Overcoming Medical Data Scarcity with GAN Generated Synthetic Images

1st Sakir Bingol
Electrical Electronics Engineering
Marmara University
Istanbul, Turkey
sakir.bingol@marmara.edu.tr

2nd Omer Sabri Emeksiz
Electrical Electronics Engineering
Marmara University
Istanbul, Turkey
omeremeksiz@marun.edu.tr

3rd Kaan Yalcin
Electrical Electronics Engineering
Marmara University
Istanbul, Turkey
kaanyalcin@marun.edu.tr

4th Furkan Ucak
Electrical Electronics Engineering
Marmara University
Istanbul, Turkey
furkanucak@marun.edu.tr

5th Bilhan Ilbey Contar
Electrical Electronics Engineering
Marmara University
Istanbul, Turkey
ilbeycontar@marun.edu.tr

6th Sude Bulut
Electrical Electronics Engineering
Marmara University
Istanbul, Turkey
sudebulut@marun.edu.tr

***Abstract*—This article underscores the pivotal role of Generative Adversarial Networks (GANs) in addressing data scarcity challenges in medical imaging. By utilizing diverse skin lesion images, our approach trying to optimize artificial intelligence models through the generation of synthetic medical images. Focused on the HAM10000 dataset, our study contributes to refining diagnostic tools, aiding medical professionals in skin disease diagnosis. We evaluate our approach with a skin detection algorithm, demonstrating GANs' effectiveness in generating realistic synthetic medical images.**

***Index Terms*—AI in Healthcare, cGANs, GANs, Medical Imaging, Synthetic Images**

I. Introduction

Medical imaging is a cornerstone of modern healthcare, offering invaluable insights into the human body's structures and functions. Medical images are classified and segmented by experts to diagnose disorders, and in some situations anticipate how diseases will progress [1]. A wide variety of methods, such as plain radiography, computed tomography (CT), magnetic resonance imaging (MRI) and positron emission tomography (PET), are used to obtain medical images that are vital for definitive diagnosis, treatment planning and overall patient care [2]. However, the manual investigation of fine patterns by experts, such as pathologists, increases the risk of errors and workload [3] [4].

The renaissance of deep learning in computer vision since 2012, as pioneered by Krizhevsky *et* al., has introduced in novel possibilities within the deep learning framework to address challenges and errors in disease diagnosis. Estimates indicate that more than 400 papers were published in prominent conference venues and journals related to medical imaging between 2016 and 2017 (Litjens *et* al., 2017). The widespread embrace of deep learning within the medical imaging community can be attributed to its proven capacity to enhance image interpretation and reinforce image representation and classification [5]. However the medical field is one of the areas where obtaining a huge number of images to train a deep learning classifier is extremely challenging and practically inefficient [6] [7]. This is because, in general, medical images are insufficient, costly, and of limited use due to legal issues (patient confidentiality). Additionally, generally available medical image datasets are inconsistent in terms of size and markup. This makes it less useful for training data-hungry artificial neural networks. This limits the development of direct medical diagnostic systems. Therefore, creating synthetic images and presenting them together with labeled images will aid medical image analysis and provide better diagnostic systems [8].

In this study, we address a critical problem using Generative Adversarial Networks (GANs), recognized for realistic medical imaging since Goodfellow *et* al. introduced them in 2014 [9]. GANs, rooted in game theory, involve a generator and discriminator, engaging in adversarial training without relying on approximate inference or Markov chains [10] [11]. GANs have gained attention in computer vision, inspiring various efficient generative modeling variations, particularly for high-quality, photorealistic natural images [12] [13].

