



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

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

ABSTRACT

Brain tumor research is essential for timely detection and appropriate treatment. Diagnosis is challenging because tumors vary in morphology, size, location, surface texture, and visual appearance across medical images. This study proposes a combined method for brain tumor detection from MRI images using three publicly available datasets containing 3,966 T1-weighted contrast-enhanced magnetic resonance images (T1-w MRI), classified as glioma, meningioma, pituitary tumor, or no tumor. The diagnostic pipeline applies preprocessing and data augmentation to improve image quality and variability. It then uses a VGG16 deep convolutional neural network to extract features and a Support Vector Machine (SVM) classifier to predict the class. The modified VGG16 output is supplied to the SVM, producing a test accuracy of 97.2%. Results from XGBoost, AdaBoost, Decision Tree, and K-Nearest Neighbors classifiers show that the SVM endpoint classifier performs best among the evaluated alternatives.

Keywords: Brain tumors ▪ Hybrid model ▪ Support Vector Machine ▪ T1-w MRI ▪ VGG16

1. INTRODUCTION

Brain tumors are an important medical concern because of their complex composition, varied growth patterns, and serious effects on survival and well-being. Tumors may develop without noticeable symptoms and remain undetected until they reach advanced stages, when treatment is more difficult. The American Cancer Society projected 18,330 deaths in the United States from malignant brain and spinal cord tumors in 2025, including 10,170 male and 8,160 female deaths [1]. Early diagnosis and appropriate treatment are therefore critical. A brain tumor is an abnormal tissue growth in the brain or spinal cord that disrupts normal neurological function. Uncontrolled cell division can produce benign tumors, which tend to have clear borders and grow slowly, or malignant tumors, which may grow rapidly and invade surrounding tissue. Malignant tumors are associated with high morbidity and mortality because they can resist conventional treatments [2].

Brain tumors are classified by cellular characteristics, anatomical location, and behavior. The main types include gliomas, meningiomas, and pituitary tumors. Gliomas originate in glial cells and include aggressive high-grade forms. Current neuro-oncology research applies artificial intelligence to tumor diagnosis, grading, molecular characterization, and treatment planning [3]. Identifying histological grade is important for predicting treatment outcomes and selecting appropriate therapy. Meningiomas arise from the meninges around the brain and spinal cord. Although they are typically benign, their growth can put pressure on brain tissue and cause neurological impairment. Pituitary tumors arise in the hormone-regulating pituitary gland and can cause endocrine abnormalities and vision problems when located near the optic chiasm [4].

Magnetic Resonance Imaging (MRI) is a primary tool for brain tumor assessment because it is non-invasive, provides high-resolution images, and distinguishes tissue types. MRI

supports tumor-type identification, boundary localization, and treatment monitoring [5]. Unlike computed tomography, MRI does not use ionizing radiation, making it suitable for repeated assessment. However, conventional MRI interpretation can be time-consuming and inconsistent between readers, and may be affected by fatigue. Computer-aided diagnosis systems assist radiologists by detecting, classifying, and segmenting tumors, thereby reducing diagnostic errors and improving workflow [6].

Artificial intelligence has shifted medical image analysis toward data-driven methods. Deep learning, particularly Convolutional Neural Networks (CNNs), learns hierarchical image representations without manually designed features [7]. CNNs have demonstrated strong performance in brain imaging, while conventional machine-learning methods remain useful in appropriate settings and can complement deep-learning systems [2, 8]. Support Vector Machines (SVMs) are effective in high-dimensional spaces and can work well with limited or moderately sized datasets. AI can support tumor diagnosis, grading, and treatment planning, although explainability and clinical validation remain important challenges [3].

This study develops a hybrid method in which a VGG16 CNN extracts deep features and an SVM classifies the resulting representations. The model distinguishes glioma, meningioma, pituitary tumor, and no-tumor MRI images. Combining VGG16 feature extraction with SVM decision boundaries is intended to improve diagnosis while controlling computational demands [9]. To improve generalization across imaging conditions and tumor presentations, the data are augmented using transformations such as rotation, flipping, and zooming [10]. The study makes three contributions:

- It combines VGG16 and SVM in a hybrid architecture for multi-class brain tumor classification.
- It evaluates the model using sensitivity, specificity, and F1-score in addition to accuracy.
- It uses public and local data with augmentation, and includes no-tumor cases to support screening of normal as well as pathological images.

LITERATURE REVIEW

Reviews of brain tumor imaging describe statistical, machine-learning, deep-learning, and hybrid computer-aided diagnosis approaches. They also note continuing challenges in comparing reported performance and moving automated systems toward clinical use [11]. CNNs are widely used for MRI analysis, while conventional classifiers remain useful in selected workflows [2, 8].

Recent studies illustrate varied CNN designs. Rahman and Islam proposed parallel CNN pathways with different window sizes to capture local and global features; with data augmentation, batch normalization, and dropout, their model reported 98.12% accuracy on a multi-class Kaggle dataset [12]. Abdusalomov et al. fine-tuned YOLOv7 for MRI-based detection and localization of glioma, meningioma, and pituitary tumors [13]. Al-Tahhan et al. evaluated 2,880 contrast-enhanced MRI scans across four classes and compared fine-tuned GoogleNet and AlexNet with AlexNet-SVM

and AlexNet-KNN; the AlexNet-KNN model achieved 98.6% validation accuracy [14].

Other studies combine classification with segmentation or transfer learning. A U-Net framework with ResNet50 as its backbone reported an Intersection over Union score of 0.9504, while NASNet reached 99.6% classification accuracy in the reported comparisons [15]. Transfer learning has also been applied to public datasets containing meningioma, glioma, and pituitary tumors, including glioma grades II, III, and IV; pre-trained ResNet50 and InceptionV3 achieved reported accuracy values of 96.13% and 98.7% [16]. A four-class hybrid CNN-LSTM model used AlexNet for feature extraction and reported 97.94% validation accuracy [17]. Grid-search hyperparameter optimization has produced reported accuracy values of 99.33% for tumor detection, 92.66% for five-class classification, and 98.14% for tumor grading [18].

Lightweight and transfer-learning models have also been investigated. One multi-class deep transfer-learning model reported 99.76% accuracy and a Matthews Correlation Coefficient of 0.929 [19]. The 13-layer BTC-fCNN model classified glioma, meningioma, and pituitary tumors, reporting average accuracy of 98.63% with transfer learning and 98.86% after five-fold retraining [20]. An improved ResNet50 model reported 99.30% accuracy for benign classification and 98.40% for malignant classification [21]. A CNN trained on the Navoneel Chakrabarty and BT-multiclass datasets achieved 91% accuracy for binary and 92% for multi-class classification [22].

