Generative Adversarial Network in Medical Imaging: A Review

Xin Yi[a,*], Ekta Walia[a,b], Paul Babyn[a]

[a]*Department of Medical Imaging, University of Saskatchewan, 103 Hospital Dr, Saskatoon, SK, S7N 0W8 Canada*
[b]*Philips Canada, 281 Hillmount Road, Markham, Ontario, ON L6C 2S3, Canada*

Abstract

Generative adversarial networks have gained a lot of attention in the computer vision community due to their capability of data generation without explicitly modelling the probability density function. The adversarial loss brought by the discriminator provides a clever way of incorporating unlabeled samples into training and imposing higher order consistency. This has proven to be useful in many cases, such as domain adaptation, data augmentation, and image-to-image translation. These properties have attracted researchers in the medical imaging community, and we have seen rapid adoption in many traditional and novel applications, such as image reconstruction, segmentation, detection, classification, and cross-modality synthesis. Based on our observations, this trend will continue and we therefore conducted a review of recent advances in medical imaging using the adversarial training scheme with the hope of benefiting researchers interested in this technique.

Keywords: Deep learning, Generative adversarial network, Generative model, Medical imaging, Review

1. Introduction

With the resurgence of deep learning in computer vision starting from 2012 (Krizhevsky et al., 2012), the adoption of deep learning methods in medical imaging has increased dramatically. It is estimated that there were over 400 papers published in 2016 and 2017 in major medical imaging related conference venues and journals (Litjens et al., 2017). The wide adoption of deep learning in the medical imaging community is due to its demonstrated potential to complement image interpretation and augment image representation and classification. In this article, we focus on one of the most interesting recent breakthroughs in the field of deep learning - generative adversarial networks (GANs) - and their potential applications in the field of medical imaging.

GANs are a special type of neural network model where two networks are trained simultaneously, with one focused on image generation and the other centered on discrimination. The adversarial training scheme has gained attention in both academia and industry due to its usefulness in counteracting domain shift, and effectiveness in generating new image samples. This model has achieved state-of-the-art performance in many image generation tasks, including text-to-image synthesis (Xu et al., 2017), super-resolution (Ledig et al., 2017), and image-to-image translation (Zhu et al., 2017a).

Unlike deep learning which has its roots traced back to the 1980s (Fukushima and Miyake, 1982), the concept of adversarial training is relatively new with significant recent progress (Goodfellow et al., 2014). This paper presents a general overview of GANs, describes their promising applications in medical imaging, and identifies some remaining challenges that need to be solved to enable their successful application in other medical imaging related tasks.

To present a comprehensive overview of all relevant works on GANs in medical imaging, we searched databases including PubMed, arXiv, proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), SPIE Medical Imaging, IEEE International Symposium on Biomedical Imaging (ISBI), and International conference on Medical Imaging with Deep Learning (MIDL). We also incorporated cross referenced works not identified in the above search process. Since there are research publications coming out every month, without losing generality, we set the cut off time of the search as January 1st, 2019. Works on arXiv that report only preliminary results are excluded from this review. Descriptive statistics of these papers based on task, imaging modality and year can be found in Figure 1.

The remainder of the paper is structured as follows. We begin with a brief introduction of the principles of GANs and some of its structural variants in Section 2. It is followed by a comprehensive review of medical image analysis tasks using GANs in Section 3 including but not limited to the fields of radiology, histopathology and dermatology. We categorize all the works according to canonical tasks: reconstruction, image synthesis, segmentation, classification, detection, registration, and others. Section 4 summarizes the review and discusses prospective applications and identifies open challenges.

*Corresponding author
Email addresses: xiy525@mail.usask.ca (Xin Yi), ewb178@mail.usask.ca (Ekta Walia), Paul.Babyn@saskhealthauthority.ca (Paul Babyn)

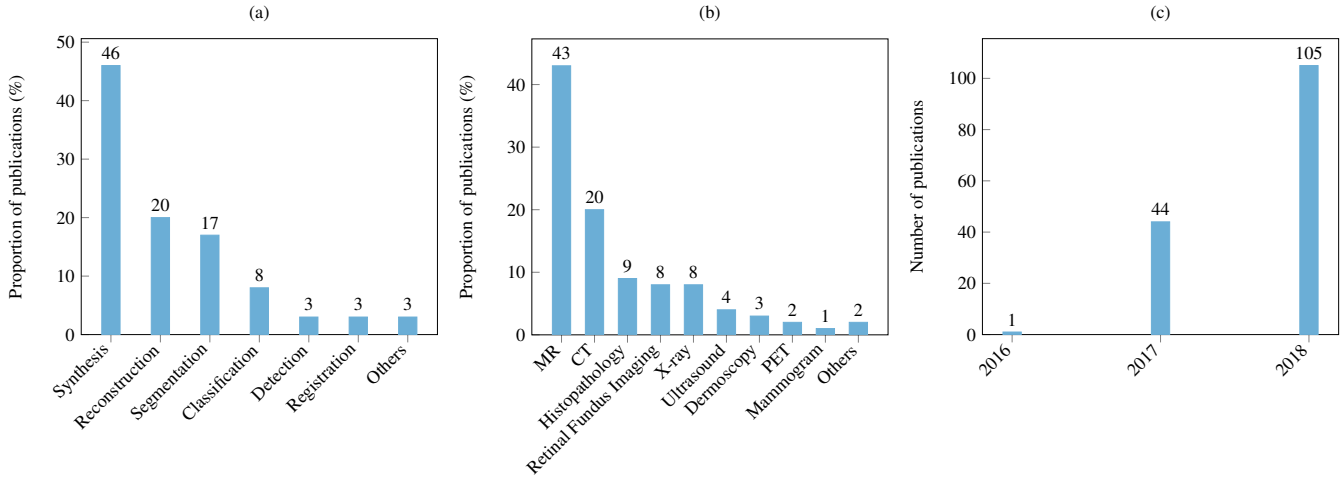

Figure 1: (a) Categorization of GAN related papers according to canonical tasks. (b) Categorization of GAN related papers according to imaging modality. (c) Number of GAN related papers published from 2014. Note that some works performed various tasks and conducted evaluation on datasets with different modalities. We counted these works multiple times in plotting these graphs. Works related to cross domain image transfer were counted based on the source domain. The statistics presented in figure (a) and (b) are based on papers published on or before January 1st, 2019.

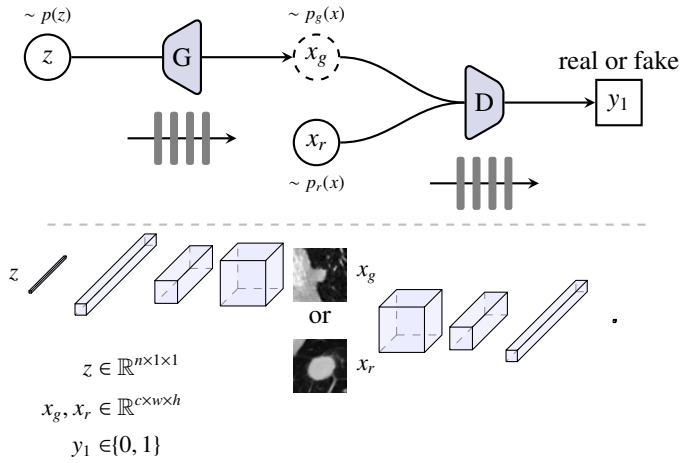

Figure 2: Schematic view of the vanilla GAN for synthesis of lung nodule on CT images. Top of the figure shows the network configuration. The part below shows the input, output and the internal feature representations of the generator G and discriminator D. G transforms a sample z from $p(z)$ into a generated nodule x_g. D is a binary classifier that differentiates the generated and real images of lung nodule formed by x_g and x_r respectively.

2. Background

2.1. Vanilla GAN

The vanilla GAN (Goodfellow et al., 2014) is a generative model that was designed for directly drawing samples from the desired data distribution without the need to explicitly model the underlying probability density function. It consists of two neural networks: the generator G and the discriminator D. The input to G, z is pure random noise sampled from a prior distribution $p(z)$, which is commonly chosen to be a Gaussian or a uniform distribution for simplicity. The output of G, x_g is expected to have visual similarity with the real sample x_r that is drawn from the real data distribution $p_r(x)$. We denote the non-linear mapping function learned by G parametrized by θ_g as $x_g = G(z; \theta_g)$. The input to D is either a real or generated sample. The output of D, y_1 is a single value indicating the probability of the input being a real or fake sample. The mapping learned by D parametrized by θ_d is denoted as $y_1 = D(x; \theta_d)$. The generated samples form a distribution $p_g(x)$ which is desired to be an approximation of $p_r(x)$ after successful training. The top of Figure 2 shows an illustration of a vanilla GAN's configuration. G in this example is generating a 2D CT slice depicting a lung nodule.

D's objective is to differentiate these two groups of images whereas the generator G is trained to confuse the discriminator D as much as possible. Intuitively, G could be viewed as a forger trying to produce some quality counterfeit material, and D could be regarded as the police officer trying to detect the forged items. In an alternative view, we can perceive G as receiving a reward signal from D depending upon whether the generated data is accurate or not. The gradient information is back propagated from D to G, so G adapts its parameters in order to produce an output image that can fool D. The training objectives of D and G can be expressed mathematically as:

$$\begin{aligned} \mathcal{L}_D^{GAN} &= \max_D \mathbb{E}_{x_r \sim p_r(x)}[\log D(x_r)] + \mathbb{E}_{x_g \sim p_g(x)}[\log(1 - D(x_g))], \\ \mathcal{L}_G^{GAN} &= \min_G \mathbb{E}_{x_g \sim p_g(x)}[\log(1 - D(x_g))]. \end{aligned} \tag{1}$$

As can be seen, D is simply a binary classifier with a maximum log likelihood objective. If the discriminator D is trained to optimality before the next generator G updates, then minimizing $\mathcal{L}_G^{GAN}$ is proven to be equivalent to minimizing the Jensen–Shannon (JS) divergence between $p_r(x)$ and $p_g(x)$ (Goodfellow et al., 2014). The desired outcome after training is that samples formed by x_g should approximate the real data distribution $p_r(x)$.

2.2. Challenges in optimizing GANs

The above GAN training objective is regarded as a saddle point optimization problem (Yadav et al., 2018) and the training is often accomplished by gradient-based methods. G and D are trained alternately from scratch so that they may evolve together. However, there is no guarantee of balance between the training of G and D with the JS divergence. As a consequence, one network may inevitably be more powerful than the other, which in most cases is D. When D becomes too strong as opposed to G, the generated samples become too easy to be separated from real ones, thus reaching a stage where gradients from D approach zero, providing no guidance for further training of G. This happens more frequently when generating high resolution images due to the difficulty of generating meaningful high frequency details.

Another problem commonly faced in training GANs is mode collapse, which, as the name indicates, is a case when the distribution $p_g(x)$ learned by G focuses on a few limited modes of the data distribution $p_r(x)$. Hence instead of producing diverse images, it generates a limited set of samples.

2.3. Variants of GANs

2.3.1. *Varying objective of D*

In order to stabilize training and also to avoid mode collapse, different losses for D have been proposed, such as f-divergence (f-GAN) (Nowozin et al., 2016), least-square (LSGAN) (Mao et al., 2016), hinge loss (Miyato et al., 2018), and Wasserstein distance (WGAN, WGAN-GP) (Arjovsky et al., 2017; Gulrajani et al., 2017). Among these, Wasserstein distance is arguably the most popular metric. As an alternative to the real/fake discrimination scheme, Springenberg (2015) proposed an entropy based objective where real data is encouraged to make confident class predictions (CatGAN, Figure 3 b). In EBGAN (Zhao et al., 2016) and BEGAN (Berthelot et al., 2017) (Figure 3 c), the commonly used encoder architecture for discriminator is replaced with an autoencoder architecture. D's objective then becomes matching autoencoder loss distribution rather than data distribution.

GANs themselves lack the mechanism of inferencing the underlying latent vector that is likely to encode the input. Therefore, in ALI (Dumoulin et al., 2016) and BiGAN (Donahue et al., 2016) (Figure 3 d), a separate encoder network is incorporated. D's objective then becomes separating joint samples (x_g, z_g) and (x_r, z_r). In InfoGAN (Figure 3 e), the discriminator outputs the latent vector that encodes part of the semantic features of the generated image. The discriminator maximizes the mutual information between the generated image and the latent attribute vector the generated image is conditioned upon. After successful training, InfoGAN can explore inherent data attributes and perform conditional data generation based on these attributes. The use of class labels has been shown to further improve generated image's quality and this information can be easily incorporated into D by enforcing D to provide class probabilities and use cross entropy loss for optimization such as used in ACGAN (Odena et al., 2016) (Figure 3 f).

2.3.2. *Varying objective of G*

In the vanilla GAN, G transforms noise z to sample $x_g = G(z)$. This is usually accomplished by using a decoder network to progressively increase the spatial size of the output until the desired resolution is achieved as shown in Figure 2. Larsen et al. (2015) proposed a variational autoencoder network (VAE) as the underlying architecture of G (VAEGAN, Figure 3 g), where it can use pixel-wise reconstruction loss to enforce the decoder part of VAE to generate structures to match the real images.

The original setup of a GAN does not have any restrictions on the modes of data it can generate. However, if auxiliary information were provided during the generation, the GAN can be driven to output images with desired properties. A GAN in this scenario is usually referred as a conditional GAN (cGAN) and the generation process can be expressed as $x_g = G(z, c)$.

One of the most common conditional inputs c is an image. pix2pix, the first general purpose GAN based image-to-image translation framework was proposed by Isola et al. (2016) (Figure 4 a). Further, task related supervision was introduced to the generator. For example, reconstruction loss for image restoration and Dice loss (Milletari et al., 2016) for segmentation. This form of supervision requires aligned training pairs. Zhu et al. (2017a); Kim et al. (2017) relaxed this constraint by stitching two generators together head to toe so that images can be translated between two sets of unpaired samples (Figure 4 b). For the sake of simplicity, we chose CycleGAN to represent this idea in the rest of this paper. Another model named UNIT (Figure 4 c) can also perform unpaired image-to-image transform by combining two VAEGANs together with each one responsible for one modality but sharing the same latent space (Liu et al., 2017a). These image-to-image translation frameworks are very popular in the medical imaging community due to their general applicability.

Other than image, the conditional input can be class labels (CGAN, Figure 3 h) (Mirza and Osindero, 2014), text descriptions (Zhang et al., 2017a), object locations (Reed et al., 2016a,b), surrounding image context (Pathak et al., 2016), or sketches (Sangkloy et al., 2016). Note that ACGAN mentioned in the previous section also has a class conditional generator.

2.3.3. *Varying architecture*

Fully connected layers were used as the building block in vanilla GAN but later on, were replaced by fully convolutional downsampling/upsampling layers in DCGAN (Radford et al., 2015). DCGAN demonstrated better training stability hence quickly populated the literature. As shown in Figure 2, the generator in DCGAN architecture works on random input noise vector by successive upsampling operations eventually generating an image from it. Two of its important ingredients are BatchNorm (Ioffe and Szegedy, 2015) for regulating the extracted feature scale, and LeakyRelu (Maas et al., 2013) for preventing dead gradients. Very recently, Miyato et al. (2018) proposed a spectral normalization layer that normalized weights in the discriminator to regulate the scale of feature response values. With the training stability improved, some works have

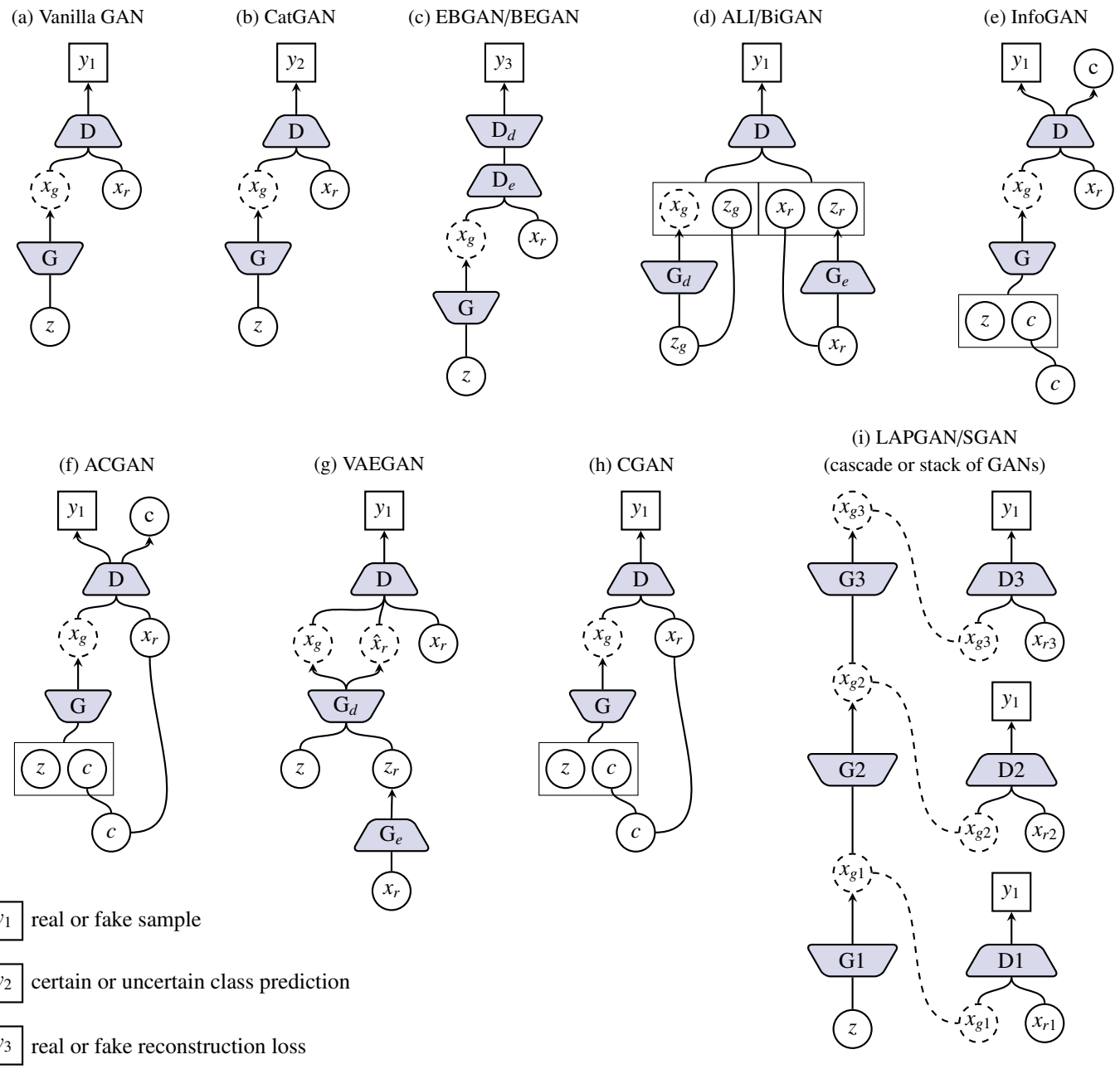

Figure 3: A schematic view of variants of GAN. c represents the conditional vector. In CGAN and ACGAN, c is the discrete categorical code (e.g. one hot vector) that encodes class labels and in InfoGAN it can also be continuous code that encodes attributes. x_g generally refers to the generated image but can also be internal representations as in SGAN.

