A Learning Based Brain Tumor Detection System

Sultan Noman Qasem\textsuperscript{1,2}, Amar Nazar\textsuperscript{3}, Attia Qamar\textsuperscript{4}, Shahaboddin Shamshirband\textsuperscript{5,6,*} and Ahmad Karim\textsuperscript{4}

Abstract: Brain tumor is one of the most dangerous disease that causes due to uncontrollable and abnormal cell partition. In this paper, we have used MRI brain scan in comparison with CT brain scan as it is less harmful to detect brain tumor. We considered watershed segmentation technique for brain tumor detection. The proposed methodology is divided as follows: pre-processing, computing foreground applying watershed, extract and supply features to machine learning algorithms. Consequently, this study is tested on big data set of images and we achieved acceptable accuracy from K-NN classification algorithm in detection of brain tumor.

Keywords: Magnetic resonance imaging, brain tumor, watershed, segmentation, K-NN classification.

1 Introduction

Brain is the most complex part of human body as well as it control multifaceted functions. It suffers from different severe conditions one of them is brain tumor. The irregular growth of tissues inside the skull causes brain tumor. A certain classification of brain tumor depend on its size and location. Some brain tumors are malignant or benign, and number of other possibilities [Naz and Hameed (2017)]. An estimated ratio of brain tumor was diagnosed in 23,880 adults (10,160 women and 13,720 men) in United States. Similarly, about 3,560 children’s are also affected by brain tumor. Moreover, brain tumor causes approximately 16,830 (7,340 women and 9,490 men) deaths in adults. Aforementioned, statistics are adapted from the National Cancer Institute, the Central Brain Tumor Registry of the United States; and Cancer facts and figures 2018 [Cancer.Net (2018)]. In this context, modern

\textsuperscript{1} Computer Science Department, College of Computer and information Sciences, Al Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, Saudi Arabia.
\textsuperscript{2} Computer Science Department, Faculty of Applied Sciences, Taiz University, Taiz, Yemen.
\textsuperscript{3} Department of Computer Science, Bahauddin Zakariya University, Pakistan.
\textsuperscript{4} Department of Information Technology, Bahauddin Zakariya University, Pakistan.
\textsuperscript{5} Department for Management of Science and Technology Development, Ton Duc Thang University, Ho Chi Minh City, Vietnam.
\textsuperscript{6} Faculty of Information Technology, Ton Duc Thang University, Ho Chi Minh City, Vietnam.
\textsuperscript{*} Corresponding Author: Shahaboddin Shamshirband. Email: shahaboddin.shamshirband@tdtu.edu.vn.
technologies are used to detect brain tumor. The image visualization techniques are MRI (magnetic resonance imaging) and CT (computed tomography) scan images become an active and operational research area [Fink, Muzi, Peck et al. (2015)].

Moreover, medical images also focused on real-time observation and examining the tumor by using established and consistent algorithms. Partitioning and nuclei of cell is one of the most disturbing problem in images diagnosis system. Similarly, undesired section and atmospheric interference exploit the strict image of the portion where tumor lies. Manual techniques in brain tumor detection are more disposed to human faults and it differs from doctor to doctor with respect to its analysis and treatment procedure. Additionally, there is a larger chance that normal brain cells can be jumbled with tumor cells and obtained directly for testing persistence. This process is very unsafe as it endangers to human health.

In this perspective, a favoured image segmentation technique is required rather than viewing the complete MRI in number of experiments. As it saves time for doctors and give more confidence to take applicable decisions. Besides, in existing researches [Sharmila and Joseph (2018); Kumar and Mathai (2017)] authors focused on clustering and classification methods, to detect brain tumor images are well explained in related work with their limitations. In the work of Er et al. [Er and Kaur (2017)] used K-mean clustering for brain tumor identification, though they achieved considerable accuracy but did not work for global cluster. In this paper, we have presented an involuntary and fast approach that help us to avoid such problems. The system is trained on MRI and it is less harmful in contrast with other approaches alike positron emission tomography (PET) scan because MRI is non-invasive. Furthermore, MRI not needed destructive isotopes that are vaccinated in body and not any surgery is needed. Some recent work target image recognitions on medical data imaging in general such as Jia et al. [Jia, Zhang and Rabczuk (2015); Pawar, Zhang, Jia et al. (2016)]. The authors presented a multilevel approach for dealing image registration drawbacks. They used B-spline basis functions to construct a spatial transformation specific development which determines control points. Further, in Pawar et al. [Pawar, Zhang, Jia et al. (2016)] the authors used hierarchical B-splines based on Finite Element Method (FEM) to recognize registration of nonrigid medical images. Apart from above mentioned approaches, the proposed methodology leads to an attractive interface in which users upload MRI scan to see results side by side.