Recent advances in medical imaging integrate GANs into various applications [14] [15]. These studies predominantly employ the image-to-image GAN technique to facilitate label-to-segmentation translations, segmentation-to-image transformations, and cross-modal medical image translations. For example, Costa *et* al. [14] harnessed a fully-convolutional network to train on retinal vessel segmentation images, subsequently extending its capacity to translate binary vessel tree maps into new retinal images. Dai and colleagues [16] focused on training GANs to generate segmentation images of lung fields and the heart using chest X-ray input. Xue *et* al. [17] introduced a dual-network GAN configuration, comprising a 'Segmentor' and a 'Critic,' to excel in translating between brain MRI images and binary brain tumor segmentation maps. In the work by Nie *et* al. [18], a patch-based GAN was

skillfully tailored for the translation between brain CT images and their corresponding MRI counterparts [19]. In a synthetic data generation study using StyleGAN and the ACDC dataset, four cardiology experts with over 15 years of experience were tasked with visually distinguishing between 100 cine MRI images 50% real and 50% synthetic. The study aimed to assess the realism of GAN-generated images. Experts achieved a 60% ($\pm$10%) accuracy in visually categorizing the images, highlighting the visual success of the synthetic images [20].

This research introduces a novel method to address medical data scarcity by leveraging GANs for creating synthetic medical images. Our primary objective is to train a GAN model to generate realistic medical images to optimizing the performance of AI and machine learning models for disease diagnosis. Our research contributes to the field of medical imaging and disease diagnosis by improving tools to assist medical professionals in their diagnosis and treatment planning.

II. System Model

A. The Principle of GANs

GANs, inspired by game theory, involve a strategic competition between a generator and a discriminator to achieve a Nash equilibrium during training. The architectural framework of GANs is depicted in Figure 1. In simple terms, the generator G's role is to create synthetic data that closely mimics the potential distribution of real data. It achieves this by using an input of an arbitrary noise vector, usually taken from an average or standard distribution. The generator G then converts this noise into a new data space, creating a synthetic sample that is represented by the multi-dimensional vector G(z). Conversely, discriminator D functions as a type of binary classifier. It assesses both authentic samples taken from the dataset and synthetic samples produced by the generator G. The likelihood that a certain sample is real rather than artificial is shown by the discriminator D's output. The training process aims for a scenario where the discriminator cannot reliably distinguish between data from the actual dataset and the generator, signaling an optimal state. This equilibrium is reached, and the result is a trained generator model G. This model has demonstrated the efficacy of the GAN framework in producing realistic synthetic samples by successfully learning the distribution of real data [21].

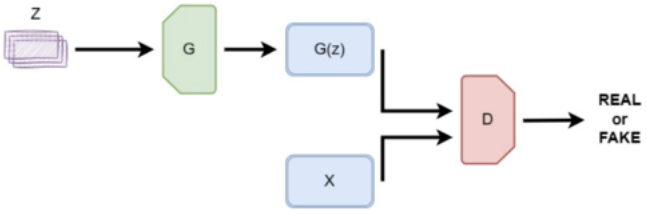

Fig. 1. The architecture of GANs [21].

Since Godfellow's initial GAN work, numerous GAN models have been created. The various GANs that have been created can be divided into two primary categories: GANs based on Objective Function Optimization and GANs based on Architecture Optimization. Within these classes, there are various models that branch out. The most commonly used of these models are listed in the Table I. There are many other GAN models such as DE-GANs (Data-Efficient GANs) [22], StyleGAN [23], HingeGAN (Geometric GAN) [24], which we do not include in the table [21].

TABLE I
Classification of GAN Models

Objective Function Optimized Based GANs	**Architecture Optimization Based GANs**		
	Convolution based GANs	***Condition based GANs***	***Autoencoder based GANs***
unrolled GAN [25]; f-GAN [26]; Mode-Regularized GAN [27]; Least-Square GAN [28]; Loss-Senstive GAN [29]; EBGAN [30]; WGAN [31]; WGAN-GP [32]; WGAN-LP [33];	DCGAN [34]	CGANs [35]; InfoGAN [36]; ACGAN [37]	AAE [38]; BiGAN [39]; ALI [40]; AGE [41]; VAEGAN [42]

