PAPER • OPEN ACCESS

A Survey Paper on Image Mining Techniques and Classification Brain Tumor

View the article online for updates and enhancements.

A Survey Paper on Image Mining Techniques and Classification Brain Tumor

Dhamea A. Jasm, Murtadha M Hamad, Azmi Tawfek Hussein Alrawi

Department of Computer Sciences, Collage of computer sciences and information Technology, University of Anbar, Ramadi-Anbar, Iraq

dhameaanwar@gmail.com
dr.mortadha61@gmail.com
azmi.alrawi@uoanbar.edu.iq

Abstract. Image mining is a method of searching and discovering the valuable information and knowledge from a set of huge image data. Image mining essentially depends on the data mining, digital image processing, machine learning, image retrieval, and artificial intelligence. Image mining is a process which is conducted to extract the hidden information such as image data and the additional pattern that could not be observed from the image. The main problems could face the mining of the collected images can be summarized in two main points: first is the image must be suitable for the mining process and second is the image's chosen objects and features in order to be treated to extract the most effective route to save the time, and to save the effort. This paper is a survey presents the steps of the image mining process and represented an intensive view on using the image mining to the classify the brain tumors. In addition, it's proposed a general scheme to accomplish the processes and to analysis the latest techniques which have been used to classify the brain tumors with comparison to the training groups and the amount of accuracy that obtained from the analysis. In addition, the paper compares the relevant and most recent published literature. The high published accuracy claim to be 98% which was obtained using the Deep convolutional neural network (DCNN).

Keywords. Data Mining, Image Mining, Brain Tumor, Image Classification.

1. Introduction

Data mining is a process is used to identify the interesting structure in form of a data. The data can be expressed in many forms such as structural designated patterns, the relationship among the collected data and the statistical or predictive models of the data. In addition, it can be used to analysis a large amount of the data in order to discover a novel and potential information [1]. The benefit of using the data mining is to accelerate the processing of the data or the rate of recovering the information from the databases, exploring or identifying the hidden or nonexplorer patterns, simplifying the data and saving the time [2]. Depending on the necessary steps, different algorithms and multiple techniques have been used to obtain the discovered knowledge. The objectives which have been discussed in this paper are divided into the following sections. Section 1 discussed the processing of the digital photos. Section 2 is a review about the most important techniques of photo exploration, while section 3 explored the

applications of the digital images. The final part of the current review is discussion to the most important research which are related to the image retrieval. The dramatic growing up on the global social network [3], the huge data are available; therefore, the data mining plays a vital role in numerous fields such as medical, satellite, and business fields. The difference between image mining and the low-level computer visions and the image processing techniques is that the image mining deals with the extracting of the patterns from the large collected images. Image mining technique treats images of an unstated data or information and reveals the relationship of the different patterns which are indirectly explored from the image. Many approaches have been used to classify the image mining. One classification is depended on extracting the information from the database or the collected images. Another classification is based on mining a mixture of the collected images and the associated alphanumeric data [2]. The typical image mining can be described in five essential phases (see Figure.1) [4].
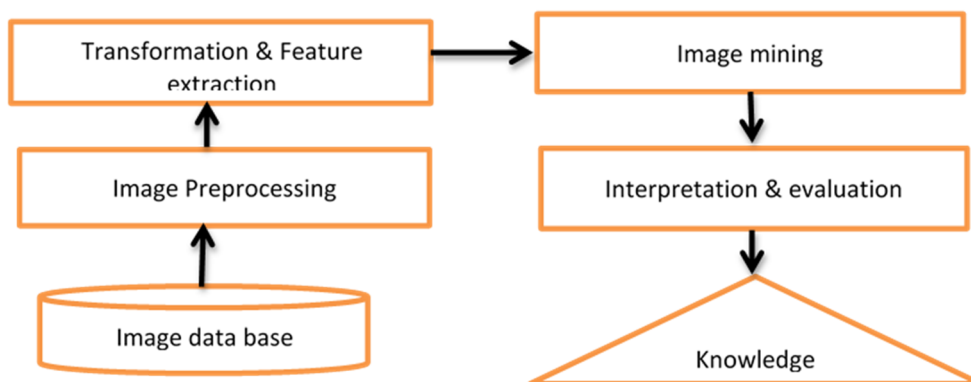

Figure 1. Scheme diagram of image mining process

Image mining consists of several steps such as refining the image database, pre-processing, transforming and extracting mining and interpreting the data and finally the Evaluation.

