

# 多尺度感知与边界引导的图像篡改检测方法

陈海鹏, 张世博, 吕颖达

(吉林大学 计算机科学与技术学院, 长春 130012)

**摘要:** 针对传统图像篡改检测方法存在边界模糊、提取特征尺度单一、忽略背景信息等问题, 本文提出多尺度感知与边界引导的图像篡改检测方法。首先, 使用改进的金字塔视觉变压器提取篡改图像的空间细节和基础特征。其次, 通过边缘感知模块探索与伪造区域边缘相关的信息, 生成边缘预测图。再次, 利用边缘引导模块突出所提取特征中的关键通道, 减少冗余通道的干扰。然后, 通过多尺度上下文感知模块, 从多个感受野学习伪造区域丰富的上下文信息。最后, 利用特征融合模块交替关注篡改图像的前景和背景, 精确分割伪造区域。将本文方法在 5 个常用的公开图像篡改检测数据集上进行定量和定性对比, 实验结果表明, 本文方法可以有效检测伪造区域, 并且优于其他方法。

**关键词:** 计算机应用; 图像篡改检测; 多尺度上下文感知; 边界引导

**中图分类号:** TP391 **文献标志码:** A **文章编号:** 1671-5497(2025)06-2114-08

**DOI:** 10.13229/j.cnki.jdxbgxb.20231027

## Multi-scale context-aware and boundary-guided image manipulation detection method

CHEN Hai-peng, ZHANG Shi-bo, LYU Ying-da

(College of Computer Science and Technology, Jilin University, Changchun 130012, China)

**Abstract:** Aiming at the problems of traditional image manipulation detection methods, such as fuzzy boundaries, single scale of extracted features, and ignoring background information, this paper proposes an image manipulation detection method based on multi-scale context-aware and boundary-guided. First, spatial details and base features of manipulated images are extracted using an improved pyramid vision transformer. Second, information related to the edge of the falsified region is explored by an edge context-aware module to generate an edge prediction map. Again, the edge guidance module is utilized to highlight the key channels in the extracted features and reduce the interference of redundant channels. Then, the rich contextual information of the manipulated region is learned from multiple sensory fields through the multi-scale context-aware module. Finally, the feature fusion module is utilized to accurately segment the manipulated region by focusing alternately on the foreground and background of the manipulated images. Comparing this paper's method quantitatively and qualitatively on five commonly used public image

**收稿日期:** 2023-09-25.

**基金项目:** 国家自然科学基金面上项目(62276112); 国家自然科学基金区域联合基金子项目(U19A2057); 吉林省科技发展计划重点研发项目(20230201088GX); 安徽高校协同创新项目(GXXT-2022-044).

**作者简介:** 陈海鹏(1978-), 男, 教授, 博士. 研究方向: 图像处理与模式识别. E-mail: chenhp@jlu.edu.cn

**通信作者:** 吕颖达(1983-), 女, 副教授, 博士. 研究方向: 图像处理与模式识别. E-mail: ydlv@jlu.edu.cn

manipulation detection datasets, the experimental results show that this paper's method can effectively detect manipulated regions and outperforms other methods.

**Key words:** computer application; image manipulation detection; multi-scale context-aware; boundary guidance

## 0 引言

图像篡改检测是多媒体取证和计算机视觉领域重要的研究课题之一<sup>[1,2]</sup>。近年来,随着数字图像处理技术的进步,图像处理已经成为一种常见的社交活动<sup>[3]</sup>。例如,一些人将处理后的图片分享到网络上,丰富了日常生活。然而,一些人利用图像处理技术进行保险欺诈、制造虚假新闻甚至散播政治谣言,对社会稳定和国家安全造成一定的负面影响<sup>[4]</sup>。因此,开发一种可靠的模型检测篡改图像中的伪造区域是十分重要的。

为了降低篡改图像带来的危害,研究人员提出了不同的图像篡改检测方法。受到传统检测方法的启发,早期的一些方法主要是通过构造手工特征捕获异常信息。例如,噪声不一致性<sup>[5]</sup>、基于离散余弦变换系数的量化效应<sup>[6]</sup>、彩色滤波阵列<sup>[7]</sup>等。由于这些方法需要大量的先验知识,很容易被有针对性地屏蔽。因此,如何自适应地提取篡改图像特征是提升模型泛化性和鲁棒性的关键。

