Panamerican Mathematical Journal
ISSN: 1064-9735
Vol 35 No. 2s (2025)

Optimizing Medical Image Classification Using Diverse Feature Extraction Methods for Brain Tumor

Aarthi D[1], Sanjayprabu S[2], Panimalar A[3], Santhosh Kumar S[1], Mohana K[4]

[1]Sri Ramakrishna Mission Vidyalaya College of Arts and Science, Coimbatore, India.

[2]Rathinam College of Liberal Arts and Science, Coimbatore, India.

[3]KGiSL Institute of Engineering, Coimbatore, India.

[4]Nirmala College for Women, Coimbatore, India.

Corresponding Author: fuzzysansrmvcas@gmail.com

Article History:

Received: *20-09-2024*

Revised: *29-10-2024*

Accepted: *14-11-2024*

Abstract:

Detecting brain tumor early and accurately is crucial for effective treatment and better patient outcomes. Three feature extraction methods are used in this study: Gray Level Run Length Matrix (GLRLM), Gray Level Co-occurrence Matrix (GLCM), and Gray Level Histogram Features (GLHF)—to classify MRI images of the brain. GLCM measures the relationship between pixel intensities to capture texture information, while GLRLM finds patterns of pixels with similar gray levels, showing areas of texture. The histogram method summarizes the overall intensity in the image, highlighting differences between normal and abnormal areas. A Support Vector Machine (SVM), a classifier designed to differentiate between brain tissue affected by tumors and normal tissue, processes these retrieved features. Using a typical MRI dataset, the research assesses how well each feature extraction method supports accurate classification by the SVM. By comparing their performance, the study identifies which technique is best for distinguishing between healthy and tumor regions in the brain. This analysis offers valuable insights into improving brain tumor detection, potentially benefiting clinical diagnosis and treatment planning.

Keywords: Gray-Level Co-occurrence Matrix, Gray-Level Run Length Matrix, Gray-Level Histogram features, Support Vector Machine.

1. Introduction

Abnormal growths in the brain that interfere with normal brain activities are called brain tumors. They might be malignant (cancerous) or benign (non-cancerous), and improving treatment results and survival rates requires early identification. Because it can image soft tissues in great detail, magnetic resonance imaging (MRI) is a key technique for identifying and evaluating brain malignancies. However, manual MRI interpretation is time-consuming and prone to errors, underscoring the need for automated tumor detection methods.

While many automated classification methods exist, accurately distinguishing tumor regions from normal tissue in MRI images remains challenging. Conventional approaches frequently depend on manually created features that might not fully reflect the intricacy of tumor formations. Texture features [11], which capture spatial pixel arrangements, and intensity-based features, which summarize pixel distribution, are promising for improving classification accuracy. Notably, GLCM [12], GLRLM, and histogram-based features have shown effectiveness in medical image analysis.

This study uses MRI images to compare the effectiveness of GLCM, GLRLM, and histogram techniques for brain tumor classification. By extracting relevant feature values and applying a SVM classifier, this study evaluates how each technique differentiates between normal and abnormal brain tissue, to identify the most effective feature extraction method for accurate classification.

This study's comparative evaluation of these three popular feature extraction methods in relation to brain tumor classifying is one of its main contributions. The study explores how texture-based (GLCM, GLRLM) and intensity-based (histogram) features impact classification accuracy when used with an SVM classifier [7]. Based on empirical results from a standard MRI dataset, this paper provides practical recommendations for selecting the most effective feature extraction method for similar medical imaging tasks.

In image processing, fuzzy logic is a powerful tool for managing uncertainty and imprecision, commonly caused by noise, low contrast, or varying light conditions. Unlike traditional methods that rely on hard boundaries for classification, fuzzy logic allows for partial membership in multiple regions. Based on Zadeh's fuzzy set theory, this flexibility enables more accurate image analysis, particularly in edge detection, segmentation, and noise reduction.

By handling uncertainty through fuzzy membership functions, fuzzy logic enables the development of adaptive and resilient algorithms, enhancing segmentation, pattern recognition, and classification in medical imaging and other precision-required fields. This approach provides a robust, mathematically sound framework for addressing the inherent ambiguity and imprecision in complex image data.

2. Literature Survey

In the paper "Segmentation and Feature Extraction in Medical Imaging: A Systematic Review" by C.L. Chowdhary and D.P. Acharjya [5], The authors offer a thorough analysis of the extraction of features and segmenting methods utilized in medical imaging., focusing on techniques applied to MRI and CT scans. The paper highlights various segmentation approaches, such as thresholding, region-based, clustering, and edge-detection methods, along with feature extraction techniques like texture analysis, wavelet transformation, and statistical methods. Their methodology involves analyzing and summarizing recent techniques to offer insights into current trends and advancements. However, a drawback noted in the paper is the limited discussion on the computational complexity and real-time applicability of these methods, which could hinder practical implementation in clinical settings.

