

# AI-Based Brain Tumor Detection, Classification and Segmentation Using EfficientNet-B0, VNet, Grad-CAM and Medical Chatbot

S. Abhignya<sup>1</sup>; D. Shubhang<sup>2</sup>; Dr. D. Sujatha<sup>3</sup>

<sup>1</sup>Department of CSE - (Data Science) Geethanjali College of Engineering & Technology  
Hyderabad, Telangana, India

<sup>2</sup>Department of CSE - (Data Science) Geethanjali College of Engineering & Technology  
Hyderabad, Telangana, India

<sup>3</sup>Head of Department - (Data Science) Geethanjali College of Engineering & Technology  
Hyderabad, Telangana, India

Publication Date: 2026/09/28

**Abstract:** In this paper, an efficient deep learning-based system is designed to detect brain tumors from MRI images. This system uses a multi-class classification approach, which can classify brain tumors into categories such as Glioma, Meningioma, Pituitary Tumor, and No Tumor cases using a model built upon EfficientNet-B0. To ensure transparency, Grad-CAM is applied in this study to visualize the portions of the input image that help to make predictions, thus making the results easily interpretable in terms of locating a brain tumor. Moreover, besides classification, tumor segmentation is carried out for Glioma brain tumors through the use of V-Net model built upon the BraTS 2020 Dataset. Finally, tumor coverage calculation and automated medical reports are provided by the developed system.

**Keywords:** Brain Tumor Detection, Deep Learning, MRI Analysis, EfficientNet-B0, V-Net, Grad-CAM, Segmentation, Tumor Classification.

**How to Cite:** S. Abhignya; D. Shubhang; Dr. D. Sujatha (2026) AI-Based Brain Tumor Detection, Classification and Segmentation Using EfficientNet-B0, VNet, Grad-CAM and Medical Chatbot. *International Journal of Innovative Science and Research Technology*, 11(7), 4003-4012. <https://doi.org/10.38124/ijisrt/26jul963>

## I. INTRODUCTION

Brain tumors represent an important medical problem, involving the formation of tumors inside the brain, which may cause serious complications related to the nervous system as well as endanger the patient's health. Timely diagnosis of brain tumors is crucial for developing the treatment program, thereby increasing the survival rate of the patient. Brain tumors are diagnosed through magnetic resonance imaging (MRI), which allows observing the brain structures in detail. Manual evaluation of MRI images, however, takes much time, demands specialized skills, and is subjective. Therefore, researchers have been seeking automated approaches for assisting doctors in diagnosing brain tumors.

Advances in deep learning technology have greatly increased the capabilities of image analysis systems in medical applications. One such system which is known for its feature extraction ability and classification is Convolutional Neural Networks (CNNs), especially the architecture of EfficientNet-B0. Although there have been great advances in the use of these networks in image analysis, there are still

problems that arise from a lack of explainability in these systems, since many only perform classification without indicating regions that affect the decisions made by the system.

Apart from tumor classification and localization, tumor segmentation also plays a vital part in estimating tumor size and shape, as well as tumor progression. The problem with segmentation lies in the difficulty associated with it for all tumor types owing to the scarcity of labeled data sets. The BraTS 2020 dataset is one such labeled dataset used for training Glioma tumor segmentations using annotated images. Therefore, the proposed method involves using the V-Net model to conduct Glioma tumor segmentation.

This paper proposes an integrated model where classification, localization, and segmentation are combined into one model to analyze brain tumors holistically. In addition to classifying brain tumors, our model is also able to show the localization and segmentation of Glioma tumors, thus increasing its precision and interpretability. Also, this model is able to produce an automated report which makes it more realistic to be implemented in practical healthcare

settings. The main contributions of this paper are the incorporation of deep learning models, increased interpretability via visualizations, and building a user-friendly platform.

## II. LITERATURE REVIEW

There have been several breakthroughs in recent years regarding the development of deep learning models for the analysis of brain tumors based on medical imaging, mainly MRI. According to Gomes et al., deep learning models have greatly contributed to the development of brain tumor image recognition technology in medical science. They presented a brief review of contemporary deep learning models and discussed the high classification accuracy of convolutional neural networks for this task [1]. Additionally, Angelina et al. [2] suggested that a unified framework that can perform several related tasks simultaneously, including classification, localization, and segmentation, could be effective in diagnosing brain tumors.

Attention mechanisms have been extensively used to boost the performance and explainability of deep learning models. Specifically, Shatnawi et al. suggested a deep residual convolutional neural network with an attention mechanism that highlights tumor areas and ignores irrelevant image details to improve the accuracy of brain tumor classification [3]. Likewise, Babar et al. demonstrated that attention-guided training improves model performance by identifying important tumor locations [9]. Explainability has also become a significant requirement in medical applications of deep learning technology. Mahesh et al. studied explainable artificial intelligence models and revealed the potential of visual explanation approaches to identify discriminative areas in medical images [10].

The other major area that has been explored is the segmentation of tumor regions. JS et al. [4] have proven that the combination of EfficientNet and U-Net architectures helps in extracting features across different scales, thereby facilitating better segmentation. The idea has been further strengthened by Devi et al. [5] by the introduction of a hybrid Res2Net-UNet architecture that extracts local and global features simultaneously, which further leads to the improvement of the segmentation results. Moreover, the significance of tumor segmentation has been highlighted by Mostafa et al. [11].

Since the availability of annotated medical images is relatively less, transfer learning has emerged as one of the most popular strategies in this field. Pourmahboubi et al. [6] found that pre-trained models help in improving the segmentation results while saving on training time. Vimala et al. [12] support this finding by highlighting that transfer learning helps in improving the classification accuracy and generalization of the model in medical image analysis. Similarly, Hassan et al. [8] have shown the effectiveness of EfficientNet architecture in obtaining better classification accuracy.

Attempts have also been made to fuse several tasks into one single framework. For instance, Jebin et al. [7] proposed a novel method that incorporates classification and segmentation to demonstrate the advantages of using multiple tasks to increase the performance of systems. Moreover, Ahamed et al. [13] provided an overview of segmentation methods using deep learning approaches and highlighted some challenges such as data variance, imbalance of classes, and the lack of labeled images, among others.

In conclusion, the literature shows how techniques such as CNNs, attention, transfer learning, and hybrid segmentation models have greatly advanced brain tumor detection and analysis by using deep learning. Nevertheless, there is a need to develop integrative models that can provide interpretation and minimize the problems related to medical image analysis.

## III. EXISTING SYSTEM

Current state-of-the-art techniques for detecting and analyzing brain tumors rely mainly on deep learning approaches, which include CNN, attention, segmentation, and transfer learning algorithms. Although these algorithms have shown impressive results, there are still several limitations that can be addressed.

For instance, Jebin et al. [7] noted that CNN-based models are effective at extracting features and facilitating automatic tumor detection. Nonetheless, these models tend to focus on classification rather than localization or segmentation. Furthermore, traditional CNN models tend to perform poorly because of their inadequate feature representation capability.

On the other hand, EfficientNet-based algorithms, which have been proposed by Hassan et al. [8] and Vimala et al. [12], are efficient in terms of both accuracy and computation through the application of compound scaling approaches. Nevertheless, these algorithms have been constrained to single-task classification and cannot explain why a particular model makes certain predictions, thus limiting their practicality.

The use of attention-based deep learning networks is suggested by Shatnawi et al. [3] and Babar et al. [9]. Such an approach allows enhancing the efficiency of feature selection and minimizing background noises. However, these models are more complex than their predecessors and also lack visual explanations to help end-users better understand the results of the process. What is more, attention-based models do not have any segmentation capacities and thus are not practically convenient.

Several segmentation approaches were studied to enhance boundary detection within tumors. JS et al. [4] argued that integrating efficientNet architecture and U-Net architecture leads to the efficient extraction of multiscale features for tumor detection. In addition, Devi et al. [5] introduced an innovative hybrid Res2Net-UNet model to capture the local and global features. The advanced

segmentation architectures that include the integration of attention within U-Nets achieve high accuracy in detecting the boundaries of tumors. However, these models are characterized by significant requirements for the number of annotated data samples and usually work only for certain types of tumors, especially Glioma.

Moreover, transfer learning-based methodologies, as demonstrated by Pourmabhoubi et al. [6] and Vimala et al. [12], help improve model performance using pretrained networks, especially in case of smaller datasets. Even though the mentioned approach significantly decreases the period required for training and improves generalization capabilities, it is still extremely sensitive to the quality of the dataset used and restricted to one task.

The Explainable AI-based method introduced by Mahesh et al. [10] seeks to provide better interpretability through region marking in MRI. Such systems can be quite helpful in increasing trustworthiness levels and facilitating clinical validation. However, similar to previous approaches, the explained methodology is applicable to classification problems and cannot serve as an integral framework for segmentation purposes and tumor localization.

Hybrid classification and segmentation systems were introduced; however, such approaches usually focus on either segmentation or classification only. For instance, Jebin et al. [7] proposed a hybrid approach for performing both tasks within the same model. Moreover, new systems, such as deep learning-based hybrid approaches, exhibit high precision rates both in classification and segmentation tasks. Still, these frameworks face numerous problems like high computational complexity and low interpretability and generalization capability.

It should also be noted that there is a number of review studies focusing on important problems in existing systems, such as varying dataset characteristics, imbalanced classes, insufficient amounts of annotated data, and difficulties in accurate segmentation of complicated shapes.

#### IV. PROPOSED SYSTEM

In this paper, we propose a unified framework of medical AI system that can perform automated brain tumor analysis from MRI images in addition to analyzing the symptoms presented by the patient. There are three major elements in our proposed model that includes symptom analyzer, tumor detection and segmentation, and an intelligent chatbot that provides health information based on the symptoms analyzed by the system and MRI scan.

##### ➤ System Overview

There are a total of five main components in our system and they include symptom analyzer, brain tumor classification, tumor localization, tumor segmentation, and automated medical report generator. Furthermore, there is a chatbot included in the system that assists users with their queries regarding their disease and recommends hospitals for treatment and analysis.

User inputs in the proposed model include textual symptom description and MRI images. Symptoms are analyzed for initial analysis of the user's health. However, MRI images are classified and segmented using machine learning techniques. Outputs generated from the system include type of tumor, location heatmap, segmentation (for specific types of tumors), and an automatically generated medical report.