In this paper, we have focused on:

- The proposed technique is based on image shape, content and texture to analyse brain images and perform precise segmentation with moderately less number of computational necessities.

- This study correctly identify brain tumor based on watershed segmentation techniques it is further leads to feature extraction and then machine learning algorithms are applied on it. Finally, this smart technique helps in exact identification of tumor.

- Using watershed we successfully overcome the issue of distorted boundaries and wrong edges. Similarly, it provides proper segmentation of identified regions including foreground and background with minimum computation cost.
• We have performed experiment on 1532 images dataset along with three different types of brain tumor images. Which is large dataset as comparison with existing researches.
• The main aim of this paper is to provide a smart technique to detect brain tumor with highest accuracy and minimum false positive rate.
• Consequently, we have achieved better detection results of brain tumor by using watershed segmentation as it required low computation time and K-NN algorithm.

The remaining paper is separated into following portions: Section 2 provides the related work that explain existing approaches for brain tumor detection. Section 3 explains a proposed methodology to detect brain tumor by watershed segmentation. Section 4 deliberate results that relates to best, average and bad and accuracy. Finally, Section 5 described the conclusion and future work of this paper.

2 Related work

In literature researchers developed automatic segmentation methods to detect brain tumor. Brain abnormalities are detected at initial stages by MR more accurately. MR technique is useful in detection of white stuff disease i.e., multifocal leukoencephalopathy, post infections, manifold sclerosis and encephalitis leukodystrophy. MRI scan process used after the brain scanning in detection of brain tumor. Moreover, this detection technique provides help to obtain location and size of brain tumor. Through segmentation we achieved most significant information from scanned brain’s MRI images. In this situation, automatic segmentation is expected as great potential in clinical medicine.

Shivakumarswamy et al. [Shivakumarswamy, Akshay, Chethan et al. (2016)] used K-NN algorithm and multistep technique in which at first step obtain an MRI scanned image. In pre-processing phase they focused on noise removal and size changes. After that it is proceeded to segmentation by using two machine learning algorithms as K-mean clustering and Fuzzy C-mean. Furthermore, tumor cells are separated from normal cells and the area of tumour is measured. Finally, results are shown to concerned people and informed about determined tumour stage. Another technique by Kalaiselvi et al. [Kalaiselvi, Nagaraja and Sriramakrishnan (2016)], which is based on thresholding that used in literature to detect brain tumor. They applied MRI technique for identification of tumor by head scans. The MRI images are pre-processed by transformation techniques and enhanced the tumor region. At next step, they checked images for abnormality using FSM (fuzzy symmetric measures). In case of abnormal images then Otsu’s thresholding is used to extract tumor region.

In the work of Kumar et al. [Kumar and Mathai (2017)] they introduced a method with modified K-mean and morphological operations for segmentation of brain tumor which is based on two significant algorithms. Through this technique, they achieved accuracy in segmentation of tumor tissue and can reproduce like manual segmentation. Moreover, in this technique we used morphological operation as it to enhance images boundary and remove noise from images. Consequently, they shows accurate results for the calculated area of tumor with the morphological centred area calculation process. Another researcher [Mukaram, Murthy and Kurian (2017)] used pillar K-mean algorithms to detect brain tumor from magnetic resonance imaging with segmentation. They divided whole process into six stages named as (a) input image (b) pre-processing phase (c) segmentation (d) post processing phase (e) feature extraction and (f) classification. They carried out only first four stages in their
study. Furthermore, input images used to read MRI brain images and pre-processing phase used to smooth and enhance image. Similarly, at segmentation phase they applied pillar K-mean algorithm that helps to effectively segment the brain tumor from MRI. At final stage, post processing phase helped them to locate correctly tumor area in the brain.

