WHAT CAUSED CHANGES? CAUSAL PHYSICS DISENTANGLEMENT NETWORK FOR REMOTE SENSING CHANGE DETECTION

Anonymous authors
Paper under double-blind review

ABSTRACT

To address sensitivity to imaging environment variation and limited physical interpretability in traditional remote sensing change detection (RSCD), we propose causal physics disentanglement network combining RS imaging physics with causal inference. Physical parameter decomposition separates intrinsic ground-object information from environmental interference. Moreover, surface reflectance difference serves as core causal evidence for change discrimination, while non-causal backdoor paths induced by illumination and atmospheric variation are explicitly blocked. Based on main RS energy transfer paths (i.e., direct solar irradiance, diffuse sky irradiance, and atmospheric path radiance), spectral radiance model of ground objects is firstly constructed. Secondly, causal inference clarifies relationships between physical variables and RSCD. Specifically, surface reflectance variation serves as core physical evidence of real ground-object change. Meanwhile, total irradiance and atmospheric transmittance are treated as non-causal imaging factors that may induce pseudo-change response. Building on this, bi-temporal physical parameter decomposition and reconstruction network is developed to disentangle input image features into total irradiance, surface reflectance, and atmospheric transmittance, etc. Moreover, orthogonality, reconstruction, and spatial smoothness constraints are jointly imposed to ensure decomposed physical parameters plausibility. Finally, reflectance guided dual-level consistency constraint is introduced. Bi-temporal reflectance difference serves as primary causal pathway for RSCD. Specifically, it guides semantic branch at feature and prediction levels toward real ground-object change. Meanwhile, non-causal backdoor paths from total irradiance and atmospheric transmittance through observed imagery are explicitly blocked. Experimental results demonstrate that proposed model outperforms existing state-of-the-art methods in both qualitative and quantitative metrics on multiple public RSCD datasets, fully verifying its effectiveness and adaptability to complex scenes.

Keywords: Remote Sensing Change Detection; Ground-Object Spectral Radiance Model; Causal Inference; Causal Physical Decoupling.

1 INTRODUCTION

RSCD aims to identify and localize changed regions by comparing images acquired over the same geographic area at different times. It has been widely applied to urban expansion analysis, disaster assessment, military reconnaissance, vegetation monitoring, and other fields Zang et al. (2025b); Wang et al. (2025b); Zhang et al. (2026b).

Currently, different illumination conditions pose significant challenges to RSCD. Since bi-temporal images are typically acquired at different times or even in different seasons, their imaging conditions are more complex and less controllable than those in natural CD scenes. To tackle illumination disturbance, the existing CD methods can be broadly divided into two categories: feature extraction Wang et al. (2025a); Yu et al. (2026); Wang et al. (2026a) and domain adaptation Ngo et al. (2024); Du et al. (2025); Zhao et al. (2026). CD methods based on feature extraction can broadly be categorized into two approaches: both traditional and deep learningXu et al. (2025). Traditional CD methods can be broadly classified into three categories, e.g., image arithmetic, image transformation, and post-classification, etc. However, these methods tend to be less generalization and robustness. Owing to powerful feature extraction and non-linear representation, deep learning methods have been extensively applied to CD tasks and have demonstrated excellent performanceKim & Kim

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

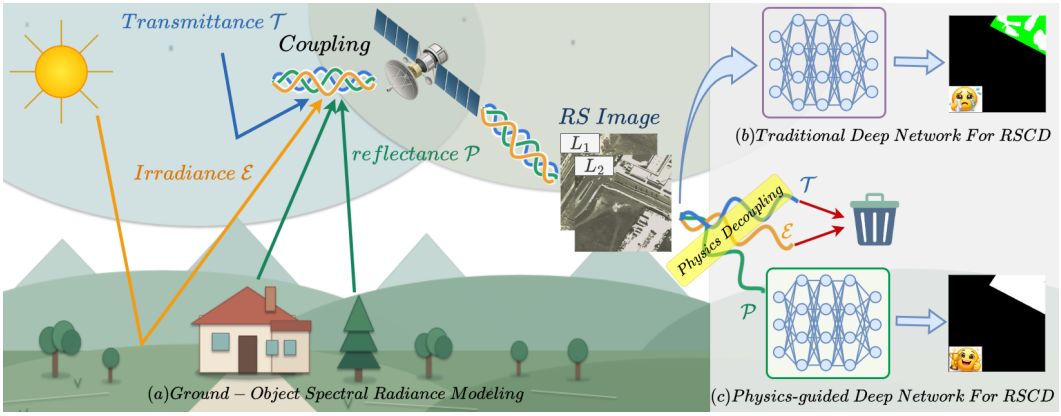


Figure 1: Two paradigms of current deep models for RSCD. (a) Ground-object spectral radiance modeling. (b) Conventional deep networks: change discrimination is directly performed on the mixed features of dual-temporal observed images. Explicit modeling of RS imaging physical priors is absent, and the networks are susceptible to interference from non-causal factors such as variations in illumination and atmospheric conditions. (c) The proposed causal physics disentanglement network: it incorporates surface radiative transfer mechanism into change interpretation process and disentangles physical parameters, including illumination, atmospheric transmittance, and surface reflectance, to suppress the influence of non-causal factors on change prediction. More reliable RSCD is achieved using reflectance differences.

(2025). Representative deep neural network frameworks for RSCD including Convolutional neural network (CNN)Zang et al. (2025a), TransformerZhang et al. (2026a), MambaParanjape et al. (2025), Denoising diffusion probabilistic model (DDPM)Chaminda Bandara et al. (2025) and their hybrid variants. However, feature extraction-based deep CD networks often exhibit limited generalization in cross-domain scenarios. Although domain adaptation methods improve model generalization to some extent, their performance still degrades substantially under complex domain shifts.

However, current mainstream RSCD methods are mostly purely data-driven paradigm, which generally fails to explore and embed RS physical nature. They directly conduct semantic feature learning on raw image, making it difficult to decouple ground-object intrinsic attribute from environmental interference, and it is prone to being disturbed by redundant information in complex scenes.

As shown in Figure 1, inspired by RS physical and causal inference, this paper proposes causal physics disentanglement network for RSCD (i.e.,CPDcd). First, we model ground-object spectral radiance based on three major energy transfer paths (i.e., direct irradiance, diffuse irradiance, and atmospheric path radiance, etc). Then, the causal relationships between physical parameters and RSCD are analyzed from the perspective of causal inference. Meanwhile, decomposition and reconstruction network separates bi-temporal imagery into total irradiance, atmospheric transmittance, and land surface reflectance, etc. Corresponding constraints encourage physically plausible parameter estimates and preserve information needed for image reconstruction. Finally, we introduce feature-level and prediction-level constraints to promote reflectance difference as primary evidence for change, while reducing spurious influence from irradiance and atmospheric transmittance.

In summary, main contributions are summarized as follows:

1) The causal physics disentanglement framework for RSCD are proposed. Based on RS radiative transfer, ground-object spectral radiance model are constructed using total irradiance, atmospheric transmittance, and ground-object reflectance as its main variables. Furthermore, causal analysis treats bi-temporal reflectance difference as primary path for representing genuine ground-object change, thereby improving model physical interpretability.

2) Based on ground-object spectral radiance model, the dedicated physical decomposition and reconstruction network is designed to achieve decoupled representation of physical quantities (i.e., total irradiance, atmospheric transmittance, and surface reflectance) from RS images. Furthermore, orthogonality, reconstruction, and smoothness constraints are utilized to improve parameter separability and stability.

