

# TopoGAN: A Topology-Aware Generative Adversarial Network

Fan Wang<sup>[0000-0002-5402-5065]</sup>, Huidong Liu, Dimitris Samaras, and Chao Chen

Stony Brook University, Stony Brook, NY 11794, USA  
{fanwang1,huidliu,samaras}@cs.stonybrook.edu  
chao.chen.1@stonybrook.edu

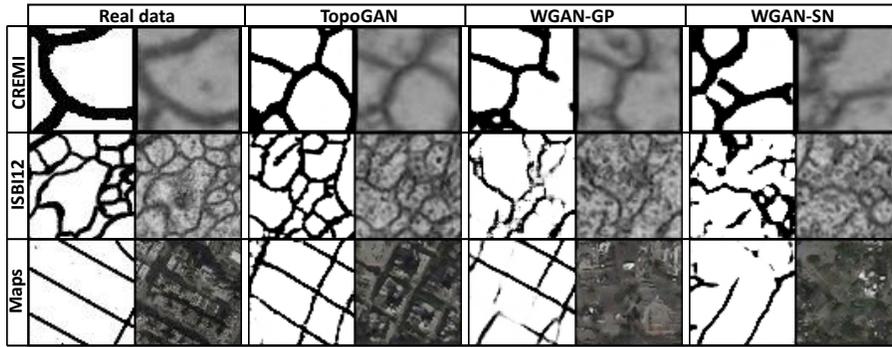
**Abstract.** Existing generative adversarial networks (GANs) focus on generating realistic images based on CNN-derived image features, but fail to preserve the structural properties of real images. This can be fatal in applications where the underlying structure (e.g., neurons, vessels, membranes, and road networks) of the image carries crucial semantic meaning. In this paper, we propose a novel GAN model that learns the topology of real images, i.e., connectedness and loopy-ness. In particular, we introduce a new loss that bridges the gap between synthetic image distribution and real image distribution in the topological feature space. By optimizing this loss, the generator produces images with the same structural topology as real images. We also propose new GAN evaluation metrics that measure the topological realism of the synthetic images. We show in experiments that our method generates synthetic images with realistic topology. We also highlight the increased performance that our method brings to downstream tasks such as segmentation.

**Keywords:** Topology, Persistent Homology, Generative Adversarial Network

## 1 Introduction

Generative adversarial networks (GANs) [20] have been very successful in generating realistic images. GANs train a generator to synthesize images that are similar to real images, and at the same time, a discriminator to distinguish these fake images from real ones. Through a minimax game, the generator converges to a network that generates synthetic images sampled from a distribution that matches the distribution of the real images.

When designing GANs, a key question is how to bridge the gap between the synthetic and real image distributions not only in appearance, but also in semantics. As shown [38, 44], widely-used GANs [20, 56, 47, 34, 5, 22] only match the first order moments of the distributions within a CNN-based image feature space. Newer methods match the synthetic/real image distributions using higher order statistics, e.g. second order statistics of the image features [44, 43]. Kosai et al. [32] explicitly add a statistical shape prior for face images into the generator. The intuition is that the more high order information a generator can learn, the more semantically realistic the synthesized images will be.



**Fig. 1.** Sample images in which the structures are neuron membranes and road networks from satellite images. From top to bottom: neuron images (CREMI [15]), neuron images (ISBI12 [4]) and satellite images (Google Maps [30]). From left to right: real images, images synthesized by TopoGAN, WGAN-GP and WGAN-SN. Each real/synthetic mask is paired with a textured image. For synthetic images, texture is added by a separately-trained pix2pix [30] network.

In this paper, we pay attention to the structural information of an image. In many applications, images contain structures with rich topology, e.g., biomedical images with membrane, neuron or vessel structures, and satellite images with road maps (Fig. 1). These structures and their topology, i.e., connectivity and loopy-ness, carry important semantic/functionality information. Structural fidelity becomes crucial if we want to use the synthetic images to train downstream methods that hinge on the structural information, e.g., diagnosis algorithms based on the structural richness of retinal vessels, navigation systems based on road network topology, or neuron classifiers based on neuron morphology and connectivity.

In this paper, we propose TopoGAN, the first GAN model that learns topology from real data. Topology directly measures structural complexity, such as the numbers of connected components and holes. This information is very difficult to learn, due to its global nature. The conventional GAN discriminator distinguishes synthetic and real images in terms of CNN-based features, but is agnostic to topological dissimilarity. Thus, the generator cannot learn real image topology. In Fig. 1, structures synthesized by conventional GANs (WGAN-GP and WGAN-SN) tend to be broken and disconnected.

Our main technical contribution is a novel *topological GAN loss* that explicitly matches the synthetic and real image distributions in terms of their topology. Based on persistent homology theory [16], we map both synthetic and real images into a topological feature space, where their topological dissimilarity can be measured as a loss. We show that our loss is differentiable and can be minimized through backpropagation. Our topological GAN loss complements the existing discriminator and teaches the generator to synthesize images that are realistic not only in CNN-based image features but also in topological features (Fig. 1). Note that TopoGAN only focuses on generating binary images (i.e. masks) de-

lineating the underlying structures. Once we have synthesized realistic topology structures, we can add texture with existing techniques such as pix2pix [30].

To the best of our knowledge, TopoGAN is *the first generative model that learns topology from real images*. We demonstrate the efficacy of TopoGAN through comprehensive experiments on a broad spectrum of biomedical, satellite and natural image datasets. We measure the success of our method in terms of a conventional GAN performance measure, FID [24]. Furthermore, we propose two novel topology-aware GAN measures, based on persistent homology and the Betti number. We show that TopoGAN outperforms baseline GAN models by a large margin in these topology-aware measures. Finally, we show that synthesized images with learnt topology can improve performance in downstream tasks such as image segmentation. In summary, our contributions are three-fold:

- We propose a topological GAN loss that measures the distance between synthetic and real image distributions in the space of topological features. Compared to previous topological loss that is applied to individual instances [28], our loss is the first to enforce topological similarity between distributions.
- We show that this loss is differentiable and incorporate it into GAN training.
- We propose novel topology-aware measures to evaluate generator performance in topological feature space.

## 2 Related work

Generative Adversarial Nets (GANs) [20] are very popular for modeling data distributions. However, GAN training is very unstable. WGAN [5], WGAN-GP [22], WGAN-TS [37], WGAN-QC [36] and others, use the Wasserstein distance to train GANs. Different gradient penalty strategies [22, 40, 60, 36] can stabilize GAN training effectively. Apart from the gradient penalty, Spectral Normalization (SN) [41] is also widely used for GAN training [7, 65]. PatchGAN [30] applies a GAN to local patches instead of the whole image in order to capture high frequency signals. Such local/high frequency signals are very useful in various generative models, such as Pix2pix [30, 63], CycleGAN [68] and SinGAN [57].

Several geometry-related GANs exploit geometric information on images. The geometricGAN [35] adopts the large margin idea from SVMs [14] to learn the discriminator and generator. The Localized GAN (LGAN) [50] uses local coordinates to parameterize the local geometry of the data manifold. The Geometry-Aware GAN (GAGAN) [32] is tailored for generating facial images using face shape priors. The Geometry-Consistent GAN (GcGAN) [18] uses a geometry-consistency constraint to preserve the image’s semantic structure. Geometric transformations are restricted to image flipping and rotation.

