Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 13, No. 1, (2022),
ISSN : **1906-9685**

DETECTION AND CLASSIFICATION OF TUMOR ON MRI IMAGES USING ANALYTICAL LEARNING ALGORITHM

[1] Pinapati Ravi Kumar, [2] Syed Naseema Sana, [3] Neelima Gottimukkala, [4]Tripuraneni Kiran Kumar
[1,2,3]Assistant Professor, [4]Student, Dept. of Electronics & Communication Engineering, Newton's Institute of Engineering, Macherla, Andhra Pradesh, India.

ABSTRACT
The brain is the most important and complex organ of the human body that controls every process that regulates the body. Currently, brain tumor cases are increasing rapidly and it leads to an increase in mortality. To cure a brain tumor and decrease mortality, a proper diagnosis method is required. A lot of methods have been introduced to cure brain tumors. Here initially an MRI image is acquired and it is pre-processed and segmented by performing K-means clustering to detect the tumor in the brain. K-means clustering is an unsupervised algorithm which groups the random dataset into different clusters. After the segmentation, the features are extracted using discrete Wavelet Transformations (DWT) and the classification of the brain tumor is done using the K-Nearest Neighbor (KNN) which is a supervised machine learning algorithm used to make highly accurate predictions. By using KNN Classifier, we achieve 97% of accuracy when compared to the state-of-the-art methods

Keywords
Brain tumor, K-means clustering, Segmentation, Discrete Wavelet Transforms, Classification, K-Nearest Neighbor.

Introduction

In our human body brain plays a vital role. The brain contains more than 100 billion nerves. Brain tumor is an uncontrolled mass of tissues in the brain. The identification of tumor in the brain is a difficult task because of its complex structure. Malignant tumor and benign tumor are two kinds of brain tumor. Benign tumor is non-cancerous as they do not spread to other parts of the brain and the body. Malignant tumors are cancerous and dangerous as they spread to the other parts of the brain and the body. If these tumors are detected and cured during its early stages then it may lead to death.

Many analytical algorithms were implemented in detection and classification of brain tumor. The neural network is used to classify the phase of brain tumor and Feature extraction by using the Gray Level Co-Occurrence Matrix (GLCM). Image recognition and image compression is done by using the Principal Component Analysis (PCA) method and Automatic brain tumor stage classification is done by using probabilistic neural network (PNN). PNN is fastest technique and also provide the good classification accuracy of 80% [1]. In comparison to the PNN method, Statistical methods are the major algorithms used and consist of few steps including pre-processing, feature extraction, segmentation, and classification. Performance of such statistical methods is an important factor for their successful adaptation. The results of these algorithms depend on the quality of images fed to the processing pipeline i.e., better the images, higher the results of accuracy of 86% [3]. Later a method to predict the grades of Gliomas using Radiomics imaging features is presented. MICCAI Brain Tumor Segmentation Challenge (BRATs 2015) training data, its segmentation ground truth and the ground truth labels were also used and features were selected through L1-norm regularization (LASSO). Gliomas were classified into low-grade glioma (LGG) or high-grade glioma (HGG) through logistic regression. This method achieves an accuracy of 89% [5]. A lot of methods have been applied in brain tumor detection ranging from image processing to signal-based analysis. Atish Chaudhary introduced a robust image processing-based method is applied using MRI images. Here, a clustering-based method is first used to segment the image and then SVM is applied for tumor detection and a total seven features were considered and 94.6% of accuracy is acquired [7]. the neuro fuzzy with binary cuckoo search optimization method is proposed for detecting tumors on MR images. The method has four stages. First a raw MR images are pre-processed and then the removal of the skull is classified. Later the functioning of singular value decomposition and principal component analysis takes place and finally, the NFBCS method is used to detect and classify tumors and the BCS algorithm optimizes the study model for better classification accuracy of 90% [9]. The obtained accuracy for detection and classification of brain tumor using adaptive k-means clustering and morphological operations, is 83.14% [11]. A deep-neural network structure, integrating DAE and RF, with a classification unit, which is used for the classification of brain MRI and obtained accuracy is 89% [13]. Finally, the segmented features are graded by the DAE with BMOA and RF. Classification of brain tumor in MRI images using a hybridized machine learning algorithm. Maximum A priori (MAP) firefly algorithm is proposed for feature selection. The accuracy obtained using MAP firefly algorithm is 88.3% [15].