The steps of the process can be summarized in the following successive steps [5]:

1: reading the image from image database.

2: image preprocessing: is important for both of users and the required application. Preprocessing involves improving the quality of the image before extracting the features. Preprocessing image is necessary for removing noise from the image that might cause unwanted results [3].

3: image transformation and feature extraction: image feature retrieving, and extracting must be based on color, texture, shape, analyzed and comparing the images.

4: image mining: mining useful knowledge from a large set of image database is the main object of the image mining. Mining performed using suitable data mining techniques in order to discover the significant patterns.

5: Interpretation and evaluation: after obtaining the resulting pattern, the pattern finally evaluated and interpreted to obtain the final knowledge, which can be applied individually or organized for a future application [4] .

2. Related Work

S. Chauhan et al. (2017), [6] investigated the detection and classification of the brain tumor, In the proposed work, the median filtering is used in pre-processed for the MRI brain imaging. In order to isolate the lesion from the image, the colored based segmentation and edge detection has been conducted. In addition, the multiple feature extraction schemes, namely histogram of oriented gradients and gray level co-occurrence matrix have been used to display the images. A transactional database has been used to store all the extracted features. IBkLG classifier (Instance based K-Nearest using Log and Gaussian weight Kernels) has been applied using WEKA 3.9 tool, to classify the tumor into normal benign or malignant. The accuracy of 86.6%. for the classification has been recorded.

S. Kumar et al. (2017), [7] reported using the hybrid approach for brain tumor classification. The approach includes using the discrete wavelet transform (DWT) to extract the features, the genetic algorithm to decrease the number of features and to support the vector machine (SVM). The yield of applying the approach declared the accuracy improvement and the minimizing on the error for the RMS compared to the state-of-the-art techniques for the same range of the context Linear accuracy 80% to 90%.

Z. Ullah et al. (2018), [8] reported an intelligent method to classify the brain Magnetic Resonance Imaging (MRI). The classification based on four stages; pre-processing, feature extraction, feature reduction, and classification. During the first stage, the noise is removed by applying a median filter on brain MRI then converting the image into RGB. In the second stage, Discrete Wavelet Transform (DWT) is conducted to extract the features of the image, in the feature reduction stage using color moments. The reduced features at the final stage are classified by the k-Nearest Neighbors (k-NN) into normal and abnormal MRI. The results confirm the better performance of the system compared to the other methods in term of reducing the number of features. The overall accuracy of this method is 94.97%.

N. A. Mazin *et al*. (2018), [9] stated using ANN from MR images can improve the automated brain tumor segmentation and identification approach without human mediation. In order to improve the MR image to detect their gray scale in the district regions, K-means clustering is utilized as a part of the principle organization in the process. Then, ANN is applied to select the correct object in the view of training phase. After which it can be extract the texture feature of brain tumor area into the division stage. The accuracy of ANN approach is 94.07%.

An automated technique for cancerous brain tumor classification has been proposed by M. Kaur and B. Prajapat. (2018), [10]. The proposition based on using Haarlet transform and probabilistic neural network. The threshold-based segmentation and binarization in addition to Haarlet transform have been used for data processing. For each of training images, the feature extraction yields twelve features which is used to train a probabilistic neural network. The accuracy of the proposed system was 96.3%

N. Abiwinanda et al. (2019), [11] attempted to train a Convolutional Neural Network (CNN) in order to to differentiate the most three common types of brain tumors (Glioma, Meningioma, and Pituitary). The simplest possible architecture of CNN has been implemented in order to identify the optimal CNN architecture (architecture 2). Architecture 2 consists of 2 layers of convolution, activation (ReLu), and maxpool, followed by a hidden layer of 64 neurons. Among all architectures, architecture 2 is demonstrated a steady state decreasing pattern in the validation loss due to the increasement in the number of epoch. The achieved training accuracy was 98.51% with a best validation accuracy of 84.19%.

F. Özyurt et al. (2019), [12] used Neutrosophy and Convolutional Neural Network (NS-CNN) to propose a hybrid method. In the first stage, the neutrosophic set–expert maximum fuzzy-sure entropy (NS-EMFSE) approach is utilized to segment the MRI images. During the classification stage, the features of the segmented brain images are obtained using the CNN and then it classified using SVM and KNN classifiers. The output of the study demonstrated that the CNN features presented a high classification performance with different classifiers. The experimental results were in consistent with SVM simulation output data of average success of 95.62% and the CNN features displayed a better classification performance.

