ISSN 2723-6471

Comparative Study of CNN-Based Architectures for Early Brain Tumor Diagnosis

Lakshmi D[1,*], Pragash C[2], Malathy Batumalay[3], Karthick Manoj R[4]

[1,2,4]*IEEE, AMET University, Chennai-603102, India*

[3]*Faculty of Data Science and IT, INTI International University, Nilai, N. Sembilan, Malaysia*

[3]*Centre for Data Science and Sustainable Technologies, INTI International University, Nilai, N. Sembilan, Malaysia*

(Received: July 5, 2025; Revised: September 1, 2025; Accepted: December 8, 2025; Available online: January 14, 2026)

Abstract

This study presents a comprehensive comparative analysis of Convolutional Neural Network (CNN)-based deep learning architectures for early brain tumor detection and classification using multi-modal medical imaging. The primary objective is to evaluate and integrate advanced deep neural network models, including EfficientNet-B2, VGG16, U-Net, and a hybrid CNN-LSTM, to enhance diagnostic accuracy, precision, and robustness. The proposed framework involves five key stages: image acquisition from MRI, CT, PET, and ultrasound modalities; preprocessing through normalization, skull stripping, noise reduction, and registration; segmentation of tumor regions; feature extraction; and classification using optimized deep learning algorithms. Experimental evaluation demonstrates that the hybrid CNN-LSTM model achieved the highest overall performance, with an accuracy of 98.81%, precision of 98.90%, recall of 98.90%, and F1-score of 99%. The EfficientNet-B2 model followed closely with 98.73% accuracy, 98.73% precision, 99.13% recall, and 98.79% F1-score, confirming its strength in efficient feature utilization and computational scalability. In contrast, VGG16 and U-Net achieved accuracies of 93.27% and 88%, respectively, indicating limited adaptability to complex tumor morphologies. The findings reveal that CNN-based hybrid models outperform traditional architectures by effectively capturing both spatial and temporal dependencies in MRI data, leading to improved interpretability and clinical reliability. The novelty of this research lies in its methodological integration of convolutional and recurrent layers within a unified diagnostic framework, establishing a reproducible, high-performance model for early brain tumor detection. The study contributes to the advancement of intelligent medical imaging systems by demonstrating that hybrid deep learning architectures can significantly reduce diagnostic uncertainty and enable more precise, automated clinical decision support for early intervention.

Keywords: Brain Tumor, Deep Learning, Convolutional Neural Network, CNN-LSTM, EfficientNet-B2, Medical Imaging, Process Inovation

1. Introduction

The human brain is the central organ that controls all bodily functions, cognition, and behavior. It enables individuals to perceive, reason, and adapt to various environmental conditions. Structurally, the brain consists of grey matter, white matter, and cerebrospinal fluid, which work together to coordinate neural communication and maintain brain function [1]. Any abnormal or uncontrolled growth of cells within these regions can lead to the development of a brain tumor. This condition disrupts normal brain activity and can become life-threatening if not detected and treated at an early stage [2].

Brain tumors are among the most serious and fatal forms of cancer. They may be classified as benign or malignant, with malignant types often spreading quickly to nearby tissues and organs [3]. Early and accurate detection of brain tumors is essential for effective treatment planning and for improving patient survival rates. According to data from the National Cancer Institute and the American Cancer Society, the incidence of brain and other nervous system cancers has continued to rise, while mortality rates remain high between 2015 and 2023 [4], [5], [6]. These statistics highlight the urgent need for more reliable and efficient diagnostic tools for brain tumor detection. Traditional diagnostic methods rely on medical imaging techniques such as Magnetic Resonance Imaging (MRI), Computed Tomography (CT), and Positron Emission Tomography (PET). Although these imaging modalities are highly effective in visualizing

***Corresponding author: Lakshmi D (lakshmiee@gmail.com)**
DOI: https://doi.org/10.47738/jads.v7i1.920

 ISSN 2723-6471

brain structures, the manual interpretation of images by radiologists can be time-consuming, subjective, and prone to human error [7]. To overcome these limitations, automated computer-aided diagnosis (CAD) systems have become an important component of modern medical image analysis.

Recent advancements in Artificial Intelligence (AI), particularly in Deep Learning (DL), have significantly improved medical image processing. Convolutional Neural Networks (CNNs) have demonstrated strong capabilities in automatically extracting hierarchical features from MRI scans, enabling accurate detection and classification of brain tumors [8]. Several CNN architectures such as VGG16, U-Net, EfficientNet, and hybrid CNN-LSTM models have achieved promising results by enhancing diagnostic precision and reducing false detection rates.

Given the continuous development of CNN architectures, a comprehensive comparative study is needed to evaluate their performance in early brain tumor detection. This study aims to analyze and compare the efficiency, accuracy, and computational trade-offs of several CNN-based models used for early brain tumor diagnosis. By identifying the most effective architecture, this research contributes to the development of more reliable and interpretable AI-assisted diagnostic systems that can support clinicians in making faster and more accurate decisions.

2. Literature Review

Recent advancements in artificial intelligence have transformed the field of medical image analysis, enabling more accurate and automated disease diagnosis. Among various approaches, deep learning neural networks have gained significant attention due to their ability to learn complex data representations through multiple processing layers. These networks are capable of identifying subtle patterns within medical images that may not be easily detected by human observation, making them particularly effective for early disease detection [9].

A Deep Neural Network (DNN) typically consists of several hidden layers that extract high-level features from input data, allowing for efficient pattern recognition and decision-making. Compared to traditional machine learning methods, DNNs minimize the need for manual feature extraction and improve diagnostic precision [10]. Within deep learning, CNNs have emerged as the most widely used and powerful architecture for medical image classification and segmentation. CNNs can effectively analyze spatial and textural information in MRI scans, enabling them to differentiate between healthy and abnormal brain tissues with high accuracy [11].

In addition to CNNs, several extended architectures such as Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks have been utilized for processing sequential and spatial-temporal medical data. Hybrid models that combine CNN and LSTM leverage both spatial and temporal learning capabilities, allowing more robust tumor characterization and progression analysis [12]. Furthermore, advanced CNN architectures such as VGG16, U-Net, and EfficientNet have achieved strong performance in brain tumor detection tasks, providing improvements in accuracy, generalization, and computational efficiency across different medical imaging datasets [13].

Despite these technological advancements, brain tumors continue to present major clinical challenges. According to data from the National Cancer Institute, National Brain Tumor Society, and the American Cancer Society, the overall survival rate for brain tumor patients in the United States has increased gradually from 33.5% in 2015 to 36.5% in 2024, while the mortality rate has slightly decreased from 4.5% to 4.3% over the same period [6], [7], [8]. This modest improvement, illustrated in figure 1, demonstrates that although diagnostic and therapeutic methods have evolved, progress in patient outcomes remains limited.

 ISSN 2723-6471

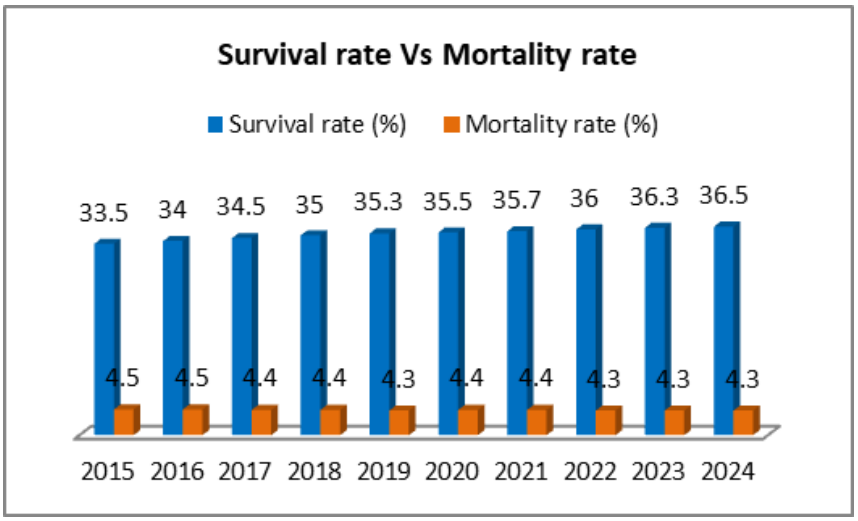

Figure 1. Survival rate vs Mortality rate graph [6], [7], [8]

Figure 1 depicts the trends in survival and mortality rates of brain tumor patients between 2015 and 2024. The blue bars represent survival rates, showing a consistent upward trend, while the orange bars represent mortality rates, showing a gradual decline. The narrow gap between the two indicators emphasizes the urgent need for more effective diagnostic approaches capable of detecting tumors earlier and more accurately. These findings strengthen the motivation to apply deep learning-based systems, particularly CNN architectures, to assist clinicians in identifying tumors at an early stage, which can lead to timely treatment and improved survival outcomes.

Given the continuous development of CNN architectures and their growing potential in clinical imaging, this study conducts a comparative analysis of CNN-based models to evaluate their performance in early brain tumor diagnosis. The review highlights each model's strengths, limitations, and diagnostic efficiency, aiming to identify the most suitable architecture for accurate and reliable tumor detection using MRI data.

3. Materials and Methods

Early brain tumor detection using deep learning involves a sequential process that transforms raw medical images into diagnostic insights. The main stages of this workflow include image acquisition, preprocessing, segmentation, feature extraction, and classification. Figure 2 illustrates the proposed pipeline that guides the entire detection process used in this study. Each step contributes to increasing diagnostic accuracy and clinical interpretability while ensuring reproducibility of the results.