When compared with above methods, our proposed methodology is pre-processing and segmentation using k-means clustering followed by features extraction using 2D-DWT and classification using k-nn classifier. We achieved 97% of accuracy, 96% of sensitivity and 100% of specificity. The K-NN classifier gives the highest accuracy when compared with other methodologies.

Preliminaries

K-means clustering [16]: It is an unsupervised machine learning algorithm that is used to segment the MRI image into regions by clustering.

$$L = \sum_{j=1}^{k} \sum_{i=1}^{N} \left\| x_i^j - c_j \right\|$$

Where,

Xi-Cj = distance of the data point from the cluster center.

L = the indicator of N data points from the cluster center.

2D- DWT: Discrete Wavelet Transform (DWT) algorithm is used for feature extraction. DWT is a wavelet transform method which provides both frequency and time information. 2D-DWT is applied in case of MRI images as they are of two dimensions. DWT operates on the filtering of low pass and high pass filters.

$$P(s) = \begin{cases} d_{ij} = \sum p(s) h * i(s - 2ij) \\ d_{ij} = \sum P(s) h * i(s - 2ij) \end{cases}$$

The coefficients di,j refers to the component attribute in signal p(s) corresponding to the wavelet function, whereas bi,j refer to the approximated components in the signal. The functions h(s) and g(s) in the equation represent high-pass and low-pass filters coefficients, respectively, while parameters i and j refer to wavelet scale and translation factors

http://doi.org/10.36893/JNAO.2022.V13N1.229-283

Metrics:

Entropy [17]: Entropy gives the information from each pixel of an image and measures the intensity distribution.

$$Entropy = -\sum_{I} P_j \log_2 P_j$$

RMSE: Root mean square error or root mean square deviation measured as square root of the mean of residuals

$$RMSE = \sqrt{\overline{(f-o)^2}}$$

Skewness: Skewness differentiate the darker and lighter part of the images. If the skewness value at a particular region is high then it is darker region and vice versa.

$$Skewness = \frac{mean - median}{standard\ deviation}$$

Kurtosis: Kurtosis gives the values of noise and resolution.

$$\text{Kurt}[X]=\text{E}\left[\left(\frac{x-\mu}{\sigma}\right)^4\right]$$

Energy: Energy is defined as the intensity of different pixels.

Energy = Intensity value of pixel - intensity value of neighboring pixel

Contrast: Contrast is defined as the ratio of Luminance difference to the Average luminance**.**

Contrast= Luminance difference/Average Luminance

Smoothness: Smoothness is defined as the extracting back the important patterns when noise is leaving.

S=2htan(x)/r

Here, h = difference of pixels, r = pixel number, and tan(x) is the reference angle.

Correlation: Correlation is a statistical measure that indicates the extent to which two or more variables fluctuate together.

$$\text{correlation} = \sum_{j=0}^{n-1} p_{ij} \frac{(\mathrm{i}-\mu)(j-u)}{\sigma^2}$$

Homogeneity: Homogeneity indicates how indistinguishable the pixels of an image are.

$$\text{Homogeneity} = \sum_{j=0}^{n-1} \frac{p_{ij}}{1+(\mathrm{i}-j)^2}$$

The classification results are shown using scalar parameters. They are:

Accuracy [18]: It is defined as a ratio between the correctly classified samples to the total number of samples as follows

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

Sensitivity: Sensitivity, True positive rate (TPR), hit rate, or recall, of a classifier represents the positive correctly classified samples to the total number of positive samples.

