

# GANs for Medical Image Analysis

Salome Kazemina<sup>a,1</sup>, Christoph Baur<sup>b,1</sup>, Arjan Kuijper<sup>c</sup>, Bram van Ginneken<sup>d</sup>, Nassir Navab<sup>b</sup>, Shadi Albarqouni<sup>b</sup>, Anirban Mukhopadhyay<sup>a</sup>

<sup>a</sup>*Department of Computer Science, TU Darmstadt, Germany*

<sup>b</sup>*Computer Aided Medical Procedures (CAMP), TU Munich, Germany*

<sup>c</sup>*Fraunhofer IGD, Darmstadt, Germany*

<sup>d</sup>*Radboud University Medical Center, Nijmegen, The Netherlands*

---

## Abstract

Generative Adversarial Networks (GANs) and their extensions have carved open many exciting ways to tackle well known and challenging medical image analysis problems such as medical image de-noising, reconstruction, segmentation, data simulation, detection or classification. Furthermore, their ability to synthesize images at unprecedented levels of realism also gives hope that the chronic scarcity of labeled data in the medical field can be resolved with the help of these generative models. In this review paper, a broad overview of recent literature on GANs for medical applications is given, the shortcomings and opportunities of the proposed methods are thoroughly discussed and potential future work is elaborated. We review the most relevant papers published until the submission date. For quick access, important details such as the underlying method, datasets and performance are tabulated. An interactive visualization categorizes all papers to keep the review alive<sup>2</sup>.

*Keywords:* Generative Adversarial Networks, Medical, Image Synthesis, Segmentation, Reconstruction, Detection, De-noising, Registration, classification

---

## 1. Introduction

From the early days of Medical Image Analysis, Machine Learning (ML) and Artificial Intelligence (AI) driven systems have been a key component for complex decision making - a brief history of which can be found in [1]. Across

---

*Email addresses:* [salome.kazemina@gris.tu-darmstadt.de](mailto:salome.kazemina@gris.tu-darmstadt.de) (Salome Kazemina), [c.baur@tum.de](mailto:c.baur@tum.de) (Christoph Baur), [arjan.kuijper@mavc.tu-darmstadt.de](mailto:arjan.kuijper@mavc.tu-darmstadt.de) (Arjan Kuijper), [bram.vanginneken@radboudumc.nl](mailto:bram.vanginneken@radboudumc.nl) (Bram van Ginneken), [navab@cs.tum.edu](mailto:navab@cs.tum.edu) (Nassir Navab), [shadi.albarqouni@tum.de](mailto:shadi.albarqouni@tum.de) (Shadi Albarqouni), [anirban.mukhopadhyay@gris.tu-darmstadt.de](mailto:anirban.mukhopadhyay@gris.tu-darmstadt.de) (Anirban Mukhopadhyay)

<sup>1</sup>The authors contributed equally to this work.

<sup>2</sup>Live tree: [http://livingreview.in.tum.de/GANs\\_for\\_Medical\\_Applications/](http://livingreview.in.tum.de/GANs_for_Medical_Applications/)

generations of development, the focus was mostly put on decision making at different granularity levels, with techniques ranging from low-level pixel processing over feature engineering combined with supervised classifier learning to the recent wave of feature learning using Convolutional Neural Networks (CNNs).

The driving focus of the machine learning-based Medical Image Analysis community has been on the supervised learning of decision boundaries, while generative tasks have been on the back seat. The unique ability of Generative Adversarial Networks (GAN) introduced in [2] by Goodfellow et al. to mimic data distributions has carved open the possibility to bridge the gap between *learning* and *synthesis*. The rapid enhancement of GANs [3] are facilitating the synthesis of realistic-looking images at unprecedented level. The reasons behind this superiority are related to two basic properties. First, GANs as an unsupervised training method aim to obtain pieces of information over data [4], in an indirect fashion of supervision. Second, GANs have shown significant performance gains in the extraction of visual features by discovering the high dimensional latent distribution of the data.

This review summarizes GAN-based architectures proposed for medical image processing applications, including synthesis, segmentation, reconstruction, detection, de-noising, registration, and classification. The distribution of papers according to this classification can be seen in Figure 1. We also provide tables to have quick access to key information like the performance of methods, metrics, datasets, modality of images and the general format of the proposed architecture. Moreover, we discuss the advantages and shortcomings of the methods and specify clear directions for future works.

In this review, we have covered medical imaging application of GAN published until end of 2018. Papers published in this time range propose using GANs in medical applications of synthesis, segmentation, reconstruction, detection, de-noising, registration, and classification. These papers were applied to different image modalities such as MRI, CT, OCT, chest X-Ray, Dermoscopy, Ultrasound, PET, and Microscopy. To find the papers we searched for keywords medical and GAN (or generative adversarial network) along with the aforementioned applications in Google Scholar, Semantic Scholar, PubMed, and CiteSeer. Also, we checked references and citations of selected papers. Since GANs are rather new, and a significant number of articles are still in the publication process of different journals and conferences, we covered pre-prints published in arXiv as well.

We thus ended up with 77 papers which we consider the most relevant ones covering a broad spectrum of applications and variety of GANs. The reminder of this paper is structured as follows. In section 3 we introduce the architecture of the GAN and its subclasses which are used in medical image applications. In section 4 different contributions of GANs in medical image processing applications (de-noising, reconstruction, segmentation, registration, detection, classification, and synthesis) are described and Section 5 provides a conclusion about the investigated methods, challenges and open directions in employing GANs for medical image processing.

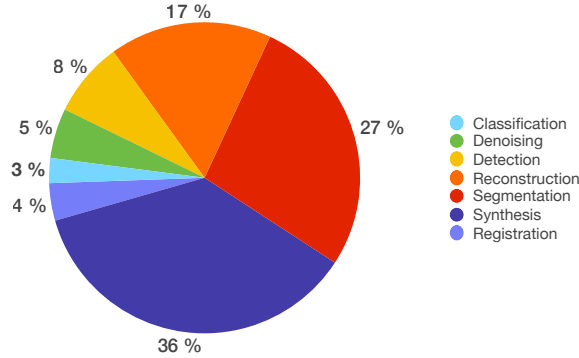


Figure 1: The distribution of papers among the different categories.

## 2. Opportunities for Medical Image Analysis

Supervised Deep Learning is currently the state-of-the-art in many Computer Vision and Medical Image Analysis tasks. However, a major limiting factor for this paradigm, not only in the context of medical applications, is its dependence on vast amounts of annotated training data. In the medical field this is particularly crucial, as the acquisition and labeling of medical images requires experts, is tedious, time-consuming and costly, which leads to a severe lack of labeled training data. In addition, in the medical field, many datasets suffer from severe class imbalance due to the rare nature of some pathologies. In this context, generative modeling can potentially act as a reliever for resolving these well known machine learning problems. Applied research of GANs for medical imaging can be broadly classified into the following categories based on the input and output: i) latent space to image, and ii) image to image.

Within the latent space to image category GANs have shown the capabilities to generate images with unprecedented realism. Under the assumption that GANs can generate meaningful samples that enhance existing datasets and carry useful information, a variety of research has already been conducted for medical image synthesis and reconstruction, which is reviewed in Subsection 4.1.1 and 4.3.

Another issue hampering the machine learning community is the necessity to handcraft similarity measures for general tasks. Traditional similarity objectives comprise pixelwise losses such as the  $\ell_1$  or  $\ell_2$ -distance, both of which induce blurry results and lack the incorporation of context. The adversarial training concept behind GANs theoretically eliminates the need to model explicit pixelwise objective functions by learning a rich similarity metric to tell real and fake data apart. This allows to optimize for concepts in images beyond the pixel-level, leading to more realistic results. This appealing property has been recently exploited for improved medical image segmentation and registration (reviewed in Subsections 4.2 and 4.6), Image-Enhancement such as Denoising

(reviewed in Subsection 4.5) and tackling the general problem of domain shift in medical images using GAN-based Image-to-Image translation techniques (reviewed in Subsections 4.1.2, 4.4). In deficiency of annotated and simultaneous abundance of unlabeled data, the paradigm of semi-supervised learning offers different frameworks for training machine learning models by ensuring similar or dissimilar model behavior for similar or dissimilar data points, where similarity needs to be defined appropriately. Again, the notion of similarity is a crucial parameter and often highly data-dependant. Under such conditions, GANs and adversarial training have also proven useful for training classifiers or dealing with domain shift in medical data, as the explicit formulation of similarity is not required (reviewed in Subsection 4.7).

### 3. Overview

In this section, we introduce the general concept behind GANs, their conditional variants as well as a variety of prominent extensions and follow-up works that have been successfully leveraged in Medical Image Analysis applications. These extensions comprise Wasserstein-GAN, conditional GAN (for example of Pix2Pix), CycleGAN, Least Squares GAN, Markovian GAN as well as Auxiliary Classifier GAN.

In the context of this work, there are three *Adversarial* concepts, which should be understood properly by their different meanings. *Adversarial attack* means to make imperceptible changes to an image such that a classifier misclassifies it, while it could classify unmodified image successfully. Usually the modified image, called adversarial image or *adversarial examples*, is not recognizable from the original image visually. *Adversarial training* proposed by "Szegedy et al." [5] is an idea that increases the robustness of neural networks against adversarial attacks by learning their characteristics. Due to the state of existing neural networks, at the time, implementing adversarial training was not a practical solution. The effectiveness of this idea becomes apparent when Goodfellow et. al employed it in GANs [2]. Sometimes GAN is mis-attributed as adversarial training, but it is necessary to differentiate between them. In reality, GAN consists of two types of networks and use the adversarial training concept, elaborated in the following section.

#### 3.1. GAN

The GAN framework [2] consists of a generator (G), and a discriminator (D) network as well as a training dataset of real data  $\mathbf{X}$  with an underlying distribution  $p_{real}$ . G, as a forger, is a multilayer network with parameters  $\theta_G$ , which aims to find a mapping  $\hat{x} = G(z; \theta_G)$  that relates latent random variables  $z \sim p_z(z)$  to fake data following the distribution  $p_\theta(\hat{x}|z)$ . By discovering the mapping, G generates fake data, which is supposed to not be distinguishable from real data, i.e.  $p_\theta(\hat{x}|z) \sim p_{real}$ . On the other hand, the discriminator  $D(x; \theta_D)$  aims to distinguish the fake samples from real ones. Thereby,  $D(x)$  is the scalar output of the discriminator network that shows the probability

that  $x$  is real rather than generated from  $p_\theta(\hat{x}|z)$  (Figure 2).  $D$  is trained to maximize the probability of correct label assignment to fake and real data, while  $G$  is trained to fool the discriminator by minimizing  $\log(1 - D(G(z)))$ . Mathematically speaking,  $D$  and  $G$  play a two-player minimax game with value function  $V(G,D)$ :

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} [\log(D(x))] + \mathbb{E}_{z \sim p_z(z)} [1 - \log(D(G(z)))] \quad (1)$$

This way, the generator is updated only through gradients back-propagated from the discriminator. "Goodfellow et al." [2] mentioned that if the generator is optimized to maximize  $\log(D(G(z)))$  instead of minimizing  $\log(1 - D(G(z)))$ , much stronger gradients can be obtained in earlier steps (iterations) of training. In general, this indirect optimization procedure prevents input components to be explicitly memorized by the generator. The main advantage of the GAN is to find similarities that map a candidate model to the distribution of real data by focusing on the underlying probability density of data. It leads to very sharp distributions around data, which can be used to degeneration of that [3].

Though GANs show such inherent advantages over discriminatively trained CNNs, there are some challenges as well: 1) *mode collapse*: when  $G$  collapses to map all latent space inputs to the same data and 2) *instability*: which leads to the generation of different outputs for same input. The main reason behind such a phenomena is vanishing gradients through the optimization procedure.

Although batch-normalization comes as a solution for the instability of the GAN, it is not enough to improve the performance of the GAN to the optimal stability. So, many subclasses of GAN have been introduced to resolve these drawbacks. Some of the most common ones are introduced here. Furthermore, many GAN-based deep networks are proposed specifically for medical image processing projects, in which different architecture and loss functions are used to enhance the reliability and accuracy of the deep networks in the necessary level of health-care CAD systems.

### 3.2. DCGAN:

To address the instability of the basic GAN architecture [2], "Radford et al." propose the Deep Convolutional GAN (DCGAN) [6], in which both the generator and discriminator follow a deep convolutional network architecture. These networks are able to extract hierarchical features of the image by learning down/up-sampling due to the location of features existence. In this way, the extracted features of objects can be used to generate new ones. Key components of the DCGAN which affect the stability of the network, are batch normalization and leaky-ReLU. Although DCGAN is more stable than the GAN, it is still prone to mode collapse.

