iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 Issue: III Month of publication: March 2022

DOI: https://doi.org/10.22214/ijraset.2022.41092

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue III Mar 2022- Available at www.ijraset.com

Machine Learning and Image Processing based Segmentation and Classification of Brain Tumor

Harsh Vats Mishra[1], Vaishnavi Singh[2], Krishna Murari[3]
[1, 2, 3]BE, Computer Science and Engineering, HKBK College of Engineering, VyalikavalSociety, HBCS Layout, Nagavara, Bengaluru, Karnataka 560045

***Abstract:** In the field of Medicine, especially in diagnostic applications, finding the right one is a daunting task requires the attention of a radiologist. Error detection early it is necessary to avoid further complications. Many the developing field in the latest technology is MRI scanning. The size of the tumor in the brain may vary different patients and minute details of the tumor, It is a complex and tedious task for radiologists to diagnose and split the tumor into a large number of images. Occasionally, cerebral fluid also appears as a mass tissue on MRI image. The project aims to be automated a system that plays an important role in assessing whether a lump (mass of tissue) in the brain can be healthy (clump thickening) or bad (sticking to the edges) in stages. The proposed model uses machine learning algorithms in sequence to improve the accuracy of the separation. The program is loaded come out with four steps that include pre-noise removal processing using a flexible median filter, segregation using Gaussian Mixture Model (GMM) for finding the place of interest, extraction element using the Gray Level Co-occurrence Matrix GLCM to extract features of different types of tumors and segregation using Neural Networks (NN) to determine and classify the plant as harmful or harmful. The test results of the proposed model show that 93.33% accuracy, 96.6% accuracy, 93.33% sensitivity as well 94.44% accuracy. For these results the proposed model it works better compared to a classical machine learning algorithms like Adaboost (Adaptive Boosting) which is divides image into different classes (Normal, OK, Dangerous) with 89.90% accuracy.*

***Keywords:** K-means, Adaptive Median Filter, GMMSegmentation, Neural Networks, GLCM.*

I. INTRODUCTION

The brain tumor is known to create a major source of rapid increase in the mortality of children, adults and especially the elderly. As the human body is made up of millions of cells, and from biology we understand that cells reproduce, grow and divide to form new cells and tissues. Due to some external factors the cells may grow out of control which leads to tumor formation. The tumors are of two types, Benign tumors are cancerous cells and are less harmful as they do not spread to other cells. Although malignant tumors are a large number of cancerous cells, they are dangerous and more likely to spread to other cells and tissues. A tumor is basically a group of cells that form uncontrolled tissue as healthy cells with an unconditional growth rate. In the last decade, statistics show that in more developed lands, more than 300 people die from the effects of a brain tumor, and the number probably increases year on year. International statistics show that in the U.S. the ever-growing growth rate of plant growth is 11% - 12% of people with cancer every year. So all of these conditions lead us to develop a brain tumor identification model. It is very important to diagnose a tumor in the brain early. The MRI method is known for the clarity of the images I can scan. On MRI the appearance of the tumor is very accurate and high, for further treatment and medication medical attention is also required. Most of the time medical examinations are performed using MRI scans as they produce better results. MRI is therefore receiving significant attention and has a wide range in the future. With the advancement of computer technology and machine learning techniques, early detection of plants is possible. Photo (Fig. 1) was taken at a hospital showing an MRI scan of the brain.

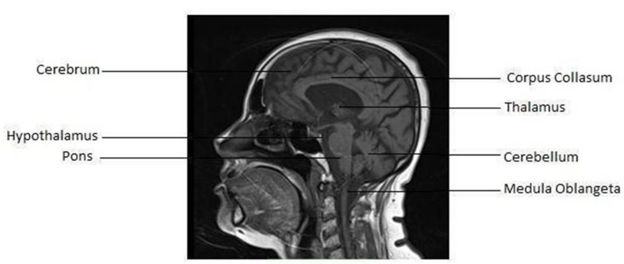


Figure 1. Image of Brain obtained after MRI scanning

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue III Mar 2022- Available at www.ijraset.com

II. LITERATURE REVIEW

Astina Minz [1] proposed a model in which image processing and machine learning methods were used to detect tumors, Median filters were used for image processing, a brain-segregation method, and the GLCM method. in order to extract the features, and finally the adaboost method is used to separate the tumor. Their results show 89.90% accuracy in tumor classification.

