Journal of Intelligent Systems and Internet of Things
Vol. 15, No. 01, PP. 01-16, 2025
DOI: https://doi.org/10.54216/JISIoT.150101

Brain Tumor Detection: Integrating Machine Learning and Deep Learning for Robust Brain Tumor Classification

Hassan Al Sukhni[1], Qusay Bsoul[2,*] , Fadi yassin Salem Al jawazneh[3], Raghad W. Bsoul[4] , Diaa Salama AbdElminaam[5,6], Magdy Abd-Elghany[7] , Yasmin Alkady[8] , Ibrahim A. Gomaa[7]

[1]Cybersecurity Department, Faculty of Science and Information Technology, Jadara University, Irbid, Jordan
[2]Cybersecurity Department, College of Computer Sciences and Informatics, Amman Arab University, Amman, Jordan
[3]Faculty of Information Technology, Applied Science Private University, Amman, Jordan
[4]Department of Clinical Pharmacy Faculty of Pharmacy, Jordan University of Science and Technology, Irbid, Jordan
[5]MEU Research Unit, Middle East University, Amman, Jordan
[6]Jadara Research Center, Jadara University, Irbid, Jordan
[7]Computer Science Department, Alobour high institute for comHigh Instituteputer and informatics, Cairo, Egypt
[8]Faculty of Computer Sciences, Misr International University, Cairo, Egypt

Emails: h.alsukhni@jadara.edu.jo; q.bsoul@aau.edu.jo; F_aljawazneh@asu.edu.jo; meramidebsoul@yahoo.com; ahmed1908194@miuegypt.edu.eg; diaa.salama@miuegypt.edu.eg; magdy_zeid83@oi.edu.eg; jassy.alkady@gmail.com; Ibraheemg@oi.edu.eg

Abstract

Accurate detection and classification of brain tumors are essential for timely diagnosis and effective treatment planning. This study presents an integrated framework leveraging both machine learning (ML) and deep learning (DL) models for brain tumor detection and classification using MRI images. Two publicly available datasets are utilized: one for binary classification (tumor vs. no tumor) and another for multiclass classification (glioma, meningioma, and pituitary tumors). Comprehensive preprocessing steps, including resizing, feature extraction using the Gray Level Co-occurrence Matrix (GLCM), and feature selection via Chi-square testing, were employed to optimize the dataset for modeling. Machine learning models such as Decision Trees, K-Nearest Neighbors (KNN), Support Vector Machines (SVM), and AdaBoost were compared with deep learning architectures like Convolutional Neural Networks (CNNs) and the pre-trained VGG16 model. Hyperparameter optimization techniques, including grid search and the Adam optimizer, were used to enhance model performance. The models were evaluated using metrics such as accuracy, precision, recall, F1-score, Mean Squared Error (MSE), and Mean Absolute Error (MAE). Results indicate that the VGG16 model consistently outperformed other approaches, achieving high validation accuracy. This study highlights the potential of integrating ML and DL techniques for accurate and efficient brain tumor detection and classification, offering valuable tools for medical diagnostics.

Received: June 20, 2024 Revised: September 15, 2024 Accepted: December 20, 2024

Keywords: Machine Learning; Deep Learning; Brain Tumors

1 Introduction

Brain tumors present a challenge in the field of medicine, affecting millions of lives worldwide. The complexity of the brain anatomy and tumor growth requires precise and efficient diagnostic methods. Traditional approaches to brain tumor detection often face limitations in accuracy and speed, leading to delayed diagnoses and sub-optimal treatment outcomes. The integration of machine learning techniques offers promising avenues for improving both discovering and classifying brain tumors.

The human brain, with its intricate network of neurons and synapses, serves as the regulatory organ for everything from basic physiological functions to complex cognitive processes. However, this organ is not immune to diseases, and one of the most formidable challenges it faces is the development of tumors. Brain tumors arise from abnormal growths of brain tissue cells, and have the potential to be benign or malignant, with varying degrees of aggressiveness and potential for spreading to surrounding tissues. The impact of brain tumors on health can be profound, leading to neurological deficits, cognitive impairments, and even life-threatening complications. Given the critical role of the brain in overall well-being, timely and precise diagnosis of brain tumors is essential for immediate management of the tumor.

Historically, methods of brain tumor detection, such as MRI and computed tomography scans, have been the cornerstone of diagnostic imaging. While these techniques provide valuable structural information, they often rely on subjective interpretation by radiologists and may overlook subtle or early signs of tumor formation. Additionally, the manual segmentation and classification of tumors from imaging data can have a lot of overhead in terms of time and resources among observers. The limitations of traditional approaches highlight the need for more sophisticated and automated methods capable of enhancing diagnostic accuracy and efficiency. Improved detection algorithms could aid in the early identification of tumors, facilitate precise detection of tumor boundaries, and guide treatment decisions. Machine learning(ML), typically used for analyzing complex datasets and extracting patterns, is a promising solution to address these challenges.

This research aims to explore the potential of AI techniques in revolutionizing the field of brain tumor detection and characterization. By using computational algorithms and large-scale imaging data, we seek to develop advanced tools that can assist clinicians in accurately diagnosing and managing brain tumors as machine learning approaches offer several advantages in addressing the challenges inherent in brain tumor detection. One key advantage is the ability to handle the randomness in tumor characteristics, including width, height, shape, and position in the brain. Traditional algorithms may struggle to adapt to these variations, whereas machine learning models can learn from diverse examples and generalize across different scenarios to adapt to new and unseen cases, making them more robust and reliable in real-world clinical settings. Furthermore, machine learning algorithms are expected to enhance tumor segmentation, which is crucial for delineating tumor boundaries and assessing tumor growth over time. By automatically identifying and outlining regions of interest within imaging data, machine learning models can assist radiologists in making more informed diagnoses and treatment plans.

Here are the contributions we have made:

- The implementation involves thorough data collection from Masoud Nickparvar's brain tumor Dataset, and Navoneel Chakrabarty's Brain MRI images Dataset both on Kaggle, providing a rich foundation of brain tumors.
- A comprehensive preprocessing phase ensures dataset refinement through exploration, Resizing, feature selection, and optimizing datasets for predictive modeling.
- The integration of hyperparameter optimization methods for both standard ML techniques and DL models enhances model performance and convergence, with grid search and Adam optimizer optimization.
- The implementation encompasses various classification algorithms to enhance tumor detection and characterization. These include K-Nearest Neighbors (KNN), AdaBoost, Decision Trees, Support Vector Machines (SVM), CNNs and VGG16.
- The contribution extends to prediction and evaluation, utilizing metrics like Accuracy, Recall, Precision, F1-Score, Mean Absolute Error, and Mean Squared Error, providing quantitative insights into the strengths and weaknesses of each algorithm.

The remainder of this paper is organized as follows: In Section 2, a comprehensive overview of related work is presented. Section 3 introduces the proposed methodology and framework for Brain Tumor classification and detection, delineating the significant steps involved in its implementation. Subsequently, Section 3.1 focuses on the critical phase of data collection. In Section 3.2, the paper delves into the intricacies of data preprocessing. Following this, Section 3.3 provides insights into the data-splitting process, showcasing tailored strategies for different algorithms. Moving forward, Section 3.4 explores the classification models based on both ML and DL algorithms. In the subsequent section, In Section 4, the experimental results and discussions unfold, presenting a detailed analysis of the performance of the models. Finally, Section 5 offers a conclusion and outlines avenues for future work.

2 Related work

Not many people know about brain tumors, but some people have studied them and found out a lot of information. We will talk about some of the papers we read to help us with our research, and we will list all of them in the references at the end.

Gokalp Cinarer and Bulent Gursel Emiroglu [1] look into classifying brain tumors using machine learning on MRI images. They compare four algorithms: KNN, RF, SVM, and LDA, with SVM getting the highest accuracy at 90%. The study uses the Rembrandt dataset, analyzing 30 features from each MR image in a dataset of 99 patients. The paper highlights how machine learning could make brain tumor diagnosis more accurate and efficient, and suggests that using a larger dataset and more features could improve the results.

Dena Nadir George, Hashem B. Jehlol, and Anwer Subhi Abdulhussein Oleiwi [2] show a method for finding brain tumors in MRI images using shape features and machine learning classifiers. They use sigma filtering, adaptive thresholding, and region detection to preprocess the images, then extract shape features like Major Axis Length, Euler Number, Minor Axis Length, Solidity, Area, and Circularity. Two classifiers, C4.5 decision tree and Multi-Layer Perceptron (MLP), classify the cases as normal or abnormal, with abnormal cases further classified into benign or five types of malignant tumors. The MLP algorithm was achieved on 174 brain MRI images with about 95% precision.

Komal Sharma, Shruti Gujral, and Shruti Gujral [3] propose an automated system for detecting brain tumors in MRI images using machine learning. Their method involves preprocessing, feature extraction using the Gray Level Co-occurrence Matrix (GLCM), and classification with algorithms like Multi-Layer Perceptron (MLP) and Naive Bayes. They achieved a classification rate of 98.6% with MLP and 97.6% with Naive Bayes, showing high accuracy in distinguishing between normal and abnormal brain MR images.

F. P. Polly, S. K. Shil, M. A. Hossain, A. Ayman, and Y. M. Jang [4] present a computerized system for detecting and classifying High-Grade Gliomas (HGG) and Low-Grade Gliomas (LGG) in brain tumors using MRI images. The system uses k-means clustering for segmentation, Discrete Wavelet Transform (DWT) and Principal Component Analysis (PCA) for feature extraction and reduction, and SVM for classification. It achieved 99% accuracy, 100% sensitivity, and 98.03% specificity in tumor detection and classification. The study highlights the importance of early detection and computerized systems in improving diagnostic accuracy and speed.

Gunasekaran Manogaran, P. Mohamed Shakeel, Azza S. Hassanein, Priyan Malarvizhi Kumar, and Gokulnath Chandra Babu [5] present an automated machine learning approach using orthogonal gamma distribution for brain tumor detection in MRI images. Their method achieves 99.55% accuracy in identifying brain tumors. They use techniques like image segmentation with partial derivatives and orthogonal gamma distribution, edge coordinate matching, and machine learning for feature analysis and classification. The study uses a benchmark medical image database for validation and emphasizes early tumor detection for treatment planning.

Ankit Vidyarthi and Namita Mittal [6] offer a survey of machine learning segmentation methods for brain MR imaging for tumor detection over a decade. They discuss challenges like low illumination in MR images and the role of CAD systems in improving image analysis. The paper categorizes segmentation methods into bi-clustered and multi-clustered algorithms, discussing their pros and cons and performance measures like DB Index and Dunn Index. They cover edge detection, region growing, fuzzy logic, k-means, and fuzzy c-means, with a focus on solutions like CLAP, ACLIME, and hierarchical clustering using probabilistic mutual information (PMI). The hierarchical approach with PMI is highlighted as the most effective, achieving the lowest DB Index (0.11) and highest Dunn Index (13.49).

Ravikumar Gurusamy and Dr. Vijayan Subramaniam [7] propose a machine learning approach for MRI brain tumor classification focusing on denoising, feature extraction, and tumor detection. They use various filters for noise removal, segment the de-noised image, and extract features using wavelet transform, which suits MRI image feature extraction. Features are classified using binary tree support vectors, compared with conventional methods like K-Means, SVM, Neural Network, and KNN. They highlight MRI's ability to show soft tissue contrasts not visible in X-rays and suggest a hybrid approach for tumor detection and classification through MRI. They achieved high classification performance with over 96

Javaria Amin and Muhammad Sharif [8] provide an overview of methods for brain tumor detection via MRI. They cover topics such as brain tumor anatomy, datasets, enhancement methods, segmentation, feature extraction, and classification, as well as the application of deep learning, transfer learning, and quantum machine learning techniques.. The paper notes challenges in accurate segmentation and classification due to tumor variations and emphasizes preprocessing like skull stripping and bias field correction for better analysis. They discuss segmentation methods like thresholding, region growing, watershed, and feature extraction techniques including GLCM, geometrical features, and texture analysis. The paper highlights deep learning models like CNN, LSTM, and U-Net for improving tumor detection and classification.

Eric Carver and Chang Liu [9] explore using a U-net neural network for automatic brain tumor segmentation and an Extreme Learning Machine (ELM) for predicting patient overall survival using preoperative MR images. The study used 210 Glioblastoma Multiforme (GBM) patients for training and 66 Low-Grade Glioma (LGG) and GBM patients for validation. The U-net model achieved Dice Similarity Coefficients (DSC) of 0.882 for the whole tumor, 0.712 for enhanced tumor, and 0.769 for the tumor core, with corresponding Hausdorff distances of 7.09, 4.46, and 9.57. The ELM algorithm showed a predictive power of 0.607 for patient overall survival. The study highlights the deep learning potential in improving the accuracy of medical analysis and clinical decision-making for brain tumor patients.

Sahan M. Vijithananda and Mohan L. Jayatilake [10] discuss brain tumor classification using MRI Apparent Diffusion Coefficient (ADC) images and machine learning. They used 1599 MRI brain ADC image slices from 195 patients, using demographic and texture features to differentiate malignant from benign tumors. They used the ANOVA f-test for feature selection, finding that skewness and GLCM homogeneity were less informative. The Random Forest classifier, with hyperparameter tuning, achieved 90.41% accuracy. The study found that patient gender was important in predicting tumor type and suggested that the model could help in clinical practices for non-invasive tumor differentiation.

B.Balakumar and P.Raviraj [11] present a two-tier classification system for brain tumor diagnosis using MRI images. They use the adaptive pillar K-means algorithm for segmentation and combine the Self-Organizing Map (SOM) neural network and Support Vector Machine (SVM) for classification. The method achieved over 89.5% accuracy and was tested on the Brainix dataset, which has 30 normal and 30 abnormal images. They used texture feature extraction via GLCM, noise removal with non-local mean filters, and edge detection using Sobel, Robert's cross, Prewitt, and Laplacian of Gaussian operators. The study shows that the two-tier classification method performs better than traditional methods in accurately classifying brain images as normal or abnormal.

S. Rinesh and K. Maheswari [12] investigate brain tumor classification using hybrid machine learning algorithms and hyperspectral imaging for accurate tumor localization. They introduce a mapping technique with k-based clustering, optimized by the firefly algorithm, and label brain regions through a multilayer feedforward neural network. Their model outperformed existing methods, achieving 96.47% accuracy, 96.32% sensitivity, and 98.24% specificity. They used k-nearest neighbor, k-means clustering, firefly algorithm optimization, and a multilayer feedforward neural network for brain tissue labeling. The study shows the effectiveness of hyperspectral imaging in improving brain tumor detection and surgical planning.

Shenbagarajan Anantharajan and Shenbagalakshmi Gunasekaran [13] introduce a new method for MRI brain tumor detection using deep learning and machine learning. Their EDN-SVM (Ensemble Deep Neural Support Vector Machine) classifier enhances MRI images by applying an Adaptive Contrast Enhancement Algorithm (ACEA) and a median filter. It then segments the images with fuzzy c-means clustering and extracts features using the Gray Level Co-occurrence Matrix (GLCM). This method achieved 97.93% accuracy, 92% sensitivity, and 98% specificity, outperforming traditional methods like CNN, RFC, ANN, and R-CNN. The study highlights the promise of integrating deep learning and machine learning techniques to improve diagnostics in medical imaging.

Lina Chato and Shahram Latifi [14] present a method for predicting the survival rate of glioma brain tumor patients using MRI images and machine learning. They used the Bra

TS 2017 dataset, which includes MRI brain images, survival times, and patient ages, labeled short-term, mid-term, and long-term survivors. Various features like volumetric, statistical, and intensity texture, histograms, and deep features were extracted and trained with ML methods like SVM, KNN, linear discriminant, tree, ensemble, and logistic regression. The best prediction accuracy was achieved using deep learning features extracted by a pre-trained CNN and trained by a linear discriminant, with an overall accuracy of 73% from the training/validation dataset. The paper suggests future research to denoise data, test histogram features, and improve classification accuracy based on deep features.

Pravin R. Kshirsagar and Anil N. Rakhonde [15] discuss using machine learning algorithms for detecting brain tumors in MRI images. They highlight the importance of computerized detection for large datasets and the complexity of tumor variance. Their work involves preprocessing MRI images, extracting texture features using GLCM, and classifying them with machine learning algorithms. They achieved 98.6% and 91.6% classification accuracy with a sample size of 212 MRI images. Techniques mentioned include Multi-Layer Perceptron and Naive Bayes algorithms. The paper suggests that accuracy could improve with larger datasets and additional features.

3 Methodology

Figure 1 , and Algorithm 1 show the proposed framework for the Brain Tumor System. We will be performing the following operations.

- Collecting Data
- Preprocessing data for model ingestion
- Splitting data between train and test

3.1 Data collection

The data collection stage is essential before conducting a thorough analysis. For our project, this involves extracting valuable insights from the two datasets: the Brain Tumor MRI Dataset, which tells us whether a tumor is present or not, and the Brain MRI scans for Brain Tumor Classification, which classifies for us different kinds of brain tumors. These datasets offer information that includes various MRI images and detailed annotations. Each dataset adds unique aspects to our understanding of brain tumor characteristics, thereby improving our capability to create effective detection and classification models.

1. **Brain Tumor MRI Dataset** : Brain Tumor MRI Dataset serves as a comprehensive repository of metadata for over 7,000 images. This extensive compilation includes details on each disease where there are three diseases: glioma, meningioma, and pituitary, and if there is no disease there will be no tumor. Each label contains certain images where we will make our preprocessing to make clean data to be prepared for machine learning and deep learning. we will add feature extraction and selection for the machine learning to get the best evaluation metrics.

Algorithm 1 Brain Tumor Detection and Classification

Input: MRI brain tumor datasets D_1, D_2, Machine Learning Models $\{ML1, ML2, ..., ML_n\}$, Deep Learning Models $\{DL1, DL2, ..., DL_n\}$
Output: Classified tumor type (glioma, meningioma, pituitary, or none)
procedure Data Collection
    Collect two datasets: Brain Tumor MRI Dataset and Brain MRI Images for Tumor Classification
end procedure
procedure Data Preprocessing
    Resize images to a consistent size (e.g., 224x224 pixels)
    Extract texture features (GLCM) such as contrast, homogeneity, energy, and correlation
    Extract shape features like area, diameter, and perimeter
    Apply feature selection (Chi-square test) to select the most relevant features
end procedure
procedure Data Splitting
    Split data into training (70%), validation (15%), and testing (15%) sets
end procedure
procedure Model Training (Machine Learning)
    for each machine learning model $ML_i \in \{ML_1, ML_2, ..., ML_n\}$ **do**
        Train ML_i on the training set
        Optimize hyperparameters using Grid Search
        Evaluate on the validation set using accuracy, precision, recall, and F1-score
    end for
end procedure
procedure Model Training (Deep Learning)
    for each deep learning model $DL_j \in \{DL_1, DL_2, ..., DL_n\}$ **do**
        Train DL_j on the training set
        Apply dropout and regularization to prevent overfitting
        Use Adam optimizer and categorical cross-entropy loss function
        Monitor validation accuracy and adjust learning rate as needed
    end for
end procedure
procedure Model Evaluation
    Generate predictions on the test set using trained models
    Evaluate models using MSE, MAE, accuracy, and F1-score
end procedure
procedure Classification and Interpretation
    Classify tumor type based on the model with the best evaluation metrics
    Provide insights into the effectiveness of the classification using feature importance
end procedure

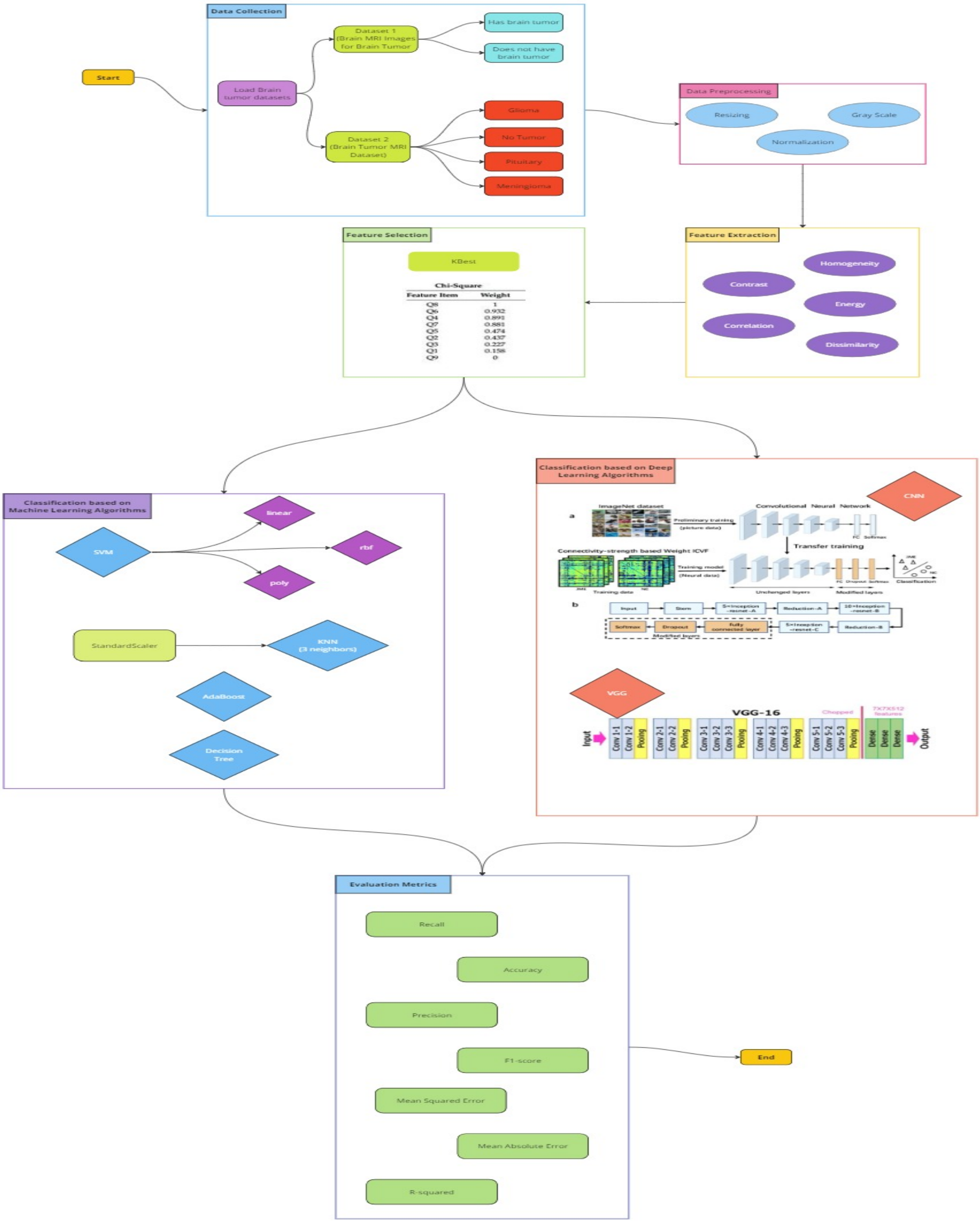

Figure 1: The Proposed System of Brain Tumor Classification.

2. **Brain MRI Images for Brain Tumor Detection** : Brain MRI Images for Brain Tumor Detection serves two labels where is yes or no so if there is brain tumor detection the output will be yes and if not the output will be no. We will make the same steps in the first dataset to get the best metrics and accurate output.

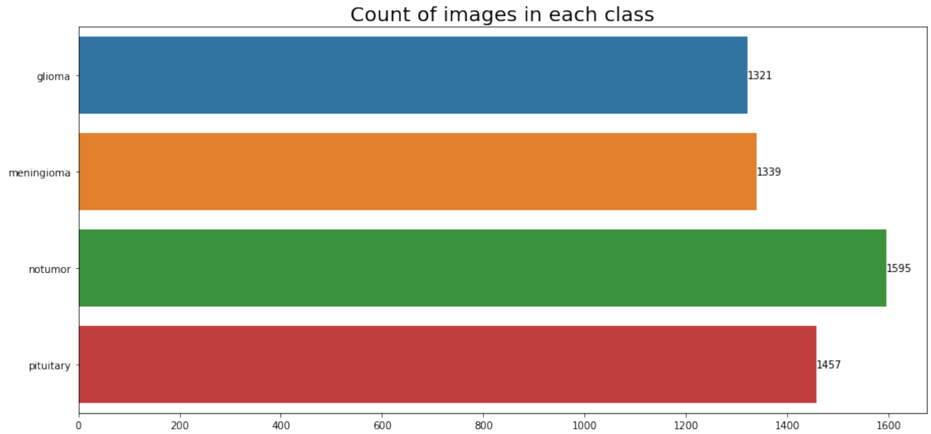

Figure 2: Distribution of classes for the first dataset.

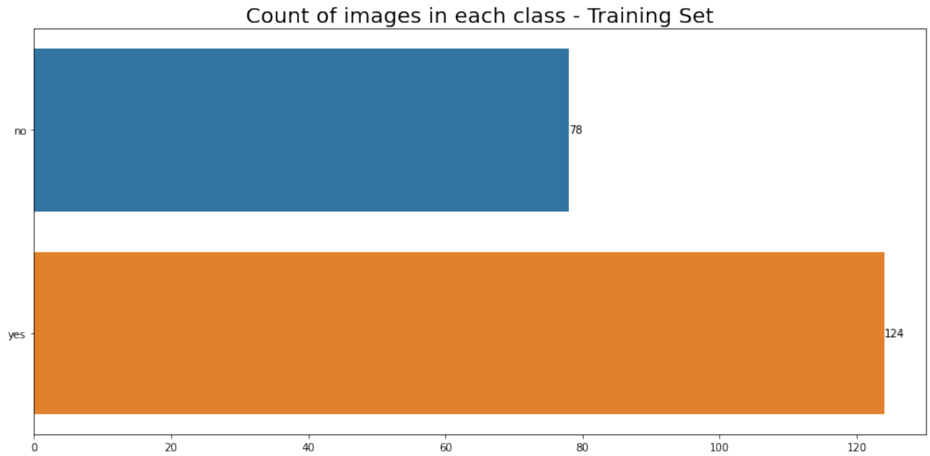

Figure 3: Distribution of classes for the second training dataset.

3.2 Data Preprocessing

In the Data Preprocessing phase, we undertook a comprehensive approach to find in the dataset that the images were not resized correctly so we applied to resize for the images we had and we made for the machine learning feature extraction where we get the best features for our dataset and then apply feature selection to select the best feature to have high accuracy.

The approach is applied to all datasets.

- **Data Exploration:** Our first step was to thoroughly examine both datasets to understand their structure. We discovered that the photos were organized into folders. In the first dataset, there are four folders, each serving as a label that we will detect using ML and DL technologies.
- **Feature Extraction:** We extracted several features from our dataset using texture features, including contrast, dissimilarity, homogeneity, energy, and correlation extracted from grey level Co-occurrence Matrix (GLCM). Additionally, we created shape features such as area, diameter, and perimeter. Although we initially considered Haralick features, they did not yield the best results, so we opted for GLCM features instead.
- **Feature Selection:** After selecting GLCM, we used a chi-square test to categorize the products into different weight categories.

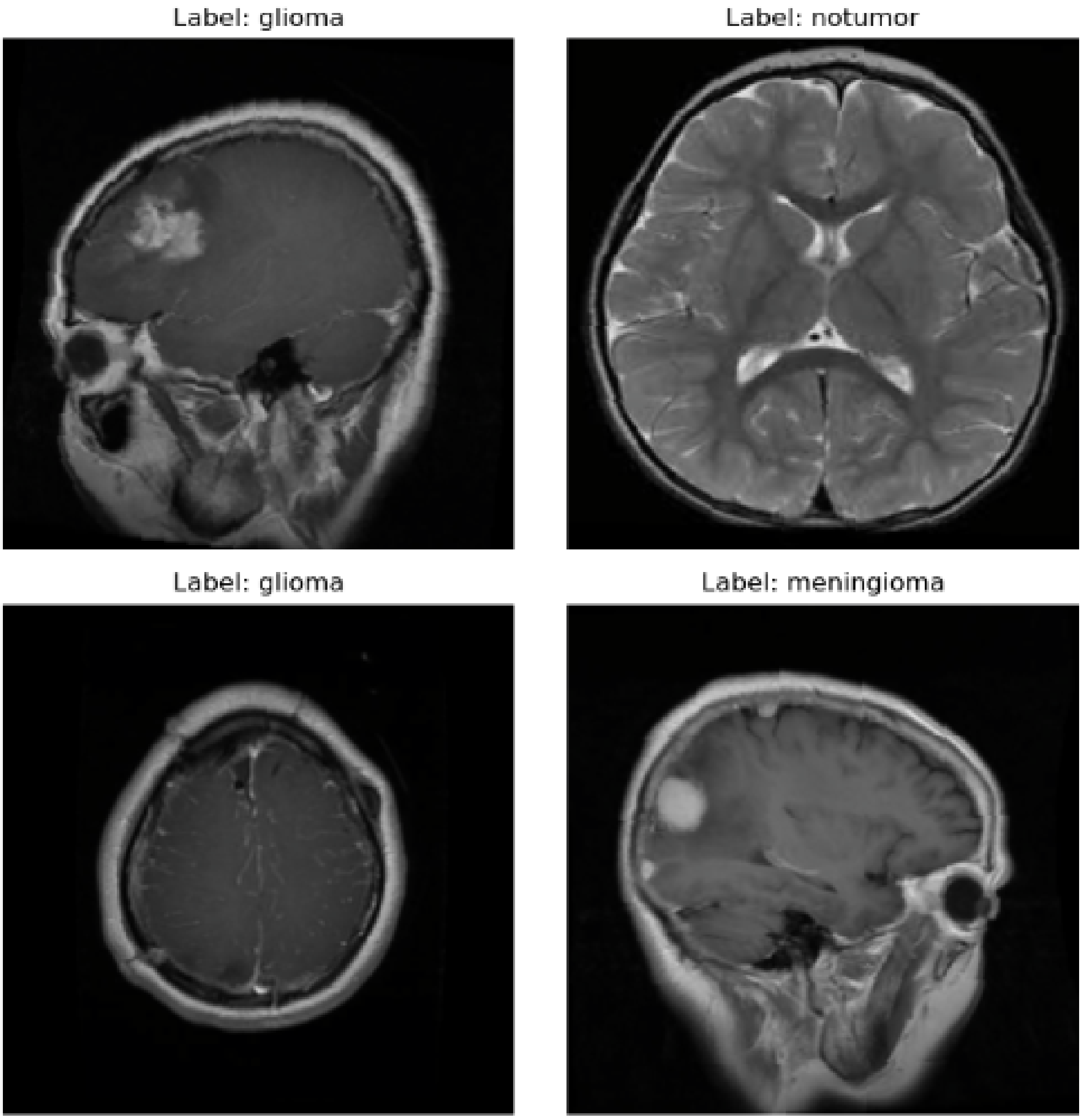

Figure 4: Brain Tumor sample Images

3.3 Data splitting

During data splitting, we use various strategies depending on the nature of the algorithms. For certain algorithms, particularly those sensitive to the training set size and benefiting from a more comprehensive exploration of different training and validation subsets, we divide the dataset into three partitions testing, training, and validation.

Conversely, for algorithms that necessitate larger training sets for effective learning, we implement a different approach. In such scenarios, we employ a holdout method, dividing the data into (70%) allocated to training, (15%) for validation of the model, and (15%) for testing the model with unseen data. This method ensures a substantial training dataset for algorithm training and a distinct validation dataset for parameter tuning. Additionally, a separate test dataset for unbiased evaluation. Combining these strategies, we can customize the data-splitting process to meet the specific needs of each algorithm, enhancing their performance and ability to generalize.

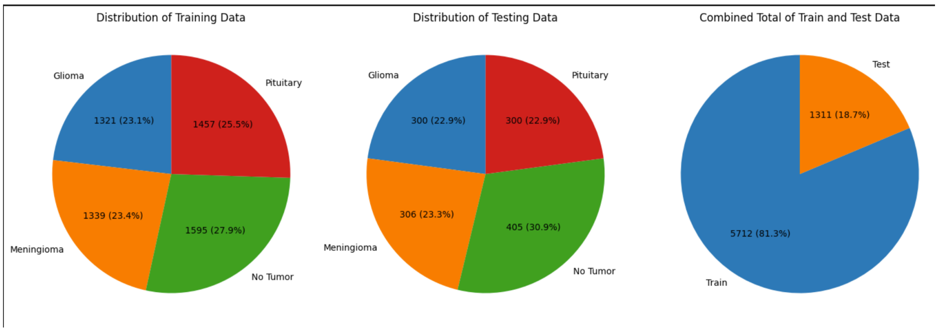

Figure 5: Training and Testing distribution for the first dataset.

3.4 Classification BASED ON AI MODELS

The Classification Systems landscape has witnessed remarkable advancements with using both machine learning and deep learning techniques, offering personalized and relevant suggestions to users in multiple domains. In this phase, we look into two distinctive ML approaches: employing a Decision Tree, KNN, Support Vector Machine, and Adaboost. Additionally, exploring deep learning, we implement the Convolutional neural network and VGG-16. These methodologies, spanning both ML and DL, are poised to unlock intricate patterns in user behavior and preferences, providing valuable insights into classification of brain tumor.

- **Machine Learning Models**
  - **K-Nearest Neighbors (KNN)** The first approach involves constructing a robust K-Nearest Neighbors (KNN) classification model to predict brain tumor images. This process begins with gathering brain tumor images and meticulously preparing the data by extracting relevant features such as GLCM and chi-square, along with the corresponding images. A crucial step is splitting the dataset into the partitions previously discussed to thoroughly evaluate the KNN model's predictive capabilities. The KNN model is then created, focusing on parameters like the number of neighbors and distance weighting schemes. This careful hyperparameter optimization, facilitated by cross-validation, identifies the optimal model configurations, highlighting the model's most effective predictive characteristics based on the following equation:

$$\hat{y} = \frac{1}{k} \sum_{i \in \mathcal{N}_k(x)} y_i \tag{1}$$

  - **Decision Trees** To enhance brain tumor detection within specific genres, we've employed a Decision Tree approach, contributing significantly to personalized content classification systems. Our methodology begins with constructing a brain tumor detection matrix, utilizing genre-wise user ratings as features, with the 'Action' genre designated as the target variable. We then partition the data into training and testing to ensure robust evaluation of our model.

    The Decision Tree classification is then initialized with a specified maximum depth, enabling control over model complexity, and trained on the training set. An essential analytical component involves extracting feature importances from the trained model, illuminating the relative significance of different genres in predicting ratings for the target genre ('Action'). This insight aids in discerning which user interactions contribute most substantially to the model's predictive capabilities.

    Moreover, we instantiate another Decision Tree model with a different maximum depth to investigate the impact of model complexity on predictive performance. We evaluate the predictive accuracy of both models on the test set and analyze the obtained predictions for insights into capturing user preferences for the target genre.

    This comprehensive methodology not only demonstrates a practical implementation for predicting genre-specific movie ratings but also offers researchers a versatile template for extending the analysis to diverse genres and optimizing model performance. The adaptable nature of this approach positions it as a valuable tool in the broader domain of content recommendation research, emphasizing interpretability through feature importance and providing a nuanced understanding of user preferences within a specific genre.

  - **Support Vector Machines** (SVM) offer a powerful method for predicting brain tumor types based on various features extracted from MRI images. In our approach, we utilize SVM with different kernels provided to us by the model's library; These include: linear, polynomial, and radial basis functions (RBF). This enables us to capture diverse patterns and relationships within the data.

    Our methodology begins with initializing SVM classifiers for each kernel type and training them on the provided dataset, which is contrived of features extracted from tumor MRI images, with the corresponding tumor types as labels. After training, the classifiers make predictions on a separate test set.

    By systematically evaluating the performance of SVM classifiers with different kernel functions, we gain valuable insights into their effectiveness in accurately predicting brain tumor types. This analysis not only facilitates the selection of the most suitable SVM model for our specific task but also provides a deeper understanding of how different kernel functions impact classification performance.

  - **AdaBoost** (Adaptive Boosting) is a powerful learning technique(known as ensemble learning) that combines multiple weak learners to create a strong classifier. In the context of brain tumor detection, AdaBoost can be utilized to effectively classify different classes of brain tumors based on features interpolated from MRI scans.

    Our methodology begins with initializing an AdaBoost classifier and training it on the provided dataset. The dataset contains features extracted from MRI scans of brain tumors, with the corresponding tumor types as labels. After training, the classifier makes predictions on a separate test set.

    By systematically evaluating the performance of the AdaBoost classifier, we gain valuable insights into its effectiveness in accurately predicting brain tumor types. This analysis not only facilitates the selection of the most suitable AdaBoost model for our specific task but also provides a deeper understanding of its strengths and limitations in the context of brain tumor detection.

- **Deep Learning Models**

  - **Convolutional Neural Network (CNN)**

    In the complex design of our CNN model is tailored for brain tumor detection, multiple layers collaborate to enhance the model's capacity to understand and predict tumor characteristics within MRI images.

    Convolutional Neural Networks are a powerful class of DL models specifically designed for processing grid data like in our use-case; images. In brain tumor detection, CNNs can effectively learn hierarchical representations of MRI images, enabling accurate classification of different tumor types.

    methodology begins with preparing the dataset by loading the images from the specified directories and resizing them to a consistent size (e.g., 224x224 pixels). We split the dataset into three datasets training and validation, with a small portion reserved for validation to monitor model performance during training.

    The CNN architecture consists of several convolutional layers followed by max-pooling layers, which help extract and learn meaningful features from the input images. Subsequent dense layers and dropout regularization are employed to further refine the learned features and decrease the chance of overfitting during training. The final dense layer of the model is a softmax activation function to output probabilities for each class of tumor in our dataser.

    We compile the model using an appropriate optimizer (Adam) and a suitable loss function such as categorical cross-entropy for multi-class classification. Additionally, we include metrics such as categorical accuracy to monitor model performance during training.

    Training the CNN involves iterating over the training dataset for a certain number of epochs while periodically evaluating the model's performance on the validation dataset. We employ callbacks such as ModelCheckpoint, ReduceLROnPlateau, and EarlyStopping to save the best-performing model, adjust the learning rate, and prevent overfitting, respectively.

By systematically training and evaluating the CNN model on the provided dataset, we aim to achieve high accuracy in the classification the brain tumors type we mentioned earlier in the report. This analysis not only facilitates the selection of the most suitable CNN architecture and hyperparameters but also enables us to understand more on the model's ability to generalize its training on unseen before data.

– **(VGG)** The methodology begins with preparing the dataset by loading images from specified directories and applying the necessary preprocessing steps. Each image is resized to a consistent size (e.g., 224x224 pixels) to ensure uniformity. The dataset is then split into training, validation, and test sets, with a small portion reserved for validation to monitor model performance during training and a separate portion for final evaluation.

  The core of our approach involves leveraging the VGG16 architecture, a pre-trained CNN model known for its effectiveness in image classification tasks. The VGG16 model, pre-trained on the ImageNet dataset, serves as the base of our model. This allows us to benefit from its learned features while adapting it to our specific classification task.

  Furthermore, we add some configurations like freezing layers, Custom Layers, and callback

- **Mean Squared Error (MSE)** Mean Squared Error computes the average of the number's squared differences between the theoretical ratings and the practical ratings. It quantifies the variance of the prediction errors and is used to measure the accuracy of the model's predictions. The formula for MSE is given by:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{2}$$

- **Mean Absolute Error** MAE computes the absolute average deviation between theoretical and practical images. Determined by calculating the mean of the absolute differences, The formula for MAE is expressed as:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{3}$$

  where n is the number of predictions, y_i represents the actual rating, and $\hat{y}_i$ is the predicted rating. Similar to MSE, a lower MAE indicates better accuracy, with zero representing a perfect prediction.

4 Experimental Results and Discussion

During the analysis and discussion phase of our study, our brain tumor detection system demonstrated strong performance across two distinct datasets: the Brain Tumor MRI Dataset and Brain MRI Images for Brain Tumor Detection. We investigated two different machine learning approaches: Decision Tree, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and AdaBoost. Additionally, we explored deep learning techniques by implementing both Convolutional-Neural-Network & VGG-16 models. To evaluate the accuracy and effectiveness of these models, we employed performance metrics such as MSE, MAE showcasing their adaptability. This comprehensive examination of various methodologies underscores their potential for further advancement in brain tumor detection research.

4.1 Case Study I (Brain Tumor MRI Dataset)

The brain tumor detection underwent a comprehensive evaluation using six different algorithms: Support vector machine (SVM), K-nearest neighbors (KNN), Decision Trees, AdaBoost, CNN, and VGG-16. The best algorithm is VGG-16 where it's validation accuracy is 0.9498 and the second is for the decision tree where its accuracy is 0.8581

The different performance metrics show us what each algorithm does well and where it struggles, stressing the importance of carefully evaluating them when picking a brain tumor detection system. In our comparison, the support vector machine (SVM) had the lowest accuracy across all its variations—linear, polynomial, and RBF—suggesting it might be better at predicting than other methods. Still, we need to look deeper, considering specific dataset details, to choose the best algorithm for spotting brain tumors effectively.

4.2 Case Study II (Brain MRI Images for Brain Tumor Detection)

We apply the same algorithms for the second dataset and the performance accuracy is lower than the first dataset as the images in it are smaller than the first one, The best performance metrics for the second dataset are for VGG-16 which comes with an accuracy of 0.8919, and the decision tree with 0.8619.

Table 1: Evaluating Metrics for Dataset 1

P.O.C	Precision	Recall	f1-score	Accuracy
K-NN	0.8171	0.8199	0.8156	0.8199
SVM-linear	0.6492	0.6681	0.6551	0.6681
SVM-poly	0.3944	0.4088	0.3217	0.4088
SVM-rbf	0.4849	0.6155	0.5399	0.6155
Adaboost	0.6693	0.6880	0.6713	0.6880
Decision Tree	0.8600	0.8518	0.8576	0.8581

Table 2: Evaluating MSE and MAE for Dataset 1

P.O.C	MSE	MAE
SVM	1.3775	0.9199
KNN	1.3775	0.9199
Adaboost	1.4019	0.9443
Decision Tree	1.4530	0.9954

Table 3: Evaluating Metrics for Dataset 2

P.O.C	Precision	Recall	f1-score	Accuracy
K-NN	0.8038	0.8077	0.8024	0.8077
SVM-linear	0.6984	0.7078	07012	0.7078
SVM-poly	0.7004	0.6262	0.6359	0.6262
SVM-rbf	0.6920	0.7055	0.6950	0.7055
Adaboost	0.6693	0.6880	0.6713	0.6880
Decision Tree	0.8636	0.8619	0.8612	0.8619

Table 4: Evaluating MSE and MAE for Dataset 2

P.O.C	MSE	MAE
SVM	1.3775	0.9199
KNN	1.3775	0.9199
Adaboost	1.4019	0.9443
Decision Tree	1.4530	0.9954

Table 5: Results Summerzation for the two datasets

P.O.C	Validation acc Dataset-1	Validation acc Dataset-2
CNN	0.9171	0.8108
VGG	0.9498	0.8919

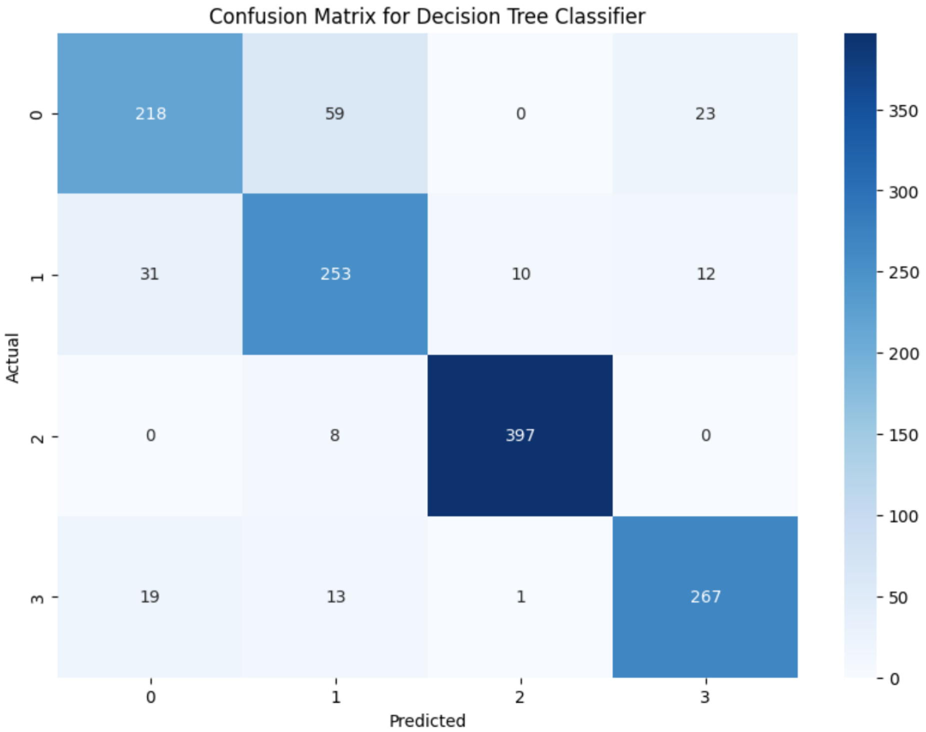

Figure 6: Best confusion Matrix is decision tree

The varied performance metrics reveal the strengths and weaknesses of each algorithm, underscoring the significance of thorough evaluation in selecting a brain tumor detection system. In our analysis, both the Support Vector Machine (SVM) and Adaboost exhibited lower accuracy across polynomial variations, suggesting potential superiority in predictive capabilities compared to other methods. However, a deeper examination of dataset specifics is necessary to determine the most suitable algorithm for effective brain tumor detection.

5 Conclusion

In our study, we conducted two distinct case studies aimed at addressing different facets of brain tumor detection and classification, each leveraging machine learning and deep learning techniques to enhance diagnostic capabilities.

5.1 Case Study I

In the first case study, we focused on utilizing the Brain Tumor MRI Dataset to determine the presence or absence of brain tumors in medical scans. This task is fundamental in clinical diagnostics, as early detection plays a crucial role in patient prognosis and treatment planning. Our objective was to develop robust machine learning and deep learning models capable of accurately classifying brain scans into two categories: tumor-positive and tumor-negative.

To achieve this, we employed a variety of algorithms, including Decision Trees, K-nearest Neighbors (KNN), Support Vector Machine (SVM), AdaBoost, Convolutional Neural Networks (CNN), and VGG-16 models. Each algorithm was trained and evaluated using the dataset to assess its effectiveness in detecting brain tumors. By analyzing the performance metrics such as accuracy, precision, recall, and F1-score, we aimed to identify the most suitable approach for reliable tumor detection.

For future work, we could involve exploring advanced deep learning architectures, such as recurrent neural networks (RNNs) or attention mechanisms, to further improve detection accuracy. Additionally, the integration of multi-modal data, such as functional MRI (fMRI) or diffusion tensor imaging (DTI), could enhance the robustness of detection models.

5.2 Case Study II

In the second case study, we shifted our focus to the classification of different types of brain tumors using the Brain MRI Images for Brain Tumor Detection dataset. Unlike the previous case study, which aimed at binary classification (tumor present or absent), this task involved categorizing tumors into specific types, such as glioma, meningioma, and pituitary tumors. This finer classification is essential for personalized treatment strategies and prognostic assessment.

Similar to Case Study 1, we employed machine learning and deep learning algorithms for classification tasks. Decision Trees, KNN, AdaBoost, SVM, CNN, and VGG-16 models were trained and evaluated using the dataset to classify brain tumors into their respective categories. Performance evaluation metrics such as F1-score, accuracy, recall, and precision were utilized to assess the efficacy of each algorithm in accurately classifying tumor types.

In future work, the dataset needs to be expanded to include a broader range of brain tumor types and sub-types that would be beneficial for developing more comprehensive classification models. Furthermore, research efforts could focus on integrating additional patient-specific data, such as genetic information or clinical history, to improve the accuracy and personalized nature of tumor classification. Deployment of these advanced models in clinical settings would also be a crucial step toward real-world application and validation. Overall, continued research and innovation in machine learning and deep learning techniques hold great promise for advancing brain tumor diagnosis and treatment.

Author contributions

All authors made equal contributions to this paper. Ahmed Mahmoud, Mohamed Amr, Kareem Amr, and Mohamed Wael participated in organizing the experiments, discussing and analyzing the results, writing the paper, and revising it. All authors read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] G. Çınarer and B. G. Emiroğlu, "Classificatin of brain tumors by machine learning algorithms," in *2019 3rd international symposium on multidisciplinary studies and innovative technologies (ISMSIT)*, pp. 1–4, IEEE, 2019.

[2] D. N. George, H. B. Jehlol, A. S. A. Oleiwi, *et al.*, "Brain tumor detection using shape features and machine learning algorithms," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 5, no. 10, pp. 454–459, 2015.

[3] K. Sharma, A. Kaur, and S. Gujral, "Brain tumor detection based on machine learning algorithms," *International Journal of Computer Applications*, vol. 103, no. 1, 2014.

[4] F. Polly, S. Shil, M. A. Hossain, A. Ayman, and Y. M. Jang, "Detection and classification of hgg and lgg brain tumor using machine learning," in *2018 International Conference on Information Networking (ICOIN)*, pp. 813–817, IEEE, 2018.

[5] G. Manogaran, P. M. Shakeel, A. S. Hassanein, P. M. Kumar, and G. C. Babu, "Machine learning approach-based gamma distribution for brain tumor detection and data sample imbalance analysis," *IEEE Access*, vol. 7, pp. 12–19, 2018.

[6] A. Vidyarthi and N. Mittal, "Brain tumor segmentation approaches: Review, analysis and anticipated solutions in machine learning," in *2015 39th National Systems Conference (NSC)*, pp. 1–6, IEEE, 2015.

[7] R. Gurusamy and V. Subramaniam, "A machine learning approach for mri brain tumor classification," *Computers, Materials and Continua*, vol. 53, no. 2, pp. 91–109, 2017.

[8] J. Amin, M. Sharif, A. Haldorai, M. Yasmin, and R. S. Nayak, "Brain tumor detection and classification using machine learning: a comprehensive survey," *Complex & intelligent systems*, vol. 8, no. 4, pp. 3161–3183, 2022.

[9] E. Carver, C. Liu, W. Zong, Z. Dai, J. M. Snyder, J. Lee, and N. Wen, "Automatic brain tumor segmentation and overall survival prediction using machine learning algorithms," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 4th International Workshop, BrainLes 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Revised Selected Papers, Part II 4*, pp. 406–418, Springer, 2019.

[10] S. M. Vijithananda, M. L. Jayatilake, B. Hewavithana, T. Gonçalves, L. M. Rato, B. S. Weerakoon, T. D. Kalupahana, A. D. Silva, and K. D. Dissanayake, "Feature extraction from mri adc images for brain tumor classification using machine learning techniques," *Biomedical engineering online*, vol. 21, no. 1, p. 52, 2022.

[11] B. Balakumar, P. Raviraj, and E. D. Devi, "Brain tumor classification using machine learning algorithms," *Elysium Journal of Engineering Research and Management*, vol. 4, no. 2, pp. 30–41, 2017.

[12] S. Rinesh, K. Maheswari, B. Arthi, P. Sherubha, A. Vijay, S. Sridhar, T. Rajendran, Y. A. Waji, *et al.*, "Investigations on brain tumor classification using hybrid machine learning algorithms," *Journal of Healthcare Engineering*, vol. 2022, 2022.

[13] S. Anantharajan, S. Gunasekaran, T. Subramanian, and R. Venkatesh, "Mri brain tumor detection using deep learning and machine learning approaches," *Measurement: Sensors*, vol. 31, p. 101026, 2024.

[14] L. Chato and S. Latifi, "Machine learning and deep learning techniques to predict overall survival of brain tumor patients using mri images," in *2017 IEEE 17th international conference on bioinformatics and bioengineering (BIBE)*, pp. 9–14, IEEE, 2017.

[15] P. R. Kshirsagar, A. N. Rakhonde, and P. Chippalkatti, "Mri image based brain tumor detection using machine learning," *Test Engineering and Management*, vol. 81, pp. 3672–3680, 2020.