

中图法分类号: TP751.1 文献标识码: A 文章编号: 1006-8961(2024)08-2188-17

论文引用格式: Jia M, Zhao Q and Lu X F. 2024. Bipartite adversarial autoencoder network for unsupervised detection of changes in heterogeneous remote sensing images. Journal of Image and Graphics, 29(08):2188-2204(贾萌, 赵秦, 鲁晓锋. 2024. 双边对抗自编码网络的无监督异质遥感图像变化检测. 中国图象图形学报, 29(08):2188-2204)[DOI:10.11834/jig.230497]

## 双边对抗自编码网络的无监督 异质遥感图像变化检测

贾萌, 赵秦, 鲁晓锋\*

西安理工大学计算机科学与工程学院, 西安 710048

**摘要:** 目的 异质遥感图像由不同类型的传感器所获取, 在数据结构、分辨率及辐射特性上均存在巨大差异。变化检测任务旨在通过分析在不同时间获取的同一目标区域的图像来检测地表覆盖物变化, 然而异质遥感图像的数据异构特性会使得变化检测过程更加困难。针对这个问题, 提出了一种嵌入聚类分析的双边对抗自编码网络来实现异质遥感图像地物变化的精确检测。方法 构造双边对抗自编码网络对异质遥感图像进行重构和风格转换, 通过结构一致性损失和对抗损失对网络训练进行约束, 迫使网络将异质图像转换到公共数据域。考虑到变化区域像素对于对抗损失函数在网络优化中的不利影响, 对映射到公共数据域的两对同质图像进行聚类分析, 基于此提出一种新的语义信息约束的对抗损失函数, 迫使网络生成具有更加一致风格的图像。结果 在4组典型的异质遥感图像数据集上对提出的变化检测网络性能进行测试, 在Italy数据集、California数据集、Tianhe数据集以及Shuguang数据集上的总体检测精度分别达到0.970 5、0.938 2、0.994 7以及0.982 6。与现有的传统以及深度学习方法对比, 提出算法在视觉及定量分析结果上均取得了较好的检测性能。结论 针对异质遥感图像变化检测所要面临的由环境、数据异构等因素造成的检测困难、错检率高的问题, 提出的基于双边对抗自编码网络的无监督异质遥感图像变化检测方法, 既可以实现变化检测过程完全无监督, 又充分利用网络特性和语义信息, 提高了变化检测性能。

**关键词:** 域转换; 聚类分析; 语义信息; 无监督变化检测; 异质遥感图像

### Bipartite adversarial autoencoder network for unsupervised detection of changes in heterogeneous remote sensing images

Jia Meng, Zhao Qin, Lu Xiaofeng\*

School of Computer Science and Engineering, Xi'an University of Technology, Xi'an 710048, China

**Abstract:** **Objective** Heterogeneous remote sensing images from different sensors are quite different in imaging mechanism, radiation characteristics, and geometric characteristics. Thus, they reflect the physical properties of the ground target at different levels. Therefore, no relationship exists between the observed values of the same object, which usually leads to "pseudo changes". As a result, the change detection task has more difficulty obtaining accurate change information of the observed ground objects. Efforts have been made toward unsupervised detection of changes in heterogeneous

收稿日期: 2023-07-20; 修回日期: 2023-10-01; 预印本日期: 2023-10-08

\* 通信作者: 鲁晓锋 luxiaofeng@xaut.edu.cn

基金项目: 国家自然科学基金项目(62076201); 陕西省重点研发计划资助(2023-YBGY-222); 碑林区科技计划项目(GX2246)

Supported by: National Natural Science Foundation of China(62076201); Key R&amp;D Program of Shaanxi Province, China(2023-YBGY-222); Beilin District Science and Technology Project(GX2246)

remote sensing images by designing various methods to obtain change information. However, traditional image difference operators based on the difference or ratio of radiation measurement is no longer applicable. Therefore, transferring the bitemporal images in a common space is a convenient way to calculate differences. Considering the excellent and flexible feature learning capability of deep neural networks, they have been widely applied in change detection tasks for heterogeneous images to effectively alleviate the influence of “pseudo changes”. Moreover, by fully utilizing the characteristics of deep neural networks, they can be designed to transform heterogeneous remote sensing images into the same feature domain. Then, the change information can be accurately represented. Inspired by the paradigm of image translation, heterogeneous images are transformed into a common domain with consistent feature representations to enable direct comparisons of data. The key point of this task is learning a suitable one-to-one mapping to build a relationship between distinct appearances of images and exclude the effect of interference factors. Therefore, this study proposes a bipartite adversarial autoencoder network with clustering (BAACL) to detect changes between heterogeneous remote sensing images. **Method** A bipartite adversarial autoencoder network is constructed to reconstruct the bitemporal images and achieve the transformation of heterogeneous images to common domain. Appropriate reconstruction loss regularization should be applied to obtain good data mapping from the constructed neural network optimization. Next, the bipartite convolutional autoencoders can be jointly trained to encode the input images and reconstruct them with high fidelity in output. Meanwhile, for the image-to-image translation task, additional structural consistency loss and adversarial loss terms are designed to constrain the network training process for converting heterogeneous images to the common data domain. The structural consistency loss term is designed to describe the internal structural relationships of images before and after translation. For this purpose, affinity matrixes are adopted to express the structural self-similarity within an image. The heterogeneous images can be conveniently compared in a new affinity space. Notably, based on the paradigm of image translation, our model can be viewed as learning two “transformation functions”, which are adopted to transform each of the heterogeneous images to the opposite data space. Furthermore, such a setup can be regarded as a special case of an “adversarial mechanism”, which is formulated by an adversarial loss to train autoencoders for matching the opposite image style. Considering the disadvantage of these changed pixels to the adversarial loss term in network optimization, the pseudo-difference image generated by two pairs of homogeneous images mapped in the common data domain is analyzed by clustering in an unsupervised manner. Thereafter, the obtained semantic information is adopted to further constrain the adversarial loss term. The overall performance of the BAACL network is illustrated on four sets of publicly available datasets of heterogeneous remote sensing images. Five popular traditional and deep learning-based detection methods for changes in heterogeneous remote sensing images are selected and compared with this method to verify its effectiveness. **Result** Results obtained from the Italy, California, Tianhe, and Shuguang datasets can also illustrate the performance of the proposed BAACL, with the overall detection accuracies of up to 0.970 5, 0.938 2, 0.994 7, and 0.982 6, respectively. Meanwhile, the proposed method is superior to the five compared methods in terms of visual results of the final change map. A set of ablation experiments is designed to verify the influence of the improved adversarial loss term on network optimization performance. the performance of the BAACL network does not change dramatically with various proportions of sample changing regions due to the semantic regularization. It verifies the effectiveness of the proposed semantic information-based adversarial loss term for network optimization. **Conclusion** The semantic information-based adversarial loss term is designed to narrow the distance of the unchanged regions of the bitemporal images for image style consistency. The reason is that, for adversarial loss term without semantic regularization, a larger proportion of the change region of the current training sample will result in a worse trained network effect. Therefore, an accurate definition of the network constraint term will result in good network optimization and further affect the change detection performance. In view of the difficulty and high false alarm rate of detection methods for changes in heterogeneous remote sensing images caused by factors such as seasons and data heterogeneity, the proposed bipartite adversarial autoencoder network for the detection of changes in heterogeneous remote sensing images not only can fully utilize network characteristics and semantic information to improve image style consistency but also can realize completely unsupervised change detection process. The change detection performance can be greatly improved as well.

**Key words:** domain transformation; cluster analysis; semantic information; unsupervised change detection; heterogeneous remote sensing images

## 0 引言

遥感图像变化检测就是对从同一地理区域在不同时间获取的双时相遥感图像进行分析,从而精确获取观测区域所发生的地物覆盖物变化(Mercier等,2008)。目前,该技术已在遥感信息处理(公茂果等,2016)、计算机视觉(Radke等,2005)、城市建筑监测(Li等,2019)、森林砍伐(Brunner等,2010)及土地覆盖变化(Griffiths等,2010;Pantze等,2010)等领域获得了广泛应用。传统的变换检测方法依赖于同质数据,即同一传感器在相同的几何形状、季节条件和记录设置下获取的一组图像。然而,对于很多实际应用中,这些限制太过严格。首先,卫星重访周期设置了监测长期趋势时的时间分辨率上限以及评估突发事件损害的响应时间下限。此外,即使以相同的配置收集两幅图像,由于其他客观影响因素的存在,例如:成像噪声、传感器分辨率、多视角成像和环境(如光照、天气、季节)等干扰因素使得双时相遥感图像对同一地物往往表现出明显不同的外观或特征,光学数据的光照条件或合成孔径雷达(synthetic aperture radar, SAR)的湿度和降水,使得图像也可能不是均质的。设计快速、自动和无监督的异质遥感图像变化检测方法具有十分重要的现实意义。然而,异质遥感图像由于其明显的特征差异而导致的“伪变化”对真实变化区域的检测带来了很大的挑战,因此,需要从跨模态异构数据中挖掘不变特征,其过程复杂,对变化检测技术提出了更高的要求(Khan等,2017;Lyu等,2022)。

研究者在无监督异质遥感图像变化检测领域进行了大量的探索与实践。传统的用于同质图像的差异图像生成算子,如:差值算子(Muchoney和Haack,1994)、比值算子(Rignot和van Zyl,1993)或改进型比值算子(Bovolo和Bruzzone,2005;Inglada和Mercier,2007),已无法适用于异质图像。因此,Liu等人(2018b)提出了同质图像像素转换(homogeneous pixel transformation, HPT),将异质图像转换到公共数据域以便于分析变化。为了抵抗各种干扰因素的影响,Sun等人(2021)提出了一种基于非局部块的图模型(nonlocal patch-based graph, NPSG)及其改进模型 INLPG (improved nonlocal patch-based graph, INPSG)(Sun等,2022),挖掘不同模态数据间的结构

不变特征实现地物变化的精确检测。

深度神经网络由于其优异灵活的特征学习能力,广泛应用于异质遥感图像变化检测任务中,以有效缓解“伪变化”的影响。一种行之有效的思路是利用深度神经网络将异质遥感图像转换到同一数据域,进而方便比较像素对差异,获取差异图像。Liu等人(2022)提出通过基于双边卷积神经网络对异质遥感图像像素对之间的差异进行联合概率建模。樊玮等人(2020)提出基于多尺度特征融合的卷积神经网络(convolutional neural network, CNN)模型,通过高层语义与低层纹理的多尺度融合来检测变化区域。Luppino等人(2024)提出通过在编码空间中构造像素对之间的相关性来建模异质数据差异信息。此外,Liu等人(2018a)提出了一种对称网络结构,将异质遥感图像转换到公共特征空间以便检测变化。王艳恒等人(2020)提出将深度学习与超像元分割算法相结合的变化检测方法。Niu等人(2019)提出了一种条件生成对抗网络(conditional generative adversarial network, cGAN)对光学图像进行转换以符合SAR图像数据的分布。Li等人(2021)提出了基于深度图像转换的异质遥感图像变化检测网络,在同一特征空间中,利用改进UNet++实现异质遥感图像地物变化的有效检测。利用深度神经网络将异质遥感图像映射到同一特征空间,难以充分对获取到的变化信息进行解释。同时,深度神经网络的学习需要训练样本,存在一定的局限性,不能很好地适用于复杂场景遥感图像的变化检测(Luppino等,2022;Lyu等,2022)。