This paper presented by Latha et al [8] gives a study on advanced feature extraction techniques aimed at improving machine learning-based classification accuracy for carotid artery ultrasound images. The authors discuss various methods, including texture-based and statistical approaches, that enhance diagnostic precision by identifying significant features for artery classifications. The study offers information about the use of these techniques for accurate and early detection of vascular issues. However, a drawback is the limited focus on computational efficiency and adaptability across different imaging modalities, which may impact the broader applicability of these techniques in real-world clinical settings.

In this paper, Bhutto et al. [3] explore feature extraction in multimodal medical images by combining deep learning with contrast enhancement to improve diagnostic clarity and information retention. The authors utilize novel deep learning models to extract features from fused images, providing better visualization and potentially enhancing clinical decision-making. Despite promising results, a drawback of this study is the lack of comprehensive analysis on the computational demands and hardware requirements of their approach, which may limit its accessibility and integration into existing medical imaging workflows, especially in resource-constrained environments.

The paper "Enhancing Multiclass Brain Tumor Diagnosis Using SVM and Innovative Feature Extraction Techniques" by Basthikodi et al. [2] explores the use of Support Vector Machines (SVM) combined with advanced feature extraction methods—such as texture, shape, and intensity-based features—for improved multiclass classification of brain tumors. By focusing on refining feature extraction to enhance SVM performance, the study achieves high accuracy in distinguishing between multiple tumor types, showing promise for diagnostic applications. However, the paper has limitations, including a lack of scalability analysis for larger datasets, minimal discussion on real-time processing and adaptability to various MRI modalities, and limited focus on interpretability, which is essential for clinical diagnostics.

The use of GLRLM, GLCM, and GLHF in feature extraction offers distinct advantages in medical imaging analysis, particularly for enhancing classification accuracy in complex datasets like brain tumor classification. These techniques capture unique aspects of image texture and intensity: GLCM provides valuable spatial relationships between pixel intensities, GLRLM identifies the frequency and length of consecutive gray-level runs, and GLHF captures distribution-based intensity features. Together, they create a comprehensive feature set that addresses limitations seen in simpler feature extraction methods by offering a richer representation of tissue textures and patterns. This combination enhances model interpretability and supports scalability across larger datasets, as these techniques provide standardized, discriminative features that are adaptable to diverse MRI modalities. Additionally, their computational efficiency and robustness in capturing complex textural details make them suitable for real-time analysis, thereby addressing concerns around scalability and clinical applicability.

3. Image Preprocessing

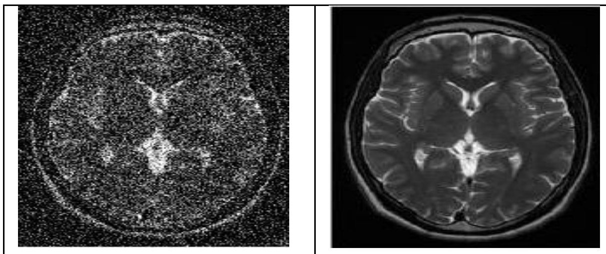
Image preprocessing is a crucial initial step in digital imaging that improves the quality and usability of images for analysis, especially in fields like medical imaging, computer vision, and machine learning. Preprocessing includes tasks such as noise reduction, contrast enhancement, and normalization, which enhance the visibility of important features and eliminate irrelevant information. Filtration techniques play a key role in this process by selectively smoothing or sharpening images to retain relevant details while reducing unwanted noise. For instance, filters like Gaussian, median, and adaptive filters are widely used to improve image clarity by removing high-frequency noise without distorting essential structures. Effective filtration helps in better feature extraction, leading to more accurate analyses and reliable model performance, as it allows algorithms to focus on relevant patterns and reduces the risk of errors due to noise or artifacts.

3.1 Filtration using Adaptive Fuzzy Gaussian 1D Filter (AFGF-1D)

This fuzzy Gaussian filter with an added derivative component offers significant advantages over traditional Gaussian filters. The inclusion of the derivative term allows it to adapt better to variations in intensity, which enhances the filter's sensitivity to edges and fine details. As a result, this filter provides superior noise reduction while preserving important features, such as edges and textures, more effectively. This adaptability to image intensity makes it especially useful in applications where edge detection and feature preservation are critical, such as medical and high-precision image processing. The Adaptive Fuzzy Gaussian 1D Filter [1] is represented mathematically by (1).

$$G(x,\sigma,m,\alpha)=\frac{1}{\sqrt{2\pi}\sigma}exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)\left[1-\alpha.\left(\frac{x-m}{\sigma^2}\right)\right] \quad (1)$$

where $m=\frac{1}{c\times d}\sum_{i=1}^{c}\sum_{j=1}^{d}A\,(i,j)$, $\sigma=\sqrt{\frac{1}{c\times d}\sum_{i=1}^{c}\sum_{j=1}^{d}(A(i,j)-m)^2}$, $A(i,j)$ is the pixel value, $\alpha=\frac{A}{\max(|A|)}$, x represents the pixel value's where it is located, $G(x,\sigma,m,\alpha)$ represent the Adaptive Fuzzy Gaussian 1D filter, $c\times d$ represent the dimension of the image. Using the above equation the image is filtered.