Hybrid deep-learning and machine-learning systems have also been explored. Saeedi et al. compared a 2D CNN, a convolutional autoencoder, and six machine-learning classifiers using 3,264 contrast-enhanced MRI images [23]. TumorDetNet, an end-to-end CNN with 48 convolutional layers, reported 99.83% detection accuracy, 100% benign/malignant classification accuracy, and 99.27% three-class accuracy [24]. Patil and Kirange fused features from a shallow CNN and VGG16 in an ensemble model, reporting classification accuracy up to 97.77% [25]. Explainable deep-learning approaches have used VGG-19, EfficientNet, Grad-CAM, Grad-CAM++, and CAM to provide visual explanations of model predictions [26].

2. METHODOLOGY

Figure 1 summarizes the proposed investigation pipeline. It begins with MRI preprocessing and data augmentation, continues with feature extraction using a pre-trained CNN backbone, and ends with a classification block using the proposed model. Each stage contributes to performance and robustness in brain tumor classification.

Preprocessing resizes, normalizes, and reduces noise in MRI images to maintain dataset consistency. It is intended to address variations in image resolution, contrast, and brightness across medical sources. Data augmentation then expands the training set using transformations such as rotation, flipping, zooming, and shifting. These operations encourage generalization and help reduce overfitting, particularly because large annotated medical datasets can be difficult to collect.

The feature-extraction stage uses a pre-trained CNN backbone to learn patterns and spatial hierarchies from the input

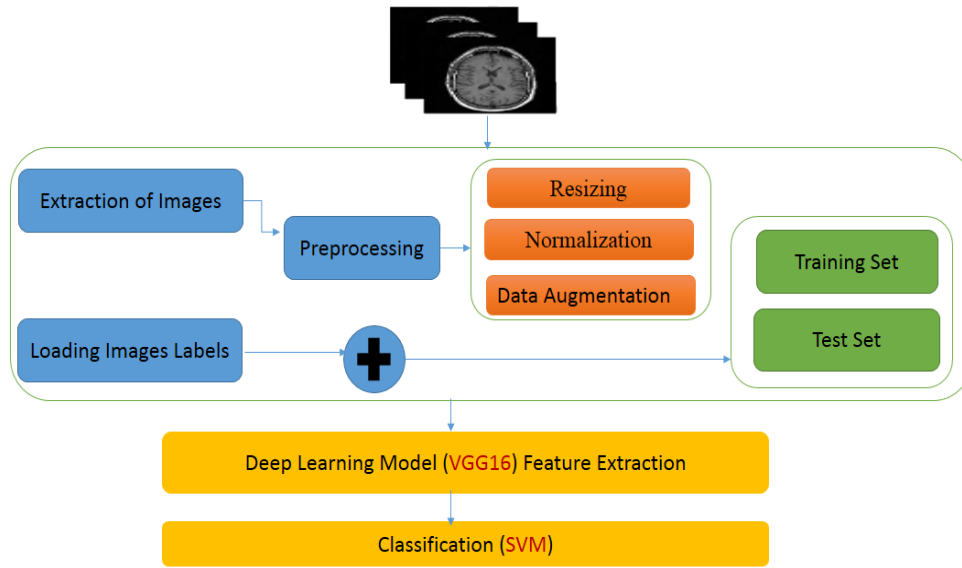


Figure 1. Full procedure of the proposed analysis pipeline.

images. Unlike conventional pipelines that use a softmax output layer, the proposed framework uses an SVM classifier after feature extraction. The SVM is selected for its margin-maximizing decision boundaries in high-dimensional spaces and its potential to separate classes that are not linearly separable. The system classifies glioma, meningioma, pituitary tumor, and no-tumor images. Its performance is compared with XGBoost, Decision Tree, K-Nearest Neighbors (KNN), and AdaBoost using quantitative metrics, confusion matrices, and ROC curves.

2.1 Dataset

The study uses the brain MRI dataset initially developed by Cheng with contributions from the Diagnostic Radiology Department at Mansoura University [27, 28]. It contains 3,504 scans representing four categories: glioma, meningioma, pituitary tumor, and no tumor. The distribution is 1,003 glioma images, 981 meningioma images, 959 pituitary tumor images, and 561 no-tumor images, as summarized in Table 1. These data provide examples for developing and evaluating machine-learning methods for brain tumor identification. The no-tumor group is included so the model can distinguish normal scans from pathological cases; it is also the smallest class in the original distribution, which makes class-sensitive evaluation important.

Table 1. Dataset distribution by class.

Class	Number of images
Glioma	1,003
Meningioma	981
Pituitary	959
No tumor	561
Total	3,504

The MRI scans cover three anatomical planes: axial, coronal, and sagittal. Figures 2–4 show example scans for the four

classes. Axial slices show cross-sectional anatomy and lesion appearance at a particular level of the brain. Coronal views provide a frontal perspective that supports assessment of vertical tumor progression and midline structures. Sagittal views show tumor depth and can help assess effects on structures such as the corpus callosum and brainstem. These different views provide complementary spatial perspectives for manual review and automated classification. Combining tumor and no-tumor cases across all three orientations provides a broader basis for model training and evaluation.

2.2 Preprocessing and Data Augmentation

Preprocessing standardizes the data and helps the model learn relevant representations rather than irrelevant patterns such as black borders. MRI data may extend beyond the 8-bit range $[0, 255]$ and may include negative intensities after scan loading. The first operation is min–max normalization, which linearly transforms pixel values to the range $[0, 1]$:

$$Y_i = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

Here, Y_i is the normalized output, X is the input pixel value, and $X_{\min}$ and $X_{\max}$ are the minimum and maximum pixel values in the dataset. Histogram equalization is then applied to spread intensity values across the available range and improve contrast between structural and pathological features. All MRI scans are resized to 256×256 pixels.

