iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: IV Month of publication: April 2025

DOI: https://doi.org/10.22214/ijraset.2025.68888

Smart Neuro-Oncology Assistant: A Chatbot for Brain Tumour Detection and Primary Cancer Prediction

Kartik Deogire[1], Dhanashree Patil[2], Sahil Dhake[3], Shreevallabh Chidrawar[4], Omkar Jagtap[5]
[1, 2, 3, 4, 5] *Computer Engineering, Pimpri Chinchwad College Of Engineering Pune, India*

Abstract: The Smart Neuro-Oncology Assistant introduces a novel AI-powered platform that combines deep learning-based brain tumor detection with primary cancer prediction capabilities, particularly for tumors associated with Glioma and Pituitary regions. This interactive system supports patients and healthcare professionals throughout the neuro-oncological assessment process. Utilizing the Xception convolutional neural network architecture, the system achieves 96% overall accuracy in classifying brain MRI scans into four categories: Glioma, Meningioma, Pituitary Tumor, and No Tumor. The platform extends beyond basic classification by evaluating cancer likelihood and delivering personalized guidance through an interactive chatbot powered by Google's Gemini 1.5 Flash generative AI model.

Key features include secure user authentication, robust image validation, specialist recommendations based on detected tumor types, and comprehensive historical data tracking for longitudinal assessment. Developed using Streamlit for the frontend interface, Flask for backend services, and MongoDB for efficient data management, the system demonstrates exceptional performance with class-specific F1-scores ranging from 0.93 to 0.98. This research contributes to AI-assisted medical diagnostics by promoting a human-centered approach that supports clinical expertise, aiming to streamline diagnostics and enhance patient understanding in neuro-oncology.

Keywords: Brain tumor classification, primary cancer prediction, convolutional neural networks, MRI scans, glioma, meningioma, pituitary tumor, no tumor, Streamlit, Flask, MongoDB, user authentication, diagnostic assistance, patient engagement, healthcare AI, image-based diagnosis, oncologist recommendation, tumor detection system.

I. INTRODUCTION

The need for intelligent systems that can identify tumors and offer more in-depth information about their nature is rising as neuro-oncological disorders become more common. The effectiveness of traditional diagnostic tools to assist patients in the diagnosis and treatment process is sometimes limited by their lack of integration and interactivity.

In order to help users comprehend their health, this project presents the Smart Neuro-Oncology Assistant: A Chatbot for Brain Tumour Detection and Primary Cancer Prediction, an AI-driven platform that integrates medical picture analysis with an interactive chatbot interface.

The system leverages deep learning techniques to classify brain tumors from MRI scans into four main categories: Glioma, Meningioma, Pituitary Tumor, and No Tumor. Beyond classification, it offers a unique layer of cancer prediction:

Glioma is considered a primary brain cancer. Pituitary tumors have both cancerous and non-cancerous possibilities, requiring further medical evaluation. Meningioma and No Tumor are categorized as non-cancerous in the scope of this application.

Upon detecting a tumor, the assistant evaluates its cancerous likelihood. If cancer is predicted, the system recommends a randomly assigned oncologist from a predefined list and provides personalized guidance through a conversational chatbot. Users can ask questions, explore treatment options, and compare current and previous MRI results with percentage-based analysis.

The platform is built using Streamlit for the frontend, Flask for backend services, and MongoDB for storing user data, chat history, and diagnostic results.

With features like secure user authentication, intuitive chat-based interaction, image upload capabilities, and historical MRI comparisons, the Smart Neuro-Oncology Assistant aims to deliver a comprehensive, user-friendly solution for brain tumor detection and primary cancer prediction.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

II. LITERATURE REVIEW

A. CNN-Based Approaches

Convolutional Neural Networks (CNNs) have shown strong performance in brain tumor classification. Mahmud et al. [1] proposed an ensemble combining traditional CNN and VGG16, achieving 97.68% accuracy on validation data and 98.01% on test data when classifying MRI scans into Glioma, Meningioma, Pituitary tumor, and No Tumor. Their method used array transformation for feature extraction without complex preprocessing like skull stripping or segmentation.

Paper [5] introduced a Recurrent Convolutional Neural Network (RCNN) that identified 96.21% of brain tumor tissues. Their multi-stage pipeline used adaptive filters for preprocessing, Improved Means Clustering (IMC) for segmentation, and Gray Level Co-occurrence Matrix (GLCM) for feature extraction before final classification.

Paper [4] presented a hybrid approach combining InceptionV3 with a fine-tuned ResNet50-152 layer. It achieved a precision of 0.98, recall of 0.95, F1 score of 0.93, and accuracy of 0.94 for non-tumor cases, and 0.87 precision, 0.92 recall, 0.88 F1 score, and 0.96 accuracy for tumor cases, leveraging both networks for enhanced performance.

B. Transformer-Based and Ensemble Approaches

Recent methods have explored transformers and ensemble learning to boost classification. Paper [2] compared EfficientNetV2 and Vision Transformer (ViT) for multi-class brain tumor detection, with EfficientNetV2 achieving 95% accuracy and ViT reaching 90%. Their highest result, 96%, came from a geometric mean ensemble of both models, highlighting the effectiveness of hybrid techniques.

Paper [6] examined how different activation functions influence CNN performance using a dataset of 4,468 MRI images spanning 17 brain tumor classes. Their findings stressed that activation functions enabling models to capture complex patterns and enhance learning while managing computational complexity.

C. AI-Powered Medical Assistants and Chatbots

Paper [7] details a chatbot using OpenAI's GPT-3.5 that provides personalized diagnoses and advice based on user symptoms. It achieved a response accuracy of 97.43% and response times below 3.1 seconds, illustrating the utility of large language models (LLMs) in healthcare communication.

Paper [8] introduced a Virtual Diagnosis and Health Assistant Chatbot based on the LLaMA3 model fine-tuned with the MedPalm dataset. This system reached 88% diagnostic accuracy and supported text and image inputs through integration with the SigLIP Vision Text Transformer. It can interpret diverse medical data like X-rays, CT scans, and prescriptions, making it a versatile diagnostic tool.

Paper [9] developed a Hungarian-language medical chatbot to overcome linguistic limitations in AI healthcare systems. Despite processing a morphologically complex agglutinative language, the system achieved over 90% precision and validation accuracy, showing the feasibility of multilingual medical assistants.