$$Sensitivity = \frac{TP}{TP+FN}$$

Specificity: Specificity, True negative rate (TNR), or inverse recall is expressed as the ratio of the correctly classified negative samples to the total number of negative samples

$$Specificity = \frac{TN}{TN + FP}$$

Methodology

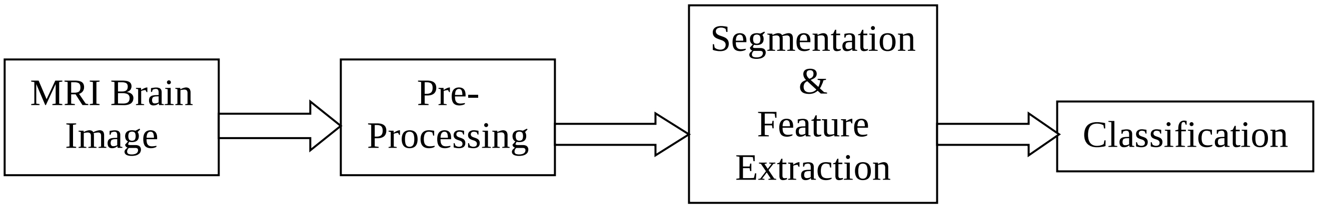

Fig 1. Block diagram of proposed methodology

The methodology was categorized into four stages as follows

a) Acquiring brain images
b) Pre-processing
c) Image segmentation and feature extraction
d) Classification

Acquisition of MRI Brain image:

The first stage follows the acquiring of MRI brain image. MRI image[19] is intrusive and it provides greater soft tissue contrast and can also alter better between the fat, water and muscle than the CT. It is also useful in distinguishing a wide variety of diseases.

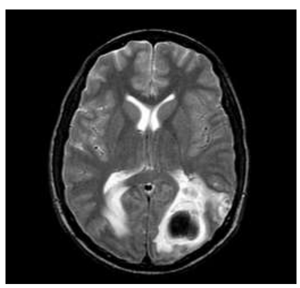

Fig 2. MRI brain image

Pre-processing:

Pre-processing is the first step in detection and classification of brain tumor. Pre-processing[20] step is performed to improve data images before the algorithmic processing and makes these images convenient for future processing. This step includes the following:

a) Image resizing
b) Gray scale conversion

Image resizing:

Image resizing [21] is used to enlarge the dimensions of MRI brain image. The image is converted to same dimensions, which representatively affects the file size and image quality. The result will be corrupted if the image is not having same dimensions. The brain image after resizing grants the best results for further process.

Gray scale conversion:

http://doi.org/10.36893/JNAO.2022.V13N1.229-283

After image resize, the resized image is transformed into a gray scale image. This process withdraws all color details, leaving only the gray scale of each pixel. The gray scale image [22] is convenient to bring out the features of brain tumor effortlessly. The RGB image is composed of pixels of dissimilar intensities. Hence it is transformed to a gray scale as it simplifies the algorithm and decreases the computational requirements.

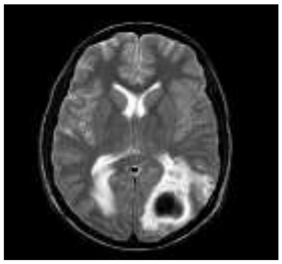

Fig 3. Preprocessed image

Image segmentation:

Image segmentation is the task of clustering parts of an image jointly that belong to the same object class. Image segmentation is performed to label the pixels of an image.

Segmentation has two objectives:

- To decompose the image into regions
- To change the representation

There are many Segmentation techniques that are used in the present day but here we used clustering-based Segmentation Algorithms.

K-means clustering is an unsupervised machine learning algorithm. K-Means clustering algorithm [23] is used as it's simple and well organized. The main motive of k-means clustering algorithm is to segment the MRI brain image into regions. The following is the representation of working of K-means algorithm,

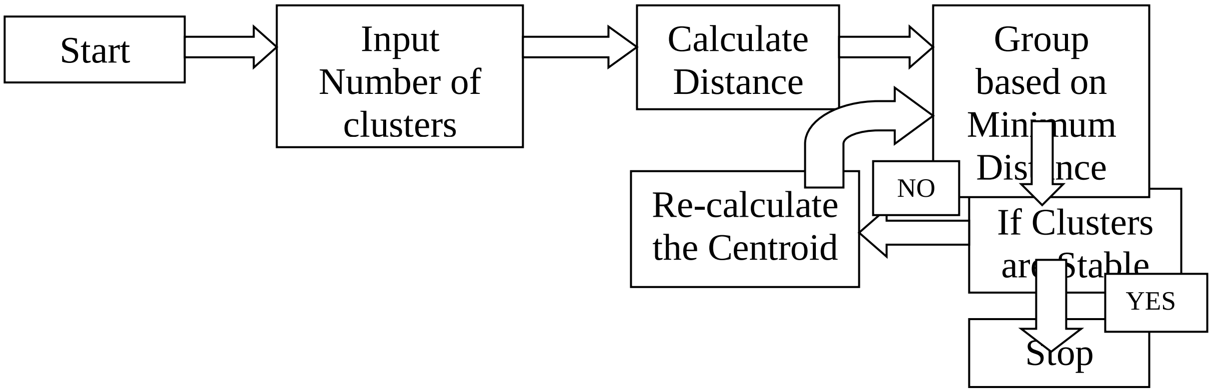

Fig 4. Block diagram of k-means clustering

Initially, the n datapoints are formed into the clusters. The clusters are formed by n datapoints differentiating with Euclidian distance i.e, the nearest data points form the clusters. The mean of the clusters forms the centroid. This procedure continues until no centroid is formed. Centroid is specified with K. If the centroid formed is 3 i.e., K=3, which means the image is segmented into three regions. The MRI brain image is segmented into tumor, non-tumor region and outer skull part of brain and the features are going to be extracted in the further process.

http://doi.org/10.36893/JNAO.2022.V13N1.229-283

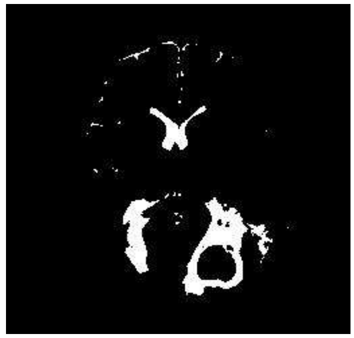

Fig 5.Segmented image

Feature extraction:

Feature extraction is a part of the dimensionality reduction process, in which an initial set of the raw details are split and decreased to more manageable groups. Wavelet transform is an effective tool for feature extraction from Segmented brain images. It gives information in both time and frequency domain. It is also having multi-resolution analytic property, which is able to examine the images at various levels of resolution.

Features are extracted from segmented image to classify the brain tumors. We apply 2D-DWT since the MRI image is two dimensional. 2D-DWT [24] operated through low pass and high pass filters on the image. The 2D discrete wavelet transform decomposes an image into four different resolutions of sub-bands, corresponding to low frequencies in the horizontal direction and high frequencies in the vertical direction LH (vertical details), high frequencies in the horizontal direction and low frequencies in the vertical direction HL (horizontal details), low frequencies in both directions LL (approximation image) and high frequencies in both directions HH (diagonal details).

The following features are extracted. They are:

a) Entropy
b) RMS
c) Skewness
d) Kurtosis
e) Energy
f) Contrast
g) Smoothness
h) Correlation
i) Homogeneity

Classification:

Classification phase is used to distinguish between benign and malignant tumor. This detection is important because prediction of tumor type at early stages can save the life of a patient. K-NN is a supervised machine learning algorithm used as classifier in real time applications. K-NN gives efficient results in classification of brain tumor. It classifies based on the features similarity. K-NN algorithm [25] consists of two stages namely

a) Training stage
b) Testing stage

In the training stage, stored data points in n space and points used to define attributes with a corresponding class. Whereas in the testing phase, calculated the distance between new extracted features and features from training data. The Euclidean distance is used to calculate distance between the training and testing stored data. In K-NN algorithm [26], initially the K value is chosen based on the available data points. Euclidian distance is measured from the data points. Based on the Euclidian distance, a new dataset is formed. Based on K value, we classify the tumor. K-NN algorithm classifies the tumor based on the features similarity. It is based on the prediction.

From the above process, we obtain 97% of accuracy in classification of brain tumor.