B. Generating Synthetic Medical Images with cGANs

Advanced GAN models known as Conditional Generative Adversarial Networks (cGANs) use Convolutional Neural Networks (CNN) and other Deep Learning (DL) techniques. In a typical GAN, the generator generates data with no particular control over its properties. In contrast, the generator in cGAN is given random noise (z) together with some pre-existing data (c). In addition, the discriminator receives the matching actual or fake data as well as the prior knowledge c [43]. The mathematical representation of the cGAN principle is as follows (1) [44]:

$$\min_G \max_D V(D,G) = \mathbb{E}_x \log(D(x|c)) + \mathbb{E}_z[1 - \log(D(G(z|c)))] \tag{1}$$

Also, the general structure of the cGAN model is shown in the image in Figure 2. In Figure 2, compared to Figure 1, the participation of vector c in the model is clearly observed.

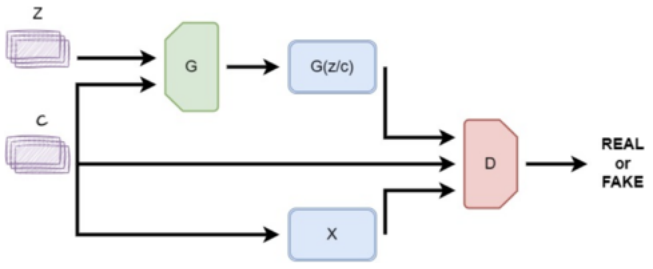

Fig. 2. The architecture of cGANs [21].

C. Methodology and Dataset Overview

Skin cancer, notably melanoma, causes over 7,000 annual deaths in the U.S., highlighting the critical need for early detection to enhance patient survival rates [45] [46]. Visual diagnosis, often initiated by patients, is vital for identifying skin lesions, but dermatologists face challenges with only 60% diagnostic accuracy due to intricate appearances [47]–[50]. In today's technological age, research focuses on overcoming these challenges, primarily using artificial intelligence. The role of AI and deep learning models in skin cancer diagnosis is crucial, contributing significantly to early identification and treatment. Despite the increasing availability of internet data for training, a challenge arises from insufficiently labeled data [51]. The HAM10000 dataset for skin cancer images has proven to be very useful. It has played a key role in creating synthetic data for various lesion categories, including melanocytic Nevi (nv), basal cell carcinoma (bcc), melanoma (mel) and benign keratosis like (bkl), as shown in Figure 3.

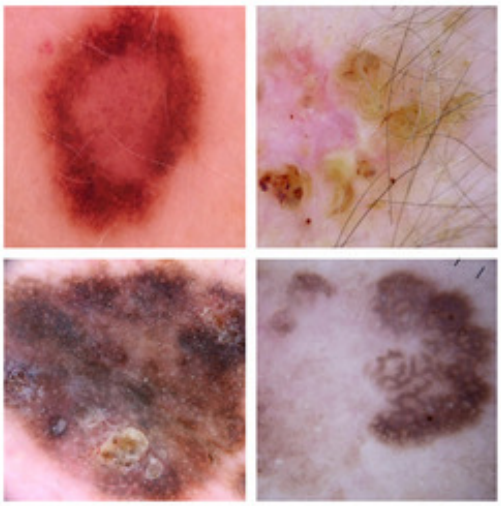

Fig. 3. HAM10000 original skin lesion example samples [52].