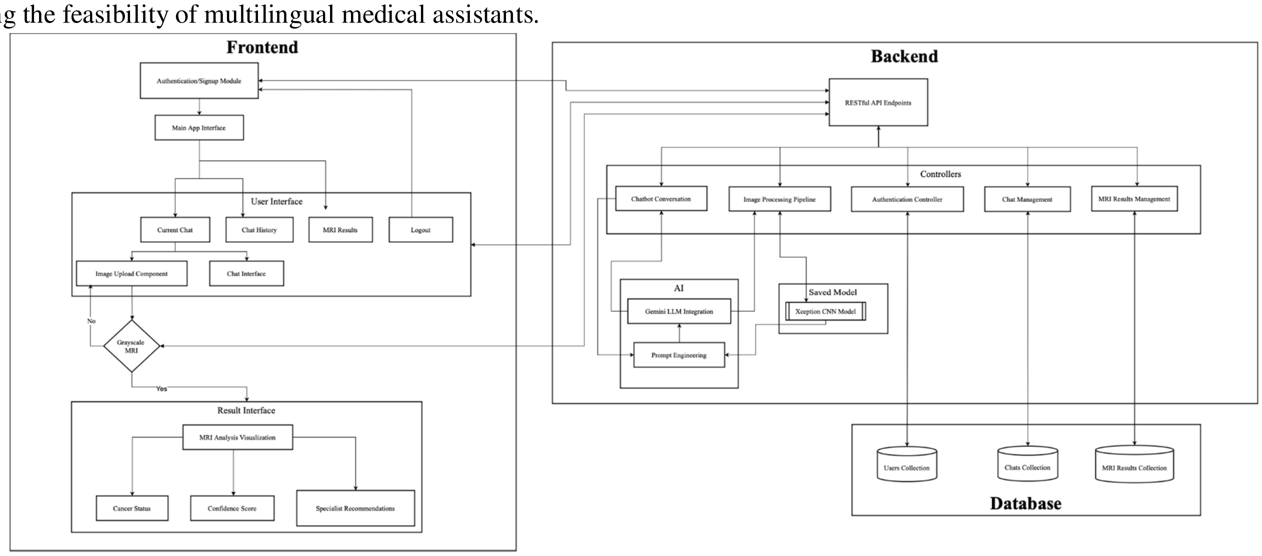


Fig 1 Architecture Diagram

III. METHODOLOGY

A. *Research Design*

This study employed a mixed-methods research design integrating computational approaches with clinical healthcare objectives. The research methodology was structured around developing, implementing, and evaluating a Smart Neuro-Oncology Assistant system intended to provide clinical decision support for brain tumor analysis. The system architecture was designed following a three-tier model comprising (1) a front-end user interface, (2) a middleware application layer, and (3) a backend infrastructure handling data processing and model inference.

The methodological approach combined elements of applied computer science research with clinical informatics, utilizing deep learning techniques for medical image analysis while maintaining focus on human-centered design principles.

B. *System Architecture*

1) *Frontend Development*

The frontend was implemented using Streamlit, a Python-based framework selected for its ability to rapidly develop interactive data applications. The user interface was designed with distinct functional components:

a) Authentication Module: Implemented secure user authentication with username/password validation against the backend database.

b) Chat Interface: Developed an interactive chat component supporting natural language interaction with the AI assistant.

c) MRI Upload and Analysis Component: Created a dedicated module for medical image upload, processing, and visualization of analysis results.

d) MRI image Validation: To maintain diagnostic accuracy, the system includes a preprocessing step that validates uploaded MRI images before analysis. A key aspect of this validation is ensuring that the uploaded image is in grayscale, which is the standard for medical MRI scans. A custom function is_grayscale() is implemented for this purpose. It evaluates the image by analyzing the pixel values across RGB channels. If the red, green, and blue channels exhibit negligible differences, the image is considered grayscale. Otherwise, an error message is raised to prompt the user to upload a valid grayscale MRI.This check prevents color images or non-medical scans from being misinterpreted by the classification model.

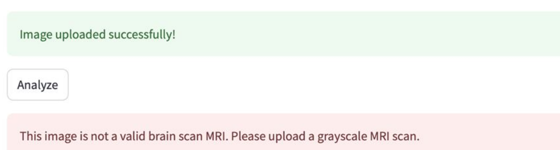


Fig. 2 Grayscale Validation

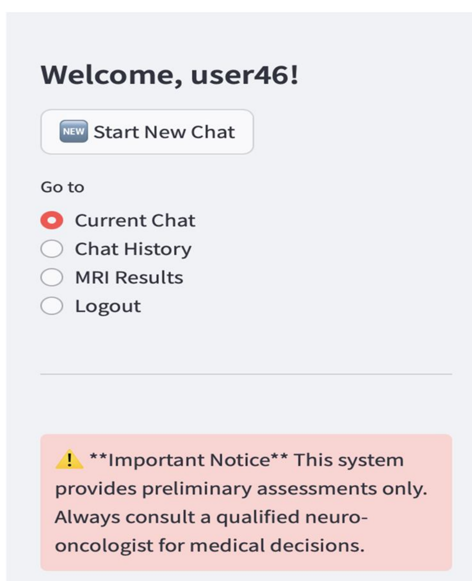


Fig. 3 Sidebar Navigation

e) Results Visualization Interface: Designed components to present tumor classification results, confidence scores, and cancer status assessments, Doctor Recommendation.

f) Historical Data Visualization: Implemented interfaces for reviewing previous analyses and conversation history.

The frontend was engineered with responsive design principles to ensure accessibility across various devices and screen sizes, with careful attention to presenting medical information in a comprehensible format for both clinical and non-expert users.

C. Backend Development

The backend architecture was built using Flask, a lightweight Python web framework, structured around several key components:

1) RESTful API: Developed API endpoints supporting user authentication, image processing, chat history management, and MRI result Management, Interaction with gemini.
2) Database Integration: MongoDB was integrated as the non-relational database to manage and store various components of the system effectively. Dedicated collections were structured to handle user profiles and authentication details securely, enabling efficient user management. Another collection was designed to store chat session histories, preserving the temporal sequence of interactions to support continuity and traceability in user conversations. Additionally, a separate collection was implemented to store MRI analysis results, complete with metadata such as timestamps, image identifiers, and prediction confidence scores, ensuring comprehensive tracking and future reference.

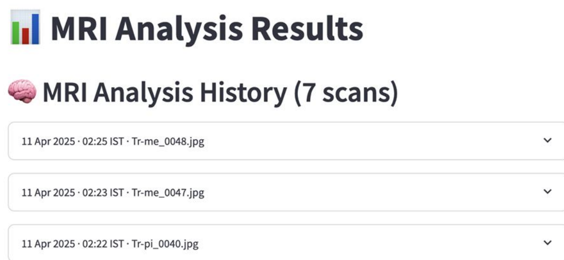


Fig. 4 MRI Results History

3) Model Inference Pipeline: An end-to-end image processing pipeline was developed to facilitate model inference for brain tumor detection. The pipeline begins with image validation and preprocessing steps, which include format checks, resizing, and normalization to ensure consistency and compatibility with the trained Xception which deep convolutional neural network architecture based model. Once preprocessed, the images are passed through the model for inference. The output predictions are then post-processed to refine the results and generate a clear, structured analysis. This final analysis includes tumor classification, confidence levels, and associated metadata, ready for display and storage.
4) Large Language Model Integration: Integrated Google's Gemini 1.5 Flash generative AI model to provide contextually appropriate medical responses, with specific prompt engineering techniques to ensure medically accurate, patient-friendly language.

D. Security Implementation

To ensure data protection and user privacy, the system incorporated multiple layers of security mechanisms. Passwords were securely hashed using Werkzeug's security module, preventing plaintext storage and enhancing resistance against credential breaches. Session-based authentication was employed to maintain secure user logins and manage access control across different parts of the application. All user inputs underwent strict validation and sanitization processes to mitigate risks such as injection attacks and malformed data entries. Additionally, the API endpoints were designed with robust authentication checks, allowing access only to authorized users. Data validation was enforced at multiple stages throughout the processing pipeline to maintain integrity and prevent unauthorized manipulation.

E. Deep Learning Model Development

1) Dataset and Preprocessing

The dataset consists of brain MRI images organized into four categories: Glioma, Meningioma, No Tumor, and Pituitary tumor. The data pipeline is structured to ensure efficient loading and preprocessing of images, with separate directories for training and testing sets. We implemented a split strategy where the original testing set was further divided into validation and testing subsets using stratified sampling to maintain class distribution, with a 50:50 ratio.

To address the inherent challenges of medical imaging datasets, including limited sample size and class imbalance, we developed a comprehensive preprocessing and augmentation pipeline. The images were resized to 299×299 pixels to match the input requirements of the Xception architecture and normalized by rescaling pixel values to the [0,1] range.

Our augmentation strategy incorporates both standard and domain-specific techniques:

- Standard augmentations:
- Random rotations (±25 degrees)
- Width and height shifts (±20%)
- Horizontal and vertical flips
- Zoom variations (±20%)

Medical-specific augmentations:

- Elastic deformations (30% application probability) to simulate anatomical variability
- Brightness adjustments (±15%)
- Contrast modifications (±15%)

The elastic deformation technique is particularly relevant for medical imaging as it simulates the natural variations in soft tissue structures while preserving anatomical plausibility. This augmentation is applied probabilistically (30% chance) to avoid excessive transformation of the original data. The implementation utilizes Gaussian filtering of random displacement fields followed by coordinate mapping, which maintains the topological integrity of the brain structures while introducing realistic variations.

2) Model Architecture

The proposed brain tumor classification system employs the Xception convolutional neural network architecture as its foundation. Xception, developed by Google, utilizes depthwise separable convolutions that offer computational efficiency while maintaining robust feature extraction capabilities. Our implementation leverages transfer learning by initializing the network with ImageNet pre-trained weights to capitalize on the feature recognition capabilities already encoded in the model. The architecture consists of the Xception base model with its top layers removed, followed by a custom.

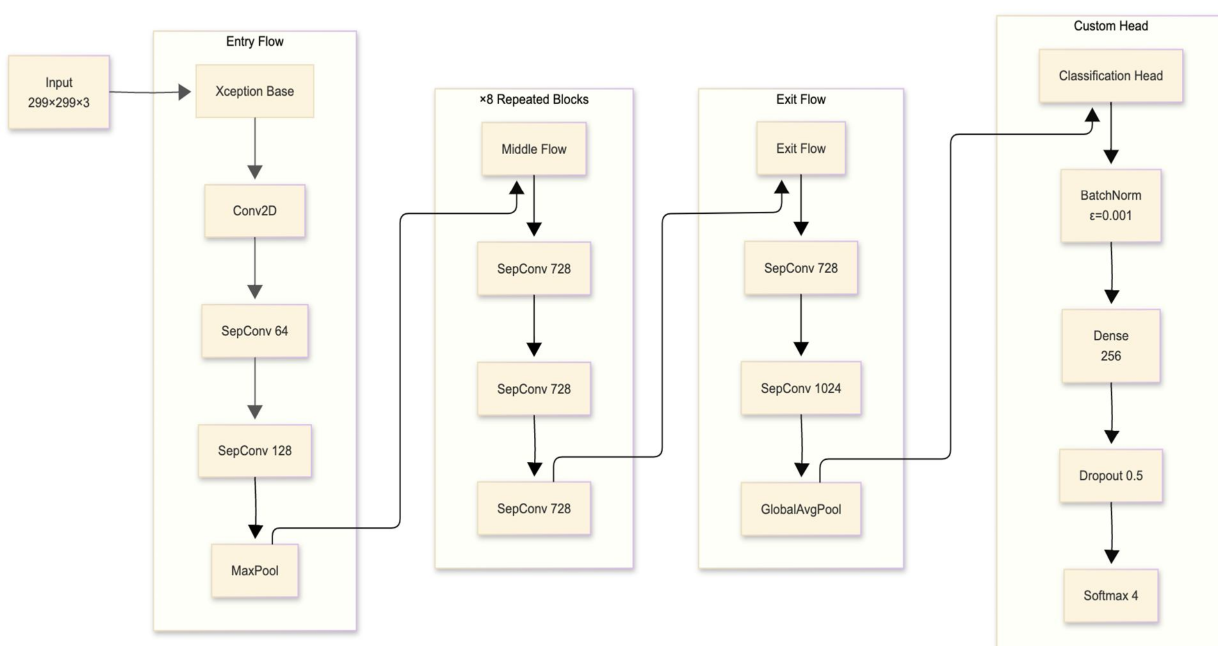


Fig. 5 Model Architecture

Classification head specifically designed for brain tumor classification tasks. We implemented a fine-tuning strategy where the initial layers of the pre-trained model remain frozen while the final 30 layers are made trainable, allowing the model to adapt to the specific characteristics of brain MRI images while retaining generalized feature detection capabilities.