### 3.3. cGAN:

Conditional GAN (cGAN) proposed by "Mirza et al." [7], have also shown that prior information can be incorporated into the GAN framework. In the cGAN, the generator is presented with random noise  $z$  as well as some prior

information  $c$  jointly. Additionally, the prior knowledge  $c$  is fed into the discriminator together with the corresponding real or fake data. Mathematically speaking, the cGAN framework is given as follows:

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} [\log(D(x|c))] + \mathbb{E}_{z \sim p_z(z)} [1 - \log(D(G(z|c)))] \quad (2)$$

By conditioning the networks, it has been shown that both training stability and output generation can be improved (Figure 3). In [4], Isola et al. propose a very successful variant of the conditional GAN named “pix2pix” for the challenging task of image-to-image translation. In this architecture, the generator and discriminator are following the U-Net [8] and MGAN (Patch-GAN) [9] networks which are demonstrated to provide a good framework for wide conditional transformation problems. In the proposed model, the  $\ell_1$  loss in combination with adversarial loss is considered to put more pressure on the generator to produce images more similar to the ground truth images.

### 3.4. MGAN

Another conditional GAN framework is Markovian GAN (MGAN) [9], proposed by “Li et al.” for fast and high quality style transfer. The MGAN, as depicted in Figure 4, heavily utilizes a pre-trained VGG19 network with fixed weights to extract high-level features for both transferring style to a target texture and simultaneously preserving the image content. In the MGAN, both discriminator and generator network are prepended with a VGG19 network to extract featuremaps. The generator transfers these featuremaps to an image with target texture, and the discriminator transforms the either input (real or texturized) image into VGG19 feature maps again, on which it finally discriminates with the help of a Fully Convolutional Network (FCN). Utilizing an FCN for classifying the input as real or fake ultimately amounts to classifying patches in VGG19 feature map space. By training the generator to fool this discriminator, it is forced to generate images which lead to realistic VGG19 feature activations as would have been obtained on real data and thus also to images with realistic style. An additional perceptual loss component (calculated using VGG) ensures that the image content does not change too much while the style is transferred.

### 3.5. cycleGAN:

“Zhu et al.” [10] propose a GAN architecture, which aims to discover the underlying relationship between two image domains through learning their definitive features from unpaired data. To achieve this goal a cycle training algorithm is used to capture main features of a domain of image for translating them to another domain. Since the map function learned by adversarial loss is not reliable to map input image to desired output, a cycle loss function is considered to reduce the space of possible mapping functions. In this way, two generators ( $G : X \rightarrow Y$  and  $F : Y \rightarrow X$ ) are considered to find the mapping from  $X$  domain to  $Y$  domain and vice versa and also two discriminators ( $D_Y$  and  $D_X$ )

to train them (Figure 5). This learning strategy stabilizes network performance and generates high quality translated images. The final loss function is defined as follows:

$$L(G, F, D_X, D_Y) = L_{GAN}(G, D_Y, X, Y) + L_{GAN}(F, D_X, Y, X) + \lambda L_{cyc}(G, F) \quad (3)$$

with

$$L_{cyc}(G, F) = \mathbb{E}_{x \sim P_{data}(x)} [\|F(G(x)) - x\|_1] + \mathbb{E}_{y \sim P_{data}(y)} [\|G(F(y)) - y\|_1] \quad (4)$$

### 3.6. AC-GAN

"Odena et al." [11] report that instead of providing both the generator and the discriminator networks with side information as shown in the cGAN, the discriminator can be tasked with reconstructing such side information. In their auxiliary classifier GAN framework (AC-GAN, Figure 6), the discriminator architecture is modified such that after a few of layers it splits into a standard sample discriminator network as well as an auxiliary classifier network, which aims at classifying samples into different categories. The authors show that this framework allows to use (partially) pre-trained discriminators and appears to stabilize training.

### 3.7. WGAN:

In the original GAN framework, the data distributions of generated and real images are compared using the Jensen-Shannon (JS) divergence. This kind of exact comparison can make the saddle-point of optimization unreachable and gradients vanishing, which leads to mode collapse and instability. So considering another, more approximate distance estimation between real and generated data distribution can be effective as a solution. "Arjovsky et al." [12] propose the Wasserstein-GAN (WGAN) architecture that uses the Earth Mover (ME) or Wasserstein-1 distance estimation instead of the JS divergence. In addition, both the generator and discriminator follow the general DCGAN architecture. WGAN provides a robust adversarial generative model through a more meaningful learning procedure, which is able to find deeper relationships between distributions. Despite these theoretical advantages, WGAN leads to a slow optimization process in practical scenarios.

### 3.8. LSGAN:

"Mao et al." propose another solution for the instability of GAN, called Least Squares GAN (LSGAN) [13]. In this architecture, some parameters are added in the loss function to avoid gradient vanishing. In this way, the fake data, which is discriminated as real but is far away from the dense distribution of real data, will be penalized due to its distance from main mode of real data. Also, the gradient will become 0 only in the case that distribution of fake data

perfectly matches the distribution of real data. The loss function for LSGAN is defined as follows:

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} [(D(x) - b)^2] + \mathbb{E}_{z \sim p_z(z)} [(D(G(z)) - a)^2] \quad (5)$$

## 4. Applications in Medical Image Processing

In this section, we summarize GAN-based methods which are proposed to solve medical imaging problems, in 7 application categories: synthesis, segmentation, reconstruction, detection, de-noising, registration, and classification. In every subsection, a table summarizes the most important details of proposed methods and the medical image modalities they are designed for.

### 4.1. Synthesis

Originally, GANs have been proposed as an entirely unsupervised generative framework, with the goal to map from random noise to synthetic, realistically looking images following the training data distribution. With the conditional GAN, the framework has also been successfully turned into a supervised generative framework by conditioning both the generator and the discriminator on prior knowledge, rather than noise alone. For clarity, we refer to the original GAN framework as the *unconditional* or *unsupervised* GAN, in contrast to the *conditional* GAN. We want to emphasize that it is very important to make a distinction between these different concepts and consequently categorize the literature accordingly.

The generative property of both frameworks has been exploited in various ways for synthesizing certain types of medical images either from noise alone (see *Unconditional Image Synthesis* 4.1.1), or from prior knowledge (see *Conditional Image Synthesis* 4.1.2) such as metadata or even image data. In the following, a broad overview on works from unconditional and conditional image synthesis will be given. In the particular case for conditional approaches, we further classify the contributions based on the image modality. For the literature on unconditional image synthesis we do not make this distinction due to the small amount of papers.

#### 4.1.1. Unconditional Image Synthesis:

A great variety of works has recently appeared in the field of unsupervised medical image generation using GANs. The synthesis of realistically looking medical images opens up many new opportunities to tackle well-known deep learning problems such as class imbalance, data augmentation [14] or the lack of labeled data. Further, it facilitates data simulation [15] and aids to gain deeper insights into the nature of data distributions and their latent structure.

Initial results have shown that GANs can be used to synthesize realistically looking patches of prostate lesions [16] or retinal images [17]. Both approaches rely on the DCGAN architecture to synthesize patches at a resolution of  $16 \times 16$



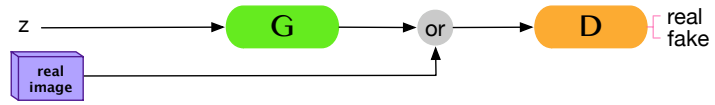


Figure 2: GAN

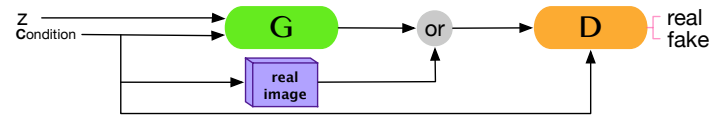


Figure 3: cGAN

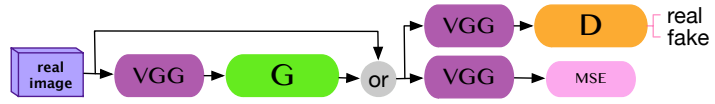


Figure 4: MGAN

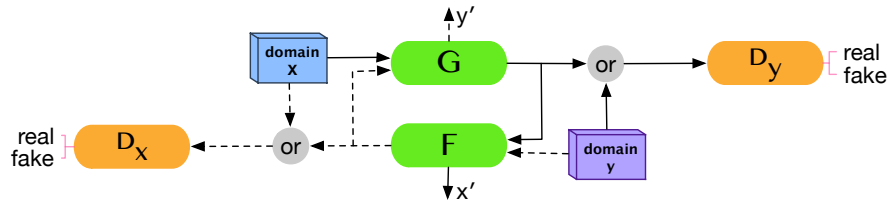


Figure 5: cycleGAN

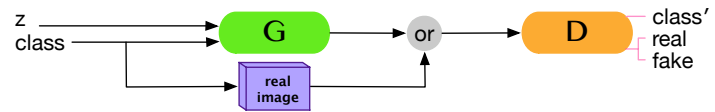


Figure 6: AC-GAN

pixels and  $64 \times 64$  pixels, respectively. In [15], the authors successfully utilize DCGANs for generating  $56 \times 56$  pixels patches of lung cancer nodules which could hardly be distinguished from real patches in a visual turing test involving two radiologists.

Frid-Adar et al. [14] make use of the DCGAN for the synthesis of focal CT liver lesion patches from different classes at a resolution of  $64 \times 64$  pixels. For each class, i.e. cysts, metastases and hemangiomas, they train a separate generative model. As the training dataset is originally quite small, they use heavily augmented data for training the GANs. In a set of experiments for liver lesion classification, the authors demonstrate that synthetic samples in addition to data augmentation can considerably improve a Convolutional Neural Network classifier.

The work in [18] has shown that the DCGAN with vanilla training is in fact also able to learn to mimic the distribution of MR data at considerably high resolution, even from a surprisingly small amount of samples. The real data distribution consisted of only 528 midline T1-weighted axial MR slices at a resolution of  $220 \times 172p$ . After training for 1500 epochs, the authors obtained visually compelling results which human observers could not reliably distinguish from real MR midline slices.

In [19], the authors utilize and compare both DCGAN, LAPGAN and modifications of the latter for the task of skin lesion synthesis at a resolution of  $256 \times 256$  pixels. Similar to [18], the training dataset was quite small, consisting of only 1,600 images. Probably due to the high variance within the training data, the small number of samples turned out not to be sufficient to train a reliable DCGAN, however the hierarchical LAPGAN and its variants showed promising synthesis results. The synthetic samples have also partially successfully been used for data augmentation when training a skin lesion classifier. In [20], the same authors employed the recently proposed concept of progressive GAN growing for synthesizing images of skin lesions and showed stunning, highly realistic synthetic images which even expert dermatologists could not reliably tell apart from real samples.

#### 4.1.2. Conditional Image Synthesis:

**CT from MR** In many clinical settings, the acquisition of CT images is required. This, however, puts the patient at risk of cell damage and cancer because of the radiation exposure, which motivates the synthesis of CT images from MR acquisitions. Nie et al. [21] synthesize CT images from corresponding MR images with the help of a cascade of 3D Fully Convolutional Networks which they train with a normal reconstruction loss, an image gradient loss and additionally with an adversarial network in order to improve realism of the synthetic CT images. The idea of utilizing a cascade of generator networks originates from the so-called Auto-Context Model, in which a network provides its output as additional input to a succeeding network in order to provide context information and allow for refinements (Figure 7). While Nie et al. require corresponding pairs of CT and MR images for training, Wolterink et al. [23] successfully utilize Cycle-GANs to transform 2D MR images to CT images without the need for

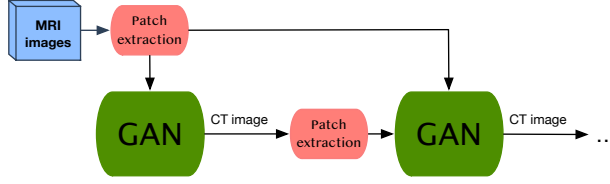


Figure 7: Proposed architecture in [22]

paired, co-registered training data. Interestingly, in contrast to training from paired, co-registered data, their training led to even better results as the model avoids to learn mappings in the presence of registration artifacts. Zhao et al. [24] utilize conditional GANs to map 3D MR data of the head to its CT counterpart for the purpose of facilitating easier segmentation of Craniomaxillofacial Bony Structures from anatomical data, as required for surgical planning, without exposing the patient to radiation. To obtain feasible image-to-image translation results, they propose so-called “deep supervision discrimination” which, similarly to the perceptual loss, utilizes the feature representations of a pretrained VGG16 model to i) tell real and synthetic CT images apart and ii) provide gradient updates to the actual generator network. Noteworthy, the discriminator leverages feature maps from different levels of the pretrained network, thus named “deep supervision”

**MR from CT** Similar to [23], Chartsias et al. [25] successfully leverage CycleGANs for unpaired image-to-image translation, however for synthesizing pairs of cardiac MR images and a segmentation mask from pairs of cardiac CT slices with the ground-truth segmentation mask. The authors have shown that the performance of a segmentation model can be improved by 16% when additionally trained with the synthetic data, and that synthetic data alone is sufficient for training a model which performs only 5% worse than a model trained on real data. A major limitation of GANs, as recently pointed out in [26], is the lack of a guarantee that tumors or lesions present in an image are preserved during image-to-image translation. To cope with this issue in CycleGANs, Jiang et al. [27] proposed a tumor-aware loss function in addition to the cycle-Consistency objective for synthesizing MR images with tumors from CT images with tumors. Once trained, they augment their originally very small MR training set for training a Unet-based tumor segmentation model and show significant improvements.