3) Causal intervention strategy is introduced with prediction consistency loss to suppress spurious paths from total irradiance or atmospheric transmittance to change prediction. Additionally, feature alignment loss encourages semantic and change-response consistency between bi-temporal deep feature and reflectance-difference feature, which strengthens reflectance difference as the primary causal evidence for RSCD.

2 Related Work

2.1 Physics-Guided RSCD

Physics-guided RSCD incorporates prior knowledge such as RS physical laws, physical properties of ground object, and radiative transfer models into deep learning-based CD, rather than relying solely on data-driven feature learning. By constraining the network to fit real physical process, it improves CD accuracy and interpretability in complex scenes. Common forms of physics guidance for RSCD can be divided into two main categories, imaging law and physical principle.

Imaging law–based RSCD incorporates physical imaging laws as prior knowledge into the multi-temporal RSCD pipeline, including sensor imaging mechanisms(Sumbul et al., 2025), radiative transfer processes(Du et al., 2026), atmospheric effects(Chen et al., 2025), and others. By modeling and removing spurious differences caused via non-real ground changes such as imaging environment variations and atmospheric attenuation, it highlights actual changes of ground object itself, thereby improving CD results accuracy and physical consistency.

Physics law–based RSCD refers to embedding objective physical laws as prior constraints into CD model, including radiative transfer(Huang et al., 2025), spectral physical properties(Zang et al., 2025a), geometrical optics(Aira et al., 2025), and spatiotemporal evolution laws of ground objects(Do et al., 2025). By modeling physically consistent relationships between multi-temporal images, it suppresses spurious differences caused by non-real land change such as illumination, atmosphere, and sensor effects, ensuring that detection results conform to physical rationality.

2.2 Causal Inference

Deep learning-based causal inference models causal relationships among variables to distinguish factors directly relevant to target task from spurious factors arising from data bias, background, or acquisition condition, etc. Furthermore, causal intervention reduces reliance on spurious correlation during prediction to improve robustness, generalization, and interpretability. Common approaches fall into two categories: structural causal modeling (SCM) and causal representation learning.

SCM-based methods use causal graph(Song et al., 2026) to describe relationships among input, confounder, and prediction. Then, they apply backdoor adjustment(Liao et al., 2026), causal intervention(Xu et al., 2026), and counterfactual reasoning(Vashishtha et al., 2026) to block non-causal paths. In contrast, causal representation learning separates stable causal information from environment-dependent information in observed data through feature disentanglement(Liu et al., 2026), domain-invariant learning(Zhang et al., 2026c), and causal regularization(Li et al., 2026). To date, these methods have been applied to image classification(Wu et al., 2026), object detection(Hong et al., 2026), and semantic segmentation(Liang et al., 2026), etc.

3 Method

RS physics suggests that bi-temporal difference in land surface reflectance can effectively indicate genuine land-cover change. Motivated by this insight, we integrate causal analysis with imaging physics and propose the reflectance-based network for RSCD. Specifically, it decomposes observed imagery into total irradiance, atmospheric transmittance, and land surface reflectance, etc. Furthermore, it uses reflectance difference to identify changed region while reducing interference from illumination and atmospheric variation. As shown in Figure 2, the framework comprises three components: ground-object spectral radiance modeling, causal analysis of physical decomposition, and reflectance-guided RSCD.

3.1 Spectral Radiance Modeling of Ground Object

In RS detection of ground objects, energy received by RS system mainly originates from reflection of solar radiation by ground objects. Solar radiation at the top of atmosphere is reflected by ground objects during atmospheric transmission and finally enters RS system. Figure 2(a) illustrates the main radiative energy transfer paths in this process. Specifically, Path A represents direct solar radiation transmitted through atmosphere, reflected by ground objects, and passing through atmosphere again, serving as main energy contribution to RS system. Path B refers to radiation scattered by atmosphere, reflected by ground objects, and then entering RS system. Path C denotes radiation directly entering RS system after atmospheric scattering. Path D corresponds to energy entering RS system after multiple reflections by ground objects and subsequent atmospheric transmission. Path E represents energy from background reflection of adjacent areas after atmospheric scattering. To simplify model, we assume energy transmission path consists mainly of A, B and C.

According to above imaging analysis, the RS system radiance can be expressed as, where $E_t(\lambda)$ denotes irradiance produced on surface by direct solar radiation, and $E_d(\lambda)$ represents irradiance

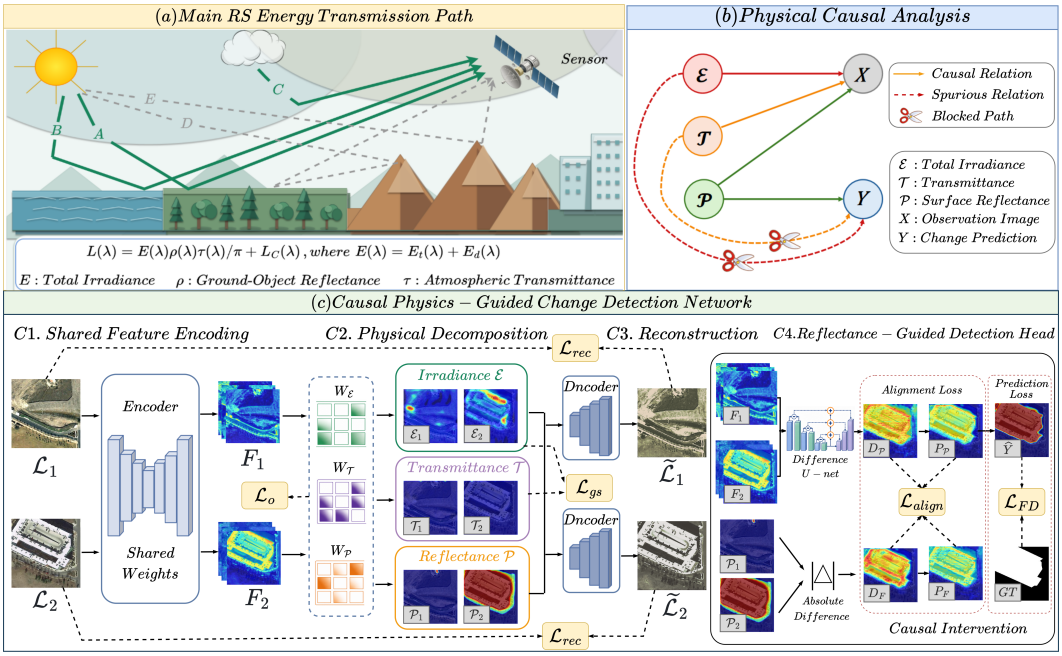


Figure 2: Overall framework of proposed causal physics-disentangled RSCD network. (a) Surface spectral radiation modeling. Based on this, RS images are decomposed into three core physical parameter features: irradiance $\boldsymbol{\mathcal{E}}$, ground-object reflectance $\boldsymbol{\mathcal{P}}$, and atmospheric transmittance $\boldsymbol{\mathcal{T}}$. (b) Physical causal analysis and intervention. The causal relationships among irradiance $\boldsymbol{\mathcal{E}}$, atmospheric transmittance $\boldsymbol{\mathcal{T}}$, surface reflectance $\boldsymbol{\mathcal{P}}$, observed image $\boldsymbol{X}$, and change result $\boldsymbol{Y}$ are modeled. Subsequently, the genuine causal path of $\boldsymbol{\mathcal{P}} \rightarrow \boldsymbol{Y}$ and the non-causal backdoor paths of $\boldsymbol{Y} \leftarrow \boldsymbol{\mathcal{T}}/\boldsymbol{\mathcal{E}} \rightarrow \boldsymbol{X}$ are identified. (c) Physics-guided parameter decomposition and RSCD. It consists of the shared encoder, decomposition and reconstruction of physical features, and CD results generated based on surface reflectance difference features.

