Science World Journal Vol. 21(No 1) 2026 https://dx.doi.org/10.4314/swj.v21i1.41
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

Full Length Research Article

CLASSIFICATION OF BRAIN TUMOR USING HYBRID MODEL OF COSINE ANNEALING AND WEIGHTED SNAPSHOT ENSEMBLES

*[1]Muttaqa Abdulhameed Jibril, [1]Salisu Aliyu, [1]Aliyu Ibrahim Tetengi , [1]Abdulaziz Saidu, [2]Isah Bello, [3]John Joshua

[1]Department of Computer Science, Ahmadu Bello University, Zaria, Nigeria
[2]School of Electrical and Information Engineering, Tianjin University, China
[3]Department of Computer Science, Federal University of Applied Sciences, Kachia, Nigeria

*Corresponding Author Email Address: mjabdulhameed@abu.edu.ng

ABSTRACT
The detection of brain tumors from MRI images is a critical initial step in brain cancer diagnosis, a task heavily dependent on the expertise of a radiologist. Consequently, there is a growing interest in developing automated diagnostic methods to assist radiologists. Hence, it minimizes the need for invasive procedures such as biopsies. Convolutional Neural Networks (CNNs) are recognized as a highly effective deep learning algorithm for accurate tumor identification and classification. While custom CNNs have proven effective in tumor classification, they often suffer from overfitting, hyperparameter sensitivity, and limited ensemble diversity, which limits their generalization performance. This study proposes a deep learning model based on a custom CNN that uses cosine annealing for learning-rate scheduling and a weighted snapshot ensemble of optimizers: Nadam, Adam, and Adamax. Cosine annealing is employed to mitigate fluctuations in validation performance during training, which can lead to overfitting, unstable training, misleading evaluation metrics, and increased risk of bias. The snapshot ensemble has been introduced to enhance the model's classification performance. Each optimizer is trained independently, with two snapshots taken after 40 epochs to ensure proper convergence. The proposed model was trained and evaluated using the publicly available Figshare Dataset. Our approach achieved exceptional performance, with an accuracy, precision, recall, and F1-score of 97.9%, respectively. These results demonstrate the potential of our model to enhance automated brain tumor detection. Hence, supporting radiologists in any clinical setting.

Keywords: Brain tumor, CNN, MRI, Cosine Annealing, Snapshot ensemble, WOA

INTRODUCTION

The brain, with its billions of cells, is one of the most intricate organs in the human body. A brain tumor develops when cells begin to divide uncontrollably, forming abnormal clusters either inside or on the surface of the brain. Although brain tumors account for just 1.8% of all new cancer cases worldwide, making them the 22nd most common type, their impact is disproportionately severe. Brain cancer may not be as prevalent as other cancers, but it has a higher death rate compared to its number of annual diagnoses (*Siegel et al., 2022)*. In 2021, the American Cancer Society projected 24,530 new brain cancer cases in the U.S. Even though brain cancer is responsible for only 3% of deaths from all cancers, a staggering 75% of those diagnosed with tumors of the brain or nervous system are likely to die. Brain cancer is also the most lethal type for men under 40 and women under 20 (*Siegel et al., 2018)*. According to the World Health Organization, brain tumors are classified into 120 types based on where they originate and how the cells behave. They are also categorized by grades that reflect their aggressiveness and rate of spread *(Louis et al., 2007)*.

According to radiologists, brain tumors are broadly categorized into two types: benign and malignant. Benign tumors are non-cancerous and do not spread to other parts of the body, while malignant tumors are cancerous and typically require surgical treatment (*Sravan et al., 2020*). Early identification of tumor types helps doctors and radiologists determine the best course of treatment. Depending on the tumor's location and severity, brain tumors are further classified into three main types: gliomas, meningiomas, and pituitary tumors (Viswa Priya, 2016; *Patil et al., 2020*). Gliomas arise from immature brain stem cells that begin to grow abnormally due to mutations (Patil & Kirange, 2023). Meningiomas form in the membranes that protect the brain and central nervous system, while pituitary tumors develop behind the eyes at the center of the head (*Patil et al., 2020*).

Each tumor type differs in severity, surrounding tissue, and location, making it essential to examine these features to classify tumors as malignant or benign. Meningiomas and pituitary tumors are usually benign, but most high-grade gliomas are malignant (Patil & Kirange, 2023). Because of these differences, accurate classification is crucial for clinical diagnosis and effective treatment planning.

Brain tumors develop from abnormal cell growth in the brain or spinal canal. Several factors, including the location of the tumor, help determine whether the tissue is benign or malignant. Detecting brain tumors early is vital to preventing their progression to a more severe stage. To diagnose tumors, radiologists use advanced techniques like biopsies, cerebrospinal fluid (CSF) analysis, and X-ray imaging (*Singh et al., 2024*). In a biopsy, a small tissue sample is surgically removed and examined to check for tumors, though this procedure carries risks such as inflammation and severe bleeding. Additionally, biopsy accuracy is only 49.1% (*Singh et al., 2024*). CSF analysis involves examining the brain's colorless fluid for signs of tumors, but this method also poses risks, such as allergic reactions and bleeding. X-ray imaging exposes the head to radiation, increasing the risk of cancer. Due to the limitations and risks of these methods, radiologists are increasingly relying on imaging modalities that offer greater accuracy and significantly lower patient risk.

Various imaging techniques are used to examine brain tumors, including magnetic resonance imaging (MRI), computed tomography (CT), positron emission tomography (PET), and single-photon emission computed tomography (SPECT). Among these, CT and MRI are the most commonly employed due to their broad availability and capability to produce high-resolution images of both normal anatomical structures and abnormal pathologies

(*Kasban et al., 2015*).

Magnetic resonance imaging (MRI) is the most widely used technique for brain tumor classification due to its high-resolution and radiation-free imaging. However, manual segmentation and classification based on MRI can be prone to human error, especially in complex cases, as the process relies heavily on a radiologist's experience (*Chhabda et al., 2016*). To pinpoint the exact tumor location, MRI images are taken in slices across different modalities. Despite its advantages, manual segmentation is time-consuming and often inaccurate, especially as patient numbers increase (*Sravan et al., 2020*).

To address these challenges, Computer-Aided Diagnosis (CAD) systems have been developed to assist in the early detection of brain tumors without human intervention. These systems analyze MRI images and can generate diagnostic reports, guiding radiologists toward more accurate conclusions (*Loffe et al., 2015*). MRI is particularly effective for visualizing brain structures such as grey matter, white matter, and ventricles, and it provides clear and precise images of brain tumors (*Ghanavati et al., 2012*). Its ability to deliver accurate results makes MRI a preferred method in clinical analysis (Sravan et al., 2020). Radiologists traditionally examine MRI scans by visually screening the slices to identify tumor tissues (*Choi et al., 2009*). In contrast, computed tomography (CT) scans, while useful for understanding the functional and structural status of brain diseases, provide less detailed soft-tissue information than MRI. CT scans, which use ionizing radiation, are typically more effective in capturing detailed images of bone structures near a tumor, such as the skull or spine. A CT scan may be used in cases where an MRI is not possible, such as when a patient has implants like a pacemaker, or in emergencies like acute hemorrhages or head trauma (*Luo et al., 2018*). Given MRI's superior ability to capture high-resolution images of brain tissues, this work focuses on exploring segmentation and classification techniques that rely solely on MRI scans for brain tumor diagnosis.

According to *Bhalodiya et al. (2022),* the application of Machine Learning (ML) and Deep Learning (DL) in medical imaging has greatly enhanced the Computer-Aided Diagnosis (CAD) process, improving the accuracy of brain tumor detection. These techniques rely on the key principles of feature extraction, feature selection, and classification (Anitha & Raja, 2013). Various feature extraction methods, such as thresholding-based, clustering-based, contour-based, and texture-based approaches, are used to isolate the tumor region from the human skull in MRI images (*Sravan et al., 2020*). Feature extraction involves isolating the most relevant details from MRI scans, followed by feature selection, where only the most important features are chosen for analysis (Jalab & Hasan, 2019). High accuracy depends on extracting features that hold strong discriminatory information. However, one potential drawback is that the feature extraction process may overlook important details from the original image, leading to a loss of crucial information (*Scherer et al., 2010*).

Deep Learning (DL) techniques address the limitations of traditional feature extraction by directly using the original image as input, without the need for manually selected features (Anitha & Raja, 2013). One popular DL model, the Convolutional Neural Network (CNN), uses multiple convolutional layers to automatically extract features from images (Yaqub et al., 2020). CNNs are particularly effective when applied to large datasets, as they can learn complex patterns from the data.

The key contributions of this study are:

1. The model employs a novel approach to convolutional kernel sizing, progressively reducing the kernel sizes from **32×32 to 3×3** across five convolutional layers. This enables efficient multi-scale feature extraction, capturing both fine-grained and high-level spatial patterns crucial for tumor classification.
2. The study introduces a **novel snapshot ensemble approach**, where two model snapshots are saved per optimizer (Adam, Adamax, and Nadam), leading to a total of **six diverse models**. This ensemble technique enhances classification robustness by leveraging the strengths of different optimization strategies.
3. To further refine classification accuracy, the **Whale Optimization Algorithm (WOA) is employed to determine the optimal weights for ensemble models**. This ensures that class probabilities are adaptively adjusted based on each model's relative importance, leading to more reliable decision-making.

Brain tumor detection and classification have become key research areas, with various segmentation techniques developed to aid diagnosis. Traditional machine learning methods rely on preprocessing, feature extraction, and classification, but struggle with feature representation due to tumor similarities (*Selvaraj et al., 2007*). Deep learning eliminates the need for manual feature extraction, enabling automatic feature learning during classification (*Litjens et al., 2017*). Studies have demonstrated the potential of AI in brain tumor detection, with some using neural networks and feature transformation techniques for MRI image classification (Sumitra & Saxena, 2013). While early research shows promise, performance metrics are often unreported.

Kaur and Gandhi (2020) introduced a deep convolutional neural network (DCNN) with transfer learning for automated brain tumor classification. They noted that CNN feature extraction depends on dataset size, with smaller datasets leading to overfitting. To address this, they evaluated pre-trained DCNN models, including AlexNet, ResNet50, GoogLeNet, VGG-16, ResNet101, VGG-19, InceptionV3, and InceptionResNetV2, modifying the final layers for classification. Using data from Harvard, clinical repositories, and the Figshare dataset, they found that AlexNet performed best, achieving 95.92% accuracy on Figshare. However, the study did not employ advanced techniques like ensemble modeling or optimization algorithms, which could have further improved accuracy. Cinar and Yildirim (2020) proposed a hybrid CNN architecture for detecting tumors in brain MRI images using transfer learning. They trained and evaluated pre-trained models, including AlexNet, ResNet50, DenseNet201, InceptionV3, and GoogLeNet, with ResNet50 achieving the highest accuracy at 97.2%. Based on its performance, ResNet50 was chosen as the base model, with the last five layers replaced by ten new layers, increasing the total from 177 to 182. The study focused on hyperparameter fine-tuning but did not assess alternative optimizers such as Adam, Adamax, and Nadam, which could have further improved performance. Mengash and Mahmoud (2021) proposed a CNN model to classify brain tumors using motion-corrected MRI images. The model distinguishes among normal tissue, astrocytoma, gliomatosis cerebri, and glioblastoma. It was tested on MRI images with motion correction using two evaluation methods. The first, k-fold cross-validation with k values of 8, 10, 12, and 14, achieved the highest accuracy of 96.26% at k = 10. The second, the hold-out testing approach, yielded 97.8% accuracy, 99.2% specificity, and 97.32%

sensitivity. However, hold-out validation may not be ideal for small or imbalanced datasets, while k-fold cross-validation incurs high computational cost and can introduce bias. For imbalanced datasets, ensemble techniques could help reduce bias and error. Deepak and Ameer (2021) proposed an automated brain tumor categorization method using CNN features and SVM to address limited data in medical imaging. They tested their system on the Figshare dataset, which includes glioma, meningioma, and pituitary tumors. By integrating a multiclass SVM with CNN features, they achieved 95.82% accuracy in fivefold cross-validation, thereby improving classification performance on small datasets. However, the method has limitations, including increased variability from testing individual data points, which can be affected by outliers. Additionally, the approach is time-intensive, making it less practical for larger datasets. Khan et al. (2022) proposed a deep CNN-based approach for accurate brain tumor detection, using two models for binary and multiclass classification. They utilized the Figshare and Harvard Medical MRI datasets, applying a 23-layer CNN to the larger Figshare dataset and transfer learning with VGG16 for the smaller Harvard dataset. The model achieved 97.8% accuracy on Figshare and 100% on Harvard. However, validation performance declined after epoch 33 and fluctuated until epoch 80 due to the absence of a learning rate scheduler. This led to overfitting, unstable training, misleading evaluation metrics, and increased bias, ultimately affecting performance. Rahman and Islam (2023) introduced a Parallel Deep Convolutional Neural Network (PDCNN) to capture both global and local features while mitigating overfitting through dropout regularization and batch normalization. The PDCNN integrates two parallel deep CNNs with different window sizes, enabling it to learn fine details and broader patterns simultaneously. It was tested on three MRI datasets: a binary tumor dataset (97.33% accuracy), the Figshare dataset (97.60% accuracy), and a multiclass Kaggle dataset (98.12% accuracy). The study also employed an Anisotropic Diffusion Filter (ADF) to reduce noise while preserving edge details, enhancing image quality, and improving classification performance. Anaya-Isaza et al. (2023) proposed a brain tumor detection and classification framework using seven pretrained deep learning models: InceptionResNetV2, InceptionV3, DenseNet121, Xception, ResNet50V2, VGG19, and EfficientNetB7. These models were trained on the Figshare dataset to classify glioma, meningioma, and pituitary tumors. Tumor detection was conducted using the Brain MRI Images for Brain Tumor Detection and the Cancer Genome Atlas Low-Grade Glioma databases. InceptionResNetV2 achieved the highest accuracy (97.22%) and an F1-score of 95.39%, while InceptionV3 had the highest sensitivity (97.81%). VGG19 and EfficientNetB7 achieved 100% specificity, with EfficientNetB7 also reaching 100% precision. The study suggests that performance could have been further improved by using metaheuristic algorithms for parameter optimization rather than manual adjustments. Patil and Kirange (2023) proposed an ensemble deep learning model to improve brain tumor detection and classification, addressing challenges in tumor localization. They combined a shallow convolutional neural network (SCNN) with a VGG16 network using T1C-modality MRI images. By fusing features from both models, they achieved 97.77% classification accuracy across glioma, meningioma, and pituitary tumors. However, the accuracy plateaued after a certain number of training epochs. So using a snapshot technique during training could have further enhanced performance. Gómez-Guzmán et al. (2023) proposed an ensemble learning approach using seven CNN models: Generic CNN, ResNet50, InceptionV3, InceptionResNetV2, Xception, MobileNetV2, and EfficientNetB0. These models were trained on the Figshare, SARTAJ, and Br35H datasets, comprising 7,023 MRI images across four classes: glioma, meningioma, pituitary tumors, and healthy brains. InceptionV3 achieved the highest accuracy at 97.12%. However, the study did not optimize the ensemble model's weights; using the Whale Optimization Algorithm (WOA) could have further improved performance by refining weight adjustments across the CNN models. Kanchanamala et al. (2023) proposed an optimization-enabled hybrid deep learning approach for brain tumor detection and classification. They introduced ExpDHO, an optimization algorithm that combines Exponential Weighted Moving Average (EWMA) with the Deer Hunting Optimization Algorithm (DHOA) to enhance training. The models were trained on the Figshare and BRAT2018 datasets, achieving accuracies of 92.9% and 91.7%, respectively. Xu and Mohammadi (2024) introduced a model for brain tumor diagnosis from MRI scans, leveraging the MobileNetV2 architecture optimized by the Contracted Fox Optimization Algorithm (CFOA). This innovative approach combines deep learning with a metaheuristic algorithm to enhance tumor detection accuracy. The researchers used MobileNetV2, a deep learning model tailored for mobile and embedded vision applications, to extract relevant features from MRI scans effectively. The results of the study are impressive, achieving an accuracy of 97.32%. However, while the lightweight MobileNet architecture enables faster processing, it generally achieves lower classification accuracy than more complex models.

MATERIALS AND METHODS

Dataset Source

The proposed model was trained, evaluated, and tested using the publicly accessible benchmark Figshare dataset. This dataset was collected from General Hospital, Tianjin Medical University, and Nanfang Hospital in China, spanning the years 2005 to 2010. It comprises a total of 3,064 T1-weighted contrast MRI slices from 233 patients, each diagnosed with one of three types of brain tumors: 708 images of meningiomas, 1,426 images of gliomas, and 930 images of pituitary tumors. To illustrate the dataset's diversity, Figure 3.1 below presents an example from each class, highlighting the tumor's location and size within the MRI images. The dataset includes MRI images viewed in three different orientations: axial, coronal, and sagittal *(Khan et al., 2022)*.

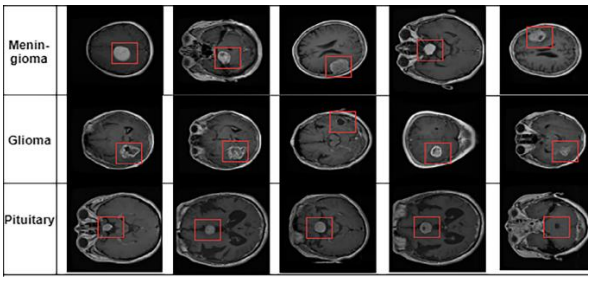


Figure 1: Different samples of brain tumors, including Glioma, Meningioma, and pituitary. With the tumor in a rectangle

Dataset distribution

The table below shows the distribution of the dataset across each class. The data contains three brain image classes: glioma,

meningioma, and pituitary tumor.

Table 1: Distribution of Brain classes.

Tumor Class	Number of patients	Number of MR slices	Percentage
Meningioma	82	708	23.1%
Glioma	91	1426	46.5%
Pituitary	60	930	30.4%
Total	233	3064	100%

Data preprocessing

In the diagnostic process for brain tumor detection using medical imaging techniques such as magnetic resonance imaging (MRI), image preprocessing plays a crucial role. This step involves a series of operations applied to raw image data to enhance quality, reduce noise, and extract important features, enabling accurate interpretation and analysis by medical professionals. In this study, several preprocessing techniques were implemented before the images were fed into the classifiers. The MRI images from the Figshare dataset are in .mat format (defined in Matlab), and each image has a resolution of 512 x 512 pixels, providing detailed information for each case. These images were resized to dimensions of 244 x 244 pixels to standardize the input for the model. Following this, all resized images were converted to NumPy arrays (in Python) to reduce storage space. Before splitting the dataset, it was shuffled to ensure that the model trained on unordered data, enhancing its ability to generalize. The shuffled dataset was then divided into three sections: approximately 70% for training, 30% for validation, and 10% for testing. This structured approach ensures that the model is effectively trained and evaluated on diverse data.

