DOI:10.11918/202201081

深度图像修复的动态特征融合取证网络

任洪昊[1],朱新山[1,2],卢俊彦[1]

(1. 天津大学 电气自动化与信息工程学院,天津 300072; 2. 数字出版技术国家重点实验室,北京 100871)

摘 要: 基于深度学习的图像修复方案在篡改后图像中遗留很少的痕迹信息给取证带来了极大的困难。目前针对深度图像修复的取证工作研究较少,并且存在篡改区域定位不准确的问题。为此,本文提出了一种动态特征融合取证网络(dynamic feature fusion forensics network,DF^{3}Net)用于定位经过深度图像修复操作的篡改区域。首先,该网络采用不同的篡改痕迹增强方式包括 SRM 滤波、空间域高通滤波和频率域高通滤波将单输入图像扩展到多输入,并提出动态特征融合模块对多种输入提取有效的修复痕迹特征后进行动态的特征融合;其次,网络采用编码器 - 解码器架构作为基础框架,并在编码器末端增加多尺度特征提取模块以获取不同尺度的上下文信息;最后,本文还设计了空间加权的通道注意力模块用于编、解码器之间的跳跃连接,以期实现有侧重地补充损失的边界细节。实验结果表明,面对不同的深度修复方案以及不同的图像数据库,DF^{3}Net 相较于现有的图像修复取证方法均可以更准确地定位篡改区域,并且对于 JPEG 压缩和高斯噪声具有较强的鲁棒性。

关键词: 图像修复取证;深度神经网络;深度修复;输入扩展;动态特征融合;痕迹特征;多尺度;注意力机制

中图分类号: TN911.73 **文献标志码:** A **文章编号:** 0367 - 6234(2022)11 - 0047 - 12

Dynamic feature fusion forensics network for deep image inpainting

REN Honghao[1], ZHU Xinshan[1,2], LU Junyan[1]

(1. School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China;
2. State Key Laboratory of Digital Publishing Technology, Beijing 100871, China)

Abstract: Deep learning-based image inpainting methods leave little trace information in the tampered image, which makes forensics extremely difficult. There are few studies on inpainting forensics, and the localization of tampered areas is inaccurate. Therefore, a dynamic feature fusion forensics network (DF^{3}Net) was proposed for locating tampered areas that have undergone deep image inpainting operations. Firstly, the network expanded single input to multi-inputs by exploiting different tamper trace enhancement methods including spatial rich model (SRM) filtering, high-pass filtering of spatial domain, and high-pass filtering of frequency domain. Then, a dynamic feature fusion module was proposed to extract effective inpainting trace features and conduct dynamic feature fusion. Secondly, the network adopted the encoder - decoder architecture as basic framework, and a multi-scale feature extraction module was added at the end of the encoder to obtain contextual information at different scales. Finally, a spatially weighted channel attention module was designed for the skip connection between encoder and decoder, so as to achieve a focused supplementation of the lost details. Experimental results show that DF^{3}Net could locate the tampered areas more accurately than existing methods on different datasets, and was robust against JPEG compression and Gaussian noise.

Keywords: image inpainting forensics; deep neural network; deep inpainting; input expansion; dynamic feature fusion; trace features; multi-scale; attention mechanism

随着互联网技术的快速发展以及智能设备的普及,数字图像已经被广泛应用于社会生活的各个领域。数字媒体时代下,由于图像处理和编辑软件的易操作性,编辑和篡改图像变得便利和简单[1];同时,篡改后的图像是不易被人眼识别的,普通人难以观察到篡改操作留下的痕迹。因此,数字图像的真实性和可靠性受到更多的关注,图像信息安全已经成为一个亟待解决的问题。图像修复取证作为取证领域的一个重要研究课题,受到了研究者的广泛关注[2-4]。

图像修复是一种根据图像已知内容对缺失或损坏区域进行修复重建的计算机视觉技术[5]。传统的图像修复可以分为基于扩散的方法[6-7]和基于样本块的方法[8-9]。但是,它们只适用于填充纹理相似的图像,不能深入理解上下文信息,对于复杂的或模式化的图像修复效果不好。为此,研究者近年来提出了基于深度学习的修复方法[10-11](简称为深度修复)。深度修复通过数据驱动的方式,很大程

收稿日期: 2022 - 01 - 17
基金项目: 国家自然科学基金(61972282,61971303)
作者简介: 任洪昊(1998—),男,硕士研究生
通信作者: 朱新山,xszhu@tju.edu.cn

度上跨越了图像底层特征与高层语义之间的语义鸿沟。它可以生成结构合理、细节丰富的图像内容,大大提升了修复效果。因此,深度修复已经成为图像修复领域的主流。

虽然图像修复技术给人们带来众多便利,但是它也可能被用于恶意的图像编辑和篡改,从而引发许多严重的社会问题。图像修复取证通过提取图像修复在操作过程中留下的篡改痕迹特征,进行篡改区域的识别和定位。因此,图像修复取证不同于一般的图像操作检测技术,它需要在图像操作层面实现像素级分类,也就是"语义分割"。

学术界针对不同的图像修复技术,提出了相应的修复取证方案。文献[12]根据通道内和通道间的局部方差构建特征集以识别基于扩散的修复方法。为了检测基于样本块的修复,文献[13]提出利用零连通特征度量图像块对相似度以识别修复区域;文献[14]在其基础上提出基于跳跃式块匹配的改进算法;文献[15]使用零连通特征识别相似块对和向量滤波的方法搜索可疑区域,并利用多区域关联来减小误警率;文献[2]采用中心像素映射的方式加快了可疑区域的搜索,并利用最大零连通特征和碎像拼接定位修复区域。

基于传统的修复方法是利用扩散或者样本块的匹配机制将非破损区域的低级非语义信息"粘贴"到破损区域。而深度修复[10-11,16]利用大规模数据学习图像的高级语义特征,可以生成更逼真的图像细节。两者不仅在操作过程的差异性很大,而且深度修复能够创造给定图像中不存在的新对象用于填补破损区域,引入了不同的修复痕迹。因此,上述传统的修复取证方案并不适用于更先进的深度修复方法。

此外,文献[1]设计了一种具有编码器-解码器结构的卷积神经网络(convolutional neural networks,CNN)[17],用于定位基于样本块的图像修复。由于深度修复生成的图像内容在视觉上是难以区分的,直接从RGB图像提取修复痕迹和学习识别特征是效果不好的;同时,CNN也倾向于学习图像主要的内容特征,而忽略修复过程中遗留的痕迹特征。这些导致该方案对深度修复的取证性能不佳。

目前,关于深度修复的取证(简称为深度修复取证)研究工作相对较少,只有文献[18]提出了一种高通预处理的全卷积取证网络。该算法利用空间高通滤波器以增强修复痕迹,实验结果证明了高通滤波对提取篡改痕迹特征的有效性。但是,深度修复方法是繁杂多样的,同一种方法遗留的痕迹也可能具有多样性和复杂性。因此,采用单一的痕迹增强方式的方案[18]有较大的局限性,鲁棒性不强。

综上所述,根据当前研究方案的不足之处,本文针对深度修复取证任务提出3个问题需要解决:

1)深度修复取证的关键是获取深度修复操作留下的微弱痕迹。由于深度修复的篡改区域与未篡改区域的视觉感知一致性[18],直接使用CNN对RGB图像的取证效果不佳。因此,利用鲁棒性强的篡改痕迹增强方式帮助痕迹特征提取是十分重要的。

2)图像修复取证任务需要对目标图像进行全局的像素级别的二分类。然而,篡改区域是具有不同的尺度大小的。因此,取证网络的决策不能仅仅局限于局部区域或目标。

3)由于取证网络中的卷积和下采样操作会导致特征图的细节信息损失,直接将高层次的特征图上采样后的输出是比较模糊的。

本文提出了一种端到端的深度修复取证网络,用于定位目标图像中经过深度修复操作的篡改区域,见图1。针对以上问题,提出了网络设计如下:采用3种痕迹增强方式将单输入扩展到多输入,通过设计的动态特征融合模块(dynamic feature fusion module, DFF)提取多种痕迹特征,并利用动态卷积[19]实现动态的特征融合。该方法结合RGB信息和CNN易忽略的细微痕迹实现有效的特征学习,具有较强的自适应能力和鲁棒性。为了弥补局部感受野导致的上下文信息不足,在编码器末端设计了多尺度特征提取模块(multi-scale feature extraction module, MFE)以扩展网络的多尺度视角。提出空间加权的通道注意力模块(spatially weighted channel attention module, SWCA)用于跳跃连接,实现有侧重地补充损失细节,避免取证结果模糊的问题。该方案能实现对篡改区域的像素级检测。在多种深度修复数据集上测试,实验结果表明该方案对于目标图像的篡改区域具有良好的定位性能。

