Vol. 06, No. 01, pp. 74 – 86 (2025)
ISSN: 2708-0757

JOURNAL OF APPLIED SCIENCE AND TECHNOLOGY TRENDS

www.jastt.org

Identification of Brain Tumors Using CNN and ML with Diverse Feature Selection Techniques

Suraj Pal Singh[1], Kamal Upreti[2], Rituraj Jain[3], Rakesh Kumar Arora*[4], Akhilesh Tiwari[5], G V Radhakrishnan[6]

[1]*Computer Science and Engineering, Chandigarh University, Mohali, India, srjsingh21@gmail.com*
[2]*Department of Computer Science, Christ University, Delhi NCR Campus, Ghaziabad, India, kamalupreti1989@gmail.com*
[3]*Department of Information Technology, Marwadi University, Rajkot, Gujarat, India, jainrituraj@yahoo.com*
[4]*Computer Science and Engineering, Dr. Akhilesh Das Gupta Institute of Professional Studies, Delhi, India, dr.rakeshkarora@gmail.com*
[5]*Department of Business and Management, Christ University, Delhi NCR Campus, Ghaziabad, India, akhileshmatiwari1@gmail.com*
[6]*Kalinga School of Management, Kalinga Institute of Industrial Technology, Bhubaneswar, India, vrkris2002@gmail.com*
* *Correspondence: dr.rakeshkarora@gmail.com*

Abstract

Early diagnosis and treatment is very essential in monitoring Brain tumor using MRI images. Convolutional Neural Networks (CNN) and Machine Learning (ML) classifiers have been widely used but there is not much work on how feature selection techniques would affect the performance of the CNN. Secondly, there is a need for investigation concerning small dataset adaptability and ML-CNN comparisons. To improve the classification accuracy, we integrate Univariate, Recursive Feature Elimination (RFE), Recursive Feature Elimination with Cross Validation (RFECV) with CNN in this study. Preprocessing, feature extraction & selection was carried out on the dataset consisting of 253 MRI images and they are classified using CNN and ML models (Logistic Regression, Decision Tree, Random Forest, Naïve Bayes). With the results 96%, CNN with Univariate Feature Selection performed better than ML classifiers, and other selection techniques. The results demonstrate that feature selection is necessary to get the best performance out of CNN models operating on small datasets. Future studies should be based on different deep learning architectures to improve classification and application in other datasets.

Keywords: brain tumor detection, CNN, feature selection, machine learning, medical imaging, MRI classification
Received: May 07th, 2025 / Revised: June 26th, 2025/ Accepted: June 28th, 2025 /Online: June 30th, 2025

I. INTRODUCTION

A brain tumor is defined as the development of a bulge or an unnatural growth of brain cells (in the brain) which many consider to be abnormal. These cells become abnormally accumulated either benign (noncancerous or malignant (cancerous). Additionally, they can be divided according to their location inside the brain, most commonly in the cerebral hemispheres, in the cerebellum or in the brainstem.

The most well-known methods of brain tumor segmentation are predominantly based on machine learning, specifically deep learning. CNNs are used in these methods to analyze medical images and determine tumor boundaries. On a labeled dataset of images in which radiologists have segmented the tumor and CNNs are trained on it, these CNNs can be used to segment new images [1, 2].

The segmentation results are usually provided as a binary mask, where the pixels corresponding to the tumor are marked as 1 and the pixels corresponding to healthy brain tissue are marked as 0. These results can be used to generate 3D models of tumor and to plan treatments of the tumor [3, 4].

Within the field of machine learning, a brain tumor is classified as a certain medical issue which can be analyzed and diagnosed using computer techniques. Models based on machine learning can be trained to identify potential patterns within medical images, such as an MRI scan which might indicate the presence of a brain tumor. Brain tumor diagnosis, segmentation, and prediction along with classification can be

doi: 10.38094/jastt61242

performed with the help of machine learning models. For example, a tumor can be segmented by virtue of being able to train a model to analyse MRI images and identify and delineate tumor boundaries with high accuracy. It allows this segmentation information to be used for surgical planning or determining the tumor's progression.

One way to classify is to train models that will be able to differentiate between different types of brain tumor, say, gliomas, meningiomas and so on, for example. These models learn from a big database of labeled brain tumor images, which help in identifying some characteristics and patterns of different types of tumors such that they can make a perfect diagnosis.

Prediction of treatment response and disease evolution can also be carried out using machine learning. Examples of models include employing historical patient data such as tumor characteristics, clinical information and treatment outcomes to learn to predict the likelihood of tumor progression, recurrence or even response to a certain therapy. For example, these predictions can aid in making informed treatment strategy and patient management decisions by the clinician.

Furthermore, machine learning techniques could be used on brain tumor data to discover meaningful features from which a better picture of how brain tumor gain cancerous properties and understand the mechanisms of disease development can be gleaned. Such knowledge can prove to be very useful in the discovery of new biomarkers that can become novel therapeutic targets or potential treatment approaches.

More generally, in machine learning, brain tumor are medical problems treated computationally. With the application of complex algorithms and models, the field of machine learning has integrated detection, diagnosis, treatment planning, and patient outcome improvements in brain tumor research and clinical practice [3-5].

Several research gaps in detecting brain tumor using ML and DL have yet to be solved. The previous studies had not investigated the role of feature selection in the CNN performance or compared the CNN with the traditional ML classifiers. Moreover, the small dataset adaptability and preprocessing problems have not been well studied.

This study is focused on evaluating feature selection techniques, comparing CNN to ML classifiers and determining which preprocessing techniques to further optimize the performance of a brain tumor classifier. This study makes an advance on existing methodology by providing the following Table I which details the key research questions, identified gaps and the contributions.

TABLE I. RESEARCH QUESTIONS, IDENTIFIED GAPS, AND CONTRIBUTIONS

Research Questions	Identified Research Gaps	Research Contributions
How does feature selection impact the accuracy of CNN-based brain tumor classification?	Limited research on how feature selection affects CNN accuracy in medical imaging.	The study applies Univariate, RFE, RFECV, and Tree-Based Feature Selection to CNN and demonstrates that Univariate Selection achieves the highest accuracy (96%).
How does CNN compare with traditional ML classifiers (Logistic Regression, Decision Tree, Random Forest, Naïve Bayes) in brain tumor detection?	Lack of comparative analysis between CNN and ML classifiers in tumor classification.	The study evaluates multiple classifiers and proves that CNN outperforms traditional ML models in accuracy and robustness.
Can CNN models be adapted to small datasets without overfitting, and how does feature selection contribute to this?	Most studies train CNNs on large datasets, leaving small dataset adaptability unexplored.	The study trains CNN on a small dataset (253 MRI images) and shows that feature selection prevents overfitting and improves performance.
What preprocessing and feature selection techniques improve CNN performance in brain tumor detection?	Prior works do not analyze preprocessing impact (e.g., normalization, augmentation) on CNN classification accuracy.	The study implements grayscale conversion, normalization, augmentation, and compares preprocessing techniques for optimal CNN performance.
Can alternative deep learning architectures (Capsule Networks, Vision Transformers) further improve CNN-based brain tumor classification?	CapsNets and Vision Transformers (ViTs) remain unexplored in tumor classification despite their advantages over CNNs.	The study suggests future research into CapsNets and ViTs to improve robustness, long-range dependencies, and spatial hierarchies in brain tumor classification.

II. RELATED WORK

With the progress of medical imaging analysis, the classification of brain tumor using ML and Deep Learning (DL) techniques has drawn significant attention. Recent advancements in the segmentation, classification, and detection of tumor have utilized CNN. However, challenges such as feature selection, small dataset adaptability, and comparative performance evaluation of CNN vs. ML classifiers remain underexplored. This section critically examines prior research, identifies gaps, and highlights the contribution of this study.

A. Machine Learning and Deep Learning in Brain Tumor Classification

CNN based approaches had been applied in several studies for brain tumor detection and classification. As seen in [5], high accuracy of 91.66% is achieved with MRI based classification of brain tumor through Faster R-CNN, using CNN for tumor identification. The study, however, did not cover other ML classifiers and did not study the feature selection techniques. Like [6], they also used a 2D CNN model for discrimination against tumor types (meningioma, glioma, pituitary tumor), with the accuracy of 91.3%. However effective their study was, they did not investigate the effects by preprocessing and the feature selection on CNN performance.

In [7] also proposed another CNN based tumor segmentation model with a dual branch architecture that extracts both global and local features to increase segmentation accuracy (Mean Dice Score: 0.92 of whole tumor, 0.87 of tumor cores). While

this approach made efficiency and feature selection more relevant through reducing the amount of overhead, it did not consider ML alternatives or feature selection strategies to improve classification performance.

A multitask CNN based classification model for tumor detection, grading, and localization is developed by [8] with accuracy of 92%. While the CNN based holistic tumor analysis was proven to benefit in this study, feature selection was not discussed. Finally, it did not compare CNN to traditional ML classifiers, which is important for analyzing the most powerful classification technique. While these studies show CNN's effectiveness, there is no comparative analysis with ML classifiers and no study on how feature selection affects CNN optimization that this research attempts.

B. Feature Selection and Preprocessing in Brain Tumor Classification

There are feature selection techniques that reduce dimensionality and improve interpretability that are necessary to improve model performance. Contrast enhancement techniques were used by [9] based on which they introduced Regularized Extreme Learning Machine (RELM) for tumor classification. However, they demonstrated that preprocessing has a large impact on classification accuracy but integrated feature selection methods. Similarly, [10] described the use of feature extraction based on Local Binary Patterns (LBP) and nLBP as preprocessing techniques which have an influence on the accuracy of the classification. However, the model performance was not evaluated regarding its different feature selection techniques.

There are several researchers that investigated the hybrid feature selection approaches. In [11], they suggested tumor segmentation by SVM and ANN based on adaptive regularized kernel based ultracened fuzzy Cmeans clustering (ARKFCM). Despite their effectiveness, they did not consider CNN based techniques. In this framework, [12] proposed a hierarchical SVM-CNN framework that achieves a better performance (Dice Score: 0.73~0.81). Nevertheless, this study did not analyse feature selection that can be used in CNN or compare with other ML classifiers. However, the effect of feature selection methods in CNN-based classification have not been investigated and therefore it is a key issue to be answered by this study.

C. Emerging Trends in Deep Learning for Brain Tumor Classification

One significant problem in classifying tumor from the brain is the scarcity of annotated medical datasets that resulted in overfitting and low generalizability of models. In particular, [13] observed that low sample sizes hinder CNN based intraocular tumor classification. However, their suggestions for solving this problem were data augmentation and low shot learning techniques, which they did not explore the use of feature selection techniques to better adapt the CNN to small datasets. [14] presented Adaptive Moving Self Organizing Map (AMSOM) and Fuzzy k means clustering (FKM) model achieved 10% better segmentation than the previous model. However, they didn't use CNN architecture and did not explore feature selection strategies in their work.

Usually, generalization problems often happen while training deep learning models on small datasets. ML models were used by [10] to classify malignant vs. benign neoplasms based on MR images of ADC images with 90.41% accuracy. While effective, the study did not compare to CNN based approaches and does not include feature selection in this ability to adapt to small dataset. To tackle these problems, this work also looks at CNN performance in a small set of data (253 MRI images) in conjunction to feature selection techniques to reduce overfitting and increase generalizability.

To reduce the time in detecting brain tumor CNN, EfficientNetB0, InceptionV3, ResNet50 and NASNETMobile have been used in [27]. They suggested a cost-sensitive model for the imbalance data and the combination of CNN and InceptionV3 gave the accuracy of 92.31%.

In [28] they collected datasets from brat's dataset and applied features were extracted using attribute aware attention (AWA) methodology. The images are segmented using visual geometry group (VGG-19) model. A relevance score of 94.86% was achieved using AWA based VGG-19 model.

III. METHODOLOGY

To detect the occurrence of a tumor in an image, the following steps are consecutively taken:

- The first step is to acquire medical images of the brain, such as MRI or CT scans.
- The next step is to preprocess the images. This includes tasks such as cropping, resizing, and normalizing the images in such a way that they are at the same scale and that the tumor is clearly visible in all images.
- The next step is to build and train a machine learning model on a labelled set of images. The model is trained during the training phase to recognize the characteristics of the brain tumor within the image such as shape, size, intensity, and more.
- After training the model, it is used to identify brain tumor in other images. Analyze these images and then inform the prediction whether there is a tumor and if yes, what is its position.
- The last stage involves assessing the effectiveness of the model. This may include metrics such as accuracy, sensitivity, and specificity.
- Depending on the results of the evaluation, the model can be fine-tuned or retrained if necessary.

Fig. 1 shows the overall procedure for modelling brain tumor using CNNs and machine learning approaches.

A. Dataset

MR images were downloaded from the Kaggle website. All the images were in jpg format, and there was a total of 253 images in which 98 images did not contain tumor and 155 images contained tumor. Fig. 2 shows the sample images that have no brain tumor, and Fig. 3 shows the sample images that have a brain tumor [15, 16].

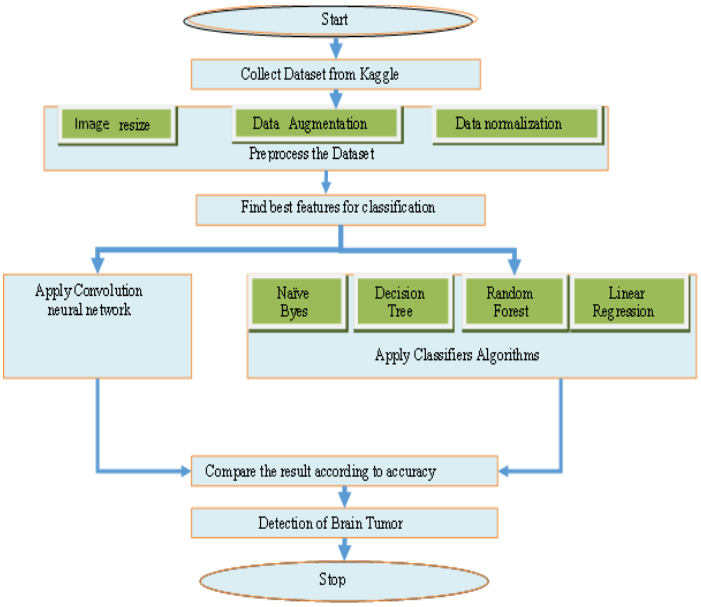

Fig. 1. CNN and machine learning based tumor classification

B. *Preprocessing of the dataset*

The dataset was downloaded from the Kaggle website, and all the images were in the jpg format, few of the images were in RGB format, and few were in grayscale format. First, all the images were converted to grayscale images using the rgb2gray method in MATLAB. Then 49 features are calculated for each image: entropy, kurtosis, mean, contrast, GM9, standard deviation, GM10, CM9, skewness, GM8, CM8, SIM6, energy, GM7, CM7, variance, GM5, GM12, CM5, Haralicks Feature 2, RIM1, GM2, SIM2, CM3, GLCM ASM Haralicks feature 1, RIM 2, CM10, GM11, SIM5, centriod y coordinate, Haralicks Feature 5, GM1, CM1, centriod x coordinate, SIM1, RIM4, Haralicks Feature4, GM4, CM4, SIM7, RIM3, Haralicks Feature 7, GM3, CM2, SIM4, Haralicks Feature 6, GM6, CM6, and SIM3. Each feature provides information about images, as a greater number of features increases the accuracy of detecting brain tumor. These features were saved in .csv format.