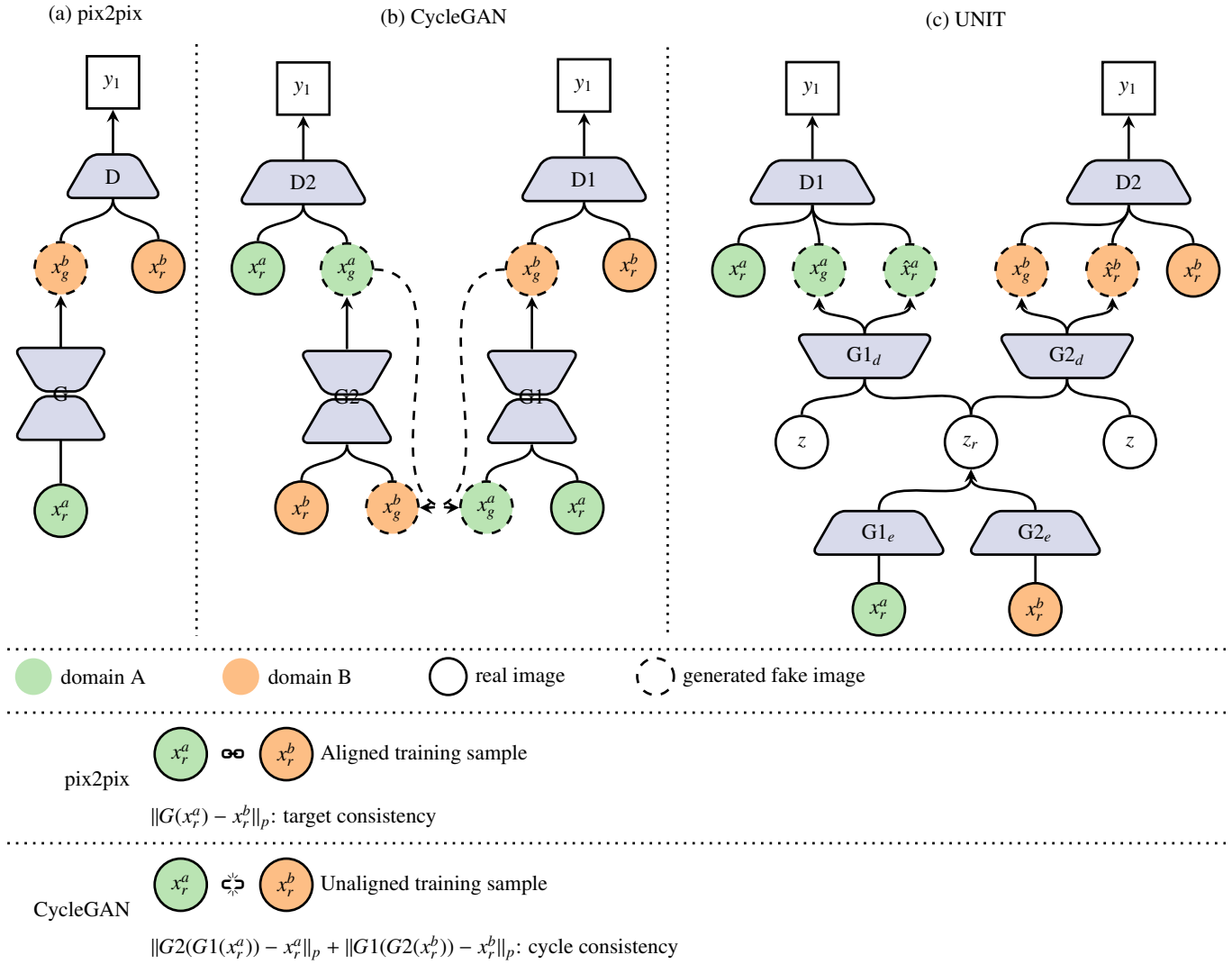

Figure 4: cGAN frameworks for image-to-image translation. pix2pix requires aligned training data whereas this constraint is relaxed in CycleGAN but usually suffers from performance loss. Note that in (a), we chose reconstruction loss as an example of target consistency. This supervision is task related and can take many other different forms. (c) It consists of two VAEGANs with shared latent vector in the VAE part.

also incorporated residual connections into both the generator and discriminator and experimented with much deeper networks (Gulrajani et al., 2017; Miyato et al., 2018). The work in Miyato and Koyama (2018) proposed a projection based way to incorporate the conditional information instead of direct concatenation and found it to be beneficial in improving the generated image's quality.

Directly generating high resolution images from a noise vector is hard, therefore some works have proposed tackling it in a progressive manner. In LAPGAN (Figure 3 i), Denton et al. (2015) proposed a stack of GANs, each of which adds higher frequency details into the generated image. In SGAN, a cascade of GANs is also used but each GAN generates increasingly lower level representations (Huang et al., 2017), which are compared with the hierarchical representations extracted from a discriminatively trained model. Karras et al. (2017) adopted an alternate way where they progressively grow the generator and discriminator by adding new layers to them rather than stacking another GAN on top of the preceding one (PGGAN). This progressive idea was also explored in conditional setting (Wang et al., 2017b). More recently, Karras et al. (2018) proposed a style-based generator architecture (styleGAN) where instead of directly feeding the latent code z to the input of the generator, they transformed this code first to an intermediate latent space and then use it to scale and shift the normalized image feature responses computed from each convolution layer. Similarly, Park et al. (2019) proposed SPADE where the segmentation mask was injected to the generator via a spatially adaptive normalization layer. This conditional setup was found to better preserve the semantic layout of the mask than directly feeding the mask to the generator.

Schematic illustrations of the most representative GANs are shown in Figure 3. They are GAN, CatGAN, EBGAN/BEGAN, ALI/BiGAN, InfoGAN, ACGAN, VAEGAN, CGAN, LAPGAN,

SGAN. Three popular image-to-image translation cGANs (pix2pix, CycleGAN, and UNIT) are shown in Figure 4. For a more in-depth review and empirical evaluation of these different variants of GAN, we refer the reader to (Huang et al., 2018; Creswell et al., 2018; Kurach et al., 2018).

3. Applications in Medical Imaging

There are generally two ways GANs are used in medical imaging. The first is focused on the generative aspect, which can help in exploring and discovering the underlying structure of training data and learning to generate new images. This property makes GANs very promising in coping with data scarcity and patient privacy. The second focuses on the discriminative aspect, where the discriminator D can be regarded as a learned prior for normal images so that it can be used as regularizer or detector when presented with abnormal images. Figure 5 provides examples of GAN related applications, with examples (a), (b), (c), (d), (e), (f) that focus on the generative aspect and example (g) that exploits the discriminative aspect. In the following subsections, in order to help the readers find applications of their interest, we categorized all the reviewed articles into canonical tasks: reconstruction, image synthesis, segmentation, classification, detection, registration, and others.

3.1. Reconstruction

Due to constraints in clinical settings, such as radiation dose and patient comfort, the diagnostic quality of acquired medical images may be limited by noise and artifacts. In the last decade, we have seen a paradigm shift in reconstruction methods changing from analytic to iterative and now to machine learning based methods. These data-driven learning based methods either learn to transfer raw sensory inputs directly to output images or serve as a post processing step for reducing image noise and removing artifacts. Most of the methods reviewed in this section are borrowed directly from the computer vision literature that formulate post-processing as an image-to-image translation problem where the conditioned inputs of cGANs are compromised in certain forms, such as low spatial resolution, noise contamination, under-sampling, or aliasing. One exception is for MR images where the Fourier transform is used to incorporate the raw K-space data into the reconstruction.

The basic pix2pix framework has been used for low dose CT denoising (Wolterink et al., 2017b), MR reconstruction (Chen et al., 2018b; Kim et al., 2018; Dar et al., 2018b; Shitrit and Raviv, 2017), and PET denoising (Wang et al., 2018b). A pretrained VGG-net (Simonyan and Zisserman, 2014) was further incorporated into the optimization framework to ensure perceptual similarity (Yang et al., 2017b; Yu et al., 2017; Yang et al., 2018a; Armanious et al., 2018c; Mahapatra, 2017). Yi and Babyn (2018) introduced a pretrained sharpness detection network to explicitly constrain the sharpness of the denoised CT especially for low contrast regions. Mahapatra (2017) computed a local saliency map to highlight blood vessels in super-resolution process of retinal fundus imaging. A similar idea was explored by Liao et al. (2018) in sparse view CT reconstruction. They compute a focus map to modulate the reconstructed output to ensure that the network focused on important regions. Besides ensuring image domain data fidelity, frequency domain data fidelity is also imposed when raw K-space data is available in MR reconstruction (Quan et al., 2018; Mardani et al., 2017; Yang et al., 2018a).

Losses of other kinds have been used to highlight local image structures in the reconstruction, such as the saliency loss to reweight each pixel's importance based on its perceptual relevance (Mahapatra, 2017) and the style-content loss in PET denoising (Armanious et al., 2018c). In image reconstruction of moving organs, paired training samples are hard to obtain. Therefore, Ravì et al. (2018) proposed a physical acquisition based loss to regulate the generated image structure for endomicroscopy super resolution and Kang et al. (2018) proposed to use CycleGAN together with an identity loss in the denoising of cardiac CT. Wolterink et al. (2017b) found that in low dose CT denoising, meaningful results can still be achieved when removing the image domain fidelity loss from the pix2pix framework, but the local image structure can be altered. Papers relating to medical image reconstruction are summarized in Table 1.

It can be noticed that the underlying methods are almost the same for all the reconstruction tasks. MR is special case as it has a well defined forward and backward operation, i.e. Fourier transform, so that raw K-space data can be incorporated. The same methodology can potentially be applied to incorporate the sinogram data in the CT reconstruction process but we have not seen any research using this idea as yet probably because the sinogram data is hard to access. The more data used, either raw K-space or image from other sequence, the better are the reconstructed results. In general, using adversarial loss produces more visually appealing results than using pixel-wise reconstruction loss alone. But using adversarial loss to match the generated and real data distribution may make the model hallucinate unseen structures. Pixel-wise reconstruction loss helps to combat this problem if paired samples are available, and if the model was trained on all healthy images but employed to reconstruct images with pathologies, the hallucination problem will still exist due to domain mismatch. Cohen et al. (2018) have conducted extensive experiments to investigate this problem and suggest that reconstructed images should not be used for direct diagnosis by radiologists unless the model has been properly verified.

However, even though the dataset is carefully curated to match the training and testing distribution, there are other problems in further boosting performance. We have seen various different losses introduced to the pix2pix framework as shown in Table 2 to improve the reconstructed fidelity of local structures. There is, however, no reliable way of comparing their effectiveness except for relying on human observer or downstream image analysis tasks. Large scale statistical analysis by human observer is currently lacking for GAN based reconstruction methods. Furthermore, public datasets used for image reconstruction are not tailored towards further medical image analysis, which leaves a gap between upstream reconstruction and downstream analysis tasks. New reference standard datasets should be created for better comparison of these GAN-based methods.

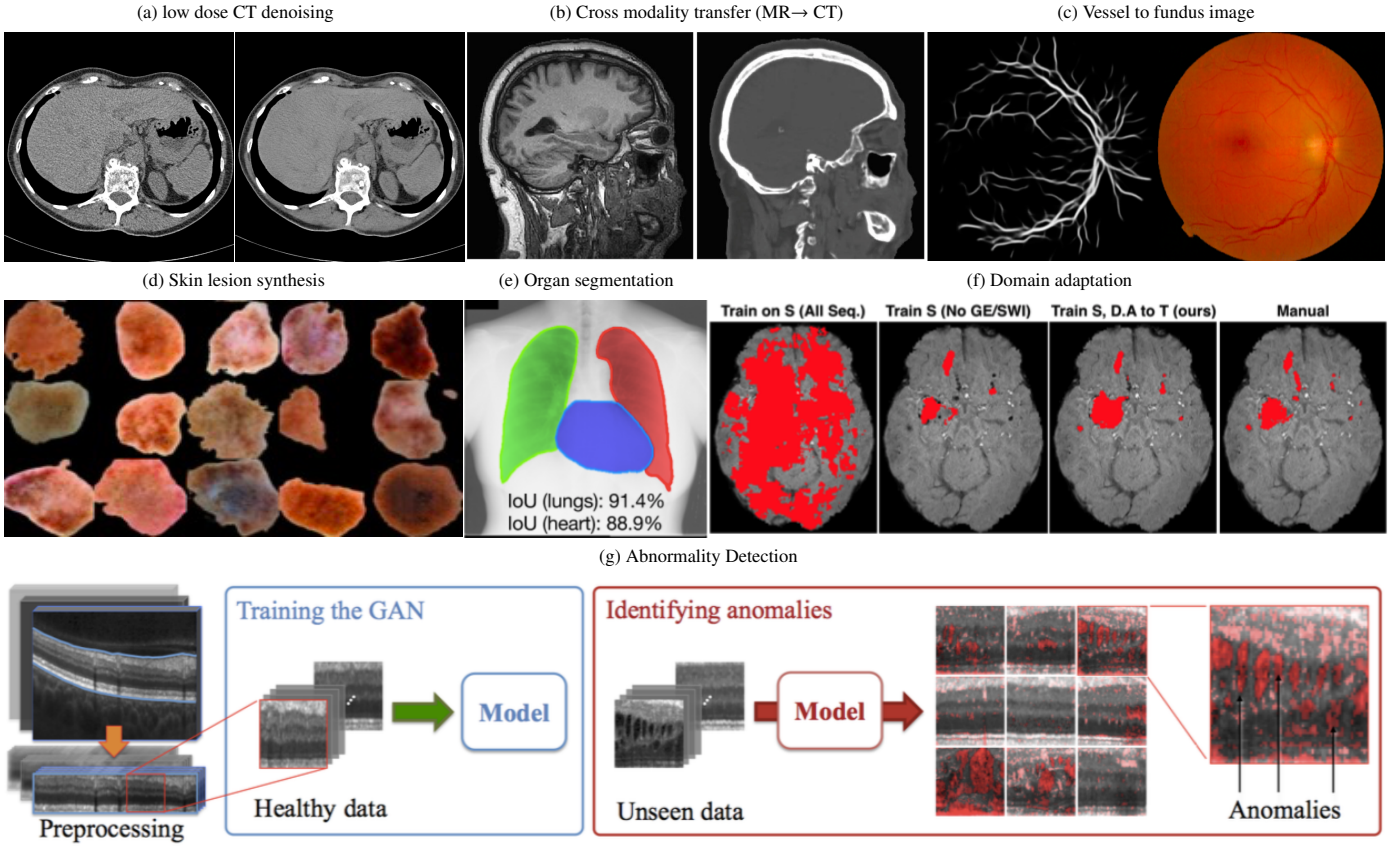

Figure 5: Example applications using GANs. Figures are directly cropped from the corresponding papers. (a) Left side shows the noise contaminated low dose CT and right side shows the denoised CT that well preserved the low contrast regions in the liver (Yi and Babyn, 2018). (b) Left side shows the MR image and right side shows the synthesized corresponding CT. Bone structures were well delineated in the generated CT image (Wolterink et al., 2017a). (c) The generated retinal fundus image have the exact vessel structures as depicted in the left vessel map (Costa et al., 2017b). (d) Randomly generated skin lesion from random noise (a mixture of malignant and benign) (Yi et al., 2018). (e) An organ (lung and heart) segmentation example on adult chest X-ray. The shapes of lung and heart are regulated by the adversarial loss (Dai et al., 2017b). (f) The third column shows the domain adapted brain lesion segmentation result on SWI sequence without training with the corresponding manual annotation (Kamnitsas et al., 2017). (g) Abnormality detection of optical coherence tomography images of the retina (Schlegl et al., 2017).

