

Generatywne sieci współzawodniczące/ przeciwstawne w obrazowaniu medycznym

Weronika Hryniewska

Generative Adversarial Network in Medical Imaging: A Review

Xin Yi^{a,*}, Ekta Walia^{a,b}, Paul Babyn^a

^a*Department of Medical Imaging, University of Saskatchewan, 103 Hospital Dr, Saskatoon, SK, S7N 0W8 Canada*

^b*Philips Canada, 281 Hillmount Road, Markham, Ontario, ON L6C 2S3, Canada*

Abstract

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

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

Generatywne sieci współzawodniczące

- Adaptacja domeny, rozszerzenie danych, translacja obrazu na obraz
- GAN-y są specjalnym rodzajem modelu sieci neuronowych, w którym trenowane są jednocześnie dwie sieci, z których jedna skupia się na generowaniu obrazu, a druga na dyskryminacji.
- koncepcja treningu współzawodniczącego/przeciwstawnego (adversarial) - rok 2014
- Zastosowania m. in. w dziedzinach radiologii, histopatologii i dermatologii

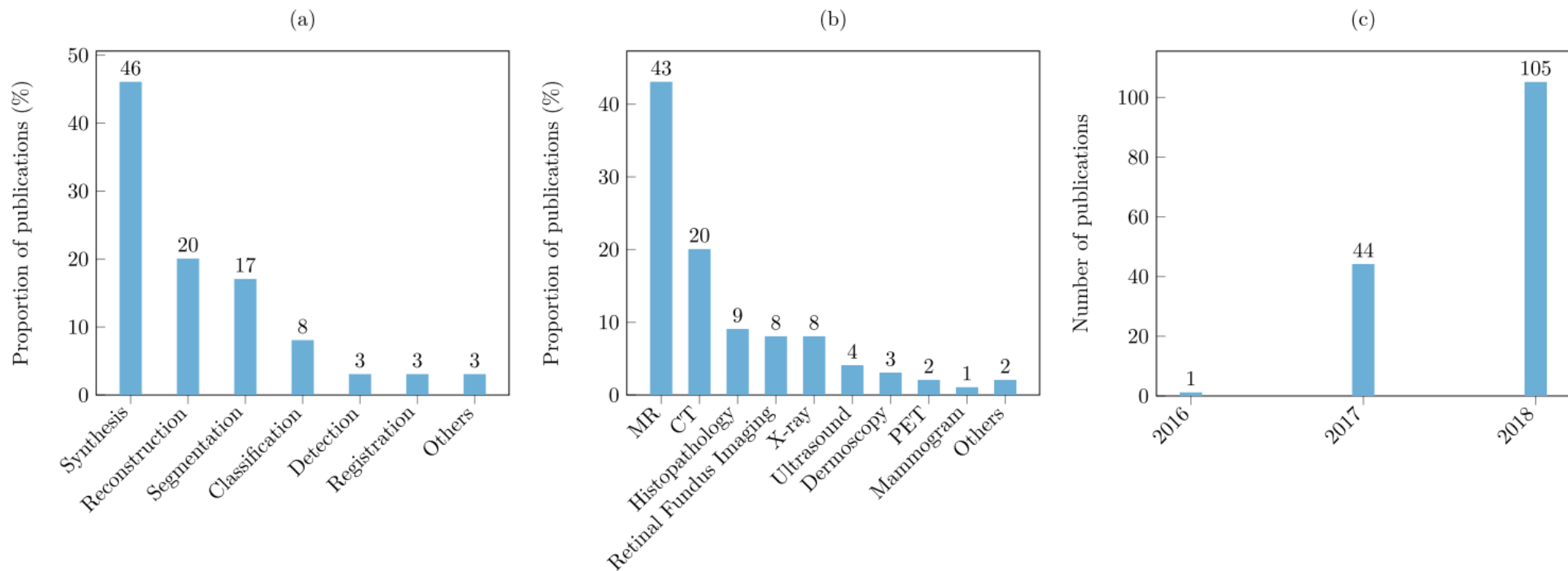


Fig. 1. (a) Categorization of GAN related papers according to canonical tasks. (b) Categorization of GAN related papers according to imaging modality. (c) Number of GAN related papers published from 2014. Note that some works performed various tasks and conducted evaluation on datasets with different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