**Retinal Image Synthesis** In [28] the authors utilize a slight modification of the adversarial training concept proposed in [4] for the challenging task of eye fundus image generation. They learn a mapping from binary images of vessel trees to new retinal images at a resolution of 512x512pixel, which look extremely realistic and rate very well in common scores for retinal image quality judgement. In follow-up work [29], the authors further introduce an adversarial autoencoder which is trained to compress vessel tree images into a multivariate normal distribution and to consecutively reconstruct them. The resulting

generative autoencoder allows to synthesize arbitrary high resolution vessel tree images by sampling from the multivariate normal distribution. The synthetic images in turn are fed into the image-to-image translation model, ultimately leading to an end-to-end framework for realistic, high resolution retinal image synthesis. Very similarly, Guibas et al. [30] propose a two-stage approach, consisting of a GAN which is trained to synthesize vessel tree images from noise, and a second conditional GAN as seen in Pix2Pix [4] to generate realistic, high resolution pairs of groundtruth vessel segmentation and the corresponding eye fundus image. In succession, they investigate the performance of a U-Net trained for segmentation using real data pairs and another model trained only on the synthetic samples, and find that training from only the synthetic data leads to an only slightly inferior model.

In [31], the authors also leverage the Pix2Pix framework for the tasks of synthesizing filamentary structured images, i.e. eye fundus images and neurons from binary segmentation masks. Compared against [28, 29], the authors also provide their framework with a reference image for style and train the generator also with the feedback from an additional VGG-network leveraged for style transfer. and show that only 10 training examples are sufficient for training such an image-to-image translation model. Opposed to Pix2Pix, they do not introduce noise with the help of dropout, but by augmenting noise to the bottleneck of the encoder-decoder network. In a set of use-case experiments on retinal image segmentation it is demonstrated that the introduction of additional synthetic images, i.e. training from both real and synthetic images, slightly improves the segmentation performance.

**PET from CT** PET images are frequently used for diagnosis and staging in oncology, and the combined acquisition of PET and CT images is a standard procedure in clinical routine. Furthermore, PET/CT imaging is becoming an important evaluation tool for new drug therapies. However, PET devices involve radioactivity and thus put patients at risk, and are expensive in general. Consequently, the medical image analysis community has been working on synthesizing PET images directly from CT data. In this context, GANs have also shown outstanding performance. Initial promising results for synthesizing liver PET images from CT data with conditional GANs have been obtained in [32]. The conditional GAN, again inspired by [4], is able to synthesize very realistic looking PET images, however at the cost of low response to underrepresented tumor regions, which leads to poor tumor detection performance in a set of use-case experiments. In contrast, the authors find that an FCN for PET image synthesis is capable of synthesizing tumors, but produces blurry images in general. By blending corresponding synthetic PET images coming from the conditional GAN and the FCN, they are able to achieve very high tumor detection performance, though. Similarly, in [33] the authors utilize a conditional GAN for synthesizing  $200 \times 200$  pixels sized PET images from pairs of CT images and binary labelmaps. While CT images alone would be sufficient as input, they note that by adding a labelmap which marks the location of a tumor, they obtain globally more realistic, synthetic output. Because of the two-channel input to the generator, they refer to their network as the multi-channel GAN. Further,

the authors validated their synthetic PET images with a tumor detection model trained on synthetic data and obtained comparable results to a model trained with real data, showing that synthetic data can in fact be beneficial when there is a lack of labeled data.

**PET from MRI** For monitoring disease progression, understanding physiopathology and evaluate treatment efficacy of Multiple Sclerosis (MS), measuring the myelination content in PET images of the human brain has recently shown to be very valuable. Unfortunately, PET imaging for MS is costly and invasive as it requires the injection of a radioactive tracer. In [34], the authors successfully utilize a cascade of two conditional GANs for synthesizing such PET images from a set of different MR modalities. Their approach operates directly on volumetric data, leveraging a 3D U-Net for the generator networks and discriminator networks with 3D convolutions. Interestingly, the authors noted that a single conditional GAN was insufficient for the task at hand as it produced blurry images. Splitting the synthesis task into smaller, more stable subproblems, seemed to drastically improve the results.

**Ultrasound** Hu et al. [35] propose a conditional GAN architecture for synthesizing 2D ultrasound images of a fetus phantom, as produced by a freehand US probe, given 3D spatial pixel locations within the anatomy. In contrast to the standard conditional GAN, the authors find it necessary to transform the pixel locations into featuremaps and to concatenate them with the produced featuremaps at each level of the generator to facilitate training. In their experiments they demonstrate the capability of simulating US images at locations unseen to the network, quantify the generation of sound images by comparing the location of clinically relevant anatomical landmarks in synthetic and real images, and verify the realism of the generated images in a usability study. The quantitative results show that anatomical landmarks are roughly synthesized at the right locations with a mean error of 6.1mm. In their usability study, the sonographer was able to mostly correctly distinguish between real and generated samples, which is due to checkerboard artifacts in the synthetic images. After blurring the images using a gaussian kernel with  $\sigma = 1.5$ , the sonographer was not able to reliably tell the difference anymore. The interested reader is also referred to the NiftyNet framework [36], in which this conditional GAN is contained. Tom et al. [37] apply GANs for intravascular ultrasound (IVUS) simulation in a multi-stage setup. A first generator conditioned on physically simulated tissue maps produces speckle images, which in turn act as the conditioning input to a second residual network based generator. The second generator maps the speckle images to low resolution, synthetic  $64 \times 64$  pixels sized US images. A third generator transforms these low resolution images into high resolution samples at a resolution of  $256 \times 256$  pixels. In a visual turing test, the synthetic images could not reliably be distinguished from real ones.

**X-ray** Mahapatra et al. [38] train conditional GANs for chest X-ray synthesis from perturbed segmentation maps and combine it with Bayesian Neural Networks for Active Learning. More precisely, they propose an uncertainty based synthetic sample selection approach for training segmentation neural networks on samples which are informative to that model.

**Stain Normalization** Conditional GANs have also been leveraged for coping with the variance in digital histopathology staining, which is well known to cause problems for CAD systems. Cho et al. [39] point out that a tumor classifier generalizes poorly on both data with staining properties different from the training set, as well as on images that have been stain-normalized with state-of-the-art methods. To overcome these issues, they propose a feature-preserving conditional GAN for stain style transfer with the particular goal to prevent a degradation in performance of CAD systems on synthetic images. First, they map histological images to a canonical gray-scale representation. In succession, they leverage a conditional GAN to transform gray-scale images into RGB images with the desired staining. By employing an additional feature-preserving loss on the hidden layers of the discriminator, they demonstrate that a tumor classifier model trained on data stemming from a certain distribution performs better on the stain-transferred images than on the original ones, and that their conditional GAN shows the smallest degradation in performance compared to other state-of-the-art stain transfer methods.

Bayramoglu et al. [40] leverage the Pix2Pix framework for virtual H&E staining on unstained hyperspectral microscopy images using  $64 \times 64$  pixels sized patches. The authors report the SSIM and MSE between synthetically stained images and the ground-truth and point out to have obtained promising result, but require expert feedback in order to draw a valid conclusion.

BenTaieb et al. [41] try to tackle the stain transfer problem with the help of a so-called Auxiliary Classifier GAN by simultaneously training a conditional GAN for stain-transfer and a task-specific network (i.e. a segmentation or classification model). The joint optimization of the generator, the discriminator and the task-specific network drives the generator to produce images with relevant features preserved for the task-specific model and overall leads to superior results in stain-normalization compared to other state-of-the-art methods.

Aformentioned methods rely on paired training data to map from a source to target staining, which is often hardly available and requires preprocessing such as co-registration. However, co-registration itself is not perfect and is prone to artifacts. Shaban et al. [42] alleviate the need for paired training data and co-registration by employing CycleGANs for the task of stain transfer. In a broad set of experiments on different datasets, they show visually much more compelling stain transfer results than previous deep-learning and non-deep learning based methods. In addition, they also show quantitatively how their approach significantly reduces domain shift which usually hampers deep learning models: A classifier trained for mitosis detection provides much better classification results on images stain-transferred with the proposed approach than on original data, and again also other stain transfer methods.

**Microscopy** Han et al. [43] propose a conditional GAN framework similar to Pix2Pix for transferring between Phase Contrast and Differential Interference Contrast (DIC) Microscopy images, however with two discriminator networks rather than one. A U-net like generator is trained to synthesize the image of a certain modality from an image of the source modality and a cell mask. Two different discriminators then either discriminate between pairs of real source

and target modality images versus pairs of real source and synthesized target modality image, or pairs of cell mask and real source versus cell mask and synthesized target images. In a set of qualitative and quantitative evaluations they rank their two-discriminator approach against the Pix2Pix framework which uses only a single discriminator. They report improved results in the metrics of SSIM and normalized RMSD when transferring from DIC image to Phase Contrast, and comparable results when trying to map from Phase Contrast to DIC. Noteworthy, the authors amount the comparable performance of the latter mapping to the details already present in Phase Contrast images, which leaves the cell mask with very little impact on the synthesis outcome.

**Blood Vessels** Machine Learning driven analysis methods for detecting atherosclerotic plaque or stenosis in coronary CT angiography (CCTA) are powerful, but data-hungry. To deal with the lack of labeled data, Wolterink et al. [44] propose to synthesize plausible 3D blood vessel shapes with the help of a Wasserstein GAN from noise and attribute vectors. To facilitate the synthesis in 3D at appropriately high resolution, the authors generate 1D parameterizations of primitives which characterize blood vessels and act as a proxy for the final vessel rendering. Magnetic Resonance Angiography (MRA) has also evolved into an important tool for visualizing vascular structures, but often times it is not acquired alongside the standard protocols. In [45], the authors propose the so-called steerable GAN for synthesizing MRA images from T1 and T2-weighted MR scans, potentially alleviating the need for additional MR scans. Their conditional, steerable GAN combines a ResNet-like generator with a PatchGAN-discriminator, an  $\ell_1$ -loss between real and synthesized image as well as a steerable filter loss to promote faithful reconstructions of vascular structures.

Tables 1, 2 and 3 give an overview of all the presented image synthesis methods. The unconditional synthesis methods are summarized in Table 1, whereas the conditional GAN variants are summarized in Table 2 and 3. In particular, we report the method, containing the underlying GAN architecture, the image modalities on which the particular method operates, the datasets which have been used and the resolution of the synthesized images. Since losses are a substantial part of the underlying GAN framework, we do not explicitly report them here. Further, we do not report any quantitative results since they i) are in many case unavailable, ii) hardly interpretable and iii) overall hardly comparable. In general, many interesting GAN-based approaches have been made for both unsupervised and conditional image synthesis. However, often the validity of the method at hand is questionable and requires more elaboration. For instance, in many visual turing tests it is fairly easy to distinguish between real and generated images [14, 15, 35] due to artifacts in synthetic samples, such as the well known checkerboard pattern. In [35, 15], the authors tackle this problem by applying anisotropic or gaussian filtering to both real and fake samples before presenting them to the raters [35, 15], which is only valid as long as blurry images still contain the required amount of information for the task at hand. Another problem is that GANs are prone to the phenomenon of mode collapse, in which the model is only able to generate samples stemming

from one or a few modes of the real data distribution, resulting in very similar looking synthetic samples. Particularly in the works of [16] and [17], where samples look fairly similar, a thorough elaboration on whether mode collapse has occurred or not would have been very interesting. In general, the community still lacks a meaningful, universal quantitative measure for judging realism of synthetic images. Regardless of the realism, aforementioned works have shown that GANs can be used successfully for data simulation and augmentation in classification and segmentation tasks. How realism, artifacts in and specific properties of generated samples affect a machine learning model when used for data augmentation also remains an open question.