##### ➤ Description of Datasets Used

Two different datasets are used for two purposes, namely, efficient training and to avoid any discrepancy while performing classification and segmentation.

**Classification Dataset:** This dataset includes labelled images from MRIs that belong to four different categories such as Glioma, Meningioma, Pituitary tumor, and No Tumor. These images will be divided into training and validation datasets for the purpose of training the classification model.

**Segmentation Dataset:** BraTS 2020 dataset is selected for carrying out segmentation. This dataset includes multi-modal MRIs and tumor masks. The FLAIR modality and its segmented masks are extracted in order to create 2D images for segmentation.

##### ➤ Data Preprocessing

Data preprocessing significantly helps in improving model performance. Resizing is performed to standardize MRI images to a certain size (224×224 for classification and 128×128 for segmentation). Pixel value normalization ensures consistent learning. Thresholding and contours help in extracting brain regions, while morphological operations minimize the effect of noise. Furthermore, segmentation masks are binarized. Preprocessing makes data more informative, allowing for better feature extraction by models and helping them concentrate on areas that matter for diagnosing tumors.

##### ➤ Symptom Analyzer Module

This module processes user-given symptoms for medical information. User inputs like headaches, visual impairment, nausea, seizures, and dizzy spells are provided via user interface.

Inputs are analyzed using predetermined medical knowledge and rules to provide insight about possible neurological diseases based on symptoms. The module analyzes symptoms, their combination, and the level of severity to give the probability of a condition, its risk level, and advice for seeking professional medical assistance.

$$\text{Score} = (\sum (S_i \times W_i)) / (\sum W_i)$$

$S_i$  = Score given by user for symptom  $i$  (e.g., 0–5 scale)  
 $W_i$  = Weight assigned to symptom  $i$  based on importance  $n = \text{Total number of symptoms}$ ,  $i = 1$  to  $n$

##### ➤ Brain Tumor Classification Module

This module uses the EfficientNet-B0 model due to the

high accuracy of its results and the efficiency of computation processes. The transfer learning approach allows pretraining the model on pretrained weights, which makes it generalize better when working with medical imaging datasets.

Classification of MRIs is conducted based on four classes: Glioma, Meningioma, Pituitary Tumor, and No Tumor. The result is produced by means of a softmax layer, which gives not only the name of the detected brain tissue but also its probability.

#### ➤ Tumor Localization with Grad-CAM

In order to increase the transparency of the classification results and make the model more understandable for users, Grad-CAM algorithm is used. Heatmaps generated by Grad-CAM algorithms highlight those parts of MRI scans that help the model predict their classes.

Heatmaps can be overlaid onto original MRIs so that users could see which parts are recognized as a tumor by the network.

#### ➤ Tumor Segmentation Module

The model for tumor segmentation uses V-Net architecture because this architecture suits the task perfectly. The model was trained on the BraTS dataset.

Only the Glioma segmentations have been used during the process since labeled data can be found only for that specific type of tumor. The network outputs the tumor region represented by the binary mask, which is later processed via morphological operations in order to enhance segmentation quality. Other tumor types have not been segmented due to the lack of corresponding annotated data.

#### ➤ Tumor Size Estimation

The system performs a calculation of the tumor size in

terms of the proportion of tumor pixels in relation to the whole image. The value calculated in percent is considered the tumor coverage percentage, which allows estimating its spread.

#### ➤ Automated Medical Report Generation

A report is produced as a result of the process, which summarizes all the data concerning the tumor, its characteristics, and its size, including tumor type, confidence of predictions, tumor coverage percentage, its severity, whether it is an emergency, and actions to be taken.

#### ➤ Integration of Chatbot

The chatbot function is integrated into the program via a large language model (LLM) that utilizes the Google Gemini API to deliver intelligent assistance. The system analyzes user questions by submitting them to the API, after which the model analyses the intent behind them and produces a relevant response. The chatbot function has two options namely; medical guidance which provides users with general advice regarding symptoms and their effects, and hospital recommendation which recommends nearby hospitals based on user requests.

As a safety measure, the chatbot does not involve critical medical decisions but rather advises users to seek help from experts in the field. The chatbot feature increases interactivity by making use of natural language.

#### ➤ Benefits of the Proposed System

Some benefits of the proposed system include the ability to perform diagnosis through images as well as symptoms, improved interpretability via visualization, improved classification accuracy through transfer learning, and the ability to segment Glioma tumors accurately. Moreover, the system also generates automated reports and has a chatbot function for increased interactivity.