(a) (b)

Fig. 1. Represents (a) Gaussian Noise Image (b) Filtered Image using AFGF-1D

4. Feature Extraction

The feature extraction [10] process transforms raw data into useful, numerical features while preserving the data source in the dataset images. It yields better results than working directly with the raw data. Feature extraction increases the accuracy of acquired models by taking features out of the input data. By removing redundant data, this fundamental idea stage reduces the dimensionality of the data. It expedites training and inference, of course. This work involves extracting of feature values from the provided images using the GLRLM.

4.1 Gray Level Run-Length Matrix (GLRLM)

Greater-order statistics characteristics of texture to be extracted, it used in the GLRLM technique [6]. A line of pixels in a specific direction with the same intensity value is known as a grey-level run. The amount, referred to as the run length, is based on how many of these pixels there are in a grey level. The grey-level run length can be calculated using the below equations:

$$N = \sum_{i=1}^{N_m}\sum_{j=1}^{N_n} P_{ij},\ \ 1 \le N_z(\theta) \le N_p$$

where, N_m and N_n are the numbers of the pixel value and run length, respectively, in the image.

S.No.	Features	Formulas
1	Short Run Emphasis	$SRE = \sum_{i=1}^{N_m}\sum_{j=1}^{N_n}\frac{P_{ij}}{j^2} / \sum_{i=1}^{N_m}\sum_{j=1}^{N_n} \mathrm{P}_{ij}$
2	Long Run Emphasis	$LRE = \sum_{i=1}^{N_m}\sum_{j=1}^{N_n} j^2 P_{ij} / \sum_{i=1}^{N_m}\sum_{j=1}^{N_n} \mathrm{P}_{ij}$
3	Gray Level Non-uniformity	$GLN = \sum_{i=1}^{N_m}\left(\sum_{j=1}^{N_n} P_{ij}\right)^2 / \sum_{i=1}^{N_m}\sum_{j=1}^{N_n} \mathrm{P}_{ij}$
4	Run Percentage	$RP = \sum_{i=1}^{N_m}\sum_{j=1}^{N_n} P_{ij}\ / N$
5	Run-Length Non-uniformity	$RLN = \sum_{j=1}^{N_m}\left(\sum_{i=1}^{N_n} P_{ij}\right)^2 / \sum_{i=1}^{N_m}\sum_{j=1}^{N_n} P_{ij}$
6	Low Gray level Run Emphasis	$LGRE = \sum_{i=1}^{N_m}\sum_{j=1}^{N_n}\frac{P_{ij}}{i^2} / \sum_{i=1}^{N_m}\sum_{j=1}^{N_n} \mathrm{P}_{ij}$
7	High Gray level Run Emphasis	$HGRE = \sum_{i=1}^{N_m}\sum_{j=1}^{N_n} i^2 P_{ij} / \sum_{i=1}^{N_m}\sum_{j=1}^{N_n} P_{ij}$

Table 1. Represents GLRLM Feature Extraction Formulas.

4.2. Gray Level Co-occurrence Matrix (GLCM)

The distribution of co-occurring image pixels at a given offset is described by the co-occurring matrix, which is a matrix constructed through an image. Medical image analysis is one of the many applications for this texture analysis technique. GLCM is the name of the a second-order statistically analysis of texture technique[15]. It estimates the regularity with which a group of pixels occurs together in an image at a particular range θ and direction d. It also analyses the coordinates between pixels. In the GLCM feature extraction technique, some of the features are calculated using the following equations:

S.No.	Features	Formulas
1	Auto Correlation	$auto\ correlation = \sum_{i=1}^{N_m}\sum_{j=1}^{N_m}(i.j)P_{ij}$
2	Contrast	$contrast = \sum_{i=1}^{N_m}\sum_{j=1}^{N_m}(i-j)^2 P_{ij}$