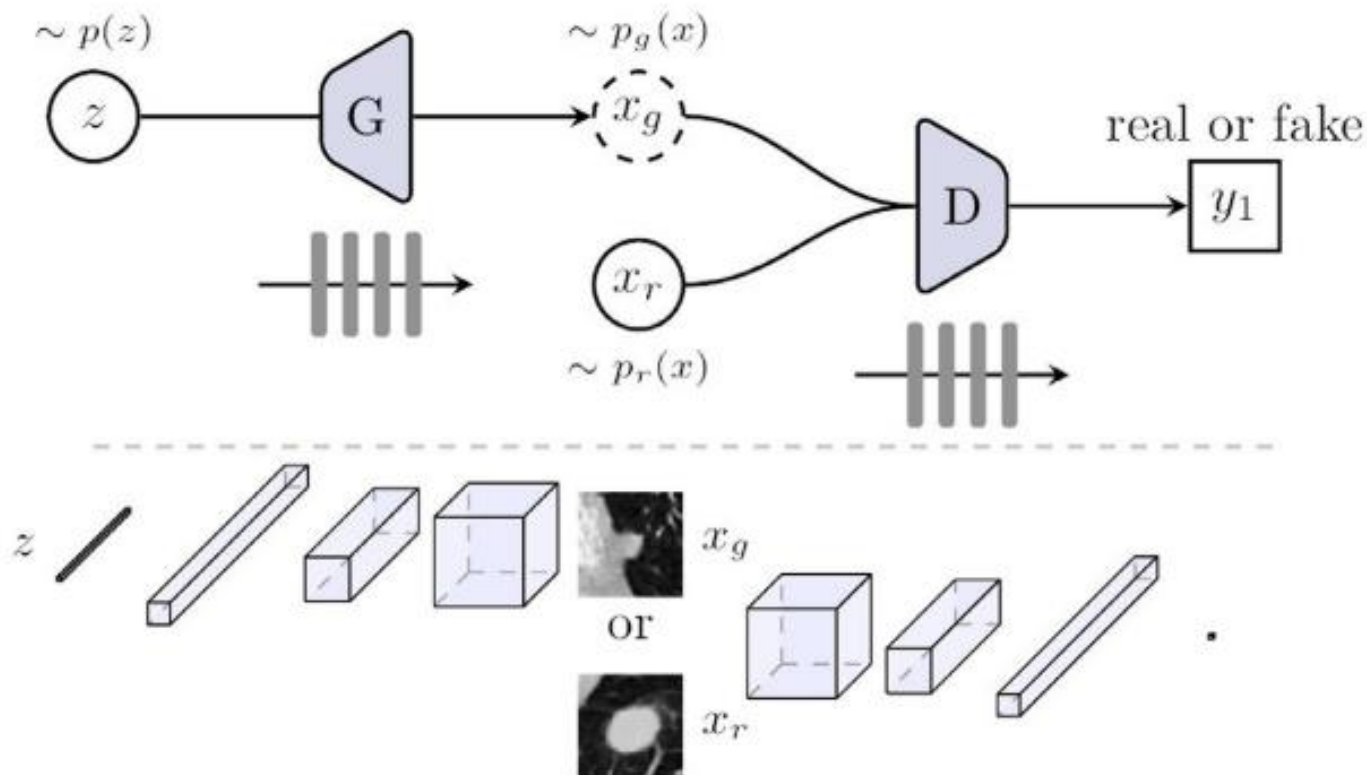
Plan artykułu

- O rodzajach sieci neuronowych
- Zastosowania w obrazowaniu medycznym
 - Rekonstrukcja
 - Synteza obrazów medycznych
 - Segmentacja
 - Klasyfikacja
 - Detekcja
 - Rejestracja
- Dyskusja - przyszłe wyzwania i zastosowania

Vanilla GAN

$$\mathcal{L}_D^{GAN} = \max_D \mathbb{E}_{x_r \sim p_r(x)} [\log D(x_r)] + \mathbb{E}_{x_g \sim p_g(x)} [\log(1 - D(x_g))],$$

$$\mathcal{L}_G^{GAN} = \min_G \mathbb{E}_{x_g \sim p_g(x)} [\log(1 - D(x_g))].$$



$$z \in \mathbb{R}^{n \times 1 \times 1}$$

$$x_g, x_r \in \mathbb{R}^{c \times w \times h}$$

$$y_1 \in \{0, 1\}$$

Problemy

- Brak gwarancji równowagi między treningiem G i D
 - W większości przypadków D jest mocniejsza.
 - Gradienty z D zbliżają się do zera, nie dając żadnych wskazówek dla dalszego treningu G
- Generowanie obrazów o wysokiej rozdzielczości
- Mode collapse

