Brain Tumor Classification Using Texture Feature Extraction

Heba Kh. Abbas¹, Nada A. Fatah¹, Haidar J. Mohamad², Ali A. Alzuky²

¹Department of Physics, College of Science for Women, University of Baghdad
²Department of Physics, College of Science, Mustansiriyah University

Email: prof.alialzuky@ouomustansiriyah.edu.iq

Abstract. Medical imaging methods are necessary for tumor treatment. Image processing increases the accuracy, location, efficiency and recognizes tumor in magnetic resonance imaging (MRI) for the human brain. There are three MRI image cases studied consist of one image of a healthy human brain, and two with brain cancer cases to extract tumor size, area, dimension, and location. We presented two steps to extract the texture properties from the three image cases. The first step consists of two methods, used to extract statistical image feature, which is a grey level co-occurrence matrix modified method and the proposed grey level Run-length matrix method. The second step consists of two methods applied to the outputs of step-one to extract the full information. These methods are segmentation based on threshold technique and supervised classification technique based on the parallel pipe. Moreover, there are twelve features studied like the image roughness, uniformity or energy, and local homogeneity extracted to show the quality difference between methods. Image histogram presented to show the small difference between the output of the features. The result shows the segmentation based on threshold technique has efficient output with higher accuracy as well as the presented methods.

Keywords. brain tumor, tumor size detection, MRI image recognition, run-length matrix, image processing algorithms.

1. Introduction

Pathological changes in living tissues are demonstrated by magnetic resonance imaging (MRI). This form of imaging relies on magnetic waves that enable us to use traditional software to detect and identify tumors (malignant and benign) [1]. As an important topic in image processing, texture analysis is used in classification and segmentation. Method of texture analysis based on intensity or variance in color within an image [2, 3]. Structural and statistical approaches are two categories of texture analysis. Whether it is constant or not, the structural method studied the texture field. The local property of the texture offers a shift in the action of the pixel if it is drastic or slow [4]. The statistical method deals with the pixel difference and measures this property of the imaging study.

For all texture images, the texture analysis properties should have the same point of view, by means of the same orientation and scale. In this case, this approach was considered to be a restriction, where it is difficult for all images with the same rotation or scaling. For all captured images [5], it is suggested to have the ideal perspective. There is a resemblance between the tumor's texture characteristics and normal tissue, and it is difficult to correctly diagnose tumor regions in this situation. This results in a false report and needs further measures to prevent this issue. Segmentation and classification to detect and recognize tumors are the suggested methods [6].

There are several previous studies in the field of tumor detection and identification in medical images. Mahmoud Khaled Abd-Allah et al. [7] used MRI images via the two-phase multi-model deep learning-
based method to detect brain tumors. The primary principle is that the precise localization of the tumor focuses on the classification of a normal and abnormal image. Using Convolutional Neural Networks (CNN) and the ECOC model, this is done by using binary support vector machine (ECOC SVM) techniques for extracting features and classifying brain tumors. Alam Shahariar et al. [8] used mean and fuzzy mean methods based on a prototype to identify brain tumors in MRI images. The value of the gray level is the initial parameter in the K-means algorithm centred on the prototype. Cluster data membership is the basis for successful outcomes in the fuzzy C-mean approach. These methods were used to detect the location of the tumor based on the membership function that based on the various tumor image characteristics, including contrast, capacity, dissimilarity, homogeneity, entropy, and correlation. Simulation findings show that under limited detachment of grey-level strength, the proposed algorithm achieves improved identification of abnormal and normal tissues in the human brain. An automated brain tumor detection and classification algorithm were suggested by Ahasan Kabir [9]. The algorithm proposed was split into five stages. First, the input MRI image was preprocessed to eliminate unwanted artefacts using the main component-based grayscale conversion coupled with the anisotropic diffusion filter. After that, to increase the image contrast, the contrast Restricted Adaptive Histogram Equalization (CLAHE) was introduced. The tumor is then segmented using the Chan-Vese algorithm and multivariate thresholding. Now, to distinguish the segmented objects, the statistical features, texture characteristics, and wavelet characteristics are measured. Finally, the necessary characteristics are chosen using a genetic algorithm and an artificial neural network is used to identify the segmented entity. Checking the performance of the proposed algorithm on publicly accessible BRATS datasets. The experimental outcome showed that the proposed algorithm achieved 99.5% accuracy for the BRATS dataset and 98.3% accuracy for the augmented BRATS dataset image.

The aim of this study is to diagnose the tumor using the modified algorithm for classification and compare it with the segmentation method for three MRI images. The classification method calculates the mean and standard deviation values of the selected block. This gives more details about the pixel classification related to the tumor area. Multi properties (features) are presented for the gray level run length matrix and gray level co-occurrence matrix to calculate the dimension of the tumor precisely.

2. Theories of Texture Feature Analysis

Using a repeated pattern, tissue arrangement, and frequent intervals [10], the texture function gives important details. In this case, as described by Guiying Li (2012), texture scale, density, form, and tissue proportion scale give texture characterization [11]. One of the basic tasks of low-level computer vision is texture analysis. In certain cases, because of the strength level, it is difficult to identify the brain picture. Further studies, such as segmentation and classification [12], are also required. The purpose of carrying out further studies is to concentrate on collecting the information in a specific area. This knowledge was derived using the proposed Co-occurrence Matrix for Gray Levels (SGLCM). Methods for the Gray Stage Run Duration Matrix (GLRLM) [13]. The SGLCM is a statistical texture test that collects second-order statistical data on pixel pairs. SGLCM is a frequency tabulation or the way the values of pixel brightness occur in an image. It just depends on the diameter of various distances and angles, and the statistical features are measured. In the textural analysis of an image, such characteristics as coarseness are a measure of roughness:

$$\text{Coarseness} = \frac{1}{2^{m+n}} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} P(i,j)$$  \hspace{1cm} (1)$$

The image homogeneity, known as uniformity or energy, is assessed by Angular Second Moment (ASM). It is the number of the squares in the SGLCM entries. When images have very strong homogeneity, or when pixels are very close, ASM is high.

$$\text{ASM} = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [P(i,j)]^2$$  \hspace{1cm} (2)$$
Local homogeneity is the Inverse Difference Moment (IDM). When the local gray level is uniform and the inverse SGLCM is high, it is high, and the inverse of the contrast weight is the weight value.

\[
IDM = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \frac{P(i,j)}{1 + (i-j)^2}
\]  

where \( i \) and \( j \) are spatial coordinate of the function, \( P(i,j) \) normalized symmetrical SGLCM, \( m \) and \( n \) represented number of gray levels in the image (gray tone).

Contrast is a statistical tool for calculating pixel intensity and comparing the studied image with its neighbor. Contrast regulates the difference in brightness and color of an object and context in the description of image processing as:

\[
\text{contrast} = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (i - j)^2 P(i,j)
\]  

Energy is the quantifiable quantity of the degree of pixel pair repetitions, the image's similarity measurement and the textural uniformity measurement. Define the following as:

\[
\text{energy} = \sqrt{\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} P^2(i,j)}
\]

One of the most commonly used dispersion indicators is Standard Deviation (SD).

\[
SD = \sqrt{\left(\frac{1}{m \times n}\right) \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} P(i,j) - \mu^2}
\]  

High Gray Level Run Focus (HGLRE) is an equivalent gray level to the focus on long runs. High gray levels are emphasized in the functionality.

\[
\text{HGLRE} = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} P(i,j)}{N_r}
\]

In the upper right quadrant of the GLRLM, where long run lengths and low grey levels are located, the Long Run Low Gray Level Run Focus (LRLGLE) focus runs.

\[
\text{LRLGLE} = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} P(i,j) \cdot j^2}{N_r}
\]

Long Runs Focus (LRE) tests long-run distribution. The LRE is highly dependent on long-term occurrence and is anticipated to be broad for coarse structural textures.

\[
\text{LRE} = \frac{\sum_{v=1}^{N_g} \sum_{w=1}^{N_r} B(v,w) \cdot w^2}{N_r}
\]

The focus of Long Run High Gray Level Concentration (LRHGLE) runs in the GLRLM's lower right quadrant, where long run lengths and high gray levels are located.

\[
\text{LRHGLE} = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} P(i,j) \cdot j^2}{N_r}
\]

Quick Run Concentrate (SRE) tracks short-run distribution. Higher value reveals fine textures

\[
\text{SRE} = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} p(i,j)}{N_r}
\]
The Run Percentage (RP) measures the fraction of the number of runs realized and the maximum number of runs probable (Galloway, 1975). A low run percentage is created by strongly linear uniform ROI volumes. Tests the homogeneity of a picture in a particular direction and the distribution of runs.

\[
\text{RP} = \frac{N_r}{\sum_{i=1}^{m} \sum_{j=1}^{n} P(i, j) \cdot j}
\]

where \(P(i, j)\) represents the number of runs with pixels of gray level intensity equal to \(i\) and length of run equal to \(j\) along a specific orientation. \(m\) the number of gray levels in the image, \(n\) the number of different run lengths that occur (so that the matrix is \(m \times n\)), and \(N_r\) total number of runs. The SGLCM Algorithm is given below.