The HAM10000 dataset, collected and organized by Harvard Dataverse, ViDIR Dataverse (Medical University of Vienna) through The International Skin Imaging Collaboration (ISIC) from 2018 Challenge, serves as a diverse and comprehensive resource for training deep learning and AI models [52] [53]. Dermatologists meticulously scrutinized and categorized the images, ensuring their accuracy. This dataset includes pigmented lesions from diverse populations, offering various examples such as melanoma, basal cell carcinoma, benign keratosis like lesions, dermatofibroma, actinic keratosis, melanocytic nevi, and vascular skin lesions. The high quality images, accompanied by meticulous annotations specifying the lesion type, make HAM10000 invaluable for advancing diagnostic capabilities, but the images in the dataset do not have a standard format, numerical classification, and labels [54]. Each dermatoscopic image in the dataset, which can serve as a training set for academic machine learning purposes, was carefully labeled and classified as shown in Table 2. The analysis reveals a predominant presence of images consistent with the diagnosis of melanocytic nevus (nv). Addressing the scarcity of labeled data on skin diseases, our study leverages the HAM10000 dataset to generate new images, aiming to overcome limitations in data and streamline the skin cancer diagnosis process.

TABLE II
HAM10000 SKIN LESION TYPE CLASSIFICATION TABLE

Skin Lesion Type	Number of Images
Melanocytic Nevi (nv)	6705
Basal Cell Carcinoma (bcc)	514
Melanoma (mel)	1113
Benign Keratosis Like (bkl)	1099
Actinic Keratosis (akiec)	327
Vascular (vasc)	142
Dermatofibroma (df)	115

Implemented in Google Colab, our project features a Conditional Deep Convolutional GAN (cDCGAN) using PyTorch to synthesize medical images from the HAM10000 dataset. The Generator, driven by the Generator class, processes Gaussian-distributed noise and one-hot encoded class labels, progressively upscaling through transposed convolutional layers. Batch normalization (momentum=0.9) and Leaky ReLU activations (slope=0.2) are applied at each layer. The generator operates in a latent noise space with 100 units and produces outputs with 7 channels. The Discriminator, implemented by the Discriminator class, distinguishes real and synthetic images through convolutional layers, utilizing batch normalization (momentum=0.9) and Leaky ReLU activations (slope=0.2). Adversarial training spans 100 epochs, leveraging an Adam optimizer (learning rate=0.0002) and binary cross-entropy loss criterion. GPU acceleration expedites training. Techniques such as spectral normalization (one power iteration) and gradient penalty (coefficient=10) enhance stability. Model checkpointing enables the continuation of training or the evaluation of the trained model on new data. After the generation process, the realistic medical images are effectively plotted to visually assess the model's performance, providing a holistic understanding of the specific numerical hyperparameters and network configurations.

III. NUMERICAL RESULTS

In our comprehensive skin cancer detection system, we evaluating the performance of generated images using our Convolutional Neural Network (CNN) based detection model. Leveraging CNN consisting of three convolutional blocks, batch normalization, and dropout, the algorithm is trained on the HAM10000 dataset. To counter class imbalance, oversampling techniques are implemented. Testing with authentic medical images from HAM10000 further substantiates its robust performance, attaining a noteworthy success accuracy of 98%. We use the model, which has achieved such high success in skin lesions taken from the HAM10000 data set, to

evaluate the realistic performance of the generated fake lesions. Knowing that the generated images are highly acc will show that these images are suitable for creating a da and can be used to train artificial intelligence models.

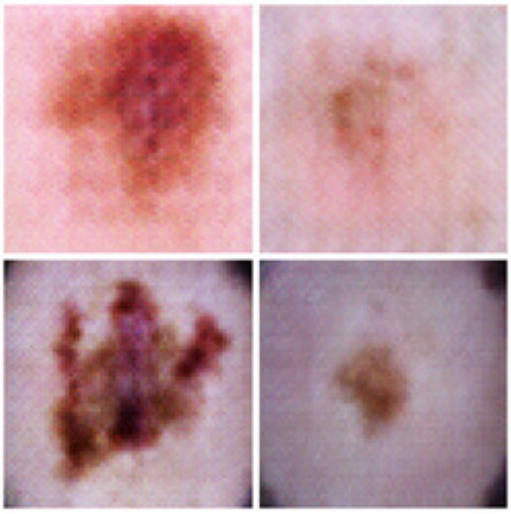

Fig. 4. Generated skin lesion example samples

Our research focused on examining the four most common types of skin lesions within the HAM10000 dataset, as detailed in Table 2. In Figure 4, melanocytic nevi (nv), basal cell carcinoma (bcc), melanoma (mel) and benign keratosis like (bkl) are shown as examples of those that were correctly detected using the detection model among a total of 400 fake skin lesions generated, respectively.

