

Article

Not peer-reviewed version

---

# Dual-Stream Contrastive Latent Learning GAN for Brain Image Synthesis and Tumor Classification

---

[Junaid Zafar](#) , [Vincent Koc](#) , [Haroon Zafar](#) \*

Posted Date: 17 March 2025

doi: 10.20944/preprints202503.1235.v1

Keywords: GAN; Tumor Classification; Contrastive Latent Learning



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

*Article*

# Dual-Stream Contrastive Latent Learning GAN for Brain Image Synthesis and Tumor Classification

Junaid Zafar <sup>1</sup>, Vincent Koc <sup>2</sup> and Haroon Zafar <sup>3,4\*</sup>

<sup>1</sup> Faculty of Engineering, Government College University, Lahore, Pakistan.

<sup>2</sup> Comet, USA.

<sup>3</sup> Lambe Institute for Translational Research, University of Galway, Ireland.

<sup>4</sup> CÚRAM- Research Ireland Centre for Medical Devices, Ireland.

\* Correspondence: all correspondence to: Dr. Haroon Zafar; 2nd Floor, Lambe Institute for Translational Research, National University of Ireland Galway, Ireland; E-mail: haroon.zafar@universityofgalway.ie

**Abstract:** Generative Adversarial Networks (GANs) prioritize pixel-level attributes over capturing the entire image distribution which is critical in image synthesis. To address this challenge, we propose (DSCLPGAN) a dual-stream generator coupled with contrastive latent projection (CLP) for the robust augmentation of MRI images. The dual-stream generator in our architecture incorporates two specialized processing pathways: one dedicated to local feature variation modeling, while the other captures global structural transformations, ensuring a more comprehensive synthesis of medical images. We used a transformer-based encoder-decoder framework for contextual coherence and the contrastive learning projection (CLP) module integrates contrastive loss into the latent space for generating diverse image samples. The generated images undergo adversarial refinement using an ensemble of specialized discriminators where discriminator 1 (D1) ensures classification consistency with real MRI images, discriminator 2 (D2) produces a probability map of localized variations and discriminator (D3) functions for preserving structural consistency. For validation, we utilize a publicly available MRI dataset which contains from 3064 T1-weighted contrast-enhanced images with three types of brain tumor: meningioma (708 slices), glioma (1426 slices), and pituitary tumor (930 slices). Experimental results demonstrate state-of-the-art performance, achieving an SSIM of 0.99, classification accuracy of 99.4% for an at an augmentation diversity level of 5 and PSNR of 34.6 dB. Our approach has the potential of generating high-fidelity augmentations for reliable AI-driven clinical decision support systems.

**Keywords:** GAN; tumor classification; contrastive latent learning

## 1. Introduction

Brain tumors (BTs) remain a significant public health concern, ranking as the 10th leading cause of death in the United States only [1]. Research indicates that brain tumors are significantly impacting patients' lives through physical, cognitive, and psychological impairments [2]. BTs can be benign or malignant, with benign tumors being slow-growing and localized, while malignant tumors are highly aggressive with metastasis [3]. Brain tumors (BTs) encompass a diverse range of types with varying degrees of aggressiveness, including glioblastoma, the most aggressive form; meningioma, which is often benign; and pituitary adenomas [4,5]. The World Health Organization (WHO) classifies brain tumors on a scale from Grade I to IV based on their degree of spread, biological behavior, and prognosis [6]. Targeted efforts are essential to enhance early detection methods and deepen the understanding of brain tumor progression, as early diagnosis expands treatment options and improves survival rates [7].

Magnetic Resonance Imaging (MRI) is a highly effective imaging modality for detecting and characterizing various aspects of brain tumors. It offers superior soft tissue contrast while minimizing

patient exposure to ionizing radiation [8]. However, BT diagnosis through MRI scans is highly time-intensive and heavily dependent on the radiologist's expertise. Accurately labeling the scans without misclassification presents challenges in both time and precision. To address this issue, various computer-aided solutions have been developed to support automated decision-making systems [9,10]. Artificial Intelligence (AI) has emerged as a promising tool for the early detection of BTs and is being leveraged to enhance diagnosis by analyzing MRI scans. Various deep learning architectures have been documented in the literature for the detection and classification of BTs using MRI scans [11–15]. In [16], an automated brain tumor classification method is proposed using an enhanced deep learning approach with DenseNet121. Transfer learning is applied, and hyper-parameter tuning optimizes the CNN model. MRI images of three distinct brain tumor types were analyzed using the DenseNet169 model for feature extraction. The extracted features were then fed into three multi-class machine learning classifiers, Random Forest (RF), Support Vector Machine (SVM), and XGBoost to enhance performance [17]. Elsewhere [18], the effectiveness of deep transfer learning has been assessed using ResNet152, VGG19, DenseNet169, and MobileNetv3 models. Similarly, [19] presented the CapsNet model for multiclass classification of neuro- degenerative diseases using a modified DenseNet-169 framework coupled with the Enhanced DeepLab V3+ model. Further, a brief summary of seven deep learning models for BTs detection including VGG-16, VGG-19, ResNet50, Inception, ResNetV2, InceptionV3, Xception, and DenseNet201 and five traditional classifiers considering SVMs, RF, Decision Trees, AdaBoost, and Gradient Boosting are documented in the literature [20]. Although these advanced architectures have achieved promising results, they largely overlook the inherent limitations of imaging datasets, particularly significant class imbalances. The challenge of acquiring large, diverse datasets encompassing patients at various disease stages constrains the full potential of deep learning networks.

To overcome the challenges posed by limited medical imaging datasets, image augmentation techniques help in enriching data diversity, mitigating class imbalance, and improve model generalization. Adversarial learning frameworks have been employed to optimize the interaction between the generator and discriminator for creating high quality synthetic medical images. In this regard, an automatic data GAN was used to learn the available annotated MRI samples of BRATS15 Challenge dataset [21] and inception, resnetv2, inceptionv3, transfer learning, and BRAIN-TUMOR-net were published in [22] for MRI images of glioma tumor, meningioma and pituitary BTs. [23] reported progressive GANs (PGGANs), in which multistage generative training was to generate BT images that were challenging for conventional GANs and further detailed reviews are available at [24–28] concerning the standard used augmentation methods and fusion deep learning models including U- Net.

