

杜智颜, 王海燕. 基于总广义变分与深度图像先验的图像修补模型[J]. 智能计算机与应用, 2025, 15(9): 168–175. DOI: 10.20169/j. issn. 2095-2163. 250926

# 基于总广义变分与深度图像先验的图像修补模型

杜智颜, 王海燕

(南京邮电大学 理学院, 南京 210023)

**摘要:** 从不完整的测量中恢复图像是具有挑战性的, 特别是当图像包含重要的细节和特征时。在现实中, 通常只有退化的图像, 而没有真实的图像。针对上述问题, 本文引入了一种新的变分模型, 该模型融合了2种先验: 总广义变分(TGV)和深度图像先验(DIP)。TGV是一种正则化方法, 可以实现在保持锐利边缘的同时减少伪影。DIP建议使用深度网络作为逆问题的调节器, 允许从卷积神经网络中学习图像恢复所需的先验信息, 消除了对大量图像对的需求。该模型结合了DIP和TGV的优点, 即在训练过程中不仅不需要大量的图像对, 而且可以有效地恢复图像的边缘和细节。采用交替方向乘子法(ADMM)求解该模型。本文的方法在图像修补任务的定量和定性结果方面超越了各种最先进的方法。

**关键词:** 图像修补; 深度图像先验; 广义全变分; 保留细节; 交替方向乘子法

中图分类号: TP391.41

文献标志码: A

文章编号: 2095-2163(2025)09-0168-08

## An image inpainting model combining total generalized variation and deep image prior

DU Zhiyan, WANG Haiyan

(College of Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

**Abstract:** Recovering images from incomplete measurements can be challenging, especially when the images contain important details and features. In reality, there usually are degraded images rather than real ones. To solve these problems, this paper introduces a new variational model, which combines two kinds of priors: Total Generalized Variational (TGV) and Depth Image Prior (DIP). TGV is a regularization method that enables the reduction of artifacts while maintaining sharp edges. DIP proposes using a deep network as a regulator for inverse problems, allowing learning a priori information required for image recovery from convolutional neural networks, eliminating the need for a large number of image pairs. The model combines the advantages of DIP and TGV, that is, it does not require a large number of image pairs during the training process, and can effectively recover the edges and details of the image. The Alternating Direction Method of Multipliers (ADMM) is used to solve the model. The proposed method exceeds various state-of-the-art methods in terms of quantitative and qualitative results for image inpainting tasks.

**Key words:** image inpainting; deep image prior; generalized total variation; keep details; alternate direction multiplier method

## 0 引言

由于成像技术的广泛采用, 高质量图像逐渐成为传输数据的重要媒介<sup>[1]</sup>。然而, 在实际应用中, 图像的生成和接收过程容易受到各种内部或外部因素的干扰, 从而不可避免地导致图像的退化, 如像素丢失造成的退化。为了从退化的图像中获得高质量的图像, 图像修补<sup>[2]</sup>在数字图像处理中得到了迅速

发展。

针对信息像素缺失导致的图像退化问题, 人们提出了多种方法, 可分为3类: 传统方法、基于深度学习的方法和基于深度图像先验(DIP)的方法。传统的图像修复技术大致可以分为2类: 基于扩散方法<sup>[3-4]</sup>和基于补丁方法<sup>[5-6]</sup>。基于扩散的方法首先从缺失区域的邻近部分选择有用的信息, 然后将其应用于缺失区域。相比之下, 基于补丁的方法通过

基金项目: 2022年江苏省研究生科研与实践创新计划项目(SJCX22\_0245, SJCX22\_0248, KYCX22\_0897)。

作者简介: 杜智颜(1999—), 女, 硕士研究生, 主要研究方向: 图像处理, 计算机视觉。Email: 1221087301@njupt.edu.cn; 王海燕(1999—), 女,

硕士研究生, 主要研究方向: 图像处理, 计算机视觉。

收稿日期: 2024-01-26

哈尔滨工业大学主办 ◆ 科技创见与应用

寻找缺失和非缺失区域之间的相似特征来填充缺失的像素。

毫无疑问, 传统方法在一定程度上促进了图像恢复的发展。然而, 这些方法的局限性在于仍然依赖于更专业的理论知识来提取特征, 并且参数调整过程耗时且繁琐。为了解决这些问题, 出现了各种基于深度学习的方法。对于图像修补任务, 神经网络得到了广泛的应用。早期的卷积神经网络能够修复小区域<sup>[7]</sup>, 而 CE<sup>[8]</sup>是第一个修复大面积缺失图像的。VGG 利用预训练网络提取特征<sup>[9]</sup>, 在图像背景中寻找相似的 patch 来填充缺失区域, 在图像背景中找到相似的 patch 来填充缺失区域。RFRNet<sup>[10]</sup>借鉴前人的经验, 设计了一个圆形的特征推理网络, 可以有效地提高图像修补的效果。

使用基于深度学习的方法在某些方面优化了传统方法, 然而训练神经网络需要大量退化和干净的图像对, 这在数据稀缺的领域提出了挑战。为了克服这个问题, 推荐了一种称为深度图像先验 (DIP)<sup>[11]</sup> 的无监督方法。DIP 是一种正则化技术, 并提出了使用深度网络本身作为反问题的正则化器, 而不是早期的监督方法。从本质上讲, DIP 摆脱了显式正则化, 并从自己的网络中学习恢复图像所需的信息。自引入以来, 该领域的研究一直不断地取得重要突破。DeepRED<sup>[12]</sup>通过添加显式先验来改进 DIP, 并结合去噪正则化 (RED) 的概念。GpDIP<sup>[13]</sup>进行的研究提供了在特定情况下 DIP 收敛于高斯过程 (GP) 的证据。DIP-TV<sup>[14]</sup> 和 DIP-VBTV<sup>[15]</sup> 分别证明了将总变分 (TV) 和矢量束总变分 (VBTV) 正则化纳入 DIP 框架可以提高 DIP 的有效性。这些新颖的基于 DIP 的方法代表了该技术在图像处理任务中的重要进步。

尽管基于 DIP 的方法取得了重大进展, 但仍有亟待改进空间, 特别是在图像的细节保存方面。为了解决这个问题, 许多基于 DIP 的方法将隐式正则化与显示正则化相结合, 例如 DIP-TV<sup>[14]</sup> 和 DIP-VBTV<sup>[15]</sup>。然而, 这些显示正则化的视觉效果并不显著, 因为会在图像的平滑区域造成阶梯效应, 导致重要细节的丢失。为了保留更多的细节, 本文建议使用 TGV 作为显示正则化。Bredies 等学者<sup>[16]</sup>首先提出了一种基于 TGV 的模型来处理高斯噪声下的模糊图像恢复问题。在恢复过程中, 该模型能有效克服阶梯效应, 保护图像的边缘结构信息。从那时起, 许多研究人员都利用总广义变分 (TGV) 正则化<sup>[17~20]</sup>来对其研发方法进行优化。

为了弥补 DIP 的方法在保留细节方面的不足, 本文提出了一种结合 DIP 和 TGV 的图像修补变分模型。本文提出方法的主要贡献可以概括如下:

(1) 建议引入显式 TGV 正则化来改进基于 DIP 的方法。

(2) 采用经典的 ADMM 算法来解决问题。

(3) 大量的实验结果表明, 该方法具有较强的泛化能力。

## 1 预备知识

### 1.1 深度图像先验

近年来, 出现了一种新的图像修复技术, 称为深度图像先验 (DIP)<sup>[11]</sup>。该方法通过利用未训练的 CNN 架构的归纳偏差, 消除了对大量数据对的需求。只给定退化的图像, DIP 就能生成相应的清晰图像。DIP 相关的优化问题可以表述如下:

$$\begin{cases} \theta = \arg \min_{\theta} \frac{1}{2} \| M(T_{\theta}(z)) - s \|_2^2, \\ u = T_{\theta}(z) \end{cases} \quad (1)$$

其中,  $M$  表示退化算子;  $z$  表示具有 3 个通道的随机图像; 作为网络  $T$  的初始输入,  $\theta$  表示网络  $T$  的参数;  $s$  表示观察到的退化图像,  $u$  为恢复图像。

DIP 在现实环境中具有优势, 因其有任何训练数据。此外, DIP 与依赖于数据的一些复杂技术在恢复质量的效果方面相当。

### 1.2 广义全变分正则

通过解决 TV 的阶梯效应问题<sup>[16]</sup>, 广义全变分 (TGV) 正则化对 TV 正则化进行了推广。结果表明, TGV 正则化比 TV 正则化保留了更多的纹理信息, 产生了更令人满意的视觉效果。 $k$  阶 TGV 定义为:

$$\text{TGV}_{\gamma}^k(u) = \sup \left\{ \int_{\Omega} u \operatorname{div}^k v dx \mid v \in C_c^k(\Omega, \operatorname{Sym}^k(\mathbb{R}^d)) \right. \\ \left. \| \operatorname{div}^l v \|_{\infty} \leq \gamma_l, l = 0, \dots, k-1 \right\} \quad (2)$$

其中,  $\operatorname{Sym}^k(\mathbb{R}^d)$  表示阶为  $k$  的对称张量的空间, 其参数为  $\mathbb{R}^d$ ;  $\gamma_l (l = 0, \dots, k-1)$  表示固定的正参数。TGV 可以被解释为“稀疏的”从第一个到第  $k$  个分配导数的最优平衡的惩罚。特别地, 当  $k = 1$  时,  $\text{TGV}_1^1$  等价于 TV。当  $k = 2$  时, 二阶 TGV 可以表示为:

$$\text{TGV}_{\gamma}^2(u) = \sup \left\{ \int_{\Omega} u \operatorname{div}^2 \omega dx \mid \omega \in C_c^2(\Omega, S^{d \times d}) \right. \\ \left. \| \omega \|_{\infty} \leq \gamma_0, \| \operatorname{div} \omega \|_{\infty} \leq \gamma_1, l = 0, \dots, k-1 \right\} \quad (3)$$

在  $k > 2$  的情况下,  $\text{TGV}_{\gamma}^k(u)$  称为高阶 TGV。

由于高阶 TGV 涉及高阶散度算子,计算复杂度较高,本文在模型中采用了二阶 TGV。

图像的数字表示大多是离散的,因此为了更好地将 TGV 正则化应用到图像处理中,本文引入了离散 TGV。当  $\operatorname{div} w = v$ ,  $U = C_c^2(\Omega, \mathbb{R})$ ,  $V = C_c^2(\Omega, \mathbb{R}^2)$ ,  $W = C_c^2(\Omega, S^{2 \times 2})$ , 离散  $\operatorname{TGV}_\gamma^2$  可以被写为:

$$\left\{ \begin{array}{l} \operatorname{TGV}_\gamma^2(u) = \max_{v \in V, w \in W} \{ \langle u, \operatorname{div} v \rangle \mid \operatorname{div} w = v \\ \|w\|_\infty \leq \gamma_0, \|v\|_\infty \leq \gamma_1 \} \end{array} \right. \quad (4)$$

$$\text{其中, } \operatorname{div} w = \begin{cases} \partial_x w_{11} + \partial_y w_{12} \\ \partial_x w_{21} + \partial_y w_{22} \end{cases}$$

闭集  $B$  的指示函数可定义为:

$$I_B = \begin{cases} 0, & x \in B \\ \infty, & \text{else} \end{cases} \quad (5)$$

根据  $I_{[0]}(\cdot) = -\min_y \langle y, \cdot \rangle$ , 则离散  $\operatorname{TGV}_\gamma^2(u)$  的表示为:

$$\begin{aligned} \operatorname{TGV}_\gamma^2(u) = & \min_{q \in V} \max_{\substack{\|w\|_\infty \leq \gamma_0, w \in W \\ \|v\|_\infty \leq \gamma_1, v \in V}} \langle u, \operatorname{div} v \rangle + \langle q, v - \operatorname{div} w \rangle = \\ & \min_{q \in V} \max_{\substack{\|w\|_\infty \leq \gamma_0, w \in W \\ \|v\|_\infty \leq \gamma_1, v \in V}} \langle -\nabla u, v \rangle + \langle q, v \rangle + \\ & \langle \bar{B}(q), w \rangle \end{aligned} \quad (6)$$

利用约束  $\{\|v\|_\infty \leq \gamma_1\}$  关于零的对称性, 将  $v$  替换为  $-v$ , 最终得到  $\operatorname{TGV}^2$  的离散形式:

$$\begin{aligned} \operatorname{TGV}_\gamma^2(u) = & \min_{q \in V} \max_{\substack{\|w\|_\infty \leq \gamma_0, w \in W \\ \|v\|_\infty \leq \gamma_1, v \in V}} \langle \nabla u - q, v \rangle + \langle \bar{B}(q), \\ & w \rangle = \min_{q \in V} \gamma_1 \|\nabla u - q\|_1 + \\ & \gamma_0 \|\bar{B}(q)\|_1 \end{aligned} \quad (7)$$

### 1.3 ADMM 算法

ADMM 算法<sup>[21]</sup>是解决具有线性约束的优化问题的经典方法。当面对复杂的目标函数时,直接找

到最优解是一项挑战。然而,利用 ADMM 算法,人们可以将一个困难的优化问题分解成一组更简单的子问题,通过迭代求解这些子问题进而获得最优解。这一过程使 ADMM 算法能够解决具有挑战性的优化问题。典型地,ADMM 算法可以解决的优化问题有:

$$\begin{cases} \min_{y,z} f_1(y) + f_2(z) \\ \text{s.t. } y + Az + B = 0 \end{cases} \quad (8)$$

其中,  $y \in \mathbb{R}^m$ ;  $z \in \mathbb{R}^q$ ;  $A \in \mathbb{R}^{m \times q}$ ;  $B \in \mathbb{R}^m$ ;  $f_1: \mathbb{R}^p \rightarrow \mathbb{R}$ ;  $f_2: \mathbb{R}^q \rightarrow \mathbb{R}$ 。根据线性约束  $y + Az + B = 0$ , 问题的拉格朗日泛函方程可表示为:

$$L_\xi(y, z, \lambda) = f_1(y) + f_2(z) + \frac{\xi}{2} \|y + Az + B\|_2^2 + \lambda^\top (y + Az + B) \quad (9)$$

其中,  $\lambda \in \mathbb{R}^k$ ,  $\lambda$  称为拉格朗日乘子,  $\xi$  表示惩罚参数。然后,变量  $y, z$  和  $\psi$  根据以下公式迭代更新:

$$\begin{aligned} y^{(k+1)} &= \arg \min_y f_1(y) + \frac{\xi}{2} \|y + Az^k + B + \psi^k\|_2^2 \\ z^{(k+1)} &= \arg \min_z f_2(z) + \frac{\xi}{2} \|y^{k+1} + Az + B + \psi^k\|_2^2 \\ \psi^{(k+1)} &= \psi^{(k)} + y^{(k+1)} + Az^{(k+1)} + B \end{aligned} \quad (10)$$

其中,  $\psi = \frac{\xi}{\lambda}$ 。最后,通过交替更新  $y, z$  和  $\psi$ ,

可以获得  $y$  和  $z$  的最优解。

## 2 提出 DIP-TGV 模型

在本文中,提出了一种结合 DIP 和 TGV 的彩色图像修补模型,本文方法流程框架如图 1 所示。

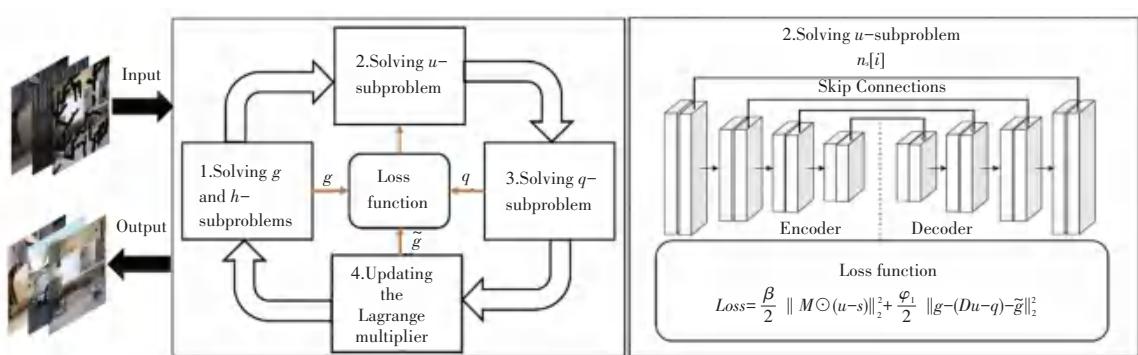


图 1 DIP-TGV 框架图

Fig. 1 The framework diagram of DIP-TGV

根据前面介绍的 DIP 和二阶 TGV 的公式, 提出了以下的一般图像恢复模型:

$$\begin{aligned} \theta &= \arg \min_{\theta} \frac{\beta}{2} \| M \odot (u - s) \|_2^2 + \text{TGV}_{\gamma}^2(u) \\ \text{TGV}_{\gamma}^2(u) &= \gamma_1 \| \nabla u - q \|_1 + \gamma_0 \| \bar{\mathbf{B}}(q) \|_1, \gamma_1 > 0, \gamma_0 > 0 \\ u &= T_{\theta}(z) \end{aligned} \quad (11)$$

其中,  $\text{TGV}_{\gamma}^2(u)$  表示二阶 TGV;  $M$  是一个二进制掩码, 并且取集合  $\{0, 1^{H \times W}\}$  中的值。文中通过  $D_1^u$  和  $D_2^u$  近似  $\nabla_1^u$  和  $\nabla_2^u$ , 其中  $D_1^u$  和  $D_2^u$  分别是沿  $x$  轴和  $y$  轴的正向有限差分算子。将  $\bar{\mathbf{B}}(q)$  近似为  $\mathbf{B}(q)$  后, 本文提出的模型可以改写为:

$$\begin{aligned} \theta &= \arg \min_{\theta} \frac{\beta}{2} \| M \odot (u - s) \|_2^2 + \text{TGV}_{\gamma}^2(u) \\ \text{TGV}_{\gamma}^2(u) &= \gamma_1 \| Du - q \|_1 + \gamma_0 \| \mathbf{B}(q) \|_1, \gamma_1 > 0, \gamma_0 > 0 \\ u &= T_{\theta}(z) \end{aligned} \quad (12)$$

其中,

$$\mathbf{B}(q) = \begin{pmatrix} \hat{\mathbf{e}}_1^T q_1 & \frac{1}{2} (D_2 q_1 + D_1 q_2) \\ \hat{\mathbf{e}}_1^T q_2 & \frac{1}{2} (D_2 q_1 + D_1 q_2) \\ \hat{\mathbf{e}}_2^T q_1 & D_2 q_2 \end{pmatrix}$$

### 3 求解 DIP-TGV 模型

ADMM 算法可以很容易地求解本文提出的模型。引入辅助变量  $\mathbf{g} = \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} \in V$ ,  $\mathbf{h} = \begin{pmatrix} h_1 & h_3 \\ h_3 & h_2 \end{pmatrix} \in$

$W$ , 并把式(12)重写为:

$$\begin{aligned} \min_{\mathbf{g}, \mathbf{h}, u, q} \frac{\beta}{2} \| M \odot (u - s) \|_2^2 + \gamma_1 \| \mathbf{g} \|_1 + \gamma_0 \| \mathbf{h} \|_1 \\ \text{s. t. } \mathbf{g} = Du - q, \mathbf{h} = \mathbf{B}(q) \end{aligned} \quad (13)$$

根据描述的 ADMM 算法, 引入拉格朗日乘子  $\tilde{\mathbf{g}}, \tilde{\mathbf{h}}$  和罚参数  $\varphi_1, \varphi_2 > 0$ , 令  $\tilde{\mathbf{g}} = \frac{\varphi_1}{\tilde{\mathbf{g}}}, \tilde{\mathbf{h}} = \frac{\varphi_2}{\tilde{\mathbf{h}}}$ 。则得

到上述最小化问题的增广拉格朗日方程如下:

$$\begin{aligned} L = \frac{\beta}{2} \| M \odot (u - s) \|_2^2 + \gamma_1 \| \mathbf{g} \|_1 + \gamma_0 \| \mathbf{h} \|_1 + \\ \frac{\varphi_1}{2} \| \mathbf{g} - (Du - q) - \tilde{\mathbf{g}} \|_2^2 + \frac{\varphi_2}{2} \| \mathbf{h} - \mathbf{B}(q) - \tilde{\mathbf{h}} \|_2^2 \end{aligned} \quad (14)$$

然后依次求解以下 5 个子问题, 得到最优解。

(1)  $g$ -子问题。研究给出的数学定义公式如下:

$$g^{n+1} = \arg \min_g \| \mathbf{g} \|_1 + \frac{\varphi_1}{2} \| \mathbf{g} - (Du^n - q^n) - \tilde{g}^n \|_2^2 \quad (15)$$

由于  $g$ -子问题是组件可分离的,  $g$ -子问题的解为:

$$g^{n+1}(l) = \gamma_1 (S_2(Du^n(l) - q^n(l) + \tilde{g}^n(l), \gamma_1/\varphi_1)), \quad l \in \Omega \quad (16)$$

其中,  $g^{n+1}(l)$  表示  $g^{n+1}$  位于  $l \in \Omega$  的部分, 各向同性收缩算子  $SH_2(a, \varphi)$  定义为:

$$SH_2(a, \varphi) = \begin{cases} 0, & a = 0, \\ (\| a \|_2 - \varphi) \frac{a}{\| a \|_2}, & a \neq 0 \end{cases} \quad (17)$$

(2)  $h$ -子问题。研究给出的数学定义公式如下:

$$h^{n+1} = \arg \min_h \| \mathbf{h} \|_1 + \frac{\varphi_2}{2} \| \mathbf{h} - (B(q^n)) - \tilde{h}^n \|_2^2$$

同样,  $h$ -子问题的解如下所示:

$$h^{n+1}(l) = \gamma_0 (S_F(B(q^n))(l) + \tilde{h}^n(l), \gamma_0/\varphi_2), \quad l \in \Omega \quad (18)$$

其中,  $h^{n+1}(l) \in S^{2 \times 2}$  表示  $h^{n+1}$  相对于  $l \in \Omega$  的组成部分且:

$$S_F(b, \varphi) = \begin{cases} 0, & b = 0 \\ (\| b \|_F - \varphi) \frac{b}{\| b \|_F}, & b \neq 0 \end{cases} \quad (19)$$

(3)  $u$ -子问题。研究给出的数学定义公式如下:

$$u^{n+1} = \arg \min_u \frac{\beta}{2} \| M \odot (u - s) \|_2^2 + \frac{\varphi_1}{2} \| g^{n+1} - (Du - q^n) - \tilde{g}^n \|_2^2 \quad (20)$$

通过 DIP 网络解决了  $u$ -子问题。DIP 的损失函数为最小化问题的目标函数。反向传播后, DIP 网络中的偏置参数和权值参数会不断更新。经过几次迭代后, 就可以得到使损失函数最小的网络输出。这个输出是  $u$ -子问题的解。子问题的解决过程可以总结如下:

$$\begin{aligned} u &= T_{\theta}(z) \\ \text{Loss} &= \frac{\beta}{2} \| M \odot (u - b) \|_2^2 + \\ &\quad \frac{\varphi_1}{2} \| g^{n+1} - (Du - q^n) - \tilde{g}^n \|_2^2 \end{aligned} \quad (21)$$

其中,  $y, q$  和  $\tilde{g}$  来自相应的子问题。

(4)  $q_1$  和  $q_2$  – 子问题。研究给出的数学定义公式如下:

$$\begin{aligned} q^{n+1} &= \arg \min_q \frac{\varphi_1}{2} \| q^{n+1} - (Du^{n+1} - q) - \tilde{g}^n \|_2^2 + \\ &\quad \frac{\varphi_2}{2} \| h^{n+1} - B(q) - \tilde{h}^n \|_2^2 \end{aligned} \quad (22)$$

对该子问题的目标函数求导并令其等于 0 后, 可以得到如下等式:

$$\begin{aligned} 0 &= \gamma_1 \varphi_1 (q_1 - D_1 u^{n+1} + g_1^{n+1} - \tilde{g}_1^n) + \\ &\quad \gamma_1 \varphi_2 (D_1^T (D_1 q_1 - h_1^{n+1} + \tilde{h}_1^n) + \frac{1}{2} D_2^T (D_2 q_1 + \\ &\quad D_1 q_2 - 2h_3^{n+1} + 2\tilde{h}_3^n)) \\ 0 &= \gamma_1 \varphi_1 (q_2 - D_2 u^{n+1} + g_2^{n+1} - \tilde{g}_2^n) + \\ &\quad \gamma_0 \varphi_2 (D_2^T (D_2 q_2 - h_2^{n+1} + \tilde{h}_2^n) + \frac{1}{2} D_1^T (D_1 q_2 + \\ &\quad D_2 q_1 - 2h_3^{n+1} + 2\tilde{h}_3^n)) \end{aligned} \quad (23)$$

经过变换和简单的 4 种基本算术运算, 可以得到  $q_1$  和  $q_2$  的解。

(5)  $\tilde{g}$  和  $\tilde{h}$  – 子问题。拉格朗日乘数更新如下:

$$\begin{cases} \tilde{g}^{n+1} = \tilde{g}^n + \mu (Du^{n+1} - q^{n+1} - g^{n+1}) \\ \tilde{h}^{n+1} = h^n + \mu (B(q^{n+1}) - h^{n+1}) \end{cases} \quad (24)$$

通过数次迭代求解  $y, z, u, p$  – 子问题和拉格朗日乘子  $\tilde{g}$  和  $\tilde{h}$  子问题, 得到最优解  $u$ 。

## 4 实验部分

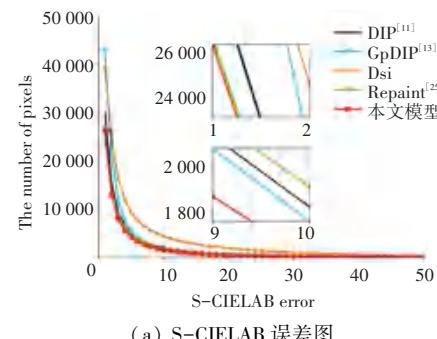
### 4.1 实验细节

由于训练不需要数据集, 本文中使用的数据集仅用于测试目的。对于图像修补, 使用 Places2<sup>[22]</sup> 和 ImageNet<sup>[23]</sup> 数据集。从每个数据集中测试了 10 张图像(共 20 张)。所有图片都是从官方网站上下载的, 并将尺寸调整为  $256 \times 256$  进行本次研究的图像修补实验。本文的方法将与基于 DIP 的方法(如 DIP<sup>[11]</sup> 和 GpDIP<sup>[13]</sup>)和基于深度学习的方法(如 AOT<sup>[24]</sup> 和 Repaint<sup>[25]</sup>)进行比较。这些方法的源代码来自本文研发的网站, 实验中使用这些方法的默认参数。为了评估本文的算法, 采用了几个定量指标, 如峰值信噪比(PSNR)、结构相似指数测量(SSIM)<sup>[26]</sup>、空间 CIELAB(S-CIELAB)误差<sup>[27]</sup>和学习感知图像斑块相似度(LPIPS)<sup>[28]</sup>。所有实验都是在 Windows 10 操作系统上配备英特尔酷睿 i7-

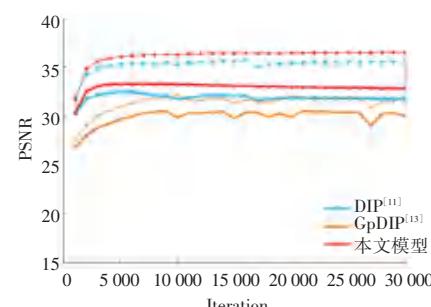
RTX 3050 的 PC 上使用 PyCharm 2021 进行的。

### 4.2 定性实验

在本小节中, 对本文的方法进行了性能分析。在图像修补领域, 基于 DIP 的方法得到了广泛的应用。如果要想以较少的迭代次数获得高 PSNR 值的修补图像, 就必须提高模型的有效性。为了说明本文使用的方法比基于 DIP 的方法效率更高, 以其中一张图像为例进行了关于 PSNR 值和迭代次数的实验。使用 3 种类型的掩码依次破坏图像, 进而生成 3 幅受损图像。本文的方法和基于 DIP 的方法先后对 3 幅受损图像进行了 5 次修复, 并记录了修复图像在迭代过程中的平均 PSNR, 结果如图 2(b)所示。由图 2(b)子图中拥有相同标志的线条有 2 条, 较粗的线为 PSNR 均值。从图 2(b)的实验结果可以看出, 本文方法修补后的图像的 PSNR 均值峰值最大, 达到峰值的时间也最早。这不仅说明了本文方法的有效性, 也进一步证实了使用 ADMM 算法的高效率。图 2(b)子图中相同标志较粗线与较细线之间的区域表示每次迭代的 PSNR 值的标准差。与基于 DIP 的方法相比, 本文的方法具有更稳定的标准差和更平滑的曲线。这表明本文所提的方法更稳定, 即在多次迭代中产生更一致的结果。此外, 无论图像受到哪种类型掩码的损坏, 本文的方法都能很好地修补受损图像。因此, 本文提出的方法比现有的基于 DIP 的方法更有效、更稳定。



(a) S-CIELAB 误差图



(b) PSNR 曲线图

Fig. 2 The graph of curves

文中分别对来自 ImageNet 和 Palaces 数据集的 10 幅图像(共 20 幅)进行了 3 种掩码类型的实验, 以说明本文提出的方法的鲁棒性。图 3 展示了各种方法修补被 3 种掩码损毁后的图像的视觉结果。图 3 中前 3 列是对 Palaces2 数据集的图像进行修补的结果, 前 3 列的掩码类型分别是“ev2li”, “nn2”和“thin”, 后 3 列是对 ImageNet 数据集的图像进行修补的结果, 后 3 列的掩码类型分别是“ev2li”, “nn2”和“thin”。对于“ev2li”掩码, 本文的方法能够保留更多的纹理信息, 如鱼鳞的纹理(第 1 列)和黄色食物的纹理(第 4 列), 这些细节是其他方法难以复原的。在恢复被名为“nn2”的掩码损坏的图像时, 基于深度学习的方法 Repaint<sup>[25]</sup> 和 AOT<sup>[24]</sup> 不能很好地

去除掩码。与基于 DIP 的方法相比, 本文方法在去除掩码的同时保留了清晰的边界和丰富的纹理信息。比如墙上的图案(第 2 列)和叶子的边缘(第 5 列)。对于命名为“thin”的不规则掩码, 本文的方法仍然可以比其他方法恢复尽可能多的信息。在图 3 的第 3 列和最后一列中, 虽然包括本文的方法都不能恢复放大区域中缺失的像素, 但是只有本文的方法可以恢复水草的边界和木板的纹理。基于 DIP 的方法恢复的图像有时过于平滑, 而基于深度学习的方法只熟悉这些掩码的 1~2 种类型。因此, 综合考虑 3 种类型的掩码, 本文得出结论, 无论图像被哪种类型的掩码损坏, 本文的方法修补的图像都能够产生最佳的视觉效果。

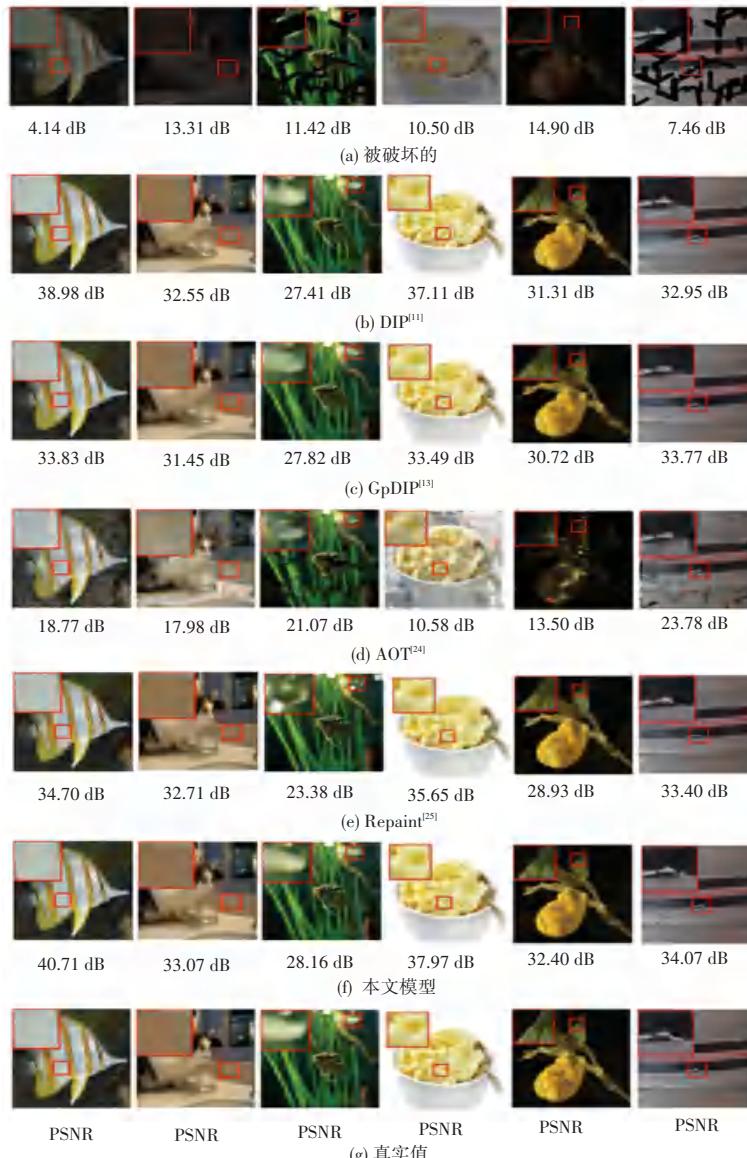


图 3 对 Palaces2 和 ImageNet 数据集的修补结果

Fig. 3 Inpainting results of the Palaces2 and ImageNet datasets

### 4.3 定量实验

测试图像的数值结果见表 1。其中,通过本文方法恢复的图像具有最高(最低)的 PSNR 和 SSIM(或 LPIPS)值。同时还观察到,使用其他方法恢复的图像存在一定的色差,本文方法通过计算所有图像的 S-CIELAB 误差值进一步说明了这一点,如图

表 1 图像修复的 PSNR、SSIM 和 LPIPS 的平均值

Table 1 Average of PSNR, SSIM, and LPIPS for image inpainting over 5 runs

| 数据集      | Mask  | PSNR                |                       |                     |                         | SSIM  |                     |                       |                     | LPIPS                   |       |                     |                       |                     |                         |       |
|----------|-------|---------------------|-----------------------|---------------------|-------------------------|-------|---------------------|-----------------------|---------------------|-------------------------|-------|---------------------|-----------------------|---------------------|-------------------------|-------|
|          |       | DIP <sup>[11]</sup> | GpDIP <sup>[13]</sup> | AOT <sup>[24]</sup> | Repaint <sup>[25]</sup> | 本文    | DIP <sup>[11]</sup> | GpDIP <sup>[13]</sup> | AOT <sup>[24]</sup> | Repaint <sup>[25]</sup> | 本文    | DIP <sup>[11]</sup> | GpDIP <sup>[13]</sup> | AOT <sup>[24]</sup> | Repaint <sup>[25]</sup> | 本文    |
| ImageNet | ev2li | 32.84               | 31.51                 | 17.08               | 32.05                   | 33.21 | 0.936               | 0.902                 | 0.427               | 0.911                   | 0.942 | 0.132               | 0.188                 | 0.525               | 0.161                   | 0.127 |
|          | nn2   | 31.04               | 30.37                 | 13.01               | 27.97                   | 31.50 | 0.902               | 0.876                 | 0.246               | 0.809                   | 0.908 | 0.169               | 0.227                 | 0.738               | 0.265                   | 0.161 |
|          | thin  | 28.95               | 29.67                 | 22.96               | 28.52                   | 29.98 | 0.854               | 0.888                 | 0.786               | 0.898                   | 0.901 | 0.225               | 0.195                 | 0.315               | 0.155                   | 0.151 |
| Palace   | ev2li | 36.91               | 34.47                 | 17.94               | 36.49                   | 37.67 | 0.967               | 0.927                 | 0.428               | 0.953                   | 0.972 | 0.029               | 0.090                 | 0.453               | 0.035                   | 0.024 |
|          | nn2   | 34.11               | 32.52                 | 12.99               | 33.48                   | 34.69 | 0.933               | 0.903                 | 0.238               | 0.920                   | 0.946 | 0.071               | 0.124                 | 0.691               | 0.064                   | 0.051 |
|          | thin  | 29.87               | 30.35                 | 21.28               | 28.34                   | 31.01 | 0.896               | 0.901                 | 0.728               | 0.922                   | 0.925 | 0.104               | 0.120                 | 0.326               | 0.062                   | 0.061 |

## 5 结束语

本文提出了一种结合 DIP 和 TGV 的图像修补模型。本次研究中采用 ADMM 算法解决问题,在一定程度上实现了传统方法与深度学习方法的有效结合。该方法在图像修补实验中保留了更多的视觉细节,优化了数值结果。

## 参考文献

- [1] LI Lin, WANG L V. The emerging role of photoacoustic imaging in clinical oncology [J]. Nature Reviews Clinical Oncology, 2022, 19(6): 365–384.
- [2] DENIS L, DALASSASSO E, TUPIN F. A review of deep-learning techniques for SAR image restoration [C]// Proceedings of 2021 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Piscataway, NJ: IEEE, 2021: 411–414.
- [3] BALLESTER C, BERTALMIO M, CASELLAS V, et al. Filling-in by joint interpolation of vector fields and gray levels [J]. IEEE Transactions on Image Processing, 2001, 10(8): 1200–1211.
- [4] BERTALMIO M, SAPIRO G, CASELLAS V, et al. Image inpainting [C]// Proceedings of the 27<sup>th</sup> Annual Conference on Computer Graphics and Interactive Techniques. New York: ACM, 2000: 417–424.
- [5] SIMAKOV D, CASPI Y, SHECHTMAN E, et al. Summarizing visual data using bidirectional similarity [C]// Proceedings of 2008 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2008: 1–8.
- [6] BARNES C, SHECHTMAN E, FINKELSTEIN A, et al. Patchmatch: A randomized correspondence algorithm for structural image editing [J]. ACM Transactions on Graphics, 2009, 28(3): 24.
- [7] KÖHLER R, SCHULER C, SCHÖLKOPF B, et al. Mask-specific inpainting with deep neural networks [C]// Proceedings of German Conference on Pattern Recognition. Cham: Springer, 2017: 523–534.
- [8] PATHAK D, KRAHENBUHL P, DONAHUE J, et al. Context encoders: Feature learning by inpainting [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2016: 2536–2544.
- [9] YANG Chao, LU Xin, LIN Zhe, et al. High-resolution image inpainting using multi-scale neural patch synthesis [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2017: 6721–6729.
- [10] LI Jingyuan, WANG Ning, ZHANG Lefei, et al. Recurrentfeature reasoning for image inpainting [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2020: 7760–7768.
- [11] ULYANOV D, VEDALDI A, LEMPITSKY V. Deep image prior [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2018: 9446–9454.
- [12] MATAEV G, MILANFAR P, ELAD M. DeepRED: Deep image prior powered by red [J]. arXiv preprint arXiv, 1903. 10176, 2019.
- [13] CHENG Zezhou, GADELHA M, MAJA S, et al. A Bayesian perspective on the deep image prior [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2019: 5443–5451.
- [14] LIU Jiaming, SUN Yu, XU Xiaojian, et al. Image restoration using total variation regularized deep image prior [C]// Proceedings of 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway, NJ: IEEE, 2019: 7715–7719.
- [15] BATARD T, HARO G, BALLESTER C. Dip-VBT: A color image restoration model combining a deep image prior and a vector bundle total variation [J]. SIAM Journal on Imaging Sciences, 2021, 14(4): 1816–1847.
- [16] BREDES K, KUNISCH K, POCK T. Total generalized variation [J]. SIAM Journal on Imaging Sciences, 2010, 3(3): 492–526.
- [17] LU Zhenbo, LI Houqiang, LI Weiping. A Bayesian adaptive

- weighted total generalized variation model for image restoration [ C ]//Proceedings of 2015 IEEE International Conference on Image Processing ( ICIP ). Piscataway, NJ: IEEE, 2015: 492–496.
- [ 18 ] ZHANG Honglu, TANG Liming, FANG Zhuang, et al. Nonconvex and nonsmooth total generalized variation model for image restoration [ J ]. Signal Processing, 2018, 143: 69–85.
- [ 19 ] ZHOU Mengmeng, ZHAO Ping. Enhanced total generalized variation method based on Moreau envelope [ J ]. Multimedia Tools and Applications, 2021, 80( 13 ): 19539–19566.
- [ 20 ] SERAFINO D D, PRAGLIOLA M. Automatic parameter selection for the tgv regularizer in image restoration under Poisson noise [ J ]. arXiv Preprint arXiv:2205.13439, 2022.
- [ 21 ] GABAY D, MERCIER B. A dual algorithm for the solution of nonlinear variational problems via finite element approximation [ J ]. Computers & Mathematics with Applications, 1976, 2( 1 ): 17–40.
- [ 22 ] ZHOU Bolei, LAPEDRIZA A, KHOSLA A, et al. Places: A 10 million image database for scene recognition [ J ]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 40( 6 ): 1452–1464.
- [ 23 ] RUSSAKOVSKY O, DENG Jia, SU Hao, et al. ImageNet large scale visual recognition challenge [ J ]. International Journal of Computer Vision, 2015, 115( 3 ): 211–252.
- [ 24 ] ZENG Yanhong, FU Jianlong, CHAO Hongyang, et al. Aggregated contextual transformations for high-resolution image inpainting [ J ]. IEEE Transactions on Visualization and Computer Graphics, 2023, 29( 7 ): 3266–3280.
- [ 25 ] LUGMAYR A, DANELLJAN M, ROMERO A, et al. Repaint: Inpainting using denoising diffusion probabilistic models [ C ]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2022: 11461–11471.
- [ 26 ] WANG Zhou, BOVIK A C, SHEIKH H R, et al. Image quality assessment: From error visibility to structural similarity [ J ]. IEEE Transactions on Image Processing, 2004, 13( 4 ): 600–612.
- [ 27 ] ZHANG X, WANDELL B A. A spatial extension of CIELab for digital color-image reproduction [ J ]. Journal of the Society for Information Display, 1997, 5( 1 ): 61–63.
- [ 28 ] ZHANG R, ISOLA P, EFROS A A, et al. The unreasonable effectiveness of deep features as a perceptual metric [ C ]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2018: 586–595.