TABLE III
GENERATED SKIN LESION DETECTION RESULTS

Skin Lesion Type	Number of Generated Images	Successful Detection (%)	Average Accuracy (%)
Melanocytic Nevi (nv)	100	%63.41	%88.34
Basal Cell Carcinoma (bcc)	100	%29.73	%99.90
Melanoma (mel)	100	%11.63	%86.20
Benign Keratosis Like (bkl)	100	%16.13	%80.21

We generate 100 synthetic skin lesion images for each of these four diseases and tested the performance of each of these images through the detection algorithm. We generated constant number of images for each disease to examine how the bias among the images within the dataset, as listed in Table 2, influenced the realism performance of the produced images. The "successful detection" column in Table 3 illustrates the percentage of correctly classified images among those generated by the detection model. As indicated in Table 3, images with a higher count in Table 2 exhibit a higher

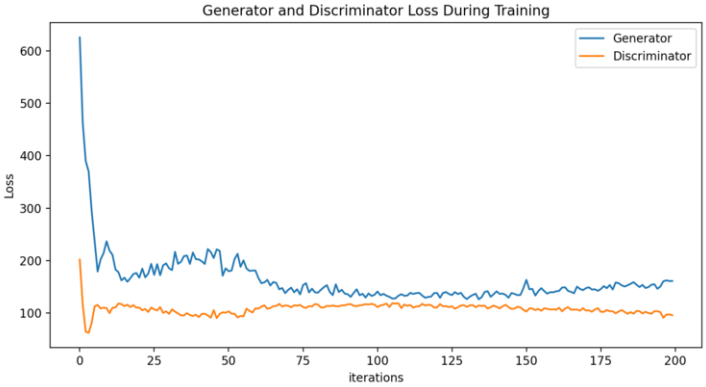

Fig. 5. Training loss of generator and discriminator

success detection rate compared to those with a lower count. The final column in Table 3, labeled "average accuracy," indicates the mean percentage of correct predictions made by the detection model for each image. Establishing a linear or comparable relationship between the average accuracy values of the generated synthetic images is not feasible. The synthesis process involves randomly choosing real images from the dataset. Consequently, the selected images for the detection model may exhibit visual patterns for which the model lacks sufficient training, or they could come from images that pose challenges in differentiation. Hence, despite the high average accuracy values, it is not viable to establish a correlation with successful detection values.

Analyzing the loss graph depicted in Figure 5 reveals that the generator's performance does not enhance as the training progresses towards completion. Enhancing the model through optimization can lead to further improvements in the generator's simplicity.

IV. CONCLUSION AND FUTURE WORK

In the end, our CNN-based model showed enough accuracy to be used for training AI models. It is proving that GANs have the potential to be used to create AI dermatology solutions that will be useful to dermatology experts and patients alike. According to the detecting algorithm's output, we were able to obtain highly accurate photos. After creating 100 photos for every type of disease, we found that the skin condition melanocytic nevi had the greatest success rate in terms of detection.

In light of the results obtained, utilizing a larger number of epochs due to hardware constraints might result in a more accurate model. Additionally, considering training the model with different dataset inputs may lead to clearer outputs. On the other hand, alternative performance analysis methods may be contemplated instead of the detection algorithm used for testing the results. Finally, fine-tuning the model can be conducted, potentially leading to more accurate outcomes.

REFERENCES

[1] Makhlouf, A., Maayah, M., Abughanam, N. et al. "The use of generative adversarial networks in medical image augmentation", Neural Comput & Applic (2023), https://doi.org/10.1007/s00521-023-09100-z.

[2] Karim Armanious , Chenming Jianga, Marc Fischer , Thomas Küstner , Tobias Heppb, Konstantin Nikolaoub, Sergios Gatidis , Bin Yanga, "Computerized Medical Imaging and Graphics,".

[3] Cooper, L.A.; Kong, J.; Gutman, D.A.; Dunn, W.D.; Nalisnik, M.; Brat, D.J. Novel genotype-phenotype associations in human cancers enabled by advanced molecular platforms and computational analysis of whole slide images. Lab. Investig. A J. Tech. Methods Pathol. 2015, 95, 366–376. [CrossRef] [PubMed].

[4] Alturkistani,H.A.;Tashkandi,F.M.;Mohammedsaleh,Z.M.Histological Stains: A Literature Review and Case Study. Glob. J. Health Sci. 2015, 8, 72–79. [CrossRef].

[5] Yi, X., Walia, E., & Babyn, " Generative adversarial network in medical imaging: A review", Medical Image Analysis, vol. 58, 101552, December 2019.

[6] M. Loey, F. Smarandache, and N. E. M. Khalifa, "Within the Lack of Chest COVID-19 X-ray Dataset: A Novel Detection Model Based on GAN and Deep Transfer Learning," Symmetry, vol. 12, no. 4, p. 651, Apr. 2020, doi: 10.3390/sym12040651.

[7] Maayan Frid-Adar, Idit Diamant, Eyal Klang, Michal Amitai, Jacob Goldberger, Hayit Greenspan, GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification, Neurocomputing, Volume 321, 2018, Pages 321-331, ISSN 0925-2312, https://doi.org/10.1016/j.neucom.2018.09.013.

[8] Talha Iqbal, Hazrat Ali, Generative Adversarial Network for Medical Images (MI-GAN), Journal of Medical Systems volume 42, Article number: 231 (2018).

[9] I. Goodfellow et al., "Generative adversarial nets," in Proc. Adv. Neural Inf. Process. Syst., 2014, pp. 2672–2680.

[10] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," Nature, vol. 323, pp. 533–536, Oct. 1986.

[11] Zhaoqing Pan, Weijie Yu, Xiaokai Yi, Asifullah Khan, Feng Yuan, Yuhui Zheng, "Recent Progress on Generative Adversarial Networks (GANs): A Survey", IEEE Access, 10.1109/ACCESS.2019.2905015, Apr. 2019.

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

[13] A. Odena, C. Olah, J. Shlens," Conditional image synthesis with auxiliary classifier gans", arXiv:1610.09585, 2016.

[14] P. Costa, A. Galdran, M.I. Meyer, M.D. Abràmoff, M. Niemeijer, A.M. Mendonça, A. Campilho, "Towards adversarial retinal image synthesis", arXiv:1701.08974, 2017.

[15] 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", arXiv: 1707.09585, 2017.

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

[17] Y. Xue, T. Xu, H. Zhang, R. Long, X. Huang, "Segan: adversarial network with multi-scale l_1 loss for medical image segmentation", arXiv:1706.01805, 2017.

[18] D. Nie, R. Trullo, C. Petitjean, S. Ruan, D. Shen, "Medical image synthesis with context-aware generative adversarial networks", arXiv:1612.05362 , 2016.

[19] Maayan Frid-Adar, Idit Diamant , Eyal Klang , Michal Amitai , Jacob Goldberger , Hayit Greenspana, "GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification", Neurocomputing , vol 321, pp. 321-331, December 2018.

[20] Y. Skandarani, P.-M. Jodoin, and A. Lalande, "GANs for Medical Image Synthesis: An Empirical Study," Journal of Imaging, vol. 9, no. 3, p. 69, Mar. 2023, doi: 10.3390/jimaging9030069.

[21] Z. Pan, W. Yu, X. Yi, A. Khan, F. Yuan and Y. Zheng, "Recent Progress on Generative Adversarial Networks (GANs): A Survey," in IEEE Access, vol. 7, pp. 36322-36333, 2019, doi: 10.1109/ACCESS.2019.2905015.