Table 1: Unconditional GANs for Medical Image Synthesis

Method	Image Modality	Dataset	Resolution
[16] Architecture: DCGAN	MRI Prostate Lesions	SPIC ProstateX Challenge 2016	16×16
[15] Architecture: DCGAN	CT Lung Cancer Nodules	LIDC-IDRI	56×56
[14] Architecture: DCGAN	focal CT liver lesion patches	non-public	64×64
[18] Architecture: DCGAN	2D axial brain MR slices	Baltimore Longitudinal Study of Aging (BLSA)	220×172
[19, 20] Architecture: DCGAN, LAPGAN, PGAN	Dermoscopic Images of Skin Lesions	ISIC2017 & ISIC2018	256×256

Table 2: Conditional GANs for Medical Image Synthesis

Method	Image Modality	Dataset	Resolution
[21] Architecture: 3D Autocontext FCN with adversarial loss, image gradient loss and $\ell_2$ -loss	MR to CT	ADNI and 22 non-public pelvic image pairs	n/a
[23] Architecture: CycleGAN	2D sagittal brain MR and CT slices	non-public	256×256
[24] Architecture: 3D cond. GAN	MR to CT	ADNI	152 × 184 × 149
[25] Architecture: CycleGAN	2D cardiac MR w. segmentation mask to cardiac CT w. segmentation mask	non-public	232×232
[27] Architecture: Tumor-Aware CycleGAN	CT to MR	NSCLC (The Cancer Imaging Archive) & non-public	256×256
[28, 29] Architecture: AAE and Pix2Pix	2D binary vessel tree images to retinal images	DRIVE, MESSIDOR	512×512
[46] Architecture: 3D cond. GAN	lung nodules 3D	LIDC	64×64×64
[30] Architecture: GAN and Pix2Pix	2D binary vessel tree images to retinal	DRIVE, MESSIDOR	512×512
[31] Architecture: Pix2Pix w. Style Transfer	eye fundus, microscopic neuronal	DRIVE, STARE, HRF, NeuB1	512×512 and higher
[32] Architecture: Pix2Pix and FCN	2D liver tumor CT to PET	non-public	n/a



Table 3: Conditional GANs for Medical Image Synthesis

Method	Image Modality	Dataset	Resolution
[33] <b>Architecture:</b> conditional multi-channel GAN	CT and segmentation pairs to PET images	non-public	200×200
[35] <b>Architecture:</b> spatially cond. GAN	2D US	non-public fetus phantom	160×120
[37] <b>Architecture:</b> multi-stage cond. GAN	simulated tissue maps to 2D Intravascular US	IVUS challenge	256×256
[38] <b>Architecture:</b> cond. GAN	Segmentation maps to synthetic X-ray images	SCR chest XRay database	512×512
[39] <b>Architecture:</b> feature-preserving conditional style- transfer GAN	Digital Histopathology	CAMELYON16	n/a
[40] <b>Architecture:</b> Pix2Pix	Hyperspectral microscopic to H&E stained	non-public	64×64
[41] <b>Architecture:</b> ACGAN	Digital Histopathology	MITOS-ATYP1A14 MICCAI16 GlaS challenge non-public ovarian carcinoma	250×250
[42] <b>Architecture:</b> CycleGAN	Digital Histopathology	MITOS-ATYP1A14 Camelyon16	256×256
[43] <b>Architecture:</b> cond. GAN with two Discriminators	DIC & Phase Contrast Microscopy	non-public	256×256
[44] <b>Architecture:</b> WassersteinGAN	Geometric parameters extracted from CCTA	non-public	n/a
[45] <b>Architecture:</b> cond. steerable GAN	MRA from T1 & T2w MRI axial slices	IXI Dataset	n/a

#### 4.2. Segmentation

Segmentation of objects and organs in medical images is an important pre-processing step for many applications, e.g. detection, classification, shape analysis etc. Tedious and time-consuming nature of manual segmentation, along with the possibility of creeping subjective bias made automatic segmentation a fundamental problem of medical image analysis. The advent of U-net has significantly advanced the field of CNN-based segmentation [1]. However, different anatomical structures still require special treatment, for example, conditional random fields [47] or statistical shape models [48] to regularize U-Net results. A GAN can possibly provide a viable alternative for segmentation - most of the GAN-based segmentation techniques so far focus on well known GAN architecture (overwhelming majority uses U-net as the generator) with combined loss functions. Unlike the previous applications relying on end-to-end GANs, segmentation applications using GAN often require pre- and post-processing steps. To emphasize the connection of segmentation methods with particular anatomies, the rest of the section focuses on GAN-based segmentation solutions for specific anatomies.

##### 4.2.1. Brain:

A major focus here is the brain tumor segmentation. Xue et al. [49] propose a U-Net GAN-based framework (SegAN) in which a multi-scale loss function is

used to learn pixel dependencies. The loss function is defined as follows:

$$\min_{\theta_G} \max_{\theta_D} L(\theta_G, \theta_D) = \frac{1}{N} \sum_{n=1}^N l_{mae}(f_C(x_n \circ S(x_n)), f_C(x_n \circ y_n)) \quad (6)$$

Where  $l_{mae}$  is Mean Absolute Error (MAE) or  $L_1$  distance,  $(x_n \circ S(x_n))$  is the input image masked with a generated segmentation mask,  $(x_n \circ y_n)$  is an input image masked by the ground-truth segmentation mask,  $f_C(x)$  shows features extracted from the input image  $x$ . Rezaei et al. [50] use a multi-class approach combining cGAN and MGAN to solve the same problem. Zeju et al. [51] propose a pipeline of preprocessing (intensity normalization and mean/distribution equalization), GAN, and post-processing (patch concatenation) for segmenting brain tumors. Moeskops et al. [52] demonstrate that using GANs training strategy in addition to DCNN methods, not only can enhance the performance of deep semantic segmentation methods, but also can bring the functionality of non-semantic segmentation methods closer to semantic ones.

Since the performance of most of the supervised segmentation methods degrades on unseen images, Kamnitsas et al. [53] propose unsupervised domain adaption for brain lesion segmentation. In this method, the generator extracts invariant features of inputs from different domains and then generates the segmentation mask. In this way, having data of a target domain corresponds to one of the input domains can lead the mapping procedure from other inputs (from different domains) to their corresponding targets.

Zhao et al. [54] show that how synthetic images information can enhance the performance of segmentation. To segment the bony structure in brain MRI images, they propose Deep-supGAN architecture in which the segmenter investigating MRI real images and their corresponding synthetic CT images (as the input) generates segmentation mask of bony parts. In addition to adversarial loss a voxel-wise loss and 3 perceptual losses due to the extracted features by a VGG network is considered to train the model.

#### 4.2.2. Chest:

Bad quality, local artifacts and the overlap of lung and heart area are the main obstacles for the segmentation procedure in chest X-Ray images. Existing approaches on this field do not provide a balance on global and local features. Dai et al. [55] propose a GAN based solution (SCAN) to enhance global consistency of segmentation and extract contours of the heart and left/right lungs. They propose a fully connected network with a VGG down-sampling path and residual blocks to use much fewer feature maps in the generator.

#### 4.2.3. Eye:

In retinal vessel segmentation, many CNN-based approaches perform even better than human experts. But segmented vessels can be blurry or contain false positives near minuscule or faint branches. Son et al. [56] replace CNN with GAN (U-Net as the generator) for segmenting such vessels. Lahiri et al. [57] modify DC-GAN for segmenting RoI patches from the background. While a

similar CNN needs a huge amount of training data to perform well, [57], using 9 times less training data, achieves comparable performance. Shankaranarayana et al. [58], on the other hand, modifies cGAN to segment optic disc and cup in 2D color fundus images. The generator is a U-net (with residual blocks) which is trained by a combination of adversarial and  $L1$  losses.

#### 4.2.4. *Abdomen:*

Varying size and shape of the spleen in abdomen MRI images lead to false labeling in deep CNN segmentation methods. Huo et al. [59] employ a cGAN (SSNet: splenomegaly segmentation network) to address this problem. In the proposed model the Generator is a novel deep network architecture inspired by the global convolutional network [60], which uses larger convolutional kernels to have better segmentation on objects with large variations.

Yang et al. [61] propose a liver segmentation method in 3D abdomen CT images in which, the generator is a convolutional encoder-decoder inspired by the U-Net architecture. Kim et al. [62] employs cycleGAN for liver and tumor segmentation. In this architecture one generator generates a segmentation mask from the input image and the other one generates a CT image from the segmentation mask. In order to enhance the performance of the model in segmenting tiny tumors, polyphase U-Net architecture is used as the generator.

#### 4.2.5. *Microscopic images:*

Automatic segmentation of microscopic images is challenging due to the variety of size, shape, and texture [63, 64]. Kecheril et al. [63] propose to use GAN with different training loss function, which considers a weight to specify which pixels in foreground/background are more important. The proposed architecture is a combination of U-net with long/short skip connections, ResNet, and multi-scale CNN. In addition, a post-processing procedure is proposed to correct the segmented area. Arbelle et al. [64] use GAN with special blocks (convolution followed by batch normalization) in discriminator for the same problem. Zhang et al. [65] propose DAN - a combination of DCAN [66] and VGG16. This model is trained with a combination of adversarial loss and multi-class cross entropy.

#### 4.2.6. *Cardiology:*

Low contrast, high level of noise, and cardiac motion challenges Left Ventricle (LV) segmentation. Dong et al. [67] propose VoxelAtlasGAN which combines an atlas-based segmentation method with cGAN architecture to segment LV in low-contrast cardiography images. In this method first, the shape and intensity of the atlas is estimated by a CNN (V-Net [68]) and then a deformation network outputs the segmented image (Figure 8). Both of the mentioned networks are placed in the generator, which combines three loss functions for training: i) adversarial loss, ii) intensity loss, and iii) label loss, which compares the intensity and shape of the segmented real image with the generated one respectively.

Since in segmentation, residual (non-RoI) information in addition to RoI features can make segmentation results more realistic, Chatsias et al. [69] and Joyce et al. [70] propose to benefit residual information by defining a loss based

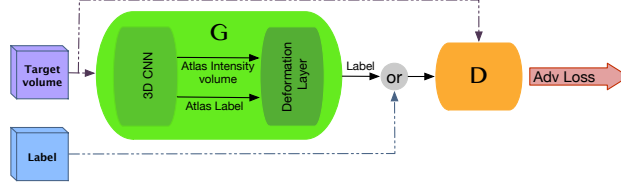


Figure 8: Proposed architecture in [67]

on it. In both papers, the input image is reconstructed from RoI and residual extracted features and reconstruction loss is considered in addition to adversarial loss. Chartsias et al. [69] propose to use cycleGAN architecture for semi-supervised segmentation and Joyce et al. [70] propose LSGAN architecture for un-supervised segmentation. Moreover, two other losses are proposed to avoid over-segmentation and large intensity variance in the segmented regions.

To segment and quantify precise myocardial infarction Xu et al. [71] introduce a GAN based architecture (MuTGAN) which can benefit from all time sequences frames information to segment infected area.

#### 4.2.7. Spine:

Vertebrae segmentation and localization is the first step for diagnosis of the vertebrae disease and surgery planning. Although machine learning based approaches achieved some success in this field, they suffer from not learning the anatomy of the region of interest. To overcome this problem Sekuboyina et al. [72] propose a butterfly shape model benefiting adversarial training to segment and localize discs in vertebra CT images. The generator is using two views of CT images to capture both the spine curve and the rib-vertebrae joints. First, in a pre-processing step, a single-shot object detection (SSD) [73] detects spine region. Then the proposed model segments discs in two views of vertebrae and finally, in a post-processing step, these results are combined for disc localization.

Tables 4 to 10 summarizes GAN-based segmentation methods. GAN-based segmentation methods mainly worked on architectural subjects to address previous methods and GANs drawbacks. It seems that from the known DNN architectures, U-Net and ResNet - due to providing general identification features - are the most popular segmentation networks.

#### 4.3. Reconstruction

Fast MR reconstruction retaining details is a core problem in medical imaging. Solving the slow acquisition and reconstruction can directly benefit lingering MR problems like presence of involuntary (for example resulted by breathing) and voluntary (for example movements due to not comfortable situation) motion artifacts. Classic compressed sensing-based solutions directly use k-space information to reconstruct images [75]. The promise of realistic looking images with fast inference makes GAN an obvious candidate for solving the MR reconstruction problem. Unlike de-noising, GAN-based MR reconstruction

Table 4: Segmentation GAN-based methods - Brain

Method	Image Modality	Dataset	Performance
<b>[49] SeGAN</b> <b>Architecture:</b> U-Net, GAN <b>Loss:</b> Adv, weighted on multiScale features	MRI	BRATS 2013 (Leadboard)	(whole, Core, Enhanced) Dice = 0.84, 0.70, 0.65 Precision = 0.87, 0.80, 0.68 Sensitivity = 0.83, 0.74, 0.72
		BRATS 2015 (Test)	Dice = 0.85, 0.70, 0.66 Precision = 0.92, 0.80, 0.69 Sensitivity = 0.80, 0.65, 0.62
<b>[50]</b> <b>Architecture:</b> c-GAN, MGAN	MRI	BRATS 2017	(Whole, core, Enhanced) Dice = 0.70, 0.55, 0.40 Sensitivity = 0.68, 0.52, 0.99 Specificity = 0.99, 0.99, 0.99
<b>[52]</b> <b>Architecture:</b> GAN <b>Loss:</b> Adv, cross entropy	MRI	MICCAI 2012 Challenge (adult)	Dice = 0.92±0.03
		MRBrainS13 challenge(elderly)	Dice = 0.85±0.01
<b>[51]</b> <b>Architecture:</b> GAN <b>Loss:</b> Adv	MRI	BRATS 2017	(Whole, Core, Enhancing) Dice = 0.87, 0.72, 0.68 sensitivity = 0.87, 0.72, 0.68
<b>[53]</b> <b>Architecture:</b> GAN, 3D-CNN <b>Loss:</b> Adv, SGD			Dice = 0.62 Recall = 0.58 Precision = 0.71
<b>[54] Deep-supGAN</b> <b>Architecture:</b> GAN, VGG16 <b>Loss:</b> Adv, Preceptual, voxel-wise	MRI 3D	ADNI	Dice = 94.46

Table 5: Segmentation GAN-based methods - Chest

Method	Image Modality	Dataset	Performance
<b>[55] SCAN</b> <b>Architecture:</b> VGG, ResNet <b>Loss:</b> Adv, pre-trained by Pixel-wise loss	X-Ray	JSRT(247)	(Lungs, Heart) Dice = 0.973, 0.927
		Montgomery(135)	IoU = 0.947, 0.866

Table 6: Segmentation GAN-based methods - eye

Method	Image Modality	Dataset	Performance
<b>[56]</b> <b>Architecture:</b> U-Net, GAN <b>Loss:</b> Adv, Cross entropy	Funduscopy (Retina)	DERIVE	Dice= 0.829 ROC=0.9803 PR=0.9149
		STARE	Dice= 0.834 ROC=0.9838 PR=0.9167
<b>[57]</b> <b>Architecture:</b> DCGAN <b>Loss:</b> Adv, L-classification	Funduscopy (Retina)	DERIVE (blood vessels)	AUC= 0.945
<b>[58]</b> <b>Architecture:</b> c-GAN, ResU-net <b>Loss:</b> Adv, L1	Funduscopy (Retina)	RIM-ONE	(Optic disc, Optic cup) F-score= 0.97, 0.94 IOU=0.89, 0.76

research has a major focus of modifying well-known architectures in addition to combination of loss functions for controlling the training procedure.

