

文章编号: 2095-2163(2022)08-0145-04

中图分类号: TP391.41

文献标志码: A

# 融合注意力与 FCN 的图像修复取证

沈万里, 张玉金, 王永琦, 胡 万, 孙 冉, 郭 静

(上海工程技术大学 电子电气工程学院, 上海 201620)

**摘要:** 图像修复检测是数字图像取证领域中一个非常具有挑战性的问题, 为了使图像在被修复后依然保持视觉上的光滑性, 常用基于样本块的修复方法来填充被篡改部分。本文提出一种融合注意力机制与全卷积神经网络(Full Convolution Network, FCN)的图像修复取证方法, 首先对修复图像进行深度特征提取, 再对最终的特征图进行权值分配, 以获取最有效的特征图, 提高检测效果。为了验证本文方法的可行性, 创建了一个修复图像数据库, 并通过不同性能指标来评估所提出取证网络的性能。

**关键词:** 数字图像取证; 注意力机制; 全卷积神经网络; 图像修复取证

## Images inpainting forensics fused with attention mechanism and FCN

SHEN Wanli, ZHANG Yujin, WANG Yongqi, HU Wan, SUN Ran, GUO Jing

(School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China)

**[Abstract]** Image inpainting forensics is a challenging problem in the field of digital image forensics. In order to maintain the visual smoothness of the image after inpainting, the inpainting method based on sample blocks is often used to fill the tampered part. This paper proposes an image inpainting forensics method that combines the attention mechanism and Full Convolutional Neural Network (FCN). First, deep feature extraction is performed on the inpainting image, and then the final feature map is weighted to obtain the most effective feature maps and improve the detection effect. In order to verify the feasibility of the method in this paper, a repaired image database is created, and the performance of the proposed forensics network is evaluated through different performance indicators.

**[Key words]** digital images forensics; attention mechanism; Full Convolutional Neural Network; images inpainting forensics

## 0 引言

图像修复可以用来填充图像内容、修复受损区域。基于区域填充的方法主要分为2类: 基于扩散的修复方法和基于样本块的修复方法<sup>[1]</sup>。其中, 基于扩散的方法主要集中在小区域修复, 例如旧照片的划痕等, 不会留下易察觉的伪影, 但该方法在修复大面积区域时效果不佳; 基于样本块的修复方法类似于图像篡改中的复制粘贴操作, 是从图像未受损区域复制图像块补丁, 填充在受损区域, 达到修复的效果。

随着图像修复技术的发展, 修复后的图像留下的可感知伪影越来越少, 无形中也增加了图像修复取证工作的难度。Wu 等人<sup>[2]</sup>提出了一种基于零连通特征和模糊隶属度的盲检测方法, 但该方法需要

人工选择可疑区域, 极大地增加了人工成本。Chang 等人<sup>[3]</sup>提出了一种可以加速搜索像素块的二阶搜索算法, 克服了上述缺点。在前文论述基础上, 本文提出了一种融合注意力机制与 FCN 的图像修复取证网络, 使用中心像素映射进行块搜索, 并结合通道注意力模块(Squeeze and Excitation Block, SE)对最终提取的特征图进行像素权重判定, 筛选出最有效的预测图。

## 1 融合 SE 与 FCN 的图像修复取证网络

基于样本块的图像修复主要依赖于图像的内容<sup>[4]</sup>, 而卷积神经网络针对图像内容具有较好的学习能力和特征提取能力, 因此, 本文设计了一种融合 SE 模块与 FCN 的图像修复取证网络。为避免网络受图像内容影响导致的检测精度降低的情况发生, 该全卷

**基金项目:** 上海市自然科学基金(17ZR1411900)。

**作者简介:** 沈万里(1997-), 男, 硕士研究生, 主要研究方向: 图像处理、深度学习和图像修复取证; 张玉金(1982-), 男, 博士, 副教授, 主要研究方向: 多媒体内容安全、图像处理与模式识别; 王永琦(1977-), 女, 硕士, 副教授, 主要研究方向: 多媒体内容安全; 胡 万(1996-), 男, 硕士研究生, 主要研究方向: 图像滤波取证; 孙 冉(1996-), 男, 硕士研究生, 主要研究方向: 图像处理; 郭 静(1996-), 女, 硕士研究生, 主要研究方向: 图像处理、计算机视觉。

**通讯作者:** 张玉金 Email: 850262106@qq.com

**收稿日期:** 2021-11-04

积神经网络的输出结果与输入图像具有相同尺寸,同时为输出图像中的每一个像素指定一个类别标签(0或者1,0代表未修复区域,1代表修复区域)。

1.1 SE 模块

Hu 等人<sup>[5]</sup>提出一种通道注意力结构 SE-net (Squeeze-and-Excitation Net),通过神经网络来自适应地学习每个特征通道的重要程度,并为每个通道赋予不同的权重系数,从而强化重要的特征,抑制非重要的特征。通过权值重分配的方式来自适应地调整通道间的特征重要性,让计算资源分配给信息中最有用的部分,又因为其额外所占用的计算成本小,很容易嵌入其他深层网络,SE 模块结构如图 1 所示。

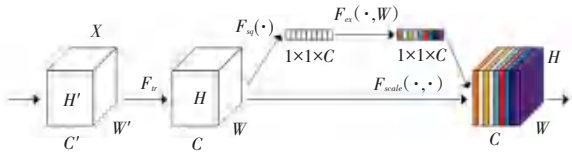


图 1 SE 模块结构图

Fig. 1 SE block structure diagram

为不同的特征通道求得对应权重是关键点,首先是 Squeeze 操作,将一个通道中整个空间特征编码为一个全局特征,通过全局池化将特征图在  $H \times W$  的空间维度上收缩,对此可表示为:

$$Z_c = F_{sq}(x_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x_c(i, j) \quad (1)$$

其中,  $F_{sq}(\cdot)$  表示全局平均池化,  $x_c$  表示特征图  $x$  的第  $c$  个通道。

得到了全局描述特征后,进行 Excitation 操作来抓取特征通道之间的关系,对此可表示为:

$$s = F_{ex}(z, w) = \sigma(\delta(F_{sq}(x_c))) \quad (2)$$

其中,  $\sigma(\cdot)$  表示 sigmoid 函数,确保  $s$  值位于  $0 \sim 1$ ,  $\delta(\cdot)$  表示 ReLU 函数。

依前处理后,下面就是 Scale 操作,即将 Excitation 操作得到的权重看作是重要信息的依据,数学定义式可写为:

$$\hat{x} = F_{scale}(s, x_c) = s \cdot x_c \quad (3)$$

其中,  $s$  作为比例因子与原通道相乘,利用一维的稀疏卷积操作来优化 SE 模块中涉及到的全连接层操作,达到大幅降低参数量并保持相当性能的目的。为了压缩参数量和计算效率,SE 模块采用“先降维-再升维”的策略,利用 2 个多层感知机来学习不同通道间的相关性,即当前的每一个特征图都与其它特征图进行交互,是一种密集型的连接。

1.2 网络结构设计

为实现对取证区域的定位,本文采用基于全卷积神经网络(FCN)的图像修复取证网络架构,总体网络结构如图 2 所示。该架构的特征提取模块由 14 个卷积层和 5 个最大池化层组成,每一个卷积层后都添加了整流线性单位(ReLus)作为激活函数。在特征提取模块的第一层,本文采用了 64 个  $3 \times 3$  的卷积核,将第一层的输出再次经过一个相同结构的卷积模块,输出  $224 \times 224 \times 64$  的特征图;使用窗口大小为  $2 \times 2$  和步长为 2 的最大池化层对第二层卷积输出特征图进行处理,得到  $112 \times 112 \times 128$  的特征图。经过 5 次不同核数的卷积层及 5 次相同窗口大小和步长的最大池化层,输出大小为  $7 \times 7$  的 4 096 个特征图。本文将第四个区块的特征图通过  $1 \times 1$  的卷积进行降维操作,同时对第五和第六个区块的特征图进行反卷积操作,分别获得  $28 \times 28 \times 2$  的特征图,再通过拼接融合方式,得到  $28 \times 28 \times 6$  的特征图。对这 6 个通道的特征图,使用 SE 模块进行权重分配,得出权重占比最大的两张特征图,分别代表修复和未修复样本的特征图。

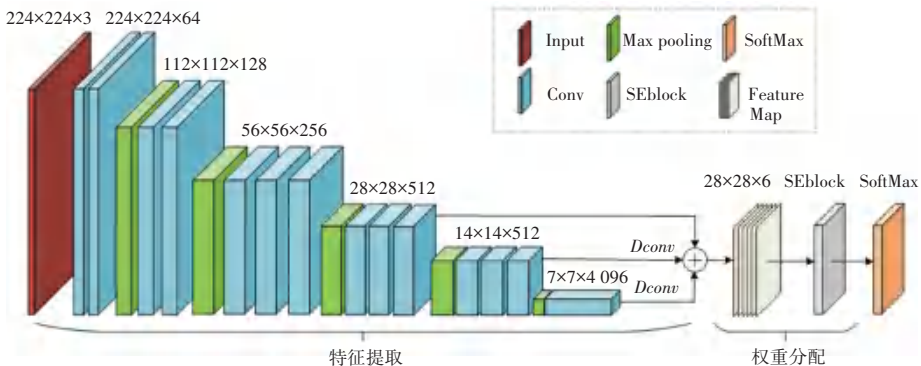


图 2 总体网络结构图

Fig. 2 Overall network structure diagram

## 2 实验

### 2.1 实验数据库的建立

本文创建了包含 2 000 个大小为  $256 \times 256$  彩色图像的修复图像数据库。该数据库中的修复区域为不规则形状,修复区域占比为 2%~40%。实验随机选择 1 600 张用于训练,400 张用于测试。本文提出的模型输入图像大小为  $224 \times 224$ ,因此先将数据库图像大小裁剪为  $224 \times 224$  作为本文使用的修复图像数据库。考虑到在训练样本中,修复区域相对于未修复区域较小,研究中采用加权交叉熵作为损失函数,用随机梯度法来迭代更新超参数。学习率设置为  $e^{-2}$ ,动量值设置为 0.9,重量衰减值设置为  $5^{-4}$ ,*batch-size* 设置为 8。

### 2.2 实验结果与分析

本文针对彩色修复图像进行了修复检测和定位,具体可视化结果如图 3 所示。由图 3 可以看出,即使不规则的修复区域面积较大,本文算法仍然能够较为准确地定位修复区域。

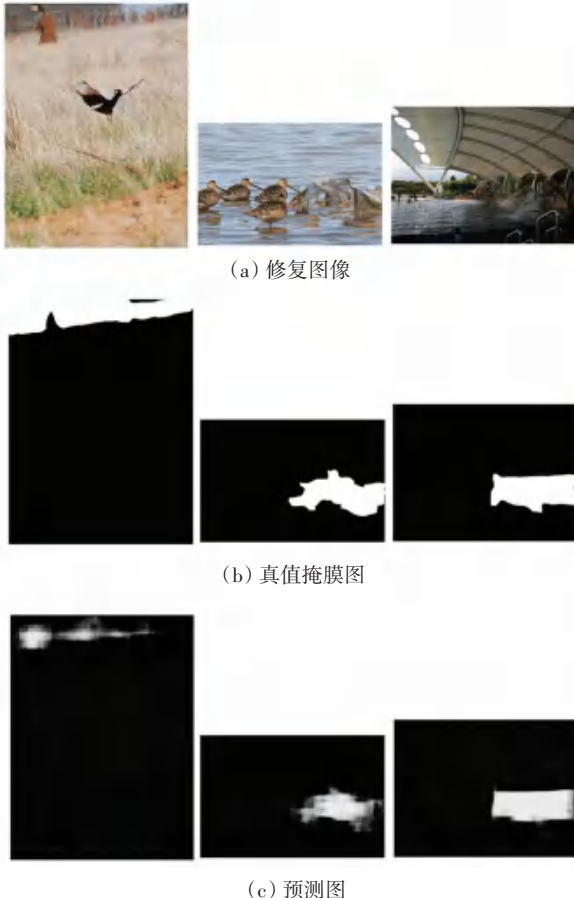


图 3 修复图像的可视化结果

Fig. 3 Visualized results of inpainting images

通过精度、召回率、 $F_1$  得分和交并比 (*IoU*) 等数据定量分析本文提出的修复图像取证网络,这些参数可由真阳样本 (*Ture Positive*, *TP*)、假阳样本 (*False Positive*, *FP*) 和假阴样本 (*False Negative*, *FN*) 计算得到。对此拟展开研究分述如下。

(1) 精度 (*P*): 定义为 *TP* 个数与 *TP* 和 *FP* 个数之和的比值,该值可由下式计算得出:

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

(2) 召回率 (*R*): 定义为 *TP* 数量与 *TP* 和 *FN* 数量之和的比值,该值可由下式计算得出:

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

(3)  $F_1$  Score: 定义为考虑 *P* 和 *R* 的调和平均值。该值可由下式计算得出:

$$F_1 = 2 \cdot \frac{Precision \times Recall}{Precision + Recall} \quad (6)$$

(4) 交并比 (*MIoU*): 是一个度量,用于计算预测输出和目标掩码之间的重叠百分比。计算方法为预测掩模和目标掩模像素数与 2 个掩模像素数并集的交集。

本文采用 2 个标签作为模型的输出;第一个标签表示被修复区域,第二个标签表示未被修复区域。由于有 2 个不同的输出标签,因此本文计算了测试结果的平均精度 (*AP*)、平均召回率 (*AR*)、平均  $F_1$  分数 (*AF<sub>1</sub> Score*) 和平均 *IoU*。本文方法在修复图像数据库上运行后的性能分析结果如图 4 所示。由图 4 可见,测试结果 *AP*、*AR*、*AF<sub>1</sub>*、*MIoU* 分别为 78.9%、71.15%、74.83% 和 70.3%,表明本文提出的图像修复取证模型具有较好的检测性能。

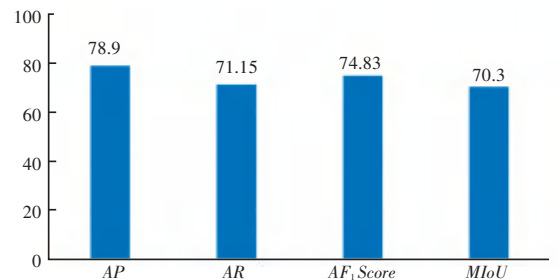


图 4 基于精度、召回率、 $F_1$  得分和交并比四个度量数据的性能分析  
Fig. 4 Performance analysis based on average precision (*AP*), average recall (*AR*), average  $F_1$  Score (*AF<sub>1</sub> Score*) and mean intersection over union (*MIoU*)

(下转第 151 页)