Neurosurgery networks in the division of tumors into normal or astrocytoma tumors in MRI images of different patients are gaining considerable attention. Paper [2] uses GLCM as a feature domain and trains emotional networks. Their results indicate that very few species should be considered for the effective classification of the plant into pilocytic (grade1), low grade (grade 2), anaplastic (grade3) and glioblastoma multiforme (grade4). The proposed system is based on the release of the GLCM feature and is used in neural networks that feed on MRI and CT images, the results of which suggest that the proposed system is 97% accurate. In this project [3] pre-processing is done to make the data sound more comfortable, no isolating is done and that is why the whole pre-processed image is used in feature extraction. Artificial intelligence techniques such as, neural networks and fuzzy logic neural networks have a major impact on this function [3]. Feature releases are performed using PCA and PNN. Conclusions suggest that PNN performance is effective and outstanding in diagnosing the tumor. System accuracy is between 73-80% depending on training data. Model that predicts a tumor to be dangerous or dangerous based on the factors provided during the training algorithm, the techniques used to predict in the proposed model are studied learning strategies that include Lineback and Obstruction [4]. These algorithms predict what type of tumor is based on the perception of features. Two factors to consider are the separation of clump thickness and marginal adhesion. The proposed system uses sigmoid and cost-effectiveness to reduce the number of errors. A decent gradient function is used to find the minimum earth data and divide the tumor. Implant imaging creates MRI images using machine learning concepts gaining great attention. Strategies used in the proposed model [2] are clear cosine modifications used to reduce image size and image purification so that it has no noise and neural networks possible to differentiate the tumor of its appropriate type. The network is trained for 20 brain tumor data samples and is 100% accurate in classification and has very little calculation time. The brain tumor detection model using the MRI images. This system revolves around the multi-model framework for detecting the presence of tumor in the brain automatically. In this system different MRI modalities are used training and evaluate the system [8]. For each MRI image features such as durability, shape, flexibility, balance and texture. These features are fed into the Adaboost browser. The adjusted results show that the system shows 90.11% accuracy. The author explains that no separation model can provide more accurate categories unless the images are processed using specific techniques to remove the sound and separate the ROI outside the image for more accurate results. In this median system filter [1], a large filter, small filters are used to maintain image clarity. In separating the two methods used, they are threshold and watershed to exclude the region of interest. Separation model that separates the tumor into different types using machine learning techniques. The proposed model uses a variety of integration techniques such as a feedback pulse integrated with the neural network to differentiate the tumor in the brain, a clear wavelength to transform the output factor and a PCA to reduce the size and divide the image into type1 or type2 tumor supply before and after broadcast of neural networks. are used. The results show that the system is able to detect vegetation with 99% accuracy which is excellent.

III. PROPOSED METHODOLOGY

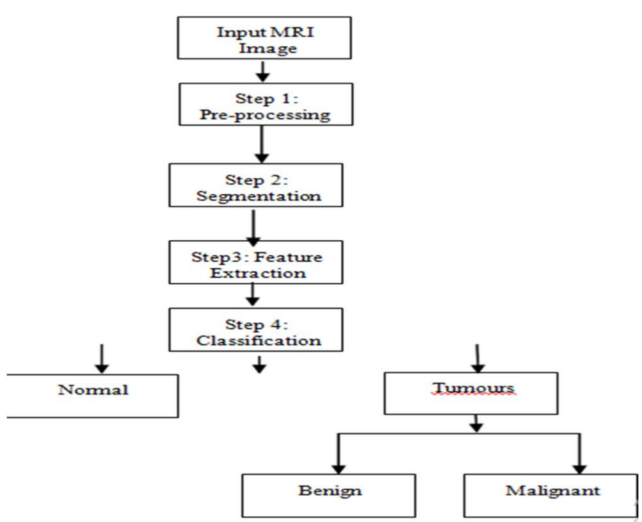


Figure 2

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue III Mar 2022- Available at www.ijraset.com

A. Input Brain MRI image

Image Acquisition is the first and foremost step of the process as shown in Fig. 2, where MR brain tumors with tumors and normal brain images are taken. The images used in this project are available in the Kaggle form. More than 60 MR image samples were collected in which 30 images were used in training and 30 images were used for experimental purposes. The figures below (3) show MRI images in dicom format.