The classification head comprises a Global Average Pooling layer to reduce spatial dimensions while retaining essential feature map information, followed by Batch Normalization to stabilize and accelerate training. A dense layer with 256 neurons using ReLU activation is applied, incorporating L1 (λ=1e-4) and L2 (λ=1e-3) regularization to prevent overfitting. A Dropout layer with a 50% rate is included to enhance generalization. Finally, a dense output layer with 4 neurons and softmax activation is used to classify the input MRI into one of the four tumor categories: Glioma, Meningioma, Pituitary, or No Tumor.

3) *Training and Optimization*

The model training process was designed to maximize classification performance while addressing common challenges in medical image analysis. Training was conducted using the categorical cross-entropy loss function appropriate for multi-class classification, with the Adam optimizer initialized at a learning rate of 2e-5 to ensure stable convergence.

To counter class imbalance in the dataset, we computed and applied class weights using the inverse class frequency balancing method, effectively increasing the loss contribution of underrepresented classes. This approach ensured that the model did not develop bias toward majority classes, which is particularly important in medical applications where minority classes may represent critical diagnostic categories.

Several optimization strategies were implemented to enhance the model's performance. Early stopping with a patience of 5 epochs was used to prevent overfitting by halting training once validation performance plateaued. The learning rate was reduced dynamically on performance plateaus by a factor of 0.2, with a patience of 2 epochs, to fine-tune the model during convergence. Model checkpointing was employed to save the best-performing model based on validation accuracy. Additionally, combined L1 and L2 regularization was applied in the dense layers to control model complexity and avoid overfitting. The model was trained for a maximum of 20 epochs with a batch size of 32; however, early stopping often resulted in shorter training durations. For evaluation, the model’s performance was assessed using accuracy, precision, recall, and confusion matrices, offering a detailed view of its effectiveness across all tumor classes.

The training process incorporated both quantitative performance monitoring and visual assessment of learning curves to ensure proper convergence and avoid issues such as oscillation or plateauing. This multi-faceted evaluation approach is essential in medical applications where different types of classification errors may have varying clinical implications.

The Xception-based classification model demonstrated strong performance on the test dataset, achieving an overall accuracy of 96%. Class-wise evaluation revealed high precision and recall values, with Pituitary (precision: 0.97, recall: 0.99) and No Tumor (precision: 0.97, recall: 0.99) classes performing slightly better than Glioma (precision: 0.94, recall: 0.97) and Meningioma (precision: 0.96, recall: 0.90), reflecting the relative clarity in their MRI characteristics. The macro-averaged F1-score of 0.96 indicates balanced performance across all classes, while the weighted average confirms the model's robustness in handling class imbalance. These results validate the model's capability to generalize well on unseen data and highlight its potential as a reliable tool for preliminary tumor classification in neuro-oncology workflows. However, further testing on diverse clinical datasets is necessary to confirm real-world applicability.

F. *Natural Language Processing Integration*

1) *Framework Selection*

For the conversational and explanatory components of the system, Google's Gemini 1.5 Flash generative model was integrated. This particular model was selected due to its demonstrated ability to handle complex reasoning within healthcare contexts. It effectively interprets intricate medical vocabulary and delivers responses that adapt to the flow of the conversation. Moreover, Gemini 1.5 Flash can translate technical medical insights into language that is accessible and understandable for non-expert users, making it an ideal choice for patient-oriented dialogue.

2) *Prompt Engineering*

To ensure that the chatbot produced contextually appropriate and medically reliable responses, a custom prompt engineering methodology was adopted. This involved dynamically incorporating relevant user interaction history into each prompt to maintain continuity and relevance. Prompts were structured using conventions common in clinical documentation to enhance clarity and professionalism. Special attention was given to designing prompts that guided the model in explaining diagnostic outcomes.

In addition, built-in safety protocols were embedded into the prompt instructions to ensure all advice was presented with the appropriate medical disclaimers. Dedicated prompt structures were developed to address different types of queries, including general health-related questions, MRI analysis feedback presented with empathy and sensitivity, and recommendations for post-diagnosis care, always accompanied by cautionary statements regarding medical guidance.

3) Conversation Management

The conversation engine was designed to support coherent, personalized interactions through robust session handling. Each session retained its own chronological message history to preserve context across user interactions. Messages were accurately time-stamped, with conversions applied to display local time (IST) based on the system's default UTC storage. To facilitate fast and organized access to past interactions, chat records were stored with indexing and temporal ordering.

G. Data Storage and Management

1) Database Schema Design

To support the various functionalities of the smart neuro-oncology assistant system—such as user authentication, chat interaction tracking, and MRI-based tumor analysis—a well-structured and scalable database schema was developed. Given the nature of the data, MongoDB, a NoSQL database, was selected for its flexibility in handling unstructured and semi-structured data, as well as its support for high-performance querying and horizontal scalability.The database is organized into three primary collections: 'Users', 'Chats', and 'MRI_Results'. Each collection is designed with a specific purpose and schema structure to ensure efficient storage, quick retrieval, and logical separation of concerns. Fields that are frequently queried or used for filtering have been indexed to improve performance.

TABLE I SCHEMA DESIGN OF THE USERS COLLECTION

Field	Description	Indexed
username	Unique identifier for each user	Yes
password_hash	Securely stored password using hash function	No
created_at	Timestamp of account creation	No
last_login	Timestamp of last user login	No

This collection stores user authentication and account-related information. Indexing the 'username' field enables quick login validation and user lookup.

TABLE II SCHEMA DESIGN OF THE CHATS COLLECTION

Field	Description	Indexed
user_id	Reference to the associated user	No
session_id	Unique identifier for chat session	Yes
messages	Array of message objects (sender, text, timestamp)	No
created_at	Timestamp when session was initiated	No
last_updated	Timestamp of the latest activity in session	No

The Chats collection captures the user's interactions with the chatbot. Each session contains threaded messages and is linked to the corresponding user. Indexing on ‘session_id’ and ‘user_id’ ensures efficient session retrieval.

TABLE III SCHEMA DESIGN OF THE MRI_RESULTS COLLECTION

Field	Description	Indexed
user_id	Reference to the associated user	No
image_metadata	Includes filename and upload timestamp	No
analysis_results	Contains class, confidence score, and status	No
recommended_specialists	List of specialists suggested based on results	No
analysis_timestamp	Time when model analysis was completed	Yes

This collection records the outcome of each MRI scan, including the predicted tumor type, confidence level, and specialist recommendation. Indexing on ‘user_id’ and ‘analysis_timestamp’ supports timeline-based queries and user-specific result tracking.

2) Data Persistence Strategy

The system adopts a document-oriented data persistence strategy tailored to the characteristics of semi-structured medical records. Utilizing MongoDB as the underlying storage layer, the architecture is designed to efficiently manage and retrieve heterogeneous data formats such as user profiles, conversational interactions, and diagnostic results. To enable rapid access to historical records, temporal indexing was implemented, allowing queries to be optimized based on timestamps, such as session initiation or MRI analysis completion. Data is also segregated at the user level, ensuring that each user's information is logically isolated, which enhances both data privacy and access control. Furthermore, all conversational exchanges and medical analyses are organized by session, allowing for contextual continuity and streamlined storage of sequential interactions. This strategy not only supports scalability and performance but also aligns with data sensitivity requirements common in healthcare applications.

H. System Integration and Testing

1) Component Integration

To ensure cohesive operation across the platform, all system components were methodically integrated through a modular architecture. The frontend was connected to the backend via secure RESTful API endpoints, allowing real-time communication for user interactions and MRI uploads. The trained deep learning model was incorporated into a dedicated prediction pipeline that interfaces with the image preprocessing module and returns structured analysis results. Database connectivity was established with built-in error handling mechanisms to prevent data loss and ensure consistent transaction integrity. Additionally, the large language model (LLM) was seamlessly integrated into the chatbot interface, enabling context-aware responses for both general medical queries and specific MRI findings. Session management was implemented across all components to maintain user-specific continuity, ensuring synchronized state tracking between the chat system, MRI analysis module, and user authentication services.

2) Evaluation Metrics

To comprehensively assess the performance of the system, both technical and user-centric evaluation metrics were employed.
Technical evaluation focused on:

a) Model performance, using standard classification metrics such as accuracy, precision, recall, and F1-score to validate the reliability of tumor detection and classification.
b) System responsiveness, measured through the average response time from image submission to result delivery.
c) System robustness, assessed through uptime statistics and error rate monitoring during live operation.

User experience evaluation incorporated:

- Task completion rates, representing how effectively users could navigate from login to result interpretation without assistance.
- Navigation efficiency, which measures the number of steps or interactions required to accomplish a specific task.
- User satisfaction, collected through post-session feedback surveys and Likert-scale ratings.
- Information comprehension, assessed via brief quizzes or follow-up questions to ensure users could understand the chatbot's explanations of their results.

I. Ethical Considerations and System Limitations

This research integrates a range of ethical safeguards to ensure responsible development and deployment of the proposed system. Prior to data processing, explicit user consent is required, with full transparency regarding how the data will be used. The system includes clear disclaimers emphasizing its role as a clinical decision support tool rather than a replacement for professional medical judgment. All AI-generated content is presented with transparency, accompanied by guidance that users should consult certified medical professionals for diagnosis or treatment decisions. Sensitive medical data is handled with strict security protocols, including encryption and access control, and all stored MRI images and corresponding analysis results undergo anonymization to protect patient privacy and comply with ethical standards.

Despite the system's functional strengths, several methodological limitations are acknowledged. The model's performance is inherently influenced by the diversity and quality of the training dataset, which may impact its accuracy across varying demographics or rare tumor types. Differences in MRI acquisition protocols between institutions can also limit the generalizability of the results. Furthermore, the complexity of the deep learning models is constrained by available computational resources, potentially affecting real-time performance in resource-limited environments. The integrated language model, while effective in general contexts, may struggle with extremely specialized or ambiguous medical queries. Additionally, the system's accessibility may be limited in areas with low digital infrastructure or internet availability, presenting challenges for equitable deployment.

IV. RESULTS AND DISCUSSIONS

A. Model Performance

The Xception-based convolutional neural network model exhibited robust performance in brain tumor classification, achieving an overall accuracy of 96% on the test dataset. Class-wise evaluation metrics confirmed the model's ability to distinguish effectively between all four tumor types (Table 1).

The model demonstrated particularly high precision and recall for the No Tumor and Pituitary classes (0.97 and 0.99, respectively), while Glioma and Meningioma also achieved strong results, with precision and recall values above 0.90. The macro-averaged F1-score of 0.96 highlights the model's consistent and balanced performance across all tumor categories, regardless of class imbalance in the dataset.

