Multi-Classification of Brain Tumor MRI Images Hybrid VGG16 Support Vector Machine Model

Asifa Iqbal[1,*]
[1]School of international languages Zhengzhou University, Henan, China
Email: asifaiqbal615@gmail.com

Abstract

Tumor brain research stands essential for detecting patients during timely periods and delivering proper treatment options. Inspecting tumors becomes difficult because tumor morphology shows diverse characteristics in terms of dimensions and placement surface texture patterns, and inconsistent visual features across various medical image types. A combined methodology will be implemented to detect brain tumors through MRI image analysis in this research. The model operated with three publicly accessible datasets containing 3,966 T1-weighted contrast-enhanced magnetic resonance images (T1-w MRI) that were split between glioma, meningioma, pituitary tumor and no tumor groups. The diagnosis pipeline starts by applying preprocessing and data augmentation steps that improve data quality alongside increasing its variability rates. The main structure of this system uses VGG16 deep convolutional neural network features alongside a Support Vector Machine (SVM) classifier to determine outputs. The modified VGG16 output became the SVM input, delivering optimal results while keeping the computational time sensible. The proposed hybrid model performs better than all existing methods analyzed in the literature according to experimental results. The test success rate of the model reached 97.2%. Test outcomes from standard machine learning methods XGBoost, AdaBoost, Decision Tree, and K-Nearest Neighbors demonstrate that using SVM as the endpoint classifier boosts achievement levels in this dataset assessment.

Keywords: Brain Tumors; Hybrid Model; Support Vector Machine; T1-w MRI; VGG16

1 Introduction

The medical community identifies brain tumors as a crucial subject because of their challenging composition, careful growth patterns, and grave influence on survival rates and patient well-being. The tumors evolve without noticeable symptoms and go undiscovered until they progress to extensive stages, where treatments prove more complex to implement. The American Cancer Society reports that brain and spinal cord malignant tumors will result in 18,330 U.S. deaths during 2025, with male and female fatalities totaling 10,170 and 8,160, respectively [1]. Patient outcomes depend heavily on achieving both early diagnosis and appropriate treatment methods because this data shows the critical nature of the situation. The term brain tumor indicates any irregular tissue development that occurs inside the spinal cord or brain, which blocks usual neurological operations. The abnormal tissue develops through uncontrolled cell division into benign noncancerous tumors and malignant cancerous tumors. Tumors of benign origin maintain clear structural borders and develop gradually, yet malignant tumors expand rapidly until they invade different regions within the brain and spinal tissue. Malignant tumors result in elevated morbidity and mortality statistics since these aggressive tumors prove resistant to typical treatment modalities [2]. Multiple criteria determine brain tumor classification, including their cellular distinctiveness, where they develop within the brain and their conduct. The three main brain tumor classifications consist of gliomas, followed by meningiomas, together with pituitary tumors. Over thirty percent of brain and CNS tumors are gliomas since these tumors originate from glial cells, while gliomas make up eighty

DOI: https://doi.org/10.54216/JAIM.090204

Received: December 18, 2024 Revised: February 25, 2025 Accepted: May 02, 2025

percent of all malignant brain tumors. According to the World Health Organization (WHO) classifications, anaplastic astrocytoma and glioblastoma belong to the particularly aggressive high-grade tumor types IV and III, respectively [3]. The prediction of treatment outcome and selection of suitable therapy depends heavily on identifying the histological grade. Meningiomas grow as benign growths from the meninges surrounding the brain and spinal cord area. The growth of these masses typically stays benign, but their expansion causes pressure on brain tissue, thus creating neurological symptoms with functional impairments. The pituitary gland serves as the origin of tumors called pituitary tumors while they form from this gland that controls hormonal actions. Near the optic chiasm, these tumors result in both hormonal system abnormalities together with vision problems [4]. Treatment of pituitary adenomas becomes necessary because of their benign nature, although they cause endocrine disturbances. Medical doctors need exact brain tumor diagnoses quickly because these tests bring direction to healthcare choices and enhance treatment success. Medical professionals use Magnetic Resonance Imaging (MRI) as their primary assessment tool since it stands as the most trustworthy and frequently implemented technique for brain tumor diagnosis. MRI displays two main benefits, resulting from its non-invasive approach and its capacity to maintain high-resolution imaging and advanced tissue differentiation powers that assist professionals in identifying tumor type while determining boundary location and effectively tracking treatment progress [5]. Because MRI prevents ionizing radiation, which CT scans produce, it becomes an ideal technology for serial tumor assessments. Overall, technological progress has not changed the fact that traditional MRI analysis requires a long duration, resulting in inconsistent interpretation between readers and becoming unstable due to human tiredness. Computer-aided diagnosis systems have become recognized as an effective method that supports radiologists. Medical imaging systems perform tumor detection in addition to classification and segmentation tasks to decrease diagnostic mistakes and enhance workflow processes according to [6]. The addition of AI technology to CAD tools changed medical imaging from being dependent on human perception to a standardized analysis through data. Medical image analysis receives exceptional benefits from Deep learning models, mainly through Convolutional Neural Networks (CNNs) [7]. The models can learn hierarchical pixel-based representations without manual feature programming. Recent research proves that CNNs perform at or better than the professional radiologist level, specifically in brain imaging applications [2]. CNNs have become a fundamental diagnostic imaging ingredient because of their ability to adapt and function at different scale levels. Standard Machine Learning (ML) approaches maintain their validity for particular settings, yet they can work optimally through dual integration with DL-based systems [8]. Support Vector Machines (SVM) demonstrate excellent capabilities in high-dimensional spaces because of their robust nature, and they maintain efficiency even with datasets of limited to moderate sizes. Hybrid systems built from CNN features and ML classification strength present a framework that resolves performance needs and interpretability requirements according to [3]. Our research develops a mixed deep learning methodology that employs a VGG16 convolutional neural network to extract deep features before the SVM classifier completes the prediction process. The proposed architecture aims to detect glioma and meningioma tumors and differentiate between pituitary tumors and no tumors from MRI brain images. The proposed model achieves better diagnosis alongside decreased computational demands through joint implementation of VGG16 feature extraction capabilities and SVM precise decision boundaries according to [9]. The model gains better generalization capability when it utilizes both local MRI scans and publicly available MRI scans. The available dataset was augmented through multiple transformation operations consisting of rotations, flips, and zooming to prevent overfitting. Training through this methodology prepares the model to deal with various imaging conditions and tumor presentation types to perform reliably in medical practice [10]. This study yields three central accomplishments that include:

- The proposed model features VGG16 and SVM to develop a mixed architecture for precise multi-class brain tumor type classification between deep and machine learning approaches.
- The proposed model demonstrates strong performance in sensitivity and specificity measures and F1-score evaluation criteria, which indicates its practical ability in tumor classification discrimination.
- The model achieves training through data collected from both public repositories and local databases and advanced data enhancement methods to maintain flexibility with diverse imaging perspectives. Including no-tumor cases within the system enhances clinical screening abilities because this capability lets the system detect standard brain images separate from pathological images.