近年来,一些基于深度学习的方法在图像篡改检测任务中彰显出巨大潜力,极大提高了模型的泛化性和鲁棒性。例如,文献[8]提出一种包括边缘预测、细化和分割的三阶段网络,实现伪造区域检测。文献[9]通过建模物体层面的视觉不一致性信息,为图像篡改检测提供了新的思路。文献[10]使用边缘特征增强模块和边缘监督策略融合篡改痕迹,进一步提高了检测性能。尽管这些方法取得了显著的成绩,但是仍然存在一些问题:①边界模糊问题。现有的方法大都忽略了边缘信息的重要性,即使实现了边界定位,也没有细化边界特征,从而在模糊的边界下提供了不理想的预测结果;②提取特征尺度单一。先前的方法仅从单一尺度进行特征学习,忽略了一些重要的上下文信息,在面对形状不规则和尺度变化较大的伪造区域时,通常无法分割完整的伪造区域;③忽略背景信息。由于伪造区域常被归类到前景中,以往的方法只关注前景信息,而忽略了背景信息,中

断了前景信息和背景信息之间的联系,在遇到前景和背景高度相似的篡改图像时,检测精度降低。

为解决上述问题,本文提出了多尺度感知与边界引导的编解码网络(Multi-scale context-aware and boundary guidance network, MCBG-Net),专注于从输入图像中学习多尺度上下文感知特征和边缘特征,以实现精准分割伪造区域。MCBG-Net的主要贡献如下:①设计了一个简单有效的边缘感知模块EAM,通过集成低层次的局部边缘信息和高层次的全局位置信息对边缘信息进行建模;②采用边缘引导模块EFM,探索关键通道减少噪声干扰;③利用多尺度上下文感知模块MCAM,从多个感受野学习伪造区域丰富的上下文信息;④设计了特征融合模块FFM,交替关注篡改图像的前景和背景,精准分割伪造区域;⑤在5个常用的公开数据集上的实验结果表明,该方法优于现有的其他方法。

## 1 本文方法

### 1.1 整体网络结构

本文方法的整体结构由5个主要部分组成:骨干网络、边缘感知模块、边缘引导模块、多尺度上下文感知模块和特征融合模块,其结构如图1所示。首先,骨干网络采用ImageNet数据集上预训练的PVTv2<sup>[11]</sup>网络模型,逐层提取篡改图像的空间细节和基础特征。其次,利用边缘感知模块探索伪造区域的边缘轮廓。再次,利用边缘引导模块突出关键通道,减少冗余通道的干扰。然后,通过多尺度上下文感知模块学习伪造区域丰富的上下文信息,以提高特征学习的表征能力。最后,利用特征融合模块逐级重构最终的检测结果,并采用多级监督的方式进行网络训练。

### 1.2 边缘感知模块

良好的边缘先验知识有助于识别伪造区域的边缘轮廓。虽然低级特征包含了丰富的边缘特征,但是也引入了许多无关信息。因此,需要高级语义信息促进对伪造区域边缘的检测。为此,本文提出了边缘感知模块EAM,结合浅层局部边缘