Dyskryminator

- Funkcje straty:
 - f-divergence (f-GAN),
 - least-square (LSGAN),
 - hinge loss,
 - **Wasserstein distance** (WGAN, WGAN-GP)

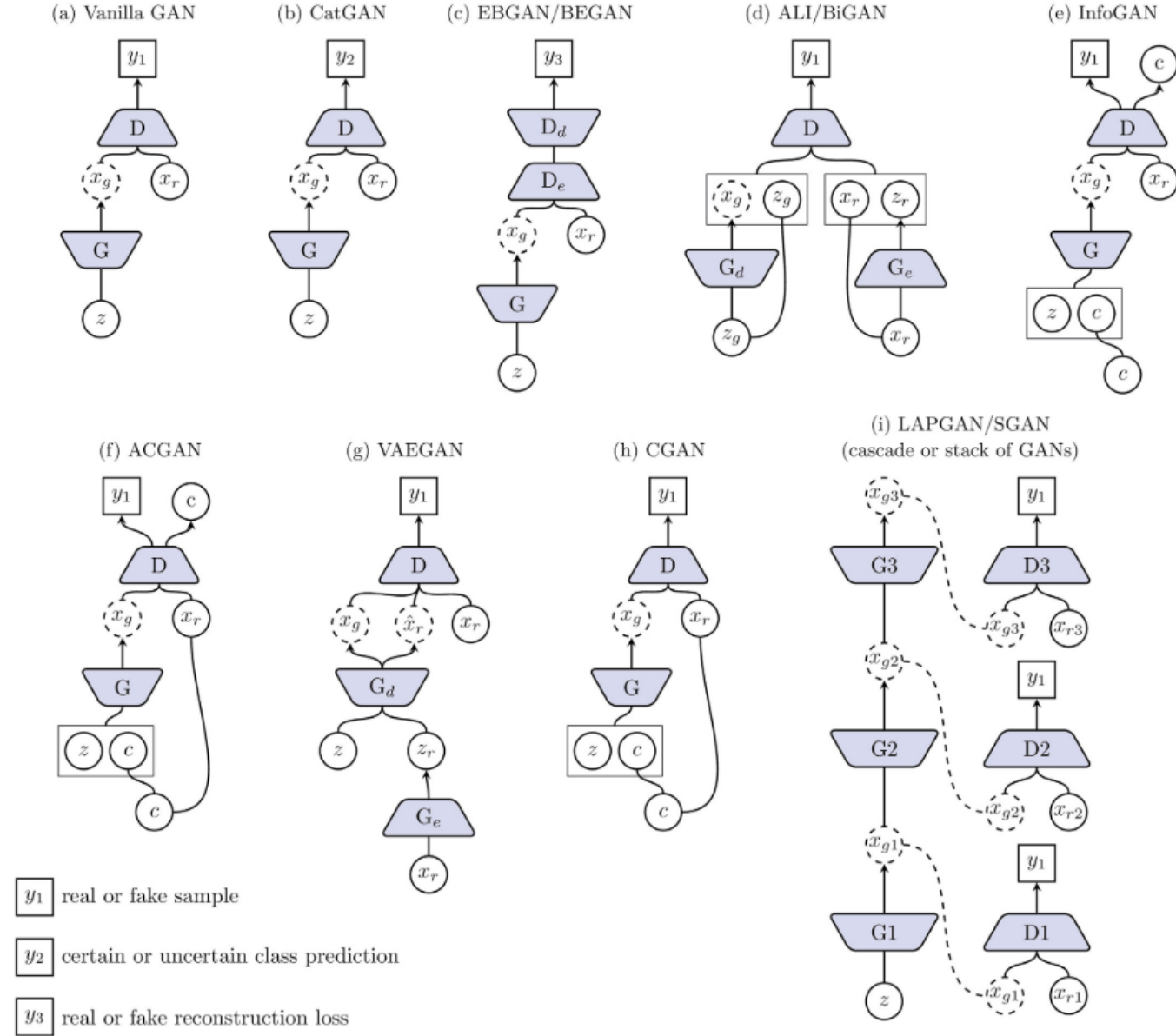


Fig. 3. A schematic view of variants of GAN. c represents the conditional vector. In CGAN and ACGAN, c is the discrete categorical code (e.g. one hot vector) that encodes class labels and in InfoGAN it can also be continuous code that encodes attributes. x_g generally refers to the generated image but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

