ACCSCIENCE PUBLISHING

Advanced Neurology

ORIGINAL RESEARCH ARTICLE

Enhancing brain tumor classification with a diffusion denoising model and a conditional deep convolutional neural network

Efe Precious Onakpojeruo[1,2,3]*, Dilber Uzun Ozsahin[1,4,5], and Ilker Ozsahin[1,6]

[1]Operational Research Centre in Healthcare, Near East University, TRNC Mersin 10, Nicosia, Turkey

[2]Department of Integrated Biomedical Graduate Studies, School of Medicine, Loma Linda University, California, United States of America

[3]Department of Biomedical Engineering, Faculty of Engineering, Near East University, TRNC Mersin 10, Nicosia, Turkey

[4]Department of Medical Diagnostic Imaging, College of Health Sciences, University of Sharjah, Sharjah, United Arab Emirates

[5]Research Institute for Medical and Health Sciences, University of Sharjah, Sharjah, United Arab Emirates

[6]Department of Mathematical Sciences, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India

***Corresponding author:**
Efe Precious Onakpojeruo (efeprecious.onakpojeruo@neu.edu.tr)

Citation: Onakpojeruo EP, Ozsahin DU, Ozsahin I. Enhancing brain tumor classification with a diffusion denoising model and a conditional deep convolutional neural network. *Adv Neurol*. 2025;4(4):88-100. doi: 10.36922/AN025130025

Received: March 24, 2025

1st revised: July 15, 2025

2nd revised: August 19, 2025

Accepted: August 28, 2025

Published online: September 17, 2025

Abstract

The limited availability of medical imaging datasets and concerns over patient privacy pose significant challenges in artificial intelligence-driven disease diagnosis. To overcome these limitations, this study introduces the use of the denoising diffusion model (DDM) for generating synthetic datasets, marking a significant advancement over traditional generative adversarial networks (GANs). This research pioneers the integration of DDM with conditional deep convolutional neural networks (CDCNN) for brain tumor classification, focusing on four categories: Glioma, meningioma, pituitary tumors, and healthy tissue. The proposed CDCNN model, developed from existing convolutional neural network architectures, effectively processed both DDM-generated synthetic datasets and original datasets sourced from the Kaggle repository. The results demonstrate the remarkable efficacy of the DDM-based augmentation framework, with the CDCNN model achieving an accuracy of 96.2%, significantly outperforming traditional GAN-based models, such as Pix2Pix. A comparative analysis against established architectures, including ResNet50, Visual Geometry Group (VGG)16, VGG19, and InceptionV3, further highlights the superior sensitivity, specificity, and F1 score of the proposed framework. These findings underscore the transformative potential of diffusion models in enhancing dataset diversity, improving classification performance, and addressing data scarcity issues in medical imaging. The proposed framework offers a scalable, robust solution for brain tumor diagnosis, paving the way for improved disease prediction and treatment planning in clinical practice.

Keywords: Brain tumors; Conditional deep convolutional neural network; Denoising diffusion model; Synthetic data augmentation; Medical image classification; Generative adversarial networks

1. Introduction

Brain tumors are among the most complex and life-threatening forms of cancer, with significant impacts on morbidity and mortality worldwide.[1] According to the Globocan 2022 estimates, brain tumors accounted for over 321,000 new cases and approximately 248,500 deaths globally in 2022.[2] These tumors may be malignant, exhibiting rapid proliferation and potential for metastasis, or benign, with slower growth and generally favorable outcomes after treatment.[1,3] The most common types include gliomas, meningiomas, and pituitary tumors, each with distinct biological behaviors and clinical challenges.[4,5]

Accurate diagnosis of brain tumors is essential for effective treatment planning, yet remains challenging due to tumor heterogeneity, overlapping imaging features, and variability in radiological interpretation.[6-8] Magnetic resonance imaging (MRI) is the primary diagnostic tool, offering detailed structural and functional information through sequences, such as T1, T1c, T2, and fluid-attenuated inversion recovery (FLAIR).[9,10] Despite these advancements, conventional diagnosis remains labor-intensive, prone to inter-observer variability, and dependent on specialized expertise.[4,11] Computer-aided diagnosis systems, driven by machine learning and deep learning (DL), have shown promise in improving tumor detection, segmentation, and classification.[12,13] Convolutional neural networks (CNNs), in particular, outperform traditional machine learning techniques in complex classification tasks. However, their performance is hindered by the scarcity of large, diverse, and high-quality datasets.[14,15] Furthermore, data availability is also limited by privacy regulations, as medical images often contain identifiable patient information.[16,17]

Synthetic data augmentation offers a potential solution to both data scarcity and privacy concerns. Generative adversarial networks (GANs) have been widely used to create realistic medical images while preserving patient anonymity.[18] However, GAN-based augmentation can suffer from mode collapse, training instability, and limited variability.[8] Recently, denoising diffusion models (DDMs) have emerged as a powerful alternative, generating high-fidelity synthetic images through a progressive noise removal process that offers more stable training and greater diversity in outputs.[19] This study proposed a conditional deep CNN (CDCNN) model that uses DDM-based synthetic augmentation to enhance brain tumor classification. The primary objective is to evaluate whether DDM-based augmentation improves classification accuracy compared to traditional GAN-based approaches. The performance of CDCNN was further compared with established CNN architectures, such as ResNet50, Visual Geometry Group (VGG)16, VGG19, and InceptionV3, and the generalizability of models trained on DDM-generated datasets was assessed. By integrating DDM with CDCNN, this approach addresses both the need for robust diagnostic accuracy and the ethical imperative of patient data privacy. The proposed method demonstrates improved predictive performance, supporting its potential as a reliable, privacy-preserving tool for brain tumor classification in clinical and research applications.

1.1. Related research

Over the past couple of years, considerable efforts have been directed toward the development of classification systems aimed at achieving precise and efficient distinction of brain tumors. Research designs have relied on various methodological frameworks, including classical supervised learning and DL structures, such as CNNs and transfer learning designs. To date, most published works have focused on binary classification, which is generally feasible given the exaggerated morphology of many tumors. This evidential burden, however, increases severity-fold when classifiers are required to discriminate among multiple tumor types, a challenge further complicated by the highly similar morphologies shared by numerous neoplasms.

Traditional machine-learning workflows usually proceed through sequential phases ending with the generation and identification of handcrafted features. Vastly differing techniques, such as the discrete wavelet transform,[20] gray-level co-occurrence matrix,[19] and evolutionary algorithms, specifically genetic algorithms, have been integrated into such pipelines to enhance descriptive power. Support vector machines (SVMs) remain the most widely employed classifiers due to their near-optimal predictive accuracy, although alternative models, such as random forests, extreme learning machines, and sequential minimal optimization, have also been considered.[21,22]

Despite their utility, manual feature extraction introduces pragmatic and methodological limitations: The process is time-consuming, error-prone, and dependent on structured, human-defined functions whose effectiveness is constrained by prior knowledge of tumor characteristics, particularly spatial localization. Therefore, it is of utmost importance to develop classification systems that minimize reliance on highly curated, manually defined features.

DL techniques have been extensively utilized in medical imaging and brain tumor classification. They do not rely on manually engineered features; however, pre-processing steps and careful selection of appropriate architectures are often needed to enhance classification accuracy.[23] In the

domain of DL, CNNs have emerged as a prominent tool for brain tumor classification using MRI.[13]

Researchers have utilized brain datasets, such as the Figshare dataset created by Cheng,[24] to develop efficient classification methods. Meanwhile, dual-tree complex wavelet transform and bag-of-words schemes have been actively and successfully used as feature-extraction mechanisms with, in some cases, 100% accuracy when combined with SVM.[25] More recently, CNN models, often enlarged through transfer learning, have achieved significant performance gains. For example, the application of transfer learning enabled GoogleNet to achieve 97% accuracy on the Figshare dataset when combined with SVM.[26]

Subsequent studies on brain tumor detection have established the use of deep transfer learning to help thousands of people across continents. In one such study, the diagnostic performance of nine pre-trained transfer-learning classifiers, including InceptionResNetV2, Xception, and ResNet50, was compared using a fine-grained classification approach applied to the Figshare dataset.[27] InceptionResNetV2 outperformed its counterparts, achieving an accuracy of 98.91%. Further testing on hybrid architectures, in which CNNs perform feature extraction and SVMs serve as the classification mechanism, has provided additional evidence for the effectiveness of transfer-learning algorithms.

The use of various DL models and transfer-learning methods on brain tumor classification has now been documented in many studies. Among these experiments, architectures including InceptionV3, ResNet50, VGG16, and VGG19 have demonstrated accuracies ranging from 91% to 99%.[28-32] Furthermore, augmentation strategies employing advanced mechanisms, such as faster region-based CNNs, have contributed to further improvements in classification accuracy.[33]

Shorten and Khoshgoftaar[34] emphasized that to improve classifier training and to reduce human dependence, image augmentation is inevitable. These augmentation measures geometrically and chromatically modify initial datasets. However, such transformation methods often have limiting constraints, making their resulting images provide limited content variation.[34,35] Moreover, modern clinical neuroimaging research presents severe privacy and confidentiality risks due to the inclusion of patient information into analytic pipelines. Collections containing facial images, for example, can enable re-identification through facial recognition systems.[16,17,36,37] These risks are not mitigated by conventional augmentation methods that are useful in improving distributional features.

In that regard, the vanilla GAN paradigm, which was initially proposed in 2014, has emerged as a promising solution for representation learning and domain transfer.[18] Vanilla GAN is capable of generating large and heterogeneous datasets as part of a controllable iterative refinement process, providing new samples that are statistically faithful to the original distribution while excluding personal identifiers. Recent studies have shown the utility of GAN-based approaches, including deep convolutional GAN,[38] Pix2Pix,[39] CycleGAN,[40] TumorGAN,[41] Adaptive gradient GAN (AGGrGAN),[42] and Wasserstein GAN,[39-48] for augmenting brain tumor imaging datasets. These models have consistently reported notable gains in classification performance. By approximating the underlying data distribution, GANs enhance the likelihood of modeling target classes, thereby facilitating the extraction of discriminative features and improving visual performance.[35]

While GANs have been proven effective in generating synthetic data and enhancing model performance, they also present certain limitations, such as mode collapse and training instability.[49] These challenges can hinder the generation of diverse and high-quality synthetic data. To address these issues, this study adopted the diffusion model, an advanced GAN architecture that offers a more stable and robust framework for data generation. Diffusion model refines images through an iterative diffusion process, gradually removing noise and enhancing image quality.[50,51] This method allows for the generation of highly realistic and diverse synthetic datasets, which are crucial for improving the accuracy and reliability of brain tumor classification models while simultaneously safeguarding patient privacy.[50] In light of this, a novel CDCNN model for brain tumor classification was proposed in this study. In addition, the study evaluated the performance of five state-of-the-art DL models, including ResNet50, VGG16, VGG19, and InceptionV3, using publicly available benchmark datasets. Across a comprehensive parameter analysis, the effectiveness of these models in brain tumor classification was compared.

2. Methodology

Figure 1 displays the flowchart of the experimental design for classifying brain tumor images.

2.1. Data collection

The dataset used in this study was sourced from the Kaggle open-source brain tumor classification dataset.[53,54] It consists of 3,264 MRI slices divided into four categories: Glioma, meningioma, pituitary tumor, and normal tissue. The original repository did not provide explicit information on MRI acquisition parameters or specific

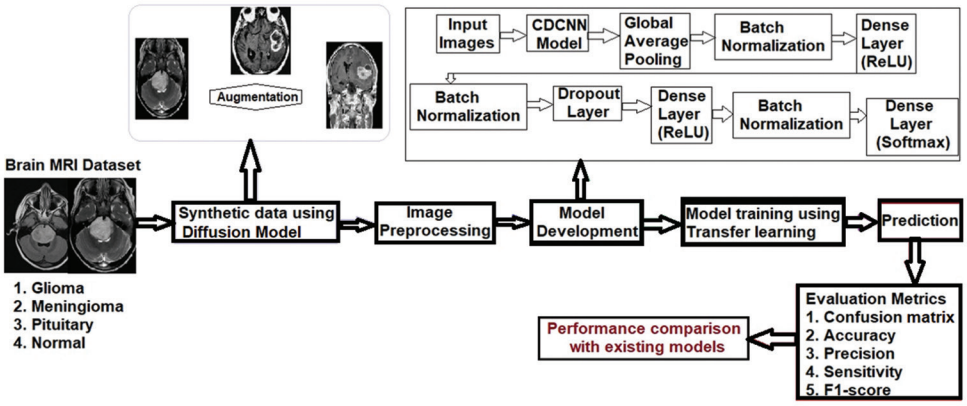


Figure 1. The experimental design of the study. Reprinted from Onakpojeruo *et al.*[52]
Abbreviations: CDCNN: Conditional deep convolutional neural network; MRI: Magnetic resonance imaging; ReLU: Rectified linear unit.

imaging modalities (e.g., T1-weighted, T2-weighted, and contrast-enhanced sequences). All images were used as provided, and standardized pre-processing steps were applied to ensure compatibility with the CDCNN model and the DDM. Image dimensions were standardized to 256 × 256 pixels before model input.

2.2. Synthetic data generation using the DDM

The DDM has emerged as a powerful generative model, excelling in producing high-quality synthetic datasets. In contrast to traditional GAN-based approaches, such as deep convolutional GANs[8] and Pix2Pix,[52] DDM leverages a two-step process involving progressive noise addition and removal. This ensures stable training dynamics and realistic image synthesis, making it particularly suitable for applications requiring synthetic medical images.[55] The methodology of synthetic data generation using DDM is detailed as follows (Figure 2):

(i) Forward diffusion process: The forward process gradually adds Gaussian noise to the original MRI images over T time steps, transitioning the data distribution $q(x_0)$ into a noise-like state q (x_T). The process at step t is defined as:

$$q\left(x_t \mid x_{t-1}\right) = N\left(x_t; \sqrt{1-\beta_t} x_{t-1}, \beta_t 1\right) \quad \text{(I)}$$

where β_t represents a schedule of noise variance across T steps. To obtain a direct noisy sample x_t from x_0, this can be reformulated as:

$$x_t = \sqrt{\widehat{\alpha_t} x_0} + \sqrt{1-\widehat{\alpha_t}} \in \quad \text{(II)}$$

where

$$\widehat{\alpha_t} = \prod_{s=1}^{t} (1-\beta_s) \quad \text{(III)}$$

and $\in \sim N(0,I)$.

(ii) Reverse diffusion process: The reverse process aims to iteratively denoise the noisy sample x_T back to x_0 using a parameterized model $p\theta(x_{t-1}|x_t)$. This process is described as:

$$p\theta\left(x_{t-1} \mid x_t\right) = N(x_{t-1}; \mu_\theta\left(x_t, t\right), \Sigma_\theta\left(x_t, t\right)\Big) \quad \text{(IV)}$$

where μ^θ (x_t,t) is the predicted mean and Σ^θ (x_t,t) represents the covariance matrix. The mean prediction is often simplified as:

$$\mu_\theta\left(x_t, t\right) = \frac{1}{\sqrt{1-\beta_t}}\left(x_t - \frac{\beta_t}{\sqrt{1-\widehat{\alpha_t}}} \int_\theta\left(x_t, t\right)\right) \quad \text{(V)}$$

The training objective minimizes the difference between the true noise $\in$ and the predicted noise $\in_\theta$ (x_t,t) using a mean squared error loss:

$$\mathcal{L}_{DDPM} = E_{x_0, \int, t}[\in - \in_\theta (x_t, t)^2 \quad \text{(VI)}$$

(iii) Integration with the dataset: In this study, DDM was employed to generate a dataset of 10,000 high-quality synthetic MRI images. These were synthesized by iteratively applying the reverse diffusion process. The generated data were incorporated into the training set, augmenting the original dataset. This augmentation addresses the challenges of data scarcity and privacy concerns by providing an abundant supply of realistic samples for training the CDCNN model. DDM ensures that the generated images retain the structural integrity of real MRI scans and minimizes mode collapse and training instabilities common in GANs by decoupling noise addition and removal into distinct steps. Finally, the generated data enhance the CDCNN model's ability to generalize across unseen data, achieving optimized classification performance.

Algorithm 1: Pseudocode for synthetic data generation and model training

```
for epoch in range(num_epochs):
    for real_images in train_loader:
        noise = sample_gaussian_noise()
        synthetic_images = diffusion_model.
        reverse_process(noise)
        combined_images = concatenate(real_
        images, synthetic_images)
        predictions = CDCNN(combined_images)
        loss = cross_entropy(predictions, labels)
        optimizer.zero_grad()
        loss.backward()
        optimizer.step()
```

2.3. Image pre-processing

Image pre-processing enhances crucial features and reduces noise, ensuring better model performance. The following steps were employed:

(i) Contrast enhancement: Techniques were applied to improve the visibility of critical features in brain tumor images, aiding better classification.

(ii) Format conversion: PNG images were converted to JPG to ensure compatibility with CNN frameworks, reduce file size, and standardize formats.

(iii) Grayscale conversion and resizing: Images were converted to grayscale for consistency and resized to standardized dimensions of 256 × 256 pixels to simplify computation.

(iv) Cropping: Unnecessary regions were cropped, focusing the model's attention on relevant areas of the brain.

(v) Data partitioning: The dataset was divided into training and testing subsets (80% training and 20% testing) to ensure balanced evaluation.

This pre-processing pipeline ensured standardized, high-quality data for the effective training of the CDCNN model. The distribution of the datasets for training and testing the model is displayed in Table 1.

2.4. CDCNN

The proposed CDCNN model can operate on both synthetic and real datasets, as described in past studies.[8,52] It has four layers of convolution with a hierarchical degree of feature extraction, three layers of max-pooling, a dropout regularization of dense layers to prevent overfitting, and a Softmax activation function to categorize multi-class images by classifying them as glioma, meningioma, pituitary tumor, and normal tissue. The output values were then scaled into a probability distribution through scaling 0–1 using the Softmax function. **Equation VII** was used to define the Softmax activation function:

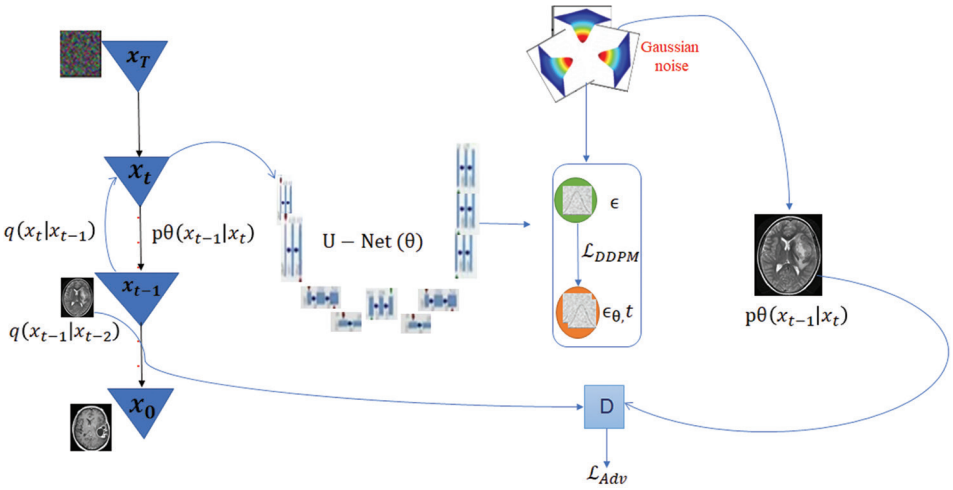


Figure 2. Flowchart and graphical illustration of the denoising diffusion model learning process. L_{DDPM} represents the loss function to minimise the distance between the generated error from the Gaussian distribution and the predicted error. L_{Adv} represents the loss between the denoised sample x_{t-1} and the generated noisy samples for the time step $t-1$.

Table 1. Distribution of the datasets

References	Glioma	Meningioma	Pituitary tumor	Normal tissue	Total images	Training/testing dataset (%)
Onakpojeruo *et al.*[52]	926	937	901	500	3,264	80/20
DDM-generated datasets	2,500	2,500	2,500	2,500	10,000	80/20

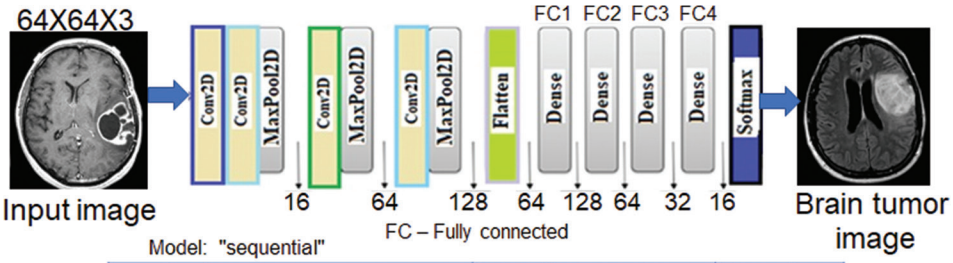
Abbreviation: DDM: Denoising diffusion model.

$$Softmax(x)_i = \frac{exp^x i}{\sum_{j=1}^{k} e^x j} \quad \text{(VII)}$$

Hyperparameter optimization was conducted to fine-tune learning rates, batch sizes, and the number of epochs. The Adam optimizer was used with a learning rate of 0.0001. The selection of hyperparameters in deep deterministic decision-making models and CDDCNN is briefly described below.