3	Correlation	$correlation\ = \sum_{i=1}^{N_m}\sum_{j=1}^{N_m}\frac{(i.j)P_{ij-}\mu_x\mu_y}{\sigma_x\sigma_y}$
4	Energy	$energy = \sum_{i=1}^{N_m}\sum_{j=1}^{N_m}P_{ij}^2$
5	Entropy	$entropy = -\sum_{i=1}^{N_m}\sum_{j=1}^{N_m}P_{ij}\ log\ P_{ij}$
6	Homogenity	$homogenity\ = \sum_{i=1}^{N_m}\sum_{j=1}^{N_m}\frac{1}{1+(i-j)^2}P_{ij}$
7	Sum average	$sum\ average = \frac{1}{N_m.N_m}\sum_{i=1}^{N_m}\sum_{j=1}^{N_m}[i.P_{ij}+j.P_{ij}]$
8	Dissimilarity	$dissimilarity = \sum_{i=1}^{N_m}\sum_{j=1}^{N_m}\lvert i-j\rvert\ P_{ij}$

Table 2. Represents GLCM Feature Extraction Formulas.

Here, N_m is the maximum discrete intensity level in the picture. P_{ij} be considered the normalized co-occurance matrix, genrealized for any direction d and angle θ.

4.3. Gray level Histogram Feature (GLHF):

These textural characteristics reveal the homogeneity, smoothness, flatness, and contrast of the image's intensity level distributions. The probabilistic distributions of the levels of intensity in the histogram [9] bins are used to calculate the statistical textural properties of mean, variance, skewness, kurtosis, energy, and smoothness.

S.No.	Features	Formulas
1	Mean	$Mean = \frac{1}{\mathcal{N}}\sum_{\mathfrak{i}}\sum_{\mathfrak{j}}\mathfrak{p}(\mathfrak{i}-\mathfrak{k},\mathfrak{j}-\mathfrak{l})$
2	Variance	$Variance = \frac{1}{\mathcal{N}}\sum_{\mathfrak{i}}\sum_{\mathfrak{j}}[\mathfrak{p}(\mathfrak{i}-\mathfrak{k},\mathfrak{j}-\mathfrak{l})-M]^2$
3	Skewness	$Skewness = \frac{1}{\mathcal{N}}\sum_{\mathfrak{i}}\sum_{\mathfrak{j}}[\mathfrak{p}(\mathfrak{i}-\mathfrak{k},\mathfrak{j}-\mathfrak{l})-M]^3$
4	Kurtosis	$Kurtosis = \frac{1}{\mathcal{N}}\sum_{\mathfrak{i}}\sum_{\mathfrak{j}}[\mathfrak{p}(\mathfrak{i}-\mathfrak{k},\mathfrak{j}-\mathfrak{l})-M]^2-3$
5	Energy	$Energy = \frac{1}{\mathcal{N}}\sum_{\mathfrak{i}}\sum_{\mathfrak{j}}\mathfrak{p}(\mathfrak{i}-\mathfrak{k},\mathfrak{j}-\mathfrak{l}).\mathfrak{p}(\mathfrak{i}-\mathfrak{k},\mathfrak{j}-\mathfrak{l})$
6	Smoothness	$Smoothness = \frac{1}{\mathcal{N}}\sum_{\mathfrak{i}}\sum_{\mathfrak{j}}[\mathfrak{p}(\mathfrak{i}-\mathfrak{k},\mathfrak{j}-\mathfrak{l}).log2\mathfrak{p}(\mathfrak{i}-\mathfrak{k},\mathfrak{j}-\mathfrak{l})]$

Table 3. Represents GLHF Feature Extraction Formulas.

The discrete image is represented as $\mathfrak{p}(\mathfrak{i},\mathfrak{j})$, where $\mathcal{N}$ is the brain region's overall pixel count.

5. Classifiers

A model of machine learning known as a classifier [13] divides data into distinct classifications, such as identifying tumorous and non-tumorous areas in medical images. When processing medical images, classifiers are crucial for accurately diagnosing conditions, aiding in early detection, and supporting treatment planning by automatically identifying abnormalities. Their accuracy directly impacts patient outcomes, making them essential for reliable medical analysis.

5.1 Support Vector Machine (SVM) Classifier

A popular supervised machine learning approach for classification tasks, such detecting brain tumors in MRI data, is the Support Vector Machine (SVM). Finding the best hyperplane to divide data points of several classes with the largest margin is how SVM [4] works. Particularly in high-dimensional feature spaces, which are frequently seen in medical imaging data, this ideal margin aids SVM [14] in achieving high classification accuracy.

For a set of training data points (x_i, y_i) where x_i represents the feature vector of the ith image, $y_i \in \{-1,1\}$ denotes the class label where 0 represents tumor and 1 represents non-tumor and $i = 1,2,\dots,n$. SVM's primary goal is to maximize the margin between the two classes to identify the ideal hyper plane, as indicated by (1), that maximally separates them.

$$w.x + b = 0 \tag{3}$$

The optimization function is represented by:

$$min_{w,b}\ \frac{1}{2}\|w\|^2 \tag{4}$$

Subject to:

$$y_i(w.x + b) \geq 1, \forall i \tag{5}$$

where , $\|w\|$ stands for the margin width, b for the bias term, and w is the weight vector perpendicular to the hyperplane.

