Unpaired Image Translations Using GANs: A Review

Komal Sheladiya[1], Dr Udesang Jaliya[2], Mahasweta Joshi[3]

[1]Student (Mtech), Department of Computer Engineering, Birla Vishvakarma Mahavidyalaya, Anand, Gujarat, India
[2]Professor, Department of Computer Engineering, Birla Vishvakarma Mahavidyalaya, Anand, Gujarat, India
[3]Professor, Department of Computer Engineering, Birla Vishvakarma Mahavidyalaya, Anand, Gujarat, India

---***---

Abstract - *A survey of image translations using Generative Adversarial Networks (GANs) is presented in this survey paper. GANs are made up of two deep networks, a generator, and a discriminator, all of which are competitively trained. GANs are capable of producing reasonable and realistic images and have demonstrated significant capacity in various image synthesis and editing applications, Deep networks have a lot of power and a competitive training method. This study reviews recent GAN papers and unpaired image-to-image translation with CycleGAN on subjects like texture creation, image imprinting, image-to-image translation, and image editing, among others. This review paper is based on different methods of GANs for different Domains such as Horse2Zebra, Winter2Summer, Medical images, Monet's images, etc..*

Key Words: **Generative adversarial networks, Cycle GAN, image editing, image-to-image translations**

1. INTRODUCTION

Deep learning has also demonstrated impressive content-generating abilities. Good fellow et al. proposed generative adversarial networks (GANs) in 2014[5]. GANs are made up of two networks: a generator and a discriminator. The discriminator tries to tell the difference between fake and real photos; the generator creates bogus images to deceive the discriminator. Both networks are competitively trained together. The result is a generator that can generate believable visuals. GAN variations have lately achieved outstanding results in a range of image synthesis and editing applications.

Paired image-to-image translation input, output strive image edges, and realistic image. edge detector and go from a practical image to those edges. but that's nice once having some formula to induce that paired training data, or have already got that paired training data for several reasons. unpaired image-to-image translation, it's not this straightforward, let's modification a horse to zebra, or have a painting by a painter. [2][3]

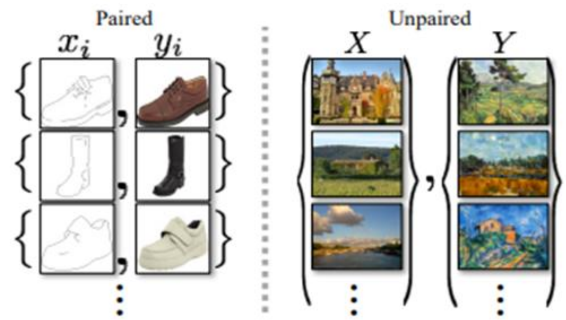

Fig -1: pair image (left side) and umpires' image (right side) [10]

Pire image the paradigm between these two image-to-image translation tasks is that for one of them, these paired images. pair up x_i and y_i so, here i is probably 0, 1, 2. All these different pairs match onto each other but do not necessarily have this correspondence all the time [3]. An unpaired image-to-image translation two piles of two different image styles, x and y. one pile which is maybe realistic photos, the other maybe its Monet or Cezanne or someone else or a pile of winter looking images versus summer scenes or a pile of horses in one or zebras in another, do not have a one-to-one correspondence at all. [1]

CycleGAN is an image from one pile, there's a real image and transform it into the style from that other pile which has that another image style going on, and then transform it back into a real image and its still fake image. It's taking in this fake image to generate another fake real image. Technically only changing styles and not the content of the image. While mapping from one pile to another and back again, Because the translation creates a cycle between these two piles, this content preservation is known as cycle consistency. Using two separate GANs is one straightforward way of generating this cycle. 1) The adversarial part of the GANs just discriminates it for both of them will ensure realism in the images, and while 2) the cycle consistency part is really in charge of getting the content to be preserved while only moving around the styles.

This survey discusses recent publications that use generative adversarial networks (GANs) for image modification. The ideas, contributions, and limitations of various networks are discussed in this survey. The following is a breakdown of the survey's structure. GANs and related variants are briefly discussed in Section 2. The CycleGAN is discussed in Section 3. Section 4 concludes with an overview of the present issues and limits of CycleGAN-based approaches.

