

## Article

# A Novel Method Based on GAN Using a Segmentation Module for Oligodendroglioma Pathological Image Generation

Juwon Kweon <sup>1</sup>, Jisang Yoo <sup>1</sup>, Seungjong Kim <sup>2</sup>, Jaesik Won <sup>2</sup> and Soonchul Kwon <sup>3,\*</sup> 

<sup>1</sup> Department of Electronic Engineering, Kwangwoon University, Seoul 01897, Korea; 02kju0203@kw.ac.kr (J.K.); jsyoo@kw.ac.kr (J.Y.)

<sup>2</sup> Molpaxbio, Daejeon 34047, Korea; jong.kim@molpax.com (S.K.); jaesik.won@molpax.com (J.W.)

<sup>3</sup> Graduate School of Smart Convergence, Kwangwoon University, Seoul 01897, Korea

\* Correspondence: ksc0226@kw.ac.kr; Tel.: +82-2-940-8637

**Abstract:** Digital pathology analysis using deep learning has been the subject of several studies. As with other medical data, pathological data are not easily obtained. Because deep learning-based image analysis requires large amounts of data, augmentation techniques are used to increase the size of pathological datasets. This study proposes a novel method for synthesizing brain tumor pathology data using a generative model. For image synthesis, we used embedding features extracted from a segmentation module in a general generative model. We also introduce a simple solution for training a segmentation model in an environment in which the masked label of the training dataset is not supplied. As a result of this experiment, the proposed method did not make great progress in quantitative metrics but showed improved results in the confusion rate of more than 70 subjects and the quality of the visual output.

**Keywords:** generative adversarial networks; pathology image synthesis; digital pathology



**Citation:** Kweon, J.; Yoo, J.; Kim, S.; Won, J.; Kwon, S. A Novel Method Based on GAN Using a Segmentation Module for Oligodendroglioma Pathological Image Generation. *Sensors* **2022**, *22*, 3960. <https://doi.org/10.3390/s22103960>

Academic Editor: Adrian Barbu

Received: 24 February 2022

Accepted: 20 May 2022

Published: 23 May 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Pathology is a branch of medicine in which the morphological changes that occur in tissue, organs, and cells are microscopically observed to investigate the causes of disease. It provides important information for making final diagnostic decisions. Latterly, pathology has used electron microscopes to magnify brain tumor, muscle, or other tissue by thousands to tens of thousands of times to observe minute morphological changes. This digitization is called digital pathology, and accordingly, digital pathology data analysis research based on deep learning is being studied using a large amount of data. In recent years, in the field of pathology, deep learning has been used to analyze and predict data and has achieved state-of-the-art performance. Through various deep learning-based algorithms, digital pathology mainly designs classifiers for diagnosing pathologist-level diseases and to assist in human decision-making processes. It has also been used for detecting or segmenting specific parts, such as nuclei, cells, and blood vessels, in slide images, normalizing H&E stained images and predicting patients' life expectancies.

For training, deep learning networks generally require tens of thousands to millions of samples in the datasets. However, collecting data from medical imaging remains difficult, and the data imbalances between classes are a problem. Data augmentation techniques and methods of synthesizing data have been actively researched to solve this problem. General data augmentation techniques, such as x-flips, y-flips, 90° rotations, color transforms, brightness control, noise addition, etc., have been used to create data to improve and generalize the performance of deep learning networks. In addition to the common data augmentation techniques, machine learning algorithms are also widely used. Representatively, there is the Synthetic Minority Oversampling TEchnique (SMOTE) [1] that uses the k-nearest neighbors (K-NN) algorithm. He et al. and Han et al. oversampled the data using both the adaptive synthetic sampling (ADASYN) [2], and Borderline-SMOTE

algorithms are based on SMOTE [3]. Deep learning algorithms also demonstrated good performance in data synthesis. Generative adversarial networks (GANs) [4] have actively studied a representative generative model for data synthesis and have produced several unanticipated applications.

In this paper, we present a method to generate oligodendroglioma (ODG) pathology images from latent space using the GAN for generating realistic data. Unlike many previous GAN algorithms, we propose the following: a method of applying a simple feature concatenation that assists in training the generator. To apply the feature concatenation technique requires a segmentation module consisting of symmetrically designed segmentation and reconstruction networks based on the U-net architecture [5]. The segmentation module consists of the segmentation network that segments important features in cellular tissue, such as nuclei and glands, and a reconstruction network that helps improve the segmentation accuracy by reconstructing the original image from the masked image, which is the result of segmentation. Similar to most segmentation tasks, a training dataset consisting of pairs of the original and labeled data are often used for supervised learning. The Cancer Genome Atlas (TCGA) LGG dataset we used does not have labeled data; therefore, for training, we provided a guideline dataset using k-means clustering [6] to our segmentation model and changed from unsupervised training to a simple method. The generator produces realistic ODG images from the random latent vector  $z$  by utilizing the features that are used in the reconstruction stage of the segmentation module. Our simple method of guiding the generator by concatenating the embedding feature maps of the reconstruction network created better quality images for the TCGA data than the existing method.

Works related to this study are described in Section 2. We propose our method in Section 3 and describe the qualitative and quantitative experimentation results in Section 4. Finally, we include the discussion and conclusion in Sections 5 and 6, respectively.

## 2. Related Work

### 2.1. Oligodendroglioma

Oligodendroglioma (ODG) are rare tumors that infiltrate a wide range but mainly occur in adult cerebral hemispheres' white matter, and they consist of cells resembling oligodendrocytes or oligodendroglia [7,8]. Histologically, ODG is characterized by round nuclei with a rounded cytoplasmic surround. Because of this perinuclear halo characteristic, it is often referred to as the "fried egg appearance". In addition, "uniformly round to oval-shaped cells", "uniform nuclei", the "chicken-wire capillary network", and "calcification" are frequently found characteristics [7,8].

According to statistics from the Central Brain Tumor Registry of the United States (CBTRUS) from 2011 to 2015, the incidence of ODG, including anaplastic oligodendroglioma, was 3.57 to per 100,000 [9]. Kleihoues reports that ODG is a brain tumor that accounts for about 5–18% of all intracranial gliomas [7]. Suh et al. reported that it accounted for approximately 3.4% of the total central nervous system gliomas [10].

### 2.2. Histology Image Synthesis

By training deep learning networks, data augmentation techniques improve the performance of deep learning models by increasing the number of training samples. Various techniques are used when the training data are insufficient. This is essential for medical data that are severely imbalanced in each class. Using methods such as x-flips, y-flips, 90° rotations, color transforms, brightness control, and noise addition increases the generalization performance of deep learning models by creating a large quantity of relevant data. Data augmentation methods may yield desired performance depending on the task, but sometimes, it degrades the network's performance. For example, for a human face dataset, such as the CelebA [11] and FFHQ [12] datasets, if color transformations are strongly applied, samples that are highly unlikely to exist can be added as the training data. Therefore, rather than simply sampling meaningless data, meaningful augmentation techniques must be applied such as increasing the number of data points in a minority layer that has

insufficient data. The SMOTE finds the nearest neighbor within the minority class and multiplies the difference between samples using a random value of between 0 and 1 to create new data samples. Consequently, it matches the quantity of data between classes [1]. He et al. proposed ADASYN, which improves SMOTE. This work applies the concept of assigning weights according to the density of the majority class when selecting a small-class sample [2]. In addition, Han et al. proposed Borderline-SMOTE, which synthesizes new data samples by applying SMOTE based on the boundary between the minority class sample set and the majority class sample set [3].

### 2.3. Segmentation

Several studies have been conducted on pixel-accurate segmentation for digital pathology. In particular, it is common to use U-Net [5] in medical imaging tasks, such as data segmentation, reconstruction, and pre-processing. U-Net is a U-shaped architecture-based model in a fully convolutional network [13]. This model can execute end-to-end learning using small datasets, resulting in efficient segmentation for biomedical imaging. Two paths, one contracting and the other expanding, function as the encoder and decoder, respectively. The two paths are connected via skip connection. This structure ensures successful segmentation results by preserving local information. In digital pathology, previous studies have used the U-Net structure to segment various cancers, such as in breast [14–18], colon [19–23], lung [24,25], and prostate tissue [26,27]. Furthermore, the structure is also applied to localized tissues such as the nuclei [24,28–32], cells [33,34] and glands [19–23], which exhibit major pathological characteristics. Currently, from the structural perspective, some relevant studies have used network-modified U-Net structures, such as Dual-U-Net [32], U2-Net [35] and R2U-Net [31]. We followed the fundamental structure of previous segmentation studies and used them to stabilize the training and contribute to improving performance by combining it with GANs.

### 2.4. GANs in Digital Pathology

Generative adversarial networks (GANs) introduced by Goodfellow are deep learning-based generative models. A GAN consists of a generator (G) that maps random noise  $z$  to the target data distribution and a discriminator (D) that estimates the probability that the data came from the target data distribution or G. The two networks increase each other's performance via the adversarial training process. GANs face critical problems during training. First, they are unstable during training and several studies have focused on solving this problem. Representatively, the DCGAN [36] suggests guidelines for stable training, and this structure is very effective. The WGAN mathematically proved that if the 1-Lipschitz constraint is satisfied, satisfactory convergence can be achieved using the Earth Mover Distance (EMD) and Wasserstein distance (WD) [37]. The WGAN-GP's training stability and mode collapse problem is resolved by applying a gradient penalty to the WGAN [38]. There are studies, such as LSGAN [39], hinge loss-based GAN [40], unrolled GAN [41], MDGAN [42], and SN-GANs [43], that improve GANs from the loss-function perspective. In addition, studies such as LAPGAN [44], PGGAN [45], StyleGAN [12], EBGAN [46], SAGAN [47], and BigGAN [48] that have improved the structure of the models. PathologyGAN [49], which generates pathological data, is a network that uses BigGAN as a baseline architecture. The BigGAN is a model that stably trained GANs with a large dataset such as the ImageNet dataset [50], which consists of 14 million images and thousands of classes by applying the Spectral Normalization (SN) and Self-Attention (SA) layers to both G and D networks. PathologyGAN used Relativistic Average Discriminator [51] and succeeded in generating high-quality fidelity data with faster convergence when hinge loss is used as the GAN's loss function. In addition, a structure that maps the latent vector  $z$ , the input of G, to  $w \sim M(z)$ , through StyleGAN's mapping network and adaptive instance normalization (AdaIN) was used. Quiros et al. used two H&E breast cancer datasets from the Netherlands Cancer Institute (NKI) and Vancouver General Hospital (VGH) cohort, and they achieved a Fréchet Inception Score (FID) of about 16.65 [49]. Deshpande et al.

generated tissue images using CGAN [52] for the CRAG (Colorectal Adenocarcinoma Gland) dataset, and they combined several small tiles to create high-resolution images with realistic pathological features such as gland appearance and nuclei structure [53]. Sondre et al. generated synthetic data from random latent vectors by combining CycleGAN [54] and ACGAN (Auxiliary classifier GAN) [55] for dermoscopic skin lesion images. Sondre et al. applied image transfer via CycleGAN to a common class (nevus) to create a rare and specific melanoma class [56]. Boyd et al. [57] generated pathological data with realistic and fine details through visual field expansion method and reported a FID of approximately 21 and 37 for CAMELYON17 [58] and CRC [59] datasets.

Digital pathology uses deep learning methods primarily for classification, segmentation, or tasks that assist medical prediction. In recent years, research on the image-generation task has increased, and pathological image generation using GANs in particular has the potential for various applications.

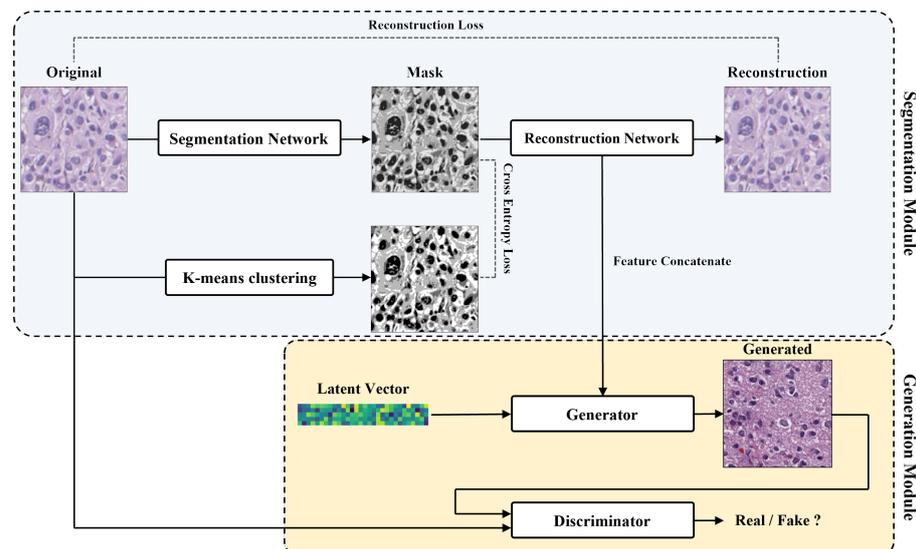
In this study, we propose a method for generating data by applying a simple method called a “embedding feature concatenation” to the conventional GAN architecture. We generated ODG images using the TCGA dataset, and the overall workflow was divided into two parts: the segmentation module and the generation module.

### 3. Method

We propose a method to synthesize oligodendroglioma (ODG) pathology data. Our model is based on the PGGAN, which among GAN models, achieves high performance. For generating realistic images, we trained our generator using latent vector  $z$  and embedding features.

As illustrated in Figure 1, our proposed method consists of two modules.

- The segmentation module creates masked images from the reference images and extracts meaningful features in the reconstruction process.
- The generation module generates pathological images from the features of the segmentation module and latent vectors by the embedding feature concatenation method.



**Figure 1.** The overall flow of our proposed model. A segmentation module (light blue box), a generation module (yellow box).

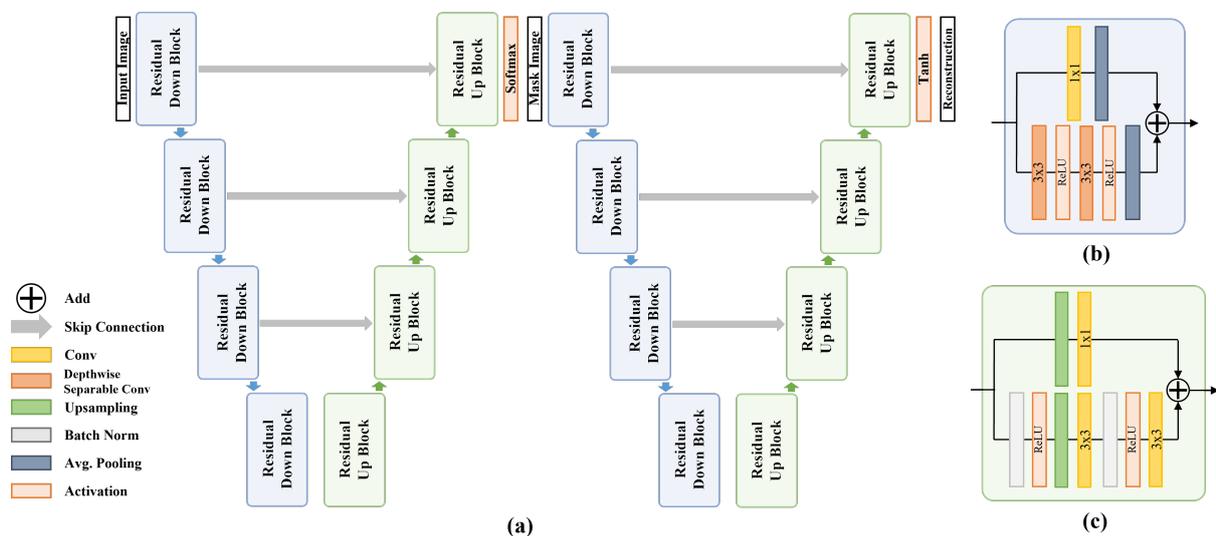
The segmentation module consists of two networks, one creates masked images from pathological images and the other reconstructs the reference images. The ODG dataset that we used has the following problem: the inability to produce segmentation labels. Therefore, k-means clustering was used to solve this problem. The generation module was designed based on the PGGAN. The PGGAN achieves good performance in general cases but often generates poor visual quality in the pathology domain. To generate visually realistic images,

we concatenate the embedding features. This method concatenates output features of the segmentation module to the corresponding scale level of generator layers. Consequently, we generated realistic, high-quality ODG data.

### 3.1. Segmentation and Reconstruction Module

In recent years, pixel-level segmentation studies have been conducted in digital pathology. Segmentation in the pathological domain distinguishes various regions such as cell nuclei, cells, and glands.

Our segmentation module aims to extract meaningful embedding features from reference images. The architecture was composed of two networks, as shown in Figure 2a. The first network is a segmentation network. The network creates masked images by discriminating blood vessels, cells, nuclei, and cytoplasm from the input pathology images. The second network reconstructs the masked results obtained by the reconstruction network into original input images of the segmentation network. Subsequently, the feature outputs of the intermediate layers of the reconstruction network are concatenated to train the generator. Additionally, we observed that the segmentation accuracy improved by adding the reconstruction network.



**Figure 2.** The segmentation module structure. (a) The structure of the segmentation network and reconstruction network. (b) The layer structure of the residual down-sampling block. (c) The layer structure of the residual up-sampling block.

Both networks were designed based on Res-UNet [60], which is a residual block applied to U-net. We designed networks based on both Res-UNet and the residual blocks applied to U-net. To capture the local information of the input accurately, both networks have a symmetrical structure. We configured the two networks in the same U-net architecture using four pairs of residual down-sampling blocks (RDB) and residual up-sampling blocks (RUB). Figure 2b,c show the structural layouts of the RDB and the RUB, respectively.

The RDB consists of two paths. One path connects two  $3 \times 3$  depth-wise separable convolution layers and an average pooling layer. The other goes through a  $1 \times 1$  convolution and an average pooling layer. Subsequently, the outputs of both paths are added. The RUB adds a path that connects an up-sampling layer and two  $3 \times 3$  convolution layers and the path of an up-sampling layer and a  $1 \times 1$  convolution layer.

We generated masked images using softmax in the last RUB, resulting in a segmentation network and obtained reconstructed images by using tanh in the reconstruction network.

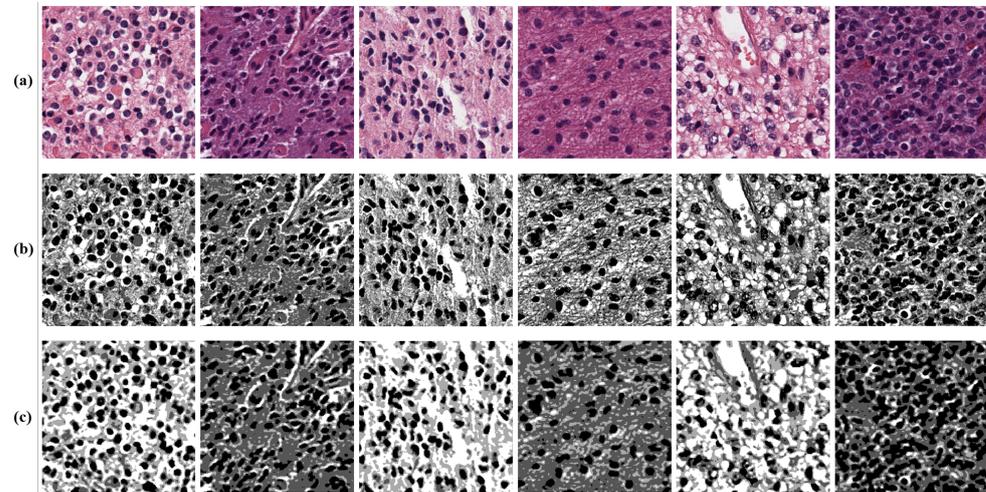
In general, the segmentation network aims to learn the function  $f : X \rightarrow Y$ , where  $X$  is the dataset of pathological images,  $Y$  is the corresponding mask label and the reconstruction network aims to learn the function  $g : Y \rightarrow X$ , which reconstructs the original input data.

Most studies on segmentation tasks use a paired dataset containing data with corresponding ground truth labels. Unfortunately, many medical datasets including the TCGA dataset we used do not provide segmentation labels.

We used k-means clustering to train the segmentation module using the TCGA dataset, because k-means clustering is a simple algorithm requiring few computational costs. We labeled groups of pixels with similar colors using k-means clustering. The TCGA dataset consisted of H&E-stained pathology images, and the images were stained with similar colors for cells, nuclei, and cytoplasm. Therefore, we conducted pixel grouping of similar colors using k-means clustering and classified the cells, nuclei, and cytoplasm with relevant labels.

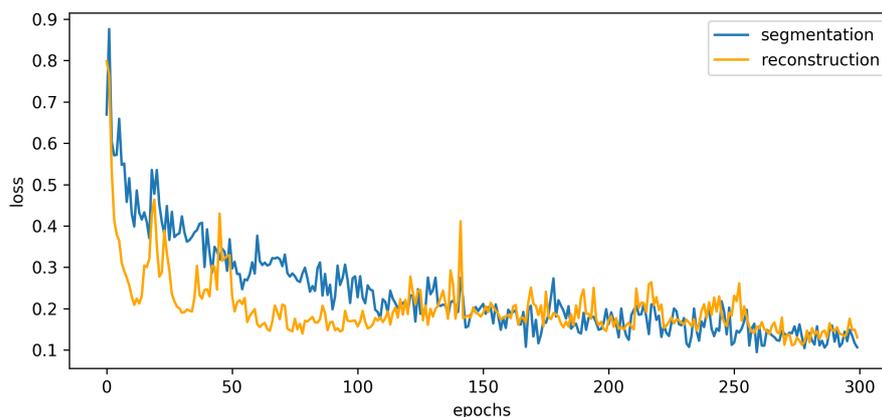
As shown in Figure 1, we used the mask image obtained by k-means clustering as ground truth labels and used the cross-entropy loss between the k-means clustering results and the segmentation results in the training process. Using this method, we addressed the problem of ground truth labels not being simply provided.

The segmentation operation results in an integer value of class per each pixel via softmax. It means that the masked data have  $k$ -channels as the output of the softmax layer, and each channel value is expressed as a probability of  $k$  classes. For visualization and training the reconstruction network, we mapped the results to the class corresponding to the highest probability. Then, we normalized the result to a value of  $[0, 1]$ . Consequently, our segmentation network generated the gray-scale masked data with a resolution of  $256 \times 256$ , as shown in Figure 3c. Figure 3a shows samples of TCGA input data, and Figure 3b shows examples of k-means clustering results.



**Figure 3.** Examples of images obtained through k-means clustering and the segmentation results. (a) The samples of the segmentation network input. (b) The samples of k-means clustering results. (c) The samples of the mask images of the segmentation module.

Figure 4 shows the loss during training of the segmentation module. The two lines are the loss of the segmentation network and the reconstruction network, respectively. Additionally, the embedding features of the reconstruction network were concatenated into the layers of the generator for each scale level. Thus, the embedding features extracted from the segmentation module improve the visual performance of the baseline.



**Figure 4.** The training loss of the segmentation module.

### 3.2. Concatenating Embedding Features in the Generation Module

This section describes how to concatenate embedding features into a general GAN structure. We generated ODG data by applying the proposed method based on a progressive growing GAN (PGGAN), which demonstrated stable training and high-quality image generation. Initially, the PGGAN maps the latent vector to the  $4 \times 4$  low-resolution samples. In addition, this model adopts the progressive training method of stacking new layers by increasing the resolution by a power of 2. Consequently, it stably generates high-resolution images. Similarly, we progressively trained our model by the proposed method concatenating the embedding features extracted from the reconstruction network of a pre-trained segmentation module of a PGGAN-based generator. As shown in Figure 2a, the reconstruction network of the segmentation module consists of 4 RDBs that constitute the encoder and 4 RUBs that constitute the decoder. We extracted the embedding feature outputs from three blocks: the last RDB, first and second RUBs. Each feature map was concatenated according to the output scale of the generator.

Therefore, the generator uses the information from the reconstruction network and supports GAN training with the process formulated as:

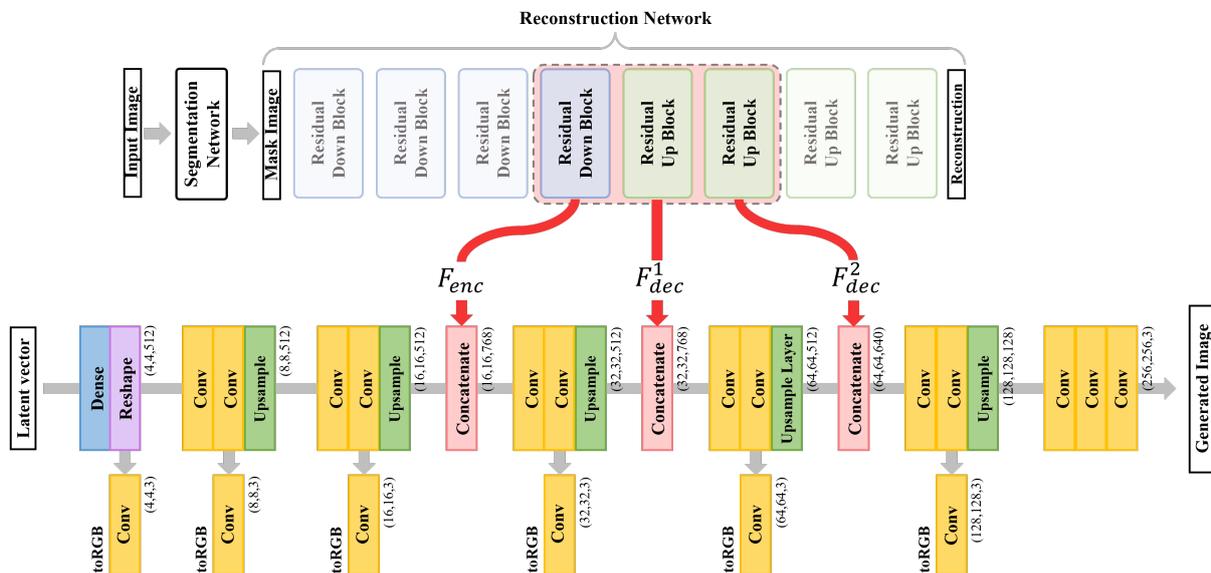
$$Img_G = G(z, \mathcal{F}_{enc}, \mathcal{F}_{dec}) \quad (1)$$

where  $G(\cdot)$  is a generator,  $z \sim \mathcal{N}(0, 1)$  is a latent vector, and  $F_{enc}$  and  $F_{dec}$  indicate the output feature maps of the encoder and decoder of the reconstruction network, respectively.

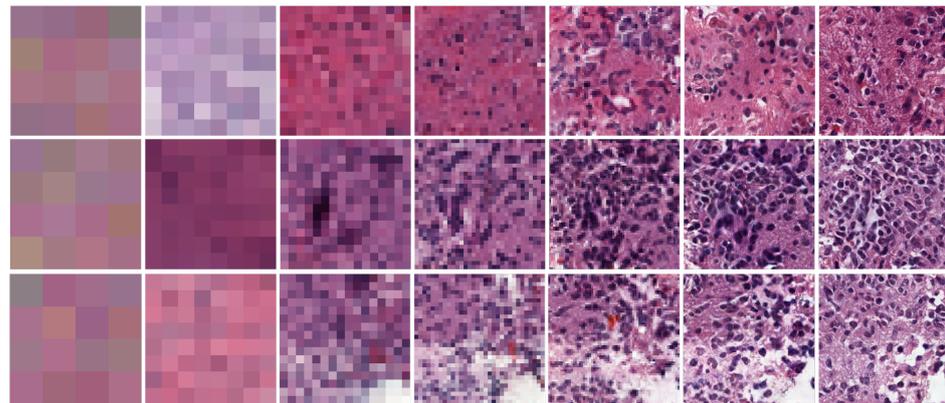
Figure 5 shows how to apply the embedding feature concatenation method during progressive training. Our training process also learns an initial generator that creates  $Img_{init} \in R^{4 \times 4}$  images from the latent vector  $z \in R^{512}$  in the same manner as the PGGAN, and the new layers fade in smoothly in subsequent steps. In the intermediate step, to generate  $16 \times 16$  resolution, the feature maps  $F_{enc} \in R^{16 \times 16 \times 512}$  are concatenated to the generator. Then,  $F_{dec}^1 \in R^{32 \times 32 \times 256}$ ,  $F_{dec}^2 \in R^{64 \times 64 \times 128}$  are concatenated in order. Consequently, we generated  $Img_G \in R^{256 \times 256 \times 3}$ , which is a high-quality pathological ODG image. Figure 6 shows the samples from low-resolution to high-resolution ODG data generated by our network.

For a more detailed description of the generation module, we used a generator and a discriminator with a symmetrical structure. In the initial generator and discriminator,  $4 \times 4$  convolutional layers are used at the beginning and end, respectively, and the kernel sizes of all other convolutional layers are  $3 \times 3$ . Both networks used the convolutional layer twice per scale for each progress and use LeakyReLU activation. The generator generates the RGB data  $Img_G$  as an output using a  $1 \times 1$  convolutional layer and tanh at the final stage, and the discriminator at the end applies a  $4 \times 4$  convolutional layer and a dense layer to distinguish synthetic and real data. The discriminator gradually add layers while using the weights of the previous training step during the training process following the PGGAN method.

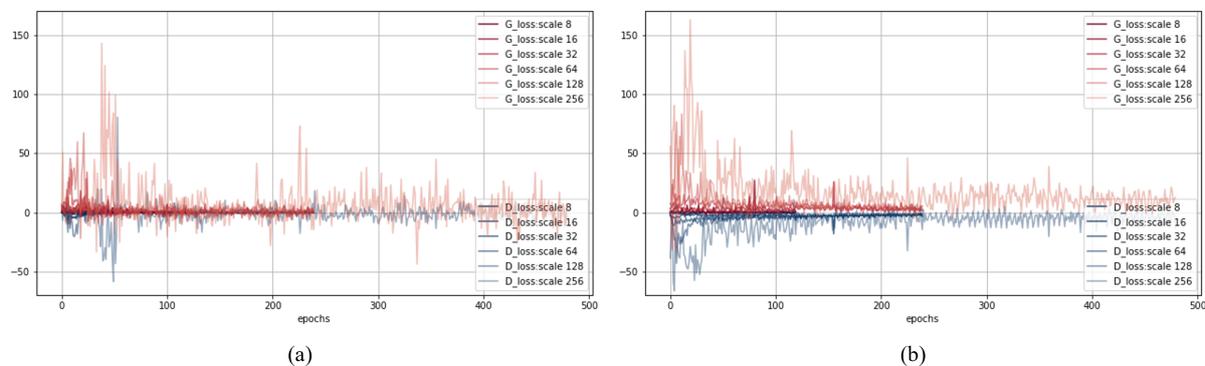
The training progressed well without applying a separate training method. Figure 7 shows samples generated by the proposed method.



**Figure 5.** The embedding feature concatenation method and the layer structure of the generator of the generation module.



**Figure 6.** Low to high-resolution data samples were progressively generated by our model. From  $4 \times 4$  (leftmost) to  $256 \times 256$  (rightmost).



**Figure 7.** The plots of losses for all resolutions during training: (a) PGGAN, (b) proposed method.

### 3.3. Training

In practice, applying the embedding feature concatenation method requires a pre-trained segmentation module. The segmentation network was trained using supervised

learning with the softmax outputs and the final layer of the network and the k-means clustering guide set. The k-means clustering results can be considered a multi-label classification problem; categorical cross-entropy (CCE) was applied as an objective function. Therefore, the segmentation loss that was used can be formulated as:

$$L_{seg} = CCE(S(\mathbf{x}), \mathbf{c}) \quad (2)$$

$$\mathbf{c} = Cluster(\mathbf{x}) \quad (3)$$

where  $S(\cdot)$  denotes the segmentation network,  $\mathbf{x}$  is the input data sample from the input dataset  $\{x_i | i = 1, \dots, N\}$ , and  $\mathbf{c}$  is the corresponding centroid from the cluster set  $\{c_i | i = 1, \dots, N\}$ , which is the result of the k-means clustering algorithm.

We used the k-means clustering guide set because our training dataset does not have corresponding ground truth labels. Unlike natural images, in which local features are important, pixel-level labels can be created based on clusters that correspond to fairly simple features in the pathological images. When the guide labels are obtained through the k-means clustering algorithm, it is not experimentally more accurate than the paired dataset with the ground truth, but it leads to reasonable segmentation performances. The model delivered good performance when  $K = 4$  or  $5$ , and it was fixed at  $K = 4$  in the experiment.

The reconstruction network,  $R(\cdot)$ , reconstructs the given input images  $\mathbf{x}$  from  $S(\mathbf{x})$ , which are the masked samples. We applied the mean squared error (MSE) to the training to reduce the the difference between the original data input  $x$  and the reconstructed data  $R(S(x))$ . Therefore, the reconstruction loss was formulated as follows:

$$L_{recon} = MSE(R(S(\mathbf{x})), \mathbf{x}) \quad (4)$$

By adding the reconstruction model and loss to a single segmentation model, we obtained more highly accurate segmentation results. Finally, the total loss of the segmentation module was utilized by adding  $L_{seg}$  and  $L_{recon}$ , as shown in Equation (5):

$$L_{total} = L_{seg} + \lambda_{recon} L_{recon} \quad (5)$$

where  $\lambda_{recon}$  is a reconstruction parameter heuristically set to 5.

We used the WGAN-GP objective function [38] as an adversarial objective function to train the generation module, and this objective function converges with that of our feature concatenation method. The GAN objective function is formulated as follows:

$$L_{gan} = \mathbb{E}_{\tilde{x} \sim \mathbb{P}_G}([D(\tilde{x})]) - \mathbb{E}_{x \sim \mathbb{P}_r}([D(x)]) + \lambda_{GP} \mathbb{E}_{\hat{x} \sim \mathbb{P}_{\hat{x}}}(\|(\nabla_{\hat{x}} D(\hat{x}) - 1)\|_2) \quad (6)$$

where  $\tilde{x}$  denotes a synthesis output from the distribution  $\mathbb{P}_G$ ,  $x$  denotes an image from the real distribution  $\mathbb{P}_r$  and  $\hat{x}$  denotes a sample image from sample distribution  $\mathbb{P}_{\hat{x}}$ . We define  $\hat{x}$  as the formula of  $\hat{x} \leftarrow \epsilon x + (1 - \epsilon)\tilde{x}$  and set the penalty coefficient  $\lambda_{gp}$  to 10 as used in [38]. Figure 7 shows the losses of the discriminator and generator according to the training epochs of PGGAN and the proposed method. Each graph shows the training results from scale 8 to 256 pixels. Figure 7a plots the training curves of PGGAN and Figure 7b plots the training curves of the proposed method.

Our networks were implemented using Tensorflow and Keras. We optimized the Generation Module using Adam with  $\beta_1$  and  $\beta_2$ , 0.9 and 0.99, respectively. The module generator and discriminator had the same learning rate (0.0001). We trained the networks using an Intel Core i7-10700 2.9 GHz and NVIDIA GPU RTX 3090 with 24 GB GPU memories. The module required approximately a week for training.

#### 4. Experimental Results

In this section, we describe the experiments conducted to measure the performance of the proposed model in two ways. First, we compared the quality of the images generated using a qualitative evaluation method. To measure qualitative scores, we used modified AMT perceptual studies [61–63]. We experimented using perceptual surveys in paired and unpaired settings. Through these experiments, we compared the confusion rates of the

generation results from the proposed method with the embedding feature concatenation method and other methods, WGAN-GP, PGGAN and Boyd et al. [57].

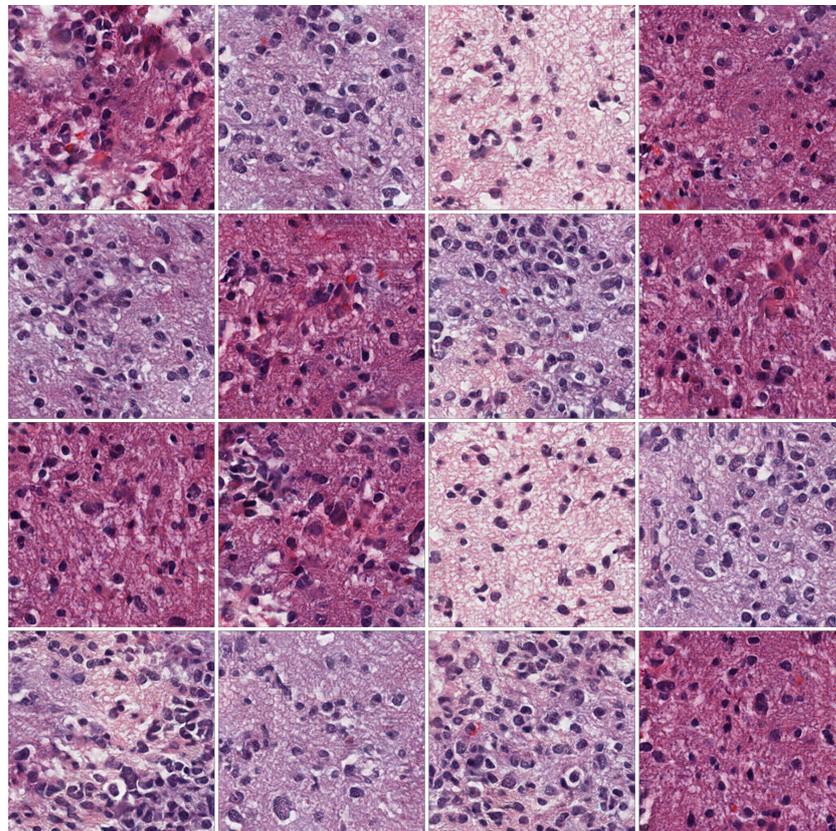
Second, to compare the two models, we calculated the Fréchet Inception Score (FID) [64] and Inception Score (IS) [65]. The score is a commonly used metric for measuring the performance of generative models. We calculated the two metrics and described the advantages of applying our method compared to the other methods.

#### 4.1. Preparing Data

We used data labeled as oligodendroglioma among the primary diagnosis classifications of TCGA-LGG as training data. The dataset was collected from <https://portal.gdc.cancer.gov/> (accessed on 20 December 2021). The TCGA-LGG ODG dataset consists of H&E stained pathology slides. We cropped whole-slide images to  $512 \times 512$  patches based on a  $40\times$  magnification. Subsequently, we resized it to  $256 \times 256$  pixels to simplify this task. We then removed the patches that were located at the boundary and that showed too little tissue because these negatively affect the training. We constructed a training dataset with a total of 120 k samples by cropping slides from the TCGA-LGG ODG dataset as patches and filtering out unnecessary data.

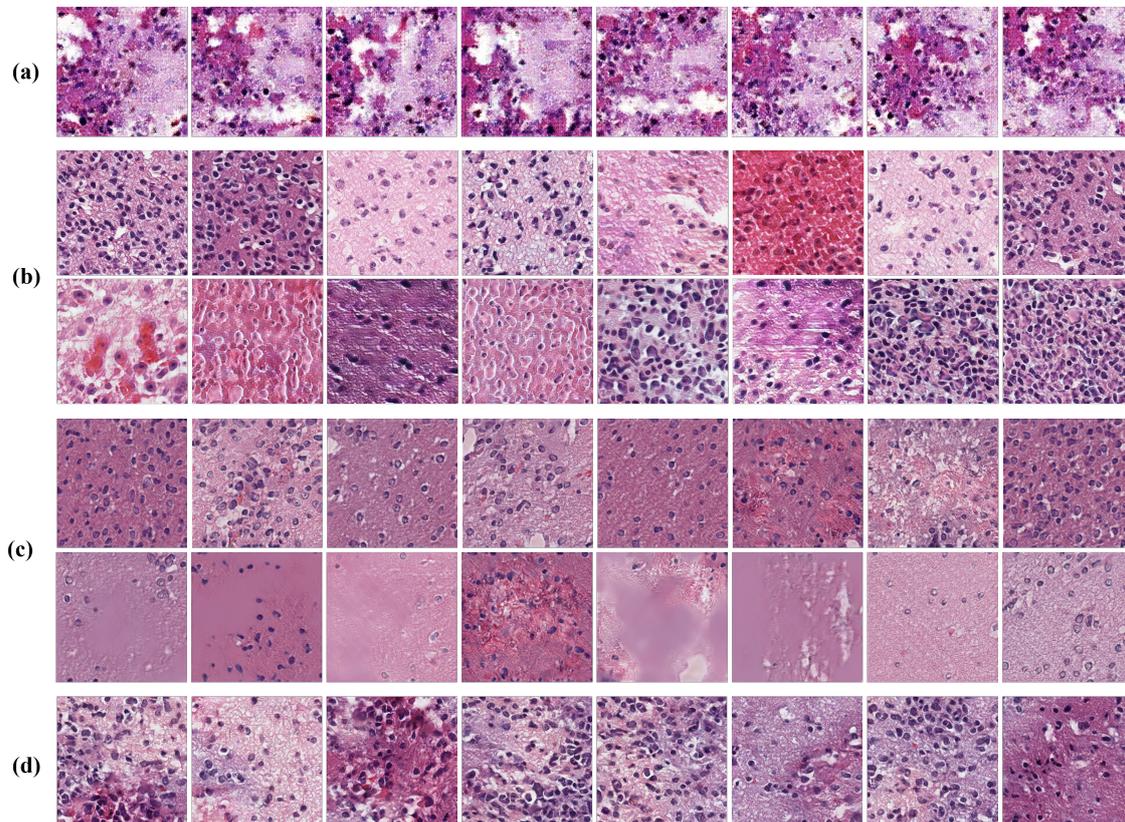
#### 4.2. Examples of the Proposed Method

We trained our network using a pre-processed ODG dataset using synthetic ODG data with a resolution of  $256 \times 256$  pixels. Figure 8 shows some of the samples generated by the proposed model. The proposed model captured the morphological features seen in pathological images such as cells, nuclei, and tissues while generating visually superior results. However, unrealistic tissue images were occasionally synthesized, indicating that the model can be improved.



**Figure 8.** Oligodendroglioma samples with a resolution of  $256 \times 256$  generated by the proposed method.

Although there is a color difference between slides in the H&E-stained tissue images, the tissues are stained with the same color in each slide because of the nature of staining. However, we observed samples in which the extracellular material's staining was not uniform in the generated patch images. As shown in Figure 9d, the color of some samples changed to a different staining color. These samples indicate a limitation of our model that must be addressed.



**Figure 9.** Generated samples from the WGAN-GP, PGGAN, Boyd et al. and our model. (a) The WGAN-GP samples generate images with a poorly uniform staining color and quality. (b) The PGGAN samples contain unrealistic patterns that do not exist in real data. (c) The samples from Boyd et al. generate images with high quality (upper row) but sometimes fail to generate (lower row). (d) The samples from our model generate images with a poorly uniform staining color.

#### 4.3. Qualitative Evaluation

We used a qualitative method to measure how distinct the synthetic ODG images we generated are from real data. We performed the perceptual experiments as a survey to measure the human confusion rate under two settings, and we designed a variation of AMT perceptual studies. The survey was conducted under two settings as follows:

1. **Paired Setting:** Subjects are shown sequences of real and synthetic image pairs for four seconds. Each trial required them to choose which they thought was the real image. The experiment consisted of 200 trials per subject, and the first 20 trials gave the correct answers as feedback.
2. **Unpaired Setting:** Subjects were randomly shown one of the WGAN-GP results, the PGGAN results, the results of Boyd et al., our results, or the real images. The subjects were asked to choose whether the images shown in each trial were either real or synthetic. This survey included a total of 250 trials, and the first 25 trials provided the correct answers as feedback.

We surveyed 72 subjects with general medical knowledge, including specialists in pathology. We configured all images used in the survey to have  $256 \times 256$  resolution

(but generated images by Boyd et al. were produced with  $224 \times 224$ ) and calculated the confusion rate as the average of the subjects' responses. The confusion rate of each subject means the rate at which the subjects answered labeled real among the fake images, and a larger value means that the synthetic data were judged as from real domain.

The confusion rates of the subjects are listed in Table 1. In a paired setting, the model increased the confusion rate by approximately 12.23% to  $55.19 \pm 11.18$  higher than the PGGAN of  $42.96 \pm 12.76$ . For the unpaired setting, the confusion rate of our model was  $51.25 \pm 13.93$ , which is a 12.5% increase over the PGGAN of  $38.75 \pm 16.26$ . As shown in Figure 9c, the method of Boyd et al. also synthesized high-quality images and recorded the second highest confusion rates of approximately 52.13% and 50.89% for each setting. Compare to the method of Boyd et al., our model showed 3.06 and 0.36% higher performance for paired and unpaired settings, respectively.

**Table 1.** The qualitative evaluation results of our study. The confusion-rate measurement result by model in the paired setting (top). The confusion-rate measurement result by model in the unpaired setting (bottom).

Survey Setting	Model	Confusion Rate (%) $\uparrow$
paired	WGAN-GP [38]	$15.28 \pm 12.63$
	PGGAN [45]	$42.96 \pm 12.76$
	Boyd et al. [57]	$52.13 \pm 10.37$
	Ours	$55.19 \pm 11.18$
unpaired	WGAN-GP [38]	$13.89 \pm 18.08$
	PGGAN [45]	$38.75 \pm 16.26$
	Boyd et al. [57]	$50.89 \pm 18.58$
	Ours	$51.25 \pm 13.93$

#### 4.4. Quantitative Evaluation Metrics

We used two metrics for the quantitative evaluation. The first is the FID score, which measures the distance between the images generated by the trained and training datasets. In addition, for comparison with our model, we measured the FID of the other methods generation results. We configured the number of data samples equal to the training data for the FID calculation and used the Inception-V3 model [66] pre-trained on ImageNet [50]. We also calculated the IS and compared its performance with that of the other methods. The IS calculates a score for the fidelity and diversity of the generated image and higher value means better performance.

The FID and IS values are listed in Table 2. Our model scored the second lowest FID score of 34.96359, which is 4.28195 higher than the PGGAN. The lower the FID, the better the performance. This means that the features of the PGGAN samples extracted from the Inception-V3 matched the real data distribution more accurately. In the case of IS, our method measured the highest score of approximately 2.91609 higher, indicating that the fidelity and diversity were higher than that of the other methods.

**Table 2.** The comparison of quantitative evaluation through FID and IS.

Model	FID $\downarrow$	IS $\uparrow$
Real Data	-	$2.85245 \pm 0.01389$
WGAN-GP [38]	283.50172	$1.45169 \pm 0.01101$
PGGAN [45]	30.68164	$2.40684 \pm 0.03606$
Boyd et al. [57]	37.11364	$2.73736 \pm 0.02308$
Ours	34.96359	$2.91609 \pm 0.03873$

## 5. Discussion

Due to the increasing interest in using deep learning for digital pathology, we conducted a study to synthesize ODG data to solve the problem of obtaining sufficient medical data.

We propose a method of concatenating the embedding features extracted from the segmentation module to the base network generator. Consequently, there was no improvement in the FID, but the IS quantitative results improved compared with the base network PGGAN. FID and IS evaluated the performance using the ImageNet-pre-trained model. Therefore, there was a difference in the domain because our dataset consisted of only pathology images. The pathology data-generation task achieved a low score compared with the general datasets in both metrics because both the base model and our generated images are from a similar domain.

Therefore, we made a subjective comparison through a qualitative evaluation. We conducted a survey, consisting of two settings as a qualitative evaluation, and our results achieved higher scores for both settings. In the paired setting, the confusion rate increased by approximately 12.23%, and in the unpaired setting, it increased by approximately 12.5%. Our results show that the confusion rates measured respectively 55.19%, 51.25% in the survey settings, indicating that the subjects could not reliably distinguish between the real and the synthetic images.

There are differences in the visual quality compared to the generation results of the PGGAN. Figure 9b shows examples of good and bad samples generated by the PGGAN. The upper row contain good samples and the lower row contain bad samples. The PGGAN synthesized high-quality pathological images; however, unnatural patterns that do not exist in the real domain are also generated. However, our proposed method created more realistic images of pathological characteristics. Our proposed method improved the visual performance by solving the problem of the baseline network that generated artifacts, i.e., strange patterns.

Although the proposed method raises the problem of normalizing the staining color, addressing it is beyond the scope of this study. However, we hope that future research will solve this problem and demonstrate the potential for pathological image generation with higher resolution. Furthermore, we think that future studies will achieve advanced performance in data augmentation and binary classification tasks on positive and negative datasets.

## 6. Conclusions

This research was intended to improve visual performance over the existing model for generating oligodendroglioma pathology data.

Unlike the existing method, the proposed method extracts embedding features with the segmentation module and uses it for training the generator. We used k-means clustering to train the segmentation module on the training dataset without labels. The reconstruction network was used to extract embedding features and concatenate them to the layers of the generator. Our generator synthesized pathological data from the linked embedding features and latent vectors.

Although the proposed method has a problem with normalizing the staining color, it consistently generated samples with high visual quality. The results were analyzed by comparison with the PGGAN used as a baseline, and low FID and high IS values were achieved. Compared to subjective evaluation by humans, the proposed method exhibited a higher confusion rate. However, the results show that our method causes more confusion in the subjective evaluation of humans. Therefore, the proposed method could visually complement the baseline network.

We expect that the proposed method will generate more realistic pathology data by solving the problem of normalizing the staining color in future studies. We hope that extended studies will synthesize meaningful data as a data augmentation technique for the binary positive and negative classification of many tumors.

**Author Contributions:** Conceptualization, J.K.; Investigation, S.K. (Seungjong Kim) and J.W.; Methodology, J.K.; Project administration, J.Y. and S.K. (Soonchul Kwon); Software, J.K. and S.K. (Seungjong Kim); Supervision, J.Y. and S.K. (Soonchul Kwon); Writing—original draft, J.K.; Writing—review and editing, S.K. (Seungjong Kim) and J.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** This work (Grants No.S3075979) was supported by Business for Startup growth and technological development(TIPS Program) funded Korea Ministry of SMEs and Startups in 2021. The results shown here are in whole or part based upon data generated by the TCGA Research Network: <https://www.cancer.gov/tcga> (accessed on 20 December 2021).

**Conflicts of Interest:** The authors declare that there is no conflict of interest.

## References

- Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [\[CrossRef\]](#)
- He, H.; Bai, Y.; Garcia, E.A.; Li, S. ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In Proceedings of the 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), Hong Kong, China, 1–8 June 2008; pp. 1322–1328.
- Han, H.; Wang, W.Y.; Mao, B.H. Borderline-SMOTE: A new over-sampling method in imbalanced data sets learning. In Proceedings of the International Conference on Intelligent Computing, Hefei, China, 23–26 August 2005; Springer: Berlin/Heidelberg, Germany, 2005; pp. 878–887.
- Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* **2014**, *27*, 2672–2680.
- Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.
- Hartigan, J.A.; Wong, M.A. Algorithm AS 136: A k-means clustering algorithm. *J. R. Stat. Soc. Ser. C (Appl. Stat.)* **1979**, *28*, 100–108. [\[CrossRef\]](#)
- Kleihoues, P. *Pathology and Genetics of Tumors of the Nervous System*; International Agency for Research Cancer: Lyon, France, 1997; pp. 78–79.
- Van den Bent, M.J.; Reni, M.; Gatta, G.; Vecht, C. Oligodendroglioma. *Crit. Rev. Oncol./Hematol.* **2008**, *66*, 262–272. [\[CrossRef\]](#)
- Ostrom, Q.T.; Gittleman, H.; Truitt, G.; Boscia, A.; Kruchko, C.; Barnholtz-Sloan, J.S. CBTRUS statistical report: Primary brain and other central nervous system tumors diagnosed in the United States in 2011–2015. *Neuro-oncology* **2018**, *20*, iv1–iv86. [\[CrossRef\]](#)
- Suh, Y.L.; Koo, H.; Kim, T.S.; Chi, J.G.; Park, S.H.; Khang, S.K.; Choe, G.; Lee, M.C.; Hong, E.K.; Sohn, Y.K.; et al. Tumors of the Central Nervous System in Korea A Multicenter Study of 3221 Cases. *J. Neuro-Oncol.* **2002**, *56*, 251–259. [\[CrossRef\]](#)
- Liu, Z.; Luo, P.; Wang, X.; Tang, X. Deep learning face attributes in the wild. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 3730–3738.
- Karras, T.; Laine, S.; Aila, T. A style-based generator architecture for generative adversarial networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 16–20 June 2019; pp. 4401–4410.
- Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.
- Gecer, B.; Aksoy, S.; Mercan, E.; Shapiro, L.G.; Weaver, D.L.; Elmore, J.G. Detection and classification of cancer in whole slide breast histopathology images using deep convolutional networks. *Pattern Recognit.* **2018**, *84*, 345–356. [\[CrossRef\]](#)
- Gu, F.; Burlutskiy, N.; Andersson, M.; Wilén, L.K. Multi-resolution networks for semantic segmentation in whole slide images. In *Computational Pathology and Ophthalmic Medical Image Analysis*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 11–18.
- Guo, Z.; Liu, H.; Ni, H.; Wang, X.; Su, M.; Guo, W.; Wang, K.; Jiang, T.; Qian, Y. A fast and refined cancer regions segmentation framework in whole-slide breast pathological images. *Sci. Rep.* **2019**, *9*, 882. [\[CrossRef\]](#)
- Lin, H.; Chen, H.; Graham, S.; Dou, Q.; Rajpoot, N.; Heng, P.A. Fast scannet: Fast and dense analysis of multi-gigapixel whole-slide images for cancer metastasis detection. *IEEE Trans. Med. Imaging* **2019**, *38*, 1948–1958. [\[CrossRef\]](#)
- Zhao, Z.; Lin, H.; Chen, H.; Heng, P.A. PFA-ScanNet: Pyramid feature aggregation with synergistic learning for breast cancer metastasis analysis. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Shenzhen, China, 13–17 October 2019; Springer: Berlin/Heidelberg, Germany, 2019; pp. 586–594.

19. BenTaieb, A.; Hamarneh, G. Topology aware fully convolutional networks for histology gland segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Athens, Greece, 17–21 October 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 460–468.
20. Xu, Y.; Li, Y.; Wang, Y.; Liu, M.; Fan, Y.; Lai, M.; Eric, I.; Chang, C. Gland instance segmentation using deep multichannel neural networks. *IEEE Trans. Biomed. Eng.* **2017**, *64*, 2901–2912.
21. Van Eycke, Y.R.; Balsat, C.; Verset, L.; Debeir, O.; Salmon, I.; Decaestecker, C. Segmentation of glandular epithelium in colorectal tumours to automatically compartmentalise IHC biomarker quantification: A deep learning approach. *Med. Image Anal.* **2018**, *49*, 35–45. [[CrossRef](#)]
22. Graham, S.; Chen, H.; Gamper, J.; Dou, Q.; Heng, P.A.; Snead, D.; Tsang, Y.W.; Rajpoot, N. MILD-Net: Minimal information loss dilated network for gland instance segmentation in colon histology images. *Med. Image Anal.* **2019**, *52*, 199–211. [[CrossRef](#)]
23. Ding, H.; Pan, Z.; Cen, Q.; Li, Y.; Chen, S. Multi-scale fully convolutional network for gland segmentation using three-class classification. *Neurocomputing* **2020**, *380*, 150–161. [[CrossRef](#)]
24. Qu, H.; Riedlinger, G.; Wu, P.; Huang, Q.; Yi, J.; De, S.; Metaxas, D. Joint segmentation and fine-grained classification of nuclei in histopathology images. In Proceedings of the 2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019), Venice, Italy, 8–11 April 2019; pp. 900–904.
25. Tokunaga, H.; Teramoto, Y.; Yoshizawa, A.; Bise, R. Adaptive weighting multi-field-of-view cnn for semantic segmentation in pathology. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 16–20 June 2019; pp. 12597–12606.
26. Bulten, W.; Pinckaers, H.; van Boven, H.; Vink, R.; de Bel, T.; van Ginneken, B.; van der Laak, J.; Hulsbergen-van de Kaa, C.; Litjens, G. Automated deep-learning system for Gleason grading of prostate cancer using biopsies: A diagnostic study. *Lancet Oncol.* **2020**, *21*, 233–241. [[CrossRef](#)]
27. Bulten, W.; Bándi, P.; Hoven, J.; van de Loo, R.; Lotz, J.; Weiss, N.; van de Laak, J.; Ginneken, B.V.; Hulsbergen-van de Kaa, C.; Litjens, G. Epithelium segmentation using deep learning in H&E-stained prostate specimens with immunohistochemistry as reference standard. *Sci. Rep.* **2019**, *9*, 864.
28. Xie, Y.; Kong, X.; Xing, F.; Liu, F.; Su, H.; Yang, L. Deep voting: A robust approach toward nucleus localization in microscopy images. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; Springer: Berlin/Heidelberg, Germany, 2015; pp. 374–382.
29. Graham, S.; Vu, Q.D.; Raza, S.E.A.; Azam, A.; Tsang, Y.W.; Kwak, J.T.; Rajpoot, N. Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. *Med. Image Anal.* **2019**, *58*, 101563. [[CrossRef](#)]
30. Xing, F.; Cornish, T.C.; Bennett, T.; Ghosh, D.; Yang, L. Pixel-to-pixel learning with weak supervision for single-stage nucleus recognition in Ki67 images. *IEEE Trans. Biomed. Eng.* **2019**, *66*, 3088–3097. [[CrossRef](#)]
31. Alom, M.Z.; Yakopcic, C.; Taha, T.M.; Asari, V.K. Nuclei segmentation with recurrent residual convolutional neural networks based U-Net (R2U-Net). In Proceedings of the NAECON 2018-IEEE National Aerospace and Electronics Conference, Dayton, OH, USA, 23–26 July 2018; pp. 228–233.
32. Li, X.; Wang, Y.; Tang, Q.; Fan, Z.; Yu, J. Dual U-Net for the segmentation of overlapping glioma nuclei. *IEEE Access* **2019**, *7*, 84040–84052. [[CrossRef](#)]
33. Xie, Y.; Xing, F.; Kong, X.; Su, H.; Yang, L. Beyond classification: Structured regression for robust cell detection using convolutional neural network. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; Springer: Berlin/Heidelberg, Germany, 2015; pp. 358–365.
34. Xie, Y.; Xing, F.; Shi, X.; Kong, X.; Su, H.; Yang, L. Efficient and robust cell detection: A structured regression approach. *Med. Image Anal.* **2018**, *44*, 245–254. [[CrossRef](#)]
35. Qin, X.; Zhang, Z.; Huang, C.; Dehghan, M.; Zaiane, O.R.; Jagersand, M. U2-Net: Going deeper with nested U-structure for salient object detection. *Pattern Recognit.* **2020**, *106*, 107404. [[CrossRef](#)]
36. Radford, A.; Metz, L.; Chintala, S. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv* **2015**, arXiv:1511.06434.
37. Arjovsky, M.; Chintala, S.; Bottou, L. Wasserstein generative adversarial networks. In Proceedings of the International Conference on Machine Learning, PMLR, Sydney, Australia, 6–11 August 2017; pp. 214–223.
38. Gulrajani, I.; Ahmed, F.; Arjovsky, M.; Dumoulin, V.; Courville, A. Improved training of wasserstein gans. *arXiv* **2017**, arXiv:1704.00028.
39. Mao, X.; Li, Q.; Xie, H.; Lau, R.Y.; Wang, Z.; Paul Smolley, S. Least squares generative adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2794–2802.
40. Lim, J.H.; Ye, J.C. Geometric gan. *arXiv* **2017**, arXiv:1705.02894.
41. Metz, L.; Poole, B.; Pfau, D.; Sohl-Dickstein, J. Unrolled generative adversarial networks. *arXiv* **2016**, arXiv:1611.02163.
42. Che, T.; Li, Y.; Jacob, A.P.; Bengio, Y.; Li, W. Mode regularized generative adversarial networks. *arXiv* **2016**, arXiv:1612.02136.
43. Miyato, T.; Kataoka, T.; Koyama, M.; Yoshida, Y. Spectral normalization for generative adversarial networks. *arXiv* **2018**, arXiv:1802.05957.
44. Denton, E.L.; Chintala, S.; Szlam, A.; Fergus, R. Deep generative image models using a laplacian pyramid of adversarial networks. *Adv. Neural Inf. Process. Syst.* **2015**, *28*, 1486–1494.

45. Karras, T.; Aila, T.; Laine, S.; Lehtinen, J. Progressive growing of gans for improved quality, stability, and variation. *arXiv* **2017**, arXiv:1710.10196.
46. Zhao, J.; Mathieu, M.; LeCun, Y. Energy-based generative adversarial network. *arXiv* **2016**, arXiv:1609.03126.
47. Zhang, H.; Goodfellow, I.; Metaxas, D.; Odena, A. Self-attention generative adversarial networks. In Proceedings of the International conference on machine learning, ICML, Long Beach, CA, USA, 10–15 June 2019; pp. 7354–7363.
48. Brock, A.; Donahue, J.; Simonyan, K. Large scale GAN training for high fidelity natural image synthesis. *arXiv* **2018**, arXiv:1809.11096.
49. Quiros, A.C.; Murray-Smith, R.; Yuan, K. PathologyGAN: Learning deep representations of cancer tissue. *arXiv* **2019**, arXiv:1907.02644.
50. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei-Fei, L. Imagenet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255.
51. Jolicoeur-Martineau, A. The relativistic discriminator: A key element missing from standard GAN. *arXiv* **2018**, arXiv:1807.00734.
52. Mirza, M.; Osindero, S. Conditional generative adversarial nets. *arXiv* **2014**, arXiv:1411.1784.
53. Deshpande, S.; Minhas, F.; Rajpoot, N. Train Small, Generate Big: Synthesis of Colorectal Cancer Histology Images. In Proceedings of the International Workshop on Simulation and Synthesis in Medical Imaging, Lima, Peru, 4 October 2020; pp. 164–173.
54. Zhu, J.Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2223–2232.
55. Odena, A.; Olah, C.; Shlens, J. Conditional image synthesis with auxiliary classifier gans. In Proceedings of the International Conference on Machine Learning, PMLR, Sydney, Australia, 6–11 August 2017; pp. 2642–2651.
56. Fossen-Romsaas, S.; Storm-Johannessen, A.; Lundervold, A.S. Synthesizing Skin Lesion Images Using CycleGANs-A Case Study. *Norsk IKT-Konferanse Forskning Utdanning*, no. 1. 2020. Available online: <https://ojs.bibsys.no/index.php/NIK/article/view/837> (accessed on 20 February 2022).
57. Boyd, J.; Liashuha, M.; Deutsch, E.; Paragios, N.; Christodoulidis, S.; Vakalopoulou, M. Self-Supervised Representation Learning using Visual Field Expansion on Digital Pathology. In Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, Montreal, BC, Canada, 11–17 October 2021; pp. 639–647.
58. Bandi, P.; Geessink, O.; Manson, Q.; Van Dijk, M.; Balkenhol, M.; Hermsen, M.; Bejnordi, B.E.; Lee, B.; Paeng, K.; Zhong, A.; et al. From detection of individual metastases to classification of lymph node status at the patient level: The camelyon17 challenge. *IEEE Trans. Med. Imaging* **2018**, *38*, 550–560. [[CrossRef](#)]
59. Kather, J.N.; Halama, N.; Marx, A. 100,000 Histological Images of Human Colorectal Cancer and Healthy Tissue. Zenodo10. 2018. Available online: <https://zenodo.org/record/1214456#.YofTvFRBxPZ> (accessed on 20 February 2022).
60. Diakogiannis, F.I.; Waldner, F.; Caccetta, P.; Wu, C. ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data. *ISPRS J. Photogramm. Remote Sens.* **2020**, *162*, 94–114. [[CrossRef](#)]
61. Isola, P.; Zhu, J.Y.; Zhou, T.; Efros, A.A. Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1125–1134.
62. Zhang, R.; Isola, P.; Efros, A.A. Colorful image colorization. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 11–14 October 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 649–666.
63. Shaham, T.R.; Dekel, T.; Michaeli, T. Singan: Learning a generative model from a single natural image. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Korea, 27 October–2 November 2019; pp. 4570–4580.
64. Heusel, M.; Ramsauer, H.; Unterthiner, T.; Nessler, B.; Hochreiter, S. Gans trained by a two time-scale update rule converge to a local nash equilibrium. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 6627–6638.
65. Salimans, T.; Goodfellow, I.; Zaremba, W.; Cheung, V.; Radford, A.; Chen, X. Improved techniques for training gans. In Proceedings of the Advances in Neural Information Processing Systems, Barcelona, Spain, 5–10 December 2016; pp. 2234–2242.
66. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the inception architecture for computer vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 2818–2826.