Rover Publications
United International Journal of Engineering and Sciences (UIJES)
An International Peer-Reviewed (Refereed) Engineering and Science Journal
Impact Factor:7.984(SJIF) Volume-4, Special Issue-3; ISSN: :2582-5887

Automated Brain Tumor Detection Through Advanced Imaging Analysis and Machine Learning

Dr. R Divya[1] Dr. Manjunath B E[2]
Assistant Professor[1], Professor[2]
Department of Electronics & Communication Engineering[1, 2]
R K College of Engineering[1, 2]
Vijayawada, India[1, 2]
sdivyareddy22@gmail.com[1], manjunathbe7666@gmail.com[2]
DOI:10.53414/UIJES:2024.43.217

Abstract – **This article explains the classification of two types of brain tumors using image filtering, segmentation, and enhancement in image processing and Support Vector Machine in machine learning, resulting in test images being classified with a minimum 62% accuracy rate.**

Keywords – **Image filtering, machine learning, tumors, segmentation, SVM.**

I. Introduction

Brain tumors are a highly dangerous illness, affecting around 12,000 people annually. The survival rate is 50% between 20-44 years old, but decreases to 5% after 65 years old [1]. Early diagnosis is crucial for rapid intervention. Machine learning and image processing methods can help detect missed diagnoses. MRI images can be transferred to digital software, which can be classified based on tumor size and location. This paper aims to detect brain images in a selected dataset using image processing and classify diagnosed tumors using machine learning techniques using MATLAB. Previous studies on brain tumor detection and classification methods have been useful for this study.

II. Literature Review

K-means clustering is a pixel-based segmentation method that detects tumor objects in magnetic resonance (MR) images by combining color translation, K-means clustering, and histogram clustering [2]. This method is easy to apply and efficient due to its minimal overhead.

Gray level co-occurrence matrix (GLCM) is used to extract second order statistical texture features for motion estimation of images using Vivado FPGA [3]. GLCM has high classification accuracy and requires less extraction time, making it one of the most efficient feature extraction methods.

The IBkLG classifier is used to classify normal, benign, and malignant tumors in brain MRI images using an automated method [4]. The system eliminates noisy images using median filtering and converts gray to RGB and l*a*b color space before transferring images to color-based segmentation for feature extraction.

DWT+PCA+SVM are a model developed to detect and classify normal and abnormal MR images of the brain [5]. DWT extracts information without loss, while PCA reduces image dimensions, improving classification accuracy but causing heavy computation burden. GRB SVM is the most accurate method for classification, with an accuracy level of 99.38%.

Skull stripping is a preprocessing method for MRI brain images, separating the brain area from the skull [6]. Morphology-based methods are used, including image thresholding, morphological operators filling gaps, edge detection, creating a binary mask, and combining the binary mask. In conclusion, these algorithms provide more efficient MRI images that are not affected by the surrounding region of the brain.

III. Dataset Definition

The dataset from Kaggle [7] contains three different tumor types: glioma, meningioma, and pituitary and normal brain images. Gliomas are tumors in the brain's glial cells, causing symptoms like headache, seizures, nausea, and vomiting (figure-1). Meningiomas are slow-growing, benign tumors that arise from brain membranes, causing symptoms by creating pressure on the brain (figure-2). Pituitary tumors are abnormal growths in the pituitary gland, causing abnormal hormone production (figure-3). The dataset is divided into training and test images, with 841 different images in the training file and 156 in the testing file. For the project, not all images were used due to different shooting angles, so MRI images were taken at the same shooting angle. The dataset was separated into training and test files for better accuracy.

Figure 1

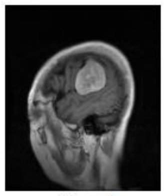

Figure 2

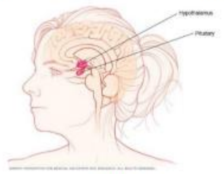

Figure 3

IV. Methodology

A. Image Processing (Pre-processing)

To begin classification of brain tumors, several Image Processing methods are employed. The skull stripping method (figure-4) removes the surrounding brain area, focusing on the tumor part. Median filtering (figure-5) removes noises from brain MR images using MATLAB's built-in function. Normalization is then applied to change the range of pixel intensity values. Image enhancement and segmentation are then used to adjust the digital images for display. The "imadjust" function is used for image enhancement, while image segmentation aims to make the image representation more meaningful. A morphology operator is applied to brain images, comparing the corresponding pixel in the input image with its neighbors (figure-6) [11]. After these steps, the MRI images are ready for GLCM to extract features.

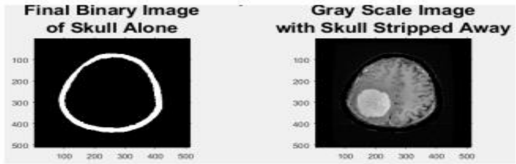

Figure 4 Skull Stripping

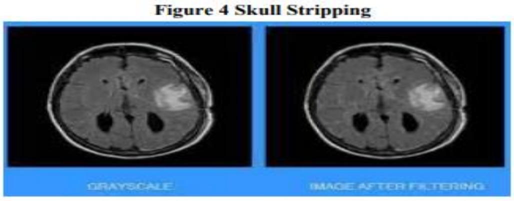

Figure 5 Median Filtering

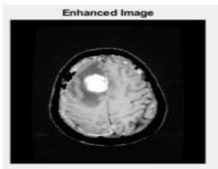

Figure 6

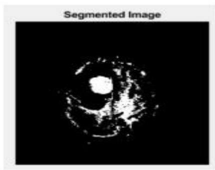

Figure 7

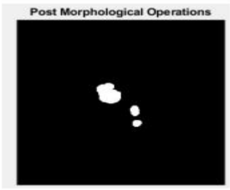

Figure 8

B. Feature Extraction (GLCM Texture)

The GLCM texture is a statistical method used to extract features from images after image processing. It considers the spatial relationship of pixels [6] and stores GLCMs in a 3-D matrix. The most important features extracted are Contrast, Correlation, Energy, and Homogeneity.

Contrast calculates local variations of the GLCM matrix, while Correlation calculates the joint probability of occurrence of pairs. Energy shows the sum of squared elements in GLCM [8], [9], and Homogeneity calculates the closeness of the distribution of elements.

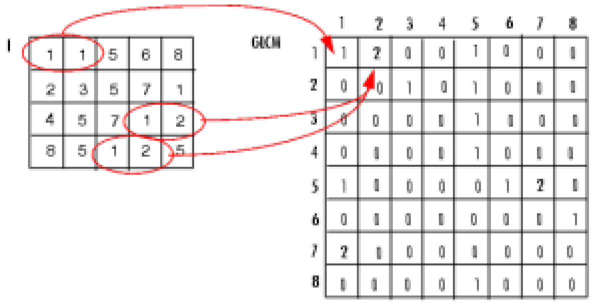

Figure 9 GLCM Algorithm

In this paper, 22 different features of each brain image were extracted, and an excel file was created to keep these values for classification. These features include Autocorrelation, Contrast, Correlation, Cluster Prominence, Cluster Shade, Dissimilarity, Energy, Entropy, Homogeneity, Maximum probability, Sum of squares, Sum average, Sum variance, Sum entropy, Difference variance, Information measure of correlation, Inverse difference, Inverse difference normalized, and Inverse difference moment normalized.

Table 1 GLCM Features Values

GLCM Features

Sample	tumour	autoc	contr	corrm	corrp	cprom	cshad	dissi	energ	entro	homom	homop
1	0	1.0857485	0.0002216	0.9431374	0.9431374	0.0294651	0.0148751	0.0002218	0.9958797	0.0156812	0.9998891	0.9998891
2	0	1.2535469	0.0044260	0.9716224	0.9716224	0.9331426	0.5064914	0.0044260	0.8296231	0.3172226	0.9977869	0.9977869
3	0	1.0722465	0.0017964	0.9622397	0.9622397	0.3418887	0.1758909	0.0017964	0.9506325	0.1256565	0.9991017	0.9991017
4	0	1.1741644	0.0045636	0.9587791	0.9587791	0.7127295	0.3786761	0.0045636	0.8847448	0.2493676	0.9977181	0.9977181
5	0	1.0040400	0.0001758	0.9360230	0.9360230	0.0206697	0.0104464	0.0001758	0.9970760	0.0116794	0.9999120	0.9999120
6	0	1.0162824	0.0005045	0.9539765	0.9539765	0.0827915	0.0418691	0.0005045	0.9885313	0.0375267	0.9997477	0.9997477
7	0	1.4370910	0.0058020	0.9768205	0.9768201	1.2274268	0.6952403	0.0058020	0.7439242	0.4505174	0.9970989	0.9970989

Sample	maxpr	sosvh	savgh	svarh	senth	dvarh	denth	inf1h	inf2h	indnc	idmnc
1	0.9979360	0.5873828	2.0039062	3.9612294	0.0155275	0.0002216	0.0020889	-0.8996464	0.1376864	0.9995261	0.9999536
2	0.9125338	0.7756333	2.1705060	3.7535571	0.3142347	0.0044260	0.0284066	-0.913443	0.6419593	0.9965246	0.9991347
3	0.9747202	0.6234539	2.0487631	3.7964817	0.1244113	0.0017964	0.0131516	-0.903838	0.4328190	0.9994011	0.9996407
4	0.9389027	0.7093365	2.1176308	3.7190973	0.2462045	0.0045636	0.0291496	-0.865233	0.5718808	0.9984787	0.9990872
5	0.9965362	0.5659299	2.0027519	3.9701678	0.0115545	0.0001758	0.0016959	-0.882652	0.1250624	0.9995413	0.9999648
6	0.9942361	0.5762788	2.0110231	3.9174887	0.0371769	0.0005045	0.0043347	-0.903489	0.2444052	0.9996318	0.9998990
7	0.8504349	0.9291806	2.2933280	3.9056022	0.4464957	0.0058020	0.0356630	-0.919807	0.7316789	0.9980659	0.9988395

C. Classification (SVM)

The classification learner app in MATLAB was used to train a dataset for this project, with the highest accuracy achieved using Kernel SVM. The fitcsvm function was used for training and cross-validation. The SVM model used Gaussian RBF (Radial Basis Function), which depends on the distance from the origin or point [10]. The common syntax for training and cross-validation is:
SVMModel=fitcsvm(X,Y,'KernelFunction','rbf',...'Standardize',true,'ClassNames',{'negClass','posClass'}).

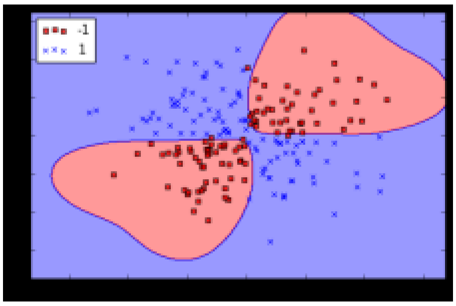

Figure 10 RBF Kernel

V. RESULT AND DISCUSSION

The algorithm was applied to Glioma and Meningioma tumor MRI images, revealing a low accuracy rate. This could be due to the high number of low-quality images in the dataset and the different perspectives of MRI images. The image processing part, such as the skull stripping algorithm, could also contribute to the low accuracy. The first K-means algorithm was used to detect tumors, but it had some disadvantages and didn't work on all MRI images. To increase accuracy, MATLAB's own algorithm, generated from 'Classification Learner', could be used. Although the Support Vector Machine algorithm has a higher accuracy rate, the SVM algorithm's results showed a lower accuracy rate. Rearranging the number of GLCM texture features could also improve the accuracy of the results.

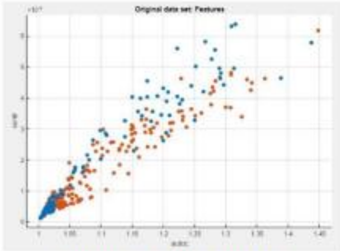

Figure 11 Spreading of Autocorrelation vs Contrast

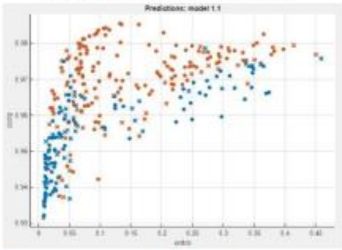

Figure 12 Spreading of Entropy vs Correlation

VII. CONCLUSION

The study successfully detected and classified various types of brain tumor images using image processing and machine learning techniques. The GLCM Texture Features method was used to extract characteristic features, while RBF Kernel SVM was used for classification. 200 training examples were used for each group, and the machine learning algorithm achieved a minimum accuracy rate of 62%.

REFERENCES

[1] https://www.medicinenet.com/brain_tumor_symptoms/views.html
[2] Ming –ni wu1 chia- chin lin2 chin-cheng-chan3, "Brain Tumor Detection Using Color-Based K-Means Clustering Segmentation" December 2007 DOI:10.1109/IIHMSP.2007.4457697, PP 1-3
[3] P. Mohanaiah1, P. Sathyanarayana2, L. GuruKumar3, "Image Texture Feature Extraction using GLCM Approach" International Journal of Scientific and Research Publications, Volume 3, Issue 5, May 2013 1 ISSN 2250-3153, PP 1-2
[4] Chauhan, S., More, A., Uikey, R., Malviya, P., & Moghe, A.A. (2017). Brain tumor detection and classification in MRI images using image and data mining. 2017 International Conference on Recent Innovations in Signal processing and Embedded Systems (RISE), 223-231.
[5] Yu-Dong Zhang and Lenan Wu, "An Mr Brain Images Classifier via Principal Component Analysis and Kernel Support Vector Machine," Progress in Electromagnetic Research, Vol. 130, 369-388, 2012.
[6] Swiebocka-Wiek J. (2016) Skull Stripping for MRI Images Using Morphological Operators. In: Saeed K., Homenda W. (eds) Computer Information Systems and Industrial Management. CISIM 2016. Lecture Notes in Computer Science, vol 9842. Springer, Cham.https://doi.org/10.1007/978-3-319-45378-1_16
[7] https://www.kaggle.com/sartajbhuvaji/brain-tumorclassification-mri?select=Training
[8] Avinash Uppuluri (2021). GLCM texture features MATLAB Central File Exchange. Retrieved June 9, 2021.
[9] https://www.mathworks.com/help/images/textureanalysis-using-the-gray-level-co-occurrence-matrixglcm.html
[10] https://towardsdatascience.com/support-vectormachines-svm-c9ef22815589
[11] https://www.mathworks.com/help/images/morphologicaldilation-and-erosion.html