DOI: https://doi.org/10.54216/JAIM.090204

Received: December 18, 2024 Revised: February 25, 2025 Accepted: May 02, 2025

Literature Review

Medical imaging professionals face essential difficulties when monitoring brain tumors, which leads to better patient recovery due to swift diagnostic methods. Doctors commonly select Magnetic Resonance Imaging (MRI) because it provides outstanding image quality without invading patients. Medical image analysis of MRI utilizes both traditional statistical approaches and machine learning mechanisms, which include linear regression, logistic regression and Bayesian regression, as well as decision trees and random forests with single-layer and multi-layer perceptrons together with deep learning frameworks like convolutional neural networks (CNNs), recurrent neural networks (RNNs), and long short-term memory (LSTM) networks. CNNs emerged as the leading method for tissue type identification because they performed best at identifying regular brain areas and different tumor categories like gliomas men, meningiomas, and pituitary tumors; thus, they became the preferred selection for practical clinical use [11]. Research teams created new versions of CNN models that specifically optimized multi-class brain tumor identification systems. Group normalization combined with max pooling and different kernel sizes in multiple convolutional layers were included in these models. Training accuracy reached 99.8%, and test accuracy hit 99.3% through the implementation of Adam optimizer and image augmentation by a single model. Deep learning models achieve superior clinical diagnostic results after organizations carefully design and train them for clinical applications [12]. A complete Ensemble Deep Learning Network (EDLNet) was developed by uniting Modified Faster R-CNN for classification with a Deep Recurrent Convolutional Neural Network (DRCNN) for tumor segmentation functions. The hybrid architecture performed tests on multiple public databases to reach exceptional classification accuracy, with results reaching 99.76% and 99.87%. By utilizing the ensemble approach, radiologists gain adequate results in clinical workflows since deep learning features from space and time merge to generate reliable classification and segmentation outcomes [13]. The technique received additional enhancement through evolutionary optimization methods. A new method uses an Evolutionary Gravitational Neocognitron Neural Network (EGNNN) and the Marine Predators Algorithm (MPA) alongside VGG16 for extracting features. The technique implemented both components of denoising alongside feature fusion to enhance task results. This model achieved better accuracy and sensitivity than the existing models AlexNet-SVM and ResNet-SGD while displaying higher precision levels because of optimized training methods according to [14]. A U-Net-based framework comprising segmentation and classification elements used ResNet50 as its backbone architecture. The model achieved an Intersection over Union measurement of 0.9504. NASNet demonstrated the best performance out of benchmarked CNNs as it reached 99.6% classification accuracy. The results demonstrate that combining segmentation and classification methodologies in real-time tumor detection systems shows promising potential, according to the study [15]. Transfer learning establishes itself as a prominent technology in this field. The researchers executed CNNs on two public databases, which labeled tumors between meningioma, glioma and pituitary and separated gliomas into grades II, III and IV. The processing models reached dual performance targets, which included 96.13% success rate and 98.7%. The research applied pre-trained ResNet50 and InceptionV3 networks to demonstrate that domain knowledge transfer between fields improves results in challenging medical diagnosis tasks [16]. Three different CNN models were developed to execute specific medical image classification tasks, including binary tumor detection followed by tumor type categorization (typical, glioma, meningioma, pituitary, metastatic) and tumor grade determination. Through grid search optimization, the implemented models reached classification accuracy levels at 99.53%. The design of customized CNN networks delivered superior results than standard neural network architectures, including AlexNet DenseNet121 and GoogleNet, as shown in [17]. Current research demonstrates that using Grid search delivers successful results for model optimization procedures. The optimization of CNN hyperparameters through this study achieved remarkable task-related accuracy of 99.33% for tumor detection together with 92.66% and 98.14% for the five-class classification and tumor grading tasks, respectively. The research findings validate that deep learning model development needs automated parameter tuning because precision requirements are essential in healthcare applications [18]. This study developed a non-hybrid lightweight CNN model, demonstrating potential for clinical environments with limited resources. A basic model achieved 99.76% accuracy in classifying samples alongside a Matthews Correlation Coefficient rating of 0.929. The network delivers high-quality results with simplified hybridization technology, which makes it suited for applications at scale in locations with minimal computational capacity [19]. The authors improved further performance by adapting the concluding layers of prominent CNN architectures, including Xception alongside DenseNet201 and InceptionResNetV2. The integrated dense blocks with softmax outputs allowed the proposed model to achieve 99.67% accuracy for three-class identification along with 95.87% for four-class categorization. The research demonstrates that pre-trained models remain versatile when applied to customized medical imaging tasks according to findings presented in reference [20]. The practice of diagnostic artificial intelligence depends heavily on binary classification systems. The improved ResNet50 architecture delivered both 99.30% accuracies for benign clas-

DOI: https://doi.org/10.54216/JAIM.090204

Received: December 18, 2024 Revised: February 25, 2025 Accepted: May 02, 2025

sification and 98.40% for malignant classification. This method demonstrates the value of improving network enhancements and training methods because it surpasses the commonly used models, including AlexNet and VGG-16 [21]. The proposed CNN-based method trained its models on data from Navoneel Chakrabarty combined with the BT-multiclass repository available through Kaggle. The proposed system obtained 91% success in binary classification while reaching 92% in multi-class classification. The CNN architecture provided performance levels that matched InceptionV3 and VGG-16 standard models while providing flexible solutions for medical clinical usage [22]. Deep CNNs, when merged with traditional machine learning classifiers, show the potential to create hybrid systems. Researchers who conducted this study employed a custom CNN for obtaining features before applying Bayesian-optimized ML models for classification purposes. The hybrid framework delivered better performance than nine competing CNN-SVM approaches. It achieved 97.15% mean classification accuracy while indicating how DL integration with traditional ML produces superior results along with enhanced training capabilities [23]. Researchers developed TumorDetNet as an end-to-end deep learning model that utilizes a deep CNN with 48 convolutional layers, LReLU activation, dropout, and average pooling to minimize overfitting. The model generated outstanding performance outcomes on six Kaggle datasets and reached 99.83% detection accuracy, 100% benign/malignant classification accuracy, and 99.27% three-class identification accuracy. The clinical validity of this system supports its practical use in medical diagnostic settings [24]. In another study, the researchers applied MobileNetV2 architecture to perform detection and classification tasks. The model achieved high generalizability by maintaining accuracy rates of 99.85% for detection tasks and 99.87% for benign/malignant classification, as well as 99.38% for glioma, pituitary and meningioma classification. The mobile health solutions would benefit directly from its efficient operations [25]. One primary obstacle to advancing AI medicine adoption is the lack of human comprehension when AI systems generate predictions. An explainable deep learning (XDL) framework was developed using VGG-19 models, including pre-trained and scratch-built configurations and EfficientNet. The combination of Grad-CAM, Grad-CAM++ and CAM visualization methods enabled the system to generate transparent visual explanations of its results. The combination of pre-trained VGG-19 with Grad-CAM achieved the best accuracy and interpretability according to test results. Establishing this approach brings critical benefits to medical practitioners seeking explanations and trust in deep learning solutions [26].

2 Methodology

Figure 1 illustrates the complete workflow of the proposed investigation pipeline, which demonstrates the analysis stages. The pipeline manages several important operations, which start with preprocessing while moving through data augmentation, then follow feature extraction using a pre-trained backbone, and ends with the classification block containing our proposed model. Each stage is essential for enhancing model performance and ensuring robustness when applied to medical imaging tasks, specifically brain tumor classification. Data

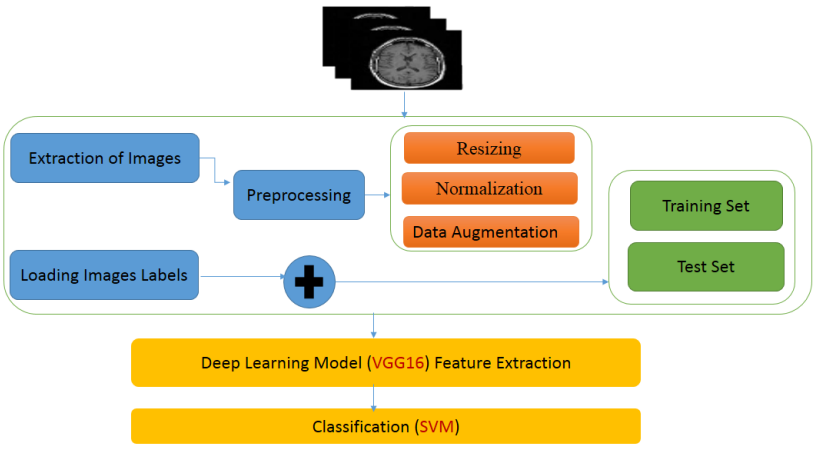


Figure 1: The full procedure of the current Analysis pipeline.

DOI: https://doi.org/10.54216/JAIM.090204

Received: December 18, 2024 Revised: February 25, 2025 Accepted: May 02, 2025

preprocessing is the initial stage that applies resizing, normalization, and noise reduction procedures to input MRI images to maintain dataset consistency. The essential preprocessing step must address the resolution con, trust, and brightness variations found in imaging data from various medical sources. After preprocessing, the data processing techniques generate artificial expansions of the training dataset. Operations from the augmentation process, like rotation and flipping alongside zooming and shifting, enable models to achieve better generalization ability while preventing overfitting. This method requires special attention since collecting large databases with annotations remains challenging for medical applications. We utilize a pre-trained CNN backbone in the feature extraction stage to capture meaningful patterns and spatial hierarchies from the input images. While traditional pipe ines employ a softmax layer as the final classification head in such pre-trained networks, our proposed framework replaces this with a Support Vector Machine (SVM) classifier integrated at the end of the feature extractor. The SVM's superior generalization boundaries are the main reason for this shift, particularly when data classes resist linear separation or display imbalances. SVM classifiers deploy a margin-maximizing strategy instead of softmax classifiers, which depend on probability distributions while operating in high-dimensional areas. This makes them more suitable in our hybrid context, as they can robustly separate classes like glioma, meningioma, pituitary tumor, and healthy (no tumor) instances, each presenting unique visual and structural features in brain MRIs. Our research incorporates systematic experiments to show how the proposed model achieves better results than XGBoost and Decision Trees, K-Nearest Neighbors, and AdaBoost alternative machine learning classifiers. These results are further reinforced by quantitative metrics and visual diagnostics such as confusion matrices and ROC curves, which are discussed in detail in later sections. The proposed investigation pipeline uses pre-trained CNN deep learning capabilities to produce an accurate brain tumor classification framework alongside the SVM classifier implementation of the classical machine learning function.

2.1 Dataset

The study adopted the dataset Cheng initially developed with local input from the Diagnostic Radiology Department at Mansoura University. The extensive brain MRI image collection maintains 3,504 scans, representing various clinically essential brain tumor types [27] [28]. The classification system offers support to diagnose four distinct categories: glioma, meningioma, and pituitary tumor, together with a new class tag named no tumor. The distribution for brain image categories includes glioma (1,003 images) followed by meningioma (981 images) and pituitary tumor (959 images) with no tumor (561 images) at the end as detailed in Table 1. This database featuring different types of brain tumors serves as an extensive resource to develop practical training and testing approaches for ML models that identify brain tumors. The model demonstrates generalization ca-

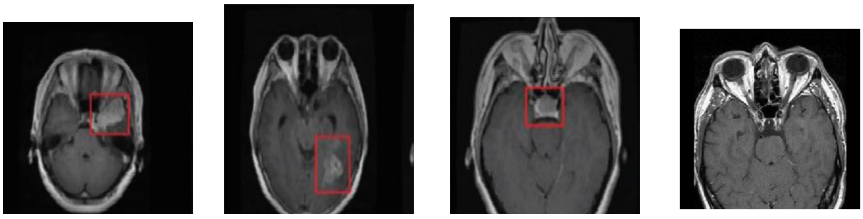
Table 1: Dataset distribution for each class.

Class	Number of Images
Glioma	1003
Meningioma	981
Pituitary	959
No tumor	561
Total	**3504**

pabilities by including MRI scans in three essential anatomical planes which are axial, coronal, and sagittal. The model analyzes tumor and non-tumor groups in all three anatomical planes to display their spatial and perspective differences regarding diagnosis. The figure provided in Figure 2 displays axial MRI slice examples from (a) meningioma to (b) glioma through (c) pituitary tumor up to (d) no tumor. Agital views serve the best purpose for tumor measurements and lateral brain structure distribution assessments. The graph in Figure 3 shows coronal sections of identical groups. The evaluation of vertical tumor progression and the assessment of midline structure involvement is possible through coronal imaging. Additional visual perspectives through these images enhance diagnostic assessment accuracy during manual and automated analysis procedures. Figure 4 shows sagittal views that enable doctors to assess the depths of tumors and their impact on structures involving the corpus callosum or brainstem. Adding sagittal slices to MRI scans makes the data collection more comprehensive because they provide a complete volumetric view of tumor structure. Combining tumor and no-tumor cases with three different tumors enables better clinical applications for the diagnostic model. Training the proposed system with pathological and normal examples from different orientations makes it more appropriate for clinical diagnostic use.

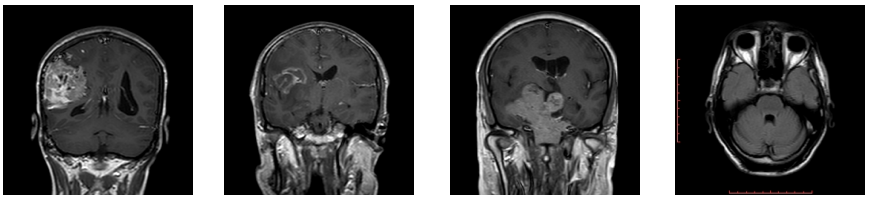
DOI: https://doi.org/10.54216/JAIM.090204

Received: December 18, 2024 Revised: February 25, 2025 Accepted: May 02, 2025



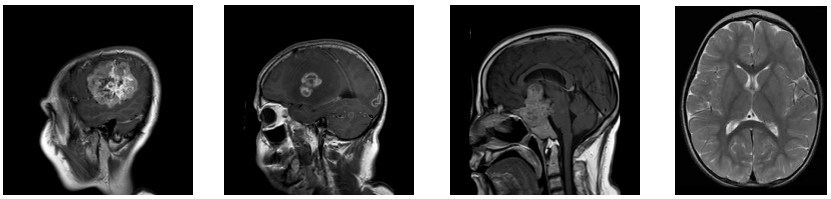
(a) Meningioma (b) Glioma (c) Pituitary tumor (d) No tumor

Figure 2: Sample of axial MRI from the dataset: (a) Meningioma, (b) Glioma, (c) Pituitary tumor, (d) No tumor



(a) Meningioma (b) Glioma (c) Pituitary tumor (d) No tumor

Figure 3: Sample of coronal MRI from the dataset: (a) Meningioma, (b) Glioma, (c) Pituitary tumor, (d) No tumor.



(a) Meningioma (b) Glioma (c) Pituitary tumor (d) No tumor

Figure 4: Sample of sagittal MRI from the dataset: (a) Meningioma, (b) Glioma, (c) Pituitary tumor, (d) No tumor.

2.2 Prepossessing and Data Augmentation

The preprocessing procedure serves as a vital step for data preparation because it helps models learn meaningful representations instead of focusing on irrelevant patterns such as black pixel borders. The preprocessing process creates standardized data that enhances both generalization potential and classification accuracy for the model. The preprocessing process assumes special importance in brain MRI since medical image data regularly extends beyond 8-bit pixel range [0, 255] with potential negative intensity values subsequent to input scan reading. The preprocessing approach begins with **min-max normalization** as its first operation. Through this method raw pixel values receive linear transformation to create a standardized value range from 0 to 1 thus making the data consistent throughout all samples. The normalization process follows this equation for its calculation:

$$Y_i = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{1}$$

The normalized output is represented by Y_i while X stands for the input pixel value and $X_{\min}$ and $X_{\max}$ represent the minimum and maximum pixel values contained in the database. A subsequent histogram equalization process modifies image contrast after normalization takes place. This enhancement technique spreads image intensity values to fill the complete intensity spectrum to enhance the distinction of medical structural and pathological features during analysis. A uniform dimension of **256 × 256 pixels** serves to normalize all brain MRI scans for entering the deep learning model. An appropriate dimensional adjustment enables the system to work seamlessly with different computational models while improving data processing speed. The models strength gets enhanced through integration with proprietary data augmentation methods which also help minimize overfitting. The augmented dataset gets synthetically expanded through modified versions of original images that represent probable variations from actual use scenarios. The model learns better generalization and achieves faster training convergence through data augmentation because it exposes the model to

DOI: https://doi.org/10.54216/JAIM.090204

Received: December 18, 2024 Revised: February 25, 2025 Accepted: May 02, 2025

diverse data during learning. Scientists in this research employed the following particular data augmentation procedures:

- **Rotation** by up to 40 degrees,
- **Width shift** and **height shift** by 8% of the image dimensions,
- **Shear transformations** with a shear intensity of 0.3,
- **Zooming** within a 10% range.

These parameters are summarized in Table 2. The augmentation techniques raised the training dataset from

Table 2: Augmentation methods.

Method	Parameters
Rotation	40 degrees
Width Shift	0.08
Height Shift	0.08
Shear Range	0.3
Zoom Range	0.1

3966 images which substantially improved its training data diversity. A single MRI scan produced different augmented images through the applied transformations as shown in Figure 5. The augmentation process creates simulated examples which reproduce clinical variations while making the model more suitable for clinical deployment.

