

中增加一个随机的、能隐藏的补丁,根据上下文信息对图像进行修复. Liu 等^[29]使用部分卷积进行图像修复,修复效果如图5所示.



图5 图像修复效果^[29]

3.1.5 视频生成和预测

在生成图像的基础上, GAN 还可以生成视频. Tukyakov 等^[30]提出一种通过分解动作和内容来生成视频的方法—MoCoGAN, 将一个随机向量的序列映射到一个视频帧序列上, 每个随机向量包含一个动作和一个内容. 视频预测是根据当前的一帧或几帧视频来预测视频接下来的内容, 比如当一个人拿起一个球的时候预测这个人接下来要干什么. Xiong 等^[31]提出一种多阶段动态生成对抗网络来生成延时视频, 首先为每一帧生成现实内容的视频, 然后进行运动建模, 使相邻帧之间的物体运动更加生动, 同时保持内容逼真. 另一种视频预测方法是 Jang 等^[32]提出的方法—通过外观和动作来指定未来, 减少不确定性, 解决了当有多个正确的、等可能的未来时, 模型不知如何选择的问题.

3.2 文本生成领域

除了生成图片之外, GAN 还能生成文本. Zhang 等^[33]用 GAN 来生成文本, 其判别器为卷积神经网络 (CNN). GAN 一般是基于连续空间, 但也可以应用于离散数据. 但是 GAN 用于生成序列存在两个问题, GAN 生成的是连续数据, 难的是直接生成离散序列; 另一个问题是, GAN 只能对整个生成序列打分, 而对于一部分生成的序列, 很难判断它现在生成序列的质量和之后生成整个序列的质量. SeqGAN^[34]能解决上述两个问题, 它结合强化学习, 能生成离散的序列, 使用策略梯度训练生成器 G 来解决输出离散数据时梯度不能返回生成模型的问题, 通过蒙特卡洛搜索得到奖励信号. WGAN 也能推广应用于离散数据空间, 在基于 WGAN 的基础上, 代威^[35]提出一种使用字符级文本构造方法的无监督文本生成方法 Lable GAN, 能适用于多种序列处理

任务, 并且能有效的干预和约束生成数据空间. SeqGAN 可以应用于语音生成、诗词生成、机器翻译、对话生成等. MaskGAN^[36]根据上下文内容来填充缺失的文本, 并引入 actor-critic 架构.

此外, GAN 还可以根据文字描述来生成图文匹配的图像. StackGAN^[37]能够根据输入文本描述生成细化的高分辨率图像. AttnGAN^[38]可以生成更细腻的图像, 通过关注自然语言描述中的相关词汇, 可以在图像的不同子区域合成出精细的细节, 此外, 还提出了一种深度注意多模态相似度模型, 用于计算用于训练生成器的细粒度图像-文本匹配损失.

3.3 其他领域

GAN 除了在视觉领域、文本生成领域外, 还在其他领域有涉及. GAN 能应用于密码破译, CipherGAN^[39]提出了一个能够用于破译密码的框架, 能够高度准确地破译使用移位和 Vigenere 密码的语言数据, 且能支持更多的词汇. GAN 还可以进行行人重识别、重定向等. GAN 除了能生成图像外, 还能生成音乐、3D 物体等.

4 存在的问题及发展趋势

4.1 存在的问题

GAN 模型是一种框架, 可以与其他方法相结合. GAN 模型包含生成器 G 和判别器 D, 在训练过程中需要保持生成器 G 和判别器 D 之间的同步, 因此容易造成训练不稳定. 虽然 WGAN 使用 EM 距离代替 JS 散度, 不需要保持生成器 G 和判别器 D 之间的同步, 解决了训练不稳定的问题, 但是, 在使用 JS 散度的模型的训练稳定性问题还是没有得到很好的解决. 此外, GAN 还存在模式崩溃问题. WGAN 能解决模式崩溃问题, 但是, 如果损失函数的参数设置不合理的话会导致梯度消失或梯度爆炸. 因此, 模式崩溃问题还没有彻底的解决.

