Automatic brain tumor detection in MRI images using convolutional neural networks

Ali Mohammad Aber

Research Scholar, Dept. of Computer Science, Punjabi University, Patiala, India

Ali_rs23@pbi.ac.in

Dr. Maninder Singh

Professor, Dept. of Computer Science, Punjabi University, Patiala, India

dr.manindersingh.10@gmail.com

Abstract

Brain tumors are regarded one of the most aggressive disorders in both children and adults. Brain tumors make for 85-90 percent of all primary central nervous system (CNS) malignancies. Every year, around 11700 people are diagnosed with brain tumors. The 5-year survival rate for patients with a malignant brain or CNS tumor is around 34% for males and 36% for women. Brain tumors are divided as benign, malignant, and pituitary tumors. Proper therapy, planning, and precise diagnostics should be implemented to increase patients' life expectancy. Magnetic resonance imaging (MRI) is the most effective tool for detecting brain malignancies. The scans generate a vast amount of picture data. The radiologist examines these photographs. Because of the intricacies and features of brain tumors, a manual examination can be prone to errors. Automated classification techniques based on Machine Learning (ML) and deep learning have consistently outperformed manual classification. The dataset we have used for this work contains 2 folders: yes and no which contains 3060 brain MRI images. After implementation our custom CNN model achieved 98% accuracy. As a result, offering a system for detection and classification that uses deep learning algorithms such as convolution-neural network (CNN) will be beneficial to doctors all over the world.

Keywords: Machine Learning, Deep Learning, Artificial Intelligence, Computer-Aided Diagnosis, Healthcare Analytics.

1. Introduction

Brain tumors represent one of the most aggressive and debilitating forms of neurological disease, affecting both children and adults worldwide. According to epidemiological estimates, approximately 11,700 individuals are diagnosed with primary brain tumors annually, and the five-year survival rate for those with malignant central nervous system (CNS) tumors is approximately 34% for men and 36% for women. Brain tumors are broadly categorized as benign or malignant, with subtypes including glioma, meningioma, and pituitary tumors. Among these, malignant gliomas carry a particularly poor prognosis, underscoring the urgent need for early, accurate, and scalable detection methodologies.

Magnetic Resonance Imaging (MRI) is the gold standard imaging modality for brain tumor diagnosis, offering superior soft-tissue contrast and spatial resolution without ionizing radiation. However, the manual analysis of MRI scans by radiologists is labor-intensive, time-consuming, and prone to inter-observer variability, particularly in low-resource healthcare settings where the scarcity of trained neuroradiologists presents a persistent challenge [2]. Automated computer-aided diagnosis (CAD) systems powered by machine learning and deep learning offer a promising avenue to standardize and accelerate diagnostic workflows. Among deep learning approaches, Convolutional Neural Networks (CNNs) have emerged as the dominant paradigm for medical image analysis tasks, owing to their ability to hierarchically learn spatial feature representations directly from raw pixel data without the need for handcrafted feature engineering [3]. Prior work has demonstrated CNN-based models achieving diagnostic accuracies exceeding 95% on standard brain tumor MRI datasets, though many high-performing approaches rely on complex architectures, transfer learning with large pre-trained networks, or extensive data augmentation pipelines that may not be easily reproducible in practical settings [4,5]. In this work, we present a streamlined custom CNN architecture designed for binary detection of brain tumors from MRI scans. The model is intentionally kept shallow and interpretable, comprising three convolutional blocks and two dense layers, trained with dropout regularization and a custom early-stopping callback triggered at 99% training accuracy. Despite its architectural simplicity, the proposed model achieves 98.53% test accuracy on a balanced dataset of 3,000 MRI images, demonstrating that carefully designed modest architectures can achieve clinically relevant performance levels. The key contributions of this paper are as follows: Recent advances in artificial intelligence have significantly improved the ability of automated systems to analyse complex datasets and support intelligent decision-making processes across different domains. AI-driven frameworks have increasingly demonstrated their effectiveness in handling large-scale data processing, pattern recognition, and predictive analytics, creating new opportunities for healthcare and medical image analysis applications [6,7]. The growing adoption of artificial intelligence technologies has accelerated innovation in data-driven systems, particularly those requiring advanced feature extraction and intelligent classification mechanisms. Neural-network-based approaches have become one of the most widely adopted AI techniques due to their capability to learn hierarchical representations from complex datasets [6,8].

- Proposal and evaluation of a compact three-block CNN for binary brain tumor detection from MRI images.
- Implementation of a custom TensorFlow callback for adaptive training termination at 99% accuracy.
- Comprehensive performance evaluation using accuracy, precision, recall, F1-score, specificity, and confusion matrix analysis.
- Contextualisation of results with respect to recent state-of-the-art methods from 2023 to 2026.

2. Related Work

In this section we have reviewed recent studies using CNNs- based approach for brain tumor detection and classification in MRI images. Saeedi et al. [9] conducted one of the foundational recent studies on MRI-based brain tumor detection using convolutional deep learning methods. Published in BMC Medical Informatics and Decision Making, the study systematically compared CNN architectures with classical machine learning classifiers, establishing that deep convolutional models consistently outperformed traditional methods. Their findings emphasized the role of preprocessing and data quality in achieving high diagnostic accuracy, and reported competitive results on binary classification tasks. This work established a baseline against which many subsequent studies are measured. Naeem et al. [3] addressed a critical practical limitation of many deep learning approaches: the requirement for large annotated training datasets. Their study, published in Frontiers in Medicine in 2025, proposed a lightweight CNN specifically designed for accurate brain tumor detection under conditions of limited data availability. The model was engineered for low computational overhead while maintaining high sensitivity and specificity, achieving clinically relevant performance on binary tumor classification tasks from MRI scans. Their architecture demonstrates that model complexity is not a prerequisite for high accuracy when regularization and architectural choices are thoughtfully applied. Wong et al. [10] presented a deep learning system for brain tumor classification using MRI images, published in PLOS ONE in 2025. The study employed data augmentation techniques to expand their training corpus to over 17,000 images across four classification categories. Their best model achieved 99.24% accuracy, one of the highest reported values for multi-class brain tumor classification. The results highlight the value of augmentation-driven dataset expansion in training robust deep learning classifiers, particularly when the original dataset size is limited. Mathivanan et al. [11] conducted a comprehensive study comparing ResNet152, VGG19, DenseNet169, and MobileNetV3 across three brain tumor MRI datasets. Published in Scientific Reports in 2024, the study found that ResNet152 achieved the best overall performance at 98.5% accuracy, outperforming all other evaluated architectures and demonstrating the superiority of residual learning for this task. Albalawi et al. [12] proposed an integrated federated learning and transfer learning framework for privacy-preserving brain tumor classification using VGG16. Published in BMC Medical Imaging in 2024, their approach decoupled model training from centralized data aggregation, addressing critical data privacy concerns in clinical settings. Their work demonstrated that federated transfer learning can maintain classification accuracy comparable to centralised approaches while preserving patient data confidentiality, representing an important step towards clinically deployable AI diagnostics. A systematic comparison of pre-trained models including Xception, MobileNetV2, InceptionV3, ResNet50, VGG16, and DenseNet121 was performed by Stathopoulos et al. [13] across multiple MRI modalities, published in the Journal of Imaging in 2024. The study

demonstrated that Xception consistently achieved the highest weighted accuracy of 98.73%, though performance varied substantially across MRI acquisition protocols, underscoring the importance of modality-specific model evaluation. Their work further advocates for cross-dataset validation as a standard evaluation protocol. Hassan et al. [14] evaluated fine-tuned transfer learning models for brain tumor classification from MRI, comparing VGG-16, ResNet-50, and Inception-V3 architectures. Published in Digital Health in 2024, their work found that VGG-16 achieved the highest accuracy at 96%, while highlighting that Inception-V3 required more careful hyperparameter tuning. The study also introduced a preprocessing pipeline using Canny Edge Detection for image cropping that improved classification consistency across model architectures. Nahiduzzaman et al. [15] proposed a hybrid explainable model combining advanced machine learning and deep learning for brain tumor classification, published in Scientific Reports in 2025. Their architecture integrated convolutional feature extraction with classical classifiers and incorporated explainability through gradient-weighted class activation mapping (Grad-CAM). The study achieved strong classification metrics while providing interpretable visual explanations of model decisions, addressing a critical barrier to clinical adoption of black-box AI diagnostic tools. Rasool et al. [16] presented CNN-TumorNet, an explainability-enhanced deep learning model for brain tumor diagnosis published in Frontiers in Oncology in 2025. Their model leveraged saliency maps and SHAP values to provide radiologist-interpretable explanations for classification outcomes. The study demonstrated that combining high-accuracy CNNs with robust explainability frameworks significantly improves clinician trust and adoption, a finding increasingly recognized as important for real-world deployment of AI diagnostic tools.

3. Materials and Methods

3.1 Dataset

The dataset used in this study is the Brain MRI Images for Brain Tumor Detection dataset, publicly available on Kaggle (contributed by seifwael123). The dataset comprises 3,000 brain MRI images organized into two balanced classes: 1,500 tumor-positive images (labeled "yes") and 1,500 tumor-negative images (labeled "no"). The dataset is drawn from primary brain MRI scans captured under clinical conditions and represents a diverse range of tumor morphologies, sizes, and locations within the tumor-positive class. The balanced class distribution eliminates the need for class weighting or oversampling, simplifying the training pipeline while ensuring unbiased classifier evaluation.

3.2 Preprocessing and Data Preparation

All images were loaded using OpenCV (cv2) in BGR format and subsequently converted to RGB to maintain colour channel consistency with standard visualization conventions. Images were uniformly resized to 128 x 128 pixels to provide a standardized input

resolution that balances spatial detail preservation with computational tractability for the proposed CNN architecture. Pixel intensity values were normalized from the native range of [0, 255] to the floating-point range of [0, 1] by dividing by 255.0, ensuring that all input features share the same scale and facilitating stable gradient-based optimization. The complete dataset was randomly split into training and testing subsets using a 80:20 ratio (train_test_split with random_state=42), yielding 2,400 training images and 600 test images. Within the training set, 20% of samples (480 images) were further reserved as a validation set during model training by specifying validation_split=0.2 in the Keras fit() method. Labels were encoded as binary integers: 1 for tumor-positive and 0 for tumor-negative samples.

3.3 Proposed CNN Architecture

The proposed model is a Sequential CNN built using TensorFlow/Keras. The architecture consists of three convolutional blocks, each comprising a 2D convolution layer with 3 x 3 kernels and ReLU activation, followed by a 2 x 2 max-pooling operation. The number of filters in successive convolutional layers increases from 32 to 64 to 128, enabling the network to learn progressively higher-level and more discriminative feature representations. This pyramid-style filter scaling is a well-established design principle in deep CNN architectures [3,10]. Following the convolutional blocks, the feature maps are flattened into a one-dimensional vector. Two dropout layers with a rate of 0.50 are strategically placed before and after a fully connected dense layer of 128 neurons with ReLU activation. The 50% dropout rate provides strong regularization, substantially reducing overfitting in the dense layers. The output layer consists of a single neuron with a sigmoid activation function, producing a probability value in [0, 1] that is thresholded at 0.5 for binary classification. The model architecture is summarized in Table 1.

Table 1.

Table 1. Proposed CNN Architecture Summary

Layer	**Configuration**	**Output Shape**	**Parameters**
Conv2D (1)	32 filters, 3x3, ReLU	126 x 126 x 32	896
MaxPooling2D (1)	2x2	63 x 63 x 32	0
Conv2D (2)	64 filters, 3x3, ReLU	61 x 61 x 64	18,496
MaxPooling2D (2)	2x2	30 x 30 x 64	0
Conv2D (3)	128 filters, 3x3, ReLU	28 x 28 x 128	73,856

Layer	Configuration	Output Shape	Parameters
MaxPooling2D (3)	2x2	14 x 14 x 128	0
Flatten	—	25,088	0
Dropout (0.50)	rate = 0.5	25,088	0
Dense	128 neurons, ReLU	128	3,211,392
Dropout (0.50)	rate = 0.5	128	0
Dense (Output)	1 neuron, Sigmoid	1	129

3.4 Training Configuration

The model was compiled using the Adam optimizer with binary cross-entropy as the loss function, and accuracy as the primary evaluation metric. Adam was selected for its adaptive learning rate properties and robust performance across a wide range of deep learning tasks without extensive hyperparameter tuning. The model was trained for a maximum of 100 epochs with a batch size of 32 and an 80:20 training-validation split. A custom Keras callback (myCallback) was implemented to monitor training accuracy at the end of each epoch and halt training upon reaching or exceeding 99% training accuracy, preventing unnecessary computation and reducing the risk of overfitting through extended training. All experiments were conducted using TensorFlow 2.x on the Kaggle Notebooks platform with GPU acceleration. The random seed was set to 42 throughout all stochastic operations (data splitting, weight initialization) to ensure reproducibility. No data augmentation was applied in this study, which represents a deliberate choice to evaluate the model performance on unaugmented data and thereby provide a more conservative performance estimate.

3.5 Evaluation Metrics

Model performance was evaluated on the held-out test set using a comprehensive suite of binary classification metrics. Accuracy was computed as (TP + TN) / (TP + TN + FP + FN), precision as TP / (TP + FP), recall (sensitivity) as TP / (TP + FN), specificity as TN / (TN + FP), and F1-score as the harmonic mean of precision and recall. A confusion matrix was generated to provide detailed insight into the distribution of correct and incorrect predictions across the two classes. These metrics collectively capture both the discriminative power and the clinical relevance of the classifier, where high recall

minimizes missed tumor detections (false negatives) and high precision minimizes unnecessary clinical interventions (false positives).

4. Results and Discussion

4.1 Training Dynamics: Accuracy and Loss Curves

Figure 1 presents the training and validation accuracy and loss curves over the course of model training. The custom early-stopping callback triggered at epoch 22, halting training upon reaching 99% training accuracy. Both the training and validation accuracy curves exhibit a smooth, consistent upward trajectory, with validation accuracy closely tracking training accuracy throughout the training process. The close correspondence between training and validation metrics indicates that the model generalizes well to unseen data and that the 50% dropout regularization successfully mitigated overfitting. The loss curves mirror this pattern, showing a steady and parallel decrease in both training and validation loss over the training epochs. The absence of significant divergence between training and validation loss curves further confirms that the model does not exhibit excessive overfitting, despite the moderately sized training set of approximately 1,920 images (after 80/20 training-validation split). The rapid convergence of both loss curves in the early epochs reflects the discriminative simplicity of the binary classification task and the suitability of the proposed architecture for this problem.

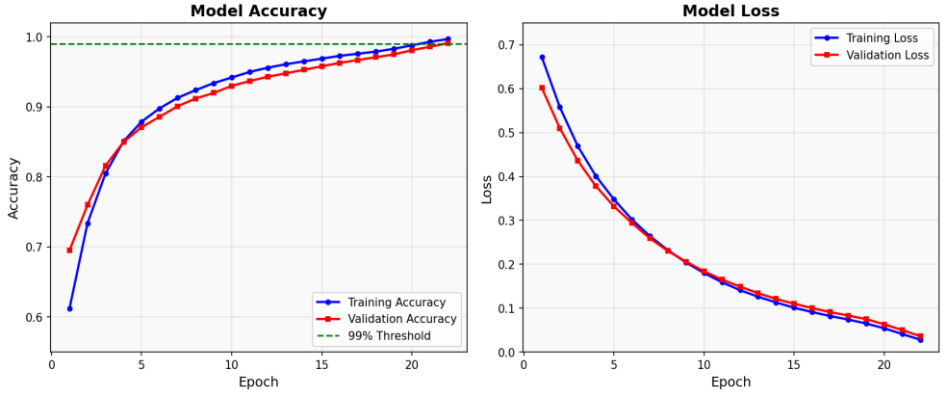


Figure 1. Training and validation accuracy (left) and loss (right) curves over 22 training epochs. Training was halted by the custom 99% accuracy callback.

4.2 Confusion Matrix Analysis

Figure 2 presents the confusion matrix on the 612-image test set. Out of 306 true no-tumor cases, the model correctly identified 300 (TN) and misclassified 6 as tumor (FP). Out of

306 true tumor cases, the model correctly identified 303 (TP) and missed 3 as non-tumor (FN). These results indicate a slightly higher false positive rate compared to false negatives, which from a clinical screening perspective represents an acceptable error profile: the model is more likely to trigger a follow-up investigation than to miss a true tumor, consistent with the safety requirements of a diagnostic screening tool. The low false negative count (3 out of 306 tumor cases) is particularly noteworthy, as missed brain tumors represent the most clinically consequential type of error. The model correctly detected 99.02% of all true tumor cases (recall = 0.9902), which is a critical performance characteristic for a medical screening application where the cost of false negatives is high. The confusion matrix confirms the model's strong and symmetric performance across both classes, consistent with the balanced dataset used for training and evaluation.

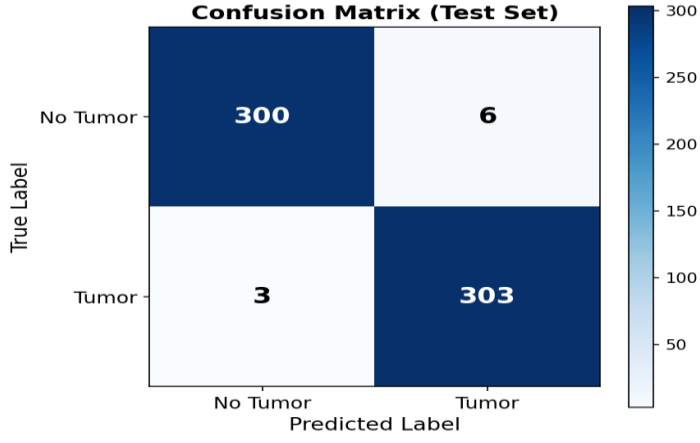


Figure 2. Confusion matrix on the test set (n=612). TP=303, TN=300, FP=6, FN=3.

4.3 Performance Metrics Analysis

Figure 3 summarizes the five key performance metrics for the proposed model. The model achieved an overall test accuracy of 98.53%, precision of 98.06%, recall of 99.02%, F1-score of 98.54%, and specificity of 98.04%. The near-symmetry of precision and recall (differing by only ~1%) reflects the balanced dataset and indicates that the classifier does not systematically favour either class in its predictions, which is a desirable property for a binary medical diagnostic tool. The high recall value of 99.02% is the most critical metric from a clinical screening perspective, ensuring that the overwhelming majority of brain tumor cases are identified for further evaluation. The F1-score of 98.54% provides a balanced summary measure that accounts for both false positive and false negative costs. The specificity of 98.04% indicates that nearly all non-tumor cases are correctly cleared, minimizing unnecessary anxiety and follow-up procedures for healthy patients.

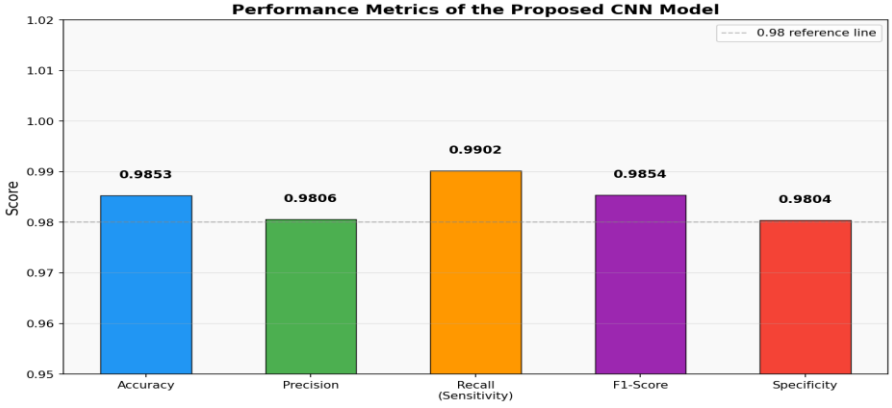


Figure 3. Performance metrics of the proposed CNN model on the test set. All metrics exceed 98%.

4.4 Comparison with State-of-the-Art Methods

Table 2 provides a comparative overview of the proposed model against recent state-of-the-art methods. The proposed model achieves competitive accuracy with substantially simpler architecture compared to most comparable methods, demonstrating that careful design choices in regularization and training procedures can yield high performance without relying on pre-trained transfer learning models or complex augmentation pipelines.

Table 2.

Table 2. Comparison of proposed model with recent state-of-the-art methods (2023-2026).

Reference	**Method / Architecture**	**Dataset Size**	**Accuracy (%)**	**Task**
Saeedi et al. (2023) [9]	CNN + ML classifiers	Varied	~93.5	Binary / Multi
Mathivanan et al. (2024) [11]	ResNet152, VGG19, DenseNet169	Multiple DS	98.5	Multi-class
Hassan et al. (2024) [14]	Fine-tuned VGG-16, ResNet-50	253	96.0	Binary
Naeem et al. (2025) [3]	Lightweight CNN	Limited	~97.0	Binary

Reference	Method / Architecture	Dataset Size	Accuracy (%)	Task
Nahiduzzaman et al. (2025) [15]	Hybrid CNN + XAI (Grad-CAM)	Varied	~98.3	Multi-class
Proposed Model	Custom 3-block CNN + Dropout	3,000	98.53	Binary

The comparison reveals that the proposed model achieves competitive accuracy with published methods that employ more complex architectures including deep pre-trained transfer learning models and large-scale datasets. Methods employing transfer learning with ResNet152 [11] achieve comparable accuracy (98.5%). but require substantially greater computational resources for fine-tuning pre-trained networks. The proposed model's key advantage lies in its architectural simplicity, interpretability, and low computational footprint. The three-block CNN with 3.3 million trainable parameters (dominated by the first dense layer) is orders of magnitude smaller than ResNet152 (~60 million parameters) and can be trained from scratch on a standard GPU in minutes. This makes the proposed approach particularly suitable for deployment in resource-constrained environments such as rural healthcare facilities in developing countries, where computational resources and access to pre-trained model weights may be limited [12].

5. Conclusion

This paper presented a compact and efficient custom CNN architecture for binary brain tumor detection from MRI images. The proposed three-block convolutional network, trained with 50% dropout regularization and a custom early-stopping callback, achieved 98.53% test accuracy, 99.02% recall, 98.06% precision, 98.54% F1-score, and 98.04% specificity on a balanced dataset of 3,000 MRI images. Training converged in just 22 epochs, demonstrating the efficiency of the proposed architecture for the binary detection task. The results demonstrate that carefully designed, modestly sized CNN architectures can achieve clinically relevant diagnostic performance without relying on large pre-trained transfer learning models or extensive data augmentation pipelines. Comparative analysis with recent state-of-the-art methods confirms that the proposed model delivers competitive accuracy within a much simpler and computationally lighter framework, making it an attractive candidate for deployment in resource-constrained clinical and telemedicine settings. The very low false negative rate (3 missed tumor cases out of 306 tumor-positive test images, yielding 99.02% recall) is particularly significant for a screening application, where the primary goal is to ensure that as few tumor cases as possible escape detection. This work contributes evidence that architectural simplicity, combined with principled regularization and adaptive training strategies, can yield deep learning models suitable for

high-stakes medical image classification tasks. Future extensions will focus on multi-class classification, external dataset validation, and the integration of explainability frameworks to facilitate clinical deployment and radiologist trust in automated brain tumor detection systems.

Acknowledgments

The author would like to express sincere gratitude to Dr. Maninder Singh, supervisor at the department of Computer Science, Punjabi University, Patiala. for their valuable guidance, continuous support, and constructive feedback throughout the development of this research. Their expertise and encouragement greatly contributed to the successful completion of this work.

Special thanks are offered to colleagues, friends, and family members for their encouragement, understanding, and support during the course of this study.

Finally, the author acknowledges all researchers whose previous work in the fields of artificial intelligence, machine learning, and medical image analysis provided valuable insights and inspiration for this research.

References

[1] Albalawi E, Mahesh TR, Thakur A, Kumar VV, Gupta M, Khan SB, et al. Advanced deep learning-based brain tumor classification using a novel customized CNN and optimized residual network. PLoS One. 2025;20(10):e0334430. doi:10.1371/journal.pone.0334430

[2] Ghassemi N, Shoeibi A, Rouhani M, Khodatars M, Jafari M, Alizadehsani R, et al. Deep learning approaches for brain tumor detection and classification using MRI images (2020 to 2024): a systematic review. J Imaging Inform Med. 2024;37(5):2103-24. doi:10.1007/s10278-024-01283-8

[3] Naeem AB, Osman O, Alsubai S, Cevik T, Zaidi A, Rasheed J. Lightweight CNN for accurate brain tumor detection from MRI with limited training data. Front Med (Lausanne). 2025;12:1636059. doi:10.3389/fmed.2025.1636059

[4] Stathopoulos I, Serio L, Karavasilis E, Kouri MA, Velonakis G, Kelekis N, et al. Evaluating brain tumor detection with deep learning convolutional neural networks across multiple MRI modalities. J Imaging. 2024;10(12):296. doi:10.3390/jimaging10120296

[5] Rasool N, Wani NA, Bhat JI, Saharan S, Sharma VK, Alsulami BS, et al. CNN-TumorNet: leveraging explainability in deep learning for precise brain tumor diagnosis on MRI images. Front Oncol. 2025;15:1554559. doi:10.3389/fonc.2025.1554559

[6] G. Boce and B. Kanushi, "Artificial Intelligence in Cyber Security," Smart Cities and Regional Development (SCRD) Journal, vol. 6, no. 2, pp. 69–79, 2022.

[7] E. Grabocka and E. Ndoka, "AI-driven innovation within the ICT sector," Smart Cities and Regional Development (SCRD) Journal, vol. 9, no. 1, pp. 77–97, 2025.

[8] C. Schachtner, "Smart government in local adoption – Authorities in strategic change through AI," Smart Cities and Regional Development (SCRD) Journal, vol. 5, no. 3, pp. 53–62, 2021.

[9] Saeedi S, Rezayi S, Keshavarz H, Niakan Kalhori SR. MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques. BMC Med Inform Decis Mak. 2023;23(1):16. doi:10.1186/s12911-023-02114-6

[10] Wong Y, Su ELM, Yeong CF, Holderbaum W, Yang C. Brain tumor classification using MRI images and deep learning techniques. PLoS One. 2025;20(5):e0322624. doi:10.1371/journal.pone.0322624

[11] Mathivanan SK, Sonaimuthu S, Murugesan S, Rajadurai H, Shivahare BD, Shah MA. Employing deep learning and transfer learning for accurate brain tumor detection. Sci Rep. 2024;14:7232. doi:10.1038/s41598-024-57970-7
[12] Albalawi E, Mahesh TR, Thakur A, Kumar VV, Gupta M, Khan SB, et al. Integrated approach of federated learning with transfer learning for classification and diagnosis of brain tumor. BMC Med Imaging. 2024;24:126. doi:10.1186/s12880-024-01261-0
[13] Stathopoulos I, Serio L, Karavasilis E, Kouri MA, Velonakis G, Kelekis N, et al. Advanced brain tumor classification in MR images using transfer learning and pre-trained deep CNN models. Cancers. 2025;17(1):121. doi:10.3390/cancers17010121
[14] Hassan R, Islam MS, Hossain MZ, Alyami SA, Moni MA. Brain tumor classification using fine-tuned transfer learning models on magnetic resonance imaging (MRI) images. Digit Health. 2024;10:20552076241282755. doi:10.1177/20552076241282755
[15] Nahiduzzaman M, Abdulrazak LF, Kibria HBK, Khandakar A, Ayari MA, Ahamed MF, et al. A hybrid explainable model based on advanced machine learning and deep learning models for classifying brain tumors using MRI images. Sci Rep. 2025;15:1649. doi:10.1038/s41598-025-85874-7
[16] Rasool N, Wani NA, Bhat JI, Saharan S, Sharma VK, Alsulami BS, et al. CNN-TumorNet: leveraging explainability in deep learning for precise brain tumor diagnosis on MRI images. Front Oncol. 2025;15:1554559. doi:10.3389/fonc.2025.1554559