1 输入扩展

由于提取修复遗留痕迹的能力直接影响到取证网络的性能。因此,为了达到理想的取证效果,要充分考虑深度修复原理和现有取证方法的局限性,从而更好地获取修复痕迹特征。现有的取证算法中有许多增强修复和篡改痕迹的方法,其中较为常见的是空域隐富模型(spatial rich model, SRM)[20]。它在隐写分析领域取得良好的效果后,被广泛应用于图像篡改取证领域中[21-23]。文献[24]提出的约束卷积类似于文献[18]使用的空间域高通滤波,对获取篡改痕迹具有积极作用。文献[25]通过引入频域信息以帮助网络定位人脸图像的篡改区域。文献[26]证明了离散小波变换有助于篡改痕迹的获取。

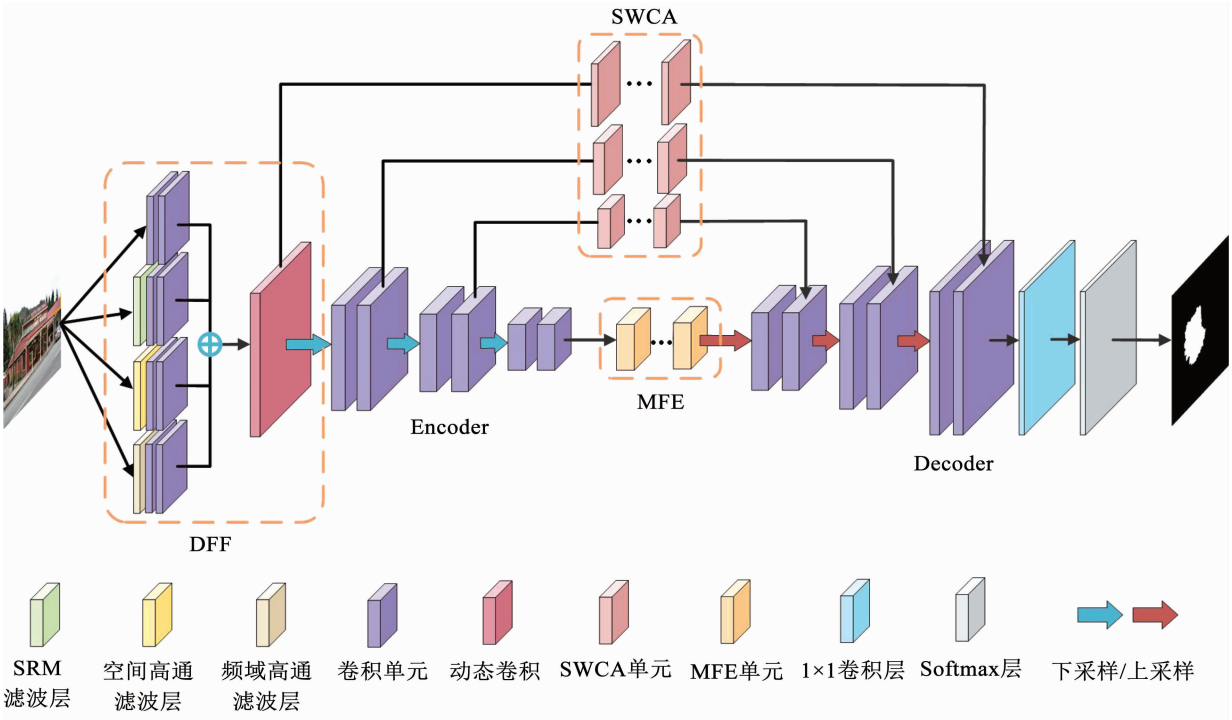


图 1　网络整体架构

Fig. 1　Architecture of the network

然而，利用上述单一的增强痕迹的方法很难学习到数据的多种特征和底层结构，导致取证网络的性能不够理想。针对这一问题，本文借鉴集成学习[27]的思想提出采用多种痕迹增强的方法用于预处理，将输入的 RGB 图像扩展到多输入，极大增加了网络对修复痕迹特征的提取能力。

虽然深度修复在视觉上残留的痕迹难以识别，但是在修复区域的形成过程中不可避免地会出现边缘不自然和纹理不连续的现象。因此，RGB 图像被直接用于检测视觉的修复痕迹。此外，为了抑制图像内容的影响，在实验测试了已有的增强修复和篡改痕迹的方法后，选择采用 SRM 滤波、空间域高通滤波和频率域高通滤波 3 种操作处理图像作为不同的输入：

1）SRM 滤波：SRM 滤波后的特征图强调的是局部噪声特征[22]。它帮助网络更加关注篡改操作遗留的噪声信息，而忽略语义信息，针对边界不明显和细节纹理丰富[5]的情况可以揭示出视觉不可见的修复痕迹。如图 2 所示，采用文献[22]选取的 3 个 SRM 滤波核作为卷积核，将 3 通道 RGB 图像映射到 3 通道的噪声特征图，得到 $\boldsymbol{X}_1 \in \boldsymbol{R}^{H\times W\times 3}$。式中，$H$ 和 W 分别为特征图或图像的高度和宽度。

2）空间域高通滤波：深度修复侧重于生成逼真的图像内容，却不能模仿原始图像中固有的不易察觉的高频信息[18]。因此，高通滤波能够揭示修复篡改后图像的异常特征，在一定程度上解决视觉差异不明显的问题。如图 3 所示，采用文献[18]初始化权重的 3 个高通滤波核作为卷积核，将 RGB 图像映射到 3 通道的残差特征图 $\boldsymbol{X}_2 \in \boldsymbol{R}^{H\times W\times 3}$。其中，卷积核参数设置为可学习的。

$$\frac{1}{4}\begin{bmatrix}0&0&0&0&0\\0&-1&2&-1&0\\0&2&-4&2&0\\0&-1&2&-1&0\\0&0&0&0&0\end{bmatrix}\quad \frac{1}{12}\begin{bmatrix}-1&2&-2&2&-1\\2&-6&8&-6&2\\-2&8&-12&8&-2\\2&-6&8&-6&2\\-1&2&-2&2&-1\end{bmatrix}\quad \frac{1}{2}\begin{bmatrix}0&0&0&0&0\\0&0&0&0&0\\0&1&-2&1&0\\0&0&0&0&0\\0&0&0&0&0\end{bmatrix}$$

图 2　3 个 SRM 滤波核

Fig. 2　Three SRM filter kernels

$$\begin{bmatrix}0&0&0\\0&-1&0\\0&1&0\end{bmatrix}\quad \begin{bmatrix}0&0&0\\0&-1&1\\0&0&0\end{bmatrix}\quad \begin{bmatrix}0&0&0\\0&-1&0\\0&0&1\end{bmatrix}$$

图 3　初始化的高通滤波核

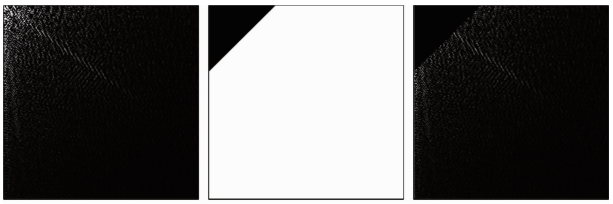
Fig. 3　Initialized high-pass filter kernels

3）频率域高通滤波：由于图像的频率特性在频率域变得直观化，更容易设计合理的滤波器以突出篡改痕迹的噪声模式。空间域和频率域不同的高通滤波方式可以实现异常特征的互相补充。本文借鉴频率域高通滤波的方式[25]抑制图像内容，放大高频的细微伪影。首先，将 RGB 图像看作 3 个独立的单通道图像 $\boldsymbol{X}_0^n \in \boldsymbol{R}^{H\times W\times 1}$，分别采用离散余弦变换（discrete cosine transform，DCT）转换到频率域。式中，$n \in \{1,2,3\}$ 为 3 个不同通道。然后，去除低频信息后反变换回空间域得到 $\boldsymbol{X}_3^n \in \boldsymbol{R}^{H\times W\times 1}$。最后，通过通道维度堆叠获取到残差特征图 $\boldsymbol{X}_3 \in \boldsymbol{R}^{H\times W\times 3}$。

整个滤波过程为

$$\boldsymbol{X}_3^n = \mathrm{IDCT}(\mathrm{DCT}(\boldsymbol{X}_0^n) \otimes \boldsymbol{\alpha}) \tag{1}$$

式中：DCT(·)为 DCT 变换，IDCT(·)为 DCT 反变换，$\boldsymbol{\alpha} \in \boldsymbol{R}^{H \times W \times 1}$为低频部分掩膜，⊗为元素相乘。如图 4 所示，由于 DCT 变换后的频域图像的低频分量位于左上角，将按照 zig-zag 顺序前 1/18 的频率点去除，即频域图像(图 4(a))与低频掩膜(图 4(b))点乘得到滤波后图像(图 4(c))。图 4(b)的黑色区域对应低频部分，设置为 0，其余白色区域为 1。



