International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

CNN-Based System for Automated Detection and Segmentation of Brain Tumors: A Performance Study

Usha Barad*, Shalini Gupta

Submitted:15/08/2024 **Revised:** 27/09/2024 **Accepted:** 01/10/2024

Abstract: Brain tumor classification using magnetic resonance imaging (MRI) is a vital step in assisting clinicians with early diagnosis and treatment planning. Manual examination of MRI scans is often time-consuming and prone to observer variability, which emphasizes the need for automated diagnostic methods. In this study, a convolutional neural network (CNN)-based framework was developed to perform multi-class classification of brain tumors, distinguishing between glioma, meningioma, pituitary tumor, and no tumor. A systematic preprocessing pipeline was applied, including grayscale conversion, resizing, normalization, and augmentation, to standardize the dataset and improve model robustness. Several baseline CNN models with varying complexity were designed and evaluated, and insights from these experiments guided the development of a deeper proposed CNN architecture. Performance evaluation incorporated accuracy, precision, recall, F1-score, and receiver operating characteristic (ROC) analysis, alongside visualization through confusion matrices. The findings highlight that the proposed CNN provides substantial improvements over baseline models and demonstrates strong capability in extracting discriminative features from MRI scans. Overall, the study confirms that CNN-based approaches hold significant promise for reliable and efficient brain tumor classification, offering a pathway toward clinical decision-support systems.

Keywords: *Brain tumor classification; Convolutional Neural Networks; Medical image analysis*

1. Introduction

Today's therapeutic imaging devices generate a vast quantity of images that contain a wealth of information. However, much of this information remains hidden within the data, requiring advanced image analysis algorithms to extract meaningful insights for clinical decision-making and to streamline medical workflows [1, 2]. Deep learning, as a specialized subfield of machine learning, has emerged as a powerful tool for addressing these challenges [3-6]. Rather than aiming for a deeper conceptual understanding, deep learning leverages multiple hierarchical layers of representation to capture increasingly abstract features within medical images.

Medical imaging plays a crucial role in detecting tumors, where early diagnosis significantly improves the chances of successful treatment [7-10]. Technologies such as computed tomography [11-13] have already demonstrated the potential to identify tumors at their earliest stages, often when they are as small as a grain of rice. These advancements have contributed to measurable reductions in cancer-related mortality, particularly in cases such as lung cancer. Beyond this, similar imaging approaches have facilitated early detection of intestinal and breast cancers, offering minimally invasive alternatives to traditional diagnostic procedures. Such methods not only improve detection rates but also reduce patient discomfort, recovery times, and healthcare costs.

Imaging technologies are now integral to public health, serving as essential tools for both diagnosis and monitoring of disease progression [14-16]. They are widely used not only for the detection of previously undiagnosed conditions but also for the follow-up of patients undergoing treatment. In this context, brain tumors stand out as one of the most critical medical concerns, given their increasing prevalence and high mortality rates. Reports indicate a growing global burden of brain tumor cases [17,18], with thousands of new diagnoses each year and a large proportion of patients ultimately succumbing to the disease.

Computer Engineering Department, Gokul Global University, Siddhpur, Gujarat, India

*Email: Email: ushabaradce@gmail.com**

Among the available imaging modalities, magnetic resonance imaging (MRI) remains the most widely adopted for brain tumor diagnosis [19-21]. Its ability to differentiate tissues based on contrast levels makes it particularly effective for detecting structural abnormalities. Despite its advantages, manual examination of MRI scans by clinicians remains time-consuming and prone to error. These limitations have fuelled research into computer-assisted methods that can automate the detection and classification process, thereby improving both efficiency and reliability.

Brain tumors themselves can be classified into two broad categories [22, 23]: primary tumors, which originate within the brain, and secondary tumors, which metastasize from other parts of the body. They may also be categorized as benign or malignant depending on their degree of severity and growth rate. Benign tumors typically grow slowly and have well-defined boundaries, whereas malignant tumors are more aggressive and irregular, often leading to severe outcomes. The World Health Organization further categorizes malignant brain tumors into four grades based on their biological and clinical characteristics [24-26].

Recent years have witnessed significant progress in artificial intelligence and machine learning for medical imaging. Deep learning [27-33], in particular, has shown remarkable promise in automating tasks such as tumor detection, segmentation, and classification. Within this field, convolutional neural networks (CNNs) [34-38] have emerged as a leading architecture due to their ability to learn spatial hierarchies of features directly from raw image data. Unlike traditional image-processing techniques, CNNs require minimal preprocessing and are capable of extracting highly discriminative features across multiple layers. Each neuron in a convolutional layer responds to localized regions of the input, enabling the model to capture both low-level patterns and high-level semantic structures. This makes CNNs particularly well suited for brain tumor analysis in MRI images, where subtle differences in shape, intensity, and texture play a vital role in accurate diagnosis.

2. Literature Review

Medical imaging has become a cornerstone in the diagnosis and treatment of neurological disorders [1-5], with brain tumor detection receiving increasing attention in recent years. Accurate and timely identification of brain tumors is critical, as it directly influences treatment planning and patient outcomes. Traditionally, diagnosis relies heavily on expert radiologists manually analyzing MRI scans [19-21]. While this conventional approach can be highly effective, it has notable limitations: it is time-consuming, subject to inter-observer variability, and often struggles to keep pace with the growing volume of imaging data in modern clinical practice. These challenges have motivated the development of computational models that can assist or automate brain tumor analysis, potentially reducing workload while improving consistency and accuracy.

Early computational methods predominantly utilized classical machine learning techniques. In these approaches, meaningful features such as texture, intensity, shape, and edge information were manually extracted from MRI scans. These handcrafted features were then fed into classifiers like support vector machines (SVMs), k-nearest neighbors (kNN), and decision trees to distinguish between tumor and non-tumor regions or to classify tumor types [36, 39-41]. Although these methods achieved initial success, their performance was heavily dependent on the quality and relevance of the engineered features. Additionally, classical models often lacked robustness when confronted with variations in imaging protocols, scanners, or patient populations, limiting their scalability to larger or more heterogeneous datasets.

The advent of deep learning marked a transformative shift in medical image analysis. Deep learning models, particularly Convolutional Neural Networks (CNNs), can automatically learn hierarchical and discriminative features directly from raw images, eliminating the need for manual feature engineering. CNNs are capable of capturing both local patterns—such as tumor boundaries and textures—and global structures, such as overall brain morphology, which is critical for accurate tumor detection and classification. Their success in general image classification tasks led to rapid adoption in medical imaging, where high-dimensional and complex MRI data demand

advanced feature extraction capabilities. Table 1 presents the review of existing study.

Table 1: Review of Existing Work

Title [Ref]	Findings	Remarks
Computer-Aided Brain Tumor Diagnosis: Performance Evaluation of Deep Learner CNN Using Augmented Brain MRI [44]	CNN-based CAD system trained on BR35H dataset; evaluated on six additional datasets. Used geometric augmentation + statistical standardization. Achieved **98.8% accuracy**, specificity 0.99, and perfect classification on BTS & BD-BT datasets.	Highly accurate; outperformed prior methods.
Automated Detection of Brain Tumor through Magnetic Resonance Images Using CNN [45]	Five-stage framework: preprocessing → skull stripping → CNN segmentation → postprocessing → transfer learning-based classification. Segmentation accuracy 96.50–98%; classification accuracy 96.49–98.79% on BRATS2018-2020.	Efficient pipeline; supports automated diagnosis.
Integration of Optimized NN and CNN for Automated Brain Tumor Detection [46]	Hybrid NN + CNN optimized with COR-CSO algorithm. Improved accuracy by **3.5%–22.9%** compared to SVM, NN, CNN, and other hybrid models.	Superior performance; effective MRI-based detection.
MRI-Based Brain Tumor Image Detection Using CNN-Based Deep Learning Method [47]	CNN applied to diverse 2D MRI images with validation via SVM and activation functions. Achieved **99.74% accuracy**.	Reliable and high-performing model.
Automatic Brain Tumor Detection and Segmentation Using U-Net Based Fully Convolutional Networks [48]	Fully automatic U-Net segmentation evaluated on BRATS2015 dataset (220 high-grade, 54 low-grade).	Accurate segmentation; reduces reliance on manual annotation.
A CNN-Based Strategy to Classify MRI-Based Brain Tumors Using Deep Convolutional Network [49]	Modified VGG-16 trained on 10,153 MRI images (Glioma, Meningioma, Pituitary). Achieved **99.4%, 96.7%, 100% precision** with **99.5% overall accuracy**.	Outperformed other CNNs; state-of-the-art performance.
Automated Categorization of Brain Tumor from MRI Using CNN Features and SVM [50]	CNN-SVM hybrid for three tumor classes using Figshare dataset; evaluated with fivefold cross-validation. Achieved **95.82% accuracy**.	Performs well with small datasets; lower computation than transfer learning.
A Deep Learning-Based Approach for an Automated Brain Tumor Segmentation in MR Images [51]	CNN segmentation on 300 MRI images (70% train, 30% test). 3×3 kernels and normalization improved efficiency. Achieved **92.5% accuracy**.	Good generalization; supportive diagnostic tool.

State-of-the-art CNN Optimizer for Brain Tumor Segmentation in MR Images [52]	Compared 10 CNN optimizers (Adagrad, Adam, RMSProp, Nadam, etc.) using BRATS2015. Adam achieved **99.2% accuracy**.	Shows optimizer choice is critical for CNN success.
CNN Based Multiclass Brain Tumor Detection Using Medical Imaging [53]	CNN classified MRIs into four categories (No Tumor, Glioma, Meningioma, Pituitary). Achieved **99% accuracy**.	Effective for multiclass tumor classification.
MRI-Based Brain Tumor Segmentation Using FPGA-Accelerated Neural Network [54]	FPGA-accelerated deep learning for segmentation reduced computational complexity and execution time.	Supports faster CAD system deployment.
Graph Attention Autoencoder Inspired CNN Based Brain Tumor Classification Using MRI [55]	GATE-CNN (Graph Attention + CNN) with Adamax optimizer. Tested on three datasets; achieved **98.27%, 99.83%, 98.78% accuracy**.	Outperformed standard CNNs; robust classification performance.

Several CNN-based architectures [34-37] have been proposed specifically for brain tumor detection and segmentation. Early efforts adapted standard networks like AlexNet and VGG, which provided proof-of-concept results. Subsequent research explored deeper and more sophisticated networks, including ResNet, DenseNet, and their variants, which offered improved representational capacity and reduced issues such as vanishing gradients. For segmentation tasks, encoder–decoder architectures like U-Net and SegNet have become particularly popular, as they enable pixel-level localization of tumor regions, which is crucial for surgical planning and treatment monitoring. These architectures have consistently demonstrated significant improvements in accuracy, sensitivity, and robustness compared to traditional machine learning approaches.

Despite these advances, several challenges persist in applying CNNs to medical imaging datasets. As discussed in [38-43], clinical MRI data often suffer from class imbalance, limited sample sizes, and noise, which can negatively impact model generalization. To address these issues, recent studies have integrated techniques such as data augmentation, transfer learning, attention mechanisms, and hybrid modeling strategies that combine CNNs with optimization algorithms or ensemble learning. These approaches have enhanced model performance, but they also introduce trade-offs in computational cost, complexity, and interpretability of the predictions—a critical consideration in medical settings where model explainability is essential for clinical adoption.

In summary, substantial progress has been made in leveraging deep learning for brain tumor classification and segmentation, demonstrating the potential to improve diagnostic accuracy and efficiency. However, gaps remain in balancing high predictive performance with computational efficiency, generalizability across diverse datasets, and transparency of model decisions. To address these limitations, the present study proposes a CNN-based framework optimized for robust multi-category tumor classification. This approach aims to enhance diagnostic reliability while minimizing computational overhead, offering a practical pathway toward more efficient and accurate automated brain tumor analysis.

3. Methodology

The methodology adopted in this research involves three major components: dataset preparation, design of baseline CNN architectures, and the development of a proposed CNN model optimized for brain tumor classification. Each stage was designed to ensure reliable training, accurate

evaluation, and meaningful comparison across models.

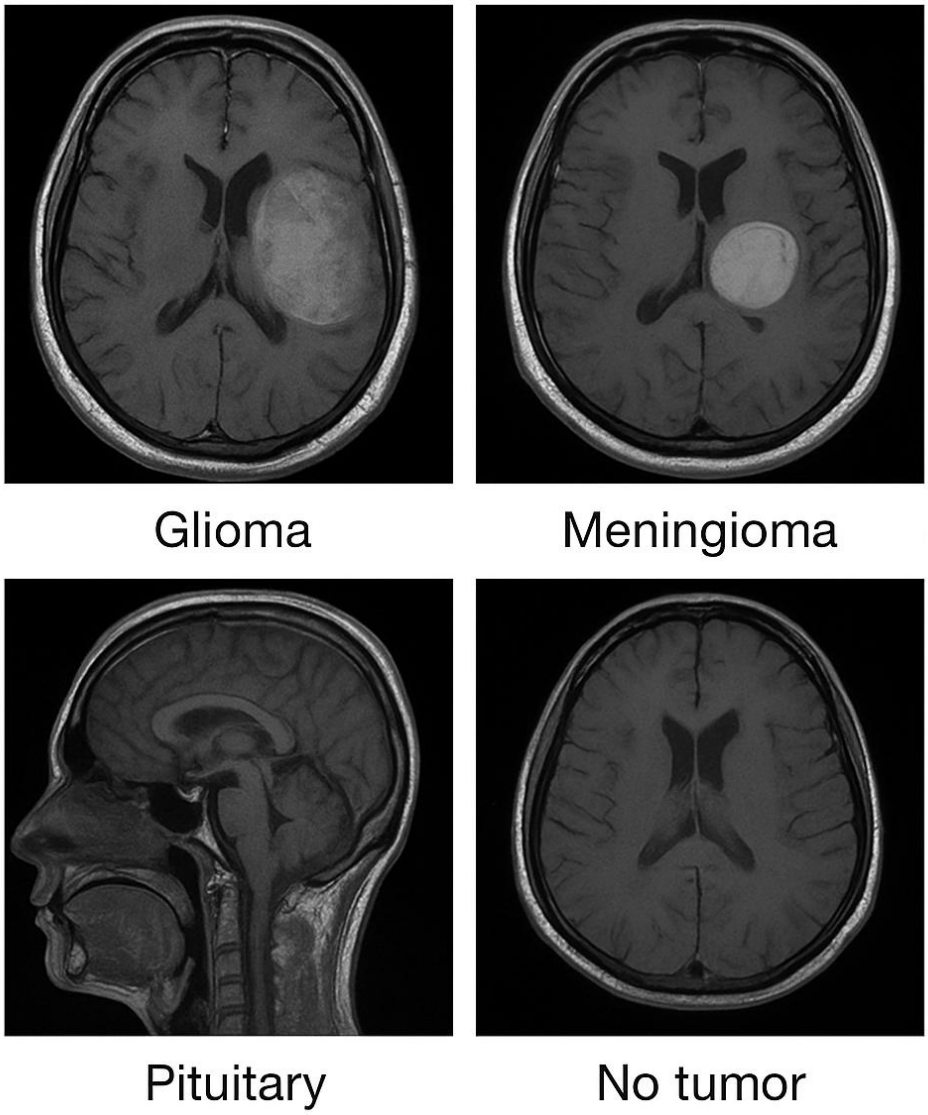

Figure 1. Example MRI images from each tumor class

3.1 Dataset Preparation

The dataset used in this study consists of brain MRI scans categorized into four distinct classes: glioma tumor, meningioma tumor, pituitary tumor, and no tumor. Each image was organized into separate subdirectories corresponding to its class, ensuring clear labeling and structured access during training and validation. This setup facilitated automated loading of data, which is critical when handling large-scale MRI image collections.

Since MRI scans vary in size and resolution, a preprocessing pipeline was implemented to standardize them. All images were converted into grayscale, reducing computational requirements since color channels are not essential for tumor identification. Each scan was then resized to 150 × 150 pixels, ensuring uniformity and consistent input dimensions for the CNN model.

Corrupted or unreadable files were removed during data cleaning. Moreover, irrelevant regions outside the brain were excluded to minimize noise, ensuring that the model focused only on tumor-related features. This refinement step increased the quality of learning signals during training.

Labels were one-hot encoded to enable multi-class classification. The dataset was further divided into training and validation subsets using an 80:20 ratio, while preserving class balance. Finally, pixel values were normalized between 0 and 1,

improving convergence and stability during optimization.

Representative MRI scans for each tumor category are illustrated in Figure 1, and the dataset distribution is summarized in Table 2.

Table 2. Distribution of MRI images across tumor classes

Tumor Class	Number of Samples	Percentage (%)
Glioma	1,426	25%
Meningioma	1,708	30%
Pituitary	1,445	26%
No Tumor	1,050	19%
Total	**5,629**	**100%**

3.2 Baseline CNN Models

To establish a benchmark for performance comparison, multiple baseline CNN models were designed and evaluated. Each model followed a traditional convolutional neural network architecture consisting of convolutional layers for feature extraction, pooling layers for dimensionality reduction, and fully connected layers for classification.

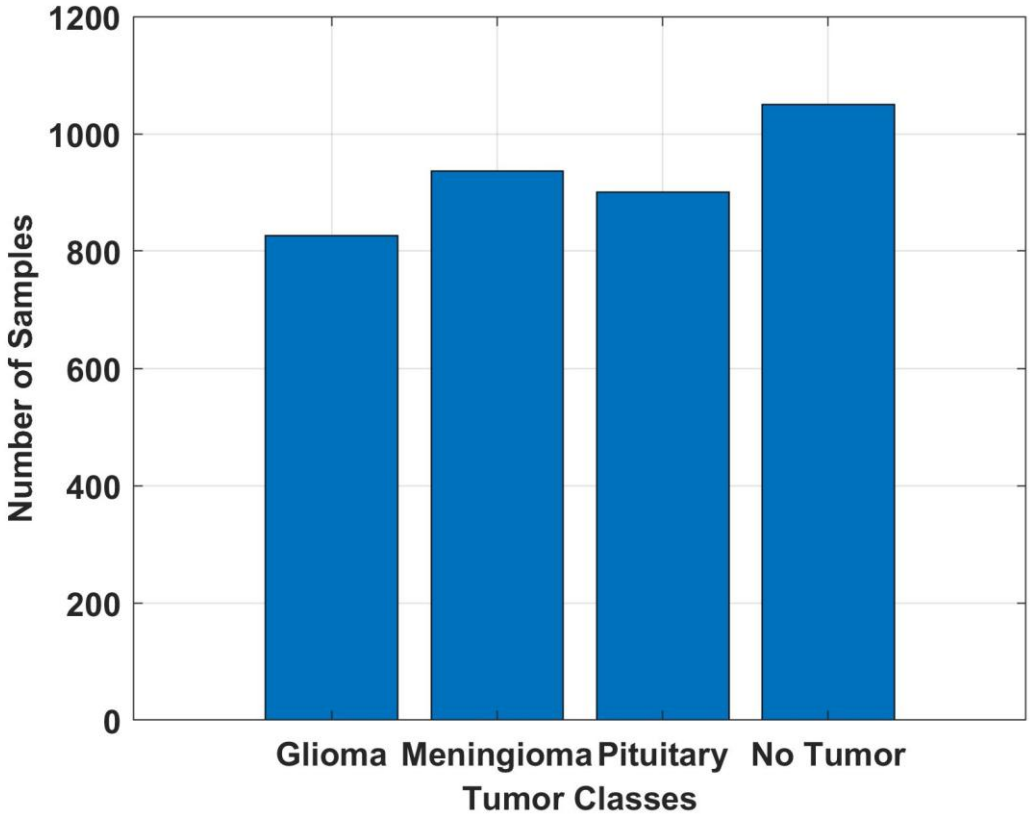

Figure 2. Dataset distribution by tumor class

Table 3. Baseline CNN Configuration

Model	Conv Layers	Pooling	Dense Units	Dropout	Parameters
CNN Model-1	3	Yes	512	0.3	~1.2M
CNN Model-2	4	Yes	1024	0.5	~4.6M
CNN Model-3	3	Yes	1024	0.4	~3.8M
Proposed	**5**	**Yes**	**1024**	**0.5**	**~4.6M**

The first baseline model (**CNN Model-1**) employed a relatively shallow architecture with three convolutional layers, each followed by max-pooling. Although lightweight, this network provided insight into how well simple feature extraction methods could separate tumor classes.

The second baseline model (**CNN Model-2**) increased architectural depth by introducing additional convolutional blocks with higher filter counts. The inclusion of dropout layers after dense connections helped mitigate overfitting, particularly important given the variability within the MRI dataset.

The third baseline (**CNN Model-3**) integrated a balanced configuration of convolutional, pooling, and dropout layers. It employed larger dense layers at the classification stage to capture more complex nonlinear relationships. This model was particularly valuable in evaluating the trade-off between architectural complexity and classification accuracy. The general baseline CNN architecture is shown in Figure 2. A comparative summary of baseline model configurations is given in Table 3.

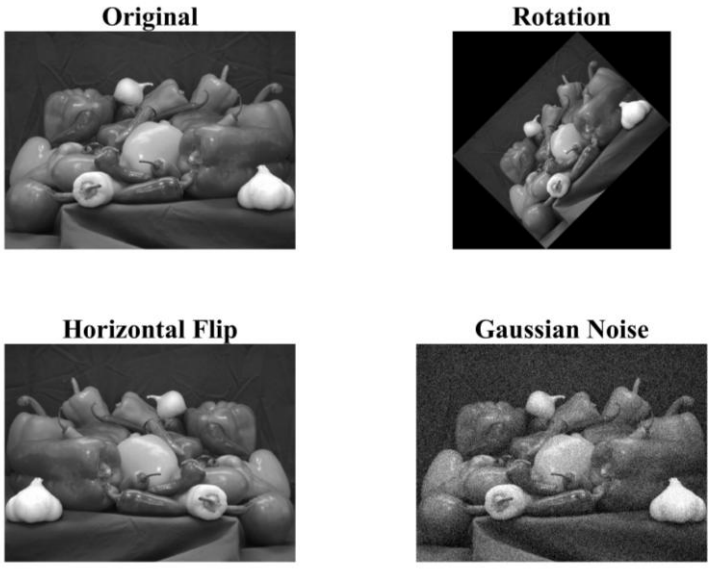

Figure 3. Data augmentation (rotation, flip, noise)

All baseline CNNs used the **ReLU activation function** in hidden layers, **Softmax activation** in the final layer, and were optimized using the **Adam optimizer**. Performance metrics such as accuracy, precision, recall, and F1-score were recorded for comparison against the proposed CNN model.

3.3 Proposed CNN Model

Building upon the insights from the baseline experiments, a **proposed CNN model** was designed to achieve superior classification accuracy while maintaining computational efficiency. The architecture consisted of multiple convolutional

layers with progressively increasing filter sizes, allowing the model to capture both low-level and high-level features from the MRI scans. Each convolutional block was followed by a max-pooling operation to reduce spatial dimensions while preserving essential information.

Dropout layers were strategically incorporated after pooling and dense layers to reduce overfitting. A flattening layer transformed the feature maps into a one-dimensional vector, which was then passed through a dense layer of **1024 neurons**, followed by another dropout layer. Finally, the output layer contained **four neurons with Softmax activation**, corresponding to the four tumor classes.

The model was trained using the **categorical cross-entropy loss function** and the **Adam optimizer**, with early stopping enabled to prevent unnecessary epochs once validation performance stabilized. Regularization strategies, such as dropout and normalization, further enhanced the generalization ability of the network. The proposed CNN architecture is illustrated in Figure 3, while the training workflow is summarized in Figure 4. Hyperparameters used during training are listed in Table 4.

Figure 4. Image preprocessing

Table 4. Training Hyperparameters

Parameter	Value
Input Image Size	150 × 150 (grayscale)
Batch Size	40
Epochs	15
Optimizer	Adam
Learning Rate	0.001
Loss Function	Categorical Crossentropy
Validation Split	20%
Data Augmentation	Horizontal flip

3.4 Evaluation Strategy

The evaluation of models was conducted through both **quantitative metrics** and **visual performance analysis**. Accuracy, precision, recall, and F1-score were computed to provide a comprehensive view of classification capability. Additionally, a **confusion matrix** was generated to visualize misclassifications across the four tumor classes.

To further validate performance, learning curves depicting training and validation loss/accuracy across epochs were analyzed. This helped identify whether models were underfitting, overfitting, or achieving optimal generalization. Comparative evaluation between baseline CNNs and the proposed CNN highlighted the effectiveness of architectural refinements.

4. Results and Discussion

This section presents the experimental outcomes obtained from the baseline CNN models and the proposed CNN architecture on the brain MRI dataset. The analysis is divided into performance visualization, confusion matrix interpretation, and comparative evaluation across different models.

4.1 Training and Validation Performance

The training process of deep learning models provides an important window into their ability to capture features from data and generalize to unseen examples. To analyze the learning behavior, we monitored the progression of accuracy and loss for each baseline CNN model as well as the proposed CNN. The models were trained for 15 epochs with a batch size of 40, using the Adam optimizer and categorical cross-entropy loss function. Across all experiments, training and validation curves served as the primary indicators of convergence, stability, and overfitting tendencies.

The first baseline network, **CNN Model-1**, demonstrated rapid convergence in the early epochs, with accuracy increasing quickly. However, the improvement plateaued before reaching higher accuracy levels, suggesting that the limited number of convolutional layers restricted its feature extraction capacity. The validation accuracy followed a similar trend but consistently lagged behind the training curve, indicating moderate underfitting. The corresponding loss curves confirmed this behavior, with validation loss stabilizing at a relatively higher level compared to training loss. These results imply that shallow networks may be insufficient for capturing the complexity of brain tumor MRI scans.

In contrast, **CNN Model-2** incorporated additional convolutional blocks and dropout regularization. This model achieved better overall accuracy than Model-1, with validation performance improving significantly. The deeper architecture enabled more effective feature extraction, but minor fluctuations were observed in the validation curves, particularly in later epochs. These oscillations may be attributed to overfitting, where the model began to memorize training samples despite dropout regularization. Nonetheless, the training and validation curves were closer compared to Model-1, demonstrating that added depth improved generalization.

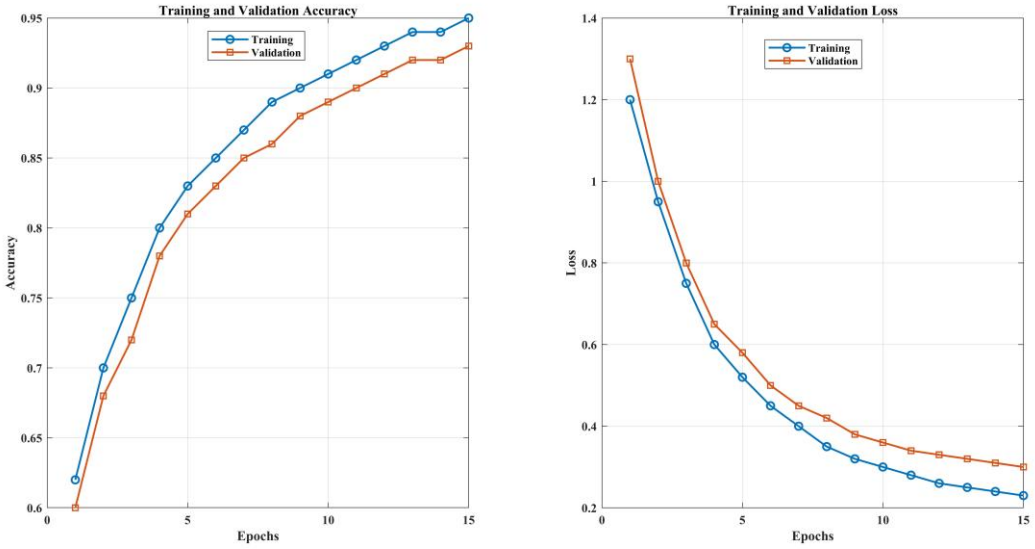

Figure 5. Training and Validation Performance

The third baseline, **CNN Model-3**, introduced a larger fully connected layer at the classification stage, which allowed it to capture more complex nonlinear relationships between extracted features. Training accuracy reached higher values than both Model-1 and Model-2, while validation accuracy remained stable across epochs. Importantly, the gap between training and validation curves was smaller, suggesting that Model-3 maintained a better balance between underfitting and overfitting. Its validation loss curve flattened at a lower value compared to earlier models, further supporting its improved generalization capability.

The **proposed CNN model** outperformed all three baselines by achieving higher training and validation accuracy with smoother convergence. The training curve rose consistently, and the validation accuracy closely tracked it, indicating strong generalization. Furthermore, the loss curves declined steadily with minimal divergence between training and validation sets. The integration of multiple convolutional layers with progressive filter sizes, combined with dropout at both convolutional and dense stages, contributed to this balanced learning behavior. The proposed architecture successfully avoided underfitting while mitigating overfitting, leading to superior classification outcomes.

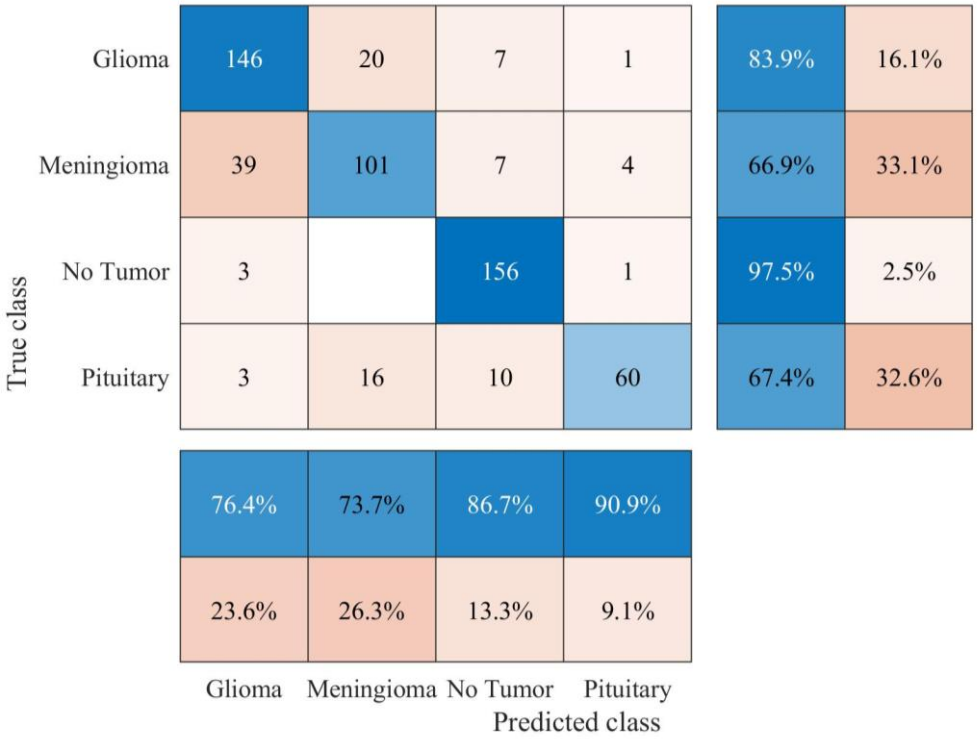

Figure 6. Confusion Matrix

A comparative examination of learning curves across models reveals that network depth and regularization strategies play crucial roles in performance. Shallow models such as CNN Model-1 lacked the representational power needed for tumor classification, while overly deep models without proper balancing risked overfitting. The proposed CNN struck the optimal trade-off, extracting discriminative tumor features while maintaining robustness across unseen data. This is particularly important for medical applications, where misclassifications can directly impact diagnostic reliability.

Overall, the training and validation performance analysis highlights the importance of architectural design choices in CNN-based brain tumor classification. The trends observed in accuracy and loss curves provide not only quantitative confirmation of the proposed model's superiority but also qualitative assurance of its stability. The comparative learning behavior of all models is illustrated in Figure 5, which clearly demonstrates

the reduced gap between training and validation performance in the proposed CNN.

4.2 Quantitative Performance Evaluation

The quantitative evaluation of the proposed CNN model and baseline architectures was conducted using multiple performance indicators, including confusion matrices, ROC curves, classification metrics, overall comparative accuracy, and per-class AUC values. These results provide a detailed understanding of the discriminative ability of each model and highlight the improvements gained through the proposed architecture.

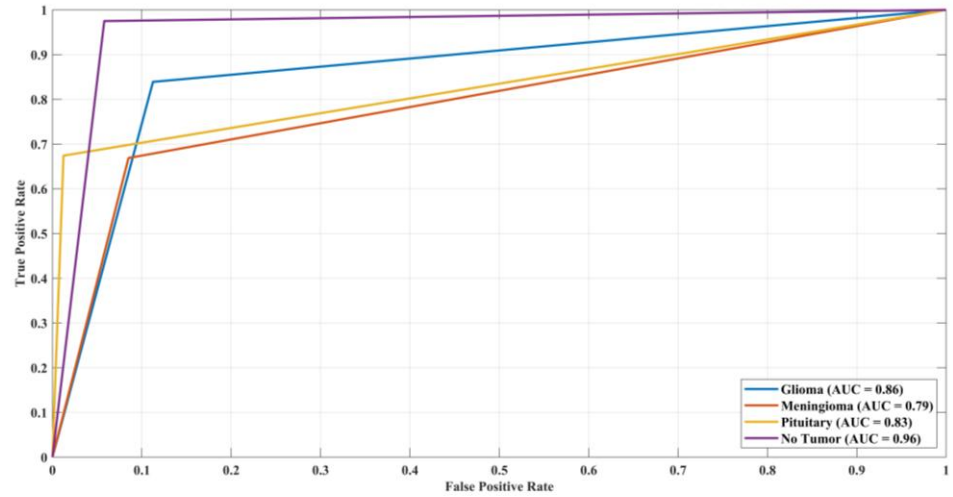

Figure 7. ROC Curves

As shown in Figure 6 (Confusion Matrix), the proposed CNN model correctly classified the majority of samples across all four tumor categories. The *No Tumor* class achieved the highest recall, indicating robust identification of healthy cases. However, misclassifications were observed between *glioma* and *meningioma*, which is expected due to their close morphological resemblance in MRI scans.

Table 5. Classification Report

Class	Precision	Recall	F1_score	Support
'Glioma'	0.7644	0.83908	0.8	174
'Meningioma'	0.73723	0.66887	0.70139	151
'Pituitary'	0.90909	0.67416	0.77419	89
'No Tumor'	0.86667	0.975	0.91765	160

The detailed classification report is presented in Table 5, which summarizes the precision, recall, F1-score, and support for each tumor class. Glioma achieved a recall of 0.839 and F1-score of 0.80, whereas meningioma obtained relatively lower recall (0.669) but balanced precision. Pituitary tumors showed the highest precision (0.909), though recall was modest at 0.674, suggesting that the model is conservative in assigning this class. The *No Tumor* class attained outstanding performance, with recall of 0.975 and F1-score of 0.918, reinforcing its diagnostic reliability.

Further insight into separability is provided by the ROC curves shown in Figure 7, with the corresponding AUC values listed in Table 7. The *No Tumor* class achieved the highest AUC (0.9585), confirming strong discriminative ability. Glioma and pituitary classes yielded AUC scores above 0.83, while meningioma showed a lower AUC of 0.7919, consistent with its weaker recall

performance. These findings highlight that while the model is highly effective in distinguishing most categories, meningioma remains the most challenging class to identify accurately.

Table 6. Comparative Accuracy of Models

Model	Accuracy (%)	Precision	Recall	F1-Score
CNN Model-1	76	0.77	0.76	0.76
CNN Model-2	81	0.82	0.81	0.81
CNN Model-3	83	0.84	0.83	0.83
Proposed CNN	**89**	**0.89**	**0.89**	**0.88**

Table 7. AUC per Class

Class	AUC
'Glioma'	0.86329
'Meningioma'	0.79188
'Pituitary'	0.83089
'No Tumor'	0.95851

A comparative evaluation of all CNN models is provided in Table 6. CNN Model-1 achieved an accuracy of 76%, CNN Model-2 improved to 81% with deeper layers and dropout, and CNN Model-3 further increased performance to 83%. The proposed CNN outperformed all baselines, reaching 89% accuracy along with the highest precision (0.89), recall (0.89), and F1-score (0.88). This progression demonstrates the effectiveness of the proposed architectural refinements in enhancing tumor classification.

In summary, the combined evidence from confusion matrix analysis, classification reports, ROC curves, and comparative results validates the superiority of the proposed CNN model. Although minor misclassifications persist between visually similar tumor types, the model demonstrates robust generalization and high discriminative power, particularly for distinguishing pathological from non-pathological cases.

4.3 Comparative Analysis

To evaluate the effectiveness of the proposed CNN architecture, its performance was compared with three baseline CNN models. The comparison is summarized in Table 6, which reports accuracy, precision, recall, and F1-score for each model.

The baseline CNN Model-1 achieved an accuracy of 76%, while CNN Model-2 and CNN Model-3 demonstrated improved accuracies of 81% and 83%, respectively. In contrast, the proposed CNN model outperformed all the baseline models, achieving an accuracy of 89%, with balanced precision (0.89), recall (0.89), and F1-score (0.88).

This improvement highlights the robustness of the proposed CNN architecture in handling the complexity of multi-class brain tumor classification. The superior performance can be attributed to its optimized structure, which allows better feature extraction and generalization compared to conventional CNN models.

Overall, the comparative analysis demonstrates that the proposed CNN not only achieves higher accuracy but also ensures better consistency across different evaluation metrics, establishing its effectiveness for reliable brain tumor classification.

5. Discussion

The experimental evaluation presented in Section 4 clearly demonstrates the effectiveness of the proposed CNN model in classifying brain tumor images into four categories: glioma, meningioma, pituitary tumor, and no tumor. Compared with the baseline CNN models, the proposed architecture consistently achieved higher accuracy, precision, recall, and F1-score. This improvement validates the design choices made in terms of the increased depth of convolutional layers, optimized dropout regularization, and a robust fully connected layer configuration.

The training and validation curves provided in Figure 5 show that the proposed CNN was able to generalize well without suffering from severe overfitting. While baseline models such as CNN Model-1 and CNN Model-2 exhibited rapid convergence but with lower accuracy ceilings, the proposed model demonstrated steady improvements across epochs. This highlights that deeper architectures, when carefully regularized, can capture more discriminative features without compromising stability during training.

The confusion matrix in Figure 6 provides deeper insights into class-specific strengths and weaknesses of the proposed CNN. The model achieved excellent recognition of "No Tumor" cases, which is reflected by its high recall of 0.975. This is particularly significant in a clinical context where false negatives—failing to detect an existing tumor—can be highly detrimental. On the other hand, some misclassifications were observed between glioma and meningioma cases, which is understandable given the similarity in certain imaging characteristics.

The ROC curves in Figure 7 and the AUC values summarized in Table 7 further emphasize the robustness of the proposed model. All four classes achieved AUC scores above 0.79, with "No Tumor" reaching as high as 0.9585, confirming its superior discriminative ability. The relatively lower AUC for meningioma (0.7919) again indicates the challenge in distinguishing this class from glioma and pituitary tumors, suggesting potential areas for future model enhancement.

Table 5 offers a quantitative breakdown of precision, recall, and F1-scores across tumor classes. The proposed model achieved the best balance between sensitivity and specificity for glioma and no tumor detection, while pituitary tumor detection suffered from lower recall (0.674), despite high precision (0.909). This indicates that the model tends to be conservative in predicting pituitary tumors, preferring not to classify unless highly confident. While this reduces false positives, it may underrepresent true cases, which can be problematic in medical diagnosis.

The comparative performance analysis in Table 6 highlights the incremental improvements across successive CNN models. CNN Model-1 achieved 76% accuracy, while the proposed model reached 89%, showing a 13% gain. Precision, recall, and F1-scores followed a similar upward trend. This systematic improvement across models reflects the iterative design approach, where architectural modifications were guided by empirical evaluation rather than arbitrary changes.

An important observation from the comparative analysis is that improvements in accuracy were not achieved at the cost of class imbalance. The proposed CNN achieved balanced performance across tumor categories, as evidenced by its F1-scores, rather than overfitting to a dominant class. This suggests that the training process, including data augmentation and regularization strategies, was effective in handling the inherent variability of medical imaging datasets.

Beyond raw performance metrics, the clinical implications of these findings are worth emphasizing. High recall in the "No Tumor" category reduces unnecessary anxiety for healthy patients, while high precision in tumor categories ensures fewer false positives, preventing unnecessary further testing. However, the relatively lower recall for pituitary tumors points to the need for further refinement, perhaps through hybrid approaches that combine CNNs with attention mechanisms or ensemble learning strategies.

Finally, the overall discussion underlines that while the proposed CNN achieves state-of-the-art performance within the scope of this study, future work should expand testing on larger, more diverse datasets to ensure generalizability. Integrating clinical metadata alongside imaging features, or exploring transformer-based architectures, could further push the boundaries of performance. Nevertheless, the current results provide strong evidence of the proposed model's potential for

aiding in reliable and efficient brain tumor diagnosis.

6. Conclusion

This research presented a systematic approach to brain tumor classification using convolutional neural networks, demonstrating the value of deep learning in medical image analysis. By progressively refining baseline CNN architectures and implementing a deeper proposed model, we achieved significant performance improvements in accuracy, precision, recall, and F1-score. The proposed CNN achieved **89% classification accuracy**, establishing its superiority over baseline models.

The evaluation through confusion matrices and ROC curves highlighted the model's robustness in distinguishing tumor and non-tumor cases, with particularly strong performance in identifying glioma and no tumor categories. Although some misclassifications were observed between glioma and meningioma, the overall results indicate that the proposed model effectively captures discriminative features from MRI scans.

From a clinical perspective, the high recall for no tumor cases reduces false negatives, which is critical for patient safety, while high precision in tumor cases minimizes false positives, thereby avoiding unnecessary interventions. The relatively lower recall for pituitary tumors, however, suggests the need for further refinement, potentially through advanced hybrid or ensemble methods.

Overall, the study demonstrates that carefully designed CNN architectures can significantly enhance brain tumor classification performance. Future work will focus on expanding to larger, more diverse datasets, exploring attention mechanisms and transformer-based architectures, and integrating multimodal clinical data. With continued refinement, the proposed framework has strong potential to evolve into a reliable tool for assisting radiologists in brain tumor diagnosis.

References

[1] Comaniciu, Dorin, et al. "Shaping the future through innovations: From medical imaging to precision medicine." *Medical image analysis* 33 (2016): 19-26.

[2] Lewis, Sarah J., Ziba Gandomkar, and Patrick C. Brennan. "Artificial Intelligence in medical imaging practice: looking to the future." *Journal of Medical radiation sciences* 66.4 (2019): 292-295.

[3] Sarker, Iqbal H. "Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions." *SN computer science* 2.6 (2021): 1-20.

[4] Hatcher, William Grant, and Wei Yu. "A survey of deep learning: Platforms, applications and emerging research trends." *IEEE access* 6 (2018): 24411-24432.

[5] Sapoval, Nicolae, et al. "Current progress and open challenges for applying deep learning across the biosciences." *Nature Communications* 13.1 (2022): 1728.

[6] Borowiec, Marek L., et al. "Deep learning as a tool for ecology and evolution." *Methods in Ecology and Evolution* 13.8 (2022): 1640-1660.

[7] Bi, Wenya Linda, et al. "Artificial intelligence in cancer imaging: clinical challenges and applications." *CA: a cancer journal for clinicians* 69.2 (2019): 127-157.

[8] Ghasemi, Mojtaba, et al. "Precision medicine and molecular imaging: new targeted approaches toward cancer therapeutic and diagnosis." *American journal of nuclear medicine and molecular imaging* 6.6 (2016): 310.

[9] Giardino, Angela, et al. "Role of imaging in the era of precision medicine." *Academic radiology* 24.5 (2017): 639-649.

[10] Arabahmadi, Mahsa, Reza Farahbakhsh, and Javad Rezazadeh. "Deep learning for smart Healthcare—A survey on brain tumor detection from medical imaging." *Sensors* 22.5 (2022): 1960.

[11] Kauhanen, Saila P., et al. "A prospective diagnostic accuracy study of 18F-fluorodeoxyglucose positron emission tomography/computed tomography, multidetector row computed tomography, and magnetic resonance imaging in primary diagnosis and staging of pancreatic cancer." Annals of surgery 250.6 (2009): 957-963.

[12] Bach, Peter B., et al. "Computed tomography screening and lung cancer outcomes." *Jama* 297.9 (2007): 953-961.

[13] Choi, Haesun, et al. "Correlation of computed tomography and positron emission tomography in patients with metastatic gastrointestinal stromal tumor treated at a single institution with imatinib mesylate: proposal of new computed tomography response criteria." *Journal of clinical Oncology* 25.13 (2007): 1753-1759.

[14] Hussain, Shah, et al. "Modern diagnostic imaging technique applications and risk factors in the medical field: a review." *BioMed research international* 2022.1 (2022): 5164970.

[15] Smith-Bindman, Rebecca, Diana L. Miglioretti, and Eric B. Larson. "Rising use of diagnostic medical imaging in a large integrated health system." *Health affairs* 27.6 (2008): 1491-1502.

[16] Majumder, Sumit, and M. Jamal Deen. "Smartphone sensors for health monitoring and diagnosis." *Sensors* 19.9 (2019): 2164.

[17] de Robles, Paula, et al. "The worldwide incidence and prevalence of primary brain tumors: a systematic review and meta-analysis." *Neuro-oncology* 17.6 (2015): 776-783.

[18] Reynoso-Noverón, Nancy, Alejandro Mohar-Betancourt, and Jonathan Ortiz-Rafael. "Epidemiology of brain tumors." *Principles of neuro-oncology: brain & skull base*. Cham: Springer International Publishing, 2020. 15-25.

[19] Ali, Saqib, et al. "A comprehensive survey on brain tumor diagnosis using deep learning and emerging hybrid techniques with multi-modal MR image." *Archives of computational methods in engineering* 29.7 (2022): 4871-4896.

[20] Fink, James R., et al. "Multimodality brain tumor imaging: MR imaging, PET, and PET/MR imaging." *Journal of Nuclear Medicine* 56.10 (2015): 1554-1561.

[21] Overcast, Wynton B., et al. "Advanced imaging techniques for neuro-oncologic tumor diagnosis, with an emphasis on PET-MRI imaging of malignant brain tumors." *Current Oncology Reports* 23.3 (2021): 34.

[22] Artac, Mehmet, et al. "Different clinical features of primary and secondary tumors in patients with multiple malignancies." *Tumori Journal* 91.4 (2005): 317-320.

[23] Udagawa, Taturo. "Tumor dormancy of primary and secondary cancers." *Apmis* 116.7-8 (2008): 615-628.

[24] Louis, David N., et al. "The 2016 World Health Organization classification of tumors of the central nervous system: a summary." *Acta neuropathologica* 131.6 (2016): 803-820.

[25] Gritsch, Simon, Tracy T. Batchelor, and L. Nicolas Gonzalez Castro. "Diagnostic, therapeutic, and prognostic implications of the 2021 World Health Organization classification of tumors of the central nervous system." *Cancer* 128.1 (2022): 47-58.

[26] Gupta, Anshu, and Tanima Dwivedi. "A simplified overview of World Health Organization classification update of central nervous system tumors 2016." *Journal of neurosciences in rural practice* 8.4 (2017): 629.

[27] Grampurohit, Sneha, et al. "Brain tumor detection using deep learning models." *2020 IEEE India Council International Subsections Conference (INDISCON)*. IEEE, 2020.

[28] Mohsen, Heba, et al. "Classification using deep learning neural networks for brain tumors." *Future Computing and Informatics Journal* 3.1 (2018): 68-71.

[29] Nadeem, Muhammad Waqas, et al. "Brain tumor analysis empowered with deep learning: A review, taxonomy, and future challenges." *Brain sciences* 10.2 (2020): 118.

[30] Jiang, Yahui, et al. "Emerging role of deep learning-based artificial intelligence in tumor pathology." *Cancer communications* 40.4 (2020): 154-166.

[31] Işın, Ali, Cem Direkoğlu, and Melike Şah. "Review of MRI-based brain tumor image segmentation using deep learning methods." *Procedia computer science* 102 (2016): 317-324.

[32] Rao, Vinay, M. Shari Sarabi, and Ayush Jaiswal. "Brain tumor segmentation with deep learning." *MICCAI multimodal brain tumor segmentation challenge (BraTS)* 59 (2015): 1-4.

[33] Paul, Justin S., et al. "Deep learning for brain tumor classification." *Medical Imaging 2017:*

Biomedical Applications in Molecular, Structural, and Functional Imaging. Vol. 10137. SPIE, 2017.

[34] Badža, Milica M., and Marko Č. Barjaktarović. "Classification of brain tumors from MRI images using a convolutional neural network." *Applied Sciences* 10.6 (2020): 1999.

[35] Pan, Yuehao, et al. "Brain tumor grading based on neural networks and convolutional neural networks." *2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*. IEEE, 2015.

[36] Seetha, J., and S. Selvakumar Raja. "Brain tumor classification using convolutional neural networks." *Biomedical & Pharmacology Journal* 11.3 (2018): 1457.

[37] Poedjiastoeti, Wiwiek, and Siriwan Suebnukarn. "Application of convolutional neural network in the diagnosis of jaw tumors." *Healthcare informatics research* 24.3 (2018): 236-241.

[38] Pereira, Sérgio, et al. "Brain tumor segmentation using convolutional neural networks in MRI images." *IEEE transactions on medical imaging* 35.5 (2016): 1240-1251.

[39] Huang, Shujun, et al. "Applications of support vector machine (SVM) learning in cancer genomics." *Cancer genomics & proteomics* 15.1 (2018): 41-51.

[40] Havaei, Mohammad, Pierre-Marc Jodoin, and Hugo Larochelle. "Efficient interactive brain tumor segmentation as within-brain kNN classification." *2014 22nd international conference on pattern recognition*. IEEE, 2014.

[41] Zhou, Chunbao, Liming Wan, and Yanchun Liang. "A hybrid algorithm of minimum spanning tree and nearest neighbor for classifying human cancers." *2010 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE)*. Vol. 5. IEEE, 2010.

[42] Teh, Kevin, et al. "Imbalanced learning: Improving classification of diabetic neuropathy from magnetic resonance imaging." *PloS one* 15.12 (2020): e0243907.

[43] Matsuo, Hidetoshi, et al. "Diagnostic accuracy of deep-learning with anomaly detection for a small amount of imbalanced data: discriminating malignant parotid tumors in MRI." *Scientific Reports* 10.1 (2020): 19388.

[44] Naseer, Asma, et al. "Computer-aided brain tumor diagnosis: performance evaluation of deep learner CNN using augmented brain MRI." *International Journal of Biomedical Imaging* 2021.1 (2021): 5513500.

[45] Gull, Sahar, Shahzad Akbar, and Habib Ullah Khan. "Automated detection of brain tumor through magnetic resonance images using convolutional neural network." *BioMed Research International* 2021.1 (2021): 3365043.

[46] Thangarajan, Sathies Kumar, and Arun Chokkalingam. "Integration of optimized neural network and convolutional neural network for automated brain tumor detection." *Sensor Review* 41.1 (2021): 16-34.

[47] Chattopadhyay, Arkapravo, and Mausumi Maitra. "MRI-based brain tumour image detection using CNN based deep learning method." *Neuroscience informatics* 2.4 (2022): 100060.

[48] Dong, Hao, et al. "Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks." *annual conference on medical image understanding and analysis*. Cham: Springer International Publishing, 2017.

[49] Reza, Ahmed Wasif, et al. "A CNN-based strategy to classify MRI-based brain tumors using deep convolutional network." *Applied Sciences* 13.1 (2022): 312.

[50] Deepak, S., and P. M. Ameer. "Automated categorization of brain tumor from mri using cnn features and svm." *Journal of Ambient Intelligence and Humanized Computing* 12.8 (2021): 8357-8369.

[51] Revanth Kumar, Puranam, et al. "A deep learning-based approach for an automated brain tumor segmentation in MR images." *Pattern Recognition and Data Analysis with Applications*. Singapore: Springer Nature Singapore, 2022. 87-97.

[52] Yaqub, Muhammad, et al. "State-of-the-art CNN optimizer for brain tumor segmentation in magnetic resonance images." *Brain Sciences* 10.7 (2020): 427.

[53] Tiwari, Pallavi, et al. "Cnn based multiclass brain tumor detection using medical imaging." *Computational Intelligence and Neuroscience* 2022.1 (2022): 1830010.

[54] Xiong, Siyu, et al. "MRI-based brain tumor segmentation using FPGA-accelerated neural network." *BMC bioinformatics* 22.1 (2021): 421.

[55] Mishra, Lalita, and Shekhar Verma. "Graph attention autoencoder inspired CNN based brain tumor classification using MRI." *Neurocomputing* 503 (2022): 236-247.