

一种改进的 YOLOv8 图像篡改检测算法

宣高媛

安徽理工大学 人工智能学院, 安徽 淮南 232001

摘要:目的 近年来,随着数字技术的发展,图像篡改已经成为一个日益严重的问题,现有的众多图像篡改和目标检测方法均存在着识别精度不足和检测效果不理想的问题。为了更有效解决这一挑战,提出一个基于改进的 YOLOv8 检测算法,希望能够实现更高的检测精度。**方法** 首先,为了捕获篡改的边缘特征,引入局部和全局注意力机制,对 YOLOv8 的骨干网络进行深度优化,这种优化结合了上下文感知和局部增强技术,大大增强了对边缘特征的识别力;考虑到篡改区域可能存在各种形态,进一步采用堆叠特征金字塔的网络结构,以捕获多尺度特征;最后,为了提高模型的计算效率和推理速度,在模型中结合深度可分离卷积和通道重排技术。**结果** 在一系列的实验中,改进的 YOLOv8 篡改检测算法在 CASIA2.0 图像篡改数据集上展现了出色的性能,与原始算法相比,其准确率高达 82.3%,明显提高了检测的效果;提出的基于改进的 YOLOv8 篡改检测算法,经过深度的网络优化和结构调整,成功地提高了图像篡改检测的准确性和效率。**结论** 文中提出的方法在图像篡改检测方面表现出较高的准确性,为图像篡改检测领域带来了重要的进展,具有深远而重要的意义。

关键词: 图像篡改; 目标检测; YOLOv8; 上下文感知; 多尺度特征; 深度可分离卷积

中图分类号: TP391.41; TP309 **文献标识码:** A **doi:** 10.16055/j.issn.1672-058X.2025.0003.012

An Improved YOLOv8 Algorithm for Image Tampering Detection

XUAN Gaoyuan

School of Artificial Intelligence, Anhui University of Science and Technology, Anhui Huainan 232001, China

Abstract: Objective With the advancement of digital technology in recent years, image tampering has emerged as a growing concern. Many existing methods for image tampering and object detection suffer from inadequate recognition accuracy and unsatisfactory detection effects. To address this challenge more effectively, a detection algorithm based on an improved YOLOv8 was proposed, aiming to achieve higher detection accuracy. **Methods** Firstly, to capture the edge features of tampering, a local and global attention mechanism was introduced to deeply optimize the backbone network of YOLOv8. This optimization combined context awareness and local enhancement techniques, significantly enhancing the recognition of edge features. Considering the diverse shapes of tampered regions, a network structure of stacked feature pyramids was further adopted to capture multi-scale features. Finally, to improve the computational efficiency and inference speed of the model, deep separable convolutions and channel shuffling were integrated into the model. **Results** In a series of experiments, the improved YOLOv8 tampering detection algorithm demonstrated excellent performance on the CASIA2.0 image tampering dataset. Compared with the original algorithm, it achieved an accuracy as high as 82.3%,

收稿日期: 2023-10-10 **修回日期:** 2023-12-11 **文章编号:** 1672-058X(2025)03-0094-08

基金项目: 国家自然科学基金资助项目(22001026); 重庆市科委面上项目(CSTB2022NSCQ-MSX1308)。

作者简介: 宣高媛(1998—),女,安徽合肥人,硕士研究生,从事图像篡改检测和数字图像处理研究。

引用格式: 宣高媛. 一种改进的 YOLOv8 图像篡改检测算法[J]. 重庆工商大学学报(自然科学版), 2025, 42(3): 94-101.

XUAN Gaoyuan. An improved YOLOv8 algorithm for image tampering detection[J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2025, 42(3): 94-101.

significantly enhancing detection effectiveness. The proposed tampering detection algorithm based on the improved YOLOv8, through deep network optimization and structural adjustments, successfully improved both the accuracy and efficiency of image tampering detection. **Conclusion** The proposed method in the paper demonstrates high accuracy in image tampering detection, representing significant advancements in this field. This study has profound and meaningful implications for the field of image tampering detection.

Keywords: image tampering; target detection; YOLOv8; context-awareness; multi-scale features; depth-separable convolution

1 引言

随着数字图像技术的快速发展,图像成为记录和传播信息的重要媒介。但是,数字图像相对容易进行修改,如图1所示,这给图像的真实性和可信度带来了挑战。一些非法行为,如图片著作权盗用、医疗诊断图片篡改^[1-2]和法律证据图片都可能因为篡改而产生误导。另一方面,技术性篡改,如图片编辑^[3]、特效加工等也日益普及。因此,如何有效检测图片是否遭过篡改,验证其真实性,乃至追踪篡改细节,已成为一个重要而迫在眉睫的问题。

过去,图像篡改检测主要依靠人工检测方法,例如对比图片的细节或与原始场景进行匹配等。但这些方法效率低下,难以应对大量图片。计算机视觉及机器学习的兴起,首次出现了基于图像统计特征的自动检测方法^[4-5]。比如通过像素点差异分布、色彩空间特征、错误线等来识别篡改痕迹。但这类方法仅能分辨简单的操作,对复杂篡改依然难以有效排查。

随着深度学习技术的兴起,许多基于卷积神经网络的端到端检测模型被提出,通过学习图像本体特征有效提升了检测效果^[6]。尽管如此,图像篡改检测技术在实践中仍然面临诸多挑战。U-Net是为图像分割而设计的神经网络结构^[7],U-Net可以从少量标注数据中进行有效的端到端训练,并在数据增强的帮助下进一步提高其性能。尽管U-Net能够快速精确地对图像中的每个像素进行分割,但在某些场景中,可能会造成因缺乏足够大范围的上下文信息而导致边界定位不准的问题。DeepLabv3+是一种为语义分割设计的深度神经网络模型^[8],它融合空间金字塔池化模块的能力来编码多尺度上下文信息和编码-解码结构的优势来清晰地捕捉物体边界。然而,由于其在处理图像时主要关注像素级的标注,可能会丢失一些较大区域的上

下文信息,导致某些篡改部分不易被检测。Zhou等^[9]提出一种创新的双流RGB-N网络。这个网络由两个分支构成:一个是RGB流,负责从RGB图像中提取特征;另一个是噪声流,它使用隐写分析的富模型进行滤波,以此来提取噪声特征,进而区分真实和篡改区域之间的噪声差异。RRU-Net是为图像拼接伪造检测而设计的深度神经网络模型^[10]。它核心的创新在于模仿人类大脑的回忆和巩固机制,通过卷积网络中的残差传播和反馈过程来加强学习方式。但针对移动复制的复杂篡改场景,检测能力提升有限。Chen等^[11]提出了MVSS-Net++,这是一种结合多视角特征学习和多尺度监督的网络。MVSS-Net++包括两个主要分支:噪声分支和边界分支。噪声分支专注于学习与语义无关的特征,从而获得更通用的特征表示;边界分支专注于学习篡改区域与真实区域在边界处的不一致性。沈万里等^[12]提出一个用于图像修复取证的U型FPN,采用VGG16和注意力机制,通过融合损失函数优化预测效果,但是该方法没能充分提取出篡改的特征信息。

尽管上述方法在图像篡改检测中取得了一定的进展,但大多数方法都存在丢失上下文信息和边界定位不准等问题。为此,本文提出一种改进的YOLOv8图像篡改检测算法,其中包括两个关键模块:C2f_CloAtt模块和VoVGSCGP模块。C2f_CloAtt模块结合上下文感知和局部增强技术,大大增强了对边缘特征的识别力;VoVGSCGP模块采用堆叠特征金字塔的网络结构,以捕获多尺度特征,结合深度可分离卷积和通道重排技术显著提升了图像篡改检测效果。

2 改进的YOLOv8检测算法

本文提出的一种改进的YOLOv8算法检测模型,用于在复杂场景中快速检测图像篡改区域,算法网络结构如图1所示。

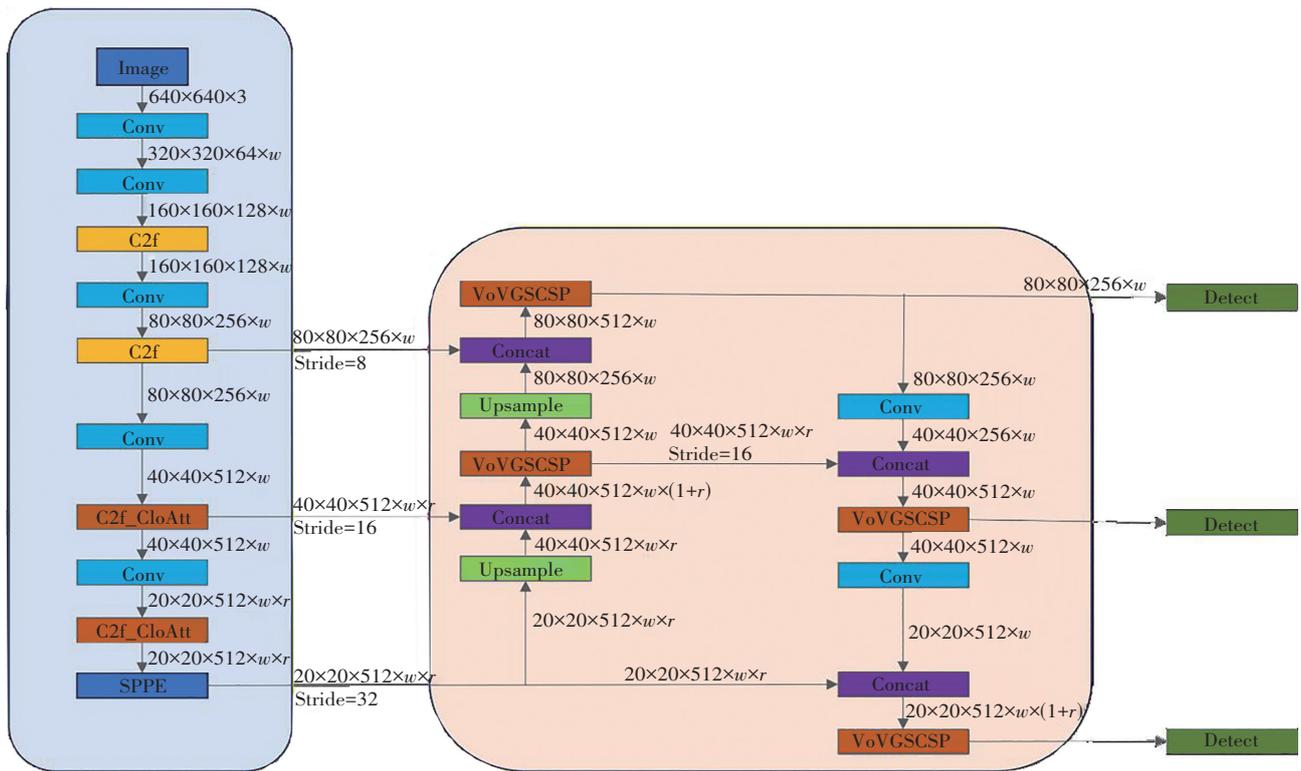


图 1 改进的 YOLOv8 算法网络结构

Fig. 1 Network structure of the improved YOLOv8 algorithm

2.1 算法简介

YOLOv8 是当前效果较好的目标检测算法,它的核心网络来源于 DarkNet-53,该网络初次在 YOLOv3^[11] 中被引入,并深受 ResNet^[12] 的影响。DarkNet-53 使用了残差机制,并连续添加了卷积模块来加强其功能性。这 53 层构造使其在性能上与 ResNet-152 相当,但其处理速度加快了一倍。基于 DarkNet-53, YOLOv8 做了进一步调整,融入了 C2f 组件,进一步加强了性能并减少了模型的大小。

本文选用轻量级的 YOLOv8n。YOLOv8n 是从 YOLOv8 算法派生出的轻量级参数结构。它包括骨干网、颈网和预测输出头。骨干网络利用卷积操作从 RGB 图像中提取各种尺度的特征。同时,颈部网络的作用是合并骨干网络提取的特征。特征金字塔结构(特征金字塔网络,FPN^[13])通常被实现将低级特征聚合高级表示。头部层负责预测目标类别,并利用 3 组不同大小的检测器来选择和检测图像内容。

本文所提模型将原来的主干网络替换为具有全局和局部特征的注意力机制模块(C2f_CloAtt),使用共享权重来聚合局部信息,并通过部署上下文感知权重来增强局部特征,更加高效和直接地捕获局部信息,使模

型能够感知高频和低频信息。其次,通过引入 VoVGSCSP 模块代替原来的 C2f 模块,该模块结合了标准卷积、深度可分离卷积和混合通道,以减少网络参数和计算量,使模型轻量化,更容易部署到端点。将输入特征图分成两个通道,并赋予不同的权重。两个通道根据通道数量进行组合和连接,这增强了算法表达非线性函数的能力,并解决了梯度消失问题。

2.2 C2f_CloAtt 模块

近年来, Vision Transformer^[16] 以其卓越的全局建模能力迅速成为流行的视觉架构。很多改进方法尝试优化 Vision Transformer 在数据效率和结构设计上的主因,而大部分通过限制注意力范围来减少计算复杂性。例如, Swin-Transformer^[17] 通过局部窗口自注意力减少计算量,同时结合全局和局部的信息建模。其他替代方法在前向传递中采用动态令牌选择,裁剪或池化含少量信息的令牌来降低计算量,这也是结合了全局和局部注意力的策略。

MobileViT 以其将卷积与注意力结合的方式,在少量参数下实现高性能而受到广泛认可; EdgeViT^[18] 采用局部-全局-局部模块来聚合信息,取得了较高的准确检测结果。在处理局部信息时,现有方法通常采用两

种途径。一种是使用具有共享权重的卷积,另一种方法使用与局部自注意力紧密相关的上下文权重。在应用到图像篡改检测中, Ren 等^[19]提出一种创新的多尺度注意力上下文感知网络(MAC-Net)。该网络引入多尺度多层次注意力模块(MMAM),它能够处理不同尺度特征不一致的问题,并自动调整输入特征的系数以维持详细特征。之后,这些融合特征会被输入到所提出的多分支全局上下文模块(MGCM)。MGCM具有3个不同的分支,能够丰富上下文信息并维护目标的详细特征。尽管这些方法都使用相似性得分作为上下文权重,但在C2f_CloAtt模块中,试图更有效地利用本地共享感知和上下文感知的权重优势。

C2f_CloAtt模块采用双分支结构。在本地分支中使用注意力卷积,有效融合共享和上下文感知权重,聚合高频局部信息(图2)。注意力卷积利用深度卷积(DWconv)具有的共享权重来提取局部表示,并部署上下文权重来增强局部特征。与通过局部自注意力生成上下文感知权重的方法不同^[20],注意力卷积使用门控机制生成上下文感知权重,这引入了比常用注意力机制更强的非线性。注意力卷积对 Q (Query)和 K (Key)应用卷积运算符以聚合局部信息,然后计算 Q 和 K 的Hadamard积,并对结果执行一系列线性或非线性变换以生成范围在 $[-1, 1]$ 内的上下文感知权重。对于全局分支,使用普通注意力,但对 K 和 V (Value)进行下采样以减少GFLOPs,这有助于模型捕获低频全局信息。通过融合局部和全局分支输出同时捕获高频和低频信息。

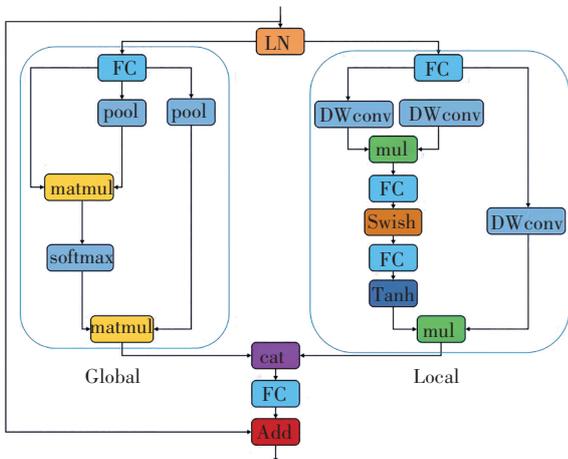


图2 C2f_CloAtt模块

Fig. 2 C2f_CloAtt module

C2f_CloAtt模块由一个局部分支和一个全局分支

组成。在全局分支中,首先对 K 和 V 进行下采样,然后对 Q 、 K 和 V 执行标准注意力过程,以提取低频全局信息:

$$X_{\text{global}} = X_{\text{attn}}(Q_g, P_{\text{pool}}(K_g), \text{POOL}(V_g)) \quad (1)$$

其中, X_{global} 是全局分支的输出, X_{attn} 是应用注意力机制后的结果, Q_g 是Query的全局表示,POOL是下采样操作。全局分支模式有效减少了注意力所需的浮点操作数(FLOP),并且获得了全局感受野。然而,虽然它有效地捕获了低频全局信息,但其处理高频局部信息的能力不足。为了解决这个限制,本文使用注意力卷积作为解决方案。

注意力卷积是使本文的模型实现高性能的关键模块,它包含了一些标准的注意操作。具体来说,在注意力卷积中,首先应用线性变换来获得 Q 、 K 和 V ,这与标准注意力相同:

$$X_{\text{attn}} = f_{\text{FC}}(X_{\text{in}}) \quad (2)$$

式(2)中, X_{attn} 是 Q 、 K 和 V 的表示, X_{in} 是注意力卷积的输入, f_{FC} 表示在全连接层的线性转换操作。在执行线性变换之后,首先在 V 上进行具有共享权重的局部特征聚合过程,然后,基于处理的 V 和 Q 、 K ,执行上下文感知的局部增强。

局部特征聚合过程中,对于 V ,使用简单的深度卷积 D_{dwconv} 来聚合局部信息。深度卷积的权重是全局共享的:

$$V_s = D_{\text{dwconv}}(V) \quad (3)$$

上下文感知本地增强过程中,在将 V 的局部信息与共享权重整合后,将 Q 和 K 组合以生成上下文感知权重。值得注意的是,本文使用的是一种与局部自我关注不同的方法,这种方法更加精细。具体来说,首先使用两个深度卷积分别聚合 Q 和 K 的本地信息;然后,计算 Q 和 K 的Hadamard乘积,并对结果进行一系列变换,以获得 -1 和 1 之间的上下文感知权重;最后,使用生成的权重来增强局部特征。整个过程可概括如下:

$$Q_l = D_{\text{dwconv}}(Q)$$

$$K_l = D_{\text{dwconv}}(K)$$

$$X_{\text{attn}} = F_{\text{fc}}(S_{\text{swish}}(\text{FC}(Q_l \odot K_l))) \quad (4)$$

$$X_{\text{attn}} = T_{\text{tanh}}\left(\frac{X_{\text{attn}}}{\sqrt{d}}\right)$$

$$X_{\text{local}} = X_{\text{attn}} \odot V_s$$

式(4)中, d 是令牌的通道数, \odot 表示Hadamard积。与普通注意力相比,本文方法引入了更强的非线性。具

体而言,局部自注意的上下文感知权重生成过程中唯一的非线性算子是 Softmax。例如, T_{\tanh} 函数和 S_{swish} 函数,更强的非线性生成更高质量的上下文感知权重。

在全球分支融合过程中,本文使用一个简单的方法来融合本地分支的输出与全局分支。具体地,在通道维度中连接两个输出,之后,将全连接层应用于通道维度:

$$\begin{aligned} X_t &= \text{Contact}(X_{\text{local}}, X_{\text{global}}) \\ X_{\text{out}} &= f_{\text{FC}}(X_t) \end{aligned} \quad (5)$$

这里, Contact 表示在通道维度上连接两个输出。

2.3 VoVGSCGP 模块

GSConv (Grouped Separable Convolutions) 通过在其结构中引入深度可分离卷积 (DSC) 层来增强非线性表示,目的是在较小的时间复杂度下维持各通道间的隐藏连接。GSConv 的设计目标是使其输出尽可能地接近信道密集卷积 (SC), 同时大幅降低计算成本。然而,如果在主网络结构中大量使用 GSConv, 模型的网络层数量会增加,从而可能导致数据流遭受阻碍和推理时间上升。但当 GSConv 应用于网络的瓶颈层时,得益于特征图在这一部分的通道尺寸最大而宽度和高度维度最小,它能够减少重复信息,同时避免不必要的压缩。

GSConv 是一种轻量级卷积方法,如图 3 所示。该结构通过添加深度卷积层和 Shuffle 来增强非线性表示,尽可能地以较小的时间复杂度保留每个通道之间的隐藏连接。该方法使卷积计算的输出尽可能接近信道密集卷积,降低了计算成本。如果在主干网中使用,则模型的网络层将加深,这将增加数据流的阻力,大大增加 GSConv 的推理时间。当 GSConv 用于颈部时,由于特征图的通道尺寸最大,宽度和高度维度最小,因此使用 GSConv 对特征图进行串行处理可以减少重复信息,并且不需要压缩。

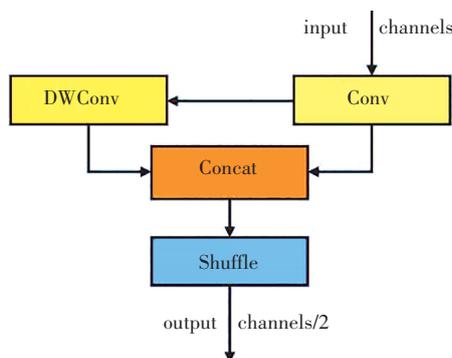


图 3 GSConv 的网络架构

Fig. 3 Network architecture of GSConv

VoVGSCSP 是在 GSConv 的基础上继续引入瓶颈层,然后采用一次性聚合方法设计一个跨阶段部分网络 (GSCSP) 模块 VoVGSCSP,如图 4 所示。该模块很好地平衡了模型的精度和速度,降低了计算和网络结构的复杂性,同时保持足够的精度和提取特征的高重用率。

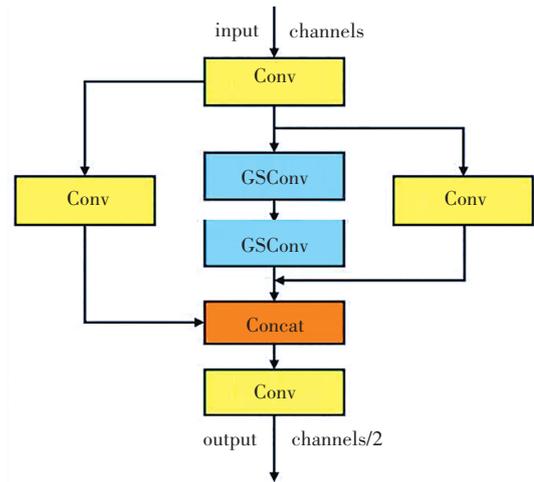


图 4 VoVGSCSP 的网络架构

Fig. 4 Network architecture of VoVGSCSP

3 仿真实验与结果分析

3.1 超参数设置

本文的所有模型都基于 Pytorch 深度学习框架和 Python 编程语言在 Ubuntu 20.04 环境中,利用 NVIDIA GeForce RTX 3090 进行实现的。在此选择随机梯度下降 (SGD) 作为优化器,并根据 hyp. yaml 的超参数配置进行相应地调整。训练持续了 200 个周期,其中前 100 个周期利用早期停止机制来防止过拟合。每个训练批次包含 64 张图片,每张图片的大小为 $640 * 640$ 。在测试阶段,为了减少重叠和冗余的检测框,采用非极大值抑制 NMS (Non-Maximum Suppression)。NMS 的阈值为 0.3,这有助于提高检测的精度,同时避免了过度抑制可能导致的目标遗漏。

3.2 损失函数

在本文方法中,回归损失采用 DF (Distribution Focal) Loss 和 CIOU (Complete Intersection over Union) Loss 的组合,以提高模型在目标定位上的准确性和鲁棒性。DF Loss 专注于改善目标边界框的精确预测,特别是对于难以定位的目标,其核心思想是通过考虑边界框坐标的概率分布来精细化预测,这可以表示为

$$L_{DF} = - \sum_{i \in Pos} (1 - P_i)^\gamma \log(P_i) \quad (6)$$

其中, P_i 表示预测边界框与真实边界框的相似度, γ 是调节因子, 用于平衡易于预测和难以预测的目标权重。

CIOU Loss 是一种先进的边界框相似度度量方法, 不仅考虑了边界框间的重叠面积, 还考虑了边界框的中心距离和长宽比。其公式可以表示为

$$L_{CIOU} = 1 - R_{IOU} + \frac{\rho^2(b, b_{gt})}{c^2} + \alpha \cdot v \quad (7)$$

其中, R_{IOU} 代表交并比, $\rho^2(b, b_{gt})$ 是预测边界框 b 和真是边界框 b_{gt} 之间的距离, c 是包含这两个边界框的最小闭合区域的对角线长度, α 和 v 是用于调整长宽比的参数。

综合上述部分, 总的回归损失为

$$L_{reg} = \lambda_{df} \cdot L_{DF} + \lambda_{ciou} \cdot L_{CIOU} \quad (8)$$

式(8)中, λ_{df} 和 λ_{ciou} 是用于平衡 DFL 和 CIOU Loss 的权重系数。

3.3 实验数据

本研究使用 CASIA2.0^[21] 公开数据集, 有 12 323 张彩色篡改图像示例, 主要包含拼接和复制移动。在 CASIA2.0 上, 伪造区域是物体, 小而精细。数据集为未压缩数据, 也是现在图像篡改领域应用最多的公共数据集。训练集和测试集按照 8 : 2 的比例进行划分, 以确保训练和测试数据中不会同时出现相同的背景和篡改对象。为防止由于数据样本较少而导致过拟合, 提高模型的训练效果, 因此对训练数据进行扩展, 通过随机高斯噪声、JPEG 压缩和随机反转进行数据增强。完成模型的预训练后, 为了评估算法的泛化能力, 使用 CASIA2.0 数据集对模型性能进行评估。

3.4 结果与讨论

根据表 1 中的实验数据, 可以得出结论: 在 CASIA 数据集上, 本文提出的改进算法相对于其他图像篡改检测方法和原始 YOLOv8 算法表现更优。由于 CASIA 数据集中小目标较多, 使用复制-移动篡改方式的样本所占比重相对较大。该方式的篡改内容与原图像来源一致, 所产生的篡改特征不太明显, 在这类检测中, 本文方法表现出更好的检测准确性。

为了进一步验证 C2f_CloAtt 模块和 VoVGSCSP 模块的有效性, 设计了一系列消融实验, 并将结果列于

表 2。从表 2 中可以看出: 对网络应用不同的检测模块后, 检测精度提升了大约 2% ~ 3%。这表明 C2f_CloAtt 模块和 VoVGSCSP 模块成功解决了针对篡改边缘难以提取篡改特征的问题, 对图像篡改检测效果有显著提升, 图 5 为可视化检测效果。

表 1 模型评估结果

方 法	准确率
DeepLabV3+	0.358
Unet	0.419
RRUnet	0.488
YOLOv8-s	0.783
Ours	0.823

表 2 不同模块的消融实验情况

C2f_CloAtt 模块	VoVGSCSP 模块	准确率
—	—	0.773
✓	—	0.796
—	✓	0.784
✓	✓	0.823

图 6 为本文方法与其他方法计算复杂度与准确率对比。如图 6 所示, 本文提出的方法在保持高准确率的同时, 展现出最低的计算复杂度。相比之下, 尽管 YOLOv8-s 在准确率上表现出竞争力, 但它的计算复杂度太高。这一对比彰显了本文方法在效率和精确度方面的显著优势。

低计算复杂度的优势意味着在计算资源受限的环境下, 本文提出的方法更具适用性。此外, 算法的优化不仅体现在降低计算资源的消耗上, 更重要的是在保持或提高检测准确性的同时, 减少对计算资源的需求。这种平衡表明, 本文方法对于实时图像篡改检测应用既高效又实用。

本文算法使得网络能够学习到篡改图像中更加微妙的特征。通过权重参数, 将不同层次的特征进行融合, 从而增强了网络对篡改特征的学习。因此, 相对于其他方法, 算法在 CASIA2.0 数据集上实现了显著提升的精确率, 证明了本文所提算法的有效性。

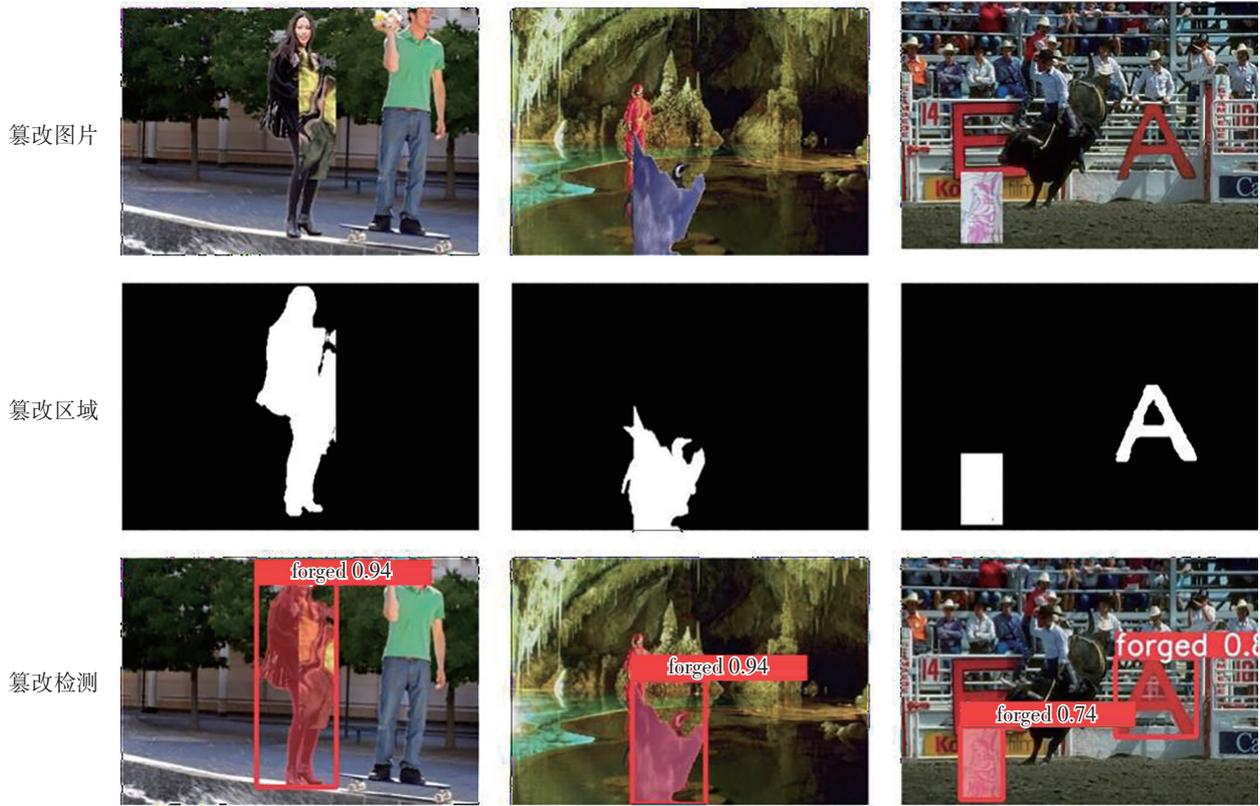


图 5 本文方法在篡改图像上的可视化检测结果

Fig. 5 The visual detection results of the proposed method on the tampered images

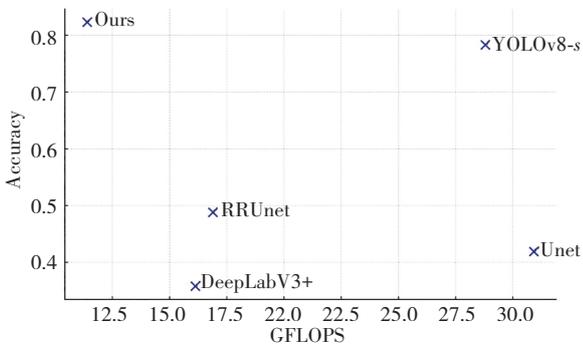


图 6 本文方法与其他图像篡改检测算法的计算复杂度与准确率对比

Fig. 6 Comparison of computational complexity and accuracy between our method and other image tampering detection algorithms

4 结论

提出一种基于改进的 YOLOv8 图像篡改检测方法。通过引入 CloAtt 骨干网络、堆叠特征金字塔的结构以及 GSConv 模块和 GS 瓶颈模型的结合,在多个方面实现了显著的优化。在 CASIA2.0 数据集上的测试结果证明了本文方法相对于其他算法具有明显的优越性,准确率达到了 82.3%,优于原始算法。与其他领先的

图像篡改检测模型相比,具有显著的竞争优势。尽管如此,该方法也存在一定的局限性。首先,在处理高度复杂的图像篡改案例时,尽管算法显示了良好的性能,但在一些极端情况下,如极其精细的篡改或高度逼真的合成图像中,检测准确性可能会受到影响。此外,当前的算法主要针对特定类型的篡改(例如拼接和复制移动篡改),对于新出现的或未知类型的篡改技术识别能力有限。在后续的工作中,会专注研究和探索针对更加复杂场景的图像篡改检测技术,以设计出更加精确和高效的检测方法。

参考文献:

[1] WU Y, ABD-ALMAGEED W, NATARAJAN P, et al. Deep matching and validation network: An end-to-end solution to constrained image splicing localization and detection [C]// International Conference on Multimedia. New York: ACM, 2017: 1480–1502.

[2] HUH M, LIU A, OWENS A, et al. Fighting fake news: Image splice detection via learned self-consistency [C]// Proceedings of the Computer Vision-ECCV 2018. New York:

- ACM, 2018: 106–124.
- [3] WU Y, ABD-ALMAGEED W, NATARAJAN P. BusterNet: Detecting copy-move image forgery with source/target localization[C]//Proceedings of the Computer Vision-ECCV 2018. Cham: Springer International Publishing, 2018: 170–186.
- [4] VERDOLIVA L. Media forensics and DeepFakes: An overview[J]. IEEE Journal of Selected Topics in Signal Processing, 2020, 14(5): 910–932.
- [5] ZHANG Y, WANG H, ZHANG R, et al. Semantic consistency-relevant multitask splicing-tempered detection [J]. Journal of Chinese Image and Graphics, 2023, 28(3): 775–788.
- [6] 张玉林, 王宏霞, 张瑞, 等. 语义一致性引导的多任务拼接篡改检测[J]. 中国图象图形学报, 2023, 28(3): 775–788.
- ZHANG Yu-lin, WANG Hong-xia, ZHANG Rui, et al. Semantic consistency-relevant multitask splicing-tempered detection[J]. Journal of Image and Graphics, 2023, 28(3): 775–788.
- [7] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention-NICCAI 2015. Cham: Springer International Publishing, 2015: 234–241.
- [8] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]//Computer Vision-ECCV 2018. Cham: Springer International Publishing, 2018: 833–851.
- [9] ZHOU P, HAN X, MORARIU V I, et al. Learning rich features for image manipulation detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2018: 1053–1061.
- [10] BI X, WEI Y, XIAO B, et al. RRU-Net: The ringed residual U-Net for image splicing forgery detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. IEEE, 2019: 30–39.
- [11] DONG C, CHEN X, HU R, et al. MVSS-net: Multi-view multi-scale supervised networks for image manipulation detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(3): 3339–3353.
- [12] 沈万里, 张玉金, 胡万. 面向图像修复取证的U型特征金字塔网络[J]. 计算机应用, 2023, 43(2): 545–551.
- SHEN Wan-li, ZHANG Yu-jin, HU Wan. U-shaped feature pyramid network for image inpainting forensics[J]. Journal of Computer Applications, 2023, 43(2): 545–551.
- [13] ZHAO L, LI S. Object detection algorithm based on improved YOLOv3[J]. Electronics, 2020, 9(3): 537.
- [14] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2016: 770–778.
- [15] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2017: 2117–2125.
- [16] ZHAO H, JIANG L, JIA J, et al. Point transformer [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. IEEE, 2021: 16259–16268.
- [17] LIU Z, LIN Y, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. IEEE, 2021: 10012–10022.
- [18] PAN J, BULAT A, TAN F, et al. EdgeViTs: Competing light-weight CNNs on mobile devices with vision transformers [C]//Computer Vision-ECCV 2022. Cham: Springer Nature Switzerland, 2022: 294–311.
- [19] REN R, NIU S, JIN J, et al. Multi-scale attention context-aware network for detection and localization of image splicing[J]. Applied Intelligence, 2023, 53(15): 18219–18238.
- [20] HASSANI A, WALTON S, LI J, et al. Neighborhood attention transformer [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2023: 6185–6194.
- [21] DONG J, WANG W, TAN T. CASIA image tampering detection evaluation database [C]//Proceedings of the IEEE China Summit and International Conference on Signal and Information Processing. IEEE, 2013: 422–426.

责任编辑:李翠薇