# 基于多层感知卷积和通道加权的图像隐写检测

叶学义 郭文风* 曾懋胜 张珂绅 赵知劲
(杭州电子科技大学通信工程学院 杭州 310018)

**摘 要**：针对目前图像隐写检测模型中线性卷积层对高阶特征表达能力有限，以及各通道特征图没有区分的问题，该文构建了一个基于多层感知卷积和通道加权的卷积神经网络(CNN)隐写检测模型。该模型使用多层感知卷积(Mlpconv)代替传统的线性卷积，增强隐写检测模型对高阶特征的表达能力；同时引入通道加权模块，实现根据全局信息对每个卷积通道赋予不同的权重，增强有用特征并抑制无用特征，增强模型提取检测特征的质量。实验结果表明，该检测模型针对不同典型隐写算法及不同嵌入率，相比Xu-Net, Yedroudj-Net, Zhang-Net均有更高的检测准确率，与最优的Zhu-Net相比，准确率提高1.95%～6.15%。

**关键词**：隐写检测；卷积神经网络；多层感知卷积；通道加权
**中图分类号**：TN911.73; TP309.2 **文献标识码**：A **文章编号**：1009-5896(2022)08-2949-08
**DOI**: 10.11999/JEIT210537

# Image Steganography Detection Based on Multilayer Perceptual Convolution and Channel Weighting

YE Xueyi GUO Wenfeng ZENG Maosheng ZHANG Keshen ZHAO Zhijin
(*School of Communication Engineering*, *Hangzhou Dianzi University*, *Hangzhou* 310018, *China*)

**Abstract**: For steganalysis, many studies have shown that convolutional neural networks have better performance than traditional artificially designed features. However, the ability of linear convolution layer to express higher-order features is limited and the feature map of each channel is not distinguished in the existing detection model which based on Convolutional Neural Networks (CNN). To solve these problems, an image steganography detection model based on Multi-layer perceptual convolution and channel weighting is constructed in this paper. The Multi-layer perceptual convolution layer (Mlpconv)is used to replace the traditional linear convolution layer to enhance the expressiveness ability of high-order features of the detection model. The channel weighting module is added to the model, which assigns different weights to each convolution channel based on global information, so that the useful features can be enhanced and the useless features can be suppressed, and the detection features extracted from the quality model can be improved. The experimental results show that the detection accuracy of proposed detection model is higher than that of Xu-Net, Yedroudj-Net, and Zhang-Net for different typical steganography algorithms and different embedding rates. And compared with the optimal Zhu-Net, the accuracy rate is increased by 1.95～6.15%.

**Key words**: Steganography detection; Convolutional Neural Networks (CNN); Multi-layer perceptual convolution (Mlpconv); Channel weighted

## 1 引 言

数字图像隐写检测阻止非法隐写通信活动、保障通信安全[1]，是信息安全领域的重要研究方向，得到了广泛的研究并且发展迅速，它通过对数字图像进行分析，从而判断图像是否嵌入了秘密信息。传统图像隐写检测方法主要通过提取基于相邻像素相关性的高阶统计特征和训练分类器完成载体/载秘图像的分类，如减法像素邻接矩阵(Subtractive Pixel Adjacency Matrix, SPAM)[2]和空域富模型(Spatial Rich Model, SRM)[3]。但是这些方法所提取的检测特征需要人工设计，特征提取耗费极高，而且随着隐写术的不断发展[4–6]，人工设计隐写检测特征的难度越来越大。

随着深度学习的兴起，研究者开始将深度学习引入图像隐写检测[7]。2014年，Tan等人[8]从一堆自

收稿日期：2021-06-08；改回日期：2021-12-21；网络出版：2022-01-13
*通信作者： 郭文风 gwf@hdu.edu.cn
基金项目：国家自然科学基金(U19B2016, 60802047)
Foundation Items: The National Natural Science Foundation of China (U19B2016, 60802047)

动编码器中使用无监督学习来训练卷积神经网络(Convolutional Neural Networks, CNN)。2015年Qian等人[9]提出了一个基于监督的5层CNN模型Qian-Net进行隐写检测，其检测准确率与SRM[3]相似。2016年Xu等人[10]构建了一个不同的5层CNN检测模型Xu-Net，不同点在于Xu等人借鉴了其他应用领域中取得成功的卷积神经网络结构单元，如绝对值激活层(ABSolute activation layer, ABS)和批归一化层(Batch Normalization, BN)。卷积层之后使用ABS层的目的是使模型学习残差图像的符号对称性；BN层的使用则避免模型在训练过程中出现梯度消失问题，且允许更大的学习率，大幅提高训练速度，实验结果表明，其模型的检测准确率稍优于SRM[3]。2017年，Ye等人[11]设计了一个10层的CNN图像隐写检测模型Ye-Net，该模型首次在预处理层使用了SRM[3]中使用到的30个高通滤波器进行滤波，有助于抑制图像内容和提取隐写信息；使用截断线性单元(Truncated linear unit, Trunc)作为第1个卷积层的激活函数，增加了模型的非线性因素，使模型可以拟合更复杂的特征，实验结果表明，其模型检测准确率远超于SRM[3]。Yedroudj等人[12]在2018年提出了一个新的5层CNN图像隐写检测模型Yedroudj-Net，其模型结合了Xu-Net[10]和Ye-Net[11]的优点，取得了更好的检测结果。同年，Zhang等人[13]提出了一个利用可分离卷积和空间金字塔池的CNN检测模型Zhang-Net，明显地提高了检测准确率。

虽然基于深度学习的图像隐写检测已经显著提高了检测准确率，但是现有模型所提取检测特征的表达能力仍然不能满足应用所需。为此本文构建了一个基于多层感知卷积和通道加权的图像隐写检测模型。在预处理层，对滤波器内核进行改进，使用SRM[3]中使用到的30个高通滤波器对预处理层进行初始化，按照尺寸大小，分别扩充为17个$3\times3$和13个$5\times5$大小，优化局部特征；在模型中使用多层感知卷积层来代替传统的卷积层，增加模型的抽象能力；为了使模型学习到更有利于检测的特征，进一步提高模型的检测准确率，在模型中加入通道加权模块。将该模型与Xu-Net[10], Yedroudj-Net[12], Zhang-Net[13]进行比较，在相同的实验条件下，小波加权( Wavelet Obtained Weights, WOW)[14]、空域-通用小波相对失真(Spatial-UNIversal WAvelet Relative Distortion, S-UNIWARD)[15]两种不同隐写算法的不同嵌入率下的实验结果表明，本模型表现出更高的检测准确率。

## 2 基于CNN图像隐写检测的最近典型模型

### 2.1 Yedroudj-Net

图1显示了Yedroudj-Net[12]的整体架构，其包含1个预处理层、5个卷积层以及1个由3个全连接层和1个Softmax[16]组成的输出层，卷积层包含线性卷积层、批归一化层(BN)、非线性激活层和平均池化层。

基于数字图像的信息隐写过程可以看作在载体图像的高频区域添加少量信息，为了提取多样性的高频信息，预处理层采用与Ye-Net[11]相同的处理方案，使用SRM[3]中的30个高通滤波器对输入图像进行滤波。以往的研究证明[9]，这种预处理很大程度上抑制了图像内容，暴露出隐写信息，如果没有这种预处理，模型很难收敛。鉴于ABS层和BN层在Xu-Net[10]模型中的良好表现，Yedroudj-Net[12]模型在搭建过程中同样使用了ABS层和BN层。

### 2.2 现有问题及分析

目前隐写算法的研究更趋向于将秘密信息嵌入载体图像的细节，尽可能小地改动载体图像，这导致检测模型必须提取更高阶的图像统计特性才能反映出载体图像在隐写前后的差异。因此，对于图像隐写检测而言，检测特征的表达对检测结果至关重要，但是现有的模型还存在以下问题：(1) 预处理

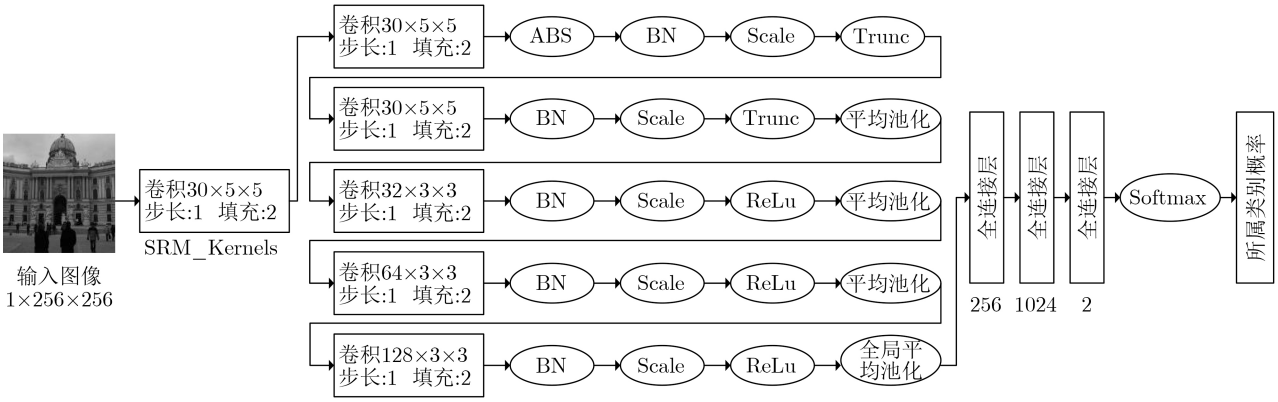



图 1 Yedroudj-Net[12]CNN架构

层需要用SRM中30个高通滤波器进行初始化，现有的处理方式是将其统一扩充为5 × 5大小，则小尺寸的高通滤波器会被填充过多的0，抑制残差图像；(2) 卷积层的线性卷积运算对高阶特征的表达能力有限；(3) 卷积得到的所有特征图以同等的权重输入下一层，没有考虑其主要性和次要性。

由此提出了基于多层感知卷积和通道加权的隐写检测模型。首先在预处理层将30个高通滤波器按照尺寸大小，分别扩充为17个3 × 3和13个5 × 5大小，避免过度补零对残差图像的抑制；其次引入Lin等人[17]提出的多层感知卷积层(Multi-layer perceptual convolution layer, Mlpconv)，它将卷积核与多层感知器连接，利用多层神经元的全连接弥补对高阶特征的表达；最后采用挤压激励(Squeeze-and-Excitation, SE)[18] 模块，实现根据全局信息对不同的特征图通道分配不同的权重。

## 3　基于多层感知卷积和通道加权的隐写检测模型

### 3.1　模型结构

如第2节所述，本文构建了一个新的隐写检测模型，整体架构如图2所示。

本模型由1个预处理层、2个多层感知卷积层(Mlpconv)、3个传统卷积层、通道加权模块SE[18]以及3层全连接层组成。

预处理层对输入图像进行滤波，滤波器采用的是SRM[3]中用到的30个高通滤波器，滤波器的值在训练过程中固定不变。滤波得到的残差图像输入多层感知卷积层，多层感知卷积层包含30个5×5大小的线性卷积核和1个多层感知器(MultiLayer Perceptron, MLP)，多层感知器由具有非线性激活函数的两个完全连接的层组成，其中非线性激活函数选择线性整流单元(Rectified Linear Unit, ReLU)，全连接层神经元的个数与线性卷积核的个数保持一致。每层卷积层的线性卷积之后都使用BN层，ABS层仅在第1个卷积层之后被使用。与Yedroudj-Net[12]相似，本模型在Mlpconv中的线性卷积之后应用Trunc激活函数，在传统卷积层应用ReLU激活函数。模型中第1层不做池化，2, 3, 4层使用平均池化，大小为5 × 5，步长为2，最后一层使用全局平均池化。通道加权模块SE[18]根据全局信息对不同通道特征图分配不同的权重，特征图完成权重重分配之后输入到下一层，SE[18]具体实现细节如3.4节所示。

