Paper ID: 1879

CROSS BRANCH FEATURE FUSION DECODER FOR CONSISTENCY REGULARIZATION-BASED SEMI-UPERVISED CHANGE DETECTION

Yan Xing[1], Qi'ao Xu[2], JingCheng Zeng[2], Rui Huang[2]*, Sihua Gao[2], Weifeng Xu[3], YuXiang Zhang[2], Wei Fan[2]

[1]College of Safety Science and Engineering, Civil Aviation University of China, Tianjin, China

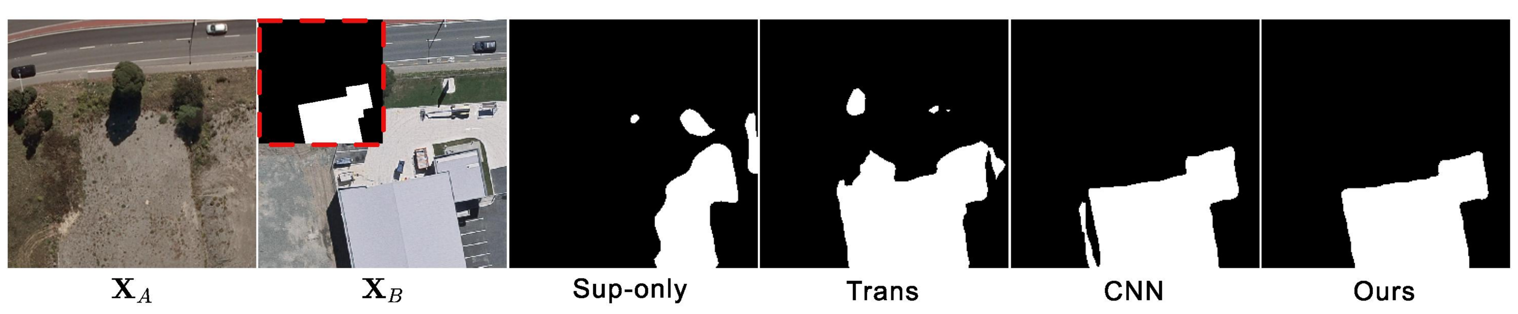
[2]College of Computer Science and Technology, Civil Aviation University of China, Tianjin, China

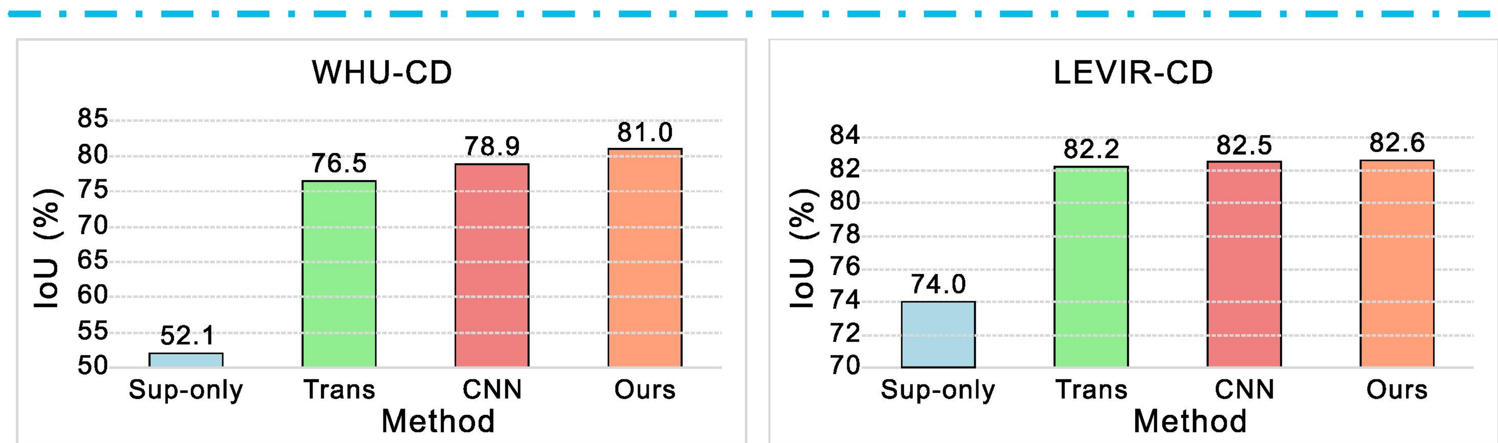
[3]Department of Computer, North China Electric Power University, Beijing, China

ICASSP 2024 KOREA

INTRODUCTION

- **Semi-supervised change detection** (SSCD) utilizes partially labeled data and a large amount of unlabeled data to detect pixel-level changes, which has wide applications in different fields.
- Existing SSCD methods primarily rely on CNN for extracting meaningful features from limited labeled data. However, transformer-based SSCD methods lag behind in performance, particularly in scenarios with scarce labeled data.
- We combine the strengths of transformer and convolution, leveraging both global and local features to enhance feature representation with limited labeled data.