The decision function is given by $f(x) = \pm(w.x + b)$. For MRI brain image,

$$f(x) = \begin{cases} (w.x + b) > 0 & tumor \\ (w.x + b) \leq 0 & non - tumor \end{cases} \tag{6}$$

5.2 Performance Metrics

The performance of classifiers is calculated using the following metrics given in Table 4

S. No.	Metrics	Formulas
1	Accuracy	$\frac{True\ Positive + True\ Negative}{Total\ Sample}$
2	Precision	$\frac{True.Positive}{True.Positive + False\ Positive}$

3	Recall	$\frac{True\ Positive}{True\ Positive + False\ Negative}$
4	F1 Score	$2 \times \frac{Precision\ \times Recall}{Precision + Recall}$
5	Sensitivity	$\frac{True\ Positive}{True\ Positive + False\ Negative}$
6	Specificity	$\frac{True\ Negative}{True\ Negative + False\ \ Positive}$
7	Error Rate	$1 - Accuracy$
8	Balanced Classifier Rate	$\frac{Sensitivity + Specificity}{2}$

Table 4. Formula for performance metrics

Pseudo Code of SVM Classifier for brain tumor classification is given as follows:

```
1. Pre-Process Data
    1.1 Load MRI image data.
    1.2 Filtration using AFGF-1D
2. Feature Extraction
    2.1 Three techniques namely GLCM, GLRLM, GHLF are used for feature extraction respectively.
    2.2 Separate dataset
        2.2.1 Features: Matrix with each row representing extracted features for MRI image.
        2.2.2 Labels: Vector where each label is 1 (for tumor) or 0 (for non-tumor)
    2.3 Devide the data into training and testing set
3. Train SVM Model
    3.1 Set Regularization parameter $C$
    3.2 For each training sample compute $f(x_i) = w.x_i + b$
    3.3 Compute weight vector using $w = \sum_{i=1}^{n} \alpha_i y_i x_i$
    3.4 Calculate the bias $b$ using support vectors
4. Classification on Test Data
    4.1 For each MRI image in the set compute $f(x) = (w.x + b)$
    4.2 Assign label based on the sign of $f(x)$. If $f(x) \geq 0$, then its tumor (label 1) or else if $f(x) < 0$, then its non-tumor (label 0)
    4.3 Calculate Performance Metrics: Confusion Matrix, Accuracy, Precision, Precision, F1 Score.
```

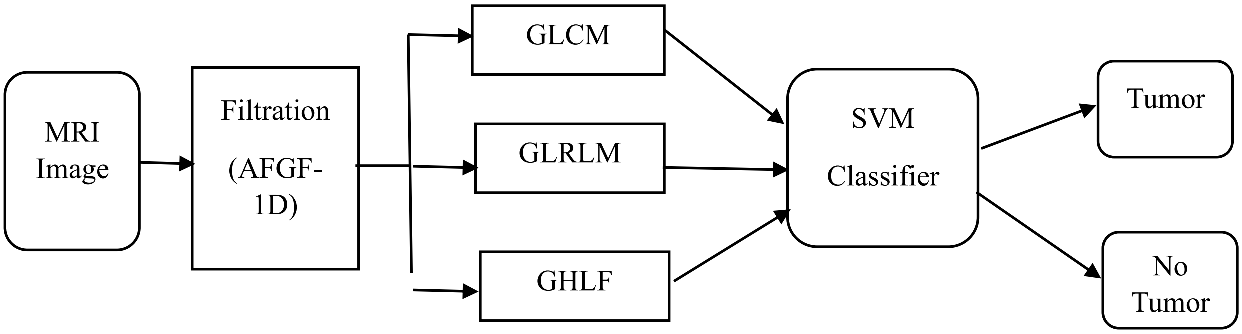


Fig. 2. Flow chart of the proposed algorithm

6. Results and Discussion

In this, an algorithm is designed such that the MRI brain image is analyzed and then classified as tumor or non-tumor. For this algorithm, three feature extraction techniques namely GLCM, GLRLM, and GLHF are used. The classifier based on the SVM is trained using the extracted features, and the classifier's performance is computed. For this process, a set of 253 labelled MRI brain images is collected from Kaggle database. Also, MATLAB 2017 is used for implementing the algorithm. Table 5, Table 6, Table 7 shows the extracted feature values of normal and abnormal images of GLRLM, GLCM and GLHF respectively.

