with LTP and LQP for feature extraction to increase the accuracy of existing systems. These shapes are used as Local Binary Pattern (LBP) variants.

Rest of the paper is segmented as: section 2 is a literature review, section 3 is an explanation of methods and techniques which we proposed in this research, section 4 comprises of results and discussions while section 5 is about the conclusion and future work.

II. LITERATURE REVIEW

In recent years the study on detection and classification of brain MRI's has been significantly focused. Many types of research presented various techniques for segmentation and classification of the tumor from MRI, few of them are as follows:

Based on the intensity threshold segmentation method, [5] research presented the comparison between Fuzzy c-means, *k*-means clustering, Otsu segmentation, and region growing. Artificial Neural Network (ANN) is used for classification of the tumor into benign and malignant. Extraction of features is done using Gray Level Co-occurrence matrix (GLCM). They use the dataset of "The Cancer Imaging Archive (TCIA)". Their results show that *k*-means clustering is more accurate and take less execution time than other methods. Otsu segmentation technique is also used in research [6]. The author also presented the technique for tumor labeling in his work.

In [7] tumor is segmented from MRI in two stages, the first stage of segmentation is super pixel zoning and secondary stage of segmentation is discriminative clustering. These two segmentation process collectively improves the correctness and clarity of tumor segmentation. Superpixels are used to reduce computational complexity and capturing redundancy. Neighbor pixels having same properties are grouped together to form a super pixel zone. These clustered superpixels are then passed to discriminative clustering method for accurate segmentation. The features are extracted using a Haar wavelet transform (HWT). It is computationally efficient and easy to implement. Next and final phase is classification. Authors used Adaboost and Random Forest (ADBRF) for classifying tumor in healthy and unhealthy class.

In [8] AdaBoost machine learning algorithm has been proposed for the classification of a brain tumor in MR images. The Research work is divided into three parts; Preprocessing, Feature Extraction and Classification. In the preprocessing stage, they removed unwanted segments from MR Image, converted RGB to grayscale, and segmented the images using threshold segmentation and median filter. In feature extraction, 22 features were extracted using GLCM technique and passed to the AdaBoost classifier. They classified the images into healthy or unhealthy brain. Further unhealthy class is classified into benign and malignant class. The accuracy of their proposed technique was 89.90%.

In [9] Region growing method has been used to detect the tumor from Brain MR Images. The first step of the region growing method is to select the seed point. After that, the Pixels with similar properties were then grouped together. Those groups keep on expanding based on the intensity of pixels. In [10] a similar kind of technique named region filling is used for noise removal. In this method, the regions are filled with pixels based on intensity similarity.

In [11] Genetic Algorithm (GA) has been used for the selection and extraction of features. These algorithms are based on natural selection. They are the subclass of an evolutionary algorithm. They used evolutionary techniques and natural selection for optimization problems. It's an iterative model worked on heuristics. By applying GA few information was missed from the neighboring segments. Curve fitting technique is used over GA to recover the missing information. Support Vector Machine (SVM) has been used for the classification of features.

The use of morphological operators like erosion and dilation has been used to detect the accurate size of the tumor and also for the extraction of some other useful information from MRI [12-14]. Morphological operators basically used the concept of mathematical structuring. In [15] morphological operators are used in post-processing. Features are extracted using the Laplacian of Gaussian (LOG)-Lindeberg method and Harris-Laplace algorithm. LOG is used to find the local maxima of an image. In Harris-Laplace, the features are extracted using Harris measure value and Laplace of an image. Researchers used Convolution Neural Network (CNN) for classification of brain MR Images.

Median filters [14-16] have been used widely by the researchers in the preprocessing stage of brain tumor detection and classification due to the property of preserving edges of the images. Researchers in [16-18] used a median filter in their proposed work for the removal of salt and pepper noise in MR Images. For segmentation, they used Fuzzy c-means and *k*-means clustering. Fuzzy c-means clustering is also known as soft clustering while *k*-means is called hard clustering. They also used particle swarm optimization and a canny edge detector. In [19-21] *k*-means and Fuzzy c-means methods are used for segmentation along with SVM and Artificial Neural Network (ANN) as classifiers. For feature extraction, the researcher used the Gray Level Co-occurrence Matrix (GLCM).

