

DOI: 10.13382/j.jemi.B2508709

基于双分支特征融合与注意力机制的变化检测网络*

周先春^{1,2,3} 倪鸿骏² 李翰哲³

(1. 元宇宙文旅场景应用技术研究江苏省文化和旅游重点实验室 南京 210044; 2. 南京信息工程大学电子与信息工程学院 南京 210044; 3. 南京信息工程大学人工智能学院 南京 210044)

摘要:近年来,深度学习在变化检测领域展现出了卓越的性能,但仍存在由于细节信息丢失和语义信息不足而导致的细粒度边界不够清晰以及存在内部空洞等不足。针对这一问题,提出一种多尺度自适应融合网络(multi-scale adaptive fusion network, MSAF-Net),该网络以 ResNet 为骨干网络,首先提出一个特征差异融合模块(feature difference fusion module, FDFM),该模块仅通过简单的操作就能从变化区域中粗略提取边界信息。然后,针对细节纹理信息丢失的问题,提出多尺度双特征融合模块(multi-scale dual-feature fusion module, MDFM),将多个感受野中的低级纹理信息与高级语义信息进行聚合。再针对复杂场景中高级语义信息不足的问题,提出一个特征协同增强模块(feature synergy enhancement module, FSEM),通过增强分支之间特征信息的融合提取更深层次的高级语义信息。网络采用双流-单支架构,在双阶段进行差异处理,结合浅层 FDFM 模块捕获细节差异与深层 MDFM 模块进行跨尺度特征融合;并使用卷积块注意模块(convolutional block attention module, CBAM)在特征融合阶段引入 CBAM 模块,自适应加权重要变化区域。实验结果表明,MSAF-Net 在 LEVIR-CD, WHU-CD 和 SYSU-CD 数据集上分别取得了 90.71%, 91.71% 和 82.71% 的 F1 值,与当前的主流方法相比分别提升了 0.14、1.2 和 0.87 个百分点。

关键词:遥感影像;变化检测;双注意力机制

中图分类号: TN911.73; TP751.1 **文献标识码:** A **国家标准学科分类代码:** 510.40

Change detection network based on dual-branch feature fusion and attention mechanism

Zhou Xianchun^{1,2,3} Ni Hongjun² Li Hanzhe³

(1. Research on Application Technology of Metaverse in Cultural Tourism Scenarios-Key Laboratory of Culture and Tourism of Jiangsu Province, Nanjing 210044, China; 2. School of Electronic and Information Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China; 3. School of Artificial Intelligence, Nanjing University of Information Science and Technology, Nanjing 210044, China)

Abstract: In recent years, deep learning has demonstrated excellent performance in the field of change detection. However, there are still shortcomings such as unclear fine-grained boundaries and internal voids caused by the loss of detailed information and insufficient semantic information in terms of false detections and missed detections. To address this issue, this paper proposes a new multi-scale adaptive fusion network (MSAF-Net), which uses ResNet as the backbone network. Firstly, it presents a feature difference fusion module (FDFM), which can roughly extract boundary information from changed regions through simple operations. Then, to tackle the problem of lost detailed texture information, a multi-scale dual-feature fusion module (MDFM) is proposed to aggregate low-level texture information and high-level semantic information from multiple receptive fields. Furthermore, to solve the problem of insufficient high-level semantic information in complex scenes, a feature synergy enhancement module (FSEM) is put forward to extract deeper high-level semantic information by enhancing the fusion of feature information between branches. The integration of the FDFM, MDFM, and FSEM modules enables efficient and accurate pixel-level change recognition. The network adopts a dual-stream-single-branch architecture and improves performance through the following key technologies: performing difference processing in two stages, combining the shallow

收稿日期:2025-09-01 Received Date: 2025-09-01

* 基金项目:国家自然科学基金(62572250)项目资助

FDFM module to capture detailed differences and the deep MDFM module to conduct cross-scale feature fusion; and introducing the convolutional block attention module (CBAM), in the feature fusion stage to adaptively weight important changed regions. The experimental results show that MSAF-Net achieved F1 values of 90.71%, 91.71%, and 82.71% on the LEVIR-CD, WHU-CD, and SYSU-CD datasets, respectively. Compared with the current mainstream methods, the performance has improved by 0.14, 1.2, and 0.87 percentage points, respectively.

Keywords: remote sensing images; change detection; dual-attention

0 引言

变化检测是通过对对比分析多时相遥感影像数据,识别地表覆盖变化的遥感核心技术^[1]。该技术方法历经4个发展阶段:1970年代基于像素差异分析,1990年代引入机器学习,2000年代结合高分辨率数据,2010年后深度学习成为主流。核心算法包括时间序列分割、光谱向量分析和卷积神经网络等,应用场景覆盖城市扩张监测、森林扰动识别及农作物变迁分析。当前面临样本制作复杂、多源数据异构性等技术挑战,未来发展方向聚焦弱监督学习与自适应模型构建。变化检测在监测和分析特定地理区域的时间变化方面起着至关重要的作用。随着遥感技术的进步,遥感卫星的成像能力和成像质量有了很大的提高。高分辨率图像提供了有关地球表面的重要信息,具有广泛的应用,包括环境监测^[2]、城市规划^[3]、灾害管理^[4]和自然资源管理^[5]。然而,由于各种因素,如角度、光照、季节变化等,基于遥感图像的变化检测是一项具有挑战性的任务。

近年来,深度学习技术的快速发展已在各个领域展现出其优越性^[6-7]。深度学习强大的特征学习和泛化能力,也使其适用于基于多源、多模态和多时相等多种遥感图像的检测任务^[8-9]。基于卷积神经网络(convolutional neural network, CNN)^[10]、递归卷积神经网络(recurrent neural networks, RNNs)^[11]、图卷积神经网络(graph convolution neural networks, GCNNs)^[12-13]以及基于Transformer的方法^[14],是当前变化检测技术中常用的方法。它们在大规模遥感数据的自动特征学习和高效处理中发挥着关键作用。与传统的变化检测方法相比,深度学习方法具有更高的准确性和鲁棒性。深度学习模型能够自动学习数据的特征表示,无需手动提取特征。包括MFIN^[15]和EGCTNet^[16]等在内的近期研究,强调了边缘信息在增强变化区域边界细粒度方面的重要性。这些研究成功地将边缘特征作为先验知识融入到检测网络中。边缘通常是图像中物体或场景之间的边界或边缘轮廓。变化区域是指在2个不同时期的遥感图像中发生变化的区域。边缘信息在变化检测中起着重要作用,因为变化通常会导致物体或场景边界的改变。通过提取和分析边缘信息,有助于识别和定位变化区域。边缘信息可以提供关于变化区域的形状、位置和边界特征的线索。因此,

准确提取和利用边缘信息有助于提高变化检测的性能,减少误检和漏检。然而,大多数方法往往忽略了边缘信息与变化信息之间的关系,仅将提取的边缘信息用作预测结果的补充,忽略了边缘信息在描述变化区域(如形状、位置和边界特征)方面的作用。此外,现有方法总是将从标签中提取的变化边界标签作为额外的输入或监督,以帮助提取边界特征,这可能会导致变化区域的一些内部特征丢失,并降低模型对变化特征的提取能力。

对于变化检测任务而言,目前主要面临2大挑战:一是由于双时相图像中光照条件差异和噪声导致的伪变化所引起的误检问题。二是漏检问题,比如因细节信息丢失和语义信息不足而导致的细粒度边界不够清晰以及存在内部空洞等情况。为更全面地利用边缘信息解决上述问题,本文提出了一种基于变化边界感知引导的检测网络。

本文提出了一种基于双分支特征融合与注意力机制的变化检测网络MSAF-Net,无需使用边缘提取算法和边界标签来辅助提取边界特征。仅通过简单的操作和常规标签,就可以激活边界特征,并引导模型取得更好的效果。针对边缘信息利用不充分以及细粒度检测边界不够清晰的问题,提出了一个FDFM模块,该模块仅通过简单的操作就能从变化区域中粗略提取边界信息。利用来自常规标签的准确变化区域信息对所获得的边界信息进行细化,确保变化区域边界的精确提取。该模块提高了边界的粒度,减轻了伪变化的影响,并为其他分支的信息聚合提供指导。针对细节纹理信息丢失的问题,提出了MDFM,将多个感受野中的低级纹理信息与高级语义信息进行聚合。该模块增强了变化区域内部的完整性,并弥补了更深层次特征中细节信息的丢失。针对复杂场景中高级语义信息不足的问题,提出了一个FSEM,通过增强分支之间特征信息的融合来提取更深层次的高级语义信息。该模块增强了网络的特征提取能力,加强了分支之间的特征交互,从而提高了解耦的效率。

1 相关工作

1.1 基于深度学习的变化检测方法

深度学习网络具有分层架构。这一特点使得孪生网络架构被应用于变化检测任务中。在这些架构中加入跳

跃连接可提高预测精度。例如,带有跳跃连接的全卷积网络(FC-EF、FC-Siam-Diff 和 FC-Siam-Conc)分别通过时空特征、特征差分特征和特征拼接等方法提取深层信息并进行预测。Fang 等^[17]提出了一种集成了密集跳跃连接的孪生网络(SNUNet),通过整合多尺度特征缓解了深层信息和局部信息的丢失问题。一些研究提出了不同的模型来提升变化检测性能。例如,Jiang 等^[18]提出了一种联合学习框架(SSANet),该框架结合了融合分支和差异提取分支,以增强上下文信息的聚合。Wen 等^[19]开发了 RS-CADM 模型,该模型结合了去噪扩散概率模型(DDPMs)和自适应校准技术,以增强对变化信息的提取。双任务约束的卷积网络模型^[20]也被提出,用于同时完成变化检测和语义分割任务。

此外,研究人员单独或联合使用卷积神经网络(CNNs)和变换器(transformers)。例如,Badhe N B 等^[21]提出了一种双时相图像变换器网络(BIT),该网络能够捕捉并对 2 个不同时相图像之间的时空上下文信息进行建模,从而理解局部内容与全局场景之间的相关性。文献[22-24]将卷积神经网络和变换器用作特征提取网络(ICIFNet、WNet 和 PASSNet),这些网络促进了局部提取的卷积神经网络特征与全局提取的变换器特征之间的交互,在保留各自特征的同时增强了细节信息。

1.2 注意力机制辅助的变化检测方法

注意力机制^[25]在包括图像融合^[26-27]、目标检测^[28-29]以及语义分割^[30]等各种视觉任务中受到了广泛关注。在变化检测(CD)任务中,传统方法主要侧重于像素级别的比较,忽略了像素之间的语义关系和上下文信息,从而导致检测精度较低。然而,注意力机制具有捕捉输入序列中位置关系的能力,这为遥感影像的变化检测带来了新的契机,它能够从全局角度有效地捕捉全局关系,并将有限的“注意力”集中在重要位置上,从而获取更有用的信息。常见的基本注意力机制有 3 种,即自注意力机制、空间注意力机制(SAM)和通道注意力机制(CAM)。目前,许多变化检测网络都是基于基本注意力机制的增强版本,以进一步突出变化特征。例如,SNUNet 提出了增强通道注意力机制(ECAM),用于自动选择并聚焦于不同组之间更有效的信息。DSIFNet^[31]交替使用通道注意力机制和空间注意力机制,从通道-空间维度对融合特征进行优化。交叉注意力机制也在变化检测中得到了广泛应用,它基于自注意力机制来考虑不同输入序列之间的关系。研究人员探索了将自注意力机制和交叉注意力机制与卷积神经网络(CNNs)相结合的方法,以便提取更准确的特征。HANet^[32]设计了一个 HAN 模块,该模块能够分别从列和行维度捕捉长距离依赖关系。DMINet^[33]将自注意力机制和交叉注意力机制整合到一个模块中,获得用于信息交互的全局注意力分布。受自注意力

机制的启发,MFIN 设计了特征交互模块,以交互方式处理 2 张图像的特征信息。SARASNet^[34]利用自注意力和交叉注意力特征,有效地检测由不同尺度物体引起的变化。HDANet^[35]引入了多分辨率并行结构,并设计了一种差分注意力模块,以保留空间信息和变化信息。此外,空间注意力机制和通道注意力机制专注于单一维度,可能缺乏对特征的细粒度关注。受神经科学理论的启发,Yang 等^[36]设计了一种无参数的三维注意力机制(SimAM),该机制通过优化能量函数来确定每个神经元的重要性,从而实现更准确的特征聚焦。woo 等^[37]提出了卷积块注意力模块(CBAM),这是一个简单而有效的前馈卷积神经网络注意力模块。给定一个中间特征图,模块按顺序推断注意力图沿着通道和空间将注意力图乘以输入特征图而进行自适应特征细化。

1.3 边缘辅助的变化检测方法

精细化的变化检测任务是当前的一个热门话题,而变化检测的细粒度边界问题仍是一个有待解决的难题。为了进一步缓解这一问题,许多领域专家针对边界特征的提取与保留开展了相关工作。常见的边缘提取算法包括 Canny 算法、Sobel 算法和拉普拉斯(Laplacian)算法。

有研究设计了一个细节特征引导模块,该模块使用拉普拉斯算子从标签中提取多尺度边缘特征,并将其插入到主干网络中以学习细节特征。

另一种方法是增强变化检测网络对边缘特征的保留能力。提出了边缘引导的变化图估计方法,该方法不仅可以预测已变化的建筑物及其边缘,还能将边缘特征直接整合到判别性特征中,从而进一步提高预测结果的质量。

2 本文方法的提出

2.1 MSAF-Net 的整体结构

如图 1 所示,本文提出的 MSAF-Net 基于编码器解码器架构,由 4 个部分组成,分别是多级特征提取模块、特征差异检测模块(FDFM)、多尺度双特征融合模块(MDFM)和特征协同增强模块(FSEM)。采用孪生网络的方法,编码器使用预训练的卷积神经网络模型(ResNet18)作为特征提取网络。解码器由 FDFM、MDFM 和 FSEM 3 个分支组成。

首先,将 2 个不同时期的遥感图像输入到特征提取网络中,得到不同尺度的特征向量 $\{F_n, f_n, n \in (1, 2, 3)\}$,如图 1 所示。设 FDFM、MDFM 和 FSEM 的输出特征向量分别为 $\{f_{fdm}^i, f_{mdm}^i, f_{fem}^i, i \in (1, 2, 3)\}$ 。将特征向量 F_3, f_3 输入到 FDFM 中,以获得边缘增强的差异特征 f_{fdm}^1 。随后,特征 $\{F_n, f_n, (n = 1, 2, 3), f_{fdm}^{i-1}\}$ 被

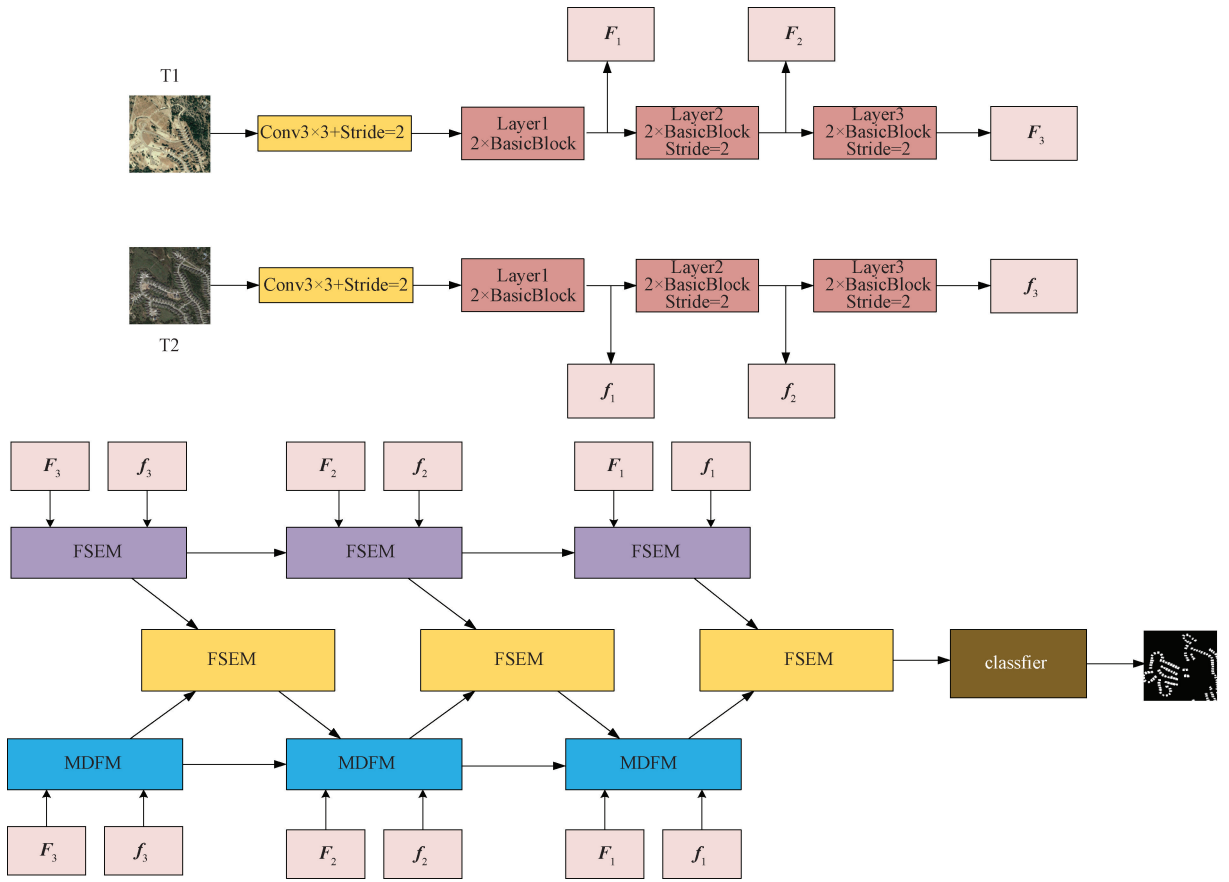


图1 MSAF-Net 网络框架

Fig. 1 Architecture of the proposed MSAF-Net

用作 FDFM 的输入,从而得到 $\{f_{fdm}^i, (i = 2, 3)\}$ 。通过 FDFM 提取的边缘增强差异特征对边缘信息具有更强的表达能力。这能够促进变化边缘信息的保留,并更好地表征变化区域。最后, FDFM 分支通过分类器得到额外的预测结果。由于 FDFM 分支中边缘特征的表达更强,额外的预测会更侧重于边缘预测。此外,它保留了变化区域的一些内部信息,与仅保留边缘信息相比,减少了有价值信息的损失。对额外的预测结果进行监督,通过损失函数计算变化区域信息与真实标签之间的差异。基于这种监督对网络进行反向传播,并更新生成的变化区域信息梯度。在训练阶段通过迭代优化,逐渐改进生成的结果,增强定位变化区域和表达变化区域特征的能力。这促使 FDFM 分支关注更准确的变化特征,引导其他分支更好地分离变化区域信息,并加深整个网络的分离能力。将 F_3, f_3 输入到 MDFM 中,以获得多感受野的纹理细节聚合特征 f_{mdfm}^1 。而用于 MDFM 的输入特征 $\{f_{mdfm}^i, i \in (2, 3)\}$ 包含特征 $\{F_n, f_n, (n = 1, 2, 3), f_{mdfm}^{i-1}, f_{mdfm}^{i-1}\}$ 。利用高级语义特征引导低级信息向高级信息的转变。对 FSEM 模块输入特征 $\{f_{fsem}^i, f_{mdfm}^i, i \in (1,$

$2, 3)\}$ 以获得高级变化语义特征 $\{f_{fsem}^i, i \in (1, 2, 3)\}$,并增强网络的特征提取能力。

2.2 FDFM 模块

边缘信息与特征差异运算的结合能够有效地描绘出变化区域的边界和轮廓。为了实现这一点,本文提出了一种 FDFM,它利用边界信息获取变化区域,并有助于保留变化区域的边界信息。该模块的结构如图 2 所示。边缘信息作为有价值的先验知识,通过捕捉物体的精细结构和复杂形状来辅助变化检测。然而,不加区分地提取边缘信息可能会在非变化区域引入不相关的信息。相反,仅考虑变化区域的边缘信息可能会导致内部特征的丢失。为了缓解这一问题,本文在进行特征差异计算之前,提取并增强双时相特征的边界信息。这种方法能够通过特征差异捕捉边界信息的差异以及差异特征。在本文方法中,将来自特征提取网络的特征 (F_i, f_i) 输入到 FDFM 中,以提取并增强双时相特征的边缘信息。

首先,使用 CBAM 注意力机制对第一级输入特征进行预注意力处理,以识别值得关注的区域。CBAM 注意

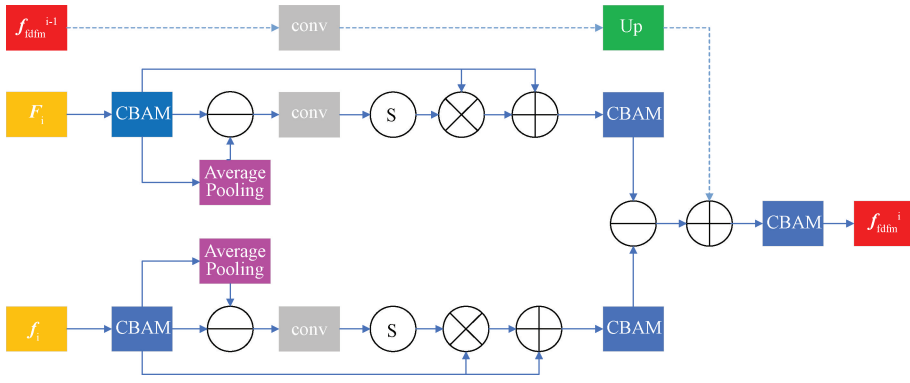


图 2 FDFM 模块

Fig. 2 Details of the FDFM

力机制是一种结合通道注意力和空间注意力的轻量级模块,由 Woo 等^[37]于 2018 年提出,它通过结合通道注意力(channel attention)和空间注意力(spatial attention)来增强 CNN 的特征表达能力,帮助模型更高效地聚焦于图像中的重要区域和通道。随后,通过池化、减法和卷积操作的组合提取边缘特征。这一设计基于边缘区域通常表现出较高梯度值的特性,而平均池化操作通过抑制高频噪声来实现特征的平滑化。减法操作通过从原始特征中减去平均池化后的特征,显式地突出了边缘区域。接着,这些中间特征通过一个 1×1 卷积层和 Sigmoid 激活函数进行处理,以增强特征的对比度和显著性。为了进一步优化边缘特征的提取,系统通过乘法和加法操作对特征进行重新增强。最后,采用 CBAM 注意力模块对显著区域进行空间-通道重校准,动态加权重要区域,从而生成最终的特征表示 $\hat{f}_e$ 。上述操作的计算流程由式(1)~(3)定义。

$$\hat{f} = \text{CBAM}(f_i) \quad (1)$$

$$f_e^t = \text{Conv}_{1 \times 1}((\hat{f} - \text{AP}(\hat{f}))) \times \hat{f} + \hat{f} \quad (2)$$

$$\hat{f}_e^t = \text{CBAM}(f_e^t) \quad (3)$$

式中:CBAM 表示注意力机制;AP 表示 3×3 平均池化层, $\text{Conv}_{1 \times 1}$ 表示使用 Sigmoid 作为激活函数的卷积块(1×1 卷积、批量归一化 BN、Sigmoid 激活函数)。特征差异作为一种直接计算差异特征的操作,容易受到噪声和信息丢失的影响。在变化检测任务中,由于双时相特征固有的相似性,特征差异可能导致变化区域的边缘模糊。为了缓解这些问题,本文在计算特征差异之前对双时相特征的边缘进行增强。这种边缘增强的特征差异提供了更丰富的上下文信息,并有助于准确定位变化区域。对不同层次的特征重复此操作,获得边缘增强的特征并计算特征差异。在每个层次上重新计算值 $(\hat{f}_e^t, \hat{f}_e^c)$ 。随后,利用 CBAM 对边缘增强的特征信息进行进一步处理。相关公式如式(4)所示。

$$f_{fdm}^i =$$

$$\begin{cases} \text{CBAM}(\text{Abs}(\hat{f}_e^t - \hat{f}_e^c)) & i = 1 \\ \text{CBAM}(\text{Abs}(\hat{f}_e^t - \hat{f}_e^c) + \text{Up}(\text{Conv}_{1 \times 1}(f_{fdm}^{i-1}))) & i = 2, 3 \end{cases} \quad (4)$$

式中: i 表示不同深度的特征;Abs 表示取绝对值操作;Up 表示上采样操作。

2.3 MDFM 模块

低层次纹理信息包含图像的基本细节和局部特征,例如纹理、颜色和灰度。这些信息能够捕捉图像中的微小变化。另一方面,双时相特征由于其固有特性,在保留丰富细节的同时也维持了空间关系。为了增强变化检测的全面性,本文提出了一种多尺度双特征融合模块(MDFM),该模块逐步结合低层次细节特征和多层次特征。该过程的示意图如图 3 所示。为了提取细节特征并保持空间关系,本文对输入的双时相特征(F_i, f_i)进行通道拼接。先利用 1 个深度 3×3 卷积,再利用 3 个尺度的深度卷积进行处理。这种方法通过不同的感受野捕捉不同尺度的变化区域,同时利用分组保持特征的空间完整性。随后,将 4 个卷积输出在通道维度上进行拼接,并使用 1×1 卷积块进行通道下采样。此步骤进一步通过 CBAM 注意力机制对特征进行优化。相关公式如式(5)~(10)所示:

$$f = \text{DWConv}_{3 \times 3}(\text{Concat}(f_i, F_i)) \quad (5)$$

$$f^{(1)} = \text{DWConv}_{5 \times 5}(f) \quad (6)$$

$$f^{(2)} = \text{DWConv}_{7 \times 7}(f) \quad (7)$$

$$f^{(3)} = \text{DWConv}_{9 \times 9}(f) \quad (8)$$

$$f_{cat} = \text{Conv}_{1 \times 1}(f^{(1)} \oplus f^{(2)} \oplus f^{(3)} \oplus f) \quad (9)$$

$$\hat{f}_{cat} = \text{CBAM}(f_{cat}) \quad (10)$$

式中:Concat 表示通道拼接操作;DWConv 为深度卷积。考虑到双时相特征中存在的共享特性,MDFM 提供了

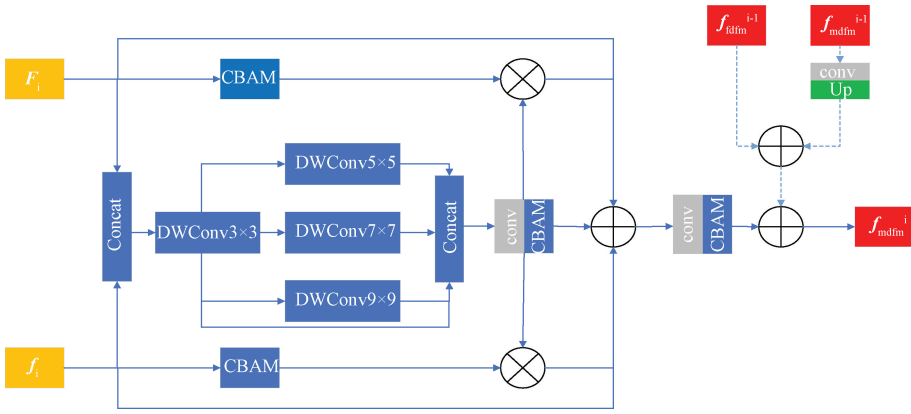


图 3 MDFM 模块

Fig. 3 Details of the MDFM

更精确的低层次纹理信息。利用 CBAM 注意力机制分别计算来自不同时相特征的像素级特征的重要性。然后,将提取的共性特征($\tilde{f}_{cat}$)与各自的时相特征(F_i , f_i)相乘,以获得它们的相似性。相关公式如式(11)和(12)所示。

$$\tilde{F}_i = \text{CBAM}(F_i) \times \tilde{f}_{cat} \quad (11)$$

$$\tilde{f}_{mdfm}^i = \begin{cases} \text{CBAM}(\text{Conv}_{3 \times 3}(\hat{F}_i + \hat{f}_i + \hat{f}_{cat})), & i = 1 \\ \text{CBAM}(\text{Conv}_{3 \times 3}(\hat{F}_i + \hat{f}_i + \hat{f}_{cat} + \tilde{f}_{fsem}^{i-1}) + \text{Up}(\text{Conv}_{1 \times 1}(\tilde{f}_{mdfm}^{i-1}))), & i = 2, 3 \end{cases}$$

最后,通过 MDFM 对多层次特征进行整合。这一整合的目的是通过结合多感受野来增强变化区域内的空间纹理信息。此外,本文将这些特征与高层次语义信息相结合,以获得对变化更细致和全面的理解。

2.4 FSEM 模块

高级语义特征在准确地定位和识别变化区域方面起着至关重要的作用。另一方面,详细的纹理特征则提供更为精确的边界和纹理信息。如图 4 所示,本文所提出的 FSEM 旨在将二者结合起来,以学习更深入的高级语义特征。

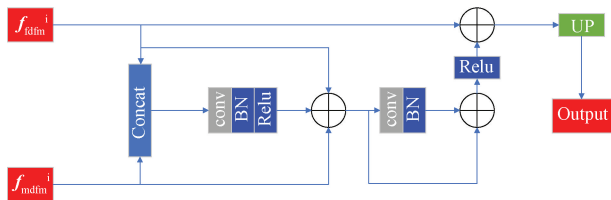


图 4 FSEM 模块

Fig. 4 Details of the FSEM

特征协同增强模块整合了从变化检测模块(FDFM)获取的变化区域信息以及从边界特征聚合模块(MDFM)获取的空间细节纹理信息。分别对来自上层变化检测模

$$\tilde{f}_i = \text{CBAM}(f_i) \times \tilde{f}_{cat} \quad (12)$$

接下来将低层次细节特征($\tilde{F}$, $\tilde{f}$, $\tilde{f}_{cat}$)与高层次变化特征($\tilde{f}_{fsem}^{i-1}$)进行求和。通过 CBAM 注意力机制,最终得到聚合后的特征。随着 CNN 模型的加深,低层次信息会逐渐丢失。为了确保即使在网络的深层也能保留足够的纹理信息,本文引入了残差连接。相关公式如式(13)。

$$i = 1 \quad (13)$$

块和边界特征聚合模块的输入特征进行拼接与求和操作。为了减少计算量,采用 1×1 卷积将通道数量减半。此外,为了保持信息的完整性,使用残差拼接将合并后的信息与之相乘。

接下来使用 3×3 卷积块提取深度特征。为了将信息损失降至最低,在应用 ReLU 激活函数之前选择采用残差拼接。

最后,将前一层变化检测模块的输出特征与深度特征相加。随后,变化边界的完整特征和高级语义特征被传入边界特征聚合模块。这确保了变化边界和高级特征能够受到聚合模块的引导。同时,通过模块内的多次残差操作,该方法将特征信息损失降至最低。具体公式如下:

$$\tilde{f}_c^i = \text{Conv}_{1 \times 1}(\text{Concat}(\tilde{f}_{mdfm}^i, \tilde{f}_{fdm}^i)) + (\tilde{f}_{mdfm}^i + \tilde{f}_{fdm}^i) \quad (14)$$

$$\tilde{f}_{fsem}^i = \text{Relu}(\text{BN}(\text{Conv}_{3 \times 3}(\tilde{f}_c^i)) + \tilde{f}_c^i) + \tilde{f}_{fdm}^i, i = 1, 2, 3 \quad (15)$$

$$\text{Output} = \text{Conv}_{3 \times 3}(\text{Relu}(\text{BN}(\text{Conv}_{3 \times 3}(\tilde{f}_{fsem}^3)))) \quad (16)$$

式中; i 表示不同深度的特征;Conv、BN 和 ReLU 分别表示卷积层、批量归一化层和 ReLU 激活函数; $\tilde{f}_c^i$ 是通道维

度变换的中间操作,而 f_{sem}^s 是提取的高层次特征。

2.5 损失函数

在变化检测领域,未变化像素与变化像素的数量存在显著不平衡。为了缓解样本不平衡的影响,本文采用混合损失函数,即加权交叉熵损失与 Dice 损失的组合,其定义为:

$$L = L_{wce} + L_{dice} \quad (17)$$

$$L_{wce} = \frac{1}{W \times H} \sum_{i=1}^{H \times W} weight[class] \cdot$$

$$\left(\lg \left(\frac{\exp(\hat{y}[i][class])}{\sum_{l=0}^1 \exp(\hat{y}[i][l])} \right) \right) \quad (18)$$

$$L_{dice} = 1 - \frac{2 \cdot S_o(\hat{Y})}{Y + S_o(\hat{Y})} \quad (19)$$

式中: L_{wce} 表示加权交叉熵损失; L_{dice} 表示 DICE 损失; $weight$ 表示权重; $class$ 的值为 157 或 0 (分别对应变化像素和未变化像素); $\hat{y}[i]$ 表示 Y 中的第 i 个点, i 和 l 是索引, $\hat{Y} = \{\hat{y}[i], i = 1, 2, \dots, W \times H\}$ 表示变化图,而 Y 表示真实值(标注真值)。

本文的网络架构由 FDFM 引导进行变化检测。因此, FDFM 分支的准确性会直接影响后续深度变化特征的提取。为了缓解这一问题,本文采用 FSEM 分支进行特征提取,并通过预测头生成预测结果。

3 实验与结果

3.1 数据集选择

1) LEVIR-CD, 该数据集由 637 个超高分辨率 (VHR, 0.5 m/pixel) 谷歌地球 (GE) 图像块对组成, 大小为 1 024×1 024 pixels。这些时间跨度为 5~14 a 的双时态图像具有显着的土地利用变化, 尤其是建筑增长。LEVIR-CD 涵盖别墅住宅、高层公寓、小型车库、大型仓库等各类建筑。本文仅关注与建筑相关的变化, 包括建筑增长(从土壤/草地/硬化地面或在建建筑到新的建筑区域的变化)和建筑衰退。这些双时态图像由遥感图像解译专家使用二进制标签(1 表示变化, 0 表示不变)进行注释。每个样本都由一个人进行注释, 然后由另一个人进行检查以生成高质量的注释。

2) WHU-CD, 该数据集主要关注 2012 年新西兰基督城地震后的建筑损坏分析。数据集由一对高分辨率 (0.2 m/pixel) 遥感图像组成, 原始图像大小为 32 507×15 354 pixels。实验中将图像分割成 256×256 pixels 大小的图像块, 分割过程中没有重复区域。训练、验证和测试的比例分别为 8:1:1。最终获得 6 096 对图像的训练集、762 对图像的验证集和 762 对图像的测试集。

3) SYSU-CD, 本文研究中使用的 SYSU 数据集由 20 000 对高分辨率 (0.5 m/pixel) 遥感图像组成。每幅图像的大小为 256×256 pixels。数据集包括香港地区的高层建筑、港口和郊区变化等地面物体。在训练过程中, 本文使用研究人员提供的标准数据分割。训练、验证和测试的比例分别为 6:2:2。最终获得 12 000 对图像的训练集、4 000 对图像的验证集和 4 000 对图像的测试集。

3.2 评估指标

使用 PyTorch 框架构建模型, 并在 8 GB 显存的 NVIDIA GeForce RTX4060 Ti GPU 上进行训练。AdamW 优化器用于最小化损失, 权重衰减设置为 1×10^{-4} 。初始学习率 ($initial_{lr}$) 为 1×10^{-4} , 步长为 8, 批量大小为 8, 150 个 epoch, 直到收敛。为了评估所提出算法的有效性, 实验重点关注精确度 (P_r)、召回率 (R_c)、F1 分数、总体准确度 (OA) 和交并比 (IoU) 等指标来评估变化检测。高精度可减少误检, 而高召回率可最大限度地减少漏检。值越高表示性能越好, 计算公式如式 (20)~(24) 所示。

$$Pr = \frac{TP}{TP + FP} \quad (20)$$

$$Rc = \frac{TP}{TP + FN} \quad (21)$$

$$F_1 = 2 \times \frac{P_r \times R_c}{P_r + P_c} \quad (22)$$

$$IOU = \frac{TP}{TP + FN + FP} \quad (23)$$

$$OA = \frac{TP + TN}{TP + TN + FN + FP} \quad (24)$$

式中: TP 表示真阳性样本数; TN 表示真阴性样本数; FN 表示假阴性样本数; FP 表示假阳性样本数。

3.3 算法对比

通过与 8 种先进的方法进行定性和定量的比较验证所提方法的有效性。这些方法涵盖了不同的技术架构, 包括纯基于 CNN 的模型: FC-EF、FC-Siam-Di、FC-Siam-Conc。结合 CNN 与注意力机制的模型: DTCDCN、SNUNet、DMINet。基于 Transformer 的模型: BIT、STANet。

1) FC-EF, 将双时相遥感图像连接起来, 然后输入到卷积神经网络进行变化检测。

2) FC-Siam-conc, 提取双时相图像的多层特征, 并通过编码流中的两个连接进行拼接以检测变化。

3) FC-Siam-diff, 利用编码流中 2 个连接的绝对差异进行特征级融合, 以实现双时相图像的变化检测。

4) STANet, 一种基于 Siamese FCN 的变化检测方法, 利用金字塔时空注意力模块来捕捉不同尺度的时空依赖关系, 从而生成更好的表示, 以适应各种大小的目标。

5) SNUNet, 多尺度特征拼接方法结合了 Siamese 网络和 NestedUNet, 用于提取多尺度特征, 并在解码器的每

一层应用通道注意力模块,以增强特征表示。

6) BIT, 提出双时相图像变换器 (bi-temporal image transformer), 用于将高层次概念的语义变化编码为紧凑的语义符号, 从而有效地建模双时相图像中的遥感上下文。

7) DTCDCSCN, 一种基于双重注意力机制和全卷积网络的模型, 添加了空间注意力和通道注意力模块, 以提取更具判别力的特征, 增强变化检测效果。

8) DMINet, 一种孪生网络, 在获得差异特征之前交互双时相特征, 结合自注意力和交叉注意力引导每个输入的全局特征分布。促进内在层次表示之间的信息耦合。

表 1~3 分别给出了在 LEVIR-CD 数据集、WHU-CD 数据集和 SYSU-CD 数据集上的定量实验结果。

表 1 LEVIR-CD 数据集上各种方法的定量比较结果
Table 1 Results of the quantitative comparison of the various methods applied on the LEVIR-CD (%)

方法	P_r	R_c	F1	IoU	OA
FC-EF	80.79	75.67	78.40	65.46	98.01
FC-Siam-Diff	87.55	81.23	84.32	70.97	98.35
FC-Siam-Conc	91.99	76.77	83.69	71.96	98.49
STANet	81.48	86.70	83.14	72.50	98.11
SNUNet	91.12	88.43	89.91	81.63	98.99
BIT	91.51	89.37	90.12	81.60	98.98
DTCDCSCN	88.42	86.56	87.79	78.15	98.77
DMINet	92.12	89.29	90.57	82.80	99.07
本文	91.23	90.55	90.71	83.90	99.13

表 2 WHU-CD 数据集上应用的各种方法的定量比较结果
Table 2 Results of the quantitative comparison of the various methods applied on the WHU-CD (%)

方法	P_r	R_c	F1	IoU	OA
FC-EF	70.05	73.24	72.24	60.14	97.51
FC-Siam-Diff	73.53	76.64	75.54	60.78	98.10
FC-Siam-Conc	72.67	75.45	74.10	61.04	98.03
STANet	71.34	67.89	69.23	65.32	97.83
SNUNet	85.55	81.63	83.71	71.23	98.69
BIT	89.64	85.36	88.13	80.01	98.78
DTCDCSCN	85.62	82.47	71.95	67.15	97.42
DMINet	93.81	88.29	90.51	85.28	99.27
本文	93.53	90.55	91.71	85.90	99.42

由表 1 可知, 在 LEVIR-CD 数据集上的 F1、IOU 指标分别达到 90.71%、83.90%, 相较于 FC-EF、FC-Sima-Diff、FC-sima-Conc、DTCDCSCN、SNUNet、BIT、DMINet 方法, F1 值分别提高 12.31、6.39、7.02、3.44、2.95、1.80、4.34 个百分点; OA 指标达到 99.13%, 比第 2 名的

BIT(98.98%)高出 0.15 个百分点。由表 2 可知, 本文方法在 F1 分数(91.71%)、IoU(85.90%)和 OA(99.42%)3 个关键指标上均取得最优结果, R_c (90.55%)显著领先其他方法, P_r (93.53%)仅次于 DMINet(93.81%), 差异仅为 0.28 个百分点。由表 2 可知, 本文方法在 WHU-CD 数据集上实现了全面而均衡的性能提升, 特别是在变化检测的完整性和定位准确性方面表现突出。虽然 P_r 略低于 DMINet, 但在 R_c 、F1 等具有明显优势, 说明本文的方法更适合要求全面检测变化区域的应用场景。表 3 为 SYSU-CD 数据集上的定量比较结果。在各模型整体表现较差的情况下, 本文方法仍在多个指标取得最优结果, IoU 达到 71.80%, 较次优方法 DMINet(69.28%)提升 2.52 个百分点, 表明模型对变化区域的空间定位更为精准, 尤其在边界细节捕捉上优势明显。 R_c (83.53%)较 DMINet(81.29%)提升 2.24 个百分点, 说明模型对真实变化区域的漏检率更低, 能够有效捕获小尺度或弱纹理变化。F1 分数(82.71%)和 OA(91.92%)均为最优, 平衡了精确率与召回率的矛盾, 且在整体场景分类中表现出更强的鲁棒性。相较于传统双流网络(如 FC-Siam-Conc), F1 分数提升 5.29 个百分点, OA 提升 2.89 个百分点, 验证了双分支特征融合与注意力机制在跨尺度特征整合中的有效性。

表 3 SYSU-CD 数据集上应用的各种方法的定量比较结果
Table 3 Results of the quantitative comparison of the various methods applied on the SYSU-CD (%)

方法	P_r	R_c	F1	IoU	OA
FC-EF	72.15	77.42	75.29	61.03	87.36
FC-Siam-Diff	78.68	70.34	72.84	61.58	87.83
FC-Siam-Conc	72.27	80.15	77.42	62.44	89.03
STANet	75.72	79.49	73.53	62.32	87.81
SNUNet	77.32	80.11	78.69	64.87	89.77
BIT	79.74	76.45	78.23	64.47	89.98
DTCDCSCN	78.12	80.17	79.95	67.15	89.92
DMINet	82.21	81.29	81.84	69.28	91.27
本文	81.53	83.53	82.71	71.80	91.92

不同方法在 LEVIR-CD、WHU-CD 和 DSIFN-CD 数据集上的可视化效果如图 4 所示。为了更好地观察, 使用不同的黑白虚线实线表示 TP(白色)、TN(黑色)、FP(虚线)和 FN(实线)。本文的方法 MSAF-Net 与其他方法相比表现出色, 有效区分了一些细微的差异, 减少了漏检。

图 5 所示为不同方法在 LEVIR-CD 数据集上的结果。这些建筑的分布较为密集, 且有部分存在粘连部分, 使边缘信息更难区分, 更容易产生误检。与其他方法相比, 本文的 MSAF-Net 网络取得了较好的预测结果, 通过 MDFM 模块和 FDFM, 通过使用双注意力机制双边特征融合的方法和特征差异检测模块, 捕捉更多的细节在整

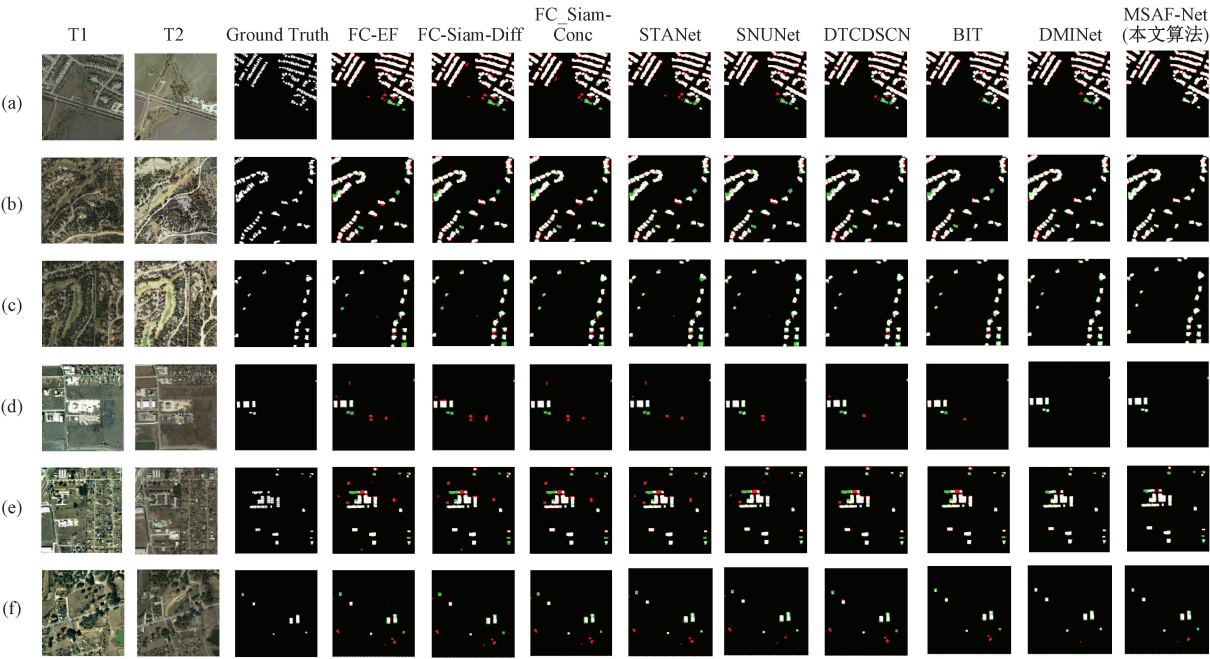


图 5 MSAF-Net 与其他方法在 LEVIR-CD 上的预测结果比较

Fig. 5 Comparison of prediction results of MSAF-Net with other methods on LEVIR-CD

体上的产生了更少的检测错误和更准确的边缘预测。

图 6 所示为不同方法在 WHU-CD 数据集上的结果。该数据集的建筑物的颜色、纹理或形状特征相似,尤其是地面上的白色物体、停车场、路面等区域,这容易导致模型错误地将这些区域判断为建筑物,更容易产生误检。

本文的 MSAF-Net 网络通过 FBFFM 模块和 FDFM,通过使用双注意力机制双边特征融合的方法和特征差异检测模块,将浅层信息和中层信息融合后通过 FBFFM 模块融合成深度信息,更好的对图像进行预测,体现了在复杂场景的应对结果。

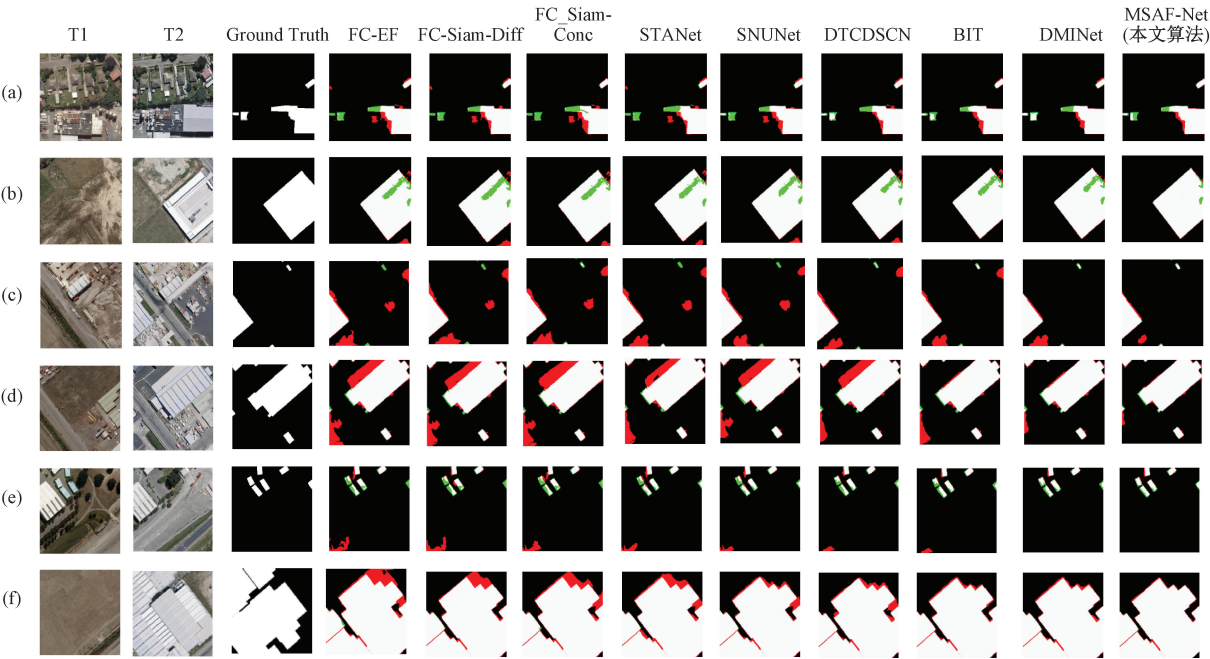


图 6 MSAF-Net 与其他方法在 WHU-CD 上的预测结果比较

Fig. 6 Comparison of prediction results of MSAF-Net with other methods on WHU-CD

图 7 所示为不同方法在 SYSU-CD 数据集上的结果。该数据集的场景更加丰富,有较多的植被变化,预测图更加复杂,各个方法的误检和漏检现象更加明显。尽管如

此,本文的 MSAF-Net 仍取得较好的结果,体现了 MSAF-Net 网络应对复杂场景下的鲁棒性。方法更好地处理仅颜色上的变化,取得了卓越的性能。

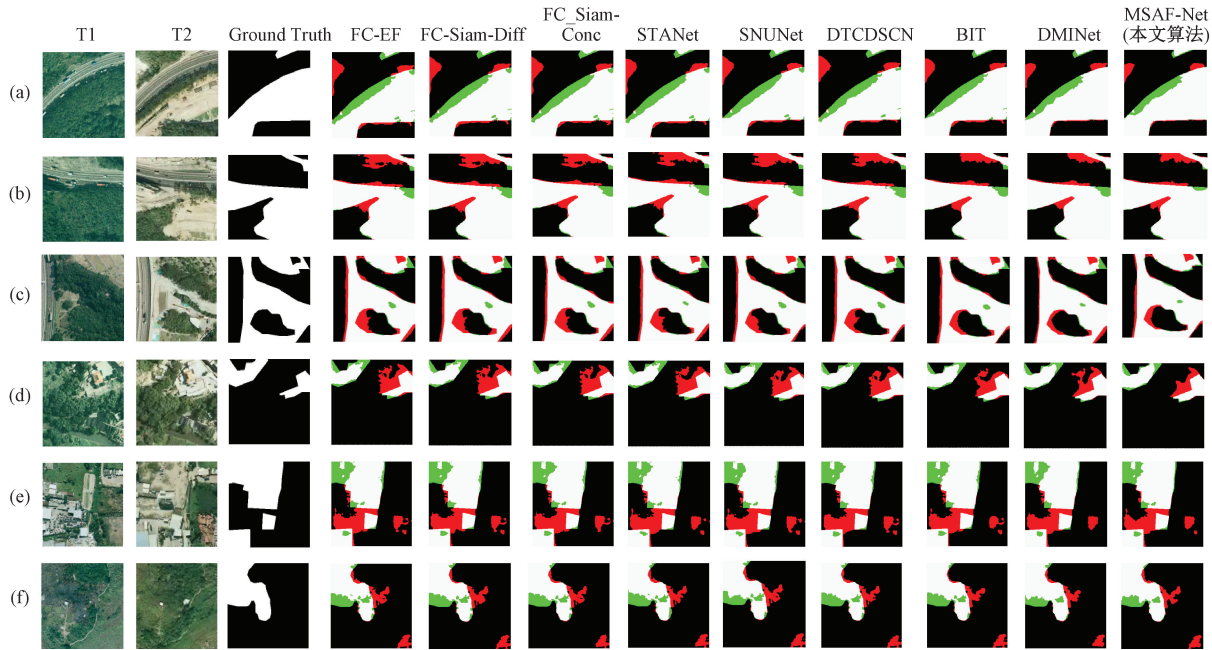


图 7 MSAF-Net 与其他方法在 SYSU-CD 上的预测结果比较

Fig. 7 Comparison of prediction results of MSAF-Net with other methods on SYSU-CD

3.4 消融实验

为了验证 MSAF-Net 网络中各模块的作用,本文在以上 3 个数据集进行了消融实验,在各个选择相同主干基线解码器包括特征差异(FD)分支作为基础网络。在方法 1、方法 2 中,将未辅助模块替换为特征拼接(FC),以评估这些模块在架构中的功效。

- 1) 基线:ResNet+FD。
- 2) 方法 1:ResNet+FDFM+FC。
- 3) 方法 2:ResNet+FDFM+MDFM+FC。
- 4) MSAF-Net:ResNet+FDFM+MDFM+FSEM。

用本文的模块逐步替换原来的特征差异和特征拼接模块,结果如表 4 所示,模型的性能有持续的增长。作为首个添加模块,在各数据集,从“基线”到“方法 1”,指标均有提升(如 LEVIR-CD 的 P_r 从 89.15% 提升到 91.10%、WHU-CD 的 P_r 从 92.76% 提升到 93.32%、SYSU-CD 的 P_r 从 78.53% 提升到 79.42%)。说明 FDFM 能有效增强特征差异提取,让模型更精准捕捉双时相遥感图像的变化信息,提升变化检测的准确性基础。从“方法 1”到“方法 2”(添加 MDFM),指标进一步上升(如 LEVIR-CD 的 F1 从 90.28% 提升到 90.60%、WHU-CD 的 F1 从 90.16% 提升到 91.41%)。MDFM 可融合多维度特征(不同分支提取的特征),补充单一 FDFM 的特征视角,优化特征表达的丰富度,使模型对复杂地物变化

(如建筑、植被变迁)的识别更全面,提升召回率与 F1 等综合指标。从“方法 2”到“MSAF-Net”(添加 FSEM),指标再创新高(如 LEVIR-CD 的 IoU 从 83.53% 提升到 83.90%、WHU-CD 的 IoU 从 84.77% 提升到 85.90%)。FSEM 借助深度可分离卷积等操作,精细化增强特征,挖掘更细粒度的变化模式(如小面积建筑改造、植被细微变化),在精确率、召回率平衡下,显著提升 F1 和 IoU,强化模型对复杂场景变化的检测能力。

4 结 论

本文设计的 MSAF-Net 网络集成 FDFM(差异检测)、MDFM(多尺度融合)、FSEM(残差增强)3 大核心模块。在融合阶段引入 CBAM 机制,强化关键特征。FDFM 能有效增强特征差异提取,让模型更精准捕捉双时相遥感图像的变化信息,提升变化检测的准确性基础,MDFM 可融合多维度特征(不同分支提取的特征),补充单一 FDFM 的特征视角,优化特征表达的丰富度,使模型对复杂地物变化(如建筑、植被变迁)的识别更全面。FSEM 借助深度可分离卷积等操作,精细化增强特征,挖掘更细粒度的变化模式(如小面积建筑改造、植被细微变化),强化模型对复杂场景变化的检测能力。在数据集 LEVIR-CD、WHU-CD 和 SYSU-CD 上的实验也证明了

表 4 不同数据集上消融实验的定量比较结果

Table 4 Quantitative comparison results of ablation experiments on different datasets								(%)
数据集	方法	FDFM	MDFM	FSEM	P_r	R_c	F1	IoU
LEVIR-CD	基线				89.15	88.67	89.76	80.70
	方法 1	✓			91.10	89.56	90.28	82.51
	方法 2	✓	✓		90.34	90.56	90.60	83.53
	MSAF-Net	✓	✓	✓	91.23	90.55	90.71	83.90
	基线				92.76	87.59	89.71	82.34
WHU-CD	方法 1	✓			93.32	86.90	90.16	83.59
	方法 2	✓	✓		93.14	88.57	91.41	84.77
	MSAF-Net	✓	✓	✓	93.53	90.55	91.71	85.90
	基线				78.53	80.56	79.52	67.85
SYSU-CD	方法 1	✓			79.42	81.58	81.05	68.02
	方法 2	✓	✓		79.72	82.67	81.69	69.14
	MSAF-Net	✓	✓	✓	81.53	83.53	82.71	71.80

MSAF-Net 模型具有很好的鲁棒性,消融实验也验证了 FDFM,MDFM,FSEM 模块的有效性。尽管 MSAF-Net 在现有的数据集上取得了较好的成果,但仍有进一步改善的空间。在复杂情况下仍然受到大变化、遮挡、阴影和复杂背景等因素的影响。模型的主干网络不够轻量化,无法满足移动端部署需求,在后续的研究中可以对模型进行改进,更好地应对更复杂多样的场景。

参考文献

[1] Parelius E J. A review of deep-learning methods for change detection in multispectral remote sensing images [J]. Remote Sensing, 2023, 15(8) : 2092.

[2] Gao F, Wang X, Gao Y, et al. Sea ice change detection in sar images based on convolutional-wavelet neural networks [J]. IEEE Geoscience and Remote Sensing Letters, 2019,16(8) :1240-1244.

[3] Vavassori A, Oxoli D, Venuti G, et al. A combined remote sensing and gis-based method for local climate zone mapping using PRISMA and Sentinel-2 imagery[J]. International Journal of Applied Earth Observation and Geoinformation, 2024, 131:103944.

[4] Bikis A, Engdaw M, Pandey D, et al. The impact of urbanization on land use land cover change using geographic information system and remote sensing: A case of Mizan Aman City Southwest Ethiopia [J]. Scientific Reports, 2025, 15:12014.

[5] Peijun D U, Xingli L I, Wen C, et al. Monitoring urban land cover and vegetation change by multi-temporal remote sensing information [J]. Mining Science and Technology (China),2010,20(6) :922-932.

[6] 顾桂梅,王小亮. 深度学习的接触网小目标缺陷识别研究 [J]. 电子测量与仪器学报, 2024, 38 (4) : 151-160.

Gu Guimei, Wang Xiaoliang. Identification of the

catenary small target defects in deep learning [J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(4) : 151-160.

[7] 王琦,张涛,徐超炜,等. 多尺度注意力融合与视觉 Transformer 方法优化的电阻抗层析成像深度学习方方法[J]. 仪器仪表学报,2024,45(7) :52-63.

Wang Qi, Zhang Tao, Xu Chaopei, et al. Optimized learning method for electrical impedance tomography with multi-scale attention fusion and vision transformer [J]. Chinese Journal of Scientific Instrument, 2024, 45(7) : 52-63.

[8] Lv Z, Huang H, Sun W, et al. Novel enhanced unet for change detection using multimodal remote sensing image[J]. IEEE Geoscience and Remote Sensing Letters, 2023, 20(000) :5.

[9] Lv Z, Huang H, Sun W, et al. Land cover change detection techniques: Very-high-resolution optical images: A review [J]. IEEE Geoscience and Remote Sensing Magazine, 2022,10(1) :44-63.

[10] Jiang X F, Xiang S, Wang M, et al. Dual-pathway change detection network based on the adaptive fusion module [J]. IEEE Geoscience and Remote Sensing Letters,2022, 19(1) :1-5.

[11] Bai B, Fu W, Lu T, et al. Edge-guided recurrent convolutional neural network for multitemporal remote sensing image building change detection [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60(3) : 1-13.

[12] Zhang X L, Gao D D, Li B J, et al. Global and local graph-based difference image enhancement for change detection[J]. Remote Sensing,2023,15(5) :1194-1194.

[13] Wang J, Gao F, Dong J, et al. Change detection from synthetic aperture radar images via graph-based knowledge supplement network [J]. IEEE Journal of

- Selected Topics in Applied Earth Observations and Remote Sensing, 2022, 15:1823-1836.
- [14] 潘成龙, 刘立程, 潘丹. 融合 CNN 和 Transformer 的三维冠状动脉 CT 图像分割[J]. 电子测量技术, 2025, 48(6):143-151.
- Pan Chenglong, Liu Licheng, Pan Dan. Integration of CNN and transformer for 3D coronary artery CT image segmentation[J]. Electronic Measurement Technology, 2025, 48(6): 143-151.
- [15] Chen Z, Chen H, Leng J, et al. VMMCD: VMamba-based multi-scale feature guiding fusion network for remote sensing change detection[J]. Remote Sensing, 2025, 17(11):1840-1840.
- [16] Xia L G, Chen J, Luo J C, et al. Building Change Detection Based on an Edge-Guided Convolutional Neural Network Combined with a Transformer[J]. Remote Sensing, 2022, 14(18):4524-4524.
- [17] Fang S, Li K, Shao J, et al. SNUNet-CD: A Densely Connected Siamese Network for Change Detection of VHR Images[J]. IEEE Geoscience and Remote Sensing Letters, 2021, PP(99):1-5.
- [18] Jiang K, Zhang W, Liu J, et al. Joint variation learning of fusion and difference features for change detection in remote sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60:1-18.
- [19] Wen Y, Ma X, Zhang X, et al. GCD-DDPM: A generative change detection model based on difference-feature guided DDPM[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62:1-16.
- [20] Xiang S, Wang M, Jiang X, et al. Dual-task semantic change detection for remote sensing images using the generative change field module[J]. Remote Sensing, 2021, 13:3336.
- [21] Badhe B N, Yele P V, Neve P R, et al. Hybrid three branch cnn and transformer with skeleton based attentions network for change detection in remote sensing images[J]. Sensing and Imaging, 2025, 26(1):109-109.
- [22] Feng Y, Xu H, Jiang J, et al. ICIF-Net: Intra-scale cross interaction and inter-scale feature fusion network for bitemporal remote sensing images change detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60:1-13.
- [23] Tang X, Zhang T, Ma J, et al. WNet: W shaped hierarchical network for remote-sensing image change detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61:1-14.
- [24] Ji R, Tan K, Wang X, et al. PASSNet: A spatial-Spectral feature extraction network with patch attention module for hyperspectral image classification[J]. IEEE Geoscience and Remote Sensing Letters, 2023, 20:1-5.
- [25] Hao M G, Xing T X, Jiang J L, et al. Attention mechanisms in computer vision: A survey[J]. Computational Visual Media, 2022, 8(3):331-368.
- [26] Mi M, Xie G, Zhang Z, et al. Smoothing filter-based panchromatic spectral decomposition for multispectral and hyperspectral image pansharpening[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2022, 15:3612-3625.
- [27] Liu C, Wei L, Zhang Z, et al. Recursive self-attention modules based network for panchromatic and multispectral image fusion[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2023, 16:10067-10083.
- [28] Zhang Z, Xia W, Xie G, et al. Fast Opium poppy detection in unmanned aerial vehicle (UAV) Imagery based on deep neural network[J]. Drones, 2023, 7(9):559-.
- [29] Shao X, Wang M, Xiao J G, et al. Cloud coverage estimation network for remote sensing images[J]. IEEE Geosci. Remote Sensing Letters, 2023, 20:1-5.
- [30] Xiang S, Xie Q, Wang M. Semantic segmentation for remote sensing images based on adaptive feature selection network[J]. IEEE Geoscience and Remote Sensing Letters, 2022, 19:1-5.
- [31] Zhou C, Peng Y, et al. A deeply supervised image fusion network for change detection in high resolution bi-temporal remote sensing images[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 166:183-200.
- [32] Han C, Wu C, Guo H M, et al. HANet: A hierarchical attention network for change detection with bi-temporal very-high-resolution remote sensing images[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2023, 16:3867-3878.
- [33] Feng Y, Jiang J, Xu H, et al. Change detection on remote sensing images using dual-branch multilevel intertemporal network[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61:1-15.
- [34] Chen C P, Hsieh J W, Chen P Y, et al. SARAS-Net: Scale and relation aware siamese network for change detection[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2023, 37(12):14187-14195.
- [35] Liu Y, Zhang X, Zhu J, et al. HDANet: Enhancing underwater salient object detection with physics-inspired multimodal joint learning[J]. IEEE Transactions Geoscience and Remote Sensing, 2025, 63:1-14.

[36] Yang L X, Zhang R Y, Li L, et al. SimAM: A simple, parameter-free attention module for convolutional neural networks[J]. PMLR, 2021, 11863-11874.

[37] Woo S, Park J, Lee J Y, et al. CBAM: Convolutional block attention module [C]. Proceedings of the European Conference on Computer Vision (ECCV), 2018.

作者简介



周先春 (通信作者), 2011 年于南京信息工程大学获得博士学位, 现为南京信息工程大学硕士生导师, 教授, 中国电子学会高级会员, 主要研究方向为信号与信息处理。
E-mail: zhoux2008@163.com

Zhou Xianchun (Corresponding author), received his Ph. D. degree from Nanjing University of Information Science and Technology in 2011. Now he is a professor and M. Sc. supervisor at Nanjing University of Information Science and Technology, he is also a senior member of China Electronics Society. His main research interests include signal and information

processing.



倪鸿骏, 2023 年于南京理工大学紫金学院获得学士学位, 现为南京信息工程大学硕士研究生, 主要研究方向为数字图像处理。
E-mail: 1390057387@qq.com

Ni Hongjun received his B. Sc. degree from Nanjing University of Science and Technology Zijin College in 2023. Now he is a M. Sc. candidate in Nanjing University of Information Science and Technology. His main research interest includes digital image processing.

李翰哲, 现为南京信息工程大学人工智能学院本科生, 主要研究方向为数字图像处理。
E-mail: 202583460093@nuist.edu.cn

Li Hanzhe is an B. Sc. student in the School of Artificial Intelligence of Nanjing University of Information Science and Technology. His main research interest includes digital image processing.