- Motivation: Comparison of SSCD with decoders of transformer, convolution, and ours by 5% labeled data. Sup-only denotes that our method utilizes only this limited labeled data for training.

RESULT

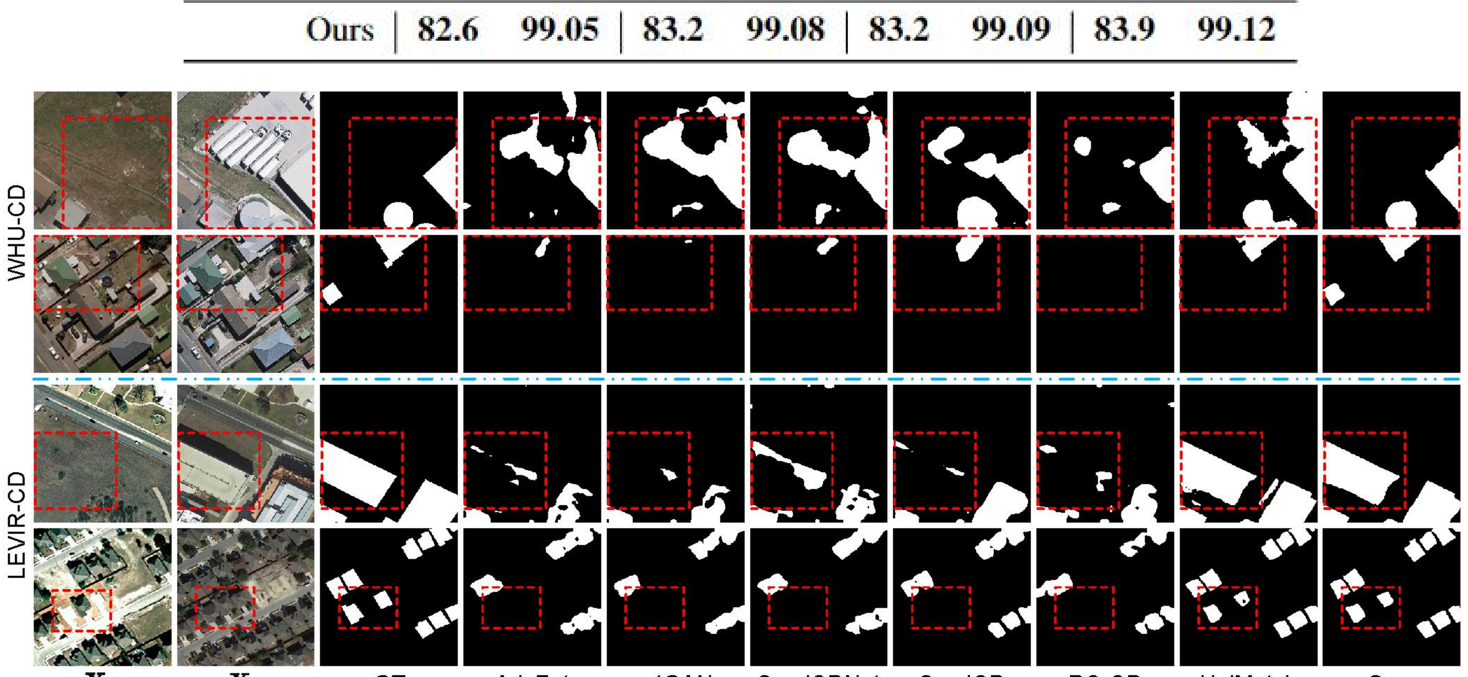
- We compare the proposed method with seven SSCD methods.
- All compared methods are implemented with PyTorch and trained with on the same training sets.
- The training set is divided into labeled and unlabeled data with ratios of [5%, 95%], [10%, 90%], [20%, 80%], and [40%, 60%].

Dataset-01: WHU-CD