H. H. Sultan et al. (2019), [13] stated use of a DL model based on a deep neural network to classify the different brain tumor types. The proposed approach used two publicly available datasets. The former database classifies tumors into (meningioma, glioma, and pituitary tumor) while the second one differentiates between the three glioma grades (Grade II, Grade III, and Grade IV). The datasets consisted of total images of 3064 and 516 on T1-weighted contrast-enhanced images from 233 and 73 patients for the first and second datasets, respectively. The overall accuracy of 96.13% has been achieved using this proposed network structure.

A. K. Anaraki et al. (2019), [14] proposed a method to noninvasively classify different grades of Glioma using magnetic resonance imaging (MRI). The proposed method based on convolutional neural

networks (CNNs) and genetic algorithm (GA). Pituitary tumor types were classified with an accuracy of 94%. The effectiveness of the proposed method has been proven to classify the brain tumor via MRI images. The flexibility of the method allows practical use of the method to assist early stage diagnosing of the brain tumors.

S. Deepak and P. M. Ameer et al. (2019) [15] investigated the 3-class classification problem to differentiate among the glioma, the meningioma and the pituitary tumors, that are form the three brain tumor prominent types. The proposed solution of the classification adopted deep transfer learning and the extracted features from the MRI imaged using a pre-trained GoogLeNet. The study concluded that the transfer learning is a useful technique when there is limiting in the accessibility to the medical images. The analytical analysis of the misclassifications has been discussed. The proposed system records a mean classification accuracy of 98%.

Table 1. Methods used in the previous related works with the proposed techniques

No	Authors	Years	Method Classification	Training	Accuracy
1	S. Chauhan et al.	2017	k-Nearest Neighbors (k-NN)	60%	86.6%
2	S. Kumar et al.	2017	hybrid approach:(DWT) & (SVM)	80%	85%
3	Z. Ullah et al.	2018	k-Nearest Neighbors (k-NN)	70%	94.974%.
4	N. A. Mazin et al.	2018	Artificial Neural Network (ANN)	70%	94.07%.
5	M.Kaur &B. Prajapat	2018	probabilistic neural network(PNN)	68%	96.3%
6	N. Abiwinanda et al.	2019	Convolutional neural network (CNN)	80%	84.19%
7	F. Özyurt et al.	2019	Neutrosophy and Convolutional Neural Network (NS-CNN)	80%	95.62%
8	H. H. Sultan et al.	2019	Deep neural network (DNN)	80%	96.13%
9	A. K. Anaraki et al.	2019	hybrid approach: (CNN) & (GA)	70%	94%
10	S. Deepak and P. M. Ameer	2019	Deep convolutional neural network(DCNN)	70%	98%

3. Image Mining Technique

Several techniques such as object recognition, image retrieval, image indexing, image classification, image clustering, association rule mining and neural network are used in order to mining the image information. Fiure.2, [4] illustrates the image mining techniques.
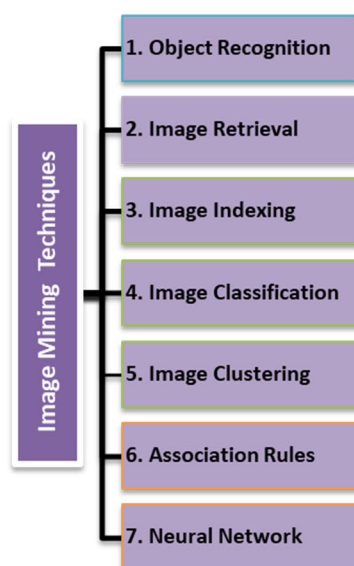

Figure 2. Image Mining Techniques

3.1. Object Recognition

This technique is an active technique in the research field of image processing. Using object models and recognition technique are utilized to find the objects from the images of the database. Useful information extrication and automatic machine learning can simply be liberated once the machine identifies and recognizes few objects. The object recognition is a supervised labeling problem based on the object models of the investigated items [4][5].

3.2. Image Retrieval

With more and more accessing into the information on the internet and on the multimedia of as become world focus on exacting and accelerating retrieval, the image retrieval is classified into three levels according to the increase in complexity [16].

Level 1: Considering low-level feature includes color, texture and shape to retrieve images.

Level 2: Considering middle -level logical features such as objects of a given type or person.

Level 3: Considering high-level feature which includes features about the meaning or scenes depicted.