Experimental setup

To analyze our performance of proposed methodology, we consider the MRI brain images from the publicly available dataset. The dataset consists of 156 abnormal and 36 normal MRI brain images. The image dimensions are 256×256. The entire experiment was implemented in MatlabR2017 version conducted in AMD Ryzen3 processor.

Results & discussion

Quantitative analysis:

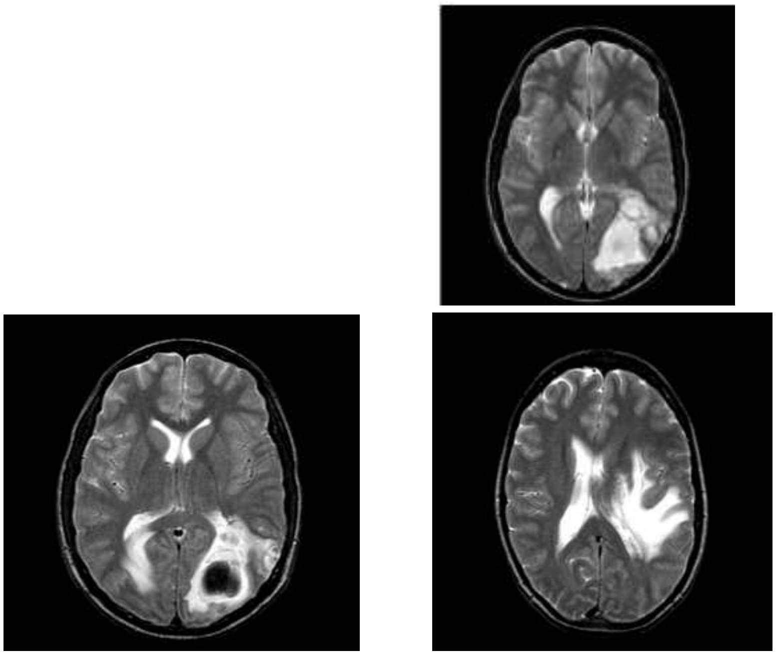

Fig a

Fig b

http://doi.org/10.36893/JNAO.2022.V13N1.229-283

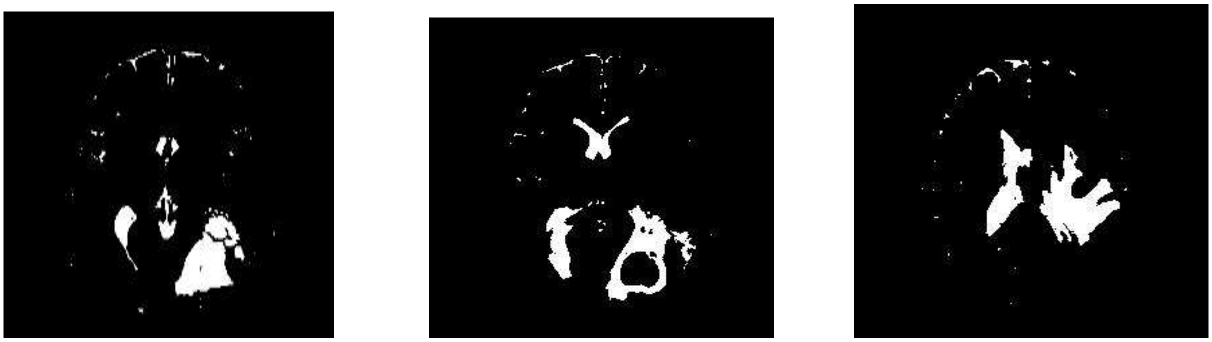

Fig c

Fig 6. a) MRI brain images b) pre-processed images c) segmented images

From the above figures, fig (a) represents the acquired MRI Images which are available publicly and these images undergoes pre-processing process. The fig (b) represents the pre-processed images which are done using MATLAB. After pre-processing process followed by segmentation using K-means clustering. Fig(c) represents segmented regions of brain image which is segmented into tumor region and non-tumor region.

