RESEARCH

Development of brain tumor radiogenomic classification using GAN-based augmentation of MRI slices in the newly released gazi brains dataset

M.M.Enes Yurtsever[1], Yilmaz Atay[2*], Bilgehan Arslan[2] and Seref Sagiroglu[2]

Abstract

Significant progress has been made recently with the contribution of technological advances in studies on brain cancer. Regarding this, identifying and correctly classifying tumors is a crucial task in the field of medical imaging. The disease-related tumor classification problem, on which deep learning technologies have also become a focus, is very important in the diagnosis and treatment of the disease. The use of deep learning models has shown promising results in recent years. However, the sparsity of ground truth data in medical imaging or inconsistent data sources poses a significant challenge for training these models. The utilization of StyleGANv2-ADA is proposed in this paper for augmenting brain MRI slices to enhance the performance of deep learning models. Specifically, augmentation is applied solely to the training data to prevent any potential leakage. The StyleGanv2-ADA model is trained with the Gazi Brains 2020, BRaTS 2021, and Br35h datasets using the researchers' default settings. The effectiveness of the proposed method is demonstrated on datasets for brain tumor classification, resulting in a notable improvement in the overall accuracy of the model for brain tumor classification on all the Gazi Brains 2020, BraTS 2021, and Br35h datasets. Importantly, the utilization of StyleGANv2-ADA on the Gazi Brains 2020 Dataset represents a novel experiment in the literature. The results show that the augmentation with StyleGAN can help overcome the challenges of working with medical data and the sparsity of ground truth data. Data augmentation employing the Style-GANv2-ADA GAN model yielded the highest overall accuracy for brain tumor classification on the BraTS 2021 and Gazi Brains 2020 datasets, together with the BR35H dataset, achieving 75.18%, 99.36%, and 98.99% on the EfficientNetV2S models, respectively. This study emphasizes the potency of GANs for augmenting medical imaging datasets, particularly in brain tumor classification, showcasing a notable increase in overall accuracy through the integration of synthetic GAN data on the used datasets.

Keywords Augmentation, Brain tumors, Classification, CNNs, GANs, Medical imaging, Transformer

Introduction

Brain tumors represent a complex and heterogeneous class of neoplasms that may arise from diverse cell types within the central nervous system. This type of disease is quite critical and it requires early diagnosis and treatment processes. Offering comprehensive insights, the CBTRUS Statistical Report delves into the occurrence, fatality rates, and relative survival rates concerning

*Correspondence:
Yilmaz Atay
yilmazatay@gazi.edu.tr
[1] Information Systems Engineering, Kocaeli University, Kocaeli 41001, Türkiye
[2] Computer Engineering, Engineering Faculty, Gazi University, Ankara 06570, Türkiye

primary malignant and non-malignant brain and other CNS tumors in the United States between 2015 and 2019 [1]. According to the CBTRUS Statistical Report, the five-year relative survival rate following diagnosis of a malignant brain and other CNS tumor was 35.7% in the United States. Survival following diagnosis with a malignant brain and other CNS tumor was highest in persons ages 0-14 years (75.1%) and ages 15-39 years (71.7%) as compared to those ages 40+ years (21.0%). To increase the survival rate of malignant brain tumors, early detection and treatment are crucial. Treatment options for brain tumors may include surgery, radiation therapy, chemotherapy, targeted therapy, or a combination of these approaches.

Tumors in the brain refer to abnormal growth of cells, and while there are different definitions in the literature regarding their classification, they can be examined basically in two types: benign, which is non-cancerous, and malignant, which is cancerous. Malignant brain tumors are of particular concern because they grow rapidly, can spread to other parts of the brain and spine, and pose a serious threat to life. Therefore, there is an urgent need for accurate and reliable brain tumor detection and classification systems that can aid in diagnosis, treatment planning, and monitoring of disease progression. Detecting a tumor at an early stage is crucial for effective treatment, and one of the primary diagnostic approaches employed by doctors is the use of various imaging methods, as they can quickly identify the presence of a tumor. Detecting cancer at an early stage is crucial for effective treatment, and one of the primary diagnostic approaches employed by doctors is the use of various imaging methods, as they can quickly identify the presence of a tumor [2].

Magnetic resonance imaging (MRI) is a widely used source to diagnose brain tumors, which can be life-threatening [3]. Misinterpretation of a brain tumor can lead to major complications and decrease a patient's chances of survival [4]. MRI is a widely employed medical imaging tool used to visualize abnormal tissues within the body. MRI is widely recognized for its numerous benefits in helping doctors detect physical brain abnormalities. During the process of acquiring MRI images, a series of 2D images can collectively represent a 3D volume of the brain. Each MRI modality serves a distinct purpose in the diagnostic process. T2 images are particularly useful for identifying areas with edema, while T1 images excel in distinguishing healthy tissues. T1-Gd images, on the other hand, are essential for delineating tumor borders. Additionally, fluid-attenuated inversion recovery (FLAIR) images prove valuable in distinguishing edematous regions from cerebrospinal fluid (CSF). The variations among the images generated by these different MRI

modalities can be harnessed to produce various types of contrast images. In standard diagnostic practice, four primary MRI modalities are employed: FLAIR, T2-weighted MRI (T2), T1-weighted MRI (T1), and T1-weighted MRI with gadolinium contrast enhancement (T1-Gd) [5].

Diverse techniques and methodologies have been devised for the segmentation and categorization of brain tumor images, and this field remains an active area of research due to the paramount importance of achieving high accuracy. Each year, it is observed that novel segmentation techniques have emerged aimed at addressing the shortcomings of earlier methods. Notably, deep learning-based approaches are currently regarded as the most effective means for identifying and extracting MRI image features for classification and segmentation purposes.

Brain tumor classification is a challenging task in the field of medical imaging, with accurate diagnosis and treatment dependent on the identification of different tumor types. As mentioned earlier, deep learning methods have shown great potential in improving the accuracy of tumor classification, but the scarcity of ground truth data presents a significant challenge for training these models. This is especially true in medical imaging, where obtaining labeled data is both difficult and time-consuming. Data augmentation has been widely used to address the issue of data scarcity, but traditional methods, such as flipping or rotating images, may not be effective for medical imaging due to the complex structures and irregular shapes of the organs. Recently, Generative Adversarial Networks (GANs) have shown great potential for data augmentation in various applications, including medical imaging.

Data augmentation plays a crucial role in medical imaging studies, particularly in tasks such as brain tumor classification and segmentation. The limited availability of annotated medical imaging datasets poses a challenge for training accurate and robust deep learning models. However, data augmentation techniques can help address this issue by artificially expanding the dataset and increasing its diversity [6]. By applying various augmentation operations, such as rotation, scaling, flipping, and cropping, to the original images, researchers can generate additional training samples that capture different variations of the tumors and their surrounding structures [7]. This augmented dataset can then be used to train deep learning models, improving their generalization and performance [6].

This paper is organized into sections as follows. An overview of related works in the field of medical imaging and brain tumor classification is provided in "Related works" section. The existing literature on various classifiers utilized in medical imaging tasks is discussed, and

Yurtsever *et al. BMC Medical Informatics and Decision Making*   (2024) 24:285

Page 3 of 24

their strengths and limitations are highlighted. Additionally, several relevant research papers in the field of brain tumor classification are selected. A comprehensive table is employed to compare the methodologies used in these papers, including information such as the methods applied, the specific problem addressed, the datasets utilized, and the results achieved by each classifier. The comparative outcomes are analyzed and discussed, emphasizing the strengths and limitations of this approach to the existing literature. In "Material and methods" section, the materials and methods employed in the study are presented. The data preprocessing steps outlined in the previous section, involving standardizing MRI slices to a dimension of 256x256 pixels and removing unnecessary slices, are described. Subsequently, the GAN model used for data augmentation, namely StyleGANv2-ADA, which has demonstrated promising results in generating synthetic brain MRI slices, is introduced. The integration of GAN-generated data into the preprocessed dataset to enhance the training process and potentially improve the overall accuracy of the CNN classifiers is explained. As part of the analysis, the impact of GAN-based data augmentation on the overall accuracy of the CNN classifiers is specifically addressed. The introduction of synthetic samples through GANs and its potential enhancement of the model's ability to learn discriminative features and improve classification performance are discussed. The observed effects are carefully evaluated and any potential challenges or caveats associated with this approach are highlighted. Finally, in the results section, conclusions are presented, summarizing the key findings from the study. The significance of GAN-based data augmentation in the context of brain tumor classification is emphasized, along with discussions on implications for future research in medical imaging. The importance of addressing the limitations of the approach is also highlighted, and potential directions for further investigation to enhance the accuracy and robustness of CNN classifiers in this domain are suggested.

Related works

MRI-based brain tumor classification and segmentation have been extensively studied in the literature. Numerous research studies have focused on developing accurate and efficient methods for analyzing brain MRI images to aid in diagnosing and treating brain tumors. In the field of brain tumor classification, various approaches have been proposed. Traditional machine learning models follow a conventional approach, involving crucial steps like feature selection, extraction, and reduction before the actual classification process. These methods have proven effective in a range of medical imaging tasks, including the classification of liver [8, 9], thyroid [10], and plaque [11].

These machine learning approaches leverage (a) feature selection in conjunction with (b) classification techniques to achieve their objectives.

MRI-based brain tumor classification has been the subject of extensive research in recent years [12–16]. These studies have focused on developing deep learning models, particularly convolutional neural networks (CNNs), for accurate and efficient classification of brain tumor images. With the advent of deep learning, CNNs have emerged as powerful tools for brain tumor classification. CNNs can automatically learn hierarchical representations from raw MRI data, enabling more accurate and robust tumor classification. These models have demonstrated superior performance compared to traditional machine learning approaches, achieving high accuracy in tumor classification tasks. This eventually led to a preference for using deep learning over traditional machine learning techniques [5]. In the research conducted by Ertosun and Rubin [17], they utilized CNNs to categorize different grades of gliomas in pathological images, specifically distinguishing between Grade II, Grade III, and Grade IV. They achieved classification accuracies of 71% for this task. [18] propose a deep learning-based method for microscopic brain tumor detection and classification. Their approach utilizes a 3D CNN and feature selection architecture for accurate tumor detection and classification. The study conducted by [19] introduces three CNN models for multi-classifying brain tumors. The first CNN model achieves 99.33% accuracy in brain tumor detection, the second classifies tumors into five types with 92.66% accuracy, and the third classifies tumors into three grades with 98.14% accuracy. These CNN models automatically optimize hyperparameters using a grid search algorithm, a novel approach in this field. The models are compared to state-of-the-art counterparts, offering satisfactory classification results, and can assist physicians and radiologists in initial brain tumor screenings.

There are also brain tumor segmentation studies that aim to improve diagnosis and treatment while ensuring patient data privacy. The researchers [20] propose a federated learning framework to overcome the limitations of traditional centralized methods, which are restricted by privacy regulations. This framework allows for collaborative learning across multiple medical institutions without sharing raw data, using a U-Net-based model architecture known for its effectiveness in segmentation tasks. Experimental results indicate that this approach significantly enhances performance metrics, including a specificity of 0.96 and a dice coefficient of 0.89, particularly with an increased number of clients. The proposed method also outperforms existing CNN- and RNN-based models, advancing medical image segmentation while maintaining data security.

Yurtsever *et al. BMC Medical Informatics and Decision Making*   (2024) 24:285

Page 4 of 24

Gliomas, the most common and deadly malignant brain tumors, require robust segmentation methods due to the limitations of manual segmentation, which is costly, time-consuming, and prone to errors based on the radiologist's experience. To address these challenges, researchers [21] have developed semi-automatic and fully automatic segmentation algorithms using machine learning (ML) techniques, including handcrafted feature-based methods and data-driven strategies such as convolutional neural networks. A novel cascaded approach has been proposed, combining the strengths of handcrafted features with CNN-based methods by intelligently incorporating prior information from ML algorithms into CNN models. This approach utilizes a Global Convolutional Neural Network (GCNN) with two parallel CNNs, CSPathways CNN (CSPCNN) and MRI Pathways CNN (MRIPCNN), to achieve high accuracy in segmenting brain tumors from BraTS datasets, resulting in a Dice score of 87%, surpassing the current state-of-the-art methods. This advancement offers significant potential to improve brain tumor segmentation, aiding in more effective diagnosis and treatment.

One of the major challenges in developing accurate brain tumor classification models is the limited availability of annotated medical imaging data. Collecting and annotating large-scale datasets is a time-consuming and expensive process, especially in the medical domain, where expert knowledge is required. This scarcity of data poses a significant obstacle to training deep learning models effectively. To address this issue, data augmentation techniques have been proposed to artificially increase the size and diversity of the training dataset. Data augmentation involves applying various transformations to the existing data, such as rotation, scaling, flipping, and adding noise, to generate new samples that retain the characteristics of the original data. By augmenting the training dataset, the models can learn from a more diverse set of examples, leading to improved generalization and performance. In the context of MRI-based brain tumor classification, data augmentation techniques effectively enhance the performance of deep learning models [22, 23]. By generating synthetic data that simulates the variability and complexity of real tumor images, these techniques can help overcome the limitations of limited annotated data. Augmentation methods such as rotation, translation, and elastic deformation have been applied to brain MRI images, resulting in improved classification accuracy. Additionally, data augmentation can help to address class imbalance issues by generating synthetic samples for underrepresented classes, thereby improving the overall performance of the models [24].

In addressing the challenges of diagnosing canine mammary tumors (CMTs) and their potential as models for human breast cancer, [25] introduces the first publicly available dataset of CMT histopathological images (CMTHis). Recognizing the tedious nature of the histopathological analysis, the researchers propose a VGG-Net-16-based framework, evaluating its performance on both the CMT dataset (CMTHis) and the human breast cancer dataset (BreakHis). The study explores the impact of data augmentation, stain normalization, and magnification on the framework's effectiveness. With support vector machines, the proposed framework achieves mean accuracies of 97% and 93% for binary classification of human breast cancer and CMT, respectively, underscoring the efficacy of the automated system in histopathological image analysis.

Recent advancements in deep learning for computer-aided medical diagnosis (CAD) systems targeting leukemia detection have highlighted the crucial role of data augmentation techniques. The study conducted by [26], focusing on acute and chronic leukemia subtypes, specifically myeloid and lymphoid leukemia, evaluates the impact of data augmentation and multilevel and ensemble configurations on CNNs. Utilizing 3,536 images across 18 datasets in five scenarios, the research demonstrates that data augmentation significantly enhances CNN performance. Ensemble configurations outperform multilevel setups in binary classification, while both exhibit comparable results in multiclass scenarios. This work provides valuable insights into optimizing CAD systems, achieving accuracies of 94.73% and 94.59% in multilevel and ensemble configurations, respectively, in a four-class scenario, thus advancing the effectiveness of deep learning-based models for accurate leukemia diagnosis.

Recent advancements in generative adversarial networks (GANs) have opened up new possibilities for data augmentation in medical imaging. GANs are deep learning models that consist of a generator network and a discriminator network [27]. The generator network learns to generate synthetic data that closely resemble the real data, while the discriminator network learns to distinguish between real and synthetic samples. GAN-based augmentation techniques have shown promise in improving the performance of brain tumor classification models. By training GANs on existing brain MRI images, synthetic tumor images can be generated, which can then be used to augment the training dataset. These synthetic images can capture the variability and complexity of real tumor images, enabling the models to learn more robust and discriminative features.

Allah et al. [28] conducted a study on the classification of brain MRI tumor images using deep learning and GAN-based augmentation. They employed a VGG19 features extractor coupled with different types of classifiers

Yurtsever *et al. BMC Medical Informatics and Decision Making*   (2024) 24:285

Page 5 of 24

to examine the efficacy of their approach. The results showed that the GAN-based augmentation approach improved the classification performance compared to traditional methods. In addition to brain tumor classification, GAN-based augmentation techniques have been applied to other medical imaging tasks as well. For example, [29] used GANs for brain MR image augmentation to improve tumor detection. They employed both noise-to-image and image-to-image GANs to handle small and fragmented datasets from multiple scanners, resulting in improved classification performance. Furthermore, GAN-based augmentation has been used in the segmentation of medical images, including brain tumor segmentation.

Iqbal et al. [30] proposed a new GAN-based approach for the generation of retinal vessel images and their segmented masks. They presented the MI-GAN model for the synthesis of retinal images, which achieves state-of-the-art performance on the STARE and DRIVE datasets. The proposed method generates precise segmented images better than existing techniques, using a predefined set of loss functions to achieve better generation and discrimination ability.

Mok et al. [31] proposed a method that uses GANs to learn augmentations for brain tumor segmentation. Their approach leverages GANs to generate diverse and realistic synthetic images, enabling more efficient learning from limited annotated samples. Data augmentation is another area where GANs have been applied to improve brain tumor classification. Kim et al. [7] synthesized brain tumor multi-contrast MR images using GANs for improved data augmentation. By generating synthetic tumor images, GANs can enhance the diversity and quantity of training data, leading to better classification performance. The combination of GANs with other deep-learning techniques has also been explored for brain tumor classification. Asiri et al. [32] proposed a multi-level deep GAN for brain tumor classification on magnetic resonance images. By integrating GANs with deep learning models, the proposed method achieved improved classification accuracy, demonstrating the potential of GAN-based approaches in brain tumor classification.

In their study, Garcea et al. [33] reviewed over 300 studies and presented a comprehensive literature review and summary of data augmentation approaches used in the training of AI models for classification, segmentation, and lesion detection in the medical field. They generally categorized data augmentation methods into two main categories: transformation and generation-based augmentation. They examined studies regarding the brain, heart, lung, breast, and various other organs. In addition to classical approaches such as affine, erasing, and elastic transformations, it has been emphasized that GANs, which are one of the generation-based methods, had a positive impact on performance in many studies. In addition to GANs, it was concluded that feature mixing, model-based, and reconstruction-based generation approaches are frequently used, and they had a positive effect on performance in three problems. Recently, it has been determined that classical transformation-based methods have been replaced by generation-based methods in the literature. It has been concluded that generation-based methods are more flexible and can be used effectively to eliminate class imbalance in training sets.

Wang et al. [34] proposed a new data augmentation approach called TensorMixup (TM) to improve glioma segmentation performance. They developed a mixing mechanism based on the Beta distribution to mix image patches and augment new synthetic data together with their segmentation masks. With the mixed image patches, new, synthetically generated images and their segmentation labels were obtained with the mixed image patches. Unlike the Mixup method, the proposed method mixes the image patches having tumors instead of mixing the whole slide images. In the same way, the masks were mixed, and then the segmentation was performed by training the 3D UNet model with the augmented data. With the proposed TM method, an increase in Dice scores of 1.53, 0.54, and 1.30 was achieved in the whole tumor, tumor core, and enhancing tumor segmentation results compared to baseline augmentation, respectively.

Alsaif et al. [35] proposed a tumor detection approach using data augmentation approaches together with CNN models such as ResNet, ALexNet, and VGG. In the data augmentation phase, classical approaches such as flipp-Newheir combinations are used. It is concluded that data augmentation approaches positively affect the performance of limited MRI datasets.

Qin et al. [36] proposed a data augmentation approach based on deep reinforcement learning and trial-and-error for the kidney segmentation problem. Their proposed approach consists of two unified modules: Dueling DQN for data augmentation and a deep learning network for segmentation. They developed a reward-penalty-based learning mechanism for data augmentation by looking at the performance obtained from the integrated model trained with the produced augmented data. With the proposed data augmentation method, an increase of 10.8 was achieved in Dice performance.

Han et al. [37] developed a brain tumor detection approach with the use of CNN models and GAN data augmentation. It has been emphasized that the distribution of the data obtained by classical data augmentation methods tends to be very similar to the original data and provides a limited increase in performance. With this

motivation, they used the PCGAN model in their studies to produce synthetic data that are completely different from the real images. In the experiments, both classical data augmentation and GAN data augmentation were used in the training phase of the ResNet-50 CNN model. The most successful results were obtained in the experiment with the use of both classical and GAN data augmentation combined in equal proportions. With the proposed method, 1.02% higher accuracy is obtained compared to the classical data augmentation.

Goceri [24] examined various data augmentation approaches used in medical imaging (brain, lung, breast, and eye) technically, comparatively, and practically. In addition to classical data augmentation methods such as rotation, noise addition, cropping, translation, and blurring GAN-based generative approaches are also examined. It has been concluded that GAN-based methods provide more successful results by increasing diversity, but the coordination between generator and discriminator structures is difficult to achieve, they are difficult to train and complex compared to classical data augmentation approaches. It has been emphasized that superior results can also be obtained by using combinations of classical data augmentation approaches other than GANs.

In the study of [38], a new generative model named TumorGAN has been proposed as a data augmentation approach for the limited data problem in HGG and LGG brain tumor segmentation. In addition, they developed new loss functions called regional perceptual loss and regional L1 loss to increase the performance in GAN training. In the results obtained, it was observed that the data augmented with TumorGAN increased the success in both single-modal and multi-modal data. The best results in the mean scores of the whole tumor, tumor core, and enhancing tumor segmentations were obtained by training the U-Net model with the data generated with the proposed TumorGAN approach.

In the [39], the authors propose a GAN-based approach to augment the dataset with synthetic liver lesion images. They train the GAN on a small dataset of real liver lesion images to generate a larger number of synthetic images. The authors report an improvement in the performance of the CNN model trained on the augmented dataset compared to the original dataset. However, their work focuses on liver lesion classification using CT images, while the proposed method in this paper focuses on brain tumor classification using MRI images. Moreover, the StyleGANv2-ADA is used in this work to generate the synthetic images, which allows for more fine-grained control over the augmentation process.

The proposed method in [40] is similar to the proposed work in this paper as it addresses the issue of insufficient training samples for deep learning-based lesion detectors by utilizing image generation techniques. However, this work focuses on delicate imaging textures, and the proposed method, TMP-GAN, employs joint training of multiple channels and an adversarial learning-based texture discrimination loss to generate high-quality images. The method also uses a progressive generation mechanism to improve the accuracy of the medical image synthesizer. Experiments on publicly available datasets demonstrate that the detector trained on the TMP-GAN augmented dataset performs better than other data augmentation methods, with an improvement in precision, recall, and F1 score.

In the study [41], the authors address the challenge of the limited availability of labeled medical imaging datasets for supervised machine learning algorithms. They propose using Generative Adversarial Networks (GANs) to generate synthetic samples that resemble real images and augment the training datasets. The authors demonstrate the feasibility of this approach in two brain segmentation tasks and show that introducing GAN-derived synthetic data improves the Dice Similarity Coefficient (DSC) by 1 to 5 percentage points under different conditions, particularly when there are fewer than ten training image stacks available. Their work demonstrates that GAN augmentation is a promising technique for improving the performance of machine learning algorithms in medical imaging tasks when the available labeled datasets are limited.

The study in [42] presents a method for generating synthetic brain MRI images with meningioma disease using a multi-scale gradient GAN (MSG-GAN). The generated images are used to augment the training set of a CNN model for a multi-class brain tumor classification problem. The evaluation of the proposed method on coronal-view images from the Figshare database shows an improvement in the classifier's performance in terms of the balanced accuracy score. The results demonstrate the potential of using GAN-based data augmentation to improve the performance of computer-aided diagnosis systems in brain tumor characterization.

The authors in [43] used two GAN models, namely DCGAN and WGAN, to see which GAN architecture is better for realistic MRI generation. The main goal of this comparison is to prevent the mode-collapsing and generating high-resolution synthetic images. The authors used the BRATS 2016 dataset. For each patient, they selected the MRI scans between number 80 and number 149 as all of the MRI scans do not carry valuable data. This cropping is also popular as it reduces the need for computational power. They processed the MRI scans by the dimensions of both 64x64 and 128x128. The study demonstrates that using WGAN

Yurtsever *et al. BMC Medical Informatics and Decision Making*   (2024) 24:285

Page 7 of 24

can generate realistic multi-sequence brain MR images, which may have practical applications in the clinical setting, such as data augmentation and physician training. The research highlights the potential of utilizing GAN-based methods for medical imaging data augmentation, specifically addressing the challenges associated with intrinsic intra-sequence variability in medical images. The comparison of the related papers is given in Table 1.

In Table 1, prominent studies on GAN and classical data augmentation in the medical field are examined and analyzed with prominent performance metrics. It can be seen that both data augmentation methods can be successfully applied in pancreatic, liver, and chest diseases other than tumors. In studies using both GAN and classical data augmentation, it can generally be concluded that GAN data augmentation provides higher performance.

In this paper, the utilization of StyleGANv2-ADA is proposed for augmenting brain MRI slices in the context of brain tumor classification. The datasets namely BraTS 2021, Gazi Brains 2020, and the BR35H are employed, and various deep-learning models are trained on them. Subsequently, a GAN-augmented model is trained on the same two datasets, and its performance is compared with that of the vanilla models. The aim is to prevent data leakage and ensure a fair comparison between the models by augmenting only the training data.

The proposed method of augmenting brain MRI slices with StyleGANv2-ADA is expected to improve the overall performance of the deep-learning models for brain tumor classification. The generated images can help to increase the diversity and quantity of the training data, thus reducing overfitting and improving the generalization ability of the models. Furthermore, the use of GANs for data augmentation has the potential to overcome the challenge of working with medical data, particularly the sparsity of ground truth data.

The key contributions of this paper are listed as follows:

- A brain tumor classification pipeline is designed, incorporating feature fusion, traditional and synthetic data augmentation, and CNN classification steps to enhance classification performance.
- Whether data augmentation with GANs is as successful and valid as well-known traditional data augmentation methods, or is a better option, is demonstrated by conducting tests with a state-of-the-art GAN model in the brain MRI domain.
- The StyleGANv2-ADA is adopted to synthesize realistic, imitated brain MRIs, augmenting the datasets, and assessing their effects on performance.
- The classification performances of different models trained on images are compared without any aug-

mentation, with traditional augmentation, and with GAN-synthesized augmentation.
- Experiments are conducted using the BraTS 2021 dataset alongside our own Gazi Brains 2020 dataset to assess whether synthetic images generated by GANs are the viable option.
- The impact of data augmentation methods on classification performance is presented by testing the proposed approach on the three datasets. The Gazi Brains 2020 Dataset comprises glioma-type brain tumors meticulously curated by six medical experts from Gazi University's Faculty of Medicine.

Material and methods

In this study, the impact of utilizing StyleGANv2-ADA for brain tumor classification via deep learning is investigated. The datasets namely the Gazi Brains 2020, BraTS 2021, and BR35H datasets are employed for augmentation and evaluation purposes. During the data pre-processing phase, excess pixels were removed, and input images were resized to fulfill model requirements. Various CNN models with DenseNet were employed for classification comparison. Additionally, the implementation of k-fold cross-validation ensured a more dependable model evaluation. The StyleGANv2-ADA GAN model was utilized to augment training data and evaluate its influence on overall model performance. Figure 1 depicts the methodology employed to assess the significance of generated synthetic data in the classification task. The provided data is first passed through the fusion stage. Later, except for classification without augmentation, the classification processes were carried out after traditional and StyleGANv2 augmentation approaches.

Brats (brain tumor segmentation) 2021 dataset

The BraTS dataset has been widely used in numerous studies for brain tumor segmentation, classification, and survival prediction using MRI images [48–51]. The dataset consists of multimodal MRI scans, including T1-weighted MRI, T1-weighted MRI with contrast enhancement, T2-weighted MRI, and Fluid Attenuated Inversion Recovery (FLAIR) for each patient. The dataset is used for various tasks, including brain tumor segmentation, classification, and survival prediction [52, 53].

One of the most commonly used versions of the dataset is BraTS 2018, which includes MRI scans from 285 training subjects and 66 validation subjects. The dataset contains both high-grade gliomas (HGGs) and low-grade gliomas (LGGs). Researchers often use the BraTS dataset to evaluate and compare the performance of different algorithms and models for brain tumor segmentation. Many advanced algorithms have been validated using the BraTS dataset, making it a benchmark for evaluating

Table 1 Comparison of methodologies used in relevant research papers on brain tumor radiogenomic classification

Study	Data	Problem	Classifier	GAN method	No augmentation	Classic/affine augmentation	Gan augmentation
[37]	Brain MRI	Tumor Detection	ResNet-50	PGGAN	Accuracy: 90.06 Sensitivity: 85.27 Specificity: 97.04	Accuracy: 90.70 Sensitivity: 88.70 Specificity: 93.62	Accuracy: 91.08 Sensitivity: 86.60 Specificity: 97.60
[40]	Pancreas CT	Lesion Detection	Faster RCNN	TMP-GAN	-	Precision: 82.64 Recall: 81.26 F1 Score: 81.94	Precision: 86.28 Recall: 85.89 F1 Score: 86.08
[41]	Brain CT	Cerebrospinal Fluid (CSF) segmentation	UNet	PCGAN	Accuracy: 88.1	-	Accuracy: 89.3
[42]	Brain MRI	Multi-class brain tumor classification	CNN	MSG-GAN	Accuracy: 90.3	-	Accuracy: 93.1
[43]	Brain MRI	Medical Image Generation	No Classifier	DCGAN, WGAN	-	-	Accuracy for DCGAN: 77 Accuracy for WGAN: 66
[44]	Brain MRI	Tumor Classification	CNN	DCGAN	Accuracy: 91.70 Sensitivity: 90.16 Specificity: 95.58 Precision: 91.17 F1-score: 90.54	-	Accuracy: 93.01 Sensitivity: 92.01 Specificity: 96.39 Precision: 92.28 F1-score: 92.10
[45]	Brain MRI	MRI Generation and Classification	CNN MobileNetV2 ResNet152V2	DCGAN	-	Accuracy: 95.84-99.09 Precision: 96.48-99.12 Recall: 95.68-99.08	Accuracy: 93.27-97.94 Precision: 93.24-96.91 Recall: 92.19-96.91
[46]	Brain MRI	Tumor Classification	Multi-stream 2D CNN	PGGAN	Accuracy: 82.45 Sensitivity: 70.51 Specificity: 87.4	-	Accuracy: 88.82 Sensitivity: 81.81 Specificity: 92.17
[47]	Liver CT	Liver Lesion Classification	CNN	DCGAN, ACGAN	Accuracy: 57	Sensitivity: 78.6 Specificity: 88.4	Sensitivity: 85.7 Specificity: 92.4

Yurtsever *et al. BMC Medical Informatics and Decision Making*   (2024) 24:285

Page 8 of 24

Yurtsever *et al. BMC Medical Informatics and Decision Making* (2024) 24:285
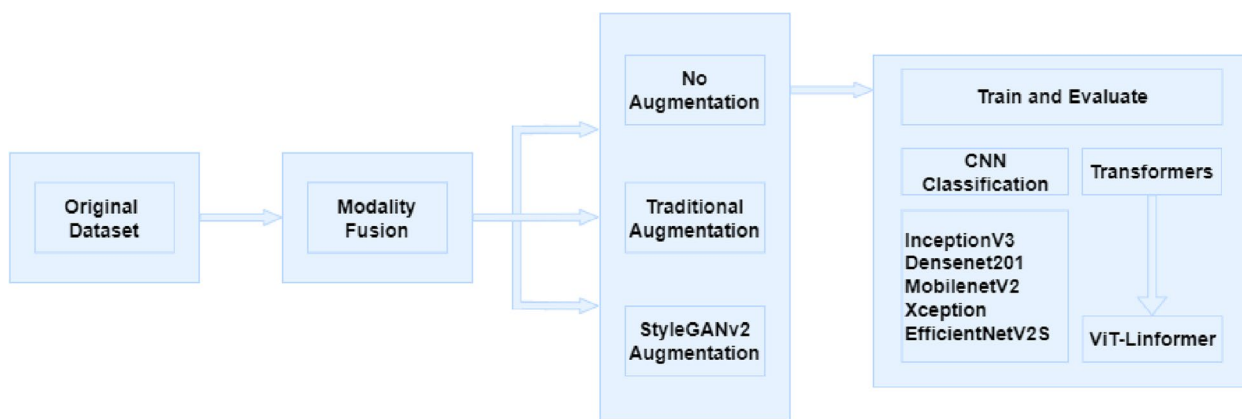
Page 9 of 24

Fig. 1 The process of StyleGANv2-ADA augmentation methodology on classification

the effectiveness of new methods. The dataset allows researchers to compare their proposed methods with existing methods and assess their performance.

The BraTS dataset is particularly valuable because it includes multimodal MRI scans, which provide different types of information about the tumors. These modalities include T1-weighted, T1-weighted with contrast enhancement, T2-weighted, and FLAIR (Fluid-Attenuated Inversion Recovery) images. The availability of multiple modalities allows researchers to develop algorithms that can leverage the complementary information provided by each modality to improve tumor segmentation accuracy. Overall, the BraTS 2021 dataset is valuable for researchers working on brain tumor segmentation and related tasks. Its large size, multi-modal nature, and availability of ground truth segmentation make it a suitable dataset for developing and evaluating algorithms for accurate and automated brain tumor analysis.

The gazi brains 2020 dataset

The Gazi Brains 2020 Dataset [54] is a collection of glioma-type brain tumors that was meticulously compiled by six medical experts from the Gazi University Faculty of Medicine. This dataset comprises brain MR images from a total of 100 patients, with 50 patients being healthy and 50 patients diagnosed with High-Grade Glioma (HGG). For each patient, the dataset includes T1-weighted, T1CE (Gadolinium), T2-weighted, and FLAIR MR images. Additionally, the dataset incorporates tumor region segmentation and 12 anatomical structure tags meticulously prepared by the experts. The segmentation labels of the Gazi Brains 2020 Dataset are given in Table 2. It further contains specific tumor findings and components of HGG patients, along with demographic characteristics like age and gender.

Table 2 Segmentation Labels of the Gazi Brains 2020 Dataset

Label	Name	Label	Name
0	Clear Label	8	Peritumor Edema Area
1	Region of Interest	9	Contrast Enhancing Part of Tumour
2	Eye	10	Necrosis of Tumour
3	Optic Nerve	11	Cavum Septum Pellicidum
4	Lateral Ventricle	12	Haemorrage
5	Third Ventricle	13	Lipoma
6	Ischemic Gliotic Changes	14	Arachnoidal Cyst
7	Contrast Non-Enhancing-Part of Tumour	15	Late Sequela Infarct Area

In the preprocessing phase of the Gazi Brains 2020 dataset; the separation of slices with and without HGG findings, density normalization, feature selection, and feature fusion steps are applied. The images obtained as a result of all preprocessing, fusion, and relabeling steps are shared in [55] as an organized dataset. In the dataset consisting of HGG and normal samples, there are segmentation masks for each patient's Whole Slide Image (WSI). There are a total of 16 labels in these masks as shown in Table 2. By taking the opinion of doctors and experts from the Gazi Medical Faculty, WSIs with 7,8,9,10, and 12 labels among all in the segmentation map are labeled as slices with HGG, otherwise, they are labeled as normal for the binary glioma classification task.

Among the modalities in the dataset, T1CE, T2, and FLAIR images were selected for fusing. The reason for using the T1CE modality instead of T1 is that it has generally known advantages over T1 in the medical field. Samples without T1CE images were eliminated. Better results can be obtained by enriching the data rather than trying to solve the problem with a single modality, as shown in [56]. Selected T1CE, T2, and FLAIR modalities

are then combined, and fused in 3 channels (RGB) to obtain a single result image. In this way, explanatory and more comprehensive data is obtained that combines the richer and different advantages of different modalities that the deep learning models can use and learn for each sample at the learning stage.

The synthetic MRI slices are generated using the StyleGANv2-ADA model. In Fig. 2, a visual comparison is presented between the original MRI slices, which have been sampled from the Gazi Brains 2020 dataset, and the MRI slices generated using the StyleGANv2-ADA architecture. Each grid features 9 MRI slices, resized to 128x128 pixels. This comparison serves to demonstrate the quality and realism of the GAN-generated MRI slices within the context of brain tumor classification. The advantages of utilizing the StyleGANv2-ADA model for synthetic MRI generation lie in its ability to capture intricate patterns and details, producing images that closely resemble real MRI scans. The model's adaptive training methodology enhances its versatility across different datasets, ensuring robust performance and generalization.

The br35h dataset

Brain tumor detection datasets have been developed to address a critical medical challenge, as these aggressive diseases significantly impact patient survival rates. The Br35h dataset [57], comprising 3060 brain MRI images, has been structured to support the development of automated systems for tumor identification. It has been divided into two main categories: 1500 tumorous and 1500 non-tumorous brain MRI scans, with an additional folder intended for prediction or testing purposes. This balanced composition has been designed to facilitate binary classification tasks, which are considered fundamental in distinguishing between healthy and tumor-affected brain tissues.

The dataset has been curated to enable various applications in the field of medical imaging analysis. CNNs and other deep learning models can be trained and tested using these images, while transfer learning techniques may be explored to enhance model performance. Although not explicitly stated in the dataset description, the possibility of tumor position segmentation has been suggested. The ultimate goal of such datasets is understood to be the development of systems that can be utilized to assist radiologists and neurosurgeons in rapidly and accurately identifying brain tumors, a capability that is particularly valued in regions where access to skilled professionals may be limited. It is expected that this resource will be leveraged by researchers to compare the efficacy of various machine learning and deep learning algorithms in the context of brain tumor detection.

Preprocessing and overall pipeline

In this section, the data preprocessing steps carried out for classifying brain tumors using medical imaging, specifically brain MRI slices are described. Steps for the data preprocessing and the general architecture of this work
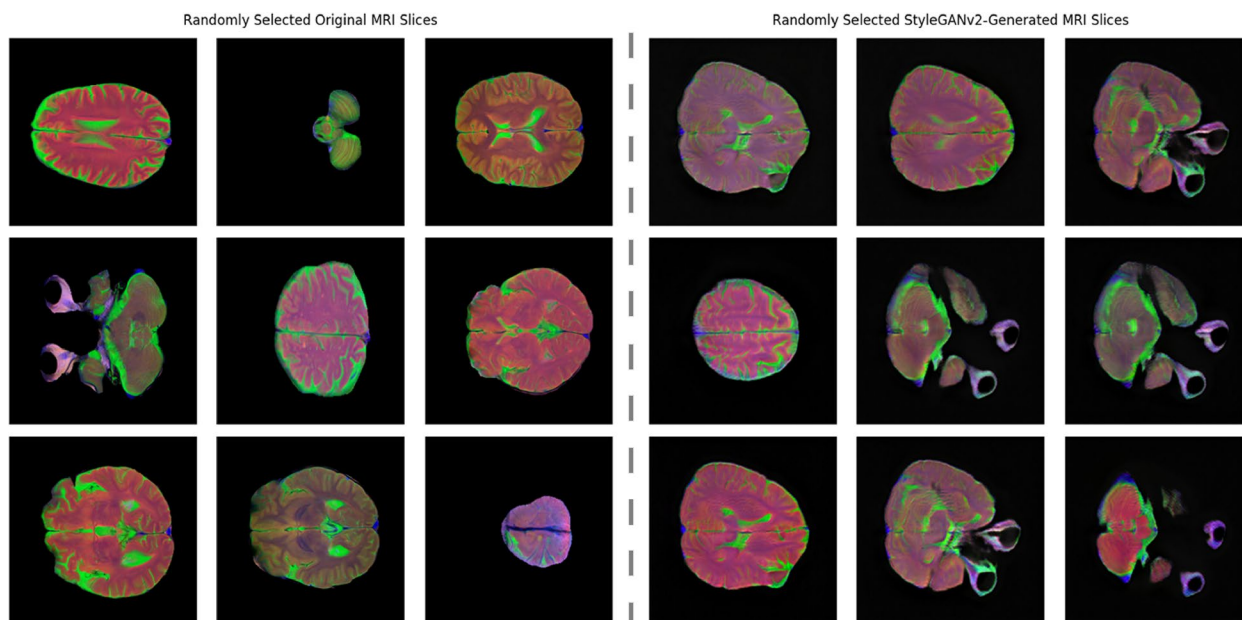

Fig. 2 Visual Comparison of Original MRI Slices and MRI Slices Generated by the StyleGANv2-ADA Model. The figure showcases a side-by-side comparison of the original MRI slices (left) and MRI slices generated using the StyleGANv2-ADA model (right)

are given in Fig. 3. The preprocessing steps aim to ensure data consistency, remove unnecessary slices, and mitigate any potential bias during the training process. Additionally, GAN-based data augmentation techniques are employed to enhance the robustness and generalizability of the classification model. As described in Fig. 3, this work mainly has three different ways to train classifier models. We first train the models with no augmentation. After that, we train the models with traditional augmentation and finally, GAN-based augmentation to observe whether the data augmentation can be considered for improving the classifier model's accuracy.

Removal of unnecessary MRI slices

Some MRI scans may contain irrelevant or redundant slices that do not contribute to the classification task. To ensure the focus is on relevant tumor-related information, we performed a thorough review of each dataset and manually removed unnecessary slices. This step aimed to reduce noise and enhance the signal-to-noise ratio of the data.

K-fold cross-validation

It is beneficial for the proposed brain tumor classification task as it helps in assessing the model's performance robustly. By iteratively training and evaluating the model on different subsets of the data, a more comprehensive understanding of the model's generalization capabilities is gained. Moreover, this technique aids in reducing the impact of data variability and ensures that the model's performance is not overly dependent on a specific subset of the data. Through k-fold cross-validation, the reliability of the proposed classification model is enhanced and makes it more adaptable to diverse patterns within the brain tumor dataset.

GAN-based data augmentation

In addition to the aforementioned preprocessing steps, the GAN-based data augmentation technique is utilized to increase the diversity and quantity of the training data. GANs are capable of generating realistic synthetic samples that can be used to augment the original dataset. By leveraging GANs, it is aimed to address the challenges posed by limited data availability and improve the model's ability to generalize to unseen brain tumor cases.

The GAN-based data augmentation involved training a GAN model on a subset of the original dataset. This GAN model learned the underlying distribution of the data and generated synthetic brain MRI slices. These synthetic samples were then combined with the original dataset to create an augmented training set, resulting in a larger and more diverse training data pool. This process effectively increased the variability in the training data, enabling the classification model to learn more robust
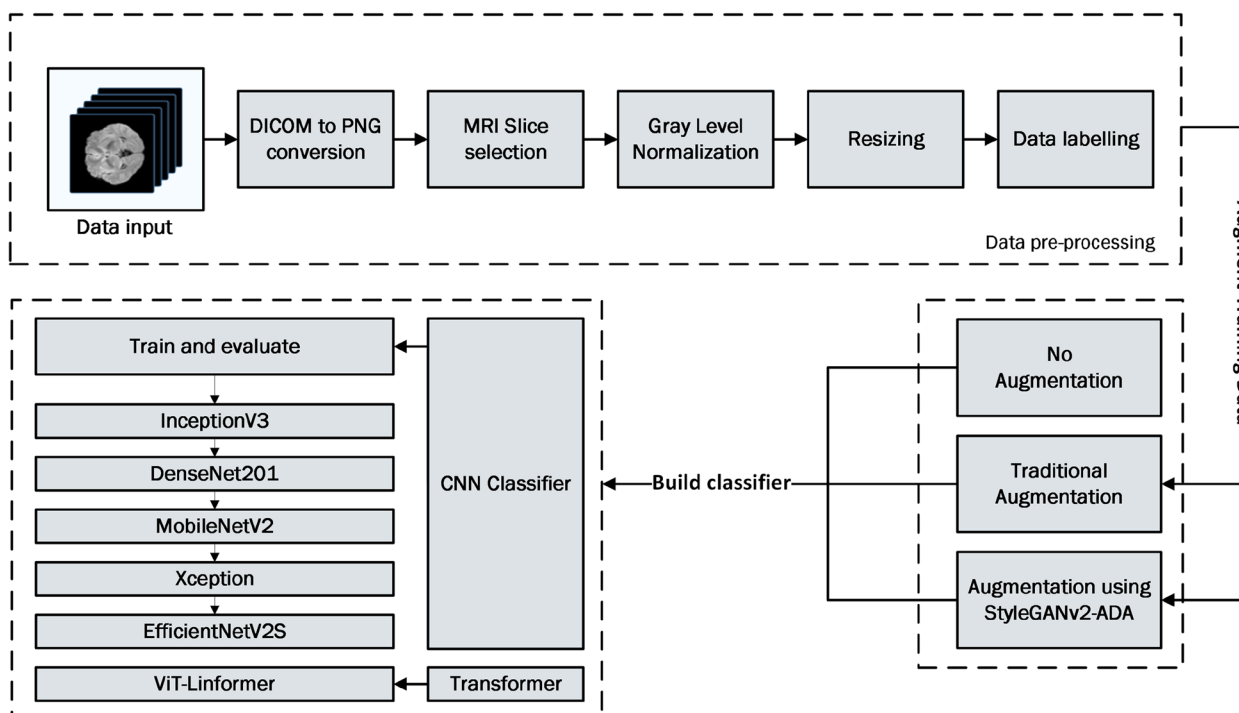

Fig. 3 Data preprocessing steps and overall architecture of the proposed method for all datasets namely the BraTS 2021, Br35h, and Gazi Brains 2020

and discriminative features for accurate brain tumor classification.

The pseudo-code, describing the main steps of the proposed method is given in Algorithm 1. In summary, the data preprocessing steps for the proposed brain tumor classification task included resizing all MRI slices to a standardized dimension of 256x256 pixels while preserving the three-channel information. A careful review and removal of unnecessary MRI slices is also conducted to enhance data quality. K-fold cross-validation is employed to prevent data bias during model training and evaluation. Furthermore, the GAN-based data augmentation technique is utilized to augment the original dataset and enhance the model's ability to generalize to unseen brain tumor cases. These preprocessing steps are crucial in ensuring reliable and unbiased classification results.

This allows the discriminator to better generalize to new examples and improves the diversity and quality of the generated images.

Stochastic discriminator augmentation: Augmentations used during training can cause unwanted artifacts to appear in generated images. Existing approaches attempt to prevent augmentation leakage by making the discriminator blind to augmentations, which is not ideal. In the proposed approach, augmentations are applied to the images shown to the discriminator, but only the augmented images are presented to it. The same augmentations are also used when training the generator, allowing for the effective incorporation of augmentations without compromising the discriminator's ability to recognize them.

Figure 4 illustrates the architecture with G representing the generator and D representing the discriminator. The augmentation probability, denoted as p and ranging between 0 and 1, controls the intensity of the augmenta-

Algorithm 1 Pseudo-code of the proposed architecture

---

1. **Read the MRI slices from the Gazi Brains 2020 Dataset**
2. **Perform image preprocessing:**
   Remove the unnecessary MRI slices.
   Resize the MRI slices to 256x256 pixels
3. **Perform Traditional Data Augmentation:**
   Increase the diversity and quantity of the training data by performing simple alterations to the MRI slices such as rotating.
4. **Perform GAN-based Data Augmentation: Enable the classification model to learn more robust and discriminative features for accurate brain tumor classification**
   To compare the impact of GAN-generated images on the classifier accuracy, create synthetic data using StyleGANv2-ADA with (A = 50%) percent of training data size.
5. **Train CNN and Transformer Models: Train five different CNN and a transformer (ViT-Linformer) models with three combinations: No augmentation, Traditional augmentation, and GAN-based augmentation.**
   Perform k-folding: mitigate any potential bias during the training and evaluation of the classifier.

Styleganv2-ADA

StyleGAN2-ADA [58] (Adaptive Discriminator Augmentation) is an improved version of the StyleGAN2 model that was introduced in 2020. The original StyleGAN2 model is a state-of-the-art generative model for synthesizing high-quality images. It uses a novel architecture that combines a generator network with a discriminator network to produce realistic images. The generator network generates images, while the discriminator network evaluates how realistic the generated images are compared to real images.

StyleGAN2-ADA introduces an additional technique called Adaptive Discriminator Augmentation (ADA) to improve the performance of the discriminator network. ADA is a regularization method that adjusts the training of the discriminator network based on the difficulty of the training examples. It works by dynamically adjusting the strength of data augmentation during training.

tions. To ensure a diverse training process, discriminator D is rarely exposed to clean images as the pipeline incorporates multiple augmentations, typically with a p-value around 0.8. During training, the generated images are augmented before being evaluated by discriminator D. By placing the Aug operation after the generation process, generator G is effectively guided to produce exclusively clean images.

Designing augmentations that do not leak: GANs can undo corruptions when trained solely on corrupted images if the augmentations used for corruption allow distinguishing between augmented image sets. For example, stochastic non-leaking augmentations, like random rotations by 0, 90, 180, or 270 degrees with 10% probability, increase the occurrence of images at 0 degrees, forcing the generator to produce correctly oriented images. Deterministic augmentations can be made non-leaking by applying them with a probability
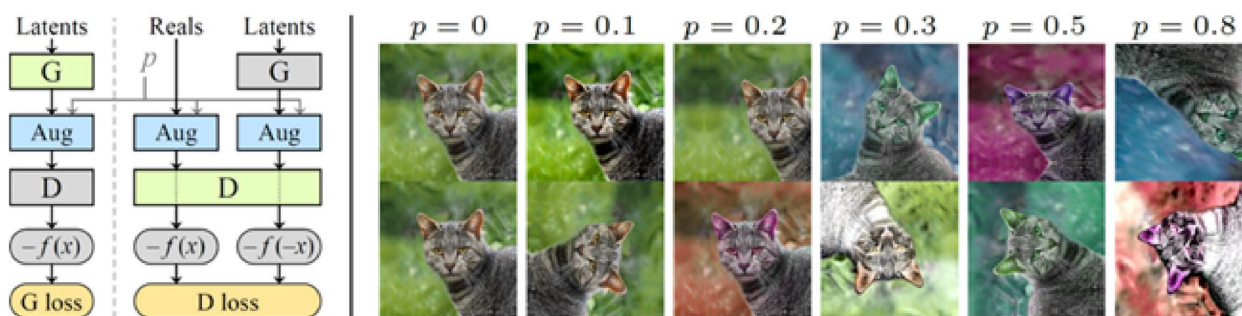
Fig. 4 Stochastic discriminator augmentation and the effect of augmentation probability *p* for the StyleGANV2-ADA [58]

(p%) below 0.8. This ensures that the augmentations do not disrupt the generator's ability to accurately replicate the original image distribution. The authors employ a total of 18 augmentations, applied in a predefined order, with equal independent probabilities. This extensive augmentation pipeline ensures that the discriminator rarely encounters an image without any augmentations. Despite this, the generator is still trained to generate clean images as long as the probability value (p) remains below a designated safe threshold.

Adaptive discriminator augmentation (ADA): To eliminate the need for manual adjustment of augmentation strengths, the authors introduce two heuristics to detect discriminator overfitting. The first heuristic compares the discriminator predictions for the validation set to those of the training set and generated images. The second heuristic measures the proportion of the training set that yields positive discriminator outputs. In practice, the value of p is initialized at 0, and after every few minibatches, the heuristics are computed. Based on the detected overfitting, p is dynamically adjusted aggressively to counteract the issue.

In the study, StyleGAN2-ADA is employed to augment brain MRI slices for training deep learning models. The application of this method is intended to enhance the overall performance of the models by offering them more diverse and realistic training examples. Augmentation is restricted solely to the training data to prevent leakage and ensure the evaluation of the model on the real test data. It is important to note that StyleGANv2-ADA is used exclusively for augmenting the training data, not the testing data. This approach ensures that the model's performance is evaluated on real, unaltered data, allowing us to accurately assess the impact of the synthetic data generated during training. By generating high-fidelity synthetic images that mirror the real data, StyleGANv2-ADA plays a crucial role in improving the robustness and accuracy of the model. The effectiveness of StyleGANv2-ADA in this context

has been thoroughly evaluated and discussed, highlighting its importance in the proposed methodology.

Classification

Based on the models that are evaluated in this study, namely MobileNetV2, Xception, and EfficientNetV2S, an inference can be made about the relationship between the parameter numbers, network depth, and the data under consideration.

MobileNetV2 [59] is designed to be lightweight and efficient, especially suitable for resource-constrained environments. Despite its lower parameter count compared to some larger models, it leverages depth-wise separable convolutions and other techniques to reduce computational requirements while maintaining reasonable accuracy.

Xception [60] is an extended version of Inception with depthwise separable convolutions, which allows it to capture fine-grained features efficiently. With a more complex architecture and increased depth, Xception is likely to have a higher number of parameters compared to Vanilla and MobileNetV2.

EfficientNetV2S [61] models are designed to achieve state-of-the-art performance while maintaining a good trade-off between accuracy and computational efficiency. The specific variant, EfficientNetV2S, signifies a smaller version of the EfficientNetV2 family. It is expected to have a moderate number of parameters, offering a balance between model complexity and performance.

ViT-Linformer [62] is a transformer-based architecture designed for vision tasks, combining the benefits of Vision Transformers (ViT) with the Linformer model's efficient attention mechanism. Unlike traditional convolutional networks, ViT-Linformer relies on global self-attention to capture dependencies across the entire image, which allows it to learn complex patterns. However, it tends to require more data for effective training and generally has a higher computational cost compared

to CNN models, despite optimizations for reduced attention complexity.

Considering the above, it can be inferred that as the model's depth and complexity increase (e.g., from MobileNetV2 to Xception and EfficientNetV2S), the number of parameters is likely to rise accordingly. However, models like MobileNetV2 and EfficientNetV2S are designed to provide a good balance between accuracy and parameter efficiency, thus offering a potentially optimal trade-off for this study.

It is important to note that the relationship between parameter numbers, network depth, and data is a complex topic, and the actual impact on model performance may vary depending on the specific dataset and task at hand. The evaluation of these models in this study provided a more concrete understanding of their performance and shed light on the relationship between parameter numbers, network depth, and the classification of brain tumor data.

Experimental studies

In this study, six deep-learning models for brain tumor classification are evaluated using the three datasets namely Gazi Brains 2020, BraTS 2021 and the BR35H. The models that are evaluated included MobileNetV2, Xception, EfficientNetV2S, and the proposed GAN-augmented model, MobileNetV2, using StyleGANv2. The 5-fold cross-validation approach is used to train and test each model, including the proposed GAN-augmented model using StyleGANv2, and their performance is evaluated using metrics such as accuracy, precision, recall, and F1-score. The hardware used for the experiments, including training and testing in this proposed GAN-augmented model using StyleGANv2, consisted of an Intel® Xeon® Gold 6336Y Processor, 512GB of RAM, and NVIDIA 4xA40 GPU with 48GB of VRAM. The TensorFlow deep learning framework is used to implement and train the models, including the proposed GAN-augmented model using StyleGANv2, and the Adam optimizer with a learning rate of 0.001 is employed. Each model's learning rate (LR) and network parameters are listed in Table 3.

Table 3 presents a comparison of various models including InceptionV3, DenseNet201, MobileNetV2, Xception, and EfficientNetV2S. These models are optimized using RMSprop with a learning rate of 1e-4. The Table 3 also displays the total number of parameters for each model, indicating their complexity. The experiment involved the evaluation of two datasets, namely the Gazi Brains 2020 Dataset and the BRaTS 2021 dataset, to assess the performance of the models with varying parameter sizes. The models were subjected to testing to ascertain whether differences in parameter counts led to

Table 3 Summary of the different CNN models that are trained on all datasets

Models	Optimizer	LR	Total number of parameters
InceptionV3	RMSprop	1×10^{-4}	31, 240, 225
DenseNet201	RMSprop	1×10^{-4}	30, 364, 481
MobileNetV2	RMSprop	1×10^{-4}	12, 744, 001
Xception	RMSprop	1×10^{-4}	37, 638, 953
EfficientNetV2S	RMSprop	1×10^{-4}	30, 817, 377

Table 4 Descriptive statistics for model accuracy are provided for three different MobileNetV2 models for the Gazi Brains 2020 dataset

Models	Std. Dev. (Acc)	Minimum (Acc)	Maximum (Acc)
The MobileNetV2	0.0082	0.9247	0.9502
MobileNetV2 with Traditional Augmentation	0.014	0.9269	0.9774
MobileNetV2 with Augmentation (StyleGANV2-ADA)	0.0168	0.9444	1

variations in performance metrics. Through this experimentation, the datasets were utilized to gauge the models' efficacy across diverse tasks and data distributions. Comparisons were made to analyze how the models, characterized by differing parameter sizes, performed in relation to the specific datasets. Results were obtained to determine any discernible patterns or correlations between parameter size and model performance across the evaluated datasets.

The presented Table 4 outlines the performance metrics of various models employed for brain tumor classification. Incorporating augmentation through StyleGANv2-ADA further raises the standard deviation to 0.0168. Notably, the minimum accuracy significantly improves to 0.9444, while the maximum accuracy reaches 1, suggesting instances of perfect classification. These results highlight the positive impact of StyleGANv2-ADA on the model's ability to generalize and classify unseen data.

Gazi Brains 2020, BraTS 2021, and BR35H datasets are initially trained on six different deep-learning models without augmentation. The evaluation metrics for the non-augmented classifier can be found in Table 5. Subsequently, traditional augmentation methods are applied to the training dataset, and the results are presented in Table 6. Finally, the GAN-augmented models are trained using six different deep-learning models, including the transformer model and a comprehensive comparison of

Yurtsever *et al. BMC Medical Informatics and Decision Making* (2024) 24:285

Page 15 of 24

Table 5 The results of models **without data augmentation** using the Gazi Brains 2020

Model	Acc.	Prec.	Recall	AUC
InceptionV3	0.9492	0.9380	0.9892	0.9559
DenseNet201	0.9484	0.9535	0.9699	0.9592
MobileNetV2	0.9349	0.9275	0.9783	0.9149
Xception	0.9524	0.9430	0.9880	0.9534
EfficientNetV2S	0.9460	0.9334	0.9892	0.9432
ViT-Linformer	0.7408	0.8003	0.8162	0.7047

Table 6 Evaluation of models for classification **with traditional data augmentation** using the Gazi Brains 2020

Model	Acc.	Prec.	Recall	AUC
InceptionV3	0.9468	0.9506	0.9711	0.9734
DenseNet201	0.9603	0.9652	0.9759	0.9799
MobileNetV2	0.9619	0.9613	0.9819	0.9551
Xception	0.9563	0.9435	0.9940	0.9811
EfficientNetV2S	0.9611	0.9520	0.9916	0.9786
ViT-Linformer	0.6573	0.6810	0.9205	0.5291

Table 7 Evaluation of all the deep learning models for brain tumor classification using the BraTS 2021 Dataset. It can be observed that the data augmentation using the StyleGANv2-ADA GAN model achieves the best overall accuracy. The accuracy given here is the average of 5-fold training. Data Augmentation A = 50% is the parameter that shows how much training data in each fold is augmented and appended only to the train data

Models	Metrics	No Aug.	Traditional Aug.	StyleGANv2-ADA A = 50%
InceptionV3	Acc.	0.7540	0.6570	0.7433
DenseNet201	Acc.	0.7306	0.6624	0.7327
MobileNetV2	Acc.	0.6240	0.5953	0.6454
Xception	Acc.	0.7316	0.7018	0.7295
EfficientNetV2S	Acc.	0.7455	0.6805	0.7518
ViT-Linformer	Acc.	0.5834	0.5312	0.6017

these methods is provided in Table 7. The primary objective is to assess whether GAN-augmented data contributes to the overall accuracy improvement of the brain tumor classifier.

The performance of the deep learning models in both Table 5 and 6 has higher recall and AUC scores than other metrics like accuracy, highlighting their effectiveness in correctly identifying positive instances, particularly crucial in tasks like brain tumor classification. The high recall values signify the models' ability to capture

most of the positive cases, essential in medical diagnosis where missing a positive instance can have severe consequences. Furthermore, the elevated AUC scores demonstrate the models' proficiency in distinguishing between classes, ensuring that true positives are ranked higher than false positives across different threshold values. These results underscore the suitability of the employed deep learning models for accurately detecting brain tumors, offering promising prospects for improved diagnostic procedures and patient care.

Tables 6 and 7 provide comprehensive evaluations of six deep-learning models for brain tumor classification using the Gazi Brains 2020 dataset under different conditions. Table 5 showcases the performance metrics of the models with no data augmentation. Notably, InceptionV3 and Xception exhibit high accuracies of 0.9492 and 0.9524, respectively, while DenseNet201 and EfficientNetV2S also demonstrate commendable results. In Table 6, the models are evaluated with traditional data augmentation. The overall accuracy improves for all models, with DenseNet201, MobileNetV2, and EfficientNetV2S reaching notably higher accuracies of 0.9603, 0.9619, and 0.9611, respectively. The use of traditional data augmentation contributes to enhanced precision, recall, and area under the curve (AUC) metrics across the models, underscoring the effectiveness of augmentation techniques in improving the deep learning models' performance for brain tumor classification in the Gazi Brains 2020 dataset.

Evaluation and analysis

In the results section of this study, the evaluation of six deep-learning models for brain tumor classification using three different datasets is presented. The performance of these models is compared with and without data augmentation using StyleGANv2-ADA. It is observed that the data augmentation technique using the StyleGANv2-ADA GAN model achieves the best overall accuracy, with an average accuracy of 94.5% across all five models on the Gazi Brains 2020 dataset.

In addition to the evaluation of the performance of the deep learning models, the impact of GAN-based data augmentation on the models' ability to learn discriminative features and improve classification performance is also analyzed. The performance of six deep learning models is presented in Table 5. It is found that the introduction of synthetic samples through GANs enhances the models' ability to learn relevant features and improves their classification performance. However, it is also acknowledged that there are potential challenges and caveats associated with this approach, such as the risk of overfitting and the need for careful selection of GAN hyperparameters. The comparison of accuracies

Table 8 Evaluation of all the deep learning models for brain tumor classification using the Gazi Brains 2020 Dataset. It can be observed that the data augmentation using the StyleGANv2-ADA GAN model achieves the best overall accuracy. The accuracy given here is the average of 5-fold training. Data Augmentation A = 50% is the parameter that shows how much training data in each fold is augmented and appended only to the train data

Models	Metrics	No Aug.	Traditional Aug.	StyleGANv2-ADA A = 50%
InceptionV3	Acc.	0.9492	0.9468	0.9936
DenseNet201	Acc.	0.9484	0.9603	0.9928
MobileNetV2	Acc.	0.9349	0.9619	0.9928
Xception	Acc.	0.9524	0.9563	0.9928
EfficientNetV2S	Acc.	0.946	0.9611	0.9936
ViT-Linformer	Acc.	0.7408	0.6573	0.6768

for three different EfficientNetV2S models on the BraTS 2021 is shown in Fig. 5a.

Tables 7 and 8 present evaluations of six deep-learning models for brain tumor classification using the BraTS 2021 and Gazi Brains 2020 datasets, respectively. In Table 7, employing the StyleGANv2-ADA GAN model for data augmentation results in the highest overall accuracy across all models. The accuracy averages of 5-fold training are provided for different augmentation scenarios, with the parameter A = 50% indicating the extent of training data augmentation in each fold. It can be observed that the data augmentation using the StyleGANv2-ADA GAN model achieves the best overall accuracy. The accuracy given here is the average of 5-fold training. Data Augmentation (A = 50%) is the parameter that shows how much training data in each fold is augmented and appended to the train data. Notably, StyleGANv2-ADA consistently outperforms both no-augmentation and traditional augmentation methods for all models. The comparison of AUC, recall and precision scores for three different EfficientNetV2S on the BraTS 2021 are presented in Figs. 6a, 7a and 8a, respectively.

Table 8 further corroborates the effectiveness of Style-GANv2-ADA in achieving the best overall accuracy in brain tumor classification using the Gazi Brains 2020 dataset. The accuracy averages, presented for different augmentation scenarios (No Aug., Traditional Aug., and StyleGANv2-ADA with A = 50%), highlight the superior performance of the GAN-based augmentation method across all models. It can be observed that the data augmentation using the StyleGANv2-ADA GAN model achieves the best overall accuracy. The accuracy given here is the average of 5-fold training. Data Augmentation (A = 50%) is the parameter that shows how much training data in each fold is augmented and

appended to the train data. These findings underscore the significant impact of StyleGANv2-ADA on enhancing the classification accuracy of deep learning models in the context of brain tumor datasets. The comparison of accuracies for three different MobileNetV2 models on the Gazi Brains 2020 is given in Fig. 5b. Also, the comparison of AUC, recall and precision scores for three different EfficientNetV2S on the BraTS 2021 are presented in Figs. 6b, 7b and 8b, respectively.

Table 9 presents the performance of six deep learning models (InceptionV3, DenseNet201, MobileNetV2, Xception, EfficientNetV2S, and ViT-Linformer) on the BR35H dataset for brain tumor classification, comparing three scenarios: no augmentation (No Aug.), traditional data augmentation (Traditional Aug.), and augmentation using the StyleGANv2-ADA model with an augmentation parameter (A) of 50%. It can be observed that the data augmentation using the Style-GANv2-ADA GAN model achieves the highest accuracy for most models, with DenseNet201 reaching an accuracy of 0.9944. The results show that DenseNet201 benefits most from the StyleGANv2-ADA augmentation, outperforming both the no-augmentation and traditional augmentation scenarios. MobileNetV2 and ViT-Linformer also see significant improvements with StyleGANv2-ADA, achieving accuracies of 0.9755 and 0.8403, respectively.

The reported accuracies are averaged across 5-fold cross-validation for all experiments conducted on the three datasets. The parameter A=50% indicates that 50% of the training data in each fold is augmented and appended to the original training set. This augmentation strategy consistently enhances the performance of the models, particularly for DenseNet201 and MobileNetV2.

Table 9 Evaluation of all the deep learning models for brain tumor classification using the BR35H Dataset. It can be observed that the data augmentation using the StyleGANv2-ADA GAN model achieves the highest accuracy for most models. The accuracy given here is the average of 5-fold training. Data Augmentation A = 50% is the parameter that shows how much training data in each fold is augmented and appended only to the train data

| Models	Metrics	No Aug.	Traditional Aug.	StyleGANv2-ADA A = 50%
InceptionV3	Acc.	0.9911	0.9944	0.9922
DenseNet201	Acc.	0.9911	0.9888	0.9944
MobileNetV2	Acc.	0.9333	0.9611	0.9755
Xception	Acc.	0.9766	0.9877	0.9877
EfficientNetV2S	Acc.	0.9911	0.9922	0.9899
ViT-Linformer	Acc.	0.8180	0.5765	0.8403
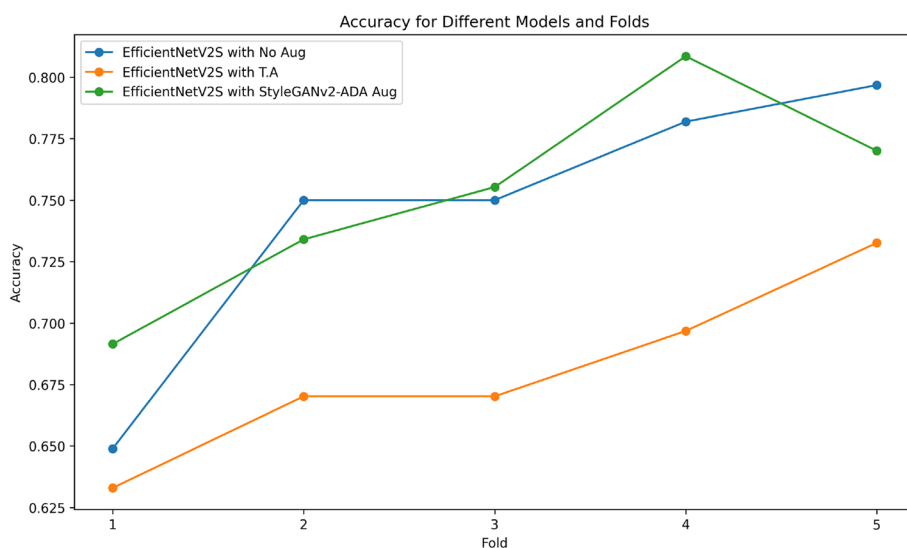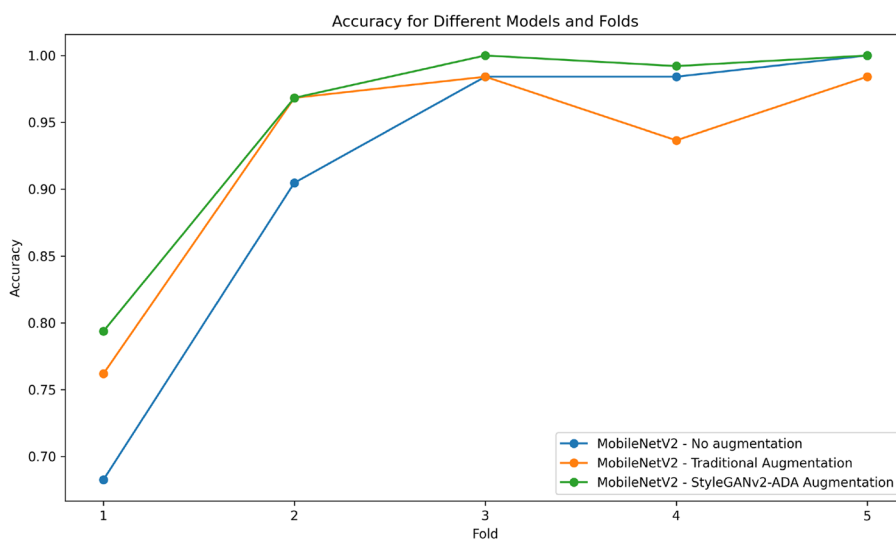
Fig. 5 The comparison of accuracies for three different EfficientNetV2S models on the BraTS 2021 (**a**), the comparison of accuracies for three different MobileNetV2 models on the Gazi Brains 2020 (**b**)

Despite the similar parameter sizes among InceptionV3, DenseNet201, Xception, and EfficientNetV2S, MobileNetV2 consistently demonstrates lower accuracy across all the datasets. This observation suggests that factors beyond parameter size, such as architectural design and model complexity, are crucial in determining model performance. The disparity in performance between MobileNetV2 and the other models with similar parameter sizes indicates that architectural differences may have

a more substantial impact on performance than parameter count alone. Furthermore, the consistent trend of MobileNetV2 underperforming compared to its counterparts across three datasets underscores the importance of architectural considerations in model selection for brain tumor classification tasks. Additionally, the effectiveness of data augmentation techniques, particularly with the StyleGANv2-ADA GAN model, highlights the potential for improving model performance irrespective

Yurtsever *et al. BMC Medical Informatics and Decision Making*   (2024) 24:285

Page 18 of 24
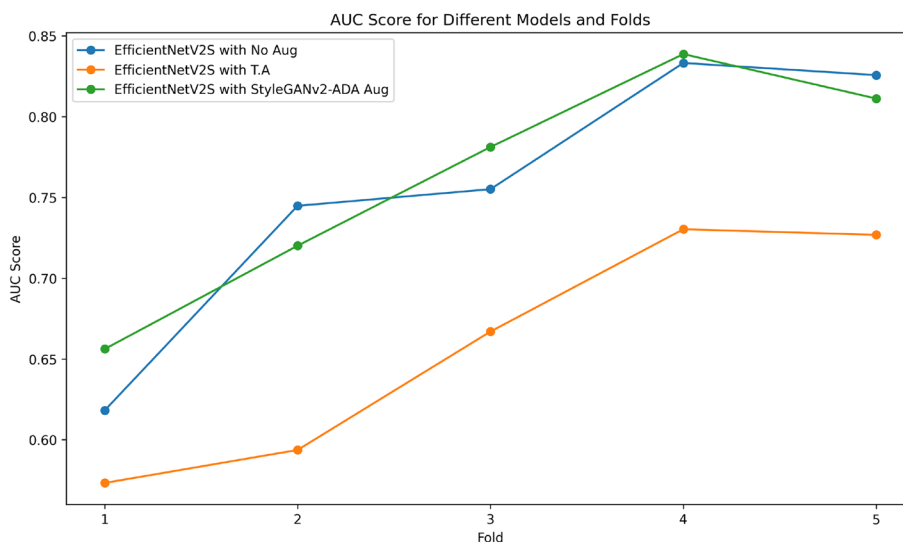
(a)
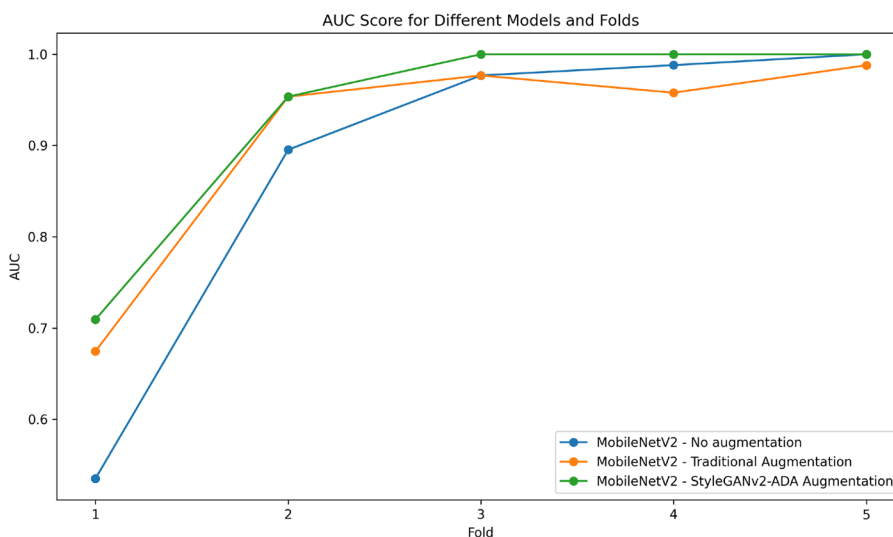
(b)

Fig. 6 The comparison of AUC scores for three different EfficientNetV2S models on the BraTS 2021 (**a**), the comparison of AUC scores for three different MobileNetV2 models on the Gazi Brains 2020 (**b**)

of parameter size, further emphasizing the multifaceted nature of model performance in deep learning tasks.

The Vision Transformer (ViT-Linformer), a transformer-based model, exhibits lower performance compared to the CNN models, especially under traditional augmentation (0.5765 accuracy for the BR35H Dataset) but shows some improvement with the StyleGANv2-ADA method (0.8403 accuracy). Transformers typically require large amounts of data to learn effectively due to their high model complexity and lack of strong inductive

biases (such as translation invariance and locality) that are inherent in CNNs. The BR35H dataset may not provide sufficient data volume or diversity to fully exploit the capabilities of transformers, leading to suboptimal performance. The results achieved in three datasets suggest that while transformer-based models like ViT-Linformer represent the state-of-the-art in many computer vision tasks, their performance is limited on some tasks due to factors such as data scarcity, lack of strong inductive biases for local feature extraction, and potential
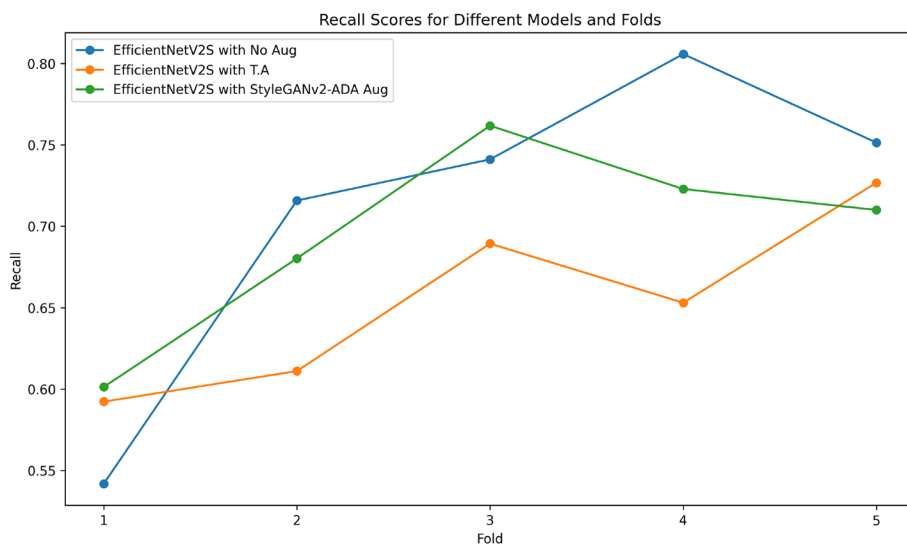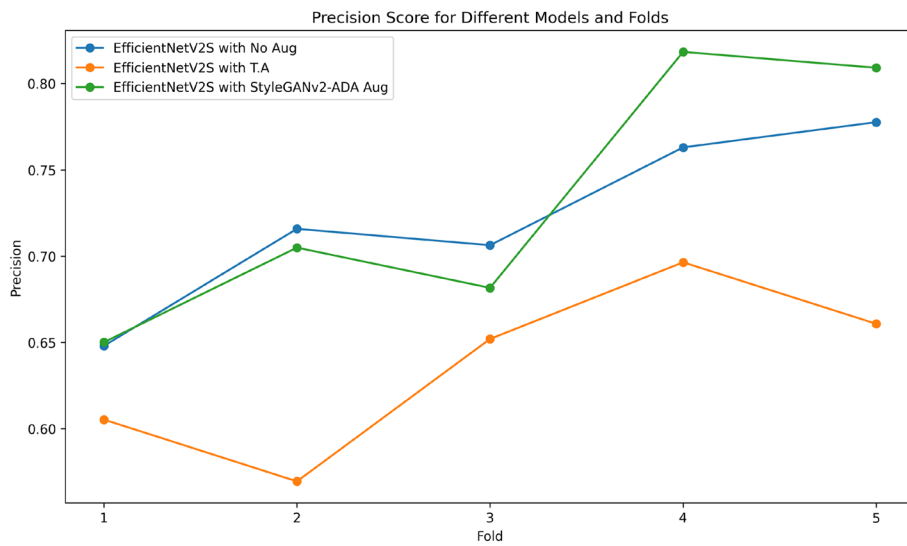
(a)
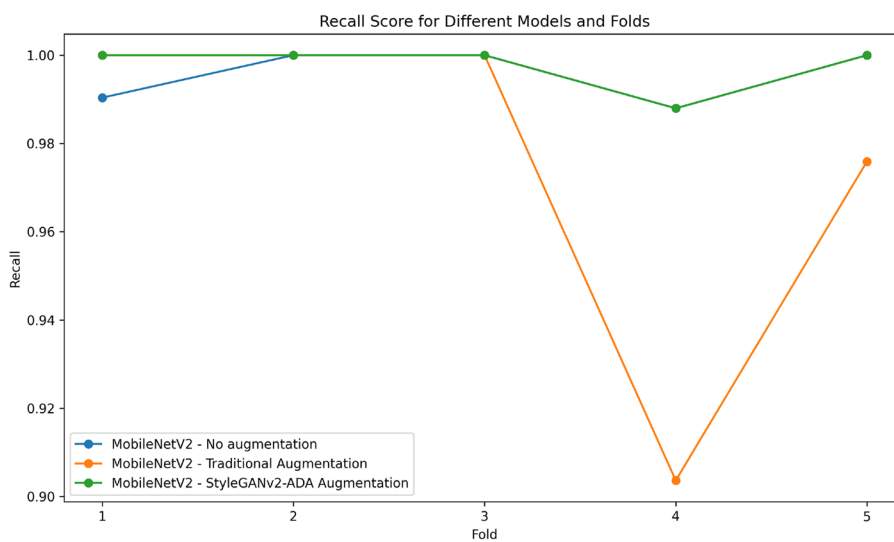
(b)

Fig. 7 The comparison of recall scores for three different MobileNetV2 models on the BraTS 2021 (**a**), the comparison of recall scores for three different MobileNetV2 models on the Gazi Brains 2020 (**b**)

optimization challenges. In contrast, CNN models demonstrate superior performance owing to their suitability for medical imaging tasks, leveraging their local feature extraction capabilities and benefiting more effectively from the provided data augmentation strategies.

As a result, some evaluations are given to present the originality of this study. A distinct strategy was employed to choose the datasets for testing the StyleGAN2-ADA approach in order to strengthen the

study's innovative aspect. Gazi Brains 2020, one of the datasets used, is obtained within the scope of the Turkish brain project and was used for the first time in this study for data augmentation purposes. On the other hand, it was observed that no data augmentation-based study was conducted on the BraTS 2021 and BR35H data sets. For this reason, it is considered that the findings from the study performed for data augmentation offer a novel gain to the literature

(a)
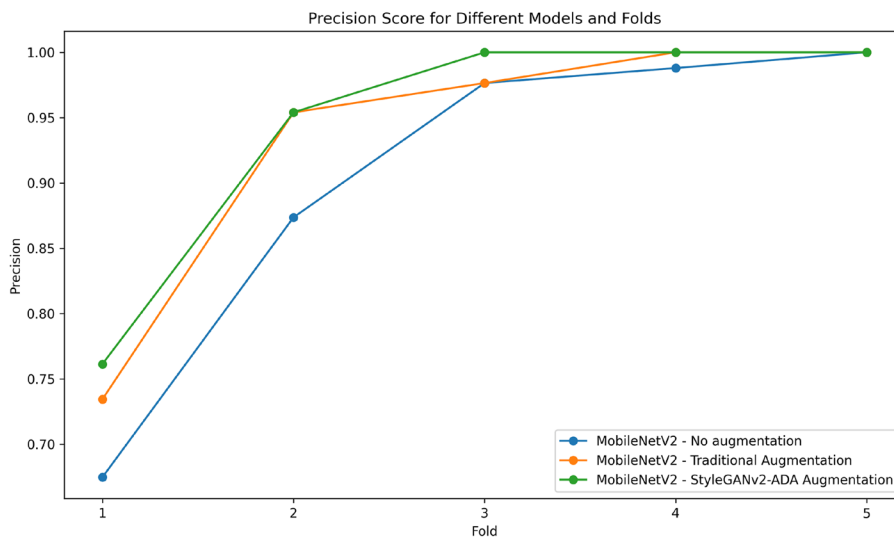
(b)

Fig. 8 The comparison of precision scores for three different EfficientNetV2S models on the BraTS 2021 (**a**), the comparison of precision scores for three different MobileNetV2 models on the Gazi Brains 2020 (**b**)

when assessed on these three data sets. A comprehensive comparison of the utilized StyleGANv2-ADA with state-of-the-art methods in terms of augmentation and classification performance is performed by using Table 1, [45, 63–65]. It can be concluded from the comparison results that using GAN-generated data can significantly improve the overall accuracy of deep learning models used for brain tumor classification.

Conclusion and discussion

This study employs a novel approach for improving brain tumor classification, wherein GAN-based augmentation of MRI slices. The results obtained indicate that the utilization of GAN-generated data can lead to a significant enhancement in the overall accuracy of deep-learning models for brain tumor classification.

Promising results have been demonstrated in addressing one of the main challenges in medical imaging, namely the sparsity of ground truth data, by employing the proposed method utilizing StyleGAN to generate synthetic MRI slices. These synthetic slices can be used to augment the training data, thereby enhancing the performance of deep learning models.

The study also highlights the importance of selecting and evaluating deep-learning models for medical imaging tasks. The performance of several popular deep learning models for brain tumor classification was compared, and it was noticed that the EfficientNetV2S model achieved the highest accuracy.

Furthermore, the adoption of StyleGANv2-ADA for synthesizing realistic, imitated brain MRIs proves to be a crucial aspect of the research. The paper rightly highlights the first-time usage of GAN-based data augmentation on the Gazi Brains 2020 dataset, showcasing the adaptability and effectiveness of the proposed method on a meticulously curated dataset of glioma-type brain tumors.

The comparative analysis of classification performances among different models, with and without augmentation, provides valuable insights into the superiority of the EfficientNetV2S model and the positive impact of GAN-synthesized augmentation. The study's successful application of StyleGANv2-ADA on all three datasets further substantiates the potential of GANs for data augmentation in medical imaging. In conclusion, this research validates the efficacy of GANs in augmenting medical imaging datasets and contributes valuable knowledge about model selection and evaluation for brain tumor classification. The main findings of this study can be listed as follows:

- A novel approach for improving brain tumor classification is proposed, involving GAN-based augmentation of MRI slices, with results indicating a significant enhancement in overall accuracy through the utilization of GAN-generated data.
- Promising results have been demonstrated in addressing one of the main challenges in medical imaging, namely the sparsity of ground truth data, by employing the proposed method utilizing StyleGAN to generate synthetic MRI slices. These synthetic slices can be utilized to augment the training data, thereby enhancing the performance of deep learning models.
- The importance of selecting and evaluating deep-learning models for medical imaging tasks is highlighted, with the performance of several popular deep-learning models for brain tumor classification compared, revealing that the EfficientNetV2S model achieved the highest accuracy.

- Furthermore, the adoption of StyleGANv2-ADA for synthesizing realistic, imitated brain MRIs proves to be a crucial aspect of the research. The paper rightly highlights the first-time usage of GAN-based data augmentation on the Gazi Brains 2020 dataset, showcasing the adaptability and effectiveness of the proposed method on a meticulously curated dataset of glioma-type brain tumors.
- The comparative analysis of classification performances among different models, with and without augmentation, provides valuable insights into the superiority of the EfficientNetV2S model and the positive impact of GAN-synthesized augmentation while achieving the highest overall accuracy for brain tumor classification on all the three datasets.
- The study's successful application of StyleGANv2-ADA on three separate datasets, namely Gazi Brains 2020, BraTS 2021 and the BR35H datasets further substantiates the potential of GANs for data augmentation in medical imaging.
- In conclusion, this research validates the efficacy of GANs in augmenting medical imaging datasets and contributes valuable knowledge about model selection and evaluation for brain tumor classification.

In future works, the integration of emerging technologies like the Internet of Things (IoT) will be explored to enhance healthcare applications. IoT devices will be utilized for real-time monitoring and data collection, providing continuous and accurate patient health information. AI algorithms will be employed to analyze the vast amounts of data collected, facilitating early diagnosis, personalized treatment plans, and improved patient outcomes. IoT and AI are expected to create a more efficient, responsive, and patient-centric healthcare system. This approach will be particularly beneficial in remote and underserved areas, where access to healthcare services is limited. By leveraging these technologies, healthcare providers will be empowered to make more informed decisions, ultimately improving the quality of care and patient satisfaction. Future work will also focus on expanding the proposed model to other types of tumors and medical conditions. While this study has concentrated on brain tumor datasets, the model's architecture and underlying methodologies are adaptable to various medical imaging tasks. By re-training the model on datasets encompassing different tumor types or other medical conditions, the model's utility could be significantly broadened. This approach would allow for a more comprehensive evaluation of the model's performance across diverse medical scenarios, ultimately contributing to improved diagnostic accuracy and clinical outcomes.

Availability of data and materials
The datasets utilized in this study, namely the Gazi Brains 2020, BRaTS 2021, and BR35H are publicly available on the Kaggle platform. The Gazi Brains 2020 dataset can be accessed at https://www.kaggle.com/datasets/gazibrains2020/gazi-brains-2020. The Kaggle website of Gazi Brains 2020 also hosts a supplementary Python notebook to work on the dataset (https://www.kaggle.com/code/gazibrains2020/training-with-pre-processed-dataset-gazibrains20). The BR35H dataset is available at https://www.kaggle.com/datasets/ahmedhamada0/brain-tumor-detection. The BRaTS 2021 dataset is accessible at https://www.kaggle.com/competitions/rsna-miccai-brain-tumor-radiogenomic-classification.

Data availability
The Gazi Brains 2020 Dataset is a collection of glioma-type brain tumors that was meticulously compiled by six medical experts from the Gazi University Faculty of Medicine. This dataset comprises brain MR images from a total of 100 patients, with 50 patients being healthy and 50 patients diagnosed with High-Grade Glioma (HGG). This data can be accessed via the following access links:
 Reference-52: Gazi University - Gazi Brains 2020 Dataset (version 2; pre-processed) (https://www.kaggle.com/datasets/gazibrains2020/gazi-brains-2020-version-2, 2020), Accessed: 2024-02-23
 Reference-58: Gazi Brains 2020 Dataset, (https://www.kaggle.com/datasets/gazibrains2020/gazi-brains-2020), Accessed: 2024-01-12.

Declarations

Ethics approval and consent to participate
All methods were carried out following relevant guidelines and regulations.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

References
1. Ostrom Q, Price M, Neff C, Cioffi G, Waite K, Kruchko C, Barnholtz-Sloan J. CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2015–2019. Neuro-oncology. 2022;24:v1–95.
2. Arabahmadi M, Farahbakhsh R, Rezazadeh J. Deep learning for smart Healthcare-A survey on brain tumor detection from medical imaging. Sensors. 2022;22:1960.
3. Shaker E, El-Hossiny A, Kandil A, Elbialy A, Afify H. Advanced imaging system for brain tumor automatic classification from MRI images using hog and bof feature extraction approaches. Int J Imaging Syst Technol. 2023;33:1661–71.
4. Tiwari P, Pant B, Elarabawy M, Abd-Elnaby M, Mohd N, Dhiman G, Sharma S. CNN based multiclass brain tumor detection using medical imaging. Comput Intell Neurosci. 2022;2022:1–8.
5. Abd-Ellah M, Awad A, Khalaf A, Hamed H. A review on brain tumor diagnosis from MRI images: Practical implications, key achievements, and lessons learned. Magn Reson Imaging. 2019;61:300–18.
6. Chlap P, Huang M, Vandenberg N, Dowling J, Holloway L, Haworth A. A review of medical image data augmentation techniques for deep learning applications. J Med Imaging Radiat Oncol. 2021;65:545–63.
7. Kim S, Kim B, Park H. Synthesis of brain tumor multicontrast MR images for improved data augmentation. Med Phys. 2021;48:2185–98.
8. Saba L, Dey N, Ashour A, Samanta S, Nath S, Chakraborty S, Sanches J, Kumar D, Marinho R, Suri J. Automated stratification of liver disease in ultrasound: an online accurate feature classification paradigm. Comput Methods Prog Biomed. 2016;130:118–34.
9. Biswas M, Kuppili V, Edla D, Suri H, Saba L, Marinhoe R, Sanches J, Suri J. Symtosis: A liver ultrasound tissue characterization and risk stratification in optimized deep learning paradigm. Comput Methods Prog Biomed. 2018;155:165–77.
10. Acharya U, Swapna G, Sree S, Molinari F, Gupta S, Bardales R, Witkowska A, Suri J. A review on ultrasound-based thyroid cancer tissue characterization and automated classification. Technol Cancer Res Treat. 2014;13:289–301.
11. Acharya U, Mookiah M, Vinitha Sree S, Afonso D, Sanches J, Shafique S, Nicolaides A, Pedro L, Fernandes J, Suri J. Atherosclerotic plaque tissue characterization in 2D ultrasound longitudinal carotid scans for automated classification: a paradigm for stroke risk assessment. Med Biol Eng Comput. 2013;51:513–23.
12. Tandel G, Biswas M, Kakde O, Tiwari A, Suri H, Turk M, Laird J, Asare C, Ankrah A, Khanna N, Madhusudhan B, Saba L, Suri J. A Review on a Deep Learning Perspective in Brain Cancer Classification. Cancers. 2019;11(1):111.
13. Badža M, Barjaktarović M. Classification of brain tumors from MRI images using a convolutional neural network. Appl Sci. 2020;10:1999.
14. Huang Z, Du X, Chen L, Li Y, Liu M, Chou Y, Jin L. Convolutional neural network based on complex networks for brain tumor image classification with a modified activation function. IEEE Access. 2020;8:89281–90.
15. Pernas F, Martínez-Zarzuela M, Antón-Rodríguez M, González-Ortega D. A deep learning approach for brain tumor classification and segmentation using a multiscale convolutional neural network. Healthcare. 2021;9:153.
16. Ullah F, Nadeem M, Abrar M, Amin F, Salam A, Alabrah A, AlSalman H. Evolutionary Model for Brain Cancer-Grading and Classification. IEEE Access. 2023;11:126182–94.
17. Ertosun M, Rubin D. Automated grading of gliomas using deep learning in digital pathology images: a modular approach with ensemble of convolutional neural networks. AMIA Annu Symp Proc. 2015;2015:1899.
18. Rehman A, Khan M, Mehmood Z, Tariq U, Noor A. Microscopic brain tumor detection and classification using 3d cnn and feature selection architecture. Microsc Res Tech. 2020;84:133–49.
19. Irmak E. Multi-classification of brain tumor MRI images using deep convolutional neural network with fully optimized framework. Iran J Sci Technol Trans Electr Eng. 2021;45:1015–36.
20. Ullah F, Nadeem M, Abrar M, Amin F, Salam A, Khan S. Enhancing Brain Tumor Segmentation Accuracy through Scalable Federated Learning with Advanced Data Privacy and Security Measures. Mathematics. 2023;11(19):4189.
21. Ullah F, Nadeem M, Abrar M. Revolutionizing brain tumor segmentation in MRI with dynamic fusion of handcrafted features and global pathway-based deep learning. KSII Trans Internet Inf Syst. 2024;18(1):105–25.
22. Kariuki P, Gikunda PK, Wandeto JM. Deep Transfer Learning Optimization Techniques for Medical Image Classification - A Survey. 2023. Authorea Preprints. https://doi.org/10.36227/techrxiv.22638937.v1.

23. Haq EU, Jianjun H, Huarong X, Li K, Weng L. A Hybrid Approach Based on Deep CNN and Machine Learning Classifiers for the Tumor Segmentation and Classification in Brain MRI. Comput Math Methods Med. 2022;2022:6446680. https://doi.org/10.1155/2022/6446680.

24. Goceri E. Medical Image Data Augmentation: Techniques, Comparisons and Interpretations. Artif Intell Rev. 2023;56(11):12561–605.

25. Kumar A, Singh S, Saxena S, Lakshmanan K, Sangaiah A, Chauhan H, Shrivastava S, Singh R. Deep feature learning for histopathological image classification of canine mammary tumors and human breast cancer. Inf Sci. 2020;508:405–21.

26. Claro M, MS Veras R, Santana A, Vogado L, Junior G, Medeiros F, Tavares J. Assessing the impact of data augmentation and a combination of CNNs on leukemia classification. Inf Sci. 2022;609:1010–29.

27. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial networks. Commun ACM. 2020;63:139–44.

28. Allah A, Sarhan A, Elshennawy N. Classification of brain MRI tumor images based on deep learning PGGAN augmentation. Diagnostics. 2021;11:2343.

29. Han C, Rundo L, Araki R, Nagano Y, Furukawa Y, Mauri G, Nakayama H, Hayashi H. Combining noise-to-image and image-to-image GANs: brain MR image augmentation for tumor detection. IEEE Access. 2019;7:156966–77.

30. Iqbal T, Ali H. Generative Adversarial Network for Medical Images (MI-GAN). J Med Syst. 2018;42:1–11.

31. Mok T, Chung A. Learning data augmentation for brain tumor segmentation with coarse-to-fine generative adversarial networks. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 4th International Workshop, BrainLes 2018, Held In Conjunction With MICCAI 2018, Granada, Spain, September 16, 2018, Revised Selected Papers, Part I 4. Cham: Springer; 2019. p. 70–80.

32. Asiri A, Shaf A, Ali T, Aamir M, Usman A, Irfan M, Alshamrani H, Mehdar K, Alshehri O, Alqhtani S. Multi-Level Deep Generative Adversarial Networks for Brain Tumor Classification on Magnetic Resonance Images. Intell Autom Soft Comput. 2023;36:127–43.

33. Garcea F, Serra A, Lamberti F, Morra L. Data augmentation for medical imaging: A systematic literature review. Comput Biol Med. 2022:106391.

34. Wang Y, Ji Y, Xiao H. A data augmentation method for fully automatic brain tumor segmentation. Comput Biol Med. 2022;149:106039.

35. Alsaif H, Guesmi R, Alshammari B, Hamrouni T, Guesmi T, Alzamil A, Belguesmi L. A novel data augmentation-based brain tumor detection using convolutional neural network. Appl Sci. 2022;12:3773.

36. Qin T, Wang Z, He K, Shi Y, Gao Y, Shen D. Automatic data augmentation via deep reinforcement learning for effective kidney tumor segmentation. ICASSP 2020-2020 IEEE International Conference On Acoustics, Speech And Signal Processing. Barcelona: IEEE; 2020. p. 1419–23.

37. Han C, Rundo L, Araki R, Furukawa Y, Mauri G, Nakayama H, Hayashi H. Infinite Brain MR Images: PGGAN-Based Data Augmentation for Tumor Detection. In: Esposito A, Faundez-Zanuy M, Morabito F, Pasero E, editors. Neural Approaches to Dynamics of Signal Exchanges, Smart Innovation, Systems and Technologies, vol 151. Singapore: Springer; 2020. p. 291–303. https://doi.org/10.1007/978-981-13-8950-4_27.

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

39. Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J, Greenspan H. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. Neurocomputing. 2018;321:321–31.

40. Guan Q, Chen Y, Wei Z, Heidari A, Hu H, Yang X, Zheng J, Zhou Q, Chen H, Chen F. Medical image augmentation for lesion detection using a texture-constrained multichannel progressive GAN. Comput Biol Med. 2022;145.

41. Bowles C, Chen L, Guerrero R, Bentley P, Gunn R, Hammers A, Dickie D, Hernández M, Wardlaw J, Rueckert D. GAN Augmentation: Augmenting Training Data Using Generative Adversarial Networks. 2018. ArXiv Preprint ArXiv:1810.10863.

42. Deepak S, Ameer PM. MSG-GAN based synthesis of brain MRI with meningioma for data augmentation. In 2020 IEEE International Conference On Electronics, Computing and Communication Technologies (CONECCT). Bangalore: IEEE; 2020. p. 1–6.

43. Han C, Hayashi H, Rundo L, Araki R, Shimoda W, Muramatsu S, Furukawa Y, Mauri G, Nakayama H. GAN-based synthetic brain MR image generation. 2018 IEEE 15th International Symposium On Biomedical Imaging (ISBI 2018). Washington, DC: IEEE; 2018. p. 734–38.

44. Ghassemi N, Shoeibi A, Rouhani M. Deep neural network with generative adversarial networks pre-training for brain tumor classification based on MR images. Biomed Signal Process Control. 2020;57:101678.

45. Alrashedy H, Almansour A, Ibrahim D, Hammoudeh M. BrainGAN: brain MRI image generation and classification framework using GAN architectures and CNN models. Sensors. 2022;22:4297.

46. Ge C, Gu I, Jakola A, Yang J. Enlarged training dataset by pairwise GANs for molecular-based brain tumor classification. IEEE Access. 2020;8:22560–70.

47. Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J, Greenspan H. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. Neurocomputing. 2018;321:321–31.

48. Pei L, Vidyaratne L, Rahman MM, Iftekharuddin K. Context Aware Deep Learning for Brain Tumor Segmentation, Subtype Classification, and Survival Prediction using Radiology Images. Sci Rep. 2020;10:19726. https://doi.org/10.1038/s41598-020-74419-9.

49. Zhang J, Lv X, Zhang H, Liu B. Aresu-net: attention residual u-net for brain tumor segmentation. Symmetry. 2020;12:721.

50. Montaha S, Azam S, Rafid A, Hasan M, Karim A, Islam A. Timedistributed-cnn-lstm: a hybrid approach combining cnn and lstm to classify brain tumor on 3d mri scans performing ablation study. IEEE Access. 2022;10:60039–59.

51. Wang G, Li W, Ourselin S, Vercauteren T. Automatic Brain Tumor Segmentation Based on Cascaded Convolutional Neural Networks with Uncertainty Estimation. Front Comput Neurosci. 2019;13:56. https://doi.org/10.3389/fncom.2019.00056.

52. Menze B, Jakab A, Bauer S, Kalpathy-Cramer J, Farahani K, Kirby J, Burren Y, Porz N, Slotboom J, Wiest R. Others, The multimodal brain tumor image segmentation benchmark (BRATS). IEEE Trans Med Imaging. 2014;34:1993–2024.

53. Bakas S, Reyes M, Jakab A, Bauer S, Rempfler M, Crimi A, Shinohara R, Berger C, Ha S, Rozycki M, Others. Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge. 2018. ArXiv Preprint ArXiv:1811.02629.

54. Gazi Brains 2020 Dataset. 2020. https://www.kaggle.com/datasets/gazibrains2020/gazi-brains-2020. Accessed 12 Jan 2024.

55. Gazi University - Gazi Brains 2020 Dataset (version 2; pre-processed). 2020. https://www.kaggle.com/datasets/gazibrains2020/gazi-brains-2020-version-2. Accessed 23 Feb 2024.

56. Özkaya Ç, Sağiroğlu Ş. Glioma Grade Classification using CNNs and Segmentation with an Adaptive Approach using Histogram Features in Brain MRIs. IEEE Access. 2023;11:52275–87.

57. Hamada A. BR35H Dataset. 2021. https://www.kaggle.com/datasets/ahmedhamada0/brain-tumor-detection. Accessed 17 Aug 2024.

58. Karras T, Aittala M, Hellsten J, Laine S, Lehtinen J, Aila T. Training Generative Adversarial Networks with Limited Data. Adv Neural Inf Process Syst. 2020;33:12104–14.

59. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L. MobileNetV2: Inverted Residuals and Linear Bottlenecks, IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City: IEEE; 2018. p. 4510–20. https://doi.org/10.1109/CVPR.2018.00474.

60. Chollet F. Xception: Deep Learning with Depthwise Separable Convolutions. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE; 2017. p. 1800–7. https://doi.org/10.1109/CVPR.2017.195.

61. Tan M, Le Q. Efficientnetv2: Smaller models and faster training. International Conference On Machine Learning. 2021. p. 10096–106.

62. Wang S, Li BZ, Khabsa M, Fang H, Ma H. Linformer: Self-attention with linear complexity. 2020. arXiv preprint arXiv:2006.04768.

63. Qi C, Chen J, Xu G, Xu Z, Lukasiewicz T, Liu Y. SAG-GAN: Semi-Supervised Attention-Guided GANs for Data Augmentation on Medical Images. arXiv e-prints, Art. no. arXiv:2011.07534. 2020. https://doi.org/10.48550/arXiv.2011.07534.

Yurtsever *et al. BMC Medical Informatics and Decision Making*   (2024) 24:285

Page 24 of 24

64. Naseer A, Yasir T, Azhar A, Shakeel T, Zafar K. Computer-aided brain tumor diagnosis: performance evaluation of deep learner CNN using augmented brain MRI. Inter J Biomed Imaging. 2021;2021:5513500.
65. Esmaeili M, Toosi A, Roshanpoor A, Changizi V, Ghazisaeedi M, Rahmim A, Sabokrou M. Generative adversarial networks for anomaly detection in biomedical imaging: a study on seven medical image datasets. IEEE Access. 2023;11:17906–21.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.