(a) 频域图像 (b) 低频部分掩膜 (c) 滤波后图像

图 4 去除低频信息

Fig. 4 Removing low-frequency information

2 取证网络模型架构

如图 1 所示，本文提出基于 CNN 的图像修复取证网络，实现对篡改区域的定位预测。由于编码器－解码器网络架构在图像语义分割等低级视觉任务中取得了极大的成功，DF^{3}Net 也采用编码器－解码器结构作为取证网络的基本框架。

2.1 动态特征融合编码器

2.1.1 动态特征融合

如图 1 所示，RGB 图像和对其不同操作得到的 3 种 3 通道特征图作为 4 组输入，采用基本卷积块将其均映射到 32 通道特征图。基本卷积块是由 2 个连续的卷积单元组成。卷积单元结构设置依次为 1 个卷积层，1 个批量归一化(batch normalization，BN)[28]层和 1 个 ReLU 激活函数。其中，卷积层的卷积核大小为 3×3，步长为 1。文中提到的基本卷积块默认为该卷积块设置。

由于不同的特征图反映了深度修复操作在不同方向和领域的遗留痕迹。为了能够充分结合多种有效的修复痕迹特征的优势，先将 4 组 32 通道特征图进行通道维度堆叠，再通过动态卷积[19]将特征图压缩到 32 通道，得到融合后的特征图 $\boldsymbol{F}_{\mathrm{r}} \in \boldsymbol{R}^{H \times W \times 32}$。其中，动态卷积的并行卷积核数设置为 3，卷积核大小为 3×3。动态卷积利用通道注意力机制[29]为多个并行的卷积核赋予权重后进行聚合。这种以非线性形式聚合多个卷积核的方式，可以实现对不同的特征图自适应的调整卷积参数，更好地对特征图进行加权和融合，极大地提升了网络的鲁棒性与取证性能。整个过程可以表示为

$$\begin{aligned}\boldsymbol{F}_{\mathrm{r}} = \mathrm{PyConv}(\mathrm{Conv}(\boldsymbol{X}_0) \oplus \mathrm{Conv}(\boldsymbol{X}_1) \oplus \\ \mathrm{Conv}(\boldsymbol{X}_2) \oplus \mathrm{Conv}(\boldsymbol{X}_3)\end{aligned} \tag{2}$$

式中：$\boldsymbol{X}_0 \in \boldsymbol{R}^{H \times W \times 3}$为输入的 RGB 图像，PyConv(·)为基本卷积块操作，Conv(·)为基本卷积块操作，⊕为通道维的堆叠操作。

2.1.2 编码器主体部分

如图 1 所示，除 DFF 和 MFE 外，编码器主体部分由下采样模块、基本卷积块组成。下采样模块通过步长为 2 的最大池化操作实现，可以有效减小特征图的尺寸，减少后续计算复杂度。基本卷积块的作用是对特征图提取修复篡改特征。通过 DFF 得到特征图 $\boldsymbol{F}_{\mathrm{r}}$ 后，编码器对其依次采取下采样和基本卷积块操作，重复相同过程 3 次，在 1/2、1/4 与 1/8 输入图像尺度下输出的特征图通道数分别为 64、128 和 256。

2.1.3 多尺度特征提取

由于篡改区域的尺度具有不确定性，为了解决网络的局部感受野引起的上下文信息不足，提高全局的分类精度，本文在编码器末端提出了 MFE 以更好地提取局部和全局的上下文信息，即获取到特征图中不同范围大小的邻域信息。如图 5 所示，该模块由 3 个部分构成。

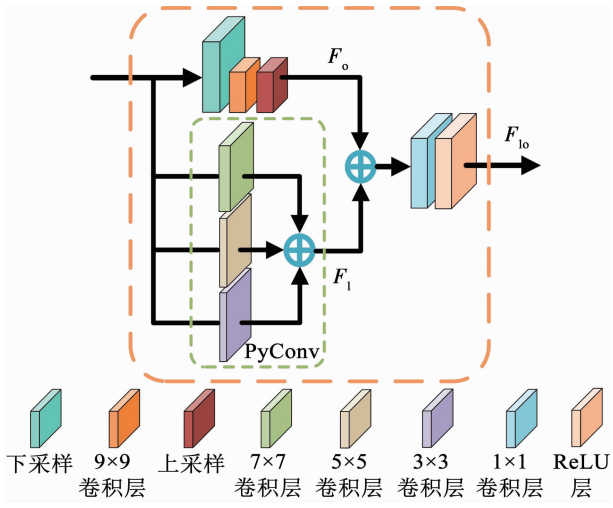


图 5 多尺度特征提取模块结构

Fig. 5 Architecture of multi-scale feature extraction module

1)局部特征提取部分：该部分引入由 3 个不同大小卷积核的卷积层并行组成的金字塔卷积(pyramidal convolution，PyConv)[30]，卷积核大小从上到下分别设置为 7×7、5×5 和 3×3，对应的特征分组数分别为 1、4 和 8，输出特征图通道数分别为输入特征图通道数的 1/2、1/4 和 1/4。最终将 3 层输出特征图进行通道维度堆叠，实现了局部特征的多尺度上下文信息的聚合，得到局部特征图 $\boldsymbol{F}_{\mathrm{l}} \in \boldsymbol{R}^{H \times W \times 256}$。此外，PyConv 在增大卷积核的同时减小

了深度，因此不会增加额外的计算成本。

2）全局特征提取部分：为了确保可以获取完整的全局特征，本文在保持合理的空间分辨率的前提下采用自适应平均池化，将输入特征图的空间大小减少到 9 × 9。然后，利用卷积核大小为 9 × 9 的卷积层提取特征。最后，使用双线性插值的方法将其映射回初始大小，得到全局特征图 $\boldsymbol{F}_{\mathrm{o}} \in \boldsymbol{R}^{H\times W\times 256}$。

3）局部和全局特征融合：将得到的局部特征和全局特征采用通道维度堆叠后，应用卷积核大小为 1 × 1 的卷积层融合不同尺度的信息。在最后增加 ReLU 激活函数，以有益于网络训练。该过程可以表示为

$$\boldsymbol{F}_{\mathrm{lo}} = \delta(\mathrm{Conv}_{1\times 1}(\boldsymbol{F}_{\mathrm{l}} \oplus \boldsymbol{F}_{\mathrm{o}})) \tag{3}$$

式中：$\boldsymbol{F}_{\mathrm{lo}} \in \boldsymbol{R}^{H\times W\times 256}$ 为局部特征与全局特征融合后的特征图，$\delta(\cdot)$ 为 ReLU 激活函数，$\mathrm{Conv}_{1\times 1}(\cdot)$ 为卷积核大小为 1 × 1 的卷积操作。

2.1.4　编码器特征可视化

图 6 展示了编码器中的 DFF（图 6（a））和 MFE（图 6（b））的特征可视化图。可视化是通过特征图同一个空间位置在通道维度相加得到的，下文中可视化处理均为该操作方式。图 6（a）中 RGB 特征（1 行3 列）、SRM 特征（2 行 1 列）、空域高通特征（2 行2 列）和频域高通特征（2 行 3 列）可视化图的差异说明利用多输入提取到的 4 组特征图反映了深度修复操作在不同方向和领域的痕迹信息。对比真实标签（1 行 1 列）和其他可视化图，对于该幅给定图像，SRM 特征（2 行 1 列）和频域高通特征（2 行 3 列）反映了较多的修复篡改痕迹，而动态融合特征（1 行 2 列）揭示的篡改区域的痕迹是最明显的。这证明 DFF 有效融合了多种痕迹特征的优势，并获取到更优异的篡改痕迹特征。对比图 6（b）的真实标签（第 4 列）和 MFE 的特征（第 1 ~ 3 列，从 1/8 图像尺度放大到 1 图像尺度）可以看出，局部特征（第 2 列）反映了篡改区域的大致定位，融合全局特征（第 1 列）后获取到的融合特征（第 3 列）进一步明确了篡改区域的位置和形状。

