Brain Tumor Detection and Classification from MRI Scans Using Deep Convolutional Neural Networks

Neelam Mary Vijaya Nirmala[1*], Kommineni Prathima[2], Myla Vinaya Sri[3], Papani Sri Lakshmi[4] and Pasam Aparna[5]

[1,2,3,4,5]Department of Information Technology, KKR & KSR Institute of Technology and Sciences, Guntur, Andhra Pradesh, INDIA
[1*]nirmala.neelam@gmail.com
[2]kommineniprathima@gmail.com
[3]vinayasrimyla68@gmail.com
[4]papanisrilakshmi151@gmail.com
[5]22jr1a1292@gmail.com

Abstract. Detection of tumors in brain from Magnetic Resonance Imaging (MRI) plays a vital role in both planning and diagnosis for the treatment. Examining these MRI images manually takes a lot of time and it depends upon the experience of the observer. This drawback motivated the significance of automated detection systems. This project presented a deep learning framework with the help of convolutional neural networks (CNN) to detect the tumors accurately and in a precise manner. The system learns features from MRI scans and improve the performance. To improve the overall capability, a multi task based learning approach is introduced which performs both tumor segmentation and tumor detection in the same model. Addition to this, a module called Explainable AI using Grad CAM is integrated into the system which gives the visual representation in the form of heat maps as the output. This method helps to develop trust on the system and improves reliability of the entire model. Experimental results describe that this model has performed well as it proved the F1 score greater than 0.91, Dice score more than 0.89 and Area Under Curve (AUC) values near to 0.98. However, the proposed model requires some more work in order to be a real viable solution in real world application Overall, this proposed system proves high accuracy, efficiency, interpretability for the automatic tumor detection and classification.

Keywords: Explainable AI, convolutional neural networks, MRI images, deep learning, Grad CAM.

1 INTRODUCTION

When the cells present in the human body grow compulsively rather than actual behavior then tumors are formed. Tumors in brain are caused by abnormal growth of cells in the brain that can form in and around the brain. These tumors can be of two types benign and malignant. Benign tumors are non-cancerous and malignant tumors are cancerous.

B. Singh et al. (eds.), *Proceedings of the International Conference on Advances in Computing Technology and Artificial Intelligence (COMPUTATIA 2026)*, Atlantis Highlights in Intelligent Systems 18,
https://doi.org/10.2991/978-94-6239-713-2_3

Benign tumors grow slowly and do not affect other parts of the body but sometimes they may cause harm to important areas of the brain whereas malignant tumors grow quickly and they may spread throughout the brain or spinal cord Tumors are graded from 1 to 4.

Tumors of grade 1 are mostly non-cancerous and slow growing and grade 2 tumors are also slow growing but sometimes they may become cancerous. Grade 3 tumors have faster growth and malignant whereas Grade 4 tumors have rapid growth, difficult to treat and cause more harm to the human body. Some examples of brain tumors include Glioma, Meningioma, Pituitary Tumors, Medullo blastoma etc. Depending upon the area of growth, tumors can be of different sizes from very small to large and do not show effect immediately. Common effects of brain tumors include headache, nausea, loss of vision, memory loss, seizures etc. At first doctors identify these symptoms in patients and undergo medical history of patients and physically examine them to detect these brain tumors, a famous method called MRI is used. MRI is an imaging technique which is used to scan the inner organs and tissues of our body. MRI of the brain shows the tissues inside clearly and it is used to detect the brain tumors. Doctors examine these MRI scans and detect the tumor and proceed with further processes depending upon the type of tumor. Generally, brain tumors are complex, harmful and require early and accurate diagnosis. Modern technology helps in detection and diagnosis of tumors very accurately and precisely. Especially brain tumors can be detected well with the help of the latest technology.

2 LITERATURE REVIEW

Detection of brain tumors by making the use of MR image of the brain requires familiarity and also it is hard to detect them sometimes because of its size and area and some other factors. In the past few years, a lot of research has been done in implementing deep learning techniques and deep learning models in detecting brain tumors. The study in [1] introduced a deep learning model called ResNet50 but it is dependent on deep architecture which can increase the complexity of the model. For example, a framework based on deep CNN was introduced in [2]. Here, the tumors are classified from MRI scans by using hierarchical feature extraction. However, it doesn't incorporate interpretability mechanisms and it is based on conventional structures of CNN. Also, another model was proposed in [3]which is combination of enhanced CNN with data augmentation techniques to decrease the occurrence of overfitting but it is not based on multi modal integration.