Data augmentation creates transformed versions of the original images to represent likely variations in practice, improve generalization, and reduce overfitting. It is particularly useful in medical imaging, where collecting large labeled datasets can be challenging. Rotation, shifting, shearing, and zooming expose the model to varied examples and are intended to improve generalization and training convergence. The transformations used in this study are listed in Table 2: rotation up to 40 degrees, width and height shifts of 0.08, shear range 0.3, and zoom range 0.1. The source reports that augmen-

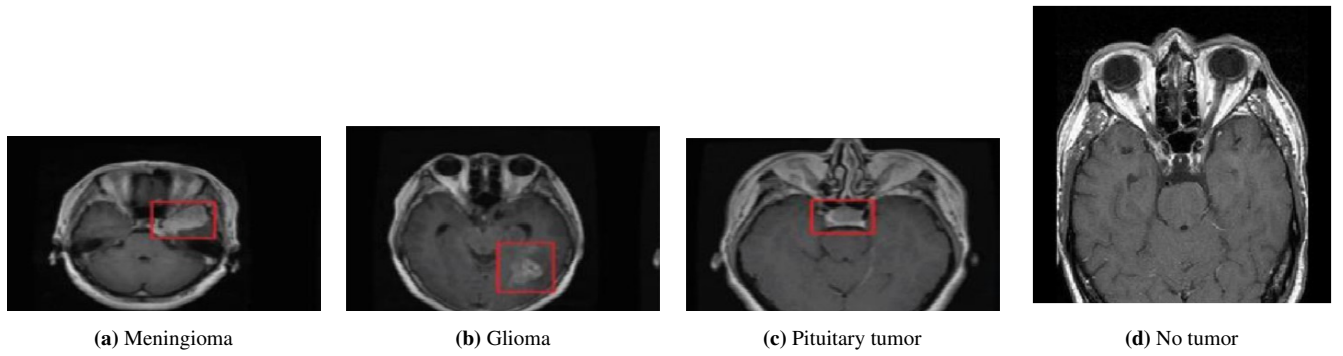


Figure 2. Representative axial MRI samples from the dataset.

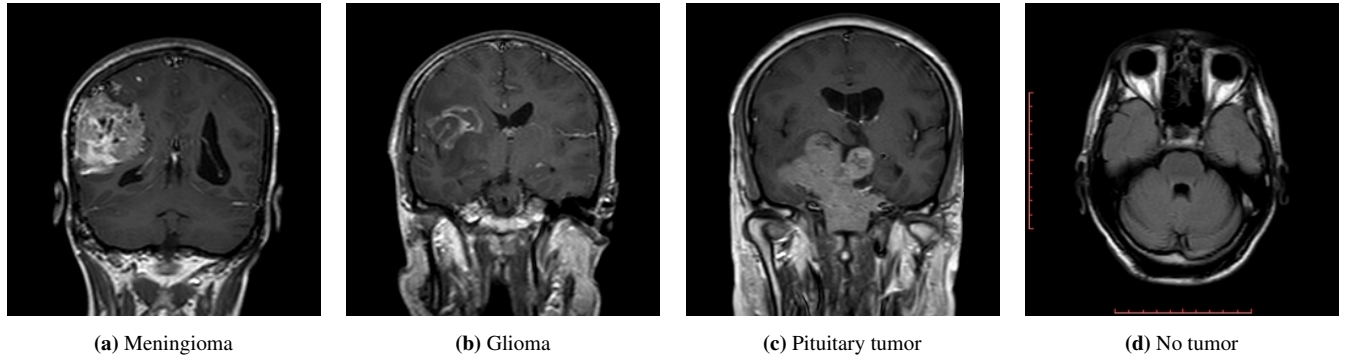


Figure 3. Representative coronal MRI samples from the dataset.

tation increased the training set to 3,966 images. Figure 5 shows example outputs, including shifts, rotation, rescaling, brightness variation, and horizontal and vertical flips.

Table 2. Data augmentation methods.

Method	Parameter
Rotation	40 degrees
Width shift	0.08
Height shift	0.08
Shear range	0.3
Zoom range	0.1

2.3 VGG16–SVM Architecture

The proposed hybrid model uses VGG16 as a feature extractor and SVM as the final classifier. After removal of the softmax layer, deep features are extracted from the penultimate fully connected layer and supplied to the SVM. The model is compared with XGBoost [29], AdaBoost, Decision Tree [30], and KNN. The proposed SVM uses an RBF kernel, which is suitable for nonlinear classification. Its performance depends on the cost parameter C and kernel parameter γ . The value of γ controls the influence range of each training example: lower values produce a wider range, while higher values produce a narrower one. The RBF kernel is

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

where z_i is a sample and z_j is a support vector. Tuning C and γ allows the SVM to determine a decision boundary that generalizes to unseen samples. Figure 6 illustrates the complete model.

2.4 Classification Head

Prior hybrid approaches have combined CNN architectures with one another or with additional deep-learning components. Such approaches may improve performance but can increase implementation complexity, computational requirements, training time, and difficulty of interpretation. This study uses VGG16 as the backbone and connects it to an SVM classifier, seeking to maintain classification performance with a simpler design.

The system classifies brain MRI images as glioma, meningioma, pituitary tumor, or no tumor using deep features extracted by VGG16. SVM is a supervised learning method that supports binary and multi-class classification. Its advantages include generalization, resistance to overfitting, and efficient memory and computational use. In image-based medical diagnosis, SVM can construct separating hyperplanes in high-dimensional feature spaces. The RBF kernel enables classification of nonlinearly separable data and can help limit noise sensitivity. These properties make SVM a useful classifier when precision and generalization are important.

2.5 Evaluation Metrics

Because class representation is unequal, the study evaluates accuracy, sensitivity, specificity, precision, and F1-score. Sensitivity measures how well tumor-bearing cases are identified among positive cases; high sensitivity is important to reduce the risk of labeling a patient with a tumor as healthy. Specificity measures the correct identification of healthy cases and helps reduce false-positive diagnoses that may lead to unnecessary concern or follow-up procedures. Precision measures the proportion of positive predictions that are correct and indicates the reliability of tumor predictions. The F1-score is the harmonic mean of precision and recall, balancing missed

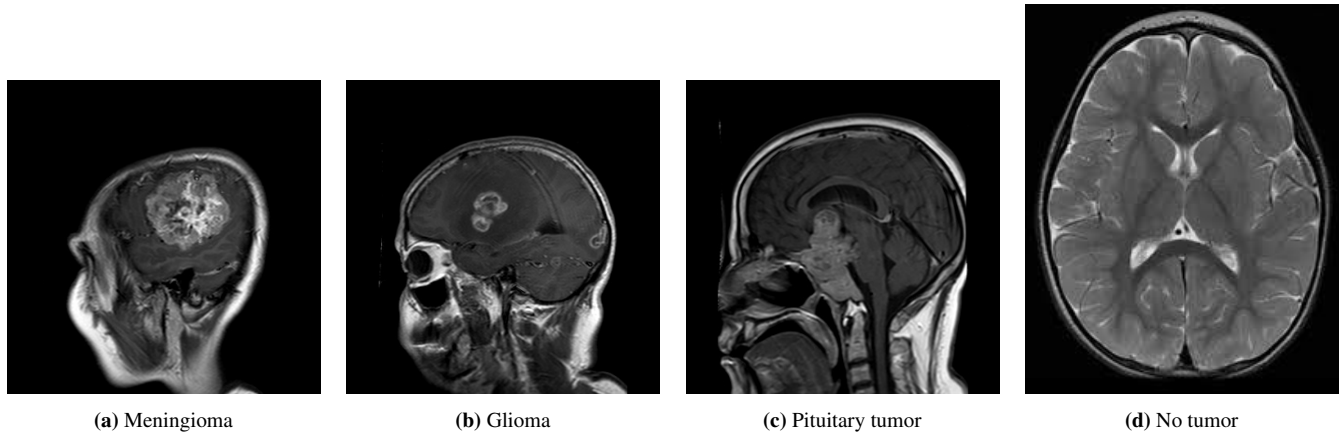


Figure 4. Representative sagittal MRI samples from the dataset.

