

# 基于 CNN 计算局部复杂度的可逆信息隐藏算法

胡润文<sup>1)</sup> 项世军<sup>1)</sup> 李晓龙<sup>2)</sup> 欧博<sup>3)</sup>

<sup>1)</sup> (暨南大学信息科学技术学院/网络空间安全学院 广州 510632)

<sup>2)</sup> (北京交通大学信息科学研究所 北京 100044)

<sup>3)</sup> (湖南大学信息科学与工程学院 长沙 410082)

**摘要** 在可逆信息隐藏领域,选择较小的预测误差有助于减少信息嵌入过程引起的失真.现有选择方法主要计算预测误差的局部复杂度,将信息嵌入到局部复杂度小的预测误差中.这些局部复杂度计算方法只使用部分相邻像素计算局部复杂度,且计算方法与预测部分使用的预测器无关,因此其选择性能有待提高.本文提出了一种基于 CNN 的局部复杂度计算方法 LCCN,该方法与预测部分的预测器相关,可以使用更多的相邻像素计算局部复杂度,提高局部复杂度与预测误差的相似性.与现有的局部复杂度计算方法相比,所提 LCCN 利用了 CNN 的多感受野特性和深度学习的全局优化能力,能使用更多相邻像素计算局部复杂度,从而选择更多数值较小的预测误差,提高嵌入算法性能.此外,我们还提出了一种适用于 LCCN 的两阶段 RDH 方案,该方案可以将 LCCN 应用于多种嵌入方法.实验结果表明,与现有的几种局部复杂度计算方法相比,本文所提方法在标准测试图像和 Kodak 数据集上嵌入性能更好,在基于 LPVO 的嵌入技术中,本文所提 LCCN 模型嵌入 10000 bits 后图像 Lena 的 PSNR 为 62.09 dB,分别高于 LV、FV、EE 和 LAE 算法 1.05 dB、1.02 dB、0.78 dB 和 0.90 dB.与现有的先进 RDH 算法比较,本文所提 LCCN 模型在不同图像上均取得较好效果.

**关键词** 可逆信息隐藏;局部复杂度;深度学习;卷积神经网络;预测误差;隐私计算

中图法分类号 TP309

DOI号 10.11897/SP.J.1016.2024.00776

## CNN Based Local Complexity Estimation for Reversible Data Hiding

HU Run-Wen<sup>1)</sup> XIANG Shi-Jun<sup>1)</sup> LI Xiao-Long<sup>2)</sup> OU Bo<sup>3)</sup>

<sup>1)</sup> (College of Information Science and Technology/College of Cyber Security, Jinan University, Guangzhou 510632)

<sup>2)</sup> (Institute of Information Science, Beijing Jiaotong University, Beijing 100044)

<sup>3)</sup> (College of Computer Science and Electronic Engineering, Hunan University, Changsha 410082)

**Abstract** In the reversible data hiding (RDH) community, how to select smaller prediction errors to reduce the embedding distortion is a crucial issue. A logical way is to use the neighboring pixels to estimate the local complexity of a target pixel for selection. Previous local complexity calculation methods often use the correlation between neighboring pixels to calculate the local complexity, in which only a small number of neighboring pixels are considered for calculation. In addition, these calculation methods are only related to the embedding methods, which are independent of the prediction methods. Due to the small calculation range and the separation from the prediction methods, the correlation between prediction errors and local complexities is limited and should be further considered. In this paper, we estimate that the local complexity is the prediction of the prediction errors. To better explain the estimation, the classical calculation method named local absolute error (LAE) is adopted, which greatly reveals that the local complexity is

收稿日期:2023-05-22;在线发布日期:2024-01-09. 本课题得到国家自然科学基金(No. 62272197, No. 62372037)、广东省基础与应用基础研究基金自然科学基金(No. 2023A1515011928)资助. 胡润文,博士研究生,主要研究领域为信息隐藏、深度学习. E-mail:runwen\_hu@qq.com. 项世军(通信作者),博士,教授,中国计算机学会(CCF)会员,主要研究领域为多媒体信息安全、可逆信息隐藏. E-mail:shijun\_xiang@qq.com. 李晓龙,博士,教授,主要研究领域为多媒体内容安全、图像处理. 欧博,博士,副教授,主要研究领域为可逆信息隐藏和相关扩展理论.

highly correlated with the prediction error. Based on the discovery, we propose a new CNN-based method titled LCCN to better calculate the local complexity of pixels. To our knowledge, we are the first to calculate the local complexity by using the deep learning-based method. Compared with previous local complexity calculation methods, the proposed LCCN can use more neighboring pixels as the context by exploiting the multi-receptive fields of CNN and the global optimization capacities of deep learning. As a bonus, satisfactory performance in selecting those smaller prediction errors for RDH can be achieved, e. g., the proposed LCCN can achieve lower mean and variance of the sorted prediction errors than the classical local complexity calculation methods, including the local variance (LV), forward variance (FV), error energy (EE), and LAE. In addition, the correlation coefficient between prediction error and local complexity is higher and the value range of the sorted prediction is more concentrated. Furthermore, we put forward a LCCN based two-stage embedding scheme, which is suitable for various embedding techniques. With several standard benchmark images and the Kodak dataset, experimental results have shown that the proposed LCCN can be applied for existing dominant embedding techniques to improve the embedding performance with higher PSNR values, e. g., the PSNR of the image Lena in the LPVO-based embedding technique can reach as high as 62.09 dB after embedding 10000 bits, which is higher than 1.05 dB, 1.02 dB, 0.78 dB, and 0.90 dB compared to LV, FV, EE, and LAE algorithms, respectively. Compared with several state-of-the-art works in eight benchmark images, the proposed method can achieve better results on average and in most cases. To further reveal the improvement of the proposed LCCN, the Kodak dataset is adopted to calculate the average PSNR with different embedding capacities, in which the proposed LCCN can achieve a higher average PSNR. For future works, we focus on improving the theory of local complexity and enhancing the similarity between local complexity and prediction error by using deep learning. In addition, the consistency relationship between prediction techniques and local complexity calculation methods should be further investigated.

**Keywords** reversible data hiding; local complexity; deep learning; convolution neural networks; prediction error; private computing

## 1 引言

可逆数据隐藏(Reversible Data Hiding, RDH)是信息隐藏研究领域中的一个重要分支,是多媒体信息安全的一个重要研究方向.发送者采用 RDH 技术将信息嵌入到多媒体文件的冗余部分,接收者能够正确提取信息并无损恢复原始多媒体文件.随着互联网技术的发展,RDH 技术逐渐被应用到各种场景,包括内容完整性认证、版权保护和秘密通信等<sup>[1-4]</sup>.在以图像为载体的信息隐藏研究领域中(如军事、医学图像和法律文书图片等),图像内容质量的精度需求往往较高,需要在传输过程中保护图像的完整性,因此 RDH 技术值得进一步研究.

在基于图像 RDH 技术的研究方向中,可以分为预测和嵌入两部分.图像预测部分旨在提高图像

像素的预测精度,减小预测过程产生的误差.目前已有较多的图像预测器被提出,包括差分预测器<sup>[5]</sup>(Difference Predictor, DP)、中值边缘检测预测器<sup>[6]</sup>(Median Edge Direction Predictor, MEDP)、梯度自适应预测器<sup>[7-8]</sup>(Gradient Adjusted Predictor, GAP)、菱形预测器<sup>[9-10]</sup>(Rhombus Predictor, RP)和多预测器<sup>[11]</sup>等.上述图像预测器利用目标像素与相邻像素之间的相关性,在预测过程中使用相邻区域内的像素计算目标像素的预测值.为了进一步提高图像预测部分的精度,研究者结合深度学习中的卷积神经网络(Convolutional Neural Network, CNN)等技术提出了新的图像预测方法.目前,CNN 在众多领域上得到了应用,包括计算机视觉<sup>[12-13]</sup>、图像恢复<sup>[14-15]</sup>、目标检测<sup>[16-17]</sup>和隐写分析<sup>[18]</sup>等.在 RDH 领域,已有使用 CNN 提高预测部分精度的相关工作.Luo 等人<sup>[19]</sup>提出一种基于

CNN 的立体图像预测方法,该方法使用立体图像的左右视图实现相互预测.在灰度图像上,Hu 等人<sup>[20]</sup>提出了基于 CNN 的预测器来提高预测部分的性能,该方法将灰度图像划分为两部分以实现相互预测.为了进一步提高预测性能,研究者提出了新的 CNN 预测器<sup>[21-22]</sup>.将 CNN 应用到图像的预测部分,RDH 技术得到了快速发展,利用深度学习提高图像预测部分的性能是目前 RDH 技术的研究方向之一.

与图像预测部分不同,图像嵌入部分将信息嵌入到图像的冗余中,旨在减少信息嵌入过程引起的失真.现有嵌入策略设计了高效的嵌入方法,主要包括差分扩展<sup>[5]</sup>、直方图平移<sup>[3]</sup>、预测误差扩展<sup>[6,23]</sup>、预测误差直方图(Prediction Error Histogram, PEH)<sup>[10]</sup>、成对预测误差扩展(Pairwise Prediction Error Expansion, PPEE)<sup>[24]</sup>、像素值排序(Pixel Value Ordering, PVO)<sup>[25-26]</sup>等方法.上述嵌入技术利用预测误差间的关系,有效降低了信息嵌入过程引起的失真,提高了算法性能.