Some of the approaches were also introduced which is an integration of classification and segmentation in [4], where the multi class brain tumor segmentation was proposed. But the major limitation is high complexity due to combined models. The study in [5] introduced a model based on CNN for tumor detection with high accuracy. But it did not include any attention mechanisms or explainability frameworks. It also lacks cross institutional validation for model stability. Also the work in [6] introduced YOLOv7, at the same time, detection and classification of brain tumors, achieving competitive detection metrics as mentioned in their study. Even though real-time detection was substantiated, the study did not fully integrate explainable components, limiting interpretability.

The research in [7] invented CNN-Tumor Net, leveraging explainability techniques such as LIME to provide interpretable decision insights. The model accomplished high accuracy and interpretability, but it was limited to simple tumor vs. no tumor classification and [8] popularized a fusion-based ensemble learning framework combining MobileNetV2 and DenseNet121 with GradCAM++ for class specific visualization. This model improved robustness and interpretability, yet depended on complex ensemble configurations, making deployment challenging in low resource environments. The study in [9] introduced an explainable Efficient NetV2 with attention enhanced MLP-Mixer architecture for MRI brain tumor classification, achieving high accuracy and interpretability on benchmark datasets; however, the integration of attention and mixer modules increased model complexity. Also in [10] research of deep models with Grad-CAM explainability for classification and visualization taken place.

According to the reviewed literature, the research gaps are as follows:

- The majority of existing deep learning models do not integrate tumor detection, multi class grading, explainability, and lightweight deployment concurrently.
- Many models require extensive computational resources, limiting usability in real-time and low-resource clinical environments.
- Further, we can continue with some techniques which add to the contrast, making the model to identify tumor regions in the scan images. Enhancement of data includes rotation, flipping of images horizontally and vertically translations.

3 METHODOLOGY

There are several steps involved in making the system which detects the tumors in brain. This section is especially to show the different phases involved in the creation of brain tumor detection and segmentation from MRI Scans using Deep Convolution Neural Networks. It consists of phases from data collection phase to the evaluation of metrics phase. All of the phases specified below together are used to develop the model achieving high accuracy when compared to baseline model.

3.1 Data Collection and Preprocessing

This is the initial step in the evolution of model. This step involves collection of dataset used for training. The dataset I have utilized is BraTS2020 which is present in Kaggle. This dataset structure is similar to the official BraTS2020 dataset but it may not contain all the metadata available in the original dataset. It has multi modal scans of 369 patients. It has 4 modalities. They are T1, T1ce, T2, FLAIR and also it has segmentation masks showing the tumor areas in scan. The input of this model is MRI image. This dataset consists of multimodal MRI images along with tumor segmentation masks. The volumes of MRI are changed into 2D slices to store them in HDF5 format to load the data efficiently. Every slice of MRI is having association with tumor mask for the identification of tumor region. The MRI images are changed to a constant input dimension which becomes suitable for the architecture of CNN. Rotations, scaling, flips are executed on the images of the dataset to upgrade robustness of model. After this intensity normalization is implemented to scale the pixel values to a fixed scope to increase the stability during training. Ultimately, the preprocessed dataset is segregated into several parts such as training, validation and testing. To prevent the data leakage this segregation is performed at the patient volume level. Validation is performed for early stopping and hyper parameter tuning. And the training is done to optimize the parameters of the model.

3.2 Proposed Model Architecture

The proposed system as shown in Fig. 1. consists of a deep convolution neural network also known as deep CNN along with a multi task framework for performing tumor detection as well as segmentation from MRI scans. This model is proposed to perform both the tasks by improving the accuracy of the model.

