Enhancing Semi-Supervised Change Detection via Fractional Differential Perturbation

Pengcheng Lu
College of Computer Science
Sichuan University
Chengdu, Sichuan 610065, China
lpcfaltings@stu.scu.edu.cn

Yifei Pu
College of Computer Science
Sichuan University
Chengdu, Sichuan 610065, China
0000-0003-2975-4976

***Abstract*—Semi-supervised change detection (SSCD) aims to learn robust pixel-level change maps from limited annotated bi-temporal images and abundant unlabeled pairs. Recent consistency-regularization frameworks such as FixMatch and UniMatch rely on strong augmentation to enforce weak-to-strong prediction consistency, but hard-occlusion methods such as Cutout and CutMix may disrupt the topology of ground objects in remote sensing scenes. We therefore propose Fractional Differential Perturbation (FDP), a structure-preserving strong augmentation strategy based on fractional calculus. FDP injects controlled high-frequency perturbations while preserving low-frequency geometric structure, thereby increasing texture-level difficulty without destroying object boundaries. Extensive experiments on the WHU-CD and LEVIR-CD datasets show that FDP consistently improves semi-supervised baselines, especially in low-label regimes (e.g., a 5.3% IoU gain on WHU-CD with 5% labels).**

***Index Terms*—Semi-supervised learning, Change detection, Fractional differential operator, Consistency regularization**

I. Introduction

Remote sensing change detection (CD) aims to identify semantic changes between bi-temporal images [1]. It underpins critical Earth observation applications, including urban expansion monitoring, land resource management, and disaster assessment [2], [3]. Beyond satellite imagery, CD techniques have also extended to industrial inspection, robotics, and street-view analysis [4]–[7].

Modern CD systems have evolved from convolutional Siamese architectures to Transformer-based models and stronger feature interaction designs [8]–[10]. Despite this progress, state-of-the-art performance depends heavily on pixel-level annotations, which motivates *semi-supervised change detection* (SSCD): learning from a small labeled set alongside abundant unlabeled data.

Consistency regularization has become the dominant SSCD paradigm, led by frameworks such as FixMatch and UniMatch [11], [12], which enforce prediction consistency between weakly and strongly augmented views. In this setting, the design of strong augmentation is crucial. However, standard techniques exhibit clear limitations in remote sensing: 1) hard-occlusion methods (e.g., Cutout and CutMix) may disrupt the topological continuity of ground objects such as roads [13], [14]; 2) simple photometric perturbations lack the structural diversity needed to simulate complex environmental shifts.

To address these issues, we introduce **Fractional Differential Perturbation (FDP)**, a structure-preserving strong augmentation tailored for SSCD. FDP uses fractional calculus [15] to inject soft perturbations that increase texture-level difficulty while retaining geometric integrity. We further adopt a pair-aware bi-temporal augmentation design that preserves spatial alignment through shared geometric preprocessing while applying symmetric strong perturbations to the two temporal images.

Our contributions are: 1) We identify the limitations of hard-masking augmentation in remote sensing CD. 2) We propose Fractional Differential Perturbation (FDP), a structure-preserving perturbation method that improves pseudo-label reliability while remaining compatible with existing consistency-based pipelines. 3) Extensive experiments demonstrate the effectiveness of FDP, especially in low-data regimes (e.g., 5% labeled data), while confirming that it introduces no additional learnable parameters and no inference-time overhead.

II. Related Work

A. Deep Learning for Change Detection

CD models have evolved from convolutional Siamese networks [8] to Transformer-based architectures and stronger feature interaction modules [9], [16], [17]. However, they still depend heavily on dense pixel-level annotations, motivating semi-supervised learning (SSL).

B. Semi-Supervised Learning

SSL leverages unlabeled data to improve generalization. Representative directions include adversarial learning [18], [19], pseudo-labeling [20], and, most notably, consistency regularization. Temporal Ensembling [21], Mean Teacher [22], FixMatch [11], and UniMatch [12] progressively strengthened weak-to-strong consistency learning. In remote sensing, SemiCDNet [23] and SemiCD [24] adapted these ideas to change detection, but standard strong augmentations such as Cutout and CutMix [13], [14] may still disrupt target topology, motivating structure-preserving perturbations.

C. Fractional Calculus in Image Processing

Fractional calculus generalizes differentiation to arbitrary orders and offers useful non-local properties. Compared with integer-order operators such as Sobel, fractional derivatives can enhance high-frequency textures while preserving low-frequency structure [25]. They have also been used in edge detection and image denoising [26]. In remote sensing, Liu *et al.* [27] showed that fractional differential masks can enhance weak texture details, motivating our use of fractional calculus for structure-aware strong augmentation.