Another researcher, Er et al. [Er and Kaur (2017)] proposed a method for brain tumor detection by using Fuzzy c-mean and mean shift that relates to clustering method and segmentation respectively. Moreover, they utilized these techniques for better classification of results that enhance image size. Accordingly, experimental results shows that segmentation method for brain tumor detection presented accurate and efficient results. Another author, Patel et al. [Patel and Rao (2017)] proposed brain tumor detection methodology with MRI which relates to thresholding techniques. They considered total 155 MRI images of brain tumor captured from different angles. Furthermore, they find out tumor location and dimension of MRI scan images. Finally, they achieved good accuracy and least time delay. Another author, Sharmila et al. [Sharmila and Joseph (2018)] presented by a method to detect brain tumor by supervised machine learning algorithms such as Naive Bayes and support vector machine. They collected dataset of 110 brain tumor images and achieved accuracy of 91.49%. However, we shows the summary of related work in Tab. 1, which elaborate proposed techniques with their results and limitations.

**Table 1: Summary of Existing Brain tumor detection approaches with results and limitations**

<table>
<thead>
<tr>
<th>Sr.no</th>
<th>Papers</th>
<th>Authors</th>
<th>Year</th>
<th>Techniques</th>
<th>Results/Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.</td>
<td>A Simple image processing approach to abnormal slices detection from MRI tumor</td>
<td>T.Kalaiselvi, P.Nagaraja and P.Sriramakrishnan</td>
<td>2016</td>
<td>Fuzzy Symmetric measures</td>
<td>It takes minimum missed alarms</td>
</tr>
<tr>
<td>3.</td>
<td>Brain Tumor Segmentation by Modified K-Mean with Morphological Operations</td>
<td>Rajeev Kumar, Dr. K. James Mathai</td>
<td>2017</td>
<td>Morphological Operators and K-mean</td>
<td>Not work for global cluster</td>
</tr>
<tr>
<td>4.</td>
<td>An Automatic Brain Tumor Detection, Segmentation and Classification Using MRI Image</td>
<td>Arbaz Mukaram Chidananda Murthy.M.V., M.Z.Kurian</td>
<td>2017</td>
<td>Classification</td>
<td>When only classification is applied, it ignores the poor quality images.</td>
</tr>
<tr>
<td>5.</td>
<td>Efficient image segmentation of brain tumor detection using fuzzy c-mean and mean-shift</td>
<td>Mandip kaur, Prabhpreet kaur</td>
<td>2017</td>
<td>Fuzzy c-mean and mean-shift</td>
<td>Neglected the use of fuzzy and region growing segmentation</td>
</tr>
</tbody>
</table>
6. Brain Tumor Detection in MRI Images with New Multiple Thresholding  
Sandeep Patel, Divyanshu Rao  
2017  
Brain Tumor Detection and Segmentation Using Histogram Thresholding  
Useful for linear image does not give accurate results

R Sharmila*1, K Suresh Joseph2  
2018  
SVM and Naïve Bayes algorithms  
It shows accuracy of 91% with SVM classification algorithm but they used only 110 brain images as dataset

3 Methodology
In this section we proposed a brain tumor detection using watershed segmentation method as it is shown in Fig. 1. The overall proposed methodology is divided into three large phases in which (a) pre-processing (b) morphological processing and (c) segmentation process through watershed technique. In watershed, topographic relief are generally grey color images and every relief is flooded from its minima. The merging of two reliefs produced a dam. Whole, presentation emulate the process of flooding. Moreover, the major advantage of watershed reduce the computation cost and sharply defines the edge detection.

![Proposed Methodology for brain tumor detection using watershed segmentation](image)

**Figure 1:** Proposed Methodology for brain tumor detection using watershed segmentation

(a) Pre-processing phase
At the very first phase, we need to remove noise and enhance the brightness and contrast ratio if it is essential. Likewise, pre-processing also minimize the chance of error by eliminating noise. In this study, we used Gaussian filter for removing noise. Furthermore, pre-processing phase involve augmented the contrast that is helpful to achieve better segmentation with upright gradient.

