Medical Image Synthesis with Generative Adversarial Networks: A Survey

Yusuf Dalva[1], Bilkent University, ID: 21602867
Ayhan Okuyan[1], Bilkent University, ID: 21601531
name[dot]surname[at]bilkent.edu.tr
Bilkent University[1], Ankara, Turkey

I. Introduction

In an effort to both overcome the scarcity of available data and learning the effective representations of the data available, Generative Adversarial Networks (GANs)[1] provided an effective framework with their proven success among various domains . As an addition to encoder-decoder architectures to generate new data samples, GANs structured themselves in a way that contains two networks, named generator and discriminator, connected via a feedback mechanism. During training, the two parts continuously try to improve themselves until they rearch an optimal point, which is possible theoretically. With this framework, GANs showed their ability of learning feature representations of a certain data distribution, despite their flaws [2] [3]. With this expressive power of GANs, many different domains adapted this method. GANs, being widely used in many fields, have become popular in Medical Imaging due to its strength in learning high fidelity and complex data representations from a target distribution. Being highly configurable, GANs are widely used in almost all sub-domains in medical imaging. Improving imaging quality by denoising CT scans [4], running segmetation, ROI detection and classfication on whole-slide biopsy images [5], creating synthetic samples from existing samples by learning domain representations [6] or transforming image domains, e.g. from CT to MRI[7] are few of many. This report investigates the the research and motivation behind using GANs in Medical Image Synthesis.

II. Generative Adversarial Networks: An Idea From Game Theory

As mentioned [1], the overall GAN framework is composed of two networks Generator (G) and Discriminator (D) where the connection between these components are established via an adversarial loss. According to this principle, the generator tries to generate new data samples from a given input (can be a noise vector or another image sample) where discriminator tries to distinguish these samples from the ones that originally belong to the data distribution. The simplest form of the adversarial loss is as follows (using Binary-Cross Entropy Loss for classification):

$$L_{ADV} = \min_{D} \max_{G} \log D(x) + \log(1 - D(G(z))) \quad (1)$$

Where z is the input to the generator and x is the data sample from a given data distribution. This loss can be interpreted as a min-max game played by generator and discriminator where both of the networks tries to outperform one another [1]. With several studies, the sensitiveness of the generator to the inputted data has been proven [2], [8]. With this verification GANs shows their potential of learning the feature representation by preserving the variance in the generated distribution.

III. Achieving Variety and Consistency in Generated Samples

Stating that GANs provide a framework on capturing the distribution of data samples where $s \epsilon S$ where S is the sample space and s is an individual sample, gaining control on the samples generated is a big challenge that is inspected by several studies [2] [9] [3] [10] [8]. The main issue here is to preserve both consistency and variety among the generated samples considering the desired environment which data generation is indented (desired state of nature for th given data). In this regard, the studies are branched into two subtopics, which in one side using state of nature information to generate context-aware samples (labeled samples) and another side attempting to control and understand the overall learning process in terms of variety and fidelity.

A. Environment-aware sample generation

In the original study for adversarial learning [1], the only information that encourages the generator to improve the samples generated is the discriminator feedback on whether the generated image is classified as fake or real. In order to generate samples that are context-aware, the corresponding class labels can be fed to the model. Here the approaches diverge into two. In one, we train separate GANs [2] while we provide the classes as a separate input in the other [9]. Originating from the second alternative, different feature maps are also attempted to be fed to the model both in general practice and in medical imaging domain. Using the context information, consistent domain translations are achieved by GANs where two-way transformations are identified by a trick named cycle-consistency loss [8]. By this way, feature representations for both $A \rightarrow B$ and $B \rightarrow A$ can successfully be learned by the model (where A and B are different domains). The corresponding component for the loss function which enables learning this two way translation [8] is as follows:

$$\begin{aligned} L_{CYCLE} = E_{x \epsilon p_{data}(x)} \| F(G(x)) - x \|_1 + \\ E_{y \epsilon p_{data}(y)} \| G(F(y)) - y \|_1 \end{aligned} \quad (2)$$

Where G is the model that performs the translation $A \rightarrow B$ and F is the model that performs the translation $B \rightarrow A$. In the studies where this loss component is applied, it is noted that the loss function both include L_{ADV} and L_{CYCLE} [8] [11]. This methodology both ensures that the generated samples are high-quality (not easily detected by discriminator) and consistent between each other (via L1 loss).
These findings contribute significantly to learning feature representations for different classes and improves image-to-image translation results overall.

B. Managing the Sensitivity of the architecture

Another research area that addresses the one of problems in the GAN framework is, the sensitivity of the method. In different studies, it is observed that training Generator and Discriminator is a sensitive process which discriminator can easily face the vanishing gradient problem [12], [3]. In order to enable GAN training regardless of the characteristics and scale of data, Lipschitz continuity is introduced to the loss function. Different than the trainability of the overall architecture, the variety and the consistency of the model has been inspected to utilize control on the model performance. Mode collapse is a common porblem identified with GANs and may effect the overall results [3] which is the the discriminator is over-optimized that generator produces output with limited variety in a rotating manner. In that regard, different weight regularization techniques are observed to hold the consistency among the generated samples while preserving the variance between resulting samples. Since this study aims to be a head start for attempting to use GANs in medical imaging domain and the details in this domain has crucial importance, more attention has been given to the stability of the model both in training and inference.

IV. Using GANs in Medical Imaging

There are many underlying motivations behind using GANs in Medical Imaging. The first and foremost one is that GANs ability to learn mappings between two domains, which enabled researchers to image one and then find its translation on the other. Since MRI is a method that induces more radiation than a CT scan, being able to convert one to the other, reduces the hypothetical radiation intake of the patient. Also, the fast evaluation process after learning how to classify regions, made the detection of ROIs in medical images and classifying them much faster than humans are capable of. This enabled the decrease the time of diagnosis given the images. Generally, GANs are used in five main sub-domains due to its configurable nature. Here we present all, but focus more on the synthesis of medical images.

A. Detection, Segmentation and Classification

There are many pathological findings that can be observed through medical imaging techniques, each targeting a different tissue structure or organ. Also, the imaging being anatomical or functional sets its purpose. In this rich environment of classification tasks, GANs are used to detect the ROIs in gigabyte-order images, and to segment various structures/organs in different domains, also to classify, binary/multi-class labels for ROIs or whole images/slides.

B. Data Augmentation

Since acquiring new samples in medical image domain is a costly process, GANs are widely used as a data augmentation technique. In the traditional data augmentation, there are certain geometrical transformations that can be applied to images (translation, rotation, scaling). With the use of GANs the augmented data gains a considerable amount of variety which tends to improve the overall performance in classification and detection tasks. From the perspective of medical imaging, augmenting different lesions with GANs is a performed practice which results in significant improvements in classifying liver lesions [13].

C. Medical Image Synthesis

Another field of that is mostly integrated with the use of GANs is **Medical Image Synthesis**. There are two technical motivations behind the this close relationship. One is that GANs are highly effective in learning the underlying mapping between two image domains that enables domain translation/adaptation. Also, by the same methods, GANs are used in the denoising of the noisy image domain for better diagnosis and segmentation.
The second most important motivation is the generative capability of GANs. Once learned the underlying distribution of the data, the network can generate synthetic examples that can't be distinguished even by pathologists, which holds exceptional value in the medical imaging domain where the data is limited and under strict conditions of release. In the upcoming sections, the improvements and state-of-the-art techniques in the field will be investigated throughly.

1) Improving Image Quality: As GANs learn the underlying representations of the respective image domains, there is research that tries to improve the research quality. For lung cancer diagnosis, a low-dose CT scan is preferred to the normal-dose CT scan since it reduces the patient to show radiation-induced lung cancer. However, this comes with reduced image quality since the reduction in the energy of x-ray photons, create a more blurred image. In [14], Wolterink et. al. trains and convolutional GAN, that uses CNN for both generator and discriminator to learn the domain specifics of the normal-dose CT images to reduce the noise in low-dose CT images by converting it to the target domain.
Furthermore, Tang et.al. approaches the same problem by using a cycle-consistent network [8]. This enables forward-backward consistency within the network, which is essential for unpaired training, meaning the sample pairs can be taken from different subjects and with different orientations. Also, this implementation comes with prior information regularization, which regularizes the weights with the L1 loss of the difference between the generator network and the prior information to ensure that the resulting images are closer to the input image in terms of structure. Furthermore, there are studies to remove artifacts such as motion artifacts from MRI scans [15].

Apart from these specific cases, GAN structure is also used to enhance image quality with the use of progressivaly-grown GANs (PGGAN),wherein each progressive training, one convolution and upsampling layer is added to both the generator and discriminator network, which enables the network to slowly learn the spatial relations regardless of size. Method first adapted by Beers et.al [16], resembles the inverse of a Gaussian pyramid structure where the convolutional parameters are not fixed but learned progressively.

2) Domain Adaptations and Translations: This section contains the translation of images from one domain to the other. The novel-paper by Nie et.al. [7], investigates this technique by extracting patches from MRI images and uses a context-aware network as described, where sequential networks are trained with the feature maps extracted from the previous iteration together with the original inputs. This method is used to translate from MRI to CT using paired images. Furthermore, the subsequent papers work on using unpaired data with the help of CycleGAN [8], or uses the same approach to translate from CT to PET scans using the class labels embedded on the input images [17]. The use of translations between different domains are not limited with different imaging techniques. The motivation that [8] provides is also used for translations in surgical images by changing different elements [11] as an augmentation technique. Furthermore, we are able to observe a general framework to run image-to image translations with MedGAN [18].

3) Learning Data Representations: While there are not many, some of the research approaches the question to all of the main tasks in medical imaging, including detection, segmentation and classification by learning the latent vector representations of the target tissue type in the underlying image domain. While this is a new and developing field of study, shows promise in learning the distribution of cancerous patches in haematoxylin and eosyn (H&E) stained breast cancer biopsy micro arrays as showed by Quiros et. al. [6]. The theoretical motivation is that, once the latent space mappings are learned, one can use it to either generate synthetic data, can even use vector operations on these vectors to generate reference images, or use the representations to classify/grade cancerous tissue microarrays.

References

[1] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative Adversarial Networks,” *arXiv:1406.2661*, 2014.

[2] A. Radford, L. Metz, and S. Chintala, “Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks,” *arXiv:1511.06434*, 2016.

[3] A. Brock, J. Donahue, and K. Simonyan, “Large scale GAN training for High Fidelity Natural Image Synthesis,” *arXiv:1809.11096v2*, 2019.

[4] J. Wolterink, T. Leiner, M. Viergever, and I. Išgum, “Generative Administrative Networks for Noise Reduction in Low-Dose CT,” *IEEE Transactions on Biomedical Engineering*, 2017.

[5] A. BenTaieb and G. Hamarneh, “Deep learning models for digital pathology,” *arXiv:1910.12329v2*, 2019.

[6] A. Quiros, R. Murray-Smith, and K. Yuan, “PathologyGAN: Learning deep representations of cancer tissue,” *arXiv:1611.07004v3*, 2018.

[7] D. Nie, R. Trullo, J. Lian, L. Wang, C. Petitjean, S. Ruan, Q. Wang, and D. Shen, “Medical Image Synthesis with Deep Convolutional Adversarial Networks,” *IEEE Transactions on Biomedical Engineering*, 2018.

[8] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks,” *arXiv:1703.10593*, 2017.

[9] M. Mirza and S. Osindero, “Conditional Generative Adversarial Nets,” *arXiv:1411.1784v1*, 2014.

[10] P. Isola, J.-Y. Zhu, T. Zhou, and Networks, “Image-to-Image Translation with Conditional Adversarial Networks,” *arXiv:1611.07004v3*, 2018.

[11] K. Lee, M.-K. Choi, and H. Jung, “DavinciGAN: Unpaired Surgical Instrument Translation for Data Augmentation,” *Proceedings of Machine Learning Research*, 2019.

[12] M. Arjovsky, S. Chintala, and L. Bottou, “Wasserstein GAN,” *arXiv:1701.07875v3*, 2017.

[13] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, and J. Goldberger, “GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification,” *Neurocomputing*, 2018.

[14] J. M. Wolterink, A. M. Dinkla, M. H. Savenije, P. R. Seevinck, C. van den Berg, and I. Isgum, “Deep MR to CT Synthesis using Unpaired Data,” *arXiv:1708.01155v1*, 2017.

[15] M. Ran, J. Hu, Y. Chen, H. Chen, H. Sun, J. Zhou, and Y. Zhang, “Denoising of 3D Magnetic Resonance Images Using a Residual Encoder-Decoder Wasserstein Generative Adversarial Network,” *arXiv:1808.03941*, 2018.

[16] A. Beers, J. Brown, K. Chang, J. Campbell, S. Ostmo, M. Chiang, and J. Kalpathy-Cramer, “High-resolution medical image synthesis using progressively grown generative adversarial networks,” *arXiv:1805.03144v2*, 2018.

[17] L. Bi, J. Kim, A. Kumar, D. Feng, and M. Fulham, “Synthesis of Positron Emission Tomography (PET) Images via Multi-channel Generative Adversarial Networks (GANs),” *arXiv:1411.1784v1*, 2014.

[18] K. Armanious, C. Jiang, M. Fischer, T. Küstner, K. Nikolaou, S. Gatidis, and B. Yang, “MedGAN: Medical Image Translation using GANs,” *arXiv:1806.06397*, 2019.