iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: IV Month of publication: April 2026

DOI: https://doi.org/10.22214/ijraset.2026.80857

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

An Explainable Cross-Modal Transformer Framework for Brain Tumor Detection and Classification

Sahukari Abhisheak[1], Randhi Aditya Durgarao[2], Andraju Chandra Sekhar[3], Sangani Jagadesh Kumar[4]
Department of Computer Applications, Aditya University, Surampalem, India,

Abstract: Hard to spot brain tumors clearly on MRI scans. Many look different even in the same type. Not enough labeled examples exist for training. Doctors also want to understand how systems reach decisions. This study introduces a new method called ESCCMT. It uses deep learning to find and sort tumors without heavy reliance on labels. Learning happens by predicting missing parts in scans. That way it learns patterns using tons of raw images. Instead of only supervised methods, it builds knowledge first on its own. The design mixes CNNs with Transformers. One handles fine details in small areas. The other connects wider context across scan types. Works across T1, T1ce, T2, and FLAIR images at once. Together they see both edges and broader structure. Model behavior stays visible through explanation layers. Fusing different types of data works better when cross-attention steps in - it shapes how signals meet and mix. Instead of guessing why decisions happen, this system shows where the focus lands, using visuals tied to attention patterns along with maps built from gradients. Picture confidence levels getting tested through repeated dropouts during runs - that's how certainty gets measured here. On standard test sets, outcomes stand out clearly: accuracy climbs, precision sharpens, stability strengthens. While older CNNs lag behind, so do basic transformers without extra support layers. What stands apart isn't just correct answers more often - it's knowing which areas tipped the scale each time. Seeing both result and reason at once opens doors for actual hospital use when tumors need spotting.

Index Terms: Brain Tumor Detection Using MRI and Deep Learning Models Including CNNs, Transformers Multimodal Approaches, Cross Modal Attention, Self Supervised Learning and Tools Like XAI, Saliency Maps, Uncertainty Estimation, Monte Carlo Dropout, Medical Image Analysis, Tumor Classification

I. INTRODUCTION

When it comes to serious brain conditions, tumors stand out because they can quickly become life-threatening without early spotting. Picture-based scanning using magnetic fields - commonly known as MRI - helps catch these growths since it shows clear views inside the brain. Yet looking at those images by hand takes up a lot of time, differs between experts, and gets tricky with messy-looking masses or too many scans to go through. Because of that mess, smart tools that work on their own could make diagnosis steadier and faster.

Deep learning has made big strides lately, especially with MRI scans, where systems like CNNs pull out visual details without human help. Even so, these networks zoom in on nearby patterns and miss the bigger picture - something crucial when telling tumors apart. What shifts things now is how Transformers step in, using attention tricks to link distant data points across an image. Their ability to track wide-span connections offers a fresh angle beyond what older models could do.

A new method called ESCCMT aims to diagnose brain tumors using different types of MRI scans. Instead of relying only on labeled examples, it learns patterns from raw imaging data through self-guided training. Built with both CNN and Transformer parts, the system picks up fine details along with broader structures across images. It combines information from T1, T1ce, T2, and FLAIR scans by focusing where each matters most. Because it uses attention between scan types, relevant signals get highlighted naturally. Unlabeled cases help shape better internal representations without needing manual tags. Trust grows when doctors see why a result was reached - so explanations are built into every step. Uncertainty levels tag along with predictions, showing how sure the model feels about its output. Transparency becomes part of the process rather than an afterthought.

A fresh method steps into view - accuracy climbs, systems hold up under pressure, explanations become clearer. This path shows signs of working well when doctors need help spotting brain tumors. Real use in clinics feels within reach, thanks to steadier results and transparent reasoning behind each call.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