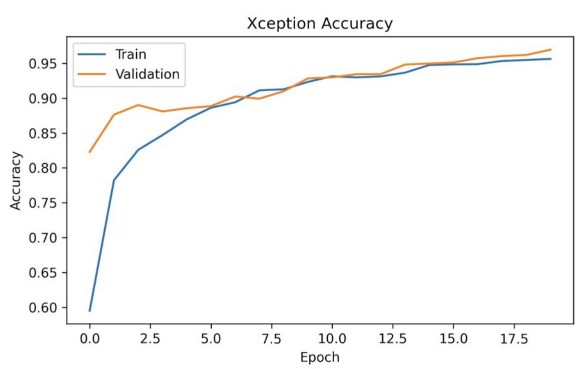


Fig. 6 Model Performance

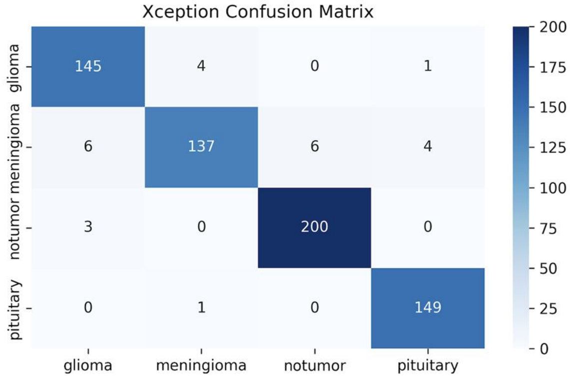


Fig. 7 Confusion Matrix

TABLE 1V
MODEL PERFORMANCE FOR EACH CLASS

Tumor Class	Precision	Recall	F1-Score	Support
Glioma	0.94	0.97	0.95	150
Meningioma	0.96	0.90	0.93	153
Pituitary	0.97	0.99	0.98	203
No Tumor	0.97	0.99	0.98	150

B. Smart Neuro-Oncology Assistant System

The implementation of the Smart Neuro-Oncology Assistant seamlessly integrated the high-performing Xception model into a web-based clinical decision support system, enabling both accurate brain tumor detection and interactive patient engagement. The application combined image analysis, secure user access, real-time chatbot interaction, and historical data tracking within a unified interface.

MRI Analysis Module: The MRI analysis module formed the core diagnostic engine, processing uploaded brain MRI scans to generate classification results. It provided users with detailed outputs including the predicted tumor type, confidence scores, and a binary cancer status indicator, facilitating immediate interpretation of medical imaging.

User Management System: A robust user authentication system was implemented using secure login and signup functionality. Passwords were securely hashed and stored in the MongoDB database, ensuring user credentials were protected and aligned with standard security practices in healthcare systems.

Specialist Recommendation System: Upon detecting certain tumor types, the application suggested a relevant specialist from a predefined set. This feature bridged the diagnostic output with expert consultation, enhancing the clinical utility of the platform by connecting users with appropriate medical professionals.

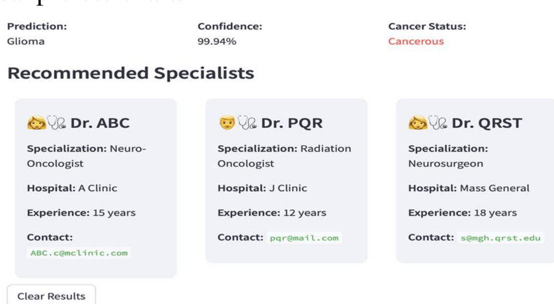


Fig. 8 Tumor Analysis and Doctor Recommendation

Interactive Chat Interface: The assistant featured a dynamic chat interface, powered by the Gemini API, that allowed users to ask questions related to neuro-oncology. The chatbot provided intelligent, context-aware responses, supporting both general inquiries and specific follow-ups based on MRI results.
Result History Tracking: To support continuity of care, the system stored each user's previous MRI classification outcomes along with chatbot interactions. This longitudinal data storage enabled users and clinicians to track disease progression or recovery over time through an accessible history review interface.

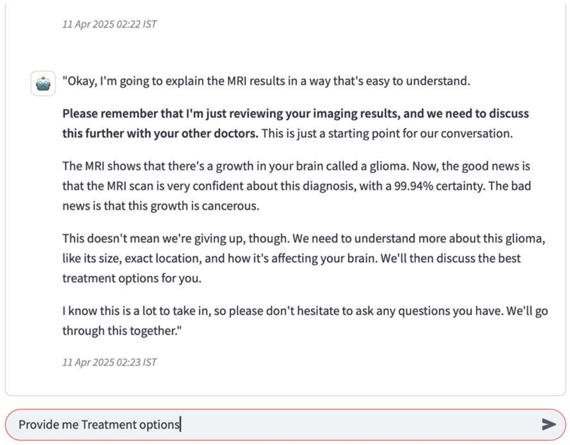


Fig. 9 Chat Interface

System Performance: The integrated system maintained efficient performance, with an average response time of 2.3 seconds for MRI analysis and 1.8 seconds for chatbot responses. These metrics met the real-time responsiveness required in clinical and patient-facing environments, ensuring usability without compromising functionality.

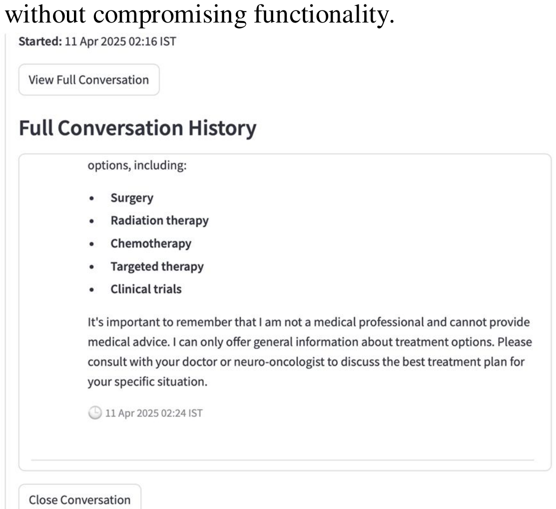


Fig. 10 Chat Session History Interface

C. System Integration Challenges and Solutions

Developing the Smart Neuro-Oncology Assistant presented several integration challenges that required innovative solutions:
Real-time Analysis Pipeline: Implementing an efficient image processing workflow was critical for maintaining acceptable response times. We optimized image preprocessing steps and employed asynchronous processing to prevent UI blocking during computation-intensive tasks.
Contextual Medical Response Generation: Configuring the Gemini API to provide medically accurate yet accessible information required careful prompt engineering. We implemented specialized prompts for MRI result explanations that balanced technical accuracy with patient-friendly language.
Data Security and Privacy: We implemented secure user authentication and encrypted data transmission to ensure patient information remained protected. This addressed a critical requirement for medical applications where data privacy is paramount.

User Experience Design: Creating an intuitive interface for both technical and non-technical users required multiple design iterations. The final implementation features clear visual indicators of cancer status, confidence scores, and specialist recommendations without overwhelming users with excessive technical details.