In [22] authors present the method for localization of brain tumor using Bounding Box (BB) and classification of a tumor using SVM. In preprocessing phase, authors used anisotropic diffusion filters for removal of noise from MR Images. In the second phase, the tumor is located using BB method. The left and right lobes of the brain are similar in nature. Bounding Box method gives better results for a symmetrical structure that's why it is used for localization of brain tumor. Finally, SVM is used for classification of brain into normal and abnormal class.

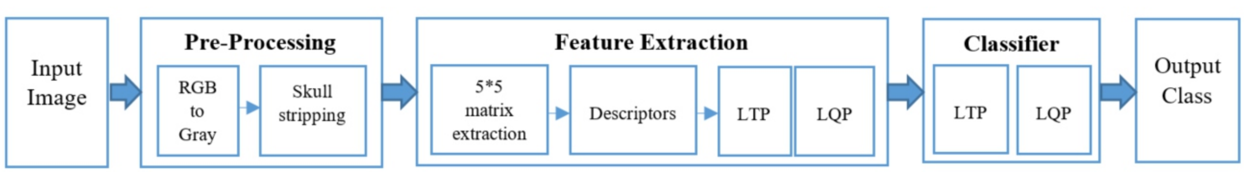


Fig. 1: Flow chart

In [23-24] Gaussian filter is used as a preprocessing stage for the removal of noise from MRI. This filtering is used for smoothness of the image. As it's a low-pass filter so used to remove the noise from the high-frequency areas of the images. The author proposed two techniques for segmentation; Level set method and watershed method. In the level set method, the difference between the pixels is calculated constantly and is based on the principle of the partial differential equation. The level set method is also used [25]. Research presented the comparison between *k*-means and level set method. The experimental results showed that the level set method is more accurate then *k*-means method. Watershed technique works on the region, so it keeps on searching the region of every pixel. This technique is also used by [26] in proposed work. Tumor labeling is also part of their proposed technique.

A pseudocolor or false color translation technique is used by [27] in their paper. In this technique, a grayscale to pseudocolor conversion is done based on the intensity values of pixels. This technique is only useful where we have a single channel image. Thermal imaging is the best example of a pseudo color technique. For segmentation researchers used *k*-means clustering technique. In *k*-means, the pixels are arranged in the clusters based on the lowest mean difference. For contrast adjustment in images, they used histogram equalization.

III. METHODOLOGY

Our proposed methodology consists of 3 stages i.e. preprocessing, feature extraction and classification. The flow chart of the proposed method is shown in Figure 1.

A. Preprocessing

In this phase, RGB images are converted into grayscale and remove the skull from MR Images using image masking technique.

B. Feature Extraction

In this phase, feature extraction is performed using geometrical descriptors using Local Ternary Patterns (LTPs) and Local Quinary Patterns (LQPs). LTP is 3-Value coded (-1, 0 and 1) and it is more sensitive to noise as compared to Local Binary Pattern (LBP) 2-Value code (0 and 1). Local Quinary Pattern (LQP) is 5-value code (-2, -1, 0, 1, 2) and it is more sensitive to noise as compared to both LTP and LBP. Equation 1, 2 and 3 represent the rules for LBP, LTP, and LQP. Here p_c is the center pixel and p is the neighbor pixel. The threshold values of LTP is represented by τ and LQP values are represented by τ_1 and τ_2.

$$f(p, p_c) = \begin{Bmatrix} 0, otherwise \\ 1, p \geq p_c \end{Bmatrix} \tag{1}$$

$$f(\tau, p, p_c) = \begin{Bmatrix} 1, p \geq pc + \tau \\ 0, pc - \tau \leq p < pc + \tau \\ -1, otherwise \end{Bmatrix} \tag{2}$$

$$f(\tau_1, \tau_2, p, p_c) = \begin{Bmatrix} 2, p \geq pc + \tau \\ 1, p_c + \tau_1 \leq p < p_c + \tau_2 \\ 0, p_c - \tau_1 \geq p < p_c + \tau_1 \\ -1, p_c - \tau_2 \leq p < p_c - \tau_1 \\ -2, otherwise \end{Bmatrix} \tag{3}$$

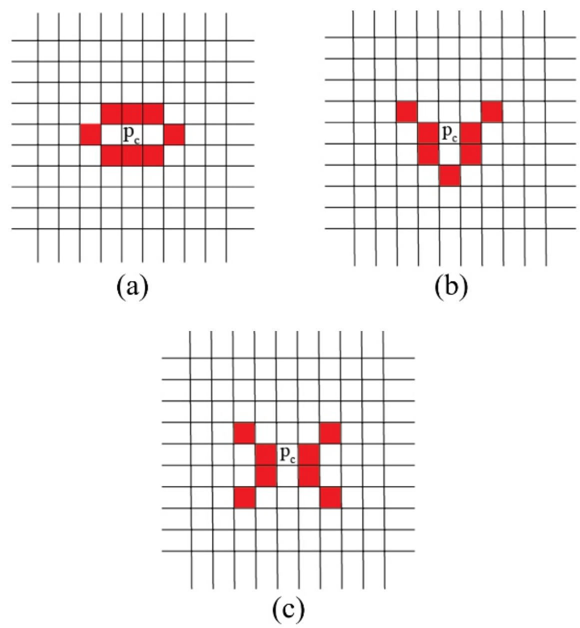


Fig. 2: Descriptors on image matrix (a) Elipse (b) Parabola (c) Hyperbola

Geometrical shape descriptors are used for calculating feature vector from brain MR Images [4]. The geometrical descriptors are elliptical, parabola and hyperbola. Figure 2 shows that neighbor pixels P which forms the geometrical shape of ellipse, parabola, and hyperbola are used in the calculation of feature vectors. The center pixel Pc is subtracted from the neighbor pixel P (shown in red) to form a respective descriptor resultant matrix.

C. Classification

The resultant matrix of each descriptor is passed to LTP and LQP separately and their histograms are calculated for feature vector minimization. These feature vectors are classified through Support Vector Machine (SVM) and K-Nearest Neighbor (KNN). SVM separate out the data items of dissimilar classes by finding hyper-plane and plot them in n-dimensional space according to their class. The $\{(a_1, b_1), (a_2, b_2) \ldots\ldots (a_n, b_n)\}$ are n-samples of the training dataset, where a_i is the data point and b_i is its corresponding label.

The task of SVM is to maximize the distance between hyperplane and data points of distinct classes so that the new data points can be plotted with confidence. The solution of maximizing hyperplane distance is represented in equation 4. Optimal weights W are calculated using Equation 5[28].

$$\max imize\left\{\sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{n}\alpha_i\alpha_j b_i b_j . k(a_i a_j)\right\} \quad (4)$$

$$W = \sum_{i=1}^{n}\alpha_i b_i a_i \quad (5)$$

Where n is the total number of data points, i and j are the i^{th} and j^{th} number of sample, a is the training data point and b is its corresponding label, α is the lagrangian multiplier and k is the scalar constant.

In KNN, the available training data items are stored, and new test data items are classified based on a similarity measure. KNN can be inefficient in some cases because each new test data item must be compared to all training samples. There are two main factors on which the performance of KNN is based, i.e., suitable similarity function and value of *k.* if the value of k is large, then small classes will be overwhelmed by big classes and if *k* is too small then the benefit of KNN algorithm will not be achieved.

The widely used strategies for KNN are given in Equation (6) and (7)[29].

$$C(g_i) = \arg\max_l \sum_{a_j \in kNN} b(a_j, c_l) \quad (6)$$

$$C(g_i) = \arg\max_l \sum_{a_j \in kNN} Sim(g_i, a_j) b(a_j, c_l) \quad (7)$$

Where g_i is the training set, $b(a_j, c_l)$ shows that aj belongs to class c_l, a_j is the neighbor in training set, $Sim(g_i, x_j)$ is the similarity between gi and xj and l is the lth number o sample. Equation (6) indicates that the prediction will be the class that has largest number of members in KNN. Equation (7) shows that class with a maximal sum of similarity will be the winner class.

