

深度图像先验

德米特里·乌里扬诺夫·斯科尔
科沃科学技术研究所,
Yandex
dmitry.ulyanov@skoltech.ru

安德里亚·维达尔迪
(Andrea Vedaldi) 牛津大学
vedaldi@robots.ox.ac.uk

维克多·伦皮茨基·斯科尔科
沃科技学院 (Skoltech)
lempitsky@skoltech.ru

Abstract

深度卷积网络已成为图像生成和恢复的流行工具。一般来说，它们的优异性能归因于它们从大量示例图像中学习真实图像先验的能力。在本文中，我们表明，相反，生成器网络的结构足以捕获大量低值图像。在任何学习之前级别图像统计。为了做到这一点，我们证明了随机初始化的神经网络可以用作手工制作的先验，在去噪、超分辨率和修复等标准逆问题中取得了优异的结果。此外，相同的先验可用于反转深层神经表示以对其进行诊断，并基于闪光-无闪光输入对恢复图像。除了其多样化的应用之外，我们的方法还强调了标准生成器网络架构捕获的归纳偏差。它还弥合了两个非常流行的图像恢复方法系列之间的差距：使用深度卷积网络的基于学习的方法和基于手工图像先验（例如自相似性）的无学习方法。

一、简介

深度卷积神经网络 (ConvNets) 目前在逆图像重建问题（例如去噪 [5, 20] 或单图像超分辨率 [19, 29, 18]）中设定了最先进的技术。卷积网络在更“奇异”的问题上也取得了巨大成功，例如根据某些深层网络中的激活或其 HOG 描述符来重建图像 [8]。更一般地说，具有类似架构的 ConvNet 如今用于使用生成对抗网络 [11]、变分自动编码器 [16] 和直接像素误差最小化 [9, 3] 等方法来生成图像。图像恢复和生成

代码和补充材料可在 https://dmitryulyanov.github.io/deep_image_prior 获取

: //

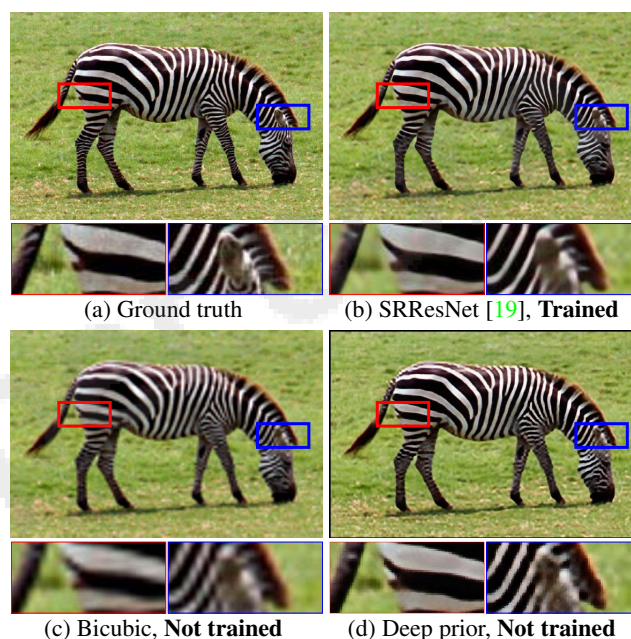


图 1：使用深度图像先验的超分辨率。我们的方法使用随机初始化的 ConvNet 对图像进行上采样，使用其结构作为图像先验；与双三次上采样类似，这种方法不需要学习，但会产生更清晰的结果和更清晰的边缘。事实上，我们的结果非常接近使用从大型数据集学习的 ConvNet 的最先进的超分辨率方法。深度图像先验对于我们可以测试的所有逆问题都很有效。

迭代几乎总是在大型图像数据集上进行训练。因此，人们可能会认为它们的出色性能归因于它们从数据中学习真实图像先验的能力。然而，仅靠学习不足以解释深度网络的良好性能。例如，[33] 的作者最近表明，在真实数据上训练时泛化良好的同一个图像分类网络在提供随机标签时也可能过度拟合。因此，泛化需要网络的结构与数据的结构。但是，那

这种相互作用的本质仍然不清楚，特别是在图像生成的背景下。在这项工作中，我们表明，与“学习对于建立良好的图像先验是必要的”这一信念相反，大量的图像统计数据是由独立于卷积图像生成器的结构捕获的对于解决各种图像恢复问题所需的统计数据尤其如此，其中需要图像先验来整合退化过程中丢失的信息。为了证明这一点，我们将未经训练的卷积网络应用于几个此类问题的解决方案。我们没有遵循在大型示例图像数据集上训练卷积网络的常见范例，而是将生成器网络拟合到单个降级图像。在该方案中，网络权重充当恢复图像的参数化。给定特定的退化图像和任务相关的观察模型，权重被随机初始化并拟合以最大化其可能性。以不同的方式表述，我们将重建视为条件图像生成问题，并表明解决该问题所需的唯一信息包含在单个降级的输入图像和用于重建的网络的手工结构中。我们表明，这种非常简单的公式对于去噪、修复和超分辨率等标准图像处理问题非常有竞争力。这是特别值得注意的，因为网络的任何方面都不是从数据中学习的；相反，网络的权重总是随机初始化的，因此唯一的先验信息存在于网络本身的结构中。据我们所知，这是第一个独立于从图像中学习网络参数而直接研究深度卷积生成网络捕获的先验的研究。除了标准图像恢复任务之外，我们还展示了我们的技术在以下方面的应用：了解深层神经网络激活中包含的信息。为此，我们考虑[21]的“自然原像”技术，其目标是通过在自然图像集上反转深层网络学习的不变量来表征它。我们证明，未经训练的深度卷积生成器可以用来取代[21]（电视范数）中使用的替代自然先验，并显着改善结果。由于新的正则化器与电视规范一样，不是从数据中学习的，而是完全手工制作的，因此产生的可视化避免了因使用强大的学习正则化器而产生的潜在偏差[8]。

2. 方法

深度网络通过学习生成器/解码器网络 $x = f_{\theta}(z)$ 将随机代码向量 z 映射到图像 x 来应用于图像生成。这种方法可用于从随机分布中采样真实图像[11]。

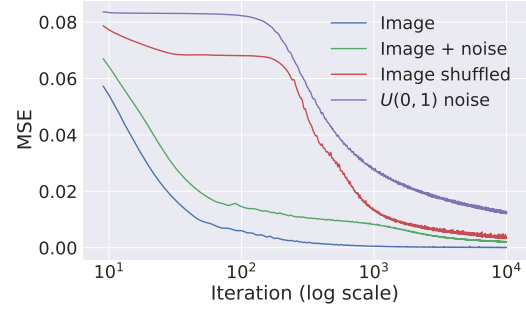


图 2：重建任务的学习曲线，使用：自然图像，相同加上独立同分布。噪声，同样随机扰频，以及白噪声。看起来自然的图像会导致更快的收敛，而噪声会被拒绝。

在这里，我们重点关注分布以损坏的观测值 x_0 为条件的情况，以解决诸如去噪 [5] 和超分辨率 [7] 之类的逆问题。我们的目标是研究由选择隐式捕获的先验信息在学习特定生成器网络结构的任何参数之前。我们通过将神经网络解释为图像 $x \in \mathbb{R}^{3 \times H \times W}$ 的参数化 $x = f_{\theta}(z)$ 来实现这一点。这里 $z \in \mathbb{R}^{C \times H \times W}$ 是代码张量/向量， θ 是网络参数。网络本身交替执行卷积、上采样和非线性激活等过滤操作。特别是，我们的大多数实验都是使用具有跳跃连接的 U-Nettype “沙漏” 架构进行的，其中 z 和 x 具有相同的空间大小。我们的默认架构有 200 万个参数 θ （有关所有使用的架构的详细信息，请参阅补充材料）。为了展示这种参数化的强大功能，我们考虑逆向任务，例如去噪、超分辨率和修复。这些可以表示为以下类型的能量最小化问题

$$x^* = \min_x E(x; x_0) + R(x), \quad (1)$$

其中 $E(x; x_0)$ 是任务相关的数据项， x_0 噪声/低分辨率/遮挡图像， $R(x)$ 是正则化器。数据项 $E(x; x_0)$ 的选择是由应用程序决定，稍后将讨论。正则化器的选择通常会捕获自然图像的通用先验，因此更加困难，并且是许多研究的主题。举一个简单的例子， $R(x)$ 可以是图像的总变分 (TV)，这鼓励解决方案包含均匀区域。在这项工作中，我们用神经网络捕获的隐式先验替换正则化器 $R(x)$ ，如下所示：

$$\theta^* = \operatorname{argmin}_{\theta} E(f_{\theta}(z); x_0), \quad x^* = f_{\theta^*}(z). \quad (2)$$

最小化器 θ^* 是使用优化器获得的，例如从参数的随机初始化开始的梯度下降。给定一个（局部）最小化器 θ^* ，恢复过程的结果为 $x^* = f_{\theta^*}(z)$ 。请注意，虽然也可以对代码 z 进行优化，但在我们的实验中我们没有这样做。因此，除非另有说明， z 是一个固定的 3D 张量，具有 32 个特征图，并且空间大小与充满均匀噪声的 x 相同。我们发现，在一些实验中，在每次迭代中随机扰动 z 会带来更好的结果（参见补充材料）。根据 (1)，由 (2) 定义的先验 $R(x)$ 是指示函数 $R(x)$ 对于所有可以通过某种架构的深度 ConvNet 从 z 生成的图像， $R(x) = 0$ ，对于所有其他信号， $R(x) = +\infty$ 。由于网络的任何方面都不是根据数据进行预先训练的，因此这种深度图像先验是有效手工制作的，就像电视规范一样。我们证明，这种手工制作的先验对于各种图像恢复任务非常有效。

具有高噪声阻抗的参数化。人们可能想知道为什么高容量网络 f_{θ} 可以用作先验。事实上，人们可能期望能够找到参数 θ 来恢复任何可能的图像 x ，包括随机噪声，以便网络不应该对生成的图像施加任何限制。我们现在表明，虽然实际上几乎任何图像都可以拟合，但网络架构的选择对如何通过梯度下降等方法搜索解空间有重大影响。特别是，我们表明网络可以抵制“糟糕”的解决方案，并更快地下降到看起来自然的图像。结果是，最小化 (2) 要么会产生漂亮的局部最优，要么至少优化轨迹接近一个。为了定量研究这种效果，我们考虑最基本的重建问题：给定目标图像 x_0 ，我们想要找到再现该图像的参数 θ^* 的值。这可以设置为 (2) 的优化，使用数据项比较生成的图像 $\text{tox}_0 : E(x; x_0) = \|x - x_0\|^2$

将其插入等式中。(2) 引出优化问题

$$\min_{\theta} \|f_{\theta}(z) - x_0\|^2 \quad (4)$$

图 2 显示了能量 $E(x; x_0)$ 的值作为图像 x_0 的四种不同选择的梯度下降迭代的函数：1) 自然图像，2) 相同图像加加性噪声，3) 相同图像随机排列像素后，4) 白噪声。从图中可以明显看出，对于情况 1) 和 2)，优化要快得多，而对于情况 3) 和 4)，参数化呈现出显着的“惯性”。因此，尽管在极限情况下参数化可以拟合非结构化噪声，它这样做是非常不情愿的。换句话说，

参数化提供了高噪声阻抗和低信号阻抗。因此，对于大多数应用，我们将优化过程 (2) 中的迭代次数限制为一定的迭代次数。所得的先验对应于一组减少的图像的投影，该图像可以由 ConvNets 从 z 生成，参数 θ 与随机初始化 θ_0 相差不太远。

3. 应用

我们现在通过实验展示所提出的先验方法如何解决各种图像重建问题。由于篇幅限制，我们提供了一些示例和数字，并在补充材料和项目网页 [30] 中包含了更多示例和数字。

去噪和通用重建。由于我们的参数化对图像噪声表现出高阻抗，因此它可以自然地用于滤除图像中的噪声。去噪的目的是从噪声观测 x_0 中恢复干净的图像 x 。有时退化模型是已知的： $x_0 = x +$ 其中遵循特定分布。然而，在盲去噪中，噪声模型更常见是未知的。这里我们在盲目性假设下工作，但可以轻松修改该方法以纳入有关噪声模型的信息。我们使用与 eqs 相同的精确公式。(3) 和 (4) 并且，给定一个有噪声的图像 x_0 ，在替换等式的最小化器 θ^* 后恢复干净的图像 $x^* = f_{\theta^*}(z)$ 。(4) 我们的方法不需要需要恢复的图像降级过程的模型。这使得它能够以“即插即用”的方式应用于图像恢复任务，其中退化过程很复杂和/或未知，并且很难获得用于监督训练的真实数据。我们通过图中的几个定性示例展示了这种能力。在图 4 和补充材料中，我们的方法使用二次能量 (3) 导出公式 (4) 来恢复因复杂且未知的压缩伪影而降级的图像。图 3（顶行）还展示了该方法在自然图像（本例中为卡通图像）之外的适用性。我们在标准数据集 1 上评估我们的去噪方法，该数据集由 9 个彩色图像组成，噪声强度为 $\sigma = 25$ 。我们实现了经过 1800 个优化步骤后，PSNR 为 29.22。如果我们对最后一次迭代中获得的恢复图像进行平均（使用指数滑动窗口），分数将提高到 30.43。如果对两次优化运行进行平均，我们的方法可进一步将 PSNR 提高到 31.00。作为参考，不需要预训练的两种流行方法 CMB3D [6] 和非局部均值 [4] 的分数分别为 31.42 和 30.26。

¹http://www.cs.tut.fi/~foi/GCF-BM3D/index.html#ref_results

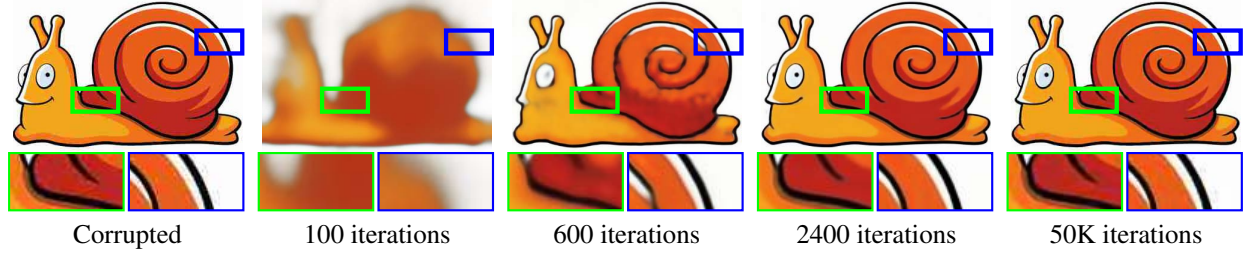


图 3: JPEG 压缩图像的盲恢复。（推荐电子放大）我们的方法可以恢复具有复杂退化的图像（在本例中为 JPEG 压缩）。随着优化过程的进展，深度图像先验可以恢复大部分信号，同时消除光晕和块效应（2400 次迭代后），然后最终过度拟合输入（50K 次迭代时）。

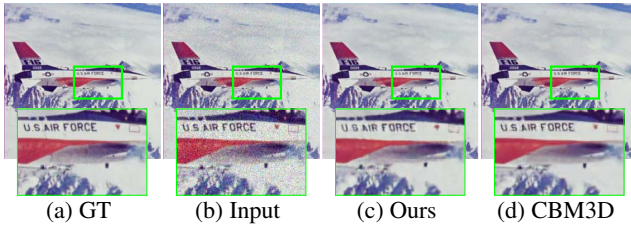


图 4: 盲图像去噪。深度先验图像成功地恢复了人造和自然模式。作为参考，显示了最先进的非学习去噪方法 [6] 的结果。

超分辨率。超分辨率的目标是采用低分辨率（LR）图像 $x_0 \in \mathbb{R}^{3 \times H \times W}$ 和上采样因子 t ，并生成相应的高分辨率（HR）版本 $x \in \mathbb{R}^{3 \times tH \times tW}$ 。为了解决这个逆问题，将（2）中的数据项设置为：

$$E(x; x_0) = \|d(x) - x_0\|^2 \quad (5)$$

其中 $d(\cdot) : \mathbb{R}^{3 \times tH \times tW} \rightarrow \mathbb{R}^{3 \times H \times W}$ 是按因子 t 调整图像大小的下采样算子。因此，问题是找到在下采样时与 LR 图像 x_0 相同的 HR 图像 x 。超分辨率是一个不适定问题，因为有无限多个 HR 图像 x 可以简化为相同的 LR 图像 x_0 （即算子 d 远非满射）。需要正则化，以便在（5）的无限最小化器中选择最合理的最小化器。（2），我们通过考虑重新参数化 $x = f_\theta(z)$ 并优化由此产生的能量来规范问题。 θ 。优化仍然使用梯度下降，利用神经网络和最常见的下采样算子（例如 Lanc-zos）可微的事实。我们使用 Set5 [2] 和 Set14 [32] 数据集评估我们方法的超分辨率能力。我们使用缩放因子 4 与其他作品进行比较，并在补充材料中显示缩放因子为 8 的结果。我们将每个图像的优化步骤数固定为 2000。

与双三次上采样和最先进的基于学习的方法 SRResNet [19]、LapSRN [29] 的定性比较如图 2 所示。5. 我们的方法可以与双三次方法进行比较，因为两种方法除了给定的低分辨率图像之外都不使用其他数据。从视觉上看，我们接近使用 MSE 损失的基于学习的方法的质量。基于 GAN [11] 的方法 SRGAN [19] 和 EnhanceNet [28]（比较中未显示）智能地幻觉图像的精细细节，这对于我们完全不使用 HR 图像世界信息的方法来说是不可能的。我们计算使用生成图像的中心裁剪的 PSNR。我们的方法在 Set5 和 Set14 数据集上分别实现了 29.90 和 27.00 PSNR。双三次上采样的得分较低，为 28.43 和 26.05，而 SRResNet 的 PSNR 为 32.10 和 28.53。虽然我们的方法仍然优于基于学习的方法，但它比双三次上采样要好得多。从视觉上看，它似乎缩小了双三次和最先进的训练 ConvNet 之间的差距（参见图 1、图 5 和 suppmat）。

修补。在图像修复中，给定一个图像 x_0 ，其中缺少像素，对应于二进制掩码 $m \in \{0, 1\}^{H \times W}$ ；目标是重建丢失的数据。相应的数据项由下式给出

$$E(x; x_0) = \|(x - x_0) \odot m\|^2, \quad (6)$$

阿达玛的产品在哪里。数据先验的必要性是显而易见的，因为该能量与缺失像素的值无关，因此如果目标直接在像素值 x 上进行优化，则在初始化后该能量永远不会改变。和之前一样，先验是通过优化数据项 w.r.t 来引入的。重新参数化（2）。在第一个示例（图 7，顶行）中，修复用于删除覆盖在图像上的文本。我们的方法与专门为修复设计的 [27] 方法进行了比较。我们的方法产生了几乎完美的结果，几乎没有伪影，而对于 [27]，文本蒙版在某些区域仍然可见。7（底部）考虑使用根据二进制伯努利分布随机采样的掩模进行修复

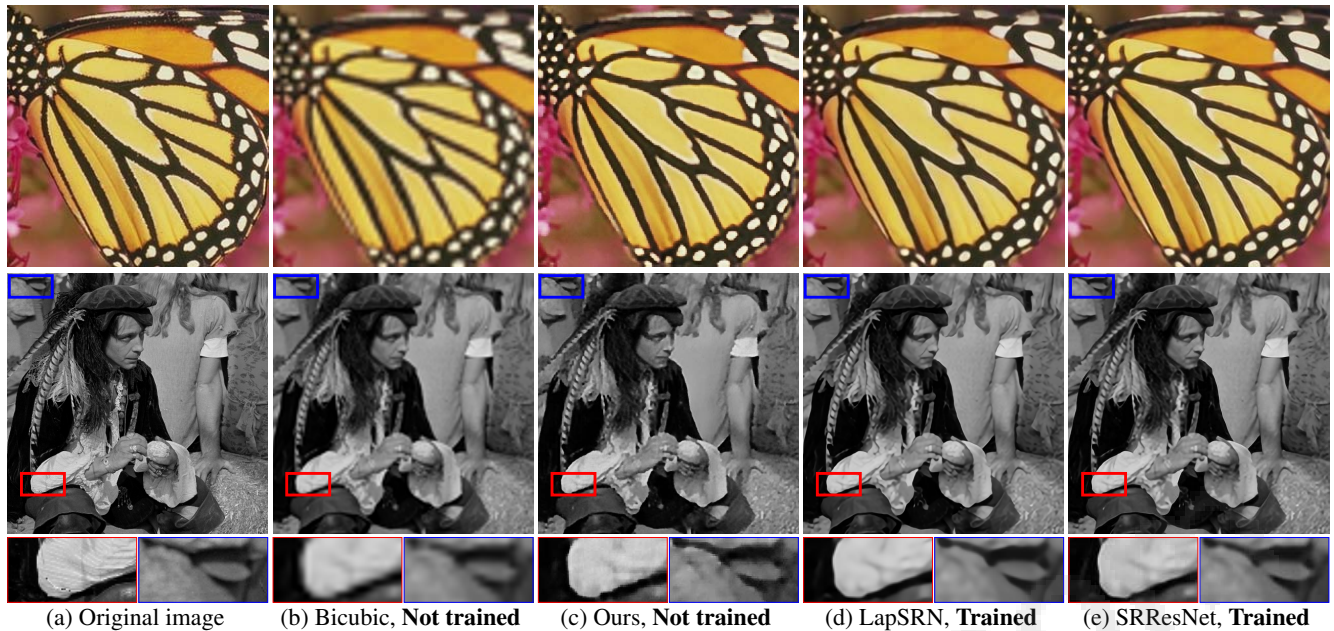


图 5: 4 倍图像超分辨率。类似于例如双三次上采样, 我们的方法永远无法访问除单个低分辨率图像之外的任何数据, 但它会产生更清晰的结果, 其锐利边缘接近最先进的超分辨率方法 (LapSRN [18], SRResNet [19]) 利用从大型数据集训练的网络。



图 6: 区域修复。在许多情况下, 深度图像先验足以成功修复大区域。尽管没有使用学习, 结果可能与[15]相当。超参数的选择很重要 (例如 (d) 展示了对学习率的敏感性), 但良好的设置对于我们尝试的大多数图像都很有效。

布丁。首先, 对掩模进行采样以随机丢弃 50% 的像素。我们将我们的方法与基于卷积稀疏编码的[25]方法进行比较。为了获得[25]的结果, 我们首先将损坏的图像 x_0 分解为与[12]类似的低频和高频分量, 并在高频部分运行他们的方法。为了公平比较, 我们使用他们方法的版本, 其中使用输入图像构建字典 (在[25]中显示表现更好)。表1给出了我们的方法在标准数据集[14]上的定量比较, 显示出所提出的方法与卷积稀疏编码相比具有强大的定量优势。在图中。图7 (底部) 我们提出了与[25]的代表性定性视觉比较。我们还将我们的方法应用于大孔的修补。

由于不可训练, 我们的方法预计无法正确用于“高度语义”的大孔修复 (例如面部修复)。然而, 它在其他情况下却出奇地有效。我们与图 15 中基于学习的方法进行比较。6. 深度图像先验利用图像的上下文, 并使用已知部分的纹理对未知区域进行插值。这种行为突出了深层图像先验和传统自相似先验之间的关系。8, 我们比较了几种架构对应的深度先验。我们在这里 (以及其他类似比较中) 的发现似乎表明, 拥有更深层次的架构是有益的, 而拥有对于识别任务 (例如语义分割) 非常有效的跳跃连接却是非常有害的。



图 7: 与最近两种修复方法的比较。顶部 – 在文本绘画示例中与 Shepard 网络 [27] 的比较。底部 – 与卷积稀疏编码 [25] 在修复 50% 丢失像素方面的比较。在这两种情况下, 我们的方法在各自论文中使用的图像上表现更好。

	Barbara	Boat	House	Lena	Peppers	C.man	Couple	Finger	Hill	Man	Montage
Papayan et al.	28.14	31.44	34.58	35.04	31.11	27.90	31.18	31.34	32.35	31.92	28.05
Ours	32.22	33.06	39.16	36.16	33.05	29.8	32.52	32.84	32.77	32.20	34.54

Table 1: Comparison between our method and the algorithm in [25]. See fig. 7 bottom row for visual comparison.

自然原像。[21]的自然原像方法是一种诊断工具, 用于研究有损函数 (例如在自然图像上运行的深度网络) 的不变性。令 Φ 为经过训练以执行图像分类等操作的神经网络的前几层。原像是导致相同表示 $\Phi(x_0)$ 的图像的集合 $\Phi^{-1}(\Phi(x_0)) = \{x \in X : \Phi(x) = \Phi(x_0)\}$ 。查看这个集合可以揭示网络丢失了哪些信息, 以及获得了哪些不变性。找到原像点可以表示为最小化数据项 $E(x; x_0) = \Phi(x) - \Phi(x_0)^2$ 。然而, 直接优化该函数可能会发现“伪影”, 即非自然图像, 网络的行为原则上是未指定的, 因此可以任意驱动它。通过将原像限制为一组 X 的自然图像 (在[21]中称为自然原像) 可以获得更有意义的可视化。

在实践中, 可以通过与上面看到的其他逆问题类似的正则化数据项来找到自然原像中的点。[21]的作者更喜欢使用电视范数, 它是弱自然图像先验, 但相对公正。相反, 诸如[8]之类的论文学习从示例中反转神经网络, 从而产生更好看的重建, 然而这可能偏向于学习数据驱动的反转先验。在这里, 我们建议使用深度图像先验 (2) 来代替。由于这是像电视规范一样手工制作的, 因此它不会偏向特定的训练集。另一方面, 它导致的反演至少与[8]的反演结果一样可解释。为了评估, 我们的方法与[22]和[8]的方法进行了比较。图 9 显示了通过考虑 AlexNet [17] 的渐进更深的子集而获得的反转表示 Φ 的结果: conv1、conv2、...、conv5、fc6、fc7 和 fc8。原像被发现

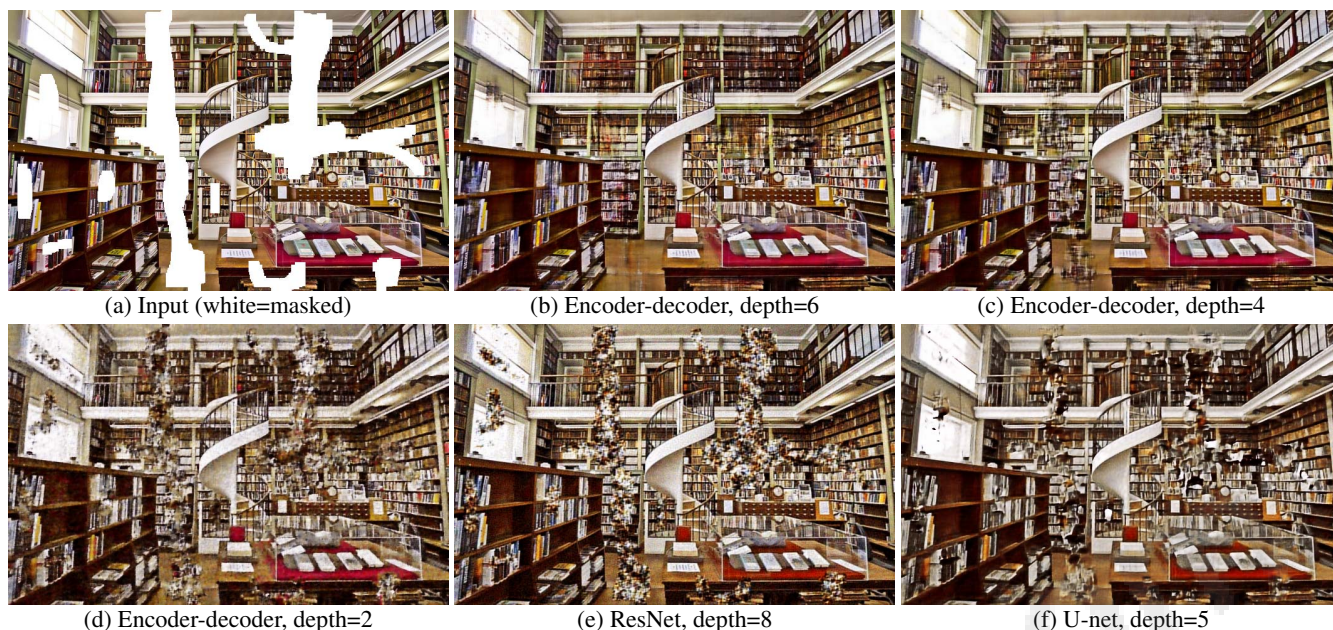


图 8: 使用不同深度和架构进行修复。该图显示, 通过使用更深的随机网络可以获得更好的修复结果。然而, 在 U-Net 中向 ResNet 添加跳跃连接是非常有害的。

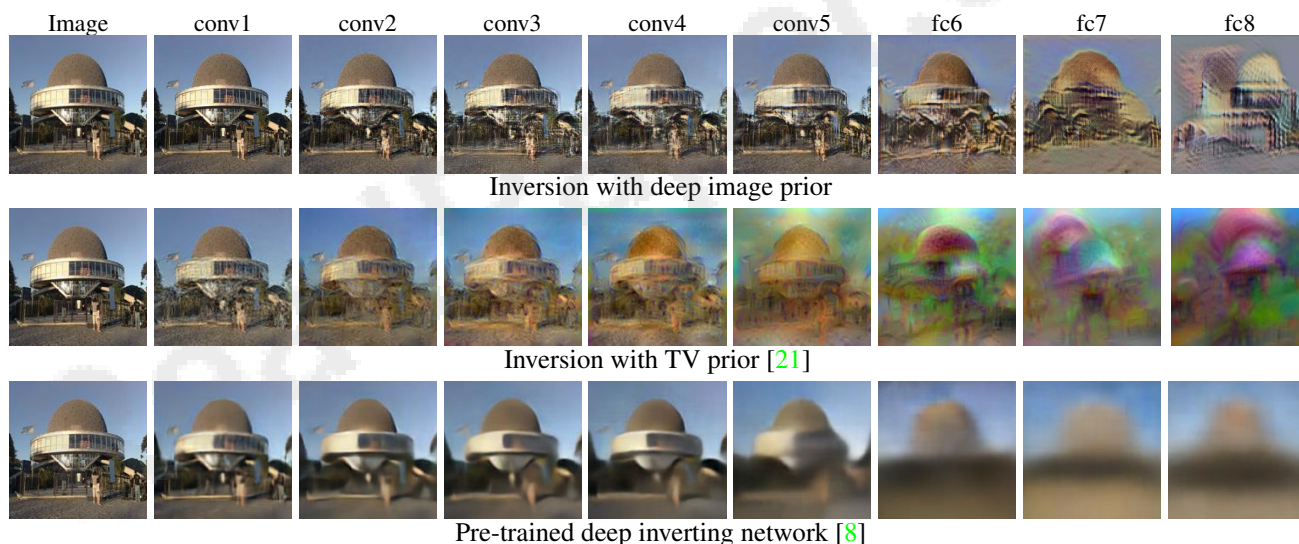


图 9: AlexNet 反演。给定左侧的图像, 我们展示了通过使用三种不同的正则化器反转 AlexNet 的不同层 (在 ImageNet ISLVR 上进行分类训练) 获得的自然原像: 深度图像先验、[21] 的 TV 范数先验和网络训练以反转保留集上的表示[8]。使用深度图像先验获得的重建在许多方面至少与[8]一样自然, 但它们不会因学习过程而产生偏差。

通过使用结构化先验优化 (2)。

如图所示。如图 9 所示, 与简单的 TV 范数相比, 我们的方法显著提高了图像清晰度。对于更深的层 (例如 fc6 和 fc7), 差异尤其显著, 其中电视范数仍然产生噪声图像, 而结构化正则化器产生的图像通常仍然是可解释的。我们的方法还比先前学习的方法产生了更多信息丰富的反演

的[8], 有明显的回归均值的趋势。

闪光-无闪光重建。虽然在这项工作中我们专注于单个图像恢复, 但所提出的方法可以扩展到多个图像的恢复任务, 例如用于视频修复的任务。因此, 我们通过定性示例总结了这组应用示例, 演示了如何将该方法应用于

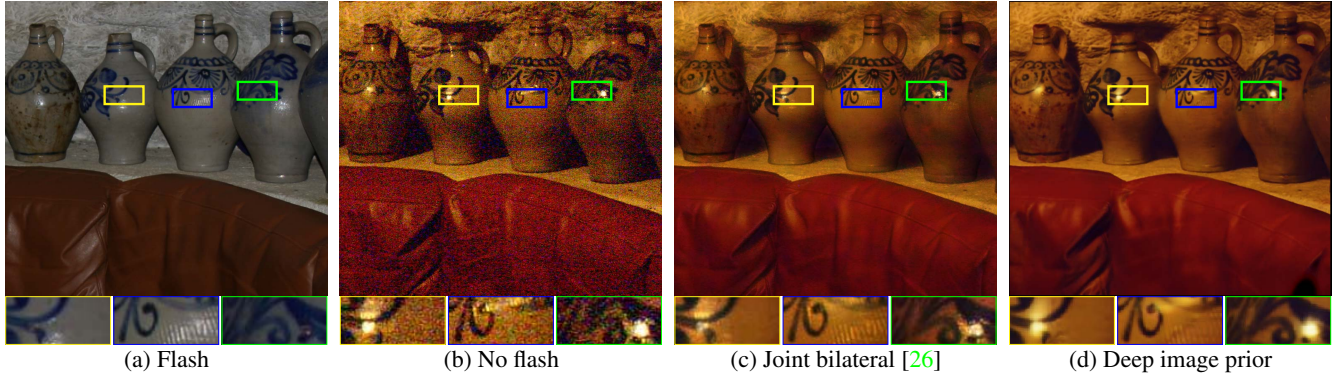


图 10: 基于闪光和无闪光图像对的重建。深度图像优先允许在照明非常接近无闪光图像的情况下获得低噪声重建。它在避免闪光灯对的照明图案“泄漏”方面比联合双边过滤更成功[26]（参见蓝色插图）。

基于图像对执行恢复。特别是，我们考虑基于闪光-无闪光图像对的恢复[26]，其目标是获得照明类似于无闪光图像的场景图像，同时使用闪光图像作为指导。一般来说，将该方法扩展到多个图像可能涉及对输入代码 z 进行一些协调优化，在我们的方法中，对于单图像任务，这些输入代码 z 通常保持固定和随机。在闪光-无闪光恢复的情况下，我们发现使用去噪公式（4）并使用闪光图像作为输入（代替随机向量 z ）可以获得良好的恢复效果。由此产生的方法可以被视为引导图像过滤的非线性推广[13]。恢复结果如图所示。10。

四、相关工作

我们的方法显然与基于可学习的ConvNets和上面引用的图像恢复和合成方法相关。同时，它与避免在保留集上进行训练的替代恢复方法组也有很大关系。该组包括基于损坏图像内相似斑块组联合建模的方法[4,6,10]，当损坏过程复杂且高度可变时（例如空间变化的模糊[1]），这些方法特别有用。该组中还有基于将字典拟合到损坏图像的补丁的方法 [23, 32] 以及基于卷积稀疏编码的方法 [31]，它们也可以将类似于浅层卷积网络的统计模型拟合到重建图像 [25]。工作 [20] 研究了将 Con-vNet 与基于自相似性的去噪相结合的模型，从而也桥接了两组方法，但仍然需要在保留集上进行训练。总体而言，先验强加于这项工作中研究的深层卷积网络似乎与基于自相似性和基于字典的先验高度相关。确实，作为

卷积滤波器的权重在图像的整个空间范围内共享，这确保了生成式 Con-vNet 可能产生的各个补丁的一定程度的自相似性。卷积网络和卷积稀疏编码之间的联系更加深入，并在[24]中在识别网络的背景下进行了研究，最近在[25]中进行了研究，其中提出了用于重建任务的单层卷积稀疏编码。然而，我们的方法与[25]（图7和表1）的比较表明，使用现代基于深度学习的方法中流行的深度ConvNet架构至少在某些情况下可能会导致更准确的恢复结果。

5. 讨论

我们研究了最近图像生成器神经网络的成功，将架构选择所施加的先验贡献与通过学习从外部图像传输的信息的贡献区分开来。作为副产品，我们已经证明，将随机初始化的卷积网络拟合到损坏的图像可以作为恢复问题的“瑞士刀”。虽然速度实际上很慢（每个图像需要几分钟的 GPU 计算），但这种方法不需要对退化过程进行建模或预训练。我们的结果与解释深度学习在图像恢复能力方面的成功的常见叙述相悖。学习而不是手工制作先验知识；相反，随机网络是更好的手工制作的先验，并且学习建立在这个基础上。这也验证了开发新的深度学习架构的重要性。

致谢。DU 和 VL 得到俄罗斯联邦教育和科学部的支持（拨款 14.756.31.0001），AV 得到 ERC677195-IDIU 的支持。

References

- [1] Y. Bahat, N. Efrat 和 M. Irani. 通过重新模糊实现非均匀盲去模糊。在过程中。CVPR, 第 3286-3294 页, 2017 年。[2] M. Bevilacqua, A. Roumy, C. Guillemot 和 M. Alberi-Morel. 基于非负邻域嵌入的低复杂度单图像超分辨率。在过程中。BMVC, 第 1-10 页, 2012 年。[3] P. Bojanowski, A. Joulin, D. Lopez-Paz 和 A. Szlam. 优化生成网络的潜在空间。CoRR, abs/1707.05776, 2017。[4] A. Buades, B. Coll 和 J.-M. 莫雷尔. 一种图像去噪的非局部算法。在过程中。CVPR, 第 2 卷, 第 60-65 页。IEEE, 2005 年。[5] H. C. Burger, C. J. Schuler 和 S. Harmeling. 图像去噪: 普通神经网络可以与 bm3d 竞争吗? 过程中。CVPR, 第 2392-2399 页, 2012 年。[6] K. Dabov, A. Foi, V. Katkovnik 和 K. Egiazarian. 通过稀疏 3 维变换域协同滤波进行图像去噪。IEEE 图像处理汇刊, 16(8):2080-2095, 2007。[7] C. Dong, C. C. Loy, K. He 和 X. Tang. 学习用于图像超分辨率的深度卷积网络。Proc.ECCV, 第 184-199 页, 2014 年。[8] A. 多索维茨基 (A. Dosovitskiy) 和 T. 布洛克斯 (T. Brox). 反转卷积网络与卷积网络一起工作。在过程中。CVPR, 2016 年。[9] A. Dosovitskiy, J. Tobias Springenberg 和 T. Brox. 学习用卷积神经网络生成椅子。过程中。CVPR, 第 1538-1546 页, 2015 年。[10] D. 格拉斯纳, S. 巴贡和 M. 伊拉尼. 单幅图像的超分辨率。在过程中。ICCV, 第 349-356 页, 2009 年。[11] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville 和 Y. Bengio. 生成对抗网络。在过程中。NIPS, 第 2672-2680 页, 2014 年。[12] S. Gu, W. Zuo, Q. Xie, D. Meng, X. Feng, and L. Zhang. 图像超分辨率的卷积稀疏编码。《ICCV》, 第 1823-1831 页。IEEE 计算机协会, 2015 年。[13] K. He, J. Sun 和 X. Tang. 引导图像过滤。T-PAMI, 35(6):1397-1409, 2013。[14] F. 海德, W. 海德里希和 G. 韦茨斯坦. 快速灵活的卷积稀疏编码。在过程中。CVPR, 第 5135-5143 页, 2015 年。[15] S. Iizuka, E. Simo-Serra 和 H. Ishikawa. 全局和局部一致的图像完成。ACM Transactions on Graphics (SIGGRAPH 报), 36(4):107:1-107:14, 2017 年。[16] D. P. Kingma 和 M. Welling. 自动编码变分贝叶斯。在过程中。国际劳工组织, 2014 年。[17] A. Krizhevsky, I. Sutskever 和 G. E. Hinton. 使用深度卷积神经网络进行图像网络分类。因 F. Pereira, C. J. C. Burges, L. Bottou 和 K. Q. Weinberger, 编辑, 《神经信息处理系统进展》25, 第 1097-1105 页。Curran Associates, Inc., 2012 年。[18] W.-S. 赖, J.-B. 黄, N. Ahuja 和 M.-H. 杨. Deeplaplacian 金字塔网络可实现快速、准确的超分辨率。在 IEEE 计算机视觉和模式识别会议 (CVPR), 2017 年 7 月。[19] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang 和 W. 施. 使用生成对抗网络实现逼真的单图像超分辨率。在 IEEE 计算机视觉和模式识别会议 (CVPR), 2017 年 7 月。[20] S. 莱夫基米蒂斯. 使用卷积神经网络进行非局部彩色图像去噪。在过程中。CVPR, 2016。[21] A. Mahendran 和 A. Vedaldi. 通过反转图像来理解深层图像表示。在过程中。CVPR, 2015。[22] A. Mahendran 和 A. Vedaldi. 使用自然原像可视化深度卷积神经网络。国际 JCV, 2016。[23] J. Mairal, F. Bach, J. Ponce 和 G. Sapiro. 矩阵分解和稀疏编码的在线学习。机器学习研究杂志, 11 (一月): 19-60, 2010。[24] V. Papan, Y. Romano 和 M. Elad. 通过卷积稀疏编码分析卷积神经网络。机器学习研究杂志, 18(83):1-52, 2017。[25] V. Papan, Y. Romano, J. Sulam 和 M. Elad. 通过本地处理进行卷积字典学习。载于 Proc.ICCV, 2017 年。[26] G. Petschnigg, R. Szeliski, M. Agrawala, M. F. Cohen, H. 霍普和 K. 富山. 使用闪光和非闪光图像对进行数码摄影。ACM 翻译. 图, 23(3):664-672, 2004。[27] J. S. J. Ren, L. Xu, Q. Yan, 和 W. Sun. Shepard 卷积神经网络。在过程中。NIPS, 第 901-909 页, 2015 年。[28] M.S.M. Sajjadi, B. Scholkopf 和 M. Hirsch. Enhancenet: 通过自动纹理合成实现单图像超分辨率。在 IEEE 国际计算机视觉会议 (ICCV), 2017 年 10 月。[29] Y. Tai, J. Yang 和 X. Liu. 通过深度递归残差网络实现图像超分辨率。在 IEEE 计算机视觉和模式识别会议 (CVPR), 2017 年 7 月。[30] D. 乌里扬诺夫, A. 维达尔迪和 V. 伦皮茨基. Deep ImagePrior 网页。https://dmitryulyanov.github.io/deep_image_prior。[31] M. D. Zeiler, D. Krishnan, G. W. Taylor 和 R. Fergus. 反卷积网络。在过程中。CVPR, 第 2528-2535 页, 2010 年。[32] R. Zeyde, M. Elad 和 M. Protter. 使用稀疏表示进行单图像缩放。《曲线与曲面》, 《计算机科学讲义》第 6920 卷, 第 711-730 页。施普林格, 2010。[33] C. 张, S. Bengio, M. Hardt, B. Recht 和 O. Vinyals. 理解深度学习需要重新思考泛化。在过程中。ICLR, 2017。