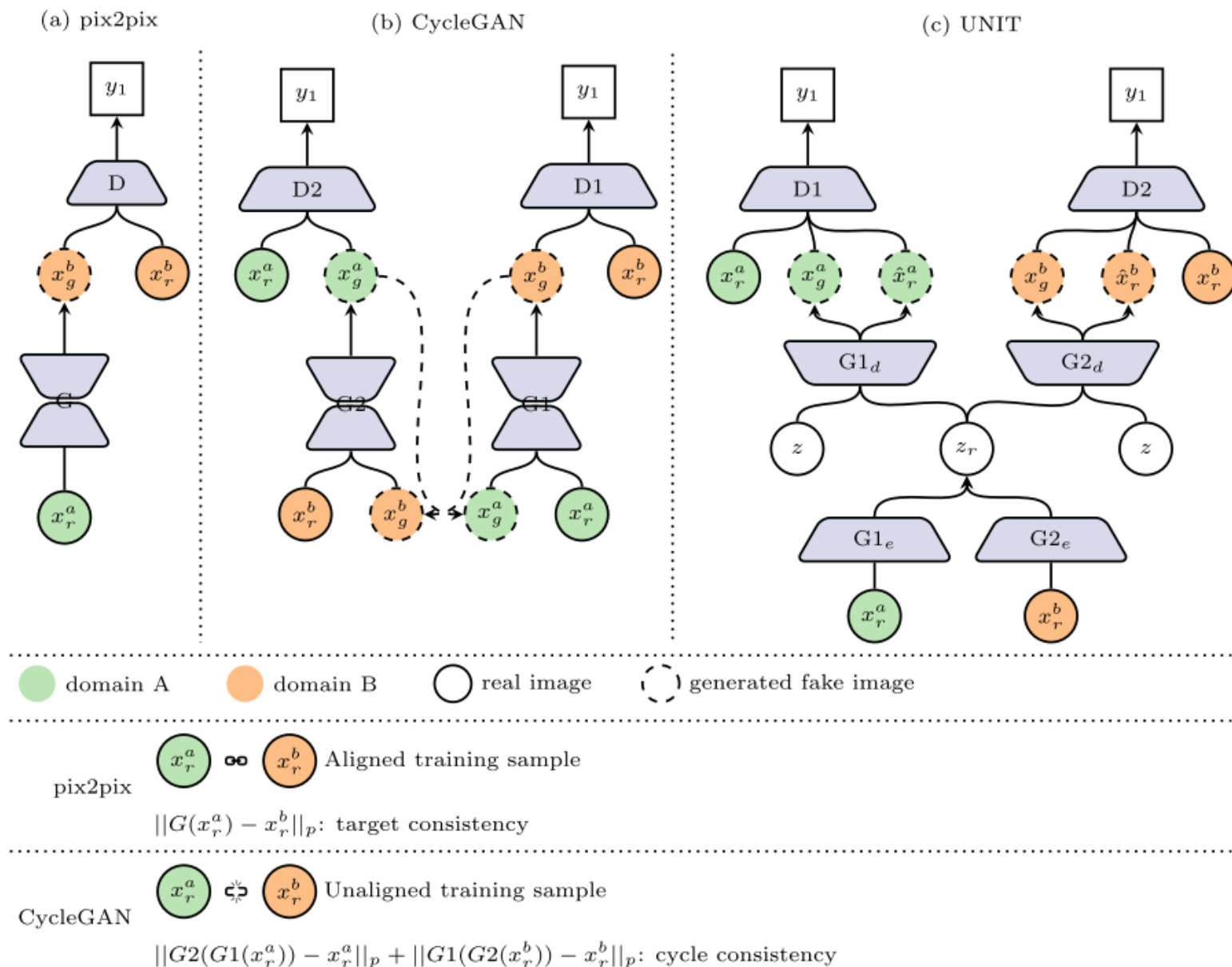


Fig. 4. cGAN frameworks for image-to-image translation. pix2pix requires aligned training data whereas this constraint is relaxed in CycleGAN but usually suffers from performance loss. Note that in (a), we chose reconstruction loss as an example of target consistency. This supervision is task related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