\[
\phi(d, \theta) = p(i, j / d, \theta), \quad 0 < i \leq \text{Ng}, \quad 0 < j \leq R_{\text{max}}
\]

where \(\text{Ng}\) is the maximum gray level and \(R_{\text{max}}\) is the maximum run length. The matrix \(p\) represents the final details of the number run length \(j\) for the gray-level \(i\) at certain direction of angle \(\theta\). To form the run length of the gray-level matrix, there are four angle directions that are \(0^\circ, 45^\circ, 90^\circ, \text{and} 135^\circ\) [14].

The extracted function may use the methods of Long Run High Gray Level Focus (LRHGLE), Short Run Focus (SRE) and Run Percentage (RP) that are extracted for each duration matrix of the gray level run [14]. The GLRLM Algorithm is explained below.

| SGLCM Algorithm |
|------------------|
| **Input:** load MRI image (img) (normal, benign and malignant) |
| **Output:** feature image, where features are (Contrast, Energy, SD, ASM, Coarseness, IDM) |
| **Step 1:** load img |
| **Step 2:** convert img to gray image (imgy) using imgy=rgb2gray(img) |
| **Step 3:** display image |
| **Step 4:** pass filter size(7×7) on gray imag(imgy) for cod: win ==img(i:i+6, j:j+6) |
| **Step 5:** apply glcm, SI = graycomatrix (win, 'NumLevels', \(2,3,8\), \[ J, 'Offset', \{row d column \(\theta\)\}] ; Where glcms: graycomatrix (win), creates a gray level co-occurrence matrix; SI: gray level co-occurrence; NumLevels: integer specifying the number of gray levels to use when scaling the gray scale values in win and Offset: integer specifying the distance (row d column \(\theta\)) between the pixel of interest and neighbor. |
| **Step 6:** work cycle for Numlevels (k) by code d(1, k)=glcm (k,k); |
| **Step 7:** apply (d) for features to obtain suggest feature image (imgF\(_3\)) |
| **End** |
3. Morphological process

The suggested method of brain tumor segmentation and classification is outlined in Figure 1.

Three distinct human brain image MRI cases were studied: (a) a healthy brain image with a size of 131x139 pixels and a bit depth of 32 [15], (b) a malignant brain image obtained from the oncology hospital for tumors with a size of 480 x 270 pixels, a bit depth of 24, and (c) a benign brain tumor image with a size of 138 x 186 pixels, a bit depth of 32 [15] shown in Figure 2.

4. Results

To display the difference and efficacy of the extraction of any tumor, the statistical performance from the algorithms can be summarized in Table 1. The characteristics are for the tumor area only and it explain the use of action in classification and segmentation.
Table 1. Statistical output for malignant and benign images using the introduced algorithms.

|         | Contrast | Energy | SD  | Coarseness | ASM | IDM |
|---------|----------|--------|-----|------------|-----|-----|
| SGLCM   | 197.02   | 179.23 | 90.71 | 86.58      | 104.46 | 92.97 |
| GLRLM   | HGLRE    | LRE    | LRHGLE | LRGLE | SRE | RP |
|         | 71.82    | 216.85 | 161.83 | 30.23 | 25.30 | 30.80 |
|         | GLRLM    | HGLRE | LRE | LRHGLE | LRGLE | SRE | RP |
|         | 127.53   | 141.10 | 7.63 | 91.98 | 78.11 | 99.04 |
|         | 107.01   | 191.57 | 135.20 | 11.43 | 36.31 | 48.64 |

The performance outcomes are pre-stepped before the classification or segmentation starts. The morphological method called cleaning through reconstruction and filling through reconstruction is this step. The outcomes of this phase seen in the data.

The selected block in the original image used to separate the tumor in the original image and feature image automatically. The algorithm results show that the histogram has directly distinguished of the tumor comparing with the image features. The image features are different for the two methods. Therefore, the distinction between all features is clear depending on the histogram data.

Figure 3 shows the results of the SGLCM adopted method and GLRLM method for the malignant image. The same opinion of the histogram data shows a good distinction between the features showing the tumor more accurately. Figure 4 presents the benign image results with SGLCM and GLRLM methods, respectively. The same procedure followed for the malignant to compare between the two images.