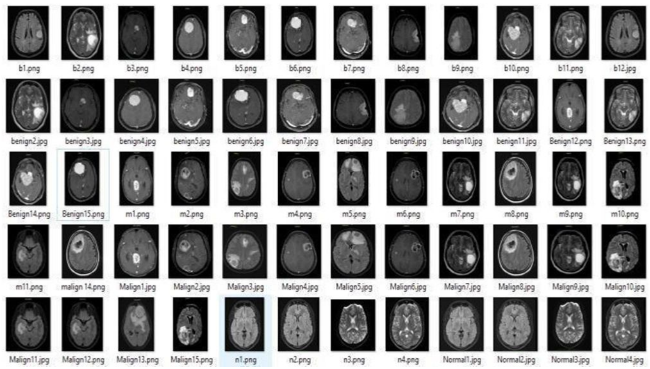


Figure 3

1) Step1: Pre-Processing

It is always necessary to have improved image quality in order to get the best results in previous steps. Pre-image processing plays an important role in bringing improved features of the image. In this proposed program to improve MRI image clarity, an effective median screening method is used. This is a very effective way to filter impulse data in images. On the other hand, in the normal medium filter, all image pixels are filtered equally i.e. both pixels are sound and image-free. Therefore a pre-processed image will result in blurred corners, deleted edges and blurred image. Many variations have therefore been introduced [1]. One such exception is the adaptive median filter, in this way the large-sized filter is used in the area of high-resolution pixels and the low-resolution filter is applied to small pixels. sound in picture input. The study [15] used this method by considering min min, max and med intensity using a window as a measure of size. The Adaptive median filter has two categories i.e. in the first stage the filter adapts to the image by distinguishing the category of noisy and noisy pixels. In the next section only sound pixels are considered and considered using a compatible median filter. In additional process no action is performed on small audio pixels and they are copied the same way into the processed image. Along the way, an adaptive median filer was used - the results of the previous processing using a modular and flexible filter as shown in fig (5,6). As can be seen in the picture (Fig. 5), the flexible middle filter is able to maintain fine lines, edges while the middle filter (fig. 4) fades the lines in the image and causes blurring.

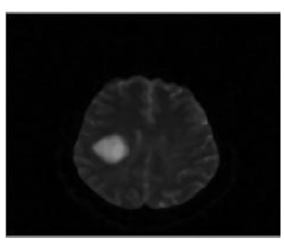


Figure 5

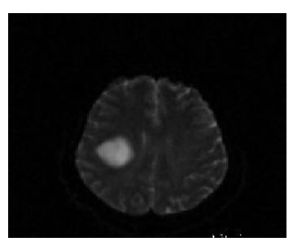


Figure 6

Processes using median filter processes using adaptive median

2) Step 2: Segmentation of ROI

Separation plays an important role in dividing an image into separate parts and only the useful part can be considered. This process should be done to reduce the amount of work to be done in additional steps i.e. the removal of the feature is easier if there is only a plant region whose features need to be removed from the whole image.

K-means is a duplicate unmodified compound. Each collection is categorized by its randomly placed centroid, in a way-it finds a minimum of costly work space and integrates. Finding the distance between centroid and Euclidean distance data points is the standard used. K-means is a solid integration method which means that a data point can belong to any collection. Here assigning a data point to the appropriate collection is based on the distance between the data point and the centroid randomly selected..

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue III Mar 2022- Available at www.ijraset.com

The proposed system uses a Gaussian hybrid model (GMM) [12] to separate the plant region without brain in the image. GMM is an advanced method over the K-means algorithm. In GMM it is a soft integration method which means that a data point can participate in both groups depending on the calculated possibilities considering the meaning and diversity combined. GMM continues to use the Expanded Expansion (EM) approach to integrate the mind into various areas and thus help to reduce data set. The results of K methods and GMM strategies are as shown in the figure (7,8). Fig Since the plant alone is a region of interest (ROI), GMM is doing a better job