Qualitative analysis:

Classifications performance using K-NN is represented by scalar values as metrics such as accuracy, sensitivity and specificity. The metric result obtained is :

Table 1: previous existing methods

Classification methods	**Accuracy**	**Sensitivity**	**Specificity**
PNN	80	80	80
DT	94	86	85
LASSO	89	88	90
SVM	90	80	80
NFBCS	95	88	99
NN	80	86	92
DNN	92	87	94
SVM	94	80	75
DNN	95	94	93
Proposed (K-NN)	97%	96%	100%

Conclusion

Brain tumor cases are rapidly increasing and mortality rate is higher due to brain tumor. So that, a proper diagnosis of brain tumor can be provided that leads to cure a brain tumor. Initially, acquired MRI brain images undergoes pre-processing step followed by segmentation process. Segmentation is done using a K-Means clustering method. K-means clustering is a region-based method. The k-means clustering is a unsupervised learning algorithm segments the MRI brain image into region of tumor and non-tumor. After the segmentation process, 2D-DWT is applied to extract the features of the tumor. The total 9 features were extracted from the segmented images using DWT. Last stage is classification process. K-NN algorithm is one of the best classifiers, so we used K-NN classifier which gives efficient results of accuracy, specificity and sensitivity. Support vector machine (SVM) and decision tree (DT) classifiers gives less accuracy when compared with K-NN classifier. Hence detection and classification of brain tumor using analytical algorithms gives high accuracy. We achieved 97% of accuracy and 96% of sensitivity and 100% of specificity. It is highest when we compared with state of art methods.

Future scope

In the future work, different classifiers can be used to increase the accuracy combining more efficient segmentation and feature extraction techniques with real- and clinical-based cases by using large dataset covering different scenarios.

References

[1]. Lavanyadevi, R., Machakowsalya, M., Nivethitha, J., & Kumar, A. N. (2017, April). Brain tumor classification and segmentation in MRI images using PNN. In 2017 IEEE International Conference on Electrical, Instrumentation and Communication Engineering (ICEICE) (pp. 1-6). IEEE.

[2]. S´ergio Pereira, Adriano Pinto, Victor Alves and Carlos A. Silva, "Brain Tumor Segmentation using Convolution Neural networks.

[3]. Tahir, B., Iqbal, S., Usman Ghani Khan, M., Saba, T., Mehmood, Z., Anjum, A., & Mahmood, T. (2019). Feature enhancement framework for brain tumor segmentation and classification. Microscopy research and technique, 82(6), 803-811.

[4]. Abbas, N., Saba, T., Rehman, A., Mehmood, Z., Kolivand, H., Uddin, M., & Anjum, A. (2018). Plasmodium life cycle stage classification based quantification of malaria parasitaemia in thin blood smears. Microscopy Research and Technique. https://doi.org/10.1002/jemt.23170

[5]. Cho, H. H., & Park, H. (2017, July). Classification of low-grade and high-grade glioma using multi-modal image radiomics features. In 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 3081-3084). IEEE.

[6]. Louis, David N., et al. "The 2007 WHO classification of tumours of the central nervous system." Acta neuropathological 114.2 (2007): 97-109.

[7]. Chaudhary, A., & Bhattacharjee, V. (2020). An efficient method for brain tumor detection and categorization using MRI images by K-means clustering & DWT. International Journal of Information Technology, 12(1), 141-148.

[8]. Shil SK, Polly FP (2017) An improved brain tumor detection and classification mechanism. In: ICTC

[9]. Arumugam, S., Paulraj, S., & Selvaraj, N. P. (2021). Brain MR image tumor detection and classification using neuro fuzzy with binary cuckoo search technique. International Journal of Imaging Systems and Technology, 31(3), 1185-1196.

[10]. Abbasi S, Tajeripour F. Detection of brain tumor in 3D MRI images using local binary patterns and histogram orientation gradient. Neurocomputing. 2017;219:526-535.

[11]. Reddy, K. R., & Dhuli, R. (2022). Segmentation and classification of brain tumors from MRI images based on adaptive mechanisms and ELDP feature descriptor. Biomedical Signal Processing and Control, 76, 103704.

[12]. D.N. Louis, H. Ohgaki, O.D. Wiestler, W.K. Cavenee, P.C. Burger, A. Jouvet, B. W. Scheithauer, The 2007 WHO classification of tumours of the central nervous system, Acta Neuropathol. 114 (2) (2007) 97–109.