同时, GAN 也面临着一个如何评估生成样本质量的问题, 也就是说 GAN 缺乏客观的评估方法. 现有对 GAN 的评价方法都是基于样本层面, 对生成样本与真实样本提取特征, 在特征空间做距离度量. Xu 等^[40]实验结果表明, 在基于样本的评价方法中, 核函数最大平均偏差 (Kernel MMD) 和 1-近邻 (1-NN) 双样本检验方法是最合适的评价方法, 这两种评价指标能够有效的区别生成样本和真实样本, 同时也能识别模式崩溃的

模式下降问题,检测过度拟合,且具有高效性。但是,对于生成效果和质量,却只能依赖于视觉上的观察,由人眼去判断生成的质量好坏。有些评价指标在数值上虽高,但是实际的生成效果却不好。因此,GAN 很难找到一个比较客观的、可量化的评价指标。

4.2 发展趋势

自从 GAN 被提出以来就受到广泛的关注,近两年来对 GAN 的研究热度也是只增不减。虽然 GAN 能实现的功能巨大,但是这个模型存在的一些问题未能彻底解决,影响其生成效果。在模型的训练稳定性和评价指标方面,仍有待研究改进。GAN 在图像领域有着卓越的应用效果,在其他领域也有着不错的应用。GAN 对于文本方面的应用是基于离散数据,效果不是很理想,在这一方面的改进空间将会很大。随着人工智能的不断发展,对无人驾驶的研究也在不断成熟,GAN 可以结合其对图像领域的生成效果,应用于无人驾驶,这也将会是很热门的一个研究方向。视频与语音在日常生活中是必需的,随着人们需求的增大,GAN 在视频与语音方面的应用也更需要去研究改进。GAN 也可以应用于图像去噪方面,现在也有一些突破性的进展。

4.2.1 图像去噪

人们对图像的反映是通过视觉直观的体现出来的,图像质量差会影响人们对图像细节的掌握,影响图像的视觉效果。噪声是图像干扰的重要原因,实际上一幅图可能会存在各种各样的噪声,所以图像噪声在视觉图像中影响很大。图像噪声一般产生于图像的采集或者传输过程中,比如使用的设备不一样,在传输过程中有压缩等等。

传统的图像去噪最初采用的是使用像素尺度的空间域滤波方法^[41],该方法中的中值滤波和均值滤波都是使用图像信号和滤波模板结合起来进行滤波。空间域滤波方法在使用简单的卷积核的时候容易导致图像信息丢失,针对这一缺点,一些学者提出基于变换域的去噪方法,包含傅里叶变换域,小波域,脊波域等等^[42,43]。随着深度学习的发展,深度卷积网络在图像去噪的应用有着不错的效果。去噪卷积网络(DnCNN)^[44]第一次结合残差网络应用于图像去噪领域,使得去噪效果大大提升。Wang^[45]采用空洞卷积代替卷积核,提高了训练运行速度,效果不亚于去噪卷积网络。

4.2.2 结合 GAN 的图像去噪

现有的去噪方法大多是基于已知噪声,例如高斯

噪声,GAN 可以学习复杂的分布,可以在 GAN 的生成器中输入含有已知噪声的图片,然后判别器输出对应无噪声的图像。张元祺^[46]结合生成对抗网络,对彩色图像进行去噪,通过神经网络将含噪图像变换到特征域进行处理,并利用多尺度特征保留图像丰富的纹理细节。

那么对于图像里包含的未知噪声信息,即对图像进行盲去噪,GAN 拥有强大的学习能力,可以学习复杂的分布,可以使用 GAN 来构建对无噪-有噪图像的训练数据集。在有未知噪声图像的情况下,一种建立成对训练数据集的方法是训练 GAN 的生成网络,学习从无噪图像到相似噪声图像的映射。生成网络可以训练学习真实噪声图像的分布,生成具有相似噪声的图像,但不能保证原始图像的纹理细节等不改变。因此,Chen 等^[47]提出一种盲去噪方法——使用 GAN 来对未知噪声图像进行噪声提取并建模,生成网络生成类似的是噪声而不是噪声图像,生成相似分布的噪声,然后使用深度卷积网络进行图像去噪。

对于噪声已知的图像,可以直接作为生成器的输入,判别器进行去噪,输出对应的无噪图像,在这个过程中,需要保留图像原始的内容,即不能损失图像的纹理细节等。如何在去噪的过程中保留图像内容也值得去研究。对于图像盲去噪,可以通过 GAN 来建模获得相似噪声分布,然后再使用深度卷积网络或者其他网络进行去噪。在这个过程中,需要使用两个或两个以上模型,如何使用 GAN 模型进行图像的盲去噪将会是一个值得研究的方向。

5 结论

生成对抗性网络 GAN 包含生成器和判别器,有很强的生成能力,在不进行限制的情况下,具有无限建模的能力,主要应用于图像视觉领域。现今 GAN 能够产生超分辨率图像,能够结合半监督、强化学习、特征学习等,在图像、视觉、文本方面有广泛应用。GAN 在图像领域的应用效果显著,虽然在其他领域也应用广泛,但实现的效果并没有图像的好。GAN 如何更好的应用于更多领域,在应用的同时不断优化性能,并发挥更大的优势,将会是有趣的研究。随着人工智能的发展,人们追求的是更加智能化,如何结合深度学习等不同技术来提高 GAN 的应用效果,使 GAN 更智能化,是 GAN 未来发展需要提升的。GAN 可以和图像去噪相结合,利用 GAN 的强学习能力,实现噪声已知的图

像去噪,也可以建模生成相似的忙噪声分布,有助于图像盲去噪。GAN对于图像去噪的应用将有待于更进一步的研究。

参考文献

- 1 Goodfellow IJ, Pouget-Abadie J, Mirza M, *et al.* Generative adversarial nets. Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal, QC, Canada. 2014. 2672–2680.
- 2 Denton E, Gross S, Fergus R. Semi-supervised learning with context-conditional generative adversarial networks. arXiv preprint arXiv: 1611.06430, 2014.
- 3 Zhu JY, Park T, Isola P, *et al.* Unpaired image-to-image translation using cycle-consistent adversarial networks. Proceedings of 2017 International Conference on Computer Vision. Venice, Italy. 2017. 2223–2232.
- 4 Hindupur A. The-gan-zoo. <https://github.com/hindupuravinash/the-gan-zoo>.
- 5 Goodfellow I. NIPS 2016 tutorial: Generative adversarial networks. arXiv preprint arXiv: 1701.00160, 2016.
- 6 Hjelm RD, Jacob AP, Che T, *et al.* Boundary-seeking generative adversarial networks. arXiv preprint arXiv: 1702.08431v4, 2018.
- 7 Che T, Li YR, Zhang RX, *et al.* Maximum-likelihood augmented discrete generative adversarial networks. arXiv preprint arXiv: 1702.07983, 2017.
- 8 Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv: 1511.06434, 2016.
- 9 Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe, NV, USA. 2012. 1097–1105.
- 10 Arjovsky M, Chintala S, Bottou L. Wasserstein GAN. arXiv preprint arXiv: 1701.07875, 2017.
- 11 Gulrajani I, Ahmed F, Arjovsky M, *et al.* Improved training of Wasserstein GANs. Proceedings of the 30th Neural Information Processing Systems. Long Beach, CA, USA. 2017. 5767–5777.
- 12 Cui SB, Jiang Y. Effective Lipschitz constraint enforcement for Wasserstein GAN training. Proceedings of the 2nd IEEE International Conference on Computational Intelligence and Applications. Beijing, China. 2017. 74–78.
- 13 Qi GJ. Loss-sensitive generative adversarial networks on Lipschitz densities. arXiv preprint arXiv: 1701.06264, 2017.
- 14 Qi GJ. Generalized LS-GAN (GLS-GAN): Now LS-GAN and WGAN are special cases of this supermodel. <https://zhuanlan.zhihu.com/p/25580027>.
- 15 Chen X, Duan Y, Houthoofd R, *et al.* InfoGAN: Interpretable representation learning by information maximizing generative adversarial nets. Proceedings of 2016 Neural Information Processing Systems. Barcelona, Spain. 2016. 2172–2180.
- 16 Odena A. Semi-supervised learning with generative adversarial networks. arXiv preprint arXiv: 1606.01583, 2016.
- 17 Springenberg JT. Unsupervised and semi-supervised learning with categorical generative adversarial networks. arXiv preprint arXiv: 1511.06390, 2015.
- 18 Mirza M, Osindero S. Conditional generative adversarial nets. arXiv preprint arXiv: 1411.1784, 2014.
- 19 Donahue J, Krähenbühl P, Darrell T. Adversarial feature learning. arXiv preprint arXiv: 1605.09782, 2016.
- 20 Ledig C, Theis L, Huszár F, *et al.* Photo-realistic single image super-resolution using a generative adversarial network. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA. 2017. 105–114.
- 21 He KM, Zhang XY, Ren SQ, *et al.* Deep residual learning for image recognition. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. 2016. 770–778.
- 22 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv: 1409.1556, 2014.
- 23 Karras T, Aila T, Laine S, *et al.* Progressive growing of GANs for improved quality, stability, and variation. arXiv preprint arXiv: 1710.10196v3, 2018.
- 24 Wang YF, Perazzi F, McWilliams B, *et al.* A fully progressive approach to single-image super-resolution. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Salt Lake City, UT, USA. 2018. 977–986.
- 25 Isola P, Zhu JY, Zhou TH, *et al.* Image-to-image translation with conditional adversarial networks. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA. 2017. 5967–5976.
- 26 Wang TC, Ming YL, Zhu JY, *et al.* High-resolution image synthesis and semantic manipulation with conditional GANs. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA.

2018. 8798–8807.
- 27 Li C, Wand M. Precomputed real-time texture synthesis with Markovian generative adversarial networks. Proceedings of the 14th European Conference on Computer Vision. Amsterdam, The Netherlands. 2016. 702–716.
- 28 Azadi S, Fisher M, Kim V, *et al.* Multi-content GAN for few-shot font style transfer. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. 2018. 7564–7573.
- 29 Liu GL, Reda FA, Shih KJ, *et al.* Image inpainting for irregular holes using partial convolutions. Proceedings of the 15th European Conference on Computer Vision. Munich, Germany. 2018. 89–105.
- 30 Tulyakov S, Liu MY, Yang XD, *et al.* MoCoGAN: Decomposing motion and content for video generation. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. 2018. 1526–1535.
- 31 Xiong W, Luo WH, Ma L, *et al.* Learning to generate time-lapse videos using multi-stage dynamic generative adversarial networks. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. 2018. 2364–2373.
- 32 Jang Y, Kim G, Song YL. Video prediction with appearance and motion conditions. Proceedings of the 35th International Conference on Machine Learning. Stockholm, Sweden. 2018. 3496–3510.
- 33 Zhang YZ, Gan Z, Carin L. Generating text via adversarial training. Proceedings of the 2016 Conference on Neural Information Processing Systems. Barcelona, Spain. 2016. 1–6.
- 34 Yu LT, Zhang WN, Wang J, *et al.* SeqGAN: Sequence generative adversarial nets with policy gradient. Proceedings of the 31st AAAI Conference on Artificial Intelligence. Phoenix, AZ, USA. 2016. 2852–2858.
- 35 代威. 基于生成式对抗网络的文本生成问题的研究[硕士学位论文]. 长春: 吉林大学, 2018.
- 36 Fedus W, Goodfellow I, Dai AM. MaskGAN: Better text generation via filling in the _____. arXiv preprint arXiv: 1801.07736, 2018.
- 37 Zhang H, Xu T, Li HS, *et al.* StackGAN: Text to photo-realistic image synthesis with stacked generative adversarial networks. Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy. 2017. 5908–5916.
- 38 Xu T, Zhang PC, Huang QY, *et al.* AttnGAN: Fine-grained text to image generation with attentional generative adversarial networks. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. 2018. 1316–1324.
- 39 Gomez AN, Huang SC, Zhang I, *et al.* Unsupervised cipher cracking using discrete GANs. arXiv preprint arXiv: 1801.04883, 2018.
- 40 Xu QT, Huang G, Yuan Y, *et al.* An empirical study on evaluation metrics of generative adversarial networks. arXiv preprint arXiv: 1806.07755, 2018.
- 41 Liu XW, Huang LH, Guo ZY. Adaptive fourth-order partial differential equation filter for image denoising. Applied Mathematics Letters, 2011, 24(8): 1282–1288. [doi: [10.1016/j.aml.2011.01.028](https://doi.org/10.1016/j.aml.2011.01.028)]
- 42 Sendur L, Selesnick IW. Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency. IEEE Transactions on Signal Processing, 2002, 50(11): 2744–2756. [doi: [10.1109/TSP.2002.804091](https://doi.org/10.1109/TSP.2002.804091)]
- 43 Luisier F, Blu T, Unser M. A new SURE approach to image denoising: Interscale orthonormal wavelet thresholding. IEEE Transactions on Image Processing, 2007, 16(3): 593–606. [doi: [10.1109/TIP.2007.891064](https://doi.org/10.1109/TIP.2007.891064)]
- 44 Zhang K, Zuo WM, Chen YJ, *et al.* Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. IEEE Transactions on Image Processing, 2017, 26(7): 3142–3155. [doi: [10.1109/TIP.2017.2662206](https://doi.org/10.1109/TIP.2017.2662206)]
- 45 Wang TY, Sun MX, Hu KN. Dilated deep residual network for image Denoising. arXiv preprint arXiv: 1708.05473, 2017.
- 46 张元祺. 基于生成对抗网络的彩色图像去噪方法[硕士学位论文]. 大连: 大连理工大学, 2018.
- 47 Chen JW, Chen JW, Chao HY, *et al.* Image blind denoising with generative adversarial network based noise modeling. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. 2018. 3155–3164.