Architektura

- Vanilla GAN – full connected layers
- DCGAN - fully convolutional downsampling/upsampling layers
- Batch normalization (G)
- LeakyRelu (G)
- Spectral normalization (D)
- Residual connection(G,D)

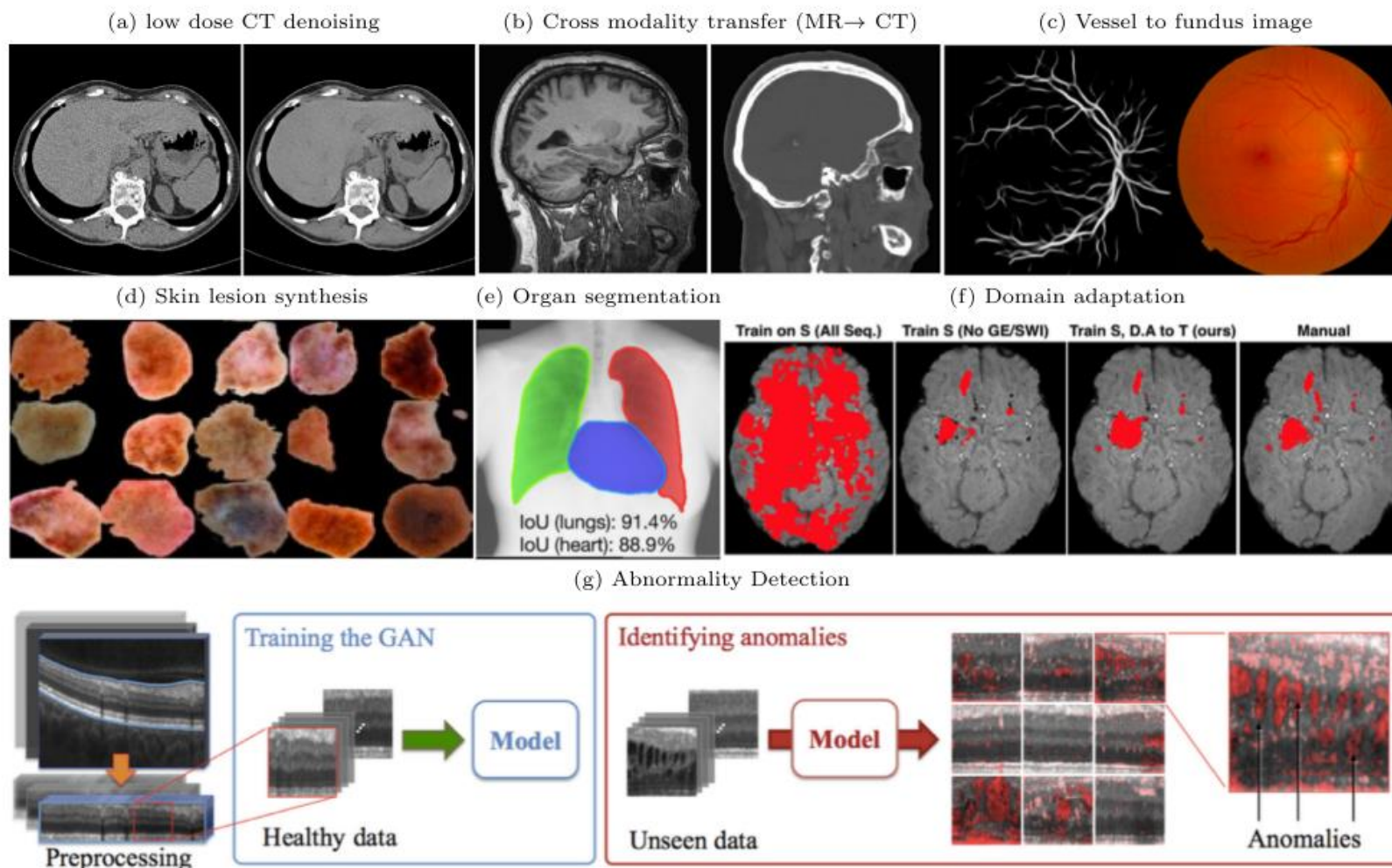


Fig. 5. Example applications using GANs. Figures are directly cropped from the corresponding papers. (a) Left side shows the noise contaminated low dose CT and right side shows the denoised CT that well preserved the low contrast regions in the liver (Yi and Babyn, 2018). (b) Left side shows the MR image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