#### 4.3.1. DAGAN-based strategies

Early research on GAN-based MR reconstruction focuses around the DAGAN architecture [76]. A series of work, refines the DAGAN architecture over time [75, 76] while refining the combination of loss functions that performs best

Table 7: Segmentation GAN-based methods - Abdominal

Method	Image Modality	Dataset	Performance
[59] <b>SSNet</b> <b>Architecture:</b> GCN, cGAN <b>Loss:</b> Adv, Dice	MRI (Splenomegaly)	unknown	Dice=0.9260
[61] <b>Architecture:</b> U-Net, encoder-decoder <b>Loss:</b> Adv, Multi-class entropy	CT 3D (Liver)	MICCAI-SLiver07	Dice=0.95 ASD=1.90
[62] <b>Architecture:</b> U-Net, cycleGAN <b>Loss:</b> cycleGAN, Cross entropy, L2	CT 3D (Liver)	LiTS2017	(liver, lesion) Dice= 0.89, 0.46 Recall=0.94, 0.5 Precision=0.86, 0.48

Table 8: Segmentation GAN-based methods - Microscopic

Method	Image Modality	Dataset	Performance
[63] <b>Architecture:</b> GAN, U-net, Res-Net, Multi scale CNN <b>Loss:</b> Adv, weighted loss	Bright-fieled Cell 2D	Columbus  MetaXpress	F-score = 0.77 Precision = 0.82 Recall = 0.73 F-score = 0.64 Precision = 0.66 Recall = 0.66
[64] <b>Architecture:</b> GAN (with rib cage) <b>Loss:</b> Adv	Cell 2D	H1299	F-score = 0.89 Precision=0.82 Recall = 0.85
[65] <b>DAN</b> <b>Architecture:</b> GAN, DCAN, VGG <b>Loss:</b> Adv, Multi-scale cross entropy	Fungus 3D	2015 MICCAI Gland Challenge	(mean of 2 part results) F-score = 0.88 Dice=0.865 ObjectHausdorff = 74.55

Table 9: Segmentation GAN-based methods - Cardiology

Method	Image Modality	Dataset	Performance
[67] <b>VoxelAtlasGAN</b> <b>Architecture:</b> cGAN, V-Net <b>Loss:</b> Adv, intensity, label	Echo- cardiography 3D	unknown	Dice=0.95 MSD=1.85 HSD=7.26 corr-of-EF=0.91 time=0.1
[69] <b>Architecture:</b> cycleGAN <b>Loss:</b> Adv, Dice, MAE	Cine MR 3D	2017 ACDC Challenge, Aprivate dataset	unlabeled data num:1200 (labeled data num, F-score) 11 , 0.415 34 , 0.678 68 , 0.731 142 , 0.767 284 , 0.771
[70] <b>DAN</b> <b>Architecture:</b> LSGAN, U-net <b>Loss:</b> Adv, Intensity Variance, Over-Seg penalty, Reconstruction	CT or MR Cardiac 2D	2017 MM-WHS Challenge	(Average on 3 Seg Results) (MR) Dice = 0.66 (CT) Dice=0.5
[71] <b>MuTGAN</b> <b>Architecture:</b> GAN, ConvLSTM, 3DConv <b>Loss:</b> Adv, MAE, Dice	3T MR Cardiac cine, DE-MR	unknown	Dice=0.90 Accuracy=96.46 Infarct size=22.3

Table 10: Segmentation GAN-based methods - Spine

Method	Image Modality	Dataset	Performance
[72] <b>Btrfly Net</b> <b>Architecture:</b> GAN, Btrfly-Net <b>Loss:</b> Adv, Btrfly-Net	CT 3D	[74]	Precision=0.84 Recall=0.83 F1-score= 0.84

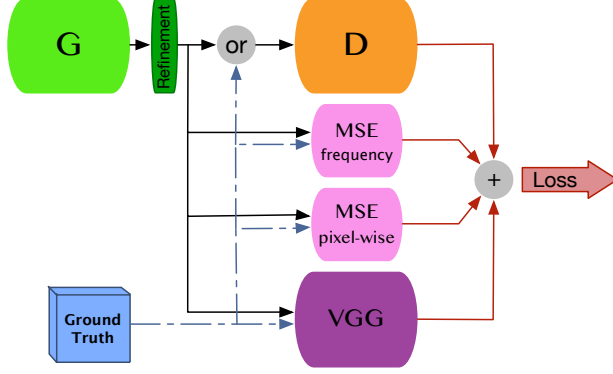


Figure 9: DAGAN architecture [76]

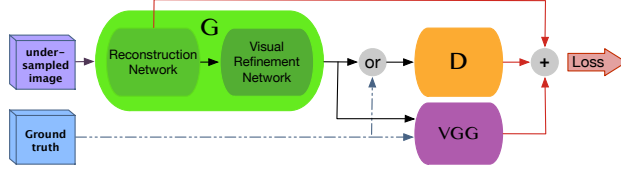


Figure 10: Proposed architecture in [77]

(Figure 9). These loss functions include pixel-wise MSE, adversarial loss, perceptual loss and frequency MSE, resulting in a combined loss function of the following form:

$$L_G = \alpha L_{image-MSE} + \beta L_{freq-MSE} + \gamma L_{VGG} + L_{GAN} \quad (7)$$

The next update of DAGAN is the refinement layer (Figure 10) introduced by Seitzer et al. [77], along with a combination of four losses as shown below:

$$L_{ref} = \frac{1}{2} \left( \frac{L_{adv}}{M} + \frac{L_{feat}}{N} \right) + \frac{L_{VGG}}{O} + \alpha L_{pen} \quad (8)$$

where  $L_{feat}$  is feature matching loss proposed in [78] and  $L_{pen}$  is a penalty to force the network to manipulate the result of MSE optimized network with the least changes, and  $L_{adv}$  and  $L_{VGG}$  are similar to the losses used in the DAGAN.

Similar design of generator architecture (Figure 11) is also proposed by Quan et al. [22]. However, their cyclic training strategy leads to two further losses as shown in the following:

$$L_G = L_{adv} + \alpha L_{freq} + \beta L_{imag} \quad (9)$$

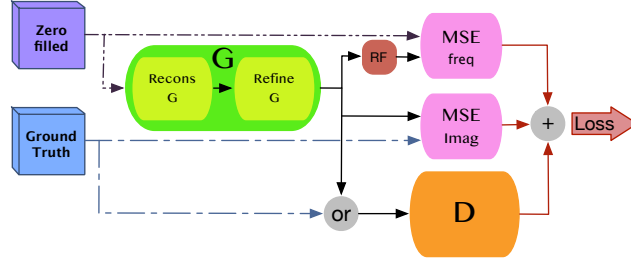


Figure 11: RefineGAN architecture [22]: The Generator G is a chain of two concatenated generators (first generator is for reconstruction and the second one is for refinement) cycle loss is calculated by MSE blocks

#### 4.3.2. 3D super-resolution strategies

3D convolutional blocks are also considered for MR reconstruction. Snchez et al. [79], adapt SRGAN [80] with 3D convolutional layers to deal with volumetric information. Their loss function combines a pixel-wise loss with a Gradient Based Loss (GDL) [81].

The problem of huge memory footprint of 3D convolution is addressed by Chen et al. [82]. Their proposed network combines WGAN and DenseNet [83] in a multi-level Densely Connected Super-Resolution Network (mDCSRN) for fast ME reconstruction.

For tomography reconstruction, Li et al. [84] propose 3DSRGAN along with a combination of four different generator losses to reconstruct thin slice tomographic 3D images from thick ones. In addition to adversarial and MSE, two other loss functions are defined for overcome over-fitting and control data interpolation.

#### 4.3.3. Other Methods

Mardani et al. [85], Shitrit et al. [86], and Zhang et al. [87] propose generator design strategies based on ResNet and added  $\ell_1 \setminus \ell_2$  losses to adversarial training loss to save more details of the image. In [87] data consistency loss and perceptual are also added to save more detailed k-space information.

Ravi et al. [88] modifies the cyclic consistency approach with two specialized losses for endomicroscopy Super Resolution (SR) application. These two specialized losses are  $l_{Vec}$  as the distance between Voronoi vectorized form of SR and LR images and  $l_{Reg}$  to regularize the training. Han et al. [89] propose to use enhanced contrast input images as inputs and add a perceptual loss to adversarial training due to diffraction pattern filters [90]. Finally, Mahapatra [91] proposes super resolved generative adversarial network for super resolution of retinal images adding image saliency loss for model training.

Table 11, 12 and 13 summarize properties of mentioned methods and their performance. It seems that GANs can provide good performance in reconstruction of medical images, by adding some manipulation in loss functions, which



highlights texture details and special features.

Table 11: Reconstruction GAN-based methods - DAGAN based

Method	Image Modality	Dataset	Performance
<b>[75]</b> <b>Arch:</b> cGAN, U-Net <b>Loss:</b> Adv, Pix-wise, Perceptual,Refinement	MRI	Brain: IXI, MICCAI (2013 Grand Challenge)	mask 30%: NMSE=0.09±0.02 PSNR=39.53±4.12 (CPU, GPU) time=0.2±0.1, 5.4±0.1(ms)
<b>[76] DAGAN</b> <b>Architecture:</b> cGAN,U-Net <b>Loss:</b> Adv, Pix-wise, Frequency, Perceptual, Refinement	MRI	Brain: IXI, MICCAI (2013 Grand Challenge)	mask 30%: NMSE=0.08±0.02 PSNR=40.20±4.07 (CPU, GPU) time= 0.2±0.1, 5.4±0.1(ms)
<b>[77]</b> <b>Architecture:</b> cGAN,U-Net <b>Loss:</b> Adv, feature matching, Perceptual, penalty	MRI Cardic	unknown	PSNR=31.82±2.28 MOS=3.24±0.63 (max=3.78±0.45) SIS(max=1)=0.94
<b>[22] RefineGAN</b> <b>Architecture:</b> Chain ofgenerator, ResNet <b>Loss:</b> Adv, Cyclic	MRI	Brain: IXI  Chest: Data Science Bowl challenge	mask 30%, time:0.16(s) SSIM=0.97±0.01 PSNR=38.71±2.57 mask 30%, time:0.18(s) SSIM=0.97±0.01 PSNR=38.64±2.76

Table 12: Reconstruction GAN-based methods - 3D super-resolution

Method	Image Modality	Dataset	Performance
<b>[79]</b> <b>Architecture:</b> SRGAN, subpixel-NN <b>Loss:</b> LSGAN, GDL, Pixel-wise	MRI(Brain)	(ANDI database)	(Scale 2, Scale 4) PSNR=39.28, 33.58 SSIM=0.98, 0.95
<b>[82] mDCSRN</b> <b>Architecture:</b> DensNet, WGAN <b>Loss:</b> MSE, WGAN	MRI(Brain)	unknown	SSIM=0.94 PSNR=35.88 NRMSE=0.0852
<b>[84]</b> <b>Architecture:</b> Res blocks, GAN <b>Loss:</b> Adv, Pixel-wise 3D total variation	MRI(Brain)	(Glioma Patients)	MSE=262.2 PSNR=24.2

#### 4.4. Detection

Computational detection of anomalies from images using supervised algorithms require a large amount of annotated training data - extremely difficult to be provided by experts. Unsupervised algorithms, such as GAN have the potential to be effective in such a setting. The main driving idea of anomaly detection is to train GAN on normal images, so that inference on diseased images would lead to incomplete reconstructions. Such incompleteness is typically associated with the regions of anomaly.