Transformers and auto-encoder architectures are imperative for MRI image augmentations due to their ability to overcome key limitations of deep learning CNN models including **long-range dependencies**, and **global contextual awareness**. A cross- transformer was published to include self-care model keys, queries, and values for the classification of BTs in MRI images and compared the results with InceptionResNetV2, InceptionV3, DenseNet121, Xception, ResNet50V2, VGG19, and EfficientNetB7 networks [29]. In [30], the possibilities of vision transformers (ViTs) in this research as a viable alternative to CNNs for brain magnetic resonance imaging was pitched by incorporating self-attention mechanisms to establish relationships between image patches for a comprehensive understanding. Likewise [31] highlighted the relevance of dual vision transformer model (DSUNET) in providing reliable and efficient differentiation between brain tumors and other brain regions by leveraging MRI BRATS 2020 dataset. Besides swin transformer was introduced into UNet++ network, in which local features of BTs were extracted by convolutional layer in UNet++ and global resolutions were captured via shift window operation of swin transformer [32]. Researchers deployed a shrinking linear time vision transformer (SL(t)-ViT) network for enhanced classification across multiple datasets of BTs [33]. Hybrid architectures combined with cross-fusion allows parallel systems to be merged between branches, resulting in reliable prediction of various tumors and similar studies are available in the literature [34,35]. In these researches, either a two-branch parallel model

that integrates the transformer module (TM) with the self-attention mechanism was used to classify brain tumors in MR images or hybrid shifted windows multi-head modules were deployed. To address the challenge of explainability in adversarial transformer models, graph attention network (GAT)-based transformer schemes were published [36]. Diffusion models on the other side **learn the actual data distribution** through likelihood-based training, making them more interpretable as compared to adversarial networks. Therefore, researchers comprehensively evaluated four GANs (progressive GAN, StyleGAN 1–3) and a diffusion model using two segmentation networks, U-Net and a Swin transformer for the task of brain tumor segmentation [37]. To address the issue of inter-class and intra-class problems, a gated global-local attention (GGLA) mechanism was evolved. The gating mechanism within the GGLA dynamically balances the contributions of global and local information, enabling the model to adaptively focus on the most relevant features for accurate classification. Additionally, an enhanced super-resolution generative adversarial network (ESRGAN) was coupled to generate images that balance the MRI image data [38]. The cutting-edge techniques in multiple stages within deep image recognition generative adversarial network (DIR-GAN) reported robustness of brain tumor detection and classification in MRI scans [39]. Similar work was published [40] where HARA-GAN was proposed by integrating residual U-net with Hybrid attention and Relative average discriminator, to mitigate noise caused by low under sampling rates. The results indicated that HARA-GAN outperforms DAGAN, RefineGAN and RSCA-GAN methods based on error maps and quantitative evaluation metrics in terms of both image quality and consistencies on MRI brain dataset. To improvise on capturing long-range dependencies and spatial variations [41] reported residual attention U-shaped network (RAUNet) for brain tumor segmentation which leverages the robust feature extraction capabilities of U-Net and the global context awareness provided by transformers to improve segmentation accuracy and [42] introduced hybrid federated adversarial MRI enhancement (FAME) by integrating advanced GAN architectures such as multi-scale convolutions, attention mechanisms, and GNNs. Furthermore, self-generating few-shot brain tumor segmentation models like CDSG-SAM are also part of the literature [43] where dynamic fuzzy support mask decoder module DFSMD were used within the to enhance the classification accuracy of BTs.

Despite the cited advancements in deep learning for brain tumor classification, a critical research gap persists in the **generation and diversity of synthetic MRI images**. Existing studies largely focus on the aspect of classification using CNNs, transformers, and hybrid architectures leading to sub-optimal generalization in terms of high quality and diverse tumor representations. The resolve of this study is to propose a dual- stream generator architecture in GANs, which incorporate contrastive learning and multi-stream feature fusion for enhancing the diversity and realism of synthetic MRI images. The present research is focus on leveraging dual-stream generator frameworks to ensure a balanced image synthesis.

The dual-stream generator and three discriminators were designed within an adversarial framework to enhance competitive learning. Traditional GAN-based augmentation suffers from mode collapse due to separate generators, a limitation addressed by our dual-stream architecture. The integration of a Contrastive Latent Projection (CLP) module preserves semantic consistency in augmented images, while contrastive learning ensures diverse yet coherent feature representations. We propose a novel objective function for DSCLPGAN, optimized for discriminative feature learning. Performance evaluation using assessment metrics demonstrates superior diversity and generalization in the generated images.

The main contributions of the paper are summarized as under;

- The proposed dual-stream augmentation framework utilizes a single generator with dual perturbations to enhance realism and diversity by effectively capturing both local and global variations in medical images.
- A rigorous mathematical formulation is developed, incorporating a CLP module to preserve semantic integrity and enhance model generalization in image augmentation tasks.



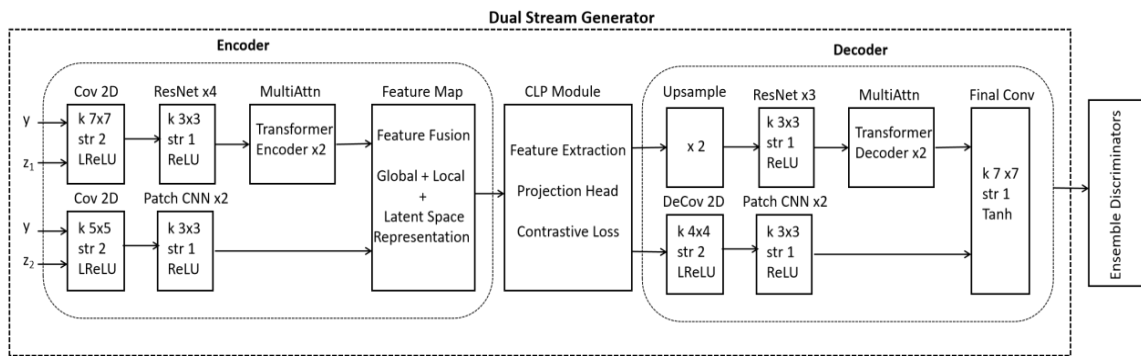
- A three-discriminator architecture is introduced, operating in parallel to assess image quality, diversity, and frequency consistency. Additionally, D1 performs classification, eliminating the need for a separate brain tumor (BT) classifier network.

## 2. Materials & Methods

A dual-stream single-generator with CLP is an efficient, robust, and medically meaningful way to augment MRI images while ensuring high-quality and diversity in synthetic images.

### 2.1. Dual Stream Generator of Our Proposed Model (DSCLPGAN)

Instead of using two separate generators, a single generator handles both augmentation streams. A single generator with dual perturbations avoids image similarity issues generated by two generators by maintaining diversity within one network. In our model, one stream applies local variations and other handles the global variations. An encoder extracts the latent representation from the input MRI scan. The latent space is split into two parallel streams with one to emphasize localized variations and the second to emphasize global changes as illustrated in Figure 1.



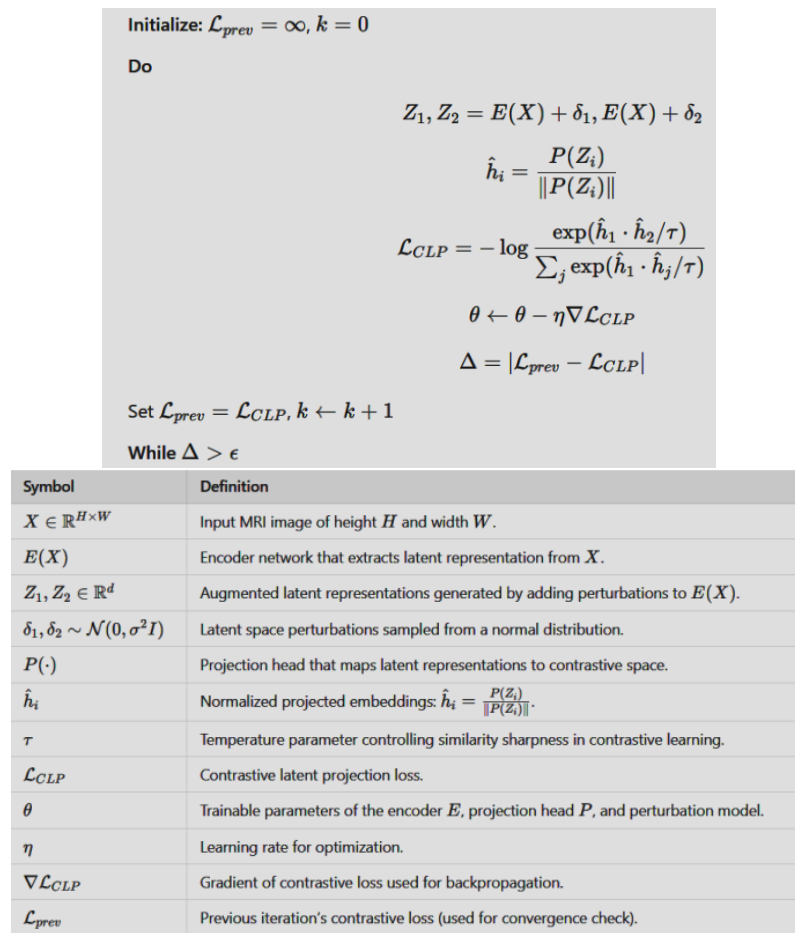
**Figure 1.** An illustration detailing the operations and blocks with an encoder, CLP module and decoder of a dual stream generator. .

As illustrated in Figure 1, the dual stream generator uses CNN/ transformer encoder for the extraction of global features and patch wise CNN for learning the local features within an image. The CLP module applies contrastive learning to project real and fake features into a contrastive latent space to ensure that the augmented images remain semantically close to the original. It applies contrastive learning using a similarity-based loss function as given in Eq. 1.

$$L_{CLP} = -\log \frac{\exp(\text{sim}(Z_1, Z_2)/\tau)}{\sum_j \exp(\text{sim}(Z_1, Z_j)/\tau)} \quad (1)$$

where  $\text{sim}(\cdot)$  is cosine similarity,  $\tau$  is temperature parameter controlling how strongly differences are penalized.

The do while pseudocode representation of CLP with input MRI image (X), encoder and projection head with perturbations is given in Figure 2 along with the description of involved parameters.

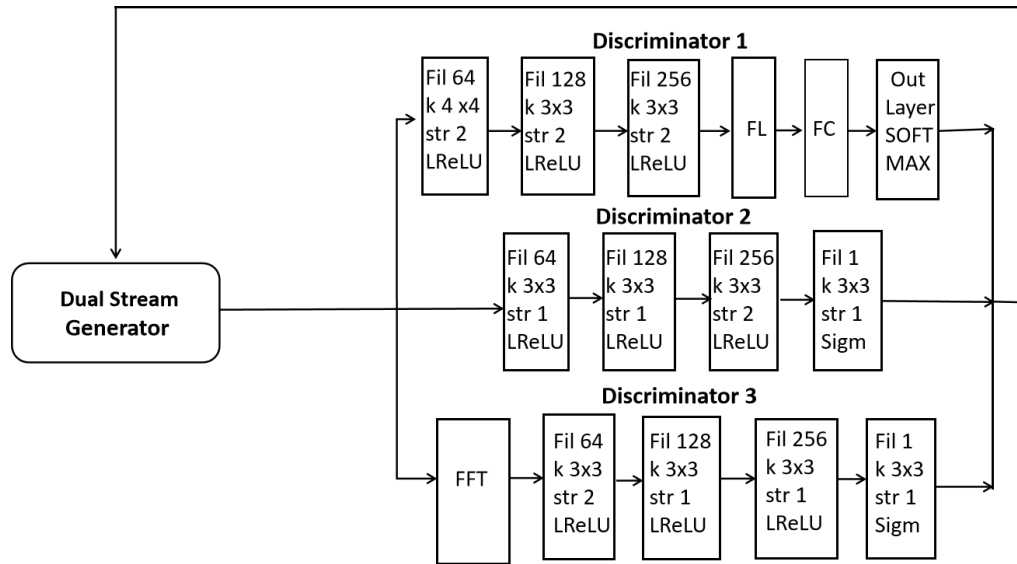


**Figure 2.** The do while code of CLP with input MRI image ( $X$ ), encoder and projection head along with the description of involved parameters.

The dual- stream decoder reconstructs the image using the global Reconstruction Path, local refinement path and skip connections as illustrated in Figure 1. The global path exploited ResNet/transformer decoder and local path involves the Deconv and CNN blocks. The final output includes high quality MRI images with structure & texture accuracy.

## 2.2. Complete Architecture of Our Proposed Model (DSCLPGAN)

In this paper, dual stream generator and triple discriminators are used for adversarial training of diverse synthetic images as presented Figure 3. The first discriminator  $D_1$  (global discriminator) is for the classification of BTs. The second discriminator  $D_2$  (local discriminator) is designed to analyze small patches within an image including specific tumor regions. The  $D_3$  converts images to frequency space for global consistency. A detailed schematic including the details of CNN and layers presented in Figure 3.



**Figure 3.** A schematic illustration of our proposed dual stream generator along with three discriminators.

### 2.3. Mathematical Formulation

The dual stream generator ( $G$ ) consists of a spatial stream  $G_s(z)$  to generate spatial features. The latent contrastive stream  $G_l(z)$  enforces latent space similarity via CLP. The generator learns to map latent noise to a realistic MRI image as given in Eq. 2.

$$x' = G(z) = G_s(z) + G_l(z) \quad (2)$$

where  $G_s(z)$  generates high resolution spatial details and  $G_l(z)$  enforces feature consistency.

$D_1$  learns to classify both the real and generated MRI images into medical conditions given in Eqs. 3- 4.

$$p(y_r|x_r) = D_1(x_r) \quad (3)$$

$$p(y'|x') = D_1(G(z)) \quad (4)$$

where  $y_r$  is the ground truth class.

The classification loss is governed by Eq. 5.

$$L_{cls} = -E_{x_r \sim P_{data}(x)} \sum_{i=1}^c y_r^i \log D_1^i(x_r) \quad (5)$$

where  $c$  represents the number of classes and  $y_r^i$  is ground truth class indicator.

The local  $D_2$  discriminator evaluates patches from real and fake MRI images and the adversarial loss is given by Eq. 6.

$$L_{D2} = -E_{x_r \sim P_{data}(x)} [\log D_2(x_r)] + E_{z \sim P_z(z)} [\log(1 - D_2(G(z)))] \quad (6)$$

MRI scans are evaluated in the Fourier domain by the  $D_3$  to ensure frequency consistency and the Fourier adversarial loss is computed using the expression in Eq. 7.

$$L_{D3} = -E_{x_r \sim P_{data}(x)} [\log D_3(F(x_r))] - E_{z \sim P_z(z)} [\log(1 - D_3(F(G(z))))] \quad (7)$$

The loss function is the aggregate loss total of all three discriminators, CLP loss and the hyper-parameters balancing each loss term and is given by Eq. 8.

$$L_D = \lambda_{cls} L_{D1} + \lambda_{adv} L_{D2} + \lambda_{freq} L_{D3} + \lambda_{clp} L_{clp} \quad (8)$$

The objective of training is to minimize the generator loss and to maximize the discriminators loss as indicated in Eq. 9.

$$G^* = \arg \min_G \max_D L_G \quad (9)$$

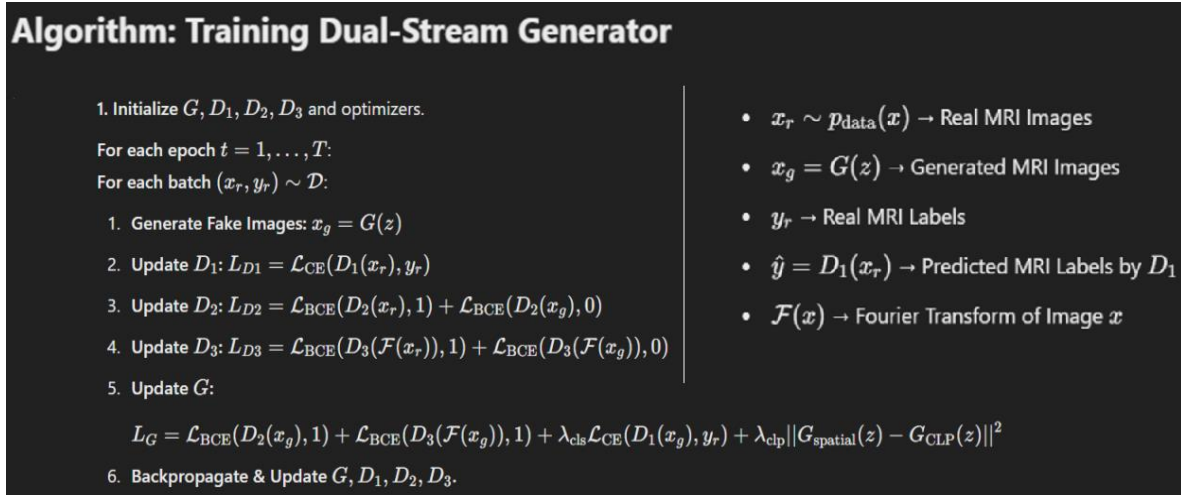
where  $L_G = \lambda_{cls} L_{D1} + \lambda_{adv} L_{D2} + \lambda_{freq} L_{D3} + \lambda_{clp} L_{clp}$

The ultimate objective is to produce MRI images that appear realistic at the global level with accurate and realistic local details and further maintain semantic consistency with the real image.

## 3. Results & Analysis

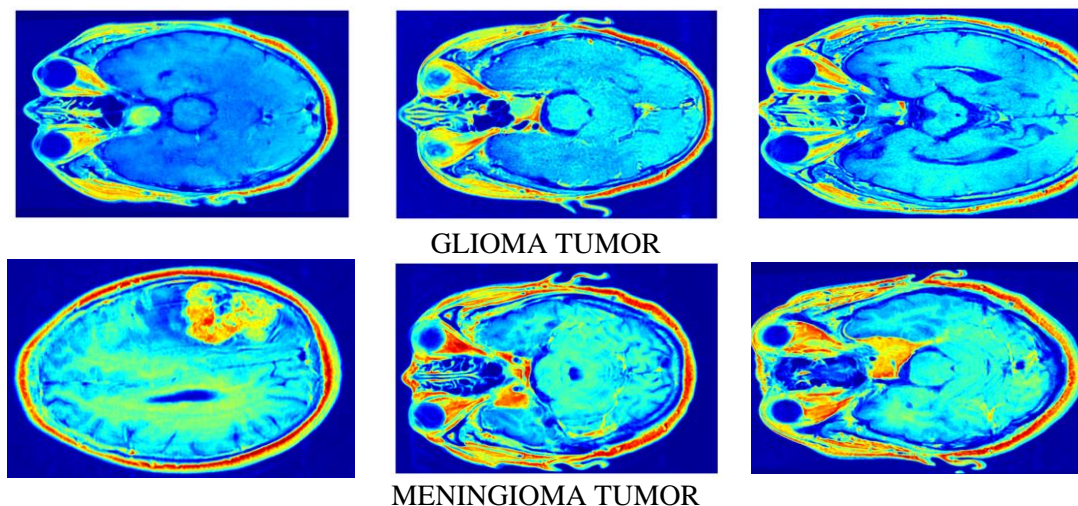
This brain tumor dataset (Figure Share 2024) contains 3064 T1-weighted contrast- enhanced images with three types of brain tumor: meningioma (708 slices), glioma (1426 slices), and pituitary

tumor (930 slices). The 5-fold cross-validation indices are also provided. Prior to augmentation and training phase, the MRI scans were subjected to a series of pre- processing stages for dataset standardization. Using our proposed **DSCLPGAN** architecture, the image data in each class is amplified by a factor 75 to increase the diversity and to validate the generalization ability of our proposed network. We test our GAN generated image data underlying objective to underscore the impact and extent of image augmentation in improvising diversity and generalization. All the models are implemented using Pytorch™. The four possible output labels are: 0- no tumor, 1- meningioma tumor, 2- glioma tumor, and 3- pituitary tumor. A latent dimension of 100 and batch size of 64 is selected for 50 epochs with learning rate of 0.0001. Adam optimizer is used for optimization with parameters  $\beta_1$  vale of 0.5 and  $\beta_2$  of 0.999. settings having The L3 **Regularization** is applied to the weights of the network to penalize large weights. The training process is exhibited in Figure 4 in which the dual-stream generator is trained by simultaneously optimizing two streams: one generating realistic augmented images and the other enforcing latent space consistency. An adversarial loss helps the generator produce high-quality outputs, while a reconstruction loss maintains fidelity to the original data.

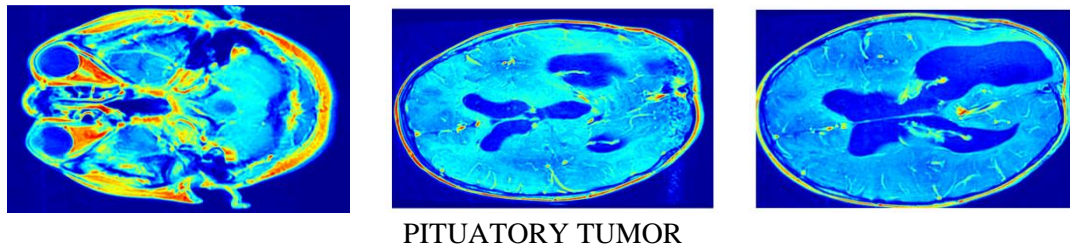


**Figure 4.** An algorithm used for training a dual stream generator in our proposed model along with details of key symbols used.

Few sample images generated by our model using the original MRI images are presented in Figure 5.

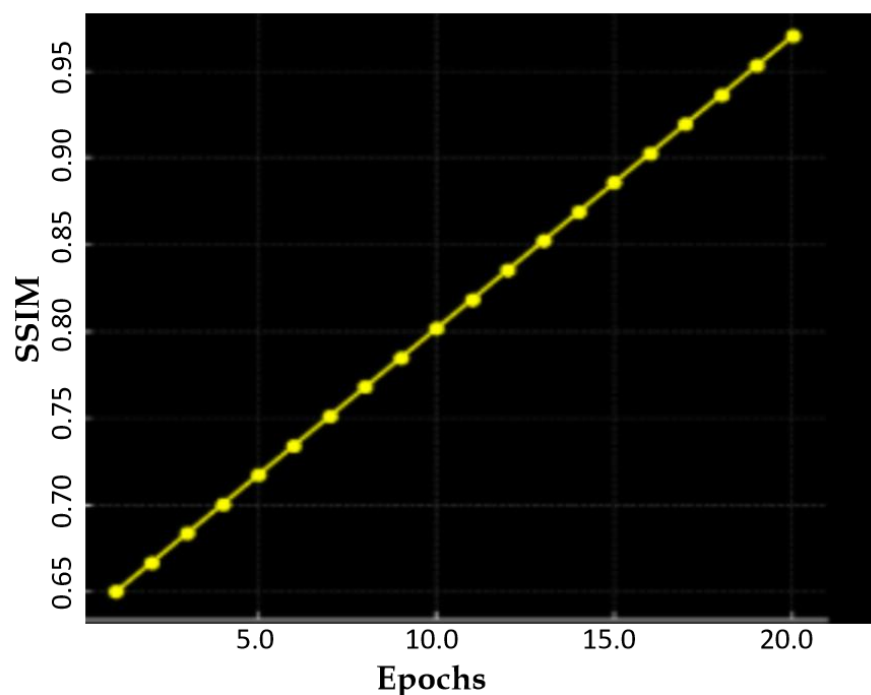






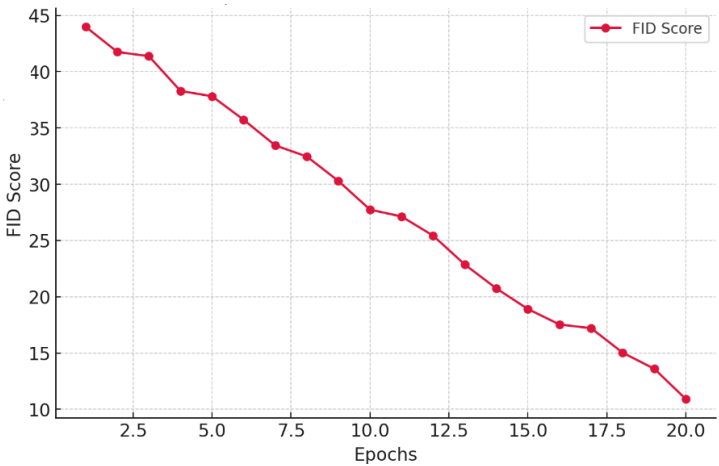
**Figure 5.** Synthetic images creation via our proposed DSCLPGAN for different brain tumors.

The Structural Similarity Index Measure (SSIM) loss shows a remarkable improvement from 0.65 to 0.99 over 20 epochs as indicated in Figure 6. This demonstrates the superior structural quality of our Dual-Stream Generator with CLP, hence ensuring that augmented MRI images remain highly realistic and clinically valuable. Such a high SSIM score reinforces the robustness of our method in generating high-quality medical images.



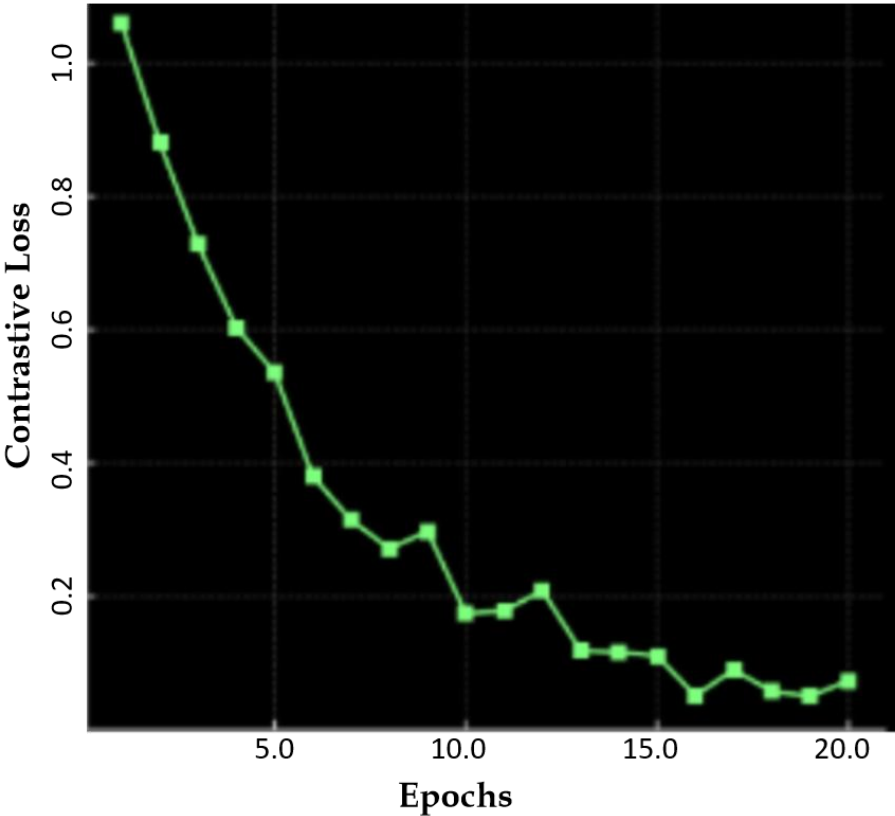
**Figure 6.** A plot of SSIM against the number of Epochs.