III. Methodology

This section presents the proposed semi-supervised change detection framework. We first outline the overall architecture based on a teacher-student consistency-regularization paradigm. We then describe the mathematical formulation of **Fractional Differential Perturbation (FDP)** and explain its role as a structure-preserving strong augmentation. Finally, we introduce the bi-temporal consistency strategy and the overall optimization objective.

A. Overview of the Framework

Our method builds on the consistency regularization framework, adapting its perturbation strategy to the change detection architecture. As illustrated in Fig. 1, the overall architecture follows a teacher-student learning paradigm.

Let $\mathcal{D}_L = \{(x_l, y_l)\}$ denote the labeled dataset containing a small set of bi-temporal pairs, and let $\mathcal{D}_U = \{u\}$ denote the abundant unlabeled dataset. The core idea is *weak-to-strong consistency*. For an unlabeled pair $u = (t_1, t_2)$, the teacher model generates a pseudo-label $\hat{y}_w$ from a weakly augmented view $u_w = \mathcal{A}_w(u)$. The student model is then trained to predict this pseudo-label from a strongly augmented view $u_s = \mathcal{A}_s(u)$.

Unlike conventional methods that rely heavily on Cutout for $\mathcal{A}_s$, which may destroy the topological continuity of remote sensing targets, we propose a texture-oriented perturbation strategy based on fractional calculus, termed FDP.

B. Fractional Differential Perturbation (FDP)

We formulate FDP as a *parameter-free* perturbation method. Unlike learnable layers that require gradient updates, FDP relies on mathematically derived fractional operators to enhance high-frequency details while retaining low-frequency structural information.

1) Fractional Differential Operator: We adopt the Grünwald-Letnikov (G-L) definition for discrete fractional differentiation, which is well-suited for digital image processing. For a 1D signal $f(t)$, the derivative of fractional order v is generally defined as:

$$D^v f(t) \approx \frac{1}{h^v} \sum_{k=0}^{n-1} (-1)^k \binom{v}{k} f(t - kh) \tag{1}$$

where h is the step size.

A digital image $f(x, y)$ is a discrete 2D signal where the minimum division interval is $h = 1$. To better handle the discrete nature of image pixels without interpolation artifacts, we adopt the **Pu-2 fractional differential operator** [15]. For a fractional order v, the partial difference in the negative x-direction and y-direction is derived as:

$$\begin{aligned} \frac{\partial^v f(x,y)}{\partial x^v} \approx\ & C_{-1} f(x+1, y) + C_0 f(x, y) \\ & + \sum_{k=1}^{n-2} C_k f(x-k, y) \end{aligned} \tag{2}$$

$$\begin{aligned} \frac{\partial^v f(x,y)}{\partial y^v} \approx\ & C_{-1} f(x, y+1) + C_0 f(x, y) \\ & + \sum_{k=1}^{n-2} C_k f(x, y-k) \end{aligned} \tag{3}$$

This formulation introduces a "look-ahead" term (C_{-1}) alongside the "history" terms (C_k), allowing the operator to capture local context more effectively than standard causal filters.

2) Calculation of Mask Coefficients: The filter coefficients $\{C_{-1}, C_0, \ldots, C_k\}$ are critical for determining the frequency response of the perturbation. Unlike integer-order operators with fixed integer values (e.g., -1, 1), these coefficients are non-linearly related to the order v via the Gamma function $\Gamma(\cdot)$. Based on the Pu-2 derivation, the first few coefficients used in our 5×5 implementation are calculated as follows:

$$\begin{aligned} C_{-1} &= \frac{v}{4} + \frac{v^2}{8} \\ C_0 &= 1 - \frac{v^2}{2} - \frac{v^3}{8} \\ C_1 &= \frac{\Gamma(2-v)}{2\Gamma(-v)} \left(\frac{v}{4} + \frac{v^2}{8} \right) + \frac{\Gamma(1-v)}{\Gamma(-v)} \left(1 - \frac{v^2}{4} \right) \\ &\quad + \left(-\frac{v}{4} + \frac{v^2}{8} \right) \end{aligned} \tag{4}$$

The subsequent coefficients C_2 and C_3 follow a similar recursive structure involving $\Gamma(n-v)$. Intuitively, the Gamma-based terms control how quickly the weights decay as the spatial distance increases: slower decay yields a wider and smoother interaction neighborhood, whereas integer-order filters concentrate most of their energy on a few adjacent pixels. As a result, FDP perturbs fine textures over a local region without collapsing into a hard edge-only response. The slowly decaying non-zero coefficients also endow the operator with a "long-term memory" property, allowing it to capture non-local textural context that differs fundamentally from local integer-order gradients.