We note that high-order structural information has been used in adversarial networks for semantic segmentation. Existing methods [39, 19, 29] use adversarial losses in the semantic segmentation space as they encode high-order structural information. However, these methods do not explicitly preserve topology.

**Topological information for image analysis.** Many methods have been proposed to directly use persistent homology as a feature extraction mechanism.

The extracted topological feature can be vectorized [3], and used as input to kernel machines [54, 33, 9] or deep neural networks [26]. For fully supervised image segmentation tasks, topological information has been explicitly used as a constraint/loss to improve segmentation quality [28, 11, 64]. Mosinska et al. [42] model topology implicitly with feature maps from pretrained VGG networks [58], but the method does not generalize to structures of unseen geometries. We also refer to methods developed for retinal vessels [23] and lung airways [51]. These methods only focus on connectivity (0-dimensional topology) and cannot generalize to high-dimensional topology. In machine learning, topological information has been used to analyze data manifold topology [10, 25, 53, 46] and to leverage advanced structural information for graph learning [66, 67].

In generative models, Khruikov and Oseledets [31] use data manifold topology to compare synthetic and real data distributions as a qualitative measure of generative models. However, their measure still focuses on the standard image feature space, and cannot really evaluate whether the generator has learned the real image topology. Brüel-Gabrielsson et al. [8] use a loss to enforce the connectivity constraint in the generated images. However, enforcing hand-crafted topological constraints (e.g., connectedness) does not help the generator to learn the true topological distribution from real data. TopoGAN is the first generative model that automatically learns topological properties from real images.

### 3 Method

Our TopoGAN matches synthetic and real image distributions for both image and topology features. For this purpose, in addition to the conventional discriminator and generator losses, we introduce a new loss term for the generator,  $L_{topo}(P_{data}, G)$ . This loss term, called the *topological GAN loss*, measures how close the images generated by  $G$  are to the real images in terms of topology. Minimizing it forces the synthetic images to have similar topology as the real images. The discriminator loss is shown in Eq. (1). The generator loss (Eq. (2)) is a sum of the conventional generator loss and the new loss. Formally, we have

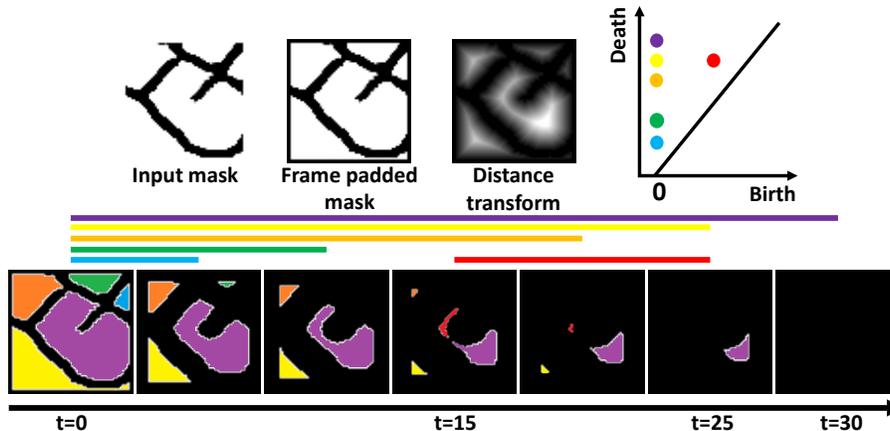
$$\arg \max_D \left[ \mathbb{E}_{x \sim P_{data}} \log D(x) + \mathbb{E}_{z \sim P_z} \log(1 - D(G(z))) \right], \quad (1)$$

$$\arg \min_G \left[ \underbrace{\mathbb{E}_{z \sim P_z} \log(1 - D(G(z)))}_{\text{conventional generator loss}} + \lambda \underbrace{L_{topo}(P_{data}, G)}_{\text{topological GAN loss}} \right], \quad (2)$$

where  $\lambda$  controls the weight of the topological GAN loss.

We focus on generating binary images, i.e., masks delineating structures such as vessels, neuron membranes, road networks, etc. The generator outputs a real-valued grey-scale image as the synthetic mask. The discriminator treats the input image (real or synthetic) as a real-valued grey-scale image ranging between 0 and 1. After mask synthesis, a separately-trained pix2pix [30] network fills in the textures based on each mask.

The rest of this section describes how to define and optimize the topological GAN loss. In Sec. 3.1, we explain how to extract the topological feature (called



**Fig. 2.** Illustration of persistent homology. Top row from left to right: the input mask, padded with a frame (so that all branches form holes), the distance transform and the output persistence diagram. Bottom row: the sequence of sublevel sets with different threshold values. Different holes are born and filled. The original holes are all born at  $t = 0$ . The almost-hole (red region, red bar, red dot) is born at a later time ( $t = 15$ ).

persistence diagram) of an input mask using the theory of persistent homology. In Sec. 3.2 and 3.3, we formalize the topological GAN loss by comparing the distributions of persistence diagrams computed from synthetic and real images respectively. Minimizing this loss practically moves a synthetic persistence diagram toward its matched real persistence diagram. This diagram modification effectively grows the structure/mask to complete almost-loops. This teaches the generator to synthesize images without incomplete loops.

As a separate technical contribution, we propose two new topology-aware metrics to compare the distributions in the topological feature space in Sec. 3.4.

### 3.1 Persistent Homology: From Images to Topological Features

We explain how to extract the topological feature of an input mask using the theory of persistent homology. We compute a *persistence diagram* capturing not only holes/loops, but also almost-holes/almost-loops (structures that almost form a hole or a loop) (Fig. 2). We describe the basic concepts, leaving technical details to supplemental material and a classic topological data analysis reference [16].

Given a topological space,  $y \subseteq \mathbb{R}^2$ , the holes and connected components are its 1- and 0-dimensional topological structures respectively. We mainly focus on 1-dimensional topology in this paper. The number of holes is the *Betti number*,  $\beta_y$ . In Fig. 2, we show a sample mask from the CREMI dataset, delineating a neuron membrane structure. We add a frame around the patch so that all structures are accounted for via 1-dimensional homology. In algebraic topology [45], we are effectively computing the relative homology.

We observe 5 holes (Betti number  $\beta_y = 5$ ) in the figure and the Betti number is only able to capture the complete holes. The dangling branch in the middle of the image almost creates a new hole. But this almost-hole is not captured by the Betti number. To effectively account for these almost-holes in our computations, we leverage the distance transform and the theory of persistent homology [16]. We review the distance transform:

**Definition 1.** *The Distance Transform (DT) [17] generates a map  $D$  for each pixel  $p$  on a binary image  $I$ :  $D(p) = \min_{q \in \Omega} \{\|p - q\| \mid I(q) = 0\}$ , in which  $\Omega$  is the image domain.*

Instead of only looking at the original function, we apply a distance transform to the mask and get a non-negative scalar function defined on the whole image domain,  $f_y : \Omega \rightarrow \mathbb{R}^+$ . We define the *sublevel set* of  $f_y$  as the domain thresholded by a particular threshold  $t$ , formally,  $\Omega_{f_y}^t = \{x \in \Omega \mid f_y(x) \leq t\}$ . We notice that one can take different sublevel sets with different thresholds. For certain threshold values, the almost-hole becomes a complete hole. The sequence of all possible sublevel sets, formally called the *filtration* induced by  $f_y$ , essentially captures the growing process of the initial mask.