本文提出了一种嵌入聚类分析的无监督双边对抗自编码网络(bipartite adversarial autoencoder network with clustering, BAACL),用于处理异质遥感图像变化检测问题。通过构造双边对抗自编码网络将异质遥感图像映射入公共数据域以方便双时相图像比较生成差异图,可完全无监督地实现异构遥感数据之间地物变化的精确检测。主要贡献为:1)在双边自编码网络框架上,基于3个损失项构造损失函数训练两对卷积自编码器,学习异构数据的非线性映射,促使网络生成两对具有同质数据结构及风格一致的图像对;2)联合对抗网络来判别真实图像与转换后图像的一致性,旨在推动网络生成图像风格更加逼真,然而,考虑到利用语义信息存在差异的样本(变化区域)构造对抗损失函数,会对网络优化性

能造成一定的不利影响,提出基于聚类分析构造具有语义约束的对抗损失函数,实现网络优化性能的进一步提高。提出算法BAACL在4组典型的异质遥感图像数据集上进行了实验并取得了优异的性能。

## 1 本文方法

异质遥感图像变化检测的目的是从不同时相、来自于不同传感器的遥感图像数据中定量分析和确定地表有意义的变化,并忽略意义不大的、低置信度的变化。然而,由噪声、辐射差异、光照变化以及天气等干扰因素所引起的“伪变化”是处理异质遥感图像变化检测问题的困难所在。当前主流的解决办法之一是异质遥感图像数据转换,即将异质图像转换到一个具有一致特征表示的公共域中,使其可以直接进行比较(Taigman等,2016;Isola等,2017;Royer等,2020)。本文方法旨在学习一个异质图像数据之间的域转换模型,在保留图像本身特征及信息的同

时,排除风格差异对变化检测任务的影响。具体地,构造双边对抗自编码网络实现图像转换,以处理异质遥感图像变化检测。给定不同时间点 $t_1$ 和 $t_2$ 获取来自同一物理区域的双时相遥感图像,经过配准等相关预处理后得到 $X \in \mathbf{R}^{W \times H \times C_1}$ 和 $Y \in \mathbf{R}^{W \times H \times C_2}$ ,图像尺寸为 $W \times H$ 。 $C_1, C_2$ 表示双时相图像的通道数。图1给出了提出方法的整体框架。该方法基于双边卷积自编码器,包含两对编—解码器及两个对应的判别网络。其中,不同损失函数的数据流用不同的颜色标记:重构损失(黑色)、结构一致性损失(黄色)和对抗损失(红色)。具体而言,在网络训练阶段,从原始异质遥感图像 $X$ 和 $Y$ 中随机选择图像块对 $I_1 \in \mathbf{R}^{\omega \times h \times C_1}$ 和 $I_2 \in \mathbf{R}^{\omega \times h \times C_2}$ ,其尺寸为 $\omega \times h$ ,分别输入双边卷积自编码器。在适当的正则化约束下,提出网络模型将首先对输入的异质遥感图像分别利用编码器进行编码,并在解码器输出端获得其高保真度重构图像,与此同时将输入转换到相对应的数据空间。

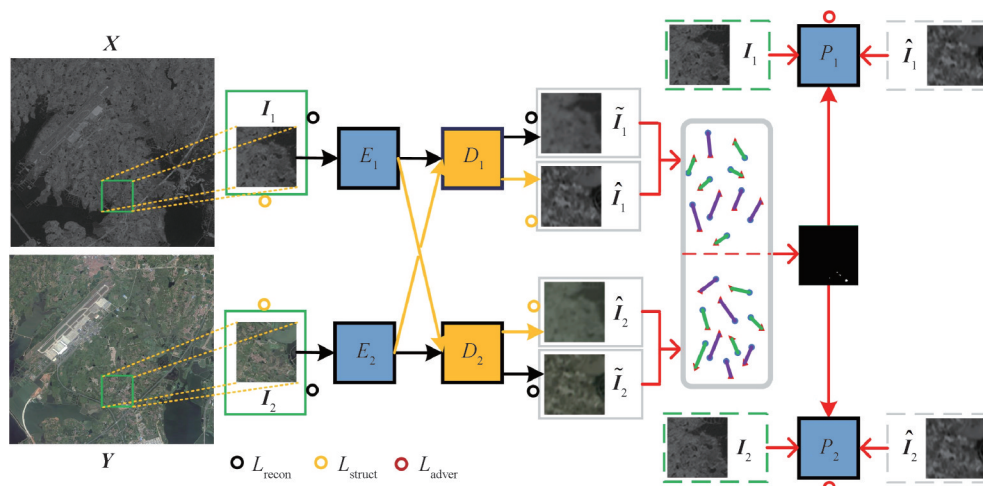


图1 BAACL方法框架

Fig. 1 The architecture of BAACL

### 1.1 算法框架

基于图像—图像(image to image, I2I)转换思想(Isola等,2017;Royer等,2020),提出方法的目标是将风格不同的两幅图像转换到具有一致特征表示的公共域,进而便于图像数据之间直接进行比较。双时相异质遥感图像可以被认为是对同一区域地物不同风格的呈现。因此,异质遥感图像变化检测这一关键任务的重点在于学习异构数据一对一的映射,以建立具有不同外观的图像之间的映射关系,在保

留图像语义、结构信息的同时排除“伪变化”对变化检测结果的影响。

BAACL实现异质遥感图像的数据转换,进而完成对变化区域的精确检测。具体地,通过将异质遥感图像 $X$ 和 $Y$ 转换至中间域 $S$ ,可以捕获一致的特征表示,利用像素级特征描述输入的双时相遥感图像。在数学上,该过程可定义为

$$S_x = F(X) \quad (1)$$

$$S_y = G(Y) \quad (2)$$

式中,  $F(\cdot)$  和  $G(\cdot)$  是将异质图像映射到中间域  $S$  的可学习的映射函数, 转换后的图像分别记为  $S_x$  和  $S_y$ 。因此, 通过度量  $S_x$  与  $S_y$  之间的像素级距离可获得差异图, 具体为

$$\Delta(i, j) = d(s_x(i, j), s_y(i, j)) \quad (3)$$

式中,  $\Delta$  为差异图, 由基于  $S$  域中  $X$  和  $Y$  的特征的像素级差分计算获得。 $(i, j)$  表示图像中像素的位置。 $s_x(i, j)$  和  $s_y(i, j)$  是  $X$  和  $Y$  在  $S$  域中位置为  $(i, j)$  的对应特征表示, 即  $S_x = \{s_x(i, j) | i = 1, \dots, W, j = 1, \dots, H\}$  和  $S_y = \{s_y(i, j) | i = 1, \dots, W, j = 1, \dots, H\}$ 。 $d(\cdot)$  为像素级的距离度量函数, 受到大规模数据计算成本的限制, 这里使用了欧几里得距离。

BAACL 异质遥感图像变化检测方法是基于两组自编码网络拟合  $P(\cdot)$  和  $Q(\cdot)$  函数分别进行对双时相遥感图像的重构, 同时拟合  $F(\cdot)$  和  $G(\cdot)$  函数实现对双时相遥感图像的转换。因此, 双边对抗自编码网络模型由两组编码器—解码器子网络组成:

1) 编码器  $E_1(X): X \rightarrow Z_x$  和相应的解码器  $D_1(Z): Z_x \rightarrow \tilde{X}$ , 即:  $\tilde{X} = P(X) = D_1(E_1(X))$ ;

2) 编码器  $E_2(Y): Y \rightarrow Z_y$  和相应的解码器  $D_2(Z): Z_y \rightarrow \tilde{Y}$ , 即:  $\tilde{Y} = Q(Y) = D_2(E_2(Y))$ 。

这里, 两组编—解码器对是由深度全卷积网络构建的, 图2给出了单边编—解码器的数据转换网络结构。 $Z_x$  和  $Z_y$  分别表示编码器  $E_1$  和  $E_2$  的编码层或隐空间。 $Z$  为  $Z_x$  和  $Z_y$  中张量的编码层表示。

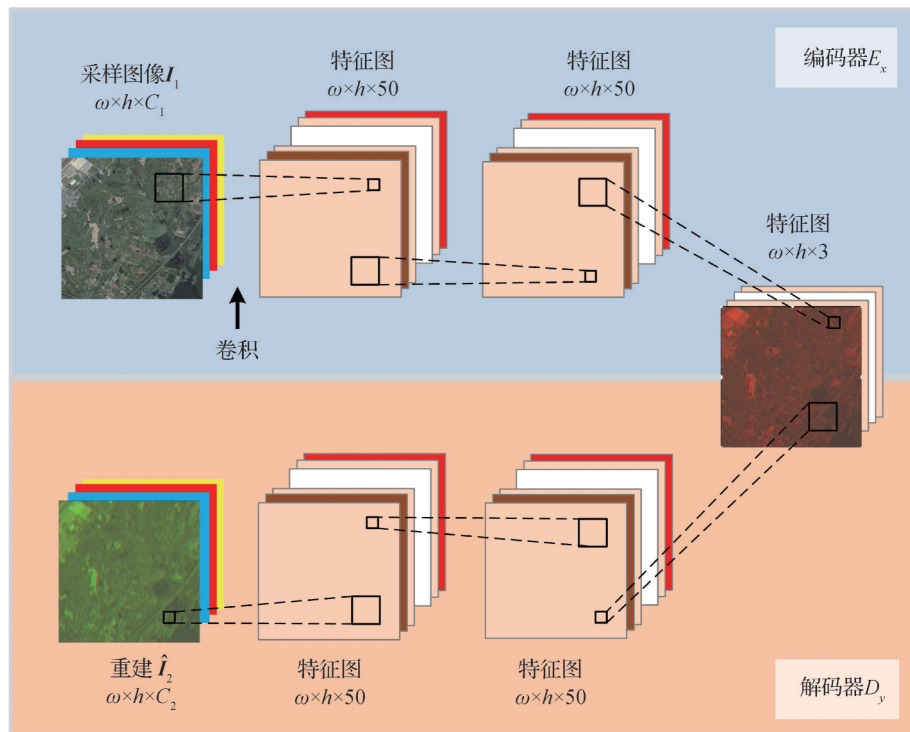


图2 BAACL中单边编—解码器网络结构

Fig. 2 Illustration of a pair of encoder-decoder structures in BAACL

因此, 在适当的正则化项的约束下, 双边自编码器可实现联合训练, 学习映射函数将输入图像映射到中间域  $S$ 。这一过程具体来说, 一方面, 通过编码器对图像进行编码, 并通过级联的解码器将编码层特征投影到原始域, 输出高保真重构; 另一方面, 图像数据被对面的解码器映射, 获得所寻求的样式转换。式(1)和式(2)给出的数据转化过程的详细数学表达为

$$S_x = F(X) = D_2(Z_x) = D_2(E_1(X)) \quad (4)$$

$$S_y = G(Y) = D_1(Z_y) = D_1(E_2(Y)) \quad (5)$$

然而, 在没有任何额外的约束项作用的情况下,  $S$  域中的表示  $S_x$  和  $S_y$  通常不会获得理想的数据样式转换结果, 不能被直接用来生成差异图像。因此, 基于图像结构信息及语义信息构建损失函数迫使网络训练实现数据对齐。也就是说, 如果在  $S$  域中实现  $S_x$  和  $S_y$  的结构及风格特征对齐, 则可完成较为精确

的变化检测。这里,映射函数 $F(\cdot)$ 和 $G(\cdot)$ 通过基于深度全卷积网络的双边自编码器实现。因此,输入维度 $(\omega \times h)$ 的图像块对,其结构及风格特征可以保留在整个自编码器的隐藏层中。双边自编码器的训练是通过最小化一个构建的关于网络参数集 $\theta$ 的损失函数 $L(\theta)$ 来实现。损失函数 $L(\theta)$ 由3个损失项组成:重构损失、结构一致性损失和对抗损失。

### 1.2 重构损失

BAACL网络目标函数的第一个正则化项为重构损失项,它的目标是强制网络编解码后获得的重构输出图像尽可能接近于输入图像。因此,重构损失项 $L_{\text{recon}}$ 可以表示为

$$L_{\text{recon}} = M(I_1, \tilde{I}_1) + M(I_2, \tilde{I}_2) \quad (6)$$

式中, $M$ 是用于度量期望输出和重构输出之间损失的均方误差(mean-square error, MSE)。

### 1.3 结构一致性损失

图2给出了单边卷积自编码器的数据转换过程。具体地,对于图像 $X$ ,对从其中随机选取的图像块 $I_1$ 使用编码器 $E_1$ 将数据转换到隐空间,再使用图像 $Y$ 对应的解码器 $D_2$ 对隐空间中图像块 $I_1$ 的特征进行解码,生成与图像 $Y$ 具有一致风格的图像块 $\hat{I}_2$ ,即: $\hat{I}_2 = F(I_1) = D_2(E_1(I_1))$ 。同样的,对于图像 $Y$ ,对从其中随机选取的图像块 $I_2$ 使用编码器 $E_2$ 将数据转换到隐空间,再使用图像 $X$ 对应的解码器 $D_1$ 对隐空间中图像块 $I_2$ 的特征进行解码,生成与图像 $X$ 具有一致风格的图像块 $\hat{I}_1$ ,即: $\hat{I}_1 = G(I_2) = D_1(E_2(I_2))$ 。为了使转换后的图像与原图像具有一致的图像结构特征,使用了如下的损失函数对数据转换过程进行约束,具体为

$$L_{\text{struct}} = \frac{1}{n^2} \left\{ \left\| B(I_2) - B(\hat{I}_1) \right\|_2 + \left\| B(I_1) - B(\hat{I}_2) \right\|_2 \right\} \quad (7)$$

为了实现异构图像数据的比较,确保图像转换前后内部结构保持不变,这里采用了 $B$ 分别计算异质图像的亲和矩阵(Wang等,2018),并利用欧氏距离度量亲和矩阵之间的距离,实现跨模态图像间的差异度量。

### 1.4 对抗损失

结构一致性损失项旨在约束图像转换前后所保留的结构信息的一致性。网络对于异质遥感图像风

格的转换能够取得理想的效果,需要额外设计损失项进行约束。也就是说,对于网络转换 $I_2$ 后生成 $\hat{I}_1$ ,网络训练的目标在于使得 $\hat{I}_1$ 与 $I_1$ 的风格尽可能接近,类似地,对于网络转换 $I_1$ 后生成 $\hat{I}_2$ ,网络训练的目标在于使得 $\hat{I}_2$ 与 $I_2$ 的风格尽可能接近,进而有利于差异图像能够更好地反映异构图像之间的变化信息,同时缓解“伪变化”对变化检测结果的影响。因此,提出一种基于语义信息的对抗损失函数来实现风格一致性约束。

#### 1.4.1 基于语义信息的对抗损失

具体来说,双边自编码网络可以看做是生成器,用于异构数据的转换,生成风格一致的图像。然而通过卷积自编码网络生成的图像是否能够达到理想的风格转换的目标,需要引入判别器对生成图像与原始图像之间的一致性进行判别,进而协同训练判别器与生成器达到理想的风格转换目标。因此,一种简单的思路就是将 $\hat{I}_1$ 和 $I_1$ 、 $\hat{I}_2$ 和 $I_2$ 作为数据对输入判别网络,通过区分转换输出图像与真实输入图像,构造对抗损失函数,迫使网络生成具有更加一致风格的图像。这一过程的数学表达为

$$L_{\text{adver}} = \min_F \max_{P_2} L_{\text{GAN},1 \rightarrow 2} + \min_G \max_{P_1} L_{\text{GAN},2 \rightarrow 1} \quad (8)$$

$$L_{\text{GAN},1 \rightarrow 2} = E_{I_2 \sim P_{\text{data}}(I_2)} \left[ \log P_2(I_2) \right] + E_{I_1 \sim P_{\text{data}}(I_1)} \left[ \log \left( 1 - P_2(\hat{I}_1) \right) \right] \quad (9)$$

$$L_{\text{GAN},2 \rightarrow 1} = E_{I_1 \sim P_{\text{data}}(I_1)} \left[ \log P_1(I_1) \right] + E_{I_2 \sim P_{\text{data}}(I_2)} \left[ \log \left( 1 - P_1(\hat{I}_2) \right) \right] \quad (10)$$

式中, $P_1$ 和 $P_2$ 表示使用全连接网络构建的判别器。然而,将 $\hat{I}_1$ 和 $I_1$ 、 $\hat{I}_2$ 和 $I_2$ 作为数据对通过判别网络进行逐像素的区分时,由于图像是来自不同时间获取的数据,存在以下两种情况:

1)当像素对属于背景区域(非变化类)时:生成网络的目标是使图像像素对风格尽可能一致,减少异构数据转换差异以及“伪变化”可能造成的虚警,因此,网络优化的目标是使得对抗损失尽可能减小。

2)当像素对属于变化区域(变化类)时:这个区域的像素本质是不同的,即使生成网络进行的数据风格转换一致,像素对之间仍然会存在较大不同,因此,在构造对抗损失时需要消除这部分像素对对网络优化的影响。

基于以上分析,提出利用像素对的语义一致性修正对抗损失函数来约束网络训练。具体来说, $\hat{I}_1$ 和 $\hat{I}_2$ 和 $I_2$ 中像素对属于不变区域时,在风格和语义上均具有一致性,需要通过网络训练拉近两者的距离。当像素对属于变化区域时,在语义上不具有 consistency,需要将这部分像素剔除,减少对网络训练的影响。因此,利用语义信息 $I_{\text{sem}}$ 实现对抗损失函数的约束,式(9)和式(10)转换为

$$L_{\text{GAN},1 \rightarrow 2} = E_{I_2 \sim P_{\text{data}}(I_2)} [\log P_2(I_2 \times I_{\text{sem}})] + E_{I_1 \sim P_{\text{data}}(I_1)} [\log (1 - P_2(\hat{I}_2 \times I_{\text{sem}}))] \quad (11)$$

$$L_{\text{GAN},2 \rightarrow 1} = E_{I_1 \sim P_{\text{data}}(I_1)} [\log P_1(I_1 \times I_{\text{sem}})] + E_{I_2 \sim P_{\text{data}}(I_2)} [\log (1 - P_1(\hat{I}_1 \times I_{\text{sem}}))] \quad (12)$$

#### 1.4.2 基于聚类分析的语义信息

上文所构建的对抗损失函数存在一个亟待解决的问题:语义信息约束 $I_{\text{sem}}$ 的构造。由于双边卷积自编码网络可学习到某种转换,将图像数据映射到公共数据域,更有利于对语义信息的获取。一个自然的想法是拉近相似像素对(背景区域)的距离,并扩大变化区域像素对的距离。然而,在无监督的方式下,缺乏类标能够训练模型对数据进行严格的划分。因此,关注于聚集具有不同语义信息的像素对。这一过程的目标函数 $L$ 可以基于类似K-均值聚类(Yu等,2020)的形式来构造语义约束 $I_{\text{sem}}$ ,具体为

$$\min_{\{c_{1,k}\}_{k=0}^1} L(I_2, I_1) = \frac{1}{n} \sum_{k \in \{0,1\}} \sum_{i=1}^n \left\| \log \frac{F(I_{2,i})}{P(I_{1,i})} - c_{1,k} \right\|^2 \quad (13)$$

$$\min_{\{c_{2,k}\}_{k=0}^1} L(I_1, I_2) = \frac{1}{n} \sum_{k \in \{0,1\}} \sum_{i=1}^n \left\| \log \frac{G(I_{1,i})}{Q(I_{2,i})} - c_{2,k} \right\|^2 \quad (14)$$

式中, $n = \omega \times h$ 为像素对总数, $i$ 表示第 $i$ 个像素点。 $c_{1,k}$ 和 $c_{2,k}$ 表示聚类中心,其取值“1”表示变化类,“0”表示非变化类,初始时采用随机中心。这里, $F, G, P$ 及 $Q$ 为通过训练双边卷积自编码器获得数据的非线性转换函数。通过分别最小化目标函数式(13)和式(14)获取两个通道图像数据的语义信息 $C_1 \in \{0, 1\}^{\omega \times h}$ 及 $C_2 \in \{0, 1\}^{\omega \times h}$ 。这样,利用网络训练获得的 $S$ 域中图像对构造语义约束,具体为

$$I_{\text{sem}} = \max_{i_{\text{sem}} \in \{0,1\}} (C_1, C_2) \quad (15)$$

基于此构建对抗损失函数,缓解变化区域对于网络训练的不利影响。图3给出了基于聚类分析的

语义信息构造流程。图像块 $I_1$ 和 $I_2$ 分别经过双边卷积自编码网络训练获得的转换函数 $F, G, P$ 及 $Q$ 进行数据映射后,通过最小化目标函数拉近相似像素对的距离,推远变化区域像素对的距离,进而获取语义信息。

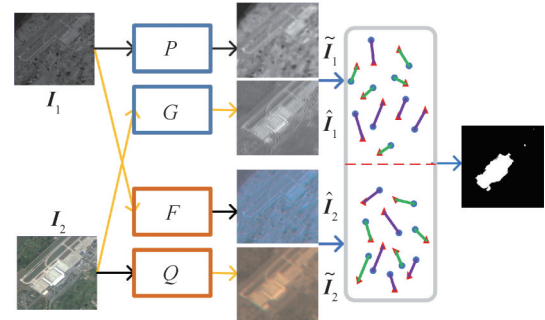


图3 基于聚类分析的语义信息构造流程

Fig. 3 Semantic information construction process based on clustering analysis

#### 1.5 网络优化

总损失函数定义为3个损失项的加权平均,具体为

$$L_{\text{total}} = L_{\text{recon}} + \alpha L_{\text{struct}} + \beta L_{\text{adver}} \quad (16)$$

式中, $\alpha$ 和 $\beta$ 是用来平衡3个损失项的权值。训练完成后,利用式(3)获得差异图像 $\Delta$ 。使用经典的Otsu方法获得最终的变化检测结果图。该阈值选择算法的关键是获得最优阈值。因此,将所有差异图的像素值与阈值进行比较。若大于阈值,则认为像素属于变化类,反之亦反。

##### 算法1:BAACL

输入:随机获取的训练图像对 $I_1$ 与 $I_2$ 。

##### 1)网络训练:

(1)构建双边卷积自编码网络及重构损失和结构一致性损失函数;

(2)将当前图像对 $I_1$ 与 $I_2$ 投入双边卷积自编码网络生成两组同质图像 $\tilde{I}_1, \hat{I}_1, \tilde{I}_2, \hat{I}_2$ 。

(3)对两组同质图像 $\tilde{I}_1, \hat{I}_1, \tilde{I}_2, \hat{I}_2$ 在其数据域中进行聚类分析,得到当前样本对的语义约束 $I_{\text{sem}}$ 。

(4)基于语义约束 $I_{\text{sem}}$ 以及 $\hat{I}_1$ 和 $I_1, \hat{I}_2$ 和 $I_2$ 构建判别器的对抗损失函数,联合重构损失函数和结构一致性损失函数训练双边自编码器网络。

2)训练结束,将图像 $X$ 和 $Y$ 投入当前网络,得到最终的差异图,并对差异图进行分析得到最终变化检测结果。

## 2 实验设置及仿真分析

### 2.1 实验数据

实验采用了4组公开的异质遥感图像数据集进行实验测试来验证提出的BAACL网络性能,分别是:Italy数据集(Sun等,2021)、California数据集(Sun等,2022)、Tianhe数据集(Luppino等,2024)以及Shuguang数据集(Sun等,2021)。图4给出了这4组数据集的双时相图像及变化参考图。

1)Italy数据集。由Landsat-5卫星在意大利撒丁岛上空采集获得,尺寸为 $412 \times 300$ 像素。其中,第1幅是1995年9月采集的近红外(near-infrared, NIR)波段图像,第2幅是1996年7月在同一地区采集的RGB波段图像。地面真实变化图显示了这段时间内撒丁岛的湖泊扩张的真实情况。

2)California数据集。该组数据集是在美国加利福尼亚州上空采集获得的,范围覆盖了萨克拉门托县、尤巴县和萨特县,尺寸为 $3\,500 \times 2\,000$ 像素,由于数据过大,在实际中将数据集下采样至 $875 \times 500$ 像素用于网络训练和验证。其中,第1幅是2017年1月5日的RGB波段图像,由Landsat 8卫星的陆地

成像仪(operational imager, OLI)和热红外传感器(thermal infrared sensor, TIRS)共同采集的11通道图像,光谱范围从深蓝到红外波段,第2幅是Sentinel-1 A卫星在2017年2月18日采集的SAR图像,以垂直发射垂直接收(vertical-vertical, VV)和垂直发射水平接收(vertical-horizontal, VH)两种极化方式获取。真实变化图反映了这段时间内洪水泛滥对该地区所造成的影响。

3)Tianhe数据集。该数据集为在中国武汉天河机场上空所采集的图像组,尺寸为 $666 \times 615$ 像素,范围覆盖 $7.28 \times 6.73 \text{ km}^2$ 的地表,空间分辨率近似为11 m/像素。其中,第1幅是2002年7月由Landsat-7卫星获取的全色图像,第2幅是2013年6月来自Google Earth的光学图像。真实变化图反映了天河机场的建筑物变化情况。

4)Shuguang数据集。采集于中国东营市曙光村的上空,图像尺寸为 $593 \times 921$ 像素,空间分辨率近似为8 m/像素。其中,第1幅是2008年6月Radar-sat-2卫星的C波段传感器所获取的SAR图像,第2幅是2012年9月来自Google Earth的光学图像。真实变化图反映了曙光村的建筑扩张情况。

### 2.2 实验设置及评价指标

提出的BAACL网络的主干部分由两对卷积自编码器组成,其中编码器由2个 $3 \times 3 \times 50$ 的卷积层加LeakyReLU激活函数,后跟1个 $3 \times 3 \times 3$ 的卷积层加Tanh激活函数组成,最终得到 $100 \times 100 \times 3$ 的隐藏层特征;解码器的最后一层是1个 $3 \times 3 \times C$ 的卷积层加Tanh激活函数,从而将特征还原到相应域空间的图像数据尺寸。LeakyReLU函数的斜率设定为0.3。学习率为 $1 \times 10^{-3}$ 。Tanh函数将输入的图像数据归一化到 $[-1, 1]$ ,以加快收敛速度(LeCun等,2012)。提出的网络模型使用TensorFlow实现,在GeForce RTX 2080 Ti显卡和11 GB内存上实现训练和验证过程。在网络训练过程中,使用自适应学习速率的随机梯度下降优化方法Adam更新误差,总共训练了200轮。在每一轮训练时,随机生成10个 $100 \times 100$ 像素的图像块对投入网络进行训练。对于对抗部分,使用全连接网络构建判别器 $P_1$ 和 $P_2$ ,神经元的数量按层数渐进依次为25、100、200、50、1。

通过以下5个定量分析指标对所提出网络的性能进行验证,包括:ROC(receiver operating characteristic)曲线下面积(area under curve, AUC),ROC曲线

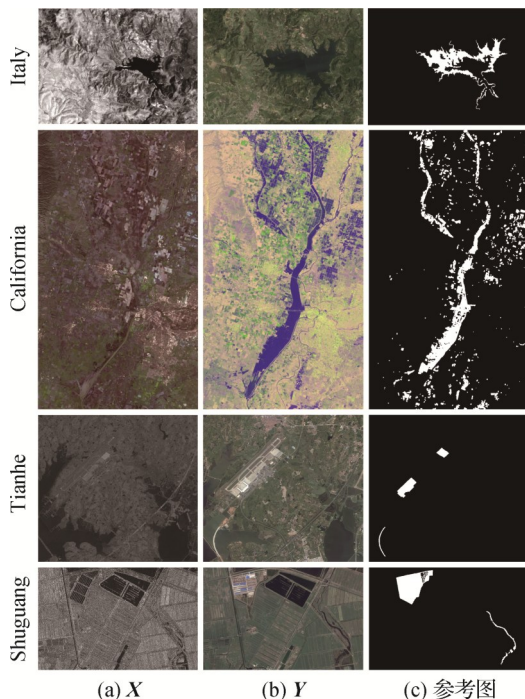


图4 4组异质遥感图像变化检测数据集

Fig. 4 Four heterogeneous remote sensing image change detection datasets ((a) X; (b) Y; (c) ground truth)

通过改变分类的阈值,进而得到一系列的(FPR, TPR)二维坐标点,之后根据阈值由小到大得到的点绘制而成;假阳性(false positives, FP),即将不变区域的像素错误判别为变化的像素数;假阴性(false negatives, FN),即将变化区域的像素错检为不变的像素数;总体精度(overall accuracy, OA),即正确分类的像素占像素总量的比例以及Kappa系数(Brennan和Prediger, 1981)。

2.3 参数分析

BAACL训练可以通过最小化总损失函数(式(16))实现。其由3个损失项组成,每个损失项指代一个特定目标,以引导网络进行训练获得期望的目标。重构损失项 $L_{recon}$ 可以促使网络重构高保真输出图像,即 $\tilde{I}_1$ 和 $\tilde{I}_2$ ;结构一致性损失项 $L_{struct}$ 和对抗损失项 $L_{adver}$ 确保网络实现异构数据的空间对齐。其中,结构一致性损失实现 $I_1$ 到 $\hat{I}_2$ , $I_2$ 到 $\hat{I}_1$ 的图像内部结构特征对齐;对抗损失实现 $I_1$ 到 $\hat{I}_1$ , $I_2$ 到 $\hat{I}_2$ 的风格特征对齐。提出的总损失函数使用权值 $\alpha$ 和 $\beta$ 来平衡

3个损失项的贡献。为了方便地确定权值 $\alpha$ 和 $\beta$ 的影响,采用了不同的权值组合 $(\alpha, \beta)$ 测试其对变化检测结果的影响,分别为(1, 2)、(2, 1)、(2, 2)以及(1, 1)。表1给出了BAACL采用不同权值组合在4个不同的数据集上获取的定量分析结果。可以看出,不同的权值组合对于最终变化检测结果的影响在4组数据集上有所不同,总体看来,当权值组合为(1, 1)时的结果在4组数据集上表现最好。同时,数据相关文献(Luppino等, 2024; Jia等, 2022)的推荐,将权值 $\alpha$ 和 $\beta$ 均设置为1。

2.4 对抗损失优化消融实验

设计的基于语义信息的生成对抗损失对网络训练进行约束是创新点之一,目的是基于更加精确的模型构建取得更好的网络优化结果,提高变化检测性能。为了验证该损失项对网络优化训练约束的有效性,设计了一组消融实验来进行对比验证。

基于语义信息的对抗损失的一大特征是对输入

表1 在4组数据集上BAACL采用不同权值组合获得的定量结果对比  
Table 1 Quantitative results obtained by BAACL with different weights on four datasets

| 数据集        | $(\alpha, \beta)$ | AUC            | OA             | Kappa          | FP           | FN            |
|------------|-------------------|----------------|----------------|----------------|--------------|---------------|
| Italy      | (1, 2)            | 0.910 1        | 0.970 1        | <b>0.768 0</b> | 1 728        | <b>1 615</b>  |
|            | (2, 1)            | 0.917 9        | 0.968 0        | 0.705 0        | <b>454</b>   | 3 086         |
|            | (0.5, 0.5)        | 0.859 5        | 0.934 5        | 0.377 4        | 3 223        | 4 817         |
|            | (1, 1)            | <b>0.918 0</b> | <b>0.970 5</b> | 0.701 2        | 626          | 3 022         |
| California | (1, 2)            | <b>0.923 3</b> | 0.933 5        | 0.503 7        | 9 519        | 19 562        |
|            | (2, 1)            | 0.915 6        | 0.928 5        | 0.508 5        | 13 653       | 17 632        |
|            | (0.5, 0.5)        | 0.880 7        | 0.916 0        | 0.476 1        | <b>5 911</b> | 22 103        |
|            | (1, 1)            | 0.897 0        | <b>0.938 2</b> | <b>0.566 5</b> | 10 750       | <b>16 242</b> |
| Tianhe     | (1, 2)            | 0.975 6        | 0.993 3        | 0.663 1        | <b>751</b>   | 1 979         |
|            | (2, 1)            | 0.987 5        | 0.988 5        | 0.625 9        | 4 007        | <b>687</b>    |
|            | (0.5, 0.5)        | 0.995 0        | 0.986 2        | 0.580 1        | 4 972        | 691           |
|            | (1, 1)            | <b>0.997 7</b> | <b>0.994 7</b> | <b>0.781 1</b> | 1 314        | 827           |
| Shuguang   | (1, 2)            | 0.894 0        | 0.972 2        | 0.557 8        | 145          | 15 048        |
|            | (2, 1)            | 0.915 7        | 0.970 8        | 0.524 7        | <b>47</b>    | 15 879        |
|            | (0.5, 0.5)        | 0.871 8        | 0.969 5        | 0.497 0        | 110          | 16 495        |
|            | (1, 1)            | <b>0.941 7</b> | <b>0.982 6</b> | <b>0.772 0</b> | 1 221        | 8 243         |

注:加粗字体表示各项指标在不同数据集上的最优结果。

图像数据使用聚类分析所获取的语义信息进行掩码操作,从而在损失函数的约束项中拉近多时相遥感图像不变区域的距离。对照组BAS(bipartite adversarial autoencoders)的对抗损失函数则没有进行该操作,其对抗损失函数的构造基于整个图像,对抗损失函数的收敛会拉进双时相图像变化区域间的距离,而这本身是要远离的,造成网络约束项定义不精确,影响变化检测性能。因此,当前训练样本的变化区域占比越大,则训练出来的网络效果将会越差。而BAACL由于进行了掩码操作,缓解变化区域对网络优化的影响,所得到的结果也会更加稳定。

为了验证语义信息约束的有效性,针对性地设计了一组消融实验。具体地,在California数据集上进行监督采样(仅该组消融实验中采用监督方式,实际训练过程为无监督方式),训练样本同样为 $100 \times 100$ 像素的图像块对,根据对应的参考图统计变化区域在整个样本区域的占比(%),将随机选取的训练样本按照变化区域不同占比划分成10个训练样本组,测试变化区域像素对网络优化的影响,分别为:(0, 10], (10, 20], (20, 30], (30, 40], (40, 50], (50, 60], (60, 70], (70, 80], (80, 90], (90, 100]。因此,在该组消融实验中,BAACL与没有语义约束的BAS形成10个对照组,并分别投入网络进行训练和验证。图5给出了利用这10个训练样本组训练网络所获得的最终变化检测结果的视觉效果对比。图5最右列为分别对应于10个训练样本组的训练样本示例,表明了训练样本组的语义差异特性,用于说明变化区域占比大小。BAACL在分别利用这10个特定训练集完成网络训练之后,将California数据集投入网络获得最终的变化检测结果(见图5(a))。为了更加有效地说明语义信息约束的有效性,在采用相同类型的训练样本的情况下,与对照组BAS所获得的变化检测结果进行对比(如图5(b)所示),可以看出,变化区域占比的不断增加对BAS性能的影响较大,而BAACL结果相对较为稳定。

采用California数据集进行采样基于以下两点:一是该组消融实验使用的变化检测数据集较小,随机采样会导致样本间相似性增大,经过多轮迭代会使网络倾向于过拟合;二是设计了10个对照组,不仅进行不同算法之间的结果比较(BAS和BAACL),还进行对照组间的比较,可以测试随着样本实际变化区域占比的增大对网络性能的影响。因此,在实

际检测中采用源自同一数据集的随机样本训练网络,使对照组间在相同的实验环境下进行网络训练,增加结果的可信度。可以看出,随着训练样本的变化区域占比增大,BAS网络的性能随之下降。而BAACL网络由于对变化区域进行了掩码操作,其性能并没有随着样本变化区域的占比的变化而剧烈变化,这验证了以上理论分析的结果,即所提出的基于语义信息的对抗损失函数对网络优化的有效性。

此外,设计了一组消融实验测试基于语义信息的对抗损失项的作用及其对最终变化检测结果的影响。在该组实验中,将仅使用重构损失 $L_{\text{recon}}$ 的网络模型作为基准,记做模型1。随后将结构一致性损失和对抗损失这两个损失项逐一添加到基准模型中,获得使用 $L_{\text{recon}} + L_{\text{adver}}$ 的模型2以及使用3个损失项 $L_{\text{recon}} + L_{\text{struct}} + L_{\text{adver}}$ 的BAACL。为了直观比较这些模型的变化检测性能,图6给出了这3个模型在4组数据集上获得的变换检测最终结果。模型2利用对抗损失函数可以提高背景区域的一致性表征,使异质图像风格更加一致,减缓“伪变化”造成的假阳性样本数量多的现象。与模型1相比,模型2在4个数据集上都取得了改进(如图6(b)所示)。在California、Tianhe以及Shuguang这3组数据集上表现尤为突出,验证了对抗损失项的有效性。此外,随着结构一致性损失项的加入,BAACL可以更好地捕捉图像结构自相似一致性表征,有效地排除干扰因素的影响。如图6(c)所示,这一损失项在Italy数据集上表现得更加明显,变化检测结果图清楚地表明了这一损失项的重要性。

## 2.5 实验对比分析

为了验证本文算法的有效性,选取了目前比较流行的5种传统及基于深度学习的异质遥感图像变化检测方法作为对比算法,这5种方法分别为:NPSG(Sun等,2021)、INLPG(Sun等,2022)、ACE(adversarial cyclic encoder)(Luppino等,2022)、CAA(code-aligned autoencoders)(Luppino等,2024)以及BAS(Jia等,2022)。这些方法对应的文献中都通过仿真实验验证了其在一定程度上表现出良好的检测性能,实验部分对于每种算法相关参数设置与原文保持不变。除了前文提到的定量分析指标结果,还给出了所有算法的运行时间,将对比较算法的计算复杂度进行了比较。

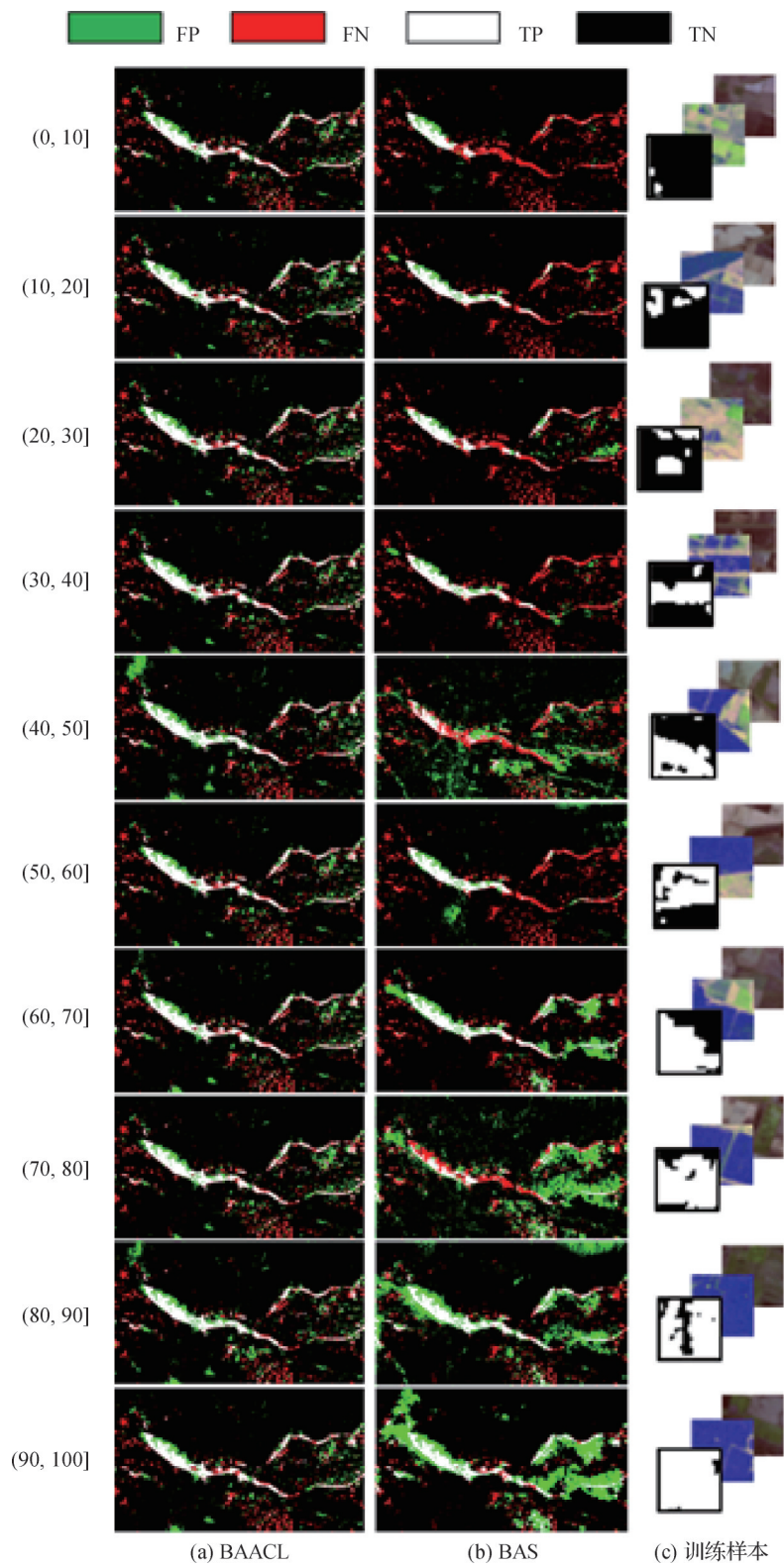


图5 基于语义信息的对抗损失效果

Fig. 5 Effectiveness of the semantic information based adversarial loss((a)BAACL;(b)BAS;(c)trained samples)

2. 5. 1 Italy 数据集

在Italy数据集上对提出方法以及5种对比算法进行实验测试。图7给出了在Italy数据集上采用6种方法得到的最终变化图,表2为Italy数据集上的

定量分析结果。Italy数据集的检测难点在于其图像的采集波段不同,并且图像获取的是崎岖不平的山区地貌,山脉高度的剧烈波动使得图像产生色差,让原本应当是不变区域的像素点容易被检测为变化区

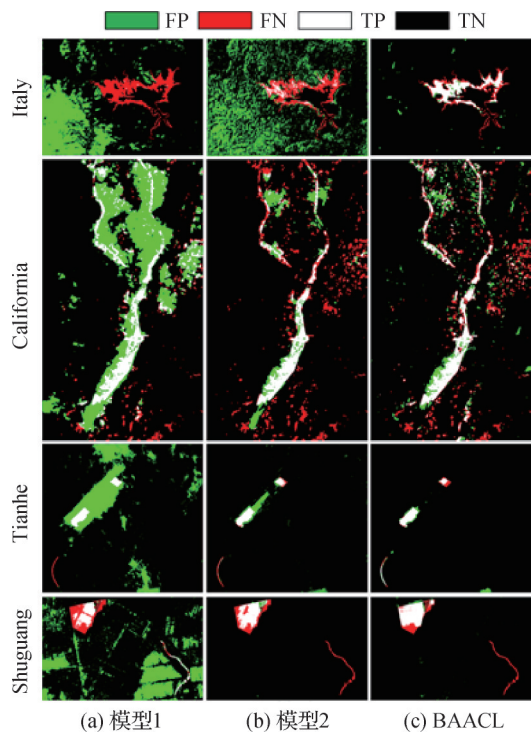


图6 在4组数据集上消融实验获得的变化检测结果  
Fig. 6 Change detection results obtained by ablation study on four datasets((a)model one;(b)model two;(c)BAACL)

域,导致较高的假阳性(FP),由变化图中的绿色标记。从图7(c)可以看出,NPSG方法所产生的总体精度最低(0.816 5),将撒丁岛的湖心误检为变化区域。同样的情况在INLPG上也有一定程度的体现。

而BAACL在FP上是最低的,其统计结果如表1所示,只有626个,表明BAACL在抑制背景噪声上与其他对比方法相比较更具优势。而假阴性(FN),即将变化区域错检为不变区域的情况,主要体现在湖泊发生扩张后湖和山地的交界处。BAACL的FN值在6种方法中处于中游位置,但是结合FP和FN进行综合分析会发现,BAACL错误检测的像素数(FP+FN)是6种方法中最低的,在背景中噪声点较少的同时能够较好地检测出变化区域,假阳性与假阴性之间取得了一个较好的平衡,因此获得了最好的Kappa结果,达到0.701 2。

BAACL以及5种对比方法在Italy数据集上的运行时间如表2所示。可以看到,INLPG的计算效率远超其他方法,仅为22.3 s。而基于传统方法的INLPG在其他数据集上的性能并没有如此突出,说明INLPG模型随着数据集的增大,其计算量将会随之大幅增加。而BAACL方法紧随其后,其计算时间是第2低的,为103.5 s。

同时,图8给出了6种方法在Italy数据集上的差异图像的ROC曲线,与ROC相关的AUC值如表2所示。可以看出,提出方法的AUC值略高于BAS的AUC值。而提出的BAACL方法的OA和Kappa结果是6种方法中最高的,相比对比方法结果有显著提升。

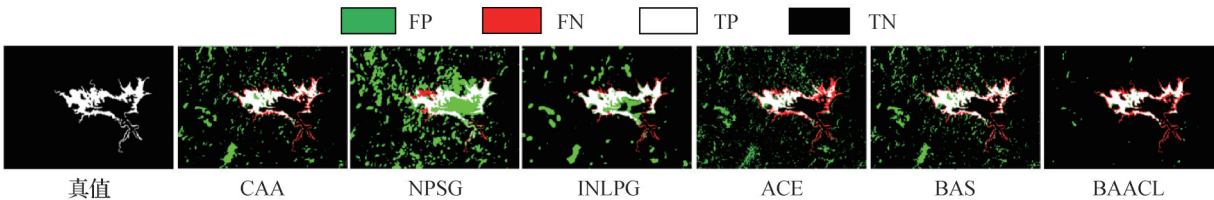


图7 Italy数据集上获取的变化图对比  
Fig. 7 Comparison of chage maps obtained on the Italy dataset

表2 Italy数据集上的定量结果  
Table 2 Quantitative results on Italy dataset

| 方法    | AUC            | OA             | Kappa          | FP         | FN           | 运行时间/s      |
|-------|----------------|----------------|----------------|------------|--------------|-------------|
| CAA   | 0.900 8        | 0.930 4        | 0.516 7        | 6 288      | 2 313        | 389.3       |
| NPSG  | 0.897 8        | 0.816 5        | 0.285 1        | 21 273     | <b>1 408</b> | 310.8       |
| INLPG | <b>0.948 6</b> | 0.930 8        | 0.552 0        | 3 022      | 6 999        | <b>22.3</b> |
| ACE   | 0.872 2        | 0.916 8        | 0.545 0        | 3 470      | 6 812        | 150.1       |
| BAS   | 0.912 0        | 0.927 5        | 0.518 0        | 6 931      | 2 035        | 103.7       |
| BAACL | 0.918 0        | <b>0.970 5</b> | <b>0.701 2</b> | <b>626</b> | 3 022        | 103.3       |

注:加粗字体表示各列最优结果。

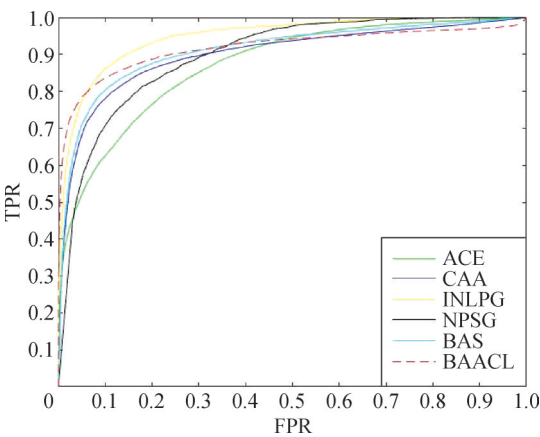


图8 Italy数据集上的差异图ROC曲线  
Fig. 8 ROC curves of the difference images on Italy dataset

2.5.2 California数据集

BAACL及对比算法在California数据集上生成的变化图如图9所示。California数据集的检测难点在于其洪灾发生在平原地区,洪水所造成的主河道的拓宽易于检测,但是河道周边点状分布的小湖泊和水潭却难以准确识别。这一点在变化图上有所体现,对于参考图中片状和点状分布的变化区域,完全有效地全部检测出来存在较大难度。尽管如此,相比于其他对比方法将点状区域大量错检为变化或者不变区域,BAACL方法在准确检测出河道受洪水

影响而拓宽的区域的基础上,对于河道周边的点状区域的错检像素数量是最少的。这一点在表3中的数据中可以看出,BAACL方法的像素错检数量(FP + FN)最少。相应地,其准确率(OA)为6种方法中最高,达到0.938 2。

6种方法在California数据集上的运行时间如表3所示。作为4组中最大的数据集,其运行时间也最高。从表3中可以看到,传统的NPSG和INLPG所花费时间远大于基于深度学习的变化检测方法。其中,NPSG的计算时间高达1 668.4 s,其改进模型INLPG也需要962.0 s。而深度学习方法的运行时间大幅度减小,其中CAA的时间最高,需要516.8 s,相比于INLPG运行时间减少了445.2 s。而BAACL的计算时间是所有方法中最低的,仅325.6 s,表明BAACL网络训练能够快速收敛并输出变化检测结果。

所有对比方法的ROC曲线如图10所示,其中INLPG的ROC曲线最接近(0, 1)点,其所对应的AUC值也是最高的。而BAACL方法的ROC曲线紧随其后,取得了较好的AUC值结果。尽管BAACL的ROC曲线逊于INLPG,但是其在抑制背景噪音方面是所有方法中最好的,这一点通过比较FP值可以看出,BAACL的FP值是最小的。对于表示检测结果均衡性的Kappa值,BAACL的结果是所有方法中最

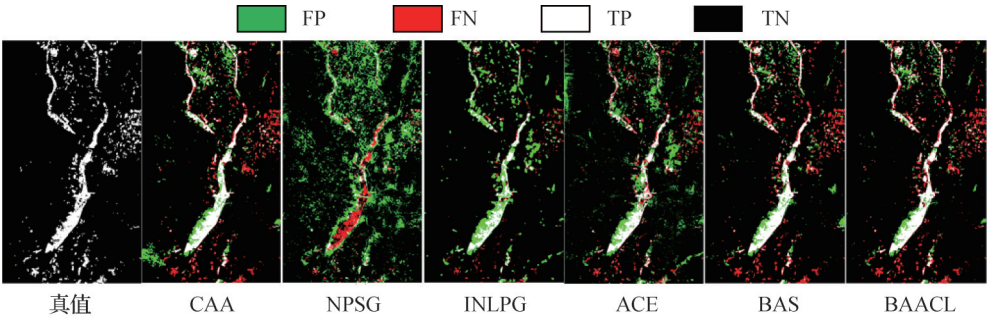


图9 California数据集上获取的变化图对比  
Fig. 9 Comparison of change maps obtained on the California dataset

表3 California数据集上的定量结果  
Table 3 Quantitative results on California dataset

| 方法    | AUC            | OA             | Kappa          | FP            | FN           | 运行时间/s       |
|-------|----------------|----------------|----------------|---------------|--------------|--------------|
| CAA   | 0.888 3        | 0.929 7        | 0.550 6        | 16 239        | 14 464       | 516.8        |
| NPSG  | 0.749 1        | 0.844 4        | 0.148 4        | 85 543        | 8 194        | 1 668.4      |
| INLPG | <b>0.939 7</b> | 0.936 2        | 0.460 9        | 22 628        | <b>5 260</b> | 962.0        |
| ACE   | 0.850 3        | 0.911 9        | 0.412 4        | 27 898        | 10 663       | 408.2        |
| BAS   | 0.882 0        | 0.935 3        | 0.558 3        | 12 390        | 15 859       | 327.2        |
| BAACL | 0.897 0        | <b>0.938 2</b> | <b>0.566 5</b> | <b>10 750</b> | 16 242       | <b>325.6</b> |

注:加粗字体表示各列最优结果。

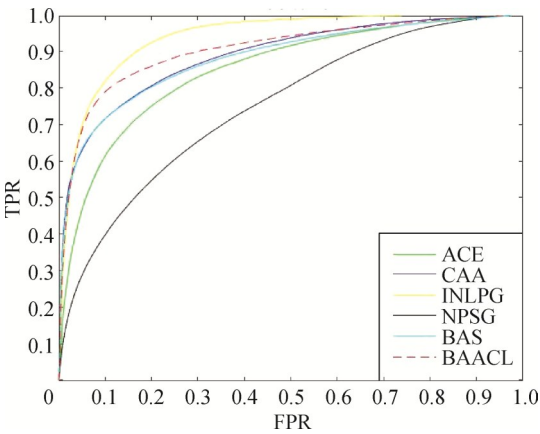


图 10 California数据集上的差异图 ROC 曲线

Fig. 10 ROC curves of the difference images on California dataset

高的,说明该方法在抑制背景噪音和检测变化区域上取得了很好的平衡。

2. 5. 3 Tianhe 数据集

图 11 给出了 BAACL 及对比方法在 Tianhe 数据集上所得到的变化图。Tianhe 数据集的变化主要为新建两处建筑和一条道路。其变化检测的难点在于两处建筑是在旧建筑的两端所进行的扩建,在检测过程中往往容易将旧建筑也误检为变化区域。加之,异构图像所造成的风格差异使得精确检测变化区域更加困难。这两点在 ACE 和 CAA 上的结果由图 11 可以看出,CAA 准确地检测出了两处新建筑,但是也将中间的旧建筑错误地判定为变化区域,同时两种方法也将周边大量的不变区域错检为变化区域。相比于 CAA 和 ACE,其他方法在 Tianhe 数据集上取得了不错的效果,均能够准确地检测出新增建

筑和道路,同时对于不变区域也没有产生大量的错判。而 BAACL 是其中效果最好的,既能准确检测出新建的两处建筑和一条道路,也没有受背景噪音的影响将其他区域错判为变化区域。

表 4 给出的评价结果从侧面佐证了上文对变化图的分析。CAA 和 ACE 的 FP 值远远高于其他方法,说明有大量的不变区域被错误地判别为变化区域。BAACL 的错检像素数量是最少的,因此具有最高的 OA 值,达到 0.994 7。图 12 为所有方法的 ROC 曲线图,可以看出六者的曲线均十分接近于理想的 (0, 1) 点,其中以 BAACL 和 INLPG 的曲线最为接近,两者曲线几乎重合。这一点从表 4 的 AUC 值也得到了印证,BAACL 的 AUC 值是所有方法中最高的,而对比方法中只有 INLPG 以 0.02% 落后于 BAACL,其他方法的 AUC 值与两者有较大差距。此外,BAACL 具有最高的 Kappa 值,表明其在背景区域和变化区域像素量极不均衡的情况下,其 OA 值并没有完全依赖于不变区域的检测精度,而是在两类的检测上取得了平衡。

所有对比算法在 Tianhe 数据集上的计算效率比较如表 4 所示。可以看到,传统方法 INLPG 的运行时间位列第 3,仅次于 BAS 和 BAACL。在所有基于深度学习的变化检测方法中,BAS 方法所花费的计算时间最短,仅为 207.7 s。BAACL 方法以 0.7 s 的差距居于第 2,这是由于 BAACL 算法中嵌入了聚类分析的方法,在一定程度上增加了算法的运行时间,而 Tianhe 数据集的变化区域较小,语义约束对于判别网络训练的计算影响较小。

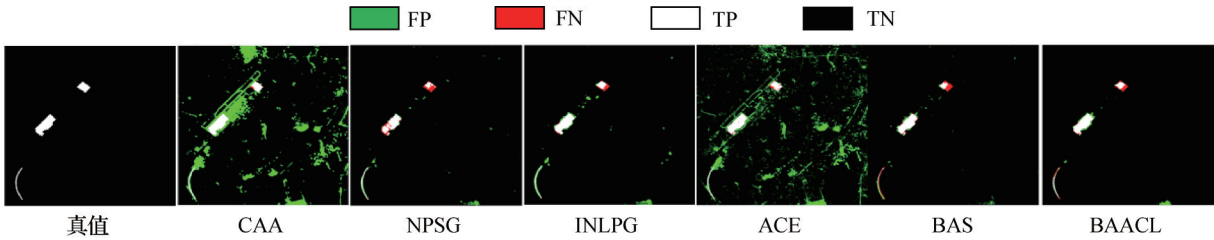


图 11 Tianhe 数据集上获取的变化图对比

Fig. 11 Comparison of change maps obtained on the Tionhe dataset

2. 5. 4 Shuguang 数据集

图 13 给出了所有方法在 Shuguang 数据集上获得的变化图。Shuguang 数据集的主要变化是左上方的农田区域新建了一片建筑,此外,右下方的一条由于地形变化引起的带状变化区域。该数据集的变化检测难点在于新建建筑的右上方有数片短条状的农

田属于不变化区域,这片农田区域受到季节、图像数据异构等因素的影响,容易误判为变化区域。此外,右下方的条状区域,由于其过于狭长,容易受到背景噪音影响而被误判为不变区域。

对图 13 给出的结果图进行对比,可以看出,CAA、NPSG、INLPG 和 ACE 的结果图建筑变化区域旁农田

表 4 Tianhe数据集上的定量结果  
Table 4 Quantitative results on Tianhe dataset

| 方法    | AUC            | OA             | Kappa          | FP         | FN         | 运行时间/s       |
|-------|----------------|----------------|----------------|------------|------------|--------------|
| CAA   | 0.989 2        | 0.915 9        | 0.199 9        | 34 260     | <b>135</b> | 439.3        |
| NPSG  | 0.993 6        | 0.993 8        | 0.731 4        | 1 202      | 1 331      | 1 118.8      |
| INLPG | 0.997 5        | 0.992 0        | 0.714 3        | <b>545</b> | 2 720      | 218.0        |
| ACE   | 0.978 3        | 0.928 3        | 0.216 6        | 29 099     | 261        | 353.5        |
| BAS   | 0.991 3        | 0.994 1        | 0.742 0        | 1 287      | 1 148      | <b>207.7</b> |
| BAACL | <b>0.997 7</b> | <b>0.994 7</b> | <b>0.781 1</b> | 1 314      | 827        | 208.5        |

注:加粗字体表示各列最优结果。

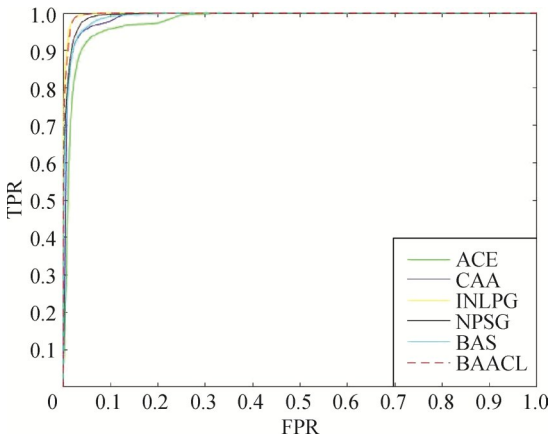


图 12 Tianhe数据集上的差异图 ROC 曲线  
Fig. 12 ROC curves of the difference images on Tianhe dataset

区域被误判为变化区域。而提出的BAACL方法准确地检测出了该区域为不变区域。这一点从表5的FP值可以看出,BAACL方法具有最低的FP值,说明其具有优秀的抑制背景噪音的能力。然而,其FN结果

较高,也就是将变化区域误判为不变区域,这样的情况下,BAACL的OA值仍然取得了最好的结果(0.982 6),说明在Shuguang数据集上,该网络整体在有效抑制背景噪音的基础上也可以准确检测出发生变化的区域,由BAACL的Kappa系数值最高得以印证,说明能够取得背景和变化区域之间较好的平衡。同时,图14给出表示网络分类性能的ROC曲线,参考与之对应的AUC,BAACL方法取得的结果一般。可以看出,虽然六者的曲线均接近理想的(0, 1)点,以NLPG和ACE的曲线的结果最好。这一点从表4的AUC值也得到了印证,ACE方法的AUC值是所有方法中最高的,而BAACL约0.2%低于这两种对比算法。

BAACL和5种对比算法在Shuguang数据集上的运行时间比较如表5所示。传统算法NPSG的计算时间最高,达到1 494.9 s,而其改进方法INLPG的

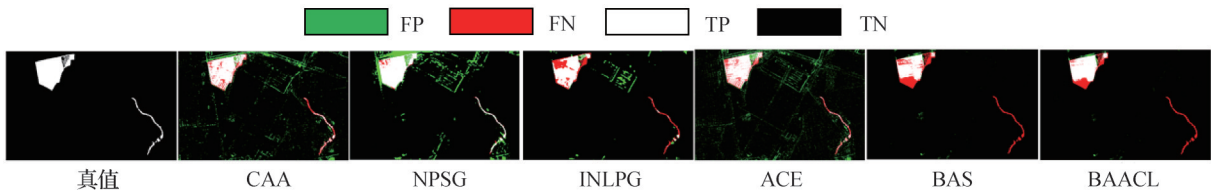


图 13 Shuguang数据集上获取的变化图对比  
Fig. 13 Comparison of change maps obtained on the shuguang dataset

表 5 Shuguang数据集上的定量结果  
Table 5 Quantitative results on Shuguang dataset

| 方法    | AUC            | OA             | Kappa          | FP           | FN           | 运行时间/s       |
|-------|----------------|----------------|----------------|--------------|--------------|--------------|
| CAA   | 0.876 2        | 0.958 6        | 0.639 7        | 13 241       | 3 699        | 377.4        |
| NPSG  | 0.965 7        | 0.955 1        | 0.638 8        | 23 304       | <b>1 235</b> | 1 494.9      |
| INLPG | 0.947 7        | 0.973 3        | 0.674 4        | 8 997        | 5 587        | 415.3        |
| ACE   | <b>0.967 9</b> | 0.950 8        | 0.576 3        | 4 870        | 22 024       | 493.6        |
| BAS   | 0.924 2        | 0.977 5        | 0.687 0        | 1 368        | 10 909       | 245.1        |
| BAACL | 0.941 7        | <b>0.982 6</b> | <b>0.772 0</b> | <b>1 221</b> | 8 243        | <b>242.9</b> |

注:加粗字体表示各列最优结果。

计算时间大幅减小,与CAA和ACE-Net的运行时间接近。所有对比算法中,BAACL的运行时间是最少的,仅花费242.9 s,表明BAACL在Shuguang数据集上能在最短的时间内优化网络获得最优参数组合,并输出最终的变化检测结果。

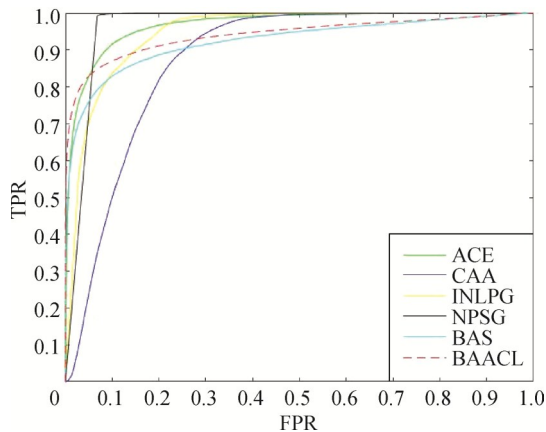


图14 Shuguang数据集上的差异图ROC曲线

Fig. 14 ROC curves of the difference images on Shuguang dataset

### 3 结 论

本文提出了一种基于嵌入聚类分析的双边对抗自编码网络实现异质遥感图像变化检测方法BAACL。双时相图像经过自编码网络实现重构的同时,充分利用双边网络结构特性,设计结构一致性损失和对抗损失约束网络优化过程,将异构数据映射到公共数据域中,实现异质图像风格转换。与此同时,在同质数据域中基于聚类分析获取语义信息,约束对抗损失来剔除其中不利于网络优化的干扰信息,可以实现异质遥感图像风格的进一步对齐。实验部分通过两组消融实验分析了基于语义信息的对抗损失函数对网络优化性能的影响。BAACL网络的整体性能在4组公开的异质遥感图像数据集上进行了验证,实验表明,与其他传统及基于深度学习的变化检测方法相比,提出的方法在最终变化图的视觉效果和定量分析结果上均优于5种对比方法。由于提出的BAACL网络的对抗损失部分减少了变化区域对网络优化的不利影响,可实现对图像转换结果的进一步提升,缓解“伪变化”及数据异构等因素的影响,从而提高检测精度,因此本算法适用于处理异质图像变化检测等相关问题。

### 参考文献(References)

- Bovolo F and Bruzzone L. 2005. A detail-preserving scale-driven approach to change detection in multitemporal SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 43(12): 2963-2972 [DOI: 10.1109/TGRS.2005.857987]
- Brennan R L and Prediger D J. 1981. Coefficient kappa: some uses, misuses, and alternatives. *Educational and Psychological Measurement*, 41(3): 687-699 [DOI: 10.1177/001316448104100307]
- Brunner D, Lemoine G and Bruzzone L. 2010. Earthquake damage assessment of buildings using VHR optical and SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 48(5): 2403-2420 [DOI: 10.1109/TGRS.2009.2038274]
- Fan W, Zhou M and Huang R. 2020. Multiscale deep features fusion for change detection. *Journal of Image and Graphics*, 25(4): 669-678 (樊玮, 周末, 黄睿. 2020. 多尺度深度特征融合的变化检测. *中国图象图形学报*, 25(4): 669-678) [DOI: 10.11834/jig.190312]
- Gong M G, Su L Z, Li H and Liu J. 2016. A survey on change detection in synthetic aperture radar imagery. *Journal of Computer Research and Development*, 53(1): 123-137 (公茂果, 苏临之, 李豪, 刘嘉. 2016. 合成孔径雷达影像变化检测研究进展. *计算机研究与发展*, 53(1): 123-137) [DOI: 10.7544/issn1000-1239.2016.20150662]
- Griffiths P, Hostert P, Gruebner O and van der Linden S. 2010. Mapping megacity growth with multi-sensor data. *Remote sensing of Environment*, 114(2): 426-439 [DOI: 10.1016/j.rse.2009.09.012]
- Inglada J and Mercier G. 2007. A new statistical similarity measure for change detection in multitemporal SAR images and its extension to multiscale change analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 45(5): 1432-1445 [DOI: 10.1109/TGRS.2007.893568]
- Isola P, Zhu J Y, Zhou T H and Efros A A. 2017. Image-to-image translation with conditional adversarial networks//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, USA: 5967-5976 [DOI: 10.1109/CVPR.2017.632]
- Jia M, Zhang C, Lyu Z Y, Zhao Z Q and Wang L. 2022. Bipartite adversarial autoencoders with structural self-similarity for unsupervised heterogeneous remote sensing image change detection. *IEEE Geoscience and Remote Sensing Letters*, 19: #6515705 [DOI: 10.1109/LGRS.2022.3201925]
- Khan S H, He X M, Porikli F and Bennamoun M. 2017. Forest change detection in incomplete satellite images with deep neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 55(9): 5407-5423 [DOI: 10.1109/TGRS.2017.2707528]
- LeCun Y A, Bottou L, Orr G B and Müller K R. 2012. Efficient Back-Prop/Montavon G, Orr G B and Müller K R, eds. *Neural Networks: Tricks of the Trade*. 2nd ed. Berlin, Heidelberg: Springer;

- 9-48 [DOI: 10.1007/978-3-642-35289-8\_3]
- Li L, Wang C, Zhang H, Zhang B and Wu F. 2019. Urban building change detection in SAR images using combined differential image and residual U-Net network. *Remote Sensing*, 11 (9) : #1091 [DOI: 10.3390/rs11091091]
- Li X H, Du Z S, Huang Y Y and Tan Z Y. 2021. A deep translation (GAN) based change detection network for optical and SAR remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 179: 14-34 [DOI: 10.1016/j.isprsjprs.2021.07.007]
- Liu J, Gong M G, Qin K and Zhang P Z. 2018a. A deep convolutional coupling network for change detection based on heterogeneous optical and radar images. *IEEE Transactions on Neural Networks and Learning Systems*, 29(3) : 545-559 [DOI: 10.1109/TNNLS.2016.2636227]
- Liu J, Zhang W H, Liu F and Xiao L. 2022. A probabilistic model based on bipartite convolutional neural network for unsupervised change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 60: #4701514 [DOI: 10.1109/TGRS.2021.3071347]
- Liu Z G, Li G, Mercier G, He Y and Pan Q. 2018b. Change detection in heterogenous remote sensing images via homogeneous pixel transformation. *IEEE Transactions on Image Processing*, 27(4) : 1822-1834 [DOI: 10.1109/TIP.2017.2784560]
- Luppino L T, Hansen M A, Kampffmeyer M, Bianchi F M, Moser G, Jenssen R and Anfinen S N. 2024. Code-aligned autoencoders for unsupervised change detection in multimodal remote sensing images. *IEEE Transactions on Neural Networks and Learning Systems*, 35(1) : 60-72 [DOI: 10.1109/TNNLS.2022.3172183]
- Luppino L T, Kampffmeyer M, Bianchi F M, Moser G, Serpico S B, Jenssen R and Anfinen S N. 2022. Deep image translation with an affinity-based change prior for unsupervised multimodal change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 60: #4700422 [DOI: 10.1109/TGRS.2021.3056196]
- Lyu Z Y, Huang H T, Li X H, Zhao M H, Benediktsson J A, Sun W W and Falco N. 2022. Land cover change detection with heterogeneous remote sensing images: review, progress, and perspective. *Proceedings of the IEEE*, 110(12) : 1976-1991 [DOI: 10.1109/JPROC.2022.3219376]
- Mercier G, Moser G and Serpico S B. 2008. Conditional copulas for change detection in heterogeneous remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 46(5) : 1428-1441 [DOI: 10.1109/TGRS.2008.916476]
- Muchoney D M and Haack B N. 1994. Change detection for monitoring forest defoliation. *Photogrammetric Engineering and Remote Sensing*, 60(10) : 1243-1251
- Niu X D, Gong M G, Zhan T and Yang Y L. 2019. A conditional adversarial network for change detection in heterogeneous images. *IEEE Geoscience and Remote Sensing Letters*, 16(1) : 45-49 [DOI: 10.1109/LGRS.2018.2868704]
- Pantze A, Fransson J E S and Santoro M. 2010. Forest change detection from L-band satellite SAR images using iterative histogram matching and thresholding together with data fusion//*IEEE International Geoscience and Remote Sensing Symposium*. Honolulu, USA: IEEE: 1226-1229 [DOI: 10.1109/IGARSS.2010.5650677]
- Radke R J, Andra S, Al-Kofahi O and Roysam B. 2005. Image change detection algorithms: a systematic survey. *IEEE Transactions on Image Processing*, 14(3) : 294-307 [DOI: 10.1109/TIP.2004.838698]
- Rignot E J M and van Zyl J J. 1993. Change detection techniques for ERS-1 SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 31(4) : 896-906 [DOI: 10.1109/36.239913]
- Royer A, Bousmalis K, Gouws S, Bertsch F, Mosseri I, Cole F and Murphy K. 2020. XGAN: unsupervised image-to-image translation for many-to-many mappings//*Singh R, Vatsa M, Patel V M and Ratha N, eds. Domain Adaptation for Visual Understanding*. Switzerland: Springer: 33-49 [DOI: 10.1007/978-3-030-30671-7\_3]
- Sun Y L, Lei L, Li X, Sun H and Kuang G Y. 2021. Nonlocal patch similarity based heterogeneous remote sensing change detection. *Pattern Recognition*, 109: #107598 [DOI: /j.patcog.2020.107598]
- Sun Y L, Lei L, Li X, Tan X and Kuang G Y. 2022. Structure consistency-based graph for unsupervised change detection with homogeneous and heterogeneous remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 60: #4700221 [DOI: 10.1109/TGRS.2021.3053571]
- Taigman Y, Polyak A and Wolf L. 2016. Unsupervised cross-domain image generation. [EB/OL]. [2023-07-20]. <http://arxiv.org/pdf/1611.02200.pdf>
- Wang X L, Girshick R, Gupta A and He K M. 2018. Non-local neural networks//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, USA: IEEE: 7794-7803 [DOI: 10.1109/CVPR.2018.00813]
- Wang Y H, Gao L R, Chen Z C and Zhang B. 2020. Deep learning and superpixel-based method for high-resolution remote sensing image change detection. *Journal of Image and Graphics*, 25(6) : 1271-1282 (王艳恒, 高连如, 陈正超, 张兵. 2020. 结合深度学习和超像元的高分遥感影像变化检测. *中国图象图形学报*, 25(6) : 1271-1282) [DOI: 10.11834/jig.190319]
- Yu H X, Wu A and Zheng W S. 2020. Unsupervised person re-identification by deep asymmetric metric embedding. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(4) : 956-973 [DOI: 10.1109/TPAMI.2018.2886878]

## 作者简介

贾萌,女,副教授,主要研究方向为遥感图像变化检测、高光谱图像分类和机器学习。E-mail: jiameng112@163.com

鲁晓锋,通信作者,男,教授,主要研究方向为计算机视觉、模式识别和智能计算。E-mail: luxiaofeng@xaut.edu.cn

赵秦,男,硕士研究生,主要研究方向为遥感图像变化检测。E-mail: 1952842955@qq.com