

基于注意力引导的多尺度特征融合的遥感图像变化检测算法研究

贾正正, 宋媛萌*, 张铭钰, 张 豪, 韩卓航

北华航天工业学院计算机学院, 河北 廊坊

收稿日期: 2025年1月2日; 录用日期: 2025年2月7日; 发布日期: 2025年2月19日

摘 要

为了解决遥感图像变化检测中变化边缘检测不清晰以及极小目标的漏检问题, 本论文提出了一种基于注意力引导的多尺度特征融合的遥感图像变化检测网络(MFM-CDNet), 旨在提高变化检测的效率和准确性。MFM-CDNet通过特征增强模块FEM和注意力引导的特征融合模块FFM, 实现了从粗到细的多尺度特征提取与融合。FEM利用双重注意力机制精确定位图像中的重要区域, 而FFM通过对多尺度特征的有效整合与优化, 增强了模型对正样本的识别能力, 并提高了边缘变化区域的定位精度。在LEVIR-CD、WHU-CD和GoogleGZ-CD三个公开数据集上的实验表明, MFM-CDNet网络综合性能优于现有的变化检测方法。消融实验进一步验证了FEM和FFM模块在提升检测性能中的关键作用。MFM-CDNet的成功应用展示了其在复杂遥感图像变化检测中的鲁棒性和泛化能力, 为未来的遥感图像变化检测研究提供了重要的参考。

关键词

遥感图像变化检测, 注意力引导, 多尺度特征融合

Research on Change Detection Algorithm for Remote Sensing Images Based on the Attention-Guided Multiscale Feature Fusion

Zhengzheng Jia, Yuanmeng Song*, Mingyu Zhang, Hao Zhang, Zhuohang Han

School of Computer Science and Engineering of North China Institute of Aerospace Engineering, Langfang Hebei

Received: Jan. 2nd, 2025; accepted: Feb. 7th, 2025; published: Feb. 19th, 2025

Abstract

In order to solve the problems of unclear detection of change edges and missed detection of very

*通讯作者。

文章引用: 贾正正, 宋媛萌, 张铭钰, 张豪, 韩卓航. 基于注意力引导的多尺度特征融合的遥感图像变化检测算法研究[J]. 计算机科学与应用, 2025, 15(2): 1-12. DOI: 10.12677/csa.2025.152028

small targets in remote sensing image change detection, this thesis proposes an attention-guided multi-scale feature fusion-based remote sensing image change detection network (MFM-CDNet), which aims to improve the efficiency and accuracy of the change detection. MFM-CDNet achieves multi-scale feature extraction and fusion from coarse to fine by means of a feature enhancement module, FEM, and an attention-guided feature fusion module, FFM. FFM achieves multi-scale feature extraction and fusion from coarse to fine. FEM utilises a dual attention mechanism to pinpoint important regions in the image, while FFM enhances the model's ability to identify positive samples and improves the localisation accuracy of edge-change regions through the effective integration and optimisation of multi-scale features. Experiments on three publicly available datasets, LEVIR-CD, WHU-CD and GoogleGZ-CD, show that the comprehensive performance of the MFM-CDNet network outperforms existing change detection methods. The ablation experiments further validate the key role of FEM and FFM modules in enhancing the detection performance. The successful application of MFM-CDNet demonstrates its robustness and generalisation ability in complex remote sensing image change detection, which provides an important reference for future research on remote sensing image change detection.

Keywords

Remote Sensing Image Change Detection, Attention Guidance, Multi-Scale Feature Fusion

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

遥感图像变化检测技术通过比较不同时期同一区域的图像, 识别地表变化, 广泛应用于城市规划[1] [2]、环境监测[3]、农业监测[4] [5]和自然灾害管理[6] [7]等领域。随着遥感技术的发展, 图像的分辨率和覆盖范围不断增加, 为变化检测提供了更加丰富的数据源, 但也带来了新的挑战。传统变化检测方法依赖于人工特征提取和简单的图像差异度量, 虽然能获得一定效果, 但在复杂遥感数据中往往表现出计算效率低和适应性差的问题。

深度学习技术因其优秀的特征学习和泛化能力, 已在图像识别[8] [9]、图像分类[10] [11]以及图像融合[12]等领域广泛应用, 尤其在遥感图像变化检测中, 通过自动学习图像的多层次特征, 减少人工干预, 提高了检测效率和准确性。然而, 遥感图像变化检测仍面临多重挑战。首先, 如何准确识别和定位地表变化区域是一个关键问题。其次, 如何有效融合不同尺度的特征, 尤其是在小目标变化检测中, 传统方法往往难以精确检测。变化检测中的边缘区域常常存在定位不准或模糊问题, 增加了模型设计的难度。尽管卷积神经网络在特征提取上表现出色, 但在变化检测任务中, 难以兼顾全局和局部信息。因此, 设计一个能够综合全局与局部信息, 并提高变化区域识别精度的网络模型, 是当前研究的重点。

因此, 为了进一步改善变化边缘检测不清晰以及极小目标的漏检问题, 本文提出了一种基于注意力引导的多尺度特征融合的遥感图像变化检测网络(简称 MFM-CDNet), 该网络通过特征增强模块、全局上下文模块和特征融合复用模块, 增强了模型对变化边缘和极小目标变化的检测能力。本文的主要贡献为:

1) 该网络新增加了全新的特征增强模块, 该模块能够捕捉到局部细节, 从而在变化检测任务中实现更精准的识别, 尤其是在识别那些在传统模型中容易被遗漏的小目标变化时表现出色。

2) 通过注意力引导以及对多尺度特征的融合和复用, 精细化处理和复用不同层次的特征信息, 优化了模型对正样本的识别能力, 显著提高了对变化边缘区域的定位精度有效减少了变化边缘检测中的不清

晰问题, 对于提高变化检测的整体性能和准确性起到了关键作用。

3) 在 LEVIR-CD、WHU-CD 和 GoogleGZ-CD 三个公开数据集上做了大量的实验和详细的消融实验, 在训练和推理方面本文提出的 MFM-CDNet 网络都表现出显著的性能, 证明了所提出的方法的有效性和优势。

2. 相关工作

随着深度学习技术的快速发展, 许多研究者开始将其应用于变化检测领域, 取得了显著的成果。在这些应用中, 卷积神经网络(CNN)和 U-Net 等架构被广泛采用, 并取得了显著效果。例如, Simonyan 和 Zisserman 提出的 VGG 网络[13]以及 He 等人提出的 ResNet [14], 在图像分类任务中表现出色, 成为变化检测领域的基础模型。Long 等人[15]提出的全卷积网络(FCN)通过用卷积层替换全连接层, 成功实现了像素级别的语义分割, 这为变化检测提供了有力的支持。U-Net 结构则凭借其对称的编码器-解码器设计, 能够高效捕捉图像中的细节信息, 尤其在医学影像分割和遥感图像变化检测中得到了广泛应用[16]。此外, Guo 等人[17]提出的动态 Siamese 网络为视觉目标跟踪提供了新的思路, 并能够扩展应用于变化检测任务。该方法通过共享权重来处理两幅图像, 并通过相似特征提取和度量生成变化图, 从而显著提升了变化检测的准确性。

特征选择与深度学习的结合进一步推动了变化检测技术的进步。例如, Zhai 等人[18]在 Siamese 网络中引入了多尺度融合机制, 提出了一种新型的融合网络。该网络通过双分支编码器和三分支解码器的设计, 提高了遥感图像变化检测的精度, 并通过融合不同层次的特征增强了模型对变化区域的识别能力。此外, Yu 等人[19]提出了一种新颖的遥感图像变化检测方法, 结合了多级特征交叉融合与 3D 卷积神经网络, 有效提高了变化检测的精度。

3. 方法

MFM-CDNet 网络包括编码器和解码器两个主要部分。该模型通过结合特征增强模块(FEM)和注意力引导的特征融合模块(FFM), 实现了从粗到细的多尺度特征提取与融合, 从而提高了变化检测的精度和鲁棒性。MFM-CDNet 模型总体结构如图 1 所示。

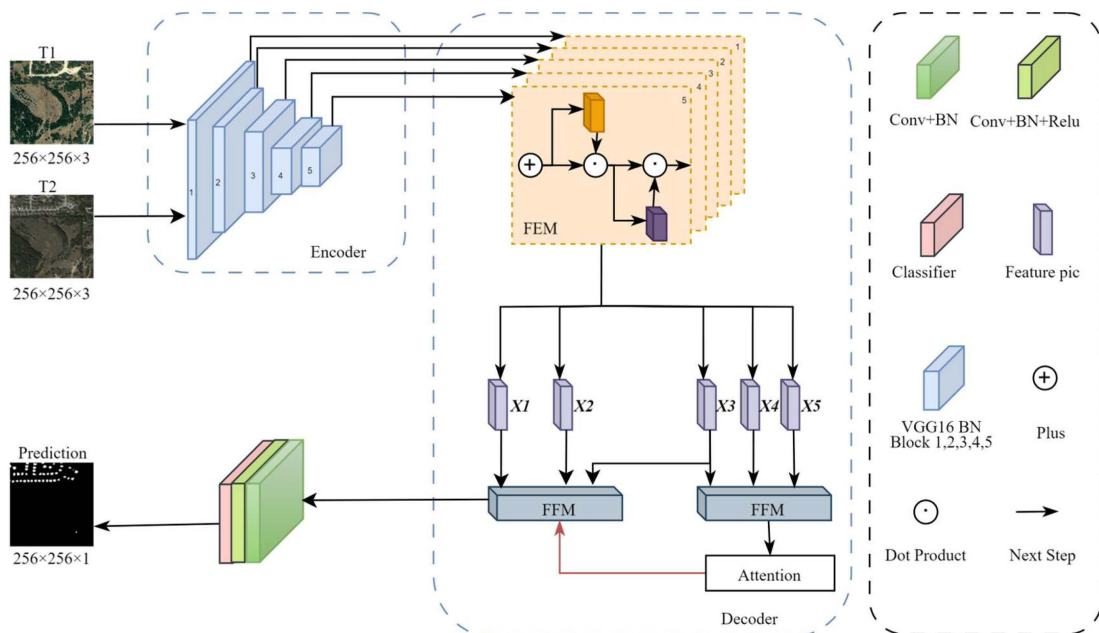


Figure 1. General architecture of MFM-CDNet model

图 1. MFM-CDNet 模型总体架构

3.1. 编码器部分

在 MFM-CDNet 的编码器部分, 采用了经典的 VGG-16 网络作为主干架构, 并对其进行了特定优化, 以更好地适应变化检测任务的需求。具体而言, 网络使用了 5 个经过预训练的 VGG16_BN 模块, 每个模块对应于 VGG16_BN 网络中的不同层段: 0~5、5~12、12~22、22~32 和 32~42 层。每个 VGG16_BN 模块都经过批量归一化处理, 这有助于加速训练过程并增强模型的泛化能力。通过这种设计, MFM-CDNet 能够在不同的阶段提取特征, 支持从粗到精的多层次特征融合。此策略使得模型能够有效地捕捉不同尺度的变化信息, 确保在变化检测任务中对细节和边缘的准确识别。

3.2. 解码器部分

1) 特征增强模块 FEM

特征增强模块(FEM)是 MFM-CDNet 中的一个关键部分, 旨在通过注意力引导精确地定位图像中的重要区域, 同时减少不必要的计算量, 从而提升模型效率和对细节变化的敏感度。FEM 采用了双重注意力机制, 包括通道注意模块和空间注意力模块。这一机制的设计使得模型能够更加专注于图像中的关键信息, 提高了变化检测的性能。

通道注意模块的作用是突出图像中最重要特征通道, 从而生成更加精炼和具有区分性的特征表示。通过这种方式, 模型能够减少对不重要特征的计算量, 提高对关键信息的关注度。通道注意模块主要解决了网络“关注什么”的问题。其主要包括自适应最大池化操作、二维卷积操作、ReLU 的激活函数、Sigmoid 激活函数等部分。通道注意力模型架构如图 2 所示。

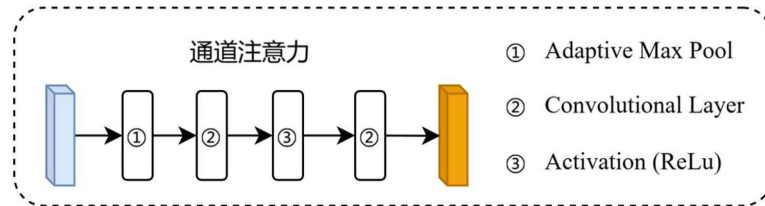


Figure 2. Channel attention module architecture
图 2. 通道注意力模块架构

自适应最大池化用于将特征图的每个通道通过最大池化方式聚合成一个标量。假设输入的特征图为 $X \in \mathbb{R}^{C \times H \times W}$, 其中 C 是通道数, H 是高度, W 是宽度。池化操作为:

$$M_c = \text{MaxPool}(X_c) \quad (1)$$

接下来, 经过自适应池化后的特征图会进行二维卷积, 以获得通道的注意力权重。假设通过卷积操作生成了一个权重矩阵 $W_{channel}$, 它与输入的特征图的每个通道进行逐元素相乘, 从而增强或抑制特定通道的响应。卷积操作为:

$$W_{channel} = \text{Conv2D}(M_c, W_{channel}) \quad (2)$$

ReLU 激活函数应用于卷积结果, 以增加非线性表示能力。具体操作为:

$$F_{channel}^{Relu} = \max(0, F_{channel}) \quad (3)$$

最后, 使用 Sigmoid 激活函数对通道的响应进行归一化, 输出每个通道的注意力权重。

$$\alpha_c = \sigma(F_{channel}^{Relu}) = \frac{1}{1 + \exp(-F_{channel}^{Relu})} \quad (4)$$

根据计算得到的通道注意力权重 α_c ，对每个通道进行加权，得到最终的增强特征图。

$$X_{enhanced} = \sum_{c=1}^C \alpha_c \cdot X_c \quad (5)$$

空间注意力模块的作用是帮助网络聚焦于图像中的关键区域，主要解决“聚焦在哪里”的问题。该模块通过行向量最大值操作和一维卷积操作来生成空间注意力图，从而指引网络注意力集中在具有变化的区域。其主要包括行向量最大值操作、一维卷积操作、Sigmoid 激活函数几个部分。值得注意的是，二维卷积运算时先水平操作，再垂直操作，而一维卷积操作只能是垂直操作。空间注意力模型架构如图 3 所示。

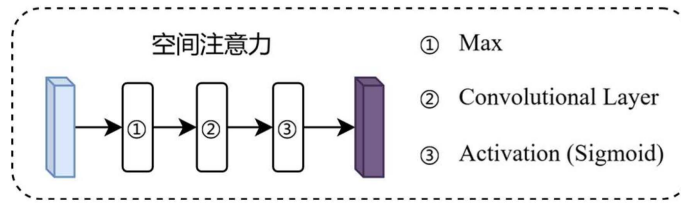


Figure 3. Spatial attention module architecture
图 3. 空间注意力模块架构

对于输入特征图 $X_{enhanced} \in \mathbb{R}^{C \times H \times W}$ ，首先沿着空间维度(即高度和宽度)对每个通道的最大值进行聚合，生成一个 $H \times W$ 的空间特征图。

$$M_{spatial}(i, j) = \max_c X_{enhanced}(c, i, j) \quad (6)$$

接下来，使用一维卷积操作对聚合后的空间特征图进行处理，计算空间注意力权重。一维卷积操作见公式(7)。

$$A_{spatial} = \text{Conv1D}(M_{spatial}, W_{spatial}) \quad (7)$$

最后，使用 Sigmoid 激活函数归一化空间注意力图。

$$\beta_{i,j} = \alpha(A_{spatial}(i, j)) \quad (8)$$

最后，将空间注意力图应用于特征图的每个像素点，使网络可以集中在图像的关键区域，增强特定空间区域的响应。

$$X_{final} = X_{enhanced} \cdot \beta_{i,j} \quad (9)$$

整个特征增强模块中，通道注意力模块和空间注意力模块共同作用使得网络能够有效地捕捉到图像中的细节信息，减少无关信息的干扰。

2) 注意力引导的特征融合模块 FFM

注意力引导的特征融合模块(简称 FFM)主要是通过通过对多尺度特征的有效整合与优化，增强模型对正样本的识别能力，并提高边缘变化区域的定位精度。FFM 不仅实现了不同层次特征的深度融合，还引入了注意力(Attention Map)机制，以突出显示图像中的重要区域，从而显著提升了变化检测的性能。FFM 模块的详细架构见图 4。

FFM 首先对最深层的特征图进行双线性上采样，使其与次深层特征图的尺寸匹配。上采样后的特征图与次深层特征图进行相乘，既保留了高层次的语义信息，又引入了低层次的细节特征。然后，对次深层特征图继续进行上采样，并与更浅层的特征图融合，逐步将所有层次的特征信息整合。最终，这些融

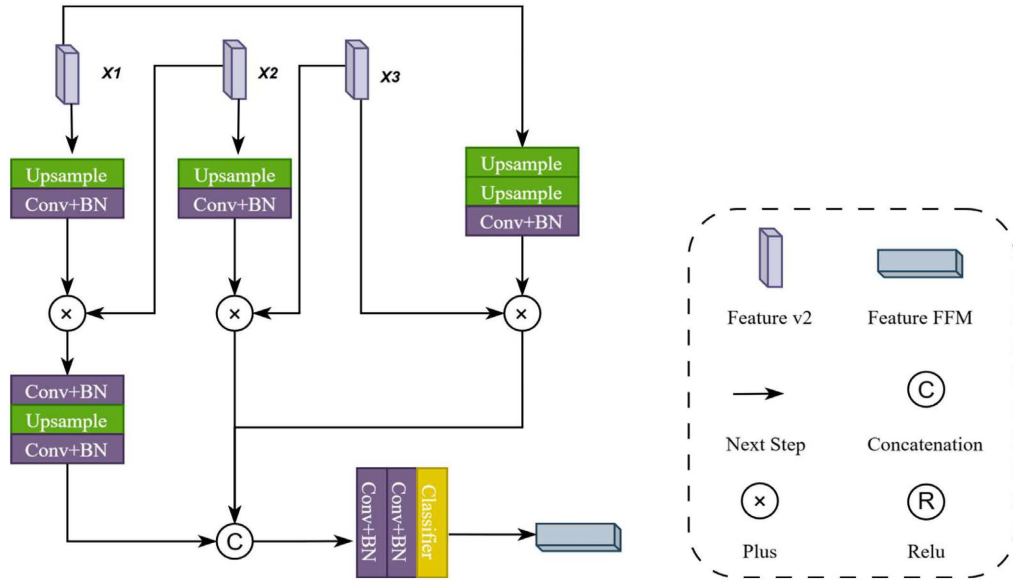


Figure 4. FFM module architecture
图 4. FFM 模块架构

合后的特征图通过水平拼接形成一个多尺度特征图，确保了不同层次特征之间的无缝衔接，增强了特征表达的丰富性与多样性。

拼接后的特征图经过两个 3×3 的二维卷积模块进行处理，生成最终的特征图，这一特征图具有更高的空间分辨率和更强的特征表达能力，有效促进了变化检测的准确性。同时，FFM 引入了注意图机制，利用高层次特征图作为引导信息，指导低层次特征图的融合过程。这使得 FFM 能够自适应地突出图像中的关键区域，减少背景噪声的干扰，从而提高了模型对变化区域和边缘细节的捕捉能力。最终，FFM 不仅提升了多尺度特征的整合效果，还显著增强了变化检测的整体性能和鲁棒性。

3.3. 损失函数

变化检测任务被视为一种特殊的二分类问题，本文使用二元交叉熵损失函数，结合了 Sigmoid 激活函数和 BCELoss。损失函数为：

$$Loss_{sup} = \{l_1, \dots, l_n\} \tag{10}$$

$$l_n = -[y_n \cdot \log(\alpha(X_n)) + (1 - y_n) \cdot \log(1 - \alpha(X_n))] \tag{11}$$

$$\alpha(X) = \frac{1}{1 + \exp(-x)} \tag{12}$$

$\alpha(X)$ 是一个 Sigmoid 激活函数，它将 x 映射到 $(0,1)$ 。

4. 实验

4.1. 数据集介绍

本次实验共使用 3 个公开数据集，分别为 LEVIR-CD、WHU-CD、GoogleGZ-CD。

LEVIR-CD 数据集主要集中在高分辨率的城市区域变化检测，注重小尺度的变化细节，适合城市建设和改造的研究。

WHU-CD 数据集提供了更广泛的变化类型，包含了城市、乡村和自然环境，数据来源和分辨率多样，

适合多种场景下的变化检测任务。

GoogleGZ-CD 数据集同样由武汉大学提供, 专注于广州郊区的长期变化检测, 偏向大范围的检测, 适用于大尺度的环境变化检测, 变化类型复杂多样。

三个数据集的训练集、验证集和测试集的样本数量详细信息如表 1 所示。

Table 1. Dataset details

表 1. 数据集详细信息

数据集	训练集样本数	验证集样本数	测试集样本数
LEVIR-CD	7120 对	1024 对	2048 对
WHU-CD	6096 对	1184 对	1910 对
GoogleGZ-CD	804 对	330 对	330 对

4.2. 实验设置

本节在 PyTorch 上实现了模型, 并在 NVIDIA RTX 4090 上进行训练和测试。对于模型的具体参数配置, 本文利用 AdamW 优化器最小化损失函数, 权重衰减系数为 0.0025, 学习率为 $5e-4$ 。由于受到 GPU 资源的限制, 并通过大量实验的验证, 本文将批量大小(batch size)设定为 16, 同时将训练周期(epoch 数)设定为 180, 以确保模型能够达到收敛。在训练过程中, 通过应用标准的数据增强技术, 对每一对遥感图像进行相同的数据增强处理, 包括随机的裁剪、旋转、翻转以及高斯模糊。

4.3. 评价指标

为了清楚地研究 MFM-CDNet 的性能, 本次实验采用三个评价指标: 精确度(Precision, P)、召回率(Recall, R)和 F1 得分。

精确度反映了模型预测的变化区域中实际发生变化的比例。它衡量的是所有被标记为变化的像素中, 真正发生变化的部分。高精确度意味着较少的误报。

召回率表示模型成功识别出的实际变化区域的比例。它衡量的是所有实际变化区域中被正确检测到的部分。高召回率意味着模型能捕捉到大多数真实变化。

F1 得分是精确度和召回率的调和平均值, 用于平衡两者。它提供了一个综合的评估标准, 常用来综合评估模型的性能。

公式分别如下所示:

$$P = \frac{TP}{TP + FP} \quad (13)$$

$$R = \frac{TP}{TP + FN} \quad (14)$$

$$F_1 = 2 \times \frac{P \times R}{P + R} \quad (15)$$

其中, TP 表示真阳性, 用于表示模型判断为变化且实际上为变化的区域; TN 表示真阴性, 表示模型判断为非变化且实际上确实不是变化的区域; FP 表示假阳性, 用于表示模型判断为变化而实际上并不是变化信息的区域; FN 表示假阴性, 用于表示模型判断为不变化而实际上为变化信息的区域。

4.4. 对比方法

本方法与几种最经典的方法进行了比较, 包括三种基于卷积的方法: FC-EF、FC-Siam-diff 和 FC-Siam-

conc; 两种基于注意力的方法: SNUNet 和 BIT。

1. FC-EF [20]: 基于 U-Net 结构的全卷积早期融合网络, 通过将两个输入图像连接形成不同通道进行早期融合。

2. FC-Siam-diff [20]: 一种全卷积暹罗差分模型, 通过计算解码器分支间差值的绝对值并进行跳跃连接来融合特征。

3. FC-Siam-conc [20]: 一种全卷积二乘模型, 通过在解码步骤中拼接两个输入图像的特征图, 增强网络在变化检测中捕捉细节的能力。

4. SNUNet [21]: 密集连接网络, 利用通道注意模块提炼多层次特征, 有效捕捉细微变化并提高检测准确性, 适用于变化检测任务。

5. BIT [22]: 基于 Transformer 的方法, 通过引入 Transformer 架构增强网络对图像长期依赖的捕捉能力, 从而提高变化检测的上下文理解和精度。

4.5. 实验结果与分析

本节分别在 LEVIR-CD、WHU-CD、GoogleGZ-CD 三个数据集上对不同方法的检测性能进行对比实验。

在 LEVIR-CD 数据集上不同方法的定量准确度结果如表 2 所示。

Table 2. Quantitative accuracy results of different methods on the LEVIR-CD dataset

表 2. LEVIR-CD 数据集上不同方法的定量准确度结果

Dataset	Method	P	R	F1
LEVIR-CD	FC-EF	86.91	80.17	83.40
	FC-Siam-diff	89.53	83.31	86.31
	FC-Siam-conc	82.68	87.52	85.03
	SNUNet	91.31	88.67	89.97
	BIT	89.24	89.37	89.30
	Ours (MFM-CDNet)	95.93	90.18	92.97

在 LEVIR-CD 数据集上的实验结果表明, MFM-CDNet 在精确度、召回率和 F1 值上均显著优于其他方法, 分别达到 95.93、90.18 和 92.97。与传统方法如 FC-EF、FC-Siam-diff 和 FC-Siam-conc 相比, MFM-CDNet 在 F1 值上提升近 10 个百分点, 尤其在变化边缘和小目标检测上表现优异。与先进方法如 SNUNet 和 BIT 相比, MFM-CDNet 也展现出更强的性能。SNUNet 的 F1 值为 89.97, 低于 MFM-CDNet, 而 BIT 的精度和召回率较为平衡, 但 MFM-CDNet 在精度上领先近 6 个百分点, 且在召回率和 F1 值上具有优势, 显示其在复杂场景中的鲁棒性。

LEVIR-CD 数据集本身包含大量小目标和边缘信息, MFM-CDNet 通过多尺度特征融合和注意力引导, 更有效地捕捉这些细节, 提升了复杂场景下的检测精度, 而其他方法在小目标识别和边缘定位上仍存在局限。

在 WHU-CD 数据集上不同方法的定量准确度结果如表 3 所示。

在 WHU-CD 数据集上的实验结果显示, MFM-CDNet 在精度、召回率和 F1 值上均显著优于其他方法, 分别达到 92.46、90.15 和 91.29。与传统方法如 FC-EF、FC-Siam-diff 和 FC-Siam-conc 相比, MFM-CDNet 在复杂城市环境下的变化区域识别能力大幅提升, 尤其在精度和召回率的平衡上表现更佳。与先进方法 SNUNet (F1 值 87.76) 和 BIT (F1 值 80.97) 相比, MFM-CDNet 在精度和召回率方面均占优, 特别是在复杂的城市环境中, 能够更准确地识别变化区域。尽管 BIT 在召回率上表现突出, 但较低的精度导致其 F1 值下降。

Table 3. Quantitative accuracy results of different methods on the WHU-CD dataset
表 3. WHU-CD 数据集上不同方法的定量准确度结果

Dataset	Method	P	R	F1
WHU-CD	FC-EF	76.49	46.77	58.05
	FC-Siam-diff	89.63	83.22	86.31
	FC-Siam-conc	72.06	57.55	63.99
	SNUNet	87.84	87.68	87.76
	BIT	74.01	89.37	80.97
	Ours (MFM-CDNet)	92.46	90.15	91.29

WHU-CD 数据集的复杂性为变化检测带来了挑战，MFM-CDNet 通过多尺度特征融合和全局上下文信息的结合，成功提升了检测精度，尤其在小目标、边缘区域和背景噪声较多的情况下，表现出色，展示了其在复杂环境中的鲁棒性和适应能力。

在 GoogleGZ-CD 数据集上不同方法的定量准确度结果如表 4 所示。

Table 4. Quantitative accuracy results of different methods on GoogleGZ-CD dataset
表 4. GoogleGZ-CD 数据集上不同方法的定量准确度结果

Dataset	Method	P	R	F1
GoogleGZ-CD	FC-EF	82.32	80.40	83.63
	FC-Siam-diff	84.21	82.79	84.27
	FC-Siam-conc	87.14	83.40	84.88
	SNUNet	78.85	68.39	73.22
	BIT	90.31	89.48	89.54
	Ours (MFM-CDNet)	90.73	89.55	90.14

在 GoogleGZ-CD 数据集上的实验表明，MFM-CDNet 在精度、召回率和 F1 值上均优于其他对比方法，特别是在精度和召回率的平衡上表现出色。MFM-CDNet 的精度(90.73)、召回率(89.55)和 F1 值(90.14)均排名第一，明显超越 FC-EF、FC-Siam-diff 和 FC-Siam-conc 等传统方法。与 SNUNet 和 BIT 相比，MFM-CDNet 展现了更强的综合性能。虽然 BIT 在召回率和精度上表现较好，但 MFM-CDNet 通过多尺度特征融合和注意力机制，进一步提升了精度，保持了较高的召回率，在 F1 值上高于 BIT。

在地物多样性和背景复杂性较强的 GoogleGZ-CD 数据集上，MFM-CDNet 凭借全局上下文信息和多尺度特征融合，显著提升了对细微变化的捕捉能力，确保了精度与召回率的平衡，展现了其在复杂遥感图像变化检测中的强大优势。

在不同数据集上的不同比较方法的定性评价结果如图 5 所示。

从上述定性评价结果图中，可以清楚地看出不同方法在变化检测任务中的实际表现。为了使可视化结果更加直观，本文使用黑色、白色、红色和浅蓝色分别表示未变化区域(TN)、变化区域(TP)、误检区域(FP)和漏检区域(FN)。黑色部分表示模型认为未发生变化的区域，白色部分表示模型成功检测到的变化区域。红色部分为误检，表示模型错误地将未变化区域识别为变化区域，而浅蓝色部分则为漏检，表示模型未能检测到实际存在的变化区域。

通过可视化结果可见，MFM-CDNet 在变化检测中表现出色，显著改善了误检和漏检问题。与其他方

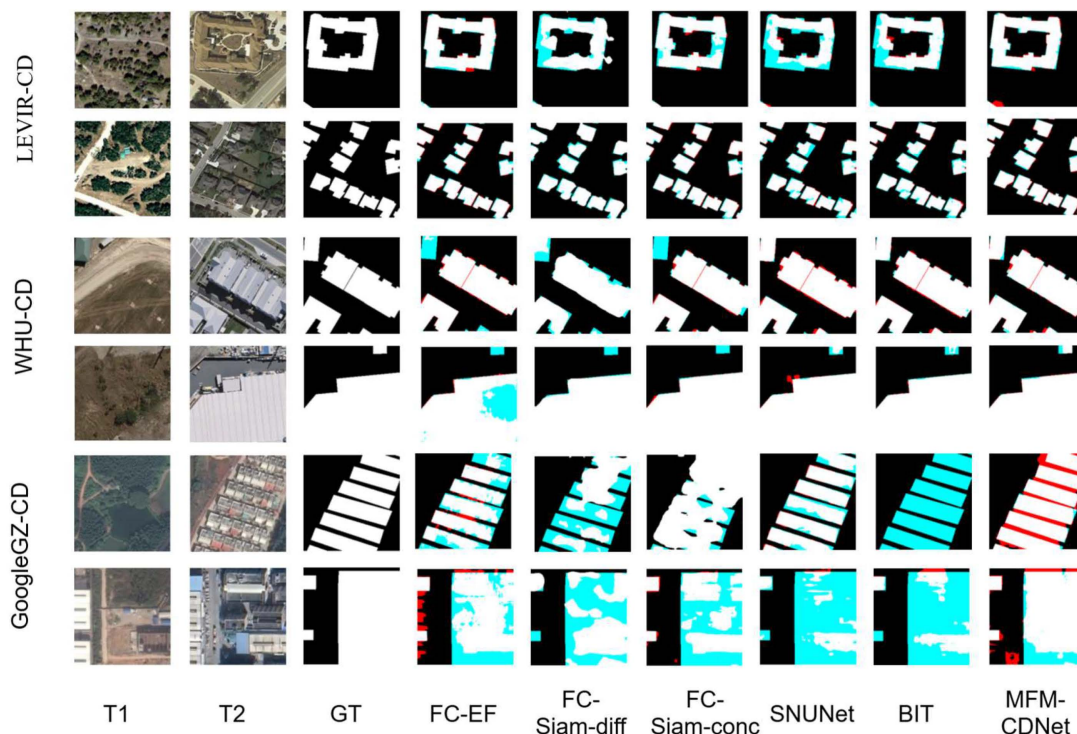


Figure 5. Qualitative evaluation results of LEVIR-CD, WHU-CD, and GoogleGZ-CD datasets
图 5. LEVIR-CD、WHU-CD、GoogleGZ-CD 数据集定性评价结果

法相比，MFM-CDNet 生成的变化边界更加清晰，细节更完整，特别是在复杂场景下，能够有效减少噪声干扰和对未变化区域的错误识别。其在边缘检测和极小目标检测方面尤为突出，成功捕捉到其他方法容易忽视的细微变化。

综上所述，MFM-CDNet 在可视化结果中展现了极其优越的性能，特别是在高分辨率遥感图像变化检测场景中展现出了强大的实用性和广泛的适用性。这些特点不仅证明了 MFM-CDNet 的先进性，也为后续的遥感图像变化检测研究提供了重要的参考与启示。

4.6. 消融实验

消融实验主要是对 MFM-CDNet 网络的特征增强模块 FEM 和注意力引导的特征融合模块 FFM 的效果进行验证，本文选取 LEVIR-CD 和 WHU-CD 两个数据集进行验证。在两个数据集上的消融实验的定量准确度结果如表 5 所示。

Table 5. Quantitative accuracy results of ablation experiments on LEVIR-CD and WHU-CD datasets
表 5. LEVIR-CD 和 WHU-CD 数据集上的消融实验定量准确度结果

Method	LEVIR-CD			WHU-CD		
	P	R	F1	P	R	F1
w/o FEM	79.03	76.87	77.94	75.27	78.35	76.78
w/o FFM	84.12	80.16	82.09	80.02	79.90	79.96
Ours (MFM-CDNet)	95.93	86.14	90.94	92.46	84.86	88.50

从表 5 的消融实验定量准确度结果可以看出，MFM-CDNet 在特征增强模块(FEM)和注意力引导的特

征融合模块(FFM)的共同作用下,在 LEVIR-CD 和 WHU-CD 数据集上的检测性能显著优于移除任一模块后的模型版本。这充分验证了 FEM 和 FFM 模块在 MFM-CDNet 网络中的重要性以及它们对整体性能提升的关键贡献。

5. 结论

本文针对遥感图像变化检测中面临的变化边缘检测不清晰和极小目标容易被漏检的问题,提出了一种基于注意力引导的多尺度特征融合的遥感图像变化检测网络(MFM-CDNet)。该网络通过引入特征增强模块(FEM)和注意力引导的特征融合模块(FFM),有效提高了模型对变化区域,尤其是边缘和小目标变化的敏感度。实验结果表明,该网络在 LEVIR-CD、WHU-CD 和 GoogleGZ-CD 三个公开数据集上均表现出优异的性能,显著提高了精确度、召回率和 F1 值,尤其在变化边缘和小目标检测方面具有显著优势。该网络在训练和推理性能上超越了其他变化检测网络。未来我们将探索更加高效的多尺度特征融合策略,以进一步增强模型的检测能力,并扩展其在不同遥感数据集和实际应用中的适应性。

基金项目

北华航天工业学院 2023 年硕士研究生科研创新项目(项目号:YKY-2023-39);北华航天工业学院 2023 年硕士研究生科研创新项目(项目号:YKY-2023-40)。

参考文献

- [1] Tison, C., Nicolas, J.-M., Tupin, F. and Maitre, H. (2004) A New Statistical Model for Markovian Classification of Urban Areas in High-Resolution SAR Images. *IEEE Transactions on Geoscience and Remote Sensing*, **42**, 2046-2057. <https://doi.org/10.1109/tgrs.2004.834630>
- [2] Stahl, B., Apfelbeck, J. and Lange, R. (2023) Classification of Micromobility Vehicles in Thermal-Infrared Images Based on Combined Image and Contour Features Using Neuromorphic Processing. *Applied Sciences*, **13**, Article 3795. <https://doi.org/10.3390/app13063795>
- [3] Lee, C., Park, S., Kim, T., Liu, S., Md Reba, M.N., Oh, J., et al. (2022) Machine Learning-Based Forest Burned Area Detection with Various Input Variables: A Case Study of South Korea. *Applied Sciences*, **12**, Article 10077. <https://doi.org/10.3390/app121910077>
- [4] Chen, Y., Duan, Y., Zhang, W., Wang, C., Yu, Q. and Wang, X. (2024) Crop Land Change Detection with MC&N-PSPNet. *Applied Sciences*, **14**, Article 5429. <https://doi.org/10.3390/app14135429>
- [5] Luo, H., Ma, Z., Wu, H., Li, Y., Liu, B., Li, Y., et al. (2023) Validation Analysis of Drought Monitoring Based on FY-4 Satellite. *Applied Sciences*, **13**, Article 9122. <https://doi.org/10.3390/app13169122>
- [6] Liu, M., Liu, J. and Hu, H. (2024) A Novel Deep Learning Network Model for Extracting Lake Water Bodies from Remote Sensing Images. *Applied Sciences*, **14**, Article 1344. <https://doi.org/10.3390/app14041344>
- [7] Affandi, E., Ng, T.F., Pereira, J.J., Ahmad, F. and Banks, V.J. (2023) Revalidation Technique on Landslide Susceptibility Modelling: An Approach to Local Level Disaster Risk Management in Kuala Lumpur, Malaysia. *Applied Sciences*, **13**, Article 768. <https://doi.org/10.3390/app13020768>
- [8] Huang, L., Yao, C., Zhang, L., Luo, S., Ying, F. and Ying, W. (2024) Enhancing Computer Image Recognition with Improved Image Algorithms. *Scientific Reports*, **14**, Article No. 13709. <https://doi.org/10.1038/s41598-024-64193-3>
- [9] Wang, M. and Liu, Z. (2024) Defense against Adversarial Attacks in Image Recognition Based on Multilayer Filters. *Applied Sciences*, **14**, Article 8119. <https://doi.org/10.3390/app14188119>
- [10] Huang, L., Chen, Y. and He, X. (2024) Spectral-Spatial Mamba for Hyperspectral Image Classification. *Remote Sensing*, **16**, Article 2449. <https://doi.org/10.3390/rs16132449>
- [11] Wang, M., Sun, Y., Xiang, J., Sun, R. and Zhong, Y. (2024) Adaptive Learnable Spectral-Spatial Fusion Transformer for Hyperspectral Image Classification. *Remote Sensing*, **16**, Article 1912. <https://doi.org/10.3390/rs16111912>
- [12] Zhang, L., Zhang, L. and Du, B. (2016) Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art. *IEEE Geoscience and Remote Sensing Magazine*, **4**, 22-40. <https://doi.org/10.1109/mgrs.2016.2540798>
- [13] Simonyan, K. (2014) Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv: 1409.1556. <https://doi.org/10.48550/arXiv.1409.1556>

-
- [14] He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 770-778. <https://doi.org/10.1109/cvpr.2016.90>
- [15] Shelhamer, E., Long, J. and Darrell, T. (2017) Fully Convolutional Networks for Semantic Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**, 640-651. <https://doi.org/10.1109/tpami.2016.2572683>
- [16] Lee, J., Wiratama, W., Lee, W., Marzuki, I. and Sim, D. (2023) Bilateral Attention U-Net with Dissimilarity Attention Gate for Change Detection on Remote Sensing Imageries. *Applied Sciences*, **13**, Article 2485. <https://doi.org/10.3390/app13042485>
- [17] Guo, Q., Feng, W., Zhou, C., Huang, R., Wan, L. and Wang, S. (2017) Learning Dynamic Siamese Network for Visual Object Tracking. 2017 *IEEE International Conference on Computer Vision (ICCV)*, Venice, 22-29 October 2017, 1781-1789. <https://doi.org/10.1109/iccv.2017.196>
- [18] Zhai, C., Wang, L. and Yuan, J. (2023) New Fusion Network with Dual-Branch Encoder and Triple-Branch Decoder for Remote Sensing Image Change Detection. *Applied Sciences*, **13**, Article 6167. <https://doi.org/10.3390/app13106167>
- [19] Yu, S., Tao, C., Zhang, G., Xuan, Y. and Wang, X. (2024) Remote Sensing Image Change Detection Based on Deep Learning: Multi-Level Feature Cross-Fusion with 3D-Convolutional Neural Networks. *Applied Sciences*, **14**, Article 6269. <https://doi.org/10.3390/app14146269>
- [20] Caye Daudt, R., Le Saux, B. and Boulch, A. (2018) Fully Convolutional Siamese Networks for Change Detection. 2018 *25th IEEE International Conference on Image Processing (ICIP)*, Athens, 7-10 October 2018, 4063-4067. <https://doi.org/10.1109/icip.2018.8451652>
- [21] Fang, S., Li, K., Shao, J. and Li, Z. (2022) SNUNet-CD: A Densely Connected Siamese Network for Change Detection of VHR Images. *IEEE Geoscience and Remote Sensing Letters*, **19**, 1-5. <https://doi.org/10.1109/lgrs.2021.3056416>
- [22] Chen, H., Qi, Z. and Shi, Z. (2022) Remote Sensing Image Change Detection with Transformers. *IEEE Transactions on Geoscience and Remote Sensing*, **60**, 1-14. <https://doi.org/10.1109/tgrs.2021.3095166>