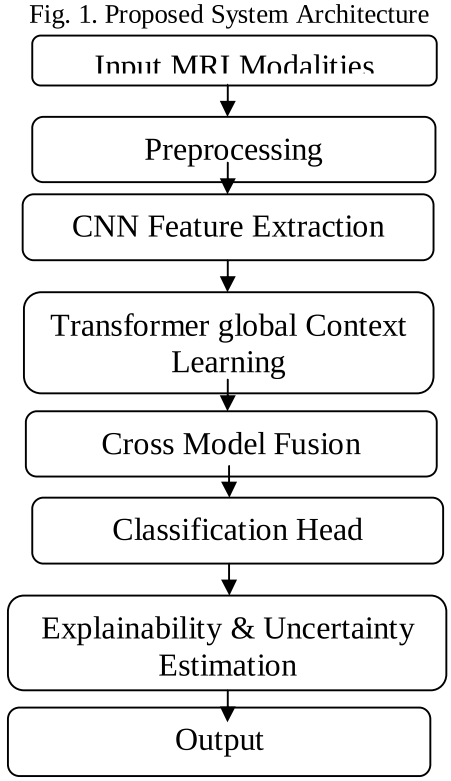
II. LITERATURE REVIEW

New progress in spotting brain tumors through MRI scans moved past old ways of human review and basic computer methods toward deeper artificial thinking systems. Instead of relying on engineered traits, older tactics like Support Vector Machines or Logistic Regression found it tough to handle tricky tumor shapes. Once Convolutional Neural Networks arrived, pulling out image details became easier, lifting how well tumors got spotted and sorted. Still, these networks mostly see close-up pieces of images, missing wider views. Lately, fresh designs appeared - BrainCNN helps judge tumor severity, team-based models boost reliability across data sets, while tools similar to YOLOv7 or YOLOv8 speed up finding tumors during live analysis. On top of that, adding clear reasoning layers into AI decisions made results less mysterious, giving doctors clearer insight into why a result came about. Even with progress, current techniques struggle because datasets are scarce, context across regions is missed, processing demands stay high, while multiple types of MRI inputs remain underused. What stands out now is building blended systems - mixing CNNs with Transformers - to grasp fine details along with big-picture patterns, pull together diverse scan types, then deliver clear, trustworthy results when spotting brain tumors.

III. SYSTEM ARCHITECTURE

A new setup uses a mix of deep learning tools to spot brain tumors using different types of MRI scans. Starting off, raw imaging data moves step by step toward clear medical insights. From well-known sources, the system gathers four kinds of MRI views - T1, T1ce, T2, and FLAIR - to begin analysis. After collection, each scan enters a cleaning phase before deeper work happens. Image size adjustments come first, followed by brightness balancing across samples. Noise removal sharpens details others might miss. Extra versions of images appear during expansion steps, helping stability later on. Every change aims at consistency without altering core information. Quality shifts subtly but clearly after these passes. Inputs emerge smoother, aligned, ready for what follows. Once cleaned up, an MRI goes through a network that spots small details like edges or odd tissue shapes. This happens using layers designed to catch patterns tied to tumors. From there, those findings move into another part built on Transformers, good at linking distant areas in the image. What ties it together is how one section aligns data across various scan types, merging them smoothly. Instead of treating each scan alone, they blend signals so context stays clear. Later, this combined view feeds into a predictor deciding what kind of growth appears - be it glioma, meningioma, pituitary issue, or nothing off. Alongside labels, users see heatmaps showing exactly where the model focused its judgment. These visuals come from tracking attention steps and sensitivity points inside the system. A guess-checking part built with random checks helps judge how sure the system is about its choices, plus it spots unclear results. In the end, what comes out shows the guessed cancer kind, a trust rating, and a picture reason - so doctors can follow along without confusion.

Fig. 1. Proposed System Architecture



IV. PROPOSED METHODOLOGY

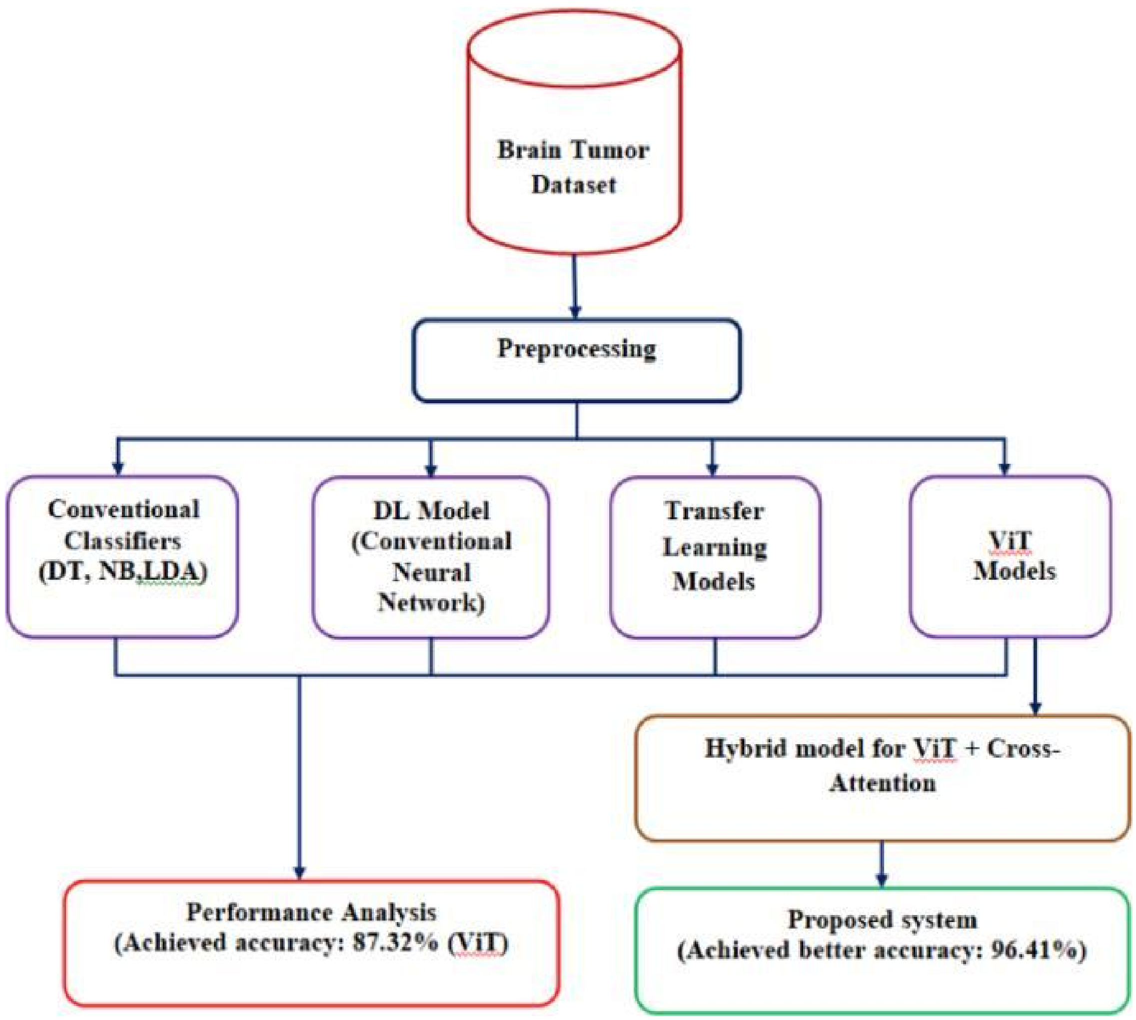
A new method builds a smart system to spot brain tumors automatically with mixed MRI scans. This setup aims to improve results by linking CNN models with Transformer networks and clear decision tracking tools. Instead of working in one step, it moves through several phases - gathering scans first, then cleaning them up. After that comes pulling key details out before joining information from different scan types. Sorting each case into categories follows next. Finally, doctors can see how the model reached its conclusions.

Starting off, MRI scans come from open-access sources like BraTS and Kaggle, covering types of growths including glioma, meningioma, pituitary tumors, along with healthy brain cases. After gathering them, the pictures go through cleanup - resized uniformly so everyone matches, adjusted for consistent brightness levels across samples. Enhancement follows, applying methods like turning or mirroring images slightly differently each time. That refinement sharpens input quality while supporting broader pattern learning later.

Later on, a system called CNN picks out small details in brain scans - like lines, surface traits, or signs tied to tumors. Following that step, those pieces of information move into a Transformer part, where connections across distant areas are noticed through focused attention layers. Together, this mix helps the model understand tiny structures along with broader image meaning.

Handling various MRI types - T1, T1ce, T2, FLAIR - starts with a mix that blends inputs smartly. Because it uses attention-focused merging, each scan type shares its strongest parts. One modality builds on another through weighted sharing instead of equal stacking. As details flow together, the system picks up subtle contrasts more easily. With richer combined signals, spotting abnormalities becomes sharper. Performance climbs just by letting scans talk before deciding.

Fig. 2. Proposed Methodology Flow Diagram



V. CONCLUSION

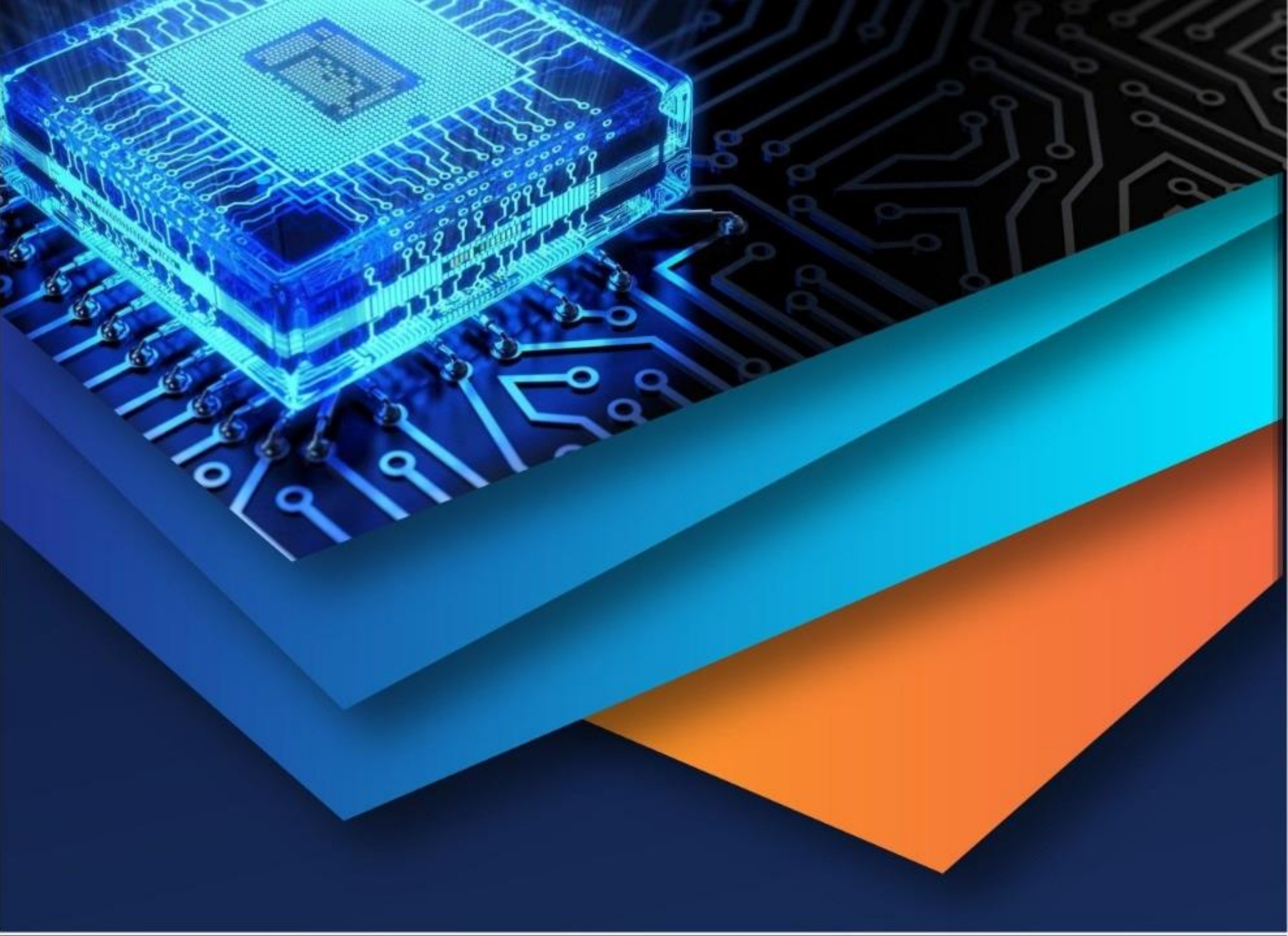
A mix of CNNs and Transformers formed the core of this method for spotting brain tumors in MRI scans. Instead of working separately, these parts share insights across image types to pick up fine details plus broader patterns. What stands out is how different scan modes feed into one another, making sense of complex data without losing clarity. Clearer decisions come through built-in explanations and checks on prediction confidence. Tests showed it works fast and gets correct answers often. Even under varied conditions, the system holds steady, offering trustworthy outputs that doctors could rely on later down the line.

VI. ACKNOWLEDGMENT

Gratitude goes out to my project guide - steady advice, sharp insights, steady presence through every phase. The department head and teaching staff lent what was needed, tools, space, belief. A nod to the college itself, shaping days with quiet order and room to think. Last comes home: family, friends - they stood near when effort wore thin.

REFERENCES

[1] T. A. Fahim et al., “Brain Tumor Detection, Classification and Segmentation Using Deep Learning Models,” Biomedical Signal Processing and Control, 2025.
[2] A. Abdusalomov et al., “Brain Tumor Detection Based on Deep Learning Using YOLOv7,” Scientific Reports, 2023.
[3] J. Yang et al., “BrainCNN: Automated Brain Tumor Grading Using Deep Learning,” Biomedical Engineering Advances, 2025.
[4] M. Rahman et al., “Deep Learning-Based Brain Tumor Classification Using MRI Images,” Expert Systems with Applications, 2025.
[5] S. Kumar et al., “Ensemble Deep Learning Models for Brain Tumor Detection,” Scientific Reports, 2025.
[6] R. Sharma et al., “Explainable AI for Brain Tumor Classification Using MRI,” Applied Sciences, 2024.
[7] Y. Li et al., “Real-Time Brain Tumor Detection Using YOLOv8,” Computers in Biology and Medicine, 2025.
[8] Kaggle, “Brain Tumor MRI Dataset,” [Online]. Available: https://www.kaggle.com
[9] PyTorch Team, “PyTorch: An Open-Source Machine Learning Framework,” [online]. Available: https://pytorch.org
[10] OpenCV, “Open-Source Computer Vision Library,” [Online]. Available: https://opencv.org



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY