基于改进CNN预测的可逆数据隐藏

Yingqiang Qiu、Wanli Peng、Xiaodan Lin、Huanqiang Zeng，高级会员，*IEEE*，以及 Zhenxing Qian，会员，*IEEE*

摘要—本文提出一种用于图像可逆数据隐藏（RDH）的改进型卷积神经网络预测器（ICNNP），由特征提取模块、像素预测模块和复杂度预测模块组成。在嵌入过程中，ICNNP 会预测每个像素的复杂度，因此所提出的方法比基于卷积神经网络预测器的方法具有更优的性能。具体而言，首先将输入图像分为两个不同的子图像，即“Dot”图像和“Cross”图像。同时，利用每个子图像来预测另一个子图像。随后，根据预测的像素复杂度对像素的预测误差进行排序。基于此，选取部分复杂度较低的已排序预测误差，并采用传统的直方图平移方案进行低失真数据嵌入。实验结果表明，在采用相同直方图平移策略的情况下，所提出的方法比卷积神经网络预测器具有更优的嵌入性能。

索引词—**卷积神经网络、多任务、可逆数据隐藏、直方图平移**

I. 引言

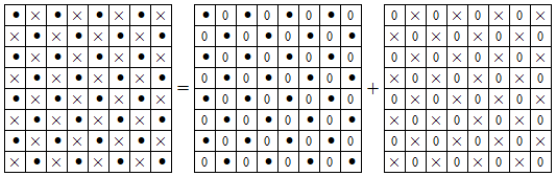
可逆数据隐藏（RDH）能够无损地恢复嵌入数据和载体介质 [1]。由于这一特性，RDH 逐渐成为信息隐藏领域的研究热点，并已广泛应用于多个实际场景 [1], , 包括医疗、军事和司法取证等领域。根据隐藏秘密消息所处的域，RDH 可分为两大分支：基于空间域的可逆数据隐藏 [2-23] 和基于 JPEG 域的可逆数据隐藏 [24-27]。基于空间域的可逆数据隐藏通常采用三种技术，即无损压缩（LC）[2-4], 、差值扩展（DE）[5-15], 和直方图平移（HS）[16-23]。而基于 JPEG 域的可逆数据隐藏主要基于离散余弦变换系数修改 [24, 25] 或霍夫曼表修改 [26, 27]。

目前，在可逆数据隐藏（RDH）领域，像素预测已成为一个关键问题，极大地影响着可逆数据隐藏算法的性能 [14]。传统预测器包括中值边缘方向（MED）预测器 [6], 插值预测器 [7], 梯度调整预测器（GAP）[8], 像素值排序（PVO）预测器 [9, 12, 22], 线性预测器 [10], 菱形预测器 [17-20], 以及岭回归预测器 [23], 等。尽管这些预测器已取得显著进展，但仍存在一个明显的弱点，即用于像素预测的邻近像素很少 [14]。如果将更多相邻像素用作参考像素，就能获得更高的预测性能。卷积神经网络（CNN）具有融合不同感受野并进行整体优化的强大能力，因此可以通过为像素预测构建非线性映射，建立并训练卷积神经网络，从而准确预测像素。鉴于此，Luo等人 [13] 利用左视图与右视图之间的相关性，提出了一种基于CNN的立体图像可逆数据隐藏方法。Hu等人 [14] 提出了一种基于CNN预测器（CNNP）的可逆数据隐藏方法：将灰度图像拆分为两个子图像，并使用CNNP交替地以其中一个子图像预测另一个子图像。此后，Hu等人 [15]将图像划分为四个部分，并使用CNNP依次以其余三个部分预测每个部分，从而获得更好的预测性能。此外，通过自适应嵌入，标记图像的视觉质量得到提升。总的来说，CNN预测器的预测性能优于传统预测器。

综上所述，为了提升性能，现有方法利用相邻像素进行像素预测。然而，这些方法 [13-15] 未利用深度学习对每个像素的复杂度进行分析，从而限制了可逆数据隐藏的性能。

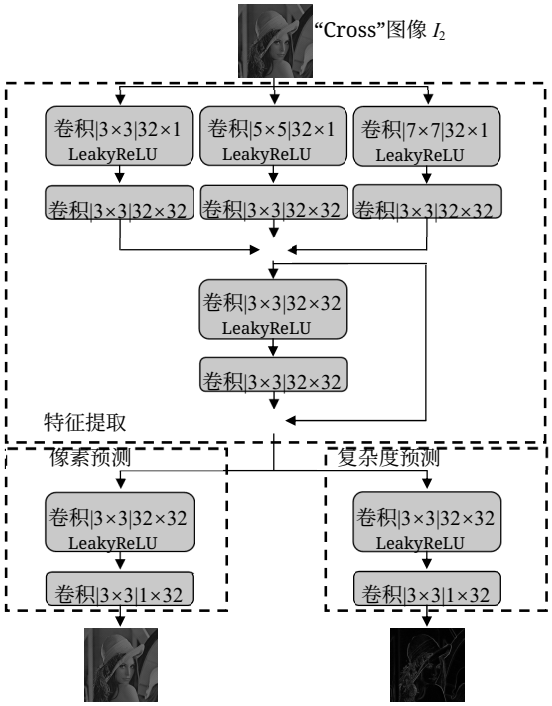
为解决上述局限，本文对[14] 中提出的CNNP进行了改进，增加了复杂度预测模块，以精确预测像素的复杂度。本文后续将其称为改进的CNNP（ICNNP）。具体而言，在数据嵌入过程中，我们首先将一幅灰度图像分割成两个子图像，其中一个子图像由另一个子图像进行预测。随后，根据预测像素的复杂度对其预测误差进行排序，并使用复杂度较低的预测误差通过经典直方图平移进行数据嵌入

本研究部分获国家重点研发计划（项目编号：2021YFE0205400）、中国厦门市自然科学基金（项目编号：3502Z20227192）和国家自然科学基金（项目编号：U20B2051、61972168、62072114、62002124、61871434）资助。通讯作者：Zhenxing Qian。Y. Qiu、X. Lin 和 H. Zeng 任职于华侨大学信息科学与工程学院（中国厦门 361021）。（电子邮箱：yqqiu@hqu.edu.cn、echo.linxd@gmail.com、zeng0043@hqu.edu.cn）。W. Peng 和 Z. Qian 任职于复旦大学计算机科学学院（中国上海 200433）。（电子邮箱：pengwanli@fudan.edu.cn、zxqian@fudan.edu.cn）。Y. Qiu 和 W. Peng 对本研究贡献相同。



(a) 原始图像 I (b) “Dot”图像 I_1　(c) “Cross”图像 I_2

图 1. 将原始图像拆分为两个子图像的示意图。



预测的“Dot”图像 $\tilde{I}_1$

策略。最后，实验结果表明，所提出方法的性能优于 [14] 中提出的 CNNP。

本文其余部分安排如下：第二节详细介绍所提出的改进的可逆数据隐藏方法，第三节给出实验结果并进行讨论。最后，第四节总结本文工作。

II. 所提出的改进方法

A. 网络架构

如图1所示，根据棋盘格上下文模型 [17], ，原始图像被划分为两个子图像，分别由“圆点”和“十字”像素组成。对于“圆点”图像，保留“圆点”像素的像素值，而将“十字”像素的像素值设为0。同时，对于“十字”图像，仅将“圆点”像素的像素值设为0。基于这两个子图像的像素相关性，利用每个子图像预测另一个子图像的像素值和复杂度。

所提出的 ICNNP 整体结构如图 2 所示。ICNNP 的架构由三部分组成，即特征提取、像素预测和复杂度预测。“Cross”图像 I_2 被输入网络，以预测“Dot”像素的值和复杂度，其中值复杂度的值被调整为 [0,255] ，以便进行良好的可视化显示。值越低，复杂度越低。特征提取由一些具有不同滤波器尺寸（3 × 3、 5 × 5、 7 × 7、 …）的卷积层组成，这些卷积层并行化后，分别与一个 3×3 卷积层连接，以从不同感受野中提取特征。随后应用残差块，进一步聚合和细化来自不同分支的学习特征。利用提取的特征，像素预测生成预测的“Dot”图像 $\tilde{I}_1$，复杂度预测生成“Dot”图像的预测复杂度 $\tilde{C}_1$ I_1。“Conv”表示卷积核尺寸为 $S \times S$ 的卷积单元，通道数为 $output \times input$。每两个卷积层之间设置一个 LeakyReLU 激活函数 [28] 。

值得注意的是，复杂度预测与像素预测类似，即使用更多相邻像素[14, 17], 而非正交相邻像素，以非线性方式预测像素区域的复杂度，从而提升可逆数据隐藏的性能。

B. 训练

在 ICNNP 中，已训练好的 CNNP 参数[14] 被载入特征提取和像素预测模块。需要注意的是，这些参数保持固定，而复杂度预测的参数则在 ICNNP 训练过程中更新。在训练中，输入为“Cross”图像I_2, ，输出为预测的“Dot”图像$\tilde{I}_1$ 以及“Dot”图像I_1的预测复杂度。由于特征提取和像素预测的滤波器参数保持固定，目标不再是“Dot”图像I_1 ，而是I_1的参考复杂度C_1 。C_1 的定义如下所述。

（1）对于“十字”像素，$C_1(i,j)$ 设为 0;

（ 2　$i=1$　$i=M$　$j=1$　$j=N$ ）对于“圆点”像素，如 $C_1(i\ j)$　0　$C_1(i\ j)(2 \le i \le M-1$ 果或或或，，则设为；否则，，，　$2 \le j \le$ $N-1)$ 计算为

$$C_1(i,j)=\frac{1}{R}\cdot\sqrt{\sum_{k=k_1}^{k_2}\sum_{l=l_1}^{l_2}(I(i+k,j+l)-I(i,j))^2}, \quad (1)$$

where,

$$\begin{cases} k_1=-1, k_2=2 & ,i=2 \\ k_1=-2, k_2=1 & ,i=M-1 \\ k_1=-2, k_2=-2 & ,2<i<M-1 \end{cases}, \quad (2)$$

$$\begin{cases} l_1=-1, l_2=2 & ,j=2 \\ l_1=-2, l_2=1 & ,j=N-1 \\ l_1=-2, l_2=-2 & ,2<j<N-1 \end{cases}, \quad (3)$$

$$R=(k_2-k_1+1)\times(l_2-l_1+1)-1. \quad (4)$$

在所提出的方法中，最大预测像素区域为 5 × 5，可准确计算像素复杂度。

与 CNNP 一样，[14], 我们采用反向传播 [29] 和 Adam算法 [30] 来优化如下定义的目标函数:

$$\text{Loss}=\frac{1}{N}\sum_{i=1}^{N}(\tilde{C}_1-C_1)^2+\lambda\|\omega\|_2^2, \quad (5)$$

其中，N表示训练样本的数量，　ω 表示网络的所有权重，λ 表示权重衰减。

C. 基于 *ICNNP* 的可逆数据隐藏

图 3 展示了基于 ICNNP 的可逆数据隐藏方法的数据嵌入架构。所采用的双重嵌入策略 [17] 结合直方图平移技术 [6] ，包括依次使用

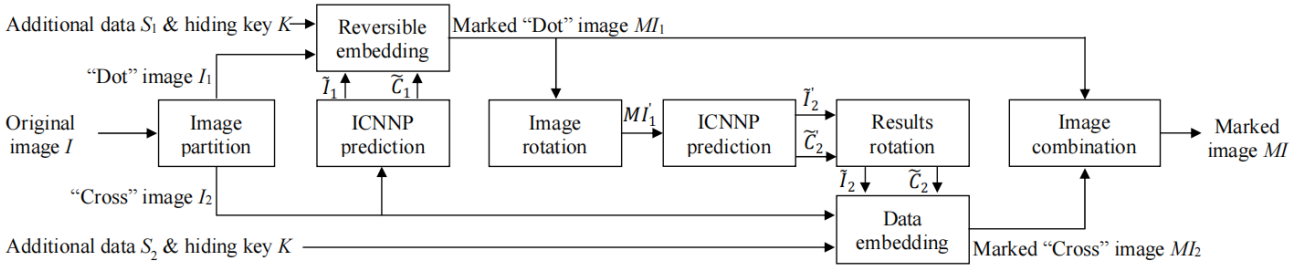


Fig. 3. The flowchart of data embedding by using the ICNNP based RDH method.