The summarized data of the tumor for the two images listed in figure 5. The size, area, and volume of the tumor presented in pixel. The related algorithm explained below.
### Supervised Classification Algorithm

**Step 1:** Load MRI and feature image to convert it to gray image (imgy, imgF) and display.

**Step 2:** Apply the classification by take block from imgy by take block \([x, y, I, rect]\) = imcrop(imgy) where \(x, y\) are dimensions of block, \(I\) is number of blocks, \(I\) intensity and \(rect\) is MATLAB function lets you select a rectangle in the current axes using the mouse.

**Step 3:** Calculate mean and standard deviation of block and compare with each pixel of imgy for code
- \(\text{if } \text{imgy}(i,j) > \mu \& \& \text{imgy}(i,j) < \sigma\) then \(\text{imO}(i,j)=1;\) else \(\text{imO}(i,j)=0;\) end. [where \(\text{imO}\) mean binary image]

**Step 4:** Apply morphological process for \(\text{imO}\) (fill and clean)

**Step 5:** Calculate the largest object area in the image and determine its location accurately, \([\text{MAX } \text{A } \text{imx } ] = \text{to find a larger object area in the image and position, idx = location largest object}\ XAXA.\]

**Step 6:** Find all object that area less than \(\text{MAX A}\) and isolated \(= \text{BW3}\)

**Step 7:** Display binary tumor image \(T=\text{imO-BW3}\)

**Step 8:** Calculate area for \(\text{imgy}\) (A) and \(\text{imgC}\) (B) by code

*• create new image \(\text{imu}\) to calculate \(A=A+1\) and \(B=B+1\)*

*where \(\text{imu}=[0](r,c,p);\) for \(i=1:c\) and for \(j=1: c,\) if \(\text{imO}(i,j)==1\), To calculate \(A\) the total original image area for tumor only.*

*\(\text{imu}(i,j,:)=\text{img}(i,j,:); A=A+1;\)*

*where \(B=B+1;\) to calculate the total image area both the original and feature for the Complement (image without tumor).*

*Total area =\(A+B\)*

**Step 9:** Calculate geometric properties for tumor

*• \(\text{Calculate geometric properties for tumor}\)*

*Step 3:** Convert tumor image to binary image for threshold=0.2 by code \(\text{BW}=\text{im2bw} (\text{imO}, 0.2);\) \([h, w]=\text{size (BW)}\); \([L, \text{Ne}]=\text{bwlabel (BW)};\)

*Where \(\text{im2bw}\) This MATLAB function converts the grayscale image \(I\) to a binary image, \(\text{imO}\) is new tumor image, 0.2 = new threshold to convert, \(h, w\) element of size matrix, \(L=\)This MATLAB function returns a label matrix, \(L\), containing labels for the connected components in \(\text{BW}, \text{Ne} = \) value of table for image \(\text{BW}\.\)

**Step 4:** Use function \(\text{regionprop}\) in MATLAB to calculate area and dimensions tumor by code

*\(\text{stat = region props (BW);}\) and determine \(n= \) length (stat);*

*• Where \(\text{region prop = measurement thr dimensions of tumor in the image, nlc = mean length of case and measures number of tumor in image.}\)*

*• Rectangle = determine position of tumor and area.\)*

*• Bounding Box = calculate length and width of tumor and recognize tumor edges.\)*

**Step 5:** \(T=[0] (2, 2); \) \(K=1;\) where \(T\) is a zero matrix contains on two row and column \((2 \times 2)\) used to measure the dimensions of the tumor image.

**Step 6:** If \(T2 > T1\) then rate = \(T1/T2\) else \(T2/T1;\) where rate calculates ratio between the tumor size to the brain size

*• If rate > 0.6 then compute size of image \(\text{TO});\) \(\text{Size } V(\text{TO} ) = \frac{4}{3} \pi R^3\)*

