

Face Image Inpainting with Evolutionary Generators

Chong Han, Junli Wang

Abstract—Recently, deep learning has become a mainstream method of image inpainting. It can not only restore the image texture, obtain high-level abstract features of images, but also restore semantic images such as human face images. Among these methods, generative adversarial networks (GANs) using autoencoder as the generator have become the promising model for image inpainting. These models implement the end-to-end image inpainting and also generate visually reasonable and clear image structures and textures. However, GANs often have problems with gradient vanishing and model collapse during training, so we propose a Generative Adversarial Network with Evolutionary Generators (EG-GAN) and apply it in face image inpainting. To stabilize the model training process, EG-GAN trains the generator network by evolution, combines two mutation functions as a training objective to update the parameter of generator networks, and produces offspring generators through crossover, using the matcher assists the discriminator to criticize the generated image. Experiments on various face image datasets such as CelebA-HQ and CelebA show that EG-GAN successfully overcomes the gradient vanishing problem, achieves stable and efficient training, and generates visually reasonable images.

Index Terms—Neuro-evolution, Autoencoder, Generative Adversarial Networks, Face Image Inpainting.

I. Introduction

Image inpainting is an important task in machine vision, which aims to fill in missing pixels in damaged images. Initially, Pathak et al. proposed an image inpainting model based on an autoencoder: Context Encoder (CE)[1]. Because CE is trained by Euclidean distance, which will inevitably lead to the blurred image[2]. Afterwards, GAN has become one of the most significant research domains because of its various applications in the field of image processing and multi-view works[3][4]. Therefore, researchers began to use GANs. For example, Raymond et al. proposed semantic image inpainting with deep generative models (SIIGAN)[5]. However, this model can't accomplish an end-to-end process, so it takes a lot of time to train. To overcome the above problems, Pathak et al. added an adversarial network into the original CE, which is applied to image inpainting.

Neuro-evolution [6] is a method that uses biological evolution theory or evolutionary computation to generate artificial neural network parameters, structures, and rules. Therefore, deep learning models based on neuro-evolution

can be divided into two categories: the first is to optimize the deep learning model by biological evolution theory. For example, Chaoyue Wang et al. proposed EGAN[7]. The second is to use evolutionary computing to optimize deep learning models. Evolutionary computing includes: genetic algorithm (GA), genetic programming (GP), evolutionary strategy (ES), and evolutionary programming (EP). For example, Masanori Suganuma et al. proposed the ES-CAE [8] which applies an evolutionary strategy to autoencoder. Moreover, these experiments prove that deep learning models optimized by neuro-evolution are easier to train and possible to generate some skip connections that are difficult to be designed by a human.

In this letter, EG-GAN is inspired by biological evolution. GAN is trained in the way of neuro-evolution and applied to face image inpainting. EG-GAN mainly consists of three parts: generator, discriminator, and matcher. The generator is regarded as an individual in the population, and the network parameters in the generator are optimized by mutation and crossover. The discriminator is utilized to criticize the quality of the entire generated image. But for image inpainting, the image quality of missing regions is the key to determine the performance of the generator. Therefore, the discriminator has certain limitations. To solve this problem, this letter proposes a matcher to assist the discriminator. The contributions of this letter are summarized as follows:

- (1) It is proposed to train the generator by evolution and combine two different mutation functions as the objective function to update the parameters in generators, avoiding the gradient vanishing and model collapse. What's more, a crossover is added to explore the commonality in excellent generator models, producing offspring population, so that elite generators will not be lost in the process of mutation.
- (2) A matcher is proposed to assist the discriminator in criticizing the generated image. The matcher can learn to obtain the relevance of each component in the image and focus on criticizing the contextual correlation of the generated regions, avoiding discriminators giving high-value feedback to images with wrong correlation, thereby misleading the generator.

- (3) EG-GAN achieves high-quality inpainting results on various face image datasets, including CelebA-HQ, CelebA, PubFig, and non-realistic high-definition face datasets generated by StyleGAN. Meanwhile, the quantitative evaluations on PSNR and SSIM are higher than some traditional and recent image inpainting models.

II. Guidelines for Graphics Preparation and Submission Guidelines

Manuscript received July 26, 2020; revised October 2, 2020. This work was supported in part by the National Key R&D Program of China(2017YFA0700602), in part by the National Natural Science Foundation of China(No.61672381); in part by the Fundamental Research Funds for the Central Universities.(Corresponding author: Junli Wang)

C. Han and J. Wang are with the Department of Electronics and Information Engineering, Tongji University, Shanghai 201804, China(e-mail: 496274966@qq.com; e-mail: junliwang@tongji.edu.cn)

III. Conclusion

A conclusion

References

- [1] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, and A. A. Efros, "Context encoders: Feature learning by inpainting," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2536–2544. I
- [2] D. Yoo, N. Kim, S. Park, A. S. Paek, and I. S. Kweon, "Pixel-level domain transfer," in *European Conference on Computer Vision*. Springer, 2016, pp. 517–532. I
- [3] P. Hu, D. Peng, Y. Sang, and Y. Xiang, "Multi-view linear discriminant analysis network," *IEEE Transactions on Image Processing*, vol. 28, no. 11, pp. 5352–5365, 2019. I
- [4] P. Hu, D. Peng, X. Wang, and Y. Xiang, "Multimodal adversarial network for cross-modal retrieval," *Knowledge-Based Systems*, vol. 180, pp. 38–50, 2019. I
- [5] R. A. Yeh, C. Chen, T. Yian Lim, A. G. Schwing, M. Hasegawa-Johnson, and M. N. Do, "Semantic image inpainting with deep generative models," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 5485–5493. I
- [6] D. Floreano, P. Dürri, and C. Mattiussi, "Neuroevolution: from architectures to learning," *Evolutionary intelligence*, vol. 1, no. 1, pp. 47–62, 2008. I
- [7] C. Wang, C. Xu, X. Yao, and D. Tao, "Evolutionary generative adversarial networks," *IEEE Transactions on Evolutionary Computation*, vol. 23, no. 6, pp. 921–934, 2019. I
- [8] M. Suganuma, M. Ozay, and T. Okatani, "Exploiting the potential of standard convolutional autoencoders for image restoration by evolutionary search," *arXiv preprint arXiv:1803.00370*, 2018. I