

计算机前沿课程复现：A fast blind zero-shot denoiser

摘要

相比于传统降噪方法，基于深度学习实现的降噪方法虽然在降噪性能上有显著优势，但其需要大量的数据集训练，获取如此庞大的数据集需要耗费大量的时间与物力。本次复现算法是近年来具有代表性的无监督深度学习降噪算法，该复现算法提出通过“奇”、“偶”图像采样与“水平方向像素平移”、“垂直方向像素平移”算法构造训练图像对，仅使用噪声图像本身即可以完成降噪过程，无需搜集大量数据进行训练。除此之外，项目使用的网络构造简洁，耗费时间短。在Set12和BSD68数据集（自然图像）上的高斯噪声降噪实验表明，复现算法可以有效去除高斯噪声。在Confocal数据集（显微镜图像）上的高斯泊松混合噪声降噪实验进一步验证了复现算法还可以去除显微镜环境下的混合噪声图像，并且处理速度较快。但在遥感数据集上的高斯噪声与条纹噪声实验表明，该复现方法在遥感场景下降噪时其性能会受到影响，尤其在处理条纹噪声时其降噪能力显著下降。在速度上，由于复现算法无法并行处理多通道图像，而遥感图像通常包含多个通道的图像数据，其需要逐通道处理，处理速度也受到显著影响。

关键词：图像降噪；无监督训练；单张噪声图像；卷积神经网络；高斯噪声；混合噪声；

1 引言

随着信息技术的飞速发展，图像作为人类社会活动中常见的信息载体，在人们日常的生活、工作和科研等各个领域都发挥着越来越重要的作用。然而，在图像的获取、传输和存储过程中，受环境因素、传感器发热、传输信道受损和存储单元失效等因素的影响，图像往往受到噪声的干扰，导致出现图像质量下降、细节丢失等问题。这些图像噪声有时还会对后续的图像分析与处理任务产生负面影响，如图像分割、图像分类、图像超分辨率和低照度图像增强算法等图像处理算法。图像去噪是图像处理领域里的一项重要的基础性工作。许多图像处理算法都需要先进行去噪处理的图像来提高性能。噪声图像一般可以用下列数学公式表示：

$$y = x + n \quad (1)$$

其中 y 表示含有噪声的高光谱图像， x 表示无失真高光谱图像， n 表示噪声。

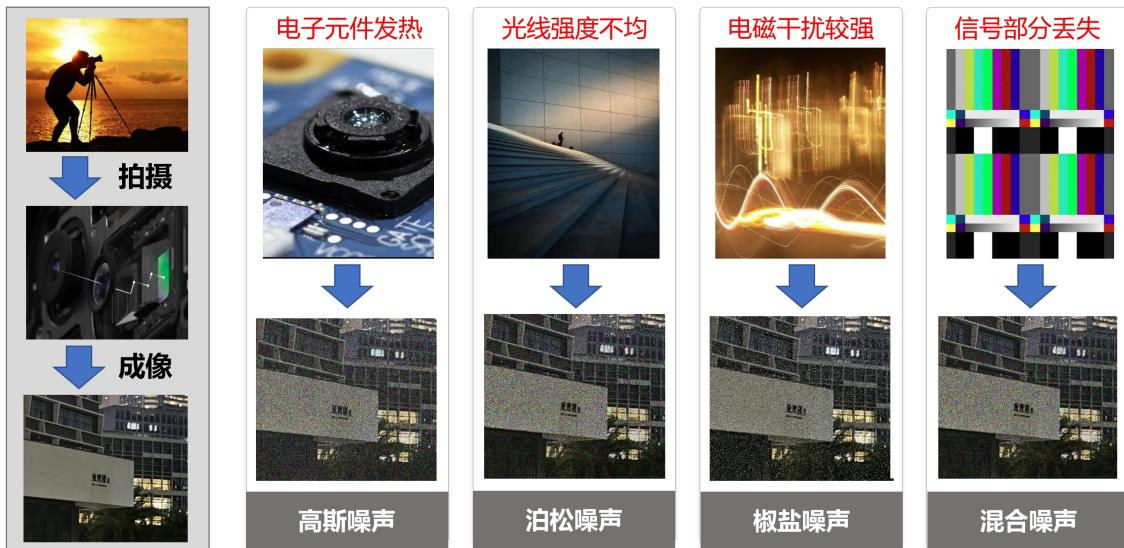


图 1. 各种噪声

图像降噪算法就是要从含有噪声的图像中恢复出关于无失真图像 x 的最佳估计。图像降噪需要经过一个非常复杂的过程，是一个典型的图像处理不适定问题 (ill posed problem)。图像中含有的噪声通常是混合的，不是单一的，其中包括了高斯噪声、泊松噪声、混合噪声等。图像降噪技术的研究和发展对于提升图像处理技术水平，推动相关领域的科技进步具有重要的意义。高效的图像降噪算法能够显著提高图像的视觉效果，还能对人工智能领域中人脸识别、目标检测等任务具有积极的促进作用。在医学成像领域，医学图像降噪算法可以提升医学图像的质量，帮助医生更准确地诊断疾病；在卫星遥感领域，降噪技术可以减弱噪声对高光谱图像的影响，提高高光谱图像分类的精度；在视频通信领域，降噪技术可以改善视频质量，提升视频观感，让视频更为清晰。对图像降噪算法的研究具有重要的理论和实际应用价值，图像降噪技术的发展对于推动数字图像处理技术的进步、提高图像应用系统的性能具有重要的意义。随着计算能力的进步、算法理论的不断完善和人们对图像质量要求的不断提高，降噪算法的研究将越来越受到关注，其应用前景也将更加广阔。

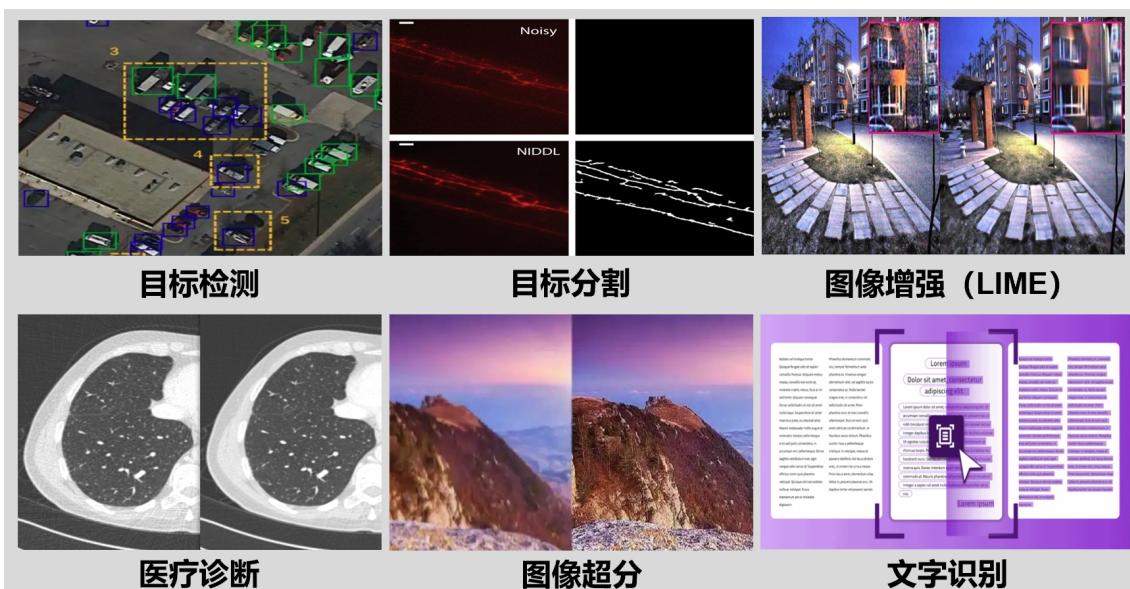


图 2. 噪声对其它计算机视觉任务的影响

2 相关工作

近年来，随着去噪算法和深度学习的发展，越来越多的去噪算法是基于深度学习实现的。基于深度学习的方法包括无监督学习和有监督学习。有监督学习方法如深度卷积神经去噪网络(DnCNN) [1]使用噪声图像与清晰图像作为输入，学习噪声图像减去清晰形象的残差。一种快速灵活的去噪卷积神经网络(FFDNet) [2]是对DnCNN方法的改进。FFDNet网络的输入包含一个可以人为调控的参数(噪声级)，因此FFDNet对不同噪声的适应性比DnCNN更好。但由于FFDNet模型用于训练的输入图像是人为调控的特殊噪声级别，因此这样的网络只能处理具有特定噪声级别的噪声图像，而对于未知噪声级别的图像往往表现较差。针对该问题，Anwar提出了一种卷积盲去噪网络(CBDNet) [3]。该方法将训练网络划分为两个子网络，一个用于估计噪声，另一个用于去噪。并且为了保证算法在不同噪声水平下都能达到良好的去噪效果，在该方法中还提出了一种非对称损失函数，用于提高网络对不同噪声水平的处理能力。基于有监督的去噪算法需要清晰图像和噪声图像对进行学习，但要获得如此大的高质量训练图像对集，需要耗费大量的人力和时间。因此，一些研究者提出了基于无监督或弱监督的深度学习去噪算法，这些算法不需要清晰图像和噪声图像对就能够训练。比如Noise2Noise(N2N) [4]，Noise2Void(N2V) [5]，Noise2Self(N2S) [6]，Self2Self(S2S) [7]，Neighbor2Neighbor [8]，Deep Image Prior (DIP) [9]。N2N提出在卷积神经网络的训练条件下，清晰图像可以被在同一图像中的噪声部分是独立分布的其它噪声图像替代。但即使是获取同一图像具有不同噪声的图像也存在一定的困难。在N2V和N2S中，可以使用一组无关的外部噪声图像上就能学习出性能良好的去噪神经网络。然而，为了达到良好的性能，用于训练的噪声图像在图像内容和噪声统计分布方面应该与被处理的噪声图像高度相关，但即使是收集这样的外部图像也是一项有困难的任务。S2S只使用需要去噪的图像本身，对训练样本没有前提条件，它生成一组对需要去噪图像进行伯努利采样的图像作为输入和目标来训练网络。Neighbor2Neighbor也需要待处理的噪声图像去噪，但需要经过复杂的采样过程与训练过程。

发表时间	方法名称	发表期刊	引用次数	方法所需数据
2018	Noise2Noise	ICML	1995	多张噪声图像对
2019	Noise2Void	CVPR	1294	单张噪声图像
2019	Noise2Self	ICML	758	单张或多张噪声图像
2020	Self2Self	CVPR	385	单张噪声图像
2021	Neighbor2Neighbor	CVPR	377	多张噪声图像
2022	Noise2Fast	Nat Mach Intell	46	单张噪声图像

图 3. 无监督降噪算法

上述的无监督深度学习方法虽然仅依靠噪声图像训练，但有的方法(如Noise2Noise)还需要大量噪声图像对训练，有的方法训练时间非常久(如Neighbor2Neighbor等)。Noise2Fast [10]方法仅需降噪噪声图像本身即可完成网络的迭代学习。其基本思想是利用像素的奇偶位置采样

得到不同的图像，称为“奇图像”和“偶图像”，然后将这些图像两两组合形成新的训练图像对。这使得该方法对降噪图像的适应性大大提高，所有的信息仅来源于需要降噪的噪声图像本身。因此只需要一个随机初始化的网络和噪声图像作为输入，将目标原始噪声图像生成的采样图像作为目标进行反复迭代式的学习，既可以采取早停策略，在某一次迭代步数中得到清晰的降噪图像。

3 本文方法

本文所复现的Noise2Fast方法仅需一张噪声图像既可以完成降噪过程，不仅不需要额外的训练数据，也不需要提前训练与保存模型，降低了存储空间成本。并且，所使用的网络非常简单，主体部分仅包含9个卷积模块，耗费时间少。同为无监督降噪方法的Sele2Self对一张512乘以512的显微镜图像降噪需要5小时，而Noise2Fast方法仅需56秒，是其耗时的0.3%。

3.1 本文方法概述

在Noise2Fast方法中，一张噪声图像首先被采样成两张图像，分别是“奇采样”图像与“偶采样”图像，再对这两张图像分别使用水平像素压缩和垂直像素压缩生成4张图像，最后将这4张图像随机两两组合成4组噪声图像对作为训练集，使用一个非常简单的卷积神经网络进行训练，得到降噪结果。本文复现的方法如图 4所示：

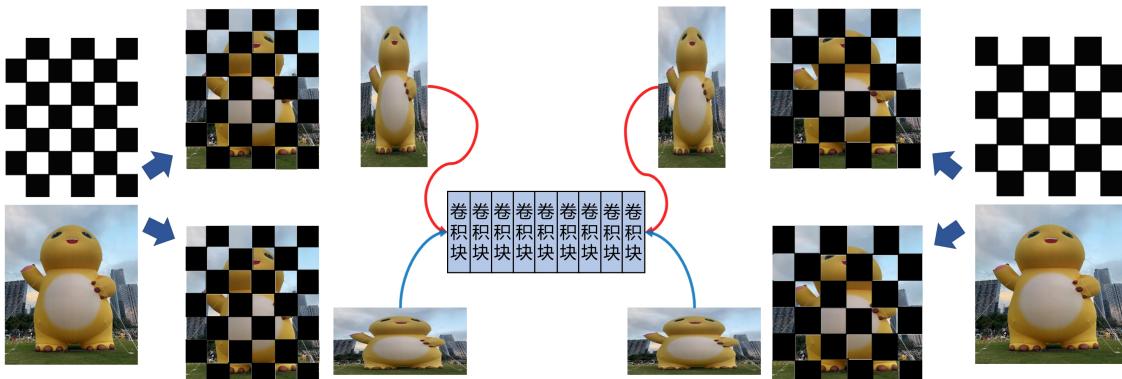


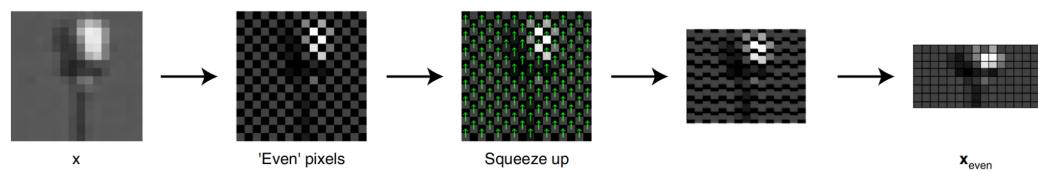
图 4. 方法示意图

由图 4可以看出，经过水平像素压缩的图像与同样经过水平像素压缩的图像构成训练对，经过垂直像素压缩的图像与同样经过垂直像素压缩的图像构成训练对，不同的是这些图像分别来源于“奇采样”图像与“偶采样”图像。图示的卷积神经网络一共包含9个卷积模块，在具体实现中，分别在第2、4、6、8个卷积模块后面紧接着一个Relu激活函数，并在最后一个卷积模块的后面使用Sigmoid激活函数对图像特征进行处理，得到最终的降噪图像。

3.2 数据处理过程

如图 5所示，在数据处理过程中，噪声图像首先被分成两张图像，分别是“奇采样”图像与“偶采样”图像，然后分别对这两张图像使用水平像素压缩和垂直像素压缩生成4张图像，以供后续训练使用。

原文数据处理过程



复现数据处理过程

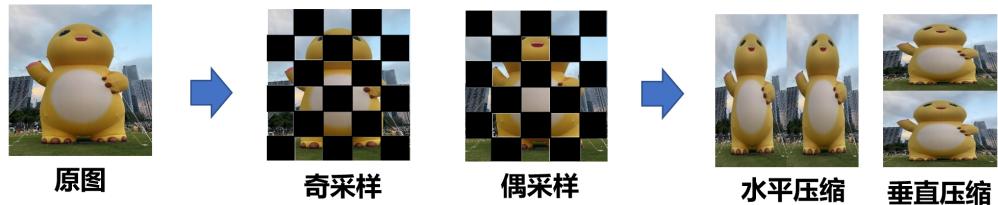


图 5. 数据处理过程图

4 复现细节

4.1 与已有开源代码对比

本文复现参考的代码来源于文献 [10] 公开的 Noise2Fast。与原文开源代码相比，本文主要做了以下额外工作：额外数据集、额外噪声类型、过程可视化展示和卷积模块个数对性能影响的探索实验。

4.2 数据集

本文复现使用的数据集除文献 [10] 公开的 Noise2Fast 采用的 Set12、BSD68、Confocal 数据集以外，还额外添加了来自谷歌地图的遥感图像和 iphone12 手机拍摄图像。

4.3 噪声类型

本文复现使用的噪声类型除文献 [10] 公开的 Noise2Fast 采用的高斯噪声、显微镜图像真实噪声以外，还包括遥感图像中的高斯噪声、条纹噪声，自然图像的高斯噪声、泊松噪声以及高斯泊松混合噪声。

4.4 训练细节

本文训练使用文献 [10] 公开的 Noise2Fast 默认训练参数。

5 实验结果分析

本次复现方法实验首先分别测试了 Set12 数据集高斯噪声水平强度为 $\sigma=15$ 、 25 和 50 ，BSD68 数据集高斯噪声水平强度为 $\sigma=25$ 和 50 ，以及 Confocal 数据集图像。还测试了真实场景图像的合成高斯噪声、泊松噪声和高斯泊松混合噪声实验，遥感图像合成高斯噪声、条纹噪声实验。最后测试了不同卷积模块个数在 Set12 数据集高斯噪声水平强度为 $\sigma=15$ 、 25 和 50 的情况下的性能变化。测试使用的图像评价指标包括峰值信噪比（Peak signal noise ratio, PSNR）和图

像相似度（Structural similarity, SSIM）指标。除了客观指标外，还采取了主观降噪结果分析，能够更直观地分析结果。

5.1 噪声类型

本文复现使用的噪声类型除文献 [10]公开的Noise2Fast采用的高斯噪声、显微镜图像真实噪声以外，还包括遥感图像中的高斯噪声、条纹噪声，自然图像的高斯噪声、泊松噪声以及高斯泊松混合噪声。

5.2 原文数据集结果

参考原文实验设置，分别对Set12、BSD68和Confocal数据集图像进行降噪处理，结果如表1、2、3所示。由表格的PSNR指标可以看出，复现结果和原文相差较小，在Set12数据集上的平均差值为+0.04，在BSD68上的平均差值为+0.01，在Confocal数据集上的插值为-0.09，基本上与原文一致。在这3个数据集上降噪后图像的视觉展示如图6、7、8、10、9所示。由图像可以看出Noise2Fast方法能够在短时间内将噪声去除，降噪前后的视觉效果对比明显。

表 1. Set12数据集高斯降噪测试PSNR结果

噪声强度	原文	复现	差值
15	31.10	31.16	+0.06
25	29.05	29.04	-0.01
50	25.82	25.91	+0.09
平均值	28.66	28.70	+0.04

表 2. BSD68数据集高斯降噪测试PSNR结果

噪声强度	原文	复现	差值
25	28.12	28.09	-0.03
50	25.23	25.28	+0.05
平均值	26.68	26.69	+0.01

表 3. Confocal数据集降噪测试PSNR结果

噪声强度	原文	复现	差值
真实噪声	36.61	36.52	-0.09



噪声图像



参考图像



Noise2Fast(1.98s)

图 6. Set12数据集CameraMan图像



噪声图像

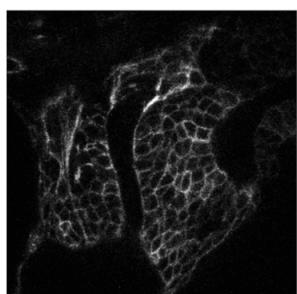


参考图像

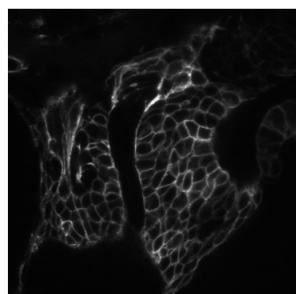


Noise2Fast(12.76s)

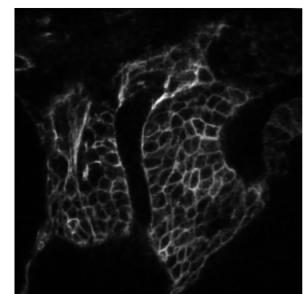
图 7. BSD68数据集House图像



噪声图像



参考图像



Noise2Fast(17.85s)

图 8. Confocal数据集Fish图像

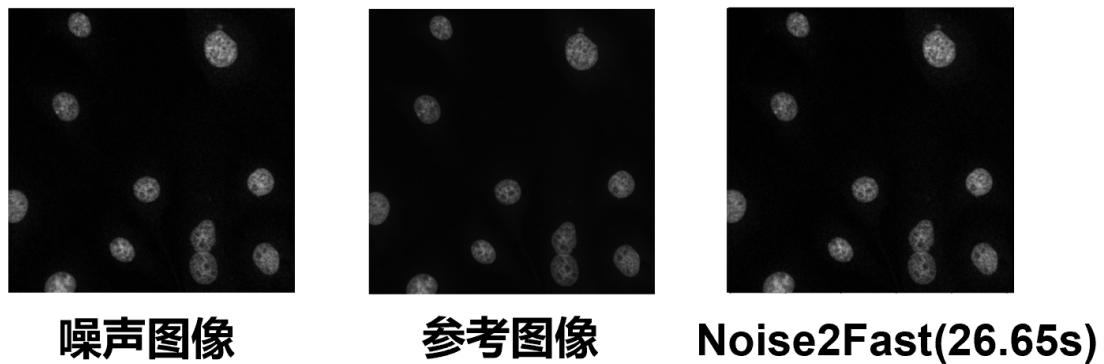


图 9. Confocal 数据集Cell图像

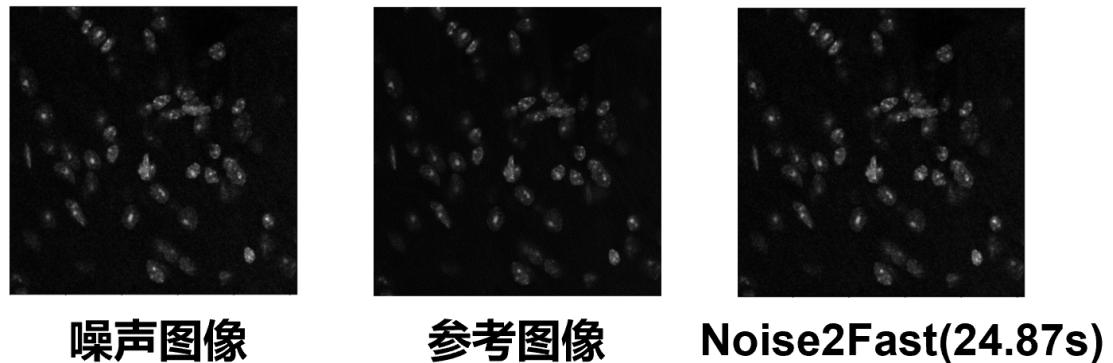


图 10. Confocal 数据集Mice图像

5.3 额外数据集结果

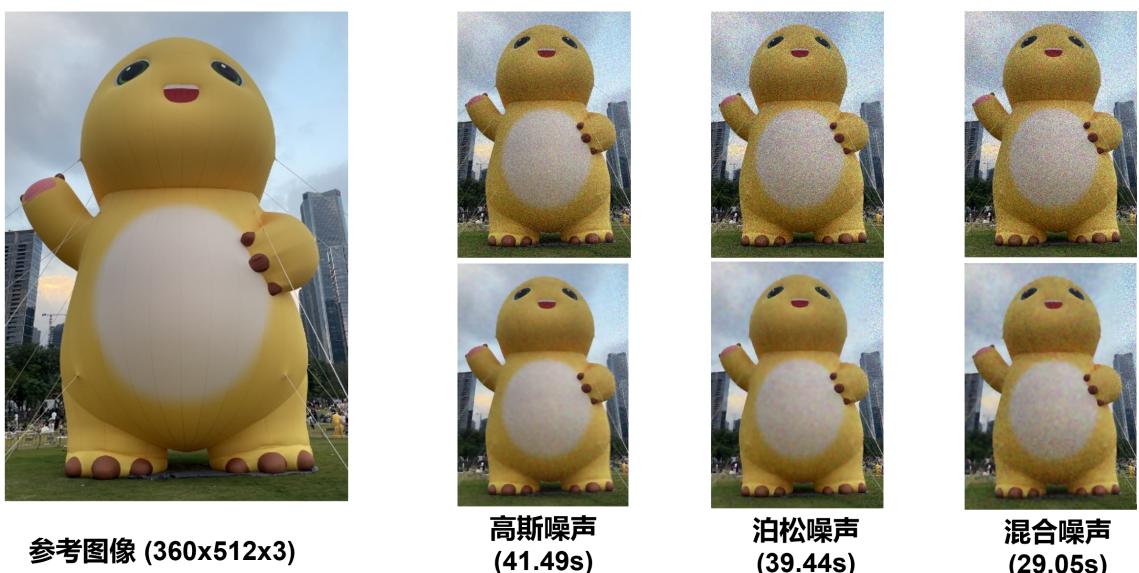


图 11. Iphone12拍摄图像合成噪声降噪实验结果

在额外数据集测试部分，选取了Iphone12手机拍摄图像和遥感图像，均是通过合成噪声的形式进行降噪实验，降噪结果如图11、12、13所示。由图像可以看出，Noise2Fast方法对于

高斯噪声降噪效果较好，对于泊松噪声降噪性能会下降，对于混合噪声降噪性能下降明显，对于条纹噪声几乎不能降噪。并且Noise2Fast只能逐通道降噪，故其在对多通道图像降噪时，耗费时间也会增多。

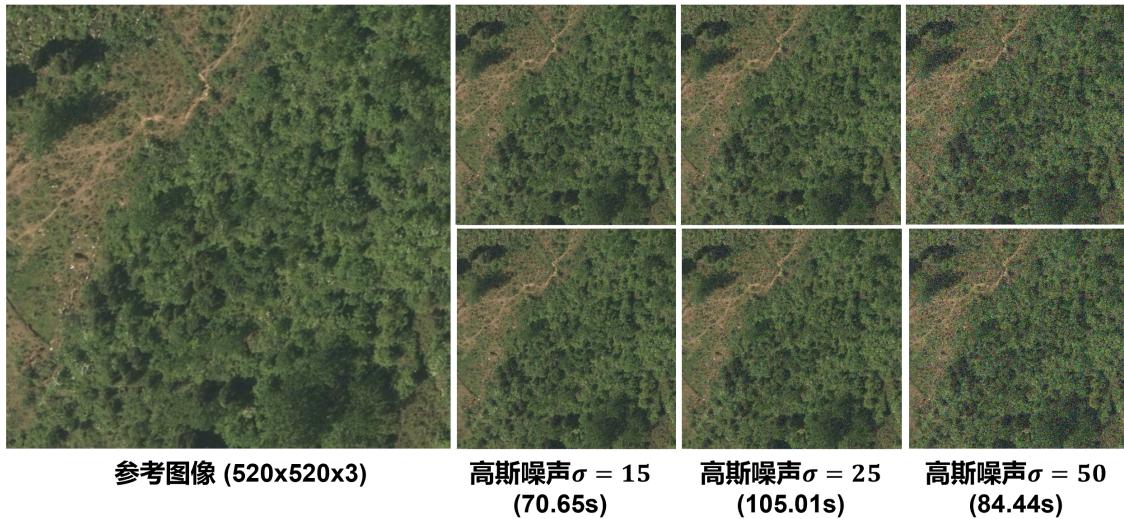


图 12. 遥感图像合成高斯噪声降噪结果

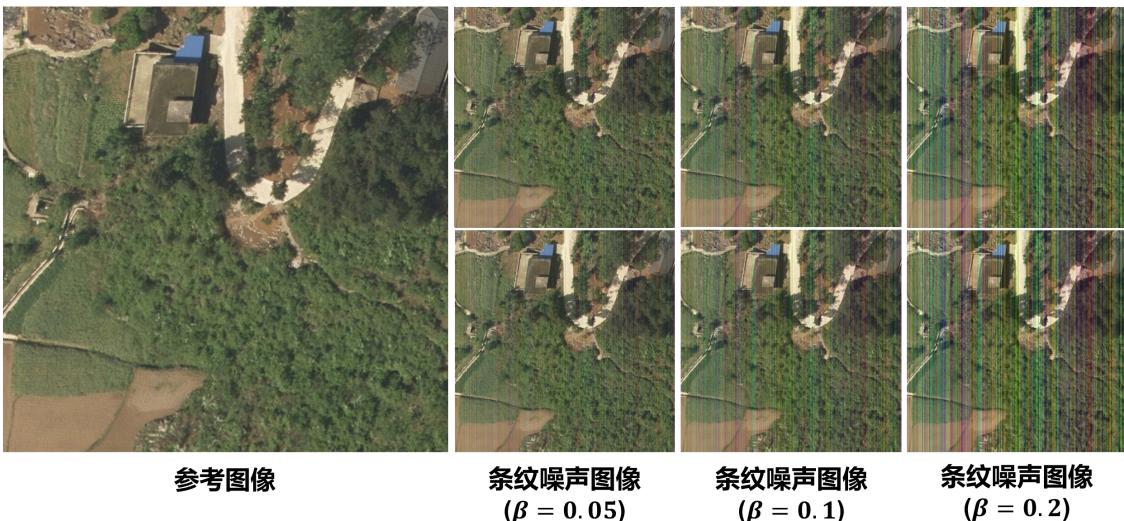


图 13. 遥感图像合成条纹噪声降噪结果

5.4 卷积模块个数对性能影响

为探究不同卷积模块个数对Noise2Fast方法性能的影响，分别使用3、5、7、9、11、13个卷积模块对Set12数据集高斯噪声进行降噪实验，结果如表4所示。可以看出，在卷积模块大于等于7个以后，PSNR增长速度放缓，甚至出现下降的趋势，但耗费时间逐渐变多。综合来看，9个卷积模块是较好的选择。图14展示了不同卷积模块个数对蝴蝶图像的降噪结果和与参考图像的残差图像，可以看出在大于等于7个卷积模块以后，残差图像基本上没有很大改变，与表格展现的PSNR结果趋势一致。

表 4. 不同卷积个数对Set12数据集高斯噪声降噪结果影响

噪声强度	3个	5个	7个	9个(原文)	9个(复现)	11个	13个
15	29.06	29.92	30.85	31.10	31.16	31.19	30.84
25	27.15	27.93	28.83	29.05	29.04	29.08	29.11
50	23.95	25.00	25.60	25.82	25.91	24.82	25.98
平均值	26.72	27.62	28.43	28.66	28.70	28.36	28.64
耗费时间(蝴蝶图像)	0.67s	1.36s	2.46s	-	1.98s	7.13s	7.81s

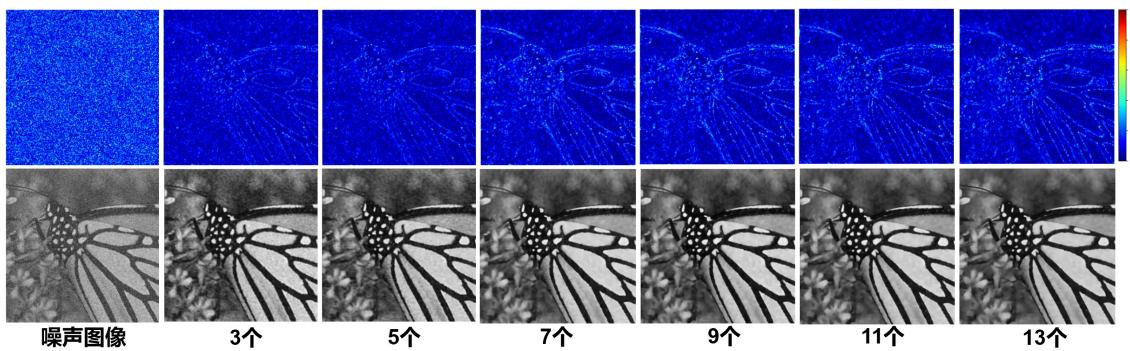


图 14. 不同卷积个数对Set12数据集蝴蝶图像高斯噪声降噪结果影响

6 总结与展望

Noise2Fast方法使用奇偶像素生成下采样图像，再经过水平和垂直方向的压缩得到不同图像进行训练，耗时短，性能也较好，能够在短时间内得到降噪结果，无需额外的数据集进行训练。缺点是对于一些类型的噪声还无法灵活适应，泛化能力有待进一步提高，在未来对该方法加入噪声信息，例如FFDNet的噪声图或基于噪声图像提取噪声特征一起输入到网络中，提高其泛化能力。

参考文献

- [1] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE Transactions on Image Processing*, 26(7):3142–3155, 2016.
- [2] K. Zhang, W. Zuo, and L. Zhang. Ffdnet: Toward a fast and flexible solution for cnn based image denoising. *IEEE Transactions on Image Processing*, pages 1–1, 2017.
- [3] Shi Guo, Zifei Yan, Kai Zhang, Wangmeng Zuo, and Lei Zhang. Toward convolutional blind denoising of real photographs. *CoRR*, abs/1807.04686, 2018.
- [4] J Lehtinen, J Munkberg, J Hasselgren, S Laine, T Karras, M Aittala, and T Aila. Noise2noise: learning image restoration without clean data. poceedings of the 35th international conference

on machine learning, icml 2018, july 10–15, 2018. *Proceedings of Machine Learning Research*, pages 2971–2980, 2018.

- [5] Alexander Krull, Tim-Oliver Buchholz, and Florian Jug. Noise2void-learning denoising from single noisy images. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2129–2137, 2019.
- [6] Joshua Batson and Loic Royer. Noise2self: Blind denoising by self-supervision. In *International Conference on Machine Learning*, pages 524–533. PMLR, 2019.
- [7] Yuhui Quan, Mingqin Chen, Tongyao Pang, and Hui Ji. Self2self with dropout: Learning self-supervised denoising from single image. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1890–1898, 2020.
- [8] Tao Huang, Songjiang Li, Xu Jia, Huchuan Lu, and Jianzhuang Liu. Neighbor2neighbor: Self-supervised denoising from single noisy images. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 14781–14790, 2021.
- [9] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Deep image prior. *International Journal of Computer Vision*, 128(7):1867–1888, mar 2020.
- [10] Jason Lequyer, Reuben Philip, Amit Sharma, Wen-Hsin Hsu, and Laurence Pelletier. A fast blind zero-shot denoiser. *Nature Machine Intelligence*, 4(11):953–963, 2022.