**Step 7:** \(\text{rad} = \sqrt{(A / \pi)};\) where radius of image

*Where \(A= \) Total image \((A+B), \) \(\pi = 3.14.\)*

**Step 8:** Use \(\text{rat}\) to calculate \(A_T\) and \(V_T\) for block

**Step 9:** tumor area \(A_T=T1*T2*\pi;\) \(TT=T1+T2/2;\) then compute tumor size \(V_T=4/3*\pi*TT^3;\)

**End algorithm**
Figure 3. The details of the studied malignant image using presented method (a) original image with 12 features, (b) complement image with its features, (c) tumor image with its features, (d) the histogram of the original image with its feature, (e) the histogram of the complement image with its feature, and (f) the histogram of the tumor image with its feature.
Figure 4. The details of the studied benign image using presented method (a) original image with 12 features, (b) complement image with its features, (c) tumor image with its features, (d) the histogram of the original image with its feature, (e) the histogram of the complement image with its feature, and (f) the histogram of the tumor image with its feature.
5. Conclusions
The tumor detection within MRI image has high complexity. Therefore, two methods presented to check the quality and efficiency of tumor extraction. Segmentation and classification methods implemented with three MRI images which contain natural, benign and malignant tumor. The method of segmentation, which based on the threshold, is very efficient in the process of detection of tumors at very high accuracy. The separation of tumor area and complement area noticeable in the histogram schemes. Histogram of the tumor or complementary image illustrates the behavior of the texture feature of each image. The presented algorithm recommended diagnosing the tumor in medical images, which improves the performance of the automated diagnosis and reliability of the comprehensive. The dimension of the tumor gives to indicate the type of tumor. The T1 and T2 if almost equal then it is benign otherwise it is malignant.

The results of the two methods show that the GLRLM method has more details and accurate in terms of histogram pixel points comparing with the original image that has zero value in the same region. This lead to the performance of GLRLM depends on the gray level value. The location, detection, and separate tumor depend on the type, size, and position of the tumor in the image. The features are different in separating the tumor for instance, the SRE has an advantage in malignant, while COARNESS has an advantage in benign. With the complement image, the GLRLM method gives better results than the GLCM method, and this agrees with the original image and features results. The complement image considered as an original image with negative values.

References
[1] Ahmed H, Chintan P, John Q, Lawrence H, Hugo J. Artificial intelligence in radiology. HHS Public Access. 2018; vol.18 (8): pp.500–510.
[2] Zhang B, Gao Y, Zhao S, Liu J. Local derivative pattern versus local binary pattern. Face recognition with high-order local pattern descriptor. IEEE Transactions on Image Processing. 2010; vol. 19(2): pp.533-544.
[3] Nagy AM, Ahmed A, Zayed H. Particle filter based on joint color texture histogram for object tracking. Paper presented at the Proceeding of the First International Conference on Image Processing. Applications and Systems Conference (IPAS). 2014; pp.1-6.

[4] Raheja J L, Kumar S, Chaudhary A. Fabric defect detection based on GLCM and Gabor filter. A comparison, Optik-International Journal for Light and Electron Optics. 2013; vol. 124(23): pp.6469-6474.

[5] Patterns, ML. Breast density classification using multiresolution local quinary patterns in Mammograms. Paper presented at the Proceeding of the 21st Annual Conference in Medical Image Understanding and Analysis. 2017; pp.365-376.

[6] Marina M, Dragan J, Aleksandar P. Thermography based breast cancer detection using texture features and minimum variance quantization. EXCLI Journal .2014; vol. 13(1611): pp.1204-1215.

[7] Mahmoud Khaled Abd-Ellah, Ali Ismail Awad, Ashraf A. M. Khalaf and Hesham F. A. Hamed, Two-phase multi-model automatic brain tumour diagnosis system from magnetic resonance images using convolutional neural networks, number (79), 2018.

[8] 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 Cogn. Comput. 2019, 3, 27.

[9] Md Ahasan Kabir, Automatic Brain Tumor Detection and Feature Extraction from MRI Image, April 2020, The Scientific World Journal 8(4):695-711.

[10] Ayea S. D., using texture characteristics in diagnosing and distinguish tumors in medical images, thesis MSc., university of Baghdad, college of science for women, 2020.

[11] Ayea S. D. Heba Kh. Abbas, Anwar H. Mahdy and Haidar J. Mohamad, Adopting Texture Features to Detect and Recognize Brain Tumors in Magnetic Resonance Images, IOP Conference Series: Materials Science and Engineering, 757(2020) 012029, doi:10.1088/1757-899X/757/1/012029.

[12] Ayea S. D. Heba Kh. Abbas, Haidar J. Mohamad, and Ali A. D. Al Zuky, Adopting Run Length Features to Detect and Recognize Brain Tumor in Magnetic Resonance Images, IEEE, 2019 First International Conference of Computer and Applied Sciences (CAS), 23 April 2020.

[13] Ali I, Cem D, Melike S. Review of MRI-based Brain Tumor Image Segmentation Using Deep Learning Methods. World Applied Sciences Journal. 2016; vol. 32(10): pp.1484-1494.

[14] Abeer A, Vanitha U, Prabhu D, Pon Nageswaran N, Sathappan R. Tumor Detection in Brain Using Morphological Image Processing. 2015; vol. 1(1): pp.131-136.

[15] Alam Sh, Rasel A, Foisal H. Tumor detection in brain MRI image using template based K-means and Fuzzy C-means clustering algorithm. American Journal of Neuroradiology.2019; vol. 24(6): pp. 162-171.