2. Generative Adversarial Networks

Generative Adversarial Networks its task to take input from the noise vector and send it forward to the Generator and then to Discriminator to identify and differentiate unique and fake inputs. The GAN is made up of two deep networks. G is a generator that makes forgeries to create realistic visuals. D is a discriminator who receives both forgeries and actual (authentic) photos and tries to distinguish between them (see Figure 2). Both are trained at the same time and in competition with one another.

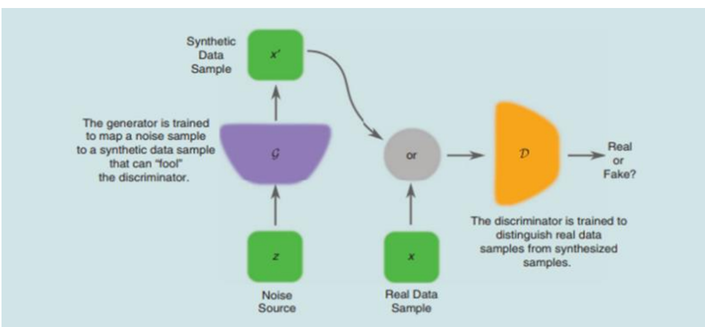

Figure 2.1: The two models that are learned during the training process for a GAN are the discriminator (D) and the generator (G). These are most commonly implemented using neural networks, although they might be built using any type of differentiable system that maps data from one space to another. [19]

3. Cycle GAN

CycleGAN and image-to-image translation model just like Pix2Pix. CycleGAN, or Cycle Generative Adversarial Network, is a method for training a deep convolutional neural network to do image-to-image translation tasks. The Network learns a mapping between input and output images using an unpaired dataset. Generating RGB imagery from SAR, multispectral photography from RGB, map routes from satellite imagery, and so on are only a few examples. the adversarial network CycleGAN also includes two parts: a generator and a discriminator. The generator's role is to generate samples from the desired distribution, while the discriminator's role is to determine if the sample is from a real distribution or one generated by the generator (fake) [11].

Cycle Consistency

The optimization issue is given a cycle consistency loss function, which indicates that if we convert a zebra image to a horse image and then back to a zebra image, we should receive the same input image. [10]

Adversarial Loss a continuously trained discriminator network defines the adversarial loss. Both mapping functions have adversarial losses [4]. The mapping function G: X → Y and its discriminator DY have the following goal:

$$LGAN\ (G, D_Y, X, Y) = E_{y\sim pdata(y)}\ [\log D_Y\ (y)] + E_{x\sim pdata(x)}\ [\log (1 - D_Y (G(x))] \quad (1)$$

Where, G seeks to generate images G(x) that resemble images from domain Y, and D_Y tries to tell the difference between translated samples G(x) and real samples y. G's goal is to minimize this objective while D wants to maximize it, i.e., minG maxD_Y LGAN (G, D_Y, X, Y). A similar adversarial loss is introduced for the mapping function F: Y→X and its discriminator D_X: minF maxD_X LGAN (F, D_X, Y, X) [10].

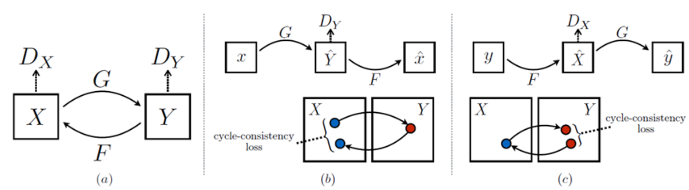

Figure 3.1: (a) CycleGAN contains two mapping functions G and F, (b) forward cycle consistency loss, and (c) backward cycle consistency loss. [10]

Cycle Consistency Loss adversarial losses alone aren't enough to ensure that the learned function can transfer a specific input x_i to the expected output Yi. The learned mapping functions are supposed to be cycle-consistent, according to the argument.

Forward cycle consistency: The image translation cycle should be able to bring each image x from domain X back to the original image. i.e.,

$$x \rightarrow G(x) \rightarrow F(G(x)) \approx x. \quad (2)$$

Likewise **Backward cycle consistency:**

$$y \rightarrow F(y) \rightarrow G(F(y)) \rightarrow y. \quad (3)$$

$L_{cyc}(G, F) = E_{x \sim pdata(x)} [kF(G(x)) - xk_1] + E_{y \sim pdata(y)} [kG(F(y)) - yk_1]$ (4)

4. Review of Image-to-Image Translation using GANs

The image-to-image translation is a category of vision and graphics problems in which the goal is to figure out how to transfer an input image to an output image. It can be used for a variety of tasks, including collection style transfer, object transformation, season transition, and photo enhancement. The image-to-image translation is a conditional generation framework for transforming photos into various styles. Taking an image and modifying it to create a new image with a distinct style while keeping the original content.

