中国科学院软件研究所学术年会’2020
暨计算机科学国家重点实验室开放周

学术论文

基于生成式对抗网络的数据增强方法综述

苏兴哲
ICNC-FSKD 2020
xingzhe2018@iscas.ac.cn, 17888823596

Introduction

Over the last decade, Deep Neural Networks (DNN) have been used in many aspects, including image classification, object detection and so on. However, large datasets are essential for these networks to get convincing results, otherwise, they are prone to overfitting. In practice, it is difficult to obtain large datasets, especially in some special fields like remote sensing and medical fields. Data augmentation is one of the common methods to help networks combat overfitting and enhance robustness. The frequently used methods are various transformations, such as rotations, flips and etc. These operations have limited effect on generating new samples and enriching the diversity of the dataset. Generative Adversarial Networks (GAN) have been continuously modified and widely used in computer vision tasks during the past few years. Data augmentation methods based on GAN also have great potential in practical applications. This paper is aimed to provide a survey on various GAN-based data augmentation methods in computer vision.

Data augmentation methods

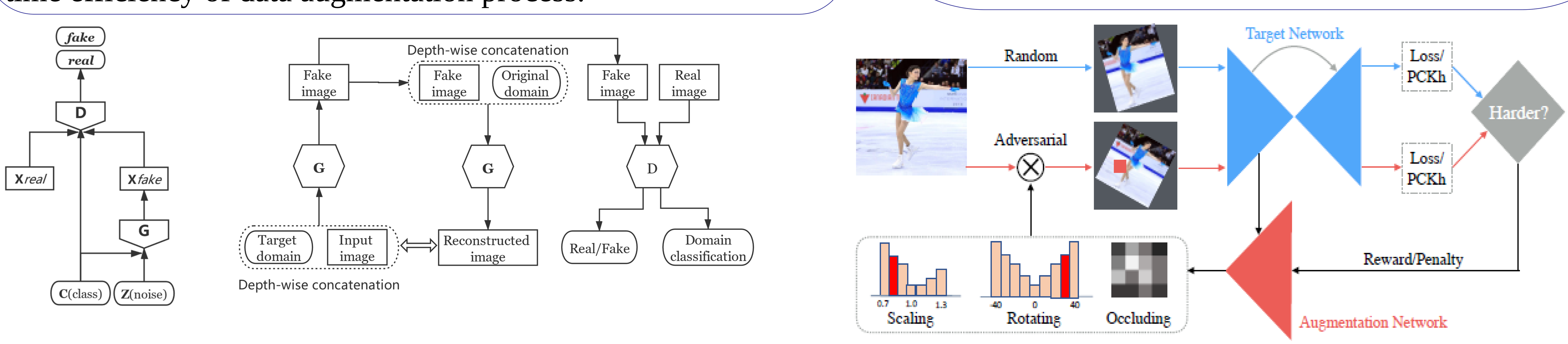
GAN-based data augmentation methods can be divided into three types according to their functions. The first is generating samples from latent space directly. CGAN has been used to extend training dataset by generating new samples which are in accordance with the original distribution. It has been used in medical field successfully. The second is image translation. StarGAN can increase the diversity of the dataset by transferring styles between different images. This method has been proven effective in person re-identification. The last produces augmentation operations utilizing adversarial strategy so as to improve the network performance[1].

Future work

Regardless of current defects, GAN-based data augmentation methods have great potential in extending and balancing dataset. For future research, I suggest a few directions. Firstly, mitigate relying on large training dataset. Secondly, introduce expert knowledge to direct data augmentation. Thirdly, jointly optimize data augmentation and target network. Lastly, improve time efficiency of data augmentation process.

Discussion

Firstly, there is a major problem that training GAN normally requires large amounts of data to learn the real distribution. This questions its practical value in data augmentation. To settle the in-sufficient data problem, semi-supervised and unsupervised conditional GANs have been explored such as CatGAN, InfoGAN and ss-InfoGAN. Secondly, the counterfeit samples generated by GAN can help improve model performance in classification, identification and so on. The performance of GAN also needs to be evaluated to avoid mode collapse, mode drop and overfitting. It still needs effort to find a better evaluation method. The training process of GAN is unstable and requires lots of time. The Nash equilibrium of the GAN game is hard to reach. Soumith Chintala [2] summarized several skills to enhance the stability of GAN training such as normalizing the inputs and using ADAM optimizer. The convergence issues of GAN need continuous attention in future.



(a) (b) (c)

Fig.1. (a)CGAN; (b)StarGAN; (c) An example of the approach proposed in [1].

References

1.Peng X., Tang Z., Yang F. and et al.: Jointly Optimize Data Augmentation and Network Training: Adversarial Data Augmentation in Human Pose Estimation. arXiv preprint arXiv: 1805.09707, 2018.
2.Soumith, How to Train a GAN? Tips and tricks to make GANs work, https://github.com/soumith/ganhacks, 2020/4/30