International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2020 IJSRCSEIT | Volume 6 | Issue 3 | ISSN : 2456-3307
DOI : https://doi.org/10.32628/CSEIT206370

Image Outpainting and Harmonization using GANs

Vignesha K. , Rabeeh Mohammed Ali
Department of Computer Science, Srinivas Institute of Technology, Mangalore, Karnataka, India

ABSTRACT

Although the inherently ambiguous task of predicting what resides on the far side all four edges of a image has rarely been explored before, we have a tendency to demonstrate that GANs hold powerful potential in manufacturing reasonable extrapolations. Two outpainting ways square measure projected that aim to instigate this line of research: the primary approach uses a context encoder inspired by common inpainting architectures and paradigms, whereas the second approach adds an extra post-processing step using a single-image generative model. This way, the hallucinated details are integrated with the design of the original image.
Keywords: Tensorflow, Deep Learning, CNN, GANS, Neural Networks.

I. INTRODUCTION

When given with an incomplete image, humans are wonderful at filling within the blanks and manufacturing a realistic clarification for what may well be missing. Image in-painting may be a well-studied drawback that replicates this behaviour, typically tasking deep neural networks with making an attempt to understanding the linguistics content of natural pictures so as to recover the missing regions of a photograph. However, the spatially inverted variant of this drawback is even more difficult and, with a little play on words, is denoted outpainting. the matter statement is shown in Figure 1; basically, the task is to extrapolate the image content instead of to interpolate among a picture. additional formally, we tend to should style a generator tt that converts a picture x with dimensions n n into a bigger image tt(x) with dimensions m m, specified the middle a part of tt(x) appearance constant as x, whereas the whole outpainted image tt(x) ought to be a plausible hypothesis of what might include the first image.

Fig. 1 : Image outpainting idea.

II. RELATED WORK

Deep neural networks have recently shown nice performance in image completion. This section discusses previous work associated with the planned strategies for outpainting.

Generative Adversarial Networks Image or video generation victimisation convolutional neural networks with associate degree adversarial loss [4] has attracted important research interest. The idea is to let a generator network produce samples consistent with some distribution, subsequently have an auxiliary network referred to as the discriminator D try and distinguish whether or not a given sample is valid or was truly generated by tt.Wherever tt attempting to fool D forces D to become higher at telling real from pretend, that successively motivates tt to synthesize more and more convincing outputs [5].

Outpainting the matter opposite to inpainting includes predicting that pixels reside on the far side the borders of a completely intact photograph, and must our data been explored few times before by the academic research community [8–10]. One approach geometrically extrapolates the field of view of a image using another broad reference image of same scene class using old-school computer vision techniques [8], whereas a newer relevant paper uses a GAN to perform horizontal outpainting [9]. [10] achieves impressive outcomes but tends to focus on limited domain datasets (such as faces only), and notes that generative models expertise difficulties attempting to suit datasets as numerous as Places2. Lastly, Google's Snapseed application includes a proprietary 'Expand' practicality that appears to pick out patches from the image and copy them to the edge [11], however a limited number of experiments suggest that this tool fails to capture the native structure of most scenes from region to region.

Single-image generative models The recently introduced SinGAN framework can train unconditional generative model on one natural photo, capturing and reproducing image statistics across various scales of the image [12]. It permits for the creation of random samples with new object configurations by ranging from a low-resolution sample at the coarsest scale, and so increasingly upsampling and refining the result through the pyramid of generators.

III. SYSTEM IMPLEMENTATION

A. Data Set

During coaching, we decide to crop pictures first before feeding them into the generator. Consequently, learning will be done a self-supervised way to the actual fact that we are now able to enforce the output to approximate the initial, uncropped image. Any sufficiently large dataset of unlabeled , natural photos will therefore suffice.A second set of experiments was performed with a dataset consisting of images scraped from WikiArt [14].

The photos can 1st be resized to 192x192 as a preprocessing step, and therefore the generator can then be tasked with increasing a crop of 128x128 back to a 192x192 image. In follow, the generator tt maps a partly covert 192x192 color image to an out-painted variant of an equivalent dimensions, with the covert masked part replaced by the model's predictions.

B. Architecture

Many study aspects and ideas are often naturally adopted from inpainting. The context encoder a part of the generator network tt repeatedly downsamples the disguised input through six convolutional layers, so as to efficiently capture the image content and object semantic within embedding area. Next, the decoder consists of a special reasonably layers known as up-convolutional or deconvolutional, which might be understood as having a third stride so as to 'undo' the downsampling performed by the encoder [7].

C. Training

Training is completed for 200 epochs, with a fixedt learning rate of α = 0.0003 and two Adam optimizers with β1 = 0.5, β2 = 0.999. The loss functions area unit as follows:

$$L_{rec} = ||x - G(x)||_1 \quad (2.1)$$

$$L_{adv} = ||D(G(x)) - 1||_2^2 \quad (2.2)$$

$$L_G = \lambda_{rec} L_{rec} + \lambda_{adv} L_{adv} \quad (2.3)$$

$$L_D = ||D(x) - 1||_2^2 + ||D(G(x)) - 0||_2^2 \quad (2.4)$$

Using an L1 reconstruction loss rather than L2 helps produce less blurry images . The weight of the adversarial loss λadv relative to the reconstruction loss λrec = 1 λadv turned to be notably tough to adjust; this factor was at first set to λadv= 0.001 as in various existing works [7, 9], although the GAN kept collapsing into a failure mode wherever the adversarial loss didn't move removed from one. this suggests that tt was unable to fool D, thus D was prior tt and will always in which tell real from fake successfully. A operating remedy concerned varied λadv(n) throughout time as a perform of the epoch n as follows:

$$\lambda_{adv}(n) = \begin{cases} 0.001, & \text{if } n \leq 10 \\ 0.005, & \text{if } 10 < n \leq 30 \\ 0.015, & \text{if } 30 < n \leq 60 \\ 0.040, & \text{otherwise} \end{cases} \qquad (2.5)$$

This will penelise the generator additional heavily for manufacturing unreasonable outputs as time progresses,instead of simply implementing an accurate pixel wise reconstruction.

