GAZİ
JOURNAL OF ENGINEERING SCIENCES

Generating Synthetic Images from Real MR Images Using Deep Learning Methods

Ercüment Güvenç [*a], **Gürcan Çetin** [b], **Mevlüt Ersoy** [c]

[a,*] Muğla Sıtkı Koçman University, Departmant of Informatic 48000 - Muğla, Türkiye
Orcid: 0000-0003-0053-9623
e mail: eguvenc@mu.edu.tr

[b] Muğla Sıtkı Koçman University, Faculty of Technology, Department of Information Systems Engineering 48000 - Muğla, Türkiye
Orcid: 0000-0003-3186-2781

[c] Suleyman Demirel University, Faculty of Engineering, Deparment of Computer Engineering 32260 – Isparta, Türkiye
Orcid: 0000-0003-2963-7729

[*]Corresponding author:
eguvenc@mu.edu.tr

ABSTRACT

One of the most important technological developments in the field of medicine is Computed Tomography and Magnetic Resonance imaging techniques. This technique allows the size and shape of tumor areas in body tissues to be determined, making it easier for specialists to determine the type of tumor as well as whether it is benign or malignant. Various deep learning-based computer software have been developed to accurately detect tumor areas in tissue. Due to the lack of image data used in deep learning studies, a limitation naturally arises in studies in this field. In order to eliminate the lack of image data in these studies, image augmentation can be performed using deep learning methods as well as data augmentation methods using various image processing techniques. In this study, Generative Adversarial Networks, a deep learning technique, were employed to duplicate brain MR images and generate synthetic images. After the resulting MR images were made usable by undergoing various pre-processing, similarity rates to real images were calculated using metrics such as Peak Signal-to-Noise Ratio, Structural similarity index and Mean Square Error, and by looking at these rates, realistic images were added to the data set and the data set was expanded.

Derin Öğrenme Yöntemleri Kullanılarak Gerçek MR Görüntülerinden Sentetik Görüntülerin Üretilmesi

ÖZ

Tıp alanındaki en önemli teknolojik gelişmelerden biri Bilgisayarlı Tomografi ve Manyetik Rezonans görüntüleme teknikleridir. Bu teknik, vücut dokularındaki tümör alanlarının boyutunun ve şeklinin belirlenmesine olanak tanıyarak, uzmanların tümörün tipinin yanı sıra iyi huylu veya kötü huylu olup olmadığını belirlemesini kolaylaştırır. Dokudaki tümör alanlarını doğru bir şekilde tespit etmek için çeşitli derin öğrenme tabanlı bilgisayar yazılımları geliştirilmiştir. Derin öğrenme çalışmalarında kullanılan görüntü verilerinin eksikliğinden dolayı bu alanda yapılan çalışmalarda doğal olarak bir sınırlılık ortaya çıkmaktadır. Bu çalışmalarda görüntü verisi eksikliğini gidermek amacıyla çeşitli görüntü işleme teknikleri kullanılarak veri büyütme yöntemlerinin yanı sıra derin öğrenme yöntemleri kullanılarak da görüntü büyütme gerçekleştirilebilmektedir. Bu çalışmada, beyin MR görüntülerini çoğaltmak ve sentetik görüntüler oluşturmak için derin öğrenme tekniği olan Üretken Çekişmeli Ağlar kullanıldı. Elde edilen MR görüntüleri çeşitli ön işlemlerden geçirilerek kullanılabilir hale getirildikten sonra Tepe Sinyal-Gürültü Oranı, Yapısal benzerlik indeksi ve Ortalama Karesel Hata gibi metrikler kullanılarak gerçek görüntülere benzerlik oranları hesaplanmış ve bu oranlara bakılarak gerçekçi sonuçlar elde edilmiştir. Veri setine görseller eklenmiş ve veri seti genişletilmiştir.

1. Introduction

With the advancement of computers in today's world, the processing of large datasets in the field of artificial intelligence and the generation of results can be achieved rapidly. Particularly in studies using deep learning models, the inadequacy of computer resources has been attempted to be addressed with the use of cloud systems in recent years. However, in conducted studies, the creation of datasets has emerged as a significant challenge, requiring considerable effort and performance, especially in studies using image data.

In studies, especially in the field of medical image processing, errors and deficiencies in the results obtained are encountered due to the insufficient availability of MR images. Various data augmentation processes have been employed in the literature to address this data deficiency, and it has been observed that research results are positively affected.