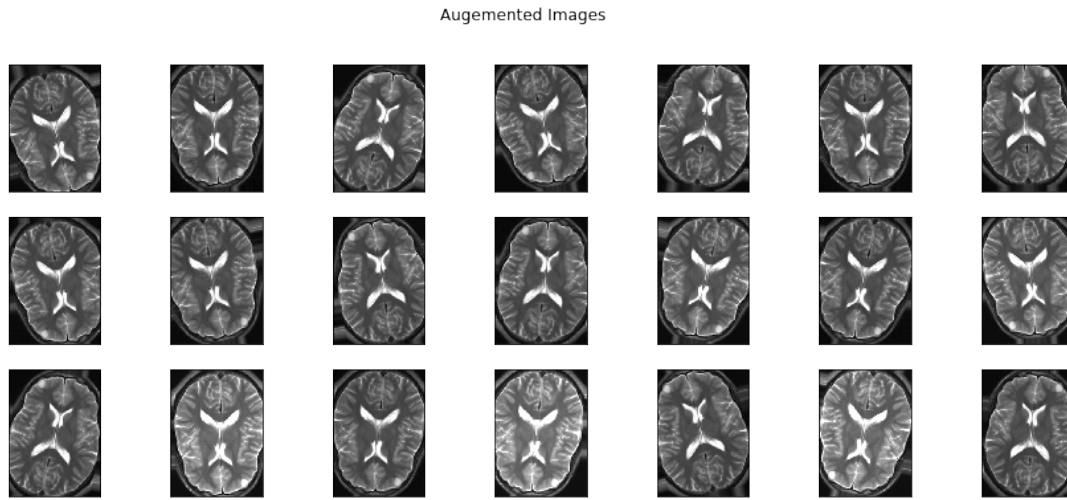


Figure 5. Example images after augmentation by width shift, rotation, height shift, rescaling, brightness adjustment, and horizontal and vertical flipping.

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	$\frac{2 \text{ Precision Recall}}{\text{Precision} + \text{Recall}}$

cases against incorrect positive predictions. It is particularly informative for imbalanced data because accuracy alone may conceal poor performance on a less frequent class. Accuracy measures correct classifications overall but does not show how performance is distributed between tumor and no-tumor cases. Table 3 gives the mathematical definitions. Here, TP and TN denote true positives and true negatives, while FP and FN denote false positives and false negatives.

2.6 Implementation Details

The model was implemented and trained on Google Colab using TensorFlow and GPU processing. The SVM classifier

was built with Scikit-learn. The reported training/testing split was 90/10. Different kernel functions and γ values were evaluated to select the model configuration. Figures 7 and 8 show the kernel comparison and tuning results; the RBF kernel achieved the highest reported test accuracy of 97.2%. Table 4 lists the selected hyperparameters.

Table 4. Hyperparameters of the proposed classifier.

Hyperparameter	Value
C	100
γ	scale
Kernel	RBF
Random state	123

3. RESULTS

The proposed model is evaluated using accuracy, sensitivity, specificity, F1-score, and processing time to provide a multi-class assessment. Table 5 compares it with other machine-learning classifiers. VGG16 feature extraction with SVM achieves 97.2% test accuracy, 100% sensitivity, 98.8% specificity, and a 97% F1-score in 319.7 seconds. The alternatives show lower accuracy and F1-score, although the Decision Tree completes processing more quickly. The results there-

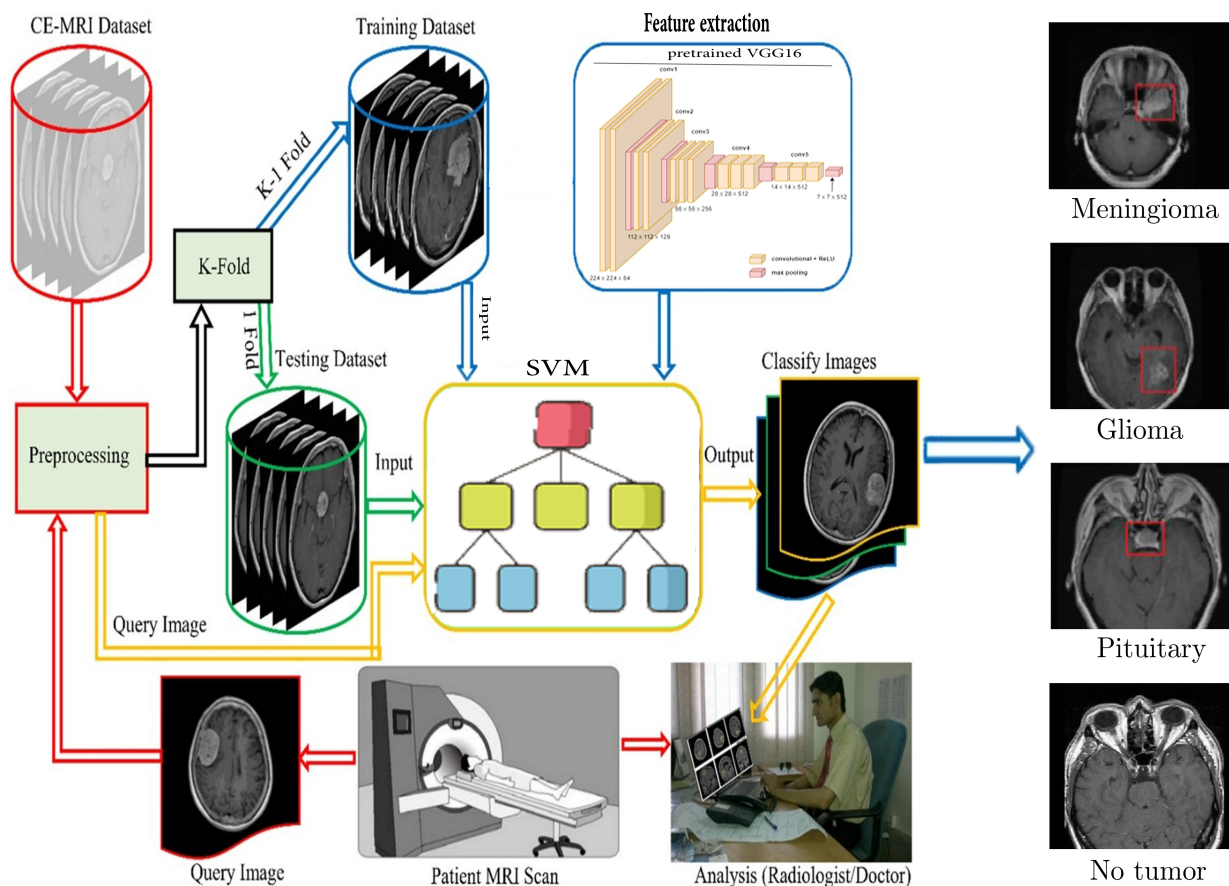


Figure 6. Proposed architecture: VGG16 feature extraction followed by SVM classification.

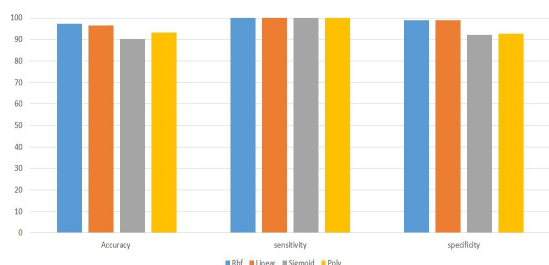


Figure 7. Proposed model performance across kernel functions.

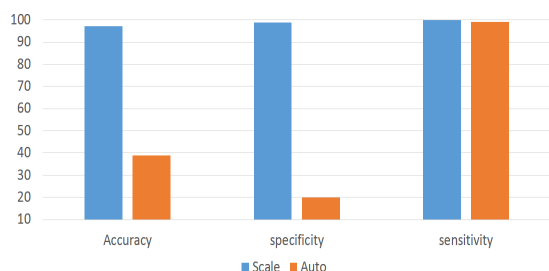


Figure 8. Tuning the γ parameter at $C = 100$.

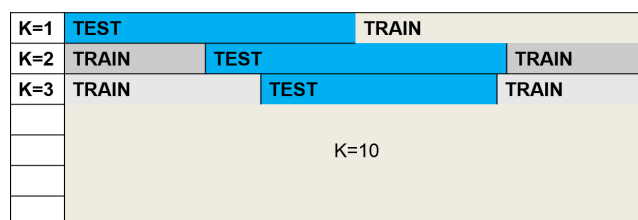


Figure 9. Ten-fold cross-validation procedure.

Table 5. Proposed model evaluation against other machine-learning classifiers. Pt denotes processing time.

Model	Acc.	Sens.	Spec.	F1	Time (s)
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%	98.8%	72%	247
VGG16+KNeighbors	88.4%	97.8%	95.4%	88%	1125

fore indicate a balance between classification performance and computation time rather than minimum runtime alone. Table 6 evaluates the model at five training/testing ratios. Its best reported results occur at 90/10, but accuracy and F1-score remain 93.7% and 94% at a 50/50 split. Sensitivity remains at or above 98.8% across the reported ratios, which the study presents as evidence that performance is maintained with less training data.

Figure 9 shows the 10-fold cross-validation procedure. The data are divided into ten parts; in each iteration, nine parts are used for training and one for validation. Each partition serves as validation data once, and the performance measures are averaged across the ten iterations.

Table 6. Performance at different training/testing ratios.

Train/test	Acc.	Sens.	Spec.	F1
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%

Table 7. Performance at a 90/10 training/testing ratio using pre-trained feature extractors with SVM.

Model	Acc.	Sens.	Spec.	F1	Time (s)
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

Table 7 compares the proposed VGG16 feature extractor with InceptionV3, MobileNetV2, Xception, VGG19, and DenseNet121, each paired with SVM. VGG16+SVM records the highest reported accuracy (97.2%) and an F1-score of 97% with a processing time of 319.7 seconds. MobileNetV2+SVM and DenseNet121+SVM reach 94.4% and 95.9% accuracy but require 636.2 and 384.6 seconds, respectively. VGG19+SVM reports 94.7% accuracy, while InceptionV3+SVM and Xception+SVM obtain F1-scores of 83% and 92%. These results favor the proposed VGG16–SVM configuration in the reported balance of accuracy and processing time.

Figure 10 compares the proposed model with VGG16+XGBoost, VGG16+KNeighbors, VGG16+Decision Tree, and VGG16+AdaBoost using confusion matrices for glioma, meningioma, pituitary tumor, and no tumor. The proposed model shows dominant diagonal values and few off-diagonal errors, indicating that most examples are classified into their correct categories. XGBoost produces adequate results in most cases but shows some glioma misclassification. KNeighbors classifies pituitary tumors effectively, although glioma and meningioma are confused

Table 8. Comparison of the proposed model with baseline VGG16 using a softmax classifier.