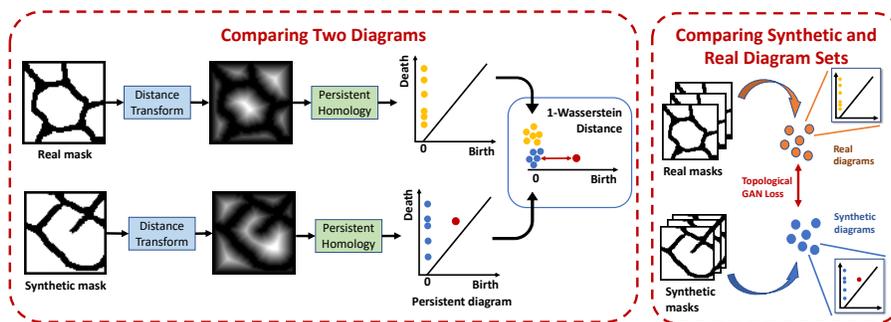
Persistent homology takes the whole filtration and inspects its topological structures (holes, connected components, and higher dimensional topological structures). Each topological structure lives during an interval of threshold values. In Fig. 2, the five original holes are born at  $t = 0$  and filled at different times, when they are filled up by the growing mask. The almost-hole (in red) is born at  $t = 15$ , when the purple hole is split into two. It dies at  $t = 25$ . All the holes (with life spans drawn as horizontal bars) are recorded as a 2D point set called a *persistence diagram*. The birth time and death time of each hole become the two coordinates of its corresponding point. In this diagram, we have 5 points with *birth* = 0 and a red point with a non-zero birth time, for the almost-hole<sup>1</sup>.

### 3.2 Distance Between Diagrams and Topological GAN Loss

In this section, we formalize our topological GAN loss. Using the distance transform and persistent homology, we transform each input binary image  $y$  into its corresponding persistence diagram,  $\text{dgm}(f_y)$ , which we call the topological feature of  $y$ . We first introduce the distance between any two persistence diagrams, which measures the topological dissimilarity between two images. Next, we define our topological GAN loss as the distance between two sets of diagrams, computed from synthetic masks and real masks respectively. We use optimal transport [62] to match the two sets of diagrams, and then define the loss as the total distance between the matched diagram pairs. An illustration of the topological GAN loss can be found in Fig. 3.

**The distance between persistence diagrams** has been well studied. One can treat two diagrams as two point sets on a 2D plane and measure their  $p$ -Wasserstein distance. This distance is well-behaved [12, 13].

<sup>1</sup> The persistence diagram definition does not require the input to be a distance transform. It can be an arbitrary scalar function defined on a topological space.



**Fig. 3.** Our topology-processing component. The input is a batch of real masks and synthetic masks. Each real or synthetic mask goes through the distance transform and persistent homology computation. We get its persistent diagram, a set of 2D points. We compare the two diagrams using the 1-Wasserstein distance only on birth times. The loss is defined as the matching distance between the two sets of diagrams (synthetic and real), computed using optimal transport.

In this paper, we use a modified version of the classic  $p$ -Wasserstein distance between diagrams. In particular, we only focus on the birth time, and drop their death time. The reason is that we are mainly focusing on the gaps one needs to close to complete an almost-hole (depending on birth time) and not particularly concerned with the size of the hole (corresponding to death time). Formally, we project all points of the two diagrams to the birth axis and compute their 1-Wasserstein distance, i.e., the optimal matching distance between the two point sets within the birth axis, as illustrated in Fig. 3. We note that points in the diagrams of the synthetic and real images are mostly paired with nearby points. The only exception is the red point corresponding to the almost-hole. The matching distance essentially measures the gap of the almost-hole. The diagram distance measures how easy it is to fix the synthetic image so it has the same number of holes as the real one. Formally, the distance between two diagrams  $\text{dgm}_1$  and  $\text{dgm}_2$  is

$$\mathcal{W}_1(\text{dgm}_1, \text{dgm}_2) = \min_{\sigma \in \Sigma} \sum_{x \in \text{dgm}_1} |b_x - b_{\sigma(x)}| = \sum_{x \in \text{dgm}_1} |b_x - b_{\sigma^*(x)}|, \quad (3)$$

in which  $\Sigma$  is the set of all possible one-to-one correspondences between the two diagrams, and  $\sigma^*$  is the optimal matching one can choose. Here  $b_x$  denotes the birth time of a point  $x$  in  $\text{dgm}_1$ . Similarly,  $b_{\sigma(x)}$  and  $b_{\sigma^*(x)}$  are the birth times of  $x$ 's match  $\sigma(x)$  and optimal match  $\sigma^*(x)$  in  $\text{dgm}_2$ . The matching may not exist when there are different numbers of points from the two diagrams. To this end, we can add infinitely many points to the diagonal line ( $b = d$ ) so the unmatched points can be matched to the diagonal line.<sup>2</sup> In practice, our

<sup>2</sup> There are more technical reasons for adding the diagonal line into the diagram, related to the stability of the metric. See [12].

algorithm for matching computation is very similar to the sliced Wasserstein distance [6] for persistence diagrams, except that we only use one of the infinitely many slices, i.e.,  $d = 0$ .

**Topological GAN loss defined via matching persistence diagrams.** Next, we define our loss, which measures the difference between two diagram distributions. The loss should be (1) simple to compute; and (2) efficient in matching the two distributions. Due to these constraints, it is not straightforward to use other approaches such as the kernel mean embedding (which is used in Sec. 3.4 to define GAN metrics). See supplemental material for discussion.

We propose a loss that is easy to compute and can be efficiently optimized. We find a pairwise matching between synthetic and real diagrams and sum up the diagram distance between all matched pairs as the loss. Let  $\mathcal{D}_{syn}$  and  $\mathcal{D}_{real}$  be the two sets of persistence diagrams generated from synthetic and real images. Suppose we have an optimal matching between the two diagram sets,  $\pi^*$ . Our loss is the total matching distance between all matched synthetic-real diagram pairs. Recall  $\mathcal{W}_1$  is the diagram distance (Eq. (3)). We have

$$L_{topo} = \sum_{\text{dgm}_i \in \mathcal{D}_{syn}} \mathcal{W}_1(\text{dgm}_i, \pi^*(\text{dgm}_i)). \quad (4)$$

To find the optimal matching  $\pi^*$  between synthetic and real diagram sets, we use the optimal transport technique. Denote  $\text{dgm}_i^s \in \mathcal{D}_{syn}$  and  $\text{dgm}_j^r \in \mathcal{D}_{real}$ . Let  $n_{syn}$  and  $n_{real}$  be the size of  $\mathcal{D}_{syn}$  and  $\mathcal{D}_{real}$ . We solve Monge-Kantorovich’s primal problem [62] to find the optimal transport plan:

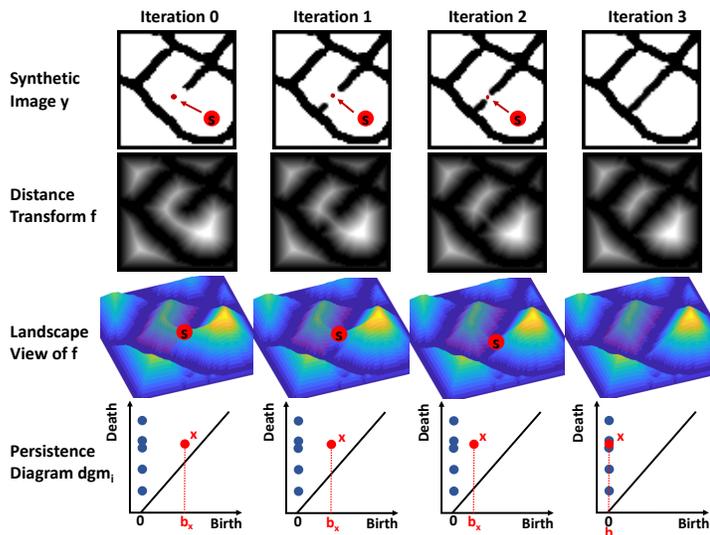
$$\gamma^* = \min_{\gamma \in \Gamma} \sum_{i=1}^{n_{syn}} \sum_{j=1}^{n_{real}} \mathcal{W}_1(\text{dgm}_i^s, \text{dgm}_j^r) \cdot \gamma_{ij} \quad (5)$$

where  $\Gamma = \{\gamma \in \mathbb{R}_+^{n_{syn} \times n_{real}} \mid \gamma \mathbf{1}_{n_{real}} = 1/n_{syn} \cdot \mathbf{1}_{n_{syn}}, \gamma^\top \mathbf{1}_{n_{syn}} = 1/n_{real} \cdot \mathbf{1}_{n_{real}}\}$ .  $\mathbf{1}_n$  is an  $n$ -dimensional vector of all ones. Denote by  $\gamma^*$  the optimal solution to Eq. (5). We compute the optimal matching ( $\pi^*$ ) by mapping the  $i$ -th synthetic  $\text{dgm}_i^s$  to the best matched real diagram w.r.t. the optimal transportation plan, i.e.,  $\text{dgm}_{h(i)}^r$  such that  $h(i) = \arg \max_j \gamma_{ij}^*$  [49]. Formally,  $\pi^*(\text{dgm}_i^s) = \text{dgm}_{h(i)}^r$ .

### 3.3 Gradient of the Loss

We derive the gradient of the topological GAN loss (Eq. (4)). The loss can be decomposed into the sum of the loss terms for individual synthetic diagrams,  $L_{topo} = \sum_i L_{topo}^i$ , in which the  $i$ -th loss term  $L_{topo}^i = \mathcal{W}_1(\text{dgm}_i, \pi^*(\text{dgm}_i))$ . Here the  $i$ -th synthetic diagram is generated from the distance transform of the  $i$ -th synthetic mask,  $y_i$ ,  $\text{dgm}_i = \text{dgm}(f_{y_i})$ . Meanwhile,  $y_i$  is a binary mask computed by thresholding the generated image  $G(z_i)$ . It suffices to calculate the gradient of  $L_{topo}^i$  with regard to the generator  $G$ .

Before deriving the gradient, we illustrate the intuition of the gradient descent in Fig. 4. For a particular synthetic image, we show how the mask is modified at different iterations and how the persistence diagram changes accordingly.



**Fig. 4.** From top to bottom: the same synthetic image being fixed at different iterations, their distance transforms, landscape views of distance transforms, and persistence diagrams. The red marker  $s$  is the saddle point, whose function value is the birth time of the almost-hole  $x$ .

As the gradient descent continues, the almost-hole in the synthetic image is slowly closed up to form a complete hole. At each iteration, the mask  $y$  grows toward the saddle point  $s$  of the distance transform  $f$ . The distance transform function value at the saddle  $s$ ,  $f(s)$ , decreases toward zero. In the persistence diagram, the corresponding dot,  $x$ , moves toward left because its birth time  $b_x = f(s)$  decreases. This reduces the 1-Wasserstein distance between the synthetic diagram and its matched real diagram.

Formally, by chain rule, we have  $\frac{\partial L_{topo}^i}{\partial G} = \frac{\partial L_{topo}^i}{\partial \text{dgm}_i} \cdot \frac{\partial \text{dgm}_i}{\partial f_{y_i}} \cdot \frac{\partial f_{y_i}}{\partial G(z_i)} \cdot \frac{\partial G(z_i)}{\partial G}$ . Next, we calculate each of the multiplicands.

**Derivative of the loss w.r.t. persistence diagrams.** Recall that by Eqs. (3) and (4), we can rewrite the  $i$ -th loss term as  $L_{topo}^i = \sum_{x \in \text{dgm}_i} |b_x - b_{\sigma^*(x)}| = \sum_{x \in \text{dgm}_i} \text{sign}(b_x - b_{\sigma^*(x)})(b_x - b_{\sigma^*(x)})$ . The equation depends on two optimal matchings,  $\pi^*$  and  $\sigma^*$ . The first one,  $\pi^* : \mathcal{D}_{syn} \rightarrow \mathcal{D}_{real}$ , is calculated by optimal transport between two sets of diagrams,  $\mathcal{D}_{syn}$  and  $\mathcal{D}_{real}$ . The second optimal matching,  $\sigma^* : \text{dgm}_i \rightarrow \pi^*(\text{dgm}_i)$ , is calculated by 1D optimal transport between points of the two matched diagrams. Without loss of generality, we assume for all  $x \in \text{dgm}_i$  and  $x' \in \pi^*(\text{dgm}_i)$ , their birth time differences  $(b_x - b_{x'})$ 's are distinct nonzero values.

While the optimal transport plan ( $\gamma^*$  in Eq. (5)) changes continuously as we change the input synthetic diagrams, the matchings  $\pi^*$  and  $\sigma^*$  only change at singularities (a measure-zero set). Within a small neighborhood of the input, we can assume constant optimal mappings  $\pi^*$  and  $\sigma^*$ , and constant  $\text{sign}(b_x - b_{\sigma^*(x)})$

and  $b_{\sigma^*(x)}$  as well. The gradient can be formally written as the partial derivative of the loss with regard to the birth and death times of each point  $x \in \text{dgm}_i$ :

$$\frac{\partial L_{\text{topo}}^i}{\partial b_x} = \text{sign}(b_x - b_{\sigma^*(x)}), \quad \frac{\partial L_{\text{topo}}^i}{\partial d_x} = 0.$$

Intuitively, the negative gradient direction  $-\frac{\partial L_{\text{topo}}^i}{\partial \text{dgm}_i} \frac{\partial \text{dgm}_i}{\partial G}$  moves each point  $x$  in the synthetic diagram  $\text{dgm}_i$  toward its matched point in the matched real diagram,  $\sigma^*(x)$ , horizontally (but not vertically). See Fig. 4 for an illustration.

**Derivative of the persistence diagram w.r.t. the distance transform.**

The derivative of the loss w.r.t. death time is zero. Therefore, we only need to care about the derivative of the birth time  $b_x$  w.r.t. the distance transform  $f_{y_i}$ ,  $\frac{\partial b_x}{\partial f_{y_i}}$ . An important observation is that the birth time of any almost-hole in a filtration is the function value of the saddle point of  $f_{y_i}$  sitting right in the middle of the gap, denoted as  $s_x$ . Formally,  $b_x = \langle \delta_{s_x}, f_{y_i} \rangle$ , in which  $\delta_{s_x}$  is the Dirac delta function at the saddle point  $s_x$ . Taking the gradient, we have  $\frac{\partial b_x}{\partial f_{y_i}} = \delta_{s_x}$ .