3.3. Image Indexing

The target of image indexing is to retrieve matching images from the image database for a certain query image. The unique feature for each image [4]. Image index can be carried out via comparing the features, the image can be similarly measured depending on the features such as the color, the intensity, the texture, the position and the shape.

3.4. Image Classification

Classification helps us to take the decisions. The information processing which is carried out during the classification helps to categorize all pixels in a digital images into various groups [3], It is a supervised learning method used to classify the images on some pre-classified results.

3.5. Image Clustering

It is a process of collecting the similar data object. Dissimilar object is another cluster. It is the way of investigating and pointing out the similarities between the data depending on their characteristic. Image clustering or (unsupervised) is a method to differentiate the images into groups according to their contents without a priori knowledge [3].

3.6. Association Rules Mining

It is an efficient tool to recognize the pattern in knowledge discovery and data mining [4]. There are two main methods represented the foundation of the normal association rule algorithm work. The first one is to discover all the significant item sets that complement the minimum support constraint. The second is the moving of the generated rules from each item which complements the minimum confidence constraint [5].

3.7. Neural Network

Neural network is a technique conducted in the retrieval and mining of the images. Basically, it is utilized in a computational system built up from simple processing units which called neurons. It is established in layers which is partially or fully connected. The major function of the neurons is receiving the input values from its neighbor neurons (their output), evaluating the output depending on its input weight and finally, transfer the output into its neighbors [5].

4. Applications

With the increase of the accessible data, the need for image mining application in many areas increased such as in the field of satellite cloud images, the field of neural network, medical image classification and others will be mentioned below.

Satellites: Remote sensors collect an abundance amount of data from the satellites. Therefore, different mining approaches are used in various satellites image application [17].

Medicare: Image classification and association rule mining techniques are the most desired Medical image mining techniques to predict diseases.

Bank: Image mining techniques are conducted in retrieving and securing the data in addition to be used in scanning the eye to prevent the frauds.

Television: These systems collect and analyses, via the fly anonymous information from the channel view, broadcasts and programing.

Traffic Image: the function of the image analysis subsystem is distinguished the motions from the image sequences and supplies the spatial and temporal characteristics such as object type, position, orientation, speed and etc. [17].

Biological data: Tissue image mining is a resourceful and powerful technique when the images of the tissue are indexed, stored and mined content [18].

5. Using Image Mining for Brain Tumors

Magnetic resonance imaging (MRI) can be expressed as a scanning using the magnetic fields and radio waves, instead of the X-rays. The Brain tumor can be defined as a generation of abnormal cells inside the brain that leads to disrupt the function of the brain. The most common types of brain tumor are Oligodendroglioma, Glioblastoma, Gli-omas and Eningiomas [19]. The brain of human is consisted of four lobes where each of them has its own function and together they build up the brain. The presence of tumor in any of these lobes would affect its function [20].

The tumors presence when the normal cells acquired errors (mutations) in their DNA. These mutations accelerate the rate of growing up and division of the cells in addition to continue living when healthy cells die. Consequently, that leads to a mass of abnormal cells, which forms a tumor. The brain tumor can be a consequence of a cancer that starts elsewhere inside the body (secondary reason) and then spreads (metastasizes) into the brain. Many symptoms indicating the brain tumors such as headaches, numbness or tingling in the arms or legs seizures, memory problems; mood and personality changes, nausea and vomiting; changes in speech, vision, or hearing [21].
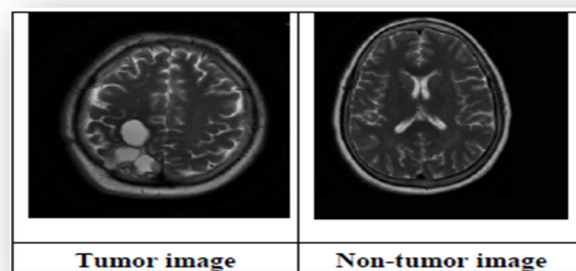

