doi:10.6041/j.issn.1000-1298.2024.05.018

# 融合注意力机制与多尺度信息的葡萄种植区变化检测

张宏鸣[1,2]　沈寅威[1]　阳　光[1]　孙志同[1]　刘康乐[1]　张二磊[1]
(1. 西北农林科技大学信息工程学院, 陕西杨凌 712100; 2. 宁夏智慧农业产业技术协同创新中心, 银川 750004)

**摘要**: 为准确获取葡萄空间变化信息,实现产业规划和可持续发展,针对葡萄种植区布局分散、面积不一,地物类型复杂,相应不同时相影像异质性较大,严重影响变化区域检测精度的问题,提出了一种融合注意力机制和多尺度信息的变化检测模型(Multiscale difference feature capture net, MDFCNet)。在 ResNet101 主干网络的基础上融合 SE(Squeeze and excitation)注意力机制,提升网络对遥感影像中变化特征提取的能力,抑制无关像素干扰。并且设计了交叉差异特征捕获(Cross difference feature capture,CDFC)模块,捕获具有密集上下文信息的差异特征来提升地物类型复杂情况下的变化检测精度,同时设计了监督集成注意力(Supervised ensemble attention,SEA)模块,逐层融合低层细节纹理特征和高层抽象语义特征来丰富多尺度特征,以此增强网络对布局分散、面积不一的种植区的检测能力。在构建的宁夏葡萄种植区变化数据集上进行实验,结果表明,相较于目前主流的 SNUNet、A2Net、DSIFN 和 ResNet-CD 变化检测模型,本文 MDFCNet 方法检测结果最优,相较于性能第 2 的模型,评价指标中交并比、召回率、F1 值和精确率分别提高 5.42、5.62、3.48、0.95 个百分点。通过消融实验也证明了融合各模块的有效性,相较于基础网络,增加 3 个模块使得交并比、召回率、F1 值和精确率分别提高 12.9、5.63、8.64、11.75 个百分点。本文模型提取出感受野更大的差异特征可为变化检测提供丰富的推断信息,融合的多尺度特征可以有效避免结果中误检测和漏检测问题,提高了变化区域的完整性和边缘细节保留,为背景复杂的大范围葡萄种植区的变化检测任务提供了解决思路。

**关键词**: 葡萄种植区; 遥感变化检测; 深度学习; ResNet; 注意力机制
**中图分类号**: TP751; S127　　**文献标识码**: A　　**文章编号**: 1000-1298(2024)05-0196-11　　**OSID**:

# Change Detection of Grape Growing Areas Based on Integrating Attention Mechanism and Multiscale Information

ZHANG Hongming[1,2]　SHEN Yinwei[1]　YANG Guang[1]　SUN Zhitong[1]　LIU Kangle[1]　ZHANG Erlei[1]
(1. *College of Information Engineering*, *Northwest A&F University*, *Yangling*, *Shaanxi* 712100, *China*
2. *Ningxia Smart Agricultural Industry Technology Collaborative Innovation Center*, *Yinchuan* 750004, *China*)

**Abstract**: Remote sensing technology for ground change detection has been widely used in the fields of agricultural planting planning and disaster situation assessment. For grapes, which is an important economic crop in China, accurately obtaining its spatial change information is crucial for industrial planning and sustainable development. Nevertheless, the dispersed arrangement of the grape growing areas, their diverse sizes, and the intricate nature of feature types, along with the heterogeneity among different temporal images, collectively contribute to a diminished accuracy in detecting areas of change. Therefore, a change detection model (Multiscale difference feature capture net, MDFCNet) based on attention mechanism and multiscale difference features was proposed. The main structure of the network adopted an encoder-decoder structure, which incorporated the squeeze and excitation (SE) attention module on the basis of ResNet101 backbone network to improve the network's ability to adequately extract change features from remote sensing images, suppressing interference from extraneous pixels. We also designed the cross difference feature capture (CDFC) module, it captured different features with dense

收稿日期: 2023-12-22　修回日期: 2024-02-27
**基金项目**: 国家重点研发计划项目(2020YFD1100601)、陕西省秦创原队伍建设项目(2023-ZDLNY-69)和陕西省重点研发计划项目(2023-YBNY-217)
**作者简介**: 张宏鸣(1979—),男,教授,博士,主要从事空间大数据管理与智慧农业研究,E-mail: zhm@nwafu.edu.cn

contextual information, thereby improving the accuracy of change detection in the case of complex feature types. While the supervised ensemble attention (SEA) module was designed to enrich multiscale features by fusing low-level detailed texture features and high-level abstract semantic features layer by layer to enhance the network's ability to detect small planting areas. Comparison and ablation experiments were conducted on the constructed change dataset of grape growing area, which was located within the city of Yinchuan, Ningxia Hui Autonomous Region. The experimental results showed that the MDFCNet method achieved the best detection results compared with the current state-of-the-art change detection methods of SNUNet, A2Net, DSIFN and ResNet-CD. Compared with the model with the 2nd highest performance (A2Net), the evaluation metrics of IoU, recall, F1 value and precision were improved by 5.42, 5.62, 3.48 and 0.95 percentage points, respectively. And the ablation experiments also demonstrated the effectiveness of fusing the modules. Compared with the base network, the addition of the three modules resulted in 12.9, 5.63, 8.64 and 11.75 percentage points increases in the evaluation metrics of IoU, recall, F1 value and precision respectively. The model extracted different features with larger sensory fields to provide rich inferential information for change detection, and the fused multiscale features can effectively avoid the problem of false detection and missed detection in the results. The extracted change areas were more complete and retain more edge detail, providing a solution to the task of change detection for the complex background of the wide range of grape growing areas.

**Key words**: grape growing areas; remote sensing change detection; deep learning; ResNet; attention mechanism

## 0　引言

葡萄作为我国重要的果树经济作物,在对外果品贸易中占据了重要地位[1]。获取葡萄种植区的空间变化信息对葡萄产业发展具有重要意义,可为产业规划提供关键的数据支持,方便管理者制定详细种植策略和优化资源配置,以适应市场的动态变化,从而确保葡萄产业稳健运行[2]。

遥感变化检测技术是使用同一地理位置不同时相的遥感影像,识别地表物体的差异信息,并提取其中的变化区域[3-4],目前已被广泛应用于生态环境检测、自然灾害评估和农业规划建设等多个领域。常见变化检测方法可以大致分为传统代数变化检测方法、机器学习变化检测方法和深度学习变化检测方法[5-6]。传统代数变化检测方法依赖变化阈值的选取[7-9],检测精度有限。而机器学习变化检测方法需要人为引入特征因子,对特征选择较为敏感[10-11]。近年来,深度学习算法在图像领域取得了巨大成功,因此也逐渐被应用于遥感领域,深度学习变化检测方法具有出色的特征提取能力,从而实现准确的检测[12]。WANG 等[13]基于卷积神经网络(Convolution neural network, CNN)叠加融合双时相影像的特征从而实现了变化检测。SHI 等[14]使用孪生 ResNet 作为网络编码器,减少了伪变化的影响。PENG 等[15]采用 UNet++的密集跳跃连接结构实现了变化检测,相较于传统的全卷积网络拥有更好的检测精度。尽管深度学习取得了显著进展,但面对背景复杂的葡萄种植区,采用单一的网络结构很难精确提取出葡萄变化区域,难点在于不同时相葡萄区影像存在异质性,像素值及饱和度差异都给葡萄区变化检测带来极大的困难,要求信息提取能力更强的模型来克服葡萄区中的干扰因素。目前深度学习变化检测方法的精度提升手段主要集中在提取双时相影像的差异特征和融合多尺度信息两方面[16]。