Schlegl et al. [17] show such an idea can be effective in detecting anomalies in optical coherence tomography images of the retina. They propose an unsupervised GAN-based architecture (AnoGAN) that during testing, given an input diseased image generates the corresponding healthy image version. The difference between generated and input images is considered as anomalies. Chen et

Table 13: Reconstruction GAN-based methods - other

Method	Image Modality	Dataset	Performance
<b>[85] GANCS</b> Architecture: ResNet, LSGAN	MRI(Chest)	contrast-enhanced MRI abdomen dataset of pediatric patients	SNR=20.48 SSIM=0.87 Time=0.02
<b>[86]</b> Architecture: ResNet, GAN Loss: Adv	MRI(Brain)	unknown	PSNR=37.95
<b>[87] GANCS</b> Architecture: ResNet, GAN, VGG	MRI(Brain 2D)	unknown	PSNR=32.32 SSIM=0.88 Time=0.37
<b>[88]</b> Architecture: [92], GAN, Cyclic, Loss: Adv, Regularization	Endomicroscopy	[93]	SSIM=0.8.7 $\Delta GCF_{HR} = 0.66$ $\Delta GCF_{LR} = 0.37$ $Tot_{cs} = 0.66$
<b>[89]</b> Architecture: [90], GAN Loss: Adv, Perceptual	Microscopy (Cell)	unknown	PSNR = 27.8591
<b>[91]</b> Architecture: ResNet, GAN Loss: Adv, CNN (weighted by SL map)	Retinal Funduscopy	unknown	(Scale 4, Scale 8) SSIM=0.89, 0.84 RMSE=6.2, 7.5 PSNR=44.3, 39db

al. [94] and Baur et al. [95] modify [17] for anomaly detection in brain MR images. Baumgartner et al. [96] modify the original AnoGAN further, introducing a map generator based on WGAN and the U-Net (VA-GAN) to detect changes of the brain from Alzheimer’s disease. Though CNNS [97] show good performance in detection of high contrast lesions from Alzheimer’s, VA-GAN focuses on the detection of low contrast lesions. In particular, VA-GAN generates a map  $M$  which converts a healthy image  $x_i$  into the corresponding diseased one ( $y_i$ ). The loss function is modified as well in the following way:

$$L_{GAN}(M, D) = \mathbb{E}_{x \sim p_d(x|c=0)}[D(x)] - \mathbb{E}_{x \sim p_d(x|c=1)}[D(x + M(x))] \quad (10)$$

A regularization term is added to the loss for inducing minimal changes, resulting in the final loss function:

$$L = L_{GAN}(M, D) + \lambda \|M(x)\|_1 \quad (11)$$

Similar modifications lead to solutions for aggressive prostate cancer detection [98], prostate landmark detection [99] and skin lesion detection [100].

Table 14 summarize these papers. Papers proposed in anomaly detection by GANs have more structural complexity in comparison with previous applications because they benefit from different aspects of GANs. In fact, the role of the discriminator is more highlighted in practice. Also, the extracted map, which defines the latent aspect of recognizing the healthy and anomaly images is used in a more perceptual way.

#### 4.5. De-noising

Image acquisition of several forms of diagnostic radiology involves the trade-off between contrast and radiation hazard. The better contrast might lead to better diagnosis but might expose the patient to unwanted excessive radiation

Table 14: Detection GAN-based methods in medical image processing.

Method	Image Modality	Dataset	Performance
[17]AnoGAN Architecture: DCGAN	SD-OCT scans	unknown	Precision= 0.8834 Recall= 0.7277 Sensitivity=0.7279 Specificity=0.8928 AUC=0.89
[94] Architecture: AnoGAN, WGAN-GP Loss: WGAN-GP, Regularization	MRI(brain)	BRATS	AUC = 0.92
[96]VA-GAN Architecture: WGAN, U-Net	MRI (brain)	ADNI	NCC = 0.27
[98] Architecture: U-Net, GAN Loss: MSE, GAN	MRI (prostate)	(NCT) Heidelberg	Specificity=0.98 $\pm$ 0.14 Dice=0.41 $\pm$ 0.28 Sensitivity=0.55 $\pm$ 0.36
[100] Architecture: cGAN, U-net	Natural Image (Skin Lesion)	unknown	(Subjective) Correct Lesion Detection = 0.914
[99] Architecture: GAN Loss: Adv, Landmark location, Contour association	Ultra-Sound (prostate)	unknown	Dice = 0.92 $\pm$ 0.3

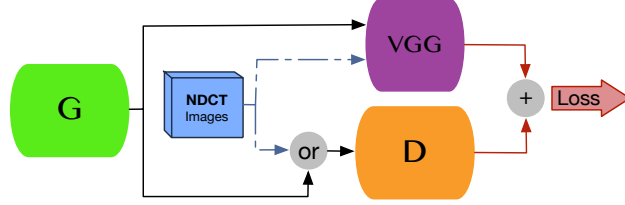


Figure 12: Proposed architecture in [103]

as well. Images acquired under low contrast, on the other hand, suffers from low Signal-to-Noise Ratio. State-of-the-art denoising algorithms suffer from blurring effect, which is taken care of by several engineering tweaks and perceptual assumptions. The ability of GANs for generating realistic looking images provides a nice alternative. In fact, several authors exploit GANs to produce de-noised high perceptual quality images. The typical workflow, however, is based mainly on GAN loss functions for better results.

For example, Wolterink et al. [101] propose to learn texture information from a small amount of paired data using a combination of voxel-wise MSE and adversarial loss. Also, Wang et al. [102] propose to use cGAN to remove metal artifacts from CT images. Yang et al. [103], On the other hand, utilize a combination of a perceptual loss and Wasserstein loss to bring in numerical stability without losing perceptual quality. The sharpness of denoised image is also the factor that Yi et al. [104] (Sharpness Aware Generative Adversarial Network - SAGAN) work on. They combine three losses in their design, namely: pixel-wise loss, adversarial loss and a sharpness mapping loss (Figure 13)

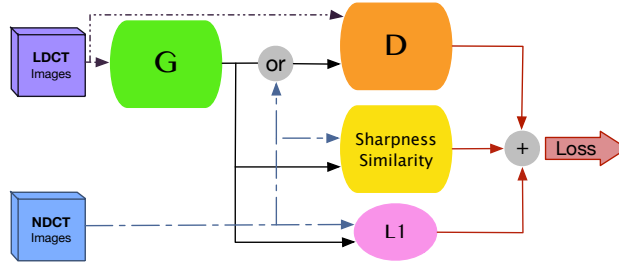


Figure 13: SAGAN architecture [104]

Table 15: De-noising GAN-based methods in medical image processing

Method	Modality	Dataset	Performance
[101] <b>Architecture:</b> CNN, GAN <b>Loss:</b> CNN, GAN	CT (phantom) and (cardiac)	unknown	Agatston Score: Median=20.7 Min=6.1 Max=145.1
[102] <b>Architecture:</b> U-net, Res blocks, cGAN <b>Loss:</b> GAN	CT of ear	unknown	P2PEs: Median=0.409 STD=0.133 Max=0.912
[103] <b>Architecture:</b> WGAN, VGG <b>Loss:</b> features distance, WGAN	CT	unknown	Noise Suppression= $3.20 \pm 0.25$ Artifact Reduction= $3.45 \pm 0.25$ Overall Quality= $3.70 \pm 0.15$
[104] <b>SAGAN</b> <b>Architecture:</b> MGAN, Sharpness detector <b>Loss:</b> Pixel-wise, MGAN, Sharpness aware	CT	CT phantom (Catphan 600)	( $N=10^4$ ) PSNR=26.77 SSIM=0.8454 ( $N=10^5$ ) PSNR=28.25 SSIM=0.87

Table 15 summarizes major GAN-based de-noising methods. It seems that an adequate objective metric to evaluate the strength of methods in preserving important medical information of the image is not available yet. As PSNR, MSE, SSIM, SD, and mean - the most commonly used metrics in the evaluation of de-noising methods - are not sensitive enough to recognize texture details, the RoI of any image should be segmented to be measured by metrics, which is an expensive procedure. So presenting a new metric to this goal can be a subject of future works. Despite this limitation, reviewed papers benefiting from the ability of GANs in learning main general features of a domain of images and also by manipulating the loss function to consider more textural features, show good performance in medical image de-noising. However, finding a fast, accurate and more stable architecture is an open direction to be worked on in the future.

#### 4.6. Registration

Image registration is another application in medical image processing which contributes to align different medical images and extract combined information of them. Also, registration models can be used as a translator to find

similar patterns in medical images [1]. While traditional registration methods suffer from task dependent parameter tuning and high volume of optimization computation [105], DL-based methods address some of these limitations. However, GAN-based architectures, providing a better general overview of image features, are recently introduced to achieve better performance in both supervised and unsupervised registration. Fan et al. [105] propose an un-supervised GAN to register a structural pattern (defined in a patch) in different brain images. Yan et al. [106] and Hu et al. [107] focusing on prostate 3D MRI and intra-procedural Transrectal Ultra Sound (TRUS) images respectively propose GAN for registration and deformation correction. While, usually in GAN-base medical architectures discriminator is only used for adversarial training, in [106] it remains in the model as a certainty evaluator.

Table 16 summarizes properties of mentioned methods and their performance. Since in registration tasks general features of the images are important, GAN-based models (manipulated by well designed training costs) are good choices to provide distributional based information.

Table 16: Registration GAN-based methods in medical image processing

Method	Image Modality	Dataset	Performance
<b>[105]</b>			
<b>Architecture:</b>		LPBA40,	Dice=71.8 $\pm$ 2.3
U-net, GAN	Brain MRI 3D	IBSR18,	Dice=57.8 $\pm$ 2.7
<b>Loss:</b>		CUMC12,	Dice=54.4 $\pm$ 2.9
Adv, Regularization		MGH10	Dice = 61.7 $\pm$ 2.1
<b>[106]</b>			
<b>Architecture:</b>	Prostate 3D		TRE = 3.84 mm
WGAN	MRI and TRUS	unknown	Dice = 0.58
<b>Loss:</b>			
Adv			
<b>[107]</b>			
<b>Architecture:</b>	Prostate 3D		TRE = 6.3 mm
3D encod-decod, GAN	MRI and TRUS	unknown	Dice = 0.82
<b>Loss:</b>			
Adv, Dice, Regularization			

#### 4.7. Classification

Due to cardiac and respiratory motions cardiac Ultra-Sound (US) images display incomplete information e.g. missing basal and/ or apical slices of Left Ventricle. To discard such unsuitable images, Zhang et al. [108] propose Semi-Coupled GAN (SCGAN) which consists of two generators and one discriminator. Initially, the generators produce new cardiac samples (with and without the basal slice) using learned high level features from both categories. The discriminator not only distinguishes between generated and real images but also classifies images into two classes: those being basal slice or not (Figure 14). To classify prostate histopathology to two Gleason grades Ren et al. [109] propose GAN to extract fixed features of every domain and then use them to classify input patches. For training in addition to the adversarial loss, a classification loss is defined which prevents the network to classify patches of the same image to be classified in the different domains.

Table 17 summarizes properties of mentioned methods and their performance. Due to lack of adequate properties of medical images, it seems that

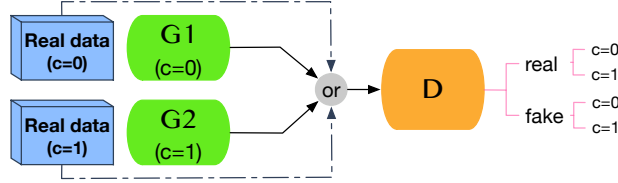


Figure 14: SCGAN architecture [108]

GANs - generating demanded samples - benefit the classification task in these images.

Table 17: Classification GAN-based methods in medical image processing

Method	Image Modality	Dataset	Performance
[108] SCGAN			(MAS - MBS)
Architecture:			Acc=92.5, 89.3
GAN	Cardiac MR	UKBB	Precision=87.6, 89.1
Loss:			Recall=90.5, 91.7
Adv			
[109]			
Architecture:	prostate	TCGA,	(TCGA , TCGA to CINJ)
[90], GAN	histopathology	CINJ	Accuracy = 0.77, 0.75
Loss:			
Adv, patch-class			

## 5. Discussion

### 5.1. Overview

GANs are receiving significant attention from the medical imaging community - this is evident by the sudden spike in the number of papers published using GANs. We found a total of 77 papers by searching Google Scholar and PubMed with 'GAN' or 'Generative Adversarial Networks' in title or keywords. Of these, we shortlisted 77 papers for review based on the innovation in key aspects of the GAN architecture. Of these 77, 28 are proposed in synthetic applications. However, the application fields are quite diverse ranging from segmentation, reconstruction all the way to de-noising - showing possible applications of GANs across many medical tasks.

### 5.2. Benefits of GANs in the medical field

GANs are capable of producing realistic looking images, providing major advantages over the more established discriminative frameworks in two challenges that are unique to medical settings:

**Scarcity of adequate dataset:** Often times, annotations are expensive and hard to come-by in medical imaging. Deploying supervised learning based deep neural networks for such problems are challenging. This leads to the possibility of semi- or un-supervised learning, but even their performance depends highly on the training data imbalances. GANs, as state of the art in generative models, show superior ability in both semi- and un-supervised learning.