Figure 3. Classification of tumor and non-tumor MRI images

Pre-processing helps improving the quality of the image in order to be suitable for further treatments. Herein, the RGB input image is converted into a grayscale. The process is also associated with many steps such as enhancing the image, contrast improvement and image sharpening. Segmentation is the process which came after the pre-processing. Segmentation algorithms are areas oriented instead of pixel oriented. Segmentation based on the dividing an image into meaningful regions which means splitting up the image into connected areas [22]. The main object of the image segmentation is to divide up the digital image into multiple segments. Feature extraction is the process by which the cluster is extracted in order to diagnose the tumor at the output of segmentation. Finally, the image mining is used to detect and to classify the tumor, diagnosed whether the mass has started, malignant or normal. Figure.4 illustrates the block diagram of the classification steps of the tumors
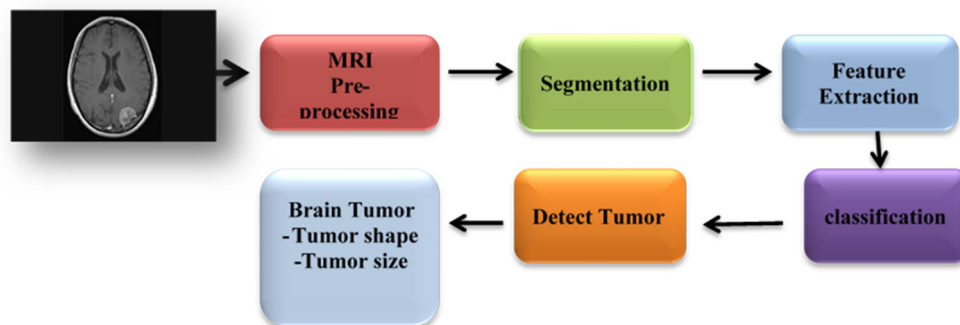
Figure 4. Block diagram

6. Analysis of Survey Studies

Studying the ways at which classification brain tumor, the measurement of training and testing groups and gathering exactness indicated the higher the training group, the more experience the image classification system will have. However, when the accuracy of the tested groups is high, the total training is small. This is an indication of the high efficiency of the proposed system. Therefore, Chauhan et al. in (2017) used k-Nearest Neighbors (k-NN) of group training (60%) and accuracy of (86.6%). However, it can be seen that the accuracy is not high. At the same time, Kumar et al. applied a hybrid approach includes discrete wavelet transform (DWT) and Genetic algorithm in order to decrease the number of features and to support the vector machine (SVM) in aim to obtain accuracy of (85%). In (2018), Z. Ullah and colleagues suggested conducted the advantages of the prementioned two approaches by integrating the k-Nearest Neighbors (k-NN) and the Discrete Wavelet Transform (DWT) techniques to reach the accuracy of (94.974%) with a group training of (70%). Later Mazin et al. utilized the Artificial Neural Network (ANN) to reach accuracy of (94.07%). The year itself, Kaur et al. used the probabilistic neural network (PNN) to obtain group training of 60% and accuracy of (96.3%). Abiwinanda et al. (2019) suggested using Convolutional neural network (CNN) to obtain accuracy of (84.19%). In (2019) Sultan et al. applied the same techniques Deep neural network (DNN) on two available datasets in aim to reach accurate up to (96.13%). In (2019) Anaraki et al. conducted a hybrid approach Convolutional neural network (CNN) and genetic algorithm (GA) to gain accuracy of 94%.Finally, the latest techniques which is suggested by Deepak and Ameer the (Deep convolutional neural network (DCNN)) proposed reaching the highest accuracy (98%) among the rest of the studies.

7. Conclusion

To sum up, the current paper pointed the necessity of using the image mining for many reasons such as the growing up in the image datasets. In addition, we have discussed most of the current image mining techniques such as object recognition, image retrieval, image indexing, image classification and clustering, association rule mining and neural network. The dynamic and future view of the image mining applications have been deeply discussed. Additionally, the current paper is an overview of the published research in the field of image exploration, and a deep reading to the researchers that used the techniques of image mining in detection the brain tumor and in classification. Finally, the analysis of the current published studies that used image mining techniques would help to propose the future work in different fields such as extending the use of image mining to deal with the video database. Conducting the image mining on global webs as a source for image database and proposing suitable indexes and retrieving knowledge.

References

[1]　R. S. J. Baker, "Data Mining for Education Data Mining for Education Advantages Relative to Traditional Educational Research Paradigms."

[2]   P. Chouhan and M. Tiwari, "Image Retrieval Using Data Mining and Image Processing Techniques," Int. J. Innov. Res. Electr. Electron. Instrum. Control Eng., vol. 3, no. 12, pp. 53–58, 2015.

[3]   S. Pandey, "A Survey Paper on Image Classification and Methods of Image Mining," Int. J. Comput. Appl., vol. 169, no. 6, pp. 10–12, 2017.