在保证可逆性的前提下,将预测部分与嵌入部分的技术结合,可以构建高效的 RDH 算法.具体而言,预测部分计算图像像素的预测值和预测误差,嵌入部分将信息嵌入到预测误差中,预测值和嵌入信息后的预测误差组合从而生成嵌入信息的图像.由于信息嵌入到数值较小的预测误差时产生的失真较小,现有 RDH 算法的嵌入部分通常选择预测误差较小的位置嵌入信息,以提高嵌入信息后图像的视觉质量.文献<sup>[27]</sup>总结了现有 RDH 算法中常用的位置选择策略,这些策略主要利用了相邻像素间的相关性,这种相关性也被称为局部复杂度(Local Complexity, LC).为了计算单个预测误差的局部复杂度,研究者提出了不同的计算方法,如局部方差<sup>[10]</sup>、前向方差<sup>[28]</sup>、误差能量估计<sup>[29]</sup>和局部绝对误差<sup>[30]</sup>等,上述局部复杂度计算方法将在章节 2.1 详细介绍.一般而言,局部复杂度越大,预测误差的数值也越大.为了降低嵌入信息引起的失真,在嵌入信息前,通常将预测误差按局部复杂度从小到大进行排序,然后将信息嵌入到局部复杂度小的预测误差中.基于此,将位置选择策略与嵌入部分技术相结合,有助于提高嵌入算法的性能.

然而,现有局部复杂度计算方法存在以下不足:(1)现有方法只使用少数相邻像素计算局部复杂度,使用更大相邻像素范围计算局部复杂度时存在无法优化的局限性;(2)现有方法只与嵌入部分的技术相关,独立于预测部分的预测器,导致局部复杂度与预

测误差的相关性不强.为了解决上述问题,本文基于 CNN 提出了新的局部复杂度计算方法,该方法与预测部分的预测器相关,能使用更多的相邻像素计算局部复杂度,有助于提高不同嵌入算法的性能.本文首先分析现有的局部复杂度计算方法,然后对局部复杂度的计算原理作假设并进行解释.根据所提假设,本文基于 CNN 提出了新的局部复杂度计算网络(Local Complexity Calculation Network, LCCN)模型.所提 LCCN 模型具有以下重要特征:(1)模型内部参考现有局部复杂度的计算方式构建残差块,用于提取输入图像的高维特征;(2)模型充分利用了 CNN 的多感受野特性和深度学习全局优化能力,能有效提高局部复杂度与预测误差之间的相关性.

本文第 2 节介绍了经典的局部复杂度计算方法,对局部复杂度的计算原理提出假设并进行解释,基于该假设,提出了新的局部复杂度计算方法 LCCN;第 3 节提出了适用于 LCCN 模型的 RDH 算法框架;第 4 节进行实验并分析结果,利用所提框架构建不同的 RDH 算法进行对比实验,比较本文所提 LCCN 模型与现有的局部复杂度计算方法在选择预测误差和提高不同嵌入算法上的性能;第 5 节对本文作总结.

## 2 局部复杂度计算

本节首先介绍现有的局部复杂度计算方法,根据这些方法,对局部复杂度的计算原理提出假设并进行解释.基于该假设,本文提出新的局部复杂度计算方法 LCCN 模型,该方法利用 CNN 的多感受野特性和深度学习的全局优化能力,可以使用更多相邻像素计算局部复杂度,从而提高局部复杂度与预测误差的相似性.

### 2.1 现有局部复杂度计算方法

现有方法通常利用预测误差对应位置相邻像素之间的相关性计算局部复杂度.局部复杂度与预测误差存在相关性,其值越大,预测误差数值也越大.基于此,RDH 的嵌入部分通常利用局部复杂度选择数值较小的预测误差嵌入信息,从而降低嵌入过程引起的失真.除了单个预测误差的局部复杂度计算方法外,还存在多个预测误差的局部复杂度计算方法,包括预测误差对<sup>[31-32]</sup>和预测误差块<sup>[26,33-34]</sup>等.由于预测误差对和预测误差块的局部复杂度可以看作是单个预测误差局部复杂度的统计值,因此本文主要研究单个预测误差的局部复杂度计算方法.现有

局部复杂度计算方法包括以下四种。

### 2.1.1 局部方差

在文献[10]和[35]中,预测误差对应位置相邻像素间的局部方差(Local Variance, LV)被用作局部复杂度.图 1(a)展示了预测误差对应的目标像素  $x$  和相邻像素  $v_1, v_2, v_3$  和  $v_4$ .局部方差 LV 的计算公式如下所示:

$$LV = \frac{1}{4} \sum_{k=1}^4 (\Delta v_k - \Delta \bar{v})^2 \quad (1)$$

其中  $\Delta v_1 = |v_1 - v_2|, \Delta v_2 = |v_2 - v_3|, \Delta v_3 = |v_3 - v_4|, \Delta v_4 = |v_4 - v_1|, \Delta \bar{v} = \sum_{k=1}^4 \Delta v_k / 4$ .

由于每个预测误差的局部方差反映了该区域内像素间的离散程度,因此可用于选择区域相关性高的预测误差.

### 2.1.2 前向方差

文献[28]采用预测误差对应像素前向方向上相邻像素的方差(Forward Variance, FV)作为局部复杂度,具体计算公式如下所示:

$$FV = \sqrt{\frac{1}{4} \sum_{k=1}^4 (v_k - \bar{v})^2} \quad (2)$$

其中  $v_1, v_2, v_3$  和  $v_4$  是目标像素  $x$  的相邻像素,其分布如图 1(b)所示.  $\bar{v}$  是四个相邻像素的平均值,其计算公式为:  $\bar{v} = \sum_{k=1}^4 v_k / 4$ .

### 2.1.3 误差能量

误差能量(Error Energy, EE)是压缩方法中的一种编码技术<sup>[36]</sup>,文献[29]采用相邻像素之间的误差能量作为局部复杂度,具体计算公式如下所示:

$$EE = (d_h + d_v) + 2 |e_{v_7}| \quad (3)$$

其中  $e_{v_7}$  是像素  $v_7$  的预测误差.  $d_h$  和  $d_v$  的计算公式如下所示:

$$d_h = |v_6 - v_7| + |v_3 - v_4| + |v_4 - v_5| \quad (4)$$

$$d_v = |v_3 - v_7| + |v_1 - v_4| + |v_2 - v_5| \quad (5)$$

其中  $v_1, \dots, v_7$  是目标像素  $x$  的相邻像素,其具体分布如图 1(c)所示.

### 2.1.4 局部绝对误差

在文献[30]中,作者改进了局部方差 LV 的计算方式,提出局部绝对误差(Local Absolute Error, LAE),具体计算公式如下所示:

$$LAE = \frac{1}{6} \sum_{k=1}^6 \Delta v_k \quad (6)$$

其中  $\Delta v_1 = |v_1 - v_2|, \Delta v_2 = |v_2 - v_3|, \Delta v_3 = |v_3 - v_4|, \Delta v_4 = |v_4 - v_1|, \Delta v_5 = |v_1 - v_3|, \Delta v_6 = |v_2 - v_4|$ .相邻像素的分布与图 1(a)一致.

由上述计算方法可知,局部复杂度计算方法利用了相邻像素间的相关性.基于这些计算方法,本文对局部复杂度计算方法的原理作如下假设:局部复杂度计算方法利用相邻像素估计预测误差,即局部复杂度是预测误差的预测值.

为了验证上述假设,本文以图 1(a)和公式(6)为例进行解释.在图 1(a)中,参考差分预测器的预测方法,四个相邻的像素( $v_1, v_2, v_3$  和  $v_4$ )被看作是目标像素  $x$  的四个不同预测值.进一步地,公式(6)中不同相邻像素之间的误差( $\Delta v_1, \Delta v_2, \Delta v_3, \Delta v_4, \Delta v_5$  和  $\Delta v_6$ )可以看作是真实预测误差的不同估计值,采用平均的方式有助于平衡不同估计值之间的差异.对于其他局部复杂度计算方法,也可以按照局部复杂度是预测误差的预测值进行合理解释.

上述解释过程参考了差分预测器的预测方式,认为现有局部复杂度计算方法直接将相邻像素作为目标像素的预测值.然而,上述解释过程存在两个问题.一方面,仅使用单个相邻像素预测目标像素存在预测精度不足问题,容易产生预测失真;另一方面,在计算局部复杂度时,分别将两个相邻像素看作目标像素和目标像素的预测值,存在多次预测问题,容易产生累计误差.此外,使用更大范围的相邻像素有助于提高局部复杂度与预测误差之间的相似性,相关结论将在章节 4.2 以实验进行验证.基于该假设,本文提出了新的局部复杂度计算方法 LCCN 模型,该方法使用 CNN 的多感受野特性和深度学习的全局优化能力,能使用更多相邻像素计算局部复杂度.此外,本文所提 LCCN 模型与预测部分使用的预测器相关,有助于进一步提高局部复杂度与预测误差之间的相似性.

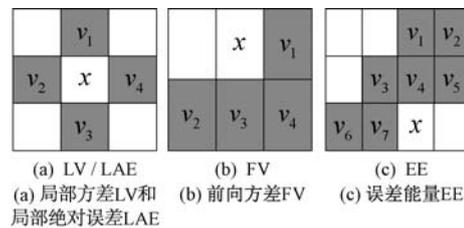


图 1 四种局部复杂度计算方法的目标像素和相邻像素的分布情况

## 2.2 本文所提出方法

根据章节 2.1,局部复杂度与预测误差的相似度越高,选择预测误差的效果越好.在训练 LCCN 模型时,将所有相邻像素和预测误差分别作为模型的输入值和目标值,通过计算模型输出的局部复杂

度与作为目标值的预测误差之间的损失来优化模型的内部参数. 完成训练后, 在测试过程中, 模型输出的局部复杂度将具有较好的选择性能.