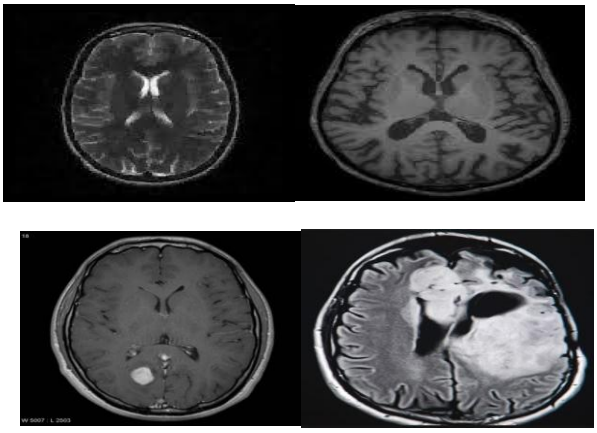

Fig. 3. Images of patients with tumor [16]

There are 49 features, and these features are calculated for each image, which is for 213 images. The statistical features that are calculated are [17]:

1. Mean: The mean is the average value of the intensity of the image.

$$\text{Mean: } \mu = \sum_{i=0}^{G-1} iP \quad (1)$$

2. Contrast: In the context of visual perception, contrast refers to the discernible distinction in luminance or color between an object and its surroundings, which allows the object to stand out. This distinction is influenced by the variation in the brightness and color of the object and other objects present in the same visual field.

$$Average Contrast: \sigma^2 = \sum_{i=0}^{G-1} (1-\mu)^2 P(i) \quad (2)$$

3. Skewness: Skewness indicates the level of asymmetry present in a real-valued random variable's probability distribution, and as such, it serves as a measure of dispersion. In other words, when skewness is absent in a distribution, it is symmetrical; however, at times, a zero value, which indicates the most frequent outcomes being bordered by the least frequent outcomes on both sides, is not always the case.

$$\text{Skewness: } \mu_3 = \sigma^{-3} \sum_{i=0}^{G-1} (1-\mu)^3 \, P(i) \quad (3)$$

4. Kurtosis: Kurtosis, in statistics, refers to the degree of flatness or peakedness of a probability distribution's histogram for a real-valued random variable. This term is mathematically linked to the fourth moment of the distribution. A distribution with high kurtosis has tails that are more pronounced and spread out, often accompanied by a sharper peak, although this is not always the case. Conversely, a low kurtosis distribution exhibits shorter, thinner tails and a more rounded peak, although exceptions to this pattern can occur.

$$\text{Kurtosis: } \mu_4 = \sigma^{-4} \sum_{i=0}^{G-1} (1-\mu)^4 P(i) - 3 \quad (4)$$

5. Energy: The measure of change in an image that can be quantified is referred to as energy, and energy measures that quantification.

$$\text{Energy: } E = \sum_{i=0}^{G-1} [P(i)]^2 \quad (5)$$

6. Entropy: Entropy is a statistical measure that enables quantification of the degree of randomness in the intensity values of an input image, and is utilized to characterize its texture. Additionally, entropy can be employed to assess the degree of variation in the distribution present in a specific region. Interestingly, the entropy values are inversely related to the uniformity measure, thereby allowing the same conclusions to be drawn from both measures.

$$\text{Entopy: } H = -\sum_{i-0}^{G-1} P(i)\log_2[P(i)] \quad (6)$$

7. Standard deviation: The standard deviation is a fundamental statistical metric that serves to quantify the extent of variation in the intensity of an image relative to its mean intensity. Standard deviation is used in image processing to track how much "dispersion" there is from a predicted or average value. A low standard deviation suggests data points are closely bunched around the mean, whereas a high standard deviation indicates distribution is much wider. Due to its ubiquity in statistics, the standard deviation (SD) represents a widely used measure of diversity or variability.

$$SD = \sqrt{\frac{\sum(x-\bar{x})^2}{n}} \quad (7)$$

8. Variance: Variance is a statistical measure used to determine the extent to which each pixel in an image deviate from neighboring pixels or a central pixel. In classification tasks that involve segmenting images into distinct regions, variance is commonly used. Variance is one of the several descriptors of a probability distribution and as such, it quantifies the extent of variation among a group of numbers. Specifically, it represents one of the moments of a distribution, and is mathematical tool that is useful for distinguishing between different types of probability distributions. However, we can perform this task with many other techniques, but the techniques based on moments are computationally and mathematically simple.

$$VARIANCE = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}(i-\mu)^2 P(i,j) \quad (8)$$

9. Gray level co-occurrence matrix (GLCM): It characterizes the distribution of prior texture of the image's cooccurring pixel values at a specified offset. The size of the GLCM is a square matrix with dimensions being same as the number of gray levels, G, in the image. A very early method in the literature of image processing was to employ statistical image features such as the GLCM. This is a technique that is used for the relationship between the two adjacent pixels. An analysis of the statistical properties of such relationships can then provide useful insights into how the image structure and content are most likely structured and what content it contains [18-21].

GLCM Contrast: It is the measure of gray level variation based on the reference pixel and its neighbors of a gray level co-occurrence matrix (GLCM). High contrast shows that there are large differences of intensity between the pixels in the GLCM.

$$CONTRAST = \sum_{n=0}^{G-1} n^2\left\{\sum_{i=1}^{G}\sum_{j=1}^{G} P(i,j)\right\}, \quad |i-j| = n \quad (9)$$

GLCM correlation: The correlation feature of a gray level co-occurrence matrix (GLCM) is an indicator of the degree of linear dependence between the gray level values in the matrix.

$$\text{CORRELATION} = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\frac{\{i*j\}*P(i,j)-\{\mu_x*\mu_y\}}{\sigma_x*\sigma_y} \quad (10)$$

GLCM Energy: The addition of the squared value to GLCM is a measure that reflects the uniformity of the distribution or, alternatively, the angular second moment.

$$\text{Energy} = \sqrt{\text{ASM}} \quad (11)$$

GLCM Homogeneity: Homogeneity quantifies how comparable the distribution of components within a GLCM is to the diagonal of the matrix. The greater the value of homogeneity, the lesser the contrast within the image.

$$\text{HOMOGENEITY} = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\frac{1}{1+(i-j)^2}P(i,j) \quad (12)$$

GLCM AsmHaralicks Feature 1: The Angular Second Moment (ASM) quantifies the degree of local homogeneity of grey scales in an image, while the sum of squared values in GLCM are calculated. When pixels in an image closely resemble each other, the resultant ASM value will be large.

$$\text{ASM} = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\{P(i,j)\}^2 \quad (13)$$

Haralick Feature 2: Entropy is a measure of randomness or disorder within an image. When all the elements within the co-occurrence matrix are identical, the entropy is maximized. Conversely, when the elements are dissimilar, the entropy is at its minimum

$$ENTROPY = -\sum_{i=0}^{G-1}\sum_{j=0}^{G-1} P(i,j) * log\big(P(i,j)\big) \quad (14)$$

Haralick Feature 3: Difference variance is a currently used measure to analyse heterogeneity and place greater weight on the more disparate intensity level pairs.

$$\text{VARIANCE} = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}(i-\mu)^2 P(i,j) \quad (15)$$

Haralick Feature 4: It is a statistical measure of dissimilarity, as known as difference average, as the mean difference in gray values of neighboring pixels in an image. The higher the dissimilarity value is, the more there is difference in intensities of neighboring pixels.

$$AVER = \sum_{i=0}^{2G-2} iP_{x+y}(i) \quad (16)$$

Haralick Feature 5: Sum Entropy is used as a statistical measure to describe the amount of uncertainty or order in the distribution of the sum of gray levels of an image. It is a measure that incorporates the probability distribution of all possible sums that can occur in the image and the randomness or predictability about those sums. An increase in the sum entropy value signifies that the image is more disordered or uncertain with regard to the distribution of the sum of the gray levels inside the image.

$$\text{SENT} = -\sum_{i=0}^{2G-2} P_{x+y}(i)\log\left(P_{x+y}(i)\right) \quad (17)$$

Haralick Feature 6: Difference entropy is a way to measure image disorder (or randomness) on the basis of the gray level difference distribution in an image.

$$\mathrm{DENT} = -\sum_{i=0}^{G-1} P_{x+y}(i)\log\left(P_{x+y}(i)\right) \quad (18)$$

Haralick Feature 7: Inertia is a statistic-based measure which determines the magnitude of local variations in an image. On the other hand, this measure of contrast is sensitive to quantities from p(i, j) that are in fact farther away from the diagonal of the co-occurrence matrix.

$$\mathrm{INERTIA} = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\{i-j\}^2 * P(i,j) \quad (19)$$

Haralick Feature 8: Cluster shade is a statistical feature to measure the asymmetry of the GLCM matrix, in other words, it is proportional to the perceptual uniformity and is HaralickFeature 8. A higher value of the Cluster Shade represents more asymmetry.

$$\mathrm{SHADE} = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\{i+j-\mu_x-\mu_y\}^3 * P(i,j) \quad (20)$$

Haralick Feature 9: Cluster prominence is a feature for measuring asymmetry in gray level co-occurrence matrix (GLCM). The greater asymmetry (higher value) the peak of the distribution shifts further away from the mean. The opposite is true when a lower value appears, the distribution is close to the mean, and is less skewed.

$$PROM = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\{i+j-\mu_x-\mu_y\}^4 * P(i,j) \quad (21)$$

10. Centre Moment: It is an application of probability theory and statistics where Centre moment is a statistical moment, that measures the properties of a probability distribution for a random variable in accordance with its mean. Specifically, it is the expected value of that particular integer power which measures the extent of deviation between this random variable and its mean [22].

If f(x,y) is the image function and is the piecewise continuous bounded function then the centre moment can be defined as

$$\mu_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\bar{x})^p\,(y-\bar{y})^y\,f(x,y)\,dxdy \quad (22)$$

where x¯=M_10/M_00 and y¯=M_01/M_00 are the centroids of the image f(x,y).

For 2D continuous function f(x,y) the moment of order (p+q) is defined as

$$M_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x^p y^q f(x,y)\,dxdy \quad (23)$$

Where p,q=0,1,2…..

$$\mathrm{CM1}: \mu_{00} = M_{00} \quad (24)$$

$$CM2: \mu_{01} = 0 \quad (25)$$

$$\mathrm{CM3}: \mu_{10} = 0 \quad (26)$$

$$\mathrm{CM4}: \mu_{11} = M_{11} - \bar{x}M_{01} = M_{11} - \bar{y}M_{10} \quad (27)$$

$$\mathrm{CM5}: \mu_{20} = M_{20} - \bar{x}M_{10} \quad (28)$$

$$CM6: \mu_{02} = M_{02} - \bar{y}M_{01} \quad (29)$$

$$CM7: \mu_{21} = M_{21} - 2\bar{x}M_{11} - \bar{y}M_{20} + 2\bar{x}^2M_{01} \quad (30)$$

$$CM8: \mu_{12} = M_{12} - 2\bar{y}M_{11} - \bar{x}M_{02} + 2\bar{y}^2M_{10} \quad (31)$$

$$CM9: \mu_{30} = M_{30} - 3\bar{x}M_{20} + 2\bar{x}^2M_{10} \quad (32)$$

$$CM10: \mu_{03} = M_{03} - 3\bar{y}M_{02} + 2\bar{y}^2M_{01} \quad (33)$$

11. Scale Invariant Moment (SIM): The scale-invariant moment, or SIM, is a technique used in computer vision for recognizing and describing local details in images. Initially, SIM key points for certain objects are collected from a group of reference images and stored in a database. For recognizing an object in a different image, each feature of the new image is compared one by one with the features stored in the database, and candidate matching features are found by using the Euclidean distance to their feature vectors. The scale invariant moments are calculated via normalization [22]. The normalized centre moments are

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}} \quad (34)$$

$$\gamma = \binom{p+q+}{2}/2, \qquad p+q = 2,3\ldots..$$

Based on the normalized central moment 7 SIM are calculated

$$\mathrm{SIM1}: \phi_1 = \eta_{20} + \eta_{02} \quad (35)$$

$$\mathrm{SIM2}: \phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (36)$$

$$\mathrm{SIM3}: \phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \mu_{03})^2 \quad (37)$$

$$\mathrm{SIM4}: \phi_4 = (\eta_{30} - 3\eta_{12})^2 + (\eta_{21} + \mu_{03})^2 \quad (38)$$

$$\mathrm{SIM5}: \phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + 3\eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} - \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{21} - \eta_{12})^2 - (\eta_{21} - \eta_{03})^2] \quad (39)$$

$$\mathrm{SIM6}: \phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} - 3\eta_{12})^2 - (\eta_{21} + \eta_{03})^2 + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \quad (40)$$

$$\mathrm{SIM7}: \phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} - \eta_{03})^2] - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} - \eta_{03})^2] \quad (41)$$

12. Centriod X coordinate: If f(x,y) is an image function then the centroid x coordinate is the center value of the x values in the image function.

$$\bar{x} = \frac{M_{10}}{M_{00}} \quad (42)$$

13. Centriod Y coordinate: If f(x,y) is an image function then the centroid y coordinate is the center value of the y values in the image function.

$$\bar{y} = \frac{M_{01}}{M_{00}} \quad (43)$$

The variables were calculated with feature individual augmentation and normalization. The dataset, with 30% allocated for testing and 70% for training the model, was then separated into test and train datasets.

C. Feature Selection

It is necessary to perform feature selection to minimize the number of irrelevant features and computation time in CNN based tumor classification. This study employs four feature selection techniques for brain tumor classification; Univariate, REF, RFECV, and Tree Based Selection. However, these methods were chosen because they have been proven able to:

- Univariate Selection: Statistical significance tests for fast removal of irrelevant features.
- RFE & RFECV: While not a supervised method, it has previously been featured in another Kaggle tutorial and provides a great approach for selecting feature importance in each dataset.
- Tree-Based Methods: Score the features within decision trees based on their impurity in determining the tree structure.

In the following, each method is discussed in detail.

1) *Univariate feature selection:*

The univariate selection was used here as the best feature selection technique because it is better, as it tests a single feature at a time, selecting the best features by correlation with the target variable. As with any method, a simple one that is nevertheless powerful makes it an appealing technique when dealing with small datasets, for which computational efficiency is critical. Furthermore, this method applies to statistical tests like ANOVA (Analysis of Variance) and Mutual Information (MI) to determine the degree of significance of each feature. Univariate Selection helps in removing the irrelevant pixel and texture-based features from the medical images which do not assist classification performance. The major point of using them is to quickly filter out irrelevant features and then use other more comprehensive selection techniques. Table II shows the features selected by this method.

2) *Recursive feature elimination (RFE):*

The reason for choosing RFE is because it is a wrapper-based method which iteratively selects the least important features to be removed depending on the model's performance. Where Univariate Selection evaluates features individually, RFE takes feature importance in the context of other features. This method identifies the optimal feature subset by training the model multiple times, starting from the complete information about the features, and removing at each step the least significant feature to stop training. Similar to a Boolean function, the main advantage of RFE is aggregation of the highest essential features, which provide an important contribution to increase the accuracy of the model. As CNN models are data intensive, use of RFE decreases dimensionality and maintains the robustness of the model. Table II shows the features selected by this method.

3) *Recursive feature elimination with cross validation (RFECV)*

RFE can be extended to RFECV first to make use of cross validation to automatically pick the optimal number of features. As opposed to choosing a fixed number of features, RFECV considers only the features that determine best validation performance. This method ensures that there would not be overfitting as a common problem in deep learning models trained on a small set of data. The cross-validation aspect makes sure that the features chosen to generalize well to new data and not simply fit too closely to the training data. And since RFECV combines recursive elimination and validation scoring, it lends itself well in such cases of small datasets in order to select an optimal feature subset well for a high accuracy. Table II shows the features selected by this method.

4) *Tree-based feature selection*

The inclusion of tree-based feature selection was because it uses decision tree-based algorithms (such as RF, Extra Trees) to rank the features based on their importance in the classification tasks. Unlike RFE and RFECV which consider ranking features based on a single model used in the pipeline, tree-based methods work on an ensemble, considering ensemble of varying number of decision trees to find the most important features. This approach is robust to nonlinear relationships and good at high dimensional datasets. Since MRI based tumor classification involves complex patterns of pixels and their intensities, tree-based methods are a more interpretable way of ranking feature importance. This method was used with the main purpose of capturing feature interactions and guaranteeing the continuance of the most informative texture-based features. Table II shows the features selected by this method.

D. Statistical Validation of Feature Selection Methods

To statistically validate the performance differences arising from various feature selection techniques, we conducted appropriate statistical analyses for each method. Univariate feature selection relied on p-values and F-scores (Table III) to identify significantly contributing features. RFE and RFECV were evaluated using logistic regression estimators and 10-fold cross-validation, with the mean accuracy and 95% confidence intervals reported in Table IV. For Tree-Based selection, feature importances were computed using Random Forest across 30 repetitions, as shown in Table V. These statistical validations confirm that the selected features significantly contribute to model accuracy and robustness and support the model's generalization capabilities.

TABLE II. CHOSEN FEATURES OBTAINED THROUGH THE PROCESS OF FEATURE SELECTION

Univariate	RFE	RFECV	TREE Based
standard deviation, variance, entropy, CM 8, SIM 3, RIM 1, GM2, Mean, CM3, SIM 1, GM5, RIM 3, Energy, kurtosis, skewness, GM11, CM 7, SIM 4, Haralicks Feature 5, GM1, CM2, SIM 2, Haralicks Feature 2, GM6, CM 9, SIM 5, Haralicks Feature 4, GM4, CM 6, Haralicks Feature 6, GM7, CM 10, RIM 2, contrast, centriod y coordinate, GM12, RIM 4, centriod x coordinate.	CM 6, Haralicks Feature 6, entropy, CM 8, Mean, CM 9, GM5, Haralicks Feature 5, variance, RIM 1, Energy, centriod y coordinate, GM2, Haralicks Feature 2, centriod x coordinate, Haralicks Feature 4, GM7, GM6, GM1, GM4.	Haralicks Feature 2, GM5, centriod y coordinate, Energy, entropy, Mean, Haralicks Feature 4, standard deviation, variance, Haralicks Feature 6, GM1, contrast, GM6, GM2, GM7, GM4, CM 10, centriod x coordinate, CM 9, Haralicks Feature 5, SIM 3.	CM1, standard deviation, RIM1, SIM 5, centriod y coordinate, SIM 7, CM 9, Mean, entropy, CM 5, CM 10, variance, Energy, CM 6, kurtosis, skewness, CM 4, GM11, 'GM9, GM12, GM10, CM3, CM2, CM 7, CM 8, SIM 6, RIM 3, RIM 2, RIM 4, centriod x coordinate

TABLE III TOP-RANKED FEATURES USING UNIVARIATE FEATURE SELECTION BASED ON F-SCORE AND P-VALUE

Feature	F-score	p-value
glcm homogeneity	1009.56144	2.4488E-148
glcm contrast	767.778365	4.2321E-122
CM 4	456.2364807	6.25016E-82
CM 5	456.0139498	6.72703E-82
CM 6	451.3846532	3.11402E-81
GLCM ASM Haralicks feature 1	272.7916659	1.32164E-53
Haralicks Feature 2	271.2077483	2.4267E-53
Haralicks Feature 5	255.5069323	1.04972E-50
Haralicks Feature 4	248.5239964	1.60456E-49
Haralicks Feature 7	246.2186789	3.96244E-49
Haralicks Feature 10	244.8966894	6.65998E-49
GM5	214.9506116	1.0123E-43
GM9	214.8191413	1.06752E-43
GM7	214.3890897	1.27012E-43
GM4	214.3784807	1.27558E-43
GM8	214.37239	1.27873E-43
GM10	214.2609255	1.33766E-43
GM6	214.0167405	1.47644E-43
GM3	212.971821	2.25312E-43
GM2	212.7574587	2.45734E-43
Haralicks Feature 6	195.8552549	2.42981E-40
CM 1	188.5366663	4.98216E-39
GM1	188.5366663	4.98216E-39
entropy	140.2225645	3.9557E-30
CM 10	130.0264937	3.39959E-28
centriod y coordinate	129.5809382	4.13443E-28
centriod x coordinate	125.181993	2.86847E-27
CM 8	123.2299242	6.79522E-27
CM 7	116.8625038	1.14645E-25
CM 9	110.9289207	1.62347E-24
glcm energy	96.07712618	1.33206E-21
contrast	63.41543991	5.13718E-15
glcm correlation	44.96370269	3.57879E-11

TABLE IV CROSS-VALIDATION ACCURACY AND CONFIDENCE INTERVAL FOR RFE AND RFECV FEATURE SELECTION

Metric	Mean Score	95% CI Lower	95% CI Upper
Accuracy	0.872	0.853	0.891

TABLE V FEATURE IMPORTANCE SCORES FROM TREE-BASED FEATURE SELECTION (RANDOM FOREST)

Feature	Mean Importance	Std Deviation
glcm contrast	0.4969	0.0258
kurtosis	0.2728	0.0234
standard deviation	0.2267	0.0253
glcm energy	0.1623	0.024
glcm correlation	0.1583	0.0234
Energy	0.0914	0.0185
variance	0.0236	0.0158
skewness	0.0221	0.0105
Mean	0.0115	0.0097

E. Model Training and Hyperparameter Optimization

Steps such as data cleansing, feature selection, model training, and evaluation were included in the processes before the brain tumor identification model was trained. The necessary dataset images were downloaded from a Kaggle repository. First, the images were converted to grayscale, then 49 statistical and textural features were extracted. To improve classification accuracy, the most discriminative features were found using various feature selection methods like REF and Univariate Feature Selection along with Tree Based Feature Selection (TFBS) as well as RFECV. To allow the model to be evaluated fairly, the data was then split into 70% training data and 30% testing data.

collection of classifiers like Logistic Regression, Decision Tree Classifier, Random Forest Classifier, and Naive Bayes Classifier was built as a part of the model training using machine learning techniques. In this case, the scikit-learn train_test_split method is what allowed me to scale these models. In addition, we trained a CNN separately with the image dataset consisting

of convolutional layers with filters that were followed by fully connected layers which classified if there was a tumor or not. These models were validated with respect to accuracy score, confusion matrix, cross validation, etc.

The performance of the model was hyperparameter tuned to optimize. Different learning rates (0.001, 0.0001 and 0.01) were tested in CNN learning. A test was run to find the proper number of estimators, those of 42 and 100 were employed and then the choice was made based on the value of the results. Then, in CNN model it was trained from 10 to 50 epochs over the batch size of 32 with categorical cross entropy as the loss function. Also, the Decision Tree Classifier depth and splitting criteria had been fine-tuned by a 10-Fold cross validation. ReLU method has been used as an activation function and for pooling 2x2 max pool has been used. Dropouts are 0.25 after convolution layer and 0.5 for dense layer. REF and Univariate Feature Selection methods led to a great performance improvement and significant impact on performance, reducing dimensionality and improving classification accuracy. Model architecture for CNN used in proposed system is shown in Fig. 4.

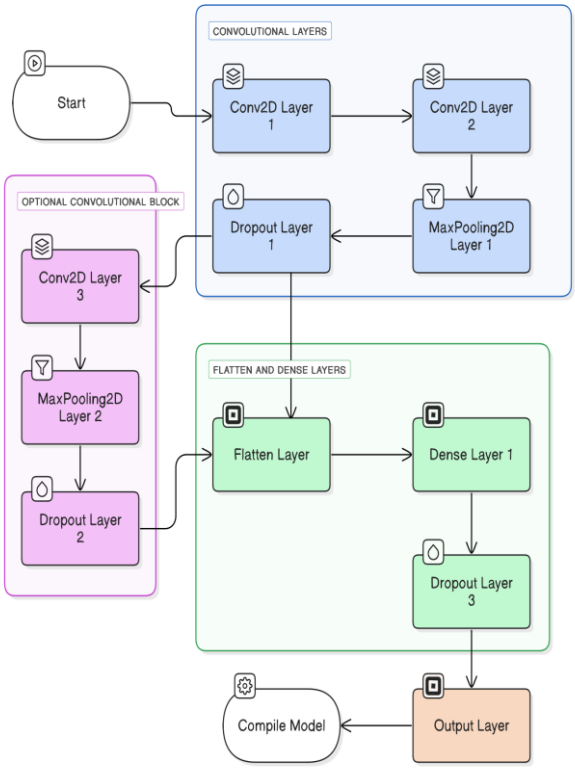

Fig. 4. CNN architecture used in proposed system

IV. RESULT

Once the most important features for brain tumor detection are identified using several methods including univariate, RFE, RFECV, TREE based, the remaining task is to decide upon the best classifier for this task. To do this, we apply a machine learning algorithm and the convolutional neural network on chosen features. We randomly selected our dataset into two subsets, one of which was made up of 70 percent of the data called 'training set' and one of the remainder or 30 percent of the data called 'testing set'. Confusion matrix is used to see its efficiency of the algorithm used. The classification itself was done using open-source Python software and features that need to be selected for the classification purpose were chosen.

Generally, a classifier is assessed in terms of accuracy, precision, recall, specificity, and the F-measure from the output of a confusion matrix. A confusion matrix contains nTP, nTN, nFP, and nFN which stand for True Positive, True Negative, False Positive, and False Negative, respectively. Once nTP, nTN, nFP, and nFN are calculated, many other metrics can be determined.

Accuracy is defined as the share of instances that have been correctly classified out of the total. The precision assesses how many of the predicted positive instances were actually true positives. Recall, also termed sensitivity or true positive rate, assesses how many of the actual positive instances were captured correctly. Specificity assesses how many of the actual negative cases were classified as true negatives. The F-measure is a metric that is calculated as 2 times precision and recall over their summation [23-26].

The accuracies of all the models, i.e., logistic regression, decision tree, random forest, naïve bayes and CNN with respect to the different feature selection algorithms, univariant, RFE, RFECV, tree based, and all the features are shown in Fig. 5. If all the available features (total 49) are provided to the different models, then the accuracy of the naïve Bayes model is low, however the accuracies of the CNN and random forest models are the same. If univariate, RFE, RFECV and the tree-based feature selection method are used with all the models and the CNN, then in all cases CNN yields better results. However, with the univariate feature selection method, the accuracy obtained from the CNN is higher than that of the other methods.

Concerning the four algorithms, logistic regression, decision tree, random forest, and naive Bayes, as well as a convolutional neural network, their accuracy for both random forest and CNN was equal. For univariate, RFE, RFECV, and tree-based feature selection methods, CNN had the best performance. When applying the CNN 10 epochs are used with each type of feature selection method.

Classifier performance in brain tumor detection relies heavily on feature selection, as depicted in Table VI. NN results were maximally accurate (96%) through Univariate selection, underscoring feature relevancy selection. Both RFE and RFECV yielded competitive results (95% with NN), as they reduced redundancies. Decision Tree (81%) and Random Forest (82%) results were both consistent due to receiving Tree-Based selection and all features, respectively. Naïve Bayes was the lowest performer (down to 39%) showcasing its vulnerability to feature sensitivity. In summation, feature selection improves the performance of the model and Tree-Based selection was the only method providing equal improvement to all classifiers. The data supports the fact that CNN is the most accurate for brain tumor detection when integrated with proper feature selection strategies.

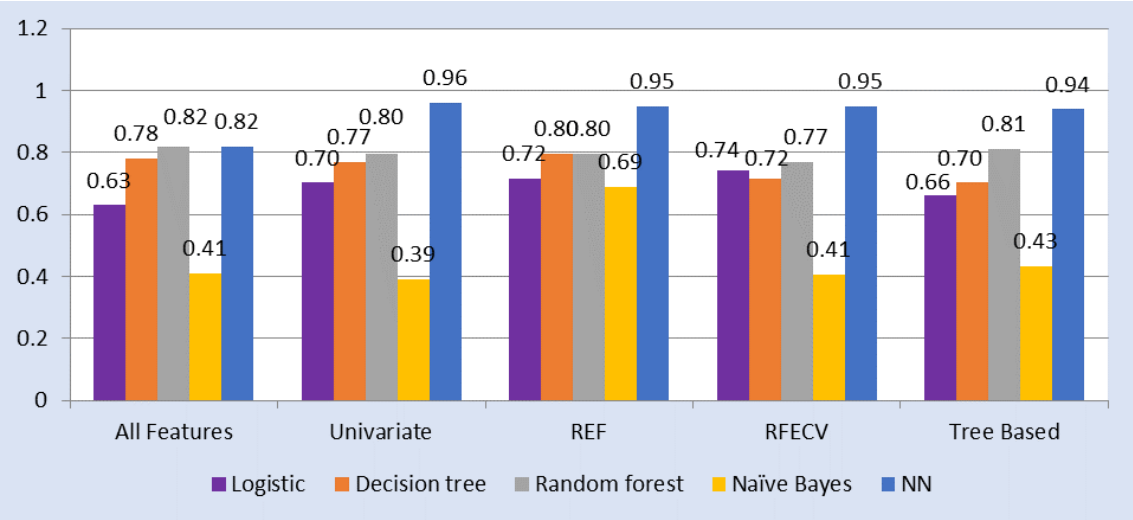

Fig. 5. Accuracy of different algorithm using different feature selection method

TABLE VI CLASSIFIER ACCURACY ACROSS DIFFERENT FEATURE SELECTION METHODS

Feature Selection	Logistic Regression	Decision Tree	Random Forest	Naïve Bayes	Neural Network (NN)
All Features	0.63	0.78	0.82	0.41	0.82
Univariate	0.7	0.8	0.77	0.39	0.96
REF	0.72	0.8	0.8	0.69	0.95
RFECV	0.74	0.77	0.72	0.41	0.95
Tree Based	0.66	0.81	0.7	0.43	0.94

Table VII provides a cross-comparison of the models regarding their training and validation accuracy, convergence epochs, and overfitting tendencies. Model which used Univariate, REF, RFECV, and Tree Based Selection of features obtained the maximum training accuracy of 90%, whereas validation accuracy stayed

steady at around 87 to 88%, showing stable generalization. Model which uses all the features for training showed lesser accuracy (75%) and mild overfitting, suggesting inadequate learning. The rest of the models exhibited overfitting after 8 epochs which indicates effective learning.

The accuracy plots shown in Fig.s 6, 7, 8, 9, and 10 illustrate the model's learning behavior over multiple epochs, showing how training and validation accuracy evolve. Fig. 6 demonstrates overfitting, where validation accuracy increases initially but starts declining after epoch 6, indicating that the model is memorizing training data rather than generalizing. This suggests the need for early stopping or regularization to improve generalization.

In contrast, Fig. 7 and 8 shows better model stability, with validation accuracy aligning closely with training accuracy. The gradual increase in accuracy suggests that the model learned effectively without excessive overfitting.

Fig.s 9 and 10 represent the best-performing models, where accuracy surpasses 90%, and the training and validation curves remain closely aligned. This indicates optimal training with well-selected features and hyperparameters, leading to superior classification performance. These trends confirm that feature selection techniques played a critical role in improving CNN performance. The steady accuracy gains across figures suggest that Univariate Feature Selection contributed to better generalization. The final model's high accuracy (>90%) proves that proper feature selection, training optimization, and regularization significantly enhance CNN's effectiveness for brain tumor classification.

Fig. 11 provides a comparison of the accuracy of our approach with that of previous approaches. [5] used CNN and achieve 91.66% accuracy, [6], used 2-D CNN and achieve 91.3% accuracy, [7] resolve over fitting then used CNN and achieve 92.03% accuracy and [8], used CNN and achieve 92% accuracy however this paper showed that highest accuracy 96% is achieved when univariate feature selection was used with the CNN.

The results obtained from the experiments indicate that feature selection significantly influences CNN performance in brain tumor classification. In this section, we analyze the implications of these findings, compare our results with previous studies, and discuss the role of feature selection in CNN optimization.

TALBE VII EPOCH-WISE MODEL ACCURACY COMPARISON

Model with different feature selection method / Epochs		1	2	3	4	5	6	7	8	9
When all features are used	Training Accuracy	0.5	0.55	0.6	0.63	0.65	0.67	0.69	0.7	0.75
	Validation Accuracy	0.55	0.63	0.65	0.7	0.75	0.75	0.75	0.74	0.72
When univariate is used	Training Accuracy	0.7	0.75	0.78	0.82	0.85	0.86	0.87	0.88	0.9
	Validation Accuracy	0.75	0.75	0.8	0.82	0.85	0.85	0.86	0.87	0.88
When RFE is used	Training Accuracy	0.7	0.75	0.78	0.82	0.85	0.86	0.87	0.88	0.9
	Validation Accuracy	0.75	0.75	0.8	0.82	0.85	0.85	0.86	0.86	0.87
When RFECV is used	Training Accuracy	0.7	0.75	0.78	0.82	0.85	0.86	0.87	0.88	0.9
	Validation Accuracy	0.75	0.75	0.8	0.82	0.85	0.85	0.86	0.87	0.88
When Tree based is used	Training Accuracy	0.7	0.75	0.78	0.82	0.85	0.86	0.87	0.88	0.9
	Validation Accuracy	0.75	0.75	0.8	0.82	0.85	0.85	0.86	0.87	0.88

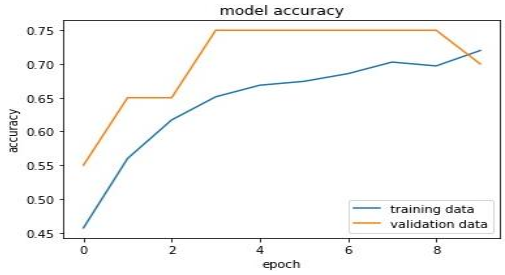

Fig 6. Model accuracy of CNN when all features are used

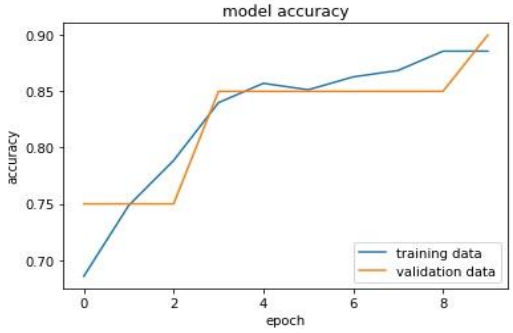

Fig 7. Model accuracy of CNN when univariate is used

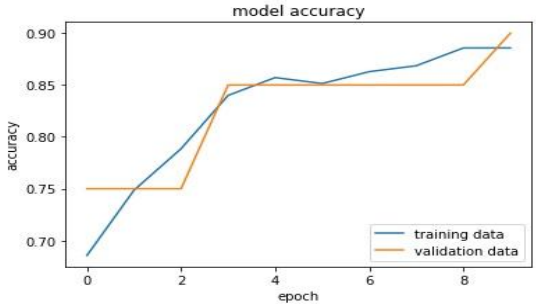

Fig 8 Model accuracy of CNN when RFE is used

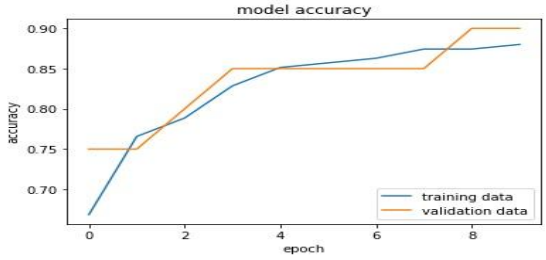

Fig 9. Model accuracy of CNN when RFECV is used

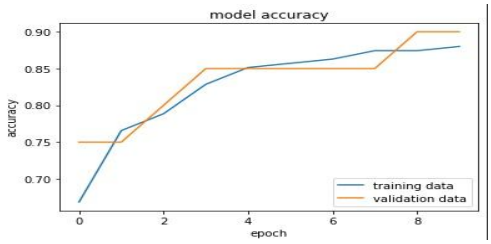

Fig 10. Model accuracy of CNN when Tree based is used

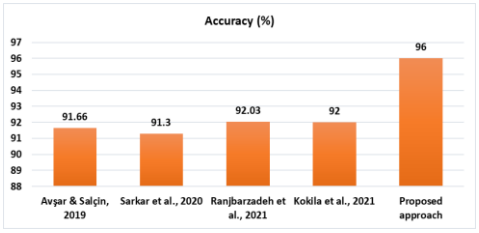

Fig 11. Comparison of accuracy with previous given approaches

V. DISCUSSION

The experimental analysis reveals that feature selection contributes heavily to the performance of CNNs in brain classification. The best result of 96% accuracy was attained from using Univariate Feature Selection with CNN as opposed to RFE, RFECV, and Tree Based Feature Selection. Furthermore, CNN surpassed all conventional classifiers including Logistic Regression, Decision Tree, Random Forest, and Naïve Bayes as was expected, further proving the efficacy of deep learning in detecting brain tumor.

Out of all the methods utilized for feature selection, Univariate Feature Selection with CNN was the most accurate, achieving an impressive 96% accuracy, which surpassed REF, Recursive Feature Elimination with Cross-Validation (RFECV) and Tree-Based Selection. Its effective performance can be attributed to multiple reasons.

One of the primary reasons for its effectiveness is flexibility towards small datasets. The dataset applied in the study was composed of 253 MRI images which, for deep learning approaches, is small. In these situations, reducing the dimensionality of the dataset while retaining the essential information becomes critical to avoid overfitting. Univariate Selection is effective in solving this problem by only selecting features that have the strongest statistical measures, which enhances the chances of improving the generalization capability of the model. On the contrary, wrapper-based approaches such as RFE and RFECV need multiple iterations of model training along with overfitting which is problematic in small datasets.

Another advantage of Univariate Selection is freedom from model dependency. Univariate Selection, unlike RFE and RFECV, does not use feature importance orders from a certain model to apply them to the selection of features, but rather it assesses them independently through their correlation to the dependent variable. This approach prevents the removal of other irrelevant features, ensuring that the CNN model contains sufficient useful information to enhance classification.

Apart from that, it is noted that Univariate Selection is more efficient in terms of computation time than other methods of feature selection. This optimization saves time for systems dependent on CNN training time as it does not require iterative steps like training RFE or RFECV. CNN-based classification is one such use case where time and resources for training a system are very critical.

RFE and RFECV selection feature selection methods achieved accuracy rates of 95, which means they manage to eliminate redundant features but lost useful ones due to the elimination process they are using. Tree-Based Selection performed with 94% accuracy, which shows that though the decision tree selection methods are effective, they do not yield the most important MRI features for CNN to perform.

This suggests that Univariate Feature Selection is much more appropriate for small medical imaging datasets because it preemptively eliminates features that may contribute to overfitting and enables CNN to perform better with out-of-the-box classification. These results indicate that there are situations when simplistic filter techniques such as Univariate Selection have an edge over more complex feature extraction methods, especially with scarce medical imaging data.

VI. CONCLUSION

This study highlights how feature selection influences the classification of brain tumor with CNNs and its difference with other conventional machine learning classifications. The study analyses the impact of integration of Univariate Feature Selection, REF, Recursive Feature Elimination with Cross-Validation (RFECV), and Tree-Based Selection on model performance using a dataset of 253 MRI images. The results indicate that CNN combined with Univariate Feature Selection outperformed other selection heuristics with an accuracy of 96%, significantly higher than that achieved with feature selection and classical ML classifiers which include Logistic Regression, Decision Tree, Random Forest, and Naïve Bayes. Research suggests Univariate Feature Selection is very useful in small datasets as it helps in preventing overfitting by including only the most significant features. When considering wrapper-based methods like RFE and RFECV, univariate selection is simpler and more efficient. This makes it easier to use in medical imaging. The study also confirms that CNN models outperform all traditional machine learning classifiers, demonstrating the ability to capture complex spatial and texture features. This approach achieves better classification accuracy compared with other studies, which further emphasizes the role of feature selection in CNN-based classification problems. The study also highlights the need to apply explainable techniques like Grad-CAM and SHAP to improve trust in AI-based models from a clinical perspective.

REFERENCES

[1] E. S. Biratu, F. Schwenker, Y. M. Ayano, and T. G. Debelee, "A Survey of Brain Tumor Segmentation and Classification Algorithms," J Imaging, vol. 7, no. 9, p. 179, Sep. 2021, doi: 10.3390/jimaging7090179.

[2] M. C. Clark, L. O. Hall, D. B. Goldgof, R. Velthuizen, F. R. Murtagh, and M. S. Silbiger, "Automatic tumor segmentation using knowledge-based techniques," IEEE Trans Med Imaging, vol. 17, no. 2, pp. 187–201, Apr. 1998, doi: 10.1109/42.700731.

[3] 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 Syst Appl, vol. 41, no. 11, pp. 5526–5545, Sep. 2014, doi: 10.1016/j.eswa.2014.01.021.

[4] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," J Big Data, vol. 6, no. 1, p. 60, Dec. 2019, doi: 10.1186/s40537-019-0197-0.

[5] E. Avşar and K. Salçin, "Detection and classification of brain tumours from MRI images using faster R-CNN," Tehnički glasnik, vol. 13, no. 4, pp. 337–342, Dec. 2019, doi: 10.31803/tg-20190712095507.

[6] S. Sarkar, A. Kumar, S. Chakraborty, S. Aich, J.-S. Sim, and H.-C. Kim, "A CNN based Approach for the Detection of Brain Tumor Using MRI Scans," Test Engineering and Management, vol. 83, pp. 16580–16586, 2020.

[7] R. Ranjbarzadeh, A. Bagherian Kasgari, S. Jafarzadeh Ghoushchi, S. Anari, M. Naseri, and M. Bendechache, "Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images," Sci Rep, vol. 11, no. 1, p. 10930, May 2021, doi: 10.1038/s41598-021-90428-8.

[8] B. Kokila, M. S. Devadharshini, A. Anitha, and S. Abisheak Sankar, "Brain Tumor Detection and Classification Using Deep Learning Techniques based on MRI Images," J Phys Conf Ser, vol. 1916, no. 1, p. 012226, May 2021, doi: 10.1088/1742-6596/1916/1/012226.

[9] A. Gumaei, M. M. Hassan, M. R. Hassan, A. Alelaiwi, and G. Fortino, "A Hybrid Feature Extraction Method With Regularized Extreme Learning

Machine for Brain Tumor Classification," IEEE Access, vol. 7, pp. 36266–36273, 2019, doi: 10.1109/ACCESS.2019.2904145.

[10] S. M. Vijithananda et al., "Feature extraction from MRI ADC images for brain tumor classification using machine learning techniques," Biomed Eng Online, vol. 21, no. 1, p. 52, Dec. 2022, doi: 10.1186/s12938-022-01022-6.

[11] T. P, B. Bhat, and K. Prakash, "Detection and Classification of Tumour in Brain MRI," International Journal of Engineering and Manufacturing, vol. 9, no. 1, pp. 11–20, Jan. 2019, doi: 10.5815/ijem.2019.01.02.

[12] M. W. Nadeem et al., "Brain Tumor Analysis Empowered with Deep Learning: A Review, Taxonomy, and Future Challenges," Brain Sci, vol. 10, no. 2, p. 118, Feb. 2020, doi: 10.3390/brainsci10020118.

[13] A. S. Chandrabhatla, T. M. Horgan, C. C. Cotton, N. K. Ambati, and Y. E. Shildkrot, "Clinical Applications of Machine Learning in the Management of Intraocular Cancers: A Narrative Review," Investigative Opthalmology & Visual Science, vol. 64, no. 10, p. 29, Jul. 2023, doi: 10.1167/iovs.64.10.29.

[14] S. Dalal et al., "An Efficient Brain Tumor Segmentation Method Based on Adaptive Moving Self-Organizing Map and Fuzzy K-Mean Clustering," Sensors, vol. 23, no. 18, p. 7816, Sep. 2023, doi: 10.3390/s23187816.

[15] A. A. Akinyelu, F. Zaccagna, J. T. Grist, M. Castelli, and L. Rundo, "Brain Tumor Diagnosis Using Machine Learning, Convolutional Neural Networks, Capsule Neural Networks and Vision Transformers, Applied to MRI: A Survey," J Imaging, vol. 8, no. 8, p. 205, Jul. 2022, doi: 10.3390/jimaging8080205.

[16] Chakrabarty Navoneel, "rain MRI Images for Brain Tumor Detection," www.kaggle.com.

[17] I. Tuncer et al., "Swin-textural: A novel textural features-based image classification model for COVID-19 detection on chest computed tomography," Inform Med Unlocked, vol. 36, p. 101158, 2023, doi: 10.1016/j.imu.2022.101158.

[18] A. K. Aggarwal, "Learning Texture Features from GLCM for Classification of Brain Tumor MRI Images using Random Forest Classifier," WSEAS TRANSACTIONS ON SIGNAL PROCESSING, vol. 18, pp. 60–63, Apr. 2022, doi: 10.37394/232014.2022.18.8.

[19] Lohar Mamata and Chorage Suvarna, "A Robust Autism Brain MRI Classification with GLCM Features and Machine Learning," Frontiers in Health Informatics, 2024, doi: 10.52783/fhi.9.

[20] P. K. Mall, P. K. Singh, and D. Yadav, "GLCM Based Feature Extraction and Medical X-RAY Image Classification using Machine Learning Techniques," in 2019 IEEE Conference on Information and Communication Technology, IEEE, Dec. 2019, pp. 1–6. doi: 10.1109/CICT48419.2019.9066263.

[21] Priyanka and D. Kumar, "Feature Extraction and Selection of kidney Ultrasound Images Using GLCM and PCA," Procedia Comput Sci, vol. 167, pp. 1722–1731, 2020, doi: 10.1016/j.procs.2020.03.382.

[22] Z. L. Chuan et al., "A Comparative of Two-Dimensional Statistical Moment Invariants Features in Formulating an Automated Probabilistic Machine Learning Identification Algorithm for Forensic Application," Malaysian Journal of Fundamental and Applied Sciences, vol. 19, no. 4, pp. 525–538, Aug. 2023, doi: 10.11113/mjfas.v19n4.2917.

[23] M. Heydarian, T. E. Doyle, and R. Samavi, "MLCM: Multi-Label Confusion Matrix," IEEE Access, vol. 10, pp. 19083–19095, 2022, doi: 10.1109/ACCESS.2022.3151048.

[24] J. Miao and W. Zhu, "Precision–recall curve (PRC) classification trees," Evol Intell, vol. 15, no. 3, pp. 1545–1569, Sep. 2022, doi: 10.1007/s12065-021-00565-2.

[25] Z. Sun, B. Jiang, X. Li, J. Li, and K. Xiao, "A Data-Driven Approach for Lithology Identification Based on Parameter-Optimized Ensemble Learning," Energies (Basel), vol. 13, no. 15, p. 3903, Jul. 2020, doi: 10.3390/en13153903.

[26] R. Yacouby and D. Axman, "Probabilistic Extension of Precision, Recall, and F1 Score for More Thorough Evaluation of Classification Models," in Proceedings of the First Workshop on Evaluation and Comparison of NLP Systems, Stroudsburg, PA, USA: Association for Computational Linguistics, 2020, pp. 79–91. doi: 10.18653/v1/2020.eval4nlp-1.9.

[27] M.T.R Shawon,. G.M.S Shibli, F Ahmed, "Explainable cost-sensitive deep neural networks for brain tumor detection from brain MRI images considering data imbalance" in Multimed Tools Appl 2025. https://doi.org/10.1007/s11042-025-20842-x.

[28] A. M. Nancy, R. Maheswari, "Brain tumor segmentation and classification using transfer learning based CNN model with model agnostic concept interpretation" in Multimed Tools Appl,vol. 84, pp. 2509–2538, 2025. https://doi.org/10.1007/s11042-024-20353-1