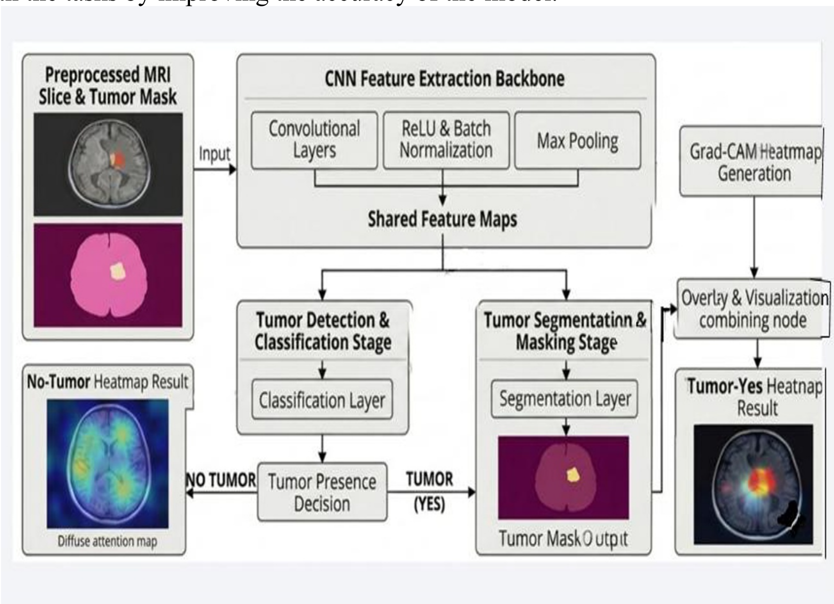

Fig. 1. Architecture of the proposed System

The input for this model is the data set which has undergone several steps like intensity normalization, resizing etc. in the previous phase. The images are then made to pass through multiple layers (convolutional layers) to learn hierarchical features and representations. This include essential features such as textures, edges and some tumor specific high level patterns. Each layer in this convolutional network has normalization and activation functions to achieve the training stabilization and improve the non-linear feature learning.

After some selected convolutional blocks, Max-pooling layers are integrated which decreases the spatial dimensions by maintaining the original essential features. It supports in reducing the computational resources and complexity and enabling the model to concentrate on important sections of the scan. The feature maps which are extracted are distributed among two specific branches. The responsibility of the first branch is detection of tumor. It performs binary classification to detect whether the tumor is present or not. The responsibility of the second branch is tumor segmentation. It generates segmentation masks to identify tumors based on the features learnt during training phase.

To increase the interpretability of model, a module named Explainable AI which is based on Gradient-weighted Class Activation Mapping shortly known as Grad CAM is included in the proposed system clearly visible in Fig1. This module gives us heat maps as the output which consists of some highlighted areas of the MRI scans that provide the predictions of model. This provides visual representations of the predictions of model and helps in the decision making by clinicians. This system is streamlined to be a lightweight model by reducing the number of parameters and utilizing productive operations. This architecture enables the model to be suitable for deployment in resource constrained environments.

3.3 Feature Extraction and Training

The proposed multi-task CNN (Convolutional Neural Networks) model learns important information directly from MRI images without needing discriminative features as hand crafted features. In the early layers of the network, the model focuses on simple visual details such as edges, corners, and textures. These details help the system recognize where a tumor begins and ends. Through the network, the model moves deeper and captures tumor shapes, sizes and intensity variations. These high level features not only help to detect the tumor but also helps to know the exact location of tumor. The feature maps which are obtained from the convolutional network layers are processed using activation functions and pooling layers. These activation functions are non-linear activation functions. These are used to highlight most significant information while those pooling layers are used for decreasing spatial dimension and computational complexity. These are the extracted features that are shared across dual branches of the network, it includes one for tumor detection and another one for tumor segmentation. By allowing the network to leverage common features for multiple tasks and improve overall generalization.

During the training period of the model, it is optimized by using loss functions which are task-specific. The tumor detection branch typically works on binary cross entropy with logits while the tumor segmentation uses the hybrid loss. Therefore, the total loss is computed as a weighted sum of individual losses. This is used to learn both tasks at a time in a framework of multi task learning. In baseline model Adam optimizer is used whereas in proposed model AdamW with cosine annealing scheduling is used for stable convergence. Batch normalization is applied to each layer of the convolutional to reduce internal covariate shift, and dropout layers are used to eliminate overfitting. During the training of the dataset, data augmentation is dynamically applied which helps to improve the performance and makes the dataset diverse. In this we use techniques that include flipping, intensity scaling, rotation etc. are applied. By integrating the deep feature extraction with optimized strategies for training the model, this system achieves more accuracy in brain tumor detection and segmentation of tumors by achieving the efficiency be fitting for real-time clinical applications.

3.4 Explainable AI

By using deep learning models, black boxes are often used for the medical image analysis which is becoming difficult to understand the predictions behind their reasoning. To overcome this, an XAI (Explainable AI) model is combined into the brain tumor detection and segmentation system which increases the interpretability and model transparency.