D. Harmonization

We first train a SinGAN model on the initial high-quality image, then propagate it forward through the outpainting generator (which produces a low-resolution output), so try to super-resolve this result by injecting it into one amongst the coarser scales of SinGAN. The hope is that the model can harmonize the outpainted elements with the style of the initial image that it had been trained on, whereas simultaneously synthesizing a finer-scale, higher-resolution variant by pushing it up through the hierarchy of multi-scale generators.

IV. RESULT

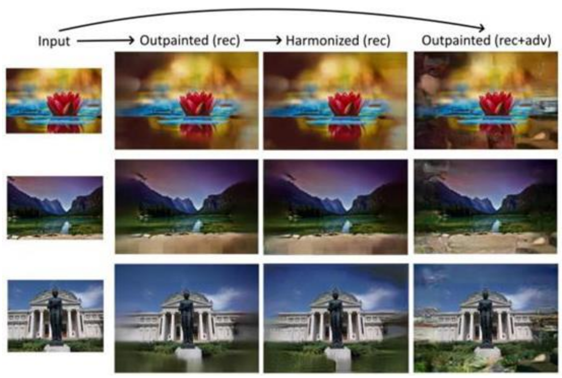

FIG. 2: Qualitative comparison of outpointing

In contrast to inpainting, outpainting doesn't have to be performed simply once; in theory, there's no limit on what number times we are able to extrapolate one image. Our outpainting introduces a further step that involves using SinGAN as described earlier. the most important downside of this approach is that it takes around associate hour to train SinGAN on one input image, in distinction to the outpainter itself that could be a model that may merely be applied anytime once it is trained on a dataset. Nevertheless, we tend to tested out the joint harmonization and super-resolution method on several pictures that demonstrate its potential, as seen in Figure 2.

V. CONCLUSION

Image outpainting is novel however exciting idea holds promise, particularly once cascaded with SinGAN to any increase the output fidelity. The models trained during this project still contain some glitches, but we believe these can be mitigated by closer investigations into the subject. Non-photorealistic images like artwork appear to supply convincing results, which we largely attribute to human judgement changing into a lot of permissive instead of to a better-performing model.

VI. REFERENCES

[1]. Shakhnarovich. Colorization as a proxy task for visual understanding. Proceedings - 30th IEEE Con- ference on Computer Vision and Pattern Recognition, CVPR 2017, 2017-Janua:840–849, 2017.

[2]. Carl Doersch, Abhinav Gupta, and Alexei A. Efros. Un- supervised visual representation learning by context pre- diction. Proceedings of the IEEE International Confer- ence on Computer Vision, 2015 Inter:1422–1430, 2015.

[3]. Mehdi Noroozi and Paolo Favaro. Unsupervised learn- ing of visual representations by solving jigsaw puzzles. Lecture Notes in Computer

Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 9910 LNCS:69–84, 2016.

[4]. Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative Adversarial Networks. pages 1–9, 2014.

[5]. Satoshi Iizuka, Edgar Simo-Serra, and Hiroshi Ishikawa. Globally and locally consistent image completion. ACM Transactions on Graphics, 36(4), 2017.

[6]. Lars Mescheder, Andreas Geiger, and Sebastian Nowozin. Which training methods for GANs do actually con- verge? 35th International Conference on Machine Learn- ing, ICML 2018, 8:5589–5626, 2018.

[7]. Deepak Pathak, Philipp Krahenbuhl, Jeff Donahue, Trevor Darrell, and Alexei A. Efros. Context Encoders: Feature Learning by Inpainting. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016-Decem:2536–2544, 2016.

[8]. Yinda Zhang, Jianxiong Xiao, James Hays, and Ping Tan. Framebreak: Dramatic image extrapolation by guided shift-maps. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 1171–1178, 2013.

[9]. Mark Sabini and Gili Rusak. Painting Outside the Box: Image Outpainting with GANs. 2018.

[10]. Yi Wang, Xin Tao, Xiaoyong Shen, and Jiaya Jia. Wide- context semantic image extrapolation. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2019-June:1399–1408, 2019.

[11]. Snapseed - apps on google play, Jun 2018.

[12]. Tamar Rott Shaham, Tali Dekel, and Tomer Michaeli. SinGAN: Learning a Generative Model from a Single Natural Image. may 2019.

[13]. Bolei Zhou, Agata Lapedriza, Aditya Khosla, Aude Oliva, and Antonio Torralba. Places: A 10 Million Image Database for Scene Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(6):1452– 1464, 2018.

[14]. Wikiart: Visual art encyclopedia.

Cite this article as :

Vignesha K., Rabeeh Mohammed Ali, "Image Outpainting and Harmonization using GANs ", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 6, Issue 3, pp.294-297, May-June-2020. Available at doi : https://doi.org/10.32628/CSEIT206370

Journal URL : http://ijsrcseit.com/CSEIT206370