generated on ground surface by atmospheric scattered radiation. They constitute surface irradiance $E(\lambda) = E_t(\lambda) + E_d(\lambda)$. Additionally, $\rho(\lambda)$ is spectral reflectance of ground object. As intrinsic property of ground object, it depends only on material composition and is not affected by disturbance such as illumination and atmosphere. $\tau(\lambda)$ is atmospheric transmittance, which quantifies attenuation effect of atmospheric medium on solar radiation. $L_C(\lambda)$ is atmospheric path radiance.

$$L(\lambda) = \left(\frac{E_t(\lambda)}{\pi} + \frac{E_d(\lambda)}{\pi}\right)\rho(\lambda)\tau(\lambda) + L_C(\lambda) \tag{1}$$

Let working wavelength band be $[\lambda_{\min}, \lambda_{\max}]$, and final spectral radiance is given by Eq.(2):

$$\int_{\lambda_{\min}}^{\lambda_{\max}} L(\lambda)d\lambda = \frac{1}{\pi}\int_{\lambda_{\min}}^{\lambda_{\max}} E_t(\lambda)\rho(\lambda)\tau(\lambda)d\lambda + \frac{1}{\pi}\int_{\lambda_{\min}}^{\lambda_{\max}} E_d(\lambda)\rho(\lambda)\tau(\lambda)d\lambda + \int_{\lambda_{\min}}^{\lambda_{\max}} L_C(\lambda)d\lambda \tag{2}$$

Let total surface irradiance be $E(\lambda) = E_t(\lambda) + E_d(\lambda)$, then transmission model is given by:

$$\int_{\lambda_{\min}}^{\lambda_{\max}} L(\lambda)d\lambda = \frac{1}{\pi}\int_{\lambda_{\min}}^{\lambda_{\max}} E(\lambda)\rho(\lambda)\tau(\lambda)d\lambda + \int_{\lambda_{\min}}^{\lambda_{\max}} L_C(\lambda)d\lambda \tag{3}$$

The final RS model is established as follows, where $\odot$ denotes element-wise product. $\boldsymbol{\mathcal{E}}$, $\boldsymbol{\mathcal{P}}$ and $\boldsymbol{\mathcal{T}}$ represent total surface irradiance, ground-object reflectance, and atmospheric transmittance, respectively. Notably, the atmospheric path radiance $\mathcal{L}_C$ is omitted in this study.

$$\mathcal{L} = \frac{\boldsymbol{\mathcal{E}}}{\pi} \odot \boldsymbol{\mathcal{P}} \odot \mathcal{T} + \mathcal{L}_C \tag{4}$$

3.2 Physics-Based Causal Disentanglement and Problem Formulation

This section analyzes how each physical variable relates to RSCD from causal perspective. First, we construct SCM for RSCD as shown in Figure 2(b). Here, $\boldsymbol{X}$ denotes input RS image, and $\boldsymbol{Y}$ denotes change prediction, respectively. Relationships among these variables are explained below.

Table 1: Performance comparison of different methods across three datasets

Methods	LEVIR-CD					SECOND					SYSU				
	P.	Re.	F1	OA	mIoU	P.	Re.	F1	OA	mIoU	P.	Re.	F1	OA	mIoU
BIT (Chen et al., 2022)	92.16	70.68	80.01	98.20	82.40	69.42	64.37	66.80	**88.72**	68.71	80.52	72.55	76.16	89.33	74.32
Change Former (Bandara & Patel, 2022)	91.17	78.57	84.40	98.52	85.74	74.61	50.58	66.79	87.95	68.21	80.83	73.68	77.09	89.67	75.11
ICIFNet (Feng et al., 2022)	76.65	81.25	78.88	98.30	81.69	**78.21**	53.96	63.86	87.79	66.61	77.37	76.35	76.86	89.16	74.59
USSFC-Net (Lei et al., 2023)	**94.85**	73.98	83.13	98.47	84.77	73.15	54.43	62.48	86.90	65.37	81.66	68.24	74.35	88.90	72.97
DMINet (Feng et al., 2023)	79.23	77.61	78.41	98.33	81.39	74.18	58.50	65.41	87.63	67.30	**86.07**	71.13	77.89	90.48	76.17
ChangeMamba (Chen et al., 2024)	94.16	75.81	84.01	98.53	85.44	69.29	63.19	66.10	87.04	67.26	83.23	76.63	79.80	90.85	66.39
STRobustNet (Zhang et al., 2025)	77.06	85.01	80.84	98.43	83.11	69.75	63.40	66.42	88.71	68.51	82.41	74.42	78.21	90.22	76.18
FCdDNet (Wang et al., 2026b)	92.05	87.13	89.52	98.96	89.97	68.42	**69.01**	68.72	87.33	68.88	74.96	77.66	76.28	88.61	73.86
Ours	91.32	**89.36**	**90.33**	**99.03**	**90.67**	71.37	67.49	**69.38**	88.08	**69.67**	82.17	**78.15**	**80.11**	**90.85**	**77.80**

All values are reported in percentage (%). Color convention: **best**, 2-nd best.

$\boldsymbol{\mathcal{E}} \rightarrow \boldsymbol{X}$: Total irradiance $\boldsymbol{\mathcal{E}}$ affects the observed image $\boldsymbol{X}$ radiometric properties through incident radiation. However, it does not directly determine ground-object change.
$\boldsymbol{\mathcal{T}} \rightarrow \boldsymbol{X}$: Atmospheric transmittance $\boldsymbol{\mathcal{T}}$ affects RS image $\boldsymbol{X}$ by altering its radiance as radiation travels from ground object to sensor. However, it does not cause actual ground-object change.
$\boldsymbol{\mathcal{P}} \rightarrow \boldsymbol{X}$: Ground-object reflectance $\boldsymbol{\mathcal{P}}$ characterizes how surface responds to incident radiation and plays major role in determining its appearance in $\boldsymbol{X}$. Furthermore, bi-temporal reflectance difference provides evidence for identifying genuine ground-object change.
$\boldsymbol{\mathcal{P}} \rightarrow \boldsymbol{Y}$: The difference of bi-temporal ground-object reflectance $\boldsymbol{\mathcal{P}}$ is utilized to generate $\boldsymbol{Y}$. Consequently, prediction relies more on ground-object change and less on in other imaging conditions.
$\boldsymbol{X} \rightarrow \boldsymbol{Y}$: Conventional RSCD methods predict $\boldsymbol{Y}$ directly from $\boldsymbol{X}$. However, $\boldsymbol{X}$ contains both ground-object information and variation caused by irradiance and atmospheric conditions. Consequently, model may mistake such variation for genuine change.
$\boldsymbol{Y} \leftarrow \boldsymbol{\mathcal{T}}/\boldsymbol{\mathcal{E}} \rightarrow \boldsymbol{X}$: Dashed lines indicate potential prediction shortcut. Irradiance $\boldsymbol{\mathcal{E}}$ and atmospheric transmittance $\boldsymbol{\mathcal{T}}$ alter observed image $\boldsymbol{X}$ appearance, which may lead model to rely on these factors when predicting change. Consequently, if model interprets these appearance differences as genuine change, it may incorrectly label unchanged regions as changed.
Therefore, effective RSCD requires distinguishing appearance variation in observed imagery from genuine ground-object change. To this end, our physical decomposition suppresses non-causal influence from irradiance and atmospheric transmittance. Meanwhile, bi-temporal ground-object reflectance difference serves as the primary evidence for change prediction.
Therefore, RSCD based on physical decoupling can be formulated as:

$$P(\widehat{\boldsymbol{Y}} \mid \boldsymbol{D}_{\boldsymbol{\mathcal{P}}}, \mathrm{do}(\boldsymbol{D}_{\boldsymbol{\mathcal{E}}}, \boldsymbol{D}_{\boldsymbol{\mathcal{T}}})) \approx P(\widehat{\boldsymbol{Y}} \mid \boldsymbol{D}_{\boldsymbol{\mathcal{P}}}). \tag{5}$$

This formulation effectively blocks backdoor paths induced by total irradiance and atmospheric transmittance (i.e., $\boldsymbol{\mathcal{T}}/\boldsymbol{\mathcal{E}} \rightarrow \boldsymbol{Y}$).