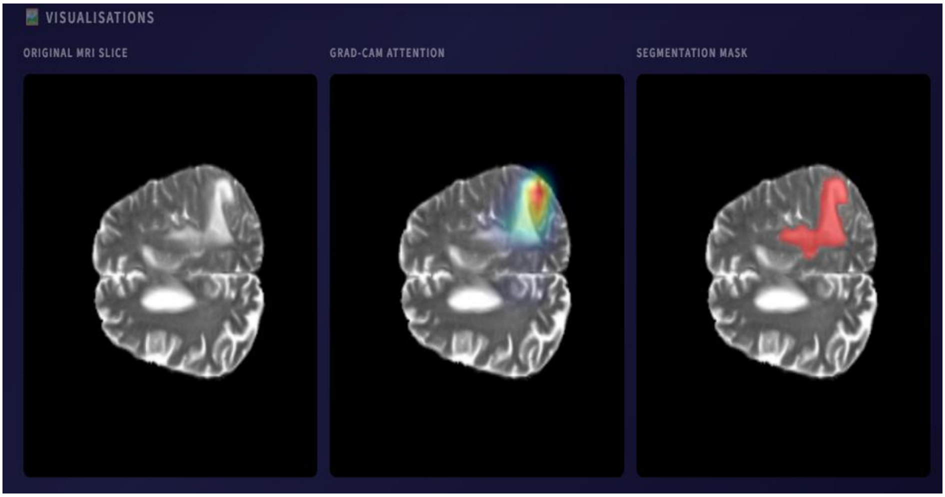

Fig. 2. Explainable AI visualization using Grad-CAM

For creating the visual explanations for the predictions of model, the proposed system utilizes the Gradient Weighted Class Activation Mapping (Grad CAM). This Grad CAM makes use of gradients of the target class which is flowing into the final convolutional layers to construct a coarse localization map, displays the model's decision by highlighting the regions in the MRI image. Grad CAM heat maps will be covered on the actual image to specify the locations of the tumor that gives tumor detection in each and every input of the image. These explanations show visually and helps to view model's decisions and increase trust in automated distinctive systems by providing intuitive, human understandable insights into the model's decision-making technique. The integration of XAI is specifically applicable in clinical applications where both interpretability and accountability are important. By furnishing these clear visual evidence of decision-making, assist radiologists and medical professionals.

Fig. 2. shows the process for a sample input of MRI slice. The MRI slice is given as input to the model. Then Grad CAM is used to produce the heat map which highlights the models decision. The model also produced the tumor segmentation masks that tells the tumor area at pixel level.

3.5 Evaluation and Performance Metrics

The performance of researched multi-task CNN based brain tumor detection and segmentation systems is estimated using standard quantitative metrics that are commonly used in medical aspects. These metrics evaluate an extensive appraisal of the model's accuracy, reliability and robustness. For tumor detection task accuracy shows general correctness, precision efforts show how many predicted tumors are truly tumors, recall indicates how well original tumors are detected. For the segmentation tasks, the metrics used are Dice Similarity Coefficient (DSC) and IoU (Intersection over Union).

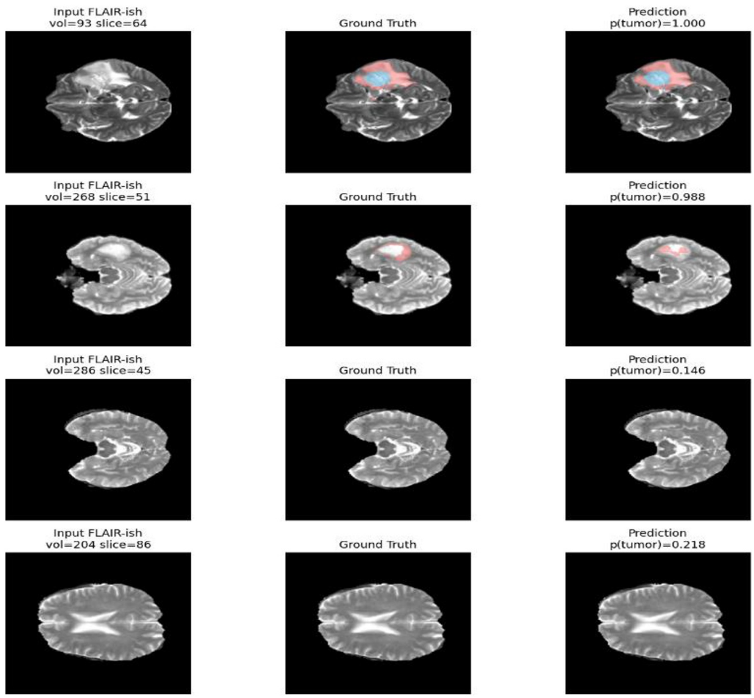

Fig. 3. Evaluation of proposed CNN

In the above Fig. 3. the first 2 rows show the detection confidence (p=1.000 and p=0.9888) and also the tumor delineation for the positive cases and the next two rows shows the negative cases. Here the model predicts less tumor probabilities (p=0.146 and p=0.218).