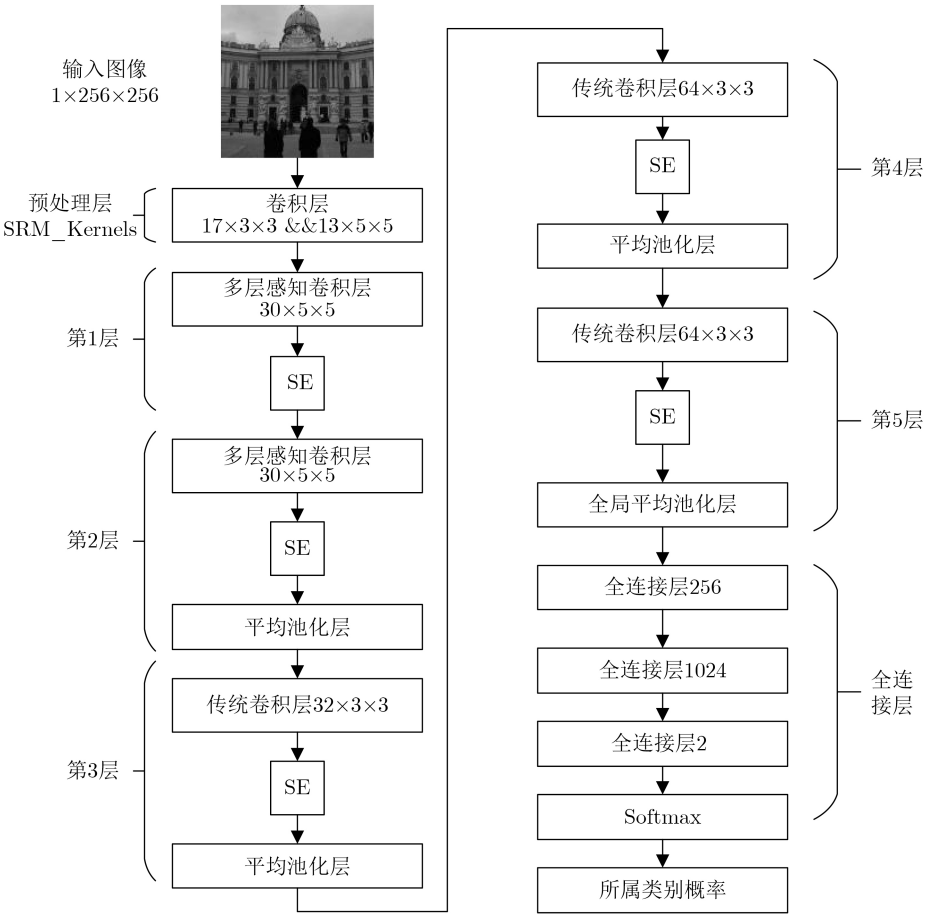



图 2 本文检测模型

### 3.2 改进预处理层

为了提取多样性的高频信息，Yedroudj-Net[12]使用SRM[3]中30个高通滤波器初始化预处理层，并将所有高通滤波器扩充为$5\times5$大小，其余元素填充为0。所提模型的预处理层中也采用这种初始化方法，不同的是并不把所有滤波器扩充为$5\times5$，已有的一些方案表明，5×5的核尺寸适合于SRM[3]中的一些高通滤波器，如“EDGE5×5”，“SQUARE5×5”和“3阶”滤波器，但对于剩余的17个高通滤波器，扩充为$5\times5$大小，将在一个较大的局部区域内建模残差元素，而较小的卷积核可以减少参数的数量，防止建模更大的区域，有效地减少计算量。所以将尺寸大于$3\times3$的扩充为$5\times5$大小，如“EDGE5×5”，“SQUARE5×5”和“3阶”共13个，而尺寸小于等于$3\times3$的扩充为$3\times3$大小，如“1阶”、“2阶”、“EDGE3×3”和“SQUARE3×3”共17个。例如其中一个2阶滤波器的扩充过程为

$$\frac{1}{2}\begin{bmatrix}1 & -2 & 1\end{bmatrix}\rightarrow\frac{1}{2}\begin{bmatrix}0 & 0 & 0\\1 & -2 & 1\\0 & 0 & 0\end{bmatrix}\tag{1}$$

扩充完后首先对所有高通滤波器进行归一化处理，使得滤波器中心元素的值为1，这样每个滤波器的参数都分布于同一量级；然后将通过高通滤波器计算得到的两部分残差图像堆叠到一起，进行阈值$T$为3的截断操作，可以有效抑制无用的图像内容，防止模型对较大的值进行建模，提高模型对特征的表达能力。

### 3.3 Mlpconv

图3所示为传统线性卷积层的示意图，对于传统的卷积层来说，一般是通过一个单层线性滤波器对输入进行线性卷积运算，然后接非线性激活函数，最终生成特征图。以ReLU激活函数为例，传统卷积层的特征图的计算公式为

$$f_{i,j,k}=\max(\boldsymbol{w}_k^{\mathrm{T}}x_{(i,j)},0)\tag{2}$$

其中，$(i,j)$表示特征图中的像素索引，$x_{(i,j)}$表示卷积窗口中以位置$(i,j)$为中心的图片块，$k$则表示特征图的通道索引。

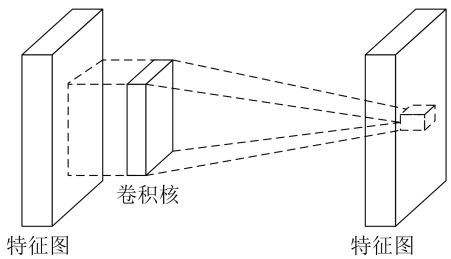



图 3 线性卷积层

众所周知，CNN通过结合较低层的特征来生成高层特征，因此在将他们生成更高层次的特征之前，对每个局部感受野进行更好的抽象是有益的。所以在本文构建的模型的低层使用了如图4所示的多层感知卷积层(Mlpconv)结构来代替传统的卷积层。

Mlpconv相当于在卷积层中包含了一个由多层感知器构成的微型网络，通过微型网络对每个局部区域进行更复杂的运算，计算更抽象的特征。Mlpconv特征图的计算公式为

$$\left.\begin{aligned}f_{(i,j,k_1)}^{1}&=\max(\boldsymbol{w}_{k_1}^{1\mathrm{T}}x_{(i,j)}+b_{k_1},0)\\&\vdots\\f_{(i,j,k_n)}^{n}&=\max(\boldsymbol{w}_{k_n}^{n\mathrm{T}}f_{(i,j)}^{n-1}+b_{k_n},0)\end{aligned}\right\}\tag{3}$$

其中，$n$表示多层感知器中的层数，ReLU函数用作多层感知器中的激活函数。当输入为多个特征图的时候，多层感知器也等效于具有多个$1\times1$卷积核的卷积层，$1\times1$的卷积核对输入特征进行加权线性重组，跨通道汇集特征信息，Mlpconv通过增加网络的非线性，以增强对高阶特征的表达能力。因为Mlpconv与卷积神经网络的结构兼容，因此可以被集成到CNN中通过反向传播进行训练。

### 3.4 通道加权

通常在设计检测模型时，会用到多个卷积核，尽可能地提取检测所需要的特征，多个卷积核提取到的特征会包含主要特征，相应地也会包含次要特征，每个特征图通道对检测结果的重要性并不完全相同。为此在模型中引入了通道加权的机制，对特征进行重新校准。SE[18]模块利用全局信息对不同的特征图通道分配不同的权重，在图像分类比赛中表现出显著的性能，通过在每层卷积层后构建SE模块[18]来完成通道加权的过程，如图5所示。

卷积得到的特征图$U$首先通过挤压(Squeeze)操作，$U$的维度为$H\times W\times C$，$C$为卷积得到的特征图通道数。Squeeze操作通过对特征图使用全局平均池化得到输出为$1\times1\times C$的特征图通道统计数据，采用全局平均池化是为了聚集全局信息，使得来自网络的全接收域的信息被它的所有层使用。计算公式为

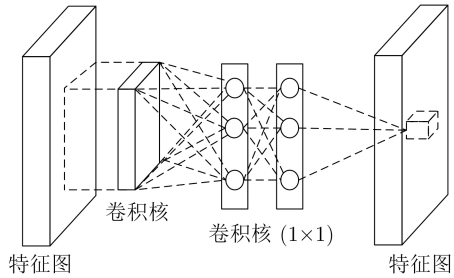



图 4 Mlpconv层

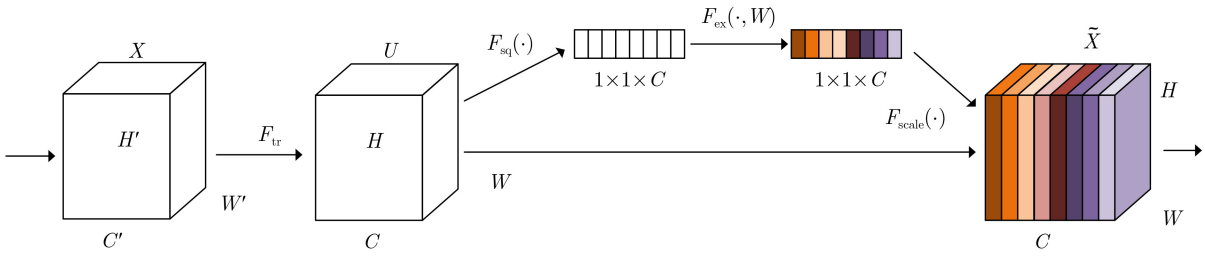



图 5 SE模块结构图

$$z_c = F_{\mathrm{sq}}(u_c) = \frac{1}{H \times W}\sum_{i=1}^{H}\sum_{j=1}^{W} u_c(i,j) \tag{4}$$

Squeeze之后是激励(Excitation)操作，将Squeeze得到的$1 \times 1 \times C$数据作为输入，输入两层具有瓶颈结构的全连接层，即具有缩减率$r$的降维层、ReLU、转换到与输入$U$的通道维度相同的升维层，目的是通过全连接层对输入进行非线性变化，在模型训练的过程中动态地学习各个通道间的非线性相互作用，然后将输入为每个通道调制权重的集合。其中缩减率$r$用来调整第1层全连接层的输出通道数，目的是降低计算量，更好地平衡检测准确率和模型复杂性。第1层全连接层的输入通道数与输出通道数的关系为

$$C_{\mathrm{out}} = C_{\mathrm{in}}/r \tag{5}$$

其中，$C_{\mathrm{out}}$表示第1层全连接层的输出通道数，$C_{\mathrm{in}}$表示第1层全连接层的输入通道数。第2层全连接层要将输出维度转换回与$U$的通道维度相等，其输入通道数等于第1层全连接层的输出通道数，输出通道数等于第1层全连接层的输入通道数即$U$的通道维度。激励的计算过程为

$$s = F_{\mathrm{ex}}(z,W) = \sigma(g(z,W)) = \sigma(W_2\delta(W_1 z)) \tag{6}$$

其中，$\delta$表示ReLU激活函数，$\sigma$表示sigmoid激活函数。最后将计算得到的权重与原先的特征图$U$相乘，对特征图进行重新校准，完成通道加权

$$\tilde{x}_c = F_{\mathrm{scale}}(u_c, s_c) = s_c u_c \tag{7}$$

其中，$\tilde{X} = [\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_c]$为权重重分配之后的特征图，大小与输入$U$相等，输出得到新的特征图直接输入后续层。通过在模型中引入通道加权机制，由SE[18]模块执行特征重新校准，实现了把重要的特征增强，不重要的特征减弱，从而让模型提取的特征指向性更强，增强模型的特征表达能力。

## 4　实验及分析

### 4.1　实验设置及数据集

所有实验均采用深度学习框架 PyTorch，在Linux操作系统下实现，图形处理器(Graphics Processing Unit, GPU)采用英伟达(NVIDIA, GTX1080)。选用数据库为BOSSBase v1.0[19]，其包含10000张大小为512×512的灰度图像，由于GPU显存限制，使用MATLAB将所有图片重采样为大小为$256 \times 256$的图像。

实验使用了两种典型的自适应隐写算法小波加权(Wavelet Obtained Weights, WOW)[14] 和空域-通用小波相对失真(Spatial-UNIversal WAvelet Relative Distortion, S-UNIWARD)[15]对载体图像进行秘密信息嵌入，由此得到10000对载体/载秘数据集，然后将10000对图像划分为两组，5000对图像分配给训练集，5000对分配给测试集，在5000对训练集中选择4000对进行训练，剩下的1000对留作验证，测试集在训练阶段保持不变。

实验中采用小批量随机梯度下降算法，权重衰减率设为0.0001，动量设置为0.95，所有层使用泽维尔 (Xavier)[20]方法初始化，训练中每一批次大小设置为16，学习率初始化为0.01，最大训练迭代数为300。实验中对比检测模型的数据直接来自对应的参考文献。

### 4.2　预处理层改进前后对比

对于隐写检测模型来说，特征提取模块的输入是预处理层所提取的残差图像，因此预处理层所提取的残差图像对于模型的检测结果至关重要。为了说明在3.2节所提对预处理层的改进对模型检测准确率的影响，用所提改进方案对Yedroudj-Net[12]预处理层进行修改，其他部分不变，与Yedroudj-Net[12]进行对比实验。在0.2 bpp和0.4 bpp两种嵌入率下，两种检测模型对WOW[14]和S-UNIWARD[15]两种隐写算法的检测准确率如表1所示。

从表1可以看出，按照3.2节提出的方案修改了预处理层之后Yedroudj-Net[12]相较原本的Yedroudj-Net[12]在两种隐写算法的两种不同的嵌入率

**表 1　Yedroudj-Net[12]修改预处理层前后准确率(%)**

| 检测模型 | WOW | | S-UNIWARD | |
|---|---|---|---|---|
| | 0.2 bpp | 0.4 bpp | 0.2 bpp | 0.4 bpp |
| Yedroudj-Net[12] | 72.20 | 85.90 | 63.30 | 77.20 |
| 改进预处理层的Yedroudj-Net | **77.21** | **87.54** | **68.81** | **81.91** |

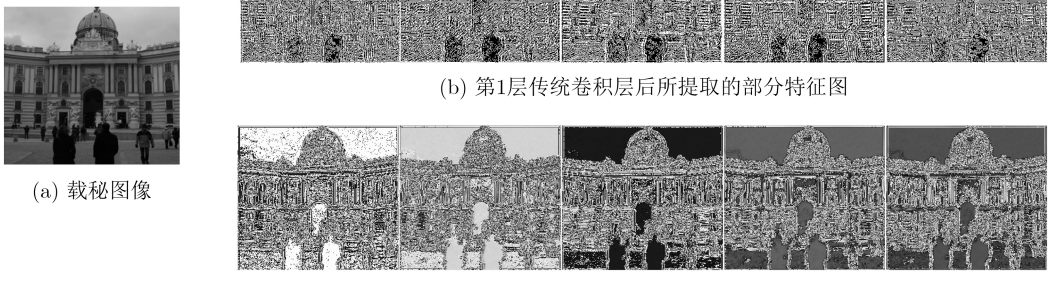
下，检测准确率都有所提升，最高提升5.51%。由此可见，改进后的预处理层所提取的残差图像对模型的检测准确率有很大的影响，更加有利于隐写检测。

### 4.3 Mlpconv对检测准确率的提升

为了说明Mlpconv对检测准确率的影响，本节实验首先对Mlpconv的抽象能力进行了分析，然后研究了Mlpconv的层数对检测结果的影响，最后与3种比较著名的检测模型进行了对比，分别是Xu-Net[10], Yedroudj-Net[12]以及Zhang-Net[13]。

#### 4.3.1 特征图可视化

特征图可以很好地描述特征提取过程，为了比较Mlpconv与传统卷积层的特征抽象能力，将经过第1层传统卷积层后的特征图与经过第1层Mlpconv之后的特征图进行了可视化处理。隐写检测模型都是使用S-UNIWARD[15]进行训练，嵌入率为0.4 bpp。选取数据集中任意一张载秘图像，模型所提取部分特征图对比如图6所示。其中图6(a)为载秘图像，图6(b)为经过第1层传统卷积层后所提取的部分特征图，图6(c)为经过第1层Mlpconv之后提取的部分特征图。实验结果表明，Mlpconv生成的特征图保留了较少的图像内容信息，并保持了较高的信噪比，在隐写检测模型中使用Mlpconv能提取出具有较强表达能力的特征。



(a) 载秘图像

(b) 第1层传统卷积层后所提取的部分特征图

(c) 第1层Mlpconv后所提取的部分特征图

图 6 模型所提取部分特征图

#### 4.3.2 Mlpconv层数对检测准确率的影响

使用Mlpconv代替传统的卷积层可以显著提高检测准确率，但过多地使用Mlpconv会带来参数量的增加，增大计算开销。为了合理使用Mlpconv，搭建性能良好的图像隐写检测模型，本节实验探究了Mlpconv层数对检测准确率的影响，分别使用1层Mlpconv、2层Mlpconv和3层Mlpconv替换传统的卷积层时，在S-UNIWARD[15]隐写算法的0.1 bpp, 0.2 bpp, 0.3 bpp以及0.4 bpp 4个不同的嵌入率的情况下做了对比实验，实验结果如图7所示。

实验结果表明，增加Mlpconv的层数并不能给模型带来检测准确率上的持续提升。当用2层Mlpconv代替传统的卷积层时，检测准确率最高，3层Mlpconv仅在0.4 bpp的嵌入率下表现较好，检测准确率为87.79%，相比较2层Mlpconv的准确率提升0.13%，并且当Mlpconv为3层时，模型所需学习的参数量相较于2层Mlpconv增长了0.6%。综合考虑，选择了使用2层Mlpconv代替传统的卷积层来搭建隐写检测模型，以达到良好的检测准确率。

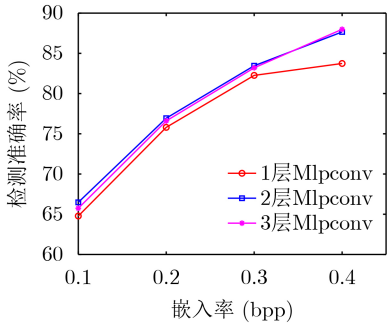



图 7 不同Mlpconv层数实验结果图

#### 4.3.3 不同方法的对比

在0.2 bpp, 0.4 bpp两种嵌入率下，Xu-Net[10], Yedroudj-Net[12], Zhang-Net[13]以及本文模型对WOW[14]和S-UNIWARD[15]两种隐写算法检测准确率如表2所示。

从表2的实验结果可以看出，针对WOW[14]隐写算法，0.2 bpp和0.4 bbp嵌入率下，本文模型检

表 2 本文模型与其他模型的对比结果(%)

| 检测模型 | WOW | | S-UNIWARD | |
|---|---|---|---|---|
| | 0.2 bpp | 0.4 bpp | 0.2 bpp | 0.4 bpp |
| Xu-Net[10] | 67.60 | 79.30 | 60.80 | 72.80 |
| Yedroudj-Net[12] | 72.20 | 85.90 | 63.30 | 77.20 |
| Zhang-Net[13] | 76.70 | 88.20 | 71.50 | 84.70 |
| 本文模型 | **80.72** | **89.59** | **76.94** | **87.66** |

测准确率与Yedroudj-Net[12]相比较提升8.52%和3.69%，与性能表现较好的Zhang-Net[13]相比，准确率提升4.02%和1.39%；隐写算法为S-UNIWARD[15]时，两种嵌入率下，所提模型与Yedroudj-Net[12]相比，准确率提升13.64%和10.46%，与Zhang-Net[13]相比，准确率提升5.44%和2.96%。实验结果表明，本模型在两种隐写算法不同嵌入率下，检测准确率均高于另外3种模型，这主要得益于Mlpconv的引入，在模型的前两层使用Mlpconv代替传统的卷积层，可以显著提升模型的检测准确率。

### 4.4 通道加权对检测准确率的提升

该模型中加入了通道加权机制，通过SE模块使模型执行特征重校准。为了验证通道加权的有效性，将原本的模型与加入SE模块之后的模型进行了对比试验，其中SE模块涉及到一个缩减率$r$的参数设置，在不同的缩减率$r$下进行了实验，实验表明根据第1层全连接层的输入通道数对$r$进行调整，保证输出通道数缩减到2时模型检测准确率表现最好，即当输入通道分别为30, 32, 64, 128时，参数缩减率$r$相对应地设置为15, 16, 32, 64时可以在准确率和复杂性之间实现良好的平衡。在WOW[14]和S-UNIWARD[15]两种隐写算法的0.2 bpp, 0.4 bpp两种嵌入率下，所提模型进行通道加权前后的对比实验结果如表3所示。

从表3可以看出，所提模型在加入通道加权的机制后，相较通道加权之前，在不同隐写算法的不同嵌入率下，检测准确率可提升0.44%～0.71%。因此，在已经设计好的模型中加入所提通道加权机制，可以进一步提升隐写检测模型的检测准确率。

## 5 结 论

基于多层感知卷积和通道加权的图像隐写检测模型，首先在预处理层中改进卷积核的扩充尺寸，将30个高通滤波器按照尺寸大小，分别扩充为17个$3\times3$和13个$5\times5$大小，避免出现因填充过多的0而抑制残差图像；其次使用多层感知卷积层代替传统的线性卷积层，增强模型对高阶特征的表达能力，提高检测准确率；最后在模型中加入通道加权的模块，对不同的卷积通道赋予不同的权重，使模型选择性地强调对检测结果更有利的信息特征，进一步提高模型的检测准确率。通过在不同隐写算法、不同嵌入率下进行的实验结果表明，与现有的CNN检测模型相比，本文模型在检测准确率上有明显的提高。

**表 3　通道加权前后模型的检测准确率(%)**

| 本文模型 | WOW | | S-UNIWARD | |
|---|---|---|---|---|
| | 0.2bpp | 0.4bpp | 0.2bpp | 0.4bpp |
| 通道加权前 | 80.72 | 89.59 | 76.94 | 87.66 |
| 通道加权后 | **81.25** | **90.15** | **77.65** | **88.10** |

## 参 考 文 献


[1] LIU Jia, KE Yan, ZHANG Zhuo, *et al*. Recent advances of image steganography with generative adversarial networks[J]. *IEEE Access*, 2020, 8: 60575–60597. doi: 10.1109/ACCESS.2020.2983175.

[2] PEVNY T, BAS P, and FRIDRICH J. Steganalysis by subtractive pixel adjacency matrix[J]. *IEEE Transactions on information Forensics and Security*, 2010, 5(2): 215–224. doi: 10.1109/TIFS.2010.2045842.

[3] FRIDRICH J and KODOVSKY J. Rich models for steganalysis of digital images[J]. *IEEE Transactions on Information Forensics and Security*, 2012, 7(3): 868–882. doi: 10.1109/tifs.2012.2190402.

[4] 付章杰, 李恩露, 程旭, 等. 基于深度学习的图像隐写研究进展[J]. 计算机研究与发展, 2021, 58(3): 548–568. doi: 10.7544/issn1000-1239.2021.20200360.

FU Zhangjie, LI Enlu, CHENG Xu, *et al*. Recent advances in image steganography based on deep learning[J]. *Computer Research and Development*, 2021, 58(3): 548–568. doi: 10.7544/issn1000-1239.2021.20200360.

[5] SHARIFZADEH M, ALORAINI M, and SCHONFELD D. Adaptive batch size image merging steganography and quantized Gaussian image steganography[J]. *IEEE Transactions on Information Forensics and Security*, 2020, 15: 867–879. doi: 10.1109/TIFS.2019.2929441.

[6] LAISHRAM D and TUITHUNG T. A novel minimal distortion-based edge adaptive image steganography scheme using local complexity[J]. *Multimedia Tools and Applications*, 2021, 80(1): 831–854. doi: 10.1007/S11042-020-09519-9.

[7] 陈君夫, 付章杰, 张卫明, 等. 基于深度学习的图像隐写分析综述[J]. 软件学报, 2021, 32(2): 551–578. doi: 10.13328/j.cnki.jos.006135.

CHEN Junfu, FU Zhangjie, ZHANG Weiming, *et al*. Review of image steganalysis based on deep learning[J]. *Journal of Software*, 2021, 32(2): 551–578. doi: 10.13328/j.cnki.jos.006135.

[8] TAN Shunquan and LI Bin. Stacked convolutional auto-encoders for steganalysis of digital images[C]. Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2014 Asia-Pacific, Siem Reap, Cambodia, 2014: 1–4. doi: 10.1109/APSIPA.2014.7041565.

[9] QIAN Yinlong, DONG Jing, WANG Wei, *et al*. Deep learning for steganalysis via convolutional neural networks[J]. *SPIE*, 2015, 9409. doi: 10.1117/12.2083479.

[10] XU Guanshuo, WU Hanzhou, and SHI Yunqing. Structural design of convolutional neural networks for steganalysis[J]. *IEEE Signal Processing Letters*, 2016, 23(5): 708–712. doi: 10.1109/LSP.2016.2548421.

[11] YE Jian, NI Jiangqun, and YI Yang. Deep learning hierarchical representations for image steganalysis[J]. *IEEE Transactions on Information Forensics and Security*, 2017, 12(11): 2545–2557. doi: 10.1109/TIFS.2017.2710946.

[12] YEDROUDJ M, COMBY F, and CHAUMONT M. Yedrouj-Net: An efficient CNN for spatial steganalysis[C]. 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, Canada, 2018: 2092–2096. doi: 10.1109/ICASSP.2018.8461438.

[13] ZHANG Ru, ZHU Feng, LIU Jianyi, *et al*. Depth-wise separable convolutions and multi-level pooling for an efficient spatial CNN-based steganalysis[J]. *IEEE Transactions on Information Forensics and Security*, 2020, 15: 1138–1150. doi: 10.1109/TIFS.2019.2936913.

[14] HOLUB V and FRIDRICH J. Designing steganographic distortion using directional filters[C]. 2012 IEEE International Workshop on Information Forensics and Security (WIFS), Costa Adeje, Spain, 2012: 234–239. doi: 10.1109/WIFS.2012.6412655.

[15] HOLUB V, FRIDRICH J, and DENEMARK T. Universal distortion function for steganography in an arbitrary domain[J]. *EURASIP Journal on Information Security*, 2014, 2014: 1. doi: 10.1186/1687-417X-2014-1.

[16] MEMISEVIC R, ZACH C, HINTON G E, *et al*. Gated softmax classification[C]. The 23rd International Conference on Neural Information Processing Systems, Red Hook, USA, 2010: 1603–1611.

[17] LIN Min, CHEN Qiang, and YAN Shuicheng. Network in network[Z]. ArXiv: 1312.4400, 2013.

[18] HU JIE, SHEN Li, ALBANIE S, *et al*. Squeeze-and-excitation networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(8): 2011–2023. doi: 10.1109/TPAMI.2019.2913372.

[19] BAS P, FILLER T, and TOMÁS PEVNÝ T. “Break Our Steganographic System”: The ins and outs of organizing BOSS[C]. Information Hiding 13th International Conference, Prague, Czech Republic, 2011: 59–70. doi: 10.1007/978-3-642-24178-9_15.

[20] GLOROT X and BENGIO Y. Understanding the difficulty of training deep feedforward neural networks[J]. *Journal of Machine Learning Research*, 2010, 9: 249–256.

叶学义：男，1973年生，副教授，研究方向为图像处理、模式识别、信息隐藏.

郭文风：女，1997年生，硕士生，研究方向为图像隐写检测.

曾懋胜：男，1998年生，硕士生，研究方向为计算机视觉.

张珂绅：男，1996年生，硕士生，研究方向为毫米波检测.

赵知劲：女，1959年生，教授，研究方向为信号处理、软件无线电.

责任编辑：余 蓉