3) Multi-Directional Feature Aggregation: To capture omnidirectional texture variations, we construct 8 fractional differential masks corresponding to the directions $d \in \{0°, 45°, \ldots, 315°\}$. Each mask is a 5×5 convolution kernel constructed by spatially arranging the coefficients $\{C_{-1}, \ldots, C_3\}$ derived in Eq. (4). For instance, the positive x-direction mask W_{x+} aligns coefficients along the center row, while diagonal masks (e.g., W_{LDD}) align them along the diagonal. Other masks are obtained by rotating or flipping these base structures. For an input image I, we compute the

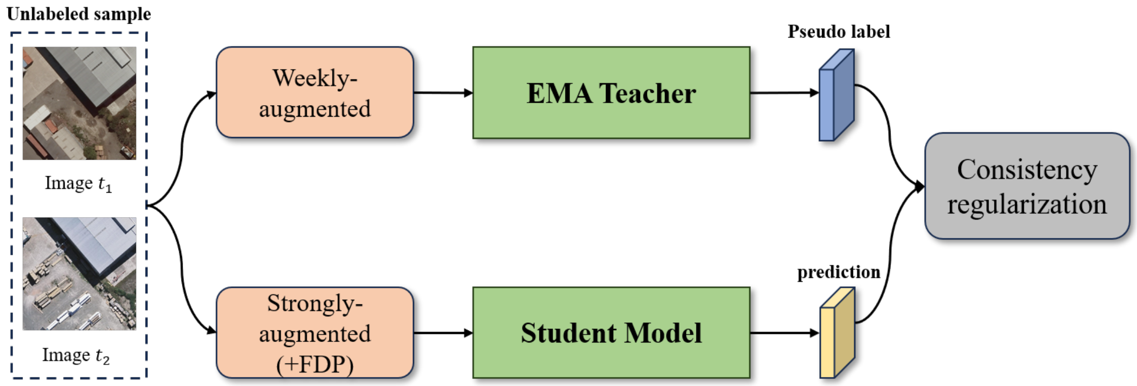

Figure 1. Pipeline of the unlabeled branch in the proposed method. FDP is inserted into the strong augmentation branch, where it uses the non-local property of fractional derivatives to inject texture perturbations while preserving the geometric structure of ground objects. The teacher and student share the same architecture, and the teacher weights are updated via EMA.

response maps $\{R_d\}$ by convolving I with all 8 masks. To ensure rotation invariance, we select the maximum response:

$$R_{max}(i,j) = \max_{d} |(I * W_d)(i,j)| \tag{5}$$

4) Residual Perturbation Injection: FDP is applied as an additive perturbation rather than a replacement. The augmented image I_{aug} is generated by superimposing the fractional response onto the original image:

$$I_{aug} = \text{Clamp}(I + \lambda \cdot R_{max}, 0, 1) \tag{6}$$

where λ is a hyperparameter controlling the perturbation intensity. This process modifies the local texture distribution, making the surface appear rougher or more complex, while preserving the global geometric boundaries of objects. This forces the model to rely on shape and structure—features that are robust for change detection—rather than overfitting to specific textures.

C. Bi-temporal Consistent Augmentation

A critical challenge in semi-supervised change detection is maintaining the semantic consistency between bi-temporal images. Naively applying random strong augmentations independently to the pre-event image t_1 and post-event image t_2 can introduce artificial differences, which the model might mistake for real land-cover changes.

To address this, we adopt a **pair-aware strong augmentation design**. In practice, the bi-temporal pair first undergoes shared geometric preprocessing, including resize, crop, and horizontal flip, to preserve spatial alignment. The two temporal images are fed into symmetric strong-augmentation branches with independently sampled random perturbations:

$$x_s = (\mathcal{A}_s(t_1;\theta_1), \mathcal{A}_s(t_2;\theta_2)) \tag{7}$$

where θ_1 and θ_2 denote independently sampled random parameters for the strong augmentation pipelines applied to t_1 and t_2, respectively. Because spatial alignment is preserved by the shared geometric preprocessing, whereas photometric and textural perturbations are sampled independently in the two branches, these perturbations force the model to ignore augmentation-induced texture noise and focus on genuine semantic changes between the two timestamps.

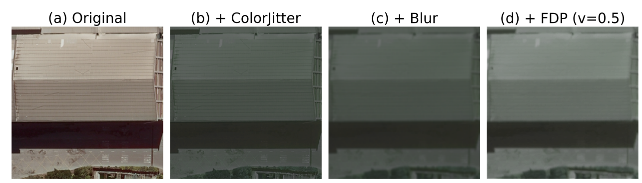

Figure 2. Proposed strong augmentation pipeline. FDP is introduced as a texture-perturbation step alongside photometric distortions (ColorJitter/Blur) and geometric deformation (CutMix).

D. Augmentation Pipeline and Optimization

The strong augmentation pipeline $\mathcal{A}_s$ models realistic appearance variations through a cascaded sequence: 1) **Photometric Distortion** using ColorJitter; 2) **Blurring** via Gaussian Blur ($p = 0.5$); 3) **Texture Perturbation** via the proposed FDP ($p = 0.4$); and 4) **Geometric Deformation** using CutMix. FDP complements CutMix: the former encourages local texture invariance, whereas the latter perturbs coarse spatial context through regional patch replacement. The network is optimized using a joint loss $\mathcal{L} = \mathcal{L}_s + \mathcal{L}_u$. The supervised loss $\mathcal{L}_s$ is the standard cross-entropy on labeled data. The unsupervised consistency loss $\mathcal{L}_u$ filters out low-confidence predictions using a threshold τ:

$$\mathcal{L}_u = \frac{1}{|\mathcal{D}_U|} \sum_{u \in \mathcal{D}_U} \mathbb{I}(\max(\hat{y}_w) \geq \tau) \cdot H(\hat{y}_w, p_s) \tag{8}$$

where p_s is the prediction on the FDP-augmented view. By minimizing this loss, the model learns to produce consistent predictions despite the significant textural perturbations.

IV. Experiments

A. Experimental Setup

1) Datasets: We evaluate on two benchmarks. **LEVIR-CD** **[2]** contains 637 pairs of 1024×1024 images (0.5m/pixel), split

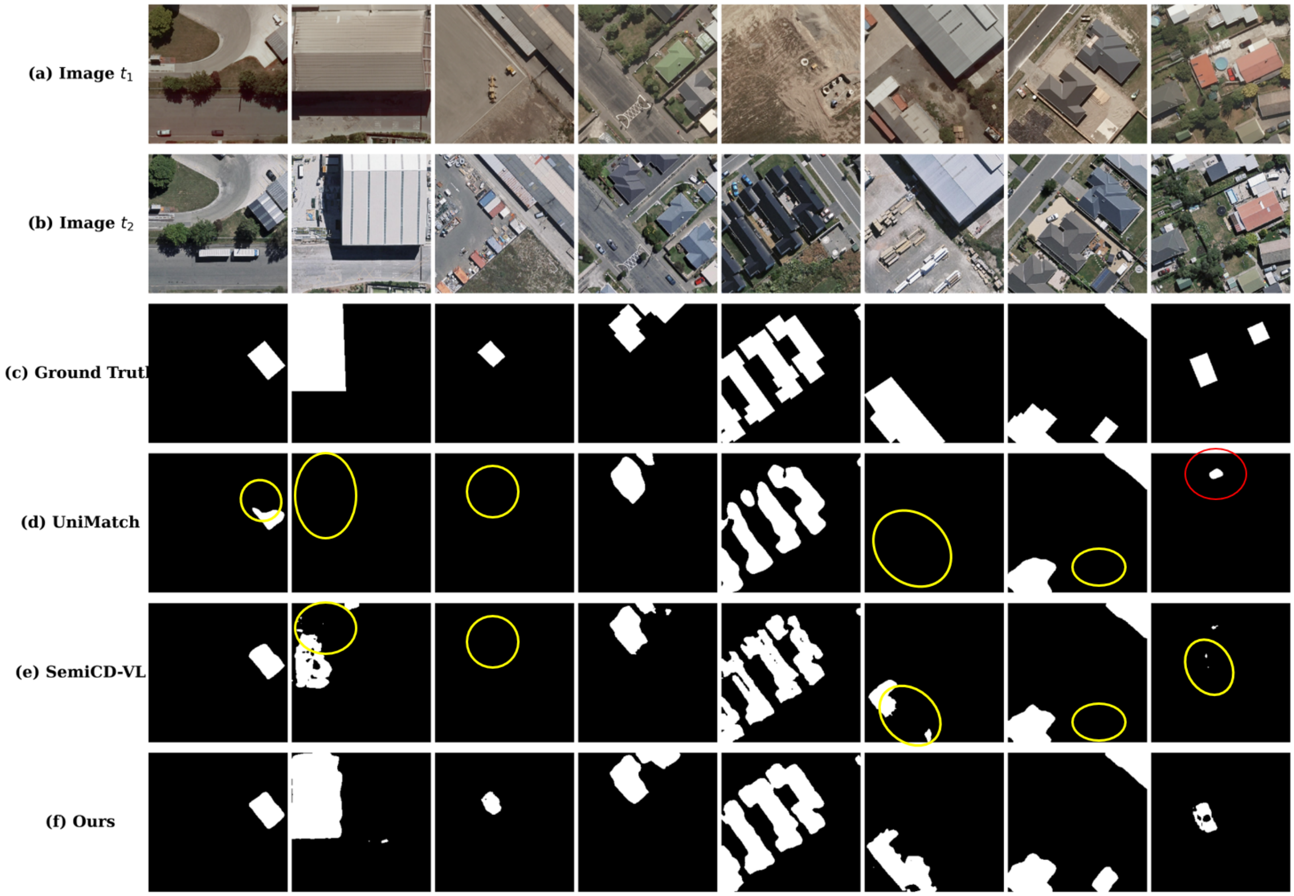

Figure 3. Visual comparison results on the WHU-CD dataset. (a) Image t_1, (b) Image t_2, (c) Ground Truth, (d) UniMatch, (e) SemiCD-VL, and (f) Ours. Red circles indicate false alarms, and yellow circles indicate missed detections.

into 445/64/128 for training/validation/testing following [24]. Each image is cropped into non-overlapping 256×256 patches, yielding 7120 training pairs. **WHU-CD [28]** comprises aerial imagery (0.2m/pixel) cropped into 5947 training and 744 testing pairs of size 256×256. Following UniMatch [12], we randomly sample 5%, 10%, 20%, and 40% of the training set as labeled data $\mathcal{D}_L$.

2) Implementation Details: We employ DPT-DINOv2-S as the backbone network for feature extraction. The models are implemented in PyTorch and trained on two Tesla T4 GPUs. We use the AdamW optimizer with a weight decay of 0.01. The encoder learning rate is set to 5×10^{-4}, the batch size is 8 labeled plus 8 unlabeled samples, and training lasts for 60 epochs. The confidence threshold τ for consistency regularization is set to 0.95. During preprocessing, all image pairs are randomly resized with a scale factor sampled from $[0.6, 1.4]$, cropped to 256×256, and horizontally flipped with probability 0.5. In the strong branch, the two temporal images follow the same augmentation policy but are perturbed with independent random draws: ColorJitter$(0.5, 0.5, 0.5, 0.25)$ is applied with probability 0.8, Gaussian blur with probability 0.5, and CutMix with probability 0.5. For FDP, the fractional order is set to $v = 0.5$, the perturbation intensity λ to 0.7, and the application probability to $p_{\text{fdp}} = 0.4$. Because FDP is used only in the training-time augmentation branch and is implemented with fixed fractional masks, it introduces no additional learnable parameters and no inference-time overhead. Unless otherwise stated, all ablation variants share the same split, backbone, optimizer, and training budget, and differ only in the perturbation method or the target FDP hyperparameter.

B. Comparison with State-of-the-Arts

We compare the proposed method against several representative semi-supervised change detection frameworks. The quantitative results in terms of Intersection over Union (IoU) and Overall Accuracy (OA) on the WHU-CD and LEVIR-CD datasets are summarized in Table I and Table II, respectively.

The compared methods cover the main semi-supervised paradigms in remote sensing. GAN-based S4GAN employs a discriminator to distinguish predicted segmentation maps from ground truth and use the discriminator's confidence to select unlabeled samples. It performs poorly in low-data regimes, reaching only 18.30% IoU on WHU-CD with 5% labels, likely because discriminator training becomes unstable under scarce supervision.

Table I
COMPARISON RESULTS ON WHU-CD DATASET. THE BEST RESULTS ARE HIGHLIGHTED IN **BOLD**.

Method	5% (297)		10% (594)		20% (1189)		40% (2378)	
	IoU	OA	IoU	OA	IoU	OA	IoU	OA
S4GAN [19]	18.30	96.69	62.60	98.15	70.80	98.60	76.40	98.96
SemiCDNet [23]	51.70	97.71	62.00	98.16	66.70	98.28	75.90	98.93
SemiCD [24]	65.80	98.37	68.10	98.47	74.80	98.84	77.20	98.96
UniMatch V1 [12]	77.50	99.06	78.90	99.10	82.90	99.26	84.40	99.32
SemiCD-VL [29]	81.60	**99.24**	82.24	99.26	84.80	99.36	85.70	99.39
UniMatch V2	79.06	98.91	82.26	99.26	86.10	**99.42**	**87.60**	**99.48**
Ours	**84.36**	99.11	**86.30**	**99.30**	**86.57**	99.30	**87.60**	99.40

Table II
COMPARISON RESULTS ON LEVIR-CD DATASET. THE BEST RESULTS ARE HIGHLIGHTED IN **BOLD**.

Method	5%		10%		20%		40%	
	IoU	OA	IoU	OA	IoU	OA	IoU	OA
S4GAN	64.00	97.89	67.00	98.11	73.40	98.51	75.40	98.62
SemiCDNet	67.60	98.17	71.50	98.42	74.30	98.58	75.50	98.63
SemiCD	72.50	98.47	75.50	98.63	76.20	98.68	77.20	98.72
UniMatch V1	75.60	98.62	79.00	98.83	79.00	98.84	78.20	98.79
SemiCD-VL	81.90	99.02	82.60	99.06	82.70	99.05	83.00	99.07
UniMatch V2	82.10	99.01	82.70	99.05	83.20	99.08	83.50	99.10
Ours	**82.36**	**99.20**	**83.13**	**99.23**	**83.75**	**99.26**	**83.97**	**99.28**

Consistency-regularization methods, including SemiCDNet, SemiCD, and UniMatch, are much stronger, although SemiCDNet and SemiCD still rely on relatively simple perturbations or feature consistency. UniMatch V1/V2 provide the strongest baseline through joint image- and feature-level consistency, while SemiCD-VL is competitive at higher computational cost.

Our approach integrates Fractional Differential Perturbation (FDP) into the consistency framework as a texture-preserving perturbation method. In the WHU-CD dataset with only 5% labeled data, our method achieves an IoU of 84.36%, outperforming the strong baseline UniMatch V2 by a margin of 5.3%. On LEVIR-CD, the gains are smaller but consistent across all label ratios. This pattern is consistent with the view that FDP stabilizes weak-to-strong consistency in low-label settings, indicating that FDP acts as a stable regularizer rather than a dataset-specific heuristic.

C. Ablation Studies

To investigate the effectiveness of the proposed components, we conduct comprehensive ablation studies on the WHU-CD dataset with 10% labeled data. Each ablation isolates one factor while keeping the remaining training configuration unchanged.

1) Effectiveness of FDP Strategy: To demonstrate that the performance gain stems from the structure-aware nature of our fractional perturbation rather than simple noise injection, we compare FDP against unstructured Gaussian noise and integer-order Sobel perturbation.

Structure-Aware vs. Random Noise: As shown in Table III, simply adding random Gaussian noise (81.70%) degrades performance compared to the baseline (82.26%). This result suggests that structure-agnostic noise corrupts semantic information without providing meaningful regularization. In contrast, our FDP achieves **86.30%**, showing that structure-aware perturbation can increase difficulty while preserving semantic integrity.

Fractional vs. Integer Order: Although Sobel improves over the baseline (+2.18%), it still lags behind FDP (+4.04%). The Sobel operator acts as a strict high-pass filter, extracting sparse edges while suppressing low-frequency context. This highlights the advantage of the fractional order ($v = 0.5$), which behaves as a flexible enhancer rather than a hard edge extractor.

Complementarity with Geometric Augmentation: Since the baseline already includes CutMix as a geometric perturbation, the significant gain brought by FDP (**+4.04%**) confirms that our texture-based perturbation is complementary to geometric augmentation. CutMix encourages robustness to coarse spatial-context replacement, whereas FDP strengthens local texture invariance, and their combination yields the best performance.

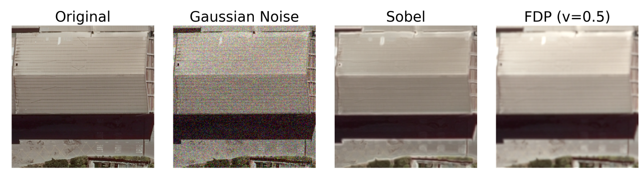

Figure 4. Visual comparison of different differential masks and augmentation. (a) Original Image. (b) Gaussian Noise. (c) Sobel Operator (Integer-order). (d) FDP (Fractional-order, $v = 0.5$): enhances rich textures while preserving semantic structure.

Table III
ABLATION STUDY OF DIFFERENT AUGMENTATION STRATEGIES ON WHU-CD (10% LABELED).

Strategy	IoU (%)		OA (%)	
	Value	Δ	Value	Δ
Baseline (UniMatch w/ CutMix)	82.26	–	99.26	–
+ Gaussian Noise (Random)	81.70	$\downarrow$0.56	99.24	$\downarrow$0.02
+ Sobel (Integer-order)	84.44	$\uparrow$2.18	99.19	$\downarrow$0.07
+ FDP (Ours)	**86.30**	**$\uparrow$4.04**	**99.30**	**$\uparrow$0.04**

2) Parameter Sensitivity Analysis: We investigate three key hyperparameters: fractional order v, intensity λ, and application probability p_{fdp}. As summarized in Table IV, performance peaks at $v = 0.5$, $\lambda = 0.7$, and $p_{\text{fdp}} = 0.4$. **1) Fractional Order v:** The results show that performance peaks at a lower fractional order of $v = 0.5$, suggesting that fractional differentiation enhances texture details without destroying the low-frequency structural context. **2) Intensity λ:** Moderate intensity ($\lambda = 0.7$) is optimal; excessive noise ($\lambda = 1.0$) degrades semantic content. **3) Probability p_{fdp}:** Applying FDP frequently but not deterministically ($p_{\text{fdp}} = 0.4$) provides the best regularization trade-off.

V. CONCLUSION

We proposed **Fractional Differential Perturbation (FDP)** to address the limitations of hard-occlusion augmentations in SSCD. FDP leverages fractional calculus to inject structure-preserving textural perturbations, increasing difficulty for the

Table IV
PARAMETER SENSITIVITY ANALYSIS ON WHU-CD (10% LABELED).

Parameter	IoU (%) values under different settings				
Order v	0.2	**0.5**	1.2	1.5	1.8
	84.82	**86.30**	84.99	85.85	85.05
Intensity λ	0.4	0.6	**0.7**	0.8	1.0
	85.31	85.60	**86.30**	85.85	84.13
Prob p_{fdp}	0.2	**0.4**	0.6	0.8	1.0
	85.69	**86.30**	85.85	85.03	84.90

student model while maintaining geometric integrity. Extensive experiments demonstrate that FDP outperforms existing methods and remains complementary to CutMix. Since FDP is inserted only in the strong augmentation branch, it does not alter the inference architecture. Future work will include more exhaustive multi-run statistical evaluation, explicit wall-clock overhead analysis during training, direct pseudo-label quality analysis, and broader validation on additional datasets or dense prediction tasks beyond remote sensing change detection.

ACKNOWLEDGMENT

The authors acknowledge the use of an AI-assisted language model (GPT, Gemini) for language editing and improving the clarity of the manuscript.

REFERENCES

[1] W. Shi, M. Zhang, R. Zhang, S. Chen, and Z. Zhan, "Change detection based on artificial intelligence: State-of-the-art and challenges," *Remote Sensing*, vol. 12, no. 10, p. 1688, 2020.

[2] H. Chen and Z. Shi, "A spatial-temporal attention-based method and a new dataset for remote sensing image change detection," *Remote sensing*, vol. 12, no. 10, p. 1662, 2020.

[3] X. X. Zhu, D. Tuia, L. Mou, G.-S. Xia, L. Zhang, F. Xu, and F. Fraundorfer, "Deep learning in remote sensing: A comprehensive review and list of resources," *IEEE geoscience and remote sensing magazine*, vol. 5, no. 4, pp. 8–36, 2017.

[4] J.-M. Park, J.-H. Jang, S.-M. Yoo, S.-K. Lee, U.-H. Kim, and J.-H. Kim, "Changesim: Towards end-to-end online scene change detection in industrial indoor environments," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2021, pp. 8578–8585.

[5] S. Salimpour, J. P. Queralta, and T. Westerlund, "Self-calibrating anomaly and change detection for autonomous inspection robots," in *2022 Sixth IEEE International Conference on Robotic Computing (IRC)*. IEEE, 2022, pp. 207–214.

[6] P. F. Alcantarilla, S. Stent, G. Ros, R. Arroyo, and R. Gherardi, "Street-view change detection with deconvolutional networks," *Autonomous Robots*, vol. 42, no. 7, pp. 1301–1322, 2018.

[7] Y. Lei, D. Peng, P. Zhang, Q. Ke, and H. Li, "Hierarchical paired channel fusion network for street scene change detection," *IEEE Transactions on Image Processing*, vol. 30, pp. 55–67, 2020.

[8] R. C. Daudt, B. Le Saux, and A. Boulch, "Fully convolutional siamese networks for change detection," in *2018 25th IEEE international conference on image processing (ICIP)*. IEEE, 2018, pp. 4063–4067.

[9] S. Fang, K. Li, and Z. Li, "Changer: Feature interaction is what you need for change detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–11, 2023.

[10] K. Li, X. Cao, and D. Meng, "A new learning paradigm for foundation model-based remote-sensing change detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1–12, 2024.

[11] K. Sohn, D. Berthelot, N. Carlini, Z. Zhang, H. Zhang, C. A. Raffel, E. D. Cubuk, A. Kurakin, and C.-L. Li, "Fixmatch: Simplifying semi-supervised learning with consistency and confidence," *Advances in neural information processing systems*, vol. 33, pp. 596–608, 2020.

[12] L. Yang, L. Qi, L. Feng, W. Zhang, and Y. Shi, "Revisiting weak-to-strong consistency in semi-supervised semantic segmentation," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2023, pp. 7236–7246.

[13] T. DeVries and G. W. Taylor, "Improved regularization of convolutional neural networks with cutout," *arXiv preprint arXiv:1708.04552*, 2017.

[14] S. Yun, D. Han, S. J. Oh, S. Chun, J. Choe, and Y. Yoo, "Cutmix: Regularization strategy to train strong classifiers with localizable features," *CoRR*, vol. abs/1905.04899, 2019. [Online]. Available: http://arxiv.org/abs/1905.04899

[15] Y.-F. Pu, J.-L. Zhou, and X. Yuan, "Fractional differential mask: a fractional differential-based approach for multiscale texture enhancement," *IEEE transactions on image processing*, vol. 19, no. 2, pp. 491–511, 2009.

[16] H. Chen, Z. Qi, and Z. Shi, "Remote sensing image change detection with transformers," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–14, 2021.

[17] W. G. C. Bandara and V. M. Patel, "A transformer-based siamese network for change detection," in *IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2022, pp. 207–210.

[18] W. Hung, Y. Tsai, Y. Liou, Y. Lin, and M. Yang, "Adversarial learning for semi-supervised semantic segmentation," *CoRR*, vol. abs/1802.07934, 2018. [Online]. Available: http://arxiv.org/abs/1802.07934

[19] S. Mittal, M. Tatarchenko, and T. Brox, "Semi-supervised semantic segmentation with high- and low-level consistency," 2019. [Online]. Available: https://arxiv.org/abs/1908.05724

[20] D.-H. Lee *et al.*, "Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks," in *Workshop on challenges in representation learning, ICML*, vol. 3, no. 2. Atlanta, 2013, p. 896.

[21] S. Laine and T. Aila, "Temporal ensembling for semi-supervised learning," *arXiv preprint arXiv:1610.02242*, 2016.

[22] A. Tarvainen and H. Valpola, "Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results," *Advances in neural information processing systems*, vol. 30, 2017.

[23] D. Peng, L. Bruzzone, Y. Zhang, H. Guan, H. Ding, and X. Huang, "Semicdnet: A semisupervised convolutional neural network for change detection in high resolution remote-sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 7, pp. 5891–5906, 2021.

[24] W. G. C. Bandara and V. M. Patel, "Revisiting consistency regularization for semi-supervised change detection in remote sensing images," *arXiv preprint arXiv:2204.08454*, 2022.

[25] B. Mathieu, P. Melchior, A. Oustaloup, and C. Ceyral, "Fractional differentiation for edge detection," *Signal Processing*, vol. 83, no. 11, pp. 2421–2432, 2003, fractional Signal Processing and Applications. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0165168403001944

[26] J. Bai and X. Feng, "Fractional-order anisotropic diffusion for image denoising," *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, vol. 16, pp. 2492–502, 11 2007.

[27] Y. Liu, "Remote sensing image enhancement based on fractional differential," in *2010 International Conference on Computational and Information Sciences*, 2010, pp. 881–884.

[28] S. Ji, S. Wei, and M. Lu, "Fully convolutional networks for multisource building extraction from an open aerial and satellite imagery data set," *IEEE Transactions on geoscience and remote sensing*, vol. 57, no. 1, pp. 574–586, 2018.

[29] K. Li, X. Cao, Y. Deng, J. Song, J. Liu, D. Meng, and Z. Wang, "Semicd-vl: Visual-language model guidance makes better semi-supervised change detector," 2024. [Online]. Available: https://arxiv.org/abs/2405.04788