在双时相影像差异特征的提取中,WANG 等[17]利用孪生 FCN 网络提取影像特征,并通过距离度量来衡量差异特征之间的变化,扩大类间差异。ZHANG 等[18]构建了一种深度监督融合网络(DSIFN),将提取的特征输入差分判别网络提取差异特征,并与多尺度深度特征融合,准确性优于其他方法。SONG 等[19]基于 U-net 设计了差异增强模块来提取变化的差异特征,生成了更为准确的检测结果。这些方法提取了双时相影像中的差异特征,但是忽略了上下文关系,并未解决葡萄种植区地物类型复杂的难点,不能为葡萄种植变化区域提供丰富推断信息,导致提取的葡萄变化区域存在大量误检测和漏检测区域;在融合多尺度信息中,樊玮等[20]提出了一种多尺度深度特征融合的变化检测网络,提升在真实场景下的检测效果。FANG 等[21]提出了一种密集连接的孪生网络用于变化检测(SNUNet),构建了集成通道注意力模块(Ensemble channel attention module, ECAM)将高低层特征融合,缓解了浅层信息丢失的问题。LI 等[22]构建了一种轻量级变化检测网络(A2Net),有效地将多级特征从高层次聚集到低层次,提高了检测准确性。可

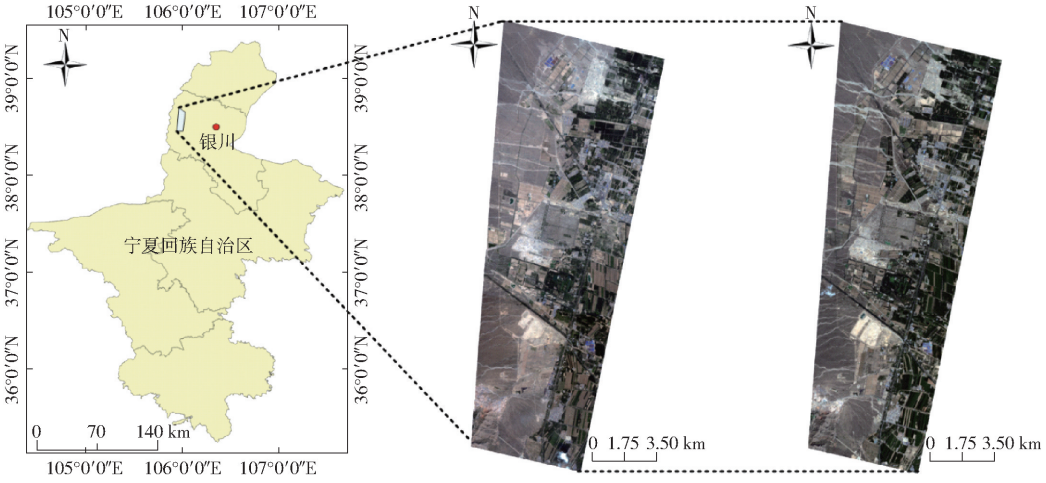
见融合多尺度信息可以保留浅层空间信息和高层语义特征,但是葡萄种植区存在地物分布不规则的难点,边界信息模糊,在葡萄差异特征融合过程中会出现噪声冗余以及语义差异等问题,需要进一步改进模型,以实现检测结果的完整性。

因此本文提出一种融合注意力机制和多尺度信息的葡萄种植区变化检测模型(Multiscale difference feature capture net, MDFCNet),由于遥感影像中含有复杂地物,因此选取性能良好的 ResNet 作为模型的主干网络,能够灵活捕捉不同尺度和形状的对象,同时在每层中融合 SE(Squeeze and excitation)注意力机制增强网络,以提取更具表征性的特征;设计交叉差异特征捕获(Cross difference feature capture, CDFC)模块生成具有丰富上下文信息的差异特征,并通过监督集成注意力(Supervised ensemble attention, SEA)模块融合葡萄种植区的低层纹理特征和高层语义特征,并减少融合过程中的噪声冗余及特征差异。在背景复杂并且类间相似性强的双时相遥感影像中,达到高精度的葡萄种植区变化检测。探索此方法与其他先进方法的检测精度,并进行消融实验验证各个模块效果,以期为葡萄种植区变化信息获取提供新的解决方法。

## 1 数据

### 1.1 研究区域

本研究区域位于东经 105°46′~106°08′,北纬 38°28′~38°44′,地处黄河上游宁夏平原中部,宁夏回族自治区银川市内,位置示意图如图 1 所示。该区域属于中带干旱气候区,地形为冲积平原,海拔1 010~1 150 m,年平均降水量 203 mm,日夜温差10~15℃,年平均日照时数2 800~3 000 h,气候适宜葡萄生长,种植区面积较大,由于其优越的自然条件使其成为世界优质的葡萄种植黄金地带之一[23]。



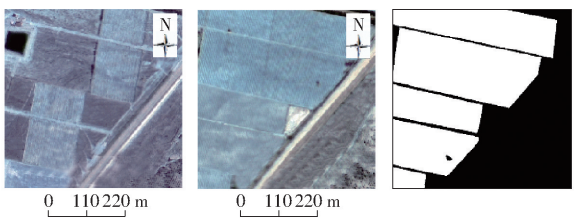

(a)研究区域 (b)2016年遥感影像 (c)2019年遥感影像

图 1 研究区示意图

Fig. 1 Study area diagrams

研究区内的葡萄种植区由于空间跨度大,其中地物类型丰富,还包含小麦、玉米、枸杞和其他类型果园等作物,会存在不同地物类型表现出相似条带纹理特征的“同谱异物”问题[24],如图 2 所示。图 2a 中包含的地块为林地,图 2b 中包含的地块为葡萄种植区,地物类型不同,但在影像中均存在竖状条纹;并且由于影像时间跨度大,同一片区域不同时相会存在像素值及饱和度差异的问题[25],葡萄种植区的光谱特征并不固定。这些问题都会严重影响后期变化检测的结果。

在研究区内,葡萄种植区的变化主要表现为与其他类型果园、林地之间的动态变化。这一变化趋势涵盖较大面积的葡萄区域,主要分布于村镇周边,并且还存在小面积葡萄区与裸地、建筑之间的变化,




(a) T1时相影像 (b) T2时相影像 (c) 标签图像

图 2 影像-标签对示意图

Fig. 2 Diagrams of image - label pair

这些变化区域呈现较为分散的特点,较难精准识别。

### 1.2 数据集构建

实验采用 2 幅高分 2 号遥感影像,分别成像于 2016 年 8 月 4 日和 2019 年 7 月 25 日,每幅影像尺寸为 14 351 像素×29 124 像素。高分 2 号卫星的主要参数如表 1 所示。

**表 1　高分 2 号卫星有效载荷参数**

**Tab. 1　GF－2 satellite payload parameters**

| 传感器 | 波段 | 谱段/m | 空间分辨率/m | 幅宽/km | 重访时间/d |
|---|---|---|---|---|---|
| 全色多光谱相机 | 全色波段 | 0.45～0.90 | 1 | 45 | 5 |
| | 蓝波段 | 0.45～0.52 | 4 | | |
| | 绿波段 | 0.52～0.59 | | | |
| | 红波段 | 0.63～0.39 | | | |
| | 近红外波段 | 0.77～0.89 | | | |

由于卫星拍摄的遥感影像在地物空间特征上会发生颜色偏差及位置偏移，需要对原始的前后时相影像进行大气校正、正射校正、影像配准和影像融合等预处理操作，使得影像更加接近地表的真实情况，可以得到空间分辨率 1 m、4 波段（红、绿、蓝、近红外）的融合影像数据，综合了丰富的光谱信息和空间信息，为后续变化检测夯实良好的数据基础，同时为使葡萄园地物特征更加明显，有助于进行解译与分析，本文图像采用假彩色呈现[26]。

为确保所构建数据集的质量，本次数据标注基于前期实地调研，使用外业精灵软件在待考察地点进行 GPS 标点，再前往实地考察，通过现场无人机影像比对，确定地块类别，后期由专业团队再对 2016 年和 2019 年遥感影像的葡萄种植区域分别进行标注，接着对生成的标签图像进行差值运算。由于本文方法属于二类变化检测[27-28]，针对特定葡萄种植区提取出变化和未变化区域，因此将葡萄园到葡萄园和非葡萄园到非葡萄园归为未变化类，对应标签像素值为 0；将非葡萄园到葡萄园和葡萄园到非葡萄园归为变化类，对应标签像素值为 255，最终得到数据标签。将其与前后时相影像配对后，通过滑动窗口裁剪成 512 像素 ×512 像素，得到最终的变化检测影像-标签对，如图 2 所示。T1 时相的非葡萄种植区变化为 T2 时相的葡萄种植区，因此两时相影像配对的标签图对应位置像素值为 255（变化）。