3.2. Medical Image Synthesis

Depending on institutional protocols, patient consent may be required if diagnostic images are intended to be used in a publication or released into the public domain (Clinical Practice Committee, 2000). GANs are widely for medical image synthesis. This helps overcome the privacy issues related to diagnostic medical image data and tackle the insufficient number of positive cases of each pathology. Lack of experts annotating medical images poses another challenge for the adoption of supervised training methods. Although there are ongoing collaborative efforts across multiple healthcare agencies aiming to build large open access datasets, e.g. Biobank, the National Biomedical Imaging Archive (NBIA), The Cancer Imaging Archive (TCIA) and Radiologist Society of North America (RSNA), this issue remains and constrains the number of images researchers might have access to.

Traditional ways to augment training sample include scaling, rotation, flipping, translation, and elastic deformation (Simard et al., 2003). However, these transformations do not account for variations resulting from different imaging protocols or sequences, not to mention variations in the size, shape, location and appearance of specific pathology. GANs provide a more generic solution and have been used in numerous works for augmenting training images with promising results.

3.2.1. Unconditional Synthesis

Unconditional synthesis refers to image generation from random noise without any other conditional information. Techniques commonly adopted in the medical imaging community include DCGAN, WGAN, and PGGAN due to their good training stability. The first two methods can handle an image resolution of up to 256×256 but if higher resolution images are desired, the progressive technique proposed in PGGAN is a choice. Realistic images can be generated by directly using the author released code base as long as the variations between images are not too large, for example, lung nodules and liver lesions. To make the generated images useful for downstream tasks, most studies trained a separate generator for each individual class; for example, Frid-Adar et al. (2018) used three DCGANs to generate synthetic samples for three classes of liver lesions (cysts, metastases, and hemangiomas); generated samples were found to be beneficial to the lesion classification task

Table 1: Medical image reconstruction publications. In the second column, * following the method denotes some modifications on the basic framework either on the network architecture or on the employed losses. A brief description of the losses, quantitative measures and datasets can be found in Table 2, 3 and 7. In the last column, symbol ✔ and ✖ denotes whether the corresponding literature used paired training data or not. All studies were performed in 2D unless otherwise mentioned.

Publications	Method	Losses	Dataset	Quantitative Measure	Remarks
CT					
Wolterink et al. (2017b)	pix2pix*	L1, 2	–	M29	[✖] [3D] Denoising
Yi and Babyn (2018)	pix2pix*	L1, 2, 6	D1	M12, 13, 24, 25	[✔] Denoising
Yang et al. (2017b)	pix2pix*	L1, 2, 8	D2	M12, 13, 25	[✔] [3D] [Abdomen] Denoising
Kang et al. (2018)	CycleGAN*	L1, 3, 19	–	M12, 25	[✖] [Coronary] Denoising CT
You et al. (2018a)	pix2pix*	L1, 2, 9	D2	M1, 11, 12, 13, 25	[✔] [3D] Denoising
Tang et al. (2018)	SGAN	L1, 2, 8	–	M32	[✔] Denoising, contrast enhance
Shan et al. (2018)	pix2pix*	L1, 8	D2	M9, 10, 12, 13	[✖] [3D] Denoising transfer from 2D
Liu et al. (2019)	pix2pix*	L1, 2, 8	–	M13, 24	[✔] Denoising, using adjacent slice
Liao et al. (2018)	pix2pix*	L1, 2, 8	–	M11, 12, 13	[✔] Sparse view CT reconstruction
Wang et al. (2018a)	pix2pix	L1, 2	–	M32	[✔] Metal artefact reduction cochlear implants
You et al. (2018b)	CycleGAN*	L1, 2, 12	D2	M12, 13, 16	[✔] Superresolution, denoising
GANs (2018)	pix2pix*	L1, 2	–	M11, 12	[✔] Sparse view CT reconstruction
Armanious et al. (2018b)	pix2pix*	L1, 2, 8, 11	–	M11, 12, 13, 15,	[✔] Inpainting
MR					
Quan et al. (2018)	pix2pix*	L1, 2, 15	D11, 12, 13	M11, 12, 13	[✔] Under-sampled K-space
Mardani et al. (2017)	pix2pix*	L1, 2, 15	–	M1	[✔] Under-sampled K-space
Yu et al. (2017)	pix2pix*	L1, 2, 8	D11, 3	M11, 12, 13, 24	[✔] Under-sampled K-space
Yang et al. (2018a)	pix2pix*	L1, 2, 8, 15	D3, 15	M11, 12, 13, 24	[✔] Under-sampled K-space
Sanchez and Vilaplana (2018)	pix2pix*	L1, 2, 4	D16	M12, 13	[✔] [3D] Superresolution
Chen et al. (2018b)	pix2pix*	L1, 2	–	M11, 12, 13	[✔] [3D] Superresolution
Kim et al. (2018)	pix2pix*	L1, 2	D19	M1, 11, 13, 26	[✔] Superresolution
Dar et al. (2018b)	pix2pix*	L1, 2	D11, 19, 22	M12, 13	[✔] Under-sampled K-space
Shitrit and Raviv (2017)	pix2pix*	L1, 2	–	M12	[✔] Under-sampled K-space
Ran et al. (2018)	pix2pix*	L1, 2, 8	D11	M12,13	[✔] [3D] Denoising
Seitzer et al. (2018)	pix2pix*	L1, 2, 8	–	M1, 12, 23	[✔] Two stage
Abramian and Eklund (2018)	CycleGAN	L1, 2, 3	D11	M13, 21	[✔] Facial anonymization problem
Armanious et al. (2018b)	pix2pix*	L1, 2, 8, 11	–	M11, 12, 13, 15	[✔] Inpainting
Oksuz et al. (2018)	pix2pix*	L1, 2	D26	M11, 12, 13	[✔] Motion correction
Zhang et al. (2018a)	pix2pix*	L1, 2, 8, 12	–	M12, 13	[✔] Directly in complex-valued k-space data
Armanious et al. (2018a)	pix2pix*	L1, 2, 8, 11	–	M13, 14, 15, 18	[✔] Motion correction
PET					
Wang et al. (2018b)	cascade cGAN	L1, 2	–	M11, 12, 27	[✔] [3D]
Armanious et al. (2018c)	pix2pix*	L1, 2, 8, 11	–	M1, 11, 12, 13, 14, 15, 18	[✔]
Retinal fundus imaging					
Mahapatra (2017)	pix2pix*	L1, 2, 8, 17	–	M11, 12, 13	[✔] Superresolution
Endomicroscopy					
Ravì et al. (2018)	pix2pix*	L1, 18, 19	–	M6, 13	[✖] Superresolution

Table 2: A brief summary of different losses used in the reviewed publications in Tables 1 and 5. The third column specifies conditions to be fulfilled in order to use the corresponding loss. L in the first column stands for loss.

Abbr.	Losses	Requirement	Remarks
L1	$\mathcal{L}_{\text{adversarial}}$	–	Adversarial loss introduced by the discriminator, can take the form of cross entropy loss, hinge loss, least square loss etc. as discussed in Section 2.3.1
L2	$\mathcal{L}_{\text{image}}$	Aligned training pair	Element-wise data fidelity loss in image domain to ensure structure similarity to the target when aligned training pair is provided
L3	$\mathcal{L}_{\text{cycle}}$	–	Element-wise loss to ensure self-similarity during cycled transformation when unaligned training pair is provided
L4	$\mathcal{L}_{\text{gradient}}$	Aligned training pair	Element-wise loss in the gradient domain to emphasize edges
L5	$\mathcal{L}_{\text{edge}}$	Aligned training pair	Similar to $\mathcal{L}_{\text{gradient}}$ but using gradient feature map as a weight to image pixels
L6	$\mathcal{L}_{\text{sharp}}$	Aligned training pair	Element-wise loss in a feature domain computed from a pre-trained network, which is expected to be the image sharpness with focus on low contrast regions
L7	$\mathcal{L}_{\text{shape}}$, $\mathcal{L}_{\text{seg}}$	Annotated pixel-wise label	Loss introduced by a segmentor to ensure faithful reconstruction of anatomic regions
L8	$\mathcal{L}_{\text{perceptual}}$	Aligned training pair	Element-wise loss in a feature domain computed from a pre-trained network which expected to conform to visual perception
L9	$\mathcal{L}_{\text{structure}}$	Aligned training pair	Patch-wise loss in the image domain computed with SSIM which claims to better conform to human visual system
L10	$\mathcal{L}_{\text{structure2}}$	Aligned pair	MIND (Heinrich et al., 2012) as used in image registration for two images with the same content from different modality
L11	$\mathcal{L}_{\text{style-content}}$	Aligned training pair	Style and content loss to ensure similarity of image style and content. Style is defined as the Gram matrix which is basically the correlation of low-level features
L12	$\mathcal{L}_{\text{self-reg}}$	–	Element-wise loss in image domain to ensure structure similarity to the input. Useful in denoising since the two have similar underlying structure
L13	$\mathcal{L}_{\text{steer}}$	Aligned training pair	Element-wise loss in a feature domain which is computed from steerable filters with focus on vessel-like structures
L14	$\mathcal{L}_{\text{classify}}$	Aligned image-wise label	Loss introduced by a classifier to get semantic information
L15	$\mathcal{L}_{\text{frequency}}$	Aligned training pair	Element-wise loss in frequency domain (K-space) used in MR image reconstruction
L16	$\mathcal{L}_{\text{KL}}$	–	Kullback–Leibler divergence which is commonly seen in variational inference to ensure closer approximation to the posterior distribution
L17	$\mathcal{L}_{\text{saliency}}$	Aligned training pair	Element-wise loss in a feature domain which is expected to be the saliency map
L18	$\mathcal{L}_{\text{physical}}$	Physical model	Loss introduced by a physical image acquisition model
L19	$\mathcal{L}_{\text{regulation}}$	–	Regulate the generated image contrast by keeping the mean value across row and column unchanged

with both improved sensitivity and specificity when combined with real training data. Bermudez et al. (2018) claimed that neuroradiologists found generated MR images to be of comparable quality to real ones, however, there were discrepancies in anatomic accuracy. Papers related to unconditional medical image synthesis are summarized in Table 4.

3.2.2. Cross modality synthesis

Cross modality synthesis (such as generating CT-like images based on MR images) is deemed to be useful for multiple reasons, one of which is to reduce the extra acquisition time and cost. Another reason is to generate new training samples with the appearance being constrained by the anatomical structures delineated in the available modality. Most of the methods reviewed in this section share many similarities to those in Section 3.1. pix2pix-based frameworks are used in cases where different image modality data can be co-registered to ensure data fidelity. CycleGAN-based frameworks are used to handle more general cases where registration is challenging such as in cardiac applications. In a study by Wolterink et al. (2017a) for brain CT image synthesis from MR image, the authors found that training using unpaired images was even better than using aligned images. This most likely resulted from the fact that rigid registration could not very well handle local alignment in the throat, mouth, vertebrae, and nasal cavities. Hiasa et al. (2018) further incorporated gradient consistency loss in

Table 3: A brief summary of quantitative measures used in the reviewed publications listed in Tables 1, 4, 5 and 6.

Abbr.	Measures	Remarks
Overall image quality without reference		
M1	Human observer	Gold standard but costly and hard to scale
M2 Breuleux et al. (2011)	Kernel density function	Estimate the probability density of the generated data and compute the log likelihood of real test data under this distribution
M3 Salimans et al. (2016)	Inception score	Measure the generated images' diversity and visual similarity to the real images with the pretrained Inception model
M4 Goodfellow et al. (2014)	JS divergence	Distance measure between two distributions (used for comparison between normalized color histogram computed from a large batch of image samples)
M5	Wasserstein distance	Distance measure between two distributions (used for comparison between normalized color histogram computed from a large batch of image samples)
M6 Matkovic et al. (2005)	GCF	Global contrast factor
M7 Köhler et al. (2013)	Q_V	Vessel-based quality metric (noise and blur) for fundus image
M8 Niemeijer et al. (2006)	ISC	Image structure clustering, a trained classifier based to differentiate normal from low quality fundus images
M9 Shan et al. (2018)	Perceptual loss	Difference of features extracted from a pre-trained VGG net
M10 Shan et al. (2018)	Texture loss	Gram matrix which is basically the correlation of low-level features, defined as style in style transfer literature
Overall image quality with respect to a groundtruth		
M11	NMSE/MAE/MSE	(Normalized) mean absolute/square error with respect to a given groundtruth
M12	PSNR/SNR	(Peak) signal to noise ratio with respect to a given groundtruth
M13 Wang et al. (2004)	SSIM	Structural similarity with respect to a given groundtruth
M14 Sheikh and Bovik (2004)	VIF	Visual information fidelity with regard to a given groundtruth
M15 Wang and Bovik (2002)	UQI	Universal quality index with regard to a given groundtruth
M16 Sheikh et al. (2005)	IFC	Information Fidelity Criterion
M17 Zhang et al. (2011)	FSIM	A low-level feature based image quality assessment metric with regard to a given groundtruth
M18 Zhang et al. (2018b)	LPIPS	Learned perceptual image patch similarity
M19 Pluim et al. (2003)	Mutual information	Commonly used in cross modality registration in evaluating the alignment of two images
M20	NMI/MI	(Normalized) median intensity, used to measure color consistancy of histology images
M21 Lee Rodgers and Nicewander (1988)	Cross correlation	Global correlation between two images
M22 Low (2010)	Clinical measure	Dose difference, gamma analysis for CT
M23 Seitzer et al. (2018)	SIS	Semantic interpretability score, essentially the dice loss of a pre-trained downstream segmentor
Local image quality		
M24	Line profile	Measure the loss of spatial resolution
M25	Noise level	Standard deviation of intensities in a local smooth region
M26	CBR	Contrast to background ratio, measure the local contrast loss
M27 Kinahan and Fletcher (2010)	SUV	Standard uptake value, a clinical measure in oncology for local interest region, should not vary too much in reconstruction
M28	NPS	Noise power spectrum
Image quality analysis by auxiliary task		
M29	Task specific statistics	Down stream task (e.g. for coronary calcium quantification)
M30	Classification	Down stream task
M31	Detection	Down stream task (e.g. for lesion/hemorrhage)
M32	Segmentation	Down stream task
M33	Cross modality registration	Down stream task
M34	Depth estimation	Down stream task

the training to improve accuracy at the boundaries. Zhang et al. (2018d) found that using only cycle loss in the cross modality synthesis was insufficient to mitigate geometric distortions in the transformation. Therefore, they employed a shape consistency loss that is obtained from two segmentors (segmentation network). Each segmentor segments the corresponding image modality into semantic labels and provides implicit shape constraints on the anatomy during translation. To make the whole system end-to-end trainable, semantic labels of training images from both modalities are required. Zhang et al. (2018c) and Chen et al. (2018a) proposed using a segmentor also in the cycle transfer using labels in only one modality. Therefore, the segmentor is trained offline and fixed during the training of the image transfer network. As reviewed in Section 2, UNIT and CycleGAN are two equally valid frameworks for unpaired cross modality synthesis. It was found that these two frameworks performed almost equally well for the transformation between T1 and T2-weighted MR images (Welander et al., 2018). Papers related to cross modality medical image synthesis are summarized in Table 5.

3.2.3. *Other conditional synthesis*

Medical images can be generated by constraints on segmentation maps, text, locations or synthetic images etc. This is useful to synthesize images in uncommon conditions, such as lung nodules touching the lung border (Jin et al., 2018b). Moreover, the conditioned segmentation maps can also be generated from GANs (Guibas et al., 2017) or from a pretrained segmentation network (Costa et al., 2017a), by making the generation a two stage process. Mok and Chung (2018) used cGAN to augment training images for brain tumour segmentation. The generator was conditioned on a segmentation map and generated brain MR images in a coarse to fine manner. To ensure the tumour was well delineated with a clear boundary in the generated image, they further forced the generator to output the tumour boundaries in the generation process. The full list of synthesis works is summarized in Table 6.

3.3. *Segmentation*

Generally, researchers have used pixel-wise or voxel-wise loss such as cross entropy for segmentation. Despite the fact that U-net (Ronneberger et al., 2015) was used to combine both low-level and high-level features, there is no guarantee of spatial consistency in the final segmentation map. Traditionally, conditional random field (CRF) and graph cut methods are usually adopted for segmentation refinement by incorporating spatial correlation. Their limitation is that they only take into account pair-wise potentials which might cause serious boundary leakage in low contrast regions. On the other hand, adversarial losses as introduced by the discriminator can take into account high order potentials (Yang et al., 2017a). In this case, the discriminator can be regarded as a shape regulator. This regularization effect is more prominent when the object of interest has a compact shape, e.g. for lung and heart mask but less useful for deformable objects such as vessels and catheters. This regulation effect can be also applied to the internal features of the segmentor to achieve domain (different scanners, imaging protocols, modality) invariance (Kamnitsas et al., 2017; Dou

Table 4: Unconditional medical image synthesis publications. A brief description of the quantitative measures and datasets can be found in Tables 3 and 7.