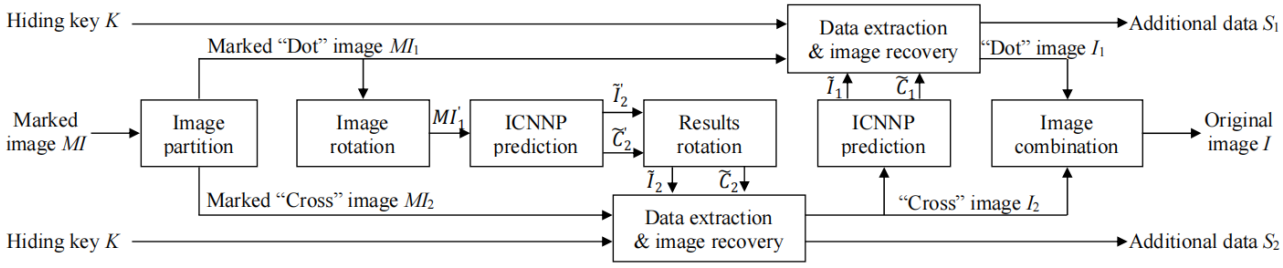


Fig. 5. The flowchart of data extraction and image recovery by using the ICNNP based RDH method.

"圆点"嵌入和"十字"嵌入，其中"十字"嵌入在"圆点"嵌入之后进行。

首先，将原始图像I划分为两个子图像，即"Dot"图像I_1和"Cross"图像I_2。接下来，利用I_2预测"Dot"图像$\tilde{I}_1$以及I_1的预测复杂度$\tilde{C}_1$，如下所示：

$$\{\tilde{I}_1, \tilde{C}_1\} = ICNNP(I_2). \quad (6)$$

然后，计算I_1的预测误差，得到

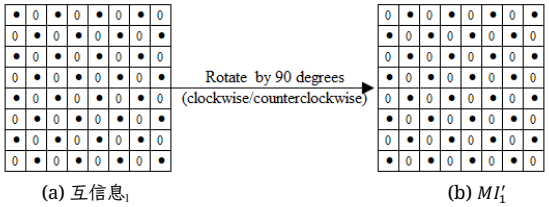
$$e_1(i,j) = I_1(i,j) - \tilde{I}_1(i,j), (i+j)\bmod 2 \equiv 0. \quad (7)$$

根据预测复杂度的大小和附加数据S_1的规模，我们选择复杂度较低的预测误差，并确定两个阈值T_{n1}（T_{n1}<0）和T_{p1}（$T_{p1}\geq 0$）用于基于 HS 的数据嵌入，其实现方式如下

$$E_1(i,j) = \begin{cases} 2e_1(i,j) + b & ,if\ e_1(i,j) \in [T_{n1}, T_{p1}] \\ e_1(i,j) + T_{p1} + 1 & ,if\ e_1(i,j) > T_{p1} \\ e_1(i,j) + T_{n1} & ,if\ e_1(i,j) < T_{n1} \end{cases}, \quad (8)$$

其中$b \in \{0,1\}$是嵌入的比特，表示加密附加数据和一些辅助数据[17]。因此，标记的"Dot"图像MI_1生成如下

$$MI_1(i,j) = \tilde{I}_1(i,j) + E_1(i,j). \quad (9)$$



(a) 互信息$_1$ (b) MI_1'

图 4. 标记的"Dot"图像的图像旋转示意图。

在"Cross"嵌入过程中，由于棋盘格模式，标记的"Dot"图像MI_1无法直接输入网络以预测"Cross"图像I_2。如图 4 所示，当高度/宽度为偶数时，将其顺时针/逆时针旋转 90 度后，所得图像MI_1'具有与"Cross"图像I_2相同的模式。使用相同的 ICNNP，如公式 (10) 所示，我们将旋转后的标记的"Dot"图像MI_1'输入网络，随后得到预测的旋转"Cross"图像$\tilde{I}_2'$及其复杂度$\tilde{C}_2'$。

$$\{\tilde{I}_2', \tilde{C}_2'\} = ICNNP(MI_1'). \quad (10)$$

然后，将$\tilde{I}_2'$和$\tilde{C}_2'$分别逆时针/顺时针旋转 90 度，得到预测的"Cross"图像$\tilde{I}_2$以及I_2的预测复杂度$\tilde{C}_2$。与"Dot"嵌入类似，另一部分附加数据S_2使用隐藏密钥K加密后嵌入I_2中，从而得到标记的"Cross"图像MI_2。

最后，我们将标记的"Dot"图像MI_1和标记的"Cross"图像MI_2合并，得到标记图像MI。

图 5 描述了数据提取/图像恢复的架构。数据提取和图像恢复是数据嵌入的逆过程，因此我们先进行"Dot"提取/恢复，再进行"Cross"提取/恢复。首先，将标记图像MI划分为两个子图像，即标记的"Dot"图像MI_1和标记的"Cross"图像MI_2。利用旋转后的标记"Dot"图像MI_1'，根据式（10）使用 ICNNP 预测$\tilde{I}_2'$和$\tilde{C}_2'$。随后，分别旋转$\tilde{I}_2'$和$\tilde{C}_2'$，得到$\tilde{I}_2$和$\tilde{C}_2$。接下来，计算I_2的标记预测误差，得到

$$E_2(i,j) = MI_2(i,j) - \tilde{I}_2(i,j), (i+j)\bmod 2 \equiv 1. \quad (11)$$

根据C_2的排序幅值，并以提取的辅助数据T_{n2}（T_{n2}<0）和T_{p2}（$T_{p2}\geq 0$）作为阈值，提取嵌入数据如下：

$$b = E_2(i,j) \bmod 2,\ E_2(i,j) \in [2T_{n2}, 2T_{p2}+1], \quad (12)$$

并恢复I_2的原始预测误差如下：

$$e_2(i,j) = \begin{cases} \lfloor E_2(i,j)/2 \rfloor & ,if\ E_2(i,j) \in [2T_{n2}, 2T_{p2}+1] \\ E_2(i,j) - T_{p2} - 1 & ,if\ E_2(i,j) > 2T_{p2}+1 \\ E_2(i,j) - T_{n2} & ,if\ E_2(i,j) < 2T_{n2} \end{cases}, \quad (13)$$

其中，$\lfloor\cdot\rfloor$是向下取整函数。我们使用隐藏密钥K解密提取的比特，得图像I_2如下：

$$I_2(i,j) = e_2(i,j) + \tilde{I}_2(i,j), (i+j)\bmod 2 \equiv 1. \quad (14)$$

同样，嵌入数据S_1被正确提取，原始"圆点"图像I_1也得以无损恢复。最后，我们



(a) Lena (b) Yacht (c) 辣椒 (d) Barbara

图 6. 四幅封面图像。

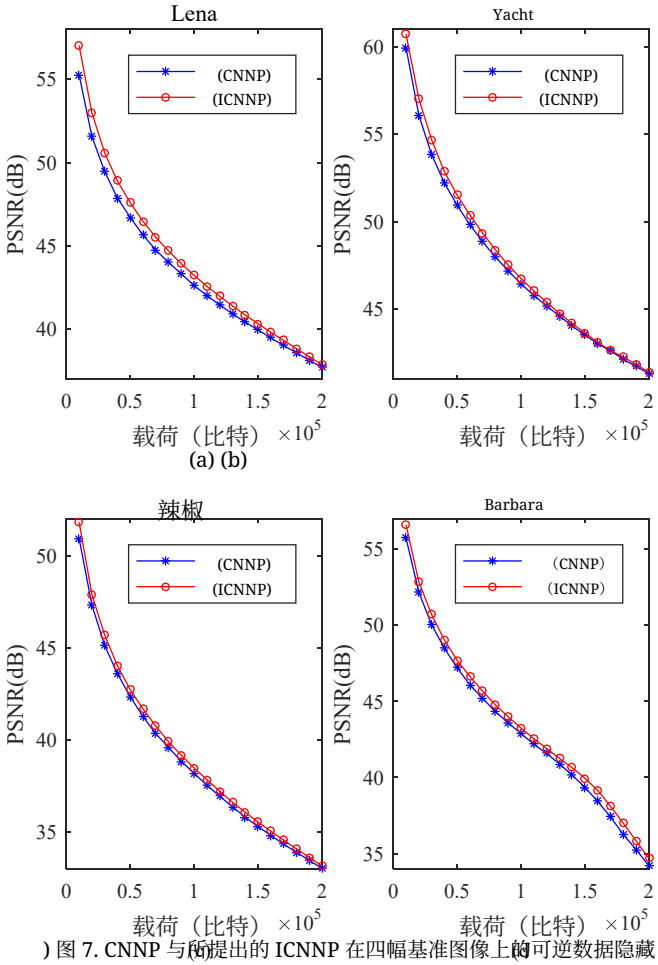


图 7. CNNP 与所提出的 ICNNP 在四幅基准图像上的可逆数据隐藏性能比较。

将恢复的“十字”图像 I_2 和“圆点”图像 I_1 合并，从而得到原始图像 I。

III. 实验结果

为评估所提出的 ICNNP 的效率，我们使用从 BOWS-2 [31] 中随机选取的 1,000 张图像来训练其复杂度预测模型的参数。所提出的 ICNNP 在配备 16 GB RAM 和 NVIDIA GeForce RTX 2060 的 Intel Core i10 CPU (3.6 GHz) 上进行训练。权重衰减 λ 设置为 1×10^{-3}，批量大小为 4，初始学习率为 1×10^{-3}。在 [14], 中，已有研究证明 CNNP 的预测准确率优于 BIP、MEDP、GAP 和 DP 等传统线性预测器；此外，采用扩展嵌入方案时，CNNP 的率失真性能优于上述预测器；采用 HS 方案时，BIP 的性能也优于采用扩展嵌入方案的 CNNP，而 HS 的性能则远优于 CNNP 配合扩展嵌入方案的性能。

因此，我们通过将 ICNNP 与采用相同直方图平移技术的 CNNP [14] 进行比较，合理地评估 ICNNP 的性能。

以图 6 中的四幅基准图像作为载体图像，并采用相同的嵌入率，我们使用原始图像与标记图像之间的峰值信噪比（PSNR）作为客观图像质量评估指标。此外，我们从 BOWS-2 [31] 中随机选取 100 幅不同于训练图像的图像，并测试不同嵌入容量下的表现，以评估 ICNNP 的通用性。

图 7 展示了四幅测试图像（Lena、Yacht、辣椒和 Barbara）在嵌入容量为 10,000 至 200,000 比特时的 PSNR 值。从图中可以看出，采用 ICNNP 的可逆数据隐藏方法的 PSNR 值高于基于 CNNP 的可逆数据隐藏方法。此外，表 I 展示了 100 幅测试图像在不同嵌入容量下的平均 PSNR 值。当嵌入容量为 10,000 比特时，所提出的基于 ICNNP 的可逆数据隐藏方法的平均 PSNR 值为 62.37 dB，比基于 CNNP 的可逆数据隐藏方法高 1.06 dB。随着嵌入容量从 20,000 比特增加到 150,000 比特，改进的可逆数据隐藏方法的平均 PSNR 值仍分别高出 0.81 dB、0.60 dB、0.55 dB、0.56 dB、0.55 dB、0.54 dB、0.53 dB、0.51 dB、0.47 dB、0.41 dB、0.37 dB、0.33 dB、0.26 dB 和 0.20 dB。

表 I 所提出的基于ICNNP的方法和基于CNNP的方法在100幅图像上的平均峰值信噪比（分贝）[14]

嵌入容量（比特）	CNNP	ICNNP
10000	61.31	**62.37**
20000	58.01	**58.82**
30000	55.98	**56.58**
40000	54.43	**54.98**
50000	53.10	**53.66**
60000	51.94	**52.49**
70000	50.86	**51.40**
80000	49.85	**50.38**
90000	48.88	**49.39**
100000	47.91	**48.38**
110000	46.95	**47.36**
120000	46.03	**46.40**
130000	45.13	**45.46**
140000	44.29	**44.55**
150000	43.47	**43.67**

IV. 结论

本文提出一种用于可逆数据隐藏的改进型卷积神经网络预测器，通过整体优化从不同感受野中提取特征，并利用更多邻近像素精准预测像素值及其复杂度。在数据嵌入过程中，将灰度图像分为两个子图像，并使用 ICNNP 交替利用一个子图像预测另一个子图像。随后，根据预测像素的复杂度对像素的预测误差进行排序，并选取复杂度较低的预测误差，采用经典直方图平移策略进行数据嵌入。正确提取嵌入数据后，可无损恢复原始图像，且数据提取与图像恢复彼此独立。实验结果表明，改进型 CNNP 与经典直方图平移策略结合所取得的性能优于 [14] 中采用相同直方图平移策略的 CNNP。

参考文献

[1] Y. -Q. Shi、X. Li、X. Zhang、H. Wu 和 B. Ma，“可逆数据隐藏：过去二十年的进展”，IEEE Access，第 4 卷，3237，2016，3210 页。–。[2] M. U. Celik、G. Sharma、A. M. Tekalp 和 E. Saber，“无损广义最低有效位数据嵌入”，IEEE 图像处理汇刊，第 14 卷，第 2 期，第 253–266 页，2005 年。[3] W. Zhang、X. Hu、X. Li 和 N. Yu，“适用于一般失真度量的可逆数据隐藏最优转移概率及其应用”，IEEE 图像处理汇刊，第 24 卷，第 1 期，第 294–304 页，2015 年。[4] D. Hou、W. Zhang、Y. Yang 和 N. Yu，“不一致失真度量下的可逆数据隐藏”，IEEE 图像处理汇刊，第 27 卷，第 10 期，第 5087–5099 页，2018 年 10 月。[5] J. Tian，“利用差值扩展实现可逆数据嵌入”，IEEE 电路与系统视频技术汇刊，第 13 卷，第 8 期，第 890–896 页，2003 年 8 月。[6] D. M. Thodi 和 J. J. Rodriguez，“用于可逆水印的扩展嵌入技术”，IEEE 图像处理汇刊，第 16 卷，第 3 期，第 721–730 页，2007 年 3 月。[7] L. Luo、Z. Chen、M. Chen、X. Zeng 和 Z. Xiong，“利用插值技术实现可逆图像水印”，IEEE 信息取证与安全汇刊，第 5 卷，第 1 期，第 187–193 页，2010 年 3 月。[8] D. Coltuc，“用于可逆水印的低失真变换”，IEEE 图像处理汇刊，第 21 卷，第 1 期，第 412–417 页，2012 年 1 月。[9] X. Li、J. Li、B. Li 和 B. Yang，“基于像素值排序和预测误差扩展的高保真可逆数据隐藏方案”，信号处理，第 93 卷，第 1 期，第 198–205 页，2013 年。[10] I.-C. Dragoi 和 D. Coltuc，“基于局部预测的差值扩展可逆水印”，IEEE 图像处理汇刊，第 23 卷，第 4 期，第 1779–1790 页，2014 年 4 月。[11] Y. Qiu、Z. Qian 和 L. Yu，“通过扩展广义整数变换实现自适应可逆数据隐藏”，IEEE 信号处理快报，第 23 卷，第 1 期，第 130–134 页，2016 年 1 月。[12] W. He 和 Z. Cai，“关于基于像素值排序预测的预测误差扩展的见解”，IEEE 电路与系统视频技术汇刊，第 15 卷，第 3859–3871 页，2020 年。[13] T. Luo、G. Jiang、M. Yu、C. Zhong、H. Xu 和 Z. Pan，《基于卷积神经网络的立体图像可逆数据隐藏方法》，视觉通信与图像表示杂志，（ 73 页，2019 年。[14] R. Hu 和 S. Xiang，“基于卷积神经网络预测的可逆数据隐藏”，IEEE 信号处理快报，第 28 卷，第 464–468 页，2021 年。[15] R. Hu 和 S. Xiang，“利用卷积神经网络预测和自适应嵌入实现可逆数据隐藏”，IEEE 模式分析与机器智能汇刊，即将发表。[16] Z. Ni、Y. Shi、N. Ansari 和 W. Su，“可逆数据隐藏”，IEEE 电路与系统视频技术汇刊，第 16 卷，第 3 期，第 354–362 页，2006 年 3 月。[17] V. Sachnev、H. J. Kim、J. Nam、S. Suresh 和 Y. -Q. Shi，“利用排序与预测的可逆水印算法”，IEEE 电路与系统视频技术汇刊，第 19 卷，第 7 期，第 989–999 页，2009 年。[18] X. Li、W. Zhang、X. Gui 和 B. Yang，“基于多直方图修改的高效可逆数据隐藏”，IEEE 信息取证与安全汇刊，第 10 卷，第 9 期，第 2016–2027 页，2015 年 9 月。[19] W. Qi、X. Li、T. Zhang 和 Z. Guo，“基于多直方图修改的最优可逆数据隐藏方案”，IEEE 电路与系统视频技术汇刊，第 30 卷，第 8 期，第 2300–2312 页，2020 年 8 月。[20] J. Wang、X. Chen、J. Ni、N. Mao 和 Y. Shi，“基于多直方图的可逆数据隐藏：框架与实现”，IEEE 电路与系统视频技术汇刊，第 30 卷，第 8 期，第 2313–2328 页，2020 年 8 月。[21] B. Ou 和 Y. Zhao，“基于多直方图修改的高容量可逆数据隐藏”，IEEE 电路与系统视频技术汇刊，第 30 卷，第 8 期，第 2329–2342 页，2020 年 8 月。[22] T. Zhang、X. Li、W. Qi 和 Z. Guo，“基于位置的像素值排序及自适应成对修改，实现高效可逆数据隐藏”，IEEE 信息取证与安全汇刊，第 15 卷，第 2306–2319 页，2020 年。[23] X. Wang、X. Wang、B. Ma、Q. Li 和 Y. -Q. Shi，《基于岭回归预测器的可逆数据隐藏高精度误差预测算法》，*IEEE* 信号处理快报，第 28 卷，第 1125-1129 页，2021 年。[24] F. Huang、X. Qu、H. J. Kim 和 J. Huang，《JPEG 图像中的可逆数据隐藏》，*IEEE* 电路与系统视频技术汇刊，第 26 卷，第 9 期，第 1610–1621 页，2016 年 9 月。[25] D. Hou、H. Wang、W. Zhang 和 N. Yu，《基于 DCT 频率和块选择的 JPEG 图像可逆数据隐藏》，信号处理，第 148 卷，第 41–47 页，2018 年。[26] Y. Qiu、Z. Qian、H. He、H. Tian 和 X. Zhang，《JPEG 比特流中的优化无损数据隐藏及其中继传输扩展》，*IEEE* 电路与系统视频技术汇刊，第 31 卷，第 4 期，第 1380-1394 页，2021 年 4 月。[27] Y. Du、Z. Yin 和 X. Zhang，《基于通用 VLC 映射的 JPEG 比特流高容量无损数据隐藏》，*IEEE* 可靠与安全计算汇刊，第 19 卷，第 2 期，第 1420-1433 页，2022 年。[28] A. L. Maas、Y. H. Awni 和 Y. N. Andrew，整流器非线性可改进神经网络声学模型。载于第*30*届国际机器学习会议论文集（*ICML*-13），2013 年。[29] Y. Lecun、L. Bottou、Y. Bengio 和 P. Haffner，《应用于文档识别的基于梯度的学习》，载于 *IEEE* 论文集，第 86 卷，第 11 期，第 2278-2324 页，1998 年 11 月。[30] D. Kingma 和 J. Ba，《Adam：一种随机优化方法》，载于国际学习表征会议论文集，2015 年。[31] P. Bas 和 T. Furon，《bows-2 图像数据库》，2017 年。 [在线]。网址：http://bows2.ec-lille.fr/。