将部分未变化的影像-标签对剔除，共得到 1 254 个影像标签对，接着通过数据增广技术生成丰富的差异化数据，对数据进行旋转变换和反转变换后，共得到 10 072 个影像-标签对，将其按照比例 9∶2∶2 划分成训练集、验证集和测试集，训练集包含 7 016 个影像-标签对，验证集和测试集包含 1 528 个影像-标签对。

## 2　方法

### 2.1　ResNet 改进

基于葡萄种植区的遥感影像包含丰富光谱和空间信息，本文采用经典 ResNet101 网络结构作为主干网络中的编码器部分，采用残差结构保护信息的完整性[29]，提取更加丰富的语义特征，有助于复杂的葡萄种植区变化检测任务，残差结构计算式为

$$H(x) = F(x, \{W_i\}) + x \tag{1}$$

其中

$$F(x, \{W_i\}) = W_2\sigma(W_1 x) \tag{2}$$

式中　$x$——输入特征　　$H(x)$——输出特征

$F$——学习的残差映射

$\sigma$——ReLU 函数

$W_i$——权重参数，$i = 1, 2$

由于输入模型的 2 幅遥感影像采集于不同年份，季节相近但时间跨度较大，彼此之间存在异质性[12]。而葡萄种植区内部纹理特征一般表现为紧密竖形条纹状，不同于其他地物类型。

因此在 ResNet 原始的瓶颈结构（Bottleneck Layer）中添加 SE 注意力模块[30]，通过压缩（Squeeze）和激励（Excitation）操作来强调与葡萄相关的纹理，同时抑制无关像素，减少原始影像中像素值及饱和度差异对条状特征提取带来的干扰，使得提取的纹理更具有表征性，能更为显著地呈现出葡萄种植区与其他地物之间的区别，其中压缩与激励操作步骤为

$$z = F_{sq}(U_c) = \frac{1}{H \times W}\sum_{i=1}^{H}\sum_{j=1}^{W} U_c(i,j) \tag{3}$$

$$F_{ex}(z, W) = \delta(W_2\sigma(W_1 z)) \tag{4}$$

式中　$F_{sq}$、$F_{ex}$——压缩和激励操作

$H$、$W$——图像高度和宽度

$U_c(i,j)$——图像通道 $c$ 上位置 $(i,j)$ 处像素

$\delta$——Sigmoid 激活函数

改进 ResNet 网络的每层如图 3 上方所示的模块。同时，为减少编码器提取复杂葡萄区特征过程中信息丢失的问题，将编码器中 Layer 1 至 Layer 4 不同层次提取的特征，跳跃连接至交叉差异特征捕获（CDFC）模块，通过 CDFC 构造出不同层次的葡萄种植区差异特征，并在后续的监督集成注意力模块（SEA）中进一步运算，从而使得不同层次的葡萄区差异特征逐层融合，针对葡萄种植区空间分布及地物类型，更好地解析变化区域，模型结构如图 3 所示。

### 2.2　交叉差异特征捕获（CDFC）模块

对葡萄种植区进行变化检测就是对遥感影像中的每个像素点进行分类，但是双时相遥感影像中存在大量与葡萄种植区纹理相似的林地与荒地，并且种植区域分散，中间零散分布着房屋、道路等各类地物，图像经过特征提取后每个像素点的感受野较小，这对于变化检测推断提供的信息十分有限，往往会

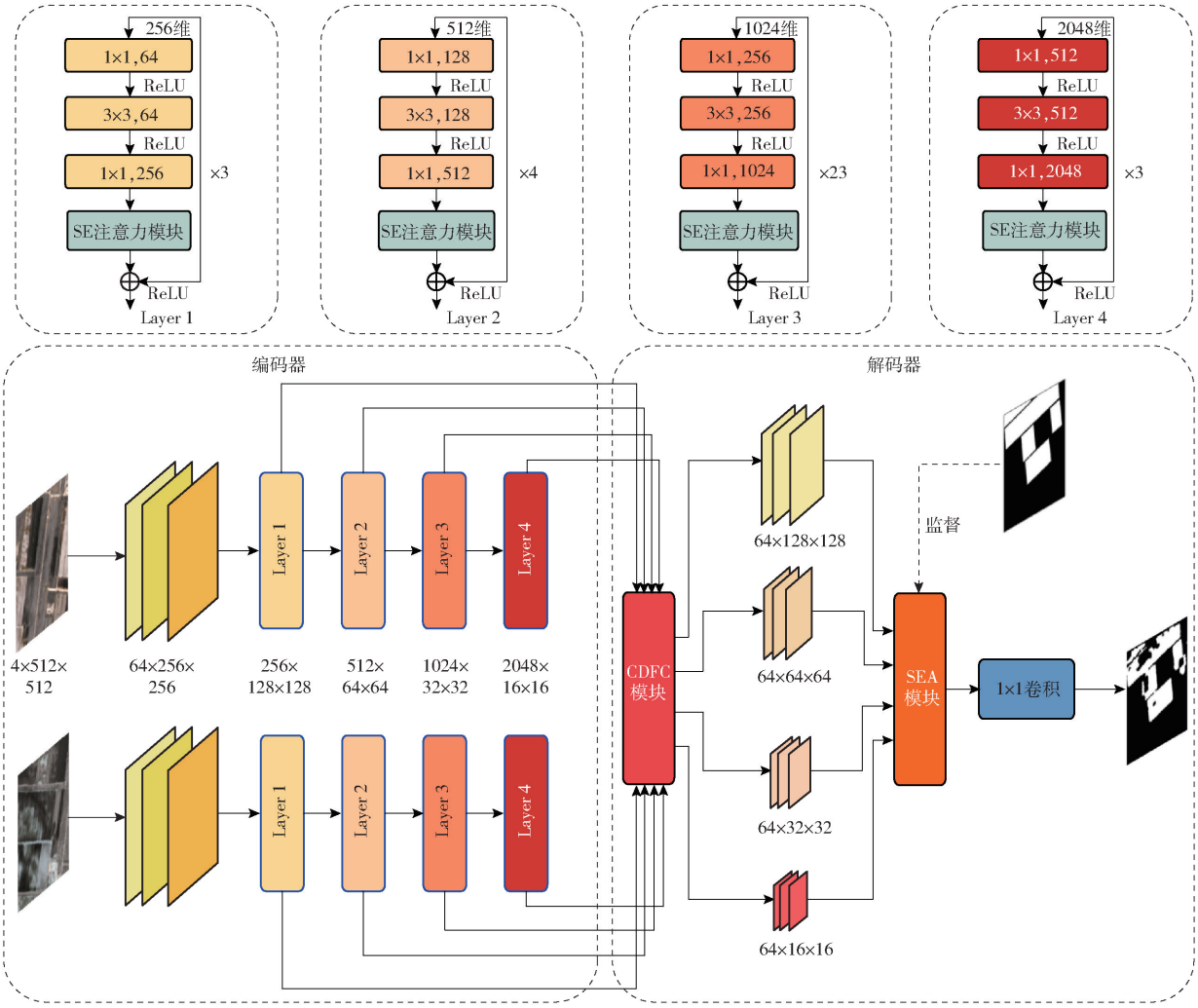



图 3 改进的网络结构示意图

Fig. 3 Diagram of improved network structure

存在大量的误检测或者漏检测问题,难以对葡萄种植区变化信息进行有效检测。

PSPNet 的金字塔池化模块[31]和 Deeplab 系列[32]的提出一定程度上解决了感受野不足问题,但是它们不能生成逐像素的上下文信息。HUANG 等[33]提出了 CCNet,利用像素点与特征图其他位置的关系进行加权,在控制计算量的同时生成了密集的上下文信息。基于此,本研究设计了适合葡萄区变化检测任务的交叉差异特征捕获(CDFC)模块,通过建立全局上下文信息,针对 2 个时相的葡萄影像特征图的每个像素点,以全局的角度考虑葡萄区域与其他地物间的关系。图 4a 为十字交叉注意力(Criss cross attention,CCA)模块,图 4b 为 CDFC 模块整体结构,与传统的十字交叉注意力模块相比,CDFC 模块更注重葡萄种植区域的特殊性。

