**数学科学学院研究生文献综述**

|  |  |  |
| --- | --- | --- |
| 姓名：何佳美 | 学号：201721100222 | 专业：计算机科学与技术 |
| 综述题目：  《基于生成对抗网络的电力设备图像扩充模型及算法研究》的文献综述 | | |
| 导师意见：  导师签字：  日期： | | |

**《基于生成对抗网络的电力设备图像扩充模型及算法研究》的文献综述**

***Abstract*-**With the rise and development of artificial intelligence, artificial intelligence has been applied to various fields. In 2017, the State Council issued the “Notice on the Development Planning of a New Generation of Artificial Intelligence”, which regards the development of artificial intelligence as a basic national strategy. The overall planning and guiding route for the development of artificial intelligence in various fields is proposed. At the same time, it is pointed out that it is necessary to vigorously develop artificial intelligence enterprises in various industries. The power equipment status detection and monitoring industry is no exception.

In response to the national master plan for artificial intelligence, the construction of the smart grid big data platform is rapidly advancing, and the core part includes the use of deep learning to achieve intelligent detection of power equipment defects. To implement the power device defect detection model by using the algorithm in deep learning requires a large amount of data sets. That is to say, the defect intelligent detection model with high accuracy requires a large number of data sets to train, verify and optimize the model, so that the model can have better generalization ability and accuracy, and then it can train the defect intelligence suitable for power equipment. The model being tested. At present, there are few image data sets for power equipment, which cannot meet the training of intelligent detection models for power equipment defects. The effect of the built-up intelligent detection model for power equipment defects is not good. If there is a good method of enhancing the data set for generating artificial samples, the accuracy of the power device defect detection model is improved. It is able to promote the construction of the smart grid big data platform, thereby promoting the development of artificial intelligence of power equipment. It is of great significance to the construction of China's smart grid and the maintenance and repair of power equipment status.

Therefore, there is a great need for a method of enhancing the data set to solve the problem of too few image data sets of the power device. Image enhancement techniques such as traditional geometry and intensity transformation can reconstruct the original image, but these reconstructed images are similar in nature to the original image, resulting in limited improvement of neural network performance, which can not effectively improve the accuracy of the defect intelligent detection model. rate. In order to solve the problem of insufficient data sets, Goodfellow et al. first proposed the Generative Adversarial Nets (GAN) in 2014, which LeCun called “the most exciting ideas in the field of machine learning in the past decade”.

The review provides a brief explanation of the power equipment detection technology; it also discusses and evaluates the classic models of generating confrontation in recent years and its application in various fields.

Generative Adversarial Networks (GANs) is an unsupervised learning method proposed by Goodfellow et al in 2014. GANs consist of two neural networks called generators and discriminators. GANs allows the generator and discriminator to learn in a mutual game: the generator randomly samples from a noise distribution as an input, and outputs an artificial sample very similar to the real sample in the training set; the input of the discriminator is a real sample or artificial The purpose of the sample is to distinguish the artificial sample from the real sample as much as possible. It can be seen that the generator should "fake" as much as possible to deceive the discriminator, and the discriminator should try to distinguish between "authentic" and "defective". During the training process, the two neural networks confront each other and constantly adjust their own parameters, and the abilities of both parties are improved. In the end, a generator with strong "falsification" ability is generated, which can generate artificial samples with false realities. GANs Currently, the main application areas are image, video, and text generation, such as generating fake images through GANs; predicting what the next frame of the video is; and generating text sequences in the field of natural language processing.

On the basis of the GANs, a number of variants of GANs have been derived. The Conditional Generative Adversalal Nets (CGAN) proposed by Mehdi Mirza et al. in 2014 solved the problem that the original GANs could not output the specified category samples by introducing the label information of the samples into the input; Alec Radford et al. Deep Convolutional Generative Adversarial Networks (DCGAN), proposed in 2015, combined with deep convolution methods to obtain a more stable training process and higher quality image samples; Jun-Yan Zhu et al. Cycle-Consistent Adversarial Networks (Cycle GAN) used a pair of GANs to train each other between two categories of data to complete the sample generation task with style migration; Arjovsky and Gulrajani et al. Two papers, for the unstable training of GANs, the lack of diversity in the generation of samples, attempted to analyze mathematically, and improved the Wasserstein GAN model.