[22] Ziqiang Li, Beihao Xia, Jing Zhang, Chaoyue Wang, Bin Li, "A Comprehensive Survey on Data-Efficient GANs in Image Generation", University of Science and Technology of China, Huazhong University of Science and Technology, The University of Sydney, JD Explore Academy, (8 Oct 2022).

[23] Karras, T.; Laine, S.; Aittala, M.; Hellsten, J.; Lehtinen, J.; Aila, T. Analyzing and Improving the Image Quality of StyleGAN. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 13–19 June 2020.

[24] 21. Lim, J.H.; Ye, J.C. Geometric GAN. arXiv 2017, arXiv:1705.02894.

[25] L. Metz, B. Poole, D. Pfau, and J. Sohl-Dickstein, "Unrolled generativeadversarial networks," in Proc. Int. Conf. Learn. Represent., 2017, pp. 1–25. [Online]. Available: https://arxiv.org/abs/1611.02163

[26] S. Nowozin, B. Cseke, and R. Tomioka, "F-GAN: Training generative neural samplers using variational divergence minimization," in Advances in Neural Information Processing Systems, D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, Eds. New York, NY, USA: Curran Associates, 2016, pp. 271–279. [Online]. Available: http://papers.nips.cc/paper/6066-f-gan-training-generative-neural-samplers-using-variational-divergence-minimization.pdf

[27] T. Che, Y. Li, A. P. Jacob, Y. Bengio, and W. Li, "Mode regularized generative adversarial networks," in Proc. Int. Conf. Learn. Represent., 2017, pp. 1–13. [Online]. Available: https://arxiv.org/abs/1612.02136

[28] X. Mao, Q. Li, H. Xie, R. Y. K. Lau, Z. Wang, and S. P. Smolley, "On the effectiveness of least squares generative adversarial networks," IEEE Trans. Pattern Anal. Mach. Intell., to be published.

[29] G.-J. Qi. (Jan. 2017). "Loss-sensitive generative adversarial networks on lipschitz densities." [Online]. Available: https://arxiv.org/abs/1701.06264.

[30] J. Zhao, M. Mathieu, and Y. LeCun, "Energy-based generative adversarial network," in Proc. Int. Conf. Learn. Represent., 2017, pp. 1–17. [Online]. Available: https://arxiv.org/abs/1609.03126

[31] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein generative adversarial networks," in Proc. Int. Conf. Mach. Learn., vol. 70, Aug. 2017, pp. 214–223. [Online]. Available: http://proceedings.mlr.press/v70/arjovsky17a.html

[32] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, "Improved training of wasserstein GANs," in Proc. 30th Adv. Neural Inf. Process. Syst., 2017, pp. 5767–5777. [Online]. Available: http://papers.nips.cc/paper/7159-improved-training-of-wassersteingans.pdf

[33] H. Petzka, A. Fischer, and D. Lukovnicov, "On the regularization of wasserstein GANs," in Proc. Int. Conf. Learn. Represent., 2018, pp. 1–24. [Online]. Available: https://arxiv.org/abs/1709.08894

[34] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," in Proc. Int. Conf. Learn. Represent., 2016, pp. 1–16. [Online]. Available: https://arxiv.org/abs/1511.06434

[35] M. Mirza and S. Osindero. (Nov. 2014). "Conditional generative adversarial nets." [Online]. Available: https://arxiv.org/abs/1411.1784

[36] X. Chen, Y. Duan, R. Houthooft, J. Schulman, I. Sutskever, and P. Abbeel, "InfoGAN: Interpretable representation learning by information maximizing generative adversarial nets," in Proc. 29th Adv. Neural Inf. Process. Syst., 2016, pp. 2172–2180. [Online]. Available: http://papers.nips.cc/paper/6399-infogan-interpretable-representationlearning-by-information-maximizing-generative-adversarial-nets.pdf