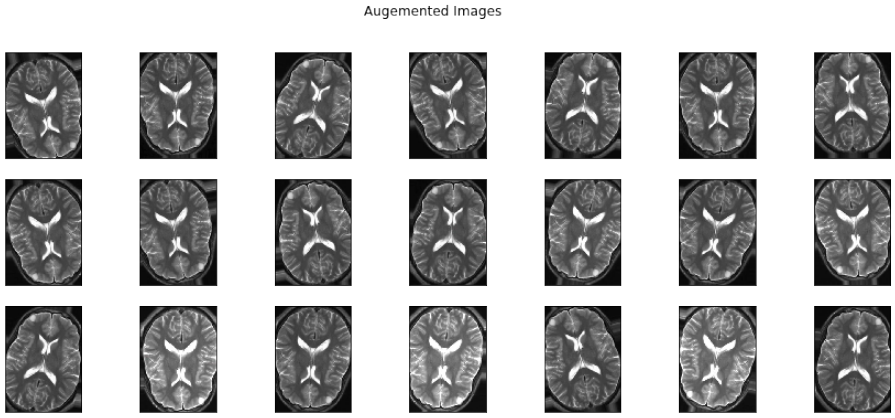


Figure 5: Example output images after augmentation by applied width shift, rotation range, height shift, rescale, brightness range, horizontal flip, and vertical flip.

Figure 6 shows our proposed hybrid model, which enhances feature extraction through deep learning combined with traditional machine learning methods for classification. Our methodology uses VGG16 as a neural network to extract features from images while implementing an SVM classifier as the standard softmax classification head alternative. This combined strategy allows VGG16 to extract abstract high-level features and enables SVMs to utilize their firm boundaries for decision-making. After removing the softmax layer from VGG16, the implementation extracts features from the penultimate fully connected layer. SVM is a prediction tool for inputting the extracted features during training. This paper evaluates SVM classification performance through four alternative machine learning models, namely Extreme Gradient Boosting (XGBoost) [29], Adaptive Boosting (AdaBoost), Decision Tree [30], and K-Nearest Neighbors (KNN). These specific models have gained popularity for their demonstrated excellence in classification while frequently used within medical imaging pipelines. SVM yields the best results when used with VGG16 based on the results discussed in the outcomes section. Our proposed architecture includes an SVM classifier with the `rbf` kernel because this kernel performs effectively when classifying non-linear problems. The RBF kernel requires tuning two critical

DOI: https://doi.org/10.54216/JAIM.090204

Received: December 18, 2024 Revised: February 25, 2025 Accepted: May 02, 2025

parameters, the C cost parameter and the γ gamma parameter. A single training example receives its influence level from the gamma parameter setting. The distance range of influence in the system depends on the gamma value, where lower values indicate a longer reach while higher values indicate a closer range. The mathematical definition of RBF kernel appears as follows:

$$M\left(z_i, z_j\right) = \exp^{\left(-\gamma \| z_i - z_j \|^2\right)} \tag{2}$$

The relationship utilizes sample point z_i together with support vector z_j. The exponential calculation allows the model to transform data into higher-dimensional space, making linear separability more possible. By adjusting γ and C parameters, SVM obtains an ideal decision boundary that transforms effectively on unknown data samples. Figure 6 visually outlines the entire pipeline. Deep features in VGG16 extract features from input MRI scans. The model feeds the extracted deep features to the SVM classifier to predict subjects diagnosed with glioma, meningioma pituitary tumor or no tumor. This system unites CNN-based representation capabilities with SVM operational strength to boost medical imaging diagnostics.

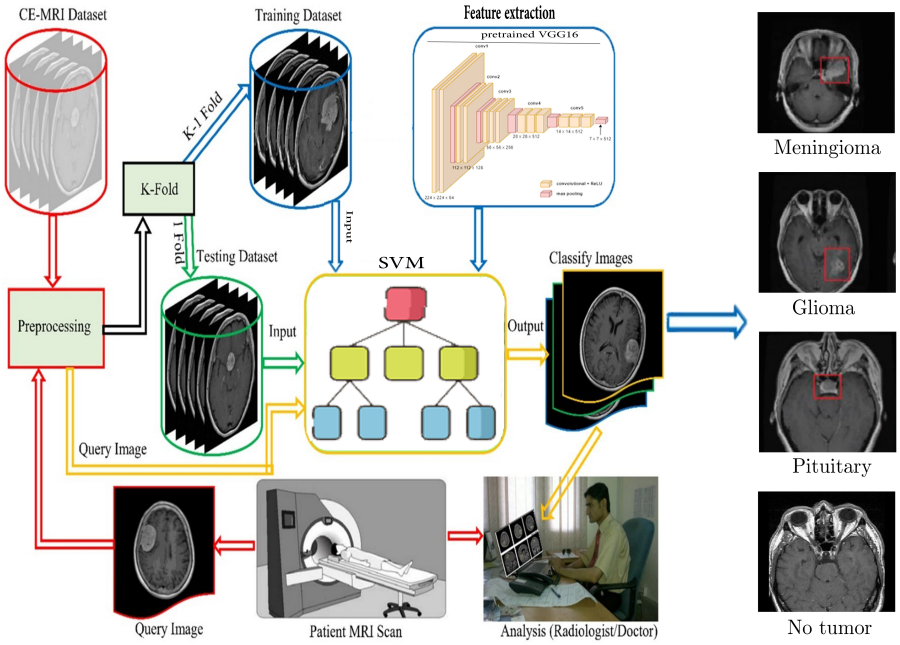


Figure 6: Proposed Architecture Model: VGG16 used as a feature extractor followed by SVM for final classification.

2.3 Classification Head

Studies before this utilized hybrid Convolutional Neural Network (CNN) models in brain tumor classification by connecting different CNN structures or uniting CNN architecture with additional deep learning structures. Hybrid CNN approaches generate minimal performance enhancements but produce complicated system implementations. Advanced neural network structures need ample computational power and extensive processing time for training and validation procedures. Detailed network architecture poses interpretation challenges to clinicians who need to understand their models' decision-making processes when using this technology clinically. A streamlined proposal uses the VGG16 model as its backbone to achieve equivalent or advanced classification outcomes to the existing system. Our model links this VGG16 CNN with an SVM classifier, simplifying design complexity while maintaining performance qualities. The basic system design shortens learning and testing durations while boosting the classification process's interpretability features and transparency elements. The brain MRI image classification system must sort these images into their tumor types among glioma, meningioma or pituitary tumor. We utilize deep features obtained from the VGG16 network to feed them into an SVM classifier. The SVM is a supervised machine learning technique whose main strength is performing binary and multiclass classification tasks. Strong generalization performance, robust protection against overfitting, and efficient memory and computing power are the principal benefits of this system. The

DOI: https://doi.org/10.54216/JAIM.090204

Received: December 18, 2024 Revised: February 25, 2025 Accepted: May 02, 2025

use of SVM for image-based medical diagnosis becomes particularly successful because it generates optimal separating hyperplanes in high-dimensional feature spaces. The radial basis function (RBF) kernel within the kernel trick enables the system to operate effectively on data that cannot be separated linearly and simultaneously resists noise interference. The combination of qualities in SVM establishes it as a dependable solution for medical image classifications when interpretability is combined with precision and generalization matter. The proposed method delivers high diagnostic accuracy through simplified operations, making it more effective for practical medical systems requiring excellent performance and efficient computing capacity.

2.4 Evaluation Metrics

When working with imbalanced datasets, a classification algorithm should use diverse evaluation metrics to assess its performance between different classes. In this research, we adopt the **F1-score** as the principal evaluation metric due to its capability to deliver a balanced assessment of model performance when dealing with classes of unequal representation. This is particularly important in medical diagnostics, where the minority class (e.g., patients with a disease) typically holds critical clinical relevance. The classification metrics include the F1-score, accuracy, Precision, sensitivity, and specificity. These various metrics provide separate insights about model behavior, producing a complete understanding of classification performance. Sensitivity measures how well the model identifies disease cases among all positive results. The model's ability to correctly identify tumor-bearing patients is expressed through sensitivity during this study. Research requires high sensitivity rates to reduce incorrect healthy classifications of tumor-bearing patients. The model displays its capability to detect healthy persons accurately through specific measurements. Specificity performance in diagnostic models helps prevent medical misdiagnoses by lowering the frequency of incorrect positive test results, thus protecting patients from pointless distress and treatments. The evaluation method for Precision counts accurate optimistic predictions among all cases the system designated as optimistic. The measure determines the reliability of tumor detection predictions made by the model. A high Precision rate ensures few false positive results, essential for medical scenarios requiring extra procedures for each incorrect alarm. The **F1-score** synthesizes Precision and recall into a single performance metric by computing their harmonic mean. The score is an important metric in unbalanced datasets because it helps achieve a fair balance between leaving out actual cases and misidentifying well patients. The F1-score gives a better assessment of model performance in minority-dominated situations because it avoids misleading results that accuracy produces in such cases. Accuracy is a mainstream metric but evaluates the complete number of appropriately identified instances. The measure counts correct classifications alongside false cases about the complete dataset to determine performance. Accuracy measurement gives overall system performance knowledge but fails to show classification accuracy for vital categories in unbalanced datasets. The evaluation metrics are formally represented through Table 3.