Rekonstrukcja

- Zamienianie w obraz wyjściowy surowych danych z sensorów
- Redukcja szumu w post-processingu
- pix2pix framework

Synteza obrazów medycznych

- Augmentacja danych
- Rodzaje
 - Bezwarunkowa
 - z szumu (DCGAN, WGAN, PGGAN)
 - 3 DCGANy do generowania trzech rodzajów zmian w wątrobie (torbiele, przerzuty i naczyniaki)
 - Pomędzy modalnościami
 - pix2pix-based frameworks - różne modalności obrazu mogą być współrejestrowane w celu zapewnienia wierności danych
 - UNIT and CycleGAN-based frameworks – w bardziej generalnych zadaniach
 - shape consistency loss
 - Inne warunkowe
 - Synteza zdjęć z nietypowymi wymaganiami (guz płucny dotykający granicy płuc)

Table 4

Unconditional medical image synthesis publications. A brief description of the quantitative measures and datasets can be found in [Tables 3](#) and [7](#).

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

Segmentacja

- Funkcja straty: pixel-wise or voxel-wise loss
- Problemem regiony o niskim kontraście
- Dyskryminator może być uznany jako regulator kształtu
 - Efekt najlepszy dla kształtów zwartych (płuca, serce)

Klasyfikacja

- Generator i dyskryminator do ekstrakcji cech; dyskryminator jako klasyfikator
- Półnadzorowany GAN może osiągnąć wydajność porównywalną z tradycyjną nadzorowaną CNN z rzędem wielkości mniej oznakowanych danych.
- Adversarial loss może zmniejszyć nadmierne dopasowanie do domeny
- Dwuetapowego proces trenowania:
 - 1. powiększanie obrazów 2. klasyfikacja
 - Zalety: zastępowalność komponentów, Wady: generowanie musi być przeprowadzone dla każdej klasy oddzielnie
- Obrazy generowane z GAN mogą służyć jako skuteczna augmentacja w reżimie średnich danych, ale mogą nie być pomocne w reżimie wysokich lub niskich danych.

Detekcja

- Dyskryminator GAN może być wykorzystany do wykrywania nieprawidłowości.
- Adversarial auto-encoder do nauki zdrowych obrazów. Potem obraz zmiany chorobowej został odwzorowany na obraz bez zmiany chorobowej. Zmiana chorobowa wyróżniona poprzez obliczenie pozostałości tych dwóch obrazów.
- Usuwanie patologii – niesparowany CycleGAN.
- Jeśli obrazy w tej samej domenie, to można wykrywać nieprawidłowości.

Inne wykorzystania

- Modelowanie rozkładu ruchu specyficznego dla pacjenta na podstawie pojedynczego obrazu przedoperacyjnego.
- Wyróżnianie regionów najbardziej odpowiedzialnych za chorobę.
- Rekoloryzacja endoskopowych danych wideo.
- Planowanie leczenia w radioterapii – przewidywanie mapy rozkładu dawki na podstawie obrazu CT.
- Modelowanie progresji choroby Alzheimera na podstawie MRI.

Dyskusja wyników przeglądu

- Około 46% z tych prac dotyczyło syntezy obrazów, przy czym synteza obrazów w różnych modalnościach jest najważniejszym zastosowaniem GANs.
- MR jest najczęstszą modalnością obrazowania badaną w literaturze związanej z GAN.
- Kolejne 37% tych badań należy do grupy rekonstrukcji i segmentacji ze względu na popularność frameworków translacji obraz-obraz.
- Kolejne 8% z tych badań dotyczy klasyfikacji.

Przyszłość GAN w medycynie

- Lepsza ocena niż MAR, PSNR, SSIM
 - Miary te jednak nie odpowiadają wizualnej jakości obrazu. - rozmyty rezultat
- Ostrzeżenie przed wykorzystywaniem wygenerowanych obrazów do bezpośredniej interpretacji przez lekarzy.
- Odtworzenie obrazów MR pozyskanych z pewnymi artefaktami, takimi jak ruch może pomóc zredukować liczbę powtarzanych badań.
- Półautomatyczne generowanie raportów - podpisywanie obrazów.
- Wykrywanie wszczepionych urządzeń na zdjęciach rentgenowskich.
- Synteza obrazów medycznych z opisów tekstowych – dla rzadkich przypadków.
- Przewidywanie progresji choroby i wpływu leku.