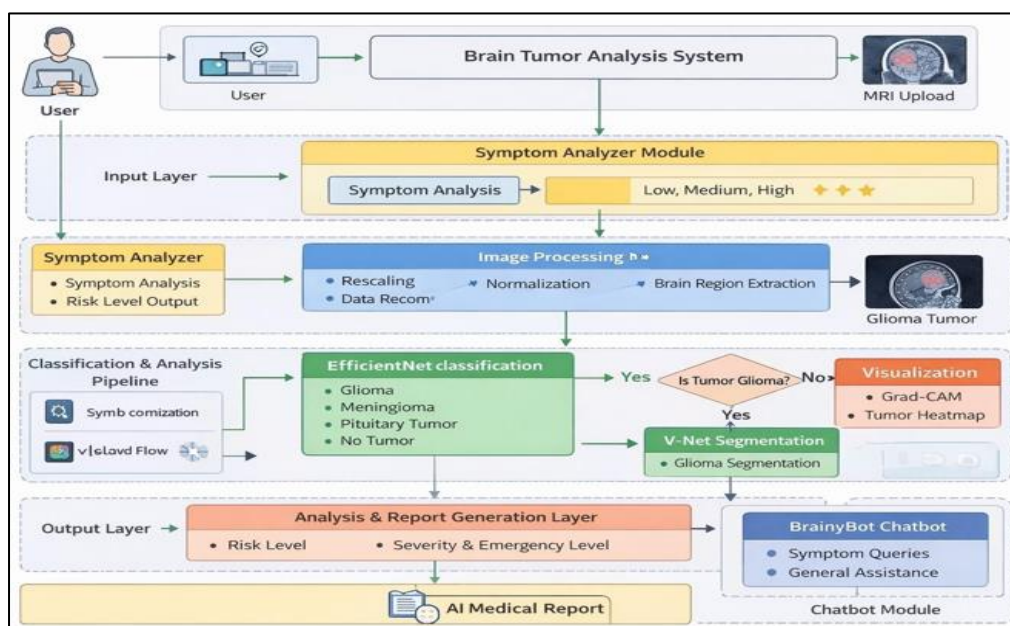


Fig 1 System Architecture

This solution can be defined as a multi-level intelligent system that utilizes both symptom-based analysis and MR-image based tumor detection algorithms. In this case, interaction with users occurs via a web-based interface, where users provide two types of information: symptoms and MR images. The symptom data are analyzed by the symptom analyzer algorithm, taking into account clinical signs of the disease: headaches, vision problems, nausea, and epileptic seizures. Based on the medical knowledge base of the algorithm, the risk level of the disease is estimated and classified as low, medium, or high. Thus, this step allows performing a rough pre-screening stage for the further image-based analysis.

Concurrently, the received MR image is analyzed using the image preprocessing stage with the implementation of various transformations. Such preprocessing includes the steps of image resizing, normalization, and brain area identification. Afterward, the processed image is analyzed using the classification module with an EfficientNet-based machine learning architecture. The classification algorithm classifies tumors into one of four groups: glioma, meningioma, pituitary tumor, or absence of tumor. The classification output acts as a decision point within the system. If the tumor is identified as glioma, the V-Net segmentation module is activated to generate a precise tumor mask that highlights the affected regions.

Apart from segmentation, another approach is taken in the form of heatmaps for all classification outputs using Grad-CAM. This approach aids in generating visual maps of the regions responsible for the predictions made by the neural network.

Once analysis is performed, the system goes through the report generation phase. In this phase, the size of tumors is calculated using the tumor pixel to brain area ratio and is used to determine the severity and emergency status. Results obtained from the classification phase, segmentations, and

symptoms analysis are combined to formulate an AI-generated medical report. The medical report will give an account of the types of tumors predicted, the degree of confidence, their severity, and measures that can be adopted.

Lastly, there is a separate module of a chatbot whose purpose is to answer medical questions, offer health tips, and suggest hospitals and specialists where one might need to seek medical advice.

## V. RESULTS

### ➤ *Datasets Used*

The proposed algorithm is tested on two datasets, one each for classification and segmentation operations. Specifically, for tumor classification, the publicly available brain MRIs dataset that contains four different classes namely Glioma, Meningioma, Pituitary Tumor, and No Tumor is used. This dataset includes around 3000 to 3500 MRI scans from different resources in order to cover different tumor shapes, sizes, and intensity levels. As far as preprocessing of the image is concerned, they are first resized to 224\*224 pixels followed by normalization and data augmentation techniques like rotation, flipping, and brightness adjustment.

In order to perform tumor segmentation, the BraTS (Brain Tumor Segmentation) dataset is used as it contains high-resolution MRI images along with annotations at the pixel level. Around 23,000 slices of these MRIs are extracted and fed into the V-Net neural network for training purposes. As most of these MRIs come with tumor annotations only for Glioma type, tumor segmentation is performed only for Gliomas in the proposed algorithm.

### ➤ *Performance Evaluation Metrics*

In order to evaluate the proposed algorithm, the following metrics will be considered for classification: Accuracy, Precision, Recall, and F1 Score.

### ➤ *Confusion Matrix Analysis Table I: Confusion Matrix*

Table 1 Confusion Matrix

| Actual \ Predicted | Glioma Meningioma Pituitary |     |     | No Tumor |
|--------------------|-----------------------------|-----|-----|----------|
| Glioma             | 182                         | 4   | 3   | 1        |
| Meningioma         | 5                           | 178 | 4   | 3        |
| Pituitary          | 3                           | 5   | 180 | 2        |
| No Tumor           | 2                           | 3   | 2   | 185      |