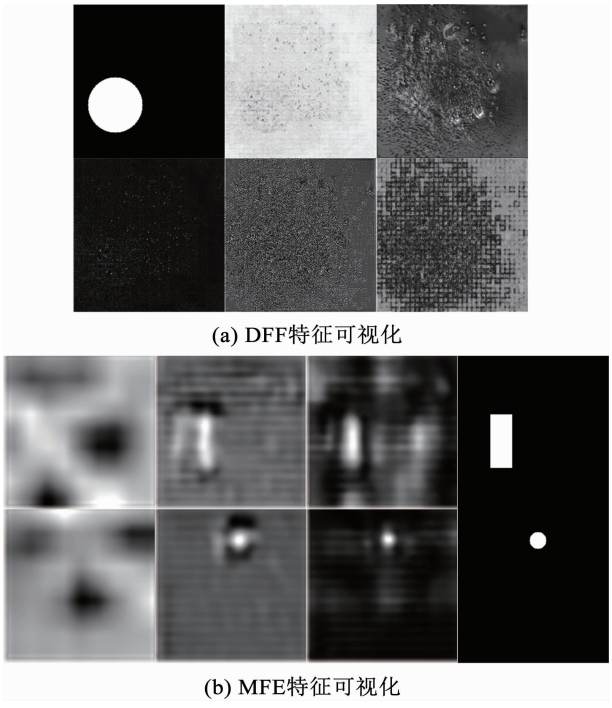
2.2　基于注意力的解码器

2.2.1　解码器主体部分

如图 1 所示，解码器主体部分由上采样模块、基本卷积块、卷积层和 Softmax 层组成。在本文中，上采样模块采用步长为 2 的转置卷积[31]将特征图逐步还原到原图大小。基本卷积块的作用是加强对篡改区域的特征表达。首先，解码器对 MFE 得到的特征图采取 3 次上采样操作，每次上采样后堆叠跳连部分得到的加权特征图，并插入基本卷积块操作。在 1/4、1/2 与 1 输入图像尺度下输出的特征图通道数分别设置为 128、64 和 32。然后，通过卷积核大小为 1 × 1 的卷积层将 32 通道的特征图映射到两通道得到 $\boldsymbol{S} \in \boldsymbol{R}^{H\times W\times 2}$。最终，利用 Softmax 分类器对其进行归一化得到输出 $\boldsymbol{Y} \in \boldsymbol{R}^{H\times W\times 2}$：

$$\boldsymbol{Y}_{kij} = \frac{\mathrm{e}^{\boldsymbol{S}_{kij}}}{\mathrm{e}^{\boldsymbol{S}_{1ij}} + \mathrm{e}^{\boldsymbol{S}_{2ij}}} \tag{4}$$

式中 $\boldsymbol{Y}_{kij}$ 和 $\boldsymbol{S}_{kij}$ 分别为 $\boldsymbol{Y}$ 和 $\boldsymbol{S}$ 第 k 维通道在坐标（i，j）处的元素值。根据 $\boldsymbol{Y}_{kij}$ 在两通道对应位置大小进行分类得到取证结果。



(a) DFF特征可视化

(b) MFE特征可视化

图 6　编码器特征图可视化

Fig. 6　Encoder feature visualization

2.2.2　基于 SWCA 的跳跃连接

对于卷积和下采样操作引入的篡改区域细节损失，编、解码器之间增加跳跃连接能够有效地补充损失的细节信息。跳跃连接的主要形式之一是将编码器与对应解码器中的特征图堆叠后进行后续处理。虽然采取这种方法达到了细节补充的效果，但是该方法平等地对待所有补充的特征，引入了许多无用信息，限制了网络的特征表达能力。

针对以上问题，本文采用跳连操作重用分辨率高的低层语义信息。同时，考虑到不同层次、不同通道的特征的重要性不同，本文在跳连部分引入了空间和通道两个维度的注意力机制，提出了基于 SWCA 的跳跃连接。如图 7 所示，该过程分为 4 个步骤：

1）特征的初步提取：通过卷积核大小为 3 × 3，步长为 1 的卷积层和 ReLU 激活函数对补充的编码器特征图 $\boldsymbol{F}_{\mathrm{e}} \in \boldsymbol{R}^{H\times W\times C}$ 进行初步的特征提取，得到

$\boldsymbol{F}_{\mathrm{p}} \in \boldsymbol{R}^{H \times W \times C}$。式中 C 为特征图的通道数。

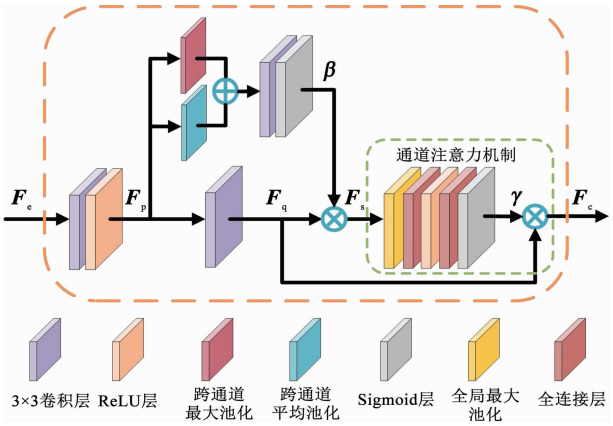


图 7 空间加权的通道注意力模块结构

Fig. 7 Architecture of spatially weighted channel attention module

2)空间位置加权:将初步提取的特征图 $\boldsymbol{F}_{\mathrm{p}}$ 利用跨通道的平均池化和跨通道的最大池化操作[32]得到两个单通道特征图,堆叠后采用卷积核大小为 3×3,步长为 1 的卷积层和 Sigmod 激活函数获取到空间权值图 $\boldsymbol{\beta} \in \boldsymbol{R}^{H \times W \times 1}$。同时,采用卷积核大小为 3×3,步长为 1 的卷积层将 $\boldsymbol{F}_{\mathrm{p}}$ 映射为 $\boldsymbol{F}_{\mathrm{q}} \in \boldsymbol{R}^{H \times W \times C}$。然后,对 $\boldsymbol{F}_{\mathrm{q}}$ 空间位置加权得到 $\boldsymbol{F}_{\mathrm{s}} = \boldsymbol{\beta} \otimes \boldsymbol{F}_{\mathrm{q}}$。其中,元素相乘过程中复制权值图与特征图保持相同维度。

3)通道维度加权:将空间加权的特征图 $\boldsymbol{F}_{\mathrm{s}} \in \boldsymbol{R}^{H \times W \times C}$ 通过文献[29]提出的通道注意力机制计算得到通道权值图 $\boldsymbol{\gamma} \in \boldsymbol{R}^{H \times W \times 1}$,对 $\boldsymbol{F}_{\mathrm{q}}$ 通道加权获取到输出特征图 $\boldsymbol{F}_{\mathrm{c}} = \boldsymbol{\gamma} \otimes \boldsymbol{F}_{\mathrm{q}}$。其中,通道注意力结构设置依次为 1 个全局最大池化层,2 个全连接层间插入 1 个 ReLU 层,1 个 Sigmod 层。

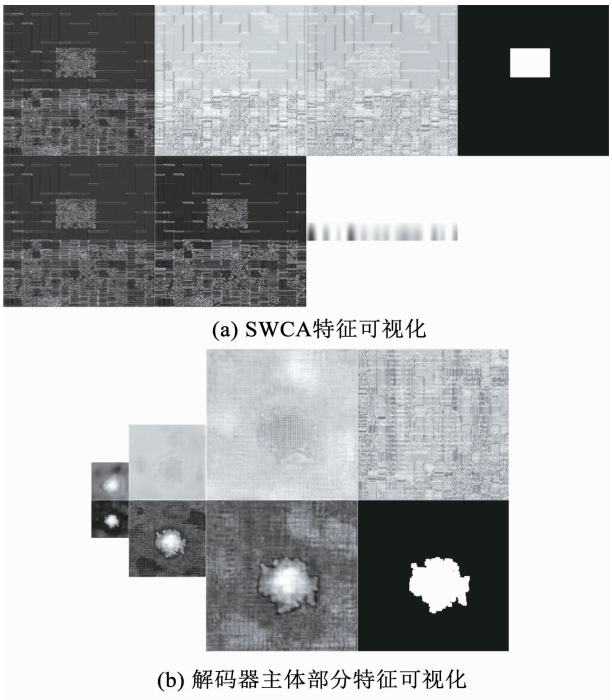
4)特征增强:对空间和通道加权后的特征图 $\boldsymbol{F}_{\mathrm{c}} \in \boldsymbol{R}^{H \times W \times C}$ 与解码器中对应尺度的特征图 $\boldsymbol{F}_{\mathrm{d}} \in \boldsymbol{R}^{H \times W \times C}$ 采用堆叠操作,得到增强后的特征图 $\boldsymbol{F}_{\mathrm{cd}} \in \boldsymbol{R}^{H \times W \times 2C}$:

$$\begin{aligned} \boldsymbol{F}_{\mathrm{cd}} &= \boldsymbol{F}_{\mathrm{c}} \oplus \boldsymbol{F}_{\mathrm{d}} = \\ &(\boldsymbol{\gamma} \otimes \boldsymbol{F}_{\mathrm{q}}) \oplus \boldsymbol{F}_{\mathrm{d}} = \\ &(\boldsymbol{\gamma} \otimes \mathrm{Conv}_{3\times3}(\boldsymbol{F}_{\mathrm{p}})) \oplus \boldsymbol{F}_{\mathrm{d}} = \\ &(\boldsymbol{\gamma} \otimes \mathrm{Conv}_{3\times3}(\delta(\mathrm{Conv}_{3\times3}(\boldsymbol{F}_{\mathrm{e}})))) \oplus \boldsymbol{F}_{\mathrm{d}} \end{aligned} \quad (5)$$

式中 $\mathrm{Conv}_{3\times3}(\cdot)$ 为卷积核大小为 3×3,步长为 1 的卷积操作。

2.2.3 解码器特征可视化

图 8(a)和(b)分别展示了解码器中 SWCA 和主体部分的特征可视化。图 8(a)以原图像尺度的基于 SWCA 的跳跃连接为例,对比真实标签(1 行 4 列),初步提取特征(1 行 1 列)能够反映大部分篡改区域的痕迹,但是存在较多空洞区域和边界模糊的情况。通过空间注意力机制的平均池化(2 行 1 列)和最大池化(2 行 2 列)突出有效信息后,得到空间加权特征(1 行 2 列)。可以看出相较于初步提取特征,空间加权特征显示的篡改区域消除了部分空洞现象,矩形篡改区域的左侧和右侧边界变得更加清晰。再利用通道注意力机制(2 行 3 列,通过放大得到)加权获取到空间加权的通道注意力特征(1 行 3 列),矩形篡改区域在保持左右侧边界清晰的基础上,下侧边界附近的空洞区域也减少。这证明 SWCA 的空间注意力和通道注意力均起到了加强特征的作用。



(a) SWCA特征可视化

(b) 解码器主体部分特征可视化

图 8 解码器特征可视化

Fig. 8 Decoder feature visualization

图 8(b)展示了解码器主体部分的特征可视化图。对比真实标签(2 行 4 列)和解码器在 3 个不同图像尺度(1/4、1/2 和 1)下的跳连前特征(1 行 1 ~ 3 列)以及对应的跳连后特征(2 行 1 ~ 3 列)可以看出,随着尺度的增大,解码器逐步解码获取到更准确的篡改区域的形状和位置。此外,在 1 图像尺度下跳连前特征(1 行 3 列)融合 SWCA 得到的空间加权的通道注意力特征(1 行 4 列)得到跳连后特征(2 行 3 列),明显看出跳连后特征比跳连前特征可以更清晰地反映篡改区域,有效增强了篡改区域的边界细节。对比其他尺度下可以得到相同的结论。因此,引入 SWCA 的跳跃连接实现了有侧重地编码器特征补充,减少了无用信息的影响,使得解码器获取到其所需要的增强信息,达到了更好的边界细节补充效果。

2.3　损失函数

由于取证问题本质上是计算机视觉中的分割问题，最常见的损失函数是标准交叉熵损失。但是一般情况下，目标图像中的篡改区域（正样本）比未篡改区域（负样本）要小得多，这将导致正负样本的不平衡。如果用标准交叉熵损失来监督网络的训练，占主导地位的负样本会贡献大部分的损失。网络模型会严重偏向负样本，导致对篡改区域分类效果较差，取证性能不佳的结果。

为了减轻类别不平衡的影响，同时减少计算量，提升模型的训练速度，本文采用加权交叉熵对输出结果的逐个像素点进行监督。损失函数可以表示为

$$E_k = \frac{-1}{W \cdot H}\sum_{j=1}^{W}\sum_{i=1}^{H}\omega_1 \boldsymbol{J}_{ij}\log(\boldsymbol{Y}_{kij}) + \omega_2(1-\boldsymbol{J}_{ij})\log(1-\boldsymbol{Y}_{kij}) \tag{6}$$

式中：ω_1 和 ω_2 分别为正样本和负样本的权重因子，$\boldsymbol{J}_{ij}$ 为 one-hot 标签 (i,j) 处的像素值，只有 1 和 0 两种取值，分别为篡改区域和非篡改区域。

3　实验结果

为了检验所提出网络的取证性能，本文利用多种深度修复方法获取实验数据集，并且将本文方法的取证结果和最先进的相关算法进行主观与客观的对比分析。同时，通过设置消融实验以证明网络不同组件的有效性。此外，将该模型应用于抗 JPEG 压缩和噪声攻击的实验以检验其鲁棒性。

3.1　图像数据集

首先，在 Place2 数据集[33]中随机选取了19 350 张 256 × 256 大小的彩色图像。然后，选择具有代表性的 3 种深度图像修复方案，分别是方案一（contextual attention，CA）[16]，方案二（globally and locally consistent，GLC）[10]和方案三（pyramid-context encoder network，PEN）[11]。利用它们训练好的网络模型分别对每幅图像进行修复篡改操作，共得到 19 350 × 3 幅修复后的图像。篡改区域的形状有圆形、矩形和不规则形状，篡改区域的面积与图像面积之比设置为 1%、5% 和 10%。每种参数情况对应的被操作图像个数是一致的，篡改位置是随机选取的。最后，随机选择 18 000 × 3 张修复篡改后的图像作为训练集，450 × 3 张图像作为验证集和 900 × 3 张图像作为测试集。图 9 为按上述规则获取到的待修复图像样本。其中，图中绿色掩膜为待修复区域。

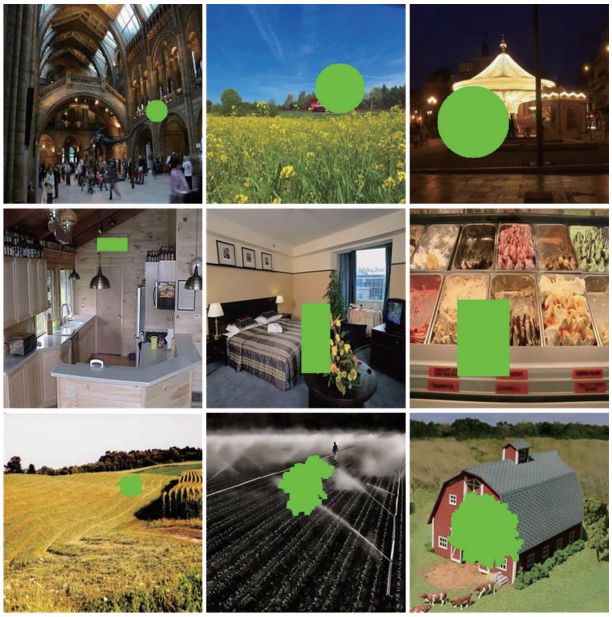


图 9　不同修复区域的图像样本

Fig. 9　Image samples of different inpainting regions

3.2　训练细节

本文所提出方法基于 PyTorch 框架实现，在 Ubuntu 环境下使用 NVIDIA 1 080 Ti GPU 完成模型的训练和测试。在网络训练时，初始学习率设置为 0.001，批尺寸设置为 8，采用动量衰减指数 β_1 = 0.9，β_2 = 0.999 的 ADAM 优化器对模型进行优化。迭代次数为 50 次，学习率每 10 次迭代衰减为之前的 0.75。权重因子 ω_1 和 ω_2 分别设置为 5 和 1。为了评估所提出方法的取证性能，本文采用现有的两种取证网络进行了对比实验，包括文献[1]提出的编、解码器结构的 CNN 和文献[18]提出的深度修复取证网络。此外，典型的语义分割算法 DeepLabv3 +[34]也被用作对比方案之一。其中，对比方案[18]采用该方案提出文献提供的训练方式，而方案[1]和 DeepLabv3 + 算法均采用和本文相同的训练方式。

本文采用以下 4 种评价指标进行取证性能评估：F1 分数，交并比（intersection over union，IoU），真阳率（true positive rate，TPR）和假阳率（false positive rate，FPR）。下文中的实验结果数据均为在测试集上测试的平均值。此外，F1 分数、IoU、TPR 和 FPR 分别表示如下：

$$\theta_{\mathrm{F1}} = \frac{2 \times |\Omega \cap \Omega_{\mathrm{d}}|}{|\Omega| + |\Omega_{\mathrm{d}}|} \tag{7}$$

$$\theta_{\mathrm{IoU}} = \frac{|\Omega \cap \Omega_{\mathrm{d}}|}{|\Omega \cup \Omega_{\mathrm{d}}|} \tag{8}$$

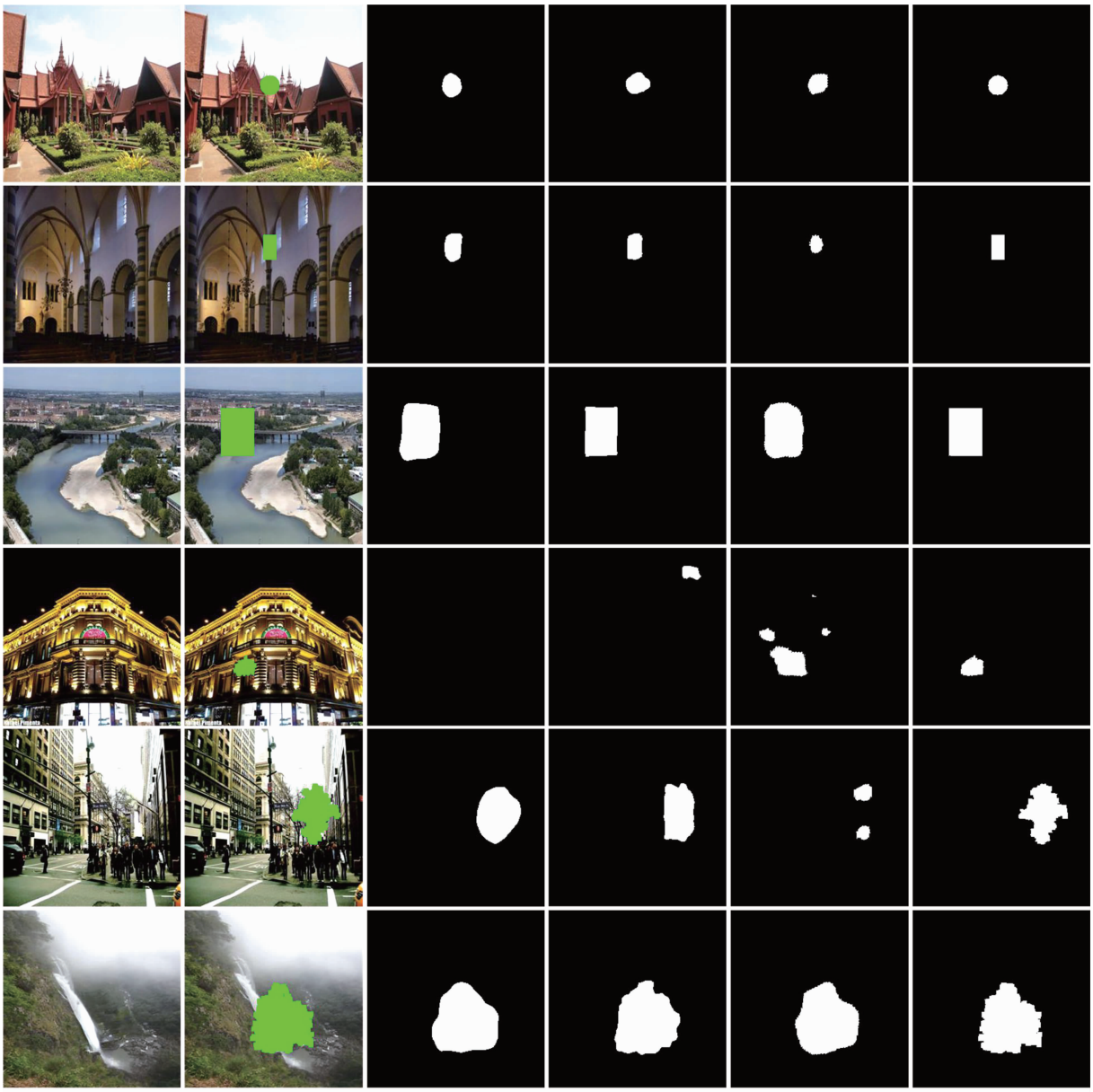
$$\theta_{\mathrm{TPR}} = \frac{|\Omega \cap \Omega_{\mathrm{d}}|}{|\Omega|} \tag{9}$$

$$\theta_{\mathrm{FPR}} = \frac{|\overline{\Omega} \cap \Omega_{\mathrm{d}}|}{|\overline{\Omega}|} \tag{10}$$

式中：$|\cdot|$为集合的势，Ω 是真实的修复区域，$\overline{\Omega}$ 是真实的未修复区域，Ω_d 是取证出的修复区域。

3.3 性能比较

图 10 展示了不同方案对深度修复方案 CA 修复篡改的 6 幅图像样本的修复取证结果。对比图像修复取证方案的定位结果(图 10(c) ~ (f))和真实篡改区域(图 10(b))可以看出，每一种方案均能够在一定程度上定位目标图像的篡改区域。对于面积较大的(如图 10 第 3 和 6 行)或者规则的(圆形和矩形)篡改区域样本(如图 10 第 1 和 2 行)，不同的方案均具有相对较好的检测性能。但是对于面积较小的不规则篡改区域的样本(如图 10 第 4 行)，即取证难度较大的一类样本。对比方案[1](图 10(c))对篡改区域产生了漏检，DeepLabv3 + (图 10(d))则是完全的误检，方案[18](图 10(e))也产生了较大范围的误检。然而，本文方法(图 10(f))仍能够相当准确地定位篡改区域。综合来说，所提出方法(图 10(f))不仅在输出结果上有极少的虚警像素，还能在位置和形状上更好地拟合真实的篡改区域。而其余对比方案(图 10(c) ~ (e))的定位结果均具有不同程度的形状失真现象。由此可见，本文所提出的取证网络能够更有效地提取修复痕迹特征，对于不同篡改区域形状和面积的变化具有很强的鲁棒性。



(a) 被修复篡改的图像　(b) 带真实篡改区域的图像　(c) 方案[1]　(d) Deeplabv3+　(e) 方案[18]　(f) 本文方法

图 10　针对 CA 的不同方案的定性比较

Fig. 10　Qualitative comparison of different methods for CA

表 1 和表 2 为各种方法的定量对比结果。以不规则篡改区域的样本为例，可以观察到，当篡改区域面积占比为 10% 时，本文所提出方法达到最高的 98.99% 的 F1 分数，98.01% 的 IoU（表 1 第11 列）以及 99.39% 的 TPR（表 2 第 11 列）。虽然随着篡改区域面积的减小，网络的取证性能略有下降，但是仍在篡改区域面积占比为 1% 时分别取得了 96.55%、93.38%（表 1 第 9 列）和 98.87%（表 2 第 9 列）的良好结果。性能下降的原因可能是更小的篡改面积导致修复操作特征不显著。同时，所提出方法的 FPR 指标也保持着极低的数值，最大值为0.16%（表 2 第 1 列），可见误警率极小，并且会随着篡改区域面积的减小得到略微改善。对于圆形或矩形篡改区域的测试图像，各项指标也得到了类似的结果。其他对比方案也具有相似的趋势变化规律，但是相较于本文方法取证性能均下降很多。以 IoU 指标为例，所提出方法在 900 张整体测试集上比方案[1]、DeepLabv3 + 和方案[18]分别高出 19.82%、15.12% 和 26.29%（表 1 第 12 列）。总体来说，对于各种形状和面积的篡改区域，本文方法在各项指标中都获得了明显优于其他对比方案的测试结果。

表 1　针对 CA 不同方案的 F1 和 IoU 比较

Tab. 1　Comparison of F1 and IoU of different methods for CA　　%

取证方案	指标	圆形			矩形			不规则形状			整体
		1	5	10	1	5	10	1	5	10	
方案[1]	F1	77.79	92.09	94.52	73.87	90.89	93.60	75.03	87.95	90.77	86.28
	IoU	67.34	85.60	89.64	63.40	83.68	88.05	63.99	79.15	83.83	78.30
Deeplabv3 +	F1	80.66	96.37	97.71	79.34	95.03	96.54	79.94	88.72	92.30	89.62
	IoU	70.43	93.08	95.61	68.97	90.63	93.57	68.80	80.17	85.75	83.00
方案[18]	F1	55.30	91.15	94.46	56.08	90.02	92.78	61.09	86.09	90.37	79.71
	IoU	46.16	84.26	89.64	46.32	82.74	87.05	50.18	76.85	83.20	71.83
DF^3 Net	F1	**98.85**	**99.49**	**99.74**	**99.39**	**99.77**	**99.81**	**96.55**	**98.70**	**98.99**	**99.03**
	IoU	**97.74**	**99.02**	**99.49**	**98.81**	**99.55**	**99.63**	**93.38**	**97.43**	**98.01**	**98.12**

注：最佳结果由粗体标出。

表 2　针对 CA 不同方案的 TPR 和 FPR 比较

Tab. 2　Comparison of TPR and FPR of different methods for CA　　%

取证方案	指标	圆形			矩形			不规则形状			整体
		1	5	10	1	5	10	1	5	10	
方案[1]	TPR	84.52	97.39	98.85	81.19	95.85	97.88	80.12	94.05	96.45	91.81
	FPR	0.22	0.71	1.16	0.26	0.78	1.26	0.23	0.99	1.67	0.81
Deeplabv3 +	TPR	84.96	97.34	98.01	81.58	95.82	96.86	83.22	91.49	92.30	91.63
	FPR	0.22	0.24	0.29	0.20	0.31	0.40	0.22	0.74	1.26	0.43
方案[18]	TPR	50.51	93.66	97.56	52.53	92.40	94.78	57.08	87.89	93.39	79.98
	FPR	0.07	0.62	1.02	0.14	0.63	0.99	0.12	0.75	1.37	0.63
DF^3 Net	TPR	**99.37**	**99.75**	**99.85**	**99.54**	**99.74**	**99.77**	**98.87**	**99.34**	**99.39**	**99.40**
	FPR	**0.02**	**0.04**	**0.04**	**0.01**	**0.02**	**0.02**	**0.05**	**0.10**	**0.16**	**0.05**

注：最佳结果由粗体标出。

特别要指出的是，对于取证难度更大的小尺寸篡改区域，所提出方法相较于对比方案具有更显著的优势。以 IoU 指标为例，本文方法对于篡改区域面积占比为 1% 的圆形、矩形和不规则样本较次优方案分别高出 27.31%、29.84% 和 24.58%（表 1 第 3、6、9 列），说明所提出方法能更好地针对困难样本。此外，根据各项指标的统计结果显示，对于圆形和矩形篡改区域样本的取证性能接近，但是对于规则篡改区域样本的取证性能优于不规则篡改区域样本。这可能是因为不规则区域的边界细节信息更丰富，而卷积和降采样操作会丢失更多细节信息导致的。

3.4　针对其他典型修复方案的取证性能测试

本文进一步在整体测试集上测试了所提出方法面对另外两种深度修复方案 GLC 和 PEN 的取证性

能,实验结果见表 3 和表 4。相较于在深度修复方案 CA 上的实验结果,本文方法对于 GLC 的测试结果是更优异的,而面对 PEN 各项性能指标则是均有所下降的。以 IoU 指标为例,在修复方法 CA、GLC 和 PEN 上分别达到 98.12%、99.21% 和 84.23%。这是由于取证网络对不同修复操作形成的篡改区域的敏感性是有差异的。但是,面对这两种修复方法,所提出方案对于 FPR、F1 分数和 IoU 指标性能也均明显优于其他对比方案。同时,FPR 也保持着较低的水平。综上所述,面对不同的深度修复方法,本文方法均优于其他对比方案,达到令人满意的取证检测性能。

表 3 针对 GLC 不同方案的性能比较

Tab. 3 Performance comparison of different methods for GLC %

取证方案	TPR	FPR	F1	IoU
方案[1]	98.96	0.74	91.37	84.38
Deeplabv3 +	98.04	0.18	96.64	93.67
方案[18]	94.23	0.08	95.92	92.42
DF^3 Net	**99.80**	**0.02**	**99.60**	**99.21**

注:最佳结果由粗体标出。

表 4 针对 PEN 不同方案的性能比较

Tab. 4 Performance comparison of different methods for PEN %

取证方案	TPR	FPR	F1	IoU
方案[1]	91.52	1.07	82.56	72.83
Deeplabv3 +	92.18	0.46	89.33	82.75
方案[18]	71.52	0.29	75.12	66.61
DF^3 Net	**95.98**	**0.54**	**90.86**	**84.23**

注:最佳结果由粗体标出。

3.5 在更多数据库上的取证性能测试

为了进一步测试网络模型在更多数据库上的取证性能,本文在 CELEBA[35]、DTD[36] 和 ImageNet[37] 数据库上随机各选取了 900 张图像,并以 PEN 修复方法为例建立了 3 个测试集。表 5 为直接使用在 Place2 数据集上训练好的模型的测试结果。可以看出,相对于模型在 Place2 数据库上取得的最佳的实验结果,各项指标除了在 DTD 数据库上有小幅度下降以外,在其他数据库上基本不变,均保持着很高的性能水平。这证明面对不同的图像数据库,本文所提出的网络仍然具有很好的取证性能。

表 5 针对 PEN 在不同数据集的测试结果

Tab. 5 Results for PEN on different datasets %

数据库	TPR	FPR	F1	IoU
Place2	95.98	0.54	90.86	84.23
CELEBA	94.76	0.55	89.75	82.79
DTD	87.89	0.61	83.96	76.29
ImageNet	95.55	0.54	90.37	83.68

3.6 鲁棒性评估

JPEG 图像压缩作为最常见的图像处理操作之一被广泛使用。为此,对根据 CA 得到原始数据集分别进行压缩因子为 95、85 和 75 的 JPEG 压缩。不同方案与前文各自的训练方式保持一致,将训练好的网络模型在篡改区域面积占比不小于 5% 的图像测试集上的测试结果见图 11。可以观察到,随着压缩因子的减小,所有方案都有着不同程度的性能下降。JPEG 压缩因子的减小代表更多的修复操作信息被移除,必然导致取证性能的下降。但是,所提出方法在不同的压缩因子下始终明显优于其他方案,说明其具有更佳的抗 JPEG 压缩性能。

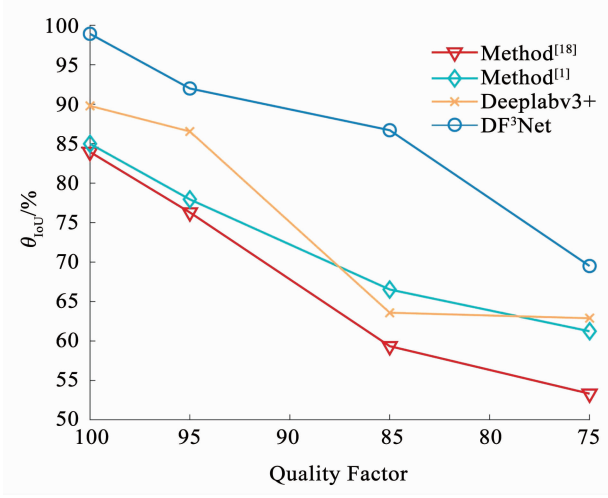


图 11 针对 CA 在不同 JPEG 压缩因子下的 IoU 指标

Fig. 11 IoU for CA with different JPEG quality factor

增加噪声作为一种修复篡改图像的常见后处理操作,常被用于对抗取证方案的检测。在 JPEG 压缩因子为 75 的条件下进行加噪测试,通过在测试集上增加信噪比分别为 50、40 dB 和 30 dB 的高斯白噪声获取到加噪测试集。表 6 为直接使用训练好的不同的网络模型的测试结果。可以看出,随着信噪比的减小即增加更多的噪声,所提出方案的 IoU 指标在加噪 50、40 dB 时基本保持不变甚至有略微上升,而在 30 dB 时不同指标均有小幅度的下降,其他方案趋势变化类似。同时,DF^3 Net 的取证性能始终明显优于其他对比方案。综上所述,证明所提出方法具有很强的抗噪声攻击的能力。

表 6 针对 CA 在不同噪声下的 IoU 指标

Tab. 6 IoU for CA with different noise

R_{SN}/dB	方案[1]/%	Deeplabv3 +/%	方案[18]/%	DF^3 Net/%
	61.57	62.88	52.41	**68.56**
50	61.76	62.90	51.42	**68.57**
40	61.81	62.96	51.73	**68.48**
30	56.58	61.35	50.12	**64.85**

注:最佳结果由粗体标出。

3.7 消融实验

本文设计了如下两组消融实验对所提出模型各个组件的有效性进行验证：一组实验设置为在本文完整网络模型（full model，FM）上分别去除 RGB 信息（removing RGB information，RR）、SRM 滤波（removing SRM filtering，RS）、空间域高通滤波（removing high-pass filtering of spatial domain，RHPS）、频域高通滤波（removing high-pass filtering of frequency domain，RHPF）以及利用普通卷积代替动态卷积（convolution instead of dynamic convolution，CIDC）。表 7 为在修复方案 CA 上得到的 F1 分数和 IoU 结果，相较于完整网络 FM，去除 DFF 中的任何一个输入以及普通卷积替换动态卷积，都会引起指标 F1 和 IoU 不同程度的下降。其中下降最明显的是去除频域高通滤波的网络 RHPF，F1 和 IoU 指标分别降低了 1.71% 和 3.22%。这说明，通过采用多输入特征和动态特征融合模块，不同的特征在相应的领域均发挥了积极作用，提取到更有效的修复痕迹特征，从而大大提升了取证性能。

表 7　消融实验（针对 CA）：组件有效性研究一

Tab.7　Ablation experiments（CA）：component validity study Ⅰ %

指标	FM	RR	RS	RHPS	RHPF	DIDC
F1	99.03	98.50	98.55	98.79	97.32	98.81
IoU	98.12	97.10	97.18	97.66	94.90	97.69

另一组实验设置为：仅带有 DFF 的网络模型（model with DFF，MD）、带有 DFF 和 MFE 的网络模型（model with DFF and MFE，MDM）和带有 DFF 和引入 SWCA 的跳跃连接的网络模型（model with DFF and skip connection introduced SWCA，MDS）。表 8 为面对修复方案 CA 的取证结果，相较于模型 MD，引入 MFE 的模型 MDM 和增加基于 SWCA 的跳跃连接的 MDS 在 IoU 和 F1 分数均有着不同程度的改进，综合使用两者的完整模型则有着更大的提升。以 IoU 为例，MDM、MDS 和 FM 较 MD 分别提升了 0.52%、2.17% 和 2.74%。因此，MFE 和引入 SWCA 的跳跃连接均能够有效提升取证网络的性能。

表 8　消融实验（针对 CA）：组件有效性研究二

Tab.8　Ablation experiments（CA）：component validity study Ⅱ %

指标	MD	MDM	MDS	FM
F1	97.57	97.89	98.71	99.03
IoU	95.38	95.90	97.55	98.12

4　结　论

本文提出了一种基于动态特征融合的深度修复取证网络 DF^{3}Net。该网络针对深度修复遗留的细微特征扩展输入后，利用 DFF 使得多种有效的痕迹特征得以被充分地获取到。此外，为进一步提升取证性能，编码器通过增加 MFE 以更好地获取上下文信息，解码器引入基于 SWCA 的跳跃连接以实现细节信息的针对性补充。实验结果表明，本文提出的取证网络对比现有方法均取得较优性能，同时面对 JPEG 压缩和噪声攻击具有较强的鲁棒性。

参考文献

[1] ZHU Xinshan，QIAN Yongjun，ZHAO Xianfeng，et al. A deep learning approach to patch-based image inpainting forensics[J]. Signal Processing：Image Communication，2018(67)：90. DOI：10.1016/j.image.2018.05.015

[2] LIANG Zaoshan，YANG Gaobo，DING Xiangling，et al. An efficient forgery detection algorithm for object removal by exemplar-based image inpainting[J]. Journal of Visual Communication and Image Representation，2015，30：75. DOI：10.1016/j.jvcir.2015.03.004

[3] ZHAO Yuqian，LIAO Miao，SHIH F Y，et al. Tampered region detection of inpainting JPEG images[J]. Optik，2013，124(16)：2487. DOI：10.1016/j.ijleo.2012.08.018

[4] LIU Qingzhong，SUNG A H，ZHOU Bing，et al. Exposing inpainting forgery in jpeg images under recompression attacks[C]//Proceedings of the 2016 15th IEEE International Conference on Machine Learning and Applications. Piscataway：IEEE，2016：164

[5] 曹承瑞，刘微容，史长宏，等. 多级注意力传播驱动的生成式图像修复方法[J]. 自动化学报，2021(45)：1
CAO Chengrui，LIU Weirong，SHI Changhong，et al. Generative image inpainting with attention propagation[J]. Acta Automatica Sinica，2021(45)：1. DOI：10.16383/j.aas.c200485

[6] BERTALMIO M，SAPIRO G，CASELLES V，et al. Image inpainting[C]//Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques. New York：ACM，2000：417. DOI：10.1145/344779.344972

[7] ELAD M，STARCK J L，QUERRE P，et al. Simultaneous cartoon and texture image inpainting using morphological component analysis[J]. Applied and Computational Harmonic Analysis，2005，19(3)：340. DOI：10.1016/j.acha.2005.03.005

[8] 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. DOI：10.1145/1531326.1531330

[9] CRIMINISI A，PÉREZ P，TOYAMA K. Region filling and object removal by exemplar-based image inpainting[J]. IEEE Transactions on Image Processing，2004，13(9)：1200. DOI：10.1109/TIP.2004.833105

[10] IIZUKA S，SIMO-SERRA E，ISHIKAWA H. Globally and locally consistent image completion[J]. ACM Transactions on Graphics，2017，36(4)：1. DOI：10.1145/3072959.3073659

[11] ZENG Yanhong，FU Jianlong，CHAO Hongyang，et al. Learning pyramid-context encoder network for high-quality image inpainting[C]//Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. [S.l.]：IEEE，2019：1486. DOI：10.1109/CVPR.2019.00158

[12] LI Haodong，LUO Weiqi，HUANG Jiwu. Localization of diffusion-

based inpainting in digital images [J]. IEEE Transactions on Information Forensics and Security, 2017, 12(12): 3050. DOI: 10.1109/TIFS.2017.2730822

[13] WU Qiong, SUN Shaojie, ZHU Wei, et al. Detection of digital doctoring in exemplar-based inpainted images [C]//Proceedings of the 2008 International Conference on Machine Learning and Cybernetics. Piscataway: IEEE, 2008: 1222. DOI: 10.1109/ICMLC.2008.4620591

[14] BACCHUWAR K S, RAMAKRISHNAN K R. A jump patch-block match algorithm for multiple forgery detection [C]//Proceedings of the 2013 International Mutli-Conference on Automation, Computing, Communication, Control and Compressed Sensing. Piscataway: IEEE, 2013: 723. DOI:10.1109/iMac4s.2013.6526502

[15] CHANG I C, YU J C, CHANG C C. A forgery detection algorithm for exemplar-based inpainting images using multi-region relation [J]. Image and Vision Computing, 2013, 31(1): 57. DOI:10.1016/j.imavis.2012.09.002

[16] YU Jiahui, LIN Zhe, YANG Jimei, et al. Generative image inpainting with contextual attention [C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 5505. DOI: 10.1109/CVPR.2018.00577

[17] ZEILER M D, FERGUS R. Visualizing and understanding convolutional networks [C]//Proceedings of the 13th European Conference on Computer Vision. Cham: Springer, 2014: 818. DOI:10.1007/978-3-319-10590-1_53

[18] LI Haodong, HUANG Jiwu. Localization of deep inpainting using high-pass fully convolutional network [C]//Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2019: 8301. DOI:10.1109/ICCV.2019.00839

[19] CHEN Yinpeng, DAI Xiyang, LIU Mengchen, et al. Dynamic convolution: attention over convolution kernels [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2020: 11030. DOI: 10.1109/CVPR42600.2020.01104

[20] FRIDRICH J, KODOVSKY J. Rich models for steganalysis of digital images [J]. IEEE Transactions on Information Forensics and Security, 2012, 7(3): 868. DOI:10.1109/TIFS.2012.2190402

[21] RAO Yuan, NI Jiangqun. A deep learning approach to detection of splicing and copy-move forgeries in images [C]//Proceedings of the 2016 IEEE International Workshop on Information Forensics and Security. Piscataway: IEEE, 2016: 1. DOI:10.1109/WIFS.2016.7823911

[22] ZHOU Peng, HAN Xintong, MORARIU V I, et al. Learning rich features for image manipulation detection [C]//Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 1053. DOI: 10.1109/CVPR.2018.00116

[23] RÖSSLER A, COZZOLINO D, VERDOLIVA L, et al. FaceForensics++: learning to detect manipulated facial images [C]//Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2019: 1. DOI:10.1109/ICCV.2019.00009

[24] BAYAR B, STAMM M C. Constrained convolutional neural networks: a new approach towards general purpose image manipulation detection [J]. IEEE Transactions on Information Forensics and Security, 2018, 13(11): 2691. DOI:10.1109/TIFS.2018.2825953

[25] CHEN Shen, YAO Taiping, CHEN Yang, et al. Local relation learning for face forgery detection [Z]. arXiv:2105.02577, 2021

[26] BI X, LIU Y, XIAO B, et al. D-Unet: a dual-encoder U-Net for image splicing forgery detection and localization [Z]. arXiv:2012.01821, 2020. DOI:10.48550/arXiv.2012.01821

[27] BREIMAN L. Bagging predictors [J]. Machine Learning, 1996, 24(2): 123. DOI:10.1023/A:1018054314350

[28] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift [C]// Proceedings of the 32nd International Conference on Machine Learning. Lille: JMLR, 2015: 448. DOI: 10.5555/3045118.3045167

[29] HU Jie, SHEN Li, ALBANIE S, et al. Squeeze-and-excitation networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 7132. DOI:10.1109/TPAMI.2019.2913372

[30] DUTA I C, LIU Li, ZHU Fan, et al. Pyramidal convolution: rethinking convolutional neural networks for visual recognition [Z]. arXiv:2006.11538, 2020. DOI:10.48550/arXiv.2006.11538

[31] ZEILER M D, KRISHNAN D, TAYLOR G W, et al. Deconvolutional networks [C]//Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2010: 2528. DOI: 10.1109/cvpr.2010.5539957

[32] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [C]//Proceedings of the 15th European Conference on Computer Vision. Cham: Springer, 2018: 3. DOI:10.1007/978-3-030-01234-2_1

[33] 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. DOI:10.1109/TPAMI.2017.2723009

[34] CHEN L C, ZHU Yukun, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]//Proceedings of the 15th European Conference on Computer Vision. Cham: Springer, 2018: 801. DOI:10.1007/978-3-030-01234-2_49

[35] LIU Ziwei, LUO Ping, WANG Xiaogang, et al. Deep learning face attributes in the wild [C]//Proceedings of the 2015 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2015: 3730. DOI:10.1109/ICCV.2015.425

[36] CIMPOI M, MAJI S, KOKKINOS I, et al. Describing textures in the wild [C]//Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2014: 3606. DOI:10.1109/CVPR.2014.461

[37] DENG Jie, DONG Wei, SOCHER R, et al. ImageNet: a large-scale hierarchical image database [C]//Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2009: 248. DOI: 10.1109/CVPR.2009.5206848

(编辑 苗秀芝)