Similarly, the **Fréchet Inception Distance (FID)** in Figure 7 demonstrates a remarkable decline from 45 to 12 over 20 epochs. This signifies that our model progressively refines image quality, making synthetic data perceptually closer to real MRI scans. Such a low FID score highlights the effectiveness of our method in preserving critical medical features while ensuring diverse and realistic augmentation.



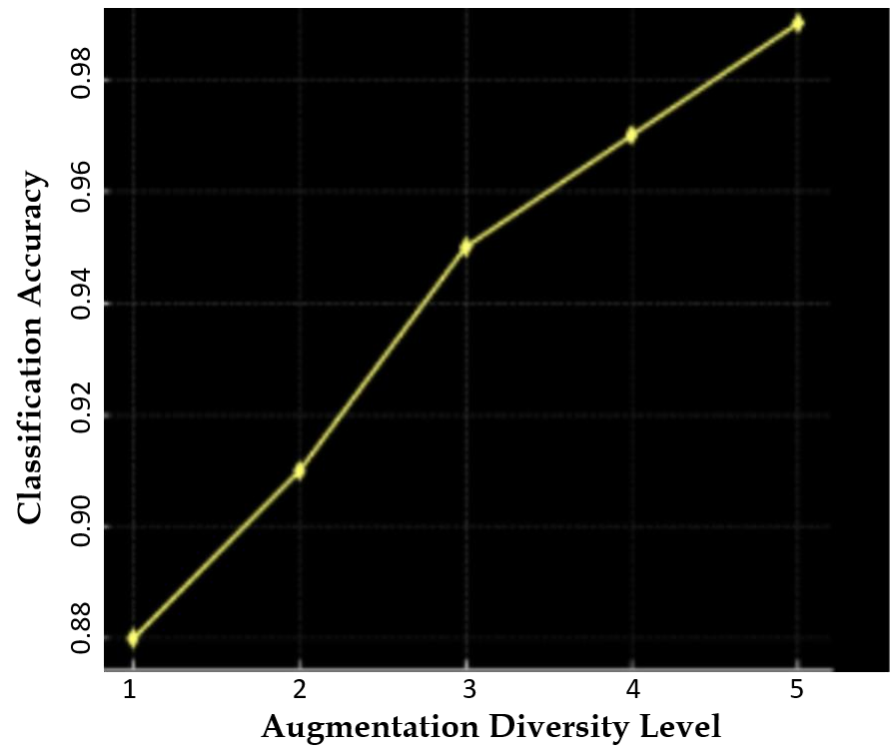
**Figure 7.** A plot of FID score against the number of Epochs.

The **Contrastive Loss** exhibits a value **0.97** after **20 epochs** as indicated in **Figure 8**. This reduction signifies improved feature consistency, ensuring that the generated images retain essential medical characteristics without losing diversity.



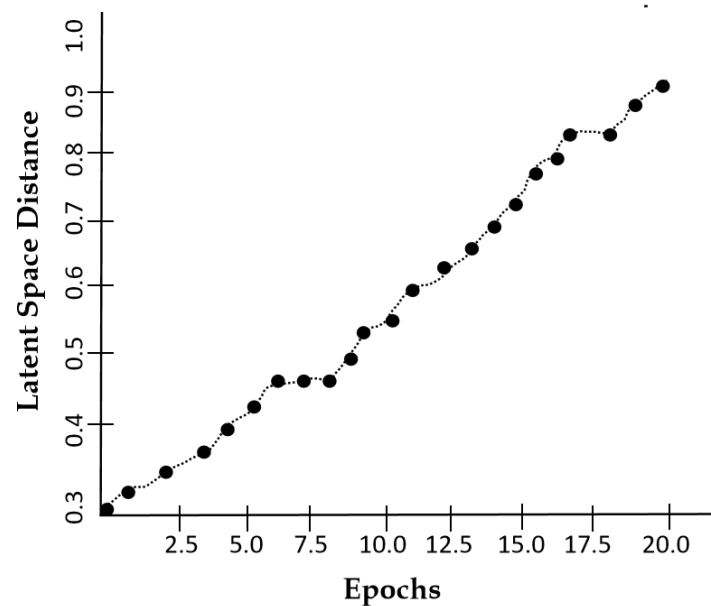
**Figure 8.** A comparison of contrastive loss of our model against the number of epochs.

Augmentation Diversity Level refers to the degree of variation introduced in synthetic data to enhance model generalization while preserving essential structural features. At an augmentation diversity level of 5, our model achieves an outstanding classification accuracy of 99.76% as illustrated in Figure 9. , proving its unparalleled ability to generate diverse yet highly realistic MRI images. This exceptional performance is attributed to the CLP module in our model for ensuring meaningful variations.

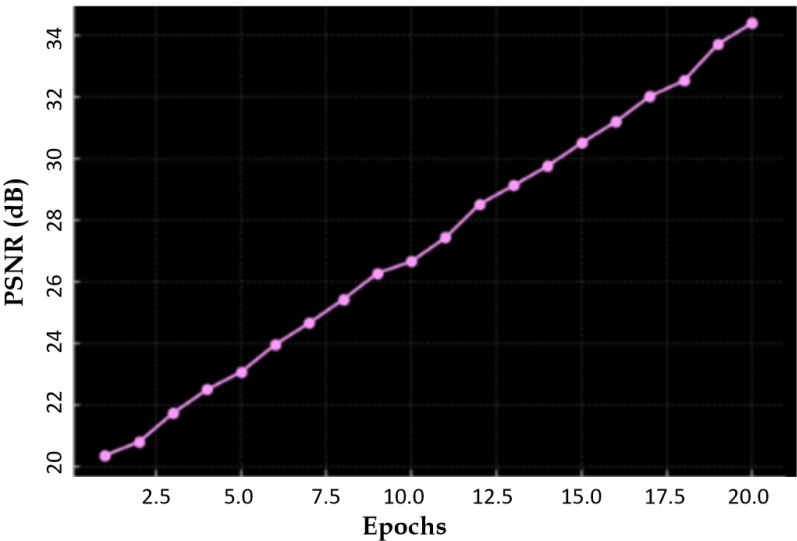


**Figure 9.** A plot of the classification accuracy of our model against augmentation diversity level.

Figure 10 suggests that the latent space distance shows an upward trend with the number of epochs which is reflecting the model’s ability to enhance feature distinction. The latent space distance ensures that the generated MRI images are not only diverse but also structurally coherent, preventing redundancy and mode collapse. Our model achieves PSNR of 34. 6 dB within 20 epochs as indicated in Figure 11, suggesting that our augmented images closely resemble the real MRI scans.



**Figure 10.** A plot of latent space distance against epochs using DSCLPGAN for feature distinction.



**Figure 11.** Peak Signal to Noise Ratio characteristics of our model against the number of epochs.

To highlight the effectiveness of the proposed DSCLPGAN, a comparative analysis is conducted against state-of-the-art GAN-based and transformer-based augmentation models and is illustrated in Table 1. The evaluation considers key performance metrics such as SSIM, PSNR and FID score. The models selected for comparison include BIGGAN [44], MAGE [45], TransGAN [46], SR TransGAN [47], CTGAN [48], StyleGANv2 [49], SFCGAN [50], VQ-GAN [51], and 3D Pix2Pix GAN [52]. An inter- comparison with such models having benchmark performance characteristics underscores the efficiency of DSCLPGAN in generating high quality and diverse medical images.

**Table 1.** A comparison of our proposed model (DSCLPGAN) with state- of- the- art methods for Image Synthesis.

Method	SSIM	FID	PSNR
BIGGAN [44]	0.7314	47.63	25.89
MAGE [45]	0.8220	45.62	27.28
TransGAN [46]	0.8376	35.45	27.66
SR TransGAN [47]	0.8504	31.29	30.28
CTGAN[48]	0.8755	29.10	26.47
StyleGANv2 [49]	0.8841	32.56	29.31
SFCGAN [50]	0.9077	28.04	29.14
VQ- GAN [51]	0.9166	26.55	31.04
3D Pix2Pix GAN [52]	0.9210	27.87	30.19
Proposed DSCLPGAN	0.9861	12	34.6

4. Conclusions

This study introduces a dual-stream GAN with CLP for MRI brain tumor image synthesis, ensuring structural fidelity and enhanced downstream performance. Experimental results demonstrate the model’s effectiveness across key metrics, including SSIM, FID, contrastive loss, and latent space consistency. The integration of triple discriminators refines image quality by evaluating realism from multiple perspectives while preserving latent space coherence. Notably, classification accuracy improves with increased augmentation diversity, highlighting the model’s superior generalization capabilities. These findings underscore its potential in data-limited medical imaging scenarios and its adaptability to broader clinical applications beyond MRI.

References



1. Kaifi R. A Review of Recent Advances in Brain Tumor Diagnosis Based on AI-Based Classification. *Diagnostics* (Basel). 2023 Sep 20;13(18):3007.
2. Fan, Y., Zhang, X., Gao, C. et al. Burden and trends of brain and central nervous system cancer from 1990 to 2019 at the global, regional, and country levels. *Arch Public Health* 80, 209 (2022).
3. Arnaout, M.M., Hoz, S., Lee, A. et al. Management of patients with multiple brain metastases. *Egypt J Neurosurg* 39, 64 (2024).
4. Louis DN, Perry A, Wesseling P, Brat DJ, Cree IA, Figarella-Branger D, Hawkins C, Ng HK, Pfister SM, Reifenberger G, Soffietti R, von Deimling A, Ellison DW. The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. *Neuro Oncol.* 2021 Aug 2;23(8):1231-1251.
5. <https://www.hopkinsmedicine.org/health/conditions-and-diseases/brain-tumor>
6. Louis, D.N., Perry, A., Reifenberger, G. et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathol* 131, 803–820 (2016).
7. Delgado-López, P.D., Corrales-García, E.M. Survival in glioblastoma: a review on the impact of treatment modalities. *Clin Transl Oncol* 18, 1062–1071 (2016).
8. Abdusalomov, A.B.; Mukhiddinov, M.; Whangbo, T.K. Brain Tumor Detection Based on Deep Learning Approaches and Magnetic Resonance Imaging. *Cancers* 2023, 15, 4172.
9. Lazli, L.; Boukadoum, M.; Mohamed, O.A. A Survey on Computer-Aided Diagnosis of Brain Disorders through MRI Based on Machine Learning and Data Mining Methodologies with an Emphasis on Alzheimer Disease Diagnosis and the Contribution of the Multimodal Fusion. *Appl. Sci.* 2020, 10, 1894.
10. Virupakshappa, Amarapur, B. Computer-aided diagnosis applied to MRI images of brain tumor using cognition based modified level set and optimized ANN classifier. *Multimed Tools Appl* 79, 3571–3599 (2020).
11. Das, S., Goswami, R.S. Advancements in brain tumor analysis: a comprehensive review of machine learning, hybrid deep learning, and transfer learning approaches for MRI-based classification and segmentation. *Multimed Tools Appl* (2024)
12. Md. Naim Islam, Md. Shafiul Azam, Md. Samiul Islam, Muntasir Hasan Kanchan, A.H.M. Shahariar Parvez, Md. Monirul Islam, An improved deep learning-based hybrid model with ensemble techniques for brain tumor detection from MRI image, *Informatics in Medicine*, Volume 47, 2024, 101483.
13. Amran, G.A.; Alsharam, M.S.; Blajam, A.O.A.; Hasan, A.A.; Alfaifi, M.Y.; Amran, M.H.; Gumaei, A.; Eldin, S.M. Brain Tumor Classification and Detection Using Hybrid Deep Tumor Network. *Electronics* 2022, 11, 3457.
14. S. Karim *et al.*, "Developments in Brain Tumor Segmentation Using MRI: Deep Learning Insights and Future Perspectives," in *IEEE Access*, vol. 12, pp. 26875-26896, 2024,
15. Sajjanar, R., Dixit, U.D. & Vagga, V.K. Advancements in hybrid approaches for brain tumor segmentation in MRI: a comprehensive review of machine learning and deep learning techniques. *Multimed Tools Appl* 83, 30505–30539 (2024).
16. Prakash, R. M., Kumari, R. S. S., Valarmathi, K., & Ramalakshmi, K. (2022). Classification of brain tumours from MR images with an enhanced deep learning approach using densely connected convolutional network. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 11(2), 266–277.
17. <https://onlinelibrary.wiley.com/doi/abs/10.1002/ima.22975>
18. Mathivanan, S.K., Sonaimuthu, S., Murugesan, S. et al. Employing deep learning and transfer learning for accurate brain tumor detection. *Sci Rep* 14, 7232 (2024).
19. S. Katkam, V. Prema Tulasi, B. Dhanalaxmi and J. Harikiran, "Multi-Class Diagnosis of Neurodegenerative Diseases Using Effective Deep Learning Models With Modified DenseNet-169 and Enhanced DeepLabV 3+," in *IEEE Access*, vol. 13, pp. 29060-29080, 2025.
20. S. Ahmad and P. K. Choudhury, "On the Performance of Deep Transfer Learning Networks for Brain Tumor Detection Using MR Images," in *IEEE Access*, vol. 10, pp. 59099-59114, 2022.
21. Mok, T.C.W., Chung, A.C.S. (2019). Learning Data Augmentation for Brain Tumor Segmentation with Coarse-to-Fine Generative Adversarial Networks. In: Crimi, A., Bakas, S., Kuijf, H., Keyvan, F., Reyes, M.,

- van Walsum, T. (eds) Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018. Lecture Notes in Computer Science(), vol 11383. Springer, Cham.
22. <https://onlinelibrary.wiley.com/doi/abs/10.1002/cpe.7031>
  23. Han, C. et al. (2020). Infinite Brain MR Images: PGGAN-Based Data Augmentation for Tumor Detection. In: Esposito, A., Faundez-Zanuy, M., Morabito, F., Pasero, E. (eds) Neural Approaches to Dynamics of Signal Exchanges. Smart Innovation, Systems and Technologies, vol 151. Springer, Singapore.
  24. Goceri, E. Medical image data augmentation: techniques, comparisons and interpretations. *Artif Intell Rev* 56, 12561–12605 (2023).
  25. Sanaat, A., Shiri, I., Ferdowsi, S. et al. Robust-Deep: A Method for Increasing Brain Imaging Datasets to Improve Deep Learning Models' Performance and Robustness. *J Digit Imaging* 35, 469–481 (2022).
  26. Gab Allah, A.M.; Sarhan, A.M.; Elshennawy, N.M. Classification of Brain MRI Tumor Images Based on Deep Learning PGGAN Augmentation. *Diagnostics* 2021, 11, 2343.
  27. <https://arxiv.org/abs/2411.00875>
  28. Asiri, A.A.; Shaf, A.; Ali, T.; Aamir, M.; Irfan, M.; Alqahtani, S.; Mehdar, K.M.; Halawani, H.T.; Alghamdi, A.H.; Alshamrani, A.F.A.; et al. Brain Tumor Detection and Classification Using Fine-Tuned CNN with ResNet50 and U-Net Model: A Study on TCGA-LGG and TCIA Dataset for MRI Applications. *Life* 2023, 13, 1449.
  29. Andrés Anaya-Isaza, Leonel Mera-Jiménez, Lucía Verdugo-Alejo, Luis Sarasti, Optimizing MRI-based brain tumor classification and detection using AI: A comparative analysis of neural networks, transfer learning, data augmentation, and the cross-transformer network, *European Journal of Radiology Open*, Volume 10, 2023, 100484.
  30. A. M. Kocharekar, S. Datta, Padmanaban and R. R., "Comparative Analysis of Vision Transformers and CNN-based Models for Enhanced Brain Tumor Diagnosis," *2024 3rd International Conference on Automation, Computing and Renewable Systems (ICACRS)*, Pudukkottai, India, 2024, pp. 1217-1223,
  31. Zakariah, Mohammed et al, Dual vision Transformer-DSUNET with feature fusion for brain tumor segmentation *Heliyon*, Volume 10, Issue 18, e37804, 2024.
  32. L. ZongRen, W. Silamu, W. Yuzhen and W. Zhe, "DenseTrans: Multimodal Brain Tumor Segmentation Using Swin Transformer," in *IEEE Access*, vol. 11, pp. 42895-42908, 2023, doi: 10.1109/ACCESS.2023.3272055.
  33. Swetha, A. V. S., Bala, M., & Sharma, K. (2024). A Linear Time Shrinking-SL(t)-ViT Approach for Brain Tumor Identification and Categorization. *IETE Journal of Research*, 70(11), 8300–8322.
  34. Sadafossadat Tabatabaei, Khosro Rezaee, Min Zhu, Attention transformer mechanism and fusion-based deep learning architecture for MRI brain tumor classification system, *Biomedical Signal Processing and Control*, Volume 86, Part A, 2023, 105119, ISSN 1746-8094.
  35. Pacal, I. A novel Swin transformer approach utilizing residual multi-layer perceptron for diagnosing brain tumors in MRI images. *Int. J. Mach. Learn. & Cyber.* 15, 3579–3597 (2024).
  36. Tehsin, S.; Nasir, I.M.; Damaševičius, R. GATransformer: A Graph Attention Network-Based Transformer Model to Generate Explainable Attentions for Brain Tumor Detection. *Algorithms* 2025, 18, 89.
  37. Usman Akbar, M., Larsson, M., Blystad, I. et al. Brain tumor segmentation using synthetic MR images - A comparison of GANs and diffusion models. *Sci Data* 11, 259 (2024).
  38. A. Saeed, K. Shehzad, S. S. Bhatti, S. Ahmed and A. T. Azar, "GGLA-NeXtE2NET: A Dual-Branch Ensemble Network With Gated Global-Local Attention for Enhanced Brain Tumor Recognition," in *IEEE Access*, vol. 13, pp. 7234-7257, 2025.
  39. Karpakam, S., Kumareshan, N. Enhanced brain tumor detection and classification using a deep image recognition generative adversarial network (DIR-GAN): a comparative study on MRI, X-ray, and FigShare datasets. *Neural Comput & Applic* (2025). <https://doi.org/10.1007/s00521-025-11031-w>
  40. L. Desalegn and W. Jifara, "HARA-GAN: Hybrid Attention and Relative Average Discriminator Based Generative Adversarial Network for MR Image Reconstruction," in *IEEE Access*, vol. 12, pp. 23240-23251, 2024

41. Lyu Y, Tian X. MWG-UNet++: Hybrid Transformer U-Net Model for Brain Tumor Segmentation in MRI Scans. *Bioengineering* (Basel). 2025 Jan 31;12(2):140. doi: 10.3390/bioengineering12020140. PMID: 40001660; PMCID: PMC11852190.
42. Ahmed, S., Feng, J., Ferzund, J. et al. FAME: A Federated Adversarial Learning Framework for Privacy-Preserving MRI Reconstruction. *Appl Magn Reson* (2025).
43. Yang Yang, Xianjin Fang, Xiang Li, Yuxi Han, Zekuan Yu, CDSG-SAM: A cross-domain self-generating prompt few-shot brain tumor segmentation pipeline based on SAM, *Biomedical Signal Processing and Control*, Volume 100, Part B, 2025, 106936, ISSN 1746-8094.
44. J. Donahue and K. Simonyan, "Large scale adversarial representation learning," *Advances in neural information processing systems*, vol. 32, 2019.
45. T. Li, H. Chang, S. K. Mishra, H. Zhang, D. Katabi, and D. Krishnan, "Mage: Masked generative encoder to unify representation learning and image synthesis," 2022.
46. <https://arxiv.org/abs/2102.07074>
47. <https://arxiv.org/abs/2312.01999>
48. Xu M, Cui J, Ma X, Zou Z, Xin Z, Bilal M. Image enhancement with art design: a visual feature approach with a CNN-transformer fusion model. *PeerJ Comput Sci*. 2024 Nov 13;10:e2417
49. Yurtsever MME, Atay Y, Arslan B, Sagioglu S. Development of brain tumor radiogenomic classification using GAN-based augmentation of MRI slices in the newly released gazi brains dataset. *BMC Med Inform Decis Mak*. 2024 Oct 4;24(1):285.
50. <https://arxiv.org/abs/2501.07055v1>
51. Zhou M, Wagner MW, Tabori U, Hawkins C, Ertl-Wagner BB, Khalvati F. Generating 3D brain tumor regions in MRI using vector-quantization Generative Adversarial Networks. *Comput Biol Med*. 2025 Feb;185:109502.
52. <https://arxiv.org/abs/2412.11849>

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.