IV. RESULTS AND DISCUSSION

The experiments are performed on Haier PC Intel core i3 with 3.2 GHz processing speed and 4 GB Ram. We used MATLAB 2018 software on Windows 10 operating system. We used Kaggle brain MRI dataset comprises of 50 MR images of healthy class and 187 images of unhealthy class. The classification is performed using hold-out validation method. The 70 percent images dataset is used for training and 30 percent used for testing. The geometrical shape descriptor performance is measured using accuracy metrics. The accuracy is information about a number of samples classified correctly by the algorithm out of the total number of samples. The accuracy is measured using equation 8.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

TP (True Positive): Positive class samples are truly predicted as positive.
TN (True Negative): Negative class samples are truly predicted as negative.
FP (False Positive): Negative class samples are falsely predicted as positive.
FN (False Negative): Positive class samples are falsely predicted as negative

Table 1 shows the accuracies for SVM and KNN classifier. In eclipse descriptor, LTP and LQP accuracy using KNN is 96.7 percent which is higher than LTP. Parabola accuracy using KNN is 97.5 percent and better than LQP. The KNN accuracy of a hyperbola with LQP is 97.1 percent which is higher than LTP.

TABLE 1: ACCURACY COMPARISON TABLE

Descriptor	Pattern	Accuracy with subspace KNN (%)	Accuracy with SVM (%)
Eclipse	LTP	96.7	94.0
	LQP	96.7	95.6
Parabola	**LTP**	**97.5**	**94.6**
	LQP	95.0	89.6
Hyperbola	LTP	94.4	94.2
	LQP	97.1	92.8

The geometrical descriptor accuracy is compared with GLCM and BWT as shown in table 2. The GLCM gives 93 percent accuracy with KNN as a classifier and BWT gives 90.54 percent accuracy using SVM classifier. The geometrical descriptor parabola gives the highest accuracy of 97.5 percent with KNN classifier.

TABLE 2 COMPARISON WITH OTHER PAPERS

Reference	Year	Technique	Classifier	Accuracy (%)
[30]	2018	Gray Level Co-occurrence Matrix (GLCM)	KNN	93.0
[31]	2017	Berkley Wavelet Transformation (BWT)	SVM	90.54
Proposed	--	Eclipse	KNN	96.7
Proposed	--	Parabola	SVM	97.5
Proposed	--	Hyperbola	SVM	97.1

ISSN:1813-1786 (Print) 2313-7770 (Online)

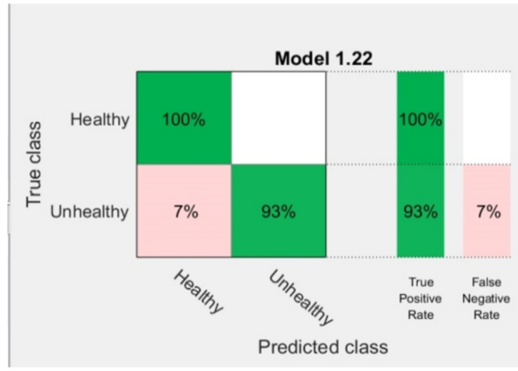
Figure 4 shows the accuracy of each class in the confusion matrix for each geometrical shape descriptors.

V. Conclusion and Future Work

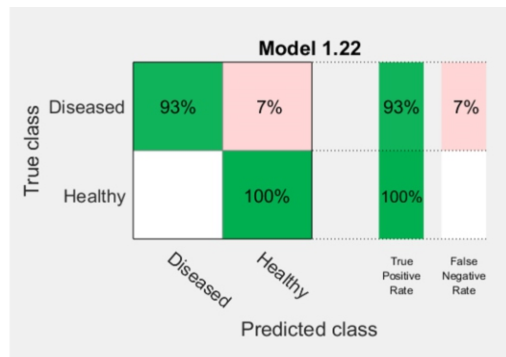
In this paper geometrical shape descriptors are used for extraction of features from brain MRI. Skull is removed from the input image in preprocessing stage to extract more precise features. The extracted features are then used to classify brain MRI into healthy and unhealthy classes. The performance of classification is measured with SVM and KNN classifier and compared with previous methods. The performance of geometrical shape descriptors is also evaluated through LTP and LQP using SVM and KNN.

The experimental results show that eclipse geometrical descriptor achieved higher accuracy as compared to hyperbola and parabola descriptors. The geometrical descriptors are efficiently used for tumor detection in MRI and give better results. Our algorithm gives the accuracy of 97.5% which shows that radiologist can use our system for computerized screening and can also take the best decisions on the basis of the results of our system

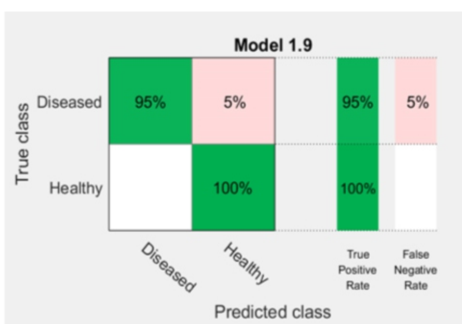
In the future, other geometrical shapes with diverse local patterns may use as a new feature descriptor to achieve more accuracy. Also, different descriptors and classifiers can be used together to achieve more accuracy.



(a)

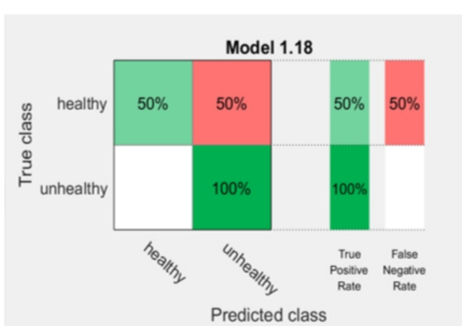


(b)



(c)

(d)



(e)

(f)

Fig. 4: Confusion matrices (a) eclipse with ltp (b) eclipse with lqp (c) parabola with ltp (d) parabola with lqp (e) hyperbola with ltp (f) hyperbola with lqp

REFERENCES

[1] Ackerman S., *Discovering the Brain*. The Development and Shaping of the Brain. 1992, Washington (DC): National Academies Press (US) Available from: https://www.ncbi.nlm.nih.gov/books/NBK234146/.

[2] Y. Du, D. d. Zhao, J. Lu, W. Zhang and Z. Li, *Intracranial lesion as onset symptom in a patient with early undifferentiated connective tissue disease: a case report*. BMC NeurologyBMC series – open, inclusive and trusted 2017, 2017.

[3] F. P. Polly, M. A. Hossain, A. Ayman, and Y. M. Jang, *Detection and Classification of HGG and LGG Brain Tumor Using Machine Learning.* IEEE, 2018.

[4] L. Nanni, S. Brahnam, *Local binary patterns variants as texture descriptors for medical image analy sis*. Elsevier, ScienceDirect, Artificial Intelligence in Medicine, 2010.

[5] A. M. Said, *Comparative Study of Segmentation Techniques for Detection of Tumors Based on MRI Brain Images.* International Journal of Bioscience, Biochemistry and Bioinformatics, January 2018. volume 8, Number 1.

[6] P. N. H. Tra, T. T. Mai, *Image Segmentation for Detection of Benign and Malignant Tumors.* IEEE, 2016: p. pp. 51 - 54.

[7] A. Panda, T. K. Mishra, V. G. Phaniharam, *Automated Brain Tumor Detection Using Discriminative Clustering Based MRI Segmentation.* Smart Innovations in Communication and Computational Sciences, Advances in Intelligent Systems and Computing, Springer Nature Singapore Pte Ltd, 2019.

[8] A. Minz, *MR Image classification using Adaboost for brain tumor type.* IEEE 7th International Advance Computing Conference (IACC) 2017.

[9] Ms. T. P. Shewale, *Detection of Brain Tumor Based On Segmentation Using Region Growing Method.* International Journal of Engineering Innovation & Research 2016. Volume 5(Issue 2).

[10] D. Deb, S. Roy, *A noble approach for noise removal from brain image using Region Filling.* IEEE International Conference on Advanced Communications Control and Computing Technologies, 2014.

[11] S. Khare, and V. Srivastava, *Optimization technique, curve fitting and machine learning used to detect Brain Tumor in MRI.* Proceedings of IEEE International Conference on Computer Communication and Systems ICCCS14, 2014.

