

# 基于风格迁移的双向孪生网络遥感变化检测<sup>①</sup>



史经业<sup>1</sup>, 罗雅露<sup>2,4</sup>, 张梦鸽<sup>3</sup>, 支瑞聪<sup>2,4</sup>, 刘吉强<sup>1</sup>

<sup>1</sup>(北京交通大学 智能交通数据安全与隐私保护技术北京市重点实验室, 北京 100044)

<sup>2</sup>(北京科技大学 计算机与通信工程学院, 北京 10083)

<sup>3</sup>(郑州大学 计算机与人工智能学院, 郑州 450001)

<sup>4</sup>(材料领域知识工程北京市重点实验室, 北京 100083)

通信作者: 支瑞聪, E-mail: [zhirc@ustb.edu.cn](mailto:zhirc@ustb.edu.cn); 刘吉强, E-mail: [jqliu2015@126.com](mailto:jqliu2015@126.com)

**摘要:** 不同时间获取的同一区域的双时相遥感图像在风格上往往具有很大差异, 大多数研究方法忽略了这个问题, 导致在风格多样的数据集上应用时, 模型的性能指标和可视化效果不尽如人意. 为此, 本文首先使用风格迁移模块针对某一时刻原始图像生成类似另一时刻风格的风格迁移图像. 其次, 提出了一种基于双向风格迁移的孪生对称差分特征金字塔网络 (symmetrical difference feature pyramid network, SDFPNet), 确定不同风格迁移方向对变化检测精度提升的影响程度. 具体来说, 将原始图像和风格迁移图像作为 SDFPNet 输入, 使用两个孪生的轻量化网络和差分特征金字塔网络 (difference feature pyramid network, DFPNet) 同时进行参数优化, 得到两个并行分支预测的变化图. 为了减少变化像素点的误判, 融合两个预测结果提升变化检测的准确性. 在 LEVIR-CD、CDD 和 SYSU-CD 这 3 个数据集上通过实验证明, 本文提出的基于双向风格迁移的 SDFPNet 在遥感变化检测任务上的评价指标优于 SOTA (state-of-the-art) 方法. 尤其是在由于季节变化, 风格差异较大的 CDD 和 SYSU-CD 数据集, 在 CDD 数据集上检测精度达到 99.37%,  $F2$  分数达到 94.19%, SYSU-CD 数据集上检测精度达到 92.31%. 有效解决了双时相图像风格差异大导致的变化检测精度不佳问题.

**关键词:** 变化检测; 遥感影像; 风格迁移; 差分特征金字塔网络

引用格式: 史经业, 罗雅露, 张梦鸽, 支瑞聪, 刘吉强. 基于风格迁移的双向孪生网络遥感变化检测. 计算机系统应用. <http://www.c-s-a.org.cn/1003-3254/9816.html>

## Bi-directional Siamese Network for Remote Sensing Change Detection Based on Style Transfer

SHI Jing-Ye<sup>1</sup>, LUO Ya-Lu<sup>2,4</sup>, ZHANG Meng-Ge<sup>3</sup>, ZHI Rui-Cong<sup>2,4</sup>, LIU Ji-Qiang<sup>1</sup>

<sup>1</sup>(Beijing Key Laboratory of Security and Privacy in Intelligent Transportation, Beijing Jiaotong University, Beijing 100044, China)

<sup>2</sup>(School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China)

<sup>3</sup>(School of Computer and Artificial Intelligence, Zhengzhou University, Zhengzhou 450001, China)

<sup>4</sup>(Beijing Key Laboratory of Knowledge Engineering for Materials Science, Beijing 100083, China)

**Abstract:** A majority of research methods neglect significant variations exhibited in the style of bi-temporal remote sensing images acquired at different times for the same area, leading to unsatisfactory model performance indexes and visualization when the model is applied to stylistically diverse datasets. To address this issue, a style transfer module is used in this article to generate an image with a style similar to that of another moment for the original image at a certain moment. Subsequently, a symmetrical difference feature pyramid network (SDFPNet) based on bi-directional style transfer is proposed to determine the influence degree of different style transfer directions on the improvement of change detection accuracy. Specifically, two lightweight Siamese networks and difference feature pyramid network (DFPNet) are

① 基金项目: 重点实验室项目 (2022-JCJQ-LA-001-080); 国家重点研发计划 (2018YFC0823002); 中央高校基本科研专项基金 (FRF-TP-20-10B); 2022 年度河南省重大科技专项 (221100210600); “十四五”重点研发计划项目 (2021YFB3901105)

收稿时间: 2024-10-10; 修改时间: 2024-10-30; 采用时间: 2024-11-12; csa 在线出版时间: 2025-02-25

used to conduct parameter optimization on the inputted original and stylized as SDFPNet, producing the change maps predicted by two parallel branches. To reduce the misclassification of changed pixels, the two prediction results are merged to improve the accuracy of change detection. Experiments on three datasets, LEVIR-CD, CDD, and SYSU-CD, demonstrate that the proposed SDFPNet based on bi-directional style transfer outperforms state-of-the-art (SOTA) methods in remote sensing change detection task, with results of CDD and SYSU-CD datasets more convincing, which have large style differences due to seasonal changes. The detection accuracy reaches 99.37% and  $F2$  score reaches 94.19% on the CDD dataset, and the detection accuracy reaches 92.31% on the SYSU-CD dataset. The proposed method in this article effectively solves the problem of poor change detection accuracy caused by large style differences in bi-temporal images.

**Key words:** change detection; remote sensing image; style transfer; difference feature pyramid network (DFPNet)

遥感影像变化检测 (remote sensing image change detection, RSICD) 的目的是对比不同时间点拍摄的同区域的遥感影像, 以识别地表发生的变化<sup>[1]</sup>. RSICD 被广泛应用于城市发展监测<sup>[2]</sup>、土地调查、生态系统监测、灾害监测评估以及军事侦察等领域<sup>[3-7]</sup>.

对于遥感图像变化信息的提取, 传统方法是依靠人工目视解译, 随着深度学习的快速发展, 基于深度学习的变化检测方法受到广泛关注, 该方法可以显著减少解译过程需要的人工和时间成本. 但是, 当获取的遥感影像风格不同时, 即使使用深度学习方法执行 RSICD 任务, 也会影响检测结果, 这是因为变化检测实质上是基于图像的内容特征做自动化分类. 由于不同年份拍摄的遥感影像存在卫星成像风格不同或季节变化, 会导致不同时相的遥感影像表现出强烈的光谱、空间、语义差异, 且这些差异很难用普通的影像预处理方法消除. 已有的变化检测方法在风格差异大的数据上往往会产生精度低、效果差的变化图, 目前仍然缺乏有效方法来解决变化检测中双时相遥感影像风格差异很大的问题.

针对以上问题, 本文提出一种基于风格迁移的双向孪生差分特征金字塔网络 (symmetrical difference feature pyramid network, SDFPNet), 解决变化检测中双时相遥感影像风格差异大的问题. 风格迁移能够在改变影像风格的同时保留影像内容特征, 利用经过风格迁移后的遥感影像进行变化检测能够减弱风格差异对检测结果的影响. 受风格迁移方法的启发, 本文在双时相影像双向风格转换的基础上, 使用对称差分孪生变化检测网络实现最终的变化检测.

## 1 相关工作

### 1.1 遥感变化检测

目前基于深度学习的方法被广泛应用在 RSICD 任务中<sup>[8,9]</sup>, 并表现出优越的特征表示能力. 该方法一般先使用特征提取器从原始多时相遥感影像中提取高级语义特征, 然后使用分类器生成变化图<sup>[10]</sup>. 常用提取器和分类器网络包括卷积神经网络 (convolutional neural network, CNN)、循环神经网络 (recurrent neural network, RNN)、Transformer 和生成对抗网络 (generative adversarial network, GAN). 对于基于 CNN 的方法, Zhan 等<sup>[11]</sup>探索基于 CNN 的孪生变化检测网络, 并使用加权对比损失来优化网络. Peng 等<sup>[12]</sup>提出一种改进的 UNet++ 模型, 该模型具有多尺度特征提取的密集跳跃连接和深度监督的多输出融合策略. 一些研究工作通过引入注意力机制来提高特征表示和辨别能力. 例如, Jiang 等<sup>[13]</sup>提出一种基于金字塔特征注意力引导的孪生网络, 其用一个注意力保持模块捕获双时相图像之间的相关性, 并设计了上下文融合策略来融合低级和高级特征. Huang 等<sup>[14]</sup>提出了一种多注意力孪生网络, 其中注意力特征融合模块用于融合孪生网络中不同层和分支的特征. 对于基于 RNN 的方法, Sun 等<sup>[15]</sup>在 UNet 中引入了传统的长短期记忆网络 (long short-term memory network, LSTM) 层 (Conv-LSTM), 以融合变化检测的空间和时间特征. Chen 等<sup>[16]</sup>使用深孪生细胞神经网络从同质和异质遥感图像中提取空间-光谱特征, 并使用多层 LSTM 将提取的特征映射到新的特征空间, 并充分挖掘变化信息. 对于基于 Transformer 的方法, Chen 等<sup>[17]</sup>提出一种双时相图像 Transformer (BIT), 其中语义标记器将图像表示为几个视觉单词, Transformer

利用它们来细化 CNN 提取的双时相特征. Zhang 等<sup>[18]</sup>提出一种纯 Transformer 变化检测网络, 名为 SwinSUNet, 具有孪生 U 形结构. 与基于 CNN 的方法的卷积单元不同, SwinSUNet 的基本单元是 Swin Transformer 块<sup>[19]</sup>. 对于基于 GAN 的方法, Chen 等<sup>[20]</sup>利用 GAN 合成新的训练样本实现实例级变化增强, 该方法包含生成的具有建筑相关变化和相应掩码的双时相图像. Zhao 等<sup>[21]</sup>提出了一种用于变化检测的注意力门生成对抗性自适应网络 (AG-GAAN). 在 AG-GAAN 模型中, 图像对被输入到生成器中以生成预测变化图, 鉴别器区分变化图和地面真实值. 生成器采用了 AGs (一种空间约束机制) 来定位变化区域并抑制不相关的干扰.

## 1.2 风格迁移

影像的风格是指影像具有的触点、纹理、色彩等特征. 影像风格迁移本质上是把影像的信息进行提取, 并分类为内容特征和风格特征信息, 保留内容特征信息的同时, 将风格特征信息和别的影像风格特征信息进行合并生成具有新风格的图像<sup>[22,23]</sup>.

传统风格迁移方法通过分析图像颜色来进行纹理合成, 基于数学方法构建风格迁移模型, 提取图像风格特征与目标图像结合, 得到风格迁移后的图像. 如 Reinhard 等<sup>[24]</sup>在保留影像内容的同时将色彩特性从一个图像转移到另一个图像. 该算法的关键是确保图像的亮度和色度解耦, 在色彩传输过程中不互相干扰. Li 等<sup>[25,26]</sup>提出 MRF 方法应用于图像特征处理. 与线性映射的方法相比, 该方法能够在一定程度上将图像的内容特征和风格特征进行分离. 但基于统计思想的模型训练过程, 需要人为参与来提高模型的性能. Huang 等<sup>[27]</sup>构建了一种能够进行实时风格迁移的模型, 该模型解决了以往模型转移风格单一的问题, 且不用进行调参, 且每张图像能够迅速地进行风格迁移. 该模型在对卷积特征进行归一化处理时, 把图像的内容特征和需要迁移的风格特征的均值和方差进行自适应匹配, 该操作大大加快了风格迁移图像的生成. 该模型不但将风格迁移过程的时间大幅度缩短, 还能够进行多种图像风格的迁移.

近年来, GAN 在计算机视觉领域被广泛应用. 从开始使用 GAN 生成模糊影像到现在生成包含高清逼真的人脸、动物和其他物体的图像, 不断表现出强大的图像生成仿真能力. 并且, GAN 在不同视觉任务中可以表现出较好的性能, 比如风格迁移任务. GAN 应

用在风格迁移任务的妙处在于无需对风格进行量化, 风格内容的抽取和学习都由 GAN 网络自己进行, 例如 CycleGAN<sup>[28]</sup>. 与之前提出的 pix2pix<sup>[29]</sup>网络相比, CycleGAN 的创新点在于引入了循环一致损失, 不需要在严格像素匹配的数据集上也可以实现很好的迁移效果. CycleGAN 的假设非常严格, 要求输入图像域和目标域之间存在双射关系, 这一点是很难满足的. CUT<sup>[30]</sup>提出了一个替代性方案, 通过最大化输入输出图像块的互信息, 使用一个对比损失函数, 指导编码器学习将对应的图像块相互联系起来, 与其他的图像块分离, 这样编码器可以专注于两个域之间共性的部分如形状, 而忽略两个域之间的差异性部分如纹理.

## 2 变化检测方法

### 2.1 网络总体架构

首先针对原始双时相遥感影像 ( $T^1$  和  $T^2$ ) 进行双向风格迁移, 即  $T^1$  迁移为与  $T^2$  同风格的迁移影像 ( $T^{1 \rightarrow 2}$ ),  $T^2$  迁移为与  $T^1$  同风格的迁移影像 ( $T^{2 \rightarrow 1}$ ). 最终得到 3 组影像对分别用于变化检测: (1)  $T^1$  与  $T^2$ ; (2)  $T^1$  与  $T^{2 \rightarrow 1}$ ; (3)  $T^2$  与  $T^{1 \rightarrow 2}$ . 但是无法确定哪个迁移方向更有利于变化检测精度的提升, 因此, 在此基础上提出一种基于双向风格迁移的 SDFPNet 变化检测架构, 如图 1 所示.

该架构包括两部分: (a) 双向风格迁移模块和 (b) SDFPNet 变化检测. 以  $T^1$  和  $T^2$  作为输入, 经过 (a) 得到  $T^1$  与  $T^{2 \rightarrow 1}$ ,  $T^2$  与  $T^{1 \rightarrow 2}$  两组影像对, 作为 (b) 的输入, 使用 SDFPNet 网络进行变化检测. SDFPNet 首先使用两个权重共享的轻量化网络进行参数优化, 分别得到两个并行分支的预测结果  $Pred^{2 \rightarrow 1}$  和  $Pred^{1 \rightarrow 2}$ . 受风格迁移生成图像的质量的影响, 变化检测网络可能在某些像素点上出现误判的情况. 为提升最终预测变化图的准确性, 使用求并策略融合两个分支生成的变化图, 得到最终预测的变化图  $FinalPred$ , 公式如式 (1):

$$FinalPred = Pred^{2 \rightarrow 1} \cup Pred^{1 \rightarrow 2} \quad (1)$$

### 2.2 双向风格迁移模块

双向风格迁移模块利用基于颜色空间的风格迁移算法和基于 GAN 的风格迁移模型作为一种预处理遥感影像方法, 将某一时相的遥感影像在保持语义内容特征不变的情况下从原始影像风格转换为另一时相遥感影像所定义的风格, 以减少双时相遥感影像之间的风格差异.

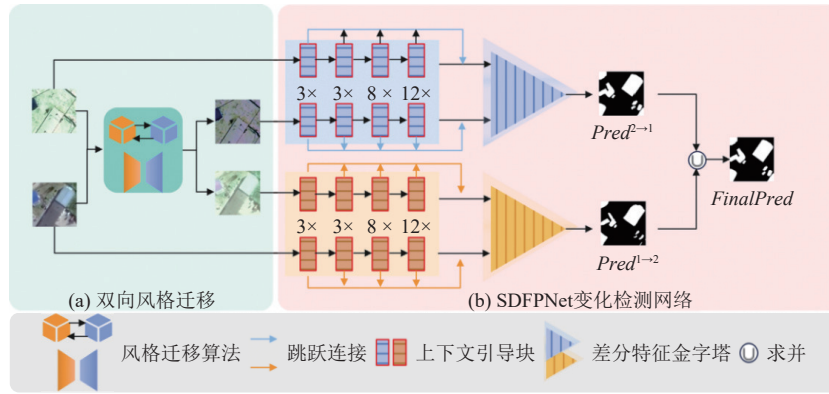


图1 基于双向风格迁移的SDFPNet架构

### 2.2.1 基于颜色空间的双向风格迁移

利用 $l\alpha\beta$ 颜色空间中各通道相关性较小, 干扰性小, 颜色能够被近似正交分离, 更有效的实现风格迁移的特点, 本小节实现了基于颜色空间的双向风格迁移算法<sup>[24]</sup>.

基于颜色空间的双向风格迁移算法的实现主要基于 $RGB$ 颜色空间、 $LMS$ 空间和 $l\alpha\beta$ 颜色空间的相互转换, 通过统计 $l\alpha\beta$ 颜色空间上的标准差、均值和梯度等信息, 建立合理的转换规则, 将原始风格图像和目标风格图像融合并为具有新色彩细节属性的新图像.

首先, 参照式(2)实现 $RGB$ 颜色空间到 $LMS$ 空间的转换, 同时通过 $L = \log L, M = \log M, S = \log S$ 计算 $LMS$ 空间的 $LMS$ 的对数空间代替原 $LMS$ 的值, 消除原始 $LMS$ 空间的数据偏差.

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.3811 & 0.5783 & 0.0402 \\ 0.1967 & 0.7244 & 0.0782 \\ 0.0241 & 0.1288 & 0.8444 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2)$$

通过式(3)将 $LMS$ 转换为一组正交基数, 利用自适应矩阵消除强关联, 完成了 $LMS$ 空间到 $l\alpha\beta$ 颜色空间的转换.

$$\begin{bmatrix} l \\ \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{6}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -2 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix} \quad (3)$$

完成 $RGB$ 颜色空间到 $l\alpha\beta$ 颜色空间的转换后, 需将 $l\alpha\beta$ 颜色空间重新转换回 $RGB$ 颜色空间. 首先利用式(4)将 $l\alpha\beta$ 颜色空间转换回 $LMS$ 空间.

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -2 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{3}}{3} & 0 & 0 \\ 0 & \frac{\sqrt{6}}{6} & 0 \\ 0 & 0 & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} l \\ \alpha \\ \beta \end{bmatrix} \quad (4)$$

转回到 $LMS$ 空间后, 再用 $L = 10^L, M = 10^M, S = 10^S$ 计算对数空间代替式(4)中原 $LMS$ 的值, 最后利用式(5)转回到 $RGB$ 颜色空间.

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 4.4679 & -3.5873 & 0.1193 \\ -1.2186 & 2.3809 & -0.1624 \\ 0.0497 & -0.2439 & 1.2045 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix} \quad (5)$$

结合式(3), 在 $l\alpha\beta$ 颜色空间上分别计算出目标风格图像和原始风格图像的3个通道的标准差表示 $(\sigma_r^l, \sigma_r^\alpha, \sigma_r^\beta)$ 、 $(\sigma_s^l, \sigma_s^\alpha, \sigma_s^\beta)$ , 均值表示 $(X_r^l, X_r^\alpha, X_r^\beta)$ 、 $(X_s^l, X_s^\alpha, X_s^\beta)$ , 利用式(6)计算出变换后的均值.

$$X^* = \frac{\sigma_r^X}{\sigma_s^X} (X_t - \bar{X}_t) + \bar{X}_t \quad (6)$$

其中,  $X$ 表示 $l\alpha\beta$ 颜色空间中的单个颜色通道,  $X_t$ 表示目标风格图像中的某一像素点在 $l\alpha\beta$ 颜色空间中的 $X$ 通道的值,  $X^*$ 表示进行颜色空间转换后的值,  $\bar{X}_t$ 是目标风格图像通道 $X$ 的均值,  $\sigma_r^X$ 是目标风格图像通道 $X$ 的标准差,  $\sigma_s^X$ 是原始风格图像通道 $X$ 的均值,  $\bar{X}_r$ 是原始风格图像通道 $X$ 的标准差. 进行颜色空间转换后得到的新图像具有与目标风格图像相同的标准差和均值, 达到风格迁移的目的.

在LEVIR-CD、CDD和SYSU-CD数据集上的风格迁移效果如图2所示. 可以看出, 迁移后的遥感影

像均具有与目标风格影像一致的整体颜色. 特别是如图 2(c) 中, 对于 CDD 数据集中的冬季场景, 风格迁移

后的影像  $T^{1 \rightarrow 2}$  与原始风格影像  $T^2$  一样都呈现出地表被白雪覆盖.

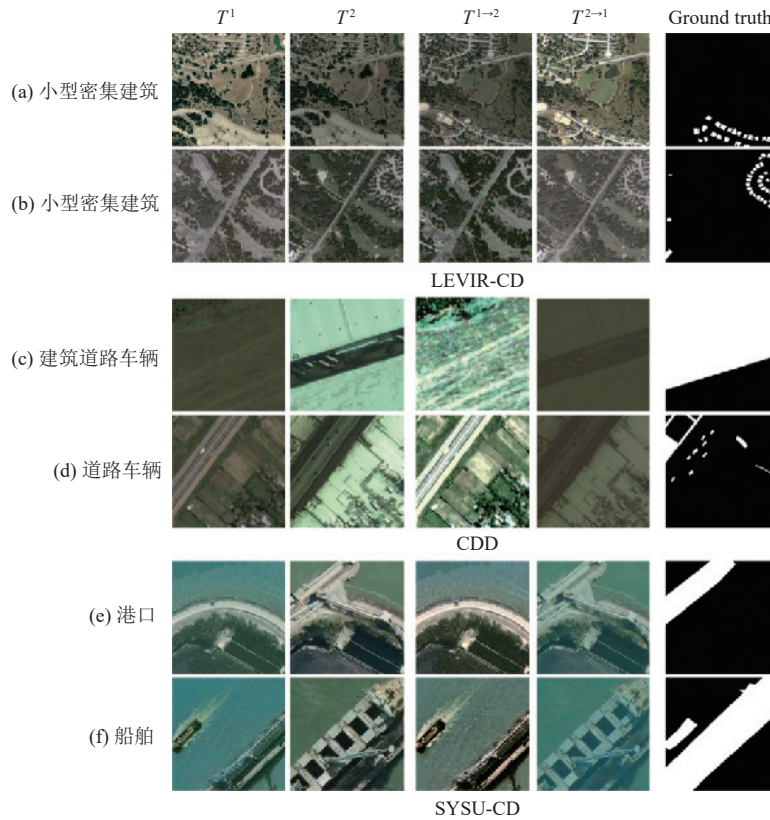


图 2 基于颜色空间的双向风格迁移算法迁移效果例子

### 2.2.2 基于 GAN 的双向风格迁移

基于 GAN 的风格迁移可以自动进行风格与内容特征的提取, 无需对风格进行量化. 通过生成器与判别器的互相博弈学习生成风格迁移图像. 本文使用基于 CUT 结构<sup>[29]</sup>的网络进行双时相遥感影像的双向风格转换 (如图 3 所示). CUT 风格迁移网络通过最大化输入输出图像块的互信息, 结合对比损失函数, 训练一个公共编码器  $G_{enc}$  来联系对应图像块, 约束编码器  $G_{enc}$  专注于两张风格图像的内容特征忽略纹理等差异性部分.

定义原始风格图像为  $X \in \mathbb{R}^{H \times W \times 3}$ , 目标风格图像为  $Y \in \mathbb{R}^{H \times W \times 3}$ ,  $H$  和  $W$  分别为图像的高和宽, 通道数均为 3. 生成的伪风格图像表示为  $\hat{y} = G(v) = G_{dec}(G_{enc}(x))$ .

在伪风格图像生成部分使用对抗损失来保证生成的伪风格图像能和目标风格图像的图像尽可能相似, 公式如式 (7):

$$\mathcal{L}_{GAN}(G, D, X, Y) = \mathbb{E}_{y \sim Y} \log(D(y)) + \mathbb{E}_{x \sim X} \log(1 - D(G(x))) \quad (7)$$

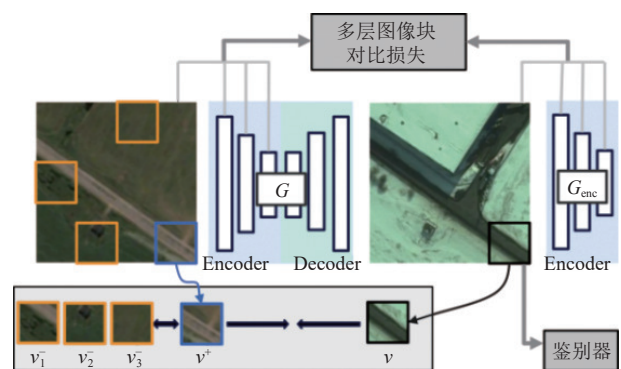


图 3 CUT 风格迁移模型架构

将目标风格图像中任意随机位置  $a$  处的图像块看作锚点, 然后将原始风格图像中的  $a$  位置的图像块看作是一个正样本, 将原始风格图像中剩余部分的图像块看作是负样本, 定义互信息交叉熵损失为式 (8):

$$\ell(v, v^+, v^-) = -\log \left[ \frac{\exp(v \cdot v^+ / \tau)}{\exp\left(v \cdot \frac{v^+}{\tau}\right) + \sum_{n=1}^N \exp(v \cdot v_n^- / \tau)} \right] \quad (8)$$

其中,  $v$  是目标风格图像中  $a$  处图像块的特征向量,  $v^+$  和  $v^-$  分别对应原始风格图像中  $a$  处和不在  $a$  处的图像块的特征向量,  $\tau$  是一个超参数. 通过最小化上式, 便可以最大化  $v$  和  $v^+$  的互信息, 而最小化  $v$  和  $v^-$  的互信息.

CUT 设置  $L$  层特征图, 将其通过  $G_{\text{enc}}$  产生一系列的特征. CUT 的目标是将原始风格图像与目标风格图像对应位置的图像块进行匹配, 同一张图像其他位置的图像块作为负样本, 最终通过对  $G_{\text{enc}}$  的每 1 层输出, 以及输出的每 1 个图像块的特征向量求互信息, 将损失记作多层图像块对比损失, 如式 (9):

$$\mathcal{L}_{\text{PatchNCE}}(G, H, X) = s \mathbb{E}_{x \sim X} \sum_{l=1}^L \sum_{s=1}^{S_l} \ell(\hat{v}_l^s, v_l^s, v_l^{s \setminus s}) \quad (9)$$

其中,  $l \in \{1, 2, \dots, L\}$ ;  $S_l$  表示第  $l$  层总的空间位置,  $s \in \{1, 2, \dots, S_l\}$ , 不同层的不同空间位置代表不同的图像块. 通过多层图像块对比损失使得编码器专注于两种风格图像之间的共性的部分如形状, 而忽略两种风格图像之间的差异性部分如纹理.

同时从数据集其他图像中提取的图像块做负样本记做  $v^-$ , 负样本字典集为  $V^-$  得到外部损失, 如式 (10):

$$\mathcal{L}_{\text{PatchEXT}}(G, H, X) = \mathbb{E}_{x \sim X, \bar{v} \sim V^-} \sum_{l=1}^L \sum_{s=1}^{S_l} \ell(\hat{v}_l^s, v_l^s, v_l^-) \quad (10)$$

最终的总损失为对抗损失, 对比损失, 和外部损失的和, 如式 (11):

$$\mathcal{L} = \mathcal{L}_{\text{GAN}}(G, D, X, Y) + \mathcal{L}_{\text{PatchNCE}}(G, H, X) + \mathcal{L}_{\text{PatchEXT}}(G, H, Y) \quad (11)$$

使用基于 GANs 的双向风格迁移算法来对 LEVIR-CD、CDD 和 SYSU-CD 数据集进行双向风格迁移, 效果如图 4 所示.

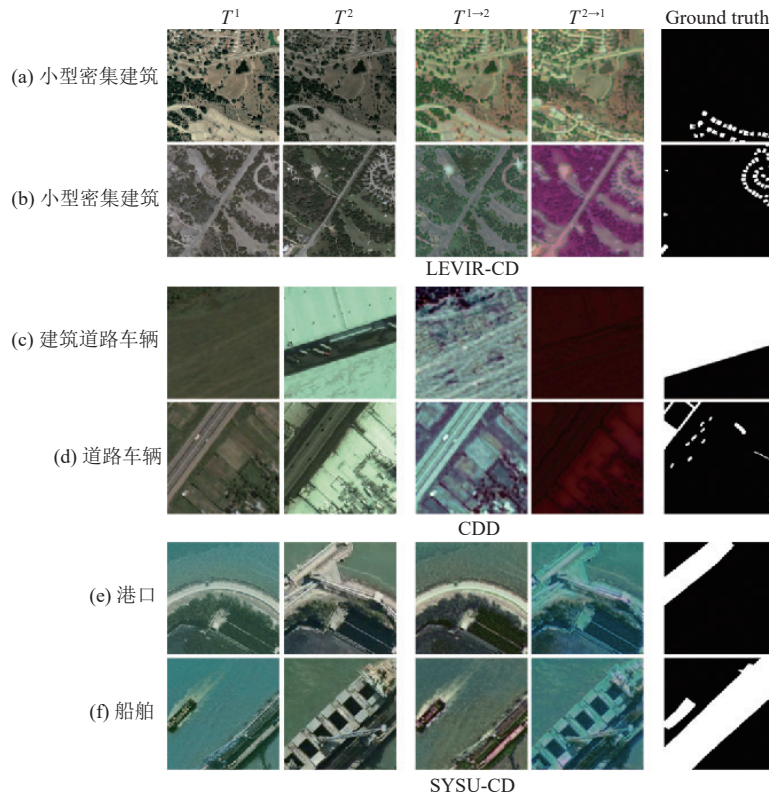


图 4 基于 GANs 的双向风格迁移算法迁移效果示例

### 2.3 对称差分特征金字塔网络

对称差分特征金字塔网络 (symmetrical difference

feature pyramid network, SDFPNet), 如图 5 所示, 包括轻量化的孪生主干模块和差分特征金字塔模块 (differ-

ence feature pyramid module, diffFPM). 主干网利用上下文引导块(context guide block, CGB)构建, 而 diffFPM 用于融合孪生特征对.

### 2.3.1 轻量级的孪生网络主干

两组影像对 ( $T^1$ 与 $T^{2 \rightarrow 1}$ ,  $T^2$ 与 $T^{1 \rightarrow 2}$ ) 经过具有共享权重的孪生网络主干, 该骨干网如图 5 所示, 由 4 个复合层组成 (从上到下, 复合层分别由 3/3/8/12 个 CGB

模块组成), 因此有 4 组特征输出.

CGB 结构如图 6 所示. 输入  $X$ , 经过并行的空洞卷积 (空洞率为 1 时为普通卷积), 以获得不同视野范围 (感受野) 的局部上下文信息. 空洞卷积以深度可分离的方式计算, 即所有的通道都被分组, 卷积只在一个独立的组中运算. 深度可分离卷积, 能大大减少计算量.

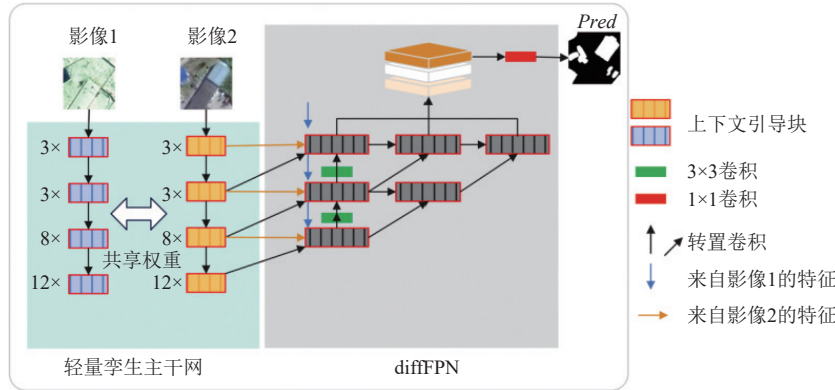


图 5 差分特征金字塔网络架构

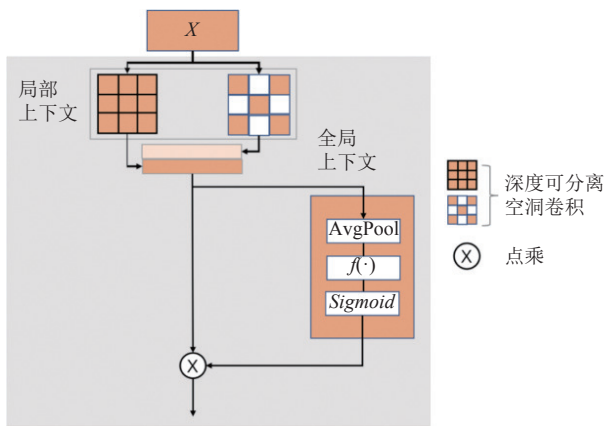


图 6 上下文引导块 (CGB) 结构示意图

注意力模块用于通道交互和全局信息提取, 由平均池化 (AvgPool), 非线性函数  $f(\cdot)$  和 Sigmoid 层组成, 可用式 (12) 表示.

$$Z_c = Sigmoid \left( f \left( \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W Y(i, j, c) \right) \right) \quad (12)$$

### 2.3.2 差分特征金字塔模块

保留每一复合层的输出 ( $T_{1,0}$ ,  $T_{1,1}$ ,  $T_{1,2}$ ,  $T_{1,3}$ ,  $T_{2,0}$ ,  $T_{2,1}$ ,  $T_{2,2}$ ,  $T_{2,3}$ ). 密集连通金字塔特征融合 (denseFPN) 方法对 3 个层次的输出  $d_{0,0}$ ,  $d_{1,0}$ ,  $d_{2,0}$  进行预测, 如图 7(a)

所示. 此结构存在两个问题: (1) 冗余连接.  $T_{1,0}$ ,  $T_{2,0}$  等浅层特征被反复输入到  $d_{0,0}$ ,  $d_{1,0}$  和  $d_{2,0}$  中, 效率很低. (2) 特征流不合理. 尽管 denseFPN 中存在冗余连接, 但输出层  $d_{0,0}$  和  $d_{1,0}$  包含来自骨干的不完整特征. 于是本文提出差分特征金字塔模块 diffFPM 如图 7(b) 所示, 计算如式 (13)–式 (15):

$$d_{0,0} = f(T_{1,0}, T_{2,0}, f(d_{0,1})) \quad (13)$$

$$d_{1,0} = f(d_{0,0}, d_{0,1}) \quad (14)$$

$$d_{2,0} = f(d_{1,0}, d_{1,1}) \quad (15)$$

在 diffFPM 中, 去掉了冗余连接, 增加了自底向上的融合路径, 使 3 个输出层都包含完整的骨干网特征, 既避免了卷积过程中信息丢失, 又实现了更高效的特征融合与差异定位.

## 2.4 损失函数

### 2.4.1 对称性损失

对称性是二进制对象变化的数学性质, 它表明二进制对象变化是无规律的, 即  $Y^{T_1 \rightarrow T_2} = Y^{T_2 \rightarrow T_1}$ . 直观地, RSICD 在双时相图像对上的输出应该遵循这一性质. 这意味着在时间对称性的约束下, SDFPNet 两个分支的输出不应因迁移方向而产生差异. 基于此, 提出了二分类对象变化检测的对称性损失, 计算如式 (16):

$$L_{\text{Symmetry}} = \frac{1}{2} \left[ \mathcal{L}_{\text{binary}}(\mathbf{F}_{\theta}(T^1, T^{2 \rightarrow 1}), \text{label}) + \mathcal{L}_{\text{binary}}(\mathbf{F}_{\theta}(T^2, T^{1 \rightarrow 2}), \text{label}) \right] \quad (16)$$

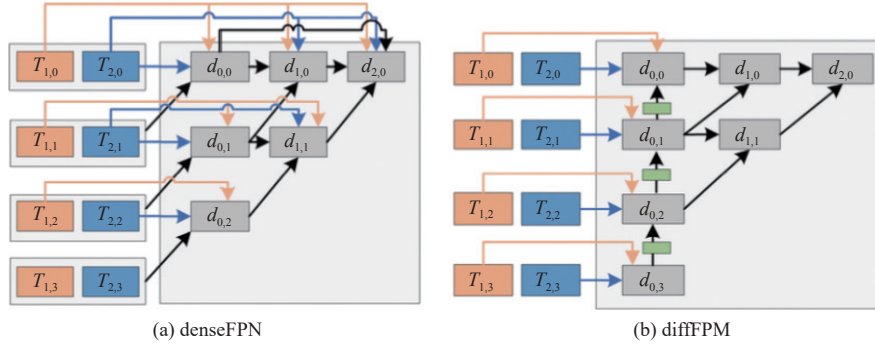


图7 孪生特征金字塔模块比较

#### 2.4.2 混合损失

变化检测是一个像素级的二元分类问题, 因此二元交叉熵损失可以直接用于变化检测任务. 二元交叉熵损失计算公式如式 (17):

$$L_{\text{BCE}} = - \sum_{i=0}^N p_i \log p_i^* + (1 - p_i) \log(1 - p_i^*) \quad (17)$$

其中,  $p_i$  为网络预测像素值 (0 或 1),  $p_i^*$  是对应的标签.

为了解决类别不平衡问题, 引入骰子损失, 定义为:

$$L_{\text{Dice}} = 1 - \frac{2 \times \text{Softmax}(\hat{Y})}{Y + \text{Softmax}(\hat{Y})} \quad (18)$$

所以总损失定义为:

$$L = L_{\text{BCE}} + L_{\text{Dice}} + L_{\text{Symmetry}} \quad (19)$$

### 3 实验结果与分析

#### 3.1 实验设计

##### 3.1.1 数据集和评价指标

本章使用 LEVIR-CD<sup>[10]</sup>、CDD<sup>[31]</sup>和 SYSU-CD<sup>[32]</sup>数据集进行方法验证.

为评估所提方法的性能, 使用 *precision*、*recall*、*F2* 分数和总体精度 (overall accuracy, *OA*) 评价指标. 在 CD 任务中, 精度值越高, 对预测结果发生的错误检测就越少, *recall* 值越大, 遗漏的预测结果就越少. 上述指标的定义如式 (20)–式 (23):

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

其中,  $\mathbf{F}_{\theta}$  表示 LSNet 变化检测过程. 使用对称性损失来约束模型从复杂的无关变化中区分真正的变化, 消除风格特征对模型变化检测的影响.

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

$$OA = \frac{TP + TN}{TP + TN + FN + FP} \quad (22)$$

其中, *TP*、*TN*、*FN*、*FP* 分别表示真阳性、真阴性、假阴性和假阳性的数量. *F2* 分数计算如式 (23):

$$f_{\beta} = (1 + \beta^2) \frac{\text{precision} \times \text{recall}}{\beta^2 \times \text{precision} + \text{recall}} \quad (23)$$

*precision* 和 *recall* 是权衡的, 它们相互限制. 在 CD 任务中, 希望保证 *precision* 值的同时尽可能地提高 *recall* 值, 因此  $\beta$  被定义为表示 *recall* 值的重要性. 以 *F2* 评分 ( $\beta = 2$ ) 作为主要评价指标, 旨在通过更关注 *recall* 值的提高来提高模型正确寻找变化样本的能力.

##### 3.1.2 实验设计

所有实验均使用 PyTorch 库, 使用 GeForce RTX 1080ti 加速模型训练. 对原始的 LEVIR-CD 数据集进行移动、缩放和旋转等数据增强, 避免由于数据缺乏而导致的潜在过拟合问题. 在实验中, 模型从初始态开始训练了 120 个 epoch, 初始学习率设置为 0.001, batch size 设置为 8. 经过 50 个 epoch 后, 学习率下降了 10%. 在每个训练阶段后进行验证, 并使用验证集上 *F2* 得分最高的最佳模型对测试集进行评价.

##### 3.1.3 对比方法

为了证明提出模型的有效性, 我们选取了一些 SOTA 方法进行对比.

UNet<sup>[33]</sup>: 主要是在医学分析领域提出的方法, 主要解决训练神经网络通常需要大量样本的问题, 它的



架构形状像字母“U”,由收缩路径 (contracting path, 也称为编码器部分) 和扩展路径 (expansive path, 也称为解码器部分) 组成.

UNet++<sup>[34]</sup>: 针对 UNet 网络的缺陷做了改进. 如重新设计了 skip connection, 使得解码器的子网络可以聚合不同尺度的特征; 利用剪枝技术提高 UNet++ 的推理速度.

DeepLabV3<sup>[35]</sup>: 采用级联或并行空洞卷积的模块、多尺度空洞率来捕获多尺度上下文. 此外, 对空间金字塔池模块进行改进, 该模块在多个尺度上探测卷积特征, 获得全局上下文的图像级特征编码, 并进一步提高性能.

DASNet<sup>[8]</sup>: 为了解决伪变化信息 (如噪声等) 的识别, 提出利用双重注意力来捕捉长距离依赖关系, 提高模型的鲁棒性. 并通过往加权损失函数来解决样本类

别不平衡.

SNUNet-CD<sup>[36]</sup>: 设计了一个密集连接的连体网络用于变化检测. 通过编码器和解码器之间、解码器和解码器之间的密集跳过连接, 保持高分辨率、细粒度的表示.

WNet<sup>[9]</sup>: 将 CNN 和 Transformer 结合起来, 提出一种“W”形双孪生分支分层网络, 解决 Transformer 忽视局部信息的问题.

### 3.2 基于风格迁移变化检测的有效性分析

本节为了验证风格迁移对变化检测结果的影响以及探究不同方向风格迁移对变化结果的影响, 在 3 个数据集上, 使用基于颜色空间的双向风格迁移算法 ( $la\beta$ ) 或基于生成对抗网络的双向风格迁移算法 (CUT) 进行双向风格迁移, DFPNet 进行变化检测, 定量结果如表 1 所示, 其中最优结果用粗体表示.

表 1 DFPNet 模块对 LEVIR-CD、CDD 和 SYSU-CD 数据集进行变化检测的量化结果

数据集	数据	precision	recall	F2	OA
LEVIR-CD	$T^1/T^2$ (数据A)	0.8914	0.8343	0.8470	<b>0.9914</b>
	$T^1/T^{2 \rightarrow 1}(la\beta)$ (数据B)	<b>0.8991</b>	<b>0.8281</b>	<b>0.8496</b>	<b>0.9914</b>
	$T^2/T^{1 \rightarrow 2}(la\beta)$ (数据C)	0.8744	0.8236	0.8333	<b>0.9914</b>
	$T^1/T^{2 \rightarrow 1}$ (CUT) (数据D)	0.7374	0.7422	0.7412	0.9094
	$T^2/T^{1 \rightarrow 2}$ (CUT) (数据E)	0.7407	0.7428	0.7424	0.9098
CDD	$T^1/T^2$ (数据A)	0.9322	0.9403	0.9387	0.9821
	$T^1/T^{2 \rightarrow 1}(la\beta)$ (数据B)	0.9309	<b>0.9452</b>	<b>0.9423</b>	<b>0.9825</b>
	$T^2/T^{1 \rightarrow 2}(la\beta)$ (数据C)	<b>0.9328</b>	0.9416	0.9398	0.9812
	$T^1/T^{2 \rightarrow 1}$ (CUT) (数据D)	0.8724	0.8891	0.8857	0.9644
	$T^2/T^{1 \rightarrow 2}$ (CUT) (数据E)	0.8803	0.8649	0.8679	0.9613
SYSU-CD	$T^1/T^2$ (数据A)	0.8022	0.7819	0.7859	0.9031
	$T^1/T^{2 \rightarrow 1}(la\beta)$ (数据B)	<b>0.9057</b>	0.7432	0.7709	0.8972
	$T^2/T^{1 \rightarrow 2}(la\beta)$ (数据C)	0.8025	<b>0.7833</b>	<b>0.7871</b>	<b>0.9034</b>
	$T^1/T^{2 \rightarrow 1}$ (CUT) (数据D)	0.7912	0.7173	0.7310	0.8559
	$T^2/T^{1 \rightarrow 2}$ (CUT) (数据E)	0.7824	0.7441	0.7514	0.8549

注: 最优结果用粗体表示.

基于  $la\beta$  算法的双向风格迁移方法在 3 个数据集上均表现最优指标. 在专注于建筑变化检测的 LEVIR-CD 数据集上, 指标提升较小, 特别是风格迁移前后 OA 值不变, 这一现象说明, 对于主要聚焦于建筑变化检测的数据集而言, 风格迁移技术并不能带来显著的效果提升. 由于季节变化, CDD 和 SYSU-CD 数据集双时相图像风格差异较大且含有多类别变化, 在 CDD 上,  $T^1/T^{2 \rightarrow 1}(la\beta)$  相比于原始影像对  $T^1/T^2$ , recall 提高了

0.49%, F2 分数提高了 0.36%, OA 提高了 0.04%. SYSU-CD 上,  $T^2/T^{1 \rightarrow 2}(la\beta)$  比  $T^1/T^2$  的 recall 提高了 0.14%, F2 分数提高了 0.12%, OA 提高了 0.03%. 将表 1 可视化为图 8 以方便观察在 OA 和 F2 分数上的提升效果. 进一步的研究证实了, 在风格差异显著的数据集上, 通过减少这些风格差异, 可以作为一种有效的预处理手段. 值得注意的是, 对 LEVIR-CD、CDD 数据集而言,  $T^1/T^{2 \rightarrow 1}(la\beta)$  方向的结果更优, 而对于 SYSU-CD

数据集而言,  $T^2/T^{1 \rightarrow 2}(\alpha\beta)$  方向的结果更优.

另外, 数据 D 和数据 E 的各项评价指标均远低于原始影像对. 这说明使用 GAN 网络生成的图片质量不如原始影像, 影响了图像本身包含的内容特征, 最终降低变化检测模型的性能, 验证了基于 CUT 算法的顾虑.

图 9 是上述实验的可视化结果. 图 9(a) 表明, 对于

小型密集建筑物 (LEVIR-CD) 变化检测, 模型主要关注的是建筑物轮廓, 所以数据 A/B/C 变化检测结果差异不大. CDD 数据集的双时相图像分别获取于冬夏季, 导致纹理、风格等差异较大, 图 9(b) 可以看出, 原始影像中道路、建筑物被积雪覆盖呈白色, 经过  $\alpha\beta$  算法能够很好地将  $T^1$  的风格迁移到  $T^2$ .

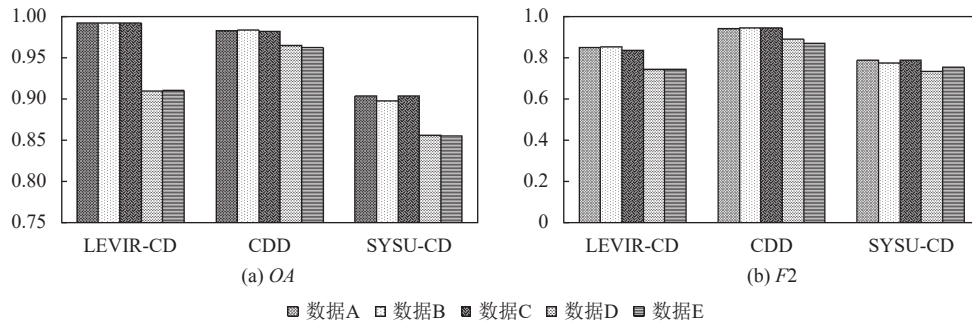


图 8 不同风格数据的变化检测结果评价指标 OA 与 F2 分数可视化比较

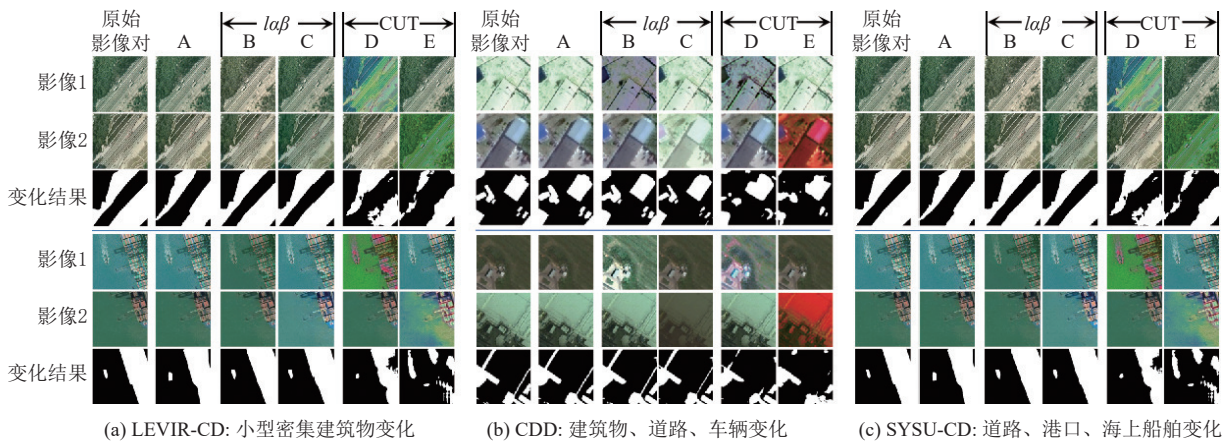


图 9 不同风格数据的变化检测结果图

针对 SYSU-CD 数据集, 选择了海上船只变化作为例子. 图 9(c) 表明对于检测轮廓清晰的船只变化, 使用风格迁移后的影像对比原始影像对进行变化检测得到的变化图错检更少, 说明风格迁移提高了模型其对颜色和纹理变化的鲁棒性.

### 3.3 SDFPNet 变化检测网络效果分析

基于第 3.2 节的实验结果, 发现需要对特定数据集进行不同方向的变化检测才能确定此数据集上哪个方向的风格迁移有效. 针对此问题, 提出了 SDFPNet, 以双向风格迁移得到的风格影像作为对称输入来进行变化检测. 图 10 表明, 相比于 DFPNet, 在 LEVIR-CD 数据集上, SDFPNet 的 recall 值得到了提升, OA 值提升不明显. 在 CDD 和 SYSU-CD 数据集上提升较明显,

特别是 CDD 数据集上 OA 值提升较大, 数据 B 的 recall 明显高于其他, SYSU-CD 数据集上数据 B 的 precision 明显高于其他. 总体而言, 风格迁移算法能作为 RSICD 的有效预处理方法.

同时, 为分析 SDFPNet 变化检测网络的性能, 在 3 个数据集上和 SOTA 方法进行比较, 结果如表 2, 其中最优的结果用粗体表示, 次优的结果用斜体和下划线表示. 与次优方法相比, SDFPNet 在 CDD 和 SYSU-CD 数据集上提升较显著, 特别是在 SYSU-CD 数据集上, recall 和 F2 分数分别提升 5.69% 和 5.22%. SDFPNet 在 LEVIR-CD、CDD 和 SYSU-CD 数据集上均获得最高的 OA 值, 和次优模型相比分别提升 0.04%、0.42%、1.35%. OA 值表示模型正确分类的像素占总像素的比

例, 所以说明 SDFPNet 网络能更有效地结合上下文信息提取判别性特征, 更好的区分变化和未变化像素, 完成变化检测任务.

将 SDFPNet 和对比模型的检测结果可视化, 如图 11 所示. 其中白色表示  $TP$  (真阳性), 黑色表示  $TN$  (真阴性), 红色表示  $FP$  (错检测), 绿色表示  $FN$  (漏检测). 从图 11(a)–(l) 可以看出, SDFPNet 的检测结果更准确, 漏检和误检更少. 对于检测建筑物变化, 如图 11(d) 和 (f), SDFPNet 检测结果比其他模型建筑物内部更完整, 漏

检更少 (绿色), 同时建筑物边界轮廓更清晰, 边界伪变化更少 (红色). CDD 数据集上, SDFPNet 对冬季白雪覆盖与夏季无覆盖的道路、车辆和建筑物的变化检测显著优于其他模型, 如图 11(g) 和 (h) 所示, 其他模型对道路的变化均存在较多漏检, 而 SDFPNet 检测结果的边界更清晰完整. 同样的, 图 11(j) 中道路变化结果内部也更完整. 同时, 如图 11(i)、(k)、(l) 所示, 对于更复杂多样的港口和海上船舶的变化, SDFPNet 检测结果中的伪变化也比其他模型少.

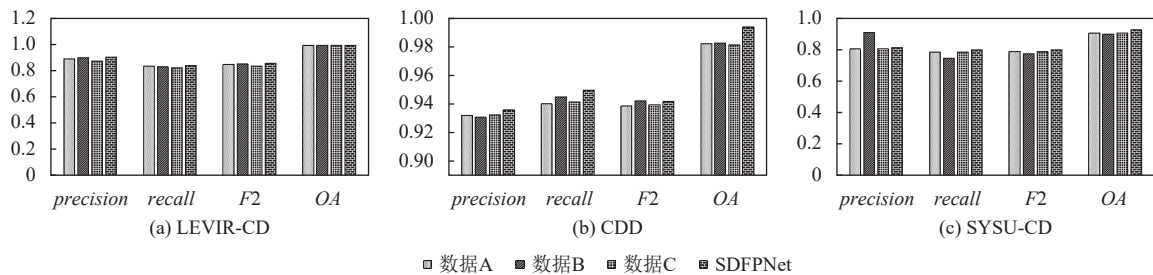


图 10 DFPNet 在不同风格上的变化检测结果与 SDFPNet 的定量比较

表 2 不同方法在 LEVIR-CD、CDD 和 SYSU-CD 数据集的量化评价结果

数据集	评价指标	UNet	UNet++	DeepLabV3	DASNet	SNUNet-CD	SwinSUNet	WNet	SDFPNet
LEVIR-CD	<i>precision</i>	0.8788	0.8682	0.8542	0.8674	0.8918	<u>0.8976</u>	0.8914	<b>0.9037</b>
	<i>recall</i>	0.8031	0.8253	0.7914	<u>0.826</u>	0.8189	0.8239	0.825	<b>0.8393</b>
	<i>F2</i>	0.8172	0.8335	0.8032	0.834	0.8325	<u>0.8377</u>	0.8374	<b>0.8562</b>
	<i>OA</i>	0.9891	0.9903	0.9872	0.9901	0.991	<u>0.9914</u>	<u>0.9914</u>	<b>0.9918</b>
CDD	<i>precision</i>	0.7806	0.7725	0.7594	0.7774	0.8374	<u>0.9344</u>	0.9245	<b>0.9362</b>
	<i>recall</i>	0.6089	0.6332	0.6147	0.672	0.8846	0.9038	<u>0.9123</u>	<b>0.9497</b>
	<i>F2</i>	0.6369	0.6569	0.6390	0.6907	0.8747	0.9097	<u>0.9147</u>	<b>0.9419</b>
	<i>OA</i>	0.9695	0.9752	0.9613	0.9827	0.9887	<u>0.9895</u>	<u>0.9895</u>	<b>0.9937</b>
SYSU-CD	<i>precision</i>	0.7562	0.7536	0.7354	0.7547	0.7749	<u>0.7804</u>	0.7678	<b>0.8116</b>
	<i>recall</i>	0.6951	0.7137	0.6924	0.7149	0.7233	0.7352	<u>0.7394</u>	<b>0.7963</b>
	<i>F2</i>	0.7065	0.7213	0.7006	0.7225	0.7331	0.7438	<u>0.7449</u>	<b>0.7971</b>
	<i>OA</i>	0.9028	0.9039	0.8985	0.904	0.9038	0.9067	<u>0.9096</u>	<b>0.9231</b>

注: 最优的结果用粗体表示, 次优的结果用斜体表示.

此外, 为了更直观地了解模型的学习动态, 我们对训练过程中的损失函数值进行了可视化展示. 如图 12 所示, 我们可以清晰地看到损失值随训练迭代次数增加而变化的趋势. 值得注意的是, 无论是在训练初期还是后期阶段, 我们的模型都能够保持相对平稳的性能表现, 这表明模型不仅快速收敛到了一个较优解, 而且在整个训练过程中展现出了高度的稳定性, 这对于确保最终模型的质量和可靠性至关重要.

#### 4 结论

风格差异较大的遥感影像变化检测是个富有挑战

性的难题, 双时相影像的光谱、纹理等特征会因实际地物变化、天气、季节等因素的影响呈现出复杂多样的差异性, 本文提出了基于双向风格迁移的 SDFPNet 用于变化检测, 将  $la\beta$  算法和 CUT 算法引入变化检测的步骤中, 来生成双向风格的新遥感数据集, 使用 DFPNet 对不同方向的影像对进行变化检测, 实验结果表明  $la\beta$  算法能作为减少遥感影像风格差异的有效预处理方法. 同时结合对称性损失来指导变化检测主干网络 SDFPNet 进行训练. 在 LEVIR-CD、CDD 和 SYSU-CD 数据集上的实验结果表明, SDFPNet 能减小双时相遥感影像的风格差异, 减少伪变化, 进一步提升变化检测的精度.

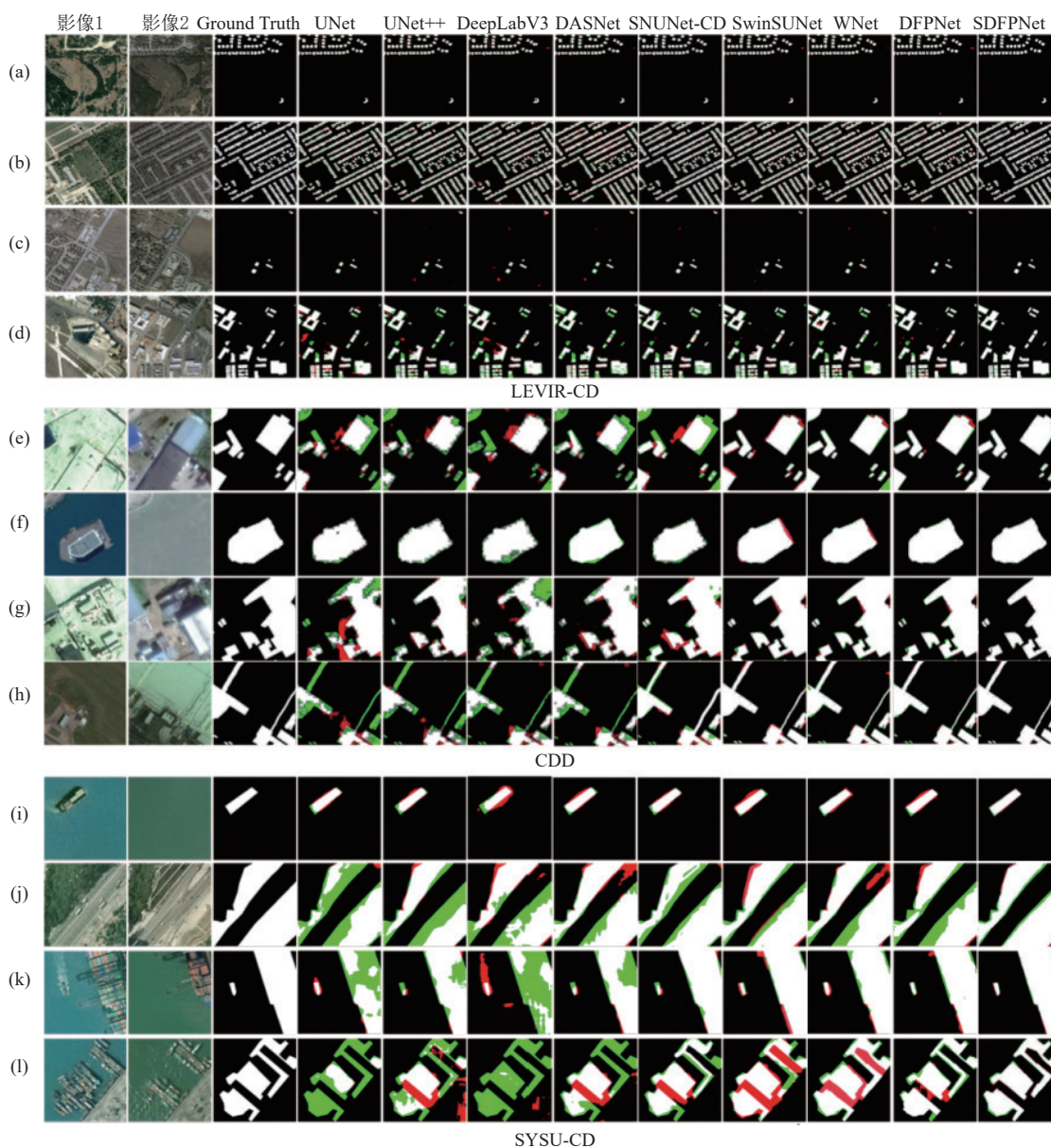


图 11 不同方法在 3 个数据集上的可视化结果

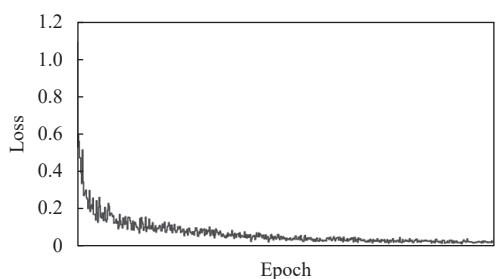


图 12 训练过程中损失收敛曲线

参考文献

1 Singh A. Review article digital change detection techniques

using remotely-sensed data. *International Journal of Remote Sensing*, 1989, 10(6): 989–1003. [doi: [10.1080/01431168908903939](https://doi.org/10.1080/01431168908903939)]

2 Liang BQ, Weng QH. Assessing urban environmental quality change of Indianapolis, United States, by the remote sensing and GIS integration. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2011, 4(1): 43–55. [doi: [10.1109/JSTARS.2010.2060316](https://doi.org/10.1109/JSTARS.2010.2060316)]

3 Li HX, Xiao PF, Feng XZ, *et al.* Using land long-term data records to map land cover changes in China over 1981–2010. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2017, 10(4): 1372–1389.

- [doi: [10.1109/JSTARS.2016.2645203](https://doi.org/10.1109/JSTARS.2016.2645203)]
- 4 Zelinski ME, Henderson J, Smith M. Use of Landsat 5 for change detection at 1998 Indian and Pakistani nuclear test sites. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2014, 7(8): 3453–3460. [doi: [10.1109/JSTARS.2013.2294322](https://doi.org/10.1109/JSTARS.2013.2294322)]
  - 5 Hu JR, Zhang YZ. Seasonal change of land-use/land-cover (LULC) detection using MODIS data in rapid urbanization regions: A case study of the Pearl River delta region (China). *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2013, 6(4): 1913–1920. [doi: [10.1109/JSTARS.2012.2228469](https://doi.org/10.1109/JSTARS.2012.2228469)]
  - 6 Vogelmann JE, Kost JR, Tolk B, *et al.* Monitoring landscape change for LANDFIRE using multi-temporal satellite imagery and ancillary data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2011, 4(2): 252–264. [doi: [10.1109/JSTARS.2010.2044478](https://doi.org/10.1109/JSTARS.2010.2044478)]
  - 7 Li ZB, Shi WZ, Myint SW, *et al.* Semi-automated landslide inventory mapping from bitemporal aerial photographs using change detection and level set method. *Remote Sensing of Environment*, 2016, 175: 215–230. [doi: [10.1016/j.rse.2016.01.003](https://doi.org/10.1016/j.rse.2016.01.003)]
  - 8 Chen J, Yuan ZY, Peng J, *et al.* DASNet: Dual attentive fully convolutional Siamese networks for change detection in high-resolution satellite images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 1194–1206. [doi: [10.1109/JSTARS.2020.3037893](https://doi.org/10.1109/JSTARS.2020.3037893)]
  - 9 Tang X, Zhang TX, Ma JJ, *et al.* WNet: W-shaped hierarchical network for remote-sensing image change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 5615814. [doi: [10.1109/TGRS.2023.3296383](https://doi.org/10.1109/TGRS.2023.3296383)]
  - 10 Khelifi L, Mignotte M. Deep learning for change detection in remote sensing images: Comprehensive review and meta-analysis. *IEEE Access*, 2020, 8: 126385–126400. [doi: [10.1109/ACCESS.2020.3008036](https://doi.org/10.1109/ACCESS.2020.3008036)]
  - 11 Zhan Y, Fu K, Yan ML, *et al.* Change detection based on deep Siamese convolutional network for optical aerial images. *IEEE Geoscience and Remote Sensing Letters*, 2017, 14(10): 1845–1849. [doi: [10.1109/LGRS.2017.2738149](https://doi.org/10.1109/LGRS.2017.2738149)]
  - 12 Peng DF, Zhang YJ, Guan HY. End-to-end change detection for high resolution satellite images using improved UNet++. *Remote Sensing*, 2019, 11(11): 1382. [doi: [10.3390/rs11111382](https://doi.org/10.3390/rs11111382)]
  - 13 Jiang HW, Hu XY, Li K, *et al.* PGA-SiamNet: Pyramid feature-based attention-guided Siamese network for remote sensing orthoimagery building change detection. *Remote Sensing*, 2020, 12(3): 484. [doi: [10.3390/rs12030484](https://doi.org/10.3390/rs12030484)]
  - 14 Huang JR, Shen Q, Wang M, *et al.* Multiple attention Siamese network for high-resolution image change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5406216. [doi: [10.1109/TGRS.2021.3127580](https://doi.org/10.1109/TGRS.2021.3127580)]
  - 15 Sun ST, Mu L, Wang LZ, *et al.* L-UNet: An LSTM network for remote sensing image change detection. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 8004505. [doi: [10.1109/LGRS.2020.3041530](https://doi.org/10.1109/LGRS.2020.3041530)]
  - 16 Chen HRX, Wu C, Du B, *et al.* Change detection in multisource VHR images via deep Siamese convolutional multiple-layers recurrent neural network. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 58(4): 2848–2864. [doi: [10.1109/TGRS.2019.2956756](https://doi.org/10.1109/TGRS.2019.2956756)]
  - 17 Chen H, Qi ZP, Shi ZW. Remote sensing image change detection with Transformers. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5607514. [doi: [10.1109/TGRS.2021.3095166](https://doi.org/10.1109/TGRS.2021.3095166)]
  - 18 Zhang C, Wang LJ, Cheng SL, *et al.* SwinSUNet: Pure Transformer network for remote sensing image change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5224713. [doi: [10.1109/TGRS.2022.3160007](https://doi.org/10.1109/TGRS.2022.3160007)]
  - 19 Liu Z, Lin YT, Cao Y, *et al.* Swin Transformer: Hierarchical vision Transformer using shifted windows. *Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision*. Montreal: IEEE, 2021. 9992–10002.
  - 20 Chen H, Li WY, Shi ZW. Adversarial instance augmentation for building change detection in remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5603216. [doi: [10.1109/TGRS.2021.3066802](https://doi.org/10.1109/TGRS.2021.3066802)]
  - 21 Zhao WZ, Chen X, Ge XS, *et al.* Using adversarial network for multiple change detection in bitemporal remote sensing imagery. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 8003605. [doi: [10.1109/LGRS.2020.3035780](https://doi.org/10.1109/LGRS.2020.3035780)]
  - 22 Portilla J, Simoncelli EP. A parametric texture model based on joint statistics of complex wavelet coefficients. *International Journal of Computer Vision*, 2000, 40(1): 49–70. [doi: [10.1023/A:1026553619983](https://doi.org/10.1023/A:1026553619983)]
  - 23 Kang JJ, Zhu ZH. A unified energy-based control framework for tethered spacecraft deployment. *Nonlinear Dynamics*, 2019, 95(3): 1117–1131. [doi: [10.1007/s11071-018-4619-x](https://doi.org/10.1007/s11071-018-4619-x)]
  - 24 Reinhard E, Adhikhmin M, Gooch B, *et al.* Color transfer between images. *IEEE Computer Graphics and Applications*, 2001, 21(5): 34–41. [doi: [10.1109/38.946629](https://doi.org/10.1109/38.946629)]

- 25 Li C, Wand M. Combining Markov random fields and convolutional neural networks for image synthesis. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016. 2479–2486. [doi: [10.1109/CVPR.2016.272](https://doi.org/10.1109/CVPR.2016.272)]
- 26 Li C, Wand M. Precomputed real-time texture synthesis with Markovian generative adversarial networks. Proceedings of the 14th European Conference on Computer Vision. Amsterdam: Springer, 2016. 702–716. [doi: [10.1007/978-3-319-46487-9\\_43](https://doi.org/10.1007/978-3-319-46487-9_43)]
- 27 Huang X, Belongie S. Arbitrary style transfer in real-time with adaptive instance normalization. Proceedings of the 2017 IEEE International Conference on Computer Vision. Venice: IEEE, 2017. 1510–1519. [doi: [10.1109/ICCV.2017.167](https://doi.org/10.1109/ICCV.2017.167)]
- 28 Zhu JY, Park T, Isola P, *et al.* Unpaired image-to-image translation using cycle-consistent adversarial networks. Proceedings of the 2017 IEEE International Conference on Computer Vision. Venice: IEEE, 2017. 2242–2251. [doi: [10.1109/ICCV.2017.244](https://doi.org/10.1109/ICCV.2017.244)]
- 29 Isola P, Zhu JY, Zhou TH, *et al.* Image-to-image translation with conditional adversarial networks. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 5967–5976. [doi: [10.1109/CVPR.2017.632](https://doi.org/10.1109/CVPR.2017.632)]
- 30 Park T, Efros AA, Zhang R, *et al.* Contrastive learning for unpaired image-to-image translation. Proceedings of the 16th European Conference on Computer Vision. Glasgow: Springer, 2020. 319–345.
- 31 Lebedev MA, Vizilter YV, Vygolov OV, *et al.* Change detection in remote sensing images using conditional adversarial networks. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2018, 42: 565–571. [doi: [10.5194/isprs-archives-XLII-2-565-2018](https://doi.org/10.5194/isprs-archives-XLII-2-565-2018)]
- 32 Shi Q, Liu MX, Li SC, *et al.* A deeply supervised attention metric-based network and an open aerial image dataset for remote sensing change detection. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5604816. [doi: [10.1109/TGRS.2021.3085870](https://doi.org/10.1109/TGRS.2021.3085870)]
- 33 Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. Proceedings of the 18th International Conference on Medical Image Computing and Computer-assisted Intervention. Munich, Germany: Springer, 2015. 234–241.
- 34 Zhou ZW, Siddiquee MR, Tajbakhsh N, *et al.* UNet++: A nested U-Net architecture for medical image segmentation. Proceedings of the 4th International Workshop on Deep Learning in Medical Image Analysis. Granada: Springer, 2018. 3–11.
- 35 Chen LC, Papandreou G, Schroff F, *et al.* Rethinking atrous convolution for semantic image segmentation. arXiv:1706.05587, 2017.
- 36 Fang S, Li KY, Shao JY, *et al.* SNUNet-CD: A densely connected Siamese network for change detection of VHR images. IEEE Geoscience and Remote Sensing Letters, 2022, 19: 8007805. [doi: [10.1109/LGRS.2021.3056416](https://doi.org/10.1109/LGRS.2021.3056416)]

(校对责编: 王欣欣)