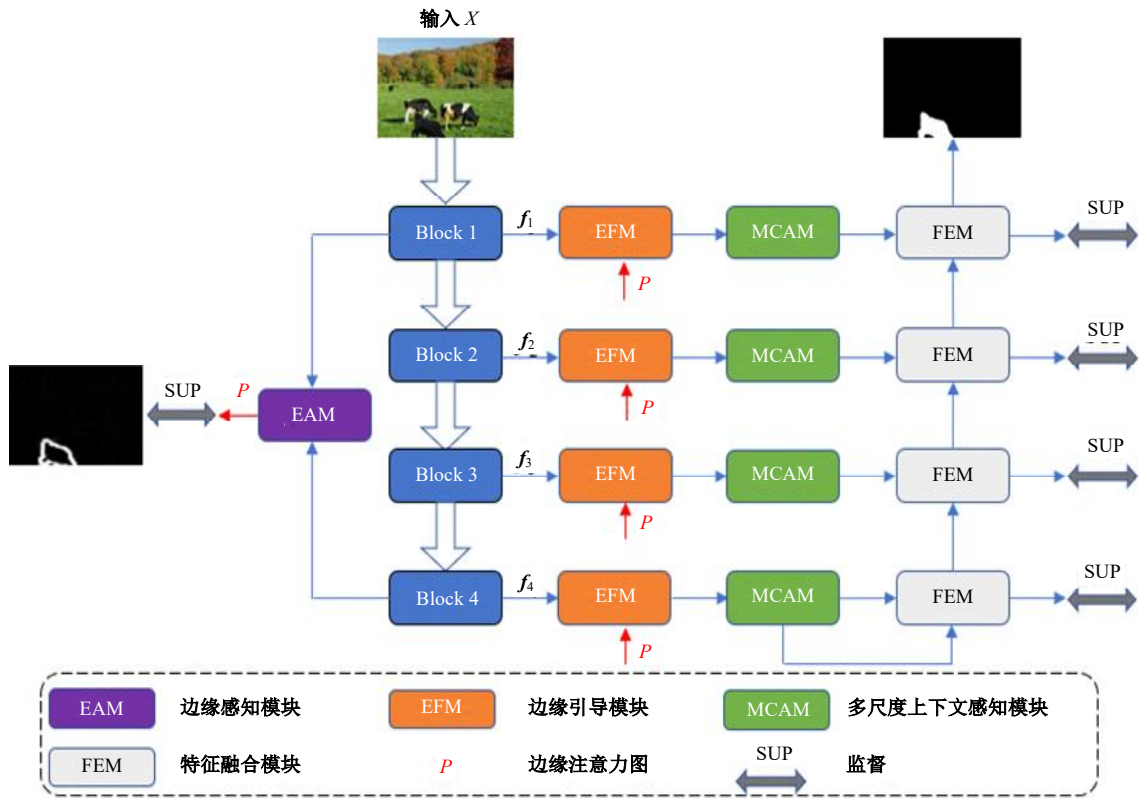


图 1 本文提出的网络整体结构图

Fig. 1 Overview of proposed network structure

信息和深层全局位置信息指导边缘特征学习,具体结构如图 2 所示。

将低级特征  $f_1$  和高级特征  $f_4$  作为此模块的输入建模与伪造区域相关的边缘信息。首先采用  $1 \times 1$  的卷积操作,将高级特征  $f_4$  的通道数更改为 256,并通过双线性插值法进行上采样,以增大特征图的尺寸。其次,沿着通道维度将两个分支的

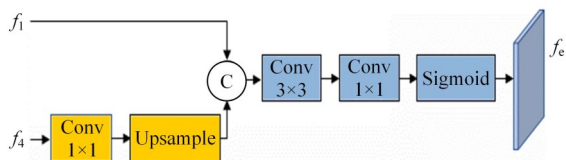


图 2 边缘感知模块结构图

Fig. 2 Architecture of the edge-aware module

特征拼接。最后,通过  $3 \times 3$  卷积操作、 $1 \times 1$  卷积操作和 Sigmoid 激活函数生成边缘注意力图  $f_e$ 。

### 1.3 边缘引导模块

通常情况下,会默认每个通道在后续任务中起到相同的作用,但在实际情况中,某些通道包含的语义信息意义并不大。为此,本文提出边缘引导模块 EFM,采用通道注意力机制生成表示各通道权重的注意力图,突出关键通道信息,以增强特征表示,具体结构如图 3 所示。

将 PVTv2 提取的基础特征  $f_i (i \in \{1, 2, 3, 4\})$  和边缘注意力图  $f_e$  作为此模块的输入。首先对  $f_e$  执行下采样操作,然后在它们之间进行像素级的乘操作,并使用额外的跳跃连接和  $3 \times 3$  的卷积操

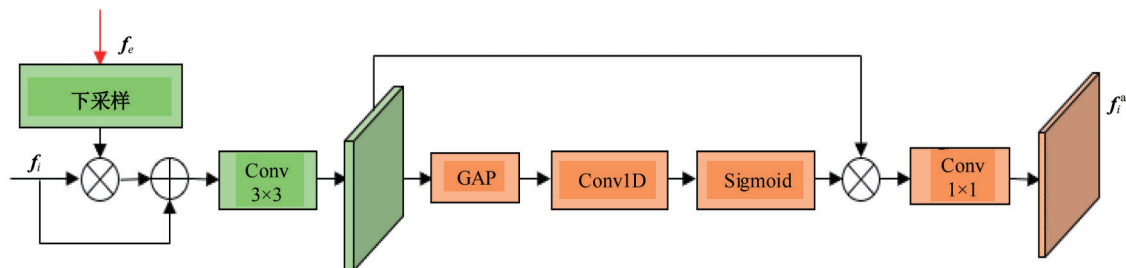


图 3 边缘引导模块结构图

Fig. 3 Architecture of the edge-guidance feature module

作,获得初始融合特征 $f_i^e$ 。接着,采用全局平均池化操作聚合 $f_i^e$ ,并通过一维卷积操作和 Sigmoid 激活函数生成通道注意力图。然后,应用像素级的乘操作将通道注意力图与 $f_i^e$ 融合,引导网络学习关键通道特征。最后,采用 $1 \times 1$ 的卷积操作减少通道数,得到输出特征 $f_i^a$ 。