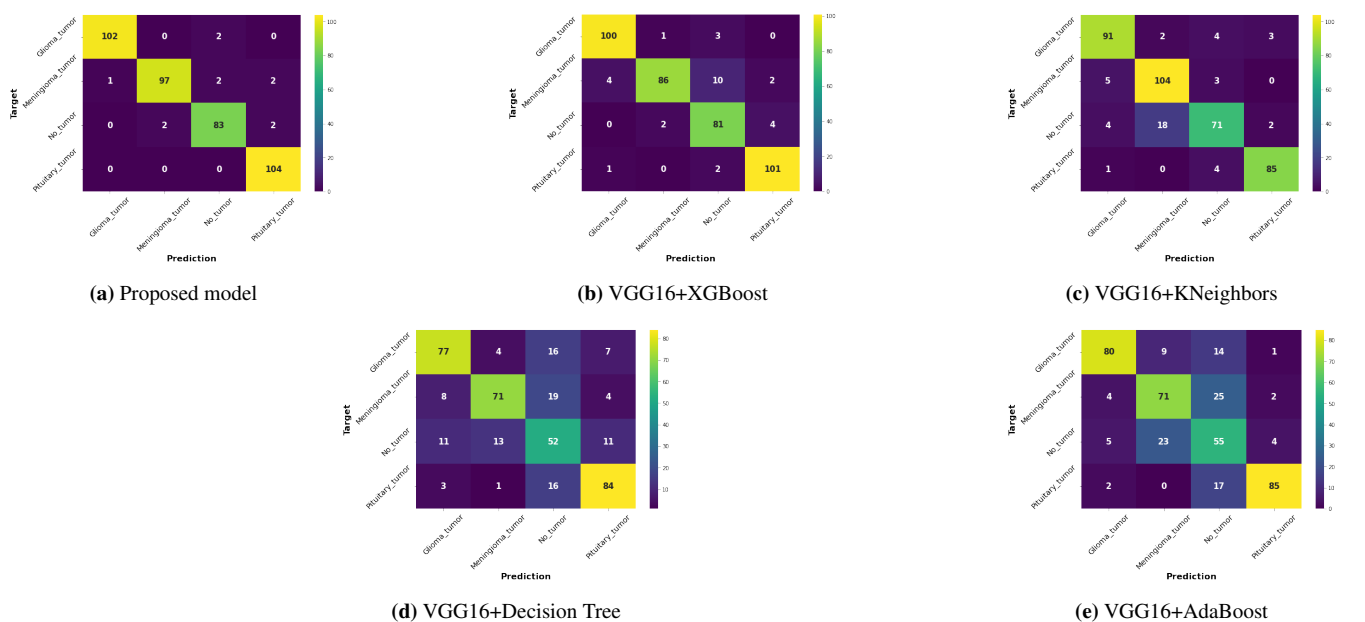
Model	Train acc.	Test acc.	Sens.	Spec.	Time (s)
Proposed model	100%	97.2%	100%	98.8%	319.7
Baseline VGG16	99.3%	94.4%	99%	98%	734.5

more often. The Decision Tree produces more incorrect classifications across tumor types, particularly between no-tumor and meningioma cases. AdaBoost performs moderately for pituitary tumors but shows lower performance for glioma and meningioma, a concern where missed tumor cases are clinically important.

Figure 11 compares the proposed model's confusion matrix with that of the baseline VGG16 softmax classifier. The softmax baseline remains practical but produces more misclassifications, particularly between glioma and meningioma. Table 8 reports test accuracy of 97.2% for the proposed model and 94.4% for the baseline; sensitivity and specificity are 100% and 98.8% for the proposed method versus 99% and 98% for the baseline. Processing time is 319.7 seconds for the proposed model and 734.5 seconds for the baseline, which the source describes as more than 50% faster.

Figure 12 shows the Receiver Operating Characteristic (ROC) curves, which compare the true-positive detection rate against the false-positive rate over different decision thresholds. The proposed model reports an Area Under the Curve (AUC) of 1.0 for glioma, meningioma, pituitary tumor, and no tumor, indicating strong class discrimination in the reported experiment. The comparison curves deviate from this result: XGBoost performs relatively well but is less consistent across classes, KNeighbors shows moderate performance, and Decision Tree and AdaBoost have lower AUC values and flatter curves. In medical diagnosis, both false positives and false negatives matter because they may respectively lead to unnecessary follow-up or missed disease. Class-wise results for the proposed model are listed in Table 9.

The proposed model achieved 100% recall for pituitary tu-

**Figure 10.** Confusion matrices for the proposed model and alternative classifiers, showing classification across tumor types.

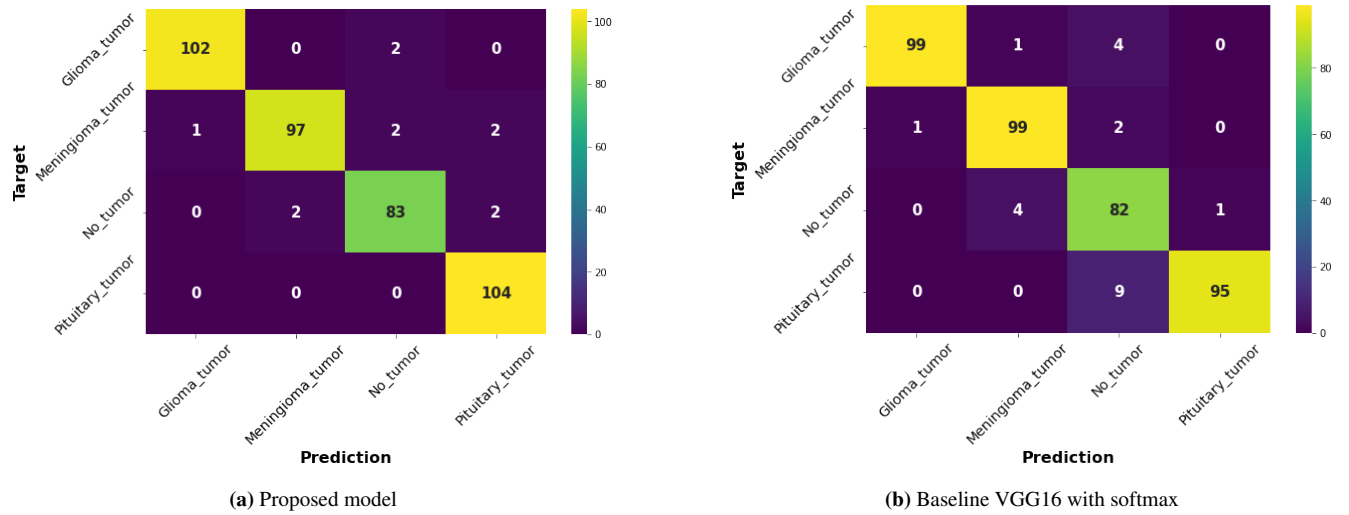


Figure 11. Confusion matrices for the proposed model and baseline VGG16 softmax classifier.