Table 3: Evaluation metrics and their mathematical definitions.

Metric	**Formula**
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$
Sensitivity (Recall)	$\frac{TP}{TP+FN}$
Specificity	$\frac{TN}{TN+FP}$
Precision	$\frac{TP}{TP+FP}$
F1-Score	$2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$

The actual positive value is TP while TN represents actual negative values, and FP shows false positive results along with FN indicating false negative values. All derived evaluation metrics use these essential components to assess diagnostic model performance. Medical diagnostic professionals strongly consider these statistical metrics as vital performance indicators. Both correctly identifying patients with diseases and correctly identifying healthy individuals without disease need to become standard in making effective and safe clinical decisions. Descriptive value assessments require multiple criteria over a single measure to achieve holistic model performance analysis at the real-world application level.

DOI: https://doi.org/10.54216/JAIM.090204

Received: December 18, 2024 Revised: February 25, 2025 Accepted: May 02, 2025

2.5 Implementation Details

The Google Colab platform using the TensorFlow library enabled the implementation and training of the proposed model with GPU processing capability. The authors built the SVM classification head through Scikit-learn. The training operation utilized 90% of the available data, followed by testing done with the remaining 10%. A summary of the model hyperparameters exists in Table 4. Different kernel functions and gamma value combinations were tried to determine the optimum model setting. The graphics in Figures 7 and 8 demonstrate

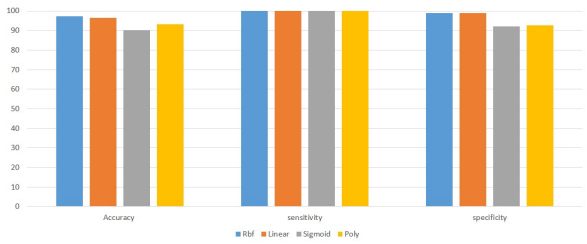


Figure 7: Performance of the proposed model across different kernel functions.

how changing kernels affects model results when gamma is set to 100. As depicted in these figures, the Radial Basis Function (RBF) kernel delivered the highest classification performance, achieving a test accuracy of 97.2%.

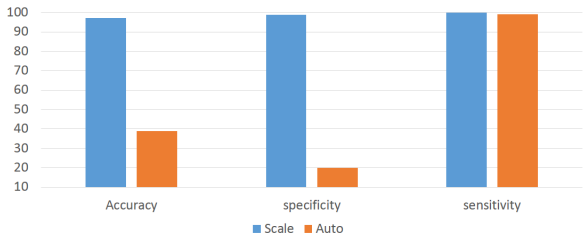


Figure 8: Tuning of gamma parameter at C=100 for optimal model performance.

Table 4: Hyper-parameters of proposed model for classification

Hyperparameter	Value
C	100
gamma	scale
kernel	rbf
Random state	123

3 Result

The proposed method uses Table 5 to assess model performance through important evaluation metrics. The model assessment utilizes accuracy alongside sensitivity, specificity, and F1-score to give a complete overview

DOI: https://doi.org/10.54216/JAIM.090204

Received: December 18, 2024 Revised: February 25, 2025 Accepted: May 02, 2025

Table 5: Proposed model evaluation against other ML classifiers. "P_t" is processing time.

	Evaluation metrics				
Model	Accuracy	Sensitivity	Specificity	F1-score	P_t(sec)
Proposed model	97.2%	100%	98.8%	97%	319.7
VGG16+XGBOOST	92.6%	99%	95.5%	93%	1169.3
VGG16+AdaBoost	73.2%	89.9%	94.6%	73%	236.5
VGG16+Decision Tree	71.5%	95%	89.8%	72%	24.7
VGG16+KNeighbors	88.4%	97.8%	95.4%	88%	1125

of forecasting capability across multiple tumor classes. The implemented model delivers superior performance than conventional machine learning models in multiple vital evaluation metrics. The VGG16-based feature extraction with SVM classification enables the hybrid system to achieve 97.2% test accuracy, complete sensitivity at 100%, and a specificity rate of 98.8%. The model demonstrates excellent performance because it combines high recall rates with precision values for optimal evaluation results in medical environments. The performance results of the proposed model for different data splits appear in Table 6. The model upholds high-performance standards throughout the training data reduction from 90% to 50%, proving its robustness and generalization abilities. The model reaches optimal performance with a 90/10 ratio but demonstrates satisfactory accuracy (97.2%) and F1-score (97%). This level of success continues even when the proportion is reduced to 50/50, where performance remains at 93.7% accuracy together with an F1-score of 94%. The model shows effective operation based on distinct training data availability levels in these results.

Table 6: Performance results with different training/testing ratios on the dataset.

	Evaluation metrics			
Training/Testing	Accuracy	Sensitivity	Specificity	F1-score
90/10	97.2%	100%	98.8%	97%
80/20	94.9%	99%	93.9%	95%
70/30	94.5%	99.6%	94.6%	95%
60/40	94%	99.5%	94.9%	94%
50/50	93.7%	98.8%	95.9%	94%

Figure 9 represents the K-Fold cross-validation process utilized throughout the training phase. The technique separates the data into 10 equivalent sections by allowing the model to learn from nine parts while validating using the remaining segment. Different partitioning is used ten times before averaging the final performance metrics. Testing every data point multiple times with this method decreases the likelihood of overfitting.

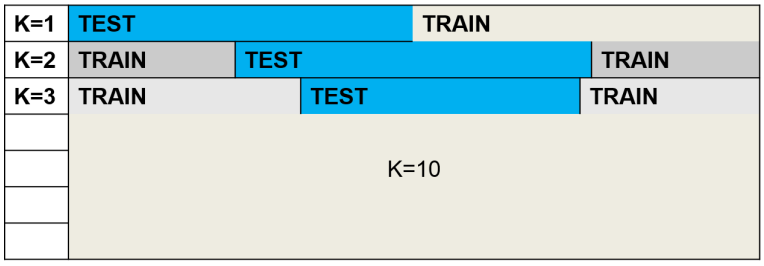


Figure 9: Technique cross-validation with a K-fold.

Table 7 presents a performance benchmark between the proposed model and several advanced pre-trained models that function as feature extractors. The proposed model delivers superior performance than the other configurations including DenseNet121, MobileNetV2, VGG19, InceptionV3, and Xception. The combination of VGG16 and SVM reaches 97.2% accuracy while performing within 319.7 seconds of processing time, demonstrating its superior characteristics over DenseNet121+SVM. SVM functions as a classification module that receives inputs from VGG16 features. A proposed model using VGG16 features as extraction backbone reached 97.2% accuracy and 100% sensitivity with 98.8% specificity alongside 97% F1-score. The model needed 319.7 seconds (P_t) to process, representing an acceptable timing period because of its superior outcome. Our proposed model outperformed MobileNetV2 and DenseNet121 in terms of balance between performance and efficiency because it reached accuracy rates of 94.4% and 95.9% while requiring 636.2s and 384.6s

DOI: https://doi.org/10.54216/JAIM.090204

Received: December 18, 2024 Revised: February 25, 2025 Accepted: May 02, 2025

Table 7: Performance results obtained with 90/10 training/testing ratio using different pre-trained models(feature extractor) and SVM(classifier).