[37] A. Odena, C. Olah, and J. Shlens, "Conditional image synthesis with auxiliary classifier GANs," in Proc. 34th Int. Conf. Mach. Learn., vol. 70, Aug. 2017, pp. 2642–2651. [Online]. Available: http://proceedings.mlr.press/v70/odena17a.html

[38] A. Makhzani, J. Shlens, N. Jaitly, and I. Goodfellow, "Adversarial autoencoders," in Proc. Int. Conf. Learn. Represent., 2016, pp. 1–16. [Online]. Available: http://arxiv.org/abs/1511.05644

[39] J. Donahue, P. Krähenbühl, and T. Darrell, "Adversarial feature learning," in Proc. Int. Conf. Learn. Represent., 2017, pp. 1–18. [Online]. Available: https://arxiv.org/abs/1605.09782

[40] V. Dumoulin et al., "Adversarially learned inference," in Proc. Int. Conf. Learn. Represent., 2017, pp.1–18. [Online]. Available: https://arxiv.org/abs/1606.00704

[41] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "It takes (only) two: Adversarial generator-encoder networks," in Proc. AAAI Conf. Artif. Intell., 2018. [Online]. Available: https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/view/16568

[42] A. B. L. Larsen, S. K. Sønderby, H. Larochelle, and O. Winther, "Autoencoding beyond pixels using a learned similarity metric," in Proc.

33rd Int. Conf. Mach. Learn., Jun. 2016, pp. 1558–1566. [Online]. Available: http://proceedings.mlr.press/v48/larsen16.html
[43] Yaxing Wang, Abel Gonzalez-Garcia, Joost van de Weijer, and Luis Herranz. 2019. SDIT: Scalable and Diverse Cross-domain Image Translation. In Proceedings of the 27th ACM International Conference on Multimedia (MM '19). Association for Computing Machinery, New York, NY, USA, 1267–1276. https://doi.org/10.1145/3343031.3351004
[44] Mirza, Mehdi, and Simon Osindero. "Conditional generative adversarial nets." arXiv preprint arXiv:1411.1784 (2014).
[45] W. Chen, R. Zheng, P.D. Baade, S. Zhang, H. Zeng, F. Bray, A. Jemal, X. Yu, J. He, Cancer statistics in China, 2015, Cancer J. Clin. 66 (2) (2016) 115–132.
[46] R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, 2017, Cancer J. Clin. 69 (1) (2019) 7–34.
[47] American Cancer Society, Cancer Facts and Figures 2018, American Cancer Society, Atlanta, 2018.
[48] M.S. Brady, S.A. Oliveria, P.J. Christos, M. Berwick, D.G. Coit, J. Katz, A.C. Halpern, Patterns of detection in patients with cutaneous melanoma, Cancer 89 (2) (2000) 342–347.
[49] M.E. Vestergaard, P.H.P.M. Macaskill, P.E. Holt, S.W. Menzies, Dermoscopy compared with naked eye examination for the diagnosis of primary melanoma: a meta-analysis of studies performed in a clinical setting, Br. J. Dermatol 159 (3) (2008) 669–676.
[50] H. Kittler, H. Pehamberger, K. Wolff, M. Binder, Diagnostic accuracy of dermoscopy, Lancet Oncol 3 (3) (2002) 159–165.
[51] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial nets. In: NIPS'14: NIPS'14: Proceedings of the 27th International Conference on Neural Information Processing Systems. 2014;2:2672–80.
[52] (2018) International Skin Imaging Collaboration Website. Available: https://challenge.isic-archive.com/landing/2018/.
[53] P. Tschandl, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," Harvard Dataverse, 2018. [Online]. Available: https://doi.org/10.7910/DVN/DBW86T.
[54] P.Tschandl, C. Rosendahl, H. Kittler, The HAM1000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions, Scientific Data, 14 August 2018.