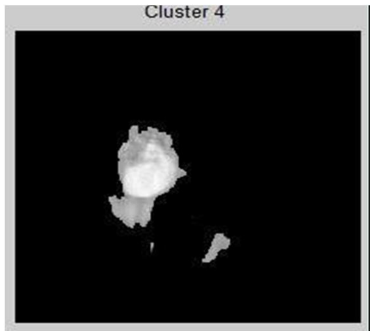


Figure 7
K-means Segmentation

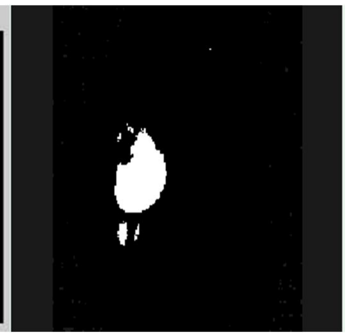


Figure 8
GMM Segmentation

3) Step 3: Feature Extraction

Feature removal is required as it reduces the representation of a large set of data. It is a way of removing elements from an image in the form of a vector.

The extracts are those that distinguish the elements between the distinct phase of the plant in which it is located. Converting data (image) into useful features that should be considered in classification is called feature extraction. The proposed model uses the Gray Level Co-Occurrence Matrix (GLCM) method for feature extraction [2]. This process is applied to the plant area divided into parts of the MR image thus extracting the textual features of the tumor. These features are released separately from the different types of tumor (harmful and harmful) and normal brain. Detection of text structures also plays an important role in extracting tumor structure structure.

GLCM is a mathematical method for extracting features. For GLCM, the matlab provides the function of determining the GLCM matrix i.e. through the function of the graycomatrix. Once the GLCM matrix has been detected using the above function, a few statistics can be deducted from the matrix.

a) Illustration of GLCM

The GLCM matrix [14] is obtained using the graycomatrix function. The function further determines how much the pixel has value i, which appears to be in the same position as the p. To calculate the GLCM matrix, the default function identifies the pixel to be processed as well as the horizontal and right pixels. Matlab also provides an opportunity for the user to define pixels with local relationships.

Each pixel in the graycomatrix effect has the value i, j, which is zero other than the number representing the number of pixels of the same intensity that occurs in the input matrix. In this way the acquisition of the graycomatrix helps to expose the properties of the image text.

The function is applied to an input image that converts to a matrix (first matrix). From fig (9) the matrix result of the graycomatrix function, the value of [1,1] is 1 as there is only one example in the input matrix where two horizontal pixels have 1 as their value respectively.

Automatic processing takes place horizontally, but can also be customized in the required directions. The next value in the matrix result i.e. [1,2] contains the value 2 because in the input matrix there it happens in two such cases when the values (1,2) occur in a horizontal way. Similarly, the third value i.e. [1,3] value is 0 as there is no [1,3] value event in the input matrix in a horizontal input. This process is continued until all the input matrix pixels find this function and perform a graycomatrix at the end of the process.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue III Mar 2022- Available at www.ijraset.com

b) Process Used to Create the GLCM

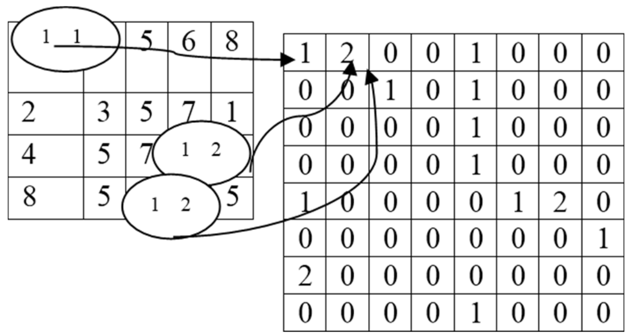


Figure 9. Process used to create GLCM matrix

There are totally 21 statistical features that are derived in thisproposed system. Few of them are shown in the below table 1

Table 1. Features extracted for a given input MR image

Statistics	Description
Contrast	Determines the difference in the intensity value between the adjacent pixels of the resultant graycomatrix
Correlation	It is one of the statistic measure used to determine how mutual and strong a relationship is between the pixels of the matrix
Energy	It identifies the sum of squared elements if thegraycomatrix which is known to be uniformity.
Homogeneity	Determines the quality, how much close the pixels of interest and adjacent are relatable to each other in the graycomatrix.

4) Step 4: Classification

In the classification process where the MRI image of the brain should be identified as normal, dangerous or dangerous, we analyze the statistical features of the input image and systematically combine the data into different categories. In the proposed system The Neural Network is used for training data (trained data) and test data (test data) and is therefore a phase 2 model. In the first section the elements of the image are separated and a different meaning is formed for each phase of the separation. In the second stage, the test data is captured and verified against the system. It therefore ensures that the images are properly classified.

The Mechanism of Machine Learning Techniques in the proposed system revolves around the idea of classifying brain scans of MRI into normal, dangerous or malignant. The main goal of machine learning is to automatically learn from training and ultimately make a wise decision with high accuracy. Neural Networks is used for segmentation. Neural networks are supervised learning algorithms using image elements extracted from the feature extraction process using GLCM. These features act as inputs to the neural network. Neural networks used for learning and based on learning divide the image into tumors or non-tumors by detecting the presence of a tumor in the brain of the MR image. In the proposed model the network input layer consists of 19 neurons representing the 19 mathematical elements released. The output layer contains 3 neurons that reflect Normal, Good, Dangerous. The hidden layer changes depending on the rules that provide the best level of recognition for each of the elements.

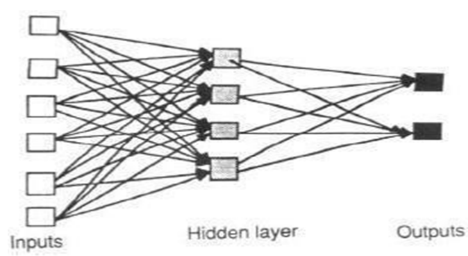


Figure 10. Example of neural network with single hidden layer

A common neural network with a single hidden layer as shown in Figure (10). Usually in neural networks output of different layers is given as input to subsequent layers using opening functions. This activation function is used to measure the output of a neural network in the correct range for better accuracy. In the proposed model there are 3 opening functions used:

a) Sigmoid Function

$$f(x) = \frac{1}{1 + e^{-x}}$$

b) Hyperbolic Tangent

$$f(x) = \frac{e^{2x} - 1}{e^{2x} + 1}$$

c) Linear Function

$$f(x) = x$$

IV. FUTURE SCOPE

The future of the project to train this algorithm on a large database and the accuracy of patient level on different databases can be made. To reduce the processing time the graphical processing unit (GPU) can be included in NN processing. The project approach will provide the basis for further research into the diagnosis of brain cancer using in-depth learning algorithms. This program will be expanded to identify different classes (Glioma, Meningioma) between Benign and Malignant cancers. The proposed method can also be tested in other medical imaging diagnoses such as Lung cancer, breast cancer and colonization. Different types of Neural networks can also be used to detect brain tumor. Accurate diagnosis of pre-cancerous growth using automated tools will help the patient to get the right treatment over time, as most cancers are curable only when diagnosed in the early stages. Once the system has achieved 100% accuracy in image classification, then this application can be distributed to improve hospitals.

V. ACKNOWLEDGEMENT

The successful completion of a project would not have been complete without the mention of people who have always guided me and supported the completion of the project. First of all, I would like to thank my director Prof. KSN Sushma, Professor, CSE Department has guided me across and focused on project improvement, supervision, support and significant contribution from the beginning of this research and provision. an unusual experience throughout the work. With the right support from Drs. Loganathan Professor and Head of Department of the CSE, I thank him for his valuable encouragement and suggestions in all the work. I would like to sincerely thank Prof. Hussain Ahmad, Principal, with kind support and permission to use the resources available at the center. Many thanks to the skills, the library staff and the CSE laboratory staff for their long-term support and communication.

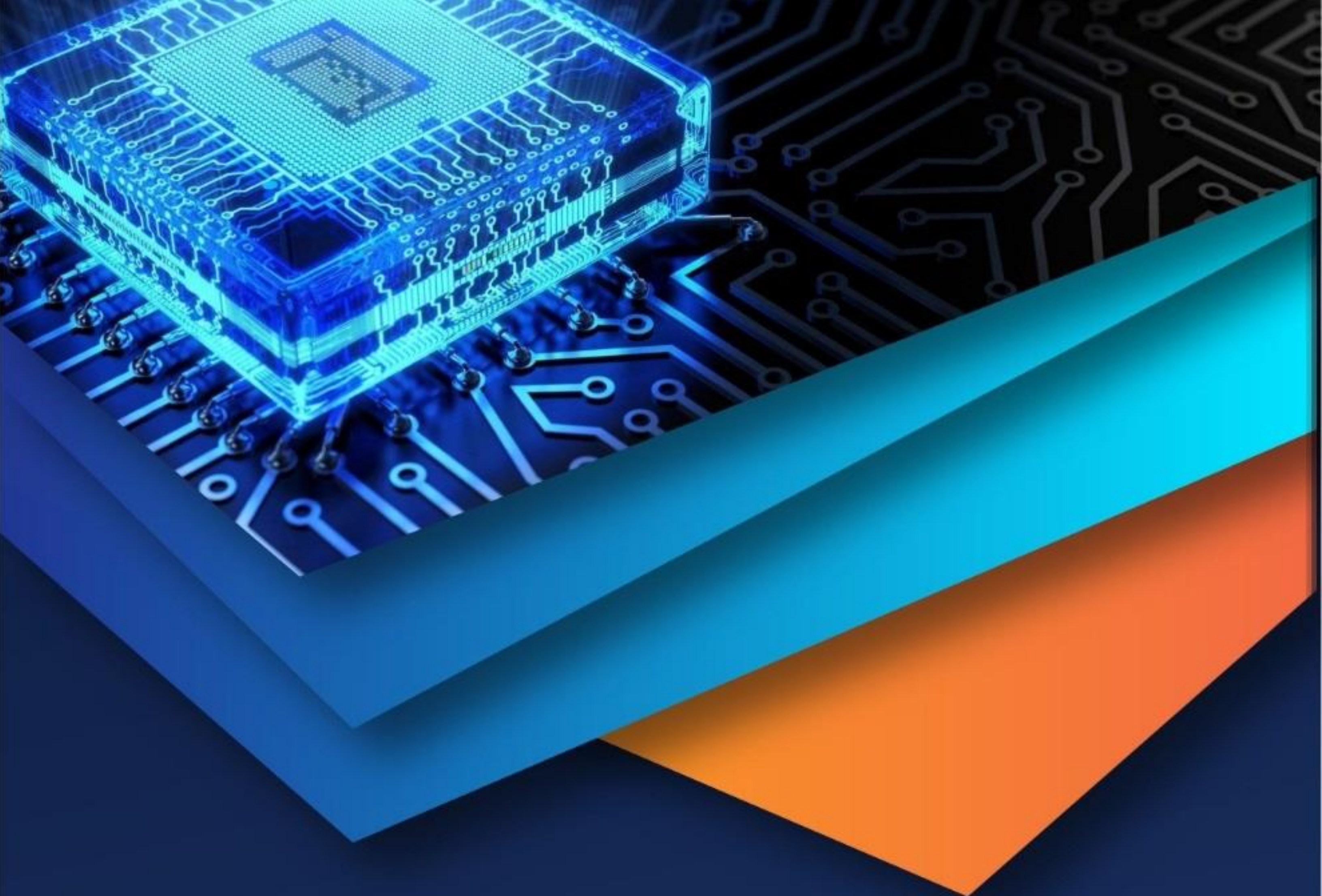
REFERENCES

[1] Astina Minz, “MR Image classification using adaboostfor brain tumor type”, 2017 IEEE 7th International Advance Computing Conference (IACC)

[2] “Segmentation of medical images for the extraction of brain tumors, A comparative study between the Hidden Markov and Deep Learning approaches” Soukaina, El Idrissi, El kaitouni and Hamid Tairi, ISCV 2020, Sept. 2020, doi:10.1109/ISCV49265.2020.9204319

[3] “Brain Tumor Detection- An Application Based on Machine Learning”, By Sharmila Gaikwad, Saarah Patel and Ajinkya Shetty INCET, 22 June 2021, doi:10.1109/INCET51464.2021.9456347

[4] “Brain Tumor Identification and Classification of MRI images using deep learning techniques” (IEEE Access, 2169-3536, 13 August 2020, doi:10.1109/ACCESS.2020.3016319)



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY