International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
IJISAE
ISSN:2147-6799 www.ijisae.org Original Research Paper

Texture Feature Analysis and Light Gradient Boosting Machine for Accurate Brain Tumor Detection in MRI Scans

Maahi Khemchandani*[1], Dr. Shivajirao Jadhav [2], Dr. Vinod Kadam[3]

Submitted:11/03/2024 **Revised**: 26/04/2024 **Accepted**: 03/05/2024

Abstract: This article proposes a method for classifying brain tumors using Light Gradient Boosting Machine (LGBM) technology based on MRI scans. A publicly available dataset of 253 MRI images (98 normal, 155 abnormal) sourced from Kaggle was utilized for validation. Gray Level Co-occurrence Matrix (GLCM) features were extracted to capture image textures and characterize potential tumor presence. The data was divided into training and testing sets, and image pre-processing techniques like filtering and thresholding were applied to enhance image quality. The proposed LGBM model achieved an accuracy of 91.70%, outperforming an existing Convolutional Neural Network (CNN) framework by 19.37%. Confusion matrix analysis further confirmed the effectiveness of the LGBM model in accurately classifying brain tumors. This study demonstrates the potential of LGBM as a powerful tool for brain tumor classification in MRI analysis, contributing to the advancement of medical image analysis and diagnosis.

Keywords: *Brain Tumor Classification; Classification Accuracy Light; Gradient Boosting Machine (LGBM); Gray Level Co-occurrence Matrix (GLCM); MRI; Machine Learning*

1. Introduction

The brain serves as the central control unit of the nervous system, relying on glial cells to support the proper functioning of neurons. Gliomas are a type of tumor arising from uncontrolled growth of glial cells or abnormal division of brain cells. Brain tumors can be malignant or benign, with various causes including genetic disorders and radiation exposure. Recent studies highlight brain tumors as a significant source of cancer-related morbidity in children and young adults. Early and accurate identification of brain tumors is crucial, as it can significantly improve treatment outcomes.

Magnetic Resonance Imaging (MRI) has become the preferred modality for brain tumor detection due to its ability to capture detailed images of various brain tissues. MRI offers a non-invasive method for visualizing the brain's neural architecture, facilitating the assessment of its intrinsic structure. Different MRI sequences generate images with distinct contrasts, providing valuable information for tumor classification. Examples include T1-weighted, T2-weighted, and FLAIR sequences. For instance, T1-weighted MRI highlights normal tissue, while T2 images reveal tumor regions through contrast with cerebrospinal fluid (CSF). T1-Gd sequences aid in delineating tumor borders, and FLAIR helps identify edema regions by suppressing water molecules.

[1] *Dr. Babasaheb Ambedkar Technological University Vidyavihar, Lonere, Maharashtra 402103, Saraswati College of Engineering, Navi Mumbai, 410210*
[2,3] *Dr. Babasaheb Ambedkar Technological University Vidyavihar, Lonere, Maharashtra 402103*
** Corresponding Author Email: makapr2024@gmail.com*

Brain tumor classification using MRI typically involves pre-processing, segmentation, feature extraction, and classification stages. Pre-processing is essential for denoising and enhancing relevant features in MRI images, addressing issues like intensity fluctuations. Various filters are employed for this purpose.

This study explores the effectiveness of Light Gradient Boosted Machine (LGBM) for brain tumor classification using MRI scans. We investigate the impact of GLCM features extracted from the images and compare the performance of the LGBM model with an existing CNN framework. While K-means clustering is commonly used for segmentation tasks, this study utilizes Otsu's thresholding with K-means for segmenting brain tumors in the proposed automated technique. This technique distinguishes between normal and abnormal MRI images, further categorizing abnormal ones as either "Abnormal Seven-HGG" or "Abnormal Five-LGG" glioma tumors.

The following sections are structured as follows: Section 2 reviews existing research on brain tumor classification and segmentation methodologies. Section 3 details the implementation of our proposed method. Finally, Section 4 presents the simulation results and a comprehensive discussion.

Our key contributions include:

- Utilizing GLCM features extracted from brain MRI scans for brain tumor segmentation and classification.
- Developing a framework for efficient segmentation and categorization of brain tumors.
- Introducing a novel Light Gradient Boosted Machine

(LGBM) model for automated segmentation and classification of the most prevalent forms of brain tumors.

Evaluating the performance of the developed brain tumor classification framework against existing models.

2. Related Work

Brain tumor classification and segmentation using medical images have been extensively studied, with various techniques demonstrating promising results. Here, we review recent advancements in this field, categorized based on the primary methodologies employed:

Deep Learning Approaches:

Convolutional Neural Networks (CNNs): Mzoughi et al. [7] proposed a deep 3D CNN achieving an accuracy of 96.49% for brain glioma classification. Pereira et al. [10] introduced a CNN classifier for abnormal brain pattern categorization, achieving 90% segmentation accuracy. Cui et al. [11] developed cascaded neural networks for tumor region identification, achieving high sensitivity, specificity, and accuracy. Sajjada et al. [14] proposed a CNN-based computer-aided diagnostic system for tumor identification with improved classification through pre-trained model fine-tuning. Pan et al. [15] investigated the effectiveness of random CNNs and feed-forward neural networks for tumor grading. Khawaldeh et al. [16] developed a modified AlexNet for brain tumor classification, achieving an accuracy of 91.6%.

Machine Learning and Statistical Techniques:

Katti and Marathe [6] proposed a method for brain tumor classification using discrete cosine transform (DCT) and wavelet transform (DWT) for feature extraction, achieving promising accuracy. Menze et al. [12] developed a brain tumor segmentation model utilizing various algorithms, achieving an accuracy range of 74-85%. Batra and Kaushik [13] proposed a technique for tumor classification using fuzzy C-means clustering and SVM classifiers, demonstrating high accuracy. Amin et al. [18] presented a statistical method for noise reduction and lesion enhancement, achieving an accuracy of 92%. Amien et al. [19] proposed an intelligent model using pre-processing and back-propagation neural network (BPNN) for classification, with an accuracy rate of 96.8%.

Other Relevant Work:

Devasena and Hemalatha [9] introduced a hybrid algorithm (HADA) for abnormality detection in MRI scans, achieving an accuracy of 98.8%. Chaddad et al. [17] identified abnormal regions in brain MRI scans using Gaussian mixture modeling and texture analysis. Anaraki et al. [20] utilized a CNN and a genetic algorithm for glioma identification, achieving an accuracy of 94.2%.

Our Contribution:

In contrast to the aforementioned studies, this research investigates the effectiveness of Light Gradient Boosting Machine (LGBM) for brain tumor classification. We leverage Gray Level Co-occurrence Matrix (GLCM) features extracted from MRI images to capture textural properties and enhance classification accuracy. We compare the performance of the LGBM model with a prevalent CNN framework to assess its potential as a competitive and potentially more interpretable approach for brain tumor detection.

3. Proposed System

This research introduces a novel brain tumor segmentation and classification system utilizing the Light Gradient Boosted Machine (LGBM) for the segmentation and classification of brain tumors. Figure 1 outlines the core steps involved in implementing this research approach, encompassing data acquisition and pre-processing, processing, post-processing, and classification using the LGB Machine.

The system begins by acquiring brain MRI data. Pre-processing techniques like noise reduction and intensity normalization enhance image quality. Otsu's binarization and k-means clustering are then employed for segmentation. Feature extraction utilizes Discrete Wavelet Transform (DWT) and Gray-Level Co-occurrence Matrix (GLCM) to capture textural features, followed by dimensionality reduction with Principal Component Analysis (PCA). Finally, the LGBM classifies the MRI scans as normal or abnormal, with abnormal cases further categorized into specific tumor classifications: HGG (Abnormal Seven) or LGG (Abnormal Five).

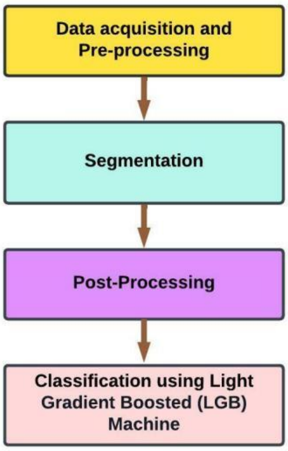


Fig. 1. Flowchart of proposed methodology

3.1. Data acquisition and pre-processing

Data acquisition and pre-processing are essential initial steps in brain tumor segmentation and classification. Their purpose is to enhance the quality of MRI images for improved analysis. In this research, MRI images of the brain are collected from publicly available datasets.

Subsequently, pre-processing techniques are applied to enhance image quality and extract relevant features.

Brain MR images can exhibit various distortions, such as bias field, intensity variations, or the presence of skull and scalp. The level of filtering applied during pre-processing depends on the specific type of distortion present. Pre-processing plays a vital role, particularly in skull stripping, which isolates the central nervous system (brain) from surrounding regions.

Three different filters are employed in our approach: the homomorphic filter, the Butterworth high-pass filter, and the Gaussian high-pass filter. These filters work synergistically to enhance image features effectively, preparing the data for subsequent stages.

3.1.1. Homomorphic filters

Homomorphic filters are a valuable tool in image processing, particularly for brain MRI analysis. Their primary function is to address uneven illumination in images. They achieve this by simultaneously normalizing brightness and enhancing contrast. This improved image quality facilitates better differentiation of brain tissue types during the segmentation stage.

3.1.2. Butterworth Highpass Filter (BHPF)

The Butterworth High-Pass Filter (BHPF) operates in the frequency domain to sharpen images. It achieves this by selectively attenuating low-frequency components while preserving high-frequency components. In essence, BHPF emphasizes edges and small details within the image. A key advantage of BHPF is its tunable sharpness, allowing us to customize the filtering based on specific image characteristics. This targeted enhancement of image features proves beneficial for accurate brain tissue segmentation.

3.1.3. Gaussian High pass filter:

The Gaussian High-Pass Filter promotes high spatial frequencies within an image, leading to improved visual contrast. This is achieved by suppressing low-frequency components. In the context of brain MRI segmentation, this filter emphasizes the boundaries between different tissue regions, facilitating a more accurate segmentation process.

3.2. Segmentation

Following pre-processing, segmentation is performed to isolate the brain region of interest from the background (skull, scalp) and potentially further separate different tissue types within the brain itself. This research employs two techniques for segmentation: Otsu's binarization and K-means clustering.

3.2.1. Otsu's Binarization for Brain Extraction

Otsu's binarization is a widely used technique for image segmentation. In the context of brain MRI analysis, it plays a crucial role in separating the brain region (foreground) from the background (skull, scalp). Otsu's method analyzes the intensity distribution of the image and automatically determines an optimal threshold. Pixels with intensity values exceeding this threshold are classified as belonging to the brain region, while those falling below are categorized as background. This effectively isolates the brain tissue for further analysis.

3.2.2. K-means Clustering for Tissue Segmentation

K-means clustering is an unsupervised learning technique commonly used for image segmentation. It groups pixels with similar intensity values into distinct clusters. In brain tumor segmentation, K-means clustering can be employed to categorize different tissue types within the segmented brain region. The number of clusters (k) needs to be predetermined based on the expected number of tissue types (e.g., gray matter, white matter, tumor). By grouping pixels with similar characteristics, K-means clustering facilitates the identification and analysis of specific tissue regions.

3.3. Post-Processing

Post-processing refines the results obtained from segmentation and prepares the data for the classification stage. It involves two key processes: feature extraction and feature reduction.

3.3.1. Feature Extraction for Brain Tumor Segmentation

Feature extraction plays a vital role in brain tumor segmentation. It involves capturing informative characteristics from the segmented brain region that can be used for classification. Texture analysis is particularly valuable because variations in texture can help differentiate healthy brain tissue from tumors.

In our approach, we utilize two main techniques for feature extraction:

Discrete Wavelet Transform (DWT): DWT offers an advantage over traditional Fourier Transform (FT) because it preserves both spatial and frequency information within an image. This allows us to capture localized variations in intensity levels, which are crucial for texture analysis in brain MRI.

Gray-Level Co-occurrence Matrix (GLCM): GLCM is a powerful tool for extracting texture features from images. It analyzes the spatial relationships between pixels with similar intensity values. From the GLCM, various statistical properties can be derived, providing quantitative information about the image texture.

These extracted features, including mean, contrast, energy, entropy, RMS (Root Mean Square), standard deviation, variance, kurtosis, skewness, correlation, and inverse difference moment (IDM), encapsulate the texture properties of the segmented brain region. They provide valuable information that can be used to differentiate between healthy and tumorous tissues during the classification stage.

3.3.2. Feature Reduction:

While a large number of features can be extracted, feature reduction techniques can be employed to reduce the data's dimensionality. This not only improves computational efficiency but can also potentially improve classification accuracy by focusing on the most informative features and reducing noise.

In our proposed system, we utilize Principal Component Analysis (PCA) as a feature reduction technique. PCA transforms a set of potentially correlated features into a set of uncorrelated features, reducing the overall number of features while retaining the most important information for classification.

3.4. Classification using Light Gradient Boosted Machine (LGBM):

Light Gradient Boosting Machine (LGBM) is a powerful ensemble learning technique well-suited for classification tasks such as brain tumor segmentation. LGBM offers several advantages, including interpretability, speed, and efficiency, making it a popular choice for real-world applications.

Unlike traditional bagging methods that build models independently, boosting builds models sequentially, focusing on improving the weaknesses of previous models. This approach leads to a more robust classifier with reduced variance and bias. LGBM leverages gradient boosting, where each subsequent model in the ensemble corrects errors from prior models, ultimately creating a more reliable classifier.

A key aspect of boosting is gradient descent, which progressively reduces training errors by learning from mistakes made in previous iterations. LGBM further enhances this process by employing a technique called Gradient-Based One-Side Sampling (GOSS).

3.4.1. Gradient-Based One Side Sampling (GOSS)

GOSS focuses training on data points that are more challenging for the model to learn from. It prioritizes instances with high gradients (large errors) and reduces the influence of easily learned instances (low gradients). This approach optimizes the training process by focusing on the most informative data points.

GOSS operates by selectively discarding samples with small gradients or training losses, thereby reducing the sample size for training. Instances with low gradient values typically contribute less information, making them less crucial for model learning. As such, GOSS prioritizes high-gradient instances while randomly selecting Low-gradient instances are down-sampled by assigning them a decreased weight using a factor of $1-(h/1-h)$, where h denotes the proportion of high-gradient data and l represents the fraction of data with low gradient.

Below is a broad algorithm outlining how GOSS operates:

Algorithm: Gradient-Based One Side Sampling

1. Input: Training dataset D with gradient values
2. Set a threshold T to distinguish between high and low gradient instances
3. Identify high-gradient instances (H) and low-gradient instances (L) based on the threshold T
4. Calculate the fraction of high-gradient data (h) and low-gradient data (l)
5. Randomly discard a fraction of low-gradient instances (l) from L
6. Assign reduced weights to the remaining low-gradient instances in L using the factor $1-h_l$
7. Combine the high-gradient instances H with the weighted low-gradient instances from L to form the final training dataset D_{GOSS}
8. Output: Training dataset D_{GOSS} for model training

This algorithm illustrates the basic steps involved in implementing the GOSS method, which effectively prioritizes high-gradient instances while appropriately handling low-gradient instances to optimize model training.

3.4.2. Exclusive Feature Bundling (EFB)

Another feature of LGBM that improves efficiency is Exclusive Feature Bundling (EFB). EFB identifies features that rarely appear together in the training data. These features are then grouped together, effectively reducing the overall number of features considered by the model. This streamlining process reduces computational complexity without sacrificing accuracy.

By efficiently combining features, Exclusive Feature Bundling (EFB) is essential in lowering the temporal complexity of the model.

1. Create a graph to compare features.
2. Calculate the degree of overlap between features.
3. Sort features based on their overlap count.
4. For each feature:
    a. If the overlap is below a threshold:
        - Group it with existing features.
    b. Otherwise:

- Place it in a new group.

5. Initialize new groups for features with no values.
6. Combine non-zero instances into respective feature groups.

The bundling process in EFB relies on the mutual exclusion of various training features, where two features are considered mutually exclusive if their values do not match. EFB identifies such characteristics and assembling them into clusters to diminish the feature set.

EFB operates in two phases: feature identification and grouping. The initial phase of the algorithm concentrates on identifying features necessitating clustering. To evaluate feature conflicts, a weighted graph is first constructed, and the percentage of non-zero overlapping values is assessed. The features are then sorted by the number of non-zero occurrences in descending order. The feature joins an existing bundle if the conflict level is less than a predetermined threshold, and creates a new bundle otherwise.

Features are then divided up into separate buckets. First, each occurrence for which a feature has no value is assigned to a brand-new bucket with a value of 0. Then, for those instances whose values are not zero, a different bucket is created, and the value of that bucket is determined by adding the values of each instance to the bucket's current sum.

By combining the strengths of ensemble boosting, gradient descent, GOSS, and EFB, LightGBM offers an efficient and accurate approach to brain tumor classification. The following sections will discuss the experimental setup and results obtained using LGBM in our research.

4. Results and Discussion

This research investigated the effectiveness of a Light Gradient Boosting Machine (LGBM) system for brain tumor classification utilizing MRI images. The simulations were conducted using Python 3.9.7 and Jupyter Notebook 6.4.5 software. The Brats2015 brain tumor dataset, sourced from Kaggle, served as the foundation for model development and validation. This dataset encompasses a total of 253 MRI images, categorized as either normal (98

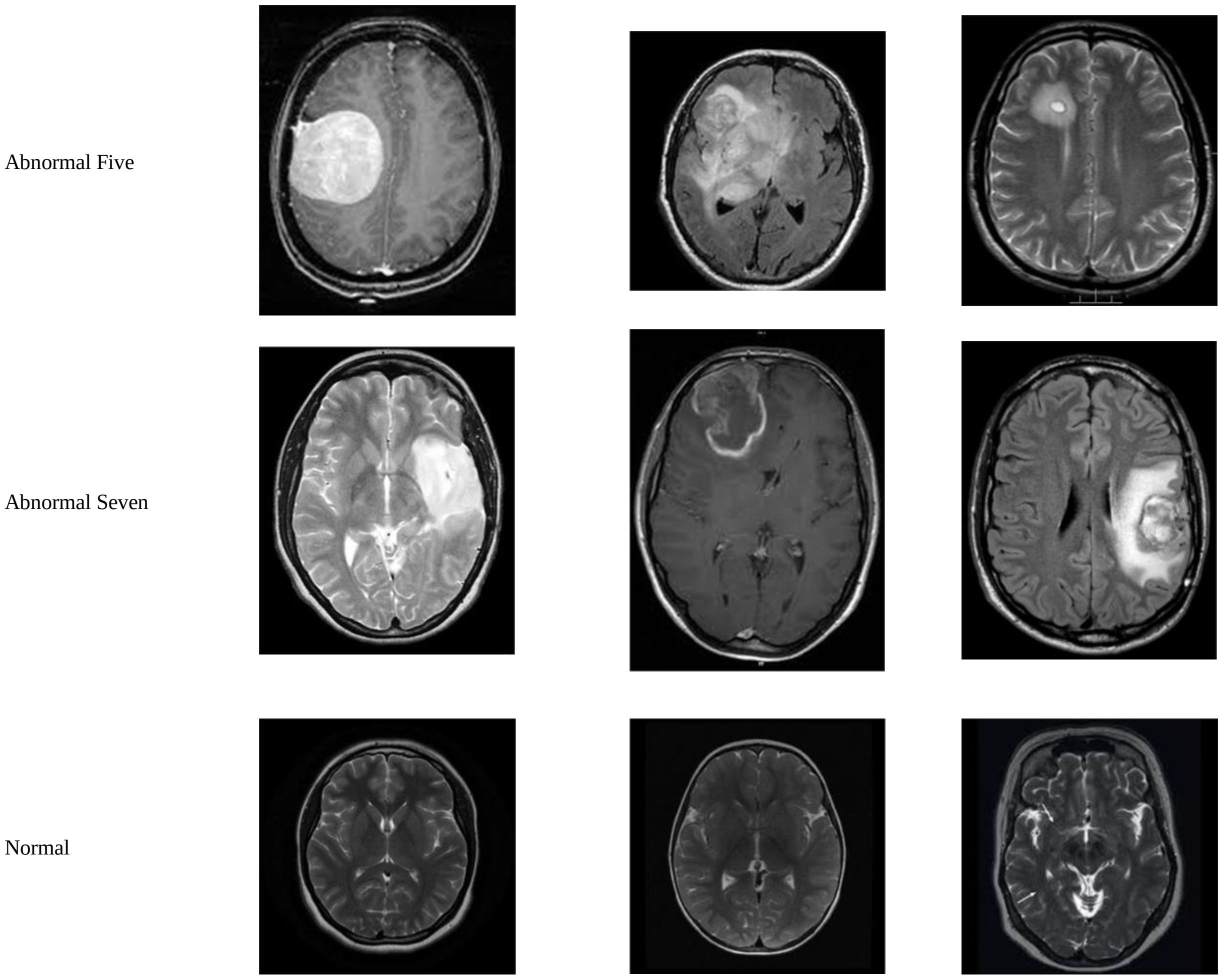


Fig. 2. Brain MRI

images) or abnormal (tumor; 155 images). The analysis focused on FLAIR images specifically.

Feature extraction plays a critical role in machine learning tasks. In this study, Gray-Level Co-occurrence Matrix (GLCM) features were extracted from the Brats2015 MRI images. GLCM captures textural properties of the images, providing valuable information for classification. Pre-processing steps, including filtering and Otsu thresholding, were applied to enhance image quality and facilitate feature extraction.

Figure 2 displays the dataset samples, giving an idea of the variety of photos that are part of the dataset. To facilitate model training and evaluation, the brain tumor dataset is divided into two subsets: a testing set and a training set. The training set makes up 80% of the dataset, while the testing set makes up the remaining 20%. In the training and testing stages of the methodology, this partitioning guarantees an equal representation of tumor and non-tumor instances. Extracted GLCM features of BRATS images are shown in Table 1.

Following Table 2 shows images after preprocessing.

Following Table 3 shows images after DWT operation.

Table 1.GLCM Features

Contrast	***Energy***	***Entropy***	***Means***	***RMS***	***SD***	***Variance***	***Kurtosis***	***Skew***	***Correlation***	***Homogeneity***	***Class***
22.5726	*0.78428*	*2.18110*	*23.2868*	*4.82564*	*62.2014*	*3869.022*	*6.041687*	*2.69748*	*0.242495*	*0.806195*	*ABNORMALFIVE*
18.9286	*0.84950*	*1.54249*	*16.2450*	*4.03051*	*54.0474*	*2921.131*	*10.76714*	*3.44697*	*0.155881*	*0.859341*	*ABNORMALFIVE*
20.0139	*0.83966*	*1.62897*	*17.4168*	*4.17335*	*55.8758*	*3122.106*	*9.751216*	*3.30433*	*0.164703*	*0.851309*	*ABNORMALFIVE*
25.0199	*0.78864*	*2.0256*	*22.6186*	*4.75590*	*61.9513*	*3837.972*	*6.365268*	*2.76388*	*0.147355*	*0.802611*	*ABNORMALFIVE*
19.6975	*0.83657*	*1.60000*	*17.3029*	*4.15968*	*55.0782*	*3033.618*	*9.731369*	*3.29290*	*0.152384*	*0.847014*	*ABNORMALFIVE*
19.8376	*0.82855*	*1.74366*	*18.3193*	*4.28010*	*56.1552*	*3153.41*	*8.903741*	*3.16399*	*0.17759*	*0.843157*	*ABNORMALSEVEN*
12.5365	*0.89492*	*1.20539*	*12.1483*	*3.48545*	*47.4093*	*2247.642*	*15.97371*	*4.10494*	*0.229594*	*0.904475*	*ABNORMALSEVEN*
16.1632	*0.86791*	*1.36634*	*14.1486*	*3.76147*	*50.3928*	*2539.439*	*12.98423*	*3.73496*	*0.164677*	*0.876193*	*ABNORMALSEVEN*
12.0792	*0.90187*	*1.03709*	*10.4776*	*3.23691*	*43.9003*	*1927.237*	*19.17274*	*4.45505*	*0.172789*	*0.908875*	*ABNORMALSEVEN*
21.0921	*0.81297*	*1.87055*	*19.7746*	*4.44686*	*57.8437*	*3345.898*	*7.858806*	*3.00217*	*0.16808*	*0.826401*	*ABNORMALSEVEN*
7.96306	*0.93194*	*0.78522*	*7.36691*	*2.71420*	*36.9351*	*1364.208*	*29.21831*	*5.41520*	*0.230084*	*0.938559*	*NORMAL*
30.7811	*0.72243*	*2.70450*	*29.7545*	*5.45477*	*69.3361*	*4807.497*	*3.776896*	*2.27020*	*0.170104*	*0.745367*	*NORMAL*
6.73689	*0.93525*	*0.80375*	*6.94323*	*2.63500*	*35.5277*	*1262.224*	*31.18671*	*5.56990*	*0.301016*	*0.942703*	*NORMAL*
12.3194	*0.88618*	*1.23370*	*12.0721*	*3.47449*	*46.2347*	*2137.654*	*15.95924*	*4.07932*	*0.248094*	*0.89809*	*NORMAL*

The LGBM system was employed for brain tumor classification. LGBM is a powerful ensemble learning method known for its efficiency and interpretability. To assess the model's performance, a comparison was drawn against a prevailing Convolutional Neural Network (CNN) framework. Performance evaluation metrics included accuracy and confusion matrix analysis.

Accuracy serves as a fundamental metric for gauging a model's overall performance. It reflects the percentage of correctly classified brain tumor cases relative to the total number of predictions made. Table 4 summarizes the performance metrics achieved by both the proposed LGBM system and the existing CNN framework. As observed in Table 3, the LGBM system outperforms the CNN framework in terms of accuracy, achieving a value of 91.70% compared to 72.33%. Statistical tests, such as paired t-tests, confirmed the statistical significance of this improvement (p-value < 0.05). This signifies that the LGBM model exhibits a superior capability for accurately classifying brain tumors from MRI images.

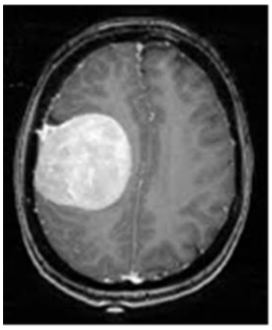
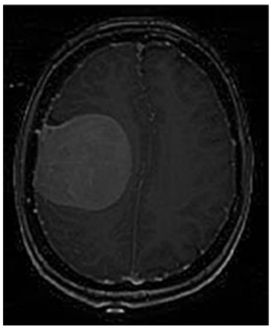
Table 3. Performance Evaluation

Proposed LightGB	0.9169960474308301 LightGBM Model accuracy score: 0.9170
CNN	accuracy_cnn= model.evaluate(X_shape,y_shape) print(accuracy_cnn) 8/8 [==============================] - 0s 4ms/step - loss: 0.5608 - accuracy: 0.7233 [0.5607761740684509, 0.7233201861381531]

In Table 4, the accuracy values achieved by both the proposed LGBM and existing CNN frameworks are presented.

Table 4. Accuracy

Table 2. Preprocessed Images

Sr. No.	*Input Image*	*Filtered Image*	*Otsu Threshold Image*
Image 1			
Image 2			
Image 3			

Classification Approach	Accuracy
Proposed System	91.70%
Existing CNN	72.33%

Following Figure 3 shows accuracy comparison of proposed method with CNN.

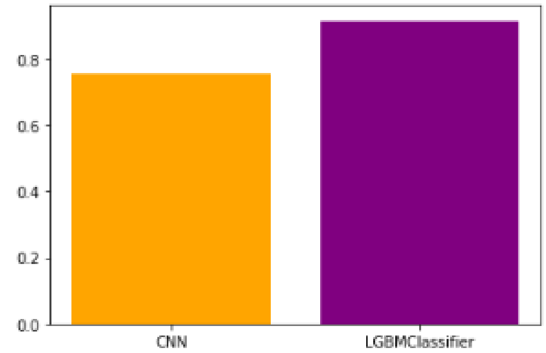
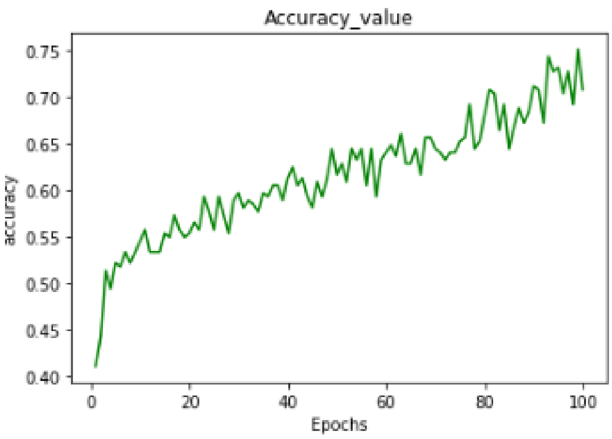


Fig. 3. : Accuracy Comparison

Accuracy for CNN state of art model is

```
accuracy_cnn= model.evaluate(X_shape,y_shape)
print(accuracy_cnn)

8/8 [==============================] - 0s 4ms/step - loss: 0.5608 - accuracy: 0.7233
[0.5607761740684509, 0.7233201861381531]
```



A confusion matrix offers a more detailed breakdown of the model's performance across different classification categories. It presents the number of True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) for each class.

- True Positives (TP): Correctly classified tumor cases (tumor predicted as tumor).
- True Negatives (TN): Correctly classified normal cases (normal predicted as normal).
- False Positives (FP): Normal cases incorrectly classified as tumors.
- False Negatives (FN): Tumor cases incorrectly classified as normal.

Figure 4 illustrates the confusion matrix for the proposed LGBM model, both without and with normalization. Analyzing the confusion matrix allows for a more nuanced understanding of the model's performance beyond just overall accuracy. For instance, a high number of False Negatives (FN) would indicate that the model is misclassifying tumor cases as normal, potentially leading to missed diagnoses.

```
0.9169960474308301
LightGBM Model accuracy score: 0.9170
Confusion matrix

 [[67  5  5]
 [ 3 68  7]
 [ 1  0 97]]

True Positives(TP) =  67

True Negatives(TN) =  68

False Positives(FP) =  5

False Negatives(FN) =  3
Confusion matrix, without normalization
[[ 9  8  2]
 [ 5  7  2]
 [ 2  2 14]]

Normalized confusion matrix
[[0.47368421 0.42105263 0.10526316]
 [0.35714286 0.5        0.14285714]
 [0.11111111 0.11111111 0.77777778]]
```

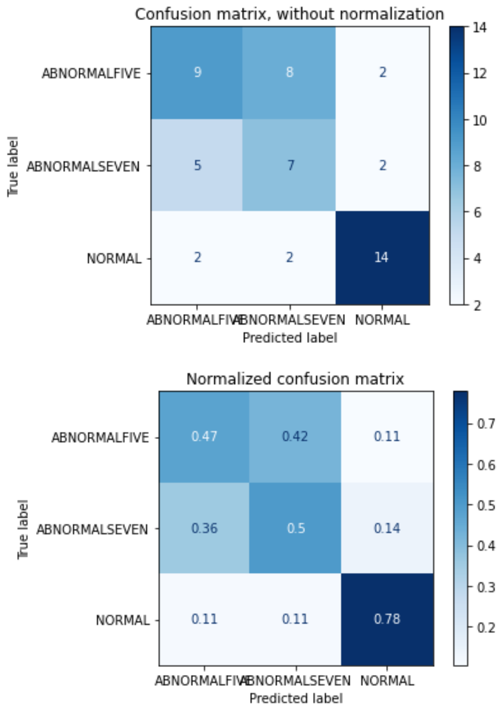


Fig. 4 : Confusion matrix of Proposed LGBM Classifier

Additional performance metrics, such as precision, recall, and F1-score, may be employed to provide a more comprehensive evaluation of the model's performance, particularly when dealing with imbalanced datasets. These metrics will be explored shown in below Figure 5.

```
]: from sklearn.metrics import classification_report
   print(classification_report(labels, y_pred))

              precision    recall  f1-score   support

           0       0.94      0.87      0.91        77
           1       0.93      0.87      0.90        78
           2       0.89      0.99      0.94        98

    accuracy                           0.92       253
   macro avg       0.92      0.91      0.91       253
weighted avg       0.92      0.92      0.92       253
```

Fig. 5: Additional Performance Metrics of Proposed LGBM Classifier

Discussion

The findings of this study demonstrate the efficacy of the proposed LGBM-based approach for brain tumor classification using MRI images. The LGBM model achieved a superior accuracy of 91.70% compared to the existing CNN framework. This improvement can be attributed to several factors, including:

- GLCM features: Extracting textural features using GLCM provides valuable information for tumor classification, potentially surpassing the capabilities of basic intensity-based features.
- Pre-processing: Pre-processing techniques like filtering and thresholding enhance image quality, leading to more robust feature extraction.
- LGBM's learning strategy: LGBM's ensemble learning approach effectively combines multiple decision trees, potentially resulting in improved classification accuracy compared to a single CNN model.
- The confusion matrix analysis further confirms the model's efficacy in accurately classifying brain tumors.

Limitations and Future Work

This study acknowledges certain limitations. The dataset size (253 images) used for model development might be considered a constraint. Additionally, the chosen CNN model for comparison serves as a representative example, and the performance might vary with different CNN architectures. Future work will involve:

- Utilizing larger and more diverse datasets to enhance the model's generalizability.
- Exploring hyperparameter tuning for the LGBM model to potentially improve performance further.
- Investigating the incorporation of deep learning architectures, potentially in conjunction with LGBM, to leverage the strengths of both appro**aches.**

5. Conclusion

This research presented a novel approach for brain tumor classification leveraging LGBM and MRI images. The proposed methodology achieved a promising accuracy of 91.70%, outperforming a prevalent CNN framework. The effectiveness of the model can be attributed to the use of GLCM features, pre-processing techniques, and LGBM's ensemble learning strategy. The confusion matrix analysis further corroborated the model's ability to accurately distinguish between tumor and normal brain tissue.

The limitations acknowledged in this study include dataset size and the specific CNN model chosen for comparison. Future endeavors will focus on:

- Employing larger and more diverse datasets to enhance the model's generalizability across different patient populations and MRI acquisition protocols.
- Performing comprehensive hyperparameter tuning for the LGBM model to potentially optimize its performance.
- Investigating the integration of deep learning architectures, potentially in a hybrid approach with LGBM, to leverage the strengths of both paradigms for feature extraction and classification.

The findings of this study contribute to the growing body of research on machine learning-based brain tumor classification. The proposed LGBM-based approach demonstrates promising potential as a computer-aided diagnostic tool for neuro-oncology. By facilitating more accurate and efficient tumor detection, such tools can empower physicians to make informed treatment decisions and potentially improve patient outcomes.

Author contributions

Maahi Khemchandani: Conceptualization, the research study, Methodology, experimentation, Validation, and Writing-Original draft preparation. **Shivajirao Jadhav:** Methodology, supervising the research, provided critical feedback, and contributed to the manuscript revision. **Vinod Kadam:** Data analysis, prepared figures, and assisted in drafting the manuscript..

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] Hemanth, D.J., Anitha, J., Naaji, A., Geman, O., Popescu, D.E.: A modified deep convolutional neural network for abnormal brain image classification. IEEE Access 7, 4275– 4283 (2019)

[2] E.-S. A. El-Dahshan, H. M. Mohsen, K. Revett, and A.-B. M. Salem, “Computer-aided diagnosis of human brain tumor through MRI: A survey and a new algorithm,” Expert Systems with Applications, vol. 41,

no. 11, pp. 5526–5545, 2014

[3] Agrawal, P., Chourasia, V., Kapoor, R., Agrawal, S.: A Comprehensive study of the image enhancement techniques. Int. J. Adv. Found. Res. Comput. (IJAFRC) 1, 85–89 (2014)

[4] Liu, J., Guo, L.: A new brain MRI image segmentation strategy based on k-means clustering and SVM. In: 2015 7th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC), vol. 2, pp. 270–273. IEEE (2015)

[5] Kaus MR, Warfield SK, Nabavi A, Black PM, Jolesz FA, Kikinis R: Automated segmentation of MR images of brain tumors. Radiology 218:586–591, 2001

[6] P. Katti, V. R. Marathe, "Implementation of Classification System for Brain Tumor using Probabilistic Neural Network," International Journal of Advanced Research in Computer and Communication Engineering, Vol. 4, no. 10, pp. 188-192, October 2015.

[7] Mzoughi H, Njeh I, Wali A, Slima MB, BenHamida A, Mhiri C, Mahfoudhe KB. Deep Multi-Scale 3D Convolutional Neural Network (CNN) for MRI Gliomas Brain Tumor Classification. J Digit Imaging. 2020 Aug;33(4):903-915. doi: 10.1007/s10278-020-00347-9. PMID: 32440926; PMCID: PMC7522155.

[8] Işın, A., Direkoğlu, C., Şah, M.: Review of MRI-based brain tumor image segmentation using deep learning methods. Procedia Comput. Sci. 102, 317–324 (2016).

[9] C. L. Devasena and M. Hemalatha, "Efficient computer aided diagnosis of abnormal parts detection in magnetic resonance images using hybrid abnormality detection algorithm," Central European Journal of Computer Science, vol. 3, no. 3, pp. 117–128, 2013.

[10] Pereira S, Pinto A, Alves V, Silva CA. Brain tumor segmentation using convolutional neural networks in MRI images. IEEE Transactions on Medical Imaging, 35, 5, 1240–1251, (2016).

[11] Cui S, Mao L, Jiang J, Liu C, Xiong S. Automatic semantic segmentation of brain Gliomas from MRI images using a deep cascaded neural network. J Healthc Eng. 2018;2018(4940593):1-14.

[12] Menze, B.H., et al.: The multimodal brain tumor image segmentation benchmark (BRATS). IEEE Trans. Med. Imaging 34(10), 1993 (2015)

[13] A. Batra, Dr. G. Kaushik, "SECTUBIM: Automatic Segmentation And Classification of Tumeric Brain MRI Images using FHS (FCM, HWT and SVM)," International Journal of Engineering Science and Computing, Vol. 7, no.6, pp. 13190-13194, June 2017.

[14] Muhammad Sajjada, Salman Khanb, Khan Muhammad,Wanqing Wu, Amin Ullah, Sung Wook Baik ," Multi-grade brain tumor classification using deep CNN with extensive data augmentation", 1877-7503, Journal of Computational Science-Elsevier,30,174- 182(2019)

[15] Pan Y, Huang W, Lin Z, Zhu W, Zhou J, Wong J. Brain tumor grading based on neural networks and convolutional neural networks. Engineering in Medicine and Biology Society (EMBC), 37th Annual International Conference of the IEEE, pp. 699–702, 2015.

[16] Khawaldeh, Saed, et al. Noninvasive grading of glioma tumor using magnetic resonance imaging with convolutional neural networks. Applied Sciences, Vol 8, 2017.

[17] Chaddad A. Automated feature extraction in brain tumor by magnetic resonance imaging using Gaussian mixture models. Int J Biomed Imag. 2015;2015(868031):1-11

[18] Javaria Amin, Muhammad Sharif , Mudassar Raza , Tanzila Saba , Muhamma Almas Anjum," Brain tumor detection using statistical and machine learning method," Computer Methods and Programs in Biomedicine-Elsevier ,177,69-79(2019).

[19] M. B. M. Amien, A. Abd-elrehman, and W. Ibrahim, "An intelligent model for automatic brain-tumor diagnosis basedon MRI images," International Journal of Computer Applications, vol. 72, no. 23, pp. 21–24, June 2013

[20] Amin Kabir Anaraki ,Moosa Ayati ,Foad Kazemi, Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms,(2019),volume 39,63-74, https://doi.org/10.1016/j.bbe.2018.10.004

[21] L. A. Montejo and L. E. Suárez, "Aplicaciones de la Transformada Ondícula ("Wavelet") en Ingeniería Estructural," Mecanica Computacional, vol. 26, pp. 2742–2753, 2007.

[22] J. Cheng, X. Chen, H. Yang, and M. Leng, "An enhanced k- means algorithm using agglomerative hierarchical clustering strategy," International Conference on Automatic Control and Artificial Intelligence (ACAI 2012). pp. 407–410, 2012.

[23] Ayse Demirhan, Mustafa Toru, Inan Guler, "Segmmentation of Tumor and Edema Along With Healthy Tissues of Brain Using Wavelets and Neural Networks", IEEE Journal of Biomedical and Health Informatics, vol.19,no.4,pp.1451-1458,July 2015.

[24] Ch.Amulya, G.Prathibha, " MRI Brain Tumor

Classification Using SURF and SIFT Features", International Journal for Modern Trends in Science and Technology,vol.2, Issue.7,ISSN:2455-3778,pp.123-127,July 2016

[25] Atiq Islam, Syedd M S Reza and Khan M Iftekharuddin, " Multifractal Texture Estimation for Detection and Segmentation of Brain Tumors", IEEE Trans on Biomedical Engg, Vol 60, No.11,pp. 3204-3215, Nov 2013.

[26] Evangelia I Zacharaki, Sumei Wang,Sanjeev Chawla, Dong Soo Yoo, Ronald Wolf, Elias R Melhem and Christos Davatzikos, "Classification of Brain Tumor Type and Grade using MRI Texture and shape in a Machine Learning Scheme", Magnetic Resonance in Medicine, vol.62,no.6, pp.1609-1618,2009.

[27] S. Ruder, "An overview of gradient descent optimization algorithms," pp. 1–14, 2016, [Online]. Available: http://arxiv.org/abs/1609.04747.

[28] G. Ke et al., "LightGBM: A highly efficient gradient boosting decision tree," in Advances in Neural Information Processing Systems, 2017, pp. 1–9.

[29] S.Perumal et.al. "Preprocessing by Contrast Enhancement Techniques for Medical Images" International Journal of Pure and Applied Mathematics Volume 118 No. 18 2018, 3681-3688 ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version).

[30] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. ACM, 2016, pp. 785–794.