Intuitively,  $-\frac{\partial L_{\text{topo}}^i}{\partial b_x} \frac{\partial b_x}{\partial f_{y_i}} \frac{\partial f_{y_i}}{\partial G}$ , the negative gradient w.r.t. the  $b_x$  of the diagram, moves the saddle point function value  $b_x = f_{y_i}(s_x)$  up or down so it gets closer to the matched real diagram point’s birth time,  $b_{\sigma^*(x)}$ . See Fig. 4.

**Derivative of the distance transform w.r.t. the synthetic image  $G(z_i)$ .**

Finally, we compute the derivative of  $f_{y_i}$  with regard to the  $i$ -th synthetic image  $G(z_i)$ . Intuitively, focusing on the saddle point  $s_x$ , to increase or decrease its distance transform  $f_{y_i}(s_x)$ , the gradient needs to grow the mask  $y_i$  at its nearest boundary point to  $s_x$ , called  $r$ . This is achieved by changing the synthetic image values of the few pixels near  $r$ . As seen in Fig. 4, as we proceed, the mask grows toward the saddle point. More derivation details are in supplemental material.

### 3.4 Topology-Aware Metrics for GAN Evaluation

We introduce two novel metrics that can evaluate GAN performance in terms of topology. Conventionally, generator quality has been evaluated by comparing synthetic and real image distributions in the space of CNN-based image features. For example, both the Inception score (IS) [56] and the Fréchet Inception distance (FID) [24] use an Inception network pre-trained on ImageNet to map images into a feature space. The topological properties of images are not guaranteed to be preserved in such CNN-based image feature space.

In this paper, for the first time, we propose metrics that directly measure the topological difference between synthetic and real image distributions. The first metric, called the *Betti score*, is directly based on the topology of the mask, measured by the Betti number. Recall the Betti number counts the number of holes in a given synthetic or real mask. A Betti score computes a histogram for all synthetic masks and another histogram for all real masks. Then it compares the two histograms using their  $\chi^2$  distance. The definition can easily extend to zero-dimensional topology, i.e., counting the number of connected components.

Our second score is based on persistence diagrams which account for both holes and almost holes. We use the kernel mean embedding method [21]. As-

sume a given kernel for persistence diagrams, we can define an implicit function,  $\Phi$ , mapping all synthetic/real persistence diagrams into a Hilbert space,  $\mathcal{H}$ . In such space, it becomes easy to compute the mean of each diagram set,  $\Phi(\mathcal{D}_{syn}) := \frac{1}{n_{syn}} \sum_{i=1}^{n_{syn}} \Phi(\text{dgm}_i^s)$  and  $\Phi(\mathcal{D}_{real}) := \frac{1}{n_{real}} \sum_{i=1}^{n_{real}} \Phi(\text{dgm}_i^r)$ . We measure the difference between the synthetic and real diagram sample sets using *maximum mean discrepancy (MMD)*,

$$\text{MMD}(\mathcal{D}_{syn}, \mathcal{D}_{real}) := \|\Phi(\mathcal{D}_{syn}) - \Phi(\mathcal{D}_{real})\|_{\mathcal{H}}.$$

It was proven that this sample-based MMD will converge to its continuous analog. We propose to use the unbiased MMD [21] (details are in the supplemental material). In terms of the kernel for persistence diagrams, there are many options [33, 9]. Here we use the Gaussian kernel based on the 1-Wasserstein distance between diagrams,  $k_{\mathcal{W}_1}(\text{dgm}_i, \text{dgm}_j) = \exp\left(-\frac{\mathcal{W}_1(\text{dgm}_i, \text{dgm}_j)}{\sigma^2}\right)$ .

Our two metrics are generally useful to evaluate GAN results w.r.t. topology. We will evaluate TopoGAN using FID, unbiased MMD and Betti score.

## 4 Experiments

TopoGAN is built on top of WGAN-GP with deep convolutional generative adversarial networks [52] (DCGANs) as backbone network architectures. Details of TopoGAN’s implementation, training, and computation cost are in Sec. B of the supplemental material. We compare TopoGAN against two baseline GANs: Wasserstein GAN with gradient penalty (WGAN-GP) and Wasserstein GAN with Spectral Normalization (WGAN-SN). These methods are best known for stabilizing GAN training and avoiding mode collapse. To demonstrate the potential of TopoGAN in practice, we showcase it in a downstream task: segmentation.

**Datasets.** TopoGAN is evaluated on five datasets: **CREMI** [15], **ISBI12** [4], **Google Maps** scraped by [30], **CMP Facade Database** [61], and **Retina** dataset. The first two are neuron image segmentation datasets and we randomly sample 7500 and 1500 patches of size  $64 \times 64$  respectively from their segmentation masks. **Google Maps** (aerial photos  $\leftrightarrow$  maps) and **CMP Facade Database** (facades  $\leftrightarrow$  labels) consist of paired RGB images. The RGB images of maps and labels are converted into grayscale images. We extract 4915 patches of size  $64 \times 64$  from the converted maps and resize all 606 facade labels to  $128 \times 128$ . The **Retina** dataset consists of 98 retina segmentations we collected from 4 datasets: **IOSTAR** (40) [2, 1], **DRIVE** (20) [59], **STARE** (20) [27], and **CHASE\_DB1** (28) [48]. All retina images are cropped and resized to  $128 \times 128$  resolution.

**Quantitative and qualitative results.** In Table 1, we report the performance of TopoGAN and two baselines w.r.t. three metrics: FID, unbiased MMD, and Betti score. TopoGAN outperforms the two baselines significantly in the two topology-aware metrics proposed in Sec. 3.4: unbiased MMD and Betti score. The superior performance of TopoGAN proves that the topological GAN loss successfully enforced the structural/topological faithfulness of the generated images, as desired. Further comparisons of the topological quality of the synthesized

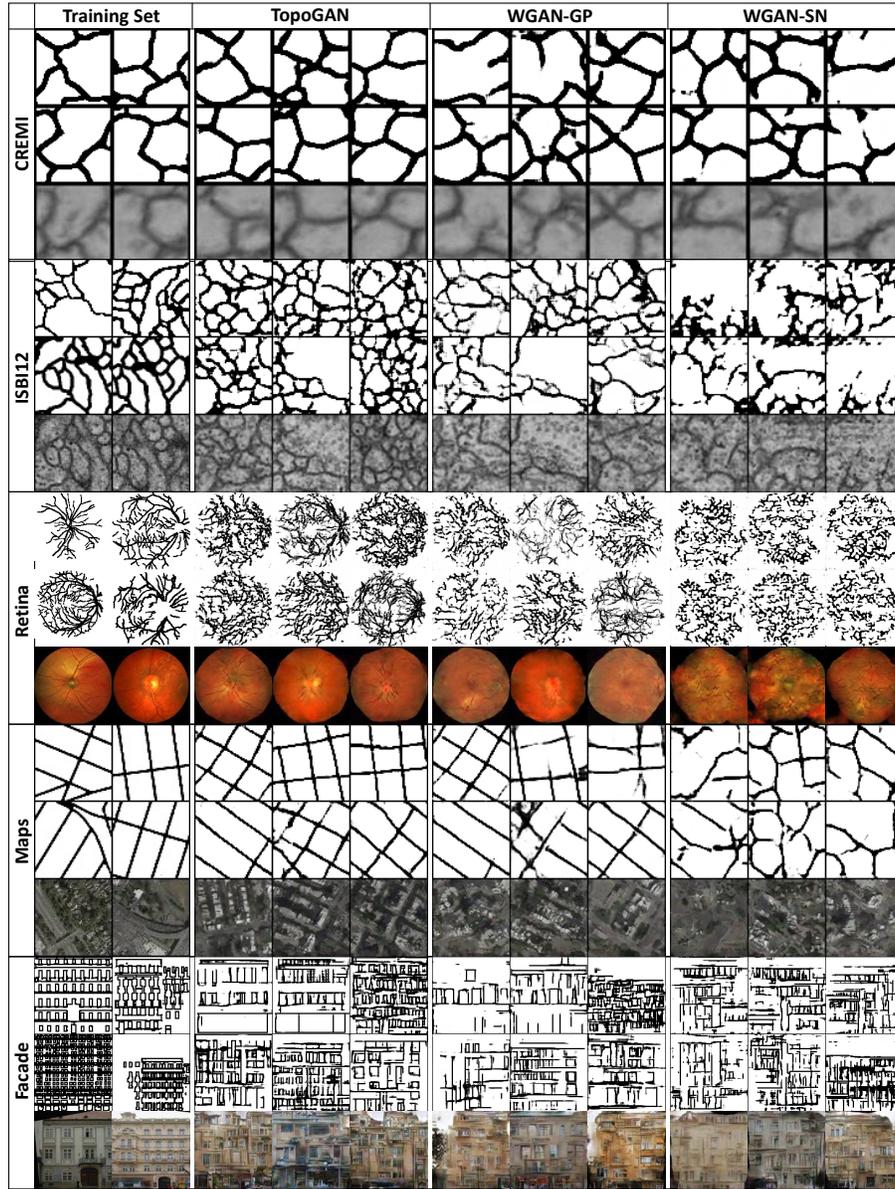
**Table 1.** Comparisons against baseline GANs on FID, unbiased MMD, and Betti score across five datasets. The standard deviations are based on 3 runs. We omit reporting unbiased MMD and Betti score of WGAN-SN on Retina as WGAN-SN fails to produce reasonable results.

	CREMI	ISBI12	Retina	Maps	Facade
FID					
WGAN-GP	21.64±0.138	83.90±0.718	179.69±19.008	72.00±0.469	122.13±0.822
WGAN-SN	34.15±0.153	78.61±0.411	269.12±2.276	175.52±0.217	126.10±1.901
TopoGAN	<b>20.96±0.195</b>	<b>31.90±0.248</b>	<b>169.21±21.976</b>	<b>60.48±0.467</b>	<b>119.11±0.874</b>
unbiased MMD					
WGAN-GP	0.142±0.014	0.558±0.010	1.735±0.050	0.482±0.007	0.137±0.004
WGAN-SN	0.326±0.016	0.602±0.006	-	0.724±0.005	0.166±0.005
TopoGAN	<b>0.134±0.019</b>	<b>0.405±0.003</b>	<b>1.602±0.114</b>	<b>0.471±0.010</b>	<b>0.080±0.002</b>
Betti score					
WGAN-GP	0.236±0.003	0.908±0.104	0.541±0.188	0.223±0.010	0.176±0.006
WGAN-SN	0.125±0.002	1.775±0.039	-	0.255±0.020	0.142±0.017
TopoGAN	<b>0.015±0.001</b>	<b>0.802±0.058</b>	<b>0.457±0.144</b>	<b>0.177±0.004</b>	<b>0.124±0.002</b>

images at different training epochs can be found in the supplemental material. Meanwhile, we observe that TopoGAN is also better in FID. This suggests that topological integrity could serve as an important visual cue when deciding image quality by human standards.

Qualitative results are in Fig. 5. For fair comparison, we use the same set of noise inputs to generate data for each GAN method. We observe that the masks produced by TopoGAN have more clear boundaries and complete cycles. They are topologically more similar to the real data (i.e., having similar Betti numbers). TopoGAN also shows better performance in texture images (details of how these textures are generated will be explained later). On the contrary, baselines WGAN-GP and WGAN-SN tend to generate broken structures. The **Retina** dataset is challenging for all GAN models. This is due to the small training set (98) and the heterogeneity of the dataset; its images are from multiple datasets with different geometry, resolutions, aspect ratios, and contrasts.

**Segmentation application.** We demonstrate that TopoGAN improves performance in a downstream binary segmentation task. For each dataset, we train a segmentation network with real training data, synthetic data, and real data augmented with synthetic data. The networks trained with synthetic data from TopoGAN are compared against networks trained with data from baseline GANs and with real training data. The segmentation networks are evaluated on test data with three segmentation metrics: (1) pixel accuracy, (2) Dice score, and (3) Adapted Rand Index (ARI). We report the results on dice score in Table 2, and leave the results on other scores to the supplemental material.



**Fig. 5.** Qualitative comparisons of TopoGAN to WGAN-GP and WGAN-SN on 5 datasets. From left to right: real masks from training set, generated masks from TopoGAN, WGAN-GP, and WGAN-SN. For each dataset, the third row shows texture images corresponding to the masks on the second row.

To produce synthetic pairs (fake masks  $\leftrightarrow$  textured masks), a pix2pix [30] network is first trained with real data pairs. The trained pix2pix network takes

**Table 2.** Dice score of segmentation networks on real test data. For each dataset, we train a total of 21 segmentation networks with real training data, synthetic data from TopoGAN and two baselines, and real data augmented with synthetic data. We report mean and standard deviation of a 3-fold cross validation.

	CREMI	ISBI12	Retina
Real data	0.896±0.004	0.932±0.011	0.883±0.010
WGAN-GP	0.820±0.018	0.927±0.005	0.891±0.012
WGAN-SN	0.827±0.019	0.902±0.008	-
TopoGAN	0.851±0.011	0.933±0.006	0.892±0.013
WGAN-GP+real data	0.897±0.008	0.943±0.007	0.899±0.010
WGAN-SN+real data	0.900±0.004	0.905±0.054	-
TopoGAN+real data	<b>0.902±0.006</b>	<b>0.944±0.008</b>	<b>0.906±0.014</b>

as inputs the GAN-generated masks and produces textured masks on which a segmentation network can be trained on. We use U-Net [55] as our segmentation network. We use a three-fold cross validation and report both the mean and standard deviation of the Dice score for all datasets. Note that only we only segment **CREMI**, **ISBI12**, and **Retina**, as the other two datasets are not segmentation datasets and have no ground truth training data.

Segmentation results are summarized in Table. 2. TopoGAN with pure synthetic data achieves comparable results to segmentation networks trained with real data on dataset **ISBI12** and **Retina**. Segmentations augmented with synthetic data always perform better than real data or synthetic data alone. TopoGAN plus real data produces the best results followed closely by WGAN-GP plus real data. Details of evaluation metrics, segmentation networks training procedure and full result table can be found in supplementary.

## 5 Conclusion

This paper proposed TopoGAN, the first GAN method explicitly learning image topology of the image from real data. We proposed a topological GAN loss and showed that this loss is differentiable and can be easily incorporated into GAN training. In addition, we proposed novel metrics to measure topological differences between synthesized and real images. Empirically, we have shown that TopoGAN generates images with better topological features than state-of-the-art GANs both quantitatively and qualitatively.

## Acknowledgement

Fan Wang and Chao Chen’s research was partially supported by NSF IIS-1855759, CCF-1855760 and IIS-1909038. Huidong Liu and Dimitris Samaras were partially supported from the Partner University Fund, the SUNY2020 Infrastructure Transportation Security Center, and a gift from Adobe.

## References

1. Abbasi-Sureshjani, S., Smit-Ockeloen, I., Bekkers, E., Dashtbozorg, B., t. H. Romeny, B.: Automatic detection of vascular bifurcations and crossings in retinal images using orientation scores. In: 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI). pp. 189–192 (April 2016). <https://doi.org/10.1109/ISBI.2016.7493241>
2. Abbasi-Sureshjani, S., Smit-Ockeloen, I., Zhang, J., Ter Haar Romeny, B.: Biologically-inspired supervised vasculature segmentation in slo retinal fundus images. In: Kamel, M., Campilho, A. (eds.) *Image Analysis and Recognition*. pp. 325–334. Springer International Publishing, Cham (2015)
3. Adams, H., Emerson, T., Kirby, M., Neville, R., Peterson, C., Shipman, P., Chepushtanova, S., Hanson, E., Motta, F., Ziegelmeier, L.: Persistence images: A stable vector representation of persistent homology. *The Journal of Machine Learning Research* **18**(1), 218–252 (2017)
4. Arganda-Carreras, I., Turaga, S.C., Berger, D.R., Cireşan, D., Giusti, A., Gambardella, L.M., Schmidhuber, J., Laptev, D., Dwivedi, S., Buhmann, J.M., Liu, T., Seyedhosseini, M., Tasdizen, T., Kamentsky, L., Burget, R., Uher, V., Tan, X., Sun, C., Pham, T.D., Bas, E., Uzunbas, M.G., Cardona, A., Schindelin, J., Seung, H.S.: Crowdsourcing the creation of image segmentation algorithms for connectomics. *Frontiers in Neuroanatomy* **9**, 142 (2015). <https://doi.org/10.3389/fnana.2015.00142>, <https://www.frontiersin.org/article/10.3389/fnana.2015.00142>
5. Arjovsky, M., Chintala, S., Bottou, L.: Wasserstein generative adversarial networks. In: *International Conference on Machine Learning*. pp. 214–223 (2017)
6. Bonneel, N., Rabin, J., Peyré, G., Pfister, H.: Sliced and radon wasserstein barycenters of measures. *Journal of Mathematical Imaging and Vision* **51**(1), 22–45 (2015)
7. Brock, A., Donahue, J., Simonyan, K.: Large scale gan training for high fidelity natural image synthesis. In: *International Conference on Learning Representations* (2018)
8. Brüel-Gabrielsson, R., Nelson, B.J., Dwaraknath, A., Skraba, P., Guibas, L.J., Carlsson, G.: A topology layer for machine learning. arXiv preprint arXiv:1905.12200 (2019)
9. Carriere, M., Cuturi, M., Oudot, S.: Sliced wasserstein kernel for persistence diagrams. In: *Proceedings of the 34th International Conference on Machine Learning—Volume 70*. pp. 664–673. JMLR. org (2017)
10. Chen, C., Ni, X., Bai, Q., Wang, Y.: A topological regularizer for classifiers via persistent homology. In: *The 22nd International Conference on Artificial Intelligence and Statistics*. pp. 2573–2582 (2019)
11. Clough, J.R., Oksuz, I., Byrne, N., Schnabel, J.A., King, A.P.: Explicit topological priors for deep-learning based image segmentation using persistent homology. In: *International Conference on Information Processing in Medical Imaging*. pp. 16–28. Springer (2019)
12. Cohen-Steiner, D., Edelsbrunner, H., Harer, J.: Stability of persistence diagrams. *Discrete & Computational Geometry* **37**(1), 103–120 (2007)
13. Cohen-Steiner, D., Edelsbrunner, H., Harer, J., Mileyko, Y.: Lipschitz functions have  $l$  p-stable persistence. *Foundations of computational mathematics* **10**(2), 127–139 (2010)
14. Cortes, C., Vapnik, V.: Support-vector networks. *Machine learning* **20**(3), 273–297 (1995)

15. Miccai challenge on circuit reconstruction from electron microscopy images, <https://cremi.org/>
16. Edelsbrunner, H., Harer, J.: Computational topology: an introduction. American Mathematical Soc. (2010)
17. Fabbri, R., Costa, L.D.F., Torelli, J.C., Bruno, O.M.: 2d euclidean distance transform algorithms: A comparative survey. *ACM Computing Surveys (CSUR)* **40**(1), 1–44 (2008)
18. Fu, H., Gong, M., Wang, C., Batmanghelich, K., Zhang, K., Tao, D.: Geometry-consistent generative adversarial networks for one-sided unsupervised domain mapping. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 2427–2436 (2019)
19. Ghafoorian, M., Nugteren, C., Baka, N., Booi, O., Hofmann, M.: El-gan: Embedding loss driven generative adversarial networks for lane detection. In: *European Conference on Computer Vision*. pp. 256–272. Springer (2018)
20. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. In: *Advances in neural information processing systems*. pp. 2672–2680 (2014)
21. Gretton, A., Borgwardt, K., Rasch, M., Schölkopf, B., Smola, A.: A kernel two-sample test. *Journal of Machine Learning Research* **13**, 723–773 (Mar 2012)
22. Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., Courville, A.C.: Improved training of wasserstein gans. In: *Advances in Neural Information Processing Systems*. pp. 5767–5777 (2017)
23. He, Y., Carass, A., Liu, Y., Jedynek, B.M., Solomon, S.D., Saidha, S., Calabresi, P.A., Prince, J.L.: Fully convolutional boundary regression for retina oct segmentation. In: Shen, D., Liu, T., Peters, T.M., Staib, L.H., Essert, C., Zhou, S., Yap, P.T., Khan, A. (eds.) *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*. pp. 120–128. Springer International Publishing (2019)
24. 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. In: Guyon, I., Luxburg, U.V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R. (eds.) *Advances in Neural Information Processing Systems 30*. pp. 6626–6637. Curran Associates, Inc. (2017), <http://papers.nips.cc/paper/7240-gans-trained-by-a-two-time-scale-update-rule-converge-to-a-local-nash-equilibrium.pdf>
25. Hofer, C., Kwitt, R., Niethammer, M., Dixit, M.: Connectivity-optimized representation learning via persistent homology. In: *International Conference on Machine Learning*. pp. 2751–2760 (2019)
26. Hofer, C., Kwitt, R., Niethammer, M., Uhl, A.: Deep learning with topological signatures. In: *Advances in Neural Information Processing Systems*. pp. 1634–1644 (2017)
27. Hoover, A.D., Kouznetsova, V., Goldbaum, M.: Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Transactions on Medical Imaging* **19**(3), 203–210 (March 2000). <https://doi.org/10.1109/42.845178>
28. Hu, X., Li, F., Samaras, D., Chen, C.: Topology-preserving deep image segmentation. In: *Advances in Neural Information Processing Systems*. pp. 5658–5669 (2019)
29. Hwang, J.J., Ke, T.W., Shi, J., Yu, S.X.: Adversarial structure matching for structured prediction tasks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 4056–4065 (2019)
30. Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. *CVPR* (2017)

31. Khrulkov, V., Oseledets, I.: Geometry score: A method for comparing generative adversarial networks. In: Dy, J., Krause, A. (eds.) *Proceedings of the 35th International Conference on Machine Learning*. *Proceedings of Machine Learning Research*, vol. 80, pp. 2621–2629. PMLR, Stockholm, Sweden (10–15 Jul 2018)
32. Kossaiji, J., Tran, L., Panagakis, Y., Pantic, M.: Gagan: Geometry-aware generative adversarial networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 878–887 (2018)
33. Kusano, G., Hiraoka, Y., Fukumizu, K.: Persistence weighted gaussian kernel for topological data analysis. In: *International Conference on Machine Learning*. pp. 2004–2013 (2016)
34. Li, C.L., Chang, W.C., Cheng, Y., Yang, Y., Póczos, B.: Mmd gan: Towards deeper understanding of moment matching network. In: *Advances in Neural Information Processing Systems*. pp. 2203–2213 (2017)
35. Lim, J.H., Ye, J.C.: Geometric gan. arXiv preprint arXiv:1705.02894 (2017)
36. Liu, H., Gu, X., Samaras, D.: Wasserstein gan with quadratic transport cost. In: *The IEEE International Conference on Computer Vision (ICCV)* (October 2019)
37. Liu, H., Xianfeng, G., Samaras, D.: A two-step computation of the exact gan wasserstein distance. In: *International Conference on Machine Learning*. pp. 3165–3174 (2018)
38. Liu, S., Bousquet, O., Chaudhuri, K.: Approximation and convergence properties of generative adversarial learning. In: *Advances in Neural Information Processing Systems*. pp. 5545–5553 (2017)
39. Luc, P., Couprie, C., Chintala, S., Verbeek, J.: Semantic segmentation using adversarial networks. arXiv preprint arXiv:1611.08408 (2016)
40. Mescheder, L., Nowozin, S., Geiger, A.: Which training methods for gans do actually converge? In: *International Conference on Machine Learning* (2018)
41. Miyato, T., Kataoka, T., Koyama, M., Yoshida, Y.: Spectral normalization for generative adversarial networks. In: *International Conference on Machine Learning* (2018)
42. Mosinska, A., Márquez-Neila, P., Koziński, M., Fua, P.: Beyond the pixel-wise loss for topology-aware delineation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2018)
43. Mroueh, Y., Sercu, T.: Fisher gan. In: *Advances in Neural Information Processing Systems*. pp. 2513–2523 (2017)
44. Mroueh, Y., Sercu, T., Goel, V.: McGAN: Mean and covariance feature matching GAN. In: Precup, D., Teh, Y.W. (eds.) *Proceedings of the 34th International Conference on Machine Learning*. *Proceedings of Machine Learning Research*, vol. 70, pp. 2527–2535. PMLR, International Convention Centre, Sydney, Australia (06–11 Aug 2017)
45. Munkres, J.R.: *Elements of algebraic topology*. CRC Press (2018)
46. Ni, X., Quadrianto, N., Wang, Y., Chen, C.: Composing tree graphical models with persistent homology features for clustering mixed-type data. In: *Proceedings of the 34th International Conference on Machine Learning-Volume 70*. pp. 2622–2631. JMLR. org (2017)
47. Nowozin, S., Cseke, B., Tomioka, R.: f-gan: Training generative neural samplers using variational divergence minimization. In: *Advances in neural information processing systems*. pp. 271–279 (2016)
48. Owen, C.G., Rudnicka, A.R., Mullen, R., Barman, S.A., Monekosso, D., Whincup, P.H., Ng, J., Paterson, C.: Measuring Retinal Vessel Tortuosity in 10-Year-Old

- Children: Validation of the Computer-Assisted Image Analysis of the Retina (CA-IAR) Program. *Investigative Ophthalmology and Visual Science* **50**(5), 2004–2010 (05 2009). <https://doi.org/10.1167/iovs.08-3018>, <https://doi.org/10.1167/iovs.08-3018>
49. Peyré, G., Cuturi, M.: Computational optimal transport foundations and trends in machine learning 11 (2019) 355 (1803)
  50. Qi, G.J., Zhang, L., Hu, H., Edraki, M., Wang, J., Hua, X.S.: Global versus localized generative adversarial nets. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1517–1525 (2018)
  51. Qin, Y., Chen, M., Zheng, H., Gu, Y., Shen, M., Yang, J., Huang, X., Zhu, Y.M., Yang, G.Z.: Airwaynet: A voxel-connectivity aware approach for accurate airway segmentation using convolutional neural networks. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 212–220. Springer (2019)
  52. Radford, A., Metz, L., Chintala, S.: Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434 (2015)
  53. Ramamurthy, K.N., Varshney, K., Mody, K.: Topological data analysis of decision boundaries with application to model selection. In: International Conference on Machine Learning. pp. 5351–5360 (2019)
  54. Reininghaus, J., Huber, S., Bauer, U., Kwitt, R.: A stable multi-scale kernel for topological machine learning. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4741–4748 (2015)
  55. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention. pp. 234–241. Springer (2015)
  56. Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., Chen, X.: Improved techniques for training gans. In: Advances in neural information processing systems. pp. 2234–2242 (2016)
  57. Shaham, T.R., Dekel, T., Michaeli, T.: Singan: Learning a generative model from a single natural image. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 4570–4580 (2019)
  58. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. In: International Conference on Learning Representations (2015)
  59. Staal, J., Abràmoff, M.D., Niemeijer, M., Viergever, M.A., Van Ginneken, B.: Ridge-based vessel segmentation in color images of the retina. *IEEE transactions on medical imaging* **23**(4), 501–509 (2004)
  60. Thanh-Tung, H., Tran, T., Venkatesh, S.: Improving generalization and stability of generative adversarial networks. In: International Conference on Learning Representations (2019)
  61. Tyleček, R., Šára, R.: Spatial pattern templates for recognition of objects with regular structure. In: German Conference on Pattern Recognition. pp. 364–374. Springer (2013)
  62. Villani, C.: Optimal transport: old and new, vol. 338. Springer Science & Business Media (2008)
  63. Wang, T.C., Liu, M.Y., Zhu, J.Y., Tao, A., Kautz, J., Catanzaro, B.: High-resolution image synthesis and semantic manipulation with conditional gans. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 8798–8807 (2018)

64. Wu, P., Chen, C., Wang, Y., Zhang, S., Yuan, C., Qian, Z., Metaxas, D., Axel, L.: Optimal topological cycles and their application in cardiac trabeculae restoration. In: International Conference on Information Processing in Medical Imaging. pp. 80–92. Springer (2017)
65. Zhang, H., Goodfellow, I., Metaxas, D., Odena, A.: Self-attention generative adversarial networks. In: International Conference on Machine Learning. pp. 7354–7363 (2019)
66. Zhao, Q., Wang, Y.: Learning metrics for persistence-based summaries and applications for graph classification. In: Advances in Neural Information Processing Systems. pp. 9859–9870 (2019)
67. Zhao, Q., Ye, Z., Chen, C., Wang, Y.: Persistence enhanced graph neural network. In: International Conference on Artificial Intelligence and Statistics. pp. 2896–2906 (2020)
68. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Computer Vision (ICCV), 2017 IEEE International Conference on (2017)