Experimental Setup

The proposed technique was implemented in Python within a Jupyter notebook hosted on Kaggle's cloud platform. The implementation utilized the Keras and TensorFlow frameworks and was supported by a GPU with 16GB of VRAM and a CPU with 29.2GB of memory. The benchmark Figshare dataset was used for the analysis.

Experimental settings and hyper-parameter selection

In the Figshare dataset, MRI images with a resolution of 512x512 are provided in '.mat' format. During data preprocessing, all images are normalized by dividing each image by 255 and resized to 244 x 244 to match the input dimensions of the Custom CNN. The dataset is then randomly divided into training, validation, and test sets in a 70:15:15 ratio. The validation set is used to monitor the learning behavior of the Custom CNN model and to fine-tune the hyperparameters. The key hyperparameters influencing the learning process were the optimizers and the initial learning rate, which was set to the default value across all optimizers. For the Custom CNN model, the selected hyperparameters include a default learning rate of 0.001 for all optimizers (Adam, Adamax, and Nadam), a batch size of 32, a maximum of 80 epochs, a dropout rate of 0.25, and categorical cross-entropy as the loss function. Table 4.1 summarizes all hyperparameters used in the study.

Table 2: Experimental hyper-parameter settings.

Hyper-parameter	Setting (Value)
Optimizers	Adam, Adamax, Nadam
Loss Function	Categorical Cross-entropy
Initial Learning Rate	0.001
Batch Size	32
Number of epochs	80
Dropout	0.25

Proposed Custom CNN architecture

The proposed Custom CNN architecture is an enhancement of the 23-layer CNN model described by Khan et al. (2022), tailored for classifying brain tumors, including meningioma, glioma, and pituitary tumors. This architecture uses MRI slices as input and passes them through a series of layers to effectively distinguish tumor types.

The architecture of this model begins with an input layer that processes MRI slices, followed by five convolutional layers. These layers use progressively smaller kernel sizes, 32x32, 22x22, 11x11, 7x7, and 3x3, to capture important details such as edges, corners, and shapes, using the dot product between the kernel and the pixel intensities of the input image. Each filter moves two pixels at a time (stride of 2), with zero-padding applied to preserve the original image size and prevent detail loss. Next, dimension reduction is handled by two pooling layers, using five max pooling layers with sizes of 4x4 and 2x2 to retain key features from each feature map. To further support training, eight batch normalization layers follow, normalizing the outputs of the convolutional and fully connected layers. This normalization is performed before the activation function, accelerating training and helping prevent covariance shift. The Leaky ReLU activation function is used consistently throughout, as it allows a small, non-zero output for negative inputs, which improves the model's learning capability. Following these, a global average pooling layer reduces multi-dimensional data to a one-dimensional vector by averaging the outputs of each feature map. Four dense layers are added next, fully connecting each input from one layer to every activation unit in the following layer, compiling features from previous layers into a final output. To prevent overfitting, a dropout layer with a 25% rate is applied before the classification layer, so that 75% of the features are dropped during each training iteration, improving the model's generalization. For the output layer, a Softmax activation function maps the processed features to class predictions. A snapshot approach is used to enhance model performance by saving multiple versions of the model during training when improvements are detected. In this study, two versions were saved for each optimizer (Adam, Adamax, and Nadam), resulting in six model versions that were later combined into an ensemble for classification. Additionally, the Whale Optimization Algorithm (WOA) is employed to find optimal weights for this ensemble technique, aiming to enhance final classification accuracy and reliability by effectively balancing the strengths of different models. This ensures that class probabilities are weighted proportionately to their relative importance.

Science World Journal Vol. 21(No 1) 2026 https://dx.doi.org/10.4314/swj.v21i1.41
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

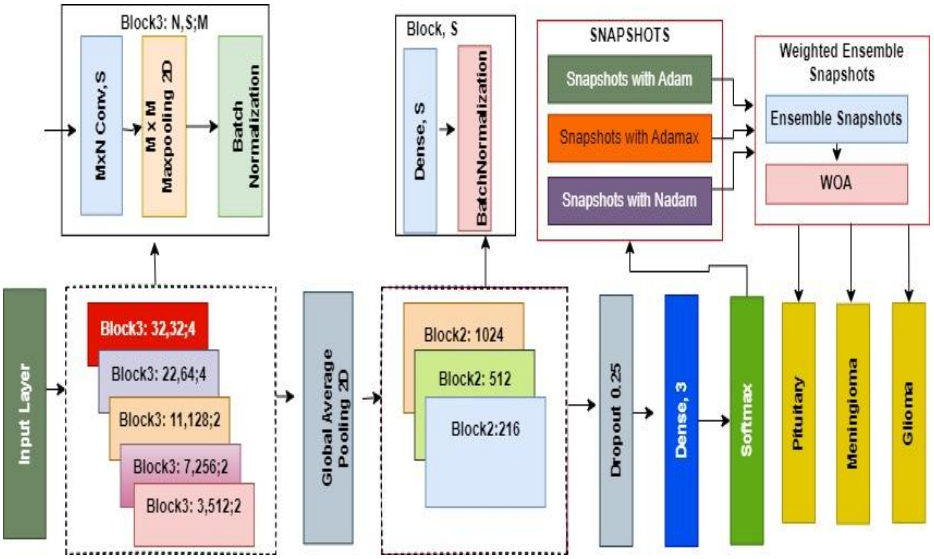


Figure 2: The proposed Custom CNN Architecture.

Ensemble weight optimization

This technique focuses on improving the performance of the proposed snapshot ensemble model by effectively integrating the outputs of multiple base models. Each of these models exhibits different training behaviors, allowing them to contribute unique predictions to the ensemble. The primary goal of ensemble weight optimization is to identify the most effective weight combinations for each model's predictions. By doing so, the final aggregated prediction becomes as accurate as possible, harnessing the strengths of each model to enhance overall performance. Suppose there are three base models: M1, M2, and M3. Ensemble weight optimization involves finding the best weights w_1, w_2, and w_3 such that the combined prediction:

Final Prediction = $w_1 \times M_1 + w_2 \times M_2 + w_3 \times M_3$

Evaluation metrics

However, the performance of the proposed model will be evaluated using four quantitative measures: accuracy, precision, recall, and F1-score. These metrics are developed from:

- TP (True Positives): MRI scans correctly identified as having a tumor
- TN (True Negatives): MRI scans correctly identified as normal
- FP (False Positives): Normal scans incorrectly predicted as having tumors
- FN (False Negatives**):** Tumor scans incorrectly predicted as normal (most dangerous case).

Predictions can be represented using the following mathematical expressions.

Accuracy

In brain tumor classification, accuracy measures how often the model correctly identifies whether an MRI scan contains a tumor (positive) or no tumor (negative). Accuracy tells us the overall correctness of the model. This reflects the number of correct predictions (TP and TN) over all predictions opentions$(TP + FP + TN + FN)$.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

Precision

Precision is important in brain tumor detection because it reflects how many MRI scans predicted as "tumor" truly contain tumors. Also, high precision means few false alarms, meaning the model rarely flags normal brains as tumors.

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

Recall

Recall is critically important in medical diagnosis, especially for brain tumors. Because it measures the model's ability to detect actual tumor cases among all tumor-containing MRIs. High recall means the model rarely misses tumors. Low recall leads to many false negatives, which is dangerous because an undetected tumor can delay treatment.

In clinical practice, recall is often more important than precision, because missing a tumor is far more harmful than a false alarm.

$$Recall = \frac{TP}{(TP+FN)} \quad (3)$$

F1-score

The F1-score balances precision and recall, making it ideal for datasets where tumor and non-tumor images are unevenly distributed. Useful when both false positives and false negatives matter. Ensures the model not only finds tumors (high recall) but also predicts tumors correctly (high precision). A high F1-score indicates a strong and reliable brain tumor classification model.

$$F1 - score = \frac{2TP}{2TP+FP+FN} \quad (4)$$

RESULTS

The proposed model in this study is structured into three stages, each using the same hyperparameters except for the optimizers. In the first stage, the model was trained with three optimizers (Nadam, Adam, and Adamax) to evaluate its performance under each. In the second stage, the Weighted Snapshot Ensemble technique was applied to enhance the model's performance. Lastly, in the third stage, each snapshot from the second stage was hybridized with optimal weight assignments to improve the model's performance further.

Science World Journal Vol. 21(No 1) 2026 https://dx.doi.org/10.4314/swj.v21i1.41
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

Performances of the Optimizers without the Snapshot Technique

This section presents the results of the model trained independently with each optimizer, without applying weights or snapshot techniques. As shown in Table 4.2, the Adamax optimizer delivered the best performance, achieving an accuracy of 97.90%, followed by Nadam at 96.90%, and finally the Adam optimizer at 96.20%.

Table 2: Results of the performances of individual optimizers without Snapshot Ensemble and Weights assignment techniques

Optimizer	Tumor type	Accuracy%	Precision%	Recall %	F1-score %
Nadam	Meningioma	92.30	92.30	94.30	93.30
	Glioma	97.20	97.20	96.20	96.20
	Pituitary	100	100	100	100
	Average	96.90	96.90	96.90	96.90
Adam	Meningioma	91.60	91.60	92.90	92.30
	Glioma	96.90	96.90	95.90	96.40
	Pituitary	98.90	98.90	99.40	99.20
	Average	96.20	96.20	96.20	96.20
Adamax	Meningioma	95.10	95.10	96.60	96.20
	Glioma	97.90	97.90	97.90	97.90
	Pituitary	100	100	100	100
	Average	**97.90**	**97.90**	**97.90**	**97.90**

Snapshot Ensemble

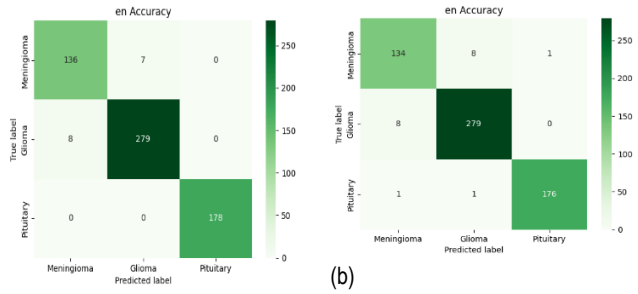
In this section, the weighted snapshot ensemble technique was applied to the model to enhance its performance. However, this technique was not considered in the earlier section 4.1. Adamax still demonstrated the best performance across all metrics during this phase. Despite this, the Adamax optimizer did not show any performance improvement with the weighted snapshot ensemble, unlike the Nadam and Adam optimizers, which showed notable improvements, as shown in Table 4.3.

Table 3: Results of the performances of individual optimizers with Snapshot Ensemble

Optimizer	Tumor type	Accuracy%	Precision%	Recall %	F1-score %
Nadam	Meningioma	95.10	95.10	94.40	94.80
	Glioma	97.20	97.20	97.60	97.40
	Pituitary	100	100	100	100
	Average	97.50	97.50	97.50	97.50
Adam	Meningioma	93.70	93.70	93.70	93.70
	Glioma	97.20	97.20	96.90	97.00
	Pituitary	99.80	98.90	99.40	99.20
	Average	96.90	96.90	96.90	96.90
Adamax	Meningioma	95.10	95.10	96.50	95.80
	Glioma	97.90	97.90	97.90	97.90
	Pituitary	100	100	100	100
	Average	**97.90**	**97.90**	**97.90**	**97.90**

.

Figure 3 presents the confusion matrices for each optimizer after applying weighted snapshot techniques. For the Nadam optimizer, the model correctly classified 136, 279, and 178 MRI slices for meningioma, glioma, and pituitary tumors, respectively, with 15 MRI slices misclassified: 7 as glioma and 8 as meningioma. For the Adam optimizer, the model accurately classified 134, 279, and 176 MRI slices for meningioma, glioma, and pituitary tumors, respectively, with 19 slices misclassified: 8 as glioma, 1 as pituitary, and 8 as meningioma. There was 1 glioma and 1 meningioma misclassification. Lastly, with the Adamax optimizer, the model correctly classified 136, 281, and 178 MRI slices as meningioma, glioma, and pituitary tumors, respectively, with only 13 misclassified slices: 5 as glioma, 6 as meningioma, and 1 each as glioma and meningioma.



(a) (b)

Figure 3: Confusion Matrix for Weighted Snapshot Ensemble for: (a) Nadam, (b) Adam, (c) Adamax

Weighted Hybridized Snapshot Ensemble

This section presents the results of the proposed model, where two snapshots were saved from each optimizer in Section 4.1.2 and hybridized with weight assignment to enhance performance. The weights of the hybridized model were optimized using the WOA. However, despite applying this technique, the model did not outperform the one trained solely with the Adamax optimizer. This is due to the nature of the weight assignment, which prevents lower-performing models from negatively impacting higher-performing ones.

The confusion matrix for the proposed model is shown in Fig. 4.2. The model successfully classified 136, 281, and 178 MRI slices for meningioma, glioma, and pituitary tumors, respectively, with only 13 MRI slices misclassified: 6 as glioma and 7 as meningioma. Other performance metrics, including accuracy, precision, and F1-score, are presented in Table 4.4. The model achieved prediction accuracies of 95.10% for meningioma, 97.90% for glioma, and 100% for pituitary tumors. Overall, the model demonstrated an average prediction accuracy of 97.9% on the Figshare dataset, with an average precision, recall, and F1-score of 97.9%, outperforming existing models. This result highlights the excellent performance of the Custom-CNN when combined with the Figshare dataset.

Science World Journal Vol. 21(No 1) 2026 https://dx.doi.org/10.4314/swj.v21i1.41
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

Table 4: Results of the Performance of the Proposed Model.

	Tumor type	Accuracy%	Precision%	Recall%	F1-score %
Proposed Model	Meningioma	95.10	95.10	95.80	95.40
	Glioma	97.90	97.90	97.90	97.90
	Pituitary	100	100	100	100
	Average	**97.90**	**97.90**	**97.90**	**97.90**

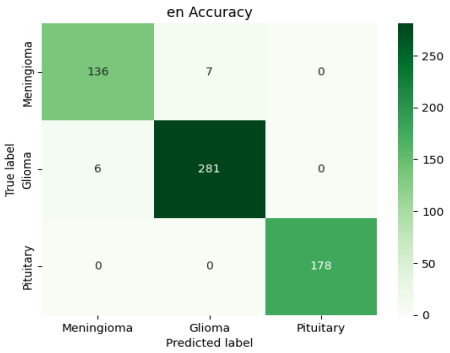
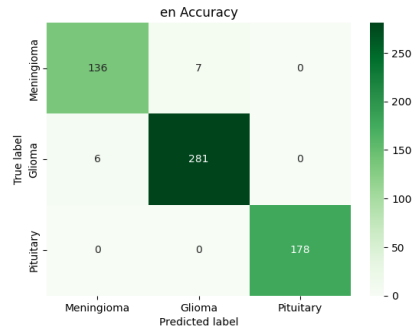
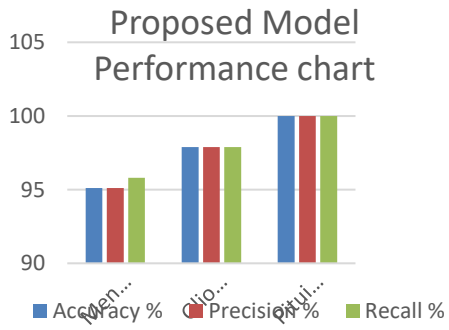


Figure 4: Confusion matrix Performance of the proposed Model.



Proposed Model Confusion Matrix



Proposed Model Performance Chart

Figure 5: Confusion matrix and Performance of the proposed Model chart

Table 5: Results of the performances of the hybridized Snapshot Ensemble with weights assignment.

Model	Tumor Type	Accuracy%	Precision%	Recall %	F1-score %
Proposed Model	Meningioma	95.10	95.10	95.80	95.40
	Glioma	97.90	97.90	97.90	97.90
	Pituitary	100	100	100	100
	Average	**97.90**	**97.90**	**97.90**	**97.90**

DISCUSSION

In this study, three optimizers, Adam, Adamax, and Nadam, were employed to diagnose multiclass brain tumors (meningioma, glioma, and pituitary) using the benchmark Figshare dataset. The proposed model was compared with existing models in the literature, all of which used the same dataset and tumor types but with different architectures. In this study, a custom CNN architecture was fine-tuned using specific hyperparameters to achieve optimal results. Leaky ReLU, which allows small non-zero outputs for negative inputs, was used as the activation function. A dynamic learning rate with an initial value of 0.001 was applied for all optimizers, along with a dropout rate of 0.25 to reduce overfitting, a batch size of 32, a maximum of 80 epochs, and categorical cross-entropy as the loss function. The proposed model demonstrated significantly improved performance compared to existing models, achieving accuracy, precision, recall, and F1-score values of 97.9%.

Proposed Model

This section presents the results of the proposed model, where two snapshots were saved from each optimizer in Section 4.3.2 and hybridized with weight assignment to enhance performance. The weights of the hybridized model were optimized using the WOA. However, despite applying this technique, the model did not outperform the one trained solely with the Adamax optimizer. This is due to the nature of the weight assignment, which prevents lower-performing models from negatively impacting higher-performing ones.

The confusion matrix for the proposed model is shown in Fig. 4. The model successfully classified 136, 281, and 178 MRI slices for meningioma, glioma, and pituitary tumors, respectively, with only 13 MRI slices misclassified: 6 as glioma and 7 as meningioma. Other performance metrics, including accuracy, precision, and F1-score, are presented in Table 4.4. The model achieved prediction accuracies of 95.10% for meningioma, 97.90% for glioma, and 100% for pituitary tumors. Overall, the model demonstrated an average prediction accuracy of 97.9% on the Figshare dataset, with an average precision, recall, and F1-score of 97.9%, outperforming existing models. This result highlights the excellent performance of the Custom-CNN when combined with the Figshare dataset.

Figure 6 shows the training and validation accuracy and loss curves for each optimizer used in the study. Both Nadam and Adam showed similar behavior across training and validation loss and accuracy. There were fluctuations in the validation loss curve, but the instability subsided after the 13th epoch, and the loss approached zero. In terms of accuracy, both Nadam and Adam showed initial fluctuations over the first 10–15 epochs, but as epochs increased, accuracy stabilized, indicating that the model successfully avoided overfitting. The final accuracy for Nadam and

Science World Journal Vol. 21(No 1) 2026 https://dx.doi.org/10.4314/swj.v21i1.41
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

Adam was 96.90% and 96.20%, respectively.
On the other hand, Adamax showed greater fluctuations in both the training and validation curves than Nadam and Adam. However, the instability in Adamax diminished after the 30th epoch, with the loss curve also approaching zero. Similar fluctuations were observed in the accuracy curve until the 30th epoch, after which accuracy remained consistent through the 80th epoch, further indicating that the model avoided overfitting. The Adamax optimizer achieved the highest accuracy at 97.90%.

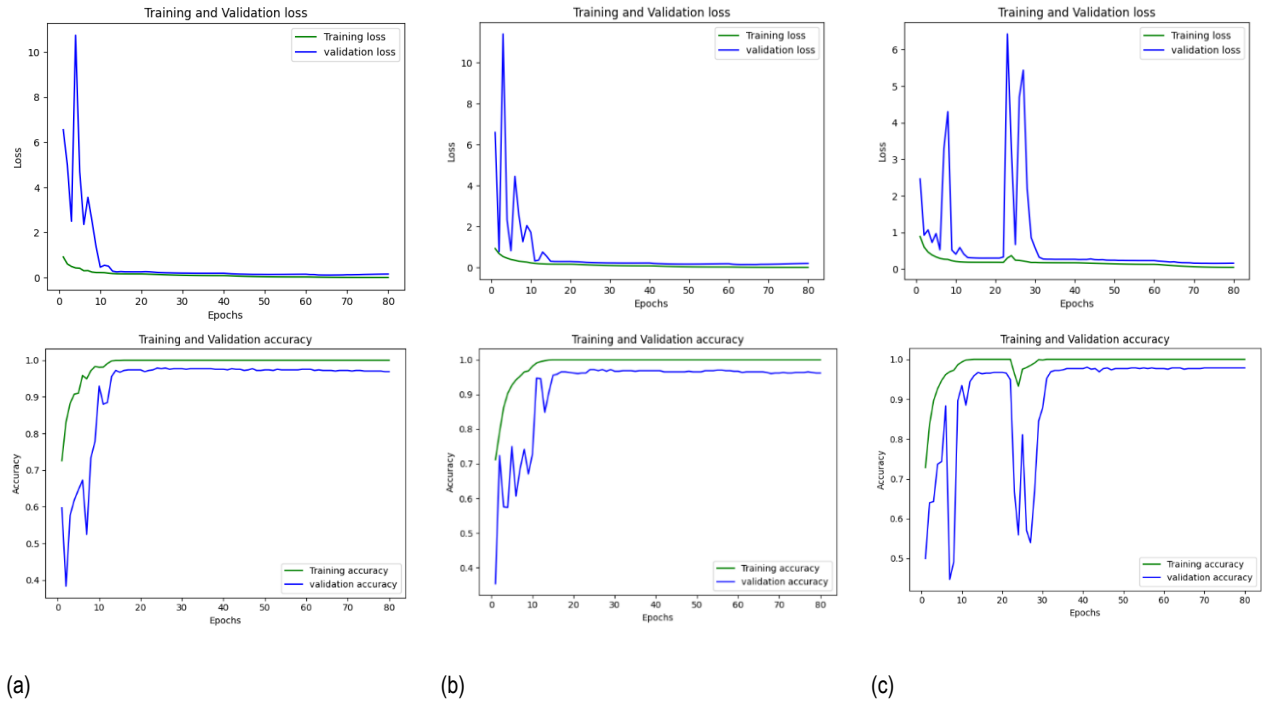


Figure 6: Training and Validation accuracy and loss for: (a) Nadam, (b) Adam, (c) Adamax

Future work

Future research should explore a wider range of optimization algorithms beyond Nadam, Adam, and Adamax, as well as alternative activation functions such as Swish or ELU to enhance model performance. Investigating heuristic algorithms such as Particle Swarm Optimization could lead to improved optimization. Incorporating advanced image preprocessing techniques, such as image enhancement and feature extraction, could reduce training time and improve efficiency.

REFERENCES

Adu, K., Yu, Y., Cai, J., Asare, I., & Quahin, J. (2022). The influence of the activation function in a capsule network for brain tumor type classification. *International Journal of Imaging Systems and Technology*, *32*(1), 123–143.

Alquran, H., Alqudah, A. M., Abu-Qasmieh, I., Al-Badarneh, A., & Almashaqbeh, S. (2019). ECG classification using higher-order spectral estimation and deep learning techniques. *Neural Network World*, *29*(4), 207–219.

Anaya-Isaza, A., Mera-Jiménez, L., Verdugo-Alejo, L., & Sarasti, L. (2023). Optimizing MRI-based brain tumor classification and detection using AI: A comparative analysis of neural networks, transfer learning, data augmentation, and the cross-transformer network. *European Journal of Radiology Open*, *10*, 100484

Bhalodiya, J. M., Lim Choi Keung, S. N., & Arvanitis, T. N. (2022). Magnetic resonance image-based brain tumour segmentation methods: A systematic review. *Digital Health*, *8*, 20552076221074122.

Chhabda, S., Carney, O., D'Arco, F., Jacques, T. S., & Mankad, K. (2016). The 2016 World Health Organization Classification of Tumours of the Central Nervous System: what the paediatric neuroradiologist needs to know. *Quantitative imaging in medicine and surgery*, *6*(5), 486.

Çinar, A., & Yildirim, M. (2020). Detection of tumors on brain MRI images using a hybrid convolutional neural network architecture: Medical Hypotheses, *139*, 109684.

Choi, H. J., Cho, B. C., Sohn, J. H., Shin, S. J., Kim, S. H., Kim, J. H., & Yoo, N. C. (2009). Brain metastases from hepatocellular carcinoma: prognostic factors and outcome: brain metastasis from HCC. *Journal of Neuro-Oncology*, *91*, 307–313.

Deepak, S., & Ameer, P. M. (2021). Automated categorization of brain tumors from MRI using CNN features and SVM. *Journal of Ambient Intelligence and Humanized Computing*, *12*(8), 8357-8369.102.

Dozat, T. (2016). Incorporating Nesterov momentum into Adam. Figshare dataset. URL: https://figshare.com/articles/braintumordataset/151242

Ghanavati, S., Li, J., Liu, T., Babyn, P. S., Doda, W., & Lampropoulos, G. (2012, May). Automatic brain tumor

detection in magnetic resonance images. In *2012 9th*

Gao, H., Yixuan, L., Geoff, P., Zhuang, L., John, H. E., & Kilian, W. Q. (2017). SNAPSHOT ENSEMBLES: TRAIN 1, GET M FOR FREE. *International Conference on Learning Representations*, (pp. 4–6). Toulon, France.

Gómez-Guzmán, M. A., Jiménez-Beristaín, L., García-Guerrero, E. E., López-Bonilla, O. R., Tamayo-Perez, U. J., Esqueda-Elizondo, J. J., ... & Inzunza-González, E. (2023). Classifying brain tumors on magnetic resonance imaging by using convolutional neural networks. *Electronics*, *12*(4), 955.

Ioffe, S., & Szegedy, C. (2015, June). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International Conference on machine learning* (pp. 448–456). pmlr.

Jalab, H. A., & Hasan, A. M. (2019). Magnetic resonance imaging segmentation techniques of brain tumors: A review. *Archives of Neuroscience*, *6*(Brain Mapping).

Kanchanamala, P., Revathi, K. G., & Ananth, M. B. J. (2023). Optimization-enabled hybrid deep learning for brain tumor detection and classification from MRI. *Biomedical Signal Processing and Control*, *84*, 104955.

Kasban, H., El-Bendary, M. A. M., & Salama, D. H. (2015). A comparative study of medical imaging techniques. *International Journal of Information Science and Intelligent Systems*, *4*(2), 37–58.

Kaur, T., & Gandhi, T. K. (2020). Deep convolutional neural networks with transfer learning for automated brain image classification. *Machine vision and applications*, *31*(3), 20.

Khan, M. S. I., Rahman, A., Debnath, T., Karim, M. R., Nasir, M. K., Band, S. S., ... & Dehzangi, I. (2022). Accurate brain tumor detection using a deep convolutional neural network. *Computational and Structural Biotechnology Journal*, *20*, 4733–4745.

Kingma, D. P. (2014). Adam: a method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical image analysis*, *42*, 60-88.

Louis, D. N., Ohgaki, H., Wiestler, O. D., Cavenee, W. K., Burger, P. C., Jouvet, A., ... & Kleihues, P. (2007). The 2007 WHO classification of tumours of the central nervous system. *Acta neuropathologica*, *114*, 97–109.

Luo, Q., Li, Y., Luo, L., & Diao, W. (2018). Comparisons of the accuracy of radiation diagnostic modalities in brain tumors: A nonrandomized, nonexperimental, cross-sectional trial. *Medicine*, *97*(31), e11256.

Malla, P. P., Sahu, S., & Alutaibi, A. I. (2023). Classification of tumors in brain MR images using a deep convolutional neural network and global average pooling. *Processes*, *11*(3), 679.

Meena, A., & Raja, R. (2013). Spatial fuzzy c-means pet image segmentation of neurodegenerative disorder. *arXiv preprint arXiv:1303.0647*.

Mengash, H. A., & Mahmoud, H. H. (2021). Brain cancer tumor classification from motion-corrected MRI images using a convolutional neural network. *Computers, Materials & Continua*, *68*(2), 1551–1563.

Mirjalili, S., & Lewis, A. (2016). The whale optimization algorithm. *Advances in engineering software*, *95*, 51–67.

Mohsen, H., El-Dahshan, E. S. A., El-Horbaty, E. S. M., & Salem, A. B. M. (2018). Classification using deep learning neural networks for brain tumors. *Future Computing and Informatics Journal*, *3*(1), 68–71.

Nayak, D. R., Padhy, N., Mallick, P. K., Zymbler, M., & Kumar, S. (2022). Brain tumor classification using dense EfficientNet. *Axioms*, *11*(1), 34.

Pathak, K., Pavthawala, M., Patel, N., Malek, D., Shah, V., & Vaidya, B. (2019, June). Classification of brain tumors using a convolutional neural network. In *2019, the 3rd International Conference on Electronics, Communication and Aerospace Technology (ICECA)* (pp. 128–132). IEEE.

Patil, S., & Kirange, D. (2023). An ensemble of deep learning models for brain tumor detection. *Procedia Computer Science*, *218*, 2468–2479.

Patil, S., Kirange, D., & Nemade, V. (2020). Predictive modelling of brain tumor detection using deep learning. *Journal of Critical Reviews*, *7*(04), 1805-1813.

Priya, V. V. (2016). An efficient segmentation approach for brain tumor detection in MRI. *Indian Journal of Science and Technology*

Rahman, T., & Islam, M. S. (2022, February). MRI brain tumor classification using a deep convolutional neural network. In *2022 International Conference on Innovations in Science, Engineering and Technology (ICISET)* (pp. 451–456). IEEE.

Scherer, D., Müller, A., & Behnke, S. (2010, September). Evaluation of pooling operations in convolutional architectures for object recognition. In *International Conference on Artificial Neural Networks* (pp. 92–101). Berlin, Heidelberg: Springer Berlin Heidelberg.

Selvaraj, H., Selvi, S. T., Selvathi, D., & Gewali, L. (2007). Brain MRI slice classification using a least-squares support vector machine. *International Journal of Intelligent Computing in Medical Sciences & Image Processing, 1*(1), 21–33.

Siegel, R. L., Miller, K. D., & Jemal, A. (2018). Cancer statistics, 2018. *CA: a cancer journal for clinicians*, *68*(1), 7–30.

Siegel, R. L., Miller, K. D., Fuchs, H. E., & Jemal, A. (2022). Cancer statistics, 2022. *CA: a cancer journal for clinicians*, *72*(1).

Singh, Y. P., & Lobiyal, D. K. (2024). A comparative analysis and classification of cancerous brain tumor detection based on classical machine learning and deep transfer learning models. *Multimedia Tools and Applications*, *83*(13), 39537–39562.

Sravan, V., Swaraja, K., Meenakshi, K., Kora, P., & Samson, M. (2020, June). Magnetic resonance images based brain tumor segmentation: a critical survey. In *2020, the 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184)* (pp. 1063–1068). IEEE.

Sumitra, N., & Saxena, R. (2013). Brain tumor classification using a back propagation neural network. *International Journal of Image, Graphics and Signal Processing, 5*,45–50.Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., et al. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis,42*, 60–88.

Tin, T. A., Aye, M. M., Khin, E. E., Oo, T., Tun, H. M., & Pradhan, D. (2024). Performance Optimization of Brain Tumor Detection and Classification Based on MRI by Using Batch Normalization Algorithms in Deep Convolution Neural Network. *Journal of Novel Engineering Science and Technology*, *3*(03), 66–72.

Xu, Y., Jia, Z., Ai, Y., Zhang, F., Lai, M., Eric, I., & Chang, C. (2015, April). Deep convolutional activation features for large-scale classification and segmentation of brain tumor histopathology images. In *2015, the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* (pp. 947–951). IEEE.

Xu, L., & Mohammadi, M. (2024). Brain tumor diagnosis from MRI based on Mobilenetv2 optimized by contracted fox optimization algorithm. *Heliyon*, *10*(1).

Yaqub, M., Feng, J., Zia, M. S., Arshid, K., Jia, K., Rehman, Z. U., & Mehmood, A. (2020). State-of-the-art CNN optimizer for brain tumor segmentation in magnetic resonance images. *Brain sciences*, *10*(7), 427.