Journal of Algorithms and Computations

journal homepage: http://jac.ut.ac.ir

Enhancing Personalized Medicine for GBM Patients through Medical Image Analysis Using Generative Models and Deep Learning

N. Naddafi[*1], P. Mazloomi[**1], T. Khatibi[†1], A. Sedighipashaki[***2]

[1] School of Industrial and Systems Engineering, Tarbiat Modares University (TMU), Tehran, Iran

[2] Assistant Professor, Department of Radio oncology, School of Medicine Cancer Research Center, Hamedan university of medical sciences, Hamedan, Iran

ABSTRACT

This study presents a personalized medicine model for glioblastoma multiforme (GBM) patients using deep learning on MRI images and clinical data. MRI scans from 23 patients were analyzed to classify them into four groups based on radiotherapy dosage. A CNN-based model was developed and tested in three scenarios: baseline (96% accuracy), with added Gaussian noise (72%), and after image denoising (94%). The results show that integrating Vision Transformer and Auto-Encoder architectures can enhance radiotherapy planning by improving accuracy and reducing noise effects in medical imaging.

Keywords: Glioblastoma multiforme, radiotherapy, Magnetic Resonance Imaging, Deep learning, Image noise reduction

AMS subject classification: ???.

ARTICLE INFO

Article history:
Research paper
Received 22 April 2025
Accepted 16 July 2025
Available Online 16 July 2025

[*] Email: Niloufar.ndf.1998@gmail.com

[**] Email: Pooya.mazllomi@gmail.com

[†] Corresponding author: T. Khatibi, Email: Toktam.khatibi@modares.ac.ir

[***] Email: sedighi@umsha.ac.ir

1 Introduction

Glioblastoma multiforme (GBM) is the most prevalent and aggressive type of primary malignant brain tumor, accounting for approximately 60% of all adult brain tumors [1]. Its incidence is estimated at fewer than 10 cases per 100,000 individuals, with a notable increase reported over the past decade [2]. GBM carries a poor prognosis, with a median survival of only 12 to 16 months and a five-year survival rate of approximately 5% [3]. The disease occurs more frequently in males than in females and is more common among Caucasians than African Americans [4].

Current treatment modalities for glioblastoma multiforme (GBM) include surgical resection, radiotherapy, chemotherapy, and the administration of temozolomide [2]. While surgery remains the primary approach for tumor removal, radiotherapy and chemotherapy are commonly employed to manage tumors that cannot be entirely excised [5]. Globally, hundreds of thousands of patients receive radiotherapy each year for primary brain tumors and brain metastases originating from extracranial sites. This modality is essential for treating most brain tumors, and standard protocols combining surgery with chemoradiotherapy have significantly improved survival outcomes in GBM patients [6].

Radiotherapy techniques can be broadly categorized into whole brain radiotherapy (WBRT) and partial brain radiotherapy (PBRT). Stereotactic radiosurgery (SRS), in particular, employs precise three-dimensional imaging to deliver targeted radiation doses while minimizing exposure to surrounding healthy tissue. These methods serve various clinical purposes, including long-term tumor control, salvage therapy, and prophylactic treatment to prevent metastasis [6].

Magnetic resonance imaging (MRI) plays a pivotal role in diagnosing GBM by providing detailed brain images. However, its diagnostic accuracy can be compromised by image artifacts, often caused by patient movement, which introduce noise and hinder image interpretation [7]. Consequently, denoising has become a critical step in medical image processing to enhance image quality and facilitate accurate analysis. Numerous methods have been developed to reduce noise while preserving key image features [8].

The integration of computer-aided diagnosis (CAD), machine learning (ML), and deep learning (DL) techniques has shown considerable promise in analyzing imaging and clinical data related to GBM. ML algorithms are especially effective for disease classification and prediction in high-dimensional datasets. Moreover, generative models have emerged as powerful tools for supporting diagnosis and informing therapeutic decisions [9].

Our study contributes to the field of personalized medicine in GBM by introducing an innovative approach that combines advanced medical image analysis with generative and deep learning models. Using MRI scans and clinical data from 23 patients, we developed a robust convolutional neural network (CNN) capable of classifying patients based on prescribed radiotherapy doses, achieving an impressive accuracy of 96%. Additionally, we tackled the challenge of image noise through two analytical scenarios: one evaluated the impact of noise on model performance, and the other utilized a hybrid Vision Transformer (ViT) and autoencoder network to enhance image quality. Our findings demonstrate not only the model's efficacy in optimizing radiotherapy

planning but also its potential in enabling personalized therapeutic strategies tailored to individual patient profiles, ultimately contributing to improved GBM outcomes.

2 Background

To provide a clearer understanding of the technical foundation supporting our medical application, we begin with an overview of Convolutional Neural Networks (CNNs).

2.1 Convolutional Neural Networks (CNN)

Convolutional neural networks (CNNs) have become a fundamental component of deep learning, particularly in image processing tasks. As a specialized class of deep learning algorithms, CNNs require significantly less pre-processing than traditional image classification techniques, eliminating the need for manual feature extraction, a process often prone to human bias [10]. This reduced dependency on prior knowledge and human intervention offers a significant advantage for the automated analysis of medical images.

The superiority of CNNs over conventional neural networks in computer vision can be attributed to several key features:

1. **Weight Sharing:** The weight-sharing mechanism inherent in CNNs reduces the number of trainable parameters, which not only enhances computational efficiency but also improves model performance. Mathematically, if W represents the weight matrix and x denotes the input feature map, the convolution operation can be expressed as:

$$y[i,j] = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} W[m,n].\, x[i+m, j+n] \tag{1}$$

where $y[i,j]$ is the output feature map at position (i,j), and M and N are the dimensions of the filter.

2. **Simultaneous Learning**: CNNs facilitate the concurrent learning of feature extraction layers and classification layers, resulting in outputs that are highly organized and closely tied to the extracted features. This dual learning approach enhances the model's ability to generalize across various datasets.
3. **Scalability**: Implementing large-scale networks using CNNs is more manageable than with other neural network architectures, allowing for deeper networks that can capture complex patterns in data [11]. The stacking of multiple convolutional layers enables CNNs to construct deep network structures that effectively learn hierarchical representations from input images.

Moreover, convolutional neural networks (CNNs) have demonstrated remarkable efficacy in medical imaging applications, including disease classification, localization of pathological regions, and image enhancement. These networks can autonomously learn features from images without manual intervention, thereby enhancing the quality and diversity of extracted representations [12]. The incorporation of pooling layers further increases robustness by introducing invariance to transformations such as translation and rotation, thus improving generalization capabilities [13]. As a result, CNNs have emerged as a transformative technology

in healthcare, addressing challenges in image interpretation and diagnostic accuracy, and contributing significantly to the advancement of personalized medicine.

2.2 Autoencoders

Recent advancements in deep learning have led to its widespread application across various domains, demonstrating significant progress in both performance and versatility. Deep learning architectures can be tailored for specific tasks, with each architecture serving distinct purposes. For instance, autoencoders have gained prominence in unsupervised learning applications, particularly for dimensionality reduction of complex datasets.

An autoencoder is a specialized type of deep neural network that operates in a self-supervised manner, effectively encoding input data into a lower-dimensional representation before reconstructing it back to its original form. Mathematically, this process can be represented as follows:

$$X^{\wedge} = f(g(X)) \tag{2}$$

where X denotes the original input, g represents the encoder function that maps the input to a latent space, and f is the decoder function that reconstructs the input from this latent representation. The objective of training an autoencoder is to minimize the reconstruction error, typically quantified using a loss function such as mean squared error (MSE):

$$Loss = \frac{1}{N}\sum_{i=1}^{N}\left|\left|X_i - X_i^{\wedge}\right|\right|^2 \tag{3}$$

where N is the number of samples and $||\cdot||^2$ denotes the L2 norm [14, 15].

By learning efficient representations of data, autoencoders facilitate various applications, including data compression, denoising, and feature extraction, making them invaluable tools in the deep learning toolkit. Their ability to capture essential features while discarding noise underscores their effectiveness in preprocessing data for subsequent machine learning tasks.

2.3 variational autoencoders

Variational autoencoders (VAEs) represent a significant advancement over traditional autoencoders by introducing a probabilistic approach to latent variable representation. In conventional autoencoders, the encoder generates a deterministic latent representation of the input data, meaning that the same input will always yield the same output. Conversely, VAEs aim to establish a mapping between input data and a probability distribution in the latent space, characterized by the mean (μ) and variance (σ^2) of a Gaussian distribution. This probabilistic framework enables VAEs to sample from the latent space, facilitating the generation of new data points that resemble the training dataset [16, 17]

The fundamental distinction of VAEs lies in their ability to learn continuous distributions for latent variables, which is particularly advantageous for generative modeling tasks. Instead of producing a single fixed vector as in traditional autoencoders, the VAE encoder outputs two vectors: one representing the mean (μ) and another representing the standard deviation (σ). This allows VAEs

to model the distribution of latent variables based on their statistical properties rather than relying on a deterministic mapping. The loss function for VAEs incorporates both reconstruction loss and Kullback-Leibler (KL) divergence, expressed mathematically as:

$$Loss = Reconstruction\ Loss + DKL(q(z \mid x) \mid\mid p(z)) \quad (4)$$

where *DKL* measures how closely the learned distribution $q(z|x)$ approximates the prior distribution $p(z)$, typically assumed to be a standard normal distribution [18, 19].

This unique approach enhances VAEs' capabilities in various applications such as feature learning, dimensionality reduction, noise removal, and notably anomaly detection due to their ability to capture complex data distributions. By leveraging their probabilistic nature, VAEs can effectively generate diverse outputs that maintain meaningful characteristics akin to the original data, thus revolutionizing generative modeling in machine learning.

2.4 GANs

Generative adversarial networks (GANs) are a powerful class of generative models that utilize a unique framework of adversarial training to produce new images by leveraging both hidden and visible features of the data. In a GAN, two neural networks, the generator and the discriminator, are trained simultaneously in a competitive setting.The generator's role is to create synthetic samples that closely resemble real data, effectively capturing the underlying probability distribution of the training dataset. Conversely, the discriminator functions as a binary classifier tasked with distinguishing between genuine samples and those generated by the generator. This adversarial relationship can be framed as a minimax game, where the generator aims to minimize the discriminator's ability to correctly classify generated samples, while the discriminator seeks to maximize its accuracy in identifying real versus fake data [20]. Mathematically, the training process of GANs can be expressed through the following optimization problem:

$$min_G max_D V(D,G) = E_{x\sim p_{data}(x)}[logD(x)] + E_{z\sim p_z(z)}[1 - \log\left(1 - D\left(G(z)\right)\right)] \quad (5)$$

where *G* represents the generator, *D* denotes the discriminator, $p_{data}(x)$ is the distribution of real data, and $p_z(z)$ is the distribution of input noise (e.g., Gaussian noise). The generator attempts to produce outputs that maximize $D(G(z))$, while the discriminator aims to maximize its ability to differentiate between real and generated samples.

Both components of GANs are typically constructed using conventional deep neural network architectures, enabling them to learn complex representations from high-dimensional data. The iterative feedback loop inherent in this setup allows for continuous improvement of both networks; as the generator becomes more adept at producing realistic images, the discriminator's task becomes increasingly challenging [21]. This dynamic ultimately leads to a Nash equilibrium where the generator successfully captures the sampling distribution of real data while maintaining high fidelity in generated outputs. GANs have revolutionized various domains, including image synthesis and data augmentation, by providing a robust framework for generating high-quality synthetic data that can be indistinguishable from real-world samples.

2.5 VIT

Vision Transformers (ViTs) have rapidly emerged as a prominent architecture in computer vision, demonstrating outstanding performance in tasks such as object recognition, image classification, and segmentation. Unlike conventional convolutional neural networks (CNNs), which operate on pixel-level data, ViTs introduce a novel paradigm by treating images as sequences of patches. This approach, inspired by the transformer architecture originally developed for natural language processing (NLP), enables ViTs to utilize self-attention mechanisms to capture complex, long-range dependencies within visual data [22].

In a ViT model, an input image is divided into fixed-size patches, each of which is flattened and linearly embedded. These embeddings function analogously to word tokens in NLP models and are fed into a transformer encoder. The use of positional encodings ensures that the spatial relationships among patches are preserved, allowing the model to effectively interpret the overall structure of the image. This architecture is particularly advantageous when trained on large datasets, where it has been shown to achieve superior accuracy across a wide range of computer vision tasks [23].

A core strength of ViTs lies in their capacity to model global context through self-attention, in contrast to the localized feature extraction typical of CNNs. This global perspective not only enhances classification performance but also benefits object detection and segmentation by enabling the model to understand relationships among distant regions in an image. Consequently, ViTs have gained traction in domains such as medical imaging, where high-resolution modalities like MRI and X-ray require precise and comprehensive visual analysis [24].

Overall, Vision Transformers mark a significant shift in computer vision methodologies, offering enhanced flexibility, scalability, and interpretability. Their effectiveness across diverse applications positions them as a transformative tool for advancing image analysis, particularly in data-intensive and high-stakes fields like healthcare and autonomous systems.

3 Methodology

In the following subsections, the main steps of the proposed method are explained.

Ethics statement

The acquisition and analysis of the data within this study was approved by ‘Research ethics committee of Tarbiat Modares University’ and they have rewarded the research ethics certificate with approval ID of IR.MODARES.REC.1402.153.

3.1 Dataset

This study is a retrospective analysis of medical records and archived samples, conducted exclusively for research purposes. No identifiable information related to individual participants was accessible during or after the data collection process, ensuring full compliance with privacy and ethical standards. The clinical data and MRI images analyzed in this research were obtained from a diagnostic and treatment center located in Hamedan, Iran. We focused on 23 patients diagnosed with glioblastoma multiforme (GBM), who presented with symptoms such as headache, vertigo, sensory impairment, and memory deficits.

To obtain comprehensive clinical and therapeutic data, we meticulously reviewed the patients' medical records in consultation with a specialist physician. Key details, including clinical history, histopathological findings, and prescribed treatment plans, were extracted and documented. MRI images were retrieved from the center's Picture Archiving and Communication System (PACS), each with a resolution of 512 × 512 pixels. Both T1-weighted and T2-weighted scans were available for each patient, encompassing three anatomical views: lateral (left), superior, and posterior.

In total, we compiled 1,187 MRI images, classified into four GBM-related categories based on treatment modalities recommended by clinical specialists. A representative set of MRI images across the three views is shown in Figure 1. This dataset forms a robust foundation for subsequent analysis and model development, aimed at enhancing diagnostic precision and optimizing therapeutic strategies for GBM.

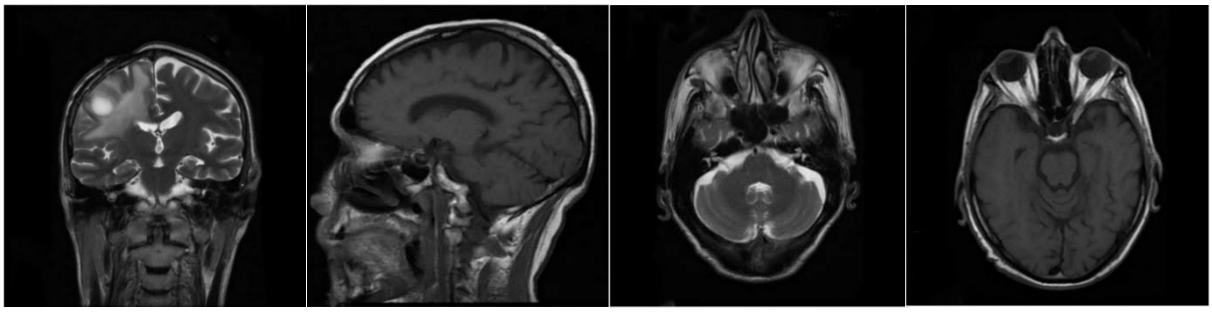

Figure 1. An example of MRI images acquired from different anatomical views

3.2 Pre-processing

Pre-processing is a critical step in enhancing image quality, refining image characteristics, and minimizing noise in medical imaging, particularly in brain MRI analysis [25]. In this study, brain scans were initially obtained in DICOM format from an imaging center and subsequently converted to JPG format to ensure compatibility with standard processing tools. As an initial measure, identifiable patient information, such as names, ages, and scan details, was manually removed from the image corners using Adobe Photoshop to protect privacy.

Due to the high volume of images, their resolution was reduced from 512×512 pixels to 128×128 pixels, which significantly accelerated the processing time and computational efficiency. Subsequently, pixel intensity values in the range [0, 255] were normalized to a [0, 1] scale. This

normalization is essential for deep learning applications, as it improves model convergence, enhances numerical stability, and ensures consistency across datasets.

In addition to these steps, several advanced pre-processing techniques can be employed to further improve image clarity and diagnostic utility. Techniques such as Gaussian smoothing, bilateral filtering, and K-means clustering are effective in reducing image noise while preserving critical anatomical details [26]. Furthermore, methods like gamma correction and window level adjustment can enhance contrast and highlight tumor regions [8]. Collectively, these pre-processing strategies contribute to producing high-quality images that facilitate more accurate diagnosis, segmentation, and treatment planning in clinical practice.

3.3 Define problem

In this study, we aimed to classify patients with glioblastoma multiforme (GBM) based on their prescribed radiotherapy dosages. The dosage range for each patient was first determined and then categorized into four distinct classes: Class 1 for doses less than 5500 cGy, Class 2 for doses between 5501 and 6000 cGy, Class 3 for doses between 6001 and 6500 cGy, and Class 4 for doses exceeding 6500 cGy. To perform this classification, we developed a Convolutional Neural Network (CNN) model specifically designed to distinguish patients among these dosage groups.

To improve the model's robustness to image noise, which is a common issue in medical imaging, we defined and analyzed two separate scenarios. In Scenario 1, we constructed a classification model and evaluated how different levels of noise influenced its accuracy. This step was essential due to the inherent variability and artifacts often found in MRI data.

Scenario 2 incorporated a hybrid architecture that combined a Vision Transformer (ViT) with an autoencoder to reduce the impact of noise and improve both image quality and classification accuracy. This design allowed us to explore the benefits of integrating advanced neural network components into the classification process.

Accurate prediction of radiotherapy dosages plays a vital role in determining treatment effectiveness and improving patient survival. Previous studies have shown the potential of deep learning models, particularly CNNs, in predicting dose distributions with high accuracy (Porkodi et al., 2023). Building on this foundation, our research introduces new strategies to refine the classification process and enhance diagnostic precision in GBM treatment planning.

3.3.1 Scenario 1

In this scenario, we developed a Convolutional Neural Network (CNN) model that utilized two types of inputs: tabular data and image data. To effectively extract features from the MRI images, we implemented two convolutional layers with 32 and 64 filters, respectively, each using a kernel size of 3x3. Following these convolutional layers, a max pooling layer was applied to reduce the dimensionality of the feature maps, thereby improving computational efficiency and helping to mitigate overfitting. The output from the convolutional layers was then flattened using a Flatten layer to prepare it for integration with the tabular data.

For the tabular data, two Dense layers with 64 and 32 neurons were designed to effectively process the structured clinical information. The outputs from both the Flatten layer (derived from image

data) and the Dense layers (from tabular data) were merged and fed into a subsequent Dense layer consisting of 128 neurons. To further prevent overfitting, a dropout layer with a dropout rate of 0.5 was incorporated, randomly deactivating half of the neurons during training to promote generalization.

The final output layer contained 4 neurons, corresponding to the four predefined dosage classes. For training, 80% of the dataset was allocated to training and the remaining 20% to testing, ensuring a robust evaluation of model performance. The CNN model was trained with a batch size of 32 over 50 epochs, optimizing the learning process.

After training, noise was introduced into the MRI images to evaluate the model's robustness and its ability to accurately classify noisy inputs. The detailed architecture and hyperparameters of the CNN model are summarized in Table 1 (a) and (b), respectively.

Table 1. (a) The architecture of our designed and developed CNN model, and (b) its hyper-parameters

(a)	
Layer Type	**Configuration**
Convolution Layer 1	32 filters, kernel size 3x3
Convolution Layer 2	64 filters, kernel size 3x3
Max Pooling Layer	Pool size 2x2
Flatten Layer	-
Dense Layer 1	64 neurons
Dense Layer 2	32 neurons
Merged Layer	-
Dense Layer 3	128 neurons
Dropout Layer	Dropout rate = 0.5
Output Layer	4 neurons (for classification)
(b)	
Parameter	**Configuration**
Feature Extraction	CNN
Optimizer	Adam
Activation Functions	ReLU (hidden layers) – Softmax (output layer)
Loss Function	Categorical Cross Entropy
Metric	Accuracy

This structured approach allowed us to evaluate the impact of image noise on classification accuracy while leveraging both image and tabular data effectively in our CNN model.

3.3.2 Scenario 2

In this scenario, we implemented a Vision Transformer (ViT) combined with an Auto Encoder model to enhance feature extraction and reduce noise in the classification task. The architecture began with the ViT-b16 model, which was used to extract features from all input images. The output from the ViT-b16 model was then passed through a Flatten layer to prepare the data for subsequent processing.

After flattening, the output was fed into a Dense layer containing 64 neurons. The encoder part of the model included a convolutional layer with 64 filters and a kernel size of 3x3. The decoder consisted of a convolutional layer with 32 filters and a kernel size of 3x3, followed by another convolutional layer with 3 filters of the same size. To ensure the output dimensions matched those of the input, an Upsampling layer with a size of 2x2 was applied at the final stage of the Auto Encoder.

The activation function for the decoder's output layer was sigmoid, while ReLU was applied in intermediate layers. The ViT-Auto Encoder model was designed to reduce noise in MRI images before classification, trained with a batch size of 16 over 50 epochs. The detailed configuration of the ViT-b16 model is summarized in Table 2.

Table 2. (a) The architecture of VIT-b16 model, and (b) its hyper-parameters

(a)	
Layer Type	**Configuration**
Feature Extraction	VIT-b16
Flatten Layer	-
Dense Layer	64 neurons
Encoder Convolution Layer	64 filters, kernel size 3x3
Decoder Convolution Layer 1	32 filters, kernel size 3x3
Decoder Convolution Layer 2	3 filters, kernel size 3x3
Upsampling Layer	Size 2x2
Activation Functions	ReLU (hidden layers) – Sigmoid (output layer)
Optimizer	Adam
Loss Function	Binary Cross Entropy
(b)	
Parameter	**Configuration**
Feature Extraction	VIT-b16
Optimizer	Adam
Activation Functions	ReLU (hidden layers) – Sigmoid (output layer)
Loss Function	Binary Cross Entropy

After training, the output of the ViT-Auto Encoder model was fed into the pre-trained CNN model from Scenario 1. This integration aimed to effectively classify noisy data by leveraging the enhanced feature representations obtained through the combined architecture. By utilizing both ViT and Auto Encoder methodologies, we sought to improve classification accuracy and robustness against noise in medical imaging data.

4 Result

This section presents the outcomes of our experiments evaluating the performance of different neural network models for classifying glioblastoma multiforme patients based on their MRI images and prescribed radiotherapy dosages.

4.1 Scenario 1

The results of the model training are illustrated in Figures 2 and 3, showing the accuracy and loss function per epoch over 50 epochs. The model achieved an impressive accuracy of **96%** when evaluated on noise-free images. However, the introduction of noise to the MRI images caused the accuracy to decrease to **72%**, highlighting the significant impact of noise on model performance. These findings emphasize the critical importance of addressing image quality in medical imaging applications.

Detailed results of the model's performance, including accuracy and loss metrics, are presented in Table 3, alongside visual representations in Figures 4 and 5. This comprehensive analysis offers valuable insights into the model's learning process and its sensitivity to noise in the input data.

Table 3.Performance measures of CNN with denoised images (left) and noisy images (right) in scenario 1

	f1-score	**recall**	**precision**		**f1-score**	**recall**	**precision**
Class 1	0.93	0.97	0.95	**Class 1**	0.60	0.17	0.29
Class 2	0.98	0.95	0.96	**Class 2**	0.47	0.80	0.64
Class 3	0.94	0.96	0.95	**Class 3**	0.60	0.70	0.82
Class 4	0.97	0.95	0.96	**Class 4**	0.97	0.99	0.98
Accuracy	0.96			**Accuracy**	0.72		
macro avg	0.96	0.96	0.96	**macro avg**	0.86	0.72	0.68
weighted avg	0.96	0.96	0.96	**weighted avg**	0.86	0.72	0.69

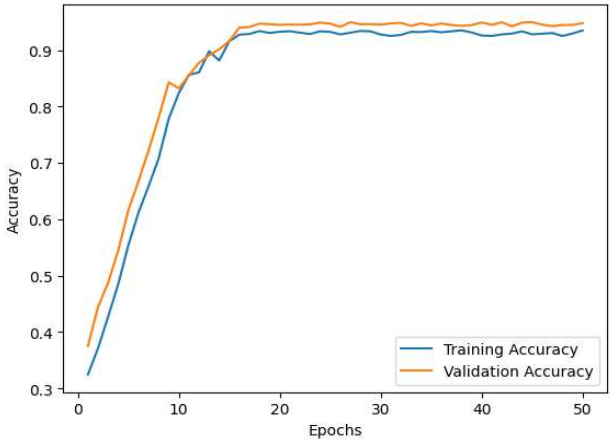

Figure 2. the diagram of accuracy per epoch for our designed CNN in Scenario 1

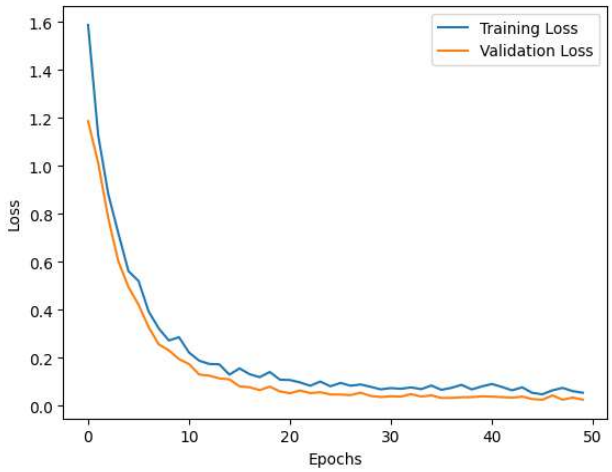

Figure 3. the diagram of loss function per epoch for our designed CNN in Scenario 1

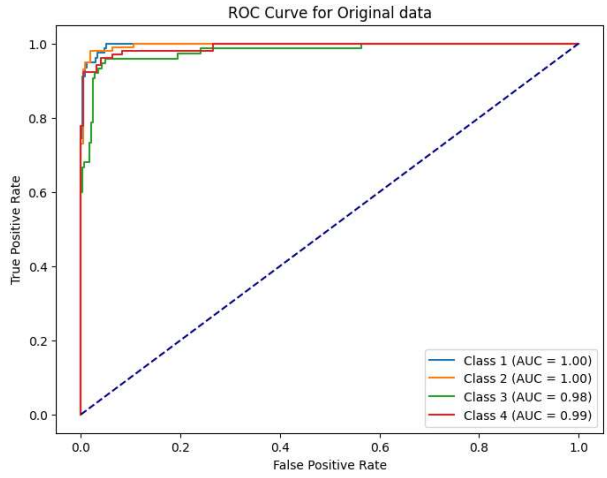

Figure 4. ROC Curve of our designed CNN in scenario 1

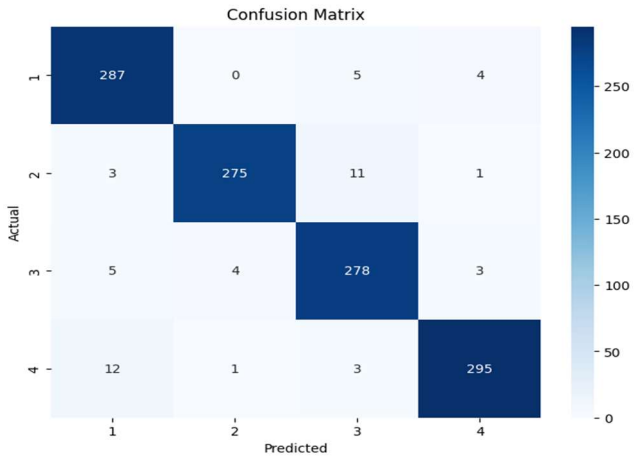

Figure 5. Confusion Matrix of our designed CNN in scenario 1

4.2 Scenario 2

By incorporating noise correction capabilities into the model, it achieved an accuracy of 94%, approaching the performance observed in Scenario 1. Table 4 presents the performance metrics of our designed ViT-based neural network in Scenario 2.

Table 4. Performance measures of our designed VIT-based Neural network in scenario 2

	f1-score	**recall**	**precision**
Class 1	0.96	0.94	0.95
Class 2	0.97	0.95	0.96
Class 3	0.87	0.96	0.92
Class 4	0.99	0.94	0.95
Accuracy		0.94	
macro avg	0.94	0.94	0.94
weighted avg	0.94	0.94	0.94

Also, the loss per epoch for MRI images before and after noise correction by the autoencoder is illustrated in Figure 6. Detailed model results for Scenario 2 are presented in Figures 7 and 8.

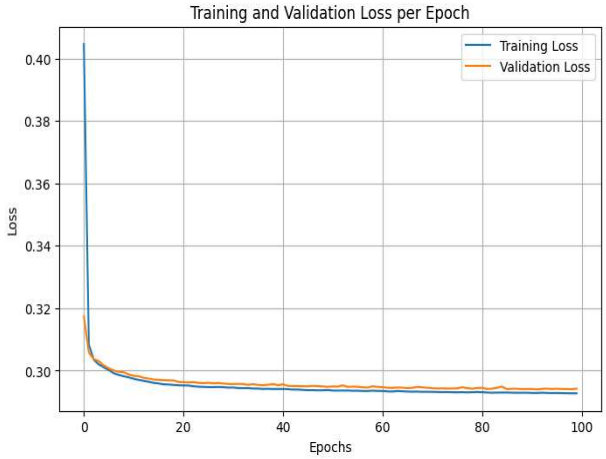

Figure 6. The diagram of loss function per epoch for our designed VIT-based Neural network in scenario 2

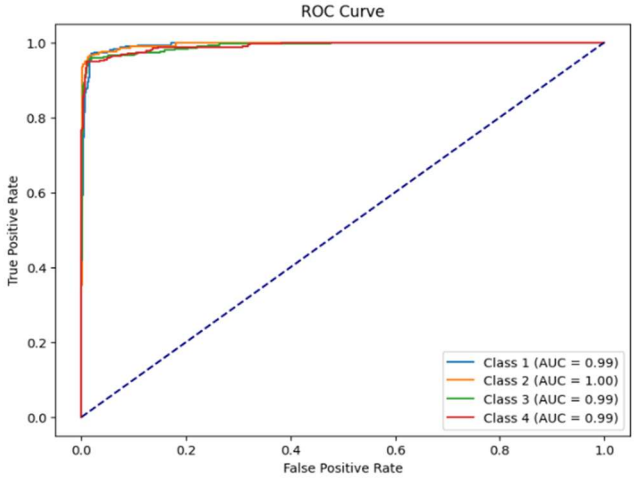

Figure 7. ROC Curve for our designed VIT-based Neural network in scenario 2

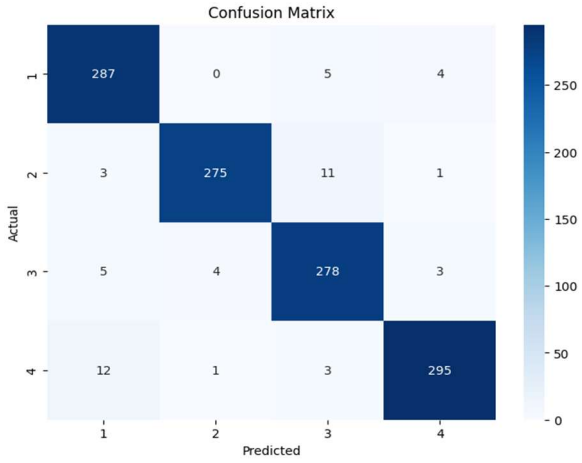

Figure 8. Confusion Matrix for our designed VIT-based Neural network in scenario 2

5 Conclusion

In this research, we successfully predicted the optimal radiotherapy dosage for patients with glioblastoma multiforme (GBM) through the analysis of MRI images using various neural network models. In Scenario 1, we developed a Convolutional Neural Network (CNN) that effectively determined the optimal radiotherapy dose based on patients' MRI images. The model achieved an accuracy of 96% when evaluated on noise-free images. However, when artificial noise was introduced to the MRI images using the OpenCV library, the model's accuracy decreased to 72%. To mitigate this degradation, we employed an Autoencoder for noise reduction, which improved the accuracy to 94% when the denoised images were classified by the original CNN model.

These results indicate that our model can serve as a valuable diagnostic aid for healthcare professionals treating GBM patients, even when handling noisy MRI data. This capability underscores the critical importance of robust image processing techniques in enhancing diagnostic accuracy and treatment planning.

For future work, we propose exploring additional brain imaging modalities to further refine dosage predictions. Moreover, extending this model to predict optimal radiotherapy doses for other types of brain tumors could significantly advance personalized treatment strategies in oncology. The findings from this study contribute to the growing body of evidence supporting the integration of deep learning models in medical imaging and radiotherapy planning, consistent with recent literature on dose prediction models (e.g., 3D U-Nets), which have demonstrated promising results in similar applications.

Declarations:

Funding:
This study was not funded by any organization.

Competing interests:
The authors declare that there are no conflicts of interest.

Ethics approval and consent to participate:
The data used for this study has been rewarded the research ethics certificate with approval ID of IR.MODARES.REC.1402.153. Our study is a retrospective study which analyzes the data collected and archived in the hospital information system previously. Therefore, no patient has participated in this study directly and there is no need for getting informed consent for participation from patients. On the other hand, data records have been anonymized before feeding to us for analysis. Thus, the need for consent to participate was waived by Research Ethics Committee of Tarbiat Modares University according to national regulations. This study adhered to the Declaration of Helsinki to this effect.

Authors' Contribution
Conceptualization: All Authors
Data curation: NN and PM
Formal analysis: NN, PM and TK

Funding acquisition: there is no funding.
Investigation: All authors
Methodology: NN, PM and TK
Project administration: TK
Software: NN and PM
Supervision: TK
Validation: TK and AS
Visualization: NN and PM
Writing – original draft: NN and PM
Writing – review & editing: All authors
All authors read and approved the manuscript.

Data availability

Availability of data and material

This is a retrospective study. Our collected dataset belongs to the third party named Mahdieh Imaging Center in Hamedan. Therefore, we are not allowed to publish dataset.

Acknowledgement

Special thanks to Mahdieh Imaging Center in Hamedan for their assistance.

Consent for publication:

Not Applicable.

References

[1] D. Nie, H. Zhang, E. Adeli, L. Liu, and D. Shen, "3D deep learning for multi-modal imaging-guided survival time prediction of brain tumor patients," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9901 LNCS, pp. 212–220, 2016, doi: 10.1007/978-3-319-46723-8_25.

[2] U. Raghavendra *et al.*, "Feature-versus deep learning-based approaches for the automated detection of brain tumor with magnetic resonance images: A comparative study," *Int. J. Imaging Syst. Technol.*, vol. 32, no. 2, pp. 501–516, 2022, doi: 10.1002/ima.22646.

[3] A. Sinha, A. R P, M. Suresh, N. R, A. D, and A. Singerji, *Brain Tumour Detection Using Deep Learning*. 2021. doi: 10.1109/ICBSII51839.2021.9445185.

[4] H. G. Wirsching, E. Galanis, and M. Weller, "Glioblastoma," *Handb. Clin. Neurol.*, vol. 134, pp. 381–397, 2016, doi: 10.1016/B978-0-12-802997-8.00023-2.

[5] M. Lara-Velazquez *et al.*, "Advances in brain tumor surgery for glioblastoma in adults," *Brain Sci.*, vol. 7, no. 12, pp. 1–16, 2017, doi: 10.3390/brainsci7120166.

[6] M. T. Makale, C. R. McDonald, J. A. Hattangadi-Gluth, and S. Kesari, "Mechanisms of radiotherapy-associated cognitive disability in patients with brain tumours," *Nat. Rev. Neurol.*, vol. 13, no. 1, pp. 52–64, 2016, doi: 10.1038/nrneurol.2016.185.

[7] I. Oksuz, "Brain MRI artefact detection and correction using convolutional neural networks," *Comput. Methods Programs Biomed.*, vol. 199, p. 105909, 2021, doi: 10.1016/j.cmpb.2020.105909.

[8] S. Ramesh, S. Sasikala, and N. Paramanandham, "Segmentation and classification of brain tumors using modified median noise filter and deep learning approaches," *Multimed. Tools Appl.*, vol. 80, no. 8, pp. 11789–11813, 2021, doi: 10.1007/s11042-020-10351-4.

[9] I. Jovčevska, "Next Generation Sequencing and Machine Learning Technologies Are Painting the Epigenetic Portrait of Glioblastoma," *Front. Oncol.*, vol. 10, no. May, pp. 1–14, 2020, doi: 10.3389/fonc.2020.00798.

[10] L. Alzubaidi *et al.*, *Review of deep learning: concepts, CNN architectures, challenges, applications, future directions*, vol. 8, no. 1. Springer International Publishing, 2021. doi: 10.1186/s40537-021-00444-8.

[11] M. Jogin, Mohana, M. S. Madhulika, G. D. Divya, R. K. Meghana, and S. Apoorva, "Feature Extraction using Convolution Neural Networks (CNN) and Deep Learning," in *2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, 2018, pp. 2319–2323. doi: 10.1109/RTEICT42901.2018.9012507.

[12] J.-E. Bibault and L. Xing, "The Role of Big Data in Personalized Medicine," in *Precision Medicine in Oncology*, 2020, pp. 229–247. doi: https://doi.org/10.1002/9781119432487.ch8.

[13] P. J. Withers *et al.*, "X-ray computed tomography," *Nat. Rev. Methods Prim.*, vol. 1, no. 1, p. 18, 2021, doi: 10.1038/s43586-021-00015-4.

[14] J. Donnelly, A. Daneshkhah, and S. Abolfathi, "Forecasting global climate drivers using Gaussian processes and convolutional autoencoders," *Eng. Appl. Artif. Intell.*, vol. 128, no. November 2023, p. 107536, 2024, doi: 10.1016/j.engappai.2023.107536.

[15] D. Bank, N. Koenigstein, and R. Giryes, *Autoencoders*. 2020.

[16] P. Naga Srinivasu, T. B. Krishna, S. Ahmed, N. Almusallam, F. Khaled Alarfaj, and N. Allheeib, "Variational Autoencoders-BasedSelf-Learning Model for Tumor Identification and Impact Analysis from 2-D MRI Images," *J. Healthc. Eng.*, vol. 2023, no. 1, p. 1566123, Jan. 2023, doi: https://doi.org/10.1155/2023/1566123.

[17] D. Papadopoulos and V. D. Karalis, "Variational Autoencoders for Data Augmentation in Clinical Studies," *Appl. Sci.*, vol. 13, no. 15, 2023, doi: 10.3390/app13158793.

[18] G. A. Noghre, A. D. Pazho, and H. Tabkhi, "An Exploratory Study on Human-Centric Video Anomaly Detection through Variational Autoencoders and Trajectory Prediction," *Proc. - 2024 IEEE Winter Conf. Appl. Comput. Vis. Work. WACVW 2024*, pp. 995–1004, 2024, doi: 10.1109/WACVW60836.2024.00109.

[19] M. Elbattah, C. Loughnane, J.-L. Guérin, R. Carette, F. Cilia, and G. Dequen, "Variational Autoencoder for Image-Based Augmentation of Eye-Tracking Data," 2021. doi: 10.3390/jimaging7050083.

[20] S. P. Porkodi, V. Sarada, V. Maik, and K. Gurushankar, "Generic image application using GANs (Generative Adversarial Networks): A Review," *Evol. Syst.*, vol. 14, no. 5, pp. 903–917, 2023, doi: 10.1007/s12530-022-09464-y.

[21] V. L. Trevisan de Souza, B. A. D. Marques, H. C. Batagelo, and J. P. Gois, "A review on Generative Adversarial Networks for image generation," *Comput. Graph.*, vol. 114, pp. 13–25, 2023, doi: https://doi.org/10.1016/j.cag.2023.05.010.

[22] A. B. Amjoud and M. Amrouch, "Object Detection Using Deep Learning, CNNs and Vision Transformers: A Review," *IEEE Access*, vol. 11, pp. 35479–35516, 2023, doi:

10.1109/ACCESS.2023.3266093.

[23] A. Dosovitskiy *et al.*, “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale,” *arXiv*, 2024.

[24] F. Shamshad *et al.*, “Transformers in medical imaging: A survey,” *Med. Image Anal.*, vol. 88, p. 102802, 2023, doi: https://doi.org/10.1016/j.media.2023.102802.

[25] D. Divyamary, S. Gopika, S. Pradeeba, and M. Bhuvaneswari, *Brain Tumor Detection from MRI Images using Naive Classifier*. 2020. doi: 10.1109/ICACCS48705.2020.9074213.

[26] N. Amiri Golilarz, H. Gao, R. Kumar, L. Ali, Y. Fu, and C. Li, “Adaptive Wavelet Based MRI Brain Image De-noising,” *Front. Neurosci.*, vol. 14, no. July, pp. 1–14, 2020, doi: 10.3389/fnins.2020.00728.