Features/Images	**Normal**	**Abnormal**
Short Run-Emphasis	0.579791	0.564737
Long Run-Emphasis	286.2814	107.7168
Gray Level Non-uniformity	5512.297	5852.888
Run Percentage	0.794724	1.084656
Run Length Non uniformity	16363.24	21525.62
Low Gray level Run Emphasis	0.134289	0.06222
High Gray level Run Emphasis	37.36023	75.03617

Table 5. Feature values of GLRLM

Features / Images	**Normal**	**Abnormal**
Contrast	16000	27352
Correlation	0.929851	0.979252
Energy	0.374171	0.153274
Entropy	1.4397	2.395131
Homogeneity	0.949861	0.913993
Sum of squares Variance	0.87349	5.048716
Sum average	3.339277	6.970895
Sum variance	3.371413	27.54369
Sum entropy	1.337865	2.221654
Difference.variance	0.481124	0.774238
Difference.entropy	0.344432	0.495559
Information measure of correlation-1	-0.70608	-0.71535
Information measure of correlation-2	0.890065	0.96466

Contrast	16000	27352
Correlation	0.929851	0.979252
Energy	0.374171	0.153274
Entropy	1.4397	2.395131
Homogeneity	0.949861	0.913993
Sum of squares Variance	0.87349	5.048716
Sum average	3.339277	6.970895
Sum variance	3.371413	27.54369
Sum entropy	1.337865	2.221654

Table 6. Feature values of GLCM

Features/Images	**Normal**	**Abnormal**
Mean	2.146194	3.853592
Variance	2.002782	6.054202
Skewness	1.066894	0.400259
Kurtosis	3.623596	2.189491
Energy	0.34704	0.182688
Smoothness	1.345367	1.923504

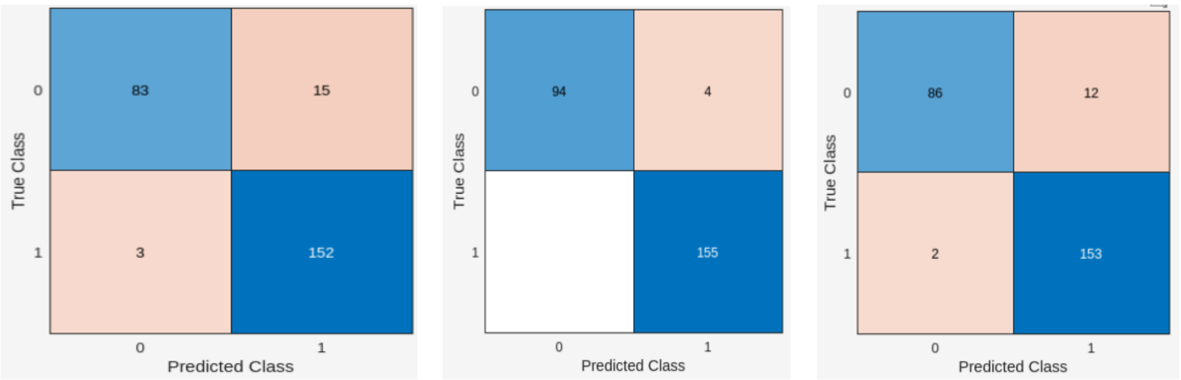
Table 7. Feature values of GLHF

The results shown in Table 8 states that GLRLM feature extraction method yields the best overall performance for MRI brain tumor classification, achieving the highest sensitivity (100%), specificity (97.48%), accuracy (98.41%), and F1 score (97.91%), indicating it effectively identifies both tumor and non-tumor cases with minimal misclassification. GLCM and GLHF also perform well but are slightly less accurate, with lower precision and error rates, suggesting some misclassifications. The Balanced Classifier Rate and MCC (Matthews Correlation Coefficient) confirm GLRLM's superiority.

Features	**GLCM**	**GLRLM**	**GLHF**
Sensitivity	96.51	**100**	97.72
Specificity	91.01	**97.48**	92.72
Accuracy	92.88	**98.41**	94.46
Error rate	7.11	**1.58**	5.53
Precision	84.69	**95.91**	87.75
F1 Score	90.21	**97.91**	92.47
Jaccard Metric	82.17	**95.91**	86
Balanced Classifier Rate	93.76	**98.74**	95.22

Table 8. Performance Analysis of Classifiers

By showing the numbers of True Positives, True Negatives, False Positives, and False Negatives, a confusion matrix is a table that assesses the effectiveness of a classification model. For MRI brain tumor detection, it shows how accurately the model identifies tumor versus non-tumor cases, highlighting areas of misclassification. Metrics like accuracy, sensitivity, specificity, and precision can be derived from it to assess the model's reliability. All things considered, the confusion matrix sheds light on the model's advantages and disadvantages when it comes to MRI image classification. In Figure 3, the confusion matrix is displayed.



(a) (b) (c)

Fig. 3. Confusion Matrix of (a) GLCM (b) GLRLM (c) GLHF

A scatter plot visualizes classification results by plotting data points based on features, with each class represented by different markers. In MRI brain tumor classification, scatter plots can illustrate how well the model separates normal from abnormal cases, showing the clustering of each class. This helps assess model performance by visually identifying areas of overlap or separation between classes. Scatter plots are especially useful for understanding patterns and improving feature selection. The results of the scatter plot of GLCM, GLRLM, and GLHF are shown in Fig. 4, 5, and 6 respectively.

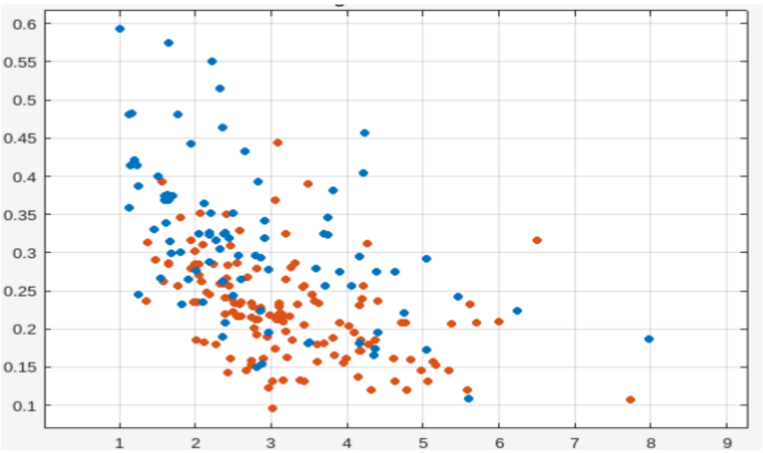


Fig.4. Scatter Plot GLCM

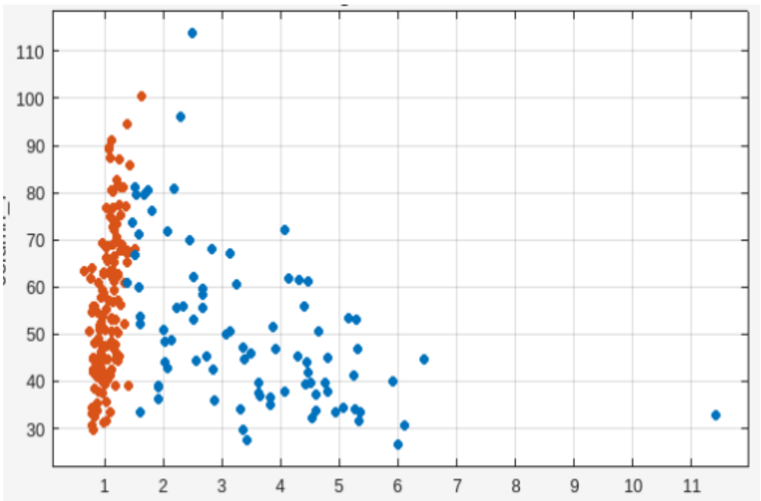


Fig.5. Scatter Plot GLRLM

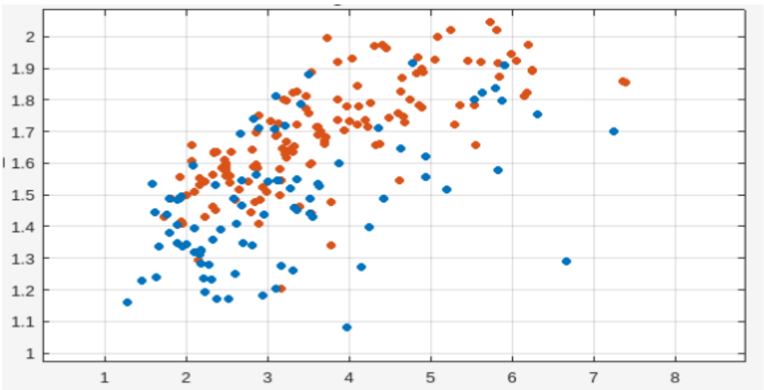


Fig. 6. Scatter Plot GLHF

7. Conclusion and Future Work

This study aims to develop an algorithm for classifying MRI brain images as normal or abnormal. To achieve this, three distinct feature extraction techniques—Gray Level Co-occurrence Matrix (GLCM), Gray Level Run Length Matrix (GLRLM), and Gray Level Histogram Features (GLHF)—were employed to extract meaningful features from the MRI images. These extracted features were then evaluated using a Support Vector Machine (SVM) classifier for classification. Based on the experimental results, it was observed that the GLRLM-based feature extraction technique outperformed the other methods, demonstrating superior classification accuracy when combined with the SVM classifier. This highlights the potential of GLRLM in enhancing the performance of brain tumor classification systems.