The F1 score the major performance metrics used to detect the tumor which is used to balance a metric that combines both precision and recall. Especially the F1-score metric useful in handling class imbalance in medical dataset. To estimate the capability of the model, both Receiver Operating Characteristic (ROC) curves and Area under the Curve (AUC) are applicable. To signify better separation between tumor and non-tumor classes, it needs higher values of AUC.

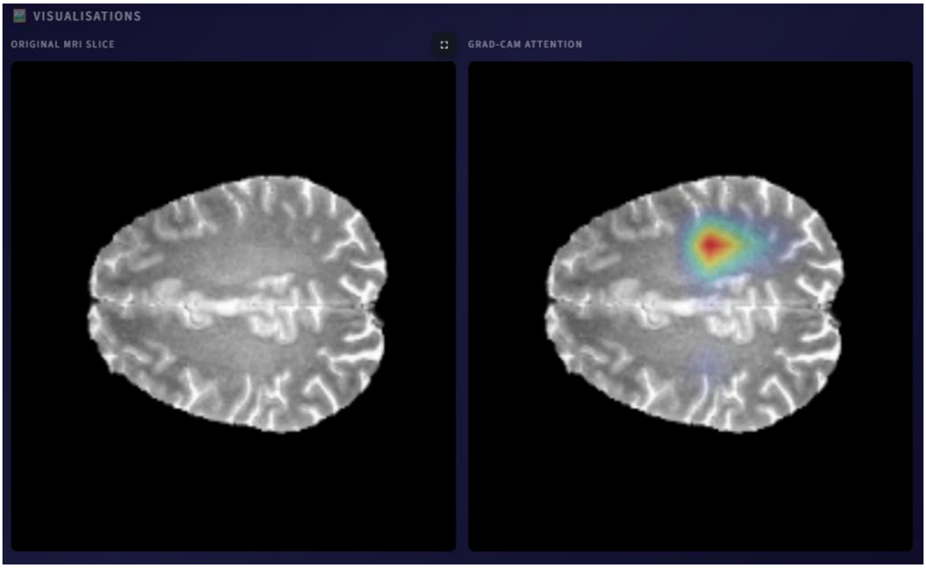

Fig. 4. Output of the proposed model when no tumor is present

The above Fig. 4 shows the output of the model when the input MRI slice doesn't contain tumor. It contains the original MRI slice and the heatmap produced for that result.

4 Results

The proposed multi task CNN model for tumor detection and segmentation is evaluated on a dataset containing multimodal MRI. The output of the model showcases the effective results of the proposed system in detection of presence of tumors and segmentation of brain tumors. This system accomplished high performance of detection and segmentation on the testing dataset as it is clearly shown in Table 1. The multi task strategy of learning improved the performance of the model as compared to the existing baseline U-Net model which is clearly shown in Table 1. Evaluating the model with the help of

F1-Score, recall, accuracy, precision, Dice Similarity Coefficient, Intersection over Union shows the overall performance of the system. The analysis of confusion matrix displays that majority of the presence or absence of tumors are identified correctly and only misclassification is occurred few times. This tells us that the model is able to learn the features which are discriminative from the MRI images.

The introduction of Explainable AI in the model provides the visualization of the predictions of model. The heat maps which are generated by the Grad CAM clearly shows the regions in the brain which are affected by tumors prompting that this model does not focus on back ground regions and aims on clinically relevant areas.

Table 1. Table captions should be placed above the tables.

Architecture	**Fold**	**Mean Region Dice**	**Detection F1**	**Detection AUC**
Existing Baseline U-Net	Fold 1	0.8355	0.8314	0.9485
	Fold 2	0.8566	0.8983	0.9710
Proposed Multi-task CNN	Fold 1	0.8915	0.9172	0.9804
	Fold 2	0.8915	0.9198	0.9798

This increases the trust on the model and also its interpretability. This is very much important for applications related to health care. Finally, the overall results of the model tell us that the proposed system has more accuracy, efficiency, computational capability and interpretability. This model can be used in detection and segmentation of brain tumors in the real world with clinical decision support.