GANs is a new initiative in deep learning in unsupervised learning. By its characteristics, it can be foreseen that GANs may play a role in the following application areas, such as generating maps from satellite photos (maps), generating color images from black and white images (old photos), and generating real photos from hand-painted images ), generating high resolution images (super resolution reconstruction) from low resolution images. However, due to the fact that GANs was not born for a long time, its architecture is still in the research stage. It still takes time to apply GANs technology in actual application scenarios.

However, as GANs research deepens, the resulting composite images become more and more realistic. In addition to the use of synthetic images for subjective evaluation, research into the use of GANs to generate artificial samples to augment data sets has emerged in recent years. Wang et al. analyzed the idea of using GANs to improve the accuracy of supervised learning, and reached a positive conclusion; Shrivastava et al. used GANs to re-optimize existing artificial samples on the MPIIGaze dataset, which improved the eyeball angle prediction and The accuracy of gesture recognition tasks. In addition, the performance evaluation between different GANs is also extremely important. Lucic et al. evaluated the original GANs and the derivative models of many GANs under the unified standard, and considered that the original GANs still have excellent performance compared with the current majority of the derived models. Ability to generate.

***Keywords***-Generative Adversarial Nets(GAN),Deep Convolutional Generative Adversarial Networks(DCGAN)

1. 前言

随着人工智能的兴起与发展，人工智能已经应用到各个领域中。2017年国务院印发了《新一代人工智能发展规划的通知》，将人工智能的发展作为基本国家战略。对各个领域的人工智能发展提出了总体规划及指导路线。同时指出要大力发展各个行业中的人工智能企业。电力设备状态检测、监测行业当然也不例外。

为了响应国家人工智能的总体规划，智能电网大数据平台的建设正在快速推进，其中核心部分就包括利用深度学习实现电力设备缺陷智能检测。要利用深度学习中的算法实现电力设备缺陷智检测模型需要大量的数据集。即要实现准确率高的缺陷智能检测模型需要大量的数据集对模型进行训练、验证、优化，才能够使得模型具有较好的泛化能力和准确率，才能够训练出适用于电力设备缺陷智能检测的模型。而目前电力设备图像数据集较少，不能够满足电力设备缺陷智能检测模型的训练。使得构建的电力设备缺陷智能检测模型的效果不好。如果有一种很好的增强数据集的方法用于生成人工样本，从而提高电力设备缺陷检测模型的准确率。就能够推进智能电网大数据平台的建设，从而促进电力设备人工智能化的发展。对于我国智能电网的建设以及电力设备状态检测维修具有重大意义。

因此，目前非常需要一种增强数据集的方法来解决电力设备图像数据集过少的问题。传统的几何和强度变换等图像增强技术能够重建原始的图像, 但是这些重建的图像在本质上和原始图像是相似的, 导致神经网络性能的提升有限, 不能够有效的提高缺陷智能检测模型的准确率。为了解决数据集不足的问题，Goodfellow等人在2014年第一次提出了生成对抗网络模型 (Generative Adversarial Nets, GAN) , 被LeCun称为“过去十年间机器学习领域最让人激动的点子”。

综述针对电力设备检测技术进行了简单的讲解；并讨论及评估了近几年来生成对抗的经典模型以及其在各个领域的应用。

**二.研究现状分析**

**1、电力设备检测研究现状**

随着计算机技术、数字化技术和图像识别技术的发展，无损检测技术被广泛的应用于电力设备检测中。无损检测技术是在不破坏监测对象的前提下进行对于检测对象的检测，检测内容是评价检测物体内部或者表面物理和机械性及各类缺陷和其他的技术参数[1]。同时，无损检测能够满足电力行业的高安全性和稳定性的要求，这使得其成为保证电力设备处于良好运行状态的技术之一。

目前而言，无损检测技术主要有超声检测、射线检测、声发射检测、红外检测、渗透检测、磁粉检测、涡流检测这几种，在电力设备检测中都有所涉及，但就其应用广泛性和发展前景而言，尤以超声检测、射线检测、声发射检测最为突出[2]。

由于无损检测技术在我国的应用时间还不是很长，并且电力设备检测的周期较长，使得目前采集的图像数据量较少，无法满足对于电力设备缺陷智能检测模型的训练。就目前而言，并未有对于电力设备图像数据集进行增强的相关研究。而在其他应用领域的基于生成对抗网络的图像生成技术发展得如火如荼。但是，对于各种生成对抗网络的应用目前还有一些问题有待解决。

**2、生成对抗网络研究现状**

机器学习方法包括两类：生成方法和判别方法，最后得到的模型称之为生成式模型（generative model）和判别式模型（discriminative model）[3]。生成方法是通过对样本数据进行学习从而得到基于样本与标签的联合概率分布。从而使得训练好的模型生成的新数据是与原始样本分布相符的。生成模型既可以是有监督的学习也可以是无监督的学习。其中无监督的学习是通过学习真实数据的本质特征，从而让模型掌握样本数据的分布特征，最后生成与原始数据高度相似的新数据。由于生成模型的参数比训练数据的量小好几个数量级，因此模型能够发现并有效内化数据的本质。生成式模型在无监督深度学习方面占据主要位置，在没有目标标签的情况下能自主的捕捉预测出对应数据的高阶相关性。深度生成模型可以通过从网络中采样来有效生成样本，近两年来流行的生成式模型主要分为三种方法：生成对抗网络（GAN）[4]，变分自动编码模型（VAE）[5][6]，自回归模型（Auto-regressive）[7]。其中，本文主要研究的是生成对抗网络（GAN）。

生成对抗网络（Generative Adversarial Networks, GANs）是由 Goodfellow 等人在 2014年提出的一种非监督式的学习方法[4]。GAN思想来源于博弈论，是由两个神经网络组成，分别包含一个生成模型（G）和一个判别模型（D）。GANs 让生成器和判别器以相互博弈的方式进行学习：生成模型学习样本的真实分布，从某种噪声分布中随机采样作为输入，生成于真实样本非常相似的人工样本。而判别模型则对输入的真实样本或人工样本进行判别，尽可能的将真实样本和人工样本区分出来。从中可以看出，生成模型的功能就是尽可能生成与真实数据类似的人工数据来欺骗判别器，而判别器则是尽力的从这些样本中将真实样本和人工样本区分开来。训练的过程中，是对两个模型的交替训练。两个神经网络相互对抗，在对抗的过程中不断调整自身的参数，双方的能力因此都得到上升。最终，产生了一个“造假”能力很强的生成模型，可以生成以假乱真的人工样本。GANs 目前主要的应用领域在图像、视频、文本生成方面，如通过 GANs 来生成以假乱真的图片[8]；预测视频的下一帧是什么[9]；在自然语言处理领域用以生成文本序列[10]。

由于GAN的学习模式太过于自由了，使得GAN的训练过程和训练结果很多时候都不太可控。为了稳定GAN，从启发式的、模型改进和理论分析的角度上后来都提出了许多训练技巧和改进方法。在基础的 GANs 上，根据实际的需求以及原始GAN的不足，衍生出了许多 GANs 的变种。

由于GAN不需要事先建模的方法太过自由，当数据集中图像的尺寸较大且包含复杂的内容时，使用简单的GAN很难控制生成人工样本的效果，对于样本的输出无法控制。Mehdi Mirza 等人于 2014 年提出的条件对抗网络（Conditional Generative Adversarial Nets, CGAN）[11]，通过在生成模型和判别模型中都引入额外的条件变量y，而这个条件变量y可以辅助引导人工样本的生成，条件变量y可以是类别标签、对图像修复有帮助的部分辅助数据等等。此时就是将GAN从纯无监督向有监督学习进行改进。从而改进了原始 GANs 对于人工样本输出类别无法指定的问题。

针对GAN训练过程中可能出现的不稳定的问题，Alec Radford 等人于2015年提出的深度卷积对抗网络（ Deep Convolutional Generative Adversarial Networks, DCGAN）[8]，将有监督学习的CNN[12]和无监督学习的GAN相结合，对GAN的生成模型和判别模型的架构进行修改，将深度卷积神经网络结构使用到GAN中，得到了更稳定的训练过程和更高质量的图像样本。但是该网络架构只是基于对生成模型和判别模型的架构进行不断实验，最终选出一种比较好的网络架构。只是从表面解决了训练不稳定的问题，并没有从原理上解决问题。

Arjovsky 与 Gulrajani 等人用了两篇论文，针对 GANs 存在的训练不稳定，生成样本缺乏多样性等通病，尝试从数学角度分析，并提出了改进后的Wasserstein GAN模型[13][14]。解决了GAN训练不稳定的问题，不需要再小心的衡量生成模型（G）和判别模型（D）的训练程度，而且最终生成的人工样本具有多样性。生成模型生成的人工样本的质量也较之前有所提高，但实验表明该方法的收敛速度较慢，同一数据集下需要多次训练才能收敛[15]。

除了上述几种对于GAN的衍生，还有其他很多方面的改进。比如，Jun-Yan Zhu 等人于2017年提出的循环一致性对抗网络（Cycle-Consistent Adversarial Networks, Cycle GAN）使用一对 GANs 在两个类别的数据之间相互训练，完成了带风格迁移效果的样本生成任务[16]； Patch GAN和Pixel GAN[17]结构用于图片到图片的转移，通过输入图片以及带条件的图片，使得生成网络最终能够生成非常接近于条件图片。Patch GAN和Pixel GAN改进网络中的目标函数，生成模型和判别模型的判别方式，将图像分成若干个小块进行判别，最终给出平均结果，相较于整张图像的判别更易于收敛[18]。

**3、生成对抗网络应用现状**

GANs是深度学习在无监督学习上一个新的创举。目前GAN可能在以下应用领域发挥作用。如由卫星照片生成地图（地图绘制）；由黑白图像生成彩色图像（老旧照片上色）；由手绘图片生成真实照片（嫌犯画像绘制）；由低分辨率图片生成高分辨率图片（超分辨率重建）[19]；医学图像合成、医学图像分割等[20]。但由于 GANs 诞生的时间不长，其架构等目前都尚处于研究阶段，想在实际的应用场景中运用 GANs 技术仍需时日。

随着GANs研究的深入，合成的人工样本图像也越来越真实。除了将合成图像用于主观性评价之外，近年来也逐渐出现了将 GANs 用于生成人工样本来扩充数据集的研究。Wang 等人对利用 GANs 提升监督学习准确率的想法进行了分析，并得出肯定结论[21]；Shrivastava 等人在 MPIIGaze 数据集上，利用GANs对已有的人工样本再进行优化，提高了眼球角度预测和手势识别任务的准确率[22]。Madani等人使用GAN来生成胸部X射线图像以增强数据集[23], 用于训练卷积神经网络, 进行心血管异常的分类。与传统数据增强方法相比, 用GAN进行增强能达到更高的分类准确度;Galbusera等人也基于所需解剖结构轮廓的简单图像, 利用GAN生成腰椎的平面X射线图像[24]。

此外，Lucic 等人对原始 GANs 和众多 GANs 的衍生模型在统一标准下进行了评价，认为原始 GANs 相比于当前多数衍生模型，仍具有优秀的生成能力[25]。从上述将GANs生成的人工样本用于扩充数据集的研究中，可以看到GANs对于生成人工样本用于扩充数据集，利用扩充后的数据集对于提高识别模型的准确率有巨大的潜力。

**总结**

随着智能电网大数据平台的推进，利用深度学习构建电力设备缺陷检测模型必不可少，而利用深度学习有关算法构建模型需要大量的数据。但是，目前电力设备图像数据集尚且不足。因此，目前非常需要一种方法来解决电力设备图像数据集不足的问题。

本文首先根据电力设备检测发展现状提出了电力设备图像数据集不足的问题以及解决这一问题的现实意义。针对这一问题，提出了用于解决数据集不足问题的生成对抗网络，然后给出了生成对抗网络的发展历程和应用现状。我们发现，生成对抗网络用于生成人工样本扩充数据集，利用扩充的数据集提高识别模型的准确率具有很大的潜力。因此，可以生成对抗网络对电力设备图像数据集进行扩增，并将扩增后的数据集用于电力设备的缺陷智能检测，以提高缺陷智能检测模型的准确率。这对于我国智能电网的建设以及电力设备状态检测维修具有重大意义和价值。

参考文献

1. 苑美实,骆令海.无损检测技术在电力系统中的应用[J].科学技术创新, 2018, (21):159-160.
2. 孙言蓓.电力设备无损检测技术研究[J].中国高新技术企业,2015,(18):29-30.
3. Wang L. Joint Probability Distribution[J]. 2013, 52(4):1063-1063.
4. Goodfellow I J, Pougetabadie J, Mirza M, et al. Generative Adversarial Networks[J].Advances in Neural Information Processing Systems, 2014, 3:2672-2680.
5. Kingma D P, Welling M. Auto-Encoding Variational Bayes[J]. 2014.
6. Rezende D J, Mohamed S, Wierstra D. Stochastic Backpropagation and Approximate Inference in Deep Generative Models[J]. Eprint Arxiv, 2014:1278-1286.
7. Chan A B, Vasconcelos N. Probabilistic Kernels for the Classification of Auto-Regressive Visual Processes[C]// IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2005:846-851.
8. Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks[J]. ar Xiv preprint ar Xiv:1511.06434,2015.
9. Mao X, Li Q, Xie H, et al. Least squares generative adversarial networks[C]. 2017 IEEE International Conference on Computer Vision (ICCV). IEEE, 2017: 2813-2821.
10. Lotter W, Kreiman G, Cox D. Unsupervised learning of visual structure using predictive generative networks[J]. ar Xiv preprint ar Xiv:1511.06380, 2015.
11. Mirza M, Osindero S. Conditional generative adversarial nets[J]. ar Xiv preprint ar Xiv:1411.1784, 2014.
12. Fukushima K.Neocognitron:A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position[J].Biological Cybernetics, 1980, 36 (4) :193-202.
13. Arjovsky M, Chintala S, Bottou L. Wasserstein gan[J]. ar Xiv preprint ar Xiv:1701.07875,2017.
14. Gulrajani I, Ahmed F, Arjovsky M, et al. Improved training of wasserstein gans[C]. Advances in Neural Information Processing Systems. 2017: 5769-5779.
15. 曹仰杰,贾丽丽,陈永霞,林楠,李学相.生成式对抗网络及其计算机视觉应用研究综述[J].中国图象图形学报,2018,23(10):1433-1449.
16. Zhu J Y, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[J]. ar Xiv preprint ar Xiv:1703.10593, 2017.
17. Isola P, Zhu J Y, Zhou T, et al. Image-to-Image Translation with Conditional Adversarial Networks[J]. 2016.
18. 肖旭.基于深度学习的图像语义分割研究[D].南昌航空大学,2017.
19. Miyato T, Koyama M. c GANs with projection discriminator[J]. ar Xiv preprint ar Xiv:1802.05637, 2018.
20. 陈锟,乔沁,宋志坚.生成对抗网络在医学图像处理中的应用[J].生命科学仪器,2018,16(Z1):71-80+91.
21. Wang J, Perez L. The effectiveness of data augmentation in image classification using deep learning[R]. Technical report, 2017.
22. Shrivastava A, Pfister T, Tuzel O, et al. Learning from simulated and unsupervised images through adversarial training[C]. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017, 3(4): 6.
23. Madani A, Moradi M, Karargyris A, et al.Chest x-ray generation and data augmentation for cardiovascular abnormality classification[C].SPIEMedical Imaging, 2018:6.
24. Galbusera F, Niemeyer F, Seyfried M, et al.Exploring the Potential of Generative Adversarial Networks for Synthesizing Radiological Images of the Spine to be Used in In Silico Trials.Frontiers in bioengineering and biotechnology.2018:53.
25. Lucic M, Kurach K, Michalski M, et al. Are GANs Created Equal? A Large-Scale Study[J]. ar Xiv preprint ar Xiv:1711.10337, 2017.