由于本文所提 LCCN 模型在训练过程中以预测误差作为目标值, 促使模型输出的局部复杂度趋向于预测误差, 因此所提 LCCN 模型与预测部分的预测器相关. 结合 CNN 的多感受野特性和深度学习的全局优化能力, 所计算局部复杂度与预测误差的相关性更强, 有助于提高不同嵌入算法的性能.

### 2.2.1 图像预处理

在 CNN 相关研究中, 模型输入和输出的形式通常为图像. 为了训练 LCCN 模型, 图像预测部分产生的预测误差用来构建目标图像, 与预测误差对应位置的相邻像素用来构建输入图像. 当深度学习模型的参数固定时, 相同的输入值会产生相同的输出值. 为了确保 RDH 算法的可逆性, 发送端和接收端的局部复杂度要保持一致, 即保持发送端和接收端中 LCCN 模型的输入图像一致. 为此, 本文将原始图像划分成两个互不重叠的子集图像, 在一个子集图像嵌入信息时, 不会影响另外一个子集图像. 由于信息提取是信息嵌入的逆过程, 采用图像划分方法保证了发送端和接收端的输入图像一致, 因此确保了算法的可逆性. 图 2 展示了本文所提图像划分方法, 其中图 2(a) 的原始图像  $I$  被划分成图 2(b) 的叉集图像  $I_C$  和图 2(c) 的点集图像  $I_D$ . 在图 2(b) 和图 2(c) 中, 空白像素块表示像素值为 0.

基于上述图像划分方法, 点集图像  $I_D$  的所有像素都是叉集图像  $I_C$  的相邻像素. 当计算图像  $I_C$  中像素对应预测误差的局部复杂度时, 图像  $I_D$  将作为 LCCN 模型的输入图像, 反之亦然. 基于此, 在训练 LCCN 模型时, 图像  $I_D$  和图像  $I_C$  的像素将分别作为模型的输入值, 相应地, 图像  $I_C$  和图像  $I_D$  中的预测误差将分别用于构建目标图像.

以图像  $I_D$  为 LCCN 模型的输入图像时, 为了构建目标图像, 图像预测部分首先计算出预测值  $\hat{x}_{i,j}$ , 然后与目标像素  $x_{i,j}$  相减以计算预测误差:

$$e_{i,j} = x_{i,j} - \hat{x}_{i,j} \quad (7)$$

根据公式(7)可以构建预测误差图像  $I_{C_e}$ , 该图像将作为 LCCN 模型训练过程的目标图像. 相似地, 叉集图像  $I_C$  和预测误差图像  $I_{D_e}$  将分别作为输入图像和目标图像. 基于所构建的输入图像和目标图像, 可以优化 LCCN 模型的内部参数, 从而提高模型输出的局部复杂度与预测误差之间的相似性.

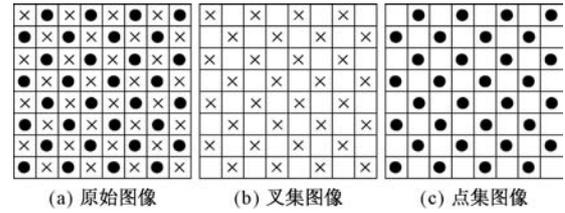


图 2 本文所提图像划分方法

### 2.2.2 LCCN 模型结构概述

本文所提模型的结构如图 3 所示. 图 3 (a) 展示了所提出 LCCN 模型的结构示意图, 其中包括多尺度特征提取部分和复杂度融合部分.

多尺度特征提取部分由多个分布平行的残差块 (Residual Block, ResBlock) 并联组成, 这种分布方式利用了 CNN 的多感受野提取图像特征<sup>[20]</sup>. 残差块的结构如图 3 (b) 所示, 其中四个卷积操作以串联形式相连, 每个卷积操作后面是激活函数 (Leaky Rectified Linear Unit, LeakyReLU). 为了利用 CNN 的多感受野提取图像特征, 第一个卷积操作中卷积核 (kernel size) 的大小被设置为  $K \times K$ , 其中  $K$  是大于 1 的奇数. 由于位于图像边缘的预测误差的相邻像素不足, 因此 ResBlock 第一个卷积操作的填充大小 (padding size) 被设置为  $\lfloor K/2 - 1 \rfloor \times \lfloor K/2 - 1 \rfloor$ , 其余填充大小被设置为  $1 \times 1$ . 此外, 在构建 LCCN 模型时参考了现有的局部复杂度计算方式, 即在 ResBlock 的第二和第三卷积操作中加入了残差运算. 具体而言, ResBlock 前两个卷积运算的输出值被视为输入值的预测值, 对应于现有局部复杂度计算中利用相邻像素作为目标像素的预测值. 预测值和输入值经过残差运算后, 输出值对应于现有局部复杂度计算中相邻像素之间的误差值. 在构建上述基本结构后, ResBlock 第二卷积操作到第四卷积操作的卷积核大小被设置为  $3 \times 3$ , 模型内部通道数被设置为  $C$ . 当输入图像经并行分布的 ResBlock 处理后, 产生不同分布的高维特征图, 这些高维特征图在空间上组合, 将作为复杂度融合部分的输入.

复杂度融合部分由两个 ResBlock 和一个卷积操作组成, 其中, 以跨接方式充分利用多尺度特征提取部分产生的高维特征图. 在复杂度融合部分, 每个卷积操作的卷积核大小和填充大小被分别设置为  $3 \times 3$  和  $1 \times 1$ , 通道数量被设置为  $C$ . 在 LCCN 模型中, 为了计算所有预测误差的局部复杂度, 多尺度特征提取部分和复杂度融合部分中所有卷积操作的步长 (stride) 被设置为  $1 \times 1$ .

LCCN 模型作为 RDH 算法的一部分, 应具有

轻量级的特性以保持较低的计算复杂度. 为此, 本文在多尺度特征提取部分中采用三个并行的 ResBlock, 其中每个 ResBlock 的第一个卷积核大小被分别设置为  $3 \times 3$ 、 $5 \times 5$  和  $7 \times 7$ . 此外, 模型内的通道数量被设置为  $C = 32$ . 采用上述超参数, 本文所提

LCCN 模型的参数量为 206497, 所占内存大小为 0.81 MB. 对于输入大小为  $512 \times 512$  的灰度图像, 其浮点数计算量 (Floating Point Operations, FLOPs) 为 53.88 GFLOPs. 基于上述设置, 本文所提 LCCN 是轻量级模型, 便于在发送端和接收端部署.

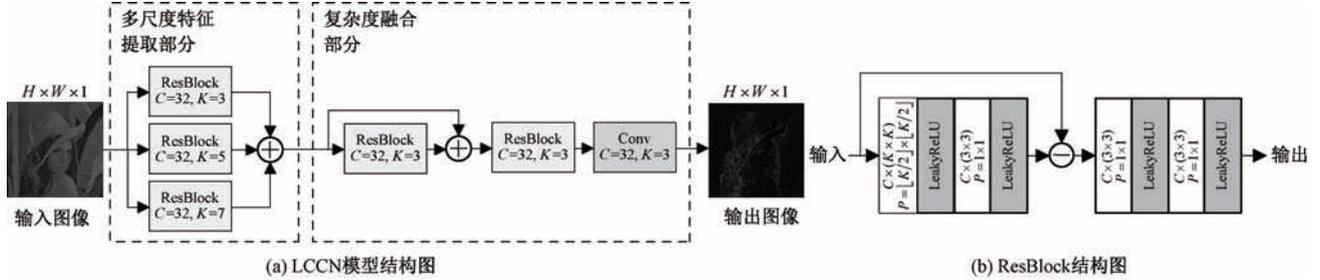


图3 所提 LCCN 模型结构示意图

在 LCCN 模型中, 多尺度特征提取部分由分布并行、卷积核大小不同的模块组成, 以提取图像在不同感受野下的高维特征; 在复杂度融合部分, 多尺度的高维特征经跨接和串联等操作后, 利用了更多的相邻像素计算局部复杂度. 结合多尺度特征提取部分和复杂度融合部分, 本文所提 LCCN 模型有效提高了局部复杂度与预测误差之间的相似性, 有助于降低信息嵌入部分引起的失真.

### 2.2.3 数据集

为了训练所提 LCCN 模型, 本文从 ImageNet<sup>[37]</sup> 中随机选取 3750 张图像, 以 8:2 的数据分布比例构建训练集 (train set) 和验证集 (validation set), 其中训练集包括 3000 张图像, 验证集包括 750 张图像. 此外, 本文将 RDH 领域中的标准测试图像用于构建测试集 (test set), 包括 8 张经典测试图像和 24 张 Kodak 数据集的图像. 具体而言, 8 张经典测试图像是大小为  $512 \times 512$  的灰度图像, 包括 Lena、Baboon、Airplane、Barbara、Lake、Boat、Elaine 和 Peppers; 24 张 Kodak 数据集中的图像是大小为  $512 \times 768$  或  $768 \times 512$  的彩色图像. 由于本文主要研究以灰度图像为载体的 RDH 算法, 因此训练集、验证集和测试集中的彩色图像均被转换为尺寸大小不变的灰度图像.

