# 超分辨率重建和GAN成像

在记录一张图像时，往往会因为成像系统本身的特性而导致空间分辨率的损失，比如光学扭曲，模糊和噪声。通过改进光学硬件比如传感器，可以提高图像分辨率，但是这种办法的代价高，且制作工艺难以改进，因此，使用信号处理技术来从低分辨率图像中获得高分辨率图像成为了研究热点，这种方法被称为超分辨率重建。

超分辨率重建中的一个重要概念叫做降质模型，它描述了自然界中的高分辨率图像转换成人眼观测到的低分辨率图像的整个过程，即高分辨率图像成像逆过程，为图像超分辨率技术提供了坚实的理论基础。该模型认为，低分辨率图像的形成过程主要由以下三个因素共同影响：运动变换、成像模糊和降采样。

根据重建算法的不同，超分辨率重建技术可分为以下三种：

1. 基于插值的超分辨率重建。此方法是最为直观的一种方法，它通过利用已有的像素点信息，拟合未知的像素点，从而达到提高分辨率的目的。Ur和Gross[1]提出了一种不均匀插值的方法，从一组空间上移位的低分辨率图像中获取一张高分辨率图像。该方法包括三个步骤：配准，插值和恢复。配准指的是运动信息的估计，不均匀插值用来从非均匀间隔的低分辨率图像中获得均匀间隔的高分辨率图像，恢复指的是移除模糊和噪声。该方法利用了Papoulis和Brown的广义多通道采样定理，而其中的线性通道对应于产生低分辨率图像（通过模糊和移位）的过程以及重建过程。Hardie等人[2]利用基于梯度的配准算法来估计所采集帧之间的偏移，并使用加权最近邻插值方法来产生一张高分辨率图像，最后使用维纳滤波来减少系统引起的模糊和噪声的影响。基于插值的方法的优点在于计算负载比较低，适合用在实时应用中。但是在这种方法中，考虑到的降质模型有限，仅仅适用于当整组低分辨率图像的模糊和噪声特性相同时，另外此方法也不能保证最优的重建算法，因为在图像恢复时忽略了插值阶段可能发生的错误。

基于重构的超分辨率重建。多图像（或经典）超分辨率算法主要是基于重构的算法，即它们试图通过模拟图像形成模型来解决由于欠采样过程而在观察到的低分辨率图像中存在的混叠伪像。此类算法从图像的降质模型出发，假定高分辨率图像是经过了适当的运动变换、模糊及噪声才得到低分辨率图像。这种方法通过提取低分辨率图像中的关键信息，并结合对未知的超分辨率图像的先验知识来约束超分辨率图像的生成。

迭代反投影法。MICHAL IRANI等人[3]从对目标高分辨率图像最初的假设出发，迭代地对此假设进行细化。这个最初的假设可以是整组低分辨率图像的平均。通过在固定的更精细网格上配准所有低分辨率图像来计算平均图像，精细网格中的每个高分辨率像素为堆叠在其上的所有低分辨率像素的平均值。之后通过不断模拟一组低分辨率图像，使得它们越来越接近观测到的那组低分辨率图像，计算得到两组图像的差异图像，将差异图像的每一个像素值反向投影到最初的假设即目标高分辨率图像中对应的感受野，以此来更新目标高分辨率图像。整个过程迭代进行，不断减少两组低分辨率图像之间的差异。

上述迭代反投影法的主要问题是最终迭代的响应可能收敛到一个解决方案，也可能在几个可行的解决方案之间振荡。这可以通过结合对未知的超分辨率图像的先验知识来约束超分辨率图像的生成。这是通过对进行优化的损失函数增加一个正则项来实现的，正则化项利用关于期望的高分辨率解决方案的一些先验信息来补偿丢失的测量信息。Sina Farsiu等人[4]使用了总变差方法（TV）作为正则化方法，并结合双边滤波器，提出了一种强大的正则化器，称为双边TV，它实现起来计算成本低，并且在重建的过程中保留了边。

