

Human Brain Tumor Detection using K-means Segmentation and Improved Support Vector Machine

Md. Khairul Islam^{1,*}, Md Shahin Ali¹, A Arjan Das¹, D U S Duranta¹, M S Alam²

¹Department of Biomedical Engineering, Islamic University, Kushtia-7003, Bangladesh

²Department of Information and Communication Technology, IU, Bangladesh

E-mail: { shahinbme.iu; abarjan00; duranta03; shahariar.ice12 } @gmail.com

*Correspondence: khairul.ice06@gmail.com

These authors contributed equally to this work.

Abstract

A tumor is the unregulated growth of abnormal cell which increase intracranial pressure within the skull. In recent decades, the human brain tumor is becoming a major cause of death of many people. Medical image processing is the most challenging and innovative field especially magnetic resonance imaging (MRI) for the detection of human brain tumors. In addition, the early stage of detection can cure severe cases and save a life. Moreover, MRI is a noninvasive imaging technique that produces high-quality MR image which is best suited to the detection of abnormal growth e.g., tumor in the brain. In this paper, we proposed a model that includes *K*-means algorithm and an improved support vector machine (ISVM) classifier for detecting brain tumors. In the proposed algorithm, firstly, the MR image is acquired from the database which is reprocessed by the median filter; secondly, the processed image is segmented using *K*-means algorithm; finally, abnormal cell e.g., tumor, is detected using ISVM algorithm. From the experimental result, the proposed algorithm achieved better accuracy when compared to other conventional schemes. In addition, this proposed algorithm reduces the execution time and can detect a tumor in a short time (in a second).

Keywords: segmentation, *K*-means algorithm, improved support vector machine, magnetic resonance imaging

1. Introduction

1.1. Motivation

Digital image processing (DIP) is an emerging field in biomedical engineering and medical science such as tumor detection and classification, cancer detection and classification, and testing and examining are the most significant parts of the human body [1][2]. The brain is the most important and complex structural part of the human body that consists of the 60-100 trillion neurons. A brain tumor is a severe threat to human life [8]. Brain tumor plays a crucial role in medical science.

A tumor is the uncontrolled growth of abnormal cell which increase intracranial pressure within the skull. The early detection and classification of brain tumors are very crucial. Presently, abnormality of brain i.e., tumor, is diagnosed using computer-aided diagnosis (CAD) systems for systematic and specific detection of brain abnormalities. A brain tumor is the unnatural growth of tissue or central spine that can interrupt the proper function of the brain. According to the report of the national cancer institute statistics (NCIS), the death rate on account of brain cancer for the USA is 12,764 per year, 1063 per month, 245 per week, and 34 per day [3][4]. so, it is most significant to diagnose brain tumors in advanced levels to save lives. Moreover, the processes of tumor detection should be done at very high speed and accuracy. This is only possible by using radio imaging techniques such as magnetic resonance (MR) images, computed tomography (CT) scan, positron emission tomography (PET) scan. MRI is a non-invasive technique that produces high-resolution cross-section image which is as usual use for the detection [13][12]. Again, MRI provides a large amount of information so that it is available for detecting tumors. The detection of a brain tumor from radio imaging

is manually done by experts.

However, some problems exist in the manual system, such as it is taking a large amount of time, and segmentation of MR image by different experts may vary significantly. Moreover, the result of tumor detection may vary based on the doctor's experienced, brightness, and contrast of image. For these reasons, the automatic brain tumors detection has become more significant. Automatic detection of brain tumors can increase the probability of survival of a tumor. .

1.2. Contribution

In this paper, we proposed an efficient and fruit full algorithm which helps in the segmentation and detection of the brain tumor as follows:

- We analyzed MRI images and segmented that image using the *K*-means algorithm which produces 8 clusters for the detection of tumors.
- We proposed an improved support vector machine algorithm which can detect brain tumor from the segmented image even if the size of the tumor is very tiny, large, and more complex.
- From the experimental result, it shows that the performance of the proposed scheme obtains better sensing and classification of the MRI image when compared to other conventional schemes like random forest method, region growing, artificial neural network (ANN)

1.3. Organization

The rest of this paper is presented as follows: Section 2. presents the related works. Section 3. presents the proposed algorithm for tumor detection. The experimental re-

sults and discussion are summarized in Sections 4., and 5. presents the conclusion and future works.

2. Related Work

Brain tumor detection is a challenging field from MRI images in medical science. In recent years, the diagnosis of tumor from MRI images is made by the radiologists. It produces inaccurate results of having high rater variability. Then, semi-automatic detection methods like Extreme Learning Machine, K-nearest neighbor (KNN), fuzzy C-means, ensemble classifier (EC) and support vector machine (SVM) were used for the detection process.

In [4], Depa et al had proposed an automated detection process that differentiates normal and abnormal tissue from MRI images. They were used wavelet transform for extracting features and support vector machine algorithms for classification. However, accuracy was good but sensitivity and specificity were poor. In [7], the authors had proposed a model that detects tumors using SVM algorithm. They had used an artificial neural network (ANN) to train the network and SVM was used to segment and detect the tumor. However, the accuracy was not good for complex and tiny regions of the tumor.

In [1], Abdullah et al, were proposed an SVM pattern recognition algorithm that detects and classification tumors based on the symmetry which is exhibited in the axial and coronal images. They got accuracy 65% which is so poor for the detection of abnormal tissue in the brain.

In [15], Usman et al, were proposed a brain tumor segmentation and classification method for the detection of tumors in multi-modality magnetic resonance imaging scans. They had used a random forest classifier for the detection in the MRI image. They detected dice overlap, tumor core region, and enhanced tumor region. However, the result would be improved and the segmentation process was not impressive for tumor detection. To mitigate, these problems an efficient tumor detection method is proposed here which gets better accuracy in the detection of the tumor.

3. Proposed Detection Model

3.1. System Model

The support vector machine is a representation of different classes in hyper-plane in multidimensional space where support vector, hyperplane, and margin are an important concept. Figure 1 shows the architecture of SVM.

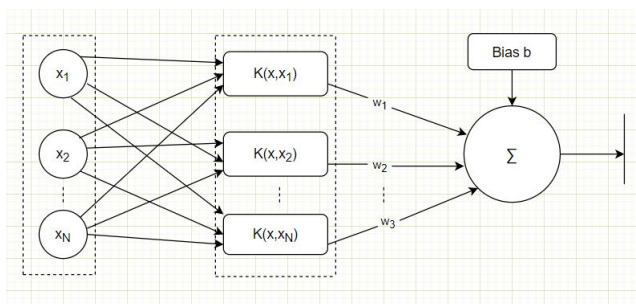


Figure 1: Architecture of support vector machine

3.2. K-means Clustering Algorithm

The K -means clustering algorithm is an unsupervised learning algorithm. It provides a simple way to classify a given data set into a specific number of clusters i.e. a set of data such as $x_1, x_2, x_3, \dots, x_n$ are grouped into K clusters. The main theme behind this algorithm is to define K centers, one for each cluster [11]. The K cluster centers is selected randomly. Distance measure plays a vital role in the performance of the algorithm. Different distance measure techniques are available for this algorithm such as Euclidean distance, Manhattan distance, and Chebychev distance, etc. A proper technique for distance calculation is entirely depend on the type of data that we are going to cluster. Moreover, we will use Euclidean distance as the distance metric because it is fast, robust, and easier to understand [5]. K -means clustering algorithm is described step by step as follows:

Algorithm 1 K -means clustering algorithm.

Assume that, $X = x_1, x_2, x_3, \dots, x_n$ be the data points set and $V = v_1, v_2, v_3, \dots, v_c$ be the set of centers.

- 1: Define number of clusters ' K '.
- 2: Randomly, define cluster centers ' c '.
- 3: Calculate the distance between each data point and cluster centers.
- 4: Data point is assigned to the cluster center whose distance from the cluster center is minimum of all the cluster centers.
- 5: Then , new cluster center is recalculated as follows.

$$V_m = \frac{1}{c_m} \sum_{j=1}^{c_m} x_m \quad (1)$$

where ' c'_m ' is the number of data points in m -th cluster.

- 6: Recalculate the distance between each data point and newly acquired cluster centers.
 - 7: If no data point was reassigned then stop, otherwise repeat steps from 3 to 6.
-

Euclidean distance is used calculated between each pixel to each cluster center. All the pixels of the image are compared individually to all cluster centers using the distance function. The pixel is lead to one of the clusters which are smaller in distance among all. Then, the center is recalculated. Then, every pixel is compared to all centroids again. This process continues until the center converges and the convergence is evaluated through a maximum number of iterations. The quality of clustering of this algorithm is optimized through repetition of K -means several times with different initialization in order to identify best centroids.

3.3. Proposed ISVM Detection Algorithm

SVM is a supervised machine learning algorithm that can be used for classification or regression problems [10]. After founding a hyperplane that completely separates the classes in our training set. We expect that when the new data comes along, it will be similar to the training data. Points that should be classified as one class or the other should lie near the points in training data with the corresponding class. It uses a technique called the kernel trick to transform your data and then based on these transformations it finds an optimal boundary between the possible outputs. The entire methodology is introduced for the detection of tumors in human brain MRI images utilizing K -means clustering AND ISVM algorithm are represented by the subsequent flowchart shown in Figure 2. At first, the acquisition of the human brain MRI image is done from the database and then the input image is processed by the median filter, and enhancement of the MRI image is also carried out. Furthermore, the window is selected and the output of the window has been segmented with the K -means clustering segmentation where 8 clusters are used. Finally, the tumor is identified and detected with a red line marked by ISVM. The proposed algorithm is presented here.

Algorithm 2 ISVM algorithm for tumor detection in Brain MRI image.

1: Define training vectors $x_i \in R^n, i = 1, \dots, l$, in two classes, and a vector $y \in R^l$ such that $y_i \in \{1, -1\}$, C -SVC

2:

$$\begin{aligned} \min_{w, b, \xi} \quad & \frac{1}{2} w^T w + C \sum_{i=1}^l \xi_i \\ & y_i(w^T \phi(x_i) + b) \geq 1 - \xi_i, \\ & \xi_i \geq 0, i = 1, \dots, l. \end{aligned} \quad (2)$$

3: Its dual is

$$\begin{aligned} \min_{\alpha} \quad & \frac{1}{2} \alpha^T Q \alpha - e^T \alpha \\ & 0 \leq \alpha_i \leq C, \quad i = 1, \dots, l, \\ & y^T \alpha = 0, \end{aligned} \quad (3)$$

where e is the vector of all ones, C is the upper bound, Q is an l by l positive semidefinite matrix, $Q_{ij} \equiv y_i y_j K(x_i, x_j)$, and $K(x_i, x_j) \equiv \phi(x_i)^T \phi(x_j)$ is the kernel. Here training vectors x_i are mapped into a higher (maybe infinite) dimensional space by the function ϕ .

4: The decision function is

$$f(x) = \text{sign}(\sum_{i=1}^l y_i \alpha_i K(x_i, x) + b).$$

5: Repeat step 2 to 4 until complete the whole process to recognized object.

6: End

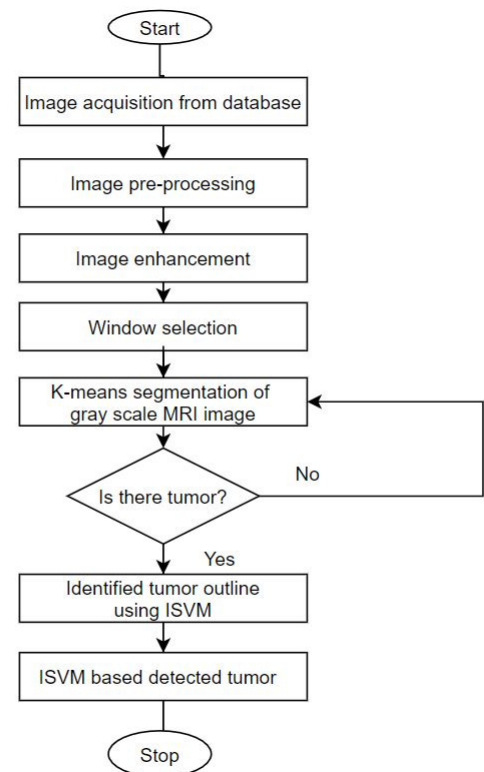


Figure 2: Flow chart of tumor detection in MRI image.

4. Experiment and Result

4.1. Dataset

A database of 20 brain tumor images is shown in Figure 3. The database has been made by collecting different complex brain tumor MRI images. We collected these images from [2], <https://www.mr-tip.com>, and pre-processed for the betterment of application in our algorithmic program. Then, we have processed these images by MATLAB 2016(a) and made the database for final use so that the tumors in these images are detected so easily.

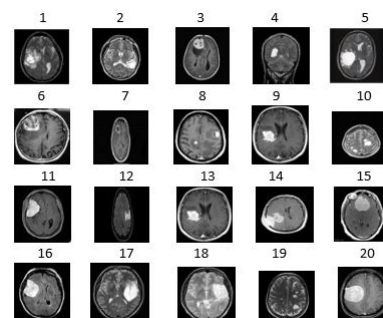


Figure 3: Database of 20 MRI images.

4.2. Results

MR image pre-processing is very significant to detect target region exactly and made easy for further processing. Usually, the dataset that contains image is so poor quality which requires filtering noise and sharpening the image. In the pre-processing step, the acquired image in the dataset is converted into a two-dimensional matrix and the image is converted into an RGB image to a grayscale image. To improve image quality and eliminate the noise, a median filter is used. Then, the enhancement of the image is done by performing the adjusted operation, histogram-based operation, and adaptive histogram-based operation. Generally, enhancement of an image means improving the contrast of the image. Then, in Figure 5(b), there is an initial segmentation of the image using a K-means algorithm which is segmented based on there gray-level intensity and temper of color where $k = 8$. After that, the tumor is filtered by the median filter again. Then, the tumor is detected and marked as a red line using the ISVM algorithm based on the Euclidean distance from the cluster center to each data point which primarily depends on the different features.

Figure 4 shows the filtered image of the input no. 5 image of the database is done by the median filter and enhanced image of the filtered image. Figure 5, shows the segmented image is done by the K-means algorithm from an enhanced image and then detected a tumor using the ISVM algorithm. Figure 6 shows the simulation process and a result of another image no. 17 of the database in incorporate in one window.

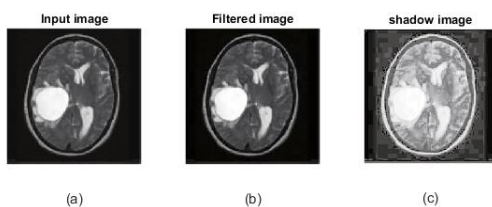


Figure 4: Filtered and gray scale image of input image no. 5 in database.

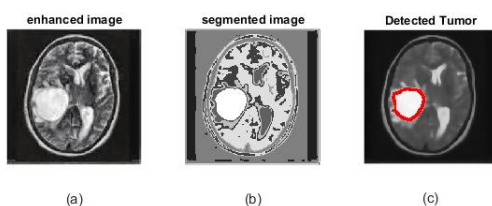


Figure 5: Segmented and detected tumor of input image no. 5 in database.

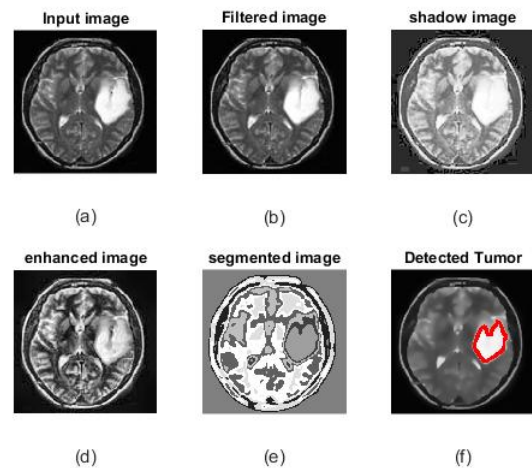


Figure 6: Tumor detection of input image no. 17 in database.

4.3. System Performance

There is some error rate based on detecting or not detecting any abnormal cell in all brain tumor MRI image. These can be measured on the basis of the value that are called true positive, false positive, true negative, and false negative. In this paper, all MRI images in the database have been tested to calculate the system performance e.g., accuracy, sensitivity and specificity. For the evaluation of the system accuracy, sensitivity and specificity the following four attributes are used in the measurement.

- TP (True Positive): The test result is positive in the existence of the objective abnormality and detected correctly.
- TN (True Negative): The test result is negative for non-existing of the objective abnormality and not detected correctly.
- FP (False Positive): The test result is positive for non-existing of the objective abnormality and detected correctly.
- FN (False Negative): The test result is negative for the existence of the objective abnormality and not detected correctly.

Sensitivity can be defined as the measurement of exact identification of the MRI image that do not contains a tumor. The sensitivity, α is defined as follows:

$$\alpha = \frac{TP}{TP + TN} \times 100 \quad (4)$$

Specificity can be defined as the measurement of exact identification of the MRI image that do not contains a tumor. The specificity, β is defined as follows:

$$\beta = \frac{TN}{TN + FP} \times 100 \quad (5)$$

Accuracy can be defined as the measurement of actual classification. The accuracy, η is defined as follows:

$$\eta = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad (6)$$

Table 1: Comparative analysis of tumor detection for the conventional schemes and the proposed ISVM algorithm.

Name of Parameter	Result
TP	18
FP	0
TN	1
FN	1
Sensitivity	94.74%
Specificity	100.00%
Accuracy	95.00%

Table 1 describes the Comparison of computational time among conventional methods and proposed TKFCM method where image size is 256 by 256, processor is Corei5 and the operating system used, is Windows10.

Table 2: Comparative analysis of tumor detection for the conventional schemes and the proposed ISVM algorithm.

SL	Title	Technique	Accuracy
1.	Tumor segmentation from single contrast MR image of human brain [14]	Random forest classification	83.00%
2.	Brain tumor detection using artificial neural networks[6]	ANN algorithm	88.24%
3.	Tumor detection and classification of MRI brain image using wavelet transform and SVM [9]	Wavelet transform and SVM	86.00%
4.	Proposed technique	K-means and ISVM algorithm	95.00%

From Table 2, we observed that the proposed ISVM model achieved a better detection accuracy when compared to other conventional tumor detection techniques.

Table 3: Comparison of computational time t among conventional methods and proposed ISVM method

Name of Algorithm	Time t
Thresholding	3 min
Region Growing	10 min
ANN	7–15 min
FCM	130–140 s
TK-means	4 min
Proposed ISVM	40–60 s

In Table 3, the average required time for different algorithms are described. From the table 3, it can be said that

most of the algorithm take very long time to compute the actual result such as thresholding, region growing and ANN algorithm takes 3, 10 and 7-15 min whereas our proposed algorithm ISVM takes only a 40–60 s.

5. Conclusion and Future Work

In this paper, the performance of the proposed ISVM algorithm acquired better compared with the conventional schemes. Moreover, this algorithm is shown that is an optimum over the conventional schemes. In brain tumor detection, the proposed ISVM algorithm has achieved an accuracy of 95.00%, the sensitivity of 94.74% and specificity of 100% which is better when compared to other conventional techniques.

For future work, we will analyze the tumor detection accuracy for a real-time dataset with extracting high dimension features. In addition, we will apply this model to reduce complexity and executive time in the future.

6. Acknowledgement

I would like to thanks for the students who help us to implement this task. This work was supported by the Department of Biomedical Engineering, Islamic University, Kushtia-7003, Bangladesh.

7. Bibliographical References

- 1 Noramalina Abdullah, Umi Kalthum Ngah, and Shalihah Azlin Aziz. Image classification of brain mri using support vector machine. In *2011 IEEE International Conference on Imaging Systems and Techniques*, pages 242–247. IEEE, 2011.
- 2 Md Shahariar Alam, Md Mahbubur Rahman, Mohammad Amazad Hossain, Md Khairul Islam, Kazi Mowdud Ahmed, Khandaker Takdir Ahmed, Bikash Chandra Singh, and Md Sipon Miah. Automatic human brain tumor detection in mri image using template-based k means and improved fuzzy c means clustering algorithm. *Big Data and Cognitive Computing*, 3(2):27, 2019.
- 3 Anneclaire J De Roos, Nat Rothman, Peter D Inskip, Martha S Linet, William R Shapiro, Robert G Selker, Howard A Fine, Peter M Black, Gary S Pittman, and Douglas A Bell. Genetic polymorphisms in gstm1, p1, t1, and cyp2e1 and the risk of adult brain tumors. *Cancer Epidemiology and Prevention Biomarkers*, 12(1):14–22, 2003.
- 4 AR Deepa et al. Mri brain tumor classification using cuckoo search support vector machines and particle swarm optimization based feature selection. In *2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI)*, pages 1213–1216. IEEE, 2018.
- 5 Nameirakpam Dhanachandra, Khumanthem Manglem, and Yambem Jina Chanu. Image segmentation using k-means clustering algorithm and subtractive clustering algorithm. *Procedia Computer Science*, 54:764–771, 2015.
- 6 Eltahir Mohammed Husein and Dalia Mahmoud Adam Mahmoud. Brain tumor detection using artificial neural networks. 2012.

- 7 T Sathies Kumar, K Rashmi, Sreevidhya Ramadoss, LK Sandhya, and TJ Sangeetha. Brain tumor detection using svm classifier. In *2017 Third International Conference on Sensing, Signal Processing and Security (ICSSS)*, pages 318–323. IEEE, 2017.
- 8 Jin Liu, Min Li, Jianxin Wang, Fangxiang Wu, Tianming Liu, and Yi Pan. A survey of mri-based brain tumor segmentation methods. *Tsinghua Science and Technology*, 19(6):578–595, 2014.
- 9 A Reema Mathew and P Babu Anto. Tumor detection and classification of mri brain image using wavelet transform and svm. In *2017 International Conference on Signal Processing and Communication (ICSPC)*, pages 75–78. IEEE, 2017.
- 10 David Meyer, Friedrich Leisch, and Kurt Hornik. The support vector machine under test. *Neurocomputing*, 55(1-2):169–186, 2003.
- 11 HP Ng, SH Ong, KWC Foong, PS Goh, and WL Nowinski. Medical image segmentation using k-means clustering and improved watershed algorithm. In *2006 IEEE southwest symposium on image analysis and interpretation*, pages 61–65. IEEE, 2006.
- 12 Jainy Sachdeva, Vinod Kumar, Indra Gupta, Niranjan Khandelwal, and Chirag Kamal Ahuja. Segmentation, feature extraction, and multiclass brain tumor classification. *Journal of digital imaging*, 26(6):1141–1150, 2013.
- 13 Kailash Sinha and GR Sinha. Efficient segmentation methods for tumor detection in mri images. In *2014 IEEE Students' Conference on Electrical, Electronics and Computer Science*, pages 1–6. IEEE, 2014.
- 14 Hui Tang, Huangxiang Lu, Weiping Liu, and Xiaodong Tao. Tumor segmentation from single contrast mr images of human brain. In *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, pages 46–49. IEEE, 2015.
- 15 Khalid Usman and Kashif Rajpoot. Brain tumor classification from multi-modality mri using wavelets and machine learning. *Pattern Analysis and Applications*, 20(3):871–881, 2017.