3.3 Physical Parameter Decomposition and Reconstruction

RS formation analysis suggests that total irradiance and atmospheric transmittance vary little within local area and have no direct relationship with land-cover class. Thus, they can introduce interference into RSCD, whereas surface reflectance varies across land-cover classes. Motivated by this distinction, we design network to decompose and reconstruct irradiance, reflectance, and atmospheric transmittance according Eq.(4). Then, reflectance difference is utilized to identify changed regions. Given bi-temporal RS image $\boldsymbol{\mathcal{L}}_i \in \mathbb{R}^{H \times W \times C}$, a shared-weight feature encoder $f(\cdot)$ is used to extract its deep feature $\boldsymbol{F}_i$:

$$\boldsymbol{F}_i = f(\boldsymbol{\mathcal{L}}_i; \theta), \quad i \in \{1, 2\} \tag{6}$$

Here, θ denotes encoder parameters. Subsequently, the features of total irradiance $\boldsymbol{\mathcal{E}}_i$, atmospheric transmittance $\boldsymbol{\mathcal{T}}_i$, and land surface reflectance $\boldsymbol{\mathcal{P}}_i$ are obtained based on deep feature $\boldsymbol{F}_i$, which can be formalized as follows:

$$\boldsymbol{\mathcal{E}}_i = \mathit{ReLU}\big[\boldsymbol{F}_i \cdot \boldsymbol{W}_{\boldsymbol{\mathcal{E}}}\big], \tag{7}$$

$$\boldsymbol{\mathcal{T}}_i = \mathit{Sigmoid}\big[\boldsymbol{F}_i \cdot \boldsymbol{W}_{\boldsymbol{\mathcal{T}}}\big], \tag{8}$$

$$\boldsymbol{\mathcal{P}}_i = \mathit{Sigmoid}\big[\boldsymbol{F}_i \cdot \boldsymbol{W}_{\boldsymbol{\mathcal{P}}}\big], \tag{9}$$

Here, $\boldsymbol{W}_{\boldsymbol{\mathcal{E}}}$, $\boldsymbol{W}_{\boldsymbol{\mathcal{T}}}$ and $\boldsymbol{W}_{\boldsymbol{\mathcal{P}}} \in \mathbb{R}^{d \times m}$ denote projection matrices for $\boldsymbol{\mathcal{E}}_i$, $\boldsymbol{\mathcal{T}}_i$ and $\boldsymbol{\mathcal{P}}_i$, respectively. They map deep semantic feature F_i into irradiance, atmospheric transmittance, and reflectance spaces, respectively. For each physical quantity, dual-temporal features F_i share the same mapping network, ensuring that their estimated physical variables lie in common representation space.

Since irradiance $\boldsymbol{\mathcal{E}}_i \geq 0$, *ReLU* activation function is adopted to ensure non-negativity. For atmospheric transmittance and ground object reflectance $(\boldsymbol{\mathcal{T}}_i, \boldsymbol{\mathcal{P}}_i) \in [0, 1]$, Sigmoid activation is used to constrain outputs within range $[0, 1]$, ensuring consistency with physical meaning.
Because physical variables $\boldsymbol{\mathcal{E}}$, $\boldsymbol{\mathcal{T}}$, and $\boldsymbol{\mathcal{P}}$ have no direct relationship, we impose an orthogonality constraint $\mathcal{L}_o$ to reduce their branches reliance on the same input feature $\boldsymbol{F}_i$.

$$\mathcal{L}_o = \overline{\boldsymbol{W}}_{\boldsymbol{\mathcal{E}}}^T \overline{\boldsymbol{W}}_{\boldsymbol{\mathcal{T}}} + \overline{\boldsymbol{W}}_{\boldsymbol{\mathcal{T}}}^T \overline{\boldsymbol{W}}_{\boldsymbol{\mathcal{P}}} + \overline{\boldsymbol{W}}_{\boldsymbol{\mathcal{P}}}^T \overline{\boldsymbol{W}}_{\boldsymbol{\mathcal{E}}} + \sum_{\boldsymbol{W} \in \{\overline{\boldsymbol{W}}_{\boldsymbol{\mathcal{E}}}, \overline{\boldsymbol{W}}_{\boldsymbol{\mathcal{T}}}, \overline{\boldsymbol{W}}_{\boldsymbol{\mathcal{P}}}\}} \left(\sum_{k=1}^{d} \boldsymbol{W}[k] - 1 \right)^2 \tag{10}$$

Where average contribution of the $k-th$ input feature to total irradiance is defined as $\overline{\boldsymbol{W}}_{\boldsymbol{\mathcal{E}}}[k] = \frac{1}{m} \sum_{j=1}^{m} |\boldsymbol{W}_{\boldsymbol{\mathcal{E}}}[k, j]|$. Similarly, we obtain other contribution vectors $\overline{\boldsymbol{W}}_{\boldsymbol{\mathcal{T}}}$ and $\overline{\boldsymbol{W}}_{\boldsymbol{\mathcal{P}}}$ for atmospheric transmittance and land surface reflectance. Additionally, a normalization constraint encourages the elements of each contribution vector to sum to 1.
Based on RS radiative transfer relationship, the $i-th$ temporal reconstruct image $\widetilde{\mathcal{L}}_i$ is acquired by decomposed physical parameters as follows:

$$\widetilde{\mathcal{L}}_i = \frac{\boldsymbol{\mathcal{E}}_i}{\pi} \odot \boldsymbol{\mathcal{P}}_i \odot \boldsymbol{\mathcal{T}}_i \tag{11}$$

Reconstruction loss $\mathcal{L}_{\text{rec}}$ is applied to ensure the completeness of physical decomposition network.

$$\mathcal{L}_{\text{rec}} = \sum_{i=1}^{2} \mathbb{E}_{\boldsymbol{\mathcal{L}}_i \sim p(\boldsymbol{\mathcal{L}}_i)} \left(\left\| \widetilde{\boldsymbol{\mathcal{L}}}_i - \boldsymbol{\mathcal{L}}_i \right\|_2^2 \right) \tag{12}$$

Since the irradiance $\boldsymbol{\mathcal{E}}$ and atmospheric transmittance $\boldsymbol{\mathcal{T}}$ are nearly identical in local space, we impose gradient smoothing constraints $\mathcal{L}_{\text{gs}}$ on them.

$$\begin{aligned} \mathcal{L}_{gs} = \frac{1}{|\Omega|} \Bigg[& \sum_{(p,q) \in \Omega} \left(|\boldsymbol{\mathcal{E}}(p+1, q) - \boldsymbol{\mathcal{E}}(p, q)| + |\boldsymbol{\mathcal{E}}(p, q+1) - \boldsymbol{\mathcal{E}}(p, q)| \right) + \\ & \sum_{(p,q) \in \Omega} \left(|\boldsymbol{\mathcal{T}}(p+1, q) - \boldsymbol{\mathcal{T}}(p, q)| + |\boldsymbol{\mathcal{T}}(p, q+1) - \boldsymbol{\mathcal{T}}(p, q)| \right) \Bigg]. \end{aligned} \tag{13}$$

Where Ω represents feature spatial domain.