凸集投影法。每张低分辨率图像都为目标高分辨率图像加上了一个先验知识，这个先验知识由一个闭凸集表示，因此K张低分辨率图像就对应有K个凸集。在迭代更新目标高分辨率图像时，都将本轮迭代的目标图像投影到每一张低分辨率图像的凸集上，投影之后的结果即是本轮更新之后的目标图像。如果一些LR图像遭受部分遮挡，或者如果他们的运动矢量估计不准确，则目标高分辨率图像将是错误的。P. Erhan Eren等人[5]使用凸集投影法，并采用有效性图和（或）分割图来解决上述问题。有效性图禁用基于具有不准确运动信息的观测的投影，以便在存在运动估计误差的情况下进行鲁棒的重建; 而分割图使得基于对象的处理成为可能，其中可以利用更精确的运动模型来改善重建图像的质量。

1. 基于学习的超分辨率重建。

以上介绍的两种算法，基于插值的和基于重构的，解决的都是多图超分辨率问题（MISR），即利用低分辨率图像序列（视频）还原一张高分辨率图像。基于多图像的超分辨率算法的准确性高度依赖于低分辨率观测图像之间运动的估计精度，但是运动估计在实际应用中变得更不稳定，因为同一场景中的不同对象可具有不同且复杂的运动，在这种情况下，基于单图像的超分辨率算法可能会取得更好的结果。针对单图超分辨率问题（SISR），基于学习的超分辨率重建算法取得了很好的成果。基于学习的方法是利用大量的训练数据，从中学习低分辨率图像和高分辨率图像之间某种对应关系，然后根据学习到的映射关系来预测低分辨率图像所对应的高分辨率图像，从而实现图像的超分辨率重建过程。

稀疏编码。稀疏信号表示（sparse signal representation）表明高分辨率信号之间的线性关系可以从它们的低分辨率投影中恢复，那么高分辨率图像也可以由低分辨率图像恢复。综合使用双字典学习和信号的稀疏表示，可以重建出高分辨率图像。杨建超老师[6]在2010年发表的”Image Super-Resolution Via Sparse Representation”一文中提出，图像可以表示为过完备字典中元素的稀疏线性组合，通过联合训练高分辨率图像块的字典和低分辨率图像块的字典，我们可以加强相对于它们各自的字典的低分辨率和高分辨率图像块对之间稀疏表示的相似性。这样，低分辨率图像的每个图像块在一个低分辨率超完备字典上的一组稀疏表示系数可以直接用在给定的高分辨率超完备字典上进行高分辨率图像重建。此方法的性能优秀，产生的高分辨率图像具有清晰的边和清晰的纹理，并且由于稀疏表示对于噪声的鲁棒性，此方法对噪声的鲁棒性更强。

深度学习方法。深度学习利用多层非线性变换提取数据的高层抽象特征。利用深度学习方法进行超分辨率重建的主要思想是用具有合适泛化能力的样本数据对深层神经网络进行训练，学习低分辨率图像跟与之对应的高分辨率图像之间的关系，从而进行重建。2014年，董超等人[7]首次将深度学习应用到图像超分辨率重建领域，他们使用一个三层的卷积神经网络学习低分辨率图像与高分辨率图像之间的映射关系，此模型被称为SRCNN。作者简单说明了传统的基于稀疏编码的超分辨率重建方法可以看成一种深度CNN结构。作者利用了一个CNN结构来进行超分辨率重建的三个重要步骤：块提取和表示，非线性映射，重建。此网络结构虽然简单，但相比当前最先进的方法其准确率更高；由于层数少，此方法的运行速度块；此网络结构是端到端的，仅有少量的预处理和后处理。可以通过采用更多的隐含层和不同的训练策略来提高性能。

2016年，董超等人[8]继续对之前提出来的SRCNN模型进行改进。首先，在非线性映射层，作者使用了更大的过滤器，增加更多的非线性映射层加深网络结构；接着，作者将此CNN模型进行扩展，使之适用于同时处理三颜色通道，并且从实验上证实了其效果优于单通道网络；论文中进行了更多的理论分析，详细从数学上介绍了块提取和表示，非线性映射，重建其实都是卷积的过程；此外，论文中还将本模型与许多其它的模型进行对比，说明了本模型在不同衡量标准下都有优异的性能。

同样在2016年，施闻哲等人[9]提出了一种新的CNN结构：先用CNN从低分辨率图像上提取特征，再用亚像素卷积层得到目标高分辨率图像，称之为ESPCN。本方法能够实时地处理图像和视频的超分辨率重建。他们认为，之前的基于DNN的超分辨率重建方法都需要在重建之前将低分辨率图像利用双立方插值等方法将其放大至目标高分辨率图像的大小，这会增加计算的复杂性。因此在本文中，作者通过一个有效的亚像素卷积层训练学习一组放大过滤器，使用这组过滤器来将最后的低分辨率特征图放大成高分辨率图像。这大大减少了计算复杂性，降低计算量提高速度。

2017年，来自推特公司的Christian Ledig等人[10]将流行的GAN结构用在超分辨率重建上，称之为SRGAN。先前许多利用CNN进行超分辨率重建的网络模型都采用均方误差（MSE）作为优化目标，但是MSE并不能很好地恢复纹理细节，导致最终生成的高分辨率图像过于平滑。作者提出了一种感知损失函数，由adversarial loss和content loss加权组成，其实就是分别利用了GAN的鉴别器网络和生成器网络。值得一提的是，本模型的生成器网络吸收了几篇优秀论文的方法，其中就包括ESPCN的亚像素卷积层。本方法最终得到的高分辨率图像具有逼真的纹理。作者进行了MOS实验，实验表明本方法得到的高分辨率图像具有更高的感知质量。

2014年，Ian J. Goodfellow等人[11]创造了生成式对抗网络（GAN），加入了一个对抗的过程来提高生成器的生成能力。GAN由两个模型组成：生成器模型和鉴别器模型，两个模型同时训练，生成器模型负责捕捉数据的分布，鉴别器模型估计一个样本来自训练数据而不是生成器的概率。作者提出了一个训练算法，并证明了该算法的收敛性。此原始GAN的缺点主要有三个方面：难以收敛，难以训练，包括梯度消失和模式崩溃，模型过于自由不可控。

由于GAN的不稳定性，并且训练时需要大量的技巧，在2016年，Radford等人[12]提出了GAN架构的升级版，称之为DCGAN模型，该模型在大多数情况下比较稳定，文中提出了稳定GAN的一些架构指导原则，在对GAN进行优化设计时遵循这些原则便可以使得GAN较为稳定。这些指导原则如下：

1.在鉴别器中用带步长的卷积代替池化，在生成器中用反卷积代替池化。

2.在鉴别器和生成器中均使用Batch Normalization。

3.为了能设计更深的网络结构，去除全连接隐藏层。

4.在生成器中除了输出层之外全部使用ReLU激活函数，输出层使用Tanh激活函数。

5.在鉴别器中所有层均使用LeakyReLU激活函数。

GAN的发展很快，有很多的变体。针对模型过于自由不可控的问题，Mehdi Mirza和Simon Osindero[13]提出了条件生成对抗网络CGAN，能够提高生成图像的质量，明确控制图像的某些方面。原始 GAN 采用的是 JS 散度来衡量两个分布之间的距离，Sebastian Nowozin等人[14]提出的f-GAN给出了通过一般的f散度来构造一般的GAN的方案，它一般化的囊括了很多GAN变种（包括原始GAN），并且可以启发我们快速地构建新的GAN变种。在所有基于f-GAN的GAN中，毛旭东等人[15]提出的LSGAN简单而且性能好，因此是最常用的。使用Wasserstein距离替换散度，理论上可以解决梯度消失问题。Martin Arjovsky等人[16]提出的WGAN既彻底解决了训练不稳定的问题，基本解决了模式崩溃的问题，也提供了一个可靠的训练进程指标，而且该指标确实与生成样本的质量高度相关。GAN的应用也非常多，作为无监督学习的代表，GAN同样也可以用在生成领域，半监督学习领域以及强化学习领域。SRGAN[10]用于超分辨率重建。Phillip Isola等人[17]将CGAN用于图像转换。Tim Salimans等人[18]将SSL-GAN用于半监督学习领域。Yuan Xue等人[19]尝试将SegAN用于医学图像分割。

Yongjun Hong等人[20]发表了一篇最新的GANs综述论文，从各个角度探讨GAN的细节，还解释了GAN如何运作以及最近提出的各种目标函数的基本含义，介绍了适用于各种任务和其他领域的GAN变体。David Bau等人[21]提出了一个分析框架来在不同的抽象级别上可视化和理解GAN，此项工作为理解GAN的内部表示第一次提供了系统分析。Chaoyue Wang等人[22]提出了基于进化算法的替代GAN框架（E-GAN），与现有的GAN采用预定义的对抗性目标函数交替训练生成器和鉴别器不同，作者利用不同的对抗训练目标作为变异操作，并演化出一群适应鉴别器环境的生成器，还利用评估机制来测量生成样本的质量和多样性，以便仅保留性能良好的生成器并用于进一步训练。E-GAN减少了训练难度，提高了生成器的性能，实验证明E-GAN提高了GAN模型的稳定性，在许多图像生成任务上性能优异。

### 参考文献

1. H.Ur and D.Gross, “Improved resolution from sub-pixel shifted pictures,” CVGIP: Graphical Models and Image Processing, vol. 54, pp. 181-186, Mar. 1992.
2. M.S. Alam, J.G. Bognar, R.C. Hardie, and B.J. Yasuda, “Infrared image registration and high-resolution reconstruction using multiple translationally shifted aliased video frames,” IEEE Trans. Instrum. Meas., vol. 49, pp. 915-923, Oct. 2000.
3. Irani, M., Peleg, S.: Super-resolution from image sequences. In: Proceedings of IEEE International Conference on Pattern Recognition, USA, pp. 115–120 (1990).
4. Farsiu, S., Robinson, D., Elad, M., Milanfar, P.: Fast and robust super-resolution. Proceedings of IEEE International Conference on Image Processing, Spain 2, 291–294 (2003)
5. Eren, P.E., Sezan, M.I., Tekalp, A.M.: Robust object-based highresolution image reconstruction from low-resolution video. IEEE Trans. Image Process. 6(10), 1446–1451 (1997)
6. Yang, J., Wright, J., Huang, T., Ma, Y.: Image super-resolution via sparse representation. TIP (2010)
7. C. Dong, C. C. Loy, K. He, and X. Tang, “Learning a deep convolutional network for image super-resolution,” in Proc. Eur. Conf. Comput. Vis., 2014, pp. 184–199.
8. C. Dong, C. C. Loy, K. He, and X. Tang. Image super-resolution using deep convolutional networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(2):295–307, 2016.
9. W. Shi, J. Caballero, F. Huszar, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang. Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 1874–1883, 2016.
10. C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. arXiv preprint arXiv:1609.04802, 2016.
11. I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In Advances in Neural Information Processing Systems (NIPS), pages 2672–2680, 2014.
12. Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434, 2015.
13. Mehdi Mirza and Simon Osindero. Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784, 2014.
14. Sebastian Nowozin, Botond Cseke, and Ryota Tomioka. f-gan: Training generative neural samplers using variational divergence minimization. In Advances in Neural Information Processing Systems, pages 271–279, 2016.
15. Xudong Mao, Qing Li, Haoran Xie, Raymond YK Lau, Zhen Wang, and Stephen Paul Smolley. Least squares generative adversarial networks. arXiv preprint ArXiv:1611.04076, 2016.
16. Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein gan. arXiv preprint arXiv:1701.07875, 2017.
17. Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. arXiv preprint arXiv:1611.07004, 2016.
18. Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, and Xi Chen. Improved techniques for training gans. In Advances in Neural Information Processing Systems, pages 2234–2242, 2016.
19. Yuan Xue, Tao Xu, Han Zhang, Rodney Long, and Xiaolei Huang. Segan: Adversarial network with multi-scale l\_1 loss for medical image segmentation. arXiv preprint arXiv:1706.01805, 2017.
20. Y. Hong, U. Hwang, J. Yoo, "How Generative Adversarial Networks and Its Variants Work: An Overview of GAN", [online] Available: https://arxiv.org/abs/1711.05914v6.
21. David Bau, Jun-Yan Zhu, Hendrik Strobelt, Bolei Zhou, Joshua B. Tenenbaum, William T. Freeman, Antonio Torralba, GAN Dissection: Visualizing and Understanding Generative Adversarial Networks, 2018, [online] Available: https://arxiv.org/abs/1811.10597
22. C. Wang, C. Xu, X. Yao, D. Tao, Evolutionary generative adversarial networks, 2018, [online] Available: https://arxiv.org/abs/1803.00657.