	Evaluation metrics				
Models	Accuracy	Sensitivity	Specificity	F1-score	P_t(sec)
InceptionV3+SVM	83.8%	98%	91.7%	83%	357.7
MobileNetV2+SVM	94.4%	99%	96.7%	94%	636.2
Xception+SVM	92.4%	96%	94.5%	92%	978.1
VGG19+SVM	94.7%	100%	98.7%	95%	113.6
DenseNet121+SVM	95.9%	98.9%	98.9%	96%	384.6
Proposed model	97.2%	100%	98.8%	97%	319.7

in computation time. Combining InceptionV3 with SVM and Xception with SVM achieved lower F1-scores (83% and 92% respectively) than our proposed method because these models demonstrated reduced classification capabilities. The visual analysis presented in Figure 10 showcases confusion matrices that compare the proposed model results and four different alternative classification models, which include VGG16+XGBoost, VGG16+KNeighbors, VGG16+Decision Tree, and VGG16+AdaBoost. The matrices demonstrate a comprehensive assessment of model predictive excellence by displaying the correct classification of glioma, meningioma, pituitary tumor and no tumor. The proposed model displays distinguished performance according to Figure 10a by showing true positives as dominant elements on the diagonal while maintaining minimal to no erroneous classifications off the diagonal. The model demonstrates an outstanding ability to differentiate tumor types and detect tumors' absence. The XGBoost-based model generates adequate outcomes across most scenarios (Figure 10b) despite showing dispersed misidentification of glioma cases. Gliomas and meningiomas exhibit high levels of confusion when subjected to the KNeighbors classifier (Figure 10c) even though this model demonstrates effective pituitary tumor classification. Separate from other models, the Decision Tree model produces inferior results because it creates multiple incorrect categorizations for different brain tumor types, particularly between no-tumor and meningioma cases (Figure 10d). According to Figure 10e, the AdaBoost classifier performs moderately with pituitary tumors. However, it demonstrates a significant decrease in accuracy when detecting gliomas and meningiomas, thus creating potential risks during medical diagnosis through potential negative results. A direct comparison between the proposed hybrid model confusion matrix and the baseline VGG16 model confusion matrix using a softmax classifier can be found in Figure 11. The softmax-based classifier proofs to be practical yet produces more misclassifications than the hybrid model proposed in this study (Figure 11b). portability and diagnosis inaccuracies are more pronounced when detecting glioma and meningioma tumors. The data in Table 8 illustrates that the proposed model outperforms the baseline by achieving a test accuracy of 94.4% and 97.2% , respectively. Diagnostic reliability improves significantly through the proposed model because it shows 100% sensitivity and 98.8% specificity. The proposed model operates with better computational efficiency because it needs 319.7 seconds for inference rather than the 734.5 seconds needed by VGG16+Softmax. This research confirms the effectiveness of uniting support vector machine (SVM) classification heads with deep feature extractors because it enhances diagnostic precision and computational speed.

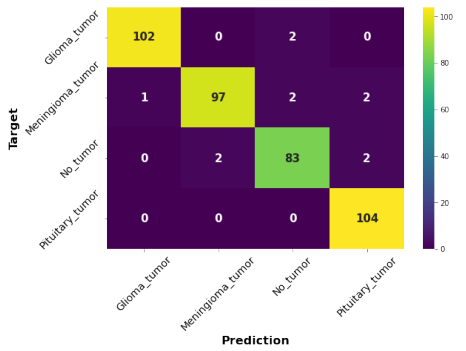
Table 8: Comparison between the proposed model and the baseline VGG16 model with softmax classifier.

	Evaluation Metrics				
Model	Train Accuracy	Test Accuracy	Sensitivity	Specificity	P_t (sec)
Proposed model	100%	97.2%	100%	98.8%	319.7
Baseline VGG16	99.3%	94.4%	99%	98%	734.5

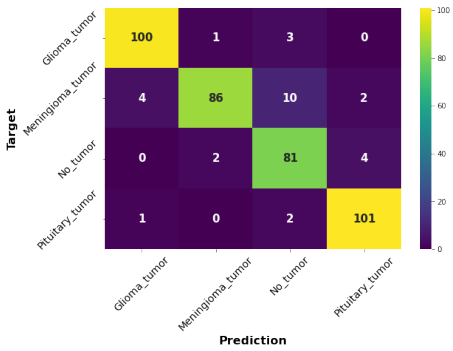
The ROC (Receiver Operating Characteristic) curves of Figure 12 present valuable assessments for measuring classification models through their separation capabilities between positive and negative instances for all classes. ROC and receiver operating characteristic (ROC) curves represent the sensitive detection rate against the false alarm rate in a two-dimensional graph that illustrates model operating thresholds. The proposed model (Figure 12a) exhibits exceptional class discriminating ability as its AUC (Area Under the Curve) reaches 1.0 for your pathologic classes of glioma, meningioma, and pituitary tumor and no tumor class. The model can differentiate various classes at an advanced level while controlling ambiguity in classification boundaries. The near-perfect ROC curve indicates minimal distribution overlap between positive and negative classes, which is vital for medical diagnostics, particularly in early-stage tumor detection of subtle features. The ROC curves of other models deviate significantly from the perfect diagonal pattern. The performance of VGG16+AdaBoost

DOI: https://doi.org/10.54216/JAIM.090204

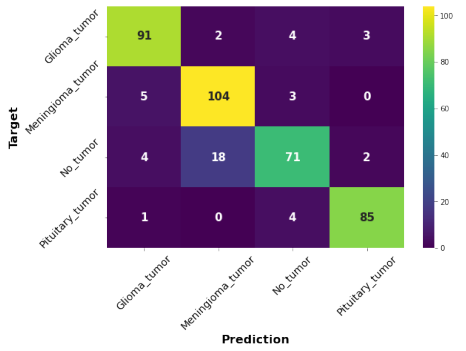
Received: December 18, 2024 Revised: February 25, 2025 Accepted: May 02, 2025



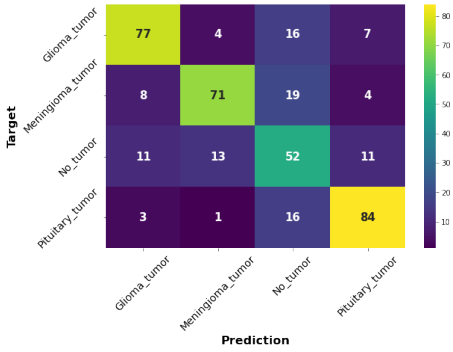
(a) Proposed model



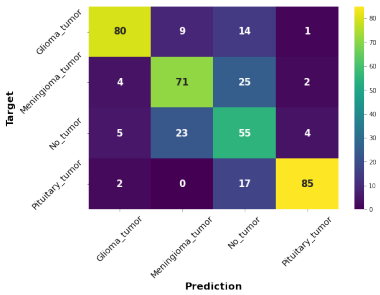
(b) VGG16 + XGBoost



(c) VGG16 + KNeighbors



(d) VGG16 + Decision Tree



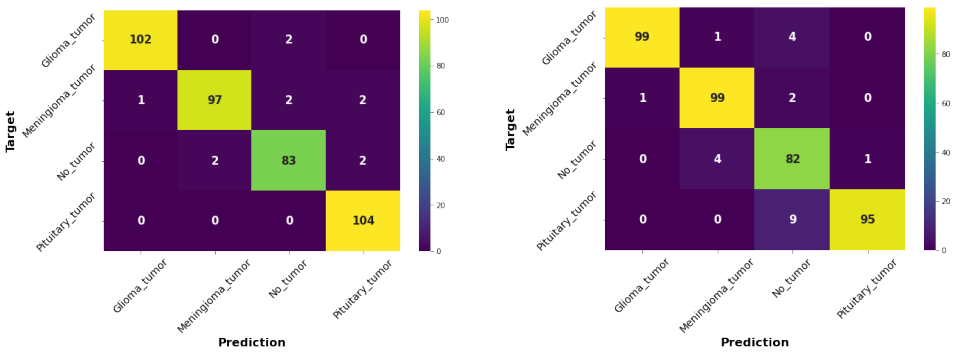
(e) VGG16 + AdaBoost

Figure 10: Confusion matrices for the (a) proposed model, (b) VGG16+XGBoost, (c) VGG16+KNeighbors, (d) VGG16+Decision Tree, and (e) VGG16+AdaBoost. These matrices illustrate classification performance across tumor types.

(Figure 12e) and VGG16+Decision Tree (Figure 12d) reflects in reduced AUC metrics as well as flatter curves which result in diminished class discrimination abilities. Clinical scenarios prefer models that produce fewer false positives and negatives while detecting cancer diagnoses. The performance of VGG16+KNeighbors classifies as moderate, although this model trails below the overall results achieved by the proposed model (Figure 12c). Implementing VGG16+XGBoost (Figure 12b) delivers impressive results but cannot replicate the same consistent class performance demonstrated by the introduced model. The proposed hybrid model has passed additional tests that prove its strong clinical use in medical situations that require precise and accurate diagnosis. The proposed model achieved results which are displayed according to class performance metrics in Table 9. The examination results demonstrate reliable high performance throughout all four diagnostic groups. In medical diagnosis, the proposed model demonstrated 100% perfect recall to detect pituitary tumors, thus ensuring complete detection of tumor cases. In predicting gliomas and meningiomas, the model achieved a 95% recall rate, and no incorrect diagnoses were found when measuring with 98% recall followed by 99% precision accuracy. The model shows excellent performance in diagnosing tumors, whether positive or negative. The measured F1-scores demonstrated acceptable reliability for clinical use since they ranged between 95% to 99% while ensuring both classes maintained their sensitivity and precision levels. The performance metrics

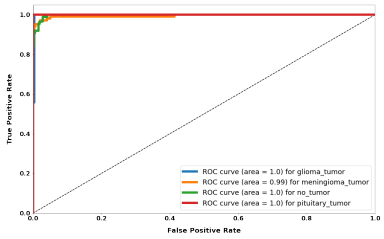
DOI: https://doi.org/10.54216/JAIM.090204

Received: December 18, 2024 Revised: February 25, 2025 Accepted: May 02, 2025

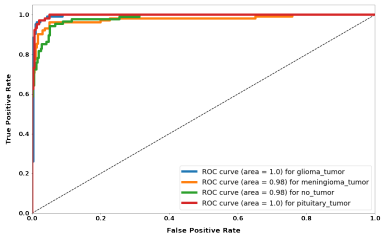


(a) (b)

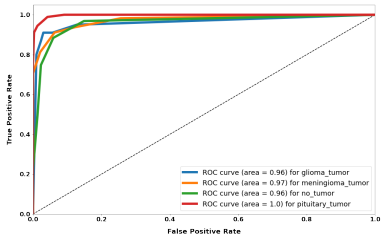
Figure 11: Confusion matrices for classification using (a) the proposed model and (b) the baseline VGG16 model with softmax classifier.



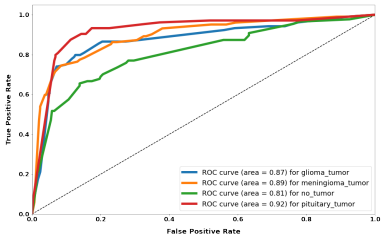
(a) Proposed model



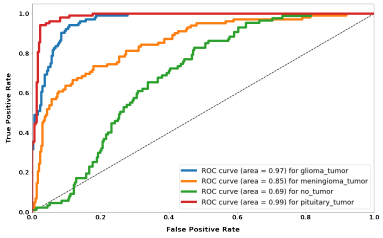
(b) VGG16 + XGBoost



(c) VGG16 + KNeighbors



(d) VGG16 + Decision Tree



(e) VGG16 + AdaBoost

Figure 12: ROC curves for different classification models. The proposed model (a) demonstrates the highest AUC values, indicating superior class discrimination capability. In contrast, models (b) through (e) show relatively lower performance distinguishing tumor types.

for the VGG16+XGBoost classifier appear in Table 10. XGBoost predictions maintain reliable accuracy rates, but its glioma diagnosis has reduced performance and generates excessive wrong pessimistic results with an 84% recall rate. The practice of recalling critical medical emergencies should not present lower diagnostic rates.

Results from this approach reached high stability levels but missed the targets set by the proposed method. Table 11 shows the metrics for VGG16+AdaBoost. The model demonstrates poorly performing results mostly among meningioma and glioma tumor cases. The F1-score for meningioma reaches only 56%, and glioma yields 69%. The results demonstrate the difficulty of achieving sufficient false positive and false adverse

DOI: https://doi.org/10.54216/JAIM.090204

Received: December 18, 2024 Revised: February 25, 2025 Accepted: May 02, 2025

Table 9: Performance metrics using proposed model

	Evaluation metrics		
Class	Precision	Recall	F1-score
No tumor	99%	98%	99%
Glioma	98%	95%	97%
Meningioma	95%	95%	95%
Pituitary	96%	100%	98%

Table 10: Performance metrics using VGG16+XGBOOST classifier

	Evaluation metrics		
Class	Precision	Recall	F1-score
No tumor	95%	96%	96%
Glioma	97%	84%	90%
Meningioma	94%	93%	89%
Pituitary	94%	97%	96%

prevention by this model. The precision value of 92% achieved by the pituitary tumor class stands out, but general performance issues prevent this model from reaching practical reliability. The classification metrics for the VGG16+Decision Tree model are presented in Table 12.

Table 11: Performance metrics using VGG16+AdaBoost classifier

	Evaluation metrics		
Class	Precision	Recall	F1-score
No tumor	88%	77%	82%
Glioma	69%	70%	69%
Meningioma	50%	63%	56%
Pituitary	92%	82%	87%

The model achieves a modest outcome by reaching an F1-score of 80% for identifying pituitary tumors. The diagnosis threshold for medical-grade identification lies outside the acceptable range for meningioma, along with every other tumor category. The results raise significant concerns because patients have a 70% chance of glioma and a 60% chance of meningioma misdiagnosis. The Decision Tree model operates quickly, although its unreliable performance makes it unsuitable for practical applications. Finally, Table 13 outlines the perfor-

Table 12: Performance metrics using VGG16+Decision Tree classifier

	Evaluation metrics		
Class	Precision	Recall	F1-score
No tumor	78%	74%	76%
Glioma	80%	70%	74%
Meningioma	50%	60%	55%
Pituitary	79%	81%	80%

mance of the VGG16+KNeighbors classifier. This model performs better than AdaBoost and Decision Tree, with relatively strong results for pituitary (94% F1) and no tumor (91% F1). However, glioma prediction is slightly inferior to the proposed model (88% F1 vs. 97%), and meningioma still presents difficulty, achieving only 80% F1-score. While better balanced than some models, it does not match our proposed method's high accuracy and precision.

The VGG16+SVM hybrid model has established its superiority through complete testing of various machine learning setups. Testing shows that this model produces superior results than standard classifiers and deep learning-based approaches across every essential measurement such as accuracy and F1-score and sensitivity and specificity. A test accuracy standing at 97.2% , sensitivity at 100%, and specificity at 98.8% make the proposed method ideal for medical diagnostic purposes requiring minimizing false negatives and false positives. The model demonstrates great generalization ability along with robust behavior because it continues

DOI: https://doi.org/10.54216/JAIM.090204

Received: December 18, 2024 Revised: February 25, 2025 Accepted: May 02, 2025

Table 13: Performance metrics using VGG16+KNeighbors classifier

	Evaluation metrics		
Class	Precision	Recall	F1-score
No tumor	90%	91%	91%
Glioma	84%	93%	88%
Meningioma	87%	75%	80%
Pituitary	94%	94%	94%

to perform well under different training-to-testing ratios documented through laboratory experiments. The model demonstrates dependable clinical readiness because it maintains equal precision and recall rates across every tumor category. The model achieves exceptional computational performance as its main advantage. The proposed model achieves classification accuracy at a moderate computational cost even though MobileNetV2 and DenseNet121 need extensive processing times for similar accuracy levels. This proposed model provides superior accuracy alongside moderate computational expenses that total 319.7 seconds, thus making it suitable for near-real-time diagnostic settings. The proposed model demonstrates its superiority by analyzing the confusion matrix and ROC curve. The visual representations of this approach show superior performance in detecting tumors and clear separation between normal and diseased tissues. The combination of SVM with the hybrid system achieves higher accuracy than baseline VGG16+Softmax and operates at speeds exceeding 50% faster. The VGG16+SVM model delivers extensive diagnostic improvements that enhance performance accuracy, system efficiency, and reliability when processing different types of tumors. This tool presents an advanced diagnostic instrument that can benefit computer-aided diagnosis systems because it offers balanced scoring with low error rates for brain tumor diagnosis in clinical practice.

4 Conclusions and future work

Our research implements a classification model that combines VGG16 as its deep feature extraction stage with a Support Vector Machine (SVM) operating as a classification header. The VGG16 network passes its learned deep features to an SVM classifier, completing the decision-making functions. This combination leverages the powerful feature representation capabilities of deep learning with the robust classification ability of traditional machine learning. The proposed model demonstrates successful detection of brain tumors with 97.2% classification accuracy and high reliability through our results. Our approach demonstrates the feasibility of integrating into existing clinical practice to deliver swift tumor diagnosis that enhances front-line healthcare support. Our approach for future work includes two methods: gathering a substantially larger dataset with more diversity and building high-quality synthetic data through GAN technology. Due to its improved research capabilities, the expanded dataset makes better subtype performance assessments possible. Additional research should concentrate primarily on developing interpretability features of the model until it achieves an acceptable standard. The current deep learning-based system operates as a black box due to its unexplainable workings, which prevents users from understanding individual prediction reasoning. The interpretability of predictive decisions will benefit from visualization tools, which include methods such as saliency maps and Grad-CAM features with attention-based mechanisms to produce supportive explanations from the model. The medical community and patients need complete explanations from AI tools that increase accuracy and provide clinical dependability and explanation for their outputs. Future advancements in this research should combine data variety with clear explanations of AI decisions to produce better medical diagnostic AI solutions that clinicians and patients can accept.

