International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0135
e ISSN: 2584-2854
Volume: 03
Issue:03 March 2025
Page No: 837-840

YOLOv10-Powered Brain Tumor Detection with AI-Based Insights

Sanket Satpute[1], Jagruti khairnar[2], Rohini Sangle[3], Kunal Shinde[4], Prof J.N. Thakur[5], Prof A.V. Gangurde[6]
[1,2,3,4] UG Scholar, Dept. of AIML, LoGMIEER, Nashik, Maharashtra, India.
[5,6]Associate Professor, Dept. of AIML, LoGMIEER, Nashik, Maharashtra, India.
Email ID: *nsksanketsatpute@gmail.com[1], jagrutikhairnar2003@gmail.com[2], rohinisangle632@gmail.com[3], kunalshinde2029@gmail.com[4]*

Abstract
This project presents an AI-powered Brain Tumor Detection and Classification System integrated with an AI Chatbot for medical query processing. The architecture consists of a backend, frontend, and user interface. The backend leverages a YOLOv10-based deep learning model for brain tumor detection, comprising image processing and tumor detection modules. Additionally, an OpenAI multimodal LLM is incorporated to process medical queries and generate captions. A SQLite database is used for data storage and retrieval. The frontend provides API endpoints for image prediction (/predict) and chatbot interactions (/chatbot), serving as a bridge between the backend and user interface. The user interface allows users to upload medical images, view detection results, and interact with the AI chatbot for medical-related inquiries. This architecture enables an automated, efficient, and user-friendly system for early brain tumor detection and medical assistance, improving accessibility and accuracy in diagnosis

Keywords: *Brain Tumor Detection, YOLOv10, Deep Learning, Image Processing, Large Language Model (LLM), AI Chatbot, SQLite Database, Medical Query Processing, API Endpoints, Healthcare AI*

1. Introduction
This research presents an advanced AI-driven system for tumor detection and medical query processing, integrating computer vision and natural language processing (NLP) techniques. The system is structured into three core components: Backend, Frontend, and User Interface, each designed to perform specific tasks while ensuring smooth interaction between users and the AI-powered functionalities. The goal of this architecture is to facilitate efficient medical image analysis and provide intelligent responses to user queries, offering a comprehensive support system for medical professionals and patients.

1.1. Backend
The backend serves as the computational core of the system, incorporating deep learning models and database management to handle image processing and user queries efficiently. At its core, the YOLOv10 model is implemented for tumor detection, utilizing advanced image-processing algorithms to analyze medical scans and identify abnormalities with high precision. This model processes uploaded images, detects potential tumors, and sends the results to the frontend for visualization. Additionally, the backend integrates an OpenAI multimodal language model (LLM) that is responsible for medical query processing and caption generation. This component enables the system to interpret user queries related to medical conditions, analyze contextual information, and generate accurate, human-like responses. The SQLite database layer is used for structured data management, storing relevant medical information, image metadata, and user interactions, ensuring data persistence and retrieval efficiency. Together, these backend components ensure the system's capability to process medical images and queries in real-time, making it an essential component in AI-assisted diagnostic applications. The seamless integration of deep learning and NLP enhances the system’s reliability, making it a valuable tool in the medical domain.

1.2. Frontend
The frontend serves as an interactive bridge between the backend and the user, managing API requests and rendering user-friendly interfaces. It is responsible for presenting the processed information, handling

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0135
e ISSN: 2584-2854
Volume: 03
Issue:03 March 2025
Page No: 837-840

user inputs, and facilitating smooth communication between different system components. This layer ensures that medical professionals and patients can seamlessly interact with the system without requiring extensive technical knowledge. One of the primary functions of the frontend is to serve HTML templates and provide API endpoints for interacting with backend functionalities. The /predict API handles tumor detection requests, sending uploaded images to the YOLOv10 model for analysis and retrieving the detection results. Similarly, the /chatbot API processes user-submitted medical queries, forwarding them to the OpenAI multimodal LLM for intelligent response generation. These endpoints ensure efficient communication between the user interface and backend processing units. Additionally, the frontend is designed to be responsive, ensuring compatibility with multiple devices, including desktops, tablets, and mobile phones. By leveraging modern web development frameworks, the system maintains a seamless, interactive experience that enhances usability and accessibility. The frontend's ability to efficiently manage data flow, present analytical results, and facilitate user interactions makes it a crucial component in the overall system architecture.

1.3. User Interface

The user interface (UI) is the primary interaction point for users, designed to provide an intuitive and accessible experience for both medical professionals and patients. It consists of three main functionalities: Image Upload, Result Display, and Chatbot, each contributing to the system's effectiveness in tumor detection and medical query processing.

Image Upload: Users can upload medical images (e.g., MRI or CT scans) through a simple and user-friendly interface. These images are then processed by the backend YOLOv10 model, which analyzes them for tumor detection.

Result Display: Once the tumor detection process is completed, the results are displayed in a visually informative manner. This component presents marked tumor regions on the uploaded scans, allowing users to interpret the analysis easily. It also provides additional insights generated by the system to assist in medical decision-making.

Chatbot: The integrated AI-powered chatbot allows users to ask medical-related questions. This feature leverages the OpenAI multimodal LLM to generate contextually accurate responses, offering real-time medical guidance and explanations related to uploaded images. (Figure 1)

2. System Architecture

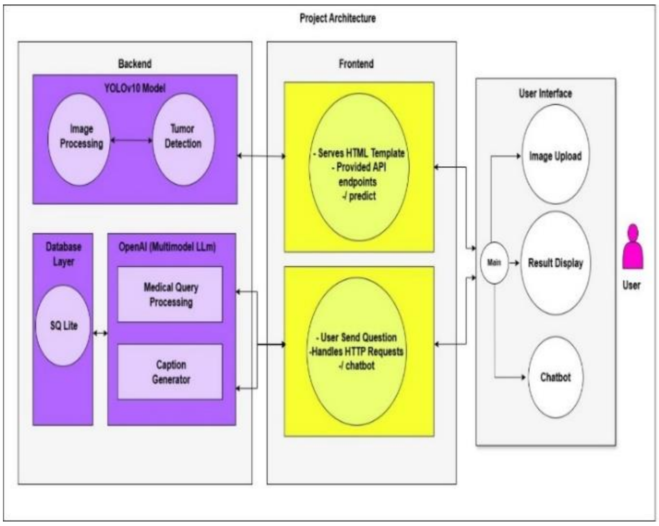

Figure 1 System Architecture

3. Workflow of the Architecture

The workflow of this architecture begins with the user uploading a medical image through the user interface. This image is sent to the backend, where the YOLOv10 model processes it for tumor detection. The image processing module analyzes the uploaded file, detects any abnormalities, and sends the results back to the frontend. Simultaneously, the SQLite database stores relevant data for future reference. If the user has a medical query, they can interact with the chatbot through the user interface. The frontend forwards the query to the OpenAI multimodal language model, which processes it and generates a meaningful response. The chatbot then displays the answer on the user interface, assisting the user with their medical concerns. Finally, the system displays the tumor detection results along with relevant explanations, ensuring that users receive both visual diagnostic insights and AI-powered medical assistance. This streamlined workflow enables an efficient, interactive, and AI-driven healthcare support system. (Figure 2)

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0135
e ISSN: 2584-2854
Volume: 03
Issue:03 March 2025
Page No: 837-840

4. Results and Discussion

4.1. Results

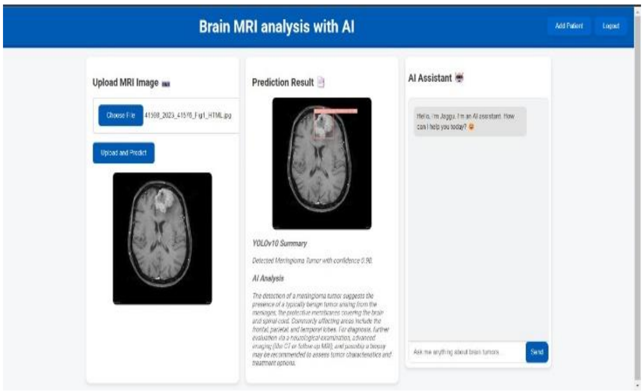

Figure 2 Result Interface

The above interface shows the Brain MRI Analysis with AI system in action. The system allows users to upload an MRI image, process it using the YOLOv10 model, and receive a tumor detection result along with an AI-generated analysis. In this case, the uploaded MRI image has been processed, and the model has detected a meningioma tumor with a confidence score of 0.90. The tumor is highlighted in the prediction result section, where a bounding box marks the affected area. The AI-generated analysis explains that meningioma is typically a benign tumor arising from the brain's protective membranes and suggests further medical evaluation through neurological exams or imaging techniques. Additionally, the interface features an AI assistant (chatbot) named "Jaggu," which is available to answer medical queries related to brain tumors. Users can interact with the assistant by typing questions and receiving AI-generated responses. Overall, the interface efficiently combines AI-driven image analysis and natural language processing to provide automated tumor detection and medical consultation in a user-friendly manner.

4.2. Discussion

The Brain MRI Analysis with AI interface presents an intuitive and interactive platform for detecting and analyzing brain tumors using artificial intelligence. It integrates computer vision and natural language processing to offer a seamless experience for medical professionals and patients seeking preliminary insights into MRI scans. The interface is divided into three main sections. The first section (left panel) allows users to upload an MRI image for analysis. Once uploaded, the system processes the image and generates predictions. The second section (middle panel) displays the prediction results, where the YOLOv10 model detects potential tumors and marks the affected area with a bounding box. In this case, the model has identified a meningioma tumor with 90% confidence and provides an AI-generated explanation about its nature, possible causes, and recommended medical follow-ups. This feature aids in preliminary diagnosis, helping radiologists and doctors make more informed decisions. The third section (right panel) consists of an AI-powered chatbot assistant named "Jaggu." This chatbot allows users to ask questions related to brain tumors and receive AI-generated responses. This enhances the user experience by providing immediate answers to medical queries, potentially reducing the need for direct consultation unless necessary. Overall, the interface successfully combines AI-driven medical imaging, data processing, and interactive chatbot support to create a comprehensive tumor analysis system. This makes it a valuable tool in telemedicine, hospitals, and research institutions, aiding in the early detection and assessment of brain tumors.

Conclusion

This AI-powered Brain Tumor Detection and Classification System revolutionizes medical diagnostics by combining the efficiency of YOLOv10 for tumor detection with the intelligence of an LLM-based AI chatbot for medical query processing. The seamless integration of deep learning, image processing, and natural language understanding ensures accurate diagnoses and enhances patient engagement. By offering an automated, fast, and user-friendly interface, this system empowers healthcare professionals and patients with early detection, instant insights, and reliable medical assistance. Its scalable architecture and real-time AI-driven predictions make it a valuable tool in modern healthcare, contributing to improved diagnostic accuracy and patient care.

Acknowledgements

We would like to express our heartfelt gratitude to our guide, Prof. Jyoti Thakur, for her valuable guidance, constant encouragement, and support throughout this project. Her insights and expertise were instrumental in the successful completion of this work. We also extend our thanks to the Department of Artificial Intelligence and Machine Learning at Loknete Gopinathji Munde Institute of Engineering Education & Research (LoGMIEER), Nashik, for providing us with the resources and infrastructure needed to conduct our research. Finally, we would like to acknowledge our friends and family for their unwavering support and understanding during the course of this project.

References

[1]. Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. arXiv preprint arXiv:1804.02767.Retrieved from https://arxiv.org/abs/1804.02767

[2]. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770-778. Retrieved fromhttps://ieeexplore.ieee.org/document/7780459

[3]. Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. A. (2017). Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. AAAI Conference on Artificial Intelligence. Retrieved from https://arxiv.org/abs/1602.07261

[4]. Touvron, H., Cord, M., Douze, M., Massa, F., Sablayrolles, A., & Jégou, H. (2021). Training data-efficient image transformers & distillation through attention. Proceedings of the International Conference on Machine Learning (ICML). Retrieved from https://arxiv.org/abs/2012.12877

[5]. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770-778. Retrieved from https:// ieeexplore. ieee.org/ document/ 7780459

[6]. Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. A. (2017). Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. AAAI Conference on Artificial Intelligence. Retrieved from https://arxiv.org/abs/1602.07261

[7]. Touvron, H., Cord, M., Douze, M., Massa, F., Sablayrolles, A., & Jégou, H. (2021). Training data-efficient image transformers & distillation through attention. Proceedings of the International Conference on Machine Learning (ICML). Retrieved from https://arxiv.org/abs/2012.12877

[8]. Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language Models are Unsupervised Multitask Learners. OpenAI GPT-2 Paper. Retrieved from https://cdn.openai.com/betterlanguage-models/language_models_are_unsupervised_multitask_learners.pdf

[9]. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 234-241. Retrieved from https://arxiv.org/abs/1505.04597

[10]. Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Amodei, D. (2020). Language Models are Few-Shot Learners. Advances in Neural Information Processing Systems (NeurIPS), 33, 1877-1901. Retrieved from https:// arxiv.org/ abs/ 2005.14165

[11]. Alkadi, R., et al. (2019). A Review of Deep Learning Methods for MRI Brain Tumor Segmentation. Journal of Imaging, 5(6), 69. Retrieved from https://www. mdpi.com/ journal/jimaging.