Jun-Yan Zhu et al. (2017) Worked on Unpaired image-to-image Translation using Cycle-Consistent Adversarial Networks Although the strategy can produce compelling results in many circumstances, the outcomes are not always favorable. In several common failure scenarios, the approach frequently succeeds on translation challenges involving color and texture alterations. Authors also attempted but failed to complete projects that required geometric adjustments. On the task of dog →cat transfiguration, for example, the learned translation devolves to making minor alterations to the input. Managing increasingly complex and radical transformations, particularly geometric shifts. [10]

Hugo Touvron et al. (2020) explain the Powers of layer for image-to-image translation Iterating a residual block to learn a complex transformation without direct supervision is what layers power is all about. With fewer parameters, the power of layers performs similarly to CycleGAN on various tasks. The common embedding space's flexibility can be utilized to vary the strength of a transformation or to combine many transformations. While the discriminator is often only utilized for training, Powers of Layers can use it at inference time to modify the transformation to the input image. [11]

Rui Zhang et al. (2019) Explain the Harmonic Unpaired image-to-image Translation The sample graph is used to ensure that the source and destination domains are smooth and consistent. The intrinsic self-consistency property of samples can be maintained by providing further regularization to enforce consistent mappings during the image-to-image translation. The researchers demonstrated that this results in a significant improvement over current state-of-the-art methods in many applications, including medical imaging, object transfiguration, and semantic labeling, through a series of quantitative, qualitative, and user evaluations. In a medical imaging challenge, in particular, our technique outperforms CycleGAN by a large margin. [12]

Lei Chen et al. (2019) Worked on Quality-aware Unpaired image-to-image Translation The problem of unpaired image-to-image translation was investigated, and a unified QGAN framework for quality-aware unpaired image translation was developed. A quality-aware loss term is explicitly implemented in the optimization function of the QGAN framework. We created two thorough quality loss implementations, QGAN A and QGAN C, that took into account the classical quality evaluation approach as well as adaptive high-level content structure information from deep networks. Extensive quantitative comparisons with past models, as well as a mean opinion score test, demonstrated that our proposed framework and two thorough implementations are of higher quality. [13]

Zengming Shen et. al. (2020) For unpaired image-to-image translation a one-to-one mapping function is used. In terms of several qualitative and quantitative criteria, the proposed one-to-one CycleGAN consistently outperforms the baseline CycleGAN model and other state-of-the-art unsupervised techniques. [14]

Table -1: Comparisons between different image translation GANs, each row gives the information of a Domain of image translation, Dataset & accuracy. [10 – 14]

Sr No.	Paper Name	Methods	Domain	Dataset	Accuracy
1	Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks	MUNIT + CycleGAN	Map → photo Map ← photo	(Mapping aerial on Google Maps) 1525	26.8% 23.2%
			Labels → photo Labels ← photo	(Cityscapes images) 3475	0.52 0.58

2	Powers of layers for image-to-image translation	Power of level (POL)	Summer → Winter Summer ← Winter	(Yosemite) 2345	46.1 44.4
			Horse → Zebra Horse ← Zebra	(CycleGAN) 2661	53.0 93.3
			Monet → Picture Monet ← Picture	(Monet) 1365	70.3 82.1
3	Harmonic unpaired image-to-image translation	Harmonic GAN Histogram VGG	(Medical images) Flair → T1 Flair ← T1	(BRATS) 1700	24.34 27.22
			Horse → zebra	(Cycle GAN) 2661	72%
4	Quality-aware Unpaired image-to-image Translation	Quality GAN	Sketch → photo	(Photo-sketch) 2378	2.71
			Label → facade	(Label-facade) 1000	3.26
			Winter → summer	(Summer-winter) 2400	3.91
5	One-to-one Mapping for Unpaired image-to-image Translation	One2one CycleGAN	Label → Photo Label ← Photo	(Cityscapes) 2528	58.2 52.7
			Horse → zebra Horse ← zebra	(Cycle GAN) 2661	75% 77%
			T1 → T2 T2 ← T1	(BRATS, MRI Data) 351 3D	22.03 18.31
			Summer → winter Summer ← winter	(Yosemite) 2345	66% 59%

5. Conclusion

A CycleGAN is created by using two GANs to transform images from two piles to and from each other. This is for non-paired image-to-image translation. PatchGANs are used as discriminators, and the generators are a combination of DCGAN and U-Net learning with additional skip connections. The approach frequently succeeds on translation problems that entail color and texture changes, such as many of those mentioned above. We've also looked into tasks that need geometric adjustments.

6. References

[1] Isola, P. (2016, November 21). Image-to-Image Translation with Conditional Adversarial Networks. arXiv.Org. https://arxiv.org/abs/1611.07004

[2] Chapter 9. CycleGAN · GANs in Action: Deep learning with Generative Adversarial Networks. (2017). Livebook.Manning.Com. https://livebook.manning.com/book/gans-in action/chapter-9/35

[3] Cohen, J. P. (2018, May 22). Distribution Matching Losses Can Hallucinate Features in Medical. . . arXiv.Org. https://arxiv.org/abs/1805.08841

[4] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in Advances in Neural Information Processing Systems 27, 2014, pp. 2672–2680.

[5] M. Mirza and S. Osindero, "Conditional generative adversarial nets," arXiv preprint arXiv:1411.1784, 2014.

[6] Radford, A. (2015, November 19). Unsupervised Representation Learning with Deep Convolutional. . . arXiv.Org. https://arxiv.org/abs/1511.06434

[7] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," in ICLR, 2016.

[8] E. L. Denton, S. Chintala, a. szlam, and R. Fergus, "Deep generative image models using a laplacian pyramid of adversarial networks," in Advances in Neural Information Processing Systems 28, 2015, pp. 1486–1494.

[9] Antonia Creswell, Tom White, Vincent Dumoulin, Kai Arulkumaran, Biswa Sengupta,and Anil A. Bharath

[10] Zhu, J. (2017, March 30). Unpaired Image-to-Image Translation using Cycle-Consistent. . . arXiv.Org. https://arxiv.org/abs/1703.10593

[11] Touvron, H. (2020, August 13). Powers of layers for image-to-image translation.arXiv.Org. https://arxiv.org/abs/2008.05763

[12] Rui Zhang, Tomas Pfister, Jia Li (2019) 'HARMONIC UNPAIRED IMAGE-TO-IMAGE TRANSLATION' ICLR, 1 – 19.

[13] Lei Chen, Le Wu, Zhenzhen Hu, Meng Wang (2019) 'Quality-aware Unpaired image-to-image Translation' IEEE, 1-12

[14] Zengming Shen, Yifan Chen, Thomas S. Huang, S. Kevin Zhou, Bogdan Georgescu, Xuqi Liu (2020) 'One-to-one Mapping for Unpaired image-to-image Translation' **WACV** Paper is the open access version, available on IEEE Xplore, 1170 – 1179

[15] Minjun Li, Haozhi Huang, Lin Ma, Wei Liu, Tong Zhang, Yugang Jiang (2018) **'**Unsupervised image-to-image Translation with Stacked Cycle-Consistent Adversarial Networks**' published** ECCV, 1 – 16.

[16] Brownlee, J. (2020, September 1). How to Develop a CycleGAN for Image-to-Image Translation with Keras. Machine Learning Mastery. https://machinelearningmastery.com/cyclegan-tutorial-with-keras/

[17] Tsang, S. (2022, January 6). Review — CycleGAN: Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks (GAN). Medium. https://shtsang.medium.com/review-cyclegan-unpaired-image-to-image-translationusing-cycle-consistent-adversarial-networks-1c2602805be2

[18] GeeksforGeeks. (2021, August 27). Cycle Generative Adversarial Network (CycleGAN). https://www.geeksforgeeks.org/cycle-generative-adversarial-network-cyclegan-2/

[19] Generative Adversarial Networks an Overview. (2018, January 1). IEEE Journals & Magazine IEEE Xplore. httpsieeexplore.ieee.orgabstractdocument825359