Unconditional GANs show significant performance improvement in generating simulated images which are not visually distinguishable from real images. In addition to addressing data imbalance problem of supervised deep neural network training (Section 4.1.1), GANs can reconstruct lost (or destroyed) components of the image (Section 4.3) as well. The idea of multi-modal image fusion for better diagnostic decision making is very well grounded. However, acquisition of such multi-modality images leads to increase in cost and associated risk factors. By synthesizing across modalities, conditional GANs provide a solution for this problem (Section 4.1.2). Moreover, GAN’s ability to learn the underlying data distribution opens up the possibility to detect unseen (healthy or disease) cases in the real datasets (Section 4.4).

**Limitation of identity feature extraction:** Learning distinctive identifier patterns in medical images plays a vital role in diagnosis of diseases. Therefore, the performance of any medical application is interconnected to its ability to detect and preserve these identical features. Pixel-wise losses which are defined for deep learning optimization might not be sufficient to discover some of general features of the image. Equipped with the adversarial training, GANs can extract semantically meaningful features of the images. This property leverages GAN-based segmentation techniques to perform segmentation (Section 4.2) and registration (Section 4.6) in an end-to-end setting, without the need of any post-processing. Hierarchical discriminative potential of GAN discriminators are utilized for classification application (Section 4.7) Finally, extracting (and preserving) main visual features of the image by GANs leads to blurring-free de-noising (Section 4.5).

### 5.3. Drawbacks

We identify three major drawbacks in the current form of GANs that might hinder its acceptance in the medical community:

**Trustability of Synthesized Data:** In healthcare, where trustability of the clinicians is the biggest challenge for any technology, images synthesized by GANs provide little comfort. The basic networks (generator and discriminator) are still deep neural networks, the mechanism of which is not well studied. In computer vision, where the overall perception is the main concern, these results are adequate. In medical images, however, intensities are typically associated with some meanings e.g. tissue types can be broadly categorized based on HU of CT data. Such an association and mapping is currently missing from the GAN reconstruction - a shortcoming severe enough for clinicians to distrust images synthesized by GAN.

**Unstable Training:** The typical GAN training is unstable because of numerical reasons pointed out in learning literature [3]. This results in situations such as mode collapse. State-of-the-art learning theory focuses on solving such numerical problems in GANs training for real images. However, in medical imaging, where the modes of images are unclear, how to identify such a problem is unclear. This leads to the question of what sort of numerical singularities could arise in medical imaging and how to address those.

**Evaluation Metric:** This is a problem, in tandem with the general computer vision community. The best possible way to evaluate reconstruction result is still unclear. In medical imaging, researchers rely mostly on traditional metrics such as PSNR or MSE to evaluate GANs reconstruction quality. This is a tricky situation in the sense that the disadvantages of such metrics were the main reason to move toward GANs. So how can we evaluate potentially better results with metrics which are not capable to understand it?

#### 5.4. Future Works

We believe GANs need to address the significant drawbacks discussed in Section 5.3 before being a technology that is trusted in healthcare. To this end, we can think of GANs as a technical building block rather than a stand-alone piece of technology for the future. For example, in the case of synthesizing CT data, enveloping GANs synthesis with a physics-based simulation might ensure realistic HU values.

The training instability issue needs to be addressed as well, which means rigorous experimentation to understand the convergence and saddle points of GAN in the medical imaging context. The question regarding metric is far trickier, going about with understanding the performance of GANs synthesized images in CAD by clinicians is a necessary first step.

In short, along with exciting results, GANs open up many possible research questions for the next few years. Proper understanding and answering those hold the key to successful GANs deployment in the real clinical scenario.

## References

- [1] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. van Ginneken, C. I. Sánchez, A survey on deep learning in medical image analysis, *Medical image analysis* 42 (2017) 60–88.
- [2] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, in: *Advances in neural information processing systems*, 2014, pp. 2672–2680.
- [3] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, A. A. Bharath, Generative adversarial networks: An overview, *IEEE Signal Processing Magazine* 35 (1) (2018) 53–65.
- [4] P. Isola, J.-Y. Zhu, T. Zhou, A. A. Efros, Image-to-image translation with conditional adversarial networks, 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2017) 5967–5976.
- [5] C. Szegedy, W. Zaremba, I. Sutskever, J. Bruna, D. Erhan, I. J. Goodfellow, R. Fergus, Intriguing properties of neural networks, *CoRR* abs/1312.6199.



- [6] A. Radford, L. Metz, S. Chintala, Unsupervised representation learning with deep convolutional generative adversarial networks, CoRR abs/1511.06434.
- [7] M. Mirza, S. Osindero, Conditional generative adversarial nets, CoRR abs/1411.1784.
- [8] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical image computing and computer-assisted intervention, Springer, 2015, pp. 234–241.
- [9] C. Li, M. Wand, Precomputed real-time texture synthesis with markovian generative adversarial networks, in: European Conference on Computer Vision, Springer, 2016, pp. 702–716.
- [10] J.-Y. Zhu, T. Park, P. Isola, A. A. Efros, Unpaired image-to-image translation using cycle-consistent adversarial networks, 2017 IEEE International Conference on Computer Vision (ICCV) (2017) 2242–2251.
- [11] A. Odena, C. Olah, J. Shlens, Conditional image synthesis with auxiliary classifier gans, in: ICML, 2017.
- [12] M. Arjovsky, S. Chintala, L. Bottou, Wasserstein gan, CoRR abs/1701.07875.
- [13] X. Mao, Q. Li, H. Xie, R. Y. K. Lau, Z. Wang, S. P. Smolley, Least squares generative adversarial networks, 2017 IEEE International Conference on Computer Vision (ICCV) (2017) 2813–2821.
- [14] M. Frid-Adar, E. Klang, M. Amitai, J. Goldberger, H. Greenspan, Synthetic data augmentation using gan for improved liver lesion classification, 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018) (2018) 289–293.
- [15] M. J. M. Chuquicusma, S. Hussein, J. R. Burt, U. Bagci, How to fool radiologists with generative adversarial networks? a visual turing test for lung cancer diagnosis, 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018) (2018) 240–244.
- [16] A. Kitchen, J. Seah, Deep generative adversarial neural networks for realistic prostate lesion mri synthesis, CoRR abs/1708.00129.
- [17] T. Schlegl, P. Seeböck, S. M. Waldstein, U. Schmidt-Erfurth, G. Langs, Unsupervised anomaly detection with generative adversarial networks to guide marker discovery, in: International Conference on Information Processing in Medical Imaging, Springer, 2017, pp. 146–157.
- [18] C. Bermudez, A. J. Plassard, L. T. Davis, A. T. Newton, S. M. Resnick, B. A. Landman, Learning implicit brain mri manifolds with deep learning, arXiv preprint arXiv:1801.01847.

- [19] C. Baur, S. Albarqouni, N. Navab, Melanogans: High resolution skin lesion synthesis with gans, CoRR abs/1804.04338.
- [20] C. Baur, S. Albarqouni, N. Navab, Generating highly realistic images of skin lesions with gans, arXiv preprint arXiv:1809.01410.
- [21] D. Nie, R. Trullo, J. Lian, C. Petitjean, S. Ruan, Q. Wang, D. Shen, Medical image synthesis with context-aware generative adversarial networks, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2017, pp. 417–425.
- [22] T. M. Quan, T. Nguyen-Duc, W.-K. Jeong, Compressed sensing mri reconstruction with cyclic loss in generative adversarial networks, CoRR abs/1709.00753.
- [23] J. M. Wolterink, A. M. Dinkla, M. H. Savenije, P. R. Seevinck, C. A. van den Berg, I. Išgum, Deep mr to ct synthesis using unpaired data, in: International Workshop on Simulation and Synthesis in Medical Imaging, Springer, 2017, pp. 14–23.
- [24] M. Zhao, L. Wang, J. Chen, D. Nie, Y. Cong, S. Ahmad, A. Ho, P. Yuan, S. H. Fung, H. H. Deng, et al., Craniomaxillofacial bony structures segmentation from mri with deep-supervision adversarial learning, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2018, pp. 720–727.
- [25] A. Chatsias, T. Joyce, R. Dharmakumar, S. A. Tsaftaris, Adversarial image synthesis for unpaired multi-modal cardiac data, in: International Workshop on Simulation and Synthesis in Medical Imaging, Springer, 2017, pp. 3–13.
- [26] J. P. Cohen, M. Luck, S. Honari, Distribution matching losses can hallucinate features in medical image translation, arXiv preprint arXiv:1805.08841.
- [27] J. Jiang, Y.-C. Hu, N. Tyagi, P. Zhang, A. Rimner, G. S. Mageras, J. O. Deasy, H. Veeraraghavan, 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, 2018, pp. 777–785.
- [28] P. Costa, A. Galdran, M. I. Meyer, M. D. Abràmoff, M. Niemeijer, A. M. Mendona, A. Campilho, Towards adversarial retinal image synthesis, CoRR abs/1701.08974.
- [29] P. Costa, A. Galdran, M. I. Meyer, M. Niemeijer, M. Abràmoff, A. M. Mendonça, A. Campilho, End-to-end adversarial retinal image synthesis, IEEE transactions on medical imaging.

- [30] J. T. Guibas, T. S. Virdi, P. S. Li, Synthetic medical images from dual generative adversarial networks, CoRR abs/1709.01872.
- [31] H. Zhao, H. Li, L. Cheng, Synthesizing filamentary structured images with gans, CoRR abs/1706.02185.
- [32] A. Ben-Cohen, E. Klang, S. P. Raskin, M. M. Amitai, H. Greenspan, Virtual pet images from ct data using deep convolutional networks: Initial results, in: International Workshop on Simulation and Synthesis in Medical Imaging, Springer, 2017, pp. 49–57.
- [33] L. Bi, J. Kim, A. Kumar, D. Feng, M. Fulham, 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, 2017, pp. 43–51.
- [34] W. Wei, E. Poirion, B. Bodini, S. Durrleman, N. Ayache, B. Stankoff, O. Colliot, Learning myelin content in multiple sclerosis from multimodal mri through adversarial training, CoRR abs/1804.08039.
- [35] Y. Hu, E. Gibson, L.-L. Lee, W. Xie, D. C. Barratt, T. Vercauteren, J. A. Noble, 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, 2017, pp. 105–115.
- [36] E. Gibson, W. Li, C. H. Sudre, L. Fidon, D. Shakir, G. Wang, Z. Eaton-Rosen, R. Gray, T. Doel, Y. Hu, T. Whyntie, P. Nachev, D. C. Barratt, S. Ourselin, M. J. Cardoso, T. Vercauteren, Niftynet: a deep-learning platform for medical imaging, in: Computer Methods and Programs in Biomedicine, 2018.
- [37] F. Tom, D. Sheet, Simulating patho-realistic ultrasound images using deep generative networks with adversarial learning, 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018) (2018) 1174–1177.
- [38] D. Mahapatra, B. Bozorgtabar, J.-P. Thiran, M. Reyes, Efficient active learning for image classification and segmentation using a sample selection and conditional generative adversarial network, arXiv preprint arXiv:1806.05473.
- [39] H. Cho, S. Lim, G. Choi, H. Min, Neural stain-style transfer learning using gan for histopathological images, CoRR abs/1710.08543.
- [40] N. Bayramoglu, M. Kaakinen, L. Eklund, J. Heikkilä, Towards virtual h&e staining of hyperspectral lung histology images using conditional generative adversarial networks, 2017 IEEE International Conference on Computer Vision Workshops (ICCVW) (2017) 64–71.

- [41] A. Bentaieb, G. Hamarneh, Adversarial stain transfer for histopathology image analysis, *IEEE Transactions on Medical Imaging* 37 (3) (2018) 792–802.
- [42] M. T. Shaban, C. Baur, N. Navab, S. Albarqouni, Staingan: Stain style transfer for digital histological images, *CoRR* abs/1804.01601.
- [43] L. Han, Z. Yin, Transferring microscopy image modalities with conditional generative adversarial networks, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2017, pp. 99–107.
- [44] J. M. Wolterink, T. Leiner, I. Isgum, Blood vessel geometry synthesis using generative adversarial networks, *arXiv preprint arXiv:1804.04381*.
- [45] S. Olut, Y. H. Sahin, U. Demir, G. Unal, Generative adversarial training for mra image synthesis using multi-contrast mri, *arXiv preprint arXiv:1804.04366*.
- [46] D. Jin, Z. Xu, Y. Tang, A. P. Harrison, D. J. Mollura, Ct-realistic lung nodule simulation from 3d conditional generative adversarial networks for robust lung segmentation, *arXiv preprint arXiv:1806.04051*.
- [47] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A. L. Yuille, Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 40 (2018) 834–848.
- [48] A. Tack, A. Mukhopadhyay, S. Zachow, Knee menisci segmentation using convolutional neural networks: data from the osteoarthritis initiative., *Osteoarthritis and cartilage* 26 5 (2018) 680–688.
- [49] Y. Xue, T. Xu, H. Zhang, L. R. Long, X. Huang, Segan: Adversarial network with multi-scale l1 loss for medical image segmentation, *Neuroinformatics* (2018) 1–10.
- [50] M. Rezaei, K. Harmuth, W. Gierke, T. Kellermeier, M. Fischer, H. Yang, C. Meinel, A conditional adversarial network for semantic segmentation of brain tumor, in: *BrainLes@MICCAI*, 2017.
- [51] Z. Li, Y. Wang, J. Yu, Brain tumor segmentation using an adversarial network, in: *International MICCAI Brainlesion Workshop*, Springer, 2017, pp. 123–132.
- [52] P. Moeskops, M. Veta, M. W. Lafarge, K. A. Eppenhof, J. P. Pluim, Adversarial training and dilated convolutions for brain mri segmentation, in: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, Springer, 2017, pp. 56–64.