[4]   A. Tripathi and H. Jangir, "A Study on Image Mining Methods and Techniques," Int. J. Innov. Res. Comput. Commun. Eng., vol. 4, no. 4, pp. 7047–7053, 2016.

[5]   V. S. Shukla, M. E. Student, and J. Vala, "A Survey on Image Mining , its Techniques and Application," Int. J. Comput. Appl., vol. 133, no. 9, pp. 12–15, 2016.

[6]   S. Chauhan, A. More, R. Uikey, P. Malviya, and A. Moghe, "Brain Tumor Detection and Classification in MRI Images using Image and Data Mining," 2017 Int. Conf. Recent Innov. Signal Process. Embed. Syst., pp. 223–231, 2017.

[7]   S. Kumar, C. Dabas, and S. Godara, "ScienceDirect Classification of Brain MRI Tumor Images : A Hybrid Approach," Procedia Comput. Sci., vol. 122, pp. 510–517, 2017.

[8]   Z. Ullah, S. Lee, M. N. Khan, M. Fayaz, and M. M. Iqbal, "Features Reductions Using Color Moments and Classification of Brain MRI Using K-NN," Tech. Journal, Univ. Eng. Technol. Taxila, Pakistan, vol. 23, no. 4, pp. 77–83, 2018.

[9]   N. A. Mazin et al., "K -Means clustering and neural network for object detecting and identifying abnormality of brain tumor," Soft Comput., 2018.

[10] M. Kaur and B. Prajapat, "Automated Classification of Cancerous Brain Tumours Using Haarlet Transform and Probabilistic Neural Network," nternational Conf. Adv. Comput. Netw. Informatics Springer, pp. 19–25, 2018.

[11] N. Abiwinanda, M. Hanif, S. T. Hesaputra, A. Handayani, and T. R. Mengko, Brain Tumor Classification Using Convolutional Neural Network. Springer Singapore, 2019.

[12] F. Özyurt, E. Sert, E. Avci, and E. Dogantekin, "Brain Tumor Detection Based on Convolutional Neural Network with Neutrosophic Expert Maximum Fuzzy Sure Entropy," Measurement, 2019.

[13] H. H. Sultan, N. M. Salem, and W. Al-atabany, "Multi-Classification of Brain Tumor Images Using Deep Neural Network," Spec. Sect. Deep Learn. Comput. Med. DIAGNOSIS, pp. 69215–69225, 2019.

[14] A. K. Anaraki, M. Ayati, and F. Kazemi, "ScienceDirect Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms," Integr. Med. Res., pp. 1–12, 2018.

[15] S. Deepak and P. M. Ameer, "Brain tumor classi fi cation using deep CNN features via transfer learning," Comput. Biol. Med., vol. 111, no. March, p. 103345, 2019.

[16] P. Nithya, B. U. Maheswari, and A. Nandini, "A STUDY ON IMAGE MINING TECHNIQUES , FRAMEWORK AND APPLICATIONS," Int. J. Eng. Sci. Res. Technol., vol. 6, no. 7, pp. 611–615, 2017.

[17] S. Yousef and S. Karunan, "Satellite Image Mining in Real-time Data Analytical Architecture," IOSR J. Comput. Eng., no. 2278–8727, pp. 49–54, 2017.

[18] R. Sudhir, "A Survey on Image Mining Techniques : Theory and Applications," Comput. Eng. Intell. Syst., vol. 2, no. 6, pp. 44–53, 2011.

[19] H. Malhotra and S. Naaz, "ANALYSIS OF MRI IMAGES USING DATA MINING FOR DETECTION OF BRAIN TUMOR," Int. J. Adv. Res. Comput. Sci., vol. 9, no. April, 2018.

[20] P. Thejaswini, M. Bhavya, and K. Prakash, "Detection and Classification of Tumour in Brain MRI," I.J. Eng. Manuf. 2019, 1, no. January, pp. 11–20, 2019.

[21] A. Mustaqeem and A. Javed, "An Efficient Brain Tumor Detection Algorithm Using Watershed & Thresholding Based Segmentation," I.J. Image, Graph. Signal Process. 2012, no. September, pp. 34–39, 2012.

[22] P. K. Saini and M. Singh, "Brain tumor detection in medical imaging using matlab," Int. Res. J. Eng. Technol., vol. 2, no. 2, pp. 191–196, 2015.