### 1.4 多尺度上下文感知模块

由于单一尺度的特征学习通常忽略了一些重要的上下文信息,影响最终的检测结果<sup>[12]</sup>。为此,本文基于人类视觉系统识别目标的过程,设计了多尺度上下文感知模块 MCAM,采用膨胀卷积和非对称卷积同时增强感受野,帮助网络学习丰富的上下文信息,具体结构如图 4 所示。

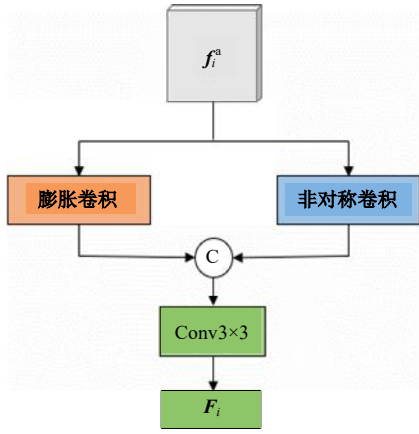


图 4 多尺度上下文感知模块结构图

Fig. 4 Architecture of multi-scale context-aware module

MCAM 共包含 2 个并行的分支,第一个分支采用膨胀率为 3 的  $3 \times 3$  卷积操作,第二个分支采用  $3 \times 3$  的卷积操作、 $1 \times 3$  和  $3 \times 1$  的非对称卷积操作。然后,将两个分支的特征沿着通道维度拼接起来。最后,通过  $3 \times 3$  的卷积操作,得到通道数为 64 的特征图。

### 1.5 特征融合模块

由于篡改手段的多样性,存在前景和背景高度相似的篡改图像,先前的网络主要关注前景信息,而忽略了背景信息,导致模型的检测效果不够理想。因此,本文提出了具有前景路径和背景路径的特征融合模块 FFM,以增强篡改图像中前景和背景的特征表示,实现互补和去噪,具体结构如图 5 所示。

将 MCAM 的输出特征  $F_i$  和上一层 FFM 的输出特征  $F_d^{i-1}$  作为此模块的输入。对于前景路

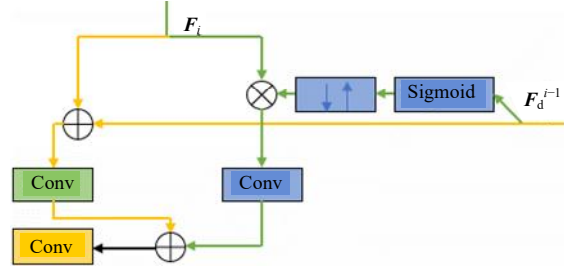


图 5 特征融合模块结构图

Fig. 5 Architecture of feature fusion module

径,沿着通道维度将输入拼接,再通过卷积操作、批量归一化和 ReLU 激活函数,得到前景特征;对于背景路径,本文设计了背景关注分量选择性地关注背景信息。此过程可表示如下:

$$F_{bg} = C\left(\left(1 - \sigma(F_d^{i-1})\right) \odot F_i\right) \quad (1)$$

式中: $C(\cdot)$ 为包含  $3 \times 3$  的卷积操作、批量归一化和 ReLU 非线性激活; $1 - \sigma(F_d^{i-1})$ 为背景注意力图。

最后,沿着通道维度将前景特征  $F_{fg}$ 、背景特征  $F_{bg}$  和上一层 FFM 的输出特征  $F_d^{i-1}$  拼接,得到最终输出特征  $F_d^i$ 。

### 1.6 损失函数

构建网络过程中,由于存在篡改图像前景和背景高度相似的情况,因此需要选择更加合适的损失函数提高检测伪造区域的精度。结合各数据集的特性和本文网络结构,实验选用两种监督方式,即伪造区域边缘监督  $G_e$  和伪造区域掩码监督  $G_r$ 。

对于  $G_e$ ,实验采用骰子损失函数处理正负样本之间的强烈不均衡;对于  $G_r$ ,因前景像素数量远小于背景像素数量时,网络可能会严重偏向于学习背景信息。为此,实验选用加权二分类交叉熵损失函数和加权 IoU 损失函数,引导网络提取伪造区域的结构特征。

由于本文采用多级监督的方式进行训练,解码器的每一层输出一张伪造区域检测图像。因此,最终的损失函数如下所示:

$$L_{total} = \sum_{i=1}^4 \left( L_{wbce}(P_i, G_r) + L_{wIoU}(P_i, G_r) \right) + \lambda L_{dice}(P_e, G_e) \quad (2)$$

式中: $\lambda$ 为一个参数,在本文实验中设置为 4; $P_e$ 为伪造区域的边缘预测图; $P_i$ 为伪造区域的掩码预测图。

## 2 实验结果及分析

### 2.1 数据集及评价指标

本文使用 5 个常用的公开图像篡改检测数据集验证网络的检测性能,分别是 CASIA<sup>[13]</sup>、NIST<sup>[14]</sup>、Columbia<sup>[15]</sup>和 COVER<sup>[16]</sup>、IMD2020<sup>[17]</sup>。其中 Columbia 不包含训练集,因此本文使用 CASIA 数据集进行训练。各数据集的详细情况如表 1 所示。

表 1 数据集细节描述

Table 1 Description of the dataset

数据集	规模		类别			后处理 操作
	训练集	测试集	拼接	复制- 粘贴	移除	
CASIA <sup>[13]</sup>	5 123	921	✓		✓	✓
NIST <sup>[14]</sup>	404	160	✓	✓	✓	✓
Columbia <sup>[15]</sup>	—	180	✓			
COVER <sup>[16]</sup>	75	25		✓		✓
IMD2020 <sup>[17]</sup>	1 610	400				✓

为了验证本文网络的检测性能,采用的评价指标包括  $F_1$  分数和 AUC(ROC 曲线下方面积大小)。 $F_1$  分数和 AUC 用来衡量每个像素的二分类性能,取值范围为 0~1,分数越高表示网络检测性能越好。

### 2.2 实验环境及参数设置

本文实验基于开源的深度学习框架 PyTorch,所有实验均在 Ubuntu 18.04 操作系统上进行,显卡为 NVIDIA GeForce RTX 3090 独立显卡。模型训练过程中使用动量为 0.9 的 Adam 优化器,初始学习率为 0.000 1。图像输入尺寸为  $256 \times 256$ ,每个 batch 包含 64 张图片。

### 2.3 结果及分析

#### 2.3.1 实验结果对比

为了验证本文方法完成图像篡改检测任务的有效性,将本文方法与最新的深度学习模型 ManTra<sup>[18]</sup>、SPAN<sup>[19]</sup>、MVSS<sup>[20]</sup>、GSRNet<sup>[8]</sup>、DenseFCN<sup>[21]</sup>、LocateNet<sup>[22]</sup>和 EMTNet<sup>[10]</sup>进行比较,实验结果如表 2 所示。由表 2 可以看出,在平均 AUC 上,本文方法与最优方法(GSRNet)相差 0.041;但在平均  $F_1$  分数上,本文方法高于次优方法(EMTNet)0.057。对图像取证而言,应确保尽可能多的疑似篡改图像被正确判定为篡改图像是至关重要的,因为某些篡改图像的恶意应用会对社会造成严重的负面影响,如新闻造假等。 $F_1$  分数较高意味着本文方法能够捕获大部分正例样本,减少错误分类的情况,更加适用于现实场景。因此,本文提出的方法具有最佳的整体性能。

伪造区域的可视化检测结果如图 6 所示。由图 6 可以看出,本文方法在图像篡改检测任务中取得了很好的检测结果,在边界模糊、小尺寸伪造区域、前景和背景高度相似等情况下,均实现了较为精准的定位。这主要得益于本文方法所选择的骨干网络能够有效提取篡改图像的基础特征,并配合方法中各模块对多尺度上下文信息进行深入挖掘以及对边缘特征进行了有效学习。

#### 2.3.2 消融实验

为了验证本文方法各模块的有效性,进行了消融实验。本文训练了多个网络,分别对边缘感知模块、边缘引导模块、多尺度上下文感知模块和特征融合模块的有效性进行了分析。将仅包含 PVTv2 的网络命名为 Baseline,在此基础上逐级添加各模块,以此评估每个模块的有效性。所有

表 2 本文方法与其他方法在 5 个数据集的定量比较

Table 2 Quantitative comparison of this paper's method with other methods on five datasets

方法	CASIA		NIST		Columbia		COVER		IMD2020		Mean	
	AUC	$F_1$	AUC	$F_1$	AUC	$F_1$	AUC	$F_1$	AUC	$F_1$	AUC	$F_1$
ManTra <sup>[18]</sup>	0.796	0.267	0.959	0.638	0.736	0.243	0.777	0.283	0.773	0.249	0.808	0.336
SPAN <sup>[19]</sup>	0.709	0.213	0.779	0.252	0.741	0.463	0.791	0.325	—	—	0.755	0.313
MVSS <sup>[20]</sup>	0.847	0.318	0.981	0.768	0.808	0.417	0.808	0.284	0.802	0.396	0.849	0.437
GSRNet <sup>[8]</sup>	0.836	0.340	0.967	0.640	0.900	0.433	0.788	0.218	—	—	0.873	0.408
DenseFCN <sup>[21]</sup>	0.809	0.203	0.979	0.812	0.761	0.257	0.754	0.185	0.715	0.272	0.804	0.346
LocateNet <sup>[22]</sup>	0.754	0.273	0.986	0.738	0.718	0.411	0.813	0.282	—	—	0.818	0.426
EMTNet <sup>[10]</sup>	0.856	0.459	0.987	0.825	0.832	0.561	0.812	0.353	—	—	0.872	0.550
Ours	0.848	0.647	0.965	0.898	0.770	0.609	0.776	0.372	0.801	0.507	0.832	0.607

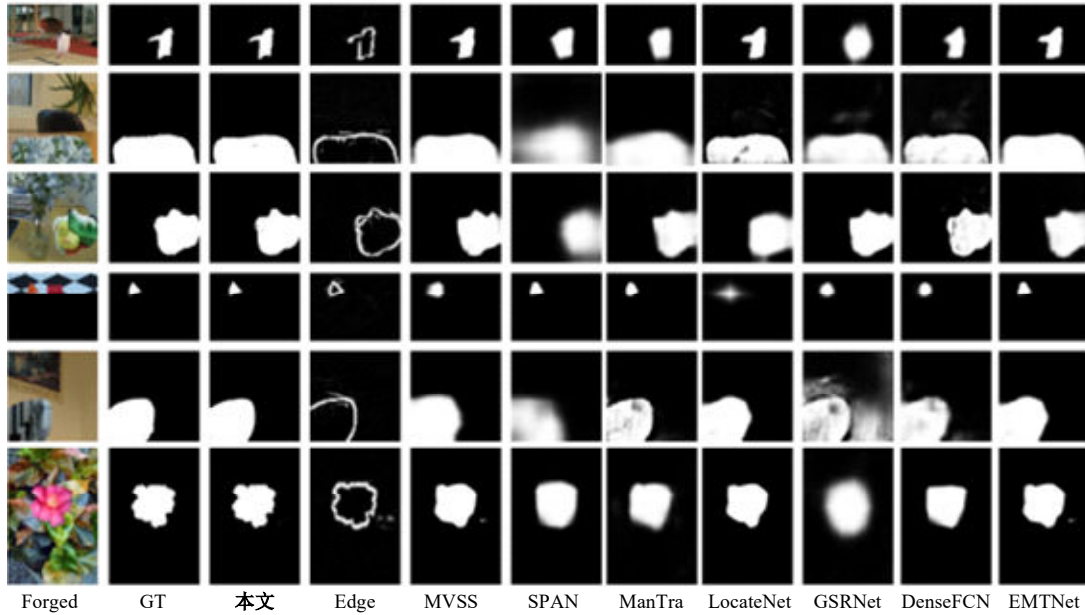


图 6 本文方法与其他方法的定性比较

Fig. 6 Qualitative comparison of this paper's method with other methods

实验均在具有挑战性的数据集 CASIA 上进行,实验结果如表 3 所示。

表 3 不同组件在 CASIA 数据集的定量结果

Table 3 Quantitative results of different components on CASIA dataset

方法	组件				AUC	$F_1$
	EA	EF	MC	FF		
	M	M	AM	M		
a.Baseline					0.800	0.595
b.Baseline+EAM	✓				0.816	0.617
c.Baseline+EAM+EFM	✓	✓			0.835	0.634
d.Baseline+EAM+EFM+MCAM	✓	✓	✓		0.841	0.639
e.Baseline+EAM+EFM+MCAM+FFM	✓	✓	✓	✓	0.848	0.647

(1)边缘感知模块的有效性分析。由表 3 可以看出,在边缘感知模块的辅助下,网络 b 在所有评价指标上均优于网络 a。例如, $F_1$ 分数由 0.595 提升至 0.617,AUC 由 0.800 提升至 0.816。证明了边缘感知模块可以有效提取边缘先验知识,以实现伪造区域的精确检测。

(2)边缘引导模块的有效性分析。通过比较网络 c 与网络 b,可以发现添加边缘引导模块后,网络的性能明显提升。例如, $F_1$ 分数由 0.617 提升至 0.634,AUC 由 0.816 提升至 0.835。这主要是因为边缘引导模块采用了通道注意力机制,减少了冗余通道的干扰,从而达到细化伪造区域边

界的目的。

(3)多尺度上下文感知模块的有效性分析。由表 3 可知,在引入多尺度上下文感知模块后,网络的性能进一步提升。例如, $F_1$ 分数由 0.634 提升至 0.639,AUC 由 0.835 提升至 0.841。这表明了多尺度上下文感知模块捕获的全局上下文信息,有利于提高特征学习的表征能力。

(4)特征融合模块的有效性分析。表 3 数据显示,在特征融合模块的帮助下,网络取得了显著的检测效果。例如, $F_1$ 分数由 0.639 提升至 0.647,AUC 由 0.841 提升至 0.848。这些数据对比证明,特征融合模块可以交替从前景和背景提取有效信息,从而精准分离篡改图像的前景和背景,使检测结果更加准确。

### 2.4 鲁棒性分析

为了验证本文方法的鲁棒性,在 NIST 上应用 3 种后处理方法(JPEG 压缩、高斯噪声、高斯模糊)检测网络的性能。鲁棒性结果如表 4 所示。

由表 4 可以看出,本文方法受 JPEG 压缩影响较小,表现出了较强的鲁棒性,而受高斯噪声和高斯模糊影响较大,这可能是因为边缘感知模块提取边缘特征时引入了额外的信息。后续的工作可以考虑通过交互学习的方式融合相邻层次的特征,以获得更准确的空间信息,进一步提升网络的鲁棒性。

表 4  $F_1$ 分数在 NIST 数据集上的鲁棒性比较  
Table 4 Robustness comparison of  $F_1$  scores on NIST dataset

后处理方法	参数值	方法							
		本文	MVSS	SPAN	ManTra	GSRNet	DenseFCN	LocateNet	EMTNet
高斯噪声	标准差=3	0.872	0.764	0.164	0.060	0.602	0.802	0.724	0.822
	标准差=9	0.854	0.758	0.165	0.051	0.603	0.749	0.707	0.809
	标准差=15	0.839	0.751	0.165	0.050	0.597	0.690	0.682	0.808
JPEG 压缩	质量因子=50	0.873	0.761	0.250	0.186	0.567	0.811	0.691	0.812
	质量因子=75	0.887	0.770	0.252	0.226	0.569	0.812	0.736	0.825
	质量因子=100	0.891	0.766	0.252	0.457	0.572	0.811	0.738	0.825
高斯模糊	卷积核=3	0.880	0.768	0.249	0.198	0.602	0.802	0.737	0.824
	卷积核=9	0.856	0.742	0.239	0.165	0.583	0.749	0.716	0.811
	卷积核=15	0.815	0.711	0.231	0.161	0.573	0.690	0.645	0.782

### 3 结束语

本文针对现有图像篡改检测方法的不足,提出了多尺度感知与边界引导的图像篡改检测方法。首先,使用骨干网络提取篡改图像的空间细节和基础特征。其次,通过边缘感知模块探索与伪造区域相关的边缘信息。再次,利用边缘引导模块突出所提取特征中的关键通道信息,从而实现有效的特征表示。然后,采用多尺度上下文感知模块学习伪造区域不同尺度的上下文信息,进而提高特征学习的表征能力。最后,利用特征融合模块交替关注篡改图像的前景和背景,生成可视化检测结果。实验结果表明,本文方法在多个图像篡改检测数据集上的检测性能稳定,并优于其他对比方法。

#### 参考文献:

- [1] Shi Z, Chen H, Zhang D. Transformer-auxiliary neural networks for image manipulation localization by operator inductions[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2023, 33(9): 4907-4920.
- [2] 钟辉, 康恒, 吕颖达, 等. 基于注意力卷积神经网络的图像篡改定位算法[J]. 吉林大学学报: 工学版, 2021, 51(5): 1838-1844.  
Zhong Hui, Kang Heng, Lyu Ying-da, et al. Image manipulation localization algorithm based on channel attention convolutional neural networks[J]. Journal of Jilin University (Engineering and Technology Edition), 2021, 51(5): 1838-1844.
- [3] 石泽男, 陈海鹏, 张冬, 等. 预训练驱动的多模态边界感知视觉 Transformer[J]. 软件学报, 2023, 34(5): 2051-2067.  
Shi Ze-nan, Chen Hai-peng, Zhang Dong, et al. Pretraining-driven multimodal boundary-aware vision transformer[J]. Journal of Software, 2023, 34(5): 2051-2067.
- [4] Xu D, Shen X, Lyu Y, et al. MC-Net: Learning mutually complementary features for image manipulation localization[J]. International Journal of Intelligent Systems, 2022, 37(5): 3072-3089.
- [5] Mahdian B, Saic S. Using noise inconsistencies for blind image forensics[J]. Image and Vision Computing, 2009, 27(10): 1497-1503.
- [6] Lin Z, He J, Tang X, et al. Fast, automatic and fine-grained tampered JPEG image detection via DCT coefficient analysis[J]. Pattern Recognition, 2009, 42(11): 2492-2501.
- [7] Popescu A C, Farid H. Exposing digital forgeries in color filter array interpolated images[J]. IEEE Transactions on Signal Processing, 2005, 53(10): 3948-3959.
- [8] Zhou P, Chen B C, Han X, et al. Generate, segment, and refine: towards generic manipulation segmentation[C]//Proceedings of the AAAI Conference on Artificial Intelligence, New York, USA, 2020: 13058-13065.
- [9] Wang J, Wu Z, Chen J, et al. Objectformer for image manipulation detection and localization[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, USA, 2022: 2364-2373.
- [10] Lin X, Wang S, Deng J, et al. Image manipulation-detection by multiple tampering traces and edge artifact enhancement[J]. Pattern Recognition, 2023, 133: 109026-109036.
- [11] Wang W, Xie E, Li X, et al. PVT v2: improved baselines with pyramid vision transformer[J]. Computational Visual Media, 2022, 8(3): 415-424.

- [12] 胡林辉, 陈保营, 谭舜泉, 等. 基于 Convnext-Upernet 的图像篡改检测定位模型[J/OL]. [2023-09-10]. <https://kns.cnki.net/kcms/detail/11.1826.tp.20230508.1605.002.html>  
Hu Lin-hui, Chen Bao-ying, Tan Shun-quan, et al. Convnext-Upernet based deep-learning model for image forgery detection and localization[J/OL]. [2023-09-10]. <https://kns.cnki.net/kcms/detail/11.1826.tp.20230508.1605.002.html>
- [13] Dong J, Wang W, Tan T. Casia image tampering detection evaluation database[C] //2013 IEEE China Summit and International Conference on Signal and Information Processing, Beijing, China, 2013: 422-426.
- [14] Guan H, Kozak M, Robertson E, et al. MFC datasets: Large-scale benchmark datasets for media for forensic challenge evaluation[C]//2019 IEEE Winter Applications of Computer Vision Workshops (WACVW), Waikoloa, USA, 2019: 63-72.
- [15] Hsu Y F, Chang S F. Detecting image splicing using geometry invariants and camera characteristics consistency[C]//2006 IEEE International Conference on Multimedia and Expo, Toronto, Canada, 2006: 549-552.
- [16] Wen B, Zhu Y, Subramanian R, et al. COVERAGE: a novel database for copy-move forgery detection[C]//2016 IEEE International Conference on Image Processing (ICIP), Phoenix, USA, 2016: 161-165.
- [17] Novozamsky A, Mahdian B, Saic S. IMD2020: a large-scale annotated dataset tailored for detecting manipulated images[C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision Workshops, Snowmass Village, USA, 2020: 71-80.
- [18] Wu Y, AbdAlmageed W, Natarajan P. Mantra-net: Manipulation tracing network for detection and localization of image forgeries with anomalous features [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, USA, 2019: 9543-9552.
- [19] Hu X, Zhang Z, Jiang Z, et al. SPAN: spatial pyramid attention network for image manipulation localization[C]//Proceedings of the European Conference on Computer Vision (ECCV), Glasgow, UK, 2020: 312-328.
- [20] Chen X, Dong C, Ji J, et al. Image manipulation detection by multi-view multi-scale supervision[C] // Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, Canada, 2021: 14165-14173.
- [21] Zhuang P, Li H, Tan S, et al. Image tampering localization using a dense fully convolutional network [J]. IEEE Transactions on Information Forensics and Security, 2021, 16: 2986-2999.
- [22] Zhuo L, Tan S, Li B, et al. Self-adversarial training incorporating forgery attention for image forgery localization[J]. IEEE Transactions on Information Forensics and Security, 2022, 17: 819-834.