### 2.2.4 LCCN 模型训练

根据章节 2.2.1, 每个原始图像均被划分为用作输入图像的叉集图像  $I_C$  和点集图像  $I_D$ , 其对应的目标图像分别是预测部分产生的预测误差图像  $I_{De}$  和  $I_{Ce}$ . 将图像  $I_C$  (或图像  $I_D$ ) 输入到 LCCN 模型后, 模型输出局部复杂度, 计算局部复杂度与预测误差图像  $I_{De}$  (或图像  $I_{Ce}$ ) 之间的损失值, 经反向传

播<sup>[38]</sup>和 Adam 优化器<sup>[39]</sup>优化模型内部参数. 本文采用均方误差作为损失函数来优化 LCCN 模型, 具体计算公式如下所示:

$$\mathcal{L} = \frac{1}{p} \sum_{i=1}^p (I_{o_i} - I_{t_i})^2 + \lambda \|\omega\|_2^2 \quad (8)$$

其中  $p$  是图像的数量,  $\lambda$  是衰减因子,  $\omega$  表示 LCCN 模型中所有的权重参数,  $I_o$  和  $I_t$  分别表示模型输出值 (局部复杂度) 和目标值 (预测误差).

为了加快模型的收敛速度以及防止出现过拟合现象,  $\lambda$  被设置为  $10^{-3}$ . 在构建数据集后, 本文所提 LCCN 模型在 NVIDIA TITAN V GPU 上完成模型训练过程. 经过 1000 个训练周期后, 模型损失值趋于收敛. 完成上述训练后, 所提 LCCN 模型用于 RDH 算法中计算局部复杂度, 以提高嵌入算法的性能.

## 3 基于 CNN 的 RDH 框架

为了将 LCCN 模型应用到 RDH 算法, 本文在两阶段框架<sup>[10]</sup>的基础上提出了基于 CNN 的 RDH 框架. 该框架主要由三部分组成: 图像预测部分、位置选择部分和信息嵌入/提取部分.

在图像预测部分, 可以使用不同的预测器计算像素的预测值和预测误差, 预测器的选择将在章节 4.1 详细介绍. 在位置选择部分, 本文所提 LCCN 模型用于计算局部复杂度, 并将预测误差按局部复杂度从小到大进行排序. 最后, 在信息嵌入部分, 采用不同的嵌入算法将信息嵌入到局部复杂度小的预测误差中, 生成信息隐藏图像. 下面小节描述了所提 RDH 框架的信息嵌入过程和信息提取过程.

### 3.1 信息嵌入过程

图 4(a) 展示了本文所提基于 CNN 的 RDH 框

架嵌入信息  $W$  和生成信息隐藏图像  $\tilde{I}$  的具体流程图. 基于章节 2.2.1 所提图像划分方法, 原始图像  $I$  被划分成叉集图像  $I_C$  和点集图像  $I_D$ . 基于二阶段嵌入框架, 信息  $W$  被均匀划分为  $W_1$  和  $W_2$ , 用于在两个阶段中嵌入信息.

在第一阶段中, 叉集图像  $I_C$  用于嵌入信息  $W_1$ , 此时点集图像  $I_D$  作为相邻像素用于计算预测值和局部复杂度. 在预测部分, 点集图像  $I_D$  中的像素用于计算预测值, 与叉集图像  $I_C$  中的像素相减得到预测误差  $e_C$ . 在位置选择部分, 点集图像  $I_D$  作为 LCCN 模型的输入图像, 用于计算预测误差  $e_C$  的局部复杂度, 并将预测误差按局部复杂度进行排序. 最后, 采用可逆嵌入算法将信息  $W_1$  嵌入到排序后的预测误差中. 嵌入信息后的预测误差与预测值相加, 生成嵌入信息的叉集图像  $\tilde{I}_C$ .

在第二阶段中, 点集图像  $I_D$  中的像素用于嵌入信息  $W_2$ . 为了确保 RDH 算法的可逆性, 第一阶段嵌入信息的叉集图像  $\tilde{I}_C$  作为相邻像素用于计算预测值和局部复杂度. 在预测部分, 基于相同的预测器, 叉集图像  $\tilde{I}_C$  用于计算图像  $I_D$  的预测值和预测误差  $e_D$ . 然后, 将图像  $\tilde{I}_C$  作为 LCCN 模型的输入, 计算局部复杂度并将预测误差  $e_D$  按局部复杂度从小到大进行排序. 在信息嵌入部分, 采用相同的可逆

嵌入算法将信息  $W_2$  嵌入到排序后的预测误差中. 嵌入信息后的预测误差与预测值相加, 生成嵌入信息的点集图像  $\tilde{I}_D$ . 完成两阶段信息嵌入后, 将图像  $\tilde{I}_C$  和  $\tilde{I}_D$  在空间上组合, 生成信息隐藏图像  $\tilde{I}$ .

### 3.2 信息提取过程

图 4(b)展示了本文所提基于 CNN 的 RDH 框架从信息隐藏图像  $\tilde{I}$  提取信息  $W$  和恢复原始图像  $I$  的流程图. 基于章节 2.2.1 所提图像划分方法, 信息隐藏图像  $\tilde{I}$  首先被划分为嵌入信息的叉集图像  $\tilde{I}_C$  和嵌入信息的点集图像  $\tilde{I}_D$ . 由于信息提取是信息嵌入的逆过程, 因此, 图像  $\tilde{I}_D$  首先用于提取信息  $W_2$  和恢复原始图像  $I_D$ , 然后, 图像  $\tilde{I}_C$  用于提取信息  $W_1$  和恢复原始图像  $I_C$ .

在第一阶段中, 从嵌入信息的点集图像  $\tilde{I}_D$  中提取信息  $W_2$ . 预测部分采用相同的预测器, 利用图像  $\tilde{I}_C$  计算预测值和嵌入信息后的预测误差  $\tilde{e}_D$ . 在位置选择部分, 图像  $\tilde{I}_C$  作为 LCCN 模型的输入, 输出与嵌入过程第二阶段一致的局部复杂度. 预测误差  $\tilde{e}_D$  按局部复杂度从小到大排序后, 采用可逆提取算法提取信息  $W_2$  并恢复原始预测误差  $e_D$ . 将预测值与恢复后的预测误差相加, 恢复原始点集图像  $I_D$ .

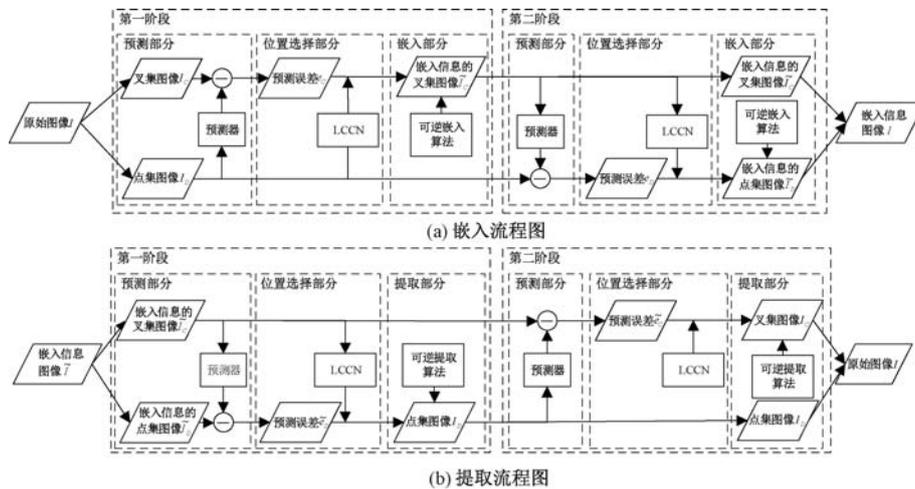


图 4 基于 CNN 的 RDH 框架

在第二阶段中, 嵌入信息的叉集图像  $\tilde{I}_C$  被用来提取信息  $W_1$ , 点集图像  $I_D$  中的像素被用来计算预测值和局部复杂度. 基于相同的预测器, 利用图像  $I_D$  计算预测值和嵌入信息后的预测误差  $\tilde{e}_C$ . 在位置选择部分, 图像  $I_D$  作为 LCCN 模型的输入计算局部复杂度. 将预测误差按局部复杂度

从小到大排序后, 采用可逆提取算法提取信息  $W_1$  并恢复预测误差  $e_C$ . 预测值与恢复后的预测误差相加, 恢复原始图像  $I_C$ . 得到恢复的图像后, 将叉集图像  $I_C$  和点集图像  $I_D$  在空间上组合以恢复原始图像  $I$ , 将提取的信息  $W_1$  和  $W_2$  组合以恢复原始信息  $W$ .

## 4 实验结果与分析

为了评估所提 LCCN 模型的性能,本文进行相关实验并分析实验结果.根据章节 3 所提 RDH 框架,在满足可逆性前提下,将不同部分组合,以构建不同的 RDH 算法.对于 RDH 框架中的图像预测部分、位置选择部分和信息嵌入部分,本文采用  $P$ 、 $L$  和  $E$  分别代表其中的技术方案.

在图像预测部分,可以采用不同的预测方法.由于本文主要研究局部复杂度计算方法,预测部分采用高精度的预测器能更好地说明所提 LCCN 模型选择预测误差的性能.为此,章节 4.1 对现有的几种预测器进行比较,并选择预测性能最好的预测器用于本文的预测部分.

在位置选择部分,本文将所提 LCCN 模型与几种经典的局部复杂度计算方法进行比较.为了说明局部复杂度与预测误差之间的相似性,章节 4.2 采用统计的方式分析了不同相邻像素范围对局部复杂度计算方法选择性能的影响.此外,章节 4.3 比较了不同局部复杂度计算方法选择预测误差的性能.上述选择性能以预测误差排序后的统计直方图、均值、方差、相关系数和数值范围进行衡量.

在信息嵌入部分,采用不同的可逆嵌入算法来说明 LCCN 模型对嵌入算法性能的提升效果.具体而言,章节 4.4 将不同的嵌入算法和位置选择策略组合以构建不同的 RDH 算法,然后比较不同的 RDH 算法嵌入信息后图像的质量.

最后,从基于 LCCN 模型所构建的 RDH 算法中选出性能较好的算法,与现有的几种先进 RDH 算法进行比较,实验结果如章节 4.5 所示.

### 4.1 预测器选择

根据章节 3 所提 RDH 框架,预测部分可以采用不同预测器.基于此,本文比较了不同图像预测器在测试集上的性能,并在后续实验中选择性能最好的预测器用于 RDH 算法的图像预测部分.

为了比较图像预测器的性能,图 5 展示了 CNNP<sup>[20]</sup>、RP<sup>[9-10]</sup>、MEDP<sup>[6]</sup>、GAP<sup>[7-8]</sup> 和 DP<sup>[5]</sup> 等预测器在测试集图像上的预测误差直方图.由图 5 可知,CNNP 的预测误差直方图的总体趋势要高于其他预测器.此外,CNNP 在范围为  $[-2, 2]$  的预测误差个数要多于其他预测器,说明 CNNP 的预测精度要高于其他预测器.因此,本文在后续实验中将 CNNP 作为预测部分使用的预测器来计算预测误差.

一般来说,在没有使用 CNN 模型作为预测器之前,RDH 程序的数据量约为 100 KB.在文献[20]中,采用轻量化 CNNP 作为预测器后的嵌入程序和提取程序数据量均约为 1.50 MB.本文所提轻量化 LCCN 模型数据量约为 0.81 MB,即结合 LCCN 模型和文献[20]的程序后数据量约为 2.31 MB.在当前带宽条件下,传输程序到接收端基本不受影响.

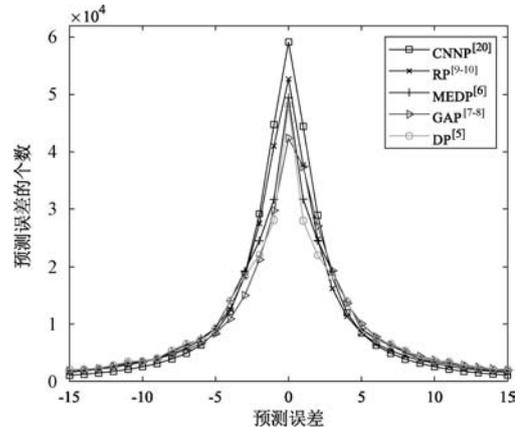


图 5 不同预测器(CNNP<sup>[20]</sup>、RP<sup>[9-10]</sup>、MEDP<sup>[6]</sup>、GAP<sup>[7-8]</sup>和 DP<sup>[5]</sup>)在测试集图像上的预测误差直方图分布情况

### 4.2 局部复杂度计算范围的影响

为了说明由不同相邻像素范围计算的局部复杂度对选择预测误差性能的影响,本文利用不同的局部复杂度计算方法进行实验,并分析所计算局部复杂度对预测误差的排序效果.

在局部方差 LV<sup>[10]</sup>和局部绝对误差 LAE<sup>[30]</sup>中,用于计算局部复杂度的相邻像素范围均为  $3 \times 3$ .为了说明不同相邻像素范围计算的局部复杂度对预测误差选择性能的影响,本文将相邻区域设定为  $3 \times 3$ 、 $5 \times 5$ 、 $7 \times 7$ 、 $9 \times 9$ 、 $11 \times 11$  和  $13 \times 13$ ,并采用 LV 和 LAE 分别计算局部复杂度.

表 1 展示了 LV 和 LAE 在不同相邻区域计算的局部复杂度选择预测误差的性能,其中,以排序后前 10000 个预测误差的均值、方差、相关系数和分布范围进行衡量.由表 1 可知,对于两种局部复杂度计算方法,随着相邻范围的增大,均值和方差逐渐变小,相关系数逐渐增大,且预测误差的范围逐渐靠近 0.然而,由于传统局部复杂度计算方法存在无法优化的局限性,导致 LV 和 LAE 在  $9 \times 9$  等更大相邻范围上出现选择性能下降的情况,如均值和方差增大,相关系数下降,分布范围增大等.根据上述实验结果可知,使用更大范围的相邻像素有助于提高局部复杂度与预测误差之间的相似性.为了实现利用更多相邻像素和建立起与预测部分的相关性,结合

具备全局优化能力的深度学习技术是可行的研究方向之一。

**表 1** 预测误差经局部复杂度计算方法 LV 和 LAE 在不同相邻像素范围计算的局部复杂度排序后的均值、方差、相关系数和数值分布范围

$L$	相邻范围	均值	方差	相关系数	分布范围
LV	$3 \times 3$	1.600	9.120	0.260	$[-64, 91]$
	$5 \times 5$	1.015	2.489	0.396	$[-21, 22]$
	$7 \times 7$	0.971	2.151	0.409	$[-21, 18]$
	$9 \times 9$	0.978	2.120	0.408	$[-21, 16]$
	$11 \times 11$	0.994	2.139	0.400	$[-21, 19]$
LAE	$13 \times 13$	1.015	2.194	0.392	$[-17, 21]$
	$3 \times 3$	1.250	5.200	0.360	$[-79, 55]$
	$5 \times 5$	0.968	2.393	0.463	$[-71, 21]$
	$7 \times 7$	0.946	2.060	0.480	$[-18, 18]$
	$9 \times 9$	0.956	2.044	0.478	$[-18, 17]$
	$11 \times 11$	0.972	2.074	0.472	$[-23, 17]$
	$13 \times 13$	0.990	2.124	0.464	$[-23, 40]$

### 4.3 LCCN 模型的有效性

在位置选择部分,局部复杂度选择预测误差的有效性体现在预测误差按局部复杂度从小到大排序后的分布情况.为了评估不同局部复杂度选择预测误差的性能,将所提 LCCN 模型与现有的局部复杂度计算方法进行比较.

由所提图像划分方法可知,测试集图像首先被划分为点集图像和叉集图像,然后使用 CNNP 计算预测值.目标像素与预测值相减得到预测误差后,分

别采用本文所提 LCCN 模型、LV<sup>[10]</sup>、FV<sup>[28]</sup>、EE<sup>[29]</sup>和 LAE<sup>[30]</sup>计算局部复杂度.此外,本文所提 LCCN 模型是基于文献[20]和[22]中的框架,为了验证 ResBlock 对 LCCN 模型性能的提升效果,训练文献[20]和[22]中的模型以计算局部复杂度,并与其结果进行比较.得到不同算法计算的局部复杂度后,将预测误差按不同的局部复杂度分别排序,比较排序后预测误差的分布情况,以评估不同局部复杂度计算方法的选择性能.

在 RDH 算法中,预测误差的大小越接近 0,嵌入信息所引起的失真越小,说明位置选择部分的性能越好.基于此,本文比较了预测误差按不同方法计算的局部复杂度排序后前  $k$  个预测误差的分布情况,其中,  $k$  的取值范围为 10000 到 40000,间隔为 10000.图 6 展示了预测误差按不同局部复杂度排序后的统计直方图,其中前  $k$  个预测误差的分布情况如图 6 (a)–(d)所示.在图 6 中,纵坐标是预测误差的数量,横坐标是预测误差的大小,其范围被设置为  $[-10, 10]$ .在前  $k$  个预测误差统计直方图中,本文所提 LCCN 模型的统计直方图分布更加集中,主要体现在排序后预测误差数值为 0 的统计个数大于其他局部复杂度计算方法.此外,对于数值范围为  $[-2, 2]$  的预测误差,经 LCCN 模型排序后预测误差的统计个数大于现有的局部复杂度计算方法 LV<sup>[10]</sup>、FV<sup>[28]</sup>、EE<sup>[29]</sup>、LAE<sup>[30]</sup>和 CNNP<sup>[20, 22]</sup>.

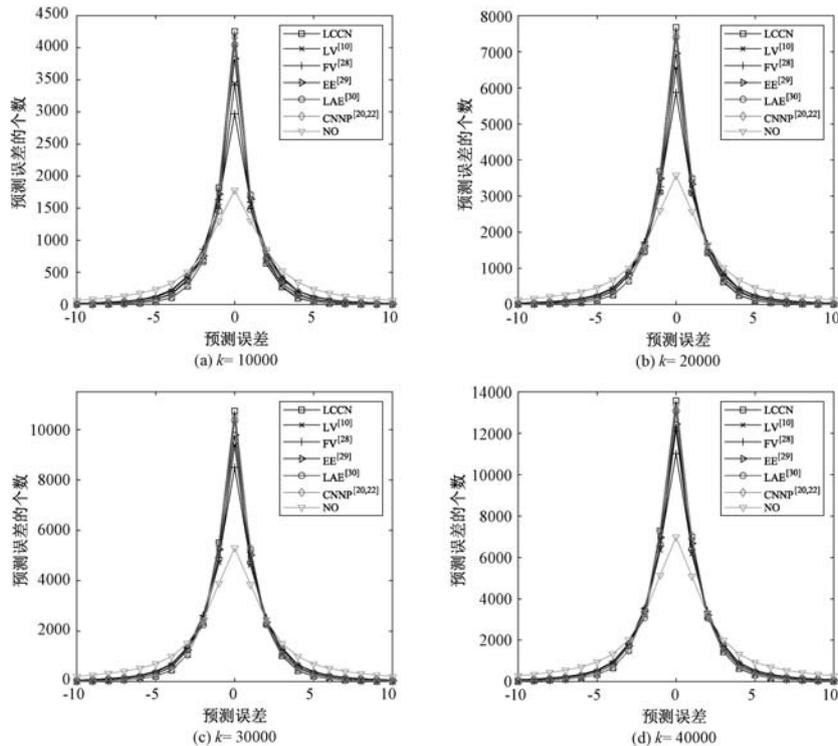


图 6 经不同局部复杂度(LCCN、LV<sup>[10]</sup>、FV<sup>[28]</sup>、EE<sup>[29]</sup>、LAE<sup>[30]</sup>和 CNNP<sup>[20, 22]</sup>)排序后前  $k$  个预测误差的统计直方图

除了以统计直方图的形式展示排序后预测误差的分布情况外,表 2 采用定量的方式展示预测误差按不同局部复杂度排序后的分布情况,即计算了排序后前 10000 个预测误差的均值、方差、相关系数和分布范围.由表 2 可知,本文所提 LCCN 模型的均值和方差为 0.92 和 1.98,均低于其他局部复杂度计算方法 LV(1.60 和 9.12),FV(1.72 和 10.26),EE(1.15 和 3.34),LAE(1.25 和 5.20)和 CNNP(0.96 和 2.17),说明预测误差按本文所提 LCCN 模型计算的局部复杂度排序后分布更加集中,其波动范围更小.相关系数反映了局部复杂度与预测误差之间的相关性强弱,其中相关系数越高,两者相关性越强.由表 2 可知,本文所提 LCCN 模型的相关系数为 0.49,均高于 LV(0.26),FV(0.23),EE(0.41),LAE(0.36)和 CNNP(0.45).此外,排序后预测误差数值的分布范围能反映局部复杂度的选择效果.本文所提 LCCN 模型的预测误差范围为 $[-16, 16]$ ,其上限和下限比其他局部复杂度计算方法的结果更加集中,说明 LCCN 模型能选择数值较小的预测误差,其选择性能优于现有局部复杂度计算方法.

表 2 预测误差按不同局部复杂度计算方法排序后的均值、方差、相关系数和分布范围

<i>L</i>	均值	方差	相关系数	分布范围
LV	1.60	9.12	0.26	$[-64, 91]$
FV	1.72	10.26	0.23	$[-98, 61]$
EE	1.15	3.34	0.41	$[-43, 37]$
LAE	1.25	5.20	0.36	$[-79, 55]$
CNNP	0.96	2.17	0.45	$[-26, 16]$
LCCN	<b>0.92</b>	<b>1.98</b>	<b>0.49</b>	<b><math>[-16, 16]</math></b>

综合上述的实验结果,本文所提 LCCN 模型可以集中更多数值接近 0 的预测误差,降低所选择的预测误差的离散程度.此外,LCCN 模型计算的局部复杂度与预测误差之间的相关系数更高,有助于提高信息嵌入部分的性能.

#### 4.4 LCCN 模型对嵌入部分性能的提升

为了说明所提 LCCN 模型对嵌入部分性能的提升效果,本文根据章节 3 所提框架进行实验比较.对于 RDH 框架中的图像预测部分、位置选择部分和信息嵌入部分,采用不同的技术方案组合以构建不同的 RDH 算法.

根据章节 4.1 的结果,本文在图像预测部分 *P* 采用 CNNP<sup>[20]</sup> 作为图像预测器,以更好地突出局部复杂度计算方法的选择性能.在位置选择部分 *L*,本文采用所提 LCCN 模型和现有的局部复杂度计算方法 LV<sup>[10]</sup>、FV<sup>[28]</sup>、EE<sup>[29]</sup> 和 LAE<sup>[30]</sup> 计算不同的

局部复杂度,然后将预测误差按局部复杂度从小到大进行排序.在嵌入部分 *E*,采用 PEH<sup>[10]</sup>、PPEE<sup>[24]</sup> 和 LPVO<sup>[26]</sup> 三种嵌入算法.在所实现的三种嵌入算法中,所有可能的溢出位置采用位置图(Location Map)进行标记,并将其作为辅助信息嵌入图像中,以确保算法的可逆性.基于上述各部分的技术,可以构建不同的 RDH 算法,如表 3 所示.在 RDH 领域中,峰值信噪比(Peak Signal to Noise Ratio, PSNR)是评估不同算法性能的重要指标.基于此,本文比较了不同 RDH 算法在测试集图像上嵌入不同容量的 PSNR,以说明 LCCN 模型对嵌入部分性能的提升作用.

表 3 根据预测部分、位置选择部分和嵌入部分所构建不同的 RDH 算法

<i>P</i>	<i>L</i>	<i>E</i>	RDH 算法
CNNP	LV	PEH	LV+PEH
	FV		FV+PEH
	EE		EE+PEH
	LAE		LAE+PEH
	LCCN		LCCN+PEH
	LV	PPEE	LV+PPEE
	FV		FV+PPEE
	EE		EE+PPEE
	LAE		LAE+PPEE
	LCCN		LCCN+PPEE
LCCN	LV	LPVO	LV+LPVO
	FV		FV+LPVO
	EE		EE+LPVO
	LAE		LAE+LPVO
LCCN	LCCN	LCCN+LPVO	

在嵌入容量相同情况下,PSNR 越高,说明 RDH 算法的性能越好.表 4 展示了所构建的不同 RDH 算法嵌入 10000 bits 到 8 张标准图像上的 PSNR 和 Kodak 数据集上 24 张图像的平均 PSNR.对于各种 RDH 算法,每张图像的最高 PSNR 以加粗的形式标记.由表 4 可知,本文所提 LCCN 模型与 PEH<sup>[10]</sup>、PPEE<sup>[24]</sup> 和 LPVO<sup>[26]</sup> 三种嵌入算法结合后在 8 张标准图像上的平均 PSNR 分别为 56.39 dB、59.27 dB 和 60.52 dB,高于 LV(55.58 dB、58.25 dB 和 59.06 dB),FV(55.50 dB、58.37 dB 和 59.23 dB),EE(55.87dB、58.58 dB 和 59.64 dB)和 LAE(55.79 dB、58.42 dB 和 59.47 dB)与三种嵌入算法组合后 RDH 算法的性能.以 LPVO 嵌入算法为例,标准测试图像 Lena 在嵌入 10000 bits 情况下,与 LCCN 模型结合的 RDH 算法的 PSNR 为 62.09 dB,分别高于与 LV、FV、EE 和 LAE 结合的 RDH 算法 1.05 dB、1.02 dB、0.78 dB 和 0.90 dB.此外,本文所提 LCCN 模型在 Kodak 数据集上的性能均优于现有局部

复杂度计算方法. 综合上述实验结果, 本文所提 LCCN 模型适用于不同嵌入算法, 有效提高了嵌入性能.

#### 4.5 与现有先进 RDH 算法的比较

根据表 4 可以知道, 本文所提 LCCN 模型在

CNNP+LCCN+LPVO 组合的 RDH 算法上取得较好的性能. 为了进一步说明上述所构建算法的性能, 与现有的几种先进 RDH 算法<sup>[24, 26-27, 30, 40-41]</sup> 在测试集图像上进行比较.

表 4 所提出的 LCCN 模型与现有局部复杂度计算方法 LV<sup>[10]</sup>、FV<sup>[28]</sup>、EE<sup>[29]</sup> 和 LAE<sup>[30]</sup> 通过使用 PEH<sup>[10]</sup>、PPEE<sup>[24]</sup> 和 LPVO<sup>[26]</sup> 嵌入技术构建的 RDH 算法在 8 张经典测试图像和 Kodak 数据集上嵌入 10000 bits 的 PSNR

RDH 算法	Lena	Baboon	Airplane	Barbara	Lake	Boat	Elaine	Peppers	Average	Kodak
LV+PEH	57.69	50.53	60.60	55.85	55.99	54.58	55.31	54.05	55.58	56.36
FV+PEH	57.69	50.32	60.72	55.65	55.94	54.47	55.19	54.03	55.50	56.07
EE+PEH	57.95	50.78	61.52	55.96	56.27	54.64	55.69	54.14	55.87	56.39
LAE+PEH	57.85	50.71	61.21	55.98	56.14	54.80	55.53	54.12	55.79	56.47
LCCN+PEH	<b>58.71</b>	<b>51.13</b>	<b>62.60</b>	<b>56.13</b>	<b>56.60</b>	<b>55.02</b>	<b>56.61</b>	<b>54.34</b>	<b>56.39</b>	<b>56.85</b>
LV+PPEE	59.44	55.08	63.06	59.29	58.26	57.07	57.54	56.25	58.25	62.00
FV+PPEE	59.43	55.10	63.26	59.27	58.58	57.18	57.69	56.41	58.37	62.05
EE+PPEE	59.65	55.26	63.44	59.44	58.82	57.37	58.14	56.51	58.58	62.11
LAE+PPEE	59.52	55.20	63.24	59.37	58.48	57.32	57.88	56.35	58.42	62.09
LCCN+PPEE	<b>60.70</b>	<b>55.53</b>	<b>64.13</b>	<b>59.76</b>	<b>59.63</b>	<b>58.12</b>	<b>59.51</b>	<b>56.74</b>	<b>59.27</b>	<b>62.57</b>
LV+LPVO	61.04	55.46	62.91	61.09	58.03	58.08	57.04	58.85	59.06	62.30
FV+LPVO	61.07	55.51	63.26	61.32	58.25	58.05	57.43	58.95	59.23	62.77
EE+LPVO	61.31	56.01	63.69	61.66	58.96	58.49	57.83	59.14	59.64	63.18
LAE+LPVO	61.19	55.94	63.35	61.61	58.39	58.58	57.67	59.04	59.47	62.99
LCCN+LPVO	<b>62.09</b>	<b>56.91</b>	<b>64.48</b>	<b>62.08</b>	<b>59.80</b>	<b>59.23</b>	<b>59.96</b>	<b>59.64</b>	<b>60.52</b>	<b>63.80</b>

为了比较不同 RDH 算法的嵌入性能, 本文采用嵌入信息后测试集图像的 PSNR 作为评估标准. 在不同嵌入容量下, 图 7 展示了测试集中 8 张标准图像在不同 RDH 算法下 PSNR 的变化趋势. 由于不同 RDH 算法在不同测试图像上的最大嵌入容量不同, 为了便于比较, 图 7 展示了各测试图像嵌入 5000 bits 到最大嵌入容量下 PSNR 的变化曲线, 其中嵌入容量的变化间隔为 1000 bits. 由图 7 (a) - (h) 可知, 在不同的嵌入容量下, 本文所提算法在大多数情况能取得更高的 PSNR. 在图像 Lena、Ba-

boon、Airplane、Barbara 和 Elaine 上均优于几种先进 RDH 算法, 而在图像 Lake、Boat 和 Peppers 等图像上性能弱于 Zhang 等人<sup>[26]</sup> 和 Xiang 等人<sup>[27]</sup> 所提算法. 为了进一步比较所构建的 RDH 算法 (CNNP+LCCN+LPVO) 的嵌入性能, 采用 Kodak 数据集进行比较. 图 8 展示了各 RDH 算法嵌入 5000 bits 到 20000 bits 的平均 PSNR 变化趋势, 其中间隔为 1000 bits. 由图 8 可知, 在不同的嵌入容量下, 基于 LCCN 模型所构建的 RDH 算法的平均 PSNR 均高于现有的先进 RDH 算法.

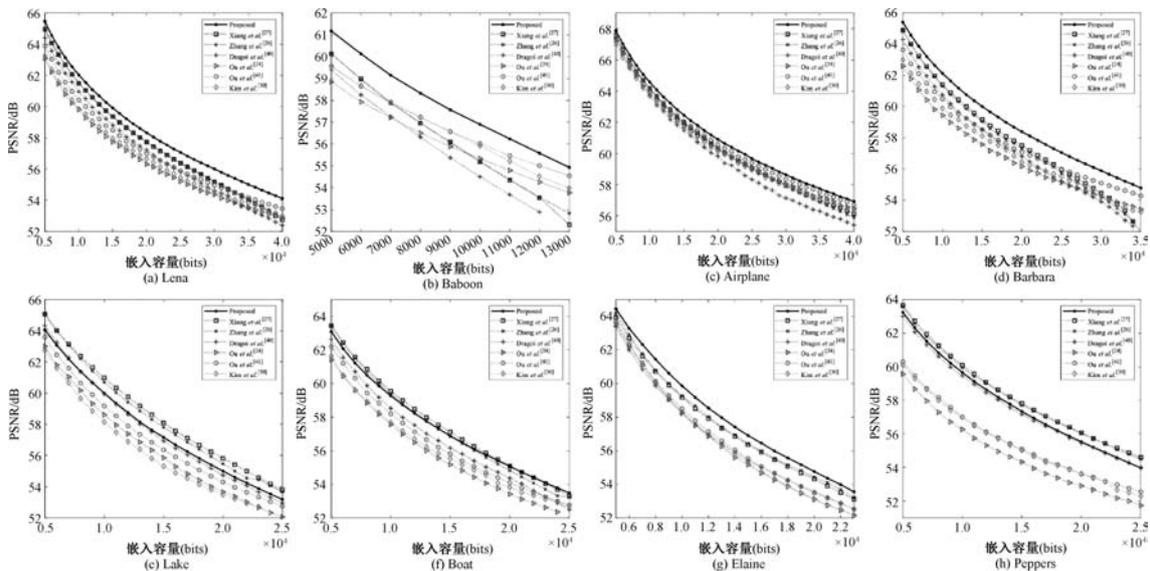


图 7 本文所提 CNNP+LCCN+LPVO 和几种先进 RDH 算法在测试集 8 张标准图像的 PSNR 变化趋势

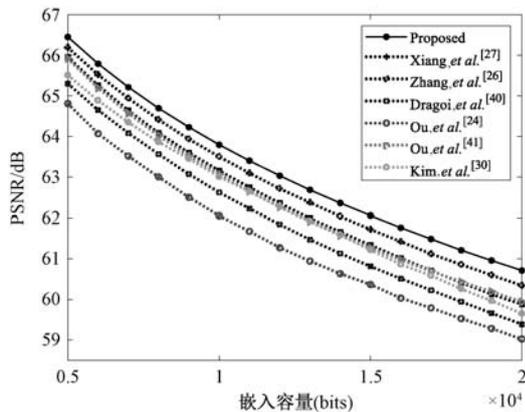


图 8 本文所提 CNNP+LCCN+LPVO 和几种先进 RDH 算法在 Kodak 数据集的平均 PSNR 变化趋势

综合本章的实验结果,本文所提出 LCCN 模型可以提高不同嵌入算法的性能;所提 LCCN 模型选择预测误差的性能优于现有局部复杂度计算方法。此外,与先进 RDH 算法相比,本文所提 LCCN 模型在不同测试图像上均能达到较好的性能,其平均性能优于现有几种先进 RDH 方法。

## 5 总 结

本文研究了 RDH 算法中的局部复杂度计算方法,根据现有的局部复杂度计算方法,提出局部复杂度是预测误差的预测值的假设,并对该假设进行合理解释。基于该假设,本文提出了新的局部复杂度计算方法 LCCN。与现有的局部复杂度计算方法相比,本文所提 LCCN 模型利用了 CNN 的多感受野特性和深度学习的全局优化能力,在计算预测误差的局部复杂度时能使用更多的相邻像素,提高了局部复杂度与预测误差之间的相似性。

在实验部分,本文采用了经典测试图像构建测试集,比较了不同局部复杂度计算方法在选择预测误差上的性能。预测误差经本文所提 LCCN 模型计算的局部复杂度排序后,其分布更加集中,表明本文所提 LCCN 模型能有效选择数值较小的预测误差,从而提高不同嵌入算法的性能。以 LPVO 嵌入算法为例,图像 Lena 在嵌入 10000 bits 情况下,与本文所提 LCCN 模型结合的 RDH 算法的 PSNR 为 62.09 dB,分别高于与 LV、FV、EE 和 LAE 结合的 RDH 算法 1.05 dB、1.02 dB、0.78 dB 和 0.90 dB。为了与现有的先进 RDH 算法比较,本文采用 CNNP+LCCN+LPVO 的 RDH 方法。根据实验结果,在不同测试图像上,本文所提算法均能达到较好

的平均 PSNR,说明本文所提 LCCN 模型能有效提高嵌入性能。综上所述,使用深度学习计算局部复杂度是具有前景的方向之一,值得进一步研究。本文针对未来局部复杂度计算工作的方向是:(1)完善局部复杂度相关理论,进一步提高基于深度学习计算的局部复杂度与预测误差之间的相似性;(2)结合预测部分计算局部复杂度,研究预测技术与局部复杂度计算方法的一致性关系。

## 参 考 文 献

- [1] Fridrich Jessica, Goljan Miroslav, Du Rui. Invertible authentication//Proceedings of the Security and Watermarking of Multimedia 2001 Contents III. San Jose, USA, 2001: 197-208
- [2] Xiang Shi-Jun, Luo Xin-Rong, Shi Shu-Xie. A novel reversible image watermarking algorithm in homomorphic encrypted domain. Chinese Journal of Computers, 2016, 39(3): 571-581 (in Chinese)  
(项世军, 罗欣荣, 石书协. 一种同态加密域图像可逆水印算法. 计算机学报, 2016, 39(3): 571-581)
- [3] Ni Zhi-Cheng, Shi Yun-Qing, Ansari Nirwan, Su Wei. Reversible data hiding. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(3): 354-362
- [4] Wu You-Qing, Guo Yu-Tang, Tang Jin, Luo Bin, Yin Zhao-Xia. Reversible data hiding in encrypted images using adaptive huffman encoding strategy. Chinese Journal of Computers, 2021, 44(4): 846-858 (in Chinese)  
(吴友情, 郭玉堂, 汤进, 等. 基于自适应哈夫曼编码的密文可逆信息隐藏算法. 计算机学报, 2021, 44(4): 846-858)
- [5] Tian Jun. Reversible data embedding using a difference expansion. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(8): 890-896
- [6] Thodi Diljith M, Rodriguez Jeffrey J. Expansion embedding techniques for reversible watermarking. IEEE Transactions on Image Processing, 2007, 16(3): 721-730
- [7] Coltuc Dinu. Improved embedding for prediction-based reversible watermarking. IEEE Transactions on Information Forensics and Security, 2011, 6(3): 873-882
- [8] Coltuc Dinu. Low distortion transform for reversible watermarking. IEEE Transactions on Image Processing, 2011, 21(1): 412-417
- [9] Luo Li-Xin, Chen Zhen-Yong, Chen Ming, Zeng Xiao, Xiong Zhang. Reversible image watermarking using interpolation technique. IEEE Transactions on Information Forensics and Security, 2009, 5(1): 187-193
- [10] Sachnev Vasiliy, Kim Hyoung Joong, Nam Jeho, Suresh Sundaram, Shi Yun-Qing. Reversible watermarking algorithm using sorting and prediction. IEEE Transactions on Circuits and Systems for Video Technology, 2009, 19(7): 989-999

- [11] Jafar Iyad F, Darabkh Khalid A, Al-Zubi Raed T, Al Na'mneh Rami A. Efficient reversible data hiding using multiple predictors. *The Computer Journal*, 2016, 59(3): 423-438
- [12] Kinli Furkan, Ozcan Baris, Kirac Furkan. Fashion image retrieval with capsule networks//Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops. Seoul, Republic of Korea, 2019
- [13] Szegedy Christian, Vanhoucke Vincent, Ioffe Sergey, Shlens Jon, Wojna Zbigniew. Rethinking the inception architecture for computer vision//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA, 2016; 2818-2826
- [14] Zamir Syed Waqas, Arora Aditya, Khan Salman, Hayat Munawar, Khan Fahad Shahbaz, Yang Ming-Hsuan, Shao Ling. Multi-stage progressive image restoration//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 14821-14831
- [15] Guo Shi, Yan Zi-Fei, Zhang Kai, Zuo Wang-Meng, Zhang Lei. Toward convolutional blind denoising of real photographs//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA, 2019; 1712-1722
- [16] Cao Jia-Le, Cholakkal Hisham, Anwer Rao Muhammad, Khan Fahad Shahbaz, Pang Yan-Wei, Shao Ling. D2det: Towards high quality object detection and instance segmentation//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020; 11485-11494
- [17] Ye Mao-Sheng, Xu Shuang-Jie, Cao Tong-Yi. Hynet: Hybrid voxel network for lidar based 3d object detection//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020; 1631-1640
- [18] Qian Yin-Long, Dong Jing, Wang Wei, Tan Tie-Niu. Deep learning for steganalysis via convolutional neural networks//Proceedings of the Media Watermarking, Security, and Forensics 2015. San Francisco, USA, 2015; 171-180
- [19] Luo Ting, Jiang Gang-Yi, Yu Mei, Zhong Cai-Ming, Xu Hai-Yong, Pan Zhi-Yong. Convolutional neural networks-based stereo image reversible data hiding method. *Journal of Visual Communication and Image Representation*, 2019, 61; 61-73
- [20] Hu Run-Wen, Xiang Shi-Jun. CNN prediction based reversible data hiding. *IEEE Signal Processing Letters*, 2021, 28; 464-468
- [21] Yang Xie, Huang Fang-Jun. New CNN-based predictor for reversible data hiding. *IEEE Signal Processing Letters*, 2022, 29; 2627-2631
- [22] Hu Run-Wen, Xiang Shi-Jun. Reversible data hiding by using CNN prediction and adaptive embedding. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 44(12); 10196-10208
- [23] Wu Han-Zhou, Wang Hong-Xia, Shi Yun-Qing. PPE-based reversible data hiding//Proceedings of the 4th ACM Workshop on Information Hiding and Multimedia Security. Vigo, Spain, 2016; 187-188
- [24] Ou Bo, Li Xiao-Long, Zhao Yao, Ni Rong-Rong, Shi Yun-Qing. Pairwise prediction-error expansion for efficient reversible data hiding. *IEEE Transactions on Image Processing*, 2013, 22(12); 5010-5021
- [25] Li Xiao-Long, Li Jian, Li Bin, Yang Bin. High-fidelity reversible data hiding scheme based on pixel-value-ordering and prediction-error expansion. *Signal Processing*, 2013, 93(1): 198-205
- [26] Zhang Tong, Li Xiao-Long, Qi Wen-Fa, Guo Zong-Ming. Location-based PVO and adaptive pairwise modification for efficient reversible data hiding. *IEEE Transactions on Information Forensics and Security*, 2020, 15; 2306-2319
- [27] Xiang Shi-Jun, Ruan Guan-Qi. Efficient PVO-based reversible data hiding by selecting blocks with full-enclosing context. *IEEE Transactions on Circuits and Systems for Video Technology*, 2021, 32(5); 2868-2880
- [28] Li Xiao-Long, Yang Bin, Zeng Tie-Yong. Efficient reversible watermarking based on adaptive prediction-error expansion and pixel selection. *IEEE Transactions on Image Processing*, 2011, 20(12); 3524-3533
- [29] Hong Wien. Adaptive reversible data hiding method based on error energy control and histogram shifting. *Optics Communications*, 2012, 285(2): 101-108
- [30] Kim Suah, Qu Xiao-Chao, Sachnev Vasily, Kim Hyoung Joong. Skewed histogram shifting for reversible data hiding using a pair of extreme predictions. *IEEE Transactions on Circuits and Systems for Video Technology*, 2018, 29(11): 3236-3246
- [31] Li Xiao-Long, Zhang Wei-Ming, Gui Xin-Lu, Yang Bin. A novel reversible data hiding scheme based on two-dimensional difference-histogram modification. *IEEE Transactions on Information Forensics and Security*, 2013, 8(7): 1091-1100
- [32] Qin Jian-Qiang, Huang Fang-Jun. Reversible data hiding based on multiple two-dimensional histograms modification. *IEEE Signal Processing Letters*, 2019, 26(6): 843-847
- [33] Peng Fei, Li Xiao-Long, Yang Bin. Improved PVO-based reversible data hiding. *Digital Signal Processing*, 2014, 25; 255-265
- [34] Ou Bo, Li Xiao-Long, Zhao Yao, Ni Rong-Rong. Reversible data hiding using invariant pixel-value-ordering and prediction-error expansion. *Signal Processing: Image Communication*, 2014, 29(7): 760-772
- [35] Hwang Hee-Joon, Kim Hyoung-Joong, Sachnev Vasily, Joo Sang-Hyun. Reversible watermarking method using optimal histogram pair shifting based on prediction and sorting. *KSII Transactions on Internet and Information Systems*, 2010, 4(4): 655-670
- [36] Wu Xiao-Lin, Memon Nasir. Context-based lossless inter-band compression-extending CALIC. *IEEE Transactions on Image Processing*, 2000, 9(6): 994-1001

- [37] Deng Jia, Dong Wei, Socher Richard, Li Li-Jia, Li Kai, Li Fei-Fei. Imagenet: A large-scale hierarchical image database//Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA, 2009: 248-255
- [38] LeCun Yann, Bottou Léon, Bengio Yoshua, Haffner Patrick. Gradient-based learning applied to document recognition, Proceedings of the IEEE, 1998, 86 (11): 2278-2324
- [39] Kingma Diederik P, Ba Jimmy. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, 2014
- [40] Dragoi Ioan Catalin, Caciula Ion, Coltuc Dinu. Improved pairwise pixel-value-ordering for high-fidelity reversible data hiding//Proceedings of the 25th IEEE International Conference on Image Processing (ICIP). Athens, Greece, 2018: 1668-1672
- [41] Ou Bo, Li Xiao-Long, Zhang Wei-Ming, Zhao Yao. Improving pairwise PEE via hybrid-dimensional histogram generation and adaptive mapping selection. IEEE Transactions on Circuits and Systems for Video Technology, 2018, 29(7): 2176-2190



**HU Run-Wen**, Ph. D. candidate. His current research interests include information hiding and deep learning.

**XIANG Shi-Jun**, Ph. D. , professor. His current research interests include multimedia information security and reversible information hiding.

**LI Xiao-Long**, Ph. D. , professor. His current research interests include multimedia content security and image processing.

**OU Bo**, Ph. D. , associate professor. His current research interests include reversible data hiding and the corresponding researches.

## Background

In recent years, reversible data hiding (RDH) in the grayscale image is a technique for hiding information in the cover image. With the capability of extracting the information correctly and recovering the original covers without any loss, RDH methods have been used widely. In the image processing aspects, a slight change in the pixel values of sensitive images like military images and medical images is undesirable. To protect the integrity of images, RDH technologies play a significant role that is worthy of research and development.

The classical RDH methods mainly focus on improving the prediction and embedding techniques. Among the prediction and embedding techniques, how to select the prediction errors is important to improve the embedding performance. In general, the previous location selection strategies mainly use the correlations between the neighboring pixels, including the local variance (LV), forward variance (FV), error energy (EE), and local absolute error (LAE). For unity, the correlations mentioned above are collectively referred to as the local complexity (LC) in this paper.

After reviewing the previous local complexity calculation methods, we estimate that the local complexities are the prediction values of the prediction errors and then explain the es-

timation. Based on the discovery, we proposed a new local complexity calculation network (LCCN), which can use more neighboring pixels to calculate the local complexities to improve the performance of different embedding techniques. The experimental results show that the proposed LCCN achieves better results, e. g. , the PSNR of the image Lena in the LPVO-based embedding technique can reach as high as 62.09 dB after embedding 10000 bits, which is higher than 1.05, 1.02, 0.78, and 0.90 dB compared to LV, FV, EE, and LAE algorithms, respectively. Compared with the previous RDH methods, the proposed one can achieve satisfactory results for several images.

The research in this paper is supported in part by the National Natural Science Foundation of China under grant 62272197 and 62372037, and in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2023A1515011928. Under these supports, we are interested in finding a new way to improve the embedding performance of RDH methods by using CNN. In this area, we have published many papers in top journals, such as Refs. [20], [22], and [27].