[12] T. S. D. Murthy, *Brain tumor segmentation using thresholding, morphological operations and extraction of features of tumor.* International Conference on Advances in Electronics Computers and Communications, 2014.

[13] Hasni, Anu, *Automatic Brain Tumor Tissue Detection in T-1 Weighted MR Images.* International Research Journal of Engineering and Technology (IRJET) Apr-2018. 05(04).

[14] Islam and M.R. Imteaz. *Detection and analysis of brain tumor from MRI by Integrated Thresholding and Morphological Process with Histogram based method*. in *2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)*. 2018. IEEE.

[15] E. F. Badran, E. G. Mahmoud, N. Hamdy, *An algorithm for detecting brain tumors in MRI images.* International Conference on Computer Engineering & Systems, 2010.

[16] N. N. Gopal, M. Karnan, *Diagnose brain tumor through MRI using image processing clustering algorithms such as Fuzzy C Means along with intelligent optimization techniques.* IEEE International Conference on Computational Intelligence and Computing Research, 2010.

[17] J.selvakumar, T.Arivoli, *Brain Tumor Segmentation and Its AreaCalculation in Brain MR Images using K-Mean Clustering and Fuzzy C-Mean Algorithm.* IEEE-International Conference On Advances In Engineering, Science And Management (ICAESM), March 30 2012.

[18] R. C. Patil, D.A.S.Balchandra, *Brain Tumour Extraction from MRI Images Using MATLAB.* International Journal of Electronics, Communication & Soft Computing Science and Engineering, April 2012.

[19] R. Ahmmed, Md. F. Hossain, and Md. A. Rafiq, *Classification of Tumors and It Stages in Brain MRI Using Support Vector Machine and Artificial Neural Network.* International Conference on Electrical, Computer and Communication Engineering (ECCE), February 16-18, 2017.

[20] S. R. Telrandhe, A. Kendhe, *Detection of Brain Tumor from MRI images by using Segmentation &SVM.* IEEE World Conference on Futuristic Trends in Research and Innovation for Social Welfare (WCFTR'16), 2016.

[21] Sharma, G. Purohit, and S. Mukherjee, *Information Retrieves from Brain MRI Images for Tumor Detection Using Hybrid Technique K-means and Artificial Neural Network (KMANN)*, in *Networking Communication and Data Knowledge Engineering*. 2018, Springer. p. 145-157.

[22] S. Polepaka, C.S.Rao, M. C. Mohan, *A Brain Tumor: Localization Using Bounding Box and Classification Using SVM.* Innovations in Electronics and Communication Engineering, Springer Nature singapore 2019.

[23] R. Dubey, S. Vasikarla, *Evaluation of Three Methods for MRI Brain Tumor Segmentation.* Eighth International Conference on Information Technology: New Generations, 2011.

[24] G. Singh, D. M. A. Ansari, *Efficient Detection of Brain Tumor from MRIs Using K-Means Segmentation and Normalized Histogram.* IEEE Issue, 2016.

[25] Mukambika, *Segmentation and Classification of MRI Brain Tumor.* International Research Journal of Engineering and Technology (IRJET), 07 July 2017. 04

[26] Ramya, *Brain Tumor Detection Based on Watershed Transformation.* IEEE-International Conference on Communication and Signal Processing, 2016: p. 0049-0054.

[27] M. N. Wu, C. C. Chang, *Brain Tumor Detection Using Color-Based K-Means Clustering Segmentation.* Third International Conference on Intelligent Information Hiding and Multimedia Signal Processing 2007.

[28] Horng, *The Construction of Support Vector Machine Classifier Using the Firefly Algorithm.* Computational Intelligence and Neuroscience Hindawi, 13 January 2015.

[29] Q. LIN., *An Adaptive k-Nearest Neighbor Text Categorization Strategy.* ACM Transactions on Asian Language Information Processing, December 2004.

[30] Hasni, Anu, *Automatic Brain Tumor Tissue Detection in T-1 Weighted MR Images.* 2018.

[31] N. B. Bahadure, H. P. Thethi, *Image Analysis for MRI Based Brain Tumor Detection and Feature Extraction Using Biologically Inspired BWT and SVM.* Hindawi International Journal of Biomedical Imaging, 2017.