[13]. Anantharajan, S., & Gunasekaran, S. (2021). Automated brain tumor detection and classification using weighted fuzzy clustering algorithm, deep auto encoder with barnacle mating algorithm and random forest classifier techniques. International Journal of Imaging Systems and Technology, 31(4), 1970-1988.

[14]. Amin J, Sharif M, Yasmin M, Fernandes S. A distinctive approach in brain tumor detection and classification using MRI. Pattern Recogn Lett. 2017;139:118-127. https://doi.org/10.1016/j.patrec.2017.10.036.

[15]. Deepa, B., Murugappan, M., Sumithra, M. G., Mahmud, M., & Al-Rakhami, M. S. (2021). Pattern Descriptors Orientation and MAP Firefly Algorithm based Brain Pathology Classification using Hybridized Machine Learning Algorithm. IEEE Access.

[16]. Kumar, A., Sinha, R., Bhattacherjee, V., Verma, D. S., & Singh, S. (2012, March). Modeling using K-means clustering algorithm. In 2012 1st International Conference on Recent Advances in Information Technology (RAIT) (pp. 554-558). IEEE.

[17]. Islam, A., Reza, S. M., & Iftekharuddin, K. M. (2013). Multifractal texture estimation for detection and segmentation of brain tumors. IEEE transactions on biomedical engineering, 60(11), 3204-3215.

[18]. Sokolova, M., Japkowicz, N., & Szpakowicz, S. (2006, December). Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation. In Australasian joint conference on artificial intelligence (pp. 1015-1021). Springer, Berlin, Heidelberg.

[19]. Naz, S., & Hameed, I. A. (2017, October). Automated techniques for brain tumor segmentation and detection: A review study. In 2017 International Conference on Behavioral, Economic, Socio-cultural Computing (BESC) (pp. 1-6). IEEE.

[20]. Bahadure, N. B., Ray, A. K., & Thethi, H. P. (2017, June). Feature extraction and selection with optimization technique for brain tumor detection from MR images. In 2017 International Conference on Computational Intelligence in Data Science (ICCIDS) (pp. 1-7). IEEE.

[21]. Mukhopadhyay, J. (2017, July). Image resizing in the compressed domain. In 2017 International Symposium on Signals, Circuits and Systems (ISSCS) (pp. 1-4). IEEE.

[22]. Papamarkou, I., & Papamarkos, N. (2013, June). Conversion of color documents to grayscale. In 21st Mediterranean Conference on Control and Automation (pp. 1609-1614). IEEE.

[23]. Kumar, A., Sinha, R., Bhattacherjee, V., Verma, D. S., & Singh, S. (2012, March). Modeling using K-means clustering algorithm. In 2012 1st International Conference on Recent Advances in Information Technology (RAIT) (pp. 554-558). IEEE.

[24]. Zhang, Z., Komazaki, N., Toda, H., Miyake, T., & Imamura, T. (2008, August). Directional selection of 2D Complex discrete wavelet transform and its application to image processing. In 2008 International Conference on Wavelet Analysis and Pattern Recognition (Vol. 1, pp. 146-151). IEEE.

[25]. Ulku, E. E., & Camurcu, A. Y. (2013, November). Computer aided brain tumor detection with histogram equalization and morphological image processing techniques. In 2013 International Conference on Electronics, Computer and Computation (ICECCO) (pp. 48-51). IEEE.

[26]. Ali, M., Jung, L. T., Abdel-Aty, A. H., Abubakar, M. Y., Elhoseny, M., & Ali, I. (2020). Semantic-k-NN algorithm: An enhanced version of traditional k-NN algorithm. Expert Systems with Applications, 151, 113374.

[27]. García, V., Mollineda, R. A., & Sánchez, J. S. (2010, August). Theoretical analysis of a performance measure for imbalanced data. In 2010 20th International Conference on Pattern Recognition (pp. 617-620). IEEE.

http://doi.org/10.36893/JNAO.2022.V13N1.229-283