3.4 Causal Intervention on Physical Parameters

Genuine land-cover change is primarily reflected in bi-temporal reflectance difference, whereas illumination and atmospheric variation can produce apparent change. Therefore, we use causal intervention to suppress spurious pathways through which total irradiance and atmospheric transmittance influence change prediction.
First, bi-temporal features $\boldsymbol{F}_i$ are fed into differential U-Net $f_{\text{D-Unet}}(\cdot)$ to extract semantic change features $\boldsymbol{D}_F$:

$$\boldsymbol{D}_F = f_{\text{D-Unet}}(\boldsymbol{F}_1, \boldsymbol{F}_2; \theta_F) \tag{14}$$

Meanwhile, we compute difference features for decomposed physical parameters, i.e., $\boldsymbol{D}_{\boldsymbol{\mathcal{P}}} = |\boldsymbol{\mathcal{P}}_2 - \boldsymbol{\mathcal{P}}_1|$, $\boldsymbol{D}_{\boldsymbol{\mathcal{E}}} = |\boldsymbol{\mathcal{E}}_2 - \boldsymbol{\mathcal{E}}_1|$ and $\boldsymbol{D}_{\boldsymbol{\mathcal{T}}} = |\boldsymbol{\mathcal{T}}_2 - \boldsymbol{\mathcal{T}}_1|$. Each feature is then fed into CD head to obtain corresponding change response $\boldsymbol{P}_F$, $\boldsymbol{P}_{\boldsymbol{\mathcal{P}}}$, $\boldsymbol{P}_{\boldsymbol{\mathcal{E}}}$ and $\boldsymbol{P}_{\boldsymbol{\mathcal{T}}}$.
To address class-imbalance in RSCD, we combine Focal Loss and Dice Loss to form primary classification prediction loss $\mathcal{L}_{FD}$:

$$\mathcal{L}_{FD}(\widehat{\boldsymbol{Y}}, \boldsymbol{Y}) = \mathcal{L}_{Focal}(\widehat{\boldsymbol{Y}}, \boldsymbol{Y}) + \mathcal{L}_{Dice}(\widehat{\boldsymbol{Y}}, \boldsymbol{Y}) \tag{15}$$

Based on causal analysis, change prediction should primarily rely on ground-object reflectance difference rather than irradiance or atmospheric transmittance variation. Accordingly, we define prediction consistency loss $\mathcal{L}_{pc}$ as follows:

$$\mathcal{L}_{pc} = \mathcal{L}_{FD}(\boldsymbol{P}_{\boldsymbol{\mathcal{P}}}, \boldsymbol{Y}) - \mathcal{L}_{FD}(\boldsymbol{P}_{\boldsymbol{\mathcal{E}}}, \boldsymbol{Y}) - \mathcal{L}_{FD}(\boldsymbol{P}_{\boldsymbol{\mathcal{T}}}, \boldsymbol{Y}). \tag{16}$$

This loss suppresses spurious influence of $\boldsymbol{D}_{\boldsymbol{\mathcal{E}}}$ and $\boldsymbol{D}_{\boldsymbol{\mathcal{T}}}$ on change prediction.

Table 2: Ablation study of loss weights

λ_1	λ_2	λ_3	λ_4	P.	Re.	F1	OA	mIoU
1	1	1	1	92.38	87.94	90.10	99.02	90.48
0.5	1	1	1	93.04	87.58	90.23	99.03	90.59
0.1	1	1	1	93.63	84.45	88.80	98.92	89.36
0.5	0.5	1	1	91.85	88.85	90.32	99.03	90.67
0.5	0.1	1	1	93.05	87.36	90.12	99.02	90.50
0.5	0.5	0.5	1	92.47	88.25	90.31	**99.04**	90.66
0.5	0.5	2	1	92.01	87.13	89.50	98.96	89.96
0.5	0.5	1	0.5	91.32	**89.36**	**90.33**	99.03	**90.67**
0.5	0.5	1	0.1	**93.86**	85.50	89.49	98.98	89.95

Furthermore, deep difference features $\boldsymbol{D}_F$ should align with reflectance $\boldsymbol{D}_{\boldsymbol{\mathcal{P}}}$ in both semantic content and change response. Accordingly, we define the following feature alignment loss $\mathcal{L}_{\text{align}}$.

$$\mathcal{L}_{align} = \sum_{i=1}^{N} \mathbb{E}_{\boldsymbol{D}_F \sim p(\boldsymbol{D}_F)} \left\| \boldsymbol{D}_{\boldsymbol{\mathcal{P}}}^{(i)} - \boldsymbol{D}_F^{(i)} \right\|_1 + \mathcal{L}_{FD}(\boldsymbol{P}_{\boldsymbol{\mathcal{P}}}, \boldsymbol{P}_F) \tag{17}$$

Here, $N \in \mathbb{R}^{H/4 \times W/4}$ denotes the number of pixels in single feature map. Under these constraints, reflectance change guides semantic branch at both feature and prediction levels, thereby reducing apparent change caused by irradiance and atmospheric transmittance variation. The constraints also reduce shared dependence on input features $\boldsymbol{F}_i$ across $\boldsymbol{\mathcal{E}}$, $\boldsymbol{\mathcal{T}}$, and $\boldsymbol{\mathcal{P}}$ branches and discourage $\boldsymbol{D}_{\boldsymbol{\mathcal{E}}}$ and $\boldsymbol{D}_{\boldsymbol{\mathcal{T}}}$ from serving as change cues.
The overall loss function is defined as follows:

$$\mathcal{L}_{total} = \mathcal{L}_{pc} + \lambda_1 \mathcal{L}_o + \lambda_2 \mathcal{L}_{rec} + \lambda_3 \mathcal{L}_{align} + \lambda_4 \mathcal{L}_{gs} \tag{18}$$

4 Experimental Results

4.1 Experimental Setup

The CPDcd model is implemented based on PyTorch framework and trained and tested on a single NVIDIA RTX 4090 GPU. The model is optimized using SGD optimizer with initial learning rate of 0.01, while a dynamic adjustment strategy is adopted for learning rate, decaying to 70% of its original value every 20 training epochs. In addition, total number of training epochs is set to 150 for all datasets with batch size of 16 and validation is performed after each training epoch, while model with optimal performance on validation set is finally used for test set evaluation.

4.2 Performental Comparison

4.2.1 Objective Indicator

To verify effectiveness of CPDcd, comparison is conducted with several state-of-the-art RSCD method, including BIT, ChangeFormer, ICIFNet, USSFC-Net, DMINet, ChangeMamba, STRobustNet and FCdDNet. For fairness, all comparative method is implemented using code released by original author, and respective parameter is set according to original paper. Table 1 compares CPDcd with representative RSCD methods on three datasets. CPDcd achieves competitive F1 and mIoU performance, demonstrating the effectiveness of physics-guided decomposition in suppressing environmental interference.

4.2.2 Visual Comparison

As shown in Figure 3, competing methods suffer from missed detections on small targets and boundary errors on large changed regions. CPDcd produces cleaner change maps with fewer false alarms and missed detections.

4.3 Ablation Experiments

To verify effectiveness of each loss and obtain optimal parameter configuration, loss function regularization is experimentally tuned on LEVIR-CD dataset. As shown in Table 2, optimal parameter combination is $\boldsymbol{\lambda} = [1, 0.5, 0.5, 1, 0.5]$. Subsequently, this configuration is applied to other two datasets, which still yields favorable performance metric. This indicates that model has good generalization ability and is not overly sensitive to regularization parameter.

4.4 Analysis of Physical Parameters

To verify rationality of network decomposition variable, total irradiance $\boldsymbol{\mathcal{E}}$, atmospheric transmittance $\boldsymbol{\mathcal{T}}$, and ground object reflectance $\boldsymbol{\mathcal{P}}$ are analyzed.

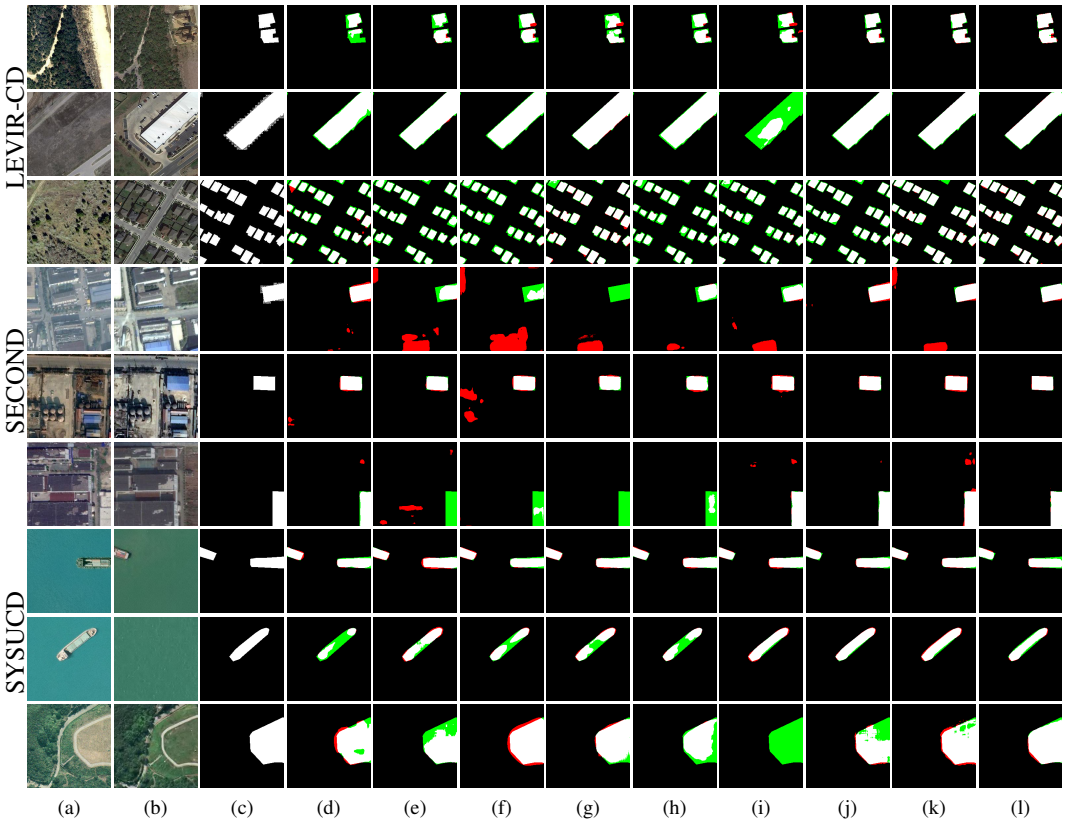


Figure 3: Qualitative results of different CD methods on LEVIR-CD, SECOND and SYSUCD datasets. Columns: (a) $\mathcal{L}_1$ Image; (b) $\mathcal{L}_2$ Image; (c) Ground Truth; (d) BIT; (e) ChangeFormer; (f) ICIFNet; (g) USSFC-Net; (h) DMINet; (i) ChangeMamba; (j) STRobustNet; (k) FCdDNet; (l) Ours.

For limited interval within single image, range of irradiance and atmospheric transmittance is relatively uniform. To this end, pixel histogram of $\mathcal{E}$ and $\mathcal{T}$ is calculate statistically, as shown in Figure 4. It can be observed that frequency of $\mathcal{E}$ concentrated in interval $[0.02, 0.08]$ reaches 98.69%, and frequency of $\mathcal{T}$ concentrated in $[0.2, 0.33]$ reaches 85.10%. This is consistent with physical law in real-world scenario, verifying physical rationality of network decoupling result.

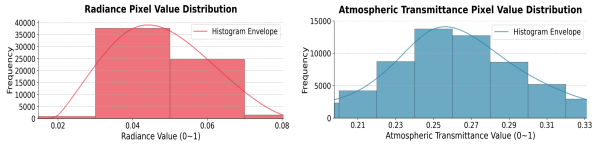


Figure 4: Pixel distribution of $\mathcal{E}$ and $\mathcal{T}$, showing consistent physical-variable distributions.

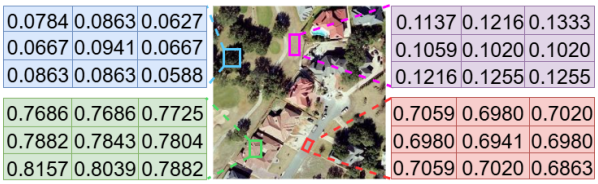


Figure 5: Local numerical visualization of $\mathcal{P}$, showing high intra-class consistency and inter-class divergence.

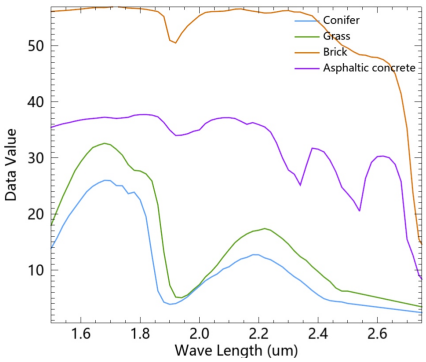


Figure 6: Real spectral response curves of four land-cover types from the ENVI spectral library.

For land surface reflectance $\mathcal{P}$, similar objects exhibit close value, whereas distinct objects show noticeable differences. As shown in Figure 5, land surface reflectance of different land cover types is extracted from decomposed variable $\mathcal{P}$, from which it can be observed that similar objects present consistent and close reflectance values. As references, the real spectral response curves of typical ground objects are extracted from ENVI, a standard RS image processing platform for global RS

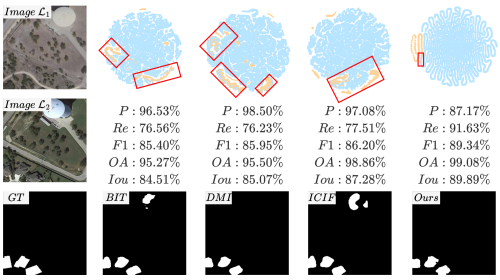


Figure 7: Visual comparison of pixel-wise classification for different models. (a) Bi-temporal RS images and GT. (b)-(e) Results of BIT, DMI, ICIF, and CPDcd, including feature visualization, quantitative metrics, and predictions.

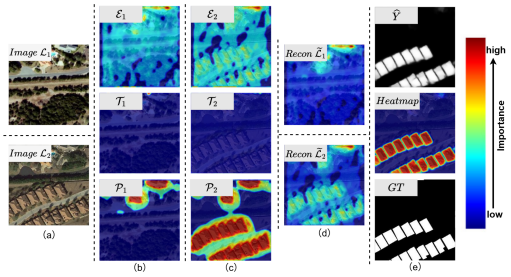


Figure 8: Visualization of key CPDcd modules. (a) Bi-temporal images $\mathcal{L}_1$ and $\mathcal{L}_2$. (b)-(c) Decomposed physical variables $\boldsymbol{\mathcal{E}}$, $\boldsymbol{\mathcal{P}}$, and $\boldsymbol{\mathcal{T}}$ from two phases images. (d) Reconstructed images. (e) Change map, GT and heat map.

and geospatial analysis. As shown in Figure 6, overall reflectance ranking of different objects is: trees $<$ grass $<$ road $<$ building, which is consistent with results presented in Figure 5.
The above analysis indicates that physical quantities derived from physics-disentangled network are consistent with prior knowledge, providing reliable CD physical basis.

4.5 Visualization

4.5.1 Visualization for Uncouple Space

To intuitively demonstrate CD performance, t-SNE dimension reduction technique is adopted to project changed and unchanged features onto 2D plane for visualization analysis. As shown in Figure 7, yellow and blue represent change and unchanged features, respectively. From visualization results, semantic boundary clarity of proposed model is significantly superior to that of others. In addition, in terms of quantitative metrics, proposed model achieves excellent performance across multiple indicators, which is consistent with visualization results. Thus, it is evident that CD model integrated with land surface reflectance can more accurately distinguish between changed and unchanged pixels, thereby effectively enhancing semantic boundary discrimination.

4.5.2 Visualization for Feature Heatmap

To facilitate more intuitive understanding of CPDcd model, representative sample from LEVIR-CD dataset is selected to visualize feature map at different stage. As observed from Figure 8(b) and (c), bi-temporal image is decomposed into physical parameter (i.e., $\boldsymbol{\mathcal{E}}$, $\boldsymbol{\mathcal{P}}$, and $\boldsymbol{\mathcal{T}}$). The visualization of different physical parameter clearly reveals difference in ground object. Furthermore, reconstructed image in Figure 8(d) is highly consistent with original image, verifying effectiveness of physical parameter reconstruction. In addition, change map and corresponding heat map generated in Figure 8(e) exhibit highly accurate target localization and fine boundary compared with GT. This demonstrates that proposed model can effectively capture changed target and improve changed region localization accuracy through physical parameter decomposition and reconstruction network.

5 Conclusion

This work integrates RS physical mechanisms with causal analysis and proposes a causal physics-disentangled network for RSCD, aiming to alleviate limited physical interpretability of data-driven methods and their sensitivity to imaging environment variations. First, based on major energy transfer pathways in RS imaging, including direct irradiance, diffuse irradiance, and atmospheric path radiance, a ground-object spectral radiance model is established. From a causal perspective, bi-temporal differences in ground-object reflectance are identified as primary evidence for genuine changes. Second, a physical parameter decomposition and reconstruction network is designed to disentangle bi-temporal RS images into total irradiance, ground-object reflectance, and atmospheric transmittance. Orthogonality, reconstruction, and spatial smoothness constraints are jointly introduced to improve parameter separability. Finally, a reflectance branch and semantic branch are constructed, where consistency constraints reduce environmental variation effects and feature alignment guides semantic features toward reflectance changes. The final change map is obtained through adaptive fusion. Experiments demonstrate that proposed method achieves superior quantitative metrics and qualitative performance on multiple datasets, fully validating its effectiveness, robustness, and adaptability to complex scenes.

AI USE STATEMENT

The authors did not use generative AI tools in the preparation of this manuscript.

REFERENCES

Luca Savant Aira, Gabriele Facciolo, and Thibaud Ehret. Gaussian splatting for efficient satellite image photogrammetry. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5959–5969, June 2025.

Wele Gedara Chaminda Bandara and Vishal M. Patel. A transformer-based siamese network for change detection. In *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*, pp. 207–210, 2022. doi: 10.1109/IGARSS46834.2022.9883686.

Wele Gedara Chaminda Bandara, Nithin Gopalakrishnan Nair, and Vishal M. Patel. Ddpm-cd: Denoising diffusion probabilistic models as feature extractors for remote sensing change detection. In *2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 5250–5262, 2025. doi: 10.1109/WACV61041.2025.00513.

Hao Chen, Zipeng Qi, and Zhenwei Shi. Remote sensing image change detection with transformers. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–14, 2022. doi: 10.1109/TGRS.2021.3095166.

Hongruixuan Chen, Jian Song, Chengxi Han, Junshi Xia, and Naoto Yokoya. Changemamba: Remote sensing change detection with spatiotemporal state space model. *IEEE Transactions on Geoscience and Remote Sensing*, 62:1–20, 2024. doi: 10.1109/TGRS.2024.3417253.

Jiuchen Chen, Xinyu Yan, Qizhi Xu, and Kaiqi Li. Tokenize image patches: Global context fusion for effective haze removal in large images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2258–2268, June 2025.

Minh Kha Do, Kang Han, Phu Lai, Khoa T. Phan, and Wei Xiang. Robsense: A robust multi-modal foundation model for remote sensing with static, temporal, and incomplete data adaptability. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7427–7436, June 2025.

S. Du, Y. Zou, J. Li, M. Liu, Y. Li, C. Shang, and Q. Shen. Pansharpening for thin-cloud contaminated remote sensing images: A unified framework and benchmark dataset. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 3696–3704, 2026. doi: 10.1609/aaai.v40i5.37369.

Yuheng Du, Sheng Yang, Lingxuan Wang, Zhenghua Hou, Chengying Cai, Zhitao Tan, Mingxia Chen, Shi-Sheng Huang, and Qiang Li. Rtmap: Real-time recursive mapping with change detection and localization. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 28021–28030, 2025. doi: 10.1109/ICCV51701.2025.02601.

Yuchao Feng, Honghui Xu, Jiawei Jiang, Hao Liu, and Jianwei Zheng. Icif-net: Intra-scale cross-interaction and inter-scale feature fusion network for bitemporal remote sensing images change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–13, 2022. doi: 10.1109/TGRS.2022.3168331.

Yuchao Feng, Jiawei Jiang, Honghui Xu, and Jianwei Zheng. Change detection on remote sensing images using dual-branch multilevel intertemporal network. *IEEE Transactions on Geoscience and Remote Sensing*, 61:1–15, 2023. doi: 10.1109/TGRS.2023.3241257.

Mingbo Hong, Feng Liu, Caroline Gevaert, George Vosselman, and Hao Cheng. Bridge: Basis-driven causal inference marries vfms for domain generalization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 31962–31973, June 2026.

Qiusheng Huang, Xiaohui Zhong, Xu Fan, and Hao Li. Fuxi-rtm: A physics-guided prediction framework with radiative transfer modeling. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 8852–8862, October 2025.

Jae-Woo Kim and Ue-Hwan Kim. Towards generalizable scene change detection. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 24463–24473, 2025. doi: 10.1109/CVPR52734.2025.02278.

Tao Lei, Xinzhe Geng, Hailong Ning, Zhiyong Lv, Maoguo Gong, Yaochu Jin, and Asoke K. Nandi. Ultralightweight spatial–spectral feature cooperation network for change detection in remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 61:1–14, 2023. doi: 10.1109/TGRS.2023.3261273.

Yulong Li, Xiwei Liu, Barrett Tang, Zhixiang Lu, Ming Hu, Yichen Li, Haochen Xue, Peixin Guo, Jionglong Su, Yutong Xie, Eran Segal, and Imran Razzak. Carl: Preserving causal structure in representation learning. In C. Vondrick, B. Hariharan, C. Raffel, L. Pinto, D. Yang, and A. Faust (eds.), *International Conference on Learning Representations*, volume 2026, pp. 132116–132176, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/file/d63cf0622eed012a17fe88fced64dcb8-Paper-Conference.pdf.

Xusheng Liang, Lihua Zhou, Nianxin Li, Miao Xu, Ziyang Song, Dong Yi, Jinlin Wu, Jiawei Ma, Hongbin Liu, Zhen Lei, and Jiebo Luo. Multimodal causality-driven representation learning for generalizable medical image segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 13670–13679, June 2026.

Zhifang Liao, Junhao Li, HaoKang Ding, and Yucheng Song. Darc: Dual adjustment reasoning with counterfactuals for trustworthy chest x-ray classification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 28234–28243, June 2026.

Bo Liu, Qiao Qin, and Qinghui He. Causalclip: Causally-informed feature disentanglement and filtering for generalizable detection of generated images. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 7069–7077, 2026. doi: 10.1609/aaai.v40i9.37642.

Ba Hung Ngo, Nhat-Tuong Do-Tran, Tuan-Ngoc Nguyen, Hae-Gon Jeon, and Tae Jong Choi. Learning cnn on vit: A hybrid model to explicitly class-specific boundaries for domain adaptation. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 28545–28554, 2024. doi: 10.1109/CVPR52733.2024.02697.

Jay N. Paranjape, Celso De Melo, and Vishal M. Patel. A mamba-based siamese network for remote sensing change detection. In *2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 1186–1196, 2025. doi: 10.1109/WACV61041.2025.00123.

Ruike Song, Zeen Song, Huijie Guo, and Wenwen Qiang. Causal reward adjustment: Mitigating reward hacking in external reasoning via backdoor correction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 33019–33027, 2026. doi: 10.1609/aaai.v40i39.40584.

Gencer Sumbul, Chang Xu, Emanuele Dalsasso, and Devis Tuia. Smarties: Spectrum-aware multisensor auto-encoder for remote sensing images. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 5569–5578, October 2025.

Aniket Vashishtha, Qirun Dai, Hongyuan Mei, Amit Sharma, Chenhao Tan, and Hao Peng. Executable counterfactuals: Improving llms'causal reasoning through code. In C. Vondrick, B. Hariharan, C. Raffel, L. Pinto, D. Yang, and A. Faust (eds.), *International Conference on Learning Representations*, volume 2026, pp. 101098–101123, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/file/a50aa557c4be35aa2bf13a471601e23f-Paper-Conference.pdf.

Bin Wang, Xiaohu Jiang, Pinle Qin, and Jianchao Zeng. Dsfdcd: Joint distribution sampling and feature decoupling deep network for remote sensing change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 63:1–12, 2025a. doi: 10.1109/TGRS.2025.3585229.

Bin Wang, Sichao Qiao, Ziyv Zhao, Pinle Qin, and Jianchao Zeng. Uafdcd: Uncertainty-aware feature decoupling network for remote sensing change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 64:5625815–5625815, 2026a. doi: 10.1109/TGRS.2026.3701174.

Bin Wang, Bo Wang, Pinle Qin, and Jianchao Zeng. FCdDNet: Feature cross-domain decoupling network for remote sensing change detection. *Pattern Recognition*, 177:113259, 2026b. ISSN 0031-3203. doi: https://doi.org/10.1016/j.patcog.2026.113259.

Zhipan Wang, Zijun Yang, and Qingling Zhang. Ia-cdnet: Change detection in adverse remote sensing image conditions with an advanced image-adaptive method. *IEEE Transactions on Geoscience and Remote Sensing*, 63:1–19, 2025b. doi: 10.1109/TGRS.2024.3518568.

W. Wu, L. Meng, Z. Qi, Z. Li, Y. Zhang, X. Yan, and X. Meng. Explicit modeling of causal factors and confounders for image classification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 10763–10771, 2026. doi: 10.1609/aaai.v40i13.38051.

Dazhi Xu, Ming Li, Yan Wu, Peng Zhang, and Xinyue Xin. Statistic-guided difference enhancement graph transformer for unsupervised change detection in polsar images. *IEEE Transactions on Circuits and Systems for Video Technology*, 35(11):11667–11684, Nov 2025. ISSN 1558-2205. doi: 10.1109/TCSVT.2025.3575082.

Shaowu Xu, Xibin Jia, Chao Fan, Junyu Gao, Jing Chang, and Qianmei Sun. Progressive cross-modal causal intervention for long-term action recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 31186–31195, June 2026.

Ruteng Yu, Lukun Wang, and Jiaming Pei. Gsag-cdgan: A generalizable small-sample attention-guided gan for remote sensing change detection (student abstract). *Proceedings of the AAAI Conference on Artificial Intelligence*, 40(48):41452–41454, Mar. 2026. doi: 10.1609/aaai.v40i48.42303. URL https://ojs.aaai.org/index.php/AAAI/article/view/42303.

Qi Zang, Dong Zhao, Shuang Wang, Dou Quan, and Zhun Zhong. Feature spectrum learning for remote sensing change detection. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 12647–12657, 2025a. doi: 10.1109/CVPR52734.2025.01180.

Qi Zang, Dong Zhao, Shuang Wang, Dou Quan, and Zhun Zhong. Feature spectrum learning for remote sensing change detection. *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 12647–12657, 2025b. URL https://api.semanticscholar.org/CorpusID:280148271.

Hong Zhang, Yuhang Teng, Haojie Li, and Zhihui Wang. Strobustnet: Efficient change detection via spatial–temporal robust representations in remote sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 63:1–15, 2025. doi: 10.1109/TGRS.2025.3540794.

Mingwei Zhang, Jingtao Hu, Qiang Li, and Qi Wang. Exploring generalizable remote sensing change detection via low-rank exchange adaptation of vision foundation model. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 12663–12671, 2026a. doi: 10.1609/aaai.v40i15.38262.

Mingwei Zhang, Jingtao Hu, Qiang Li, and Qi Wang. Exploring generalizable remote sensing change detection via low-rank exchange adaptation of vision foundation model. *Proceedings of the AAAI Conference on Artificial Intelligence*, 40(15):12663–12671, Mar. 2026b. doi: 10.1609/aaai.v40i15.38262. URL https://ojs.aaai.org/index.php/AAAI/article/view/38262.

Y. Zhang, Y. Zhang, Y. Zheng, B. Raducanu, and D. Liu. Causal-tune: Mining causal factors from vision foundation models for domain generalized semantic segmentation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 12916–12924, 2026c. doi: 10.1609/aaai.v40i15.38290.

Jiaqi Zhao, Jianpeng Xie, Yong Zhou, Wen-Liang Du, Hancheng Zhu, and Rui Yao. Causal decoupling domain generalization for remote sensing change detection. *Proceedings of the AAAI Conference on Artificial Intelligence*, 40(15):13135–13143, Mar. 2026. doi: 10.1609/aaai.v40i15.38314. URL https://ojs.aaai.org/index.php/AAAI/article/view/38314.