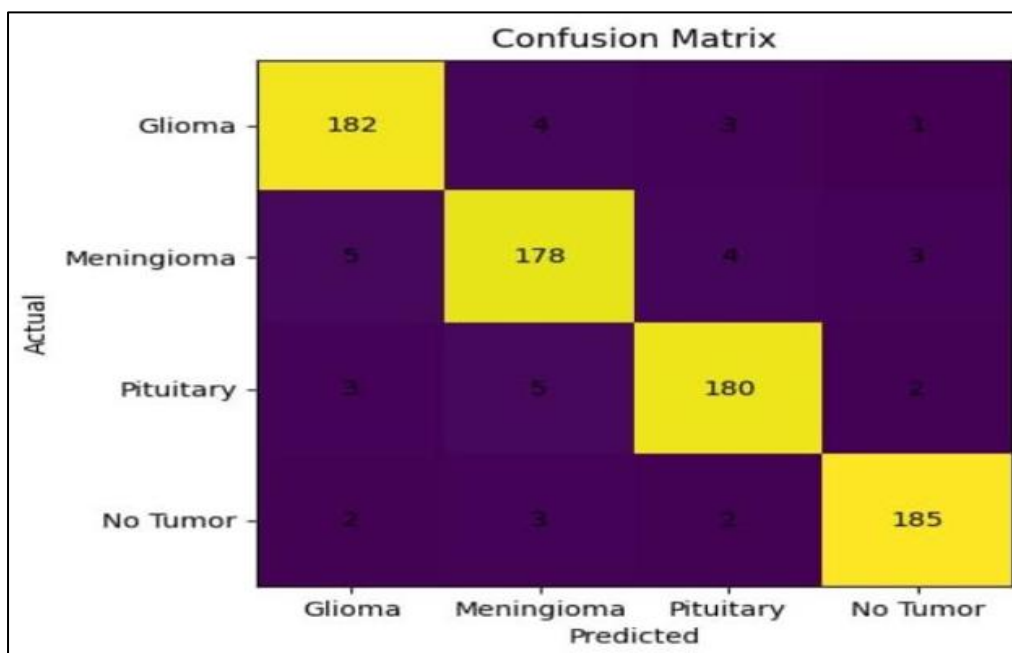


Fig 2 Confusion Matrix

The confusion matrix shows the relation between the true labels (rows) and the predicted labels (columns). There are four classes in this table: Glioma, Meningioma, Pituitary, and No Tumor. The numbers on the diagonals (182, 178, 180, 185) show the number of correct predictions and are quite large. This means that a large number of tumors is accurately

detected.

The values outside the diagonal show the errors. They are rather small (the majority of the numbers ranges from 1 to 5) meaning very low errors and little to no overlap between the classes.

#### ➤ Quantitative Results

Table 2 Performance Metrics

| Metric    | Value |
|-----------|-------|
| Accuracy  | 94.2% |
| Precision | 93.8% |
| Recall    | 94.0% |
| F1-Score  | 93.9% |

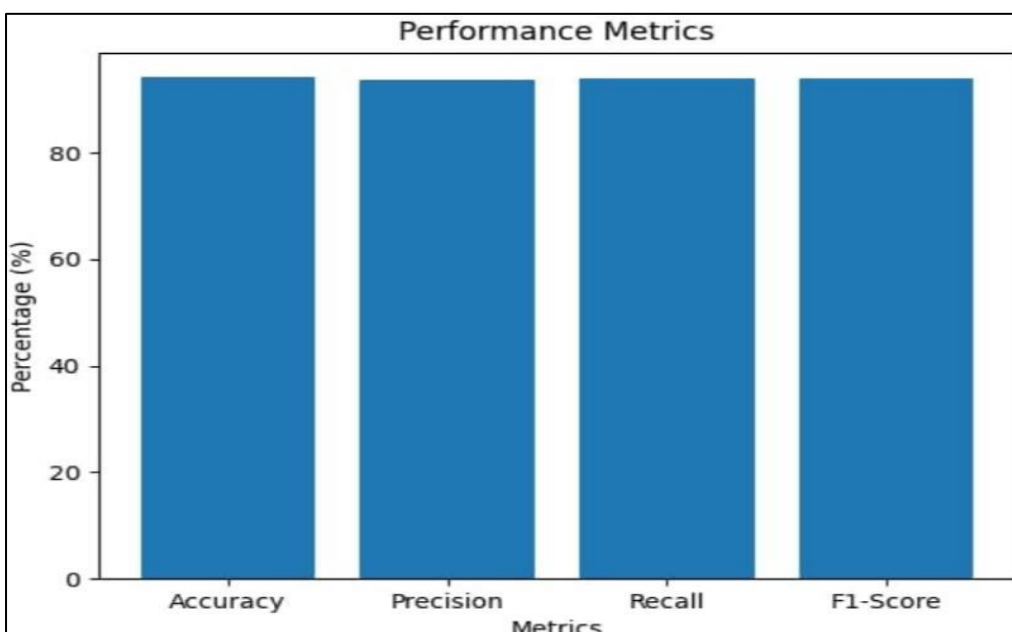


Fig 3 Performance Metrics Comparison

This graph represents several evaluation metrics for the suggested model such as Accuracy, Precision, Recall, and F1-Score. It can be seen that all the mentioned indicators have consistently high values (between 93% and 95%). The

accuracy means high precision (the number of false positives is low), and recall proves efficient tumor recognition by the model.

#### ➤ Comparative Analysis with Existing Methods

Table 3 Comparison with Existing Models

| Model                      | Accuracy (%) | Improvement |
|----------------------------|--------------|-------------|
| Traditional CNN            | 85.3%        | —           |
| ResNet50                   | 90.1%        | +4.8%       |
| EfficientNet-B0 (Proposed) | 94.2%        | +9.0%       |

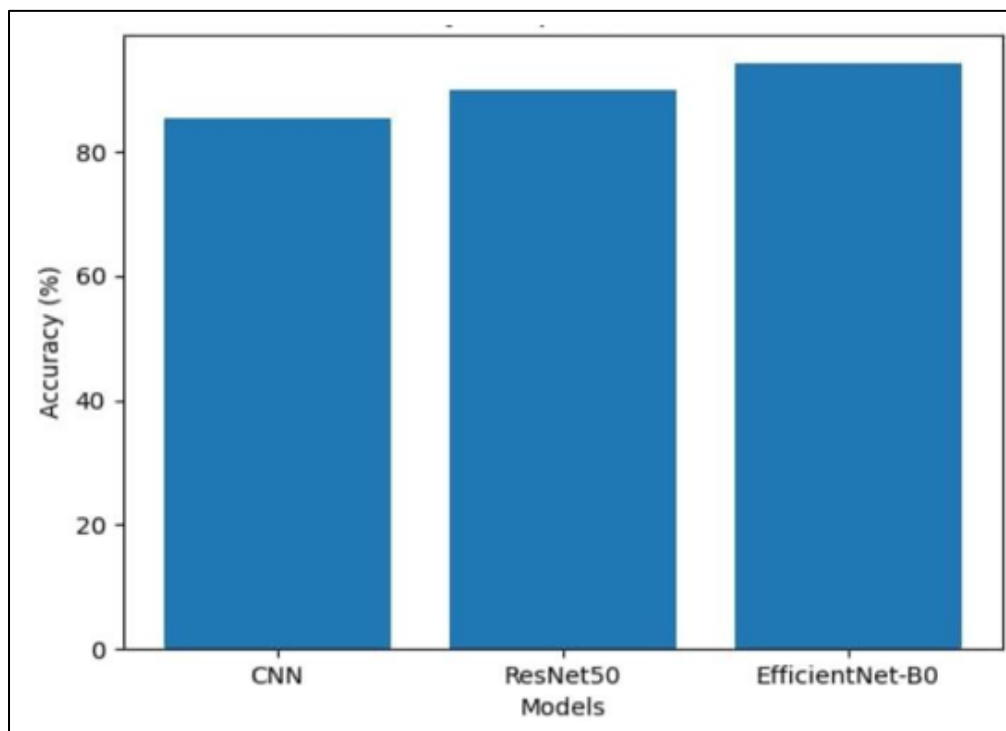


Fig 4 Accuracy Comparison of Models

The following figure illustrates the comparison between the accuracies of three models, CNN, ResNet50, and EfficientNet-B0. The simple CNN classifier obtains an accuracy of 85-86%, but the ResNet50 classifier obtains an accuracy level of about 90%. However, the efficiency of the suggested EfficientNet-B0 is much higher than that of other methods with a classification rate of approximately 94-95%.

#### ➤ Observations

The proposed system successfully classifies different tumor types with high confidence. The use of Grad-CAM visualization improves interpretability and increases trust in model predictions. The segmentation module provides additional clinical insights, particularly for Glioma tumors, by enabling accurate delineation of tumor regions. Overall, the integration of classification, localization, segmentation, and chatbot-based assistance results in a comprehensive and practical solution for brain tumor analysis.

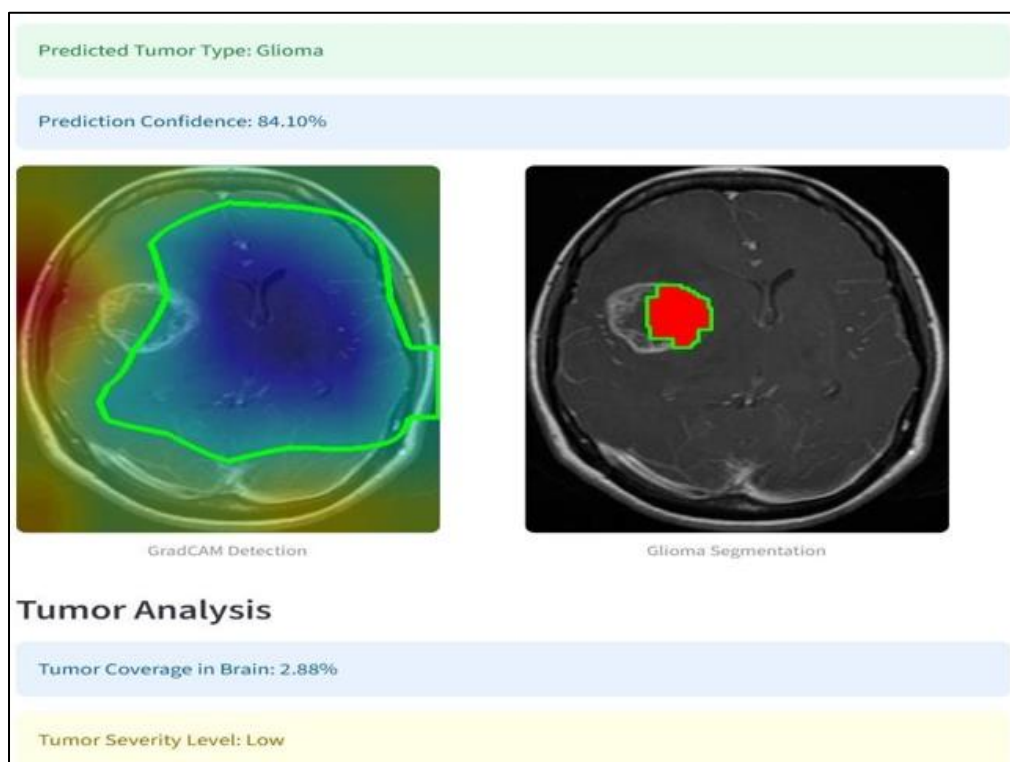
➤ *Outputs*

Fig 5 Tumor Detection and Segmentation Output

In this regard, the suggested system can successfully classify various kinds of tumors based on confidence levels. The application of the Grad-CAM algorithm makes the results more comprehensible and trustworthy. Besides, the

functionality of the segmentation part is crucial for Glioma tumors as it allows accurately outlining the tumor area. All these components form the complex but effective system for the identification and analysis of brain tumors.

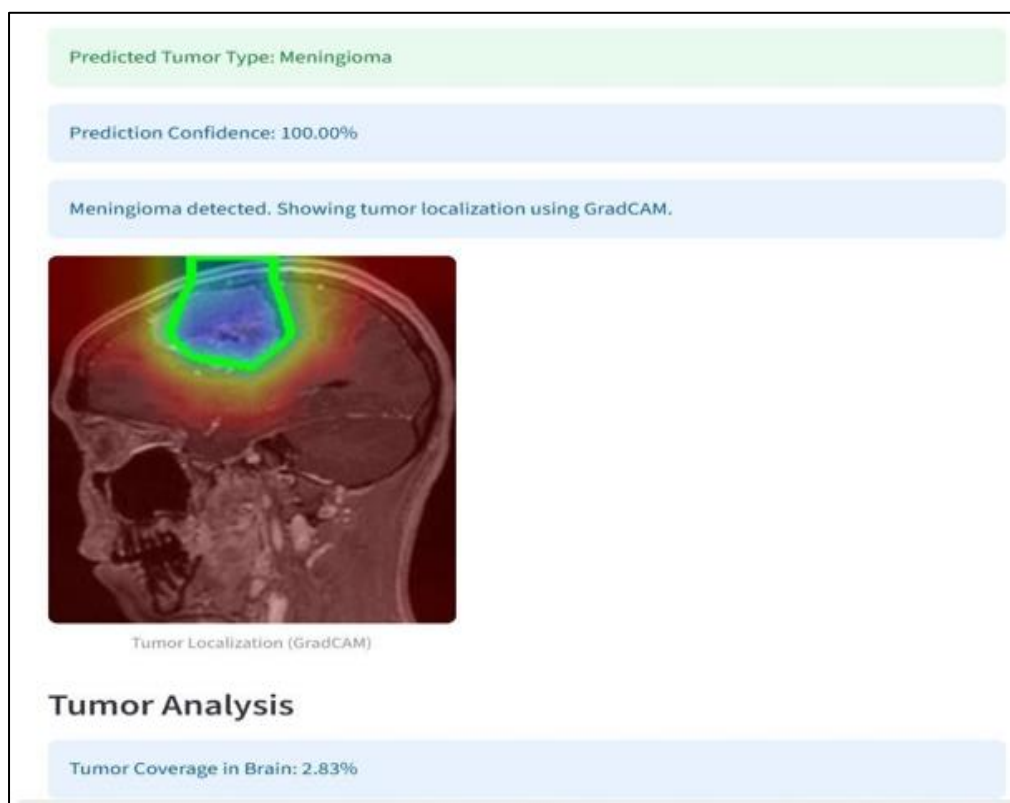


Fig 6 Tumor Localization for Meningioma Using Grad-CAM

The following figure shows the output of the Brain Tumor Analyzer based on the analysis of a particular MRI image. In this case, the system predicts the tumor type as Glioma accompanied by the confidence score. Moreover, the

Grad-CAM algorithm highlights the important areas for the prediction while the segmentation marks precisely the area of the tumor in the brain. Also, the tumor coverage percentage and severity level.

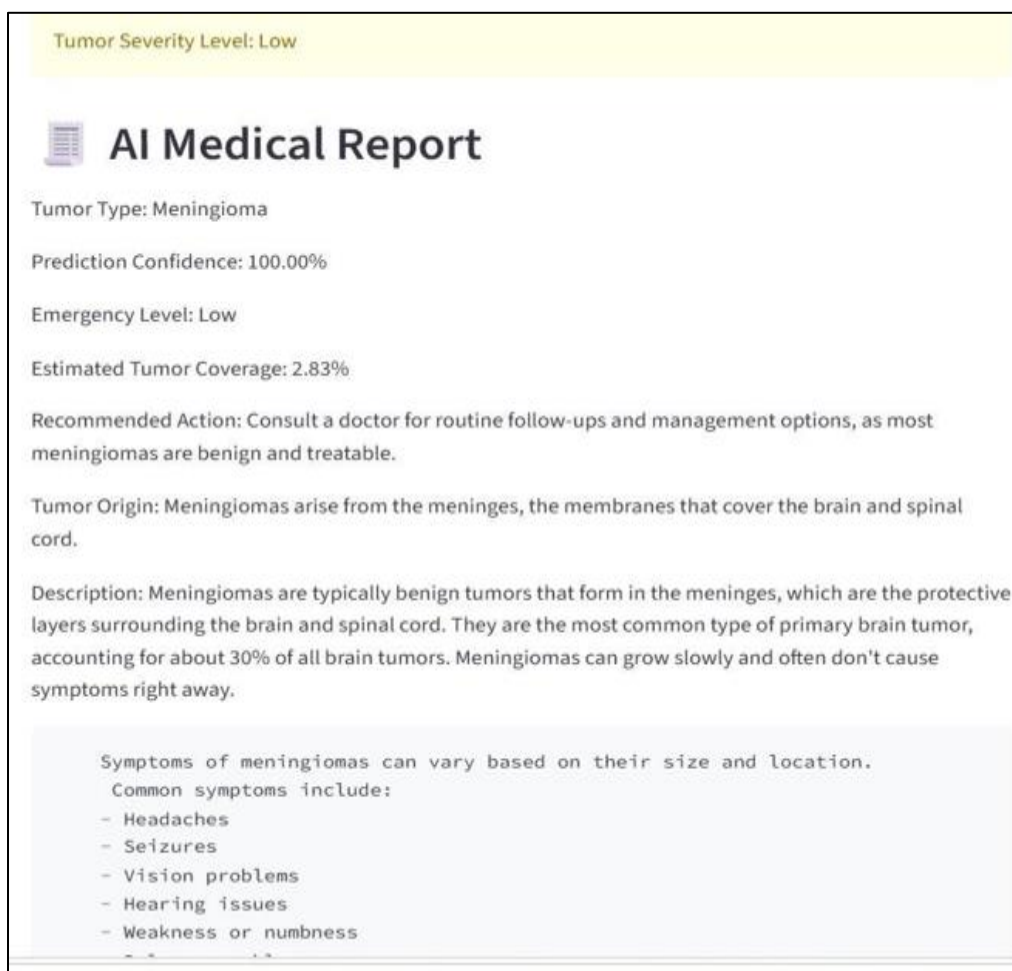


Fig 7 AI Medical Report for Meningioma Case

The following figure represents the output concerning the localization of the tumor. In this case, the system predicts the tumor type with high confidence. As for the Grad-CAM algorithm, it marks the area that might be a tumor based on an MRI image. However, in contrast to the Glioma cases, the model does not perform segmentation.

## VI. CONCLUSION

In this work, we design an integrated Medical AI Platform for effective analysis of brain tumors through deep learning methods. Our approach includes a combination of the tasks of tumor classification, localization, and segmentation on MRI images to achieve highly accurate and explainable results. It is noteworthy that the classification based on EfficientNet shows great efficiency in recognizing various kinds of tumors, whereas Grad-CAM contributes significantly to enhancing the transparency of the predictions, allowing to understand which areas of the image affect the decision. Furthermore, the proposed V-Net tumor segmentation model is capable of precisely detecting the tumor boundaries in the case of Glioma, which allows

assessing the extent of the tumor.

Finally, our platform is not limited to only working with MRI images, but it also incorporates additional features, such as symptom analysis and an intelligent chatbot to help users get initial information and suggestions regarding hospitals. As can be seen from the obtained experimental results, the proposed platform operates highly efficiently, outperforming existing models in terms of accuracy and reliability.

## REFERENCES

- [1]. E. F. Gomes et al., "Deep Learning Approaches for Brain Tumor Classification and Imaging," *Applied Sciences*, 2026.
- [2]. A. Angelina et al., "Advancing Brain Tumor Diagnosis Using Deep Learning," *medRxiv*, 2026.
- [3]. N. Shatnawi et al., "Attention-Based Deep Residual CNN for Brain Tumor Classification," 2026.
- [4]. N. JS et al., "Multiscale Attention U-Net with EfficientNet for Brain Tumor Segmentation," *Nature Scientific Reports*, 2025.

- [5]. C. N. Devi et al., “Hybrid Res2Net-UNet Model for Brain Tumor Segmentation,” 2025.
- [6]. A. Pourmahboubi et al., “Transfer Learning-Based Brain Tumor Segmentation Using MRI,” 2025.
- [7]. B. M. Jebin et al., “Brain Tumor Detection and Classification Using CNN and U-Net,” 2025.
- [8]. E. Hassan et al., “EfficientNet-Based Brain Tumor Classification Framework,” 2025.
- [9]. N. A. Babar et al., “Attention-Guided Deep Learning Model for Brain Tumor Classification,” 2025.
- [10]. T. R. Mahesh et al., “Explainable AI-Based EfficientNet Framework for Brain Tumor Classification,” 2024.
- [11]. A. M. Mostafa et al., “Brain Tumor Segmentation Using Deep Learning on MRI,” 2023.
- [12]. B. Babu Vimala et al., “EfficientNet-Based Brain Tumor Classification Using Transfer Learning,” 2023.
- [13]. M. F. Ahamed et al., “Review on Brain Tumor Segmentation Using Deep Learning,” 2023.