$\boldsymbol{f}_Q \times \boldsymbol{f}_K$ 将特征图通道方向的每个特征向量与其同行同列的特征向量相乘,经过激活函数生成 $\boldsymbol{f}_M$,它记录了特征图中每个像素点和同行同列像素的空间依赖度,利用其对 $\boldsymbol{f}_V$ 进行像素加权,并与原始特

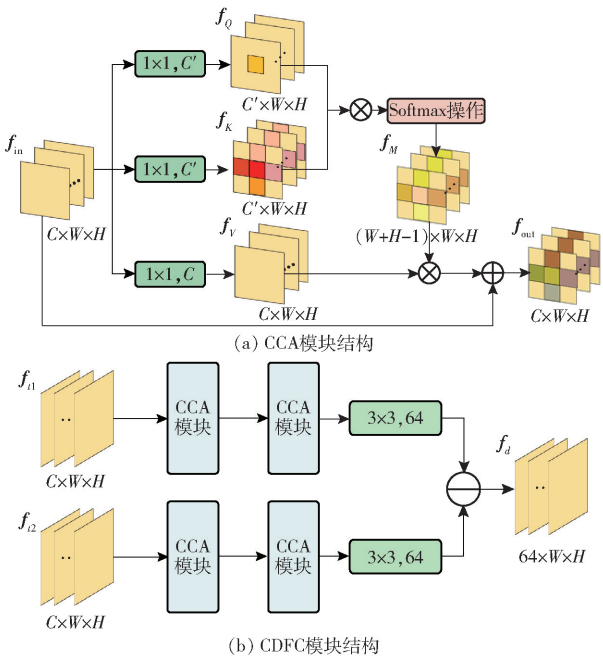



图 4 交叉差异特征捕获模块(CDFC)示意图

Fig. 4 Schematics of cross difference feature capture (CDFC) module

征图相加，得到的特征图每个像素点都包含了密集的上下文信息，使得葡萄区的像素点中不仅包含相关的纹理特征，还包含与周边地物之间的空间关系。模块详细描述为

$$\boldsymbol{f}_Q = \boldsymbol{f}_K = \mathrm{conv}_{1\times1}^{C'}(\boldsymbol{f}_{\mathrm{in}}) \tag{5}$$

$$\boldsymbol{f}_V = \mathrm{conv}_{1\times1}^{C}(\boldsymbol{f}_{\mathrm{in}}) \tag{6}$$

$$\boldsymbol{f}_M = \varepsilon(\boldsymbol{f}_Q \times \boldsymbol{f}_K) \tag{7}$$

$$\boldsymbol{f}_{\mathrm{out}} = \boldsymbol{f}_{\mathrm{in}} + \boldsymbol{f}_M \times \boldsymbol{f}_V \tag{8}$$

式中　$\boldsymbol{f}_Q$、$\boldsymbol{f}_K$、$\boldsymbol{f}_V$、$\boldsymbol{f}_M$——图中对应位置特征图

$\boldsymbol{f}_{\mathrm{in}}$、$\boldsymbol{f}_{\mathrm{out}}$——输入、输出特征

$\mathrm{conv}_{1\times1}^{C}$——卷积核个数为 $C$，大小为 1×1 的卷积函数

$\varepsilon$——softmax 函数

接着对两时相特征分别进行 CCA 模块运算，再通过特征融合捕获具有丰富信息的差异特征，使得特征像素点具有较大的感受野和长距离的空间依赖，包含周边地物特征与葡萄区纹理特征的区别，相较于单一的纹理特征，更有利于后阶段的变化区域推断，详细步骤为

$$\boldsymbol{f}_{ti}^{*} = \mathrm{conv}_{3\times3}^{64}(F_c(F_c(\boldsymbol{f}_{ti}))) \quad (i=1,2) \tag{9}$$

$$\boldsymbol{f}_d = |\boldsymbol{f}_{t1}^{*} \ominus \boldsymbol{f}_{t2}^{*}| \tag{10}$$

式中　$\boldsymbol{f}_{ti}^{*}$——经过 CCA 模块处理的第 $i$ 个时相的特征图

$\boldsymbol{f}_{ti}$——第 $i$ 个时相的特征图

$F_c$——CCA 模块函数

$\ominus$——元素减法运算

## 2.3　监督集成注意力(SEA)模块

前文所述的 CDFC 模块分别对 ResNet 提取的各层特征进行处理，生成了深层和浅层的差异特征。深层差异特征具有丰富的语义信息，而浅层差异特征具有较多的空间细节信息，但是由于研究区域背景复杂并且包含大量小面积的葡萄种植区，仅凭借单尺度差异特征进行变化检测会导致边缘细节信息丢失以及准确率低等问题。

SNUNet 的稠密连接结构和 A2Net 的渐进特征聚合模块都为高低层特征信息的融合提供了解决思路。基于这两项研究，对其进行改进，设计出更适用于葡萄种植区多尺度差异特征融合的监督集成注意力(SEA)模块，将高层抽象特征和低层细节特征逐层融合，在完整表达葡萄种植区变化信息的同时，保留细节边缘信息，提升模型对小面积葡萄区变化的完整检测能力。

$\boldsymbol{f}_{d1}$中包含的是葡萄区中明显的纹理以及边缘特征，$\boldsymbol{f}_{d4}$中包含的是葡萄区中抽象语义特征，通过将$\boldsymbol{f}_{d4}$经过相应卷积和激活操作后，生成初步的葡萄变化区域，但此时内部仍缺失区域的细节信息，因此将其与自身元素相乘，去除和葡萄变化区域无关的噪声和冗余特征后，再逐层地向上融合，直到融合至最低层的$\boldsymbol{f}_{d1}$特征。SEA 模块结构如图 5 所示，模块详细描述为

$$\boldsymbol{f}_{di}^{*} = \begin{cases} \boldsymbol{f}_{di} + \mathrm{up}(\boldsymbol{f}_{di+1}^{*} \times \mathrm{conv}_{1\times1}(\delta(\mathrm{conv}_{1\times1}(\boldsymbol{f}_{di+1}^{*})))) & (i=1,2,3) \\ \mathrm{up}(\mathrm{conv}_{1\times1}(\boldsymbol{f}_{di})) & (i=4) \end{cases} \tag{11}$$

式中　$\boldsymbol{f}_{di}$——经过 CDFC 模块运算的第 $i$ 层差异特征

$\mathrm{up}(\,)$——上采样操作

$\boldsymbol{f}_{di}^{*}$——融合了深层信息的第 $i$ 层差异特征

但由于上述运算均在空间层面进行，因此添加通道注意力机制在通道维度进一步辅助高低层差异特征融合，计算式为

$$F_{\mathrm{cam}}(\boldsymbol{f}) = \boldsymbol{f} \times \delta(W_1(W_2(\mathrm{AvgPool}(\boldsymbol{f}))) + W_1(W_2(\mathrm{MaxPool}(\boldsymbol{f})))) \tag{12}$$

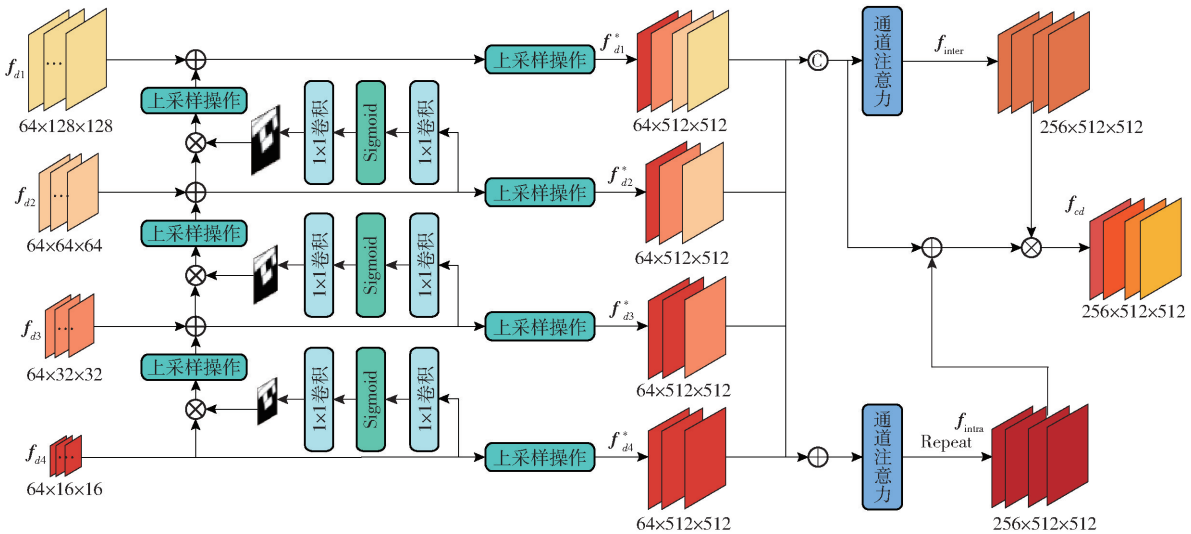



图 5　改进的监督集成注意力模块(SEA)示意图

Fig. 5　Schematic of improved supervised ensemble attention (SEA) module

$$\boldsymbol{f}_{\text{intra}} = F_{\text{repeat}}\left(F_{\text{cam}}\left(\sum_{i=1}^{4}\boldsymbol{f}_{di}^{*}\right)\right) \tag{13}$$

$$\boldsymbol{f}_{\text{inter}} = F_{\text{cam}}([\boldsymbol{f}_{d1}^{*}, \boldsymbol{f}_{d2}^{*}, \boldsymbol{f}_{d3}^{*}, \boldsymbol{f}_{d4}^{*}]) \tag{14}$$

$$\boldsymbol{f}_{cd} = ([\boldsymbol{f}_{d1}^{*}, \boldsymbol{f}_{d2}^{*}, \boldsymbol{f}_{d3}^{*}, \boldsymbol{f}_{d4}^{*}] + \boldsymbol{f}_{\text{intra}}) \times \boldsymbol{f}_{\text{inter}} \tag{15}$$

式中 AvgPool、MaxPool——平均池化、最大池化

$F_{\text{cam}}$——通道注意力模块运算

$F_{\text{repeat}}$——复制拼接操作

[ ]——拼接操作

将高低层差异特征相加和拼接,通过通道注意力提取通道关系后与自身相乘,得到特征图 $\boldsymbol{f}_{\text{intra}}$ 和 $\boldsymbol{f}_{\text{inter}}$,分别表示葡萄区高低层差异特征之间的组间关系及组内关系,与原始融合特征相乘,得到了克服高低层语义差异和定位差异的特征图 $\boldsymbol{f}_{cd}$,它既保留了葡萄区低层细节边缘信息,也包含了其内部的丰富语义特征,极大提高了葡萄种植区变化检测的完整性和细节性。

### 2.4 实验环境与指标

#### 2.4.1 实验运行环境

本文模型编码器部分采用迁移学习,使用 ImageNet 数据集上的预训练网络,初始化主干网络权重。实验运行的硬件平台为 Intel(R) Xeon(R) Gold 6226R CPU @ 2.90 GHz,操作系统为 Ubuntu 20.04.4 LTS,搭载 2 块 NVIDIA GeForce RTX 3090 显卡,使用 Pytorch 深度学习框架实现模型构建与调试。

实验超参数设置如下:训练批量大小为 8,使用双卡并行训练,训练最大轮数为 150,采用 Poly 策略动态调整学习率,初始学习率设置为 0.001,优化算法采用 Adam。

损失函数采用 BCE 损失函数和 Dice 损失函数,具体计算式为

$$L_{\text{BCE}} = Y\log_n\hat{Y} + (1-Y)\log_n(1-\hat{Y}) \tag{16}$$

$$L_{\text{Dice}} = 1 - \frac{2Y\hat{Y}}{Y+\hat{Y}} \tag{17}$$

式中 $Y$——真值 $\hat{Y}$——预测值

$n$——对数函数底数

#### 2.4.2 实验评价指标

实验采用精确率(Precision)、召回率(Recall)、F1 值和交并比(Intersection over union, IoU)4 种变化检测任务中常用的指标来评价模型性能。

## 3 结果与分析

### 3.1 实验方案

为了验证本研究提出的 MDFCNet 在葡萄种植区变化检测中的优越性,选取目前变化检测领域中最先进和最常用的检测模型进行对比,对 SNUNet[21]、A2Net[22]、DSIFN[18]和 ResNet-CD[29]进行训练和测试,将其识别结果与本文 MDFCNet 模型进行对比,通过变化检测定性分析和定量分析分别验证本研究模型的性能。

本文模型融合 SE 模块、CDFC 模块和 SEA 模块以提升模型在葡萄种植区变化检测的性能,为了探寻每个模块对于最终结果的提升效果,进行消融实验,分别移除 SE 模块(MDFCNet without SE,简称模型 b)、移除 CDFC 模块(MDFCNet without CDFC,简称模型 c)、移除 SEA 模块(MDFCNet without SEA,简称模型 d)以及移除所有模块(MDFCNet without All,简称模型 a)测试性能,以此验证各模块的设计意义。同时本文也探究不同主干网络(MDFCNet with ResNet50,简称模型 e)的检测效果。

### 3.2 对比实验

本文提出的 MDFCNet 方法与其他 4 种方法在本文数据集上的变化检测结果如图 6 所示。通过种植区域背景复杂、布局分散和面积细小等 8 个不同的场景展现出各方法变化区域的预测效果,实验评价指标如表 2 所示,以此阐释本文方法相较于目前主流变化检测方法的优势。

#### 3.2.1 变化检测定性分析

通过可视化展示出 MDFCNet 模型与其他 4 种主流模型检测结果对比(图 6),白色和黑色部分表示模型正确检测的区域,而红色和青色部分表示模型误检测和漏检测的部分。从整体来看,本文方法红色和青色部分最少,更接近于图像标签,对于葡萄种植区变化部分检测最为准确。

由于双时相影像存在异质性,大量像素值及饱和度差异会影响变化检测的结果,如图 6 场景 1、2 所示,SNUNet 和 A2Net 未能很好地检测出葡萄区主体的变化区域(内部存在大量青色区域),并且会将研究区域中像素值发生变化的田地错误检测成变化的葡萄区(周围存在凌乱红色区域),而本文方法可以很好地避免这些影响,提取出更有表征性的变化特征,准确完整地检测出变化的葡萄种植区。并且,影像中存在许多与葡萄种植区纹理相似的田地和荒地,这些会导致检测结果中存在大量的误检测和漏检测,如图 6 场景 5、6 所示,ResNet-CD、SNUNet、A2Net 和 DSIFN 都将一些纹理相似的田地和田埂误检测成变化区域(存在大块红色区域),并且会漏检测部分变化区域,而本文的 MDFCNet 由于获取了密集的上下文信息,增大了感受野,捕获的差异特征包含丰富的变化信息,使得检测结果更为准确。同时,ResNet-CD、SNUNet、A2Net 和 DSIFN 在小面积葡

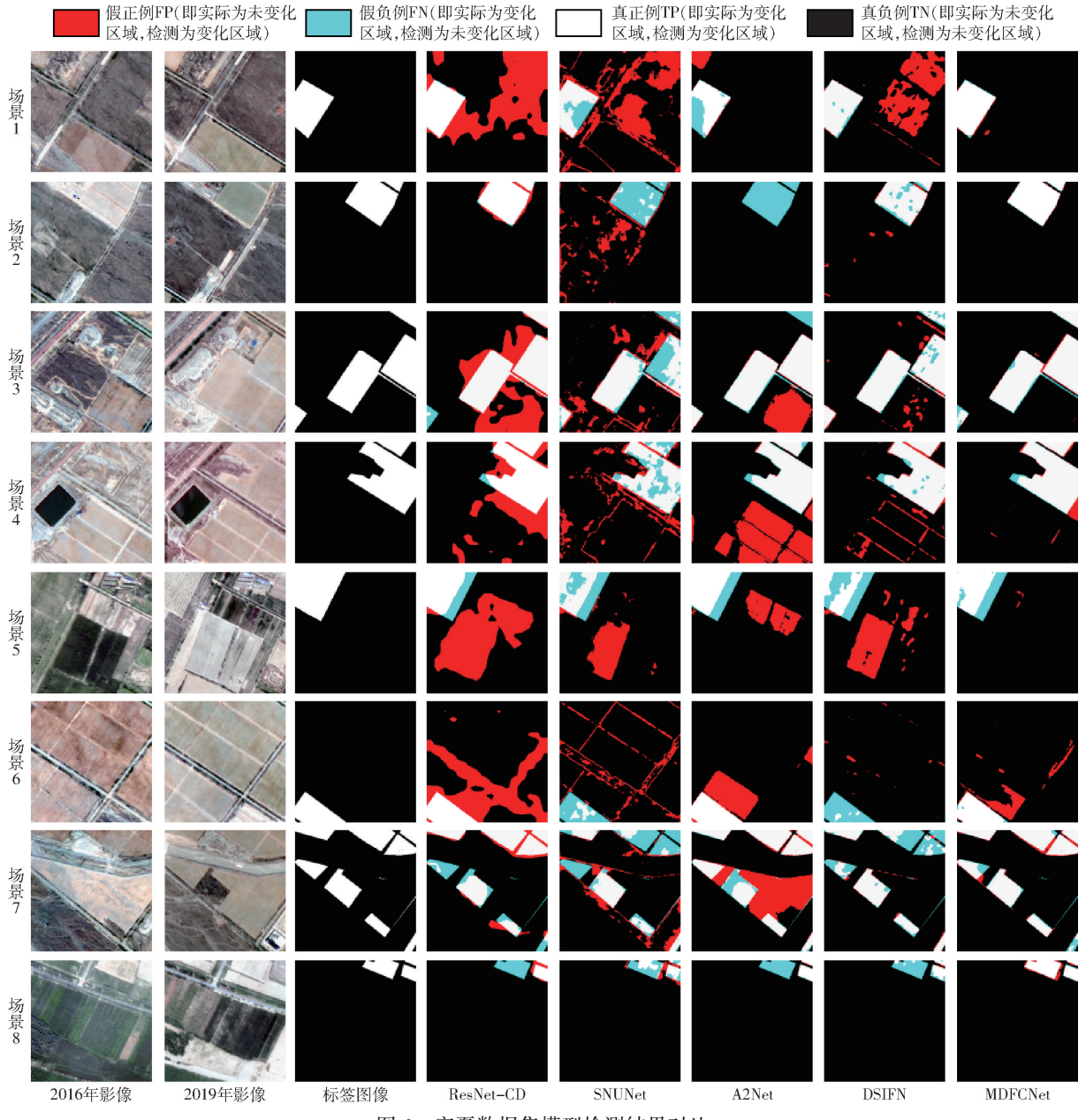



图 6　宁夏数据集模型检测结果对比

Fig. 6　Comparison of model test results on Ningxia Dataset

**表 2　宁夏数据集不同方法对比实验结果**

**Tab. 2　Comparison test results of different methods on Ningxia Dataset**

| 模型 | 年份 | 交并比/% | 召回率/% | F1 值/% | 精确率/% | 显存占用量/MB |
|---|---|---|---|---|---|---|
| A2Net | 2023 | 73.95 | 80.57 | 85.02 | 89.99 | 3.78 |
| SNUNet | 2022 | 68.29 | 82.03 | 81.15 | 80.29 | 27.10 |
| DSIFN | 2020 | 70.99 | 80.02 | 83.03 | 86.28 | 50.70 |
| ResNet－CD | 2015 | 66.47 | 80.56 | 79.86 | 79.19 | 51.60 |
| MDFCNet | 2023 | 79.37 | 86.19 | 88.50 | 90.94 | 58.70 |

注：年份表示方法发表年份。

萄区变化的检测上存在大量的漏检测(存在大量小块青色区域)，如图 6 场景 7、8 所示，可见单一尺度特征不能很好地保留细小地物特征，本文 MDFCNet 将高层语义特征和低层空间信息逐层融合，得到了最好的检测结果，同时由于在逐层融合时增加了注意力机制，克服了存在的语义差异，使得 MDFCNet 的检测结果更加完整，如图 6 场景 3、4 所示，其余方法的检测结果都存在边缘细节较差以及内部出现孔洞的问题，本文 MDFCNet 在葡萄种植区变化检测中表现更为优异。

同时，为进一步验证本文 MDFCNet 方法的鲁棒性，选取 2016 年 8 月 4 日和 2020 年 5 月 20 日两时相的非标签葡萄种植区进行检测对比，尺寸为 1 024 像素 ×1 024 像素，如图 7 所示。其中，白色表示预测为变化的区域，黑色表示预测为未变化的区域，可见 SNUNet 和 DSIFN 的检测效果较差，部分葡萄变化地块并未很好地识别出来，ResNet－CD 和 A2Net 虽然将大部分变化区域提取出来，但边缘并

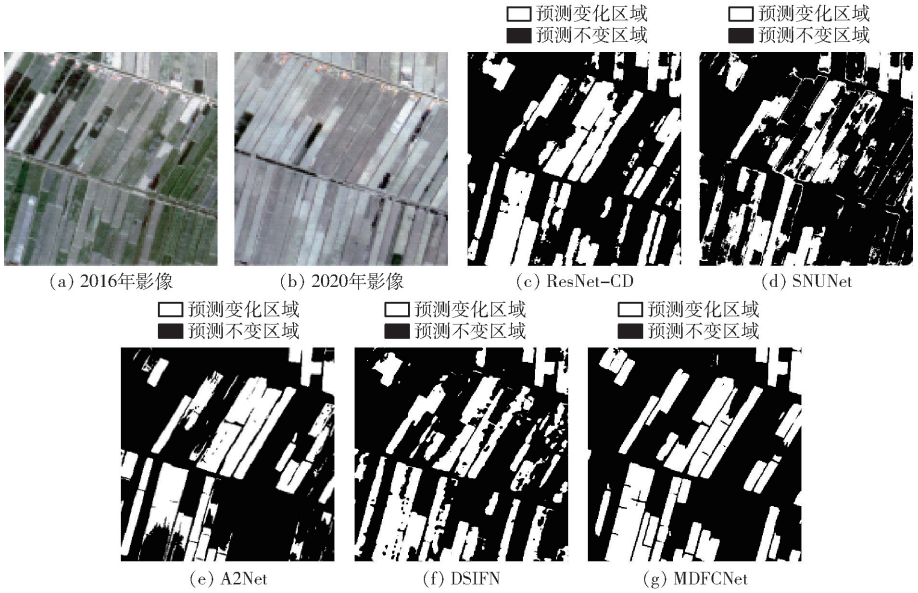



图 7 宁夏非标签区域检测结果对比

Fig. 7 Comparison of test results on no-label areas in Ningxia

不连续,存在细节丢失的问题,而本文 MDFCNet 方法提取的变化区域更加完整,边缘细节也更为精确,证明了本文方法出色的检测性能与鲁棒性。

### 3.2.2 变化检测定量分析

通过 A2Net、SNUNet、DSIFN、ResNet - CD 和本文 MDFCNet 在葡萄种植区变化检测数据集上的定量评估比较,计算了各个模型显存占用量,如表 2 所示。从表 2 可知,本文所提出的 MDFCNet 相较于现有变化检测方法在葡萄种植区上取得了最优结果。

对比模型中,A2Net 的显存占用量最少,但是由于采用的轻量级网络并不能很好地提取出表征性特征,很难在此基础上提高检测的精度与细节。本文 MDFCNet 相较于 A2Net,交并比、召回率、F1 值、精确率分别提高 5.42、5.62、3.48、0.95 个百分点。SNUNet、DSIFN 和 ResNet - CD 的显存占用量依次增多,但检测结果反而是 DSIFN 效果最好,交并比为 70.99%、召回率为 80.02%、F1 值为 83.03%、精确率为 86.28%。基础的 ResNet - CD 最差,交并比为 66.17%、召回率为 80.56%、F1 值为 79.86%、精确率为 79.19%,可见复杂的模型在提取丰富差异特征后,会由于冗余信息干扰和语义差异反而导致检测效果不足,本文改进后的 MDFCNet 很好地避免了这种情况,在显存占用量仅提升 7.1 MB 的情况下,达到了最好的检测结果,交并比为 79.37%、召回率为 86.19%、F1 值为 88.50%、精确率为 90.94%,相较于 DSIFN 的结果,交并比、召回率、F1 值、精确率分别提升 8.38、6.17、5.47、4.66 个百分点,精度提升明显,说明了改进的 MDFCNet 在葡萄种植区变化检测上具有一定的优越性。

## 3.3 消融实验

为验证本文提出的各个模块的有效性,进行了消融实验,分别去除各个模块进行精度定量分析,并探究不同主干网络对于最终结果的影响,消融实验定性结果如图 8 所示,定量结果如表 3 所示。

SE 模块通过抑制与变化无关的像素,增加了表征性特征提取的能力。为验证其有效性,进一步评估了去除瓶颈结构中的 SE 模块的性能,如图 8 的场景 1 所示,去除 SE 模块的模型 b 将对比度范围不同的葡萄区识别成变化区域(存在大块红色区域),检测结果明显不如 MDFCNet,表明在编码器阶段增加 SE 模块可以有效解决双时相影像异质性问题,有助于变化检测精度的提升,如表 3 中模型 b 所示。

CDFC 模块在捕获差异特征的同时,构建了密集的上下文信息,为解码器阶段提供了丰富推断信息。如图 8 的场景 2 所示,去除 CDFC 模块的模型 c 由于忽略了上下文信息,导致出现大范围的误检测(存在大面积红色区域),MDFCNet 充分利用了空间信息,有效避免了此类问题,相较于模型 c,其 4 个指标分别提升 3.71、3.19、2.36、1.41 个百分点,在 3 个模块中对精度提升最高,如表 3 中模型 c 所示,证明了 CDFC 模块在本文方法中的有效性。

SEA 模块通过融合高低层差异特征信息,避免了冗余噪声和语义差异的影响,如图 8 的场景 3 所示,去除 SEA 模块的模型 d 由于未充分利用多尺度信息,检测结果的主体内部存在大量孔洞(不规则的青色区域),并且边缘不连续,而 MDFCNet 的识别结果更加完整,相较于模型 d,4 个指标分别提升 3.2、2.31、2.03、1.71个百分点,证明了多尺度特征

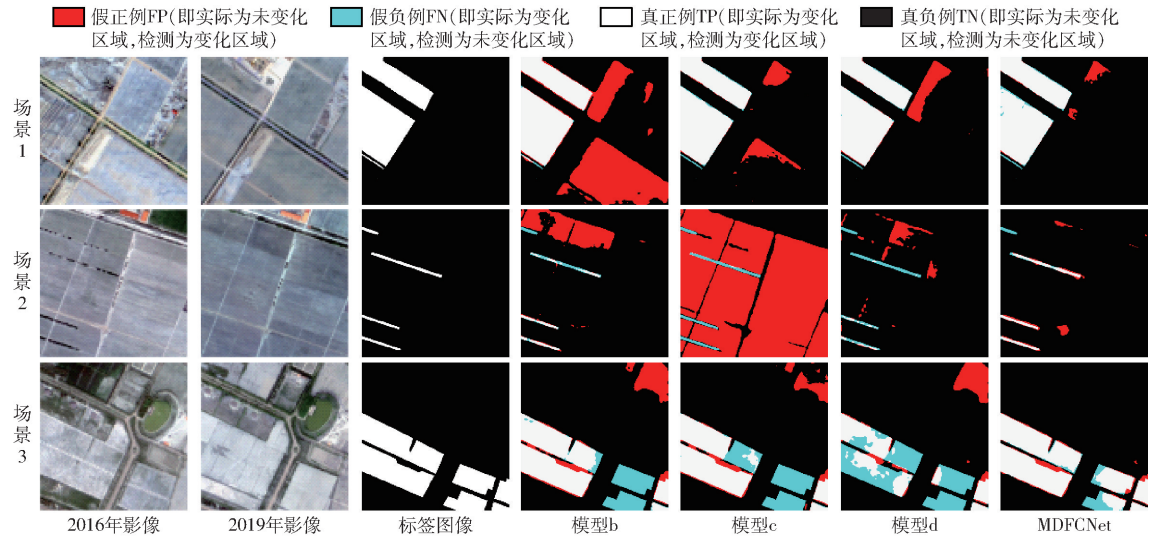



图 8　宁夏数据集消融实验对比图

Fig. 8　Comparison of ablation experiments on Ningxia Dataset

**表 3　消融实验结果**

**Tab. 3　Results of ablation experiments**

| 模型 | 交并比/% | 召回率/% | F1 值/% | 精确率/% | 显存占用量/MB |
|---|---|---|---|---|---|
| 模型 a | 66.47 | 80.56 | 79.86 | 79.19 | 51.6 |
| 模型 b | 76.68 | 87.01 | 86.80 | 86.60 | 53.9 |
| 模型 c | 75.66 | 83.00 | 86.14 | 89.53 | 56.7 |
| 模型 d | 76.17 | 83.88 | 86.47 | 89.23 | 58.6 |
| 模型 e | 73.95 | 80.42 | 85.02 | 90.19 | 37.4 |
| MDFCNet | 79.37 | 86.19 | 88.50 | 90.94 | 58.7 |

在变化检测中的重要性。

在去除所有模块后，检测精度达到最低，4 个指标相较于 MDFCNet 分别下降 12.9、5.63、8.64、11.75 个百分点，如表 3 中模型 a 所示，充分证明 3 个模块融合进原始模型的有效性，增加了检测精度，检测效果也大幅提升。与此同时，本文也进行了不同主干网络的探究实验，将主干网络中 101 层卷积加全连接（ResNet101）替换成 50 层卷积加全连接（ResNet50）后，虽然模型显存占用量下降了，但各项指标也受到了很大程度的影响，如表中模型 e 所示，证明了浅层神经网络并不能很好地提取出有效特征信息，不适用于特征丰富的遥感影像变化检测任务，本文提出的 MDFCNet 取得了最好的效果。

## 4　结论

（1）提出了一种适用于葡萄种植区的融合注意力机制和多尺度信息的变化检测网络（MDFCNet），在 ResNet 的基础上添加注意力机制增强网络的特征提取能力，在具有异质性的葡萄种植区双时相影像上表现优异，有效减少了由于像素值及饱和度差异带来的误差，同时在解码器中利用交叉差异特征捕获（CDFC）模块构建具有全局上下文信息的差异特征，极大程度避免了检测中的误检测和漏检测问题，并且通过监督集成注意力（SEA）模块，融合高低层特征，增强网络对于小面积种植区变化的检测能力，并且保留更多的边缘和细节，相较于未添加 SEA 的原始模型，交并比、召回率、F1 值和精确率分别提高 3.2、2.31、2.03、1.71 个百分点。

（2）本文模型在宁夏葡萄种植区数据集上进行实验，实验结果交并比为 79.37%、召回率为 86.19%、F1 值为 88.50%、精确率为 90.94%，相较于性能第 2 的模型，交并比、召回率、F1 值、精确率分别提高 5.42、5.62、3.48、0.95 个百分点。证明了本文 MDFCNet 在葡萄种植区变化检测中均优于目前最先进的模型，为实现葡萄种植区变化高精度检测任务提供了较好的解决思路。同时可将本文方法用于其他场景下的变化检测任务，具有优异的普适性，本文方法在其他场景下的变化检测数据集进行了评估，均取得了不错的检测精度。

（3）本研究进行了葡萄区的二类变化检测，但受限于单一类别，未引入更多类别标签，应用领域相对有限，同时模型主干网络层数较多，导致模型参数量较大，运算速率有待提高。

## 参　考　文　献

[1]　田野，陈冠铭，李家芬，等. 世界葡萄产业发展现状[J]. 热带农业科学，2018，38(6)：96－101.
TIAN Ye，CHEN Guanming，LI Jiafen，et al. Present development of grape industry in the world[J]. Chinese Journal of Tropical Agriculture，2018，38(6)：96－101.（in Chinese）

[2]　张宏鸣，张国良，朱珊娜，等. 基于 U－Net 的葡萄种植区遥感识别方法[J]. 农业机械学报，2022，53(4)：173－182.

ZHANG Hongming, ZHANG Guoliang, ZHU Shanna, et al. Remote sensing recognition method of grape planting regions based on U-Net[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(4): 173-182. (in Chinese)

[3] ZHU Z. Change detection using Landsat time series: a review of frequencies, preprocessing, algorithms, and applications[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2017, 130(9): 370-384.

[4] 孙中平,白金婷,史园莉,等. 基于高分影像的面向对象土地利用变化检测方法研究[J]. 农业机械学报, 2015, 46(增刊): 297-303.
SUN Zhongping, BAI Jinting, SHI Yuanli, et al. Object-oriented detection of land use changes based on high spatial resolution remote sensing image[J]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(Supp.): 297-303. (in Chinese)

[5] WEN D, HUANG X, BOVOLO F, et al. Change detection from very-high-spatial-resolution optical remote sensing images: methods, applications, and future directions[J]. IEEE Geoscience and Remote Sensing Magazine, 2021, 9(4): 68-101.

[6] 周旭,陈元鹏,刘岩涛,等. 基于云平台和 BFAST 算法的地表变化检测方法[J]. 农业机械学报, 2022, 53(7): 179-186.
ZHOU Xu, CHEN Yuanpeng, LIU Yantao, et al. Land surface change detection method based on cloud platform and BFAST algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(7): 179-186. (in Chinese)

[7] 张晗,倪维平,严卫东,等. 利用分形和多尺度分析的中低分辨率 SAR 图像变化检测[J]. 武汉大学学报(信息科学版), 2016, 41(5): 642-648.
ZHANG Han, NI Weiping, YAN Weidong, et al. Mid and low resolution SAR image change detection based on fractal and multi-scale analysis[J]. Geomatics and Information Science of Wuhan University, 2016, 41(5): 642-648. (in Chinese)

[8] FERRARIS V, DOBIGEON N, WEI Q, et al. Detecting changes between optical images of different spatial and spectral resolutions: a fusion-based approach[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(3): 1566-1578.

[9] DENG J S, WANG K, DENG Y H, et al. PCA-based land-use change detection and analysis using multitemporal and multisensor satellite data[J]. International Journal of Remote Sensing, 2008, 29(16): 4823-4838.

[10] 冯文卿,眭海刚,涂继辉,等. 高分辨率遥感影像的随机森林变化检测方法[J]. 测绘学报, 2017, 46(11): 1880-1890.
FENG Wenqing, SUI Haigang, TU Jihui, et al. Change detection method for high resolution remote sensing images using random forest[J]. Acta Geodaetica et Cartographica Sinica, 2017, 46(11): 1880-1890. (in Chinese)

[11] HABIB T, INGLADA J, MERCIER G, et al. Support vector reduction in SVM algorithm for abrupt change detection in remote sensing[J]. IEEE Geoscience and Remote Sensing Letters, 2009, 6(3): 606-610.

[12] 眭海刚,冯文卿,李文卓,等. 多时相遥感影像变化检测方法综述[J]. 武汉大学学报(信息科学版), 2018, 43(12): 1885-1898.
SUI Haigang, FENG Wenqing, LI Wenzhuo, et al. Review of change detection methods for multi-temporal remote sensing imagery[J]. Geomatics and Information Science of Wuhan University, 2018, 43(12): 1885-1898. (in Chinese)

[13] WANG Q, YUAN Z, DU Q, et al. GETNET: a general end-to-end 2-D CNN framework for hyperspectral image change detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(1): 3-13.

[14] SHI Q, LIU M, LI S, et al. A deeply supervised attention metric-based network and an open aerial image dataset for remote sensing change detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60:1-16.

[15] PENG D, ZHANG Y, GUAN H. End-to-end change detection for high resolution satellite images using improved UNet plus [J]. Remote Sensing, 2019, 11(11):1382.

[16] DU Y, ZHONG R, LI Q, et al. TransUNet plus plus SAR: change detection with deep learning about architectural ensemble in SAR images[J]. Remote Sensing, 2023, 15(1):6.

[17] WANG Z, PENG C, ZHANG Y, et al. Fully convolutional siamese networks based change detection for optical aerial images with focal contrastive loss[J]. Neurocomputing, 2021, 457: 155-167.

[18] ZHANG C, YUE P, TAPETE D, 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.

[19] SONG D, DONG Y, LI X. Context and difference enhancement network for change detection[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2022, 15: 9457-9467.

[20] 樊玮,周末,黄睿. 多尺度深度特征融合的变化检测[J]. 中国图象图形学报, 2020, 25(4): 669-678.
FAN Wei, ZHOU Mo, HUANG Rui. Multiscale deep features fusion for change detection[J]. Journal of Image and Graphics, 2020, 25(4): 669-678. (in Chinese)

[21] 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, 2022, 19: 1-5.

[22] LI Z, TANG C, LIU X, et al. Lightweight remote sensing change detection with progressive feature aggregation and supervised attention[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 1-12.

[23] 孙志同,朱珊娜,高郑杰,等. 基于波段增强的 DeepLabv3+多光谱影像葡萄种植区识别[J]. 农业工程学报, 2022, 38(7): 229-236.
SUN Zhitong, ZHU Shanna, GAO Zhengjie, et al. Recognition of grape growing areas in multispectral images based on band enhanced DeepLabv3+[J]. Transactions of the CSAE, 2022, 38(7): 229-236. (in Chinese)

(下转第 234 页)

[26] LIU L Y, JIANG H M, HE P C, et al. On the variance of the adaptive learning rate and beyond[J]. arXiv preprint arXiv: 1908.03265, 2019, 8.
[27] ILYA L, FRANK H. Decoupled weight decay regularization[J]. arXiv preprint arXiv: 1711.05101, 2017, 11.
[28] 宋怀波,华志新,马宝玲,等. 基于 SimCC - ShuffleNetV2 的轻量化奶牛关键点检测方法[J]. 农业机械学报, 2023, 54(10): 275 - 281,363.
SONG Huaibo, HUA Zhixin, MA Baoling, et al. Lightweight keypoint detection method of dairy cow based on SimCC - ShuffleNetV2[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(10): 275 - 281, 363. (in Chinese)
[29] 陈丽,王世勇,高思莉,等. Sentinel - 2 卫星的多光谱轻量级船舶目标检测算法[J]. 光谱学与光谱分析,2022,42(9): 2862 - 2869.
CHEN Li, WANG Shiyong, GAO Sili, et al. Multispectral lightweight ship target detection algorithm for Sentinel - 2 satellite [J]. Spectroscopy and Spectral Analysis, 2022, 42(9): 2862 - 2869. (in Chinese)
[30] 赵辉,黄镖,王红君,等. 基于改进 YOLO v7 的农田复杂环境下害虫识别算法研究[J]. 农业机械学报,2023,54(10): 246 - 254.
ZHAO Hui, HUANG Biao, WANG Hongjun, et al. Pest identification method in complex farmland environment based on improved YOLO v7[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(10): 246 - 254. (in Chinese)

(上接第 206 页)

[24] WU J, LI B, NI W, et al. An adaptively weighted multi-feature method for object-based change detection in high spatial resolution remote sensing images[J]. Remote Sensing Letters, 2020, 11(4): 333 - 342.
[25] NIU X, GONG M, ZHAN T, et al. A conditional adversarial network for change detection in heterogeneous images[J]. IEEE Geoscience and Remote Sensing Letters, 2019, 16(1): 45 - 49.
[26] PATRA S K, SHEKHER M, SOLANKI S S, et al. A technique for generating natural colour images from false colour composite images[J]. International Journal of Remote Sensing, 2006, 27(14): 2977 - 2989.
[27] 张祖勋,姜慧伟,庞世燕,等. 多时相遥感影像的变化检测研究现状与展望[J]. 测绘学报, 2022, 51(7): 1091 - 1107.
ZHANG Zuxun, JIANG Huiwei, PANG Shiyan, et al. Review and prospect in change detection of multi-temporal remote sensing images[J]. Acta Geodaetica et Cartographica Sinica, 2022, 51(7): 1091 - 1107. (in Chinese)
[28] ZHU Q Q, GUO X, LI Z Q, et al. A review of multi-class change detection for satellite remote sensing imagery[J]. Geo-spatial Information Science, 2022, 27(1): 1 - 15.
[29] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 770 - 778.
[30] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2011 - 2023.
[31] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 6230 - 6239.
[32] CHEN L, PAPANDREOU G, KOKKINOS I, et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834 - 848.
[33] HUANG Z, WANG X, WEI Y, et al. CCNet: criss-cross attention for semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(6): 6896 - 6908.