References

[1] James H Park, Adrian Lopez Garcia de Lomana, Diego M Marzese, Tiffany Juarez, Abdullah Feroze, Parvinder Hothi, Charles Cobbs, Anoop P Patel, Santosh Kesari, Sui Huang, et al. A systems approach to brain tumor treatment. *Cancers*, 13(13):3152, 2021.

DOI: https://doi.org/10.54216/JAIM.090204

Received: December 18, 2024 Revised: February 25, 2025 Accepted: May 02, 2025

[2] Karrar Neamah, Farhan Mohamed, Myasar Mundher Adnan, Tanzila Saba, Saeed Ali Bahaj, Karrar Abdulameer Kadhim, and Amjad Rehman Khan. Brain tumor classification and detection based dl models: a systematic review. *IEEE Access*, 12:2517–2542, 2023.

[3] Sirvan Khalighi, Kartik Reddy, Abhishek Midya, Krunal Balvantbhai Pandav, Anant Madabhushi, and Malak Abedalthagafi. Artificial intelligence in neuro-oncology: advances and challenges in brain tumor diagnosis, prognosis, and precision treatment. *NPJ precision oncology*, 8(1):80, 2024.

[4] Muhammed Yildirim, Emine Cengil, Yeşim Eroglu, and Ahmet Cinar. Detection and classification of glioma, meningioma, pituitary tumor, and normal in brain magnetic resonance imaging using deep learning-based hybrid model. *Iran journal of computer science*, 6(4):455–464, 2023.

[5] Cara Marie Rogers, Pamela S Jones, and Jeffrey S Weinberg. Intraoperative mri for brain tumors. *Journal of neuro-oncology*, 151:479–490, 2021.

[6] Ardhendu Sekhar, Soumen Biswas, Ranjay Hazra, Arun Kumar Sunaniya, Amrit Mukherjee, and Lixia Yang. Brain tumor classification using fine-tuned googlenet features and machine learning algorithms: Iomt enabled cad system. *IEEE journal of biomedical and health informatics*, 26(3):983–991, 2021.

[7] Isselmou Abd El Kader, Guizhi Xu, Zhang Shuai, Sani Saminu, Imran Javaid, and Isah Salim Ahmad. Differential deep convolutional neural network model for brain tumor classification. *Brain Sciences*, 11(3):352, 2021.

[8] Parvathy Jyothi and A Robert Singh. Deep learning models and traditional automated techniques for brain tumor segmentation in mri: a review. *Artificial intelligence review*, 56(4):2923–2969, 2023.

[9] Asaf Raza, Huma Ayub, Javed Ali Khan, Ijaz Ahmad, Ahmed S. Salama, Yousef Ibrahim Daradkeh, Danish Javeed, Ateeq Ur Rehman, and Habib Hamam. A hybrid deep learning-based approach for brain tumor classification. *Electronics*, 11(7):1146, 2022.

[10] Shagun Sharma, Kalpna Guleria, Sunita Tiwari, and Sushil Kumar. A deep learning based convolutional neural network model with vgg16 feature extractor for the detection of alzheimer disease using mri scans. *Measurement: Sensors*, 24:100506, 2022.

[11] Alice Oh, Inyoung Noh, Jian Choo, Jihoo Lee, Justin Park, Kate Hwang, Sanghyeon Kim, and Soo Min Oh. Machine learning approach to brain tumor detection and classification. *arXiv preprint arXiv:2410.12692*, 2024.

[12] Rohith Krishnan, PG Gokul, Gopikrishnan Sujith, T Anjali, and S Abhishek. Enhancing brain tumor diagnosis: A cnn-based multi-class classification approach. In *2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)*, volume 2, pages 1–6. IEEE, 2024.

[13] Surendra Reddy Vinta, Phaneendra Varma Chintalapati, Gurujukota Ramesh Babu, Rajyalakshmi Tamma, and Gunupudi Sai Chaitanya Kumar. Edlnet: ensemble deep learning network model for automatic brain tumor classification and segmentation. *Journal of Biomolecular Structure and Dynamics*, pages 1–13, 2024.

[14] A Lakshmi, Manjunathan Alagarsamy, A Anbarasa Pandian, and Dinesh Paramathi Mani. Evolutionary gravitational neocognitron neural network optimized with marine predators optimization algorithm for mri brain tumor classification. *Electromagnetic Biology and Medicine*, 43(1-2):1–18, 2024.

[15] Tariq Sadad, Amjad Rehman, Asim Munir, Tanzila Saba, Usman Tariq, Noor Ayesha, and Rashid Abbasi. Brain tumor detection and multi-classification using advanced deep learning techniques. *Microscopy research and technique*, 84(6):1296–1308, 2021.

[16] Hossam H Sultan, Nancy M Salem, and Walid Al-Atabany. Multi-classification of brain tumor images using deep neural network. *IEEE access*, 7:69215–69225, 2019.

[17] Saravanan Srinivasan, Divya Francis, Sandeep Kumar Mathivanan, Hariharan Rajadurai, Basu Dev Shivahare, and Mohd Asif Shah. A hybrid deep cnn model for brain tumor image multi-classification. *BMC Medical Imaging*, 24(1):21, 2024.

DOI: https://doi.org/10.54216/JAIM.090204

Received: December 18, 2024 Revised: February 25, 2025 Accepted: May 02, 2025

[18] Emrah Irmak. Multi-classification of brain tumor mri images using deep convolutional neural network with fully optimized framework. *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, 45(3):1015–1036, 2021.

[19] Sohaib Asif, Ming Zhao, Fengxiao Tang, and Yusen Zhu. An enhanced deep learning method for multi-class brain tumor classification using deep transfer learning. *Multimedia Tools and Applications*, 82(20):31709–31736, 2023.

[20] Bentahar Heythem, Mohamad Djerioui, Tawfiq Beghriche, Azzedine Zerguine, and Azeddine Beghdadi. Customized cnn for multi-class classification of brain tumor based on mri images. *Arabian Journal for Science and Engineering*, 49(12):16903–16918, 2024.

[21] Sandeep Kumar, Shilpa Choudhary, Arpit Jain, Karan Singh, Ali Ahmadian, and Mohd Yazid Bajuri. Brain tumor classification using deep neural network and transfer learning. *Brain topography*, 36(3):305–318, 2023.

[22] Cherifi Dalila, Cherifi Zakaria, and Belkadi Wassim. Brain tumor classification using convolutional neural networks and transfer learning. In *International Conference on Artificial Intelligence in Renewable Energetic Systems*, pages 37–48. Springer, 2022.

[23] Muhammed Celik and Ozkan Inik. Development of hybrid models based on deep learning and optimized machine learning algorithms for brain tumor multi-classification. *Expert Systems with Applications*, 238:122159, 2024.

[24] Naeem Ullah, Ali Javed, Ali Alhazmi, Syed M Hasnain, Ali Tahir, and Rehan Ashraf. Tumordetnet: A unified deep learning model for brain tumor detection and classification. *Plos one*, 18(9):e0291200, 2023.

[25] Baiju BV, Ananda Kumar S, Sandeep Kumar Mathivanan, and Mohd Asif Shah. Efficient brain tumor grade classification using ensemble deep learning models. *BMC Medical Imaging*, 24(1):297, 2024.

[26] Tahir Hussain and Hayaru Shouno. Explainable deep learning approach for multi-class brain magnetic resonance imaging tumor classification and localization using gradient-weighted class activation mapping. *Information*, 14(12):642, 2023.

[27] Jun Cheng. Brain tumor dataset. figshare. dataset, 2018.

[28] Jun Cheng. Brain tumor dataset. figshare. dataset (2017), 2017.

[29] Recep Can, Sultan Kocaman, and Candan Gokceoglu. A comprehensive assessment of xgboost algorithm for landslide susceptibility mapping in the upper basin of ataturk dam, turkey. *Applied Sciences*, 11(11):4993, 2021.

[30] Omer Sagi and Lior Rokach. Approximating xgboost with an interpretable decision tree. *Information Sciences*, 572:522–542, 2021.

DOI: https://doi.org/10.54216/JAIM.090204

Received: December 18, 2024 Revised: February 25, 2025 Accepted: May 02, 2025