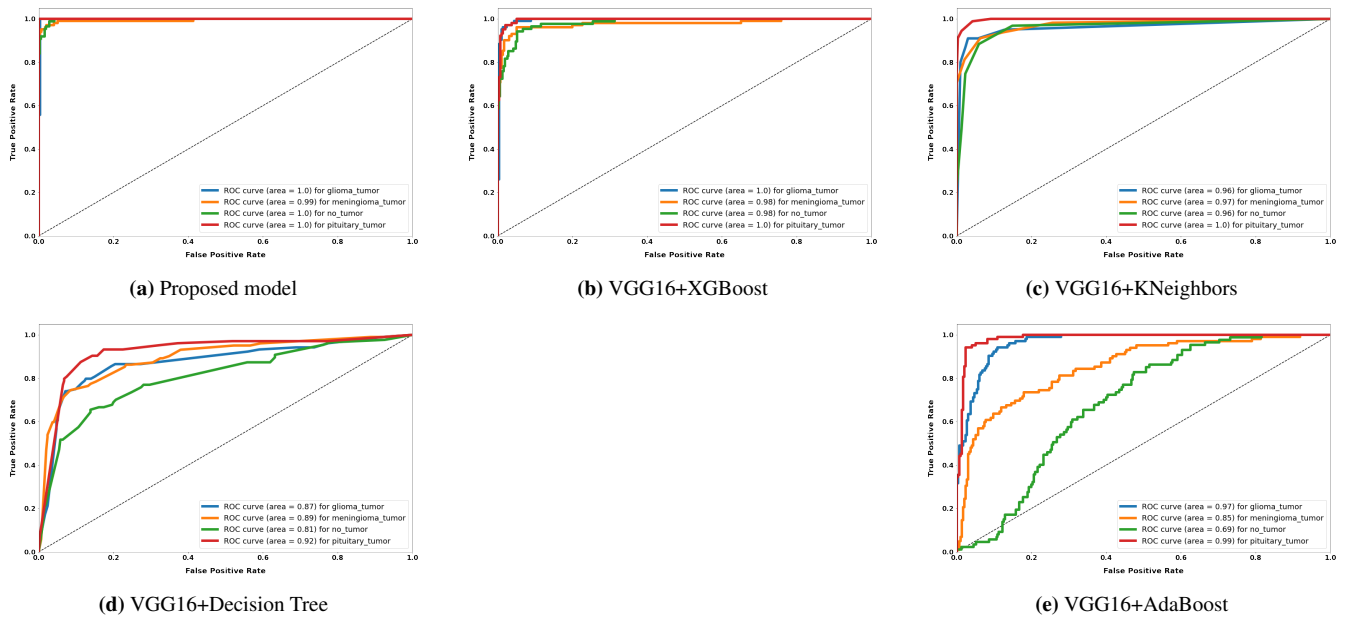


Figure 12. ROC curves for the proposed model and comparison classifiers.

mors and 95% recall for glioma and meningioma. The source reports F1-scores between 95% and 99% across the four classes, with 99% F1-score for no tumor, 97% for glioma, 95% for meningioma, and 98% for pituitary tumor. Tables 10–13 show class-wise results for the alternative classifiers. XGBoost achieved 84% recall for glioma. AdaBoost had its weakest results for meningioma and glioma, with F1-scores of 56% and 69%, respectively; its pituitary precision was 92%, but overall performance remained uneven. Decision Tree results were also limited, particularly for meningioma, which had an F1-score of 55%. KNeighbors performed better than AdaBoost and Decision Tree for pituitary and no-tumor cases, but its F1-scores for glioma and meningioma (88% and 80%) remained below those of the proposed model.

Overall, the VGG16–SVM hybrid outperformed the evaluated classifiers and the VGG16 softmax baseline on the reported test accuracy, F1-score, sensitivity, specificity, and processing time. It maintained performance across different training/testing ratios and achieved 97.2% test accuracy, 100% sensitivity, and 98.8% specificity with a processing time of 319.7 seconds. The comparisons with other feature extractors likewise favor the proposed VGG16–SVM configuration in the balance of accuracy and processing time. Confusion matrices and ROC curves show strong class separation, while the class-wise tables reveal that the alternative classifiers have particular difficulty with glioma and meningioma. These findings support the potential value of the hybrid method for computer-aided diagnosis, while remaining specific to the

Table 9. Proposed model class metrics.

Class	Prec.	Recall	F1
No tumor	99%	98%	99%
Glioma	98%	95%	97%
Meningioma	95%	95%	95%
Pituitary	96%	100%	98%

Table 10. VGG16+XGBoost class metrics.

Class	Prec.	Recall	F1
No tumor	95%	96%	96%
Glioma	97%	84%	90%
Meningioma	94%	93%	89%
Pituitary	94%	97%	96%

Table 11. VGG16+AdaBoost class metrics.

Class	Prec.	Recall	F1
No tumor	88%	77%	82%
Glioma	69%	70%	69%
Meningioma	50%	63%	56%
Pituitary	92%	82%	87%

Table 12. VGG16+Decision Tree class metrics.

Class	Prec.	Recall	F1
No tumor	78%	74%	76%
Glioma	80%	70%	74%
Meningioma	50%	60%	55%
Pituitary	79%	81%	80%

Table 13. VGG16+KNeighbors class metrics.

Class	Prec.	Recall	F1
No tumor	90%	91%	91%
Glioma	84%	93%	88%
Meningioma	87%	75%	80%
Pituitary	94%	94%	94%

datasets and evaluation conditions used in the study; further validation is needed before clinical use.

4. CONCLUSION

This study presents a brain tumor classification model that combines VGG16 for deep feature extraction with an SVM classifier. The learned VGG16 features are passed to SVM for final prediction, combining deep-learning representations with a traditional classifier. The proposed model achieved 97.2% classification accuracy in the reported experiments and may support rapid tumor diagnosis within computer-aided clinical workflows.

Future work includes collecting a larger and more diverse dataset and generating high-quality synthetic images using Generative Adversarial Networks (GANs). A larger dataset would allow more reliable assessment of tumor subtypes. Further work should also improve interpretability, since the current deep-learning system is difficult to explain. Visualization methods such as saliency maps, Grad-CAM, and attention mechanisms could provide clinicians with explanations for model predictions. Combining broader data coverage with interpretable predictions may help improve the accuracy, reliability, and acceptance of AI-based diagnostic tools.

REFERENCES

- [1] J. H. Park, A. L. Garcia de Lomana, D. M. Marzese, T. Juarez, A. Feroze, P. Hothi, C. Cobbs, A. A. Patel, S. Kesari, S. Huang *et al.*, "A systems approach to brain tumor treatment," *Cancers*, vol. 13, no. 13, p. 3152, 2021.
- [2] K. Neamah, F. Mohamed, M. M. Adnan, T. Saba, S. A. Bahaj, K. A. Kadhim, and A. R. Khan, "Brain tumor classification and detection based DL models: A systematic review," *IEEE Access*, vol. 12, pp. 2517–2542, 2023.
- [3] V. Nakhate and L. N. Gonzalez Castro, "Artificial intelligence in neuro-oncology," *Frontiers in Neuroscience*, vol. 17, p. 1217629, 2023.
- [4] M. Yildirim, E. Cengil, Y. Eroglu, and A. Cinar, "Detection and classification of glioma, meningioma, pituitary tumor, and normal in brain magnetic resonance imaging using deep learning-based hybrid model," *Iranian Journal of Computer Science*, vol. 6, no. 4, pp. 455–464, 2023.
- [5] C. M. Rogers, P. S. Jones, and J. S. Weinberg, "Intraoperative MRI for brain tumors," *Journal of Neuro-Oncology*, vol. 151, pp. 479–490, 2021.
- [6] A. Sekhar, S. Biswas, R. Hazra, A. K. Sunaniya, A. Mukherjee, and L. Yang, "Brain tumor classification using fine-tuned GoogLeNet features and machine learning algorithms: IoMT enabled CAD system," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 3, pp. 983–991, 2021.
- [7] I. Abd El Kader, G. Xu, Z. Shuai, S. Saminu, I. Javaid, and I. S. Ahmad, "Differential deep convolutional neural network model for brain tumor classification," *Brain Sciences*, vol. 11, no. 3, p. 352, 2021.
- [8] P. Jyothi and A. R. Singh, "Deep learning models and traditional automated techniques for brain tumor segmentation in MRI: A review," *Artificial Intelligence Review*, vol. 56, no. 4, pp. 2923–2969, 2023.
- [9] A. Raza, H. Ayub, J. A. Khan, I. Ahmad, A. S. Salama, Y. I. Daradkeh, D. Javeed, A. Ur Rehman, and H. Hamam, "A hybrid deep learning-based approach for brain tumor classification," *Electronics*, vol. 11, no. 7, p. 1146, 2022.
- [10] S. Sharma, K. Guleria, S. Tiwari, and S. Kumar, "A deep learning based convolutional neural network model with VGG16 feature extractor for the detection of alzheimer disease using MRI scans," *Measurement: Sensors*, vol. 24, p. 100506, 2022.
- [11] U. Raghavendra, A. Gudigar, A. Paul, T. S. Goutham, M. A. Inamdar, A. Hegde, A. Devi, C. P. Ooi, R. C. Deo, P. D. Barua, F. Molinari, E. J. Ciaccio, and U. R. Acharya, "Brain tumor detection and screening using artificial intelligence techniques: Current trends and future perspectives," *Computers in Biology and Medicine*, vol. 163, p. 107063, 2023.
- [12] T. Rahman and M. S. Islam, "Mri brain tumor detection and classification using parallel deep convolutional neural networks," *Measurement: Sensors*, vol. 26, p. 100694, 2023.
- [13] A. B. Abdusalomov, M. Mukhiddinov, and T. K. Whangbo, "Brain tumor detection based on deep learning approaches and magnetic resonance imaging," *Cancers*, vol. 15, no. 16, p. 4172, 2023.
- [14] F. E. Al-Tahhan, G. A. Khouqeer, S. Saadi, A. Elgarayhi, and M. Sallah, "Refined automatic brain tumor classification using hybrid convolutional neural networks for mri scans," *Diagnostics*, vol. 13, no. 5, p. 864, 2023.
- [15] T. Sadad, A. Rehman, A. Munir, T. Saba, U. Tariq, N. Ayesha, and R. Abbasi, "Brain tumor detection and multi-classification using advanced deep learning techniques," *Microscopy Research and Technique*, vol. 84, no. 6, pp. 1296–1308, 2021.
- [16] H. H. Sultan, N. M. Salem, and W. Al-Atabany, "Multi-classification of brain tumor images using deep neural network," *IEEE Access*, vol. 7, pp. 69 215–69 225, 2019.

- [17] S. K. Rajeev, M. P. Rajasekaran, K. Ramaraj, G. Vishnuvarthanan, T. Arunprasath, and V. Muneeswaran, "A hybrid CNN-LSTM network for brain tumor classification using transfer learning," 2023 9th International Conference on Smart Computing and Communications (ICSCC), IEEE, pp. 77–82, 2023.
- [18] E. 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*, vol. 45, no. 3, pp. 1015–1036, 2021.
- [19] S. Asif, M. Zhao, F. Tang, and Y. Zhu, "An enhanced deep learning method for multi-class brain tumor classification using deep transfer learning," *Multimedia Tools and Applications*, vol. 82, no. 20, pp. 31 709–31 736, 2023.
- [20] B. S. Abd El-Wahab, M. E. Nasr, S. Khamis, and A. S. Ashour, "BTC-fCNN: Fast convolution neural network for multi-class brain tumor classification," *Health Information Science and Systems*, vol. 11, p. 3, 2023.
- [21] S. Kumar, S. Choudhary, A. Jain, K. Singh, A. Ahmadian, and M. Y. Bajuri, "Brain tumor classification using deep neural network and transfer learning," *Brain Topography*, vol. 36, no. 3, pp. 305–318, 2023.
- [22] C. Dalila, C. Zakaria, and B. Wassim, "Brain tumor classification using convolutional neural networks and transfer learning," *International Conference on Artificial Intelligence in Renewable Energetic Systems*, Springer, pp. 37–48, 2022.
- [23] S. Saeedi, S. Rezayi, H. Keshavarz, and S. R. Nikaan Kalhori, "Mri-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques," *BMC Medical Informatics and Decision Making*, vol. 23, no. 1, p. 16, 2023.
- [24] N. Ullah, A. Javed, A. Alhazmi, S. M. Hasnain, A. Tahir, and R. Ashraf, "TumorDetNet: A unified deep learning model for brain tumor detection and classification," *PLOS ONE*, vol. 18, no. 9, p. e0291200, 2023.
- [25] S. Patil and D. Kirange, "Ensemble of deep learning models for brain tumor detection," *Procedia Computer Science*, vol. 218, pp. 2468–2479, 2023.
- [26] T. Hussain and H. Shouno, "Explainable deep learning approach for multi-class brain magnetic resonance imaging tumor classification and localization using gradient-weighted class activation mapping," *Information*, vol. 14, no. 12, p. 642, 2023.
- [27] J. Cheng, "Brain tumor dataset," Figshare dataset, 2018.
- [28] J. Cheng, W. Huang, S. Cao, R. Yang, W. Yang, Z. Yun, Z. Wang, and Q. Feng, "Enhanced performance of brain tumor classification via tumor region augmentation and partition," *PLOS ONE*, vol. 10, no. 10, p. e0140381, 2015.
- [29] R. Can, S. Kocaman, and C. Gokceoglu, "A comprehensive assessment of XGBoost algorithm for landslide susceptibility mapping in the upper basin of ataturk dam, turkey," *Applied Sciences*, vol. 11, no. 11, p. 4993, 2021.
- [30] O. Sagi and L. Rokach, "Approximating XGBoost with an interpretable decision tree," *Information Sciences*, vol. 572, pp. 522–542, 2021.