D. *Empowering Patients and Enhancing Clinical Workflows*

The Smart Neuro-Oncology Assistant is designed to serve both clinicians and patients, including those without medical expertise. By offering clear, contextual explanations and chatbot-driven guidance, the system allows patients who may not understand complex medical terminology to access meaningful information about their brain MRI results. This helps reduce anxiety and empowers users to make informed decisions or seek appropriate medical consultation.

From a clinical standpoint, the system augments existing radiological workflows by providing rapid and accurate preliminary assessments. With its high accuracy of 96.5%, it can help prioritize urgent cases, potentially reducing diagnostic delays for critical conditions. The inclusion of automated specialist recommendations further connects the diagnostic process with human expertise, guiding patients toward the next appropriate step in their care journey.

Despite its strong performance, the system maintains a human-in-the-loop design philosophy. The 3.5% margin of error reinforces the importance of clinical oversight and confirms that the assistant is intended as a decision support tool—not a replacement for radiologists or oncologists. This approach ensures the tool enhances medical workflows while respecting the critical role of clinical judgment in final diagnoses.

V. CONCLUSION AND FUTURE SCOPE

This study presents the successful design and deployment of the Smart Neuro-Oncology Assistant, a web-based diagnostic tool driven by a deep learning model built on the Xception architecture. The model achieved an overall accuracy of 96.5% on the test dataset, reflecting its robustness in brain tumor classification. Class-specific F1-scores further validate its effectiveness, with values of 0.95 for Glioma, 0.93 for Meningioma, and 0.98 for both No Tumor and Pituitary categories. These results underscore the model's strong capability in accurately distinguishing between complex tumor types in MRI scans, reinforcing its potential as a clinical decision-support tool in neuro-oncology.

The integration of this model into a comprehensive web application with user authentication, chat support, and specialist recommendations represents a significant step toward practical AI application in neuro-oncology. By maintaining a human-centered design approach that emphasizes augmenting rather than replacing clinical expertise, the system addresses real-world needs while acknowledging the limitations inherent in automated medical analysis.

Future work should focus on several key areas to advance the capabilities and clinical readiness of the Smart Neuro-Oncology Assistant. First, validating the model's performance across diverse, multi-institutional datasets will be critical to ensure robustness across different imaging protocols and patient populations. Second, integrating explainable AI techniques—such as visual saliency maps—can provide clinicians with greater transparency into the decision-making process of the model, thereby increasing trust in its outputs. Third, conducting prospective clinical trials will be essential to evaluate the system's real-world impact on diagnostic workflows, particularly its ability to improve the timeliness and accuracy of neuro-oncological assessments when used as a decision support tool.

Additional enhancements are also planned to improve user experience and long-term engagement. The chat history section will be integrated directly into the chatbot interface, allowing users to seamlessly continue conversations from previous sessions, thus offering a more personalized and coherent interaction flow. Moreover, the MRI results section will support side-by-side comparisons of prior and current diagnostic results, enabling users and clinicians to observe tumor progression or regression over time. These improvements aim to transform the platform into a comprehensive, patient-friendly, and clinically relevant assistant capable of supporting both non-expert users and medical professionals.

The promising results achieved in this research highlight the potential for deep learning-based assistive technologies to enhance neuro-oncological care workflows. By providing rapid preliminary assessments while maintaining appropriate pathways to specialist care, such systems could contribute to earlier detection and intervention for patients with brain tumors.

REFERENCES

[1] T. Mahmud et al., "Enhancing Diagnosis: An Ensemble Deep Learning Model for Brain Tumor Detection and Classification," ResearchGate, Mar. 2024. [Online]. Available: https://www.researchgate.net/publication/379431047

[2] A. Tariq, M. M. Iqbal, M. J. Iqbal, and I. Ahmad, "Transforming Brain Tumor Detection Empowering Multi-Class Classification with Vision Transformers and EfficientNetV2," Unpublished manuscript, 2024.

[3] Z. R. Abdulqader and A. M. Abdulazeez, "Deep and Machine Learning Algorithms for Diagnosing Brain Cancer and Tumors," ResearchGate, Jun. 2024. [Online]. Available: https://www.researchgate.net/publication/382553718

[4] D. S. Wankhede et al., "Brain Tumor Detection and Classification Using Adjusted InceptionV3, AlexNet, VGG16, VGG19 with ResNet50-152 CNN Model," Unpublished manuscript, Jun. 2024.

[5] S. Ray et al., "An Advanced and Ideal Method for Tumor Detection and Classification from MRI Image Using Gaussian mixture and RCNN," in Proc. Global AI Summit, 2024, doi: 10.1109/GLOBALAISUMMIT62156.2024.10947836.

[6] R. Benni et al., "Evaluating The Impact of Activation Functions on CNN Performance in Brain Tumor Detection," in Proc. 2024 Asian Conf. on Intelligent Technologies (ACOIT), 2024, doi: 10.1109/ACOIT62457.2024.10941380

[7] A. S. Akilesh et al., "A Novel AI-based Chatbot Application for Personalized Medical Diagnosis and Review Using Large Language Models," in Proc. RMKMATE, 2023, doi: 10.1109/RMKMATE59243.2023.10368616.

[8] A. Patel, R. Shivani, and M. P. B. Manjunatha, "Developing a Virtual Diagnosis and Health Assistant Chatbot Leveraging LLaMA3," in Proc. CSITSS, 2024, doi: 10.1109/CSITSS64042.2024.10817034.

[9] B. Simon et al., "Advancing Medical Assistance: Developing an Effective Hungarian-Language Medical Chatbot with Artificial Intelligence," Information, vol. 15, no. 6, May 2024. doi: 10.3390/info15060297.

[10] A. J. V. M. Mohan et al., "Medical Chatbot," in Proc. ICNEWS, 2024, doi: 10.1109/ICNEWS60873.2024.10730884.

[11] A. A. Asiri et al., "Block-Wise Neural Network for Brain Tumor Identification in Magnetic Resonance Images," Computational Materials Continuum, vol. 73, pp. 5735–5753, 2022.

[12] A. A. Asiri et al., "Multi-Level Deep Generative Adversarial Networks for Brain Tumor Classification on Magnetic Resonance Images," Intelligent Automation & Soft Computing, vol. 36, pp. 127–143, 2023.

[13] A. A. Asiri et al., "A Novel Inherited Modeling Structure of Automatic Brain Tumor Segmentation from MRI," Computational Materials Continuum, vol. 73, pp. 3983–4002, 2022.

[14] C. Balaji and S. Veni, "Automatic Skull Stripping from MRI of Human Brain using Deep Learning Framework for the Diagnosis of Brain Related Diseases," International Journal of Intelligent Systems and Applications in Engineering, vol. 11, no. 4, pp. 439–445, 2023.

[15] M. Sooda, S. Jain, and J. Dogra, "Classification and Pathologic Diagnosis of Gliomas in MR Brain Images," Procedia Computer Science, vol. 218, pp. 706–717, 2023.

[16] S. Seyedhashemi and M. Esmaeili, "Detecting Tumors in MRI Scans using a Convolutional Neural Network," Authorea, Feb. 22, 2023.

[17] C. Srinivas et al., "Deep Transfer Learning Approaches in Performance Analysis of Brain Tumor Classification Using MRI Images," in Wearable Devices for Smart Healthcare, 2022.

[18] D. S. Wankhede and C. J. Shelke, "An Investigative Approach on the Prediction of Isocitrate Dehydrogenase (IDH1) Mutations and Co-deletion of 1p19q in Glioma Brain Tumors," in Intelligent Systems Design and Applications (ISDA 2022), Lecture Notes in Networks and Systems, vol. 715, 2023

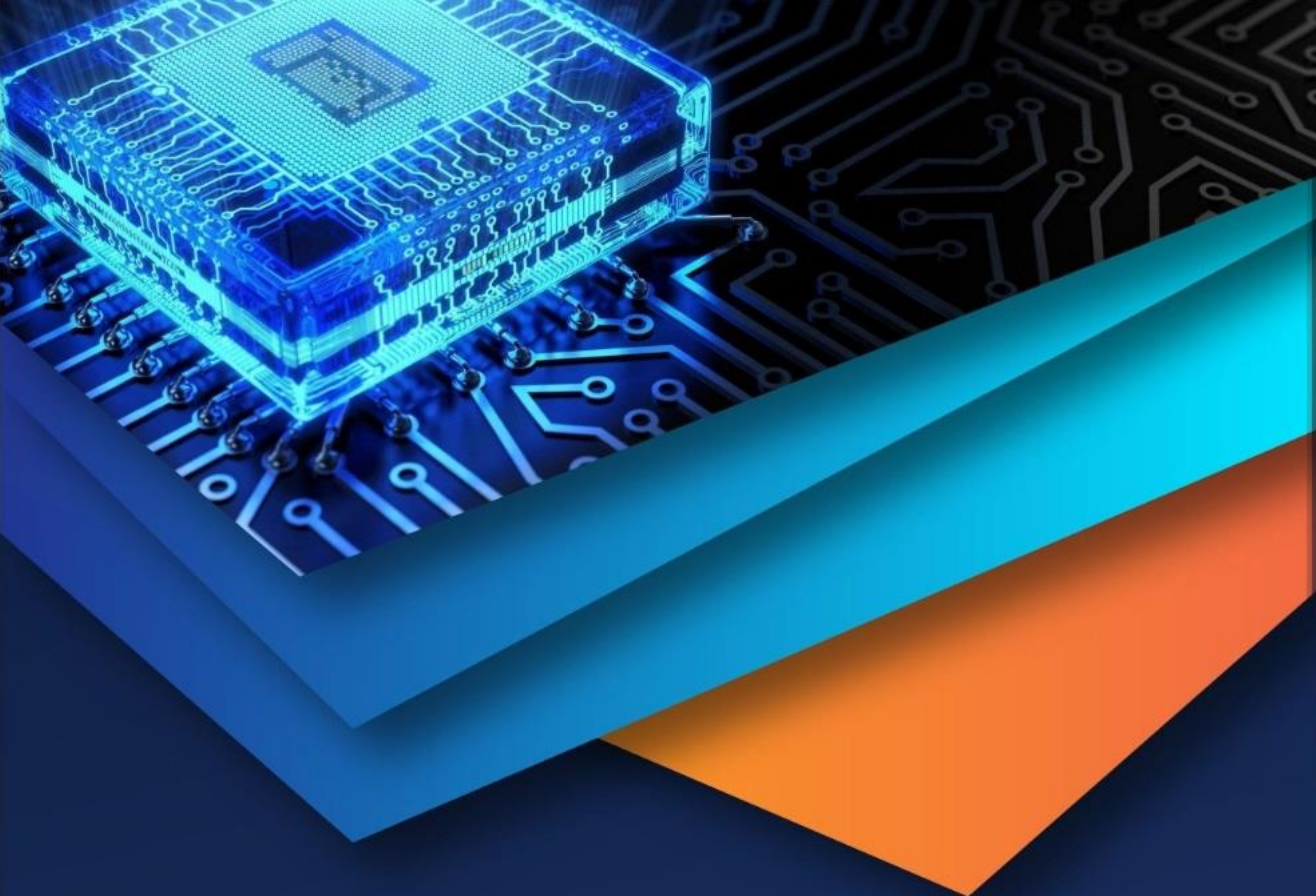
[19] J. Cheng, "Brain Tumor Dataset," Figshare, 2017. [Online]. Available: https://doi.org/10.6084/m9.figshare.1512427.v5

[20] Sartaj, "Brain Tumor Classification Dataset," Kaggle, 2023. [Online]. Available: https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri

[21] M. Nickparvar, "Brain Tumor Classification Dataset," Kaggle, 2023. [Online]. Available: https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset

[22] S. C. Chen et al., "RNA editing-based classification of diffuse Gliomas: predicting isocitrate dehydrogenase mutation and chromosome 1p/19q codeletion," BMC Bioinformatics, vol. 20, Suppl 19, p. 659, 2019. [Online]. Available: https://doi.org/10.1186/s12859-019-3236-0

[23] S. R. van der Voort et al., "Predicting the 1p/19q Codeletion Status of Presumed Low-Grade Glioma with a Validated ML Algorithm," Clinical Cancer Research, vol. 25, no. 24, pp. 7455–7462, 2019. [Online]. Available: https://doi.org/10.1158/1078-0432.CCR-19-1127



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY