Bitlis Eren Üniversitesi Fen Bilimleri Dergisi
BİTLİS EREN UNIVERSITY JOURNAL OF SCIENCE
ISSN: 2147-3129/e-ISSN: 2147-3188
VOLUME: 13 NO: 4 PAGE: 1194-1204 YEAR: 2024
DOI: 10.17798/bitlisfen.1540373

DentaGAN: GAN-Based Synthetic Individual Dental Data Generation in Radiographic Images

Buse Yaren KAZANGİRLER[1*], Caner ÖZCAN[2]

[1] *Karabuk University, Department of Computer Engineering, Karabük, Türkiye*
[2] *Karabuk University, Department of Software Engineering, Karabük, Türkiye*
(ORCID: 0000-0002-8690-2042) (ORCID: 0000-0002-2854-4005)

Keywords: Generative adversarial networks, Dental image generation, Synthetic data augmentation, Two-stage neural network, Panoramic images

Abstract
Panoramic radiographs are a low radiation exposure type often used as a data source for many deep learning algorithms. On the other hand, the operational structure of a traditional deep learning algorithm requires a large amount of data, which is a major problem for many researchers. It is aimed to overcome this problem through deep GAN models, many versions of which have been developed recently. The main purpose of the study is to generate a two-stage GAN model for data with the same image dimensions. The study is carried out in the form of inputting panoramic images containing a whole view, as well as single tooth data whose performance is desired to be measured, to the architecture. The generator model created for each tooth object in all panoramic radiographs generates new tooth objects that the model has yet to encounter in the dataset. Fréchet Inception Distance was used as a performance metric by measuring the distance for the Inception-v3 activation distributions for the real samples in the generated and training set. Thus, the statistical similarity of these two groups obtained from the experimental results was observed in the part of the experimental results. The cropped individual tooth classes were much more successful than the entire panoramic dataset.

1. Introduction

Dental radiographs, widely used in medical imaging, are continuously used to detect tooth loss, tooth material loss, and many health problems that are not detected by visual examination. Various digital radiographs and scan results containing data are essential in medicine, especially in forensic dentistry [1]. In addition, digital imaging has become the most widely used imaging technique in dentistry, the most common field of digital radiography [2]. Although intraoral and extraoral radiographic techniques have advantages or disadvantages relative to each other, both types of imaging contribute negatively to the health of the patient. Radiographs have extremely impressive impacts on radiation, whether a device placed in the mouth or imaging the whole oral region. Therefore, apart from using radiographs to treat the patient during the examination, there should be a limited acquisition process for the datasets intended to contribute to the academic literature [1]. The use of various Artificial Learning (AL) techniques is more appropriate when sufficient data cannot be obtained. If the amount of data in the studies mentioned above is small, it would only be appropriate for the individuals who will use the data to reproduce it manually with an expert. It is also possible to misinterpret the types of radiographs taken in clinics and healthcare institutions depending on the quality of the device and the need for more expert interpretation [2]. It is obvious that the learning-based world goes beyond manual programming and produces solutions to these problems [3]. For this reason, the age of Artificial Intelligence (AI) has rapidly developed in medicine, health, science, and many other important fields [4].

Deep Learning (DL), which has made a name for itself as a sub-branch of AI, makes things much

*Corresponding author: tekinbuseyaren@gmail.com
Received: 29.08.2024, Accepted: 06.12.2024

easier compared to classical machine learning techniques. In the last decade, it has been observed that the frequency of use of DL models has increased significantly, especially in the field of health and medicine, due to their use in the detection of important diseases such as cancer, tumor cells, and pneumonia [5]. The function of using AI in dentistry has been inevitable due to the reasons explained and the development of the technology. In the literature, many qualified scientific studies, both old and new, have taken their place in this field. One of the DL methods, Convolutional Neural Networks (CNNs), is that many conventional studies in dentistry identify teeth by non-manual detection or segmentation [6].

CNNs, which are very effective learning and recognition models, include the building blocks of algorithms used to detect objects in data [7]. Thanks to the use of CNN models, it is possible to detect teeth in radiographic images, but sometimes satisfactory performance is not achieved when the existing images are limited. Augmenting the existing images externally or with an automatic AI model is necessary. Although data augmentation for preprocessing is done in classification studies, it has been observed that the purpose of targeted data duplication has yet to be achieved. The need for more data is a disadvantage, especially for deep learning studies with a small dataset. For this reason, synthetic data generation is provided using the Generative Adversarial Networks (GANs) introduced by Goodfellow et al. [8]. GANs are models that can generate data using two different Artificial Neural Networks (ANNs) that compete in their internal structure. In particular, GAN models have been widely used for synthetic data generation in dental data. Thus, small datasets will be expanded by GAN generation. Therefore, these models have been created to eliminate the data generation problem and generate synthetic data that is very similar to reality [9]. The purpose of generative models is to analyze the given training examples to produce the most similar fake data and, thus, to analyze the probability distributions [8]. It is concluded that the performance of a Machine Learning (ML) based algorithm positively depends on the size, cleanliness, quality, and diversity of the input data [10], [11], [12],[13]. The most significant reason for choosing the model predicted for this study as GAN is the lack of an adequate dataset and the fact that radiographic data harms human health, so X-ray data cannot be obtained again.

The literature has many types of neural networks and different data generation and augmentation theories. Although the performance of the data increased by transformation with CNN models is acceptable, the augmented versions of the cropped and zoomed data do not lead to generating different data. A great deal of the previous research into the small sample problem has focused on the "generative classification" paradigm, proposing a semi-supervised framework [7], [8]. While most studies have used GAN models on intraoral and extraoral radiographs, some studies have used high-quality dental computed tomography data. The study by Hu et al. [12], it was stated that more increased quality results would be obtained because the noise, such as low-dose artifacts and blurring found in dental computed tomography data compared to other panoramic data is less. While it is mentioned in the study that low-dose artifacts and soft tissues are detected with high quality, it should be noted that dental objects in extraoral radiographs will be used to produce high-quality data.

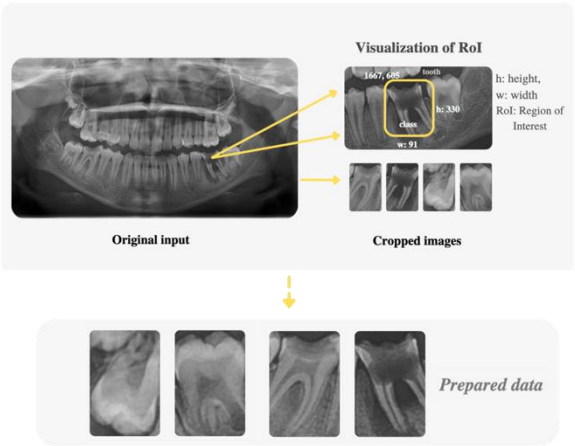

Figure 1. Automatically according to the coordinates of objects in radiographs.

Figure 1 shows that dental objects in a whole panoramic radiographic data are evaluated separately by cropping from their individually labeled bounding boxes. Although GANs can generate new objects by using the features in data, the most remarkable reason for using them individually is because each tooth class exhibits very different features in complex data, such as a molar, premolar, and incisor. The method proposed in the DentaGAN is passed through convolutional filters for conventional deep learning models and generates fake images with a two-stage neural network model. The data generated by the generator network is discriminated by the discriminator and passed through the two-stage structure. The accuracy of the discriminator network in separating spurious samples depends on optimizing the model during its natural training. Therefore, it is ensured that the discriminator neural network is sufficiently trained so that the model can distinguish between real and fake samples. The tooth classes for

the common notation used by dentists were determined and cropped, and each was given to the GAN model as a separate class. The main contribution of our work can be summarized as follows:

- This paper presents a GAN-based approach, which is an innovative algorithm in this field, especially for radiographic images containing individual dental data. This method offers synthetic data as an alternative, especially when real data sets are limited. In addition, this study is the first to implement conventional GAN networks in the field of dental radiography, making it an innovative contribution to the literature. In the literature, GANs have generally been used in medical image processing to expand limited datasets. However, these studies have generally focused on general medical imaging fields such as brain images, eye vessel to eye fundus images in MRI, CT, etc. data.
- The dataset used in the study was evaluated as a total of three separate studies, including all uncropped panoramic images, cropped tooth data from panoramic images, and a separate class according to Dental Federation Notation (FDI) each of which was evaluated.
- The proposed model has effectively increased the diversity and quality of dental datasets. This strategy is crucial for improving the accuracy of dental image analysis models.
- The study showed that traditional GANs are more challenging for generating dental objects than other object classes.
- Since the new radiographic data is close to real-world data and can be used in different neural networks, all caries, damaged, or healthy dental objects are included in the training set.

2. Material and Method

2.1. Generative Adversarial Networks

GANs provide an AL method without the need for annotation files and pieces of information developed as an alternative to learning image spaces. Today, the unique data produced by GANs are used in many different applications, such as image classification, object detection, style transfer, and semantic processing. Thanks to the two different neural network structures it contains, it ensures that the image data produced is statistically indistinguishable from the images in the training set and is new and similar [13]. GANs are also known as generative algorithms, which fall under both generative and discriminative algorithms in ML. For this reason, it has become prevalent nowadays [14]. In the GAN architecture concept, random latent vectors (noise) are passed through the generator to synthesize new data from scratch. Note that there is no direct image input, as the generator only uses the latent vector structure to generate images. Furthermore, since the new images generated from this structure are noise-based, the images in the dataset are only used to check the discriminator for real and fake data. During training, the discriminator is taught to accurately discriminate between real and generated images, ensuring that the generator gets better at generating more realistic images. These two networks are trained in tandem, with the generator continuously adapting to fool the discriminator as the discriminator gets increasingly better at identifying synthetic data.

2.2. Image Generation

Image generation means generating data similar to the input data by adding noise. As shown in Figure 2, it consists of five block layers. The semantic feature vector extracted from the input data is generated in the last dense layer. When the architecture of the generator is examined, the tanh activation unit is used to normalize the last layer from [-1, +1]. Then, the generated vector is given to the next block, the discriminator network [15].

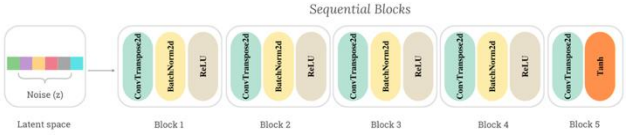

Figure 2. Generator module architecture responsible for image generation for GANs.

In the generator module, the images are passed through certain pre-processing steps to be given to the neural network in the same format. Images are given to the model architecture as “Single Tooth” classes. This architecture input only points to an example model. Otherwise, the model is given per class such as “13”, “48”, or the entire panoramic data.

2.3. Image Discrimination

In the discriminator module, the similarity of the data to the original data is calculated by a discriminator as a result of the combination of real and fake data. In this case, while the generator neural network generates the data in the decoder structure, the generated data is in the decoded position while being transmitted to the discriminator neural network. The crucial task is the discriminator, which tries to reach

the maximum likelihood of predicting real and fake data [16]. As seen in Figure 3, it consists of five dense layers. The sigmoid function was found suitable for the activation function, which will be used in the probability estimation in the last layer. In addition, the reason for using Leaky ReLU activation instead of Rectified Linear Unit (ReLU) is due to the faster convergence of the discriminator network. Thus, we have taken precautions to avoid overfitting the model.

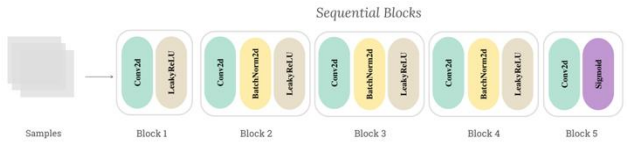

Figure 3. The architecture of the discriminator module responsible for the fidelity of images generated for GANs.

The loss value needs to be backpropagated for the weights updated in the predictions made by the discriminator. Hence, the layers are frozen, and the backpropagation is done with the generator in the weight update. The main method of training the network in GANs is to select the appropriate generator to enable the discriminator to perform classification with maximum accuracy [17]. During the training of the generator network, the optimization phase of the value function in [1] is included [18].

$$V(D,G) = E_{x\sim px}[logD(x)] + Ez\sim pz\left[\log\left(1D\left(G(z)\right)\right)\right] \quad (1)$$

where px and pz indicate original data (x) and generated data distributions (z), respectively [17]. The generator is represented by mapping from the noise space to the data space. $D(x)$ indicates the likelihood that input x comes from data rather than generator G. At the same time, G is trained to minimize $[log(1 - D(G(z)))]$.

2.4. Dental dataset and preparation

The dental dataset used in the study, 565 panoramic radiographs, was annotated by dentists to determine the bounding boxes of dental objects. DentiAssist [19] was used as an object annotation software specifically designed for dental radiographic data. Panoramic radiographs for this study were initially $G(x)$ to be generated as a "Single Tooth" class. Then, considering that the tooth structures in panoramic radiographs are very different, they were cropped by looking at the spatial information in the label files.

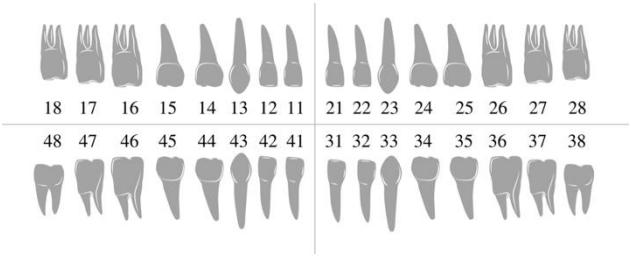

Figure 4. Teeth numbering notation in data labeling and preparation step.

In the FDI notation shown in Figure 4, the superset digits are ordered from left to right, and the subset digits are from right to left. The first digits are named (1 - 2 - 3 - 4) and consist of four separate digits. The reason for using this notation is that it is a universal system of notation in the world. These 32 classes, also labeled in the panoramic radiographs, were cropped from their bounding boxes, as shown in Figure 1. In this way, it also provided the individual evaluation of the teeth. The cropped classes were prepared as "Single Tooth" to be given as input to the generator in the GAN model. In addition, the 32-class tooth categories in the GAN model were considered separately in a complete panoramic class and the single tooth class. Figure 5 shows the data numbers of individual dental datasets in 32 classes. In Figure 5, the names of the individual tooth classes are indicated by the number (11). The percentages written below the numbers represent their contribution to the data set. The difference in the data numbers in the figure is that the molar teeth with the second digit of eight (impacted) are very likely to be extracted in individuals. From this conclusion, the clusters of teeth belonging to the 18-28-38-48 classes are expected to be slightly lower than the discriminator networks. Apart from these situations, the small number of data compared to other datasets is due to the dental structures of the patients who underwent radiographs.

This study uses test data for evaluation after the training phase. A specifically allocated validation cycle needs to be allocated during the training process. It means that the test set is only used for post-training evaluation. Considering that the generator and discriminator performance is continuously evaluated against each other, no direct validation and test sets are defined in the study. On average, 10% of each dataset is reserved for testing and 90% for training. For example, 509 train data were allocated for the training of the panoramic dataset, while 56 were selected for testing. For instance, for a single tooth, this was 13,497 train data and 1,499 test data, while for an example, individual class (class 47) was 421 train data and 46 test data.

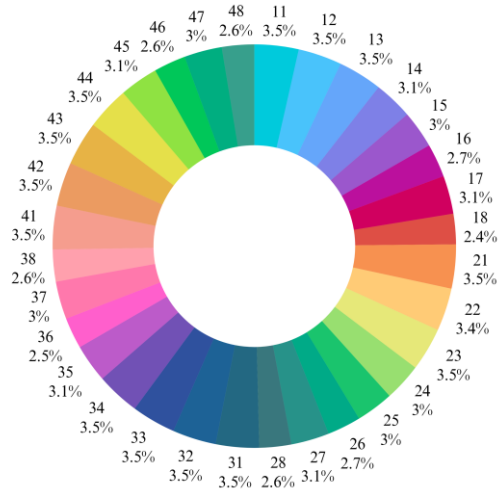

Figure 5. Data distribution statistics for different dataset groups.

During image preprocessing in the generator, batch size 16 was selected, and the image size was 64x64. The scaling, size reduction, and normalization processes are completed before the input data is given to the generator architecture. Then, it passed the relevant data loader and augmentation stage and duplicated the data with certain pre-processing steps.

2.5. Dental dataset and preparation

The architecture of the two-stage neural network structure in dental radiographs is shown in Figure 6. As the first step in the architecture, all panoramic radiographs were given input, while simultaneously cropped panoramic images in 224x224 dimensions were used as input. The original input images are passed through the appropriate pre-processing steps, and the data generated by the generator through latent space and added noises are given to the discriminator. The generated fake and real images of the data set are controlled by a discriminator using the binary cross-entropy loss function. During the training procedure, the output unit should be a number between 0-1 when calculating the loss value for the generator and the discriminator. For this reason, we will calculate the loss value with the binary cross-entropy calculation [20]. The purpose of the function at this stage is to determine whether the generated image is real or fake. The binary cross-entropy loss minimizes the mean probability error between the target and the estimation label for all pixel values found [19].

As seen in the [2], $log(y\hat{}'x)$ for $logD(x\hat{}i)$ wants likelihood close to 1 by capturing satisfactory predictions in real images with gradient ascent. The y in the equation represents the original images, while y' tries to be 1 in this case. Also, for $1 - D(G(z\hat{}i))$, the fake images must be well-estimated, so a probability close to 0 is intended [17].

$$L(w) = -y^i \log\left(y'^{(i)}\right)(1 - y^{(i)})\log\,(1 - y'^{(i)}) \qquad (2)$$

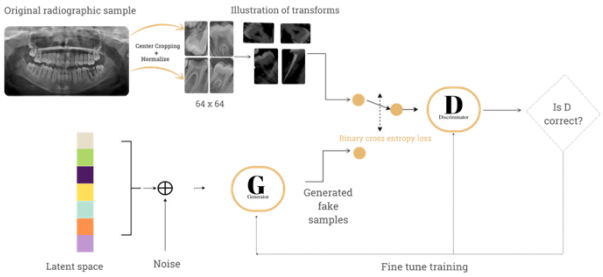

Figure 6. Two-stage generation and discrimination architecture of dental images [8].

Gradient descent in neural networks descends on a static loss surface. In GANs, on the other hand, each downhill step changes the entire surface by a certain amount. The training of the model takes place under difficult conditions, as it is a dynamic system that searches for the experiment between two forces instead of searching for minimum values [17]. Therefore, it was necessary to carefully consider the parameters of the GAN model to be created. In the selected parameters, different learning rates were chosen as the size of the hidden vector z that is, 64 as the size of the generator input, 128 as the size of the feature maps in the builder, ndf: 64 as the size of the feature maps in the discriminator, glr: 0.001 (learning rate of the generator) and dlr: 0.0005 (learning rate of the discriminator) for the optimizers. The same parameters were chosen for comparison in the appropriate epoch training of all the datasets mentioned in the study. One of the hyperparameters that should be used for Adam optimizers, ($beta_1$) was determined as 0.5. In addition, the real label value is 0.9, and the fake label is 0. Following the determined parameters, the original images in the dataset are trained with the discriminator, and the weights are updated during the training. In the figure, starting from the first layer 512 of the created generator neural network, the 2-dimensional transpose convolution layer is used to the last layer 32. Figure 7 is an example of the generator and discriminator pipeline on whole panoramic images. Other datasets are also given as inputs to the generator and discriminator networks. The three output channels from the ConvTranspose2d unit of the generator network are given to the discriminator network in the next section. Then it is transmitted to the 64, 128, 256, and 512 layers, respectively, as shown in the figure. Both weight updates and feature extractions take place in this step. Pixel distance and feature distance are compared in images for diversity and fidelity. Therefore, evaluating GANs is a challenging task [21], [22].

Algorithm 1. Pseudocode of the training procedure for the proposed approach

Input: *maxepoch, dataloader,generator,discriminator, classifier,cof ,adw*
for *epoch in range(maxepoch)* **do**
for *(realdata, reallabel, embedding)* in *dataloader* **do**
batchsize = len(realdata)
#step 1: generate fake images and train the discriminator with real data
genimg=generator(random((batch, 128, 1, 1)), embedding)
predreal=discriminator(realdata), BCE(predreal, ones((batchsize, 1))).backward()
#step 2: train the discriminator with generated data
predfake=discriminator(genimg), BCE(predfake, zeros((batchsize, 1))).backward()
#step 3: train the generator and opt. discriminator
predgen=discriminator(genimg), BCE(pred−gen, zeros((batchsize, 1))).backward()
#step 4: train the classifier with real data
outreal=classifier(realdata), classification−loss(outreal, reallabel)
#step 5: train the classifier and data generated in the sample control
fakeprob=classifier(genimg)
outlabel=argmax(fakeprob, dim = 1)
if *softmax(fake)[reallabel]* $\geq$ *cof* **then**
loss'=adw $*$ *classificationloss(selectout, selectlabel)*
loss'.backward()
end
end
end
Output: Generated synthetic data (radiographic image)
Input: *maxepoch, dataloader,generator,discriminator, classifier,cof ,adw*
for *epoch in range(maxepoch)* **do**
for *(realdata, reallabel, embedding)* in *dataloader* **do**

The algorithm of the training procedure for the GAN model is as follows. The input image and the label are generated for each epoch, creating fake images in the specified batch size. The data loader parameter is central to loading the training data. It provides iterability over the dataset with the shuffle parameter: "True" option. When n epochs are performed for Algorithm 1 the images created by the generator are kept. Then, the discriminator model takes the real data and provides a backward prediction for binary cross-entropy loss. The total error is propagated back and updated. The generator and the discriminator are trained for original and fake samples that provide sequences of zeros and ones, the result to be detected in the output [22]. Thus, predicted data is generated. The classifier in steps 5 and 6 of Algorithm 1 represents a general neural network. Therefore, the classifier refers to both the generator and the discriminator networks. Also, a classical GAN structure has only a generator and a discriminator. However, this algorithm has an additional classifier. The task of the classifier is to support the learning process of the model by classifying real and generated data. In other words, the generator generates fake images using random input and embedding, the discriminator is the network that tries to distinguish whether these images are real or fake, and the classifier is the structure that helps the model to better learn the difference between real and fake data. As a result, the classifier acts as an additional control mechanism for the GAN. A classification process is performed between the input images and the new images generated, and they are checked with softmax condition. Accordingly, the loss value is updated. Because GAN models mostly consist of two models that can be tuned to each other, this training demonstrates the struggle between the two neural networks [23]. In summary, we maximize D(discriminator) and $log(D(x)) + log(1 - D(G(z)))$ by maximizing $log(D(G(z)))$ G(generator) is updated. Evaluation is an open area in generator model research. The FID metric in [3] is the most popular for evaluating the success of the data generated by the generator. The Inception-v3 network

measures the performance of the fake images generated. The purpose of using this initial network is to extract features from the intermediate. The data is distributed using a multivariate Gaussian distribution with mean and covariance for statistical purposes [16], [23].

$$FID = |\mu_1 - \mu_2| + T_r(\sigma_1 + \sigma_2 - 2\sqrt[2]{\sigma_1 * \sigma_2}) \quad (3)$$

where Tr sums all the diagonal elements, $\mu 1$ and $\sigma 1$ are the mean and covariance of the training data, and $\mu 2$ and $\sigma 2$ are the mean and covariance of the test data [16]. It calculates the distance of the curves between the fake and the real embeddings. The closer these statistics are to each other, the closer the fake embeddings model is to the real embeddings. Therefore, a smaller value means that features in the reals and fakes are more similar, so the lower the FID, the better and closer distributions.

3. Experimental Results and Discussion

To obtain the experimental results, a device with NVIDIA Geforce RTX 3050 graphics card support was used in the GAN model training. GANs with the same parameters and hyperparameters were generated for 34 dental datasets mentioned in the paper. The experimental findings obtained under the specified conditions are given in Table 1. While training with the generator continues, the high threshold value set by the discriminator ensures that the generator obtains a thriving fake image. When the table is examined, it is revealed that different datasets are used. Taking into account the possible errors that can occur when generating panoramic images, one of the main objectives of the study was to properly trim the data from the bounding boxes. One of the most significant reasons for this is that although GANs generate successful images, radiographic data is more complex than simple data has a disadvantage. As a first step in the study, panoramic data were generated, and due to the high FID value, even if changed the parameter, it was necessary to find another solution. Thus, the teeth were cropped according to the FDI notation, and all the crops were used as a given dataset. In this study, 500 epoch training sessions were performed on dental radiographs and single tooth images to ensure a fair approach. The average training time of the generative and discriminative networks simultaneously per dataset is 7466 seconds, about 124 minutes for the individual class (41 classes). The single tooth class of combined teeth trained in the study: 14,996 images, and the panoramic dataset contains 565 images. Also, the number of individually cropped tooth images varies depending on the panoramic images, as shown in Figure 5. This situation is mainly since sometimes cases such as impacted teeth and molars are not included depending on the age of the extracted individual.

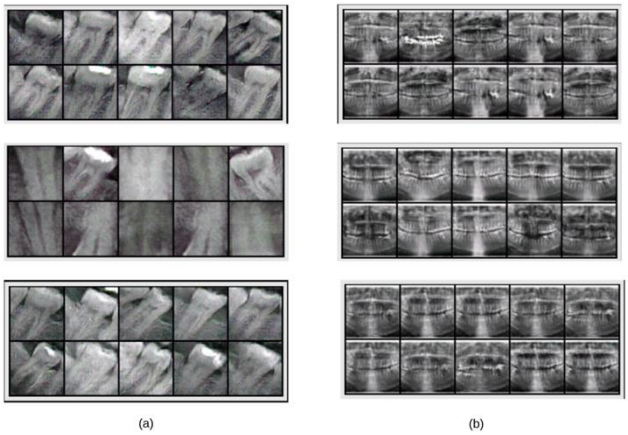

Figure 7. Fake images generated because of training GANs. (a) the results of cropped images, (b) the result of an entire panoramic image.

Nevertheless, the total amount of data generated for all classes reached 10,000. Considering the results in Table 1, the single tooth class has a large amount of data, which leads to a lower FID result. The training time for the single tooth class was about 3,420 minutes, or 57 hours, because it contains a lot of data. In addition, other individual tooth classes produced different results depending on their number and complexity in the dataset. The panoramic dataset contains 565 images, but it also contains a wider range of mouth views and patterns, which increases the time for the neural networks to scan the data. Therefore, producing the entire panoramic image takes 9,300 seconds, or 155 minutes. During the training of this data, the loss of the generator reached 2.9414, while the loss of the discriminator reached 1.2484.

The fake images generated by neural network are shown in Figure 7, and according to this figure, it was observed that the individual tooth classes were generated by GAN much better than all the panoramic images. However, more was needed to restrict all images generated to one class due to the differences in the mouth's molar, premolar, and incisor classes. For these reasons, one-sided label softening was performed while adjusting the parameter. For these reasons, one-sided label softening was performed when adjusting the parameter. An excessively high confidence value can cause some problems with deep neural networks. If the discriminator depends on the small feature map while recognizing the actual images, the generator generates these features to

benefit from the discriminator. In this case, the optimization performed can be very greedy. To avoid such problems in real image estimation, the discriminator is penalized when it exceeds $0.9(D(real > 0.9))$. The target label value has been updated to 0.9 instead of 1.0.

Since FID measures the distributional similarity between the generated and real images, the test set provides a realistic example of this distribution. The results of the test data are shown in Table 1 below. In contrast, the discriminator loss value is excessively high, causing overfitting due to the imbalance between the generator and the discriminator. Overfitting is recognized as one of the major challenges in training GANs. Suffering from such difficulties, the experimental results in this study may not have achieved the expected performance in some individual tooth classes. However, the overall panoramic view was worse in all conditions than in other individual tooth classes. On the other hand, it is expected that the single tooth class created with samples from single tooth classes will have high performance due to the high diversity of data.

Table 1. Findings of neural networks generated and discriminated for different dataset.

Generated class	Loss G	Loss D	Fréchet Inception Distance	Generated class	Loss G	Loss D	Fréchet Inception Distance
Single tooth	2.9414	1.2484	124.8310	Tooth 31	3.4235	0.5093	102.8283
Tooth 11	4.7493	0.4203	187.8658	Tooth 32	3.8033	0.4238	114.4682
Tooth 12	5.0410	0.3903	178.6688	Tooth 33	4.2498	0.6059	144.1210
Tooth 13	3.3994	0.4371	127.4457	Tooth 34	4.0968	0.3835	153.1053
Tooth 14	4.3447	0.3905	146.7246	Tooth 35	4.0497	0.4592	158.1447
Tooth 15	4.6472	0.4127	173.7276	Tooth 36	4.1037	0.5133	213.4831
Tooth 16	3.8004	1.1397	160.5275	Tooth 37	3.1822	0.4515	127.4164
Tooth 17	6.7436	0.6928	297.8457	Tooth 38	4.0272	0.4591	175.5649
Tooth 18	3.8478	0.4241	138.1094	Tooth 41	4.9409	0.3572	151.7190
Tooth 21	3.8511	0.4163	169.2506	Tooth 42	3.1018	0.7503	190.0691
Tooth 22	3.8585	0.3945	181.6382	Tooth 43	4.1703	0.6607	131.0531
Tooth 23	4.2090	0.7617	126.5523	Tooth 44	5.3700	0.3606	150.3511
Tooth 24	5.1406	0.3366	179.1189	Tooth 45	3.1822	0.4515	127.4164
Tooth 25	5.2565	1.4472	215.2094	Tooth 46	4.8492	0.9850	171.0522
Tooth 26	3.0496	1.7622	163.5350	Tooth 47	4.4303	0.4255	257.3137
Tooth 27	4.4674	0.6183	215.6518	Tooth 48	4.6258	0.3531	226.9374
Tooth 28	4.8688	0.7251	135.4039	Panoramic	3.9198	0.4527	223.4762

On the other hand, our study adapts existing GAN algorithms to dental radiography, enabling the synthetic generation of individual tooth images. However, datasets with more complex and individual characteristics, such as dental structures, are being studied. This requires more detailed modeling than general medical imaging studies. As a result, in terms of the applicability of GANs to radiographic images and the solution to data limitations in dentistry, this study makes an important contribution to the literature. In particular, it is necessary to accurately generate anatomical details and anomalies that may be present in radiographic data, such as periodontal structures, root canals, or differences due to anomalies such as caries and restorations. However, the current architecture of conventional GANs is insufficient to maintain such a precise level of detail.

4. Conclusion and Suggestions

This paper implements a novel method for generating and augmenting synthetic individual dental data in radiographic images using GANs. The results show that the recommended strategy is effective in increasing the diversity and quality of dental datasets, both of which are necessary to improve the accuracy of dental image analysis models. Synthetic data alleviates the problems caused by the lack of real-world datasets and provides a reliable alternative for building robust DL models in dental radiography. Results have shown that the AI generated by the DentaGAN model produces synthetic data that is as high quality as actual dental radiographs, with the added benefit of being adaptable to a range of clinical settings. This capability has significant potential in the field of dental informatics as it provides larger and more

representative training datasets, leading to more robust models and, therefore, better patient outcomes. However, when the FID findings for each class were analyzed, it was observed that evaluating alternative GAN methods in future research may be appropriate. The findings suggest that the performance can be improved in the future by experimenting with different GANs with larger-sized images.

In future studies, although the individual tooth classes are balanced in the dataset, the data of the tooth classes that are difficult to learn will be increased. As a result of experimental studies and findings, we have concluded that generating dental objects with the traditional GAN model is more challenging than other object classes. The results of the study indicate that conventional GANs are more difficult to generate dental objects than other object classes. This finding suggests that experimenting with different GANs on larger images may improve performance in the future. Therefore, in the next stage of the study, it will be possible to increase the possible performance by working with different types of GAN models with large image sizes.

Contributions of the authors

B. Y. Kazangirler: Methodology, visualization, wrote the paper, performed the experiments, interpretation of the data
C. Ozcan: Conceived and designed the analysis, investigation, wrote the paper, revised the manuscript

Conflict of Interest Statement

There is no conflict of interest between the authors.

Statement of Research and Publication Ethics

The study is complied with research and publication ethics.

References

[1] J. Park *et al.*, "Deep learning on time series laboratory test results from electronic health records for early detection of pancreatic cancer," *J. Biomed. Inform.*, vol. 131, p. 104095, Jul. 2022, doi: 10.1016/j.jbi.2022.104095.

[2] S. Vinayahalingam *et al.*, "Automated chart filing on panoramic radiographs using deep learning," *J. Dent.*, vol. 115, p. 103864, Dec. 2021, doi: 10.1016/j.jdent.2021.103864.

[3] M. J. Cardoso, N. Houssami, G. Pozzi, and B. Séroussi, "Artificial intelligence (AI) in breast cancer care - Leveraging multidisciplinary skills to improve care," *Breast Off. J. Eur. Soc. Mastology*, vol. 56, pp. 110–113, Dec. 2020, doi: 10.1016/j.breast.2020.11.012.

[4] B. Y. Tekin, C. Ozcan, A. Pekince, and Y. Yasa, "An enhanced tooth segmentation and numbering according to FDI notation in bitewing radiographs," *Comput. Biol. Med.*, vol. 146, p. 105547, 2022.

[5] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017, doi: 10.1016/j.media.2017.07.005.

[6] D. Frejlichowski and R. Wanat, "Application of the Laplacian Pyramid Decomposition to the Enhancement of Digital Dental Radiographic Images for the Automatic Person Identification," in *Image Analysis and Recognition*, A. Campilho and M. Kamel, Eds., Berlin, Heidelberg: Springer, 2010, pp. 151–160. doi: 10.1007/978-3-642-13775-4_16.

[7] Y. Lin *et al.*, "DHI-GAN: Improving Dental-Based Human Identification Using Generative Adversarial Networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 12, pp. 9700–9712, Dec. 2023, doi: 10.1109/TNNLS.2022.3159781.

[8] S. Tian *et al.*, "DCPR-GAN: Dental Crown Prosthesis Restoration Using Two-Stage Generative Adversarial Networks," *IEEE J. Biomed. Health Inform.*, vol. 26, no. 1, pp. 151–160, Jan. 2022, doi: 10.1109/JBHI.2021.3119394.

[9] R. Havale, B. S. Sheetal, R. Patil, R. Hemant Kumar, R. T. Anegundi, and K. R. Inushekar, "Dental notation for primary teeth: a review and suggestion of a novel system," *Eur. J. Paediatr. Dent.*, vol. 16, no. 2, pp. 163–166, Jun. 2015.

[10] N. K. Singh and K. Raza, "Medical Image Generation Using Generative Adversarial Networks: A Review," in *Health Informatics: A Computational Perspective in Healthcare*, R. Patgiri, A. Biswas, and P. Roy, Eds., Singapore: Springer, 2021, pp. 77–96. doi: 10.1007/978-981-15-9735-0_5.

[11] "U-Patch GAN: A Medical Image Fusion Method Based on GAN | Journal of Imaging Informatics in Medicine." Accessed: Nov. 29, 2024. [Online]. Available: https://link.springer.com/article/10.1007/s10278-022-00696-7

[12] Z. Hu *et al.*, "Artifact correction in low-dose dental CT imaging using Wasserstein generative adversarial networks," *Med. Phys.*, vol. 46, no. 4, pp. 1686–1696, Apr. 2019, doi: 10.1002/mp.13415.

[13] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath, "Generative Adversarial Networks: An Overview," *IEEE Signal Process. Mag.*, vol. 35, no. 1, pp. 53–65, Jan. 2018, doi: 10.1109/MSP.2017.2765202.

[14] J. Gui, Z. Sun, Y. Wen, D. Tao, and J. Ye, "A Review on Generative Adversarial Networks: Algorithms, Theory, and Applications," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 4, pp. 3313–3332, Apr. 2023, doi: 10.1109/TKDE.2021.3130191.

[15] A. Borji, "Pros and cons of GAN evaluation measures: New developments," *Comput. Vis. Image Underst.*, vol. 215, p. 103329, 2022.

[16] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen, "Improved techniques for training gans," *Adv. Neural Inf. Process. Syst.*, vol. 29, 2016, Accessed: Aug. 24, 2024. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2016/hash/8a3363abe792db2d8761d6403605aeb7-Abstract.html

[17] A. Radford, L. Metz, and S. Chintala, "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks," Jan. 07, 2016, *arXiv*: arXiv:1511.06434. doi: 10.48550/arXiv.1511.06434.

[18] A. Figueira and B. Vaz, "Survey on synthetic data generation, evaluation methods and GANs," *Mathematics*, vol. 10, no. 15, p. 2733, 2022.

[19] A. Karaoglu, C. Ozcan, A. Pekince, Y. Yasa, B. Tekin, and D. Ozdemir, "Automatic dental segmentation module supported by artificial intelligence for dentistry students education," *Artif Intell Theory Appl*, vol. 1, pp. 180–190, 2021.

[20] U. Ruby and V. Yendapalli, "Binary cross entropy with deep learning technique for image classification," *Int J Adv Trends Comput Sci Eng*, vol. 9, no. 10, 2020, Accessed: Aug. 24, 2024. [Online]. Available: https://www.researchgate.net/profile/Vamsidhar-Yendapalli/publication/344854379_Binary_cross_entropy_with_deep_learning_technique_for_Image_classification/links/5f93eed692851c14bce1ac68/Binary-cross-entropy-with-deep-learning-technique-for-Image-classification.pdf

[21] N. Shibuya, "Understanding generative adversarial networks," *Retrieved Internet*, 2017.

[22] M. Lucic, K. Kurach, M. Michalski, S. Gelly, and O. Bousquet, "Are gans created equal? a large-scale study," *Adv. Neural Inf. Process. Syst.*, vol. 31, 2018, Accessed: Aug. 24, 2024. [Online]. Available: https://proceedings.neurips.cc/paper/2018/hash/e46de7e1bcaaced9a54f1e9d0d2f800d-Abstract.html

[23] J. Liu, C. Gu, J. Wang, G. Youn, and J.-U. Kim, "Multi-scale multi-class conditional generative adversarial network for handwritten character generation," *J. Supercomput.*, vol. 75, no. 4, pp. 1922–1940, Apr. 2019, doi: 10.1007/s11227-017-2218-0.