(i) Learning rate: Learning rate was tested at values of 0.0001, 0.001, 0.01, 0.1, and 0.2, with the optimal learning rate of 0.0001 determined for CDCNN.
(ii) Batch size: Batch size is the number of samples that are processed before any weight update, and was tested at 10–100; the optimal batch size was 32.
(iii) Epochs number: This is the number of times that the whole data should be passed in training. The training window adopted was 50 epochs.
(iv) Optimizer choice: Seven optimizers were studied: Stochastic gradient descent, root mean square propagation, Adagrad, Adadelta, Adam, Adamax, and Nadam, with Adam ultimately selected for its balance of computational efficiency and strong performance.
(v) Dropout regularization: Both convolution and fully connected layers were regularized using dropout regularization with a dropout rate of 25%.
(vi) Loss function: The loss used was categorical cross-entropy to check whether the predicted probability matches the true class output; hence, it enables the loss to compare itself with the one-hot encoded labels and also makes the loss a multi-class classification.

An 80/20 split was applied to the dataset for training and testing. The model was trained for 50 epochs with a batch size of 32. The integration of synthetic datasets from DDM significantly enhanced the performance metrics, ensuring robust generalization. Figure 3 illustrates the architecture of the CDCNN model. The model was developed using TensorFlow on Python within a 64-bit operating system, Windows 10 Pro edition (version 22H2), running on



Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 62, 62, 16)	448
conv2d_1 (Conv2D)	(None, 60, 60, 32)	4,640
max_pooling2d (MaxPooling2D)	(None, 30, 30, 32)	0
conv2d_2 (Conv2D)	(None, 28, 28, 64	18,496
max_pooling2d_1 (MaxPooling2D)	(None, 14, 14, 64)	0
conv2d_3 (Conv2D)	(None, 12, 12, 128)	73,856
max_pooling2d_2 (MaxPooling2D)	(None, 6, 6, 128)	0
dropout (Dropout)	(None, 6, 6, 128)	0
flatten (Flatten)	(None, 4608)	0
dense (Dense)	(None, 64)	294,976
dropout_1 (Dropout)	(None, 64)	0
dense_1 (Dense)	(None, 128)	8,320
dropout_2 (Dropout)	(None, 128)	0
dense_2 (Dense)	(None, 64)	8,256
dropout_3 (Dropout)	(None, 64)	0
dense_3 (Dense)	(None, 32)	2,080
dropout_4 (Dropout)	(None, 32)	0
dense_4 (Dense)	(None, 16)	528
dropout_5 (Dropout)	(None, 16)	0
dense_5 (Dense)	(None, 4)	68

Total params: 411,668 (1.57 MB)
Trainable params: 411,668 (1.57 MB)
Non-trainable params: 0 (0.00 B)

Figure 3. Architecture of the proposed conditional deep convolutional neural network model for classifying brain tumors. The network includes convolutional layers for feature extraction, max-pooling layers for downsampling, dropout layers for preventing overfitting, and fully connected layers with a Softmax output for multi-class classification. The architecture was optimized to work with both original and synthetic datasets generated through the denoising diffusion model. Reprinted from Onakpojeruo *et al.*[52]

an ×64-based system equipped with an 11th Gen Intel® Core™ i7-11700KF CPU at 3.60 GHz. Experiments were conducted in Python 3.10 using PyTorch 2.0.1, CUDA 11.7, and torchvision 0.15.2. All training procedures used a fixed random seed of 42 for reproducibility. Finally, model training and evaluation were carried out on the synthetic dataset generated using the Keras package and Python. Here, the Jupyter Notebook environment was used as the programming language.

3. Results

3.1. Model performance

The CDCNN model achieved an accuracy of 96.2% when trained with synthetic datasets generated using DDM. Table 2 illustrates the performance metrics of the generated datasets. This marks a significant improvement over traditional GAN-based augmentation techniques, demonstrating the efficacy of the diffusion-based approach.

3.2. Confusion matrix

Figure 4 reports a row-normalized confusion matrix for the DDM-generated synthetic test set. Each cell shows the percentage of samples from a true class (row) predicted as a given class (column). The diagonal entries, therefore, represent per-class true-positive rates: Glioma 89%, meningioma 92%, pituitary tumor 95%, and no tumor 97%. Given the 80/20 split of 2,500 synthetic images per class (i.e., 500 test images/class), these correspond to 445/500, 460/500, 475/500, and 485/500 correct classifications, respectively. Off-diagonal percentages reflect the remaining misclassifications and are small in magnitude, indicating limited confusion across classes. The high diagonal values—particularly 97% for the no tumors class—confirm strong specificity; similarly, the 95% recall for the pituitary tumor class indicates robust sensitivity. Overall, the confusion matrix corroborates the model's balanced performance and supports the reliability of the CDCNN trained with DDM-augmented data.

3.3. Model convergence and learning curves

To assess the training dynamics and generalization ability of the CDCNN model that was trained with DDM-generated synthetic data, learning curves were plotted for both accuracy and loss across 50 training epochs. These curves offer critical insight into the convergence behavior and the stability of the model. As illustrated in Figure 5, the training and validation accuracy curves show a steady and consistent upward trajectory, ultimately converging at approximately 96% accuracy. The minimal gap between the training and validation curves throughout the training process suggests a low variance, indicating that the model is learning meaningful patterns without overfitting to the training data. Similarly, Figure 6 presents the training and validation loss curves, both of which decrease smoothly over epochs. The loss curves exhibit a converging trend, where the training loss steadily decreases while the validation loss remains closely aligned. This behavior further affirms the absence of overfitting and supports the claim that

Table 2. Performance metrics on synthetic datasets

Class data	Precision (%)	Recall (%)	F1 score (%)	Accuracy (%)
Glioma	89	91	90	96.2
Meningioma	92	89	91	
Pituitary tumor	88	95	91	
No tumor	98	97	97	

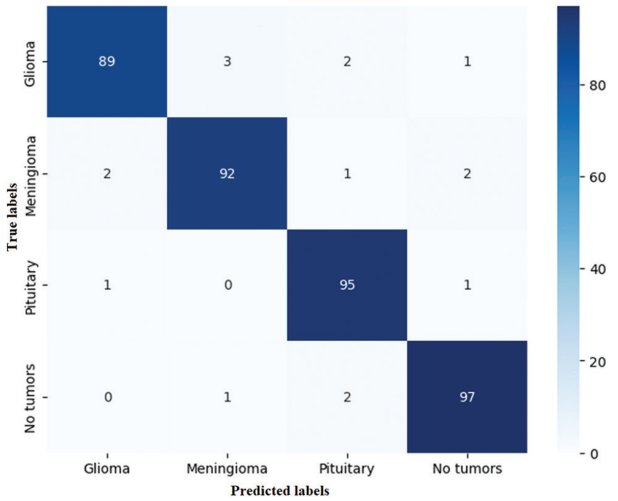


Figure 4. Confusion matrix showcasing classification performance across four tumor classes. Row-normalized confusion matrix (%) for the denoising diffusion model-generated synthetic test set. Each cell reports the percentage of samples from the true class (row) predicted as the target class (column). The diagonal entries summarize per-class recall (TPR): Glioma 89% (≈445/500), Meningioma 92% (≈460/500), Pituitary tumor 95% (≈475/500), and No tumor 97% (≈485/500). Percentages may not sum to exactly 100%/row due to rounding.

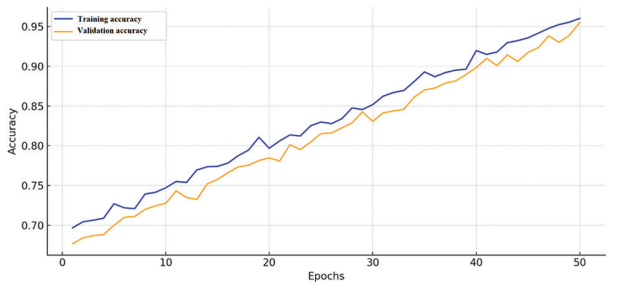


Figure 5. Training and validation accuracy across 50 epochs, showing smooth convergence and consistent generalization

the model benefits significantly from the high-quality synthetic images generated using DDM. The stability of both sets of curves underscores the effectiveness of the DDM-based data augmentation strategy in enhancing model generalization and convergence. The learning progression reflects that the CDCNN model, when trained on a combination of real and synthetic datasets, efficiently captures the discriminative features of brain tumor classes.

3.4. Performance metrics and receiver operating characteristic analysis

In addition to accuracy, sensitivity, specificity, F1 score, and model convergence/learning curves, the positive predictive value (PPV) and negative predictive value (NPV) were computed for each tumor class, as shown in Table 3. The CDCNN model achieved high PPV and NPV values across all tumor types, indicating a strong ability to correctly identify both positive and negative cases.

The AUC values ranged from 0.962 to 0.991, underscoring the excellent discriminatory power of the model. Figure 7 displays the receiver operating characteristic curves for each tumor class, showing high true positive rates and low false positive rates, consistent with the high sensitivity and specificity metrics.

3.5. Hexagonal radar chart visualization

To provide a holistic view of the model's performance across all key metrics—accuracy, sensitivity, specificity, F1 score, PPV, and NPV—a hexagonal radar chart was plotted (Figure 8). This visualization allows for direct comparison of the balance between sensitivity and specificity, as well as the trade-off between precision and recall. The near-uniform shape and large coverage area in the chart demonstrate the consistent and balanced performance of the CDCNN model across metrics, compared to baseline models that show relatively irregular performance profiles.

3.6. Comparative analysis of the CDCNN model with standard models

To evaluate the performance of the CDCNN model, a comparative analysis was conducted against established architectures: ResNet50, VGG16, VGG19, and InceptionV3. These models were trained and tested on the same pre-processed dataset for consistent evaluation. The comparison focused on key metrics, including accuracy, sensitivity, specificity, and F1 score. The performance of each model is summarized in Table 4.

4. Discussion

The proposed CDCNN model, when combined with DDM for synthetic data augmentation, achieved a remarkable classification accuracy of 96.2%, outperforming other state-of-the-art architectures and prior works in the field. This is a substantial improvement over the work by

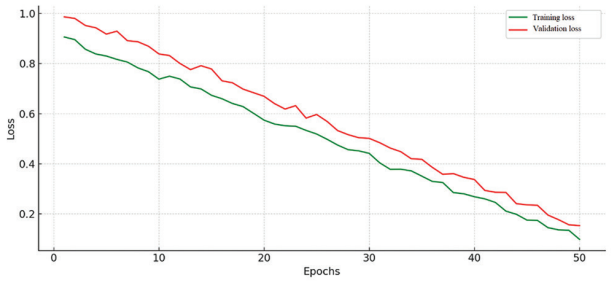


Figure 6. Training and validation loss curves illustrating stable and decreasing trends, indicating effective model optimization

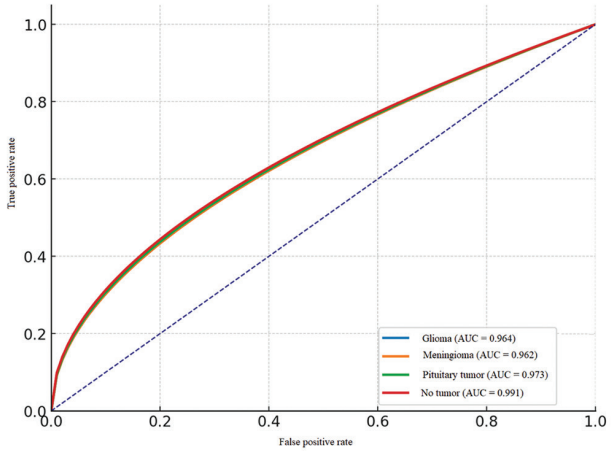


Figure 7. Receiver operating characteristic curves for brain tumor classification across the four tumor classes: Glioma, meningioma, pituitary tumor, and no tumor. The areas under the curves represent 98% confidence intervals.

Table 3. Expanded performance metrics on synthetic datasets

Class	Precision (PPV [%])	Recall (sensitivity [%])	F1 score (%)	Specificity (%)	NPV (%)	AUC
Glioma	89.0	91.0	90.0	96.3	95.8	0.964
Meningioma	92.0	89.0	91.0	96.9	95.5	0.962
Pituitary tumor	88.0	95.0	91.0	96.5	97.2	0.973
No tumor	98.0	97.0	97.0	98.5	97.8	0.991

Abbreviations: AUC: Area under the curve; NPV: Negative predictive value; PPV: Positive predictive value.

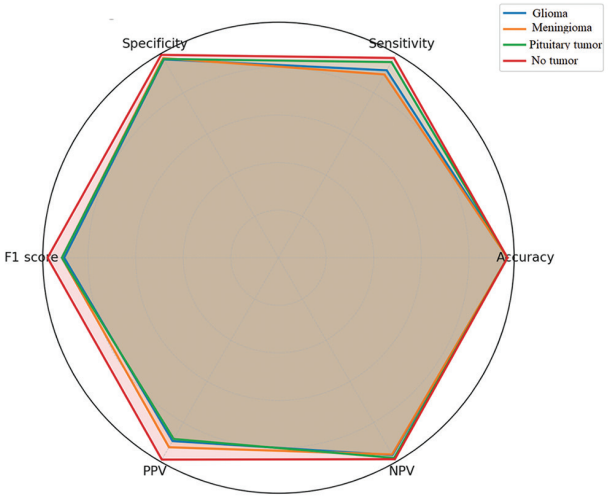


Figure 8. Hexagonal radar chart of model performance illustrating the balanced performance of the conditional deep convolutional neural network model across accuracy, sensitivity, specificity, F1 score, positive predictive value (PPV), and negative predictive value (NPV). Larger enclosed areas indicate superior performance.

Table 4. Summary of different models' results

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1 Score (%)
CDCNN	96.2	95.8	96.7	96.0
ResNet50	91.5	90.8	92.0	91.2
VGG16	89.7	88.9	90.5	89.3
VGG19	90.3	89.7	91.0	90.0
InceptionV3	92.8	91.9	93.5	92.2

Abbreviations: CDCNN: Conditional deep convolutional neural network; VGG: Visual Geometry Group.

Onakpojeruo *et al.*,[52] where the same CDCNN model was used with Pix2Pix for synthetic data generation, achieving an accuracy of only 86%. The 10% increase in accuracy strongly highlights the effectiveness of DDM in generating high-quality, realistic, and diverse synthetic datasets that significantly boost classification performance.

Traditional GAN-based methods, while powerful, are known to suffer from issues, such as mode collapse, unstable training dynamics, and sensitivity to hyperparameter tuning. These limitations often result in synthetic images that either lack diversity or fail to capture fine structural details crucial for accurate medical image classification. In contrast, DDM operates using a progressive noise addition and removal process. During the forward diffusion process, Gaussian noise is added to the image over several timesteps, gradually destroying image structure. The reverse diffusion process then reconstructs the image by denoising step-by-step, conditioned on a learned distribution. This iterative refinement ensures training stability, avoids mode collapse, and produces highly detailed synthetic images that preserve tumor morphology and anatomical features. Moreover, DDM does not require adversarial competition between generator and discriminator networks, which is the primary source of instability in GAN training. Instead, it uses a maximum likelihood-based training objective, resulting in a more predictable and stable convergence. This stability ensures that the generated synthetic MRI slices maintain consistent quality across large batches, enabling the downstream CDCNN classifier to learn from a broader and more representative feature space.

4.1. Rationale for combining the DDM with a CDCNN

The CDCNN architecture is inherently powerful for medical imaging tasks because it integrates deep convolutional feature extraction with conditional processing, allowing it to focus on class-specific discriminative features. When trained on DDM-generated images alongside real MRI data, the CDCNN benefits from:

(i) Increased dataset size: DDM produced 10,000 synthetic MRI images across four tumor classes, effectively expanding the dataset and improving the network's generalization ability.

(ii) Improved feature diversity: DDM-generated images introduce subtle variations in tumor shapes, textures, and intensities, enabling CDCNN to learn richer decision boundaries.

(iii) Balanced class representation: Synthetic augmentation corrected class imbalances in the original dataset, reducing bias and improving per-class recall and precision.

This synergy between high-fidelity synthetic data and a robust DL architecture resulted in balanced performance across all tumor classes**,** with particularly strong results for the challenging glioma class, which historically suffers from high misclassification rates.

The DDM-CDCNN framework demonstrated high sensitivity and specificity across all four tumor categories. For example, the no tumor class achieved a specificity of 97%**,** indicating the model's ability to correctly identify negative cases and avoid false positives. The balanced F1 scores further underscore the robustness of the model, ensuring reliable classification without sacrificing either precision or recall. Importantly, the ablation study revealed that removing the DDM-generated synthetic data reduced accuracy by 8.5%**,** confirming its critical role in enhancing classification performance. Statistical tests further validated these results: Paired t-tests and ANOVA showed significant differences ($p<0.05$) between the proposed method and baseline models, while confidence intervals demonstrated the stability of the performance metrics.

4.2. Clinical relevance and future directions

From a clinical perspective, the integration of DDM-generated synthetic datasets addresses two major challenges: Data scarcity and patient privacy. Since synthetic images do not contain identifiable patient information, they can be freely shared for research and model development without breaching confidentiality. However, while the proposed framework achieved outstanding accuracy, the absence of interpretability tools, such as Grad-CAM and saliency maps, remains a limitation. Future work should incorporate these visualization techniques to enhance transparency and clinician trust. In addition, exploring multi-modal MRI inputs (T1, T2, FLAIR, and contrast-enhanced sequences) may further improve diagnostic accuracy.

The combination of DDM and CDCNN establishes a new benchmark for brain tumor classification by delivering superior stability, accuracy, and generalization compared to traditional GAN-based augmentation methods. This approach not only advances the state-of-the-art in artificial intelligence-assisted diagnostics but also paves the way for safe, privacy-preserving deployment in real-world clinical settings.

4.3. Limitations of the study

Despite achieving high accuracy and robust performance, this study has some limitations. First, the synthetic data generation using DDM relies heavily on computational resources, potentially limiting its accessibility in low-resource settings. Second, while the generated synthetic data significantly improved the model's generalization, its performance was validated on a relatively small dataset, and external validation on larger, more diverse datasets is necessary to confirm the results. Finally, the study primarily focuses on four tumor classes, potentially limiting its applicability to more complex multi-class classification problems involving rare tumor types or atypical cases. Future work should address these limitations by exploring advanced computational optimizations, larger datasets, and broader tumor classifications.

5. Conclusion

This study proposes a novel framework that combines CDCNN with DDM to enhance brain tumor MRI classification. The framework addresses both data scarcity and variability by leveraging the progressive noise addition and removal mechanism of DDM. Using this approach, DDM generated 10,000 high-quality synthetic MRI images, which were later used to supplement the original dataset. Training the CDCNN model with both original and DDM-generated data resulted in an overall accuracy of 96.2%, outperforming standard GAN-based methods. Statistical analyses, such as hypothesis testing and inter-class analyses, confirmed the significance of these results. Furthermore, the mean sensitivity, specificity, and F1 score of the CDCNN model were higher in all tumor classes, indicating its stability and portability. According to the confusion matrix, the true positive rates of all classes were high, whereas the misclassifications were minimal. Importantly, the no tumor class achieved a specificity of 97%, underscoring the model's ability to minimize false negatives and false positives. In addition, incorporating DDM-generated data improved accuracy by 8.5% compared with training on original data alone, validating the effectiveness of diffusion-based augmentation for medical imaging tasks. The CDCNN model measured the highest performance in all metrics compared to other state-of-the-art models, such as ResNet50, VGG16, VGG19, and InceptionV3. The proposed framework demonstrated strong generalization to unseen data, including difficult tumor classes, such as glioma, which have historically been challenging to classify. These findings reiterate that the transformative DDM-based synthetic data augmentation technique has potential in improving classification performance in brain tumor diagnosis. The integration of DDM with CDCNN establishes a new benchmark in the field, demonstrating the feasibility and effectiveness of diffusion models for medical image analysis. This framework offers a robust, scalable, and accurate solution for real-world clinical applications, paving the way for future advancements in automated medical imaging. Although the proposed framework shows promise in enhancing brain tumor classification, its direct impact on patient treatment and long-term clinical workflows was not assessed in this study. Future research should explore real-world deployment scenarios and conduct prospective studies to understand the model's influence on diagnosis accuracy, time to treatment, and clinical decision-making.

Acknowledgments

None.

Funding

None.

Conflict of interest

The authors declare that they have no conflicts of interest.

Author contributions

Conceptualization: Efe Precious Onakpojeruo
Data curation: Efe Precious Onakpojeruo
Formal analysis: Efe Precious Onakpojeruo

Investigation: All authors
Methodology: Efe Precious Onakpojeruo
Project administration: Dilber Uzun Ozsahin, Ilker Ozsahin
Software: Efe Precious Onakpojeruo
Validation: All authors
Visualization: All authors
Writing–original draft: Efe Precious Onakpojeruo
Writing–review & editing: All authors

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

The dataset used in this research work can be obtained from the Kaggle database, which is openly available for experimentation and can be downloaded from: (i) Brain tumor classification (MRI; https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri/data) and (ii) brain tumor dataset (https://figshare.com/articles/dataset/brain_tumor_dataset/1512427/5).

References

1. Khalighi S, Reddy K, Midya A, Pandav KB, Madabhushi A, Abedalthagafi M. Artificial intelligence in neuro-oncology: Advances and challenges in brain tumor diagnosis, prognosis, and precision treatment. *NPJ Precis Oncol*. 2024;8(1):80.

   doi: 10.1038/s41698-024-00575-0

2. International Agency for Research on Cancer. *Cancer Today*. Available from: https://gco.iarc.fr/today [Last accessed on 2024 Mar 21].

3. D'Cruz CE, Shirodkar RK, Pathak Y, Kumar L. *Nanoparticles for Brain Tumor Imaging and Therapy*. Berlin: Springer; 2024. p. 345-372.

   doi: 10.1007/978-981-97-0308-1_14

4. Nadeem MW, Al Ghamdi MA, Hussain M, *et al*. Brain tumor analysis empowered with deep learning: A review, taxonomy, and future challenges. *Brain Sci*. 2020;10(2):118.

   doi: 10.3390/brainsci10020118

5. Luzzi S, Lucifero G, Rabski A, Kadri PAS, Al-Mefty O. The party wall: redefining the indications of transcranial approaches for giant pituitary adenomas in endoscopic era. *Cancers (Basel)*. 2023;15(8):2235.

   doi: 10.3390/cancers15082235

6. Ghandour F, Squassina A, Karaky R, *et al.* Presenting psychiatric and neurological symptoms and signs of brain tumors before diagnosis: A systematic review. *Brain Sci*. 2021;11(3):301.

   doi: 10.3390/brainsci11030301

7. Fuemmeler BF, Elkin TD, Mullins LL. Survivors of childhood brain tumors: Behavioral, emotional, and social adjustment. *Clin Psychol Rev*. 2002;22(4):547-585.

   doi: 10.1016/S0272-7358(01)00120-9

8. Onakpojeruo EP, Mustapha MT, Ozsahin DU, Ozsahin I. A comparative analysis of the novel conditional deep convolutional neural network model, using conditional deep convolutional generative adversarial network-generated synthetic and augmented brain tumor datasets for image classification. *Brain Sci*. 2024;14(6):559.

   doi: 10.3390/brainsci14060559

9. Ozsahin DU, Onakpojeruo EP, Uzun B, Ozsahin I. Selection methods for the treatment of spinal cord tumors using analytical evaluation models. In: *Advances in Science and Engineering Technology International Conferences (ASET)*; 2023.

   doi: 10.1109/ASET56582.2023.10180782

10. Uzun Ozsahin D, Onakpojeruo EP, Uzun B. Hydrogel-based drug delivery nanoparticles with conventional treatment approaches for cancer tumors; a comparative study using MCDM technique. In: *Advances in Science and Engineering Technology International Conferences (ASET)*; 2023.

    doi: 10.1109/ASET56582.2023.10180659

11. Hussain S, Mubeen I, Ullah N, *et al.* Modern diagnostic imaging technique applications and risk factors in the medical field: A review. *Biomed Res Int*. 2022;2022:5164970.

    doi: 10.1155/2022/5164970

12. Piorkowski A, Obuchowicz R, Najjar R. Redefining radiology: A review of artificial intelligence integration in medical imaging. *Diagnostics (Basel)*. 2023;13(17):2760.

    doi: 10.3390/diagnostics13172760

13. Uzun Ozsahin D, Onakpojeruo EP, Uzun B, Mustapha MT, Ozsahin I. Mathematical assessment of machine learning models used for brain tumor diagnosis. *Diagnostics (Basel)*. 2023;13(4):618.

    doi: 10.3390/diagnostics13040618

14. Abdusalomov AB, Mukhiddinov M, Whangbo TK. Brain tumor detection based on deep learning approaches and magnetic resonance imaging. *Cancers (Basel)*. 2023;15(16):4172.

    doi: 10.3390/cancers15164172

15. Neamah K, Mohamed F, Adnan MM, *et al*. Brain tumor classification and detection based DL models: A systematic review. *IEEE Access*. 2024;12:251-2542.

    doi: 10.1109/ACCESS.2023.3347545

16. Lotan E, Tschider C, Sodickson DK, *et al*. Medical imaging and privacy in the era of artificial intelligence: Myth, fallacy,

and the future. *J Am Coll Radiol*. 2020;17(9):1159.

doi: 10.1016/j.jacr.2020.04.007

17. Paul M, Maglaras L, Ferrag MA, Almomani I. Digitization of healthcare sector: A study on privacy and security concerns. *ICT Express*. 2023;9(4):571-588.

doi: 10.1016/j.icte.2023.05.004

18. Goodfellow IJ, Pouget-Abadie J, Mirza M, *et al.* Generative adversarial nets. In: *Advances in Neural Information Processing Systems*. Proceedings of the 27th International Conference on Neural Information Processing Systems: Canada; 2014. p. 2672-2680.

19. Muhammad A, Salman Z, Lee K, Han D. Harnessing the power of diffusion models for plant disease image augmentation. *Front Plant Sci*. 2023;14:1280496.

doi: 10.3389/fpls.2023.1280496

20. Ahmed MS, Hossain S, Haque MN, *et al.* MRI based automated detection of brain tumor using DWT, GLCM, PCA, ensemble of SVM and PNN in sequence. In: *Lecture Notes in Computer Sciences, LNICST*. V*ol.* 490. Berlin: Springer; 2023. p. 267-279.

doi: 10.1007/978-3-031-34619-4_22

21. Kang J, Ullah Z, Gwak J. MRI-based brain tumor classification using ensemble of deep features and machine learning classifiers. *Sensors (Basel)*. 2021;21(6):2222.

doi: 10.3390/s21062222

22. Ramakrishnan T, Sankaragomathi B. A professional estimate on the computed tomography brain tumor images using SVM-SMO for classification and MRG-GWO for segmentation. *Pattern Recognit Lett*. 2017;94:163-171.

doi: 10.1016/j.patrec.2017.03.026

23. Eze MC, Vafaei LE, Eze CT, Tursoy T, Ozsahin DU, Mustapha MT. Development of a novel multi-modal contextual fusion model for early detection of Varicella Zoster Virus skin lesions in human subjects. *Processes*. 2023;11(8):2268.

doi: 10.3390/pr11082268

24. Cheng J. *Brain Tumor Dat*aset. Figshare; 2017.

doi: 10.6084/m9.figshare.1512427

25. Ayadi W, Elhamzi W, Charfi I, Atri M. A hybrid feature extraction approach for brain MRI classification based on bag-of-words. *Biomed Signal Process Control*. 2019;48:144-152.

doi: 10.1016/j.bspc.2018.10.010

26. Rehman A, Naz S, Razzak MI, Akram F, Imran M. A deep learning-based framework for automatic brain tumors classification using transfer learning. *Circuits Syst Signal Process*. 2020;39(2):757-775.

doi: 10.1007/s00034-019-01246-3

27. Ullah N, Khan JA, Khan MS, *et al.* An effective approach to detect and identify brain tumors using transfer learning. *Appl Sci*. 2022;12(11):5645.

doi: 10.3390/app12115645

28. Priyadarshini P, Khairuzzaman AKM, Kanungo P. Brain tumor detection and classification from MRI images using cascaded deep neural networks. In: *Lecture Notes in Electrical Engineering*. V*ol.* 976. Singapore: Springer Nature Singapore; 2023. p. 301-311.

doi: 10.1007/978-981-99-0412-9_26

29. Ahmmed S, Podder P, Mondal MRH, *et al.* Enhancing brain tumor classification with transfer learning across multiple classes: An in-depth analysis. *BiomedInformatics*. 2023;3(4):1124-1144.

doi: 10.3390/biomedinformatics3040068

30. Jain J, Kubadia M, Mangla M, Tawde P. Comparison of transfer learning techniques to classify brain tumours using MRI images. *Eng Proc*. 2024;59(1):144.

doi: 10.3390/engproc2023059144

31. Veni N, Manjula J. VGG-16 architecture for MRI brain tumor image classification. In: *Lecture Notes in Electrical Engineering*. V*ol.* 966. Singapore: Springer Nature Singapore; 2023. p. 319-328.

doi: 10.1007/978-981-19-8338-2_26

32. Srinivas C, Nandini NP, Zakariah M, *et al.* Deep transfer learning approaches in performance analysis of brain tumor classification using MRI images. *J Healthc Eng*. 2022;2022:3264367.

doi: 10.1155/2022/3264367

33. Suryawanshi S, Patil SB. Brain tumor detection using YOLOv5 and Faster R-CNN. *Int J Intell Syst Appl Eng*. 2023;12(4s):335-342.

doi: 10.1109/ViTECoN58111.2023.10157773

34. Kesana A, Nallola J, Bootapally RT, Amaraneni S, Reddy GS. Brain tumor detection using YOLOv5 and Faster R-CNN. In: 2023 2nd International Conference on Vision Towards Emerging Trends in Communication and Networking Technologies (ViTECoN). *IEEE*; 2023. p. 1-6.

doi: 10.1109/ViTECoN58111.2023.10157773

35. Chlap P, Min H, Vandenberg N, *et al.* A review of medical image data augmentation techniques for deep learning applications. *J Med Imaging Radiat Oncol*. 2021;65(5):545-563.

doi: 10.1111/1754-9485.13261

36. Mangaokar N, Pu J, Bhattacharya P, Reddy CK, Viswanath B. Jekyll: Attacking medical image diagnostics using deep generative models. In: *Proceedings of EuroS&P* 2020. Genoa, Italy; 2020 p. 39-157.

doi: 10.1109/EuroSP48549.2020.00017

37. Salvi M, Branciforti F, Molinari F, Meiburger KM. Generative

models for color normalization in digital pathology and dermatology: Advancing the learning paradigm. *Expert Syst Appl*. 2024;245:123105.

doi: 10.1016/j.eswa.2023.123105

38. Kaur J, Singh AK, Jindal N. Systematic survey on generative adversarial networks for brain tumor segmentation and classification. *Concurrency Comput Pract Exper*. 2023;35(27):e7850.

doi: 10.1002/cpe.7850

39. Myronenko A. 3D MRI brain tumor segmentation using autoencoder regularization. In: *Lecture Notes in Computer Science*. Vol. 11384. Cham, Switzerland: Springer Nature; 2019. p. 311-320.

doi: 10.1007/978-3-030-11726-9_28

40. Huang P, Li D, Jiao Z, *et al*. Common feature learning for brain tumor MRI synthesis by context-aware generative adversarial network. *Med Image Anal*. 2022;79:102472.

doi: 10.1016/j.media.2022.102472

41. Cirillo MD, Abramian D, Eklund A. Vox2Vox: 3D-GAN for brain tumour segmentation. In: *Lecture Notes in Computer Science*. Vol. 12658. Cham, Switzerland: Springer Nature; 2021. p. 274-284.

doi: 10.1007/978-3-030-72084-1_25

42. Aljohani A, Alharbe N. Generating synthetic images for healthcare with novel deep Pix2Pix GAN. *Electronics*. 2022;11(21):3470.

doi: 10.3390/electronics11213470

43. Mehmood M, Alshammari N, Alanazi SA, *et al*. Improved colorization and classification of intracranial tumor expanse in MRI images via hybrid scheme of Pix2Pix-cGANs and NASNet-large. *J King Saud Univ Comp Inf Sci*. 2022;34(7):4358-4374.

doi: 10.1016/j.jksuci.2022.05.015

44. Narayanan SJ, Anil AS, Ashtikar C, Chunduri S, Saman S. Automated brain tumor segmentation using GAN augmentation and optimized U-Net. In: *Lecture Notes in Networks and Systems*. Vol. 519. Singapore: Springer Nature; 2023. p. 635-646.

doi: 10.1007/978-981-19-5191-6_51

45. Li Q, Yu Z, Wang Y, Zheng H. TumorGAN: A multi-modal data augmentation framework for brain tumor segmentation. *Sensors (Basel)*. 2020;20(15):4203.

doi: 10.3390/s20154203

46. Mukherjee D, Saha P, Kaplun D, Sinitca A, Sarkar R. Brain tumor image generation using an aggregation of GAN models with style transfer. *Sci Rep*. 2022;12:9380.

doi: 10.1038/s41598-022-12646-y

47. Azni HM, Afsharchi M, Allahverdi A. Improving brain tumor segmentation performance using CycleGAN-based feature extraction. *Multimed Tools Appl*. 2023;82(12):18039-18058.

doi: 10.1007/s11042-022-14174-3

48. Kim S, Kim B, Park H. Synthesis of brain tumor MR images for learning data augmentation. *Med Phys*. 2020;48(5):2185-2198.

doi: 10.1002/mp.14701

49. Saxena D, Cao J. Generative adversarial networks (GANs). *ACM Comput Surv*. 2021;54(3):1-38.

doi: 10.1145/3446374

50. Akbar MU, Larsson M, Blystad I, Eklund A. Brain tumor segmentation using synthetic MR images-a comparison of GANs and diffusion models. *Sci Data*. 2024;11:159.

doi: 10.1038/s41597-024-03073-x

51. Dhariwal P, Nichol A. Diffusion models beat GANs on image synthesis. *Adv Neural Inf Process Syst*. 2021;34:8780-8794.

52. Onakpojeruo EP, Mustapha MT, Ozsahin DU, Ozsahin I. Enhanced MRI-based brain tumour classification with a novel Pix2pix generative adversarial network augmentation framework. *Brain Commun*. 2024;6(6):fcae372.

doi: 10.1093/braincomms/fcae372

53. *Brain Tumor Classification (MRI) Dataset*. Kaggle; 2024. Available from: https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri [Last accessed on 2024 Mar 13].

54. Cheng J. *Brain Tumor Dataset*. Figshare; 2017. Available from: https://figshare.com/articles/dataset/brain_tumor_dataset/1512427/5 [Last accessed on 2024 Mar 13].

55. Yu J, Oh H, Lee Y, Yang J. Denoising diffusion model with adversarial learning for unsupervised anomaly detection on brain MRI images. *Pattern Recognit Lett*. 2024;186:229-235.

doi: 10.1016/j.patrec.2024.10.007