(b) Morphological operations
At the second phase, we applied morphological operations that target the specific shape and forms. Image Pixels are in groped form based on common trait and it is a desirable
watershed segmentation that used ridges for high intensities and continuous regions. Moreover, it helps to group image pixels and separate them from neighbouring area.

(c) Watershed Segmentation and Classification

The third phase which involves segmentation through watershed algorithm [Steve Eddins (2018)], edge detection and feature extraction and machine learning classification steps. In image processing edge is base step that help us to identifying sharp and sudden change in the intensity values of pixels. Correspondingly, it finds valuable in registration, recognition, segmentation and identification. Furthermore, features are selected for model construction. After segmentation, features are extracted from explicit area which specify to test existence of tumor. Features are used as training data and fed it into machine learning classifier. In this experimental setup, we applied K-NN (K Nearest Neighbour) classification which is significant for pattern recognition and regression. Finally, it presented results based on learning and classify two states either brain tumor is present or not.

4 Results and discussions

In this section we described experimental results that relates to best, average and bad results. Moreover, in Subsection 4.7 we explained the dataset and experimental setup of this work. Additionally, accuracy is calculated by K-NN machine learning algorithm. In Fig. 2, we shows the image data with watershed algorithm output and intermediate results. (a) It shows the original image as it is in grey scale. In pre-processing phase we converted the image it into greyscale and then segmentation and morphological operations are applied on it. (b) It displays the gradient image which is achieved by applying Sobel filter. It shows high gradient value at edges while lower at inside the image. (c) Computer markers are applied and the watershed segmentation to the gradient magnitude. (d) There are two morphological techniques are used “opening by reconstruction” and “closing by reconstruction” to compute markers and hold overall shape of the object. (e) and (f) indicate result of thresh holding, watershed segmentation and separated area of brain tumor.

(a) Original Image. It indicates original image which is already converted into grey scale from RGB

(b) Gradient Magnitude. Sobel operator is used for calculation of gradient magnitude and it works on 3*3 kernels that take place on x and y axis.

(c) Watershed transformation of gradient magnitude. It results in over segmentation.
A Learning Based Brain Tumor Detection System

(d) Applying two morphological operations. Opening and closing reconstruction operations are applied.

(e) Regional Maxima. This area is segmented and based on regional maxima. Regional Maxima.

(f) Watershed on region: maxima. Watershed algorithm is applied on the maxima.

Figure 2: Intermediate to final result with watershed algorithm from (a) to (f) brain tumor images result

In Tab. 2, we described evaluation parameters that help us to calculate best, average and bad results. Furthermore, a machine learning K-NN algorithms used and its accuracy is mentioned in confusion matrix Tab. 3.

Table 2: Evaluation parameters for brain tumor detection images

<table>
<thead>
<tr>
<th>Evaluation Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best Results</td>
</tr>
<tr>
<td>Average Results</td>
</tr>
<tr>
<td>Bad Results</td>
</tr>
<tr>
<td>True Positive (TP)</td>
</tr>
<tr>
<td>True Negative (TN)</td>
</tr>
<tr>
<td>False Positive (FP)</td>
</tr>
<tr>
<td>False Negative (FN)</td>
</tr>
<tr>
<td>Accuracy</td>
</tr>
</tbody>
</table>

Best result indicated that best localisation and thin similarity with real edge.

Average results also have good localisation of the brain tumor. But in this case edge are not fully matched but it is near about 90%.

In bad results it wrongly identify the segmented area that leads to healthy area.

The true positive rate shows correctly identified brain tumor cases.

The number of cases that are wrongly detected brain tumor and predicted as healthy.

Number of cases that are inaccurately identified as patient.

Number of cases that are imperfectly recognised as healthy.

It counts overall accuracy of the approach and shows detection rate.
4.1 Best results

In Fig. 3, we presented best results from (a) to (f) as for brain tumor detection. Best results relates to true edges detection, localization, exactly identify the boundaries and real segmented area as depicted in Fig. 3 with (a) to (f).

(a) Edges are truly detected to real tumor
(b) localization of the tumor is quite accurate
(c) Boundaries are of the tumor are accurately identified
(d) Margins of tumor is truly identified
(e) Truly detected tumor area and localisation
(f) Boundaries are effectively closed to tumor area

Figure 3: Best results for brain tumor detection (a) to (f)

(a) Best Results: The detected edges are accurate to the real tumour. The noise has been removed from image in pre-processing phase to avoid intrusion in detection. Moreover, edges are closely map to the actual tumor.
(b) Best Result: This is classified as a best result because localization of the tumor is absolutely accurate. Edges resemble to the actual edges of the tumor. In this fig. (b) It shows no extra details or spaces as tumors. Similarly, only actual tumor is localized and identified.

(c) Best Result: The boundaries are of the tumor are accurately identified. The area of the tumor is resemble to actual tumor. Furthermore, there is no extra space nearby area. The recognized area relates to actual tumor.

(d) Best Result: The margins of tumor is identified which is close to 90% of real boundary. It ignored some area of real tumor but still reflected as best result due to correct localization.

(e) Best Result: The result shows truly detected tumor area and localisation is also tremendous. Likewise, no extra region is determined as a part of tumor and boundaries are truly correspond to definite tumor.

(f) Best Result: It presented best identified brain tumor as its boundaries are effectively closed to tumor area.

4.2 Average results
Average results take place among best and bad results. In Fig. 4, we described average results for brain tumor detection as it shows least difference from best results. This least difference is occurred due to noise in images. In Fig. 4, average results are presented as (a), (b) and (c) with their description.

(a) Tumor is localised but edges are not truly identified
(b) Tumor is not exactly identified and 40% area is left over

(c) Tumor is recognized but nearby tumor area not truly identify

**Figure: 4** Average results for brain tumor detection (a), (b) and (c)

(a) Average Result: It shows average because tumor is localised and boundaries are not correctly identified. The boundaries are not fully recognized the actual tumor. Also, some area of tumor is lost in recognized region.

(b) Average Result: In this image tumor is not exactly identified. Nearly 40% area is left from desired area. Moreover, it is not fully enclosed and that’s why it is considered in average results.

(c) Average Result: In this image tumor is identified but it also consider the nearby area and not correctly identified.

**4.3 Bad results**

Although, we applied number of pre-processing procedures before watershed but still few images in our dataset belongs to bad quality and not detected truly. In Fig. 5, we presented some bad results in which normal part of brain is recognized as tumor and not properly detected in (a), (b) and (c).
(a) Tumor is not correctly identified

(b) Normal brain area is as considered tumor area

(c) Normal brain area is marked as tumorous area

Figure 5: Bad results for brain tumor detection (a) to (c)

(a) Bad Result: The given results are not appropriate because it consider the large area of tumor form actual size and shows wrong results.
(b) Bad result: It shows wrong identification of tumor and normal brain area is also considered in brain area.
(c) Bad Result: Tumor is not localized and healthy brain area is considered as tumor area. The exact tumor area is not detected and normal area is marked as tumorous area.

### 4.4 Accuracy of K-NN algorithm

In this sub section we elaborate truly and wrongly detected brain tumor area as true positive and false positive individually. Moreover, other evaluation results are also defined as true negative and false negative. In Tab. 3, we presented result of 1532 samples from which 1325 samples are truly detected as brain tumor. Furthermore, we defined accuracy with formula and results. It shows 86% accuracy with K-NN (K-nearest neighbour) machine learning classification algorithms.

#### Table 3: Confusion Matrix

<table>
<thead>
<tr>
<th></th>
<th>Predicted No</th>
<th>Predicted Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual No</td>
<td>TN = 724</td>
<td>FP = 108</td>
</tr>
<tr>
<td>Actual Yes</td>
<td>FN = 99</td>
<td>TP = 601</td>
</tr>
</tbody>
</table>

K-NN accuracy is defined mathematically as in given below Eq. (1).

\[
\% \text{ Accuracy} = \frac{TP + TN}{(TP + TN + FP + FN)} \tag{1}
\]

The description of TP, TN, FP and FN is defined in Tab. 3. With the help of Tab. 4 and Eq. (1), we calculate the accuracy as follows:

\[
\% \text{ Accuracy} = \frac{601 + 724}{1532} \times 100 = 86 \%
\]

Consequently, we achieved 86% accuracy in detection of brain tumor from K-NN classification algorithm.