5 DISCUSSIONS

The final results of the model show that the proposed system addresses the limitations of the tumor detection and segmentation framework using MR images of the brain. The performance of the model can be measured by shared learning feature representations and optimizing the brain tumor prediction and segmentation tasks in parallel. This strategy of multi task learning decreases the redundancy and increases the overall performance when compared to the traditional tumor detection models. The integration of the Explainable AI module using Grad CAM is the major advantage of this model. The visual representations given by the heat maps tells that the model concentrates on only important regions and that regions are clinically meaningful. Hence, the trust on the system is increased and clinicians can view the predictions of the model. This is very much important in medical based applications because this system also justifies the decisions made by the system. These obtained features plays a major role in locating the segment of the brain tumor in the MRI scan if the tumor exists.

Architecture of the model is proposed to be light weight such that there would be an increase of usage in the real world applications even by maintaining high accuracy levels. This model uses multimodal data. The dataset used here has 4 modalities. Regardless of its performance this model has some limitations. The efficiency of the model is effected by diversity and quality of datasets used for training and also this model uses 2D slices of MRI scans not but use fully 3D spatial context of MRI scans.

6 CONCLUSIONS

This research resulted in a framework based on deep learning for the detection and segmentation of brain tumors from MRI images with the support of deep convolutional neural network. This proposed model efficiently combines the preprocessing of images, extracting features from images, Explainable AI, multitask learning. By performing the two tasks in parallel, that is tumor detection and segmentation, this system aims to decrease the redundancy and improve the overall performance of the model. Based on the study, we found that the proposed approach possesses more performance in F1 score, DSC, IoU etc. These results show the strength of the model in identification as well as classification of tumors. The output of the model established that these deep CNN based models can also be utilized for medical imaging applications after performing necessary validation. Overall, this system provides a stable and clear method for automatic diagnosis of tumors in brain and this model also helps clinicians especially radiologists to save time. It can also be extended to deploy in clinics real world.

References

1. M., M.M., R., M.T., V., V.K., Guluwadi, S.: Enhancing brain tumor detection in MRI images through explainable AI using Grad-CAM with ResNet50. BMC Med. Imaging 24(1), 107 (2024). https://doi.org/10.1186/s12880-024-01292-7
2. Wong, Y., Su, E.L.M., Yeong, C.F., Holderbaum, W., Yang, C.: Brain tumor classification using MRI images and deep learning techniques. PLoS ONE 20(5), e0322624 (2025). https://doi.org/10.1371/journal.pone.0322624
3. Kumar, K., Jyoti, K., Kumar, K.: Machine learning for brain tumor classification: Evaluating feature extraction and algorithm efficiency. Discov. Artif. Intell. 4(1), 112 (2024). https://doi.org/10.1007/s44163-024-00214-4
4. Ali, A., Li, X., Mashwani, W.K., Abiad, M., Karim, F.K., Mostafa, S.M.: Multi-class brain tumor MRI segmentation and classification using deep learning and machine learning approaches. Cancer Imaging 25, 131 (2025). https://doi.org/10.1186/s40644-025-00953-2
5. Xie, Y.: Brain tumour MRI detection and classification based on the convolutional neural network. Appl. Comput. Eng. 15, 223–228 (2023). https://doi.org/10.54254/2755-2721/15/20230839
6. Nimmagadda, R., Devi, P.K.: A deep learning approach for brain tumour classification and detection in MRI images using YOLOv7. Front. Oncol. 15 (2025). https://doi.org/10.3389/fonc.2025.1508326
7. Rasool, N., et al.: CNN-TumorNet: Leveraging explainability in deep learning for precise brain tumor diagnosis on MRI images. Front. Oncol. 15 (2025). https://doi.org/10.3389/fonc.2025.1554559

8. Filvantorkaman, M., Piri, M., Torkaman, M.F., Zabihi, A., Moradi, H.: Fusion-based brain tumor classification using deep learning, explainable AI, and rule-based reasoning. arXiv preprint arXiv:2508.06891 (2025). https://doi.org/10.48550/arXiv.2508.06891
9. Yurdakul, M., Taşdemir, Ş.: MRI-based brain tumor detection through an explainable EfficientNetV2 and MLP-Mixer-attention architecture. arXiv preprint arXiv:2509.06713 (2025). https://doi.org/10.48550/arXiv.2509.06713
10. Patil, S., Kirange, D.: Ensemble of deep learning models for brain tumor detection. Procedia Comput. Sci. 218, 2468–2479 (2023). https://doi.org/10.1016/j.procs.2023.01.222