Publications	Method	Dataset	Measures	Remarks
CT				
Chuquicusma et al. (2017)	DCGAN	D4	M1	[Lung nodule]
Frid-Adar et al. (2018)	DCGAN /ACGAN	–	M30	[Liver lesion] Generating each lesion class separately (with DCGAN) is than generating all classes at once (using ACGAN)
Bowles et al. (2018a)	PGGAN	–	M32	[Brain] Joint learning of image and segmentation map
MR				
Calimeri et al. (2017)	LAPGAN	–	M1, 2, 3	[Brain]
Zhang et al. (2017b)	Semi-Coupled-GAN	–	M30	[Heart] Two generators coupled with a single discriminator which outputted both a distribution over the image data source and class labels
Han et al. (2018a)	WGAN	D20	M1	[Brain]
Beers et al. (2018)	PGGAN	D21	-	[Brain]
Bermudez et al. (2018)	DCGAN	D23	M1	[Brain]
Mondal et al. (2018)	DCGAN*	D18, 25	M32	[Brain] Semi-supervised training with labeled, unlabeled, generated data
Bowles et al. (2018a)	PGGAN	–	M32	[Brain] Joint learning of image and segmentation map
X-ray				
Salehinejad et al. (2017)	DCGAN	–	M30	[Chest] Five different GANs to generate five different classes of chest disease
Madani et al. (2018b)	DCGAN	D34	M30	[Chest] Semi-supervised DCGAN can achieve performance comparable with a traditional supervised CNN with an order of magnitude less labeled data
Madani et al. (2018a)	DCGAN	D34	M30	[Chest] Two GANs to generate normal and abnormal chest X-rays separately
Mammography				
Korkinof et al. (2018)	PGGAN	–	–	–
Histopothology				
Hu et al. (2017a)	WGAN+infoGAN	D42	M30, M32	Cell level representation learning
Retinal fundus imaging				
Beers et al. (2018)	PGGAN	–	–	–
Lahiri et al. (2017)	DCGAN	D43	M30	Semi-supervised DCGAN can achieve performance comparable with a traditional supervised CNN with an order of magnitude less labeled data
Lahiri et al. (2018)	DCGAN	D43, 44	M30	Extend the above work by adding an unsupervised loss into the discriminator
Dermoscopy				
Baur et al. (2018b)	LAPGAN	D28	M4, 11	–
Baur et al. (2018a)	PGGAN	D29	M1	–
Yi et al. (2018)	CatGAN+ WGAN	D27, 30	M30	Semi-supervised skin lesion feature representation learning

et al., 2018). The adversarial loss can also be viewed as a adaptively learned similarity measure between the segmented outputs and the annotated groundtruth. Therefore, instead of measuring the similarity in the pixel domain, the discriminative network projects the input to a low dimension manifold and measures the similarity there. The idea is similar to the perceptual loss. The difference is that the perceptual loss is computed from a pre-trained classification network on natural images whereas the adversarial loss is computed from a network that trained adaptively during the evolvement of the generator.

Xue et al. (2018) used a multi-scale L_1 loss in the discriminator where features coming from different depths are compared. This was demonstrated to be effective in enforcing the multi-scale spatial constraints on segmentation maps and the system achieved state-of-the-art performance in the BRATS 13 and 15 challenges. Zhang et al. (2017c) proposed to use both annotated and unannotated images in the segmentation pipeline. The annotated images are used in the same way as in (Xue et al., 2018; Son et al., 2017) where both element-wise loss and adversarial loss are applied. The unannotated images on the other hand are only used to compute a segmentation map to confuse the discriminator. Li and Shen (2018) combined pix2pix with ACGAN for segmentation of fluorescent microscopy images of different cell types. They found that the introduction of the auxiliary classifier branch provides regulation to both the discriminator and the segmentor.

Unlike these aforementioned segmentation works where adversarial training is used to ensure higher order structure consistency on the final segmentation maps, the adversarial training scheme in (Zhu et al., 2017b) enforces network invariance to small perturbations of the training samples in order to reduce overfitting on small dataset. Papers related to medical image segmentation are summarized in Table 8.

3.4. Classification

Classification is arguably one of the most successful tasks where deep learning has been applied. Hierarchical image features can be extracted from a deep neural network discriminatively trained with image-wise class labels. GANs have been used for classification problems as well, either using part of the generator and discriminator as a feature extractor or directly using the discriminator as a classifier (by adding an extra class corresponding to the generated images). Hu et al. (2017a) used combined WGAN and InfoGAN for unsupervised cell-level feature representation learning in histopathology images whereas Yi et al. (2018) combined WGAN and CatGAN for unsupervised and semi-supervised feature representation learning for dermoscopy images. Both works extract features from the discriminator and build a classifier on top. Madani et al. (2018b), Lahiri et al. (2017) and Lecouat et al. (2018) adopted the semi-supervised training scheme of GAN for chest abnormality classification, patch-based retinal vessel classification and cardiac disease diagnosis respectively. They found that the semi-supervised GAN can achieve performance compara-

Table 5: Cross modality image synthesis publications. In the second column, * following the method denotes some modifications on the basic framework either on the network architecture or on the employed losses. A brief description of the losses, quantitative evaluation measures and datasets can be found in Tables 2, 3 and 7. In the last column, symbol ✔ and ✖ denotes whether the corresponding literature used paired training data or not.

Pulications	Method	Loss	Dataset	Measures	Remarks
MR → CT					
Nie et al. (2017, 2018)	Cascade GAN	L1, 2, 4	D16	M11, 12	[✔]Brain; Pelvis
Emami et al. (2018)	cGAN	L1, 2	–	M11, 12, 13	[✔]Brain
CT → MR					
Jin et al. (2018a)	CycleGAN	L1, 2, 3	–	M11, 12	[✖] Brain
Jiang et al. (2018)	CycleGAN*	L1, 2, 3, 7, 8	D8	M32	[✖] Lung
MR ↔ CT					
Chartsias et al. (2017)	CycleGAN	L1, 3	D9	M32	[✖] Heart
Zhang et al. (2018d)	CycleGAN*	L1, 3, 7	–	M32	[✖][3D] Heart
Huo et al. (2017)	CycleGAN*	L1, 3, 7	–	M32	[✖] Spleen
Chartsias et al. (2017)	CycleGAN	L1, 3	–	M32	[✖] Heart
Hiasa et al. (2018)	CycleGAN*	L1, 3, 4	–	M19, 32	[✖] Musculoskeletal
Wolterink et al. (2017a)	CycleGAN	L1, 3	–	M11, 12	[✖] Brain
Huo et al. (2018b)	CycleGAN	L1, 3, 7	–	M32	[✖] Abdomen
Yang et al. (2018b)	CycleGAN*	L1, 2, 3, 10	–	M11, 12, 13	[✖] Brain
Maspero et al. (2018)	pix2pix	L1, 2	–	M11, 22	[✔] Pelvis
CT → PET					
Bi et al. (2017)	cGAN	L1, 2	–	M11, 12	[✔] Chest
Ben-Cohen et al. (2018)	FCN+cGAN	L1, 2	–	M11, 12, 31	[✔] Liver
PET → CT					
Armanious et al. (2018c)	cGAN*	L1, 2, 8, 11	–	M11, 12, 13, 14, 15, 18	[✔] Brain
MR → PET					
Wei et al. (2018)	cascade cGAN	L1, 2	–	M29	[✔] Brain
Pan et al. (2018)	3D CycleGAN	L1, 2, 3	D16	M30	[✔] Brain
PET → MR					
Choi and Lee (2017)	pix2pix	L1, 2	D16	M13, 29	[✔] Brain
Synthetic → Real					
Hou et al. (2017)	synthesizer+cGAN	L1, 2, 7	D35, 36	M1, 32	[✔] Histopathology
Real → Synthetic					
Mahmood et al. (2017)	cGAN	L1, 12	–	M34	[✖] Endocsocpy
Zhang et al. (2018c)	CycleGAN*	L1, 3, 7	–	M32	[✖] X-ray
Domain adaption					
Chen et al. (2018a)	CycleGAN*	L1, 3, 7	D32, 33	M32	[✖] X-ray
T1 ↔ T2 MR					
Dar et al. (2018a)	CycleGAN	L1, 3	D11, 19, 22	M12, 13	[✖] Brain
Yang et al. (2018c)	cGAN	L1, 2	D19	M11, 12, 19, 32, 33	[✖] Brain
Welander et al. (2018)	CycleGAN, UNIT	L1, 2, 3	D24	M11, 12, 19	[✖] Brain
Liu (2018)	CycleGAN	L1, 2, 3	D14	M32	[✖] Knee
T1 → FLAIR MR					
Yu et al. (2018)	cGAN	L1, 2	D19	M11, 12, 32	[✔] [3D] Brain
T1, T2 → MRA					
Olut et al. (2018)	pix2pix*	L1, 2, 13	D11	M12, 32	[✔] Brain
3T → 7T MR					
Nie et al. (2018)	Cascade GAN	L1, 2, 4	–	M11, 12	[✔] Brain
Histopathology color normalization					
Bentaieb and Hamarneh (2018)	cGAN+classifier	L1, 5, 14	D37, 38, 39	M30	[✖]
Zanjani et al. (2018)	InfoGAN	L1, 2, 12, 16	–	M20	[✖]
Shaban et al. (2018)	CycleGAN	L1, 2, 3	D37, 40	M12, 13, 17, 30	[✖]
Hyperspectral histology → H&E					
Bayramoglu et al. (2017a)	cGAN	L1, 2	D41	M12, 13	[✔] Lung

Table 6: Other conditional image synthesis publications categorized by imaging modality. * following the method denotes some modifications on the basic framework either on the network architecture or on the employed losses. A brief description of the losses, quantitative evaluation measures and datasets can be found in Tables 2, 3 and 7

Publications	Conditional information	Method	Dataset	Evaluation
CT				
Jin et al. (2018b) (lung nodule)	VOI with removed central region	[3D] pix2pix* ($\mathcal{L}_1$ loss considering nodule context)	D2	M32
MR				
Mok and Chung (2018)	Segmentation map	Coarse-to-fine boundary-aware	D19	M32
Shin et al. (2018)	Segmentation map	pix2pix	D16, 19, 21	M32
Gu et al. (2018)	MR	CycleGAN	D24	M13, 21
Lau et al. (2018)	Segmentation map	Cascade cGAN	-	M32
Hu et al. (2018)	Gleason score	cGAN	-	-
Ultrasound				
Hu et al. (2017b) (fetus)	Probe location	cGAN	–	M1
Tom and Sheet (2018)	Segmentation map	cascade cGAN	D52	M1
Retinal fundus imaging				
Zhao et al. (2017)	Vessel map	cGAN	D41, 43, 45	M32
Guibas et al. (2017)	Vessel map	Dual cGAN	D43	M7, 32
Costa et al. (2017a)	Vessel map	Segmentor+pix2pix	D43	M7, 8
Costa et al. (2017b)	Vessel map	Adversarial VAE+cGAN	D43, 46	M8
Appan and Sivaswamy (2018)	Vessel map; Lesion map	cGAN	D46, 47, 48	M7, 31
Iqbal and Ali (2018)	Vessel map	cGAN	D43, 44	M32
Histopathology				
Senaras et al. (2018)	Segmentation map	pix2pix	–	M1
X-ray				
Galbusera et al. (2018)	Different view; segmentation map	pix2pix/CycleGAN	–	–
Mahapatra et al. (2018b)	segmentation map+X-ray	pix2pix* (content loss encourage dissimilarity)	D33	M30, 32
Oh and Yun (2018)	X-ray (for bone supression)	pix2pix* (Haar wavelet decomposion)	–	M11, 12, 13, 28

Table 7: Common datasets used in the reviewed literature. In the first column, D stands for Dataset.

Abbre.	Dataset	Purpose	Anatomy	Modality
D1 Yi and Babyn (2018)	Piglet	Denoising	Whole body	CT
D2 McCollough et al. (2017)	LDCT2016	Denoising	Abdomen	CT
D3	MICCAI2013	Organ segmentation	Abdomen, Pelvis	CT
D4 Armato III et al. (2015)	LIDC-IDRI	Lung cancer detection and diagnosis	Lung	CT
D5 Yan et al. (2018a)	DeepLesion	Lesion segmentation	–	CT
D6	LiTS2017	Liver tumor segmentation	Liver	CT
D7 Glocker et al. (2013)	Spine	Vertebrate localization	Spine	CT
D8 Aerts et al. (2015)	NSCLC-Radiomics	Radiomics	Lung	CT
D9 Zhuang and Shen (2016)	MM-WHS	Whole heart segmentation	Heart	CT, MR
D10 Pace et al. (2015)	HVSMR 2016	Whole heart and great vessel segmentation	Heart, Vessel	MR
D11	IXI	Analysis of brain development	Brain	MR
D12	DSB2015	End-systolic/diastolic volumes measurement	Heart	MR
D13	Mridata	MRI reconstruction	Knee	MR
D14	Ski10	Cartilage and bone segmentation	Knee	MR
D15 Crimi et al. (2016)	BrainLes	Lesion segmentation	Brain	MR
D16	ADNI	Alzheimers disease neuroimaging Initiative	Brain	MR, PET
D17	MAL	Brain structure segmentation	Brain	MR
D18	BRATS2013	Gliomas segmentation	Brain	MR
D19	BRATS2015	Gliomas segmentation	Brain	MR
D20	BRATS2016	Gliomas segmentation	Brain	MR
D21	BRATS2017	Gliomas segmentation, overall survival prediction	Brain	MR
D22 Bullitt et al. (2005)	MIDAS	Assessing the effects of healthy aging	Brain	MR
D23 Resnick et al. (2003)	BLSA	Baltimore longitudinal study of aging	Brain	MR
D24 Van Essen et al. (2012)	HCP	Human connectome project	Brain	MR
D25 Wang et al. (2019)	iSeg2017	Infant brain tissue segmentation	Brain	MR
D26	UK Biobank	Health research	Brain, Heart, Body	MR
D27 Gutman et al. (2016)	ISIC2016	Skin lesion analysis	Skin	Dermoscopy
D28 Codella et al. (2018)	ISIC2017	Skin lesion analysis	Skin	Dermoscopy
D29	ISIC2018	Skin lesion analysis	Skin	Dermoscopy
D30 Mendonca et al. (2015)	PH2	Skin lesion analysis	Skin	Dermoscopy
D31 Ballerini et al. (2013)	Dermofit	Skin lesion analysis	Skin	Dermoscopy
D32 Jaeger et al. (2014)	Montgomery	Pulmonary disease detection	Chest	X-Ray
D33 Shiraishi et al. (2000)	JSRT	Pulmonary nodule detection	Chest	X-Ray
D34	NIH PLCO	Cancer screening trial for Prostate, lung, colorectal and ovarian (PLCO)	-	X-ray; Digital pathology
D35	CBTC2015	Segmentation of nuclei	Nuclei	Digital pathology
D36	CPM2017	Segmentation of nuclei	Nuclei	Digital pathology
D37	MITOS-ATYPIA	Mitosis detection; Nuclear atypia score evaluation	Breast	Digital pathology
D38 Sirinukunwattana et al. (2017)	GlaS	Gland segmentation	Colon	Digital pathology
D39 Köbel et al. (2010)	OCHD	Carcinoma subtype prediction	Ovary	Digital pathology
D40	Camelyon16	Lymph node metastases detection	Breast	Digital pathology
D41 Bayramoglu et al. (2017b)	Neslihan	Virtual H&E staining	Lung	Digital pathology
D42 Kainz et al. (2015)	CellDetect	Cell detection	Bone marrow	Digital pathology
D43 Staal et al. (2004)	DRIVE	Blood vessels segmentation	Eye	Fundus imaging
D44	STARE	Structural analysis of the retina	Eye	Fundus Imaging
D45 Budai et al. (2013)	HRF	Image quality assessment, segmentation	Eye	Fundus Imaging
D46 Decencière et al. (2014)	Messidor	Segmentation in retinal ophthalmology	Eye	Fundus Imaging
D47 Prentasic et al. (2013)	DRiDB	Diabetic retinopathy detection	Eye	Fundus Imaging
D48 Kälviäinen and Uusitalo (2007)	DIARETDB1	Diabetic retinopathy detection	Eye	Fundus Imaging
D49 Fumero et al. (2011)	RIM-ONE	Optic nerve head segmentation	Eye	Fundus Imaging
D50 Hobson et al. (2015)	I3A	HEp-2 cell classification	Skin	Fluorescent microscopy
D51	MIVIA	HEp-2 cell segmentation	Skin	Fluorescent microscopy
D52 Balocco et al. (2014)	IVUS	Vessel inner and outer wall border detection	Blood Vessel	Ultrasound
D53 Moreira et al. (2012)	INbreast	Mass segmentation	Breast	Mammography
D54 Heath et al. (1998)	DDSM-BCRP	Mass segmentation	Breast	Mammography

ble with a traditional supervised CNN with an order of magnitude less labeled data. Furthermore, Madani et al. (2018b) have also shown that the adversarial loss can reduce domain overfitting by simply supplying unlabeled test domain images to the discriminator in identifying cardiac abnormalities in chest X-ray. A similar work in addressing domain variance in whole slide images (WSI) has been conducted by Ren et al. (2018).

Most of the other works that used GANs to generate new training samples have been already mentioned in Section 3.2.1. These studies applied a two stage process, with the first stage learned to augment the images and the second stage learned to perform classification by adopting the traditional classification network. The two stages are trained disjointedly without any communication in between. The advantage is that these two components can be replaced easily if more advanced unconditional synthesis architectures are proposed whereas the downside is that the generation has to be conducted for each class separately (N models for N classes), which is not memory and computation efficient. A single model that is capable of performing conditional synthesis of multiple categories is an active research direction (Brock et al., 2018). Surprisingly, Frid-Adar et al. (2018) found that using separate GAN (DCGAN) for each

Table 8: Segmentation publications. A brief description of the datasets can be found in Table 7.

Publications	Dataset	Remarks
CT		
Yang et al. (2017a)	–	[3D] [Liver] Generator is essentially a U-net with deep supervisions
Dou et al. (2018)	D9	Ensure that the feature distribution of images from both domains (MR and CT) are indistinguishable
Rezaei et al. (2018a)	D6	Additional refinement network, patient-wise batchNorm, recurrent cGAN to ensure temporal consistancy
Sekuboyina et al. (2018)	D7	Adversarial training based on EBGAN; Butterfly shape network to combine two views
MR		
Xue et al. (2018)	D18, 19	A multi-scale L_1 loss in the discriminator where features coming from different depth are compared
Rezaei et al. (2017)	D21	The generator takes heterogenous MR scans of various contrast as provided by BRATS 17 challenge
Rezaei et al. (2018b)	D10	A cascade of cGANs in segmenting myocardium and blood pool
Li et al. (2017)	D21	The generator takes heterogenous MR scans of various contrast as provided by BRATS 17 challenge
Moeskops et al. (2017)	D17, 18	–
Kohl et al. (2017)	–	[Prostate] Improved sensitivity
Huo et al. (2018a)	–	[Spleen] Global convolutional network (GCN) with a large receptive field as the generator
Kamnitsas et al. (2017)	–	Regulate the learned representation so that the feature representation is domain invariant
Dou et al. (2018)	D9	Ensure that the feature distribution of images from both domains (MR and CT) are indistinguishable
Rezaei et al. (2018a)	D21	Additional refinement network, patient-wise batchNorm, recurrent cGAN to ensure temporal consistency
Xu et al. (2018)	–	Joint learning (segmentation and quantification); convLSTM in the generator for spatial-temporal processing; Bi-LSTM in the discriminator to learn relation between tasks
Han et al. (2018b)	–	Local-LSTM in the generator to capture spatial correlations between neighbouring structures
Zhao et al. (2018)	D16	Deep supervision; Discriminate segmentation map based on features extracted from a pre-trained network
Retinal fundus imaging		
Son et al. (2017)	D43, 44	Deep architecture is better for discriminating whole images and has less false positives with fine vessels
Zhang et al. (2017c)	D38	Use both annotated and unannotated images in the segmentation pipeline
Shankaranarayana et al. (2017)	D49	–
X-ray		
Dai et al. (2017b)	D32, 33	Adversarial loss is able to correct the shape inconsistency
Histopothology		
Wang et al. (2017a)	–	Basal membrane segmentation
fluorescent microscopy		
Li and Shen (2018)	D50, 51	pix2pix + ACGAN; Auxiliary classifier branch provides regulation to both the discriminator and the segmentor
Dermoscopy		
Izadi et al. (2018)	D31	Adversarial training helps to refine the boundary precision
Mammography		
Zhu et al. (2017b)	D53, 54	Enforce network invariance to small perturbations of the training samples in order to reduce overfitting on small size dataset
Ultrasound		
Tuysuzoglu et al. (2018)	–	Joint learning (landmark localization + prostate contour segmentation); Contour shape prior imposed by the discriminator

lesion class resulted in better performance in lesion classification than using a unified GAN (ACGAN) for all classes. The underlying reason remains to be explored. Furthermore, Finlayson et al. (2018) argue that images generated from GANs may serve as an effective augmentation in the medium-data regime, but may not be helpful in a high or low-data regime.

3.5. Detection

The discriminator of GANs can be utilized to detect abnormalities such as lesions by learning the probability distribution of training images depicting normal pathology. Any image that falls out of this distribution can be deemed as abnormal. Schlegl et al. (2017) used the exact idea to learn a manifold of normal anatomical variability and proposed a novel anomaly scoring scheme based on the fitness of the test image's latent code to the learned manifold. The learning process was conducted in an unsupervised fashion and effectiveness was demonstrated by state-of-the-art performance of anomaly detection on optical coherence tomography (OCT) images. Alex et al. (2017) used GAN for brain lesion detection on MR images. The generator was used to model the distribution of normal patches and the trained discriminator was used to compute a posterior probability of patches centered on every pixel in the test image. Chen and Konukoglu (2018) used an adversarial auto-encoder to learn the data distribution of healthy brain MR images. The lesion image was then mapped to an image without a lesion by exploring the learned latent space, and the lesion could be highlighted by computing the residual of these two images. We can see that all the detection studies targeted for abnormalities that are hard to enumerate.

In the image reconstruction section, it has been observed that if the target distribution is formed from medical images without pathology, lesions within an image could be removed in the CycleGAN-based unpaired image transfer due to the distribution matching effect. However, it can be seen here that if the target and source domain are of the same imaging modality differing only in terms of normal and abnormal tissue, this adverse effect can actually be exploited for abnormality detection Sun et al. (2018).

3.6. Registration

cGAN can also be used for multi-modal or uni-modal image registration. The generator in this case will either generate transformation parameters, e.g. 12 numbers for 3D affine transformation, deformation field for non-rigid transformation or directly generate the transformed image. The discriminator then discriminates aligned image pairs from unaligned image pairs. A spatial transformation network (Jaderberg et al., 2015) or a deformable transformation layer (Fan et al., 2018) is usually plugged in between these two networks to enable end-to-end training. Yan et al. (2018b) performed prostate MR to transrectal ultrasound (TRUS) image registration using this framework. The paired training data was obtained through manual registration by experts. Yan et al. (2018b) employed a discriminator to regularize the displacement field computed by the generator and found this approach to be more effective than the other regularizers in MR to TRUS registration. Mahapatra et al. (2018a) used CycleGAN for multi-modal (retinal) and uni-modal (MR) deformable registration where the generator produces both the transformed image and the deformation field. Mahapatra et al. (2018c) took one step further and explored the idea of joint segmentation and registration with CycleGAN and found their method performs better than the separate approaches for lung X-ray images. Tanner et al. (2018) employed CycleGAN for deformable image registration between MR and CT by first transforming the source domain image to the target domain and then employing a mono-modal image similarity measure for the registration. They found this method can achieve at best similar performance with the traditional multi-modal deformable registration methods.

3.7. Other works

In addition to the tasks described in the aforementioned sections, GANs have also been applied in other tasks discussed here. For instance, cGAN has been used for modelling patient specific motion distribution based on a single preoperative image (Hu et al., 2017c), highlighting regions most accountable for a disease (Baumgartner et al., 2017) and re-colorization of endoscopic video data (Ross et al., 2018). In (Mahmood et al., 2018) pix2pix was used for treatment planning in radiotherapy by predicting the dose distribution map from CT image. WGAN has also been used for modelling the progression of Alzheimer's disease (AD) in MRI. This is achieved by isolating the latent encoding of AD and performing arithmetic operation in the latent space (Bowles et al., 2018b).

4. Discussion

In the years 2017 and 2018, the number of studies applying GANs has risen significantly. The list of these papers reviewed for our study can be found on our [1]GitHub repository.

About 46% of these papers studied image synthesis, with cross modality image synthesis being the most important application of GANs. MR is ranked as the most common imaging modality explored in the GAN related literature. We believe one of the reasons for the significant interest in applying GANs for MR image analysis is due to the excessive amount of time spent on the acquisition of multiple sequences. GANs hold the potential to reduce MR acquisition time by faithfully generating certain sequences from already acquired ones. A recent study in image synthesis across different MR sequences using CollaGAN shows the irreplaceable nature of exogenous contrast sequence, but reports the synthesis of endogenous contrast such as T1, T2, from each other with high fidelity (Lee et al., 2019). A second reason for the popularity of GANs in MR might be because of large number of publicly available MR datasets as shown in Table 7.

[1]https://github.com/xinario/awesome-gan-for-medical-imaging

Another 37% of these studies fall into the group of reconstruction and segmentation due to the popularity of image-to-image translation frameworks. Adversarial training in these cases imposes a strong shape and texture regulation on the generator's output which makes it very promising in these two tasks. For example, in liver segmentation from 3D CT volumes, the incorporation of adversarial loss significantly improves the segmentation performance on non-contrast CT (has fuzzy liver boundary) than graph cut and CRF (Yang et al., 2017a).

Further 8% of these studies are related to classification. In these studies, the most effective use case was to combat domain shift. For the studies that used GAN for data augmentation in classification, most focused on generating tiny objects that can be easily aligned, such as nodules, lesions and cells. We believe it is partly due to the relatively smaller content variation of these images compared to the full context image which makes the training more stable with the current technique. Another reason might be related to the computation budget of the research since training on high resolution images requires a lot of GPU time. Although there are studies that applied GAN on synthesizing whole chest-X-ray (Madani et al., 2018a,b), the effectiveness has only been shown on fairly easy tasks, e.g. cardiac abnormality classification and on a medium size data regime, e.g. a couple of thousand images. With the advent of large volume labeled datasets, such as the CheXpert (Irvin et al., 2019), it seems there is diminishing return in the employment of GANs for image generation, especially for classification. We would like to argue that GANs are still useful in the following two cases. First, nowadays the training of a deep neural network heavily relies on data augmentation to improve the network's generalizability on unseen test data and reduce overfitting. However, existing data augmentation operations are all manually designed operations, e.g. rotation, color jittering, and can not cover the whole variation of the data. Cubuk et al. (2018) recently proposed to learn an augmentation policy with reinforcement learning but the search space still consisted of basic hand-crafted image processing operations. GANs, however, can allow us to sample the whole data distribution which offers much more flexibility in augmenting the training data (Bowles et al., 2018a). For example, styleGAN, is able to generate high resolution realistic face images with unprecedented level of details. This could be readily applied to chest X-ray datasets to generate images of a pathology class that has sufficient number of cases. Second, it is well known that medical data distribution is highly skewed with its largest mass centered on common diseases. It is impossible to accumulate enough training data for rare diseases, such as rheumatoid arthritis, sickle cell disease. But radiologists have been trained to detect these diseases in the long tail. Thus, another potential of GANs will be in synthesizing uncommon pathology cases, most likely through conditional generation with the conditioned information being specified by medical experts either through text description or hand drawn figures.

The remaining studies pertaining to detection, registration and other applications are so limited that it is hard to draw any conclusion.

4.1. Future challenges

Alongside many positive utilities of GANs, there are still challenges that need to be resolved for their employment to medical imaging. In image reconstruction and cross modality image synthesis, most works still adopt traditional shallow reference metrics such as MAE, PSNR, or SSIM for quantitative evaluation. These measures, however, do not correspond to the visual quality of the image. For example, direct optimization of pixel-wise loss produces a suboptimal (blurry) result but provides higher numbers than using adversarial loss. It becomes increasingly difficult to interpret these numbers in horizontal comparison of GAN-based works especially when extra losses as shown in Table 2 are incorporated. One way to alleviate this problem is to use down stream tasks such as segmentation or classification to validate the quality of the generated sample. Another way is to recruit domain experts but this approach is expensive, time consuming and hard to scale. Recently, Zhang et al. (2018b) proposed learned perceptual image path similarity (LPIPS), which outperforms previous metrics in terms of agreement with human judgements. It has been adopted in MedGAN (Armanious et al., 2018c) for evaluation of the generated image quality but it would be interesting to see its effectiveness for different types of medical images as compared to subjective measures from experienced human observers in a more extensive study. For natural images, the unconditional generated sample quality and diversity is usually measured by inception score (Salimans et al., 2016), the mean MS-SSIM metric among randomly chosen synthetic sample pairs (Odena et al., 2016), or Fréchet Inception distance (FID) (Heusel et al., 2017). The validity of these metrics for medical images remains to be explored.

Cross domain image-to-image translation can be achieved with both paired and unpaired training data and it offers many prospective applications in medical imaging as has already been seen in section 3.2.2. Unpaired training does not have the data fidelity loss term therefore there is no guarantee of preservation of small abnormality regions during the translation process. Cohen et al. (2018) warn against the use of generated images for direct interpretation by doctors. They observe that trained CycleGAN networks (for unpaired data) can be subject to bias due to matching the generated data to the distribution of the target domain. This system bias comes into being when target domain images in the training set have an over or under representation of certain classes. As an example of exploitation of this effect, Mirsky et al. (2019) demonstrate the possibility of malicious tampering of 3D medical imaging using 3D conditional GANs to remove and inject solitary pulmonary nodule into patient's CT scans. This system bias also exists in paired cross domain image-to-image translation with the data fidelity loss but only happens when the model was trained on normal images but tested on abnormal images. Cautions should be taken in training of the translation model and new methods should be proposed to faithfully preserve local abnormal regions.

4.2. Interesting future applications

Similar to other deep learning neural network models, various applications of GANs demonstrated in this paper have di-

rect bearing on improving radiology workflow and patient care. The strength of GANs however lies in their ability to learn in an unsupervised and/or weakly-supervised fashion. In particular, we perceive that image-to-image translation achieved by cGANs can have various other useful applications in medical imaging. For example, restoration of MR images acquired with certain artifacts such as motion, especially in a pediatric setting, may help reduce the number of repeated exams.

Exploring GANs for image captioning task (Dai et al., 2017a; Shetty et al., 2017; Melnyk et al., 2018; Fedus et al., 2018) may lead to semi-automatic generation of medical imaging reports (Jing et al., 2017) potentially reducing image reporting times. Success of adversarial text classification (Liu et al., 2017b) also prompts potential utility of GANs in improving performance of such systems for automatic MR protocol generation from free-text clinical indications (Sohn et al., 2017). Automated systems may improve MRI wait times which have been on the rise (CIHI, 2017) as well as enhance patient care. cGANs, specifically CycleGAN applications, such as makeup removal (Chang et al., 2018), can be extended to medical imaging with applications in improving bone x-ray images by removal of artifacts such as casts to facilitate enhanced viewing. This may aid radiologists in assessing fine bony detail, potentially allowing for enhanced detection of initially occult fractures and helping assess the progress of bone healing more efficiently. The success of GANs in unsupervised anomaly detection (Schlegl et al., 2017) can help achieve the task of detecting abnormalities in medical images in an unsupervised manner. This has the potential to be further extended for detection of implanted devices, e.g. staples, wires, tubes, pacemaker and artificial valves on X-rays. Such an algorithm can also be used for prioritizing radiologists' work lists, thus reducing the turnaround time for reporting critical findings (Gal Yaniv, 2018). We also expect to witness the utility of GANs in medical image synthesis from text descriptions (Bodnar, 2018), especially for rare cases, so as to fill in the gap of training samples required for training supervised neural networks for medical image classification tasks. The recent work on styleGAN shows the capability to control (Karras et al., 2018) the high level attributes of the synthesized image by manipulating the scale and bias parameters of the AdaIN layer (Huang and Belongie, 2017). Similarly, the SPADE (Park et al., 2019) controls the semantic layout of the synthesized image by a spatially adaptive normalization layer. Imagine in the future the desired attribute can be customized and specified in prior and manipulated in a localized fashion. We may then be able to predict the progression of disease, measure the impact of drug trial as suggested in Bowles et al. (2018b) but with more fine-grained controls.

Different imaging modalities work by exploiting tissue response to a certain physical media, such as x-rays or a magnetic field, and thus can provide complementary diagnostic information to each other. As a common practice in supervised deep learning, images of one modality type are labelled to train a network to accomplish a desired task. This process is repeated when switching modalities even if the underlying anatomical structure is the same, resulting in a waste of human effort. Adversarial training, or more specifically unpaired cross modality translation, enables reuse of the labels in all modalities and opens new ways for unsupervised transfer learning (Dou et al., 2018; Ying et al., 2019).

Finally, we would like to point out that, although there have many promising results reported in the literature, the adoption of GANs in medical imaging is still in its infancy and there is currently no breakthrough application as yet adopted clinically for GANs-based methods.

References

Abramian, D., Eklund, A., 2018. Refacing: reconstructing anonymized facial features using gans. arXiv preprint arXiv:1810.06455 .

Aerts, H., Rios Velazquez, E., Leijenaar, R.T., Parmar, C., Grossmann, P., Carvalho, S., Lambin, P., 2015. Data from nsclc-radiomics. The cancer imaging archive .

Alex, V., KP, M.S., Chennamsetty, S.S., Krishnamurthi, G., 2017. Generative adversarial networks for brain lesion detection, in: SPIE Medical Imaging, International Society for Optics and Photonics. pp. 101330G–101330G.

Appan, P., Sivaswamy, J., 2018. Retinal image synthesis for cad development, in: International Conference Image Analysis and Recognition, Springer. pp. 613–621.

Arjovsky, M., Chintala, S., Bottou, L., 2017. Wasserstein gan. arXiv preprint arXiv:1701.07875 .

Armanious, K., Küstner, T., Nikolaou, K., Gatidis, S., Yang, B., 2018a. Retrospective correction of rigid and non-rigid mr motion artifacts using gans. arXiv preprint arXiv:1809.06276 .

Armanious, K., Mecky, Y., Gatidis, S., Yang, B., 2018b. Adversarial inpainting of medical image modalities. arXiv preprint arXiv:1810.06621 .

Armanious, K., Yang, C., Fischer, M., Küstner, T., Nikolaou, K., Gatidis, S., Yang, B., 2018c. Medgan: Medical image translation using gans. arXiv preprint arXiv:1806.06397 .

Armato III, S.G., McLennan, G., Bidaut, L., McNitt-Gray, M.F., Meyer, C.R., Reeves, A.P., Clarke, L.P., 2015. Data from lidc-idri. the cancer imaging archive.

Ballerini, L., Fisher, R.B., Aldridge, B., Rees, J., 2013. A color and texture based hierarchical k-nn approach to the classification of non-melanoma skin lesions, in: Color Medical Image Analysis. Springer, pp. 63–86.

Balocco, S., Gatta, C., Ciompi, F., Wahle, A., Radeva, P., Carlier, S., Unal, G., Sanidas, E., Mauri, J., Carillo, X., et al., 2014. Standardized evaluation methodology and reference database for evaluating ivus image segmentation. Computerized medical imaging and graphics 38, 70–90.

Baumgartner, C.F., Koch, L.M., Tezcan, K.C., Ang, J.X., Konukoglu, E., 2017. Visual feature attribution using wasserstein gans. arXiv preprint arXiv:1711.08998 .

Baur, C., Albarqouni, S., Navab, N., 2018a. Generating highly realistic images of skin lesions with gans, in: OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis. Springer, pp. 260–267.

Baur, C., Albarqouni, S., Navab, N., 2018b. Melanogans: High resolution skin lesion synthesis with gans. arXiv preprint arXiv:1804.04338 .

Bayramoglu, N., Kaakinen, M., Eklund, L., Heikkila, J., 2017a. Towards virtual h&e staining of hyperspectral lung histology images using conditional generative adversarial networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 64–71.

Bayramoglu, N., Kaakinen, M., Eklund, L., Heikkila, J., 2017b. Towards virtual h&e staining of hyperspectral lung histology images using conditional generative adversarial networks, in: International Conference on Computer Vision.

Beers, A., Brown, J., Chang, K., Campbell, J.P., Ostmo, S., Chiang, M.F., Kalpathy-Cramer, J., 2018. High-resolution medical image synthesis using progressively grown generative adversarial networks. arXiv preprint arXiv:1805.03144 .

Ben-Cohen, A., Klang, E., Raskin, S.P., Soffer, S., Ben-Haim, S., Konen, E., Amitai, M.M., Greenspan, H., 2018. Cross-modality synthesis from ct to pet using fcn and gan networks for improved automated lesion detection. arXiv preprint arXiv:1802.07846 .

Bentaieb, A., Hamarneh, G., 2018. Adversarial stain transfer for histopathology image analysis. IEEE transactions on medical imaging 37, 792–802.

Bermudez, C., Plassard, A.J., Davis, L.T., Newton, A.T., Resnick, S.M., Landman, B.A., 2018. Learning implicit brain mri manifolds with deep learning, in: Medical Imaging 2018: Image Processing, International Society for Optics and Photonics. p. 105741L.

Berthelot, D., Schumm, T., Metz, L., 2017. Began: boundary equilibrium generative adversarial networks. arXiv preprint arXiv:1703.10717 .

Bi, L., Kim, J., Kumar, A., Feng, D., Fulham, M., 2017. Synthesis of positron emission tomography (pet) images via multi-channel generative adversarial networks (gans), in: Molecular Imaging, Reconstruction and Analysis of Moving Body Organs, and Stroke Imaging and Treatment. Springer, pp. 43–51.

Bodnar, C., 2018. Text to image synthesis using generative adversarial networks. arXiv preprint arXiv:1805.00676 .

Bowles, C., Chen, L., Guerrero, R., Bentley, P., Gunn, R., Hammers, A., Dickie, D.A., Hernández, M.V., Wardlaw, J., Rueckert, D., 2018a. Gan augmentation: Augmenting training data using generative adversarial networks. arXiv preprint arXiv:1810.10863 .

Bowles, C., Gunn, R., Hammers, A., Rueckert, D., 2018b. Modelling the progression of alzheimer's disease in mri using generative adversarial networks, in: Medical Imaging 2018: Image Processing, International Society for Optics and Photonics. p. 105741K.

Breuleux, O., Bengio, Y., Vincent, P., 2011. Quickly generating representative samples from an rbm-derived process. Neural computation 23, 2058–2073.

Brock, A., Donahue, J., Simonyan, K., 2018. Large scale gan training for high fidelity natural image synthesis. arXiv preprint arXiv:1809.11096 .

Budai, A., Bock, R., Maier, A., Hornegger, J., Michelson, G., 2013. Robust vessel segmentation in fundus images. International journal of biomedical imaging 2013.

Bullitt, E., Zeng, D., Gerig, G., Aylward, S., Joshi, S., Smith, J.K., Lin, W., Ewend, M.G., 2005. Vessel tortuosity and brain tumor malignancy: a blinded study1. Academic radiology 12, 1232–1240.

Calimeri, F., Marzullo, A., Stamile, C., Terracina, G., 2017. Biomedical data augmentation using generative adversarial neural networks, in: International Conference on Artificial Neural Networks, Springer. pp. 626–634.

Chang, H., Lu, J., Yu, F., Finkelstein, A., 2018. Pairedcyclegan: Asymmetric style transfer for applying and removing makeup, in: 2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Chartsias, A., Joyce, T., Dharmakumar, R., Tsaftaris, S.A., 2017. Adversarial image synthesis for unpaired multi-modal cardiac data, in: International Workshop on Simulation and Synthesis in Medical Imaging, Springer. pp. 3–13.

Chen, C., Dou, Q., Chen, H., Heng, P.A., 2018a. Semantic-aware generative adversarial nets for unsupervised domain adaptation in chest x-ray segmentation. arXiv preprint arXiv:1806.00600 .

Chen, X., Konukoglu, E., 2018. Unsupervised detection of lesions in brain mri using constrained adversarial auto-encoders. arXiv preprint arXiv:1806.04972 .

Chen, Y., Shi, F., Christodoulou, A.G., Xie, Y., Zhou, Z., Li, D., 2018b. Efficient and accurate mri super-resolution using a generative adversarial network and 3d multi-level densely connected network, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 91–99.

Choi, H., Lee, D.S., 2017. Generation of structural mr images from amyloid pet: Application to mr-less quantification. Journal of nuclear medicine: official publication, Society of Nuclear Medicine .

Chuquicusma, M.J., Hussein, S., Burt, J., Bagci, U., 2017. How to fool radiologists with generative adversarial networks? a visual turing test for lung cancer diagnosis. arXiv preprint arXiv:1710.09762 .

CIHI, 2017. Wait times for priority procedures in canada. https://secure.cihi.ca/free_products/wait-times-report-2017_en.pdf. Accessed: 2018-10-17.

Clinical Practice Committee, D.S.o.t., 2000. Informed consent for medical photographs. Genetics in Medicine 2, 353.

Codella, N.C., Gutman, D., Celebi, M.E., Helba, B., Marchetti, M.A., Dusza, S.W., Kalloo, A., Liopyris, K., Mishra, N., Kittler, H., et al., 2018. Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic), in: Biomedical Imaging (ISBI 2018), 2018 IEEE 15th International Symposium on, IEEE. pp. 168–172.

Cohen, J.P., Luck, M., Honari, S., 2018. Distribution matching losses can hallucinate features in medical image translation. arXiv preprint arXiv:1805.08841 .

Costa, P., Galdran, A., Meyer, M.I., Abràmoff, M.D., Niemeijer, M., Mendonça, A.M., Campilho, A., 2017a. Towards adversarial retinal image synthesis. arXiv preprint arXiv:1701.08974 .

Costa, P., Galdran, A., Meyer, M.I., Niemeijer, M., Abràmoff, M., Mendonça, A.M., Campilho, A., 2017b. End-to-end adversarial retinal image synthesis. IEEE Transactions on Medical Imaging .

Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., Bharath, A.A., 2018. Generative adversarial networks: An overview. IEEE Signal Processing Magazine 35, 53–65.

Crimi, A., Menze, B., Maier, O., Reyes, M., Handels, H., 2016. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: First International Workshop, Brainles 2015, Held in Conjunction with MICCAI 2015, Munich, Germany, October 5, 2015, Revised Selected Papers. volume 9556. Springer.

Cubuk, E.D., Zoph, B., Mane, D., Vasudevan, V., Le, Q.V., 2018. Autoaugment: Learning augmentation policies from data. arXiv preprint arXiv:1805.09501 .

Dai, B., Lin, D., Urtasun, R., Fidler, S., 2017a. Towards diverse and natural image descriptions via a conditional gan. arXiv preprint arXiv:1703.06029 .

Dai, W., Doyle, J., Liang, X., Zhang, H., Dong, N., Li, Y., Xing, E.P., 2017b. Scan: Structure correcting adversarial network for chest x-rays organ segmentation. arXiv preprint arXiv:1703.08770 .

Dar, S.U.H., Yurt, M., Karacan, L., Erdem, A., Erdem, E., Çukur, T., 2018a. Image synthesis in multi-contrast mri with conditional generative adversarial networks. arXiv preprint arXiv:1802.01221 .

Dar, S.U.H., Yurt, M., Shahdloo, M., Ildız, M.E., Çukur, T., 2018b. Synergistic reconstruction and synthesis via generative adversarial networks for accelerated multi-contrast mri. arXiv preprint arXiv:1805.10704 .

Decencière, E., Zhang, X., Cazuguel, G., Lay, B., Cochener, B., Trone, C., Gain, P., Ordonez, R., Massin, P., Erginay, A., Charton, B., Klein, J.C., 2014. Feedback on a publicly distributed database: the messidor database. Image Analysis & Stereology 33, 231–234. URL: http://www.ias-iss.org/ojs/IAS/article/view/1155, doi:10.5566/ias.1155.

Denton, E.L., Chintala, S., Fergus, R., et al., 2015. Deep generative image models using a laplacian pyramid of adversarial networks, in: Advances in neural information processing systems, pp. 1486–1494.

Donahue, J., Krähenbühl, P., Darrell, T., 2016. Adversarial feature learning. arXiv preprint arXiv:1605.09782 .

Dou, Q., Ouyang, C., Chen, C., Chen, H., Heng, P.A., 2018. Unsupervised cross-modality domain adaptation of convnets for biomedical image segmentations with adversarial loss. arXiv preprint arXiv:1804.10916 .

Dumoulin, V., Belghazi, I., Poole, B., Lamb, A., Arjovsky, M., Mastropietro, O., Courville, A., 2016. Adversarially learned inference. arXiv preprint arXiv:1606.00704 .

Emami, H., Dong, M., Nejad-Davarani, S.P., Glide-Hurst, C., 2018. Generating synthetic cts from magnetic resonance images using generative adversarial networks. Medical physics .

Fan, J., Cao, X., Xue, Z., Yap, P.T., Shen, D., 2018. Adversarial similarity network for evaluating image alignment in deep learning based registration, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 739–746.

Fedus, W., Goodfellow, I., Dai, A.M., 2018. Maskgan: Better text generation via filling in the _. arXiv preprint arXiv:1801.07736 .

Finlayson, S.G., Lee, H., Kohane, I.S., Oakden-Rayner, L., 2018. Towards generative adversarial networks as a new paradigm for radiology education. arXiv preprint arXiv:1812.01547 .

Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J., Greenspan, H., 2018. Gan-based synthetic medical image augmentation for increased cnn performance in liver lesion classification. arXiv preprint arXiv:1803.01229 .

Fukushima, K., Miyake, S., 1982. Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition, in: Competition and cooperation in neural nets. Springer, pp. 267–285.

Fumero, F., Alayón, S., Sanchez, J.L., Sigut, J., Gonzalez-Hernandez, M., 2011. Rim-one: An open retinal image database for optic nerve evaluation, in: 2011 24th international symposium on computer-based medical systems (CBMS), IEEE. pp. 1–6.

Gal Yaniv, Anna Kuperberg, E.W., 2018. Deep learning algorithm for optimizing critical findings report turnaround time, in: SIIM.

Galbusera, F., Niemeyer, F., Seyfried, M., Bassani, T., Casaroli, G., Kienle, A., Wilke, H.J., 2018. Exploring the potential of generative adversarial networks for synthesizing radiological images of the spine to be used in in silico trials. Frontiers in Bioengineering and Biotechnology 6, 53.

GANs, W., 2018. Sparse-view ct reconstruction using. Machine Learning for Medical Image Reconstruction 11074, 75.

Glocker, B., Zikic, D., Konukoglu, E., Haynor, D.R., Criminisi, A., 2013. Vertebrae localization in pathological spine ct via dense classification from sparse annotations, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 262–270.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative adversarial nets, in: Advances in neural information processing systems, pp. 2672–2680.

Gu, X., Knutsson, H., Eklund, A., 2018. Generating diffusion mri scalar maps from t1 weighted images using generative adversarial networks. arXiv preprint arXiv:1810.02683 .

Guibas, J.T., Virdi, T.S., Li, P.S., 2017. Synthetic medical images from dual generative adversarial networks. arXiv preprint arXiv:1709.01872 .

Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., Courville, A., 2017. Improved training of wasserstein gans. arXiv preprint arXiv:1704.00028 .

Gutman, D., Codella, N.C., Celebi, E., Helba, B., Marchetti, M., Mishra, N., Halpern, A., 2016. Skin lesion analysis toward melanoma detection: A challenge at the international symposium on biomedical imaging (isbi) 2016, hosted by the international skin imaging collaboration (isic). arXiv preprint arXiv:1605.01397 .

Han, C., Hayashi, H., Rundo, L., Araki, R., Shimoda, W., Muramatsu, S., Furukawa, Y., Mauri, G., Nakayama, H., 2018a. Gan-based synthetic brain mr image generation, in: Biomedical Imaging (ISBI 2018), 2018 IEEE 15th International Symposium on, IEEE. pp. 734–738.

Han, Z., Wei, B., Mercado, A., Leung, S., Li, S., 2018b. Spine-gan: Semantic segmentation of multiple spinal structures. Medical image analysis 50, 23–35.

Heath, M., Bowyer, K., Kopans, D., Kegelmeyer, P., Moore, R., Chang, K., Munishkumaran, S., 1998. Current status of the digital database for screening mammography, in: Digital mammography. Springer, pp. 457–460.

Heinrich, M.P., Jenkinson, M., Bhushan, M., Matin, T., Gleeson, F.V., Brady, M., Schnabel, J.A., 2012. Mind: Modality independent neighbourhood descriptor for multi-modal deformable registration. Medical image analysis 16, 1423–1435.

Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., Klambauer, G., Hochreiter, S., 2017. Gans trained by a two time-scale update rule converge to a nash equilibrium. arXiv preprint arXiv:1706.08500 .

Hiasa, Y., Otake, Y., Takao, M., Matsuoka, T., Takashima, K., Prince, J.L., Sugano, N., Sato, Y., 2018. Cross-modality image synthesis from unpaired data using cyclegan: Effects of gradient consistency loss and training data size. arXiv preprint arXiv:1803.06629 .

Hobson, P., Lovell, B.C., Percannella, G., Vento, M., Wiliem, A., 2015. Benchmarking human epithelial type 2 interphase cells classification methods on a very large dataset. Artificial intelligence in medicine 65, 239–250.

Hou, L., Agarwal, A., Samaras, D., Kurc, T.M., Gupta, R.R., Saltz, J.H., 2017. Unsupervised histopathology image synthesis. arXiv preprint arXiv:1712.05021 .

Hu, B., Tang, Y., Chang, E.I., Fan, Y., Lai, M., Xu, Y., et al., 2017a. Unsupervised learning for cell-level visual representation in histopathology images with generative adversarial networks. arXiv preprint arXiv:1711.11317 .

Hu, X., Chung, A.G., Fieguth, P., Khalvati, F., Haider, M.A., Wong, A., 2018. Prostategan: Mitigating data bias via prostate diffusion imaging synthesis with generative adversarial networks. arXiv preprint arXiv:1811.05817 .

Hu, Y., Gibson, E., Lee, L.L., Xie, W., Barratt, D.C., Vercauteren, T., Noble, J.A., 2017b. Freehand ultrasound image simulation with spatially-conditioned generative adversarial networks, in: Molecular Imaging, Reconstruction and Analysis of Moving Body Organs, and Stroke Imaging and Treatment. Springer, pp. 105–115.

Hu, Y., Gibson, E., Vercauteren, T., Ahmed, H.U., Emberton, M., Moore, C.M., Noble, J.A., Barratt, D.C., 2017c. Intraoperative organ motion models with an ensemble of conditional generative adversarial networks, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 368–376.

Huang, H., Yu, P.S., Wang, C., 2018. An introduction to image synthesis with generative adversarial nets. arXiv preprint arXiv:1803.04469 .

Huang, X., Belongie, S., 2017. Arbitrary style transfer in real-time with adaptive instance normalization, in: Proceedings of the IEEE International Conference on Computer Vision, pp. 1501–1510.

Huang, X., Li, Y., Poursaeed, O., Hopcroft, J.E., Belongie, S.J., 2017. Stacked generative adversarial networks., in: CVPR, p. 3.

Huo, Y., Xu, Z., Bao, S., Assad, A., Abramson, R.G., Landman, B.A., 2017. Adversarial synthesis learning enables segmentation without target modality ground truth. arXiv preprint arXiv:1712.07695 .

Huo, Y., Xu, Z., Bao, S., Bermudez, C., Plassard, A.J., Liu, J., Yao, Y., Assad, A., Abramson, R.G., Landman, B.A., 2018a. Splenomegaly segmentation using global convolutional kernels and conditional generative adversarial networks, in: Medical Imaging 2018: Image Processing, International Society for Optics and Photonics. p. 1057409.

Huo, Y., Xu, Z., Moon, H., Bao, S., Assad, A., Moyo, T.K., Savona, M.R., Abramson, R.G., Landman, B.A., 2018b. Synseg-net: Synthetic segmen-

tation without target modality ground truth. IEEE transactions on medical imaging .

Ioffe, S., Szegedy, C., 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167 .

Iqbal, T., Ali, H., 2018. Generative adversarial network for medical images (mi-gan). Journal of medical systems 42, 231.

Irvin, J., Rajpurkar, P., Ko, M., Yu, Y., Ciurea-Ilcus, S., Chute, C., Marklund, H., Haghgoo, B., Ball, R., Shpanskaya, K., et al., 2019. Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison. arXiv preprint arXiv:1901.07031 .

Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A., 2016. Image-to-image translation with conditional adversarial networks. arXiv preprint arXiv:1611.07004 .

Izadi, S., Mirikharaji, Z., Kawahara, J., Hamarneh, G., 2018. Generative adversarial networks to segment skin lesions, in: Biomedical Imaging (ISBI 2018), 2018 IEEE 15th International Symposium on, IEEE. pp. 881–884.

Jaderberg, M., Simonyan, K., Zisserman, A., et al., 2015. Spatial transformer networks, in: Advances in neural information processing systems, pp. 2017–2025.

Jaeger, S., Candemir, S., Antani, S., Wáng, Y.X.J., Lu, P.X., Thoma, G., 2014. Two public chest x-ray datasets for computer-aided screening of pulmonary diseases. Quantitative imaging in medicine and surgery 4, 475.

Jiang, J., Hu, Y.C., Tyagi, N., Zhang, P., Rimner, A., Mageras, G.S., Deasy, J.O., Veeraraghavan, H., 2018. Tumor-aware, adversarial domain adaptation from ct to mri for lung cancer segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 777–785.

Jin, C.B., Jung, W., Joo, S., Park, E., Saem, A.Y., Han, I.H., Lee, J.I., Cui, X., 2018a. Deep ct to mr synthesis using paired and unpaired data. arXiv preprint arXiv:1805.10790 .

Jin, D., Xu, Z., Tang, Y., Harrison, A.P., Mollura, D.J., 2018b. Ct-realistic lung nodule simulation from 3d conditional generative adversarial networks for robust lung segmentation. arXiv preprint arXiv:1806.04051 .

Jing, B., Xie, P., Xing, E., 2017. On the automatic generation of medical imaging reports. arXiv preprint arXiv:1711.08195 .

Kainz, P., Urschler, M., Schulter, S., Wohlhart, P., Lepetit, V., 2015. You should use regression to detect cells, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 276–283.

Kälviäinen, R., Uusitalo, H., 2007. Diaretdb1 diabetic retinopathy database and evaluation protocol, in: Medical Image Understanding and Analysis, Citeseer. p. 61.

Kamnitsas, K., Baumgartner, C., Ledig, C., Newcombe, V., Simpson, J., Kane, A., Menon, D., Nori, A., Criminisi, A., Rueckert, D., et al., 2017. Unsupervised domain adaptation in brain lesion segmentation with adversarial networks, in: International Conference on Information Processing in Medical Imaging, Springer. pp. 597–609.

Kang, E., Koo, H.J., Yang, D.H., Seo, J.B., Ye, J.C., 2018. Cycle consistent adversarial denoising network for multiphase coronary ct angiography. arXiv preprint arXiv:1806.09748 .

Karras, T., Aila, T., Laine, S., Lehtinen, J., 2017. Progressive growing of gans for improved quality, stability, and variation. arXiv preprint arXiv:1710.10196 .

Karras, T., Laine, S., Aila, T., 2018. A style-based generator architecture for generative adversarial networks. arXiv preprint arXiv:1812.04948 .

Kim, K.H., Do, W.J., Park, S.H., 2018. Improving resolution of mr images with an adversarial network incorporating images with different contrast. Medical physics .

Kim, T., Cha, M., Kim, H., Lee, J.K., Kim, J., 2017. Learning to discover cross-domain relations with generative adversarial networks. arXiv preprint arXiv:1703.05192 .

Kinahan, P.E., Fletcher, J.W., 2010. Positron emission tomography-computed tomography standardized uptake values in clinical practice and assessing response to therapy, in: Seminars in Ultrasound, CT and MRI, Elsevier. pp. 496–505.

Köbel, M., Kalloger, S.E., Baker, P.M., Ewanowich, C.A., Arseneau, J., Zherebitskiy, V., Abdulkarim, S., Leung, S., Duggan, M.A., Fontaine, D., et al., 2010. Diagnosis of ovarian carcinoma cell type is highly reproducible: a transcanadian study. The American journal of surgical pathology 34, 984–993.

Kohl, S., Bonekamp, D., Schlemmer, H.P., Yaqubi, K., Hohenfellner, M., Hadaschik, B., Radtke, J.P., Maier-Hein, K., 2017. Adversarial networks for the detection of aggressive prostate cancer. arXiv preprint arXiv:1702.08014 .

Köhler, T., Budai, A., Kraus, M.F., Odstrčilik, J., Michelson, G., Hornegger, J., 2013. Automatic no-reference quality assessment for retinal fundus images using vessel segmentation, in: Computer-Based Medical Systems (CBMS), 2013 IEEE 26th International Symposium on, IEEE. pp. 95–100.

Korkinof, D., Rijken, T., O'Neill, M., Yearsley, J., Harvey, H., Glocker, B., 2018. High-resolution mammogram synthesis using progressive generative adversarial networks. arXiv preprint arXiv:1807.03401 .

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks, in: Advances in neural information processing systems, pp. 1097–1105.

Kurach, K., Lucic, M., Zhai, X., Michalski, M., Gelly, S., 2018. The gan landscape: Losses, architectures, regularization, and normalization .

Lahiri, A., Ayush, K., Biswas, P.K., Mitra, P., 2017. Generative adversarial learning for reducing manual annotation in semantic segmentation on large scale miscroscopy images: Automated vessel segmentation in retinal fundus image as test case, in: Conference on Computer Vision and Pattern Recognition Workshops, pp. 42–48.

Lahiri, A., Jain, V., Mondal, A., Biswas, P.K., 2018. Retinal vessel segmentation under extreme low annotation: A generative adversarial network approach. arXiv preprint arXiv:1809.01348 .

Larsen, A.B.L., Sønderby, S.K., Larochelle, H., Winther, O., 2015. Autoencoding beyond pixels using a learned similarity metric. arXiv preprint arXiv:1512.09300 .

Lau, F., Hendriks, T., Lieman-Sifry, J., Sall, S., Golden, D., 2018. Scargan: chained generative adversarial networks to simulate pathological tissue on cardiovascular mr scans, in: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, pp. 343–350.

Lecouat, B., Chang, K., Foo, C.S., Unnikrishnan, B., Brown, J.M., Zenati, H., Beers, A., Chandrasekhar, V., Kalpathy-Cramer, J., Krishnaswamy, P., 2018. Semi-supervised deep learning for abnormality classification in retinal images. arXiv preprint arXiv:1812.07832 .

Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., et al., 2017. Photo-realistic single image super-resolution using a generative adversarial network. arXiv preprint .

Lee, D., Moon, W.J., Ye, J.C., 2019. Which contrast does matter? towards a deep understanding of mr contrast using collaborative gan. arXiv preprint arXiv:1905.04105 .

Lee Rodgers, J., Nicewander, W.A., 1988. Thirteen ways to look at the correlation coefficient. The American Statistician 42, 59–66.

Li, Y., Shen, L., 2018. cc-gan: A robust transfer-learning framework for hep-2 specimen image segmentation. IEEE Access 6, 14048–14058.

Li, Z., Wang, Y., Yu, J., 2017. Brain tumor segmentation using an adversarial network, in: International MICCAI Brainlesion Workshop, Springer. pp. 123–132.

Liao, H., Huo, Z., Sehnert, W.J., Zhou, S.K., Luo, J., 2018. Adversarial sparse-view cbct artifact reduction, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 154–162.

Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., van der Laak, J.A., van Ginneken, B., Sánchez, C.I., 2017. A survey on deep learning in medical image analysis. arXiv preprint arXiv:1702.05747 .

Liu, F., 2018. Susan: segment unannotated image structure using adversarial network. Magnetic resonance in medicine .

Liu, M.Y., Breuel, T., Kautz, J., 2017a. Unsupervised image-to-image translation networks, in: Advances in Neural Information Processing Systems, pp. 700–708.

Liu, P., Qiu, X., Huang, X., 2017b. Adversarial multi-task learning for text classification. arXiv preprint arXiv:1704.05742 .

Liu, Z., Bicer, T., Kettimuthu, R., Gursoy, D., De Carlo, F., Foster, I., 2019. Tomogan: Low-dose x-ray tomography with generative adversarial networks. arXiv preprint arXiv:1902.07582 .

Low, D.A., 2010. Gamma dose distribution evaluation tool, in: Journal of Physics-Conference Series, p. 012071.

Maas, A.L., Hannun, A.Y., Ng, A.Y., 2013. Rectifier nonlinearities improve neural network acoustic models, in: Proc. icml, p. 3.

Madani, A., Moradi, M., Karargyris, A., Syeda-Mahmood, T., 2018a. Chest x-ray generation and data augmentation for cardiovascular abnormality classification, in: Medical Imaging 2018: Image Processing, International Society for Optics and Photonics. p. 105741M.

Madani, A., Moradi, M., Karargyris, A., Syeda-Mahmood, T., 2018b. Semi-supervised learning with generative adversarial networks for chest x-ray classification with ability of data domain adaptation, in: Biomedical Imaging (ISBI 2018), 2018 IEEE 15th International Symposium on, IEEE. pp. 1038–1042.

Mahapatra, D., 2017. Retinal vasculature segmentation using local saliency maps and generative adversarial networks for image super resolution. arXiv preprint arXiv:1710.04783 .

Mahapatra, D., Antony, B., Sedai, S., Garnavi, R., 2018a. Deformable medical image registration using generative adversarial networks, in: Biomedical Imaging (ISBI 2018), 2018 IEEE 15th International Symposium on, IEEE. pp. 1449–1453.

Mahapatra, D., Bozorgtabar, B., Thiran, J.P., Reyes, M., 2018b. Efficient active learning for image classification and segmentation using a sample selection and conditional generative adversarial network. arXiv preprint arXiv:1806.05473 .

Mahapatra, D., Ge, Z., Sedai, S., Chakravorty, R., 2018c. Joint registration and segmentation of xray images using generative adversarial networks, in: International Workshop on Machine Learning in Medical Imaging, Springer. pp. 73–80.

Mahmood, F., Chen, R., Durr, N.J., 2017. Unsupervised reverse domain adaptation for synthetic medical images via adversarial training. arXiv preprint arXiv:1711.06606 .

Mahmood, R., Babier, A., McNiven, A., Diamant, A., Chan, T.C., 2018. Automated treatment planning in radiation therapy using generative adversarial networks. arXiv preprint arXiv:1807.06489 .

Mao, X., Li, Q., Xie, H., Lau, R.Y., Wang, Z., 2016. Least squares generative adversarial networks. arXiv preprint ArXiv:1611.04076 .

Mardani, M., Gong, E., Cheng, J.Y., Vasanawala, S., Zaharchuk, G., Alley, M., Thakur, N., Han, S., Dally, W., Pauly, J.M., et al., 2017. Deep generative adversarial networks for compressed sensing automates mri. arXiv preprint arXiv:1706.00051 .

Maspero, M., Savenije, M.H., Dinkla, A.M., Seevinck, P.R., Intven, M.P., Jurgenliemk-Schulz, I.M., Kerkmeijer, L.G., van den Berg, C.A., 2018. Dose evaluation of fast synthetic-ct generation using a generative adversarial network for general pelvis mr-only radiotherapy. Physics in Medicine & Biology 63, 185001.

Matkovic, K., Neumann, L., Neumann, A., Psik, T., Purgathofer, W., 2005. Global contrast factor-a new approach to image contrast. Computational Aesthetics 2005, 159–168.

McCollough, C.H., Bartley, A.C., Carter, R.E., Chen, B., Drees, T.A., Edwards, P., Holmes, D.R., Huang, A.E., Khan, F., Leng, S., et al., 2017. Low-dose ct for the detection and classification of metastatic liver lesions: Results of the 2016 low dose ct grand challenge. Medical physics 44.

Melnyk, I., Sercu, T., Dognin, P.L., Ross, J., Mroueh, Y., 2018. Improved image captioning with adversarial semantic alignment. arXiv preprint arXiv:1805.00063 .

Mendonca, T., Celebi, M., Mendonca, T., Marques, J., 2015. Ph2: A public database for the analysis of dermoscopic images. Dermoscopy Image Analysis .

Milletari, F., Navab, N., Ahmadi, S.A., 2016. V-net: Fully convolutional neural networks for volumetric medical image segmentation, in: 3D Vision (3DV), 2016 Fourth International Conference on, IEEE. pp. 565–571.

Mirsky, Y., Mahler, T., Shelef, I., Elovici, Y., 2019. Ct-gan: Malicious tampering of 3d medical imagery using deep learning. arXiv preprint arXiv:1901.03597 .

Mirza, M., Osindero, S., 2014. Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784 .

Miyato, T., Kataoka, T., Koyama, M., Yoshida, Y., 2018. Spectral normalization for generative adversarial networks. arXiv preprint arXiv:1802.05957 .

Miyato, T., Koyama, M., 2018. cgans with projection discriminator. arXiv preprint arXiv:1802.05637 .

Moeskops, P., Veta, M., Lafarge, M.W., Eppenhof, K.A., Pluim, J.P., 2017. Adversarial training and dilated convolutions for brain mri segmentation, in: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, pp. 56–64.

Mok, T.C., Chung, A.C., 2018. Learning data augmentation for brain tumor segmentation with coarse-to-fine generative adversarial networks. arXiv preprint arXiv:1805.11291 .

Mondal, A.K., Dolz, J., Desrosiers, C., 2018. Few-shot 3d multi-modal medical image segmentation using generative adversarial learning. arXiv preprint arXiv:1810.12241 .

Moreira, I.C., Amaral, I., Domingues, I., Cardoso, A., Cardoso, M.J., Cardoso, J.S., 2012. Inbreast: toward a full-field digital mammographic database. Academic radiology 19, 236–248.

Nie, D., Trullo, R., Lian, J., Petitjean, C., Ruan, S., Wang, Q., Shen, D., 2017. Medical image synthesis with context-aware generative adversarial networks, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 417–425.

Nie, D., Trullo, R., Lian, J., Wang, L., Petitjean, C., Ruan, S., Wang, Q., Shen, D., 2018. Medical image synthesis with deep convolutional adversarial networks. IEEE Transactions on Biomedical Engineering .

Niemeijer, M., Abramoff, M.D., van Ginneken, B., 2006. Image structure clustering for image quality verification of color retina images in diabetic retinopathy screening. Medical image analysis 10, 888–898.

Nowozin, S., Cseke, B., Tomioka, R., 2016. f-gan: Training generative neural samplers using variational divergence minimization, in: Advances in Neural Information Processing Systems, pp. 271–279.

Odena, A., Olah, C., Shlens, J., 2016. Conditional image synthesis with auxiliary classifier gans. arXiv preprint arXiv:1610.09585 .

Oh, D.Y., Yun, I.D., 2018. Learning bone suppression from dual energy chest x-rays using adversarial networks. arXiv preprint arXiv:1811.02628 .

Oksuz, I., Clough, J., Bustin, A., Cruz, G., Prieto, C., Botnar, R., Rueckert, D., Schnabel, J.A., King, A.P., 2018. Cardiac mr motion artefact correction from

k-space using deep learning-based reconstruction, in: International Workshop on Machine Learning for Medical Image Reconstruction, Springer. pp. 21–29.

Olut, S., Sahin, Y.H., Demir, U., Unal, G., 2018. Generative adversarial training for mra image synthesis using multi-contrast mri. arXiv preprint arXiv:1804.04366 .

Pace, D.F., Dalca, A.V., Geva, T., Powell, A.J., Moghari, M.H., Golland, P., 2015. Interactive whole-heart segmentation in congenital heart disease, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 80–88.

Pan, Y., Liu, M., Lian, C., Zhou, T., Xia, Y., Shen, D., 2018. Synthesizing missing pet from mri with cycle-consistent generative adversarial networks for alzheimer's disease diagnosis, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 455–463.

Park, T., Liu, M.Y., Wang, T.C., Zhu, J.Y., 2019. Semantic image synthesis with spatially-adaptive normalization. arXiv preprint arXiv:1903.07291 .

Pathak, D., Krahenbuhl, P., Donahue, J., Darrell, T., Efros, A.A., 2016. Context encoders: Feature learning by inpainting, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2536–2544.

Pluim, J.P., Maintz, J.A., Viergever, M.A., 2003. Mutual-information-based registration of medical images: a survey. IEEE transactions on medical imaging 22, 986–1004.

Prentasic, P., Loncaric, S., Vatavuk, Z., Bencic, G., Subasic, M., Petkovic, T., Dujmovic, L., Malenica-Ravlic, M., Budimlija, N., Tadic, R., 2013. Diabetic retinopathy image database (dridb): a new database for diabetic retinopathy screening programs research, in: Image and Signal Processing and Analysis (ISPA), 2013 8th International Symposium on, IEEE. pp. 711–716.

Quan, T.M., Nguyen-Duc, T., Jeong, W.K., 2018. Compressed sensing mri reconstruction using a generative adversarial network with a cyclic loss. IEEE transactions on medical imaging 37, 1488–1497.

Radford, A., Metz, L., Chintala, S., 2015. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434 .

Ran, M., Hu, J., Chen, Y., Chen, H., Sun, H., Zhou, J., Zhang, Y., 2018. Denoising of 3-d magnetic resonance images using a residual encoder-decoder wasserstein generative adversarial network. arXiv preprint arXiv:1808.03941 .

Ravì, D., Szczotka, A.B., Shakir, D.I., Pereira, S.P., Vercauteren, T., 2018. Adversarial training with cycle consistency for unsupervised super-resolution in endomicroscopy .

Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B., Lee, H., 2016a. Generative adversarial text to image synthesis, in: Proceedings of The 33rd International Conference on Machine Learning.

Reed, S.E., Akata, Z., Mohan, S., Tenka, S., Schiele, B., Lee, H., 2016b. Learning what and where to draw, in: Advances in Neural Information Processing Systems, pp. 217–225.

Ren, J., Hacihaliloglu, I., Singer, E.A., Foran, D.J., Qi, X., 2018. Adversarial domain adaptation for classification of prostate histopathology whole-slide images. arXiv preprint arXiv:1806.01357 .

Resnick, S.M., Pham, D.L., Kraut, M.A., Zonderman, A.B., Davatzikos, C., 2003. Longitudinal magnetic resonance imaging studies of older adults: a shrinking brain. Journal of Neuroscience 23, 3295–3301.

Rezaei, M., Harmuth, K., Gierke, W., Kellermeier, T., Fischer, M., Yang, H., Meinel, C., 2017. A conditional adversarial network for semantic segmentation of brain tumor, in: International MICCAI Brainlesion Workshop, Springer. pp. 241–252.

Rezaei, M., Yang, H., Meinel, C., 2018a. Conditional generative refinement adversarial networks for unbalanced medical image semantic segmentation. arXiv preprint arXiv:1810.03871 .

Rezaei, M., Yang, H., Meinel, C., 2018b. Whole heart and great vessel segmentation with context-aware of generative adversarial networks, in: Bildverarbeitung für die Medizin 2018. Springer, pp. 353–358.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical image computing and computer-assisted intervention, Springer. pp. 234–241.

Ross, T., Zimmerer, D., Vemuri, A., Isensee, F., Wiesenfarth, M., Bodenstedt, S., Both, F., Kessler, P., Wagner, M., Müller, B., et al., 2018. Exploiting the potential of unlabeled endoscopic video data with self-supervised learning. International journal of computer assisted radiology and surgery , 1–9.

Salehinejad, H., Valaee, S., Dowdell, T., Colak, E., Barfett, J., 2017. Generalization of deep neural networks for chest pathology classification in x-rays using generative adversarial networks. arXiv preprint arXiv:1712.01636 .

Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., Chen, X., 2016. Improved techniques for training gans, in: Advances in Neural Information Processing Systems, pp. 2226–2234.

Sanchez, I., Vilaplana, V., 2018. Brain mri super-resolution using 3d generative adversarial networks .

Sangkloy, P., Lu, J., Fang, C., Yu, F., Hays, J., 2016. Scribbler: Controlling deep image synthesis with sketch and color. arXiv preprint arXiv:1612.00835 .

Schlegl, T., Seeböck, P., Waldstein, S.M., Schmidt-Erfurth, U., Langs, G., 2017. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery, in: International Conference on Information Processing in Medical Imaging, Springer. pp. 146–157.

Seitzer, M., Yang, G., Schlemper, J., Oktay, O., Würfl, T., Christlein, V., Wong, T., Mohiaddin, R., Firmin, D., Keegan, J., et al., 2018. Adversarial and perceptual refinement for compressed sensing mri reconstruction, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 232–240.

Sekuboyina, A., Rempfler, M., Kukačka, J., Tetteh, G., Valentinitsch, A., Kirschke, J.S., Menze, B.H., 2018. Btrfly net: Vertebrae labelling with energy-based adversarial learning of local spine prior. arXiv preprint arXiv:1804.01307 .

Senaras, C., Niazi, M.K.K., Sahiner, B., Pennell, M.P., Tozbikian, G., Lozanski, G., Gurcan, M.N., 2018. Optimized generation of high-resolution phantom images using cgan: Application to quantification of ki67 breast cancer images. PloS one 13, e0196846.

Shaban, M.T., Baur, C., Navab, N., Albarqouni, S., 2018. Staingan: Stain style transfer for digital histological images. arXiv preprint arXiv:1804.01601 .

Shan, H., Zhang, Y., Yang, Q., Kruger, U., Kalra, M.K., Sun, L., Cong, W., Wang, G., 2018. 3-d convolutional encoder-decoder network for low-dose ct via transfer learning from a 2-d trained network. IEEE transactions on medical imaging 37, 1522–1534.

Shankaranarayana, S.M., Ram, K., Mitra, K., Sivaprakasam, M., 2017. Joint optic disc and cup segmentation using fully convolutional and adversarial networks, in: Fetal, Infant and Ophthalmic Medical Image Analysis. Springer, pp. 168–176.

Sheikh, H.R., Bovik, A.C., 2004. Image information and visual quality, in: Acoustics, Speech, and Signal Processing, 2004. Proceedings.(ICASSP'04). IEEE International Conference on, IEEE. pp. iii–709.

Sheikh, H.R., Bovik, A.C., De Veciana, G., 2005. An information fidelity criterion for image quality assessment using natural scene statistics. IEEE Transactions on image processing 14, 2117–2128.

Shetty, R., Rohrbach, M., Hendricks, L.A., Fritz, M., Schiele, B., 2017. Speaking the same language: Matching machine to human captions by adversarial training, in: Proceedings of the IEEE International Conference on Computer Vision (ICCV).

Shin, H.C., Tenenholtz, N.A., Rogers, J.K., Schwarz, C.G., Senjem, M.L., Gunter, J.L., Andriole, K.P., Michalski, M., 2018. Medical image synthesis for data augmentation and anonymization using generative adversarial networks, in: International Workshop on Simulation and Synthesis in Medical Imaging, Springer. pp. 1–11.

Shiraishi, J., Katsuragawa, S., Ikezoe, J., Matsumoto, T., Kobayashi, T., Komatsu, K.i., Matsui, M., Fujita, H., Kodera, Y., Doi, K., 2000. Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules. American Journal of Roentgenology 174, 71–74.

Shitrit, O., Raviv, T.R., 2017. Accelerated magnetic resonance imaging by adversarial neural network, in: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, pp. 30–38.

Simard, P.Y., Steinkraus, D., Platt, J.C., 2003. Best practices for convolutional neural networks applied to visual document analysis, in: null, IEEE. p. 958.

Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 .

Sirinukunwattana, K., Pluim, J.P., Chen, H., Qi, X., Heng, P.A., Guo, Y.B., Wang, L.Y., Matuszewski, B.J., Bruni, E., Sanchez, U., et al., 2017. Gland segmentation in colon histology images: The glas challenge contest. Medical image analysis 35, 489–502.

Sohn, J.H., Trivedi, H., Mesterhazy, J., Al-adel, F., Vu, T., Rybkin, A., Ohliger, M., 2017. Development and validation of machine learning based natural language classifiers to automatically assign mri abdomen/pelvis protocols from free-text clinical indications, in: SIIM.

Son, J., Park, S.J., Jung, K.H., 2017. Retinal vessel segmentation in fundoscopic images with generative adversarial networks. arXiv preprint arXiv:1706.09318 .

Springenberg, J.T., 2015. Unsupervised and semi-supervised learning with categorical generative adversarial networks. arXiv preprint arXiv:1511.06390 .

Staal, J., Abramoff, M., Niemeijer, M., Viergever, M., van Ginneken, B., 2004. Ridge based vessel segmentation in color images of the retina. IEEE Transactions on Medical Imaging 23, 501–509.

Sun, L., Wang, J., Ding, X., Huang, Y., Paisley, J., 2018. An adversarial learning approach to medical image synthesis for lesion removal. arXiv preprint arXiv:1810.10850 .

Tang, Y., Cai, J., Lu, L., Harrison, A.P., Yan, K., Xiao, J., Yang, L., Summers, R.M., 2018. Ct image enhancement using stacked generative adversarial networks and transfer learning for lesion segmentation improvement, in: International Workshop on Machine Learning in Medical Imaging, Springer. pp. 46–54.

Tanner, C., Ozdemir, F., Profanter, R., Vishnevsky, V., Konukoglu, E., Goksel, O., 2018. Generative adversarial networks for mr-ct deformable image registration. arXiv preprint arXiv:1807.07349 .

Tom, F., Sheet, D., 2018. Simulating patho-realistic ultrasound images using deep generative networks with adversarial learning, in: Biomedical Imaging (ISBI 2018), 2018 IEEE 15th International Symposium on, IEEE. pp. 1174–1177.

Tuysuzoglu, A., Tan, J., Eissa, K., Kiraly, A.P., Diallo, M., Kamen, A., 2018. Deep adversarial context-aware landmark detection for ultrasound imaging. arXiv preprint arXiv:1805.10737 .

Van Essen, D.C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T., Bucholz, R., Chang, A., Chen, L., Corbetta, M., Curtiss, S.W., et al., 2012. The human connectome project: a data acquisition perspective. Neuroimage 62, 2222–2231.

Wang, D., Gu, C., Wu, K., Guan, X., 2017a. Adversarial neural networks for basal membrane segmentation of microinvasive cervix carcinoma in histopathology images, in: Machine Learning and Cybernetics (ICMLC), 2017 International Conference on, IEEE. pp. 385–389.

Wang, J., Zhao, Y., Noble, J.H., Dawant, B.M., 2018a. Conditional generative adversarial networks for metal artifact reduction in ct images of the ear, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 3–11.

Wang, L., Nie, D., Li, G., Puybareau, É., Dolz, J., Zhang, Q., Wang, F., Xia, J., Wu, Z., Chen, J., et al., 2019. Benchmark on automatic 6-month-old infant brain segmentation algorithms: the iseg-2017 challenge. IEEE transactions on medical imaging .

Wang, T.C., Liu, M.Y., Zhu, J.Y., Tao, A., Kautz, J., Catanzaro, B., 2017b. High-resolution image synthesis and semantic manipulation with conditional gans. arXiv preprint arXiv:1711.11585 .

Wang, Y., Yu, B., Wang, L., Zu, C., Lalush, D.S., Lin, W., Wu, X., Zhou, J., Shen, D., Zhou, L., 2018b. 3d conditional generative adversarial networks for high-quality pet image estimation at low dose. NeuroImage 174, 550–562.

Wang, Z., Bovik, A.C., 2002. A universal image quality index. IEEE signal processing letters 9, 81–84.

Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., 2004. Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing 13, 600–612.

Wei, W., Poirion, E., Bodini, B., Durrleman, S., Ayache, N., Stankoff, B., Colliot, O., 2018. Learning myelin content in multiple sclerosis from multimodal mri through adversarial training. arXiv preprint arXiv:1804.08039 .

Welander, P., Karlsson, S., Eklund, A., 2018. Generative adversarial networks for image-to-image translation on multi-contrast mr images-a comparison of cyclegan and unit. arXiv preprint arXiv:1806.07777 .

Wolterink, J.M., Dinkla, A.M., Savenije, M.H., Seevinck, P.R., van den Berg, C.A., Išgum, I., 2017a. Deep mr to ct synthesis using unpaired data, in: International Workshop on Simulation and Synthesis in Medical Imaging, Springer. pp. 14–23.

Wolterink, J.M., Leiner, T., Viergever, M.A., Isgum, I., 2017b. Generative adversarial networks for noise reduction in low-dose ct. IEEE Transactions on Medical Imaging .

Xu, C., Xu, L., Brahm, G., Zhang, H., Li, S., 2018. Mutgan: Simultaneous segmentation and quantification of myocardial infarction without contrast agents via joint adversarial learning, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 525–534.

Xu, T., Zhang, P., Huang, Q., Zhang, H., Gan, Z., Huang, X., He, X., 2017. Attngan: Fine-grained text to image generation with attentional generative adversarial networks. arXiv preprint .

Xue, Y., Xu, T., Zhang, H., Long, L.R., Huang, X., 2018. Segan: Adversarial network with multi-scale l 1 loss for medical image segmentation. Neuroinformatics 16, 383–392.

Yadav, A., Shah, S., Xu, Z., Jacobs, D., Goldstein, T., 2018. Stabilizing adversarial nets with prediction methods .

Yan, K., Wang, X., Lu, L., Zhang, L., Harrison, A.P., Bagheri, M., Summers, R.M., 2018a. Deep lesion graphs in the wild: relationship learning and organization of significant radiology image findings in a diverse large-scale lesion database, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 9261–9270.

Yan, P., Xu, S., Rastinehad, A.R., Wood, B.J., 2018b. Adversarial image registration with application for mr and trus image fusion. arXiv preprint arXiv:1804.11024 .

Yang, D., Xu, D., Zhou, S.K., Georgescu, B., Chen, M., Grbic, S., Metaxas, D., Comaniciu, D., 2017a. Automatic liver segmentation using an adversarial

image-to-image network, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 507–515.

Yang, G., Yu, S., Dong, H., Slabaugh, G., Dragotti, P.L., Ye, X., Liu, F., Arridge, S., Keegan, J., Guo, Y., et al., 2018a. Dagan: Deep de-aliasing generative adversarial networks for fast compressed sensing mri reconstruction. IEEE transactions on medical imaging 37, 1310–1321.

Yang, H., Sun, J., Carass, A., Zhao, C., Lee, J., Xu, Z., Prince, J., 2018b. Unpaired brain mr-to-ct synthesis using a structure-constrained cyclegan, in: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, pp. 174–182.

Yang, Q., Li, N., Zhao, Z., Fan, X., Chang, E.I., Xu, Y., et al., 2018c. Mri image-to-image translation for cross-modality image registration and segmentation. arXiv preprint arXiv:1801.06940 .

Yang, Q., Yan, P., Zhang, Y., Yu, H., Shi, Y., Mou, X., Kalra, M.K., Wang, G., 2017b. Low dose ct image denoising using a generative adversarial network with wasserstein distance and perceptual loss. arXiv preprint arXiv:1708.00961 .

Yi, X., Babyn, P., 2018. Sharpness-aware low-dose ct denoising using conditional generative adversarial network. Journal of digital imaging , 1–15.

Yi, X., Walia, E., Babyn, P., 2018. Unsupervised and semi-supervised learning with categorical generative adversarial networks assisted by wasserstein distance for dermoscopy image classification. arXiv preprint arXiv:1804.03700 .

Ying, X., Guo, H., Ma, K., Wu, J., Weng, Z., Zheng, Y., 2019. X2ct-gan: Reconstructing ct from biplanar x-rays with generative adversarial networks. arXiv preprint arXiv:1905.06902 .

You, C., Yang, Q., Shan, H., Gjesteby, L., Guang, L., Ju, S., Zhang, Z., Zhao, Z., Zhang, Y., Cong, W., et al., 2018a. Structure-sensitive multi-scale deep neural network for low-dose ct denoising. arXiv preprint arXiv:1805.00587 .

You, C., Zhang, Y., Zhang, X., Li, G., Ju, S., Zhao, Z., Zhang, Z., Cong, W., Saha, P.K., Wang, G., 2018b. Ct super-resolution gan constrained by the identical, residual, and cycle learning ensemble (gan-circle). arXiv preprint arXiv:1808.04256 .

Yu, B., Zhou, L., Wang, L., Fripp, J., Bourgeat, P., 2018. 3d cgan based cross-modality mr image synthesis for brain tumor segmentation, in: Biomedical Imaging (ISBI 2018), 2018 IEEE 15th International Symposium on, IEEE. pp. 626–630.

Yu, S., Dong, H., Yang, G., Slabaugh, G., Dragotti, P.L., Ye, X., Liu, F., Arridge, S., Keegan, J., Firmin, D., et al., 2017. Deep de-aliasing for fast compressive sensing mri. arXiv preprint arXiv:1705.07137 .

Zanjani, F.G., Zinger, S., Bejnordi, B.E., van der Laak, J.A., et al., 2018. Histopathology stain-color normalization using generative neural networks .

Zhang, H., Xu, T., Li, H., Zhang, S., Huang, X., Wang, X., Metaxas, D., 2017a. Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks, in: IEEE Int. Conf. Comput. Vision (ICCV), pp. 5907–5915.

Zhang, L., Gooya, A., Frangi, A.F., 2017b. Semi-supervised assessment of incomplete lv coverage in cardiac mri using generative adversarial nets, in: International Workshop on Simulation and Synthesis in Medical Imaging, Springer. pp. 61–68.

Zhang, L., Zhang, L., Mou, X., Zhang, D., et al., 2011. Fsim: a feature similarity index for image quality assessment. IEEE transactions on Image Processing 20, 2378–2386.

Zhang, P., Wang, F., Xu, W., Li, Y., 2018a. Multi-channel generative adversarial network for parallel magnetic resonance image reconstruction in k-space, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 180–188.

Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O., 2018b. The unreasonable effectiveness of deep features as a perceptual metric. arXiv preprint .

Zhang, Y., Miao, S., Mansi, T., Liao, R., 2018c. Task driven generative modeling for unsupervised domain adaptation: Application to x-ray image segmentation. arXiv preprint arXiv:1806.07201 .

Zhang, Y., Yang, L., Chen, J., Fredericksen, M., Hughes, D.P., Chen, D.Z., 2017c. Deep adversarial networks for biomedical image segmentation utilizing unannotated images, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 408–416.

Zhang, Z., Yang, L., Zheng, Y., 2018d. Translating and segmenting multimodal medical volumes with cycle-and shape-consistency generative adversarial network. arXiv preprint arXiv:1802.09655 .

Zhao, H., Li, H., Cheng, L., 2017. Synthesizing filamentary structured images with gans. arXiv preprint arXiv:1706.02185 .

Zhao, J., Mathieu, M., LeCun, Y., 2016. Energy-based generative adversarial network. arXiv preprint arXiv:1609.03126 .

Zhao, M., Wang, L., Chen, J., Nie, D., Cong, Y., Ahmad, S., Ho, A., Yuan, P., Fung, S.H., Deng, H.H., et al., 2018. Craniomaxillofacial bony structures segmentation from mri with deep-supervision adversarial learning, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 720–727.

Zhu, J.Y., Park, T., Isola, P., Efros, A.A., 2017a. Unpaired image-to-image translation using cycle-consistent adversarial networks. arXiv preprint arXiv:1703.10593 .

Zhu, W., Xiang, X., Tran, T.D., Hager, G.D., Xie, X., 2017b. Adversarial deep structured nets for mass segmentation from mammograms. arXiv preprint arXiv:1710.09288 .

Zhuang, X., Shen, J., 2016. Multi-scale patch and multi-modality atlases for whole heart segmentation of mri. Medical image analysis 31, 77–87.