Future work could focus on developing a more refined classifier using sophisticated machine learning methods to improve the accuracy of tumor classification, such as deep learning models. Additionally, implementing a hybrid model that combines multiple feature extraction methods or classifiers may further improve sensitivity and specificity. This approach could leverage the strengths of each model, offering a robust solution for MRI brain tumor classification.

References

[1] Aarthi D, Panimalar A, Santhosh Kumar S, Anitha K, Development of adaptive gaussian filter based denoising as an image enhancement technique. Mathematical Modelling of Engineering Problems, Vol. 11, No. 10, pp. 2715-2725, 2024. https://doi.org/10.18280/mmep.111013

[2] Basthikodi, M., Chaithrashree, M., Ahamed Shafeeq, B.M. *et al.* Enhancing multiclass brain tumor diagnosis using SVM and innovative feature extraction techniques. *Sci Rep* 14, 26023 (2024). https://doi.org/10.1038/s41598-024-77243-7

[3] Bhutto, J.A., Guosong, J., Rahman, Z. *et al.* Feature extraction of multimodal medical image fusion using novel deep learning and contrast enhancement method. *Appl Intell* 54, 5907–5930 (2024). https://doi.org/10.1007/s10489-024-05431-z

[4] Cano Lengua M A and E. A. Papa Quiroz, A Systematic Literature Review on Support Vector Machines Applied to Classification, IEEE Engineering International Research Conference, Lima, Peru, pp. 1-4, 2020. https://doi.org/10.1109/EIRCON51178.2020.9254028.

[5] Chiranji Lal Chowdhary, D.P. Acharjya, Segmentation and Feature Extraction in Medical Imaging: A Systematic Review, Procedia Computer Science, Volume 167, 2020, Pages 26-36, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2020.03.179.

[6] Dash, S., Senapati, M.R. Gray level run length matrix based on various illumination normalization techniques for texture classification, Evolutionary Intelligence, Vol. 14, 217–226 ,2021. https://doi.org/10.1007/s12065-018-0164-2

[7] Jair Cervantes, Farid Garcia-Lamont, Lisbeth Rodríguez-Mazahua, Asdrubal Lopez, A comprehensive survey on support vector machine classification: Applications, challenges and trends, Neurocomputing, Volume 408, Pages 189-215, 2020. https://doi.org/10.1016/j.neucom.2019.10.118.

[8] Latha S, Muthu P, Dhanalakshmi, Samiappan, Kumar R, Lai Khin Wee, Wu, Xiang, Emerging Feature Extraction Techniques for Machine Learning- Based Classification of Carotid Artery Ultrasound Images, Computational Intelligence and Neuroscience, 1847981, 14 pages, 2022. https://doi.org/10.1155/2022/1847981

[9] Ling-Hui Liu and X. Luan, A novel feature extraction method using histogram-based sparsity, International Conference on Machine Learning and Cybernetics Jeju, Korea (South), 2016, pp. 1061-1065, https://doi.org/10.1109/ICMLC.2016.7873026.

[10] Mutlag, Warmidh and Ali, Shaken and Mosad, Zahoor and Ghrabat, Bahaa Hussein, Feature Extraction Methods: A Review, Journal of Physics Conference series, 1591 issue 1:012028,2020. https://doi.org/101088/1742- 6596/1591/1/012028

[11] Sabry, E.S., Elagooz, S., El-Samie, F.E.A. *et al.* Assessment of various feature extraction methods for object discrimination in different scenarios. Journal of Optics, Vol. **53**, 49–69 (2024). https://doi.org/10.1007/s12596-023-01132-x

[12] Sanjayprabu S, Sathish Kumar R, Kavitha Somasundaram, Karthikamani R, Brain Tumor detection using GLCM features and classifiers, AIP Conference Proceedings, Vol. 2725, No. 1, 2023. https://doi.org/10.1063/5.0125260

[13] Sanjayprabu Sivakumar, Sathish Kumar Rangasamy, Saeid Jafari, Karthikamani Ramamoorthy, On Performance Analysis of Diabetic Retinopathy Classification, Electric Letters on Computer Vision and Image Analysis, 22 (2), pp 12-25, 2023. https://doi.org/10.5565/rev/elcvia.1677

[14] Siemers, F.M., Bajorath, J., Differences in learning characteristics between support vector machine and random forest models for compound classification revealed by Shapley value analysis. Sci Rep 13, 5983 (2023). https://doi.org/10.1038/s41598-023-33215-x

[15] Yi Lu, Changbao Yang, Influence of GLCM texture parameters on lithological mapping using sentinel-1 imagery, Taylor and Francis, Geocarto International, Vol. 39, Issue no. 1, https://doi.org/10.1080/10106049.2024.2425183.