#### 4.5 Time required to produce results

In our study, we have tried out dissimilar procedures for edge detection in brain tumor. The mentioned Fig. 6, proves that morphological operator gives best result in least time consuming as comparison with other operators such as sobal, ropert and prewitt etc. Furthermore, cellular automata operator taken maximum time with 1.8 sec to produce results and morphological operator shows least latency rate among various mentioned operators.
4.6 Comparison with existing approaches

In Tab. 4, we comparatively describe our proposed method with existing approaches according to their dataset and type of images. The accuracy of existing approaches [Patel and Rao (2017); Sharmila and Joseph (2018)] is largely effected by various factors such as, dataset size, image properties, and method adopted. The aforementioned approaches [Patel and Rao (2017)] use brain tumor datasets (155 samples only) with left/right angles only with histogram thresholding method. Whereas Sharmila et al. [Sharmila and Joseph (2018)] adopted machine learning classifiers (SVM and Naïve Bayes) to classify dataset of 110 sample images. Apart from the above-mentioned approaches, we proposed methodology that uses Watershed segmentation along with machine learning classifier (KNN) to identify brain tumor from diverse nature of tumors i.e, for Meningioma, Pituitary and Glioma. As a result, our system’s accuracy is acceptable (i.e., 86%) as it largely depends on various other factors that ancestors are not following.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Method</th>
<th>Dataset</th>
<th>Type of Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patel and Rao (2017)</td>
<td>Histogram Thresholding</td>
<td>155</td>
<td>Brain tumor images captured from various angles (left, right)</td>
</tr>
<tr>
<td>Sharmila and Joseph (2018)</td>
<td>Machine Learning Classifiers (SVM and Naïve Bayes)</td>
<td>110</td>
<td>N/A</td>
</tr>
<tr>
<td>Our Proposed Methodology</td>
<td>Watershed segmentation and KNN classification</td>
<td>1532</td>
<td>Meningioma, Pituitary and Glioma</td>
</tr>
</tbody>
</table>
4.7 Dataset and experimental setup

MRI images are collected for healthy and tumor brains as we used T1, T2 and flair all sequence. A total dataset contained 1532 samples in our experiment. We split dataset into training and testing in which 30 volumes randomly particular for training and 70 volumes for testing set. For validation purpose, we used images of real data which is gathered from figshare website (https://figshare.com) with different tumor types such as glioma, pituitary and meningioma as shown in Tab. 5. Furthermore, for experimental setup of brain tumor images dataset we used MATLAB R2018a [Matlab (2018)] and WEKA [WEKA] tool for further evaluation process.

<table>
<thead>
<tr>
<th>Source</th>
<th>dataset of images</th>
<th>No. of patients</th>
<th>Types of tumor</th>
<th>Total samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figshare</td>
<td>3064 T1-weighted contrast-enhanced images, T2, flair all sequence</td>
<td>233 patients</td>
<td>Meningioma, glioma, pituitary</td>
<td>1532</td>
</tr>
</tbody>
</table>

5 Conclusion and future work

In this paper we have presented technique for brain tumor detection which is based on watershed segmentation. This process proves that less computational cost as compared to other techniques as it is divided into three main phases (a) pre-processing (b) morphological processing and (c) segmentation process through watershed technique and applied machine learning classification. Additionally, we tested out our proposed methodology on big dataset of images and obtained precise results. As in results section, we achieved best, average, bad and accuracy of 86% by using K-NN classification algorithm. However, in future we will use larger dataset with variety that leads to bone and lungs tumor. Furthermore, we can use other methods like fuzzy C mean, wavelet to enhance the precision and litheness.

References


Kumar, R.; Mathai, K. J. (2017): Brain tumor segmentation by modified k-mean with morphological operations. *Brain*, vol. 6, no. 8.


