使用深先验的神经盲反卷积

一、摘要

盲解卷积是一个经典但具有挑战性的低层视觉问题，在许多实际应用中都有应用。

传统的基于最大后验概率（MAP）的方法严重依赖于固定的和手工制作的先验知识，这些先验知识在刻画清晰图像和模糊核（PSF）方面肯定是不够的，并且通常采用专门设计的交替最小化来避免琐碎的解。

相比之下，现有的深度运动去模糊网络从大量的训练图像中学习到干净图像或模糊核的映射，但在处理各种复杂和大尺寸的模糊核方面受到限制。

为了将MAP模型和deep模型联系起来，本文提出了两种生成网络，分别对清晰图像和模糊核的deep先验进行建模，并提出了一种无约束的神经优化盲反卷积方法。特别地，我们采用了一种带跳跃连接的非对称自动编码器来生成潜在的清晰图像，并采用了一种全连接网络（FCN）来生成模糊核。此外，在FCN的输出层应用非线性SoftMax以满足非负和等式约束。

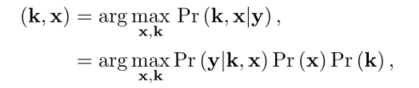
神经优化过程可以解释为生成网络的一种“zero-shot”自监督学习，因此我们提出的方法被称为SelfDeblur。实验结果表明，与基准数据集和真实世界模糊图像上最先进的盲反卷积方法相比，我们的SelfDeblur可以实现显着的定量增益以及视觉上更合理的去模糊结果。

二、简介

曝光过程中的相机抖动不可避免地会产生模糊图像，这是一个长期存在的恼人问题。从模糊图像中去除失真，即图像去模糊，是低水平视觉中一个典型的不适定问题，受到了相当多的研究关注。当模糊核具有空间不变性时，也称为盲解卷积，其中模糊图像可以表示为

y=k⊗x+n，⊗表示二维卷积算子，xis表示潜在的干净图像，kis表示模糊核，nis表示噪声级为σ的加性高斯白噪声（AWGN）。可以看出，blind deconvolution应该从模糊的图像y中估计出k和x，经过几十年的研究，这仍然是一个非常具有挑战性的问题。

大多数传统的盲反褶积方法都基于最大后验（MAP）框架。



其中，pr（y | k,x）是对应于保真度项的可能性，pr（x）和pr（k）分别建模干净图像和模糊核的先验。虽然对于X[2,15,27,56]和K[19,22,24,29,29,34,41,50,56]提出了许多先验知识，但它们通常是手工制作的，在描述干净的图像和模糊的内核方面肯定是不够的。此外，基于MAP的模型的非凸性也增加了优化的难度。

贡献如下：

提出了一种神经盲去卷积方法,i.e.,SelfDeblur, DIP和FCN分别用来捕获干净图像和模糊内核的优先级。将非线性SoftMax应用于FCN的输出层满足非负和等式约束。

•提出联合优化算法解决无约束神经盲反卷积模型用于估计模糊核和生成潜在干净的图像，使非盲反卷积对于我们的 SelfDeblur 不是强制性选择。

•大量实验表明，我们的 SelfDeblur 在定量和定性评估方面优于现有的基于 MAP 的方法。据我们所知，SelfDeblur首次尝试应用深度学习来产生最先进的盲反卷积性能。

在本节中，我们简要介绍了相关工作，包括基于优化的盲反卷积和基于深度学习的盲去模糊方法。

1）基于优化的盲反卷积

传统的基于优化的盲反卷积方法可以进一步分为两类，即基于变分贝叶斯 (VB) 和基于 MAP 的方法。 基于 VB 的方法 [20] 在理论上很有前途，但计算成本很高。 对于基于 MAP 的方法，已经提出了许多先验方法来对干净图像和模糊内核进行建模。

在 [2] 的开创性工作中，Chan 等人。引入总变差 (TV) 正则化来模拟盲反卷积中的潜在干净图像，并基于梯度的先验激发几个变体，例如，'0-norm [27] 和'p-norm [56]。其他专门设计的正则化，例如，'1/'2-norm [15]、基于补丁的先验 [24,41]、低秩先验 [34] 和暗通道先验 [29,50] 也被提出来识别和保留有利于模糊核估计的显着边缘。

最近，提出了一种判别先验[21]来区分干净图像和模糊图像，但仍然严重依赖'0范数正则化器来获得最先进的性能。对于 Pr(k)，通常采用梯度稀疏先验 [19,29,56] 和光谱先验 [22]。为了解决基于 MAP 的模型，投影交替最小化算法引入了一些技巧，包括模糊核的延迟归一化 [32]、多尺度实现 [15] 和时变参数[56]。

在模糊核估计之后，需要非盲反卷积来恢复具有精细纹理细节的潜在清洁图像[15]。 因此，先验应该有利于自然图像，例如超拉普拉斯[14]、GMM [55]、非局部相似性[5]，例如RTF [37]、CSF [36]和CNN[16,53]， 这与模糊核估计中使用的那些完全不同。 我们的 SelfDeblur 可以看作是一种特殊的基于 MAP 的方法，但是采用了两个生成网络，即 DIP 和 FCN，分别捕获干净图像和模糊核的深层先验。 此外，联合优化算法对于估计模糊核和生成清晰图像是有效的，使得非盲反卷积不是SelfDeblur的强制选择。

2）图像去模糊中的深度学习

许多研究将深度学习（DL）应用于盲去模糊。 例如，DL 可用于帮助学习映射到模糊核。通过模仿基于优化的方法中的交替最小化步骤，Schuleret al。 [38] 设计用于模糊核估计的深度网络架构。通过研究模糊图像的光谱特性，建议使用深度 CNN 来预测傅里叶系数 [1]，然后可以将其投影以估计模糊核。 在 [40] 中，CNN用于预测运动模糊图像的参数模糊核。

对于动态场景去模糊，已经开发了深度 CNN 来学习直接映射到潜在的干净图像 [17,25,40,43,51]。受盲反卷积中的多尺度策略的启发，提出了多尺度 CNN [25] 和尺度循环网络 [43] 来直接从模糊图像中估计潜在的干净图像。为了更好地恢复运动去模糊中的纹理细节，还引入了对抗性损失 [17]。此外，通过利用相邻帧之间的时间信息，深度网络也被应用于视频运动去模糊[9,26,30]。然而，由于大尺寸和复杂的模糊核导致的严重不适定性，现有的基于深度学习的方法在盲反卷积方面仍然无法超越传统的基于优化的方法。

最近，DIP [45] 和 Double-DIP [6] 被引入来捕获图像统计数据，并已部署到许多低级视觉任务，如超分辨率、绘画、去雾、透明分离等。尽管如此，DIP 网络在捕获模糊核的先验方面受到限制，并且 Double-DIP 在盲反卷积方面仍然表现不佳。据我们所知，我们的 SelfDeblur 首次尝试应用深度网络来产生最先进的盲反卷积性能。

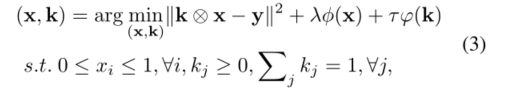
3.提出的方法

在本节中，我们首先介绍了基于MAP的盲反卷积的一般公式，然后介绍了我们提出的神经盲反卷积模型以及联合优化算法。

3.1基于MAP的盲反卷积的一般公式

根据 Eqn(1)，我们将保真度定义为，−log (Pr (y|k,x)) =kk⊗x−yk2。

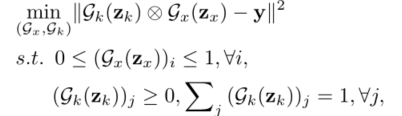
我们进一步引入了两个正则化项-log(Pr(x)) =φ(x) 和-log(Pr(k)) =φ(k)，分别用于对潜在干净图像和模糊核的先验建模。 Eqn(2) 中基于 MAP 的盲反卷积模型可以重新表述为，



其中λ和τ是权衡正则化参数。 除了这两个正则化项外，我们进一步引入了模糊核 k[29,32,41,56] 的非负和等式约束，并且像素点也被限制在范围 [0,1] 内。其中R（x）是捕获自然图像统计信息的正则化器，与φ（x）有很大不同。

3.2神经盲反卷积

受 DIP [45] 和 Double-DIP [6] 成功的启发，我们提出了神经盲反卷积模型，采用生成网络 GxandGk 来捕获 xandk 的先验。 通过用 Gx 和 Gk 替换 x 和 k 并去除正则化项 φ(x) 和 φ(k)，神经盲反卷积可以表示为：



其中 zx 和 zk 是从均匀分布中采样的，(·)i 和 (·)j 表示第 i 个和第 j 个元素。 我们注意到zk 为一维向量，对 Gk(zk) 进行整形得到二维矩阵模糊内核。

然而，神经盲反卷积仍有几个问题需要解决。

(i) DIP 网络 [45] 旨在捕获低级图像统计数据，并且在捕获模糊内核的先验方面受到限制。 结果，我们凭经验发现 Double-DIP [6] 在盲反卷积方面表现不佳（参见第 4.1.2 节中的结果）。

(ii) 由于非负和等式约束，方程中的结果模型。 (5) 是一个受约束的神经优化问题，难以优化。

(iii) 尽管生成网络 Gx 和 Gk 对图像噪声具有高阻抗，但 DIP 的去噪性能在很大程度上依赖于最后一次迭代和不同优化运行的额外平均 [45]。然而，这种启发式解决方案既带来了更多的计算成本，也不能直接用于处理模糊和嘈杂的图像。

在下文中，我们通过设计适当的生成网络 Gx 和 Gk 来提出解决问题 (i) 和 (ii) 的解决方案。至于问题 (iii)，我们引入了一个额外的 TV 正则化器和一个正则化参数，以明确考虑神经盲反卷积模型中的噪声水平。

**生成网络 Gx**。潜在的干净图像通常包含显着的结构和丰富的纹理，这需要生成网络 Gx 具有足够的建模能力。幸运的是，自从引入生成对抗网络 [8] 以来，在生成高质量自然图像方面取得了巨大进展 [45]。对于 x 建模，我们采用 DIP 网络，即在 [45] 中具有跳跃连接的非对称自动编码器 [35]，作为 Gx。如图 2 所示，前 5 层编码器跳过连接到后 5 层解码器。最后，卷积输出层用于生成潜在的干净图像。为了满足范围约束，Sigmoid 非线性应用于输出层。有关 Gx 的更多架构细节，请参阅补充文件。

**生成网络 Gk**。一方面，DIP 网络 [45] 旨在捕获自然图像的统计信息，但在模糊核的先验建模方面表现有限。 另一方面，blur kernelk通常包含的信息比latent clean imagex少得多，并且可以通过更简单的生成网络很好地生成。 因此，我们简单地采用全连接网络（FCN）作为

克。 如图 2 所示，FCN Gk 以 200 维的一维噪声 zk 作为输入，具有 1000 个节点的隐藏层和 K2 个节点的输出层。 为了保证始终满足非负和等式约束，将 Soft Max非线性应用于Gk的输出层。 最后，将 K2 个条目的 1D 输出重新整形为 2DK×K 模糊核。 请参阅增刊。 更多Gk 的架构细节。

**具有 TV 正则化的无约束神经盲反卷积。**通过上述生成网络 Gx 和 Gk，我们可以将神经盲反卷积制定为无约束的优化形式。 然而，生成的模型与噪声水平无关，使其在具有不可忽略噪声的模糊图像上表现不佳。 为了解决这个问题，我们结合 Gx 和 TV 正则化来捕获图像先验，然后我们的神经盲反卷积模型可以写成



其中 λ 表示由噪声水平σ 控制的正则化参数。尽管生成网络 Gx 更强大，但 Gx 和另一个图像先验的结合通常有利于反卷积性能。此外，引入与噪声水平相关的正则化参数 λ 可以大大提高处理具有各种噪声水平的模糊图像的鲁棒性。

特别是，我们在我们的实现中经验性地设置λ= 0.1×σ，并且可以使用[54]估计噪声水平σ。

3.3. 优化算法

Eqn(6) 的优化过程可以解释为一种“零样本”自我监督学习 [39]，其中生成网络 Gk和Gx仅使用测试图像（即模糊图像）进行训练，没有提供真实干净的图像。 因此，我们的方法被称为 SelfDeblur。 在下文中，我们提出了两种用于 SelfDeblur 的算法，即交替优化和联合优化。

**交替优化**。 类似于传统盲反卷积中的交替最小化步骤[2,27,29,41,56]，Gk 和 Gx 的网络参数也可以交替优化。 正如算法 1 中总结的那样，Gk 的参数通过 ADAM [13] 通过固定 Gx 来更新，反之亦然。 特别是，梯度 w.r.t. Gx 或 Gk 都可以使用自动微分导出 [31]。

**联合优化**。在传统的基于MAP的框架中，交替最小化允许使用投影算子来处理非负和等式约束，并修改优化细节以避免琐碎求解，因此被广泛采用。对于我们的神经盲反卷积，Eqn(6) 中的模型是无约束优化的，Gk 和 Gxis 强大的建模能力有利于避免琐碎的delta 核解。我们还注意到，无约束的神经盲反卷积是高度非凸的，交替优化可能会卡在鞍点 [44]。因此，对于 SelfDeblur，联合优化比交替优化更受青睐。使用自动微分技术 [31]，可以导出 w.r.t.Gk 和 Gx 的梯度。算法2总结了联合优化算法，其中Gk和Gx的参数可以使用ADAM算法联合更新。我们在第 4.1.1 节中的实证研究还表明，联合优化通常会收敛到比交替优化更好的解决方案。

当达到 T 次迭代时，交替优化和联合优化算法都会停止。 然后，可以分别使用 k=GTk(zk) 和 x=GTx(zx) 生成估计的模糊核和潜在的干净图像。 受益于 Gx(zx) 的建模能力，估计的 x 具有视觉上有利的纹理，我们的 SelfDeblur 采用另一种非盲反卷积方法来生成最终的去模糊结果并不是强制性的选择。

4.实验结果

本节首先进行消融研究，分析优化算法和网络架构的效果。 然后，我们的 SelfDeblur 在两个基准数据集上进行评估，并与最先进的盲反卷积方法进行比较。 最后，我们报告了 SelfDeblur 在几个真实世界模糊图像上的结果。

我们的 SelfDeblur 是使用 Pytorch [31] 实现的。 实验在配备一个 NVIDIA Titan V GPU 的 PC 上进行。 除非特别说明，否则实验遵循相同的设置，即 T = 5,000，并且噪声 zx 和 zk 是从具有固定随机种子 0 的均匀分布中采样的。在 [45] 之后，我们在每次迭代中进一步随机扰动 zx。 初始学习率设置为 0.01，并在达到 2,000、3,000 和 4,000 次迭代时乘以0.5 衰减。

4.1消融研究

Levinet 等人对数据集进行了消融研究[19]，这是一个流行的盲反卷积基准，由4个干净的图像和8个模糊内核组成。 使用[54]，数据集中模糊图像的平均估计噪声水平为σ≈1×10−5。 因此，我们在这个数据集上简单地采用λ= 1×10−6。

4.1.1 交替优化与联合优化

我们首先使用交替优化 (SelfDeblur-A) 和联合优化 (SelfDeblur-J) 来评估 SelfDeblur 的性能。 表 1 报告了平均 PSNR 和 SSIM 值。 在定量指标方面，SelfDeblur-J 明显优于 SelfDeblur-A，证明了联合优化的优越性。 在补充文件中，我们提供了SelfDeblur-A的几个失败案例。其中 SelfDeblur-A 可能会收敛到 delta 内核和更差的解决方案，而 SelfDeblur-J 在这些情况下表现良好。因此，在接下来的实验中，采用联合优化作为默认的自去模糊方法。

4.1.2 Gk的网络架构

在这个实验中，我们通过考虑四种网络架构来比较结果：（i）SelfDeblur，（ii）Double-DIP [6]（Gx和Gk都具有跳过连接的非对称自动编码器），（iii）SelfDeblurk-（去除隐藏层fromGk）和（iv）SelfDeblurk+（为Gk添加一个额外的隐藏层）。从表 2 和图 4 可以看出，SelfDeblur 在估计模糊核和潜在图像方面明显优于 Double-DIP。

结果表明 DIP 网络仅限于捕捉模糊核的先验，简单的 FCN 可以作为 Gk 的一个很好的选择。我们进一步将 SelfDeblur 与 SelfDeblurk− 和 SelfDeblurk+ 进行比较。可以看出，没有隐藏层的 FCN（即 SelfDeblurk−）也成功地估计了模糊核和干净图像（见图 4），但性能远不如 SelfDeblur。此外，三层 FCN（即SelfDeblurk+）优于SelfDeblurk−，但不如SelfDeblur。综上所述，SelfDeblur 是先验模糊核建模的不错选择。

4.1.3 中间结果的可视化

使用来自 Levinet 的数据集的图像。 [19]，图 3 显示了在迭代 t=1,20,100,600,2,000 和 5,000 时估计的模糊核和干净图像的中间结果，以及 k 的 MSE 曲线和 x 的 PSNR 曲线。 当迭代 t=20 时，x 的中间结果主要包含显着图像结构，这与传统方法中显着边缘对于初始模糊核估计至关重要的观察结果一致。 随着迭代次数的增加，Gx和Gk开始在x和k中生成更精细的细节。 与传统方法不同，SelfDeblur 在迭代 t≥20 时有效地同时估计模糊核和恢复潜在的干净图像，使得非盲反卷积不是 SelfDeblur 的强制选择。

图像的模糊有多种原因，噪声

非盲去卷积

预处理：图像复原

图像复原算法可以分类线性和非线性两大类，其中非线性耗费时间长，考虑到花卉识别过程中的实时性，采用线性

去模糊

失焦模糊和运动模糊

噪声：光线、云雾、灰尘

微医 195

联易融 4500

奇安信 280