In recent years, machine learning and its subfield, deep learning, have become quite popular in artificial intelligence studies in various fields. Artificial intelligence mimics the logic of the human brain to solve problems related to decision-making and prediction [1]. Therefore, deep learning, like artificial intelligence, is based on a series of algorithms inspired by the structure and functioning of the brain. Deep learning has become a popular area in data science, being used in various fields such as medical, robotics, image recognition, image generation, autonomous vehicles, natural language processing, space, and defense industries [2].

Deep neural networks, which use large-scale data, have demonstrated better performance in extracting semantic information from images compared to classical image processing techniques. The approaches proposed in deep learning have also shown significant success in image generation. In this study, Generative Adversarial Networks (GAN), introduced by Ian Goodfellow in 2014 and widely used in various fields today, were employed for the generation of synthetic MR images.

Challenges in obtaining data for training in artificial intelligence studies, the laborious nature of real-world data collection, and the requirement for expertise from different disciplines in the labeling processes are limitations in this field. For instance, difficulties in obtaining MR images in a real hospital environment include the presence of different types of tumors, limited numbers of individuals with the same disease, or the inability to capture a standard image quality. Therefore, research is often conducted using ready-made datasets, such as ImageNet, CIFAR10/100, and BRaTS. However, upon examining these datasets, it becomes evident that datasets containing MR images are limited for high-performance deep learning models.

Synthetic data generation processes are often preferred in many studies as a solution to performance issues arising from the lack of images in datasets in deep learning studies using medical images or insufficient images in face recognition models. The inadequacy of deep learning models used due to the lack of image data in training is a common issue in the foundations of these studies. Therefore, the deficiency of image data in the datasets used during the training of these models has been attempted to be addressed through the generation of synthetic images, leading to significant improvements in model performance. In this study, some methods used for generating synthetic image data to address the deficiency of image data are discussed, and methods preferred for calculating the similarity ratios of generated images to real images are mentioned.

2. Method

2.1. Dataset

In the conducted study, MR images used in the segmentation of brain tumors were preferred. It is crucial for artificial intelligence-based studies that the tumor regions on MR images are accurately labeled. This is because the goal of artificial intelligence studies is to predict results closest to expert opinions. The calculation of these results and the ability to make predictions are directly related to the training of the deep neural network. For this training to be conducted accurately, the data in the dataset must be correctly labeled.

Especially in deep learning studies involving image data, ready-made datasets are preferred due to the difficulty of accessing labeled image data [3]. Ready-made datasets are prepared by various experts at universities worldwide, and competitions are organized in specified fields after the labeling of images is completed. One such dataset is the BraTS dataset, known as the "Multimodal Brain Tumor Segmentation

Challenge." In the conducted study, images from the Flair sequence of the BRATS dataset were chosen as the source of synthetic image data to be generated. Images from different sequences in the dataset are shown in Figure 1.
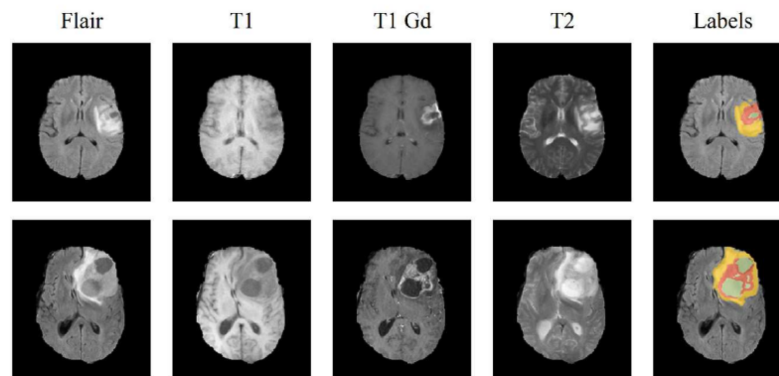

Figure 1. MR images in different sequences in the dataset

2.2. Image augmentation methods

Upon reviewing the conducted studies, it is observed that, in addition to image processing methods, various deep learning-based methods have also been employed for image data augmentation. Image processing techniques such as Translation, Cropping, Rotation, and Color Space are preferred, and in deep learning-based methods, the Generative Adversarial Network (GAN) method is among the preferred techniques for generating synthetic image data.

2.2.1. Image processing methods

Translation: Horizontal axis translation is much more common than vertical axis rotation. This data augmentation method is one of the easiest to apply, and it has yielded successful results in datasets such as CIFAR-10 and ImageNet [4].

Color Space: Enlargements in the color channels by performing simple matrix operations to increase or decrease the brightness of an image is another practical method for data augmentation. RGB values can be easily modified with this method [4].

Cropping: Image cropping is a practical image processing technique that involves cropping the central part of each image for image data with mixed height and width dimensions. In addition, the random cropping method can be used to achieve an effect similar to the rotation process. In random cropping, changes are made to the size of the image in pixels [4].

Rotation: The rotation method is performed by rotating the image to the right or left on an axis between 1° and 359°. The rotation degree parameter is crucial for data augmentation using the rotation method. Mild rotations between 1 and 20 or -1 and -20, for example, can be useful in tasks such as step recognition, as seen in some ready-made datasets. However, as the rotation degree increases, distortions in the labels of the data may occur [4].

Generative Adversarial Netwoks - GAN: Deep neural networks that utilize large-scale data have demonstrated better performance in extracting semantic information from images compared to classical image processing techniques. The approaches proposed in deep learning have also shown significant success in image generation. The first model developed for generating images using deep neural networks was proposed by Ian Goodfellow in 2014. The Generative Adversarial Network (GAN) model is heavily used in various fields such as style transfer, tabular data generation, image enhancement, sound transfer, and artificial face generation.

The GAN model has been successfully employed for tasks such as generating entirely new faces that have never existed before, in addition to its widespread use in style transfer, tabular data generation, image enhancement, and voice transfer [5].
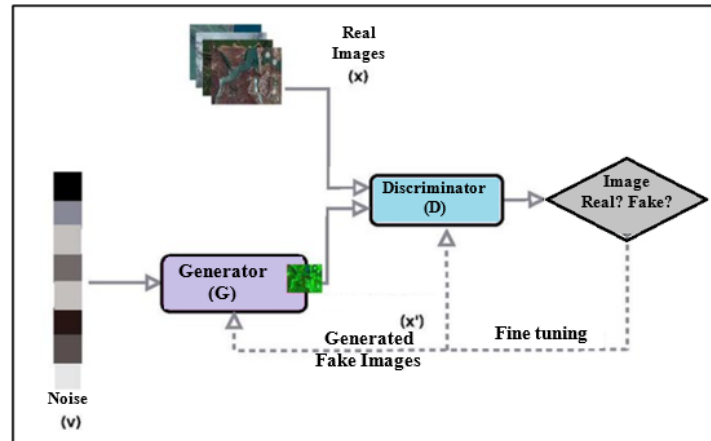
Figure 2. GAN architecture

The structure of the GAN model consists of two distinct deep networks, as seen in Figure 2: the generator and the discriminator. These two networks in the GAN structure work adversarially, learning from each other in a competitive manner.

The discriminator, a deep network, attempts to distinguish generated images by comparing them to real images, while the generator, another deep network, tries to produce images that are as close as possible to the real images based on the input noise signals. In this way, the discriminator and generator networks are competitively trained. The training process continues until a certain number of iterations are reached, and after the training of the networks is completed, the generator network produces new images resembling real images [6].

Generative Adversarial Networks (GANs) have become popular in recent years for data augmentation processes. In the study titled "GAN-Based Synthetic Brain MR Image Generation," Changhee Han et al. successfully generated synthetic MR images using specific slices from MR images in the BRATS 2016 dataset with DCGAN and WGAN models. They demonstrated that these models could be used for generating medical images [7].

The GAN model creates artificial data from a dataset while preserving features similar to the original dataset. The successful performance of Generative Adversarial Networks in image generation has made it an important model among data augmentation methods [4].

2.3. Synthetic Image Data Generation

The DCGAN model (Deep Convolutional Generative Adversarial Network) used in the study is structurally similar to the standard GAN model. In this model, deep convolutional neural networks have been added to the standard GAN architecture. This feature allows DCGAN to process larger and more complex images more efficiently, leading to more stable results. In the standard GAN architecture, the output size from the previous layer is adjusted to be equal to the input size of the previous layer, whereas the DCGAN architecture can modify these dimensions using image slices. As a result, the model used is made more suitable for processing larger and more complex images. The architectural structure of the DCGAN model used in the application is illustrated in Figure 3.
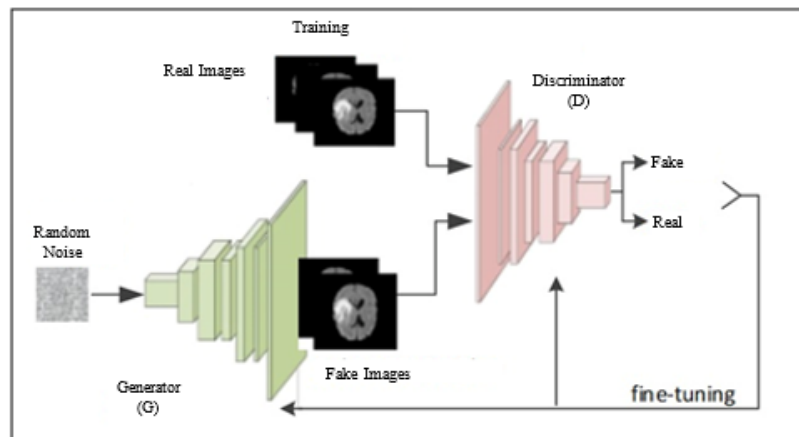
Figure 3. The structure of the DCGAN model used in the application

According to Figure 3, the Generator network (G) takes a random noise vector and then performs the "up-sampling" process to create feature maps at different scales. The goal of training the Generator network is to learn real examples and generate fake examples to deceive the Discriminator network (D). This process is achieved by minimizing the distribution distance between generated and real examples. Each "up-sampling" operation is followed by a batch normalization layer and an activation function. All activation functions, except the tanh function in the final layer, are applied as ReLU. The Discriminator network takes both generated and real examples, and its training objective is to distinguish generated examples from real images. The objective function value of the Discriminator network represents the likelihood of the input example being real, making it a binary classification problem. This training process continues until the Discriminator network cannot distinguish input examples from real images (loss function value = 0) [8].

In the conducted study, 144 MR images from 210 patients in the dataset, where the tumor regions were identified as clear by expert doctors, were used as the training dataset for the DCGAN model. The prepared DCGAN model, trained on Colab Pro, generated approximately 250,000 images over an approximate working time of 8-10 hours. The similarities of the generated images to the real images were compared based on the processing steps of real images, and the number of model iterations (epochs) was determined. Some sample images from the DCGAN model at specific iterations are shown in Figure 4.
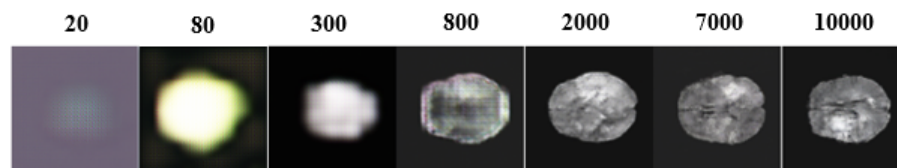

Figure 4. The images generated by the DCGAN model
at specific iterations are as follows

The generated images are created in the same dimensions as the real MR images, with a size of 240x240 pixels. Due to the large number of generated images and the need to distinguish those that are similar to MR images, expert doctor opinions are required. However, to overcome the challenges of time and labor, an attempt has been made to establish an automatic classifier by calculating image similarity coefficients mentioned in the literature.

To determine whether the synthetic image data meets the desired criteria, it is essential to calculate the similarity of the generated images to real images or obtain an expert opinion. Considering the challenges of consulting an expert and the time that would be lost, the importance of automatically calculating similarities in a computer environment for faster results becomes significant. Various methods are used to perform these calculations, including Mean Square Error (MSE), Structural Similarity Index (SSIM), and Peak Signal-to-Noise Ratio (PSNR).

Mean Square Error (MSE): Mean Square Error (MSE) is a commonly used similarity measure in image processing studies. This metric calculates the average of the squared differences between pixel values of two

images. MSE is expressed by the following formula:

$$\mathrm{MSE}(I, J) = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} (I(i,j) - J(i,j))^2 \tag{1}$$

Here, I and J represent the two compared images, and m and n denote the dimensions of the images. The closer the value of MSE is to zero, the more similar the two images are considered.

The use of MSE is prevalent, especially to quantitatively express how similar or different one image is compared to another. However, MSE has some disadvantages. For instance, small differences in pixel values can lead to large errors and may exhibit inconsistency with differences perceived by the human eye. Therefore, when used alone, MSE might not adequately reflect perceptual similarity [9].

In this context, the use of MSE is often preferred in combination with other similarity measures to obtain more reliable results.

Structural Similarity Index (SSIM): The Structural Similarity Index (SSIM) is a similarity measure used in image processing to evaluate the structural similarity between two images. It is designed to better mimic the perception of the human eye and often outperforms simple error metrics like Mean Square Error (MSE) in many scenarios.

SSIM provides a similarity index incorporating three fundamental components: brightness, contrast, and structure. In general, SSIM is expressed by the following formula:

$$\mathrm{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_x\sigma_y + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{2}$$

Here, x and y represent the two compared images, μx and μy are the mean values of the images, σ_x^2 and σ_y^2 are the variances, σxy is the covariance, and C1 and C2 are small constants preventing division by zero errors.

The value of SSIM typically ranges between -1 and 1, where values closer to 1 indicate a higher similarity between the two images. SSIM can be more sensitive to contrast reductions and structural changes, leading to results that are closer to the perception of the human eye.

Among image similarity metrics, SSIM is a preferred metric, especially in applications such as medical imaging, video quality assessment, and similar fields where a more nuanced evaluation of image similarity is required [10].
.
Peak Signal-to-Noise Ratio (PSNR): Peak Signal-to-Noise Ratio (PSNR) is a metric commonly used in image processing, especially for assessing image quality. It measures the similarity between the original and processed images and is generally considered as an error metric. A higher PSNR value indicates increased similarity between two images.

PSNR is associated with Mean Square Error (MSE) and is typically expressed by the following formula:

$$\mathrm{PSNR} = 10 \cdot \log_{10} \left(\frac{\text{Max Pixel Value}^2}{\text{MSE}} \right) \tag{3}$$

In this formula, the Max Pixel Value represents the maximum value a pixel can take, depending on the image format. MSE expresses the mean value of the squared errors between the original and processed images. PSNR is usually measured in decibels.

One advantage of PSNR is its simple and fast computation. However, it is known that PSNR does not fully reflect the perception of the human eye and may be inadequate in certain situations. Particularly, when errors have different effects on different regions of the image, the performance of PSNR may decrease.

Therefore, PSNR is generally not used alone for image quality assessment; it is often used in conjunction with

other similarity metrics and human eye evaluations to provide a more comprehensive evaluation [9].

These metrics are used to objectively assess how close the generated images are to the real ones. However, each metric has its own advantages and limitations; therefore, it is generally recommended to use multiple metrics together.

Acceptable value ranges for metrics are general guidelines and can vary depending on specific use cases. The interpretation of values should take into account visual quality and similarity, and acceptable similarity ranges should be determined based on a specific application.

3. Results and Discussion

In the conducted study, the DCGAN model was employed to generate synthetic images. For synthetic image generation with the DCGAN model, 120 images were selected as training data, and synthetic images were generated until reaching 10,000 epochs. A method was developed to select synthetic MR images that resemble the real MR images from the generated 10,000 synthetic images. According to this method, samples selected from the real dataset images were compared with the generated synthetic MR images using MSE, SSIM, and PSNR algorithms to reveal similarity ratios. Thus, the automatic separation of numerous images within synthetic MR images was achieved, facilitating the image separation process compared to the expert-based image separation process in previous studies.

In the study, 12 randomly selected real MR images from the dataset were compared individually with the generated 10,000 synthetic MR images. The automatically separated synthetic MR images, based on the valid values in the similarity algorithms, were copied to different folders, aiming to ease the workload of medical experts. The valid boundaries for similarity ratios were determined separately for each of the three algorithms. The synthetic MR images that met these criteria were separated into different folders. For the MSE algorithm, the threshold value was set at 0.0145, determined as the value closest to 0, and images below this threshold were grouped in a folder as real MR images. The same process was repeated for 12 randomly selected real MR images. This process was conducted for the SSIM algorithm, with a threshold of 0.11, and the automatically selected synthetic MR images were copied to a different folder created for the SSIM algorithm. Finally, for the PSNR algorithm, an acceptable similarity ratio range of 60-80 dB, commonly preferred for 16-24-bit grayscale images, was determined. The results obtained by the algorithms are presented in Table 1 after performing these processes for three different algorithms.

Table 1. Number of Images Automatically Separated According to Algorithms

MR Index	Compared Real MR Image	Number of Synthetic MR Images Separated Based on Valid Similarity Ratios		
		MSE	SSIM	PSNR
7	**Brats18_2013_7_1_flair.ni_z036**	182	295	8391
8	**Brats18_CBICA_AAB_1_flair.ni_z033**	319	155	8097
12	**Brats18_CBICA_AME_1_flair.ni_z023**	75	67	8175
62	**Brats18_TCIA01_411_1_flair.ni_z005**	19	40	8712
72	**Brats18_TCIA02_171_1_flair.ni_z020**	357	88	8266
79	**Brats18_TCIA02_321_1_flair.ni_z021**	285	109	8324
87	**Brats18_TCIA02_605_1_flair.ni_z015**	4	53	8623
92	**Brats18_TCIA03_199_1_flair.ni_z027**	10	140	7885
94	**Brats18_TCIA03_498_1_flair.ni_z024**	14	62	8635
96	**Brats18_TCIA04_328_1_flair.ni_z020**	95	227	8368
110	**Brats18_TCIA08_205_1_flair.ni_z030**	177	169	7804
119	**Brats18_TCIA08_469_1_flair.ni_z019**	27	88	8620

In Table 1, the index numbers and names of 12 real MR images randomly selected from the data set are given, and the numbers of synthetic MR images that are automatically separated as a result of comparing these images with 10000 MR images produced using the determined algorithms are given. When the comparison

was made using the MSE algorithm in Table 1, the most similar image among the synthetic MR images was the image with index number 72. As a result of comparing the same image with the SSIM algorithm, 88 images were selected, and as a result of comparing it with PSNR, 8266 images were selected. In the comparison made using the SSIM algorithm, the most similar image was the image with index number 7, while 182 images were selected as a result of comparing this image with the MSE algorithm, and 8391 images were selected as a result of the comparison with PSNR. While the most similar image selection using the PSNR algorithm was achieved with the image with index number 62, 19 images could be selected as a result of comparing this image with the MSE algorithm, and 40 images could be selected as a result of the comparison with SSIM. According to the table, 1564 images were parsed with the MSE algorithm, 1493 with the SSIM algorithm and a total of 99900 images were parsed with the PSNR algorithm.

Table 2 shows the number of MR images that can be added to the training data as a result of the verification of these separated synthetic MR images by the specialist doctor and their ratios to the separated images.

Table 2. Number of synthetic MR images verified by the expert

Number and ratio of images separated by the expert									
MR Index	Compared Real MR Image	MSE			SSIM			PSNR		
7	**Brats18_2013_7_1_flair.ni_z036**	182	13	7.14%	295	69	23.39%	8391	868	10.34%
8	**Brats18_CBICA_AAB_1_flair.ni_z033**	319	216	67.71%	155	59	38.06%	8097	861	10.63%
12	**Brats18_CBICA_AME_1_flair.ni_z023**	75	14	18.67%	67	8	11.94%	8175	871	10.65%
62	**Brats18_TCIA01_411_1_flair.ni_z005**	19	3	15.79%	40	5	12.50%	8712	878	10.08%
72	**Brats18_TCIA02_171_1_flair.ni_z020**	357	122	34.17%	88	12	13.64%	8266	873	10.56%
79	**Brats18_TCIA02_321_1_flair.ni_z021**	285	46	16.14%	109	14	12.84%	8324	867	10.42%
87	**Brats18_TCIA02_605_1_flair.ni_z015**	4	1	25.00%	53	11	20.75%	8623	869	10.08%
92	**Brats18_TCIA03_199_1_flair.ni_z027**	10	4	40.00%	140	39	27.86%	7885	877	11.12%
94	**Brats18_TCIA03_498_1_flair.ni_z024**	14	6	42.86%	62	12	19.35%	8635	915	10.60%
96	**Brats18_TCIA04_328_1_flair.ni_z020**	95	13	13.68%	227	34	14.98%	8368	867	10.36%
110	**Brats18_TCIA08_205_1_flair.ni_z030**	177	136	76.84%	169	60	35.50%	7804	865	11.08%
119	**Brats18_TCIA08_469_1_flair.ni_z019**	27	7	25.93%	88	8	9.09%	8620	978	11.35%

When the values in Table 2 are examined, it is observed that the images with high rates of verification are the images selected using the MSE algorithm. According to Table 2, the real MR image in which the most correct image was selected was the image with index number 110 with an accuracy of 76.84%. When this image was compared with the MSE algorithm, 177 similar images were found and 76.84% of them were determined to be the desired type of MR image. When the values of the same image for other algorithms were examined, it was determined that there were 169 images separated using the SSIM algorithm and 35.5% of these images were the desired type of MR image. In the PSNR algorithm, 865 of 7804 images were determined to be suitable.

When the ratios of verified images to parsed images according to the algorithms in Table 2 were examined, it was determined that the images parsed by the MSE algorithm and SSIM algorithms were more similar to the real MR image. While it has been determined that in these two algorithms, attention is paid to the similarity of the images separated from the produced synthetic images to the compared MR image, the PSNR algorithm pays attention to depth, resolution, etc. apart from the similarity of the two images in the comparison. Since it includes features such as, it considered the images in almost the entire data set to be similar.

Figure 5 shows the ratio of the separated images and expert-verified images as a result of comparing the algorithms with the synthetic MR images produced according to each image selected from the data set. According to Figure 5, it was seen that the images parsed with the MSE algorithm were more similar to the selected images.
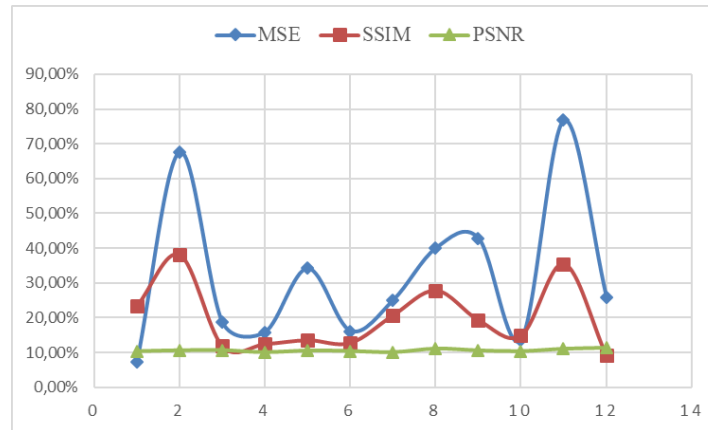
Figure 5. Rates of images verified according to algorithms

Some studies in the literature show that the production of medical images using the GAN method helps improve the performance of classification or clustering algorithms. In their study for the detection of abnormalities in the brain, Raja and Vijayachitra developed a different method using Gan, K-means and MobileNet algorithms and models. In this study, they used the Gan model to produce synthetic images from real MR images. According to the results of the study, the proposed model achieved the best scores among different models with 98% image similarity rate, 99% accuracy rate and 98% F1-score values [11].

In another study, Kazuhiro and his colleagues created a training data set from MR images of the T1 sequence and used the DCGAN method to obtain new images from these images. In the study, images were examined by 5 different radiologists to distinguish the images produced by DCGAN. As a result, it has been revealed that the images produced by the DCGAN model can be used in the enlargement of the data set [12].

In this study, a method was applied to eliminate the lack of data sets that are frequently encountered in deep learning studies carried out with medical image data. In this method, it is aimed to prevent loss of labor and time in the verification phase of MR images produced with GAN while eliminating the missing training data in studies based on the processing of medical images. At this point, this study carries out an automatic image selection and parsing process based on the similarities of real and synthetic images with MSE, SSIM and PSNR algorithms, taking into account the type of synthetic images produced. In other words, instead of selecting thousands of synthetic images produced one by one, it is aimed to automatically select those that are similar to the sample images and then consult an expert opinion.

When we look at the numbers in Table 1 based on the results, it is observed that the MSE and SSIM algorithms produce values close to each other, while the images selected by the PSNR algorithm are almost all of the images produced. Here, since MSE and SSIM algorithms are similar algorithms, similar results were obtained as a result of the comparison processes. As a result of the parsing processes performed with the MSE algorithm, a total of 1564 images were selected from 10000 synthetic MR images, and when these images were parsed by the specialist doctor, it was determined that 581 of them could be added to the training data as MR images. Here, the ratio of images selected by the MSE algorithm to verified images was 37.15%. The SSIM algorithm selected 1493 images and 331 of these images were verified by the specialist doctor. The success rate of the SSIM algorithm was calculated as 22.17%. Since the PSNR algorithm makes comparisons by focusing more on the quality of the image, the results here show that it is not correct to use the PSNR algorithm for the purpose of this study. Because when we look at the comparison results obtained with the PSNR algorithm, it is seen that almost 80% of 10000 synthetic images are selected for each real image and that there are many images that do not resemble MR images. In the parsing process performed with the PSNR algorithm, a total of 99900 images were selected in 12 comparisons, and as a result of the verification of these images, 10589 images were selected, resulting in a low verification rate of 10%.

In line with these results, while MSE and SSIM algorithms are preferable algorithms in the stage of parsing medical images by comparing them according to their similarities and determining the data to be added to the training data set, it has been clearly seen that the PSNR algorithm is not suitable for the purpose of the study.

It turns out that the MSE and SSIM algorithms used in the study are preferable image processing algorithms in order to reduce the time lost and labor spent in the post-reproduction stage of the images in different studies where images similar to the biomedical image types preferred in this study are used and in studies where synthetic images or similar images in different data sets are used. has been placed.

Conflict of Interest Statement

The authors declare that there is no conflict of interest.

References

[1] A.G. Eker and N. Duru, "Medikal görüntü işlemede derin öğrenme uygulamaları," *Acta Infologıca*, vol. 5, no. 2, pp. 459-474, 2021. doi:10.26650/acin.927561

[2] U. Kaya, A. Yılmaz and Y. Dikmen, "Sağlık alanında kullanılan derin öğrenme yöntemleri," *Avrupa Bilim ve Teknoloji Dergisi*, vol. 16, pp. 792-808, August 2019. doi:10.31590/ejosat.573248

[3] H. Chen, Z. Qin, Y. Ding, L. Tian and Z. Qin, "Brain tumor segmentation with deep convolutional symmetric neural network," *Neurocomputing*, vol. 392, pp. 305-313, 2020. doi: 10.1016/j.neucom.2019.01.111

[4] G. Çelik and M. F. Talu, "Çekişmeli üretken ağ modellerinin görüntü üretme performanslarının incelenmesi," *BAUN Fen Bil. Enst. Dergisi*, vol. 22, pp. 181-192, 2020. doi: 10.25092/baunfbed.679608

[5] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, Aaron Courville and Yoshua Bengio, "Generative adversarial nets," *Proceedings of the 27th International Conference on Neural Information Processing Systems, 2014*, Available: NeurIPS Proceedings, https://proceedings.neurips.cc/paper_files/paper/2014/hash/5ca3e9b122f61f8f06494c97b1afccf3-Abstract.html. [Accessed: 20 August. 2023].

[6] C. Shorten and T. M. Khoshgoftaar, "A survey on ımage data augmentation for deep learning", *J Big Data*, vol. 6, pp. 60, 2019. doi: 10.1186/s40537-019-0197-0

[7] Y. Liu, J. Zhang, T. Zhao and Z. Wang, "Reconstruction of the meso-scale concrete model using a deep convolutional generative adversarial network (DCGAN)", *Construction and Building Materials*, vol. 370, 2023. doi:10.1016/j.conbuildmat.2023.130704

[8] R. Gonzalez and R. Woods, *Digital Image Processing*, Edition. 3, NJ: Pearson, 2008.

[9] Z. Wang, A. Bovik, H. Sheikh and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612. doi:10.1109/TIP.2003.819861.

[10] M. Raja and S. Vijayachitra, "A hybrid approach to segment and detect brain abnormalities from MRI scan", *Expert Systems with Applications*, vol. 16, pp. 1-9, 2023. doi:10.1016/j.eswa.2022.119435

[11] K. Kazuhiro, R. A. Werner, F. Toriumi, M. S. Javadi, M. G. Pomper, L. B. Solnes, F. Verde, T. Higuchi and S. P. Rowe, "Generative adversarial networks for the creation of realistic artificial brain magnetic resonance images", *Tomography*, vol. 4, no. 4, pp. 159–163, 2018. doi:10.18383/j.tom.2018.00042

[12] C. Han, H. Hayashi, L. Rundo, R. Araki, W. Shimoda, S. Muramatsu, Y. Furukawa, G. Mauri and H. Nakayama, "GAN-based synthetic brain MR image generation," *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA*, 2018, pp. 734-738. doi:10.1109/ISBI.2018.8363678.

_____