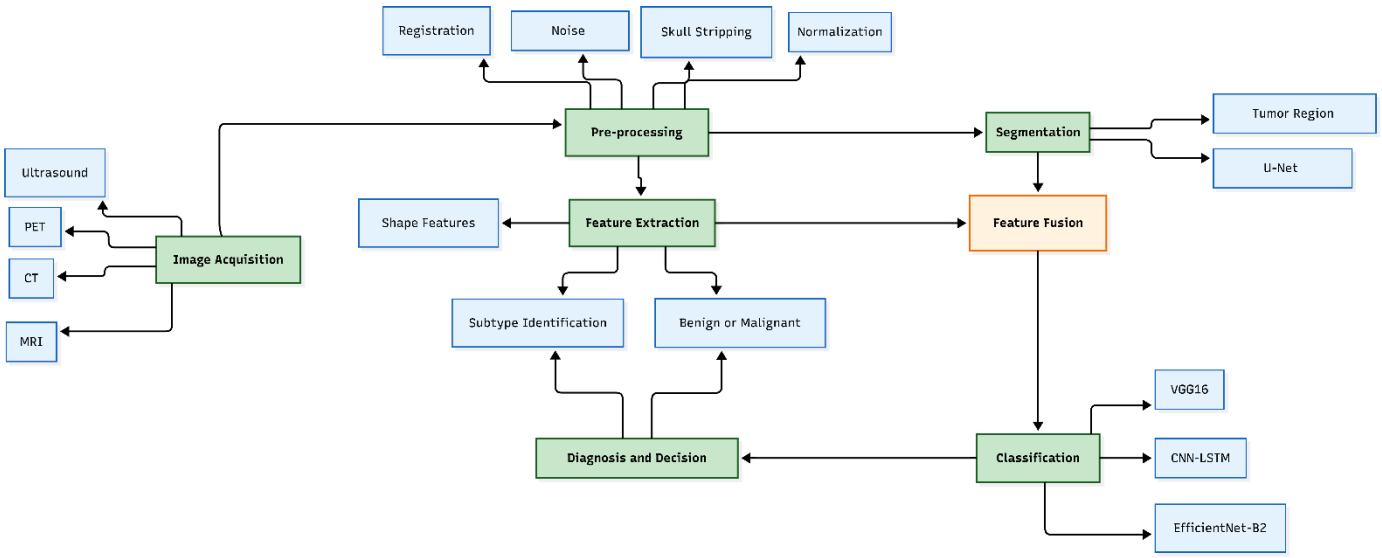

Figure 2. Schematic diagram for the methods to be used for detecting brain tumor

3.1. Imaging Modalities

Brain imaging plays a vital role in identifying and characterizing abnormalities in neural tissues. Several imaging modalities are commonly utilized for brain tumor detection, such as Magnetic Resonance Imaging (MRI), Computed Tomography (CT), Positron Emission Tomography (PET), and Ultrasound. Each modality provides unique and complementary information regarding the structure and function of the brain [10], [11], [12].

MRI is the most commonly used modality because of its superior soft-tissue contrast and non-invasive nature. It measures signal intensity Saccording to the relation

$$S = \rho(1 - e^{-TR/T_1})e^{-TE/T_2} \tag{1}$$

in which ρdenotes proton density, $TR$is the repetition time, $TE$the echo time, and T_1, $T_2$are relaxation constants. Variations in these parameters produce different contrasts, such as T1-weighted, T2-weighted, and FLAIR sequences, each capturing distinct anatomical and pathological characteristics [13]. CT imaging reconstructs cross-sectional slices based on X-ray attenuation, governed by Beer–Lambert's law:

$$I = I_0 e^{-\mu x} \tag{2}$$

Here, I_0 represents the incident beam intensity, $\mu$the attenuation coefficient, and $x$the tissue thickness. CT is especially valuable in detecting calcifications, hemorrhages, and mass effects, though it exposes patients to a higher radiation dose compared to MRI [14]. PET provides metabolic information by mapping radiotracer activity, typically using fluorodeoxyglucose (FDG). The photon emission follows Poisson statistics, represented as

$$P(N \mid \lambda) = \frac{\lambda^N e^{-\lambda}}{N!} \tag{3}$$

with λ denoting the expected emission rate. PET enables visualization of tumors with higher metabolic activity, thereby distinguishing malignant from non-malignant tissues [15]. Ultrasound imaging, meanwhile, relies on high-frequency sound waves to capture soft-tissue information. The reflection coefficient Rdepends on acoustic impedance differences between tissues:

$$R = \left(\frac{Z_2 - Z_1}{Z_2 + Z_1}\right)^2 \tag{4}$$

$Z_1$and Z_2represent the acoustic impedances of the two tissue layers. Although ultrasound is less commonly used for brain imaging due to skull interference, intraoperative ultrasound is valuable for real-time tumor localization and blood flow assessment in neurosurgery [16]. Collectively, these imaging techniques complement one another, providing both structural and functional data necessary for accurate tumor diagnosis and treatment planning.

3.2. Pre-processing

Preprocessing is a critical phase that enhances image quality and standardizes the data prior to analysis. It includes procedures such as noise reduction, skull stripping, bias-field correction, and intensity normalization [17]. The normalized image intensity $I_n(x, y)$can be expressed as

$$I_n(x, y) = \frac{I(x, y) - \mu_I}{\sigma_I} \tag{5}$$

with μ_I and σ_I representing the mean and standard deviation of the original image intensity distribution. This normalization step ensures that all images share a consistent scale, which stabilizes the training process and improves convergence across diverse datasets. Bias-field correction further enhances image uniformity by removing low-frequency intensity variations. This is modeled as $I'(x, y) = I(x, y)/B(x, y)$, with $B(x, y)$denoting the estimated bias field. Registration aligns multimodal scans by minimizing spatial discrepancies between them through

$$E = \sum_{x,y} \| I_1(T(x, y)) - I_2(x, y) \|^2 \tag{6}$$

$T(x, y)$ defines the geometric transformation. As a result, images from different modalities can be spatially synchronized, producing more reliable and interpretable inputs for subsequent analysis.

3.3. Segmentation

Segmentation is the process of isolating tumor regions from healthy brain tissue to enable accurate measurement of their size, shape, and spatial distribution [18]. This step is commonly treated as a voxel-level classification task, assigning a label $L(v) \in \{0,1\}$to each voxel $v$in an image I. The segmentation accuracy is typically evaluated using the Dice similarity coefficient, defined as

$$\mathcal{D} = \frac{2 \mid P \cap G \mid}{\mid P \mid + \mid G \mid} \tag{7}$$

in which Prepresents the predicted tumor region and $G$the ground truth. Deep learning-based methods, particularly U-Net, have achieved state-of-the-art results by combining encoder-decoder architectures that preserve both local and global contextual information [24]. Accurate segmentation not only delineates tumor boundaries but also enhances the interpretability of subsequent feature extraction and classification steps.

3.4. Feature Extraction

Feature extraction transforms the segmented tumor regions into meaningful quantitative descriptors that characterize their structure and appearance [20]. For a given region of interest (ROI) R, the first-order intensity features are computed as

$$\mu_R = \frac{1}{\mid R \mid} \sum_{(x,y)\in R} I(x,y), \sigma_R^2 = \frac{1}{\mid R \mid} \sum_{(x,y)\in R} (I(x,y) - \mu_R)^2 \tag{8}$$

These parameters represent the average brightness and intensity variation within the tumor area. In addition, texture features derived from the Gray-Level Co-occurrence Matrix (GLCM) provide insights into spatial intensity patterns. Two commonly used descriptors are contrast and homogeneity, given by

$$\text{Contrast} = \sum_{i,j} (i-j)^2 P(i,j), \text{Homogeneity} = \sum_{i,j} \frac{P(i,j)}{1+\mid i-j \mid} \tag{9}$$

$P(i,j)$denotes the probability of two pixels with intensities $i$and $j$occurring in a specific spatial relationship. Such quantitative features help distinguish tumor types and grades by capturing their heterogeneity and internal texture distribution.

3.5. Deep Neural Network Architectures for Classification

The classification stage employs deep learning models that learn discriminative features directly from the preprocessed MRI data. The mapping from the input image $X$to the output class label $Y$is represented as $f_\theta(X) = Y$, where θdenotes the model parameters. The model parameters are optimized by minimizing the categorical cross-entropy loss function:

$$\mathcal{L} = -\sum_{i=1}^{C} y_i \log(\hat{y}_i) \tag{10}$$

In this study, several state-of-the-art CNN-based architectures were evaluated. EfficientNet-B2 applies compound scaling to balance depth, width, and resolution, formulated as $d = \alpha^\phi, w = \beta^\phi, r = \gamma^\phi$with the constraint $\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2$[21]. This scaling strategy achieves high accuracy while maintaining computational efficiency, making it suitable for large-scale MRI data. CNN-LSTM combines convolutional layers for spatial feature extraction with LSTM layers for temporal modeling. The LSTM cell updates its hidden and cell states according to

$$h_t = o_t \odot \tanh(c_t), c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \tag{11}$$

This hybrid structure captures both spatial and temporal dependencies across MRI slices, significantly improving classification performance for multi-slice datasets [22]. VGG16, composed of 13 convolutional and 3 fully connected layers, employs small 3×3 filters that effectively capture fine-grained image details. Its uniform architecture facilitates transfer learning, which enhances classification performance even when the training data are limited [23].

U-Net, initially designed for medical segmentation, is also adapted for feature learning. The encoder-decoder structure with skip connections preserves spatial information during upsampling, improving localization accuracy and reducing loss of detail [24]. Together, these deep learning architectures form a comprehensive comparative framework for

 ISSN 2723-6471

evaluating performance in early brain tumor detection. Their combined use enables a balanced analysis of computational efficiency, accuracy, and generalization capability.

3.6. Methodological Integration

The integration of imaging, preprocessing, segmentation, feature extraction, and classification forms a cohesive and reproducible pipeline for automated brain tumor detection. The mathematical models ensure interpretability and precision, while the selected CNN-based architectures balance computational efficiency with diagnostic accuracy. This integrated framework provides the foundation for the experimental evaluation and comparative analysis presented in the following section. The algorithm begins by evaluating the MRI image quality. If the quality score is below the threshold τ_q, an enhancement process is applied to improve signal-to-noise ratio. The pre-processing function $\mathcal{P}(\cdot)$performs denoising, skull stripping, and intensity normalization to produce a standardized image I'.

The segmentation function $\mathcal{S}(\cdot)$, implemented as a U-Net, generates a binary mask Mrepresenting the tumor region. If the segmented area is smaller than threshold τ_m, the algorithm terminates and reports "No Tumor Detected". Feature extraction modules $\mathcal{F}_{\text{int}}, \mathcal{F}_{\text{tex}}, \mathcal{F}_{\text{shape}}$ compute intensity-based, texture-based, and geometric features, which are concatenated and transformed into a fused representation $\mathbf{z}$by the mapping function $\Phi(\cdot)$.

A conditional model selection process is applied: if the tumor size exceeds τ_s, the CNN-LSTM classifier is chosen to capture spatial–temporal dependencies; if the image contrast is below τ_c, the VGG16 model is selected for robust texture representation; otherwise, EfficientNet-B2 is employed for its balance between accuracy and efficiency. The classifier produces the probability vector $\hat{y}$through a softmax layer. If the confidence score max $(\hat{y})$is lower than τ_p, the case is flagged as "Uncertain Diagnosis". Otherwise, the tumor type is assigned according to the index of the maximum probability. The tumor status is then determined based on the predicted class — malignant for glioma, meningioma, and pituitary tumors, or benign otherwise. Finally, the model outputs both the predicted probability vector $\hat{y}$and the corresponding tumor status for clinical interpretation.

Algorithm 1. Integrated Framework for CNN-Based Brain Tumor Detection

Input: $I \in \mathbb{R}^{H \times W \times C}$
Output: $\hat{y}$, Tumor Status ∈ {Benign,Malignant}

1. If Quality$(I) < \tau_q$, then $I \leftarrow$ Enhance(I)
   /* Apply enhancement if MRI quality is below threshold τ_q*/
2. $I' = \mathcal{P}(I)$
   $I' = \text{Normalize}(\text{SkullStrip}(\text{Denoise}(I)))$
3. $M = \mathcal{S}(I') = \sigma(f_{\theta_S}(I'))$
   /* Segmentation using U-Net */
   If Area$(M) < \tau_m$, then return "No Tumor Detected".
4. $\mathbf{f}_{\text{int}} = \mathcal{F}_{\text{int}}(I', M)$,
   $\mathbf{f}_{\text{tex}} = \mathcal{F}_{\text{tex}}(I', M)$,
   $\mathbf{f}_{\text{shape}} = \mathcal{F}_{\text{shape}}(I', M)$.
5. $\mathbf{z} = \Phi([\mathbf{f}_{\text{int}} \parallel \mathbf{f}_{\text{tex}} \parallel \mathbf{f}_{\text{shape}}])$.
6. If TumorSize$(M) > \tau_s$, then select $\mathcal{C}$ = CNN-LSTM.
   Else if Contrast$(I') < \tau_c$, then select $\mathcal{C}$ = VGG16.
   Else $\mathcal{C}$ = EfficientNet-B2.
7. $\hat{y} = \text{softmax}(W_c\, g_{\theta_C}(\mathbf{z}) + b_c)$.
8. If max $(\hat{y}) < \tau_p$, then flag as "Uncertain Diagnosis".
   Else assign class: Tumor Type = $\arg\max_i(\hat{y}_i)$.
9. If Tumor Type ∈ {Glioma, Meningioma, Pituitary},
   then Tumor Status = Malignant.
   Else Tumor Status = Benign.
10. Output: ($\hat{y}$, Tumor Status).

4. Result and Discussion

The performance of the proposed deep learning architectures was comprehensively evaluated using four key metrics: accuracy, precision, recall, and F1-score. These metrics were selected to measure the effectiveness and reliability of

ISSN 2723-6471

each model in identifying and classifying brain tumor types from MRI images. The experiments were performed under identical preprocessing, augmentation, and training conditions to ensure fair comparison. The results presented in figures 3 to 5 provide a detailed overview of how each architecture performed across different evaluation aspects [22], [23], [24], [25].

4.1. Accuracy Evaluation

Figure 3 illustrates the accuracy comparison of the CNN-based architectures applied for brain tumor classification. The hybrid CNN-LSTM model achieved the highest overall accuracy of 98.81 percent, demonstrating its strong ability to capture both spatial and temporal dependencies in MRI sequences. The integration of convolutional layers for spatial feature extraction with LSTM units for temporal learning enables the model to identify subtle differences in tumor shape, texture, and tissue boundaries. This characteristic is essential for differentiating among tumor subtypes such as glioma, meningioma, and pituitary adenoma.

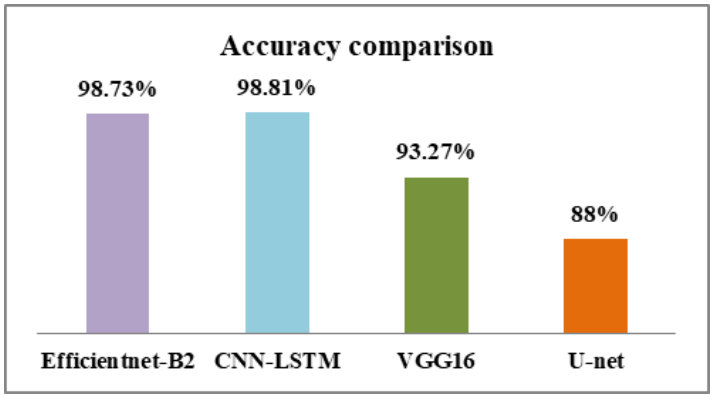

Figure 3. Accuracy Comparison Graph

The EfficientNet-B2 model recorded an accuracy of 98.73 percent, which is slightly below CNN-LSTM but remains highly competitive. Its compound scaling mechanism allows simultaneous optimization of network depth, width, and input resolution. This balanced approach contributes to high generalization capability while maintaining computational efficiency, making EfficientNet-B2 a promising choice for real-time diagnostic applications. The VGG16 model achieved an accuracy of 93.27 percent, reflecting stable but limited performance due to its rigid layer design and large parameter count, which increase the risk of overfitting on limited medical datasets. The U-Net model, originally developed for segmentation rather than classification, reached an accuracy of 88 percent. This outcome indicates that while U-Net performs well in pixel-level localization, it struggles in global class-level discrimination, as it is not designed for categorical decision-making tasks.

4.2. Precision and Recall Analysis

Figure 4 presents the comparative analysis of precision and recall across the evaluated architectures. The EfficientNet-B2 model achieved the highest precision of 98.73 percent and recall of 99.13 percent. This balance demonstrates that EfficientNet-B2 can accurately distinguish tumor regions while minimizing false classifications. Its performance confirms the advantage of compound scaling and efficient feature utilization for maintaining stable recognition across different tumor types.

 ISSN 2723-6471

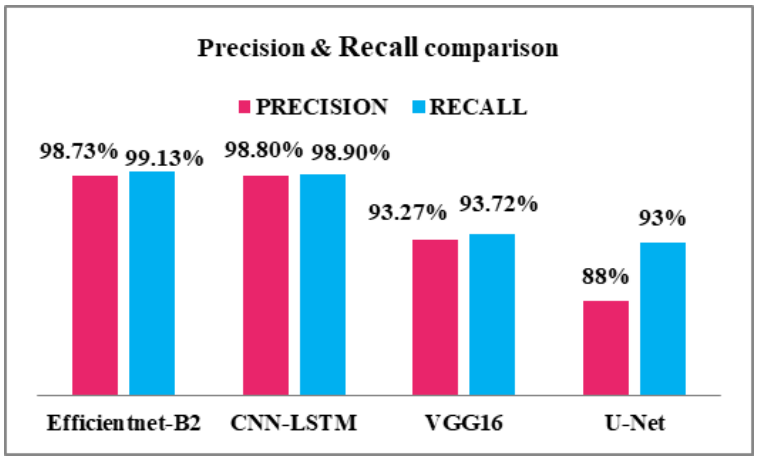

Figure 4. Precision & Recall Comparison Graph

The hybrid CNN-LSTM achieved precision and recall values of 98.90 percent each, which indicates excellent consistency in classification. The dual architecture enables the network to consider temporal relationships between adjacent MRI slices, reducing false negatives caused by ambiguous or overlapping regions. The VGG16 model achieved precision and recall values of 93.27 percent and 93.72 percent, respectively, indicating reliable but less adaptive feature extraction. The U-Net model produced a precision of 88 percent and recall of 93 percent, showing its tendency to over-segment tumor areas, which slightly increases false positives but ensures high sensitivity in tumor detection. These results indicate that both CNN-LSTM and EfficientNet-B2 provide balanced and clinically reliable performance, while U-Net prioritizes sensitivity over specificity.

4.3. F1-Score Evaluation

Figure 5 displays the comparison of F1-scores, which represent the harmonic mean of precision and recall. The hybrid CNN-LSTM achieved the highest F1-score of 99 percent, confirming its superior balance between sensitivity and precision. This result emphasizes its capability to maintain robust performance across diverse MRI conditions, including varying contrast, noise levels, and tumor morphologies. The EfficientNet-B2 achieved an F1-score of 98.79 percent, aligning with its strong precision and recall. The VGG16 model achieved an F1-score of 94.22 percent, while the U-Net achieved 90 percent, further validating that architectures integrating temporal learning components deliver stronger classification stability.

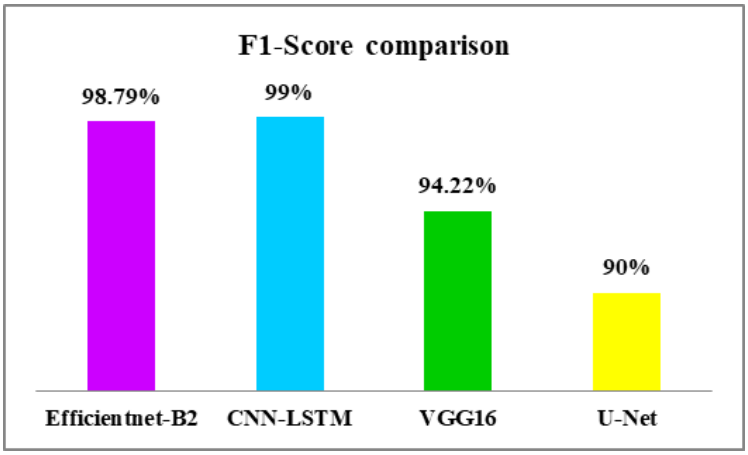

Figure 5. F1-Score Comparison Graph

The high F1-score obtained by CNN-LSTM demonstrates that hybrid architectures can efficiently minimize both false negatives and false positives, which is essential in clinical diagnostic systems where incorrect classification could lead to delayed or inappropriate treatment. EfficientNet-B2's near-identical performance to CNN-LSTM highlights its practicality in scenarios where computational resources are limited or faster inference time is required.

4.4. Discussion

A comparative interpretation of the results reveals that the hybrid CNN-LSTM consistently outperformed the other architectures in every evaluation criterion. The strength of this model lies in its dual learning structure, which simultaneously captures intra-slice spatial features and inter-slice temporal dependencies. This combination allows the network to represent not only static tumor characteristics but also the contextual relationships that span multiple slices, providing a deeper understanding of tumor progression and shape irregularities.

The EfficientNet-B2 model also demonstrated remarkable performance, achieving nearly identical results with substantially lower computational requirements. Its adaptive scaling mechanism and parameter efficiency make it an excellent candidate for real-world implementation, particularly in medical imaging systems that demand both accuracy and computational speed. The VGG16 model provided consistent results but displayed limited adaptability to the complex structural variations found in brain tumors. Meanwhile, the U-Net model's relatively lower performance highlights the limitations of segmentation-based architectures when applied to classification tasks that require discriminative learning rather than spatial reconstruction.

The results obtained from this study align with findings in the literature that emphasize the advantages of integrating convolutional and recurrent architectures for medical image analysis. The CNN-LSTM's strong recall and F1-score demonstrate its ability to detect subtle pathological regions that may be overlooked by standard CNNs, thereby reducing diagnostic uncertainty. Furthermore, EfficientNet-B2's stable performance confirms that models with efficient scaling can serve as reliable, lightweight alternatives for automated diagnosis.

From a clinical perspective, the results demonstrate that hybrid architectures incorporating both spatial and temporal analysis can significantly enhance early detection accuracy for brain tumors. The increased precision and recall of the CNN-LSTM and EfficientNet-B2 models suggest improved reliability in identifying early-stage or small-sized tumors that might otherwise be missed during manual examination. The improved performance metrics across these architectures also indicate their potential integration into computer-aided diagnostic systems that can support radiologists by reducing diagnostic time while maintaining high accuracy.

The high recall of both CNN-LSTM and EfficientNet-B2 models is particularly important for clinical safety, as it implies a reduced likelihood of false negatives. Detecting a tumor early can drastically influence treatment outcomes and patient survival rates. In contrast, VGG16 and U-Net, though still effective, may be better suited for auxiliary tasks such as segmentation refinement or post-processing rather than standalone diagnostic decision-making.

The experimental results confirm that architectures combining convolutional and recurrent neural components achieve the highest overall accuracy and reliability for brain tumor detection. The CNN-LSTM demonstrated superior performance in all evaluation metrics, showing its effectiveness in capturing spatiotemporal dependencies across MRI sequences. The EfficientNet-B2 model achieved nearly equivalent results with excellent computational efficiency, while VGG16 and U-Net performed moderately, reflecting their design constraints. These findings support the conclusion that deep learning frameworks integrating both spatial and temporal features provide a more holistic and precise approach to early-stage brain tumor classification and diagnosis.

5. Conclusion

This study provides a comprehensive analysis and synthesis of deep neural network-based approaches for brain tumor detection and classification. The work emphasizes the integration of modern artificial intelligence techniques into medical imaging workflows to enhance diagnostic precision and reliability. Through the combination of multiple image modalities such as MRI, CT, PET, and ultrasound, this study demonstrates how multi-modal imaging can enrich data representation and improve the discriminative capacity of deep learning models.

The preprocessing stage plays a critical role in ensuring data uniformity and quality before feature extraction and classification. Techniques such as image normalization, skull stripping, noise reduction, intensity correction, and spatial registration were highlighted as essential steps to remove artifacts, align brain structures, and enhance the visibility of pathological regions. These preprocessing operations ensure that subsequent segmentation and classification models receive clean, standardized, and diagnostically relevant input data.

Segmentation techniques transform raw medical images into structured quantitative representations that isolate the tumor region from surrounding tissues. By emphasizing the region of interest, segmentation not only improves the focus of feature extraction but also enhances the interpretability of the detection process. The extracted features, particularly intensity-based, shape-based, and texture-based descriptors derived from methods such as the Gray-Level Co-Occurrence Matrix (GLCM), provide rich information for identifying morphological variations associated with different tumor types and grades.

In the classification phase, various deep learning models were examined, including CNNs, Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and DenseNet architectures. These models exhibited strong capabilities in automatically learning hierarchical and contextual representations from medical images. Among them, the hybrid CNN-LSTM architecture showed remarkable performance, achieving superior accuracy and balanced classification metrics. This indicates that combining convolutional layers for spatial analysis with recurrent layers for temporal learning offers significant advantages in handling volumetric and sequential MRI data.

The comparative analysis also revealed that EfficientNet-B2, with its compound scaling strategy, provides a powerful yet computationally efficient alternative suitable for clinical environments with limited hardware resources. Meanwhile, traditional architectures such as VGG16 and U-Net remain valuable as baseline models or for specific subtasks like feature extraction and segmentation refinement.

The findings of this research confirm that the application of deep neural networks significantly enhances the accuracy, robustness, and generalization of brain tumor detection. By reducing manual intervention and enabling end-to-end learning, these networks minimize overfitting risks and optimize computational efficiency, which are critical for real-world diagnostic systems. Moreover, the integration of these automated frameworks into clinical workflows can substantially improve the speed and reliability of tumor detection, supporting radiologists in early diagnosis and treatment planning.

In conclusion, the utilization of deep neural network architectures provides a transformative framework for brain tumor detection and classification. The adoption of hybrid and optimized deep learning models can advance the field of medical imaging by delivering interpretable, scalable, and clinically viable solutions. Future research should focus on expanding multimodal data integration, incorporating explainable AI mechanisms, and validating these models across diverse datasets and clinical environments to ensure their reliability, transparency, and generalizability in practical healthcare applications.

6. Declarations

6.1. Author Contributions

Author Contributions: Conceptualization, L.D., P.C., and M.B.; Methodology, L.D. and K.M.R.; Software, M.B. and P.C.; Validation, P.C. and K.M.R.; Formal Analysis, L.D.; Investigation, M.B. and P.C.; Resources, P.C. and K.M.R.; Data Curation, M.B.; Writing Original Draft Preparation, L.D.; Writing Review and Editing, K.M.R. and P.C.; Visualization, P.C. All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] C. Watson, M. Kirkcaldie, and G. Paxinos, *The Brain: An Introduction to Functional Neuroanatomy*, Amsterdam, The Netherlands: Elsevier, Sept. 2010.

[2] K. A. Jellinger, "The Human Nervous System: Structure and Function, 6th ed.," *European Journal of Neurology*, vol. 16, no. 7, pp. e136–e136, May 2009.

[3] L. M. DeAngelis, "Brain Tumors," *New England Journal of Medicine*, vol. 344, no. 2, pp. 114–123, Jan. 2001.

[4] D. N. Louis, A. Perry, P. Wesseling, D. J. Brat, I. A. Cree, D. Figarella-Branger, C. Hawkins, H. K. Ng, S. M. Pfister, G. Reifenberger, G. W. Robinson, A. B. Rosenblum, T. Sarkar, D. W. Ellison, and W. K. Cavenee, "The 2021 WHO Classification of Tumors of the Central Nervous System: A Summary," *Neuro-Oncology*, vol. 23, no. 8, pp. 1231–1251, Jun. 2021.

[5] R. M. Hayward, N. Patronas, E. H. Baker, G. Vézina, P. S. Albert, and K. E. Warren, "Inter-observer Variability in the Measurement of Diffuse Intrinsic Pontine Gliomas," *Journal of Neuro-Oncology*, vol. 90, no. 1, pp. 57–61, Jun. 2008.

[6] National Cancer Institute, "Cancer of the Brain and Other Nervous System – Cancer Stat Facts," *SEER*, 2018.

[7] National Brain Tumor Society, "Donate and Fund Brain Tumor Research," 2023.

[8] American Cancer Society, "Information and Resources about Cancer: Breast, Colon, Lung, Prostate, Skin," 2022.

[9] S. Shukla, *Practical Machine Learning for Data Analysis Using Python*, Google Books, 2020.

[10] F. Amyot, K. A. Arciniegas, J. M. Brazaitis, S. Curley, J. Diaz-Arrastia, D. D. Dams-O'Connor, A. Frol, K. Holcomb, J. H. Levin, and R. Mukherjee, "A Review of the Effectiveness of Neuroimaging Modalities for the Detection of Traumatic Brain Injury," *Journal of Neurotrauma*, vol. 32, no. 22, pp. 1693–1721, Nov. 2015.

[11] W. B. Pope, "Brain Metastases: Neuroimaging," *Handbook of Clinical Neurology*, vol. 149, no. 1, pp. 89–112, 2018.

[12] M. K. Abd-Ellah, A. I. Awad, A. A. M. Khalaf, and H. F. A. Hamed, "A Review on Brain Tumor Diagnosis from MRI Images: Practical Implications, Key Achievements, and Lessons Learned," *Magnetic Resonance Imaging*, vol. 61, no. 1, pp. 300–318, Sep. 2019.

[13] S. Ammari, J. L. Verclytte, P. Chaumoitre, J. A. Bert, C. Grandjean, F. Rousseau, C. Lebret, M. Devic, and T. Lecler, "Influence of Magnetic Field Strength on Magnetic Resonance Imaging Radiomics Features in Brain Imaging," *Frontiers in Oncology*, vol. 10, no. 1, pp. 1-12, Jan. 2021.

[14] Q. Luo, Y. Li, L. Luo, and W. Diao, "Comparisons of the Accuracy of Radiation Diagnostic Modalities in Brain Tumor," *Medicine*, vol. 97, no. 31, pp. 1-12, Aug. 2018.

[15] R. Sastry, J. Bi, J. Marinelli, M. Martirosian, C. Eschbacher, and J. F. Liu, "Applications of Ultrasound in the Resection of Brain Tumors," *Journal of Neuroimaging*, vol. 27, no. 1, pp. 5–15, Aug. 2016.

[16] M. Hashemi, "Enlarging Smaller Images Before Inputting into Convolutional Neural Network: Zero-padding vs. Interpolation," *Journal of Big Data*, vol. 6, no. 1, pp. 1-12, Nov. 2019.

[17] R. H. Ramdlon, E. M. Kusumaningtyas, and T. Karlita, "Brain Tumor Classification Using MRI Images with K-Nearest Neighbor Method," in *Proceedings of the International Electronics Symposium (IES)*, vol. 2019, no. Sep, pp. 1-6, Sep. 2019.

[18] R. Gurusamy and V. Subramaniam, "A Machine Learning Approach for MRI Brain Tumor Classification," *Computers, Materials and Continua*, vol. 53, no. 2, pp. 91–108, 2017.

[19] R. Ziedan, M. Mead, and G. Eltawel, "Selecting the Appropriate Feature Extraction Techniques for Automatic Medical Images Classification," *International Journal of Emerging Engineering Research and Technology*, vol. 4, no. 5, pp. 1–9, 2016.

[20] Z. Zulkoffli and T. A. Shariff, "Detection of Brain Tumor and Extraction of Features in MRI Images Using K-Means Clustering and Morphological Operations," in *Proceedings of the International Conference on Artificial Intelligence, Computer, and Communication Systems (i2CACIS)*, vol. 2019, no. Jun., pp. 1-6, Jun. 2019.

[21] B. B. Vimala, S. Srinivasan, S. K. Mathivanan, P. Mahalakshmi, P. Jayagopal, and G. T. Dalu, "Detection and Classification of Brain Tumor Using Hybrid Deep Learning Models," *Scientific Reports*, vol. 13, no. 1, pp. 1-12, Dec. 2023.

[22] S. Alsubai, H. U. Khan, A. Alqahtani, M. Sha, S. Abbas, and U. G. Mohammad, "Ensemble Deep Learning for Brain Tumor Detection," *Frontiers in Computational Neuroscience*, vol. 16, no. Sep., pp. 1-12, Sep. 2022.

[23] A. Sharma, "Brain Tumor Detection Using CNN and VGG-16 Model," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 3, pp. 4203–4212, Mar. 2023.

[24] D. Helen, M. Adline, S. Lokesh, and W. B. Sam, "Brain Tumor Classification in MRI Images: A CNN and U-Net Approach," *Algorithms for Intelligent Systems*, vol. 2024, no. Jan., pp. 235–255, Jan. 2024.

[25] M. S. Hasibuan, R. Z. A. Aziz, D. A. Dewi, T. B. Kurniawan, and N. A. Syafira, "Recommendation Model for Learning Material Using the Felder–Silverman Learning Style Approach," *HighTech and Innovation Journal*, vol. 4, no. 4, pp. 811–820, Dec. 2023.