- [53] K. Kamnitsas, C. Baumgartner, C. Ledig, V. Newcombe, J. Simpson, A. Kane, D. Menon, A. Nori, A. Criminisi, D. Rueckert, et al., Unsupervised domain adaptation in brain lesion segmentation with adversarial networks, in: International Conference on Information Processing in Medical Imaging, Springer, 2017, pp. 597–609.
- [54] M. Zhao, L. Wang, J. Chen, D. Nie, Y. Cong, S. Ahmad, A. S. P. Ho, P. Yuan, S. H. Fung, H. H. Deng, J. J. Xia, D. Shen, Craniomaxillofacial bony structures segmentation from mri with deep-supervision adversarial learning, in: MICCAI, 2018.
- [55] W. Dai, J. Doyle, X. Liang, H. Zhang, N. Dong, Y. Li, E. P. Xing, Scan: Structure correcting adversarial network for organ segmentation in chest x-rays, 2017.
- [56] J. Son, S. J. Park, K.-H. Jung, Retinal vessel segmentation in fundoscopic images with generative adversarial networks, CoRR abs/1706.09318.
- [57] A. Lahiri, K. Ayush, P. K. Biswas, P. Mitra, Generative adversarial learning for reducing manual annotation in semantic segmentation on large scale microscopy images: Automated vessel segmentation in retinal fundus image as test case, 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) (2017) 794–800.
- [58] S. M. Shankaranarayana, K. Ram, K. Mitra, M. Sivaprakasam, Joint optic disc and cup segmentation using fully convolutional and adversarial networks, in: Fetal, Infant and Ophthalmic Medical Image Analysis, Springer, 2017, pp. 168–176.
- [59] Y. Huo, Z. Xu, S. Bao, C. Bermudez, A. J. Plassard, J. Liu, Y. Yao, A. Assad, R. G. Abramson, B. A. Landman, Splenomegaly segmentation using global convolutional kernels and conditional generative adversarial networks, in: Medical Imaging 2018: Image Processing, Vol. 10574, International Society for Optics and Photonics, 2018, p. 1057409.
- [60] C. Peng, X. Zhang, G. Yu, G. Luo, J. Sun, Large kernel matters improve semantic segmentation by global convolutional network, 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2017) 1743–1751.
- [61] D. Yang, D. Xu, S. K. Zhou, B. Georgescu, M. Chen, S. Grbic, D. Metaxas, D. Comaniciu, Automatic liver segmentation using an adversarial image-to-image network, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2017, pp. 507–515.
- [62] B. Kim, J. C. Ye, Cycle-consistent adversarial network with polyphase u-nets for liver lesion segmentation.

- [63] S. K. Sadanandan, J. Karlsson, C. Wählby, Spheroid segmentation using multiscale deep adversarial networks, 2017 IEEE International Conference on Computer Vision Workshops (ICCVW) (2017) 36–41.
- [64] A. Arbelle, T. Riklin-Raviv, Microscopy cell segmentation via adversarial neural networks, 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018) (2018) 645–648.
- [65] Y. Zhang, L. Yang, J. Chen, M. Fredericksen, D. P. Hughes, D. Z. Chen, Deep adversarial networks for biomedical image segmentation utilizing unannotated images, in: MICCAI, 2017.
- [66] H. Chen, X. Qi, L. Yu, P.-A. Heng, Dcan: Deep contour-aware networks for accurate gland segmentation, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016) 2487–2496.
- [67] S. Dong, G. Luo, K. Wang, S. Cao, A. Mercado, O. Shmuelovich, H. Zhang, S. Li, Voxelatlasgan: 3d left ventricle segmentation on echocardiography with atlas guided generation and voxel-to-voxel discrimination, CoRR abs/1806.03619.
- [68] F. Milletari, N. Navab, S.-A. Ahmadi, V-net: Fully convolutional neural networks for volumetric medical image segmentation, 2016 Fourth International Conference on 3D Vision (3DV) (2016) 565–571.
- [69] A. Chartsias, T. Joyce, G. Papanastasiou, S. Semple, M. Williams, D. E. Newby, R. Dharmakumar, S. A. Tsaftaris, Factorised spatial representation learning: Application in semi-supervised myocardial segmentation, in: MICCAI, 2018.
- [70] S. A. T. Thomas Joyce, Agisilaos Chartsias, Deep multi-class segmentation without ground-truth labels, 2018.
- [71] C. Xu, L. Xu, G. Brahm, H. Zhang, S. Li, Mutgan: Simultaneous segmentation and quantification of myocardial infarction without contrast agents via joint adversarial learning, in: MICCAI, 2018.
- [72] A. Sekuboyina, M. Rempfler, J. Kukacka, G. Tetteh, A. Valentinič, J. S. Kirschke, B. H. Menze, Btrfly net: Vertebrae labelling with energy-based adversarial learning of local spine prior, CoRR abs/1804.01307.
- [73] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, A. C. Berg, Ssd: Single shot multibox detector, in: European conference on computer vision, Springer, 2016, pp. 21–37.
- [74] B. Glocker, D. Zikic, E. Konukoglu, D. R. Haynor, A. Criminisi, Vertebrae localization in pathological spine ct via dense classification from sparse annotations, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2013, pp. 262–270.

- [75] S. Yu, H. Dong, G. Yang, G. G. Slabaugh, P. L. Dragotti, X. Ye, F. Liu, S. R. Arridge, J. Keegan, D. N. Firmin, Y. Guo, Deep de-aliasing for fast compressive sensing mri, CoRR abs/1705.07137.
- [76] G. Yang, S. Yu, H. Dong, G. Slabaugh, P. L. Dragotti, X. Ye, F. Liu, S. Arridge, J. Keegan, Y. Guo, et al., Dagan: Deep de-aliasing generative adversarial networks for fast compressed sensing mri reconstruction, IEEE Transactions on Medical Imaging.
- [77] M. Seitzer, G. Yang, J. Schlemper, O. Oktay, T. Würfl, V. Christlein, T. Wong, R. Mohiaddin, D. N. Firmin, J. Keegan, D. Rueckert, A. Maier, Adversarial and perceptual refinement for compressed sensing mri reconstruction, CoRR abs/1806.11216.
- [78] T. Salimans, I. J. Goodfellow, W. Zaremba, V. Cheung, A. Radford, X. Chen, Improved techniques for training gans, in: NIPS, 2016.
- [79] I. Sánchez, V. Vilaplana, Brain mri super-resolution using 3d generative adversarial networks, 2018.
- [80] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. P. Aitken, A. Tejani, J. Totz, Z. Wang, W. Shi, Photo-realistic single image super-resolution using a generative adversarial network, 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2017) 105–114.
- [81] M. Mathieu, C. Couprie, Y. LeCun, Deep multi-scale video prediction beyond mean square error, CoRR abs/1511.05440.
- [82] Y. Chen, F. Shi, A. G. Christodoulou, Z. Zhou, Y. Xie, D. Li, Efficient and accurate mri super-resolution using a generative adversarial network and 3d multi-level densely connected network, CoRR abs/1803.01417.
- [83] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A. L. Yuille, Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs, IEEE Transactions on Pattern Analysis and Machine Intelligence 40 (2018) 834–848.
- [84] Z. Li, Y. Wang, J. Yu, Reconstruction of thin-slice medical images using generative adversarial network, in: International Workshop on Machine Learning in Medical Imaging, Springer, 2017, pp. 325–333.
- [85] M. Mardani, E. Gong, J. Y. Cheng, S. S. Vasanawala, G. Zaharchuk, M. T. Alley, N. Thakur, S. Han, W. J. Dally, J. M. Pauly, L. Xing, Deep generative adversarial networks for compressed sensing automates mri, CoRR abs/1706.00051.
- [86] O. Shitrit, T. Riklin-Raviv, Accelerated magnetic resonance imaging by adversarial neural network, in: DLMIA/ML-CDS@MICCAI, 2017.

- [87] P. Zhang, F. Wang, W. Xu, Y. Li, 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, 2018, pp. 180–188.
- [88] D. Ravi, A. B. Szczotka, D. I. Shakir, S. P. Pereira, T. Vercauteren, Adversarial training with cycle consistency for unsupervised super-resolution in endomicroscopy, 2018.
- [89] L. Han, Z. Yin, A cascaded refinement gan for phase contrast microscopy image super resolution, in: MICCAI, 2018.
- [90] H. Su, Z. Yin, S. Huh, T. Kanade, Cell segmentation in phase contrast microscopy images via semi-supervised classification over optics-related features, Medical image analysis 17 7 (2013) 746–65.
- [91] D. Mahapatra, B. Bozorgtabar, S. Hewavitharanage, R. Garnavi, Image super resolution using generative adversarial networks and local saliency maps for retinal image analysis, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2017, pp. 382–390.
- [92] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. P. Aitken, A. Tejani, J. Totz, Z. Wang, et al., Photo-realistic single image super-resolution using a generative adversarial network., in: CVPR, Vol. 2, 2017, p. 4.
- [93] B. André, T. Vercauteren, A. M. Buchner, M. B. Wallace, N. Ayache, A smart atlas for endomicroscopy using automated video retrieval, Medical image analysis 15 (4) (2011) 460–476.
- [94] X. Chen, E. Konukoglu, Unsupervised detection of lesions in brain mri using constrained adversarial auto-encoders, CoRR abs/1806.04972.
- [95] C. Baur, B. Wiestler, S. Albarqouni, N. Navab, Deep autoencoding models for unsupervised anomaly segmentation in brain mr images, arXiv preprint arXiv:1804.04488.
- [96] C. F. Baumgartner, L. M. Koch, K. C. Tezcan, J. X. Ang, E. Konukoglu, Visual feature attribution using wasserstein gans, CoRR abs/1711.08998.
- [97] R. Shwartz-Ziv, N. Tishby, Opening the black box of deep neural networks via information, CoRR abs/1703.00810.
- [98] S. Kohl, D. Bonekamp, H.-P. Schlemmer, K. Yaqubi, M. Hohenfellner, B. Hadaschik, J.-P. Radtke, K. H. Maier-Hein, Adversarial networks for the detection of aggressive prostate cancer, CoRR abs/1702.08014.
- [99] A. Tuysuzoglu, J. Tan, K. Eissa, A. P. Kiraly, M. Diallo, A. Kamen, Deep adversarial context-aware landmark detection for ultrasound imaging, CoRR abs/1805.10737.



- [100] A. Udrea, G. D. Mitra, Generative adversarial neural networks for pigmented and non-pigmented skin lesions detection in clinical images, in: Control Systems and Computer Science (CSCS), 2017 21st International Conference on, IEEE, 2017, pp. 364–368.
- [101] J. M. Wolterink, T. Leiner, M. A. Viergever, I. Išgum, Generative adversarial networks for noise reduction in low-dose ct, IEEE transactions on medical imaging 36 (12) (2017) 2536–2545.
- [102] J. Wang, Y. Zhao, J. H. Noble, B. M. Dawant, 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, 2018, pp. 3–11.
- [103] Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi, X. Mou, M. K. Kalra, Y. Zhang, L. Sun, G. Wang, Low dose ct image denoising using a generative adversarial network with wasserstein distance and perceptual loss, IEEE Transactions on Medical Imaging.
- [104] X. Yi, P. Babyn, Sharpness-aware low-dose ct denoising using conditional generative adversarial network, Journal of Digital Imaging (2018) 1–15.
- [105] J. Fan, X. Cao, Z. Xue, P.-T. Yap, D. Shen, Adversarial similarity network for evaluating image alignment in deep learning based registration, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2018, pp. 739–746.
- [106] P. Yan, S. Xu, A. R. Rastinehad, B. J. Wood, Adversarial image registration with application for mr and trus image fusion, in: MLMI@MICCAI, 2018.
- [107] Y. Hu, E. Gibson, N. Ghavami, E. Bonmati, C. M. Moore, M. Emberton, T. Vercauteren, J. A. Noble, D. C. Barratt, Adversarial deformation regularization for training image registration neural networks, in: MICCAI, 2018.
- [108] L. Zhang, A. Gooya, A. F. Frangi, 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, 2017, pp. 61–68.
- [109] J. Ren, I. Hacihaliloglu, E. A. Singer, D. J. Foran, X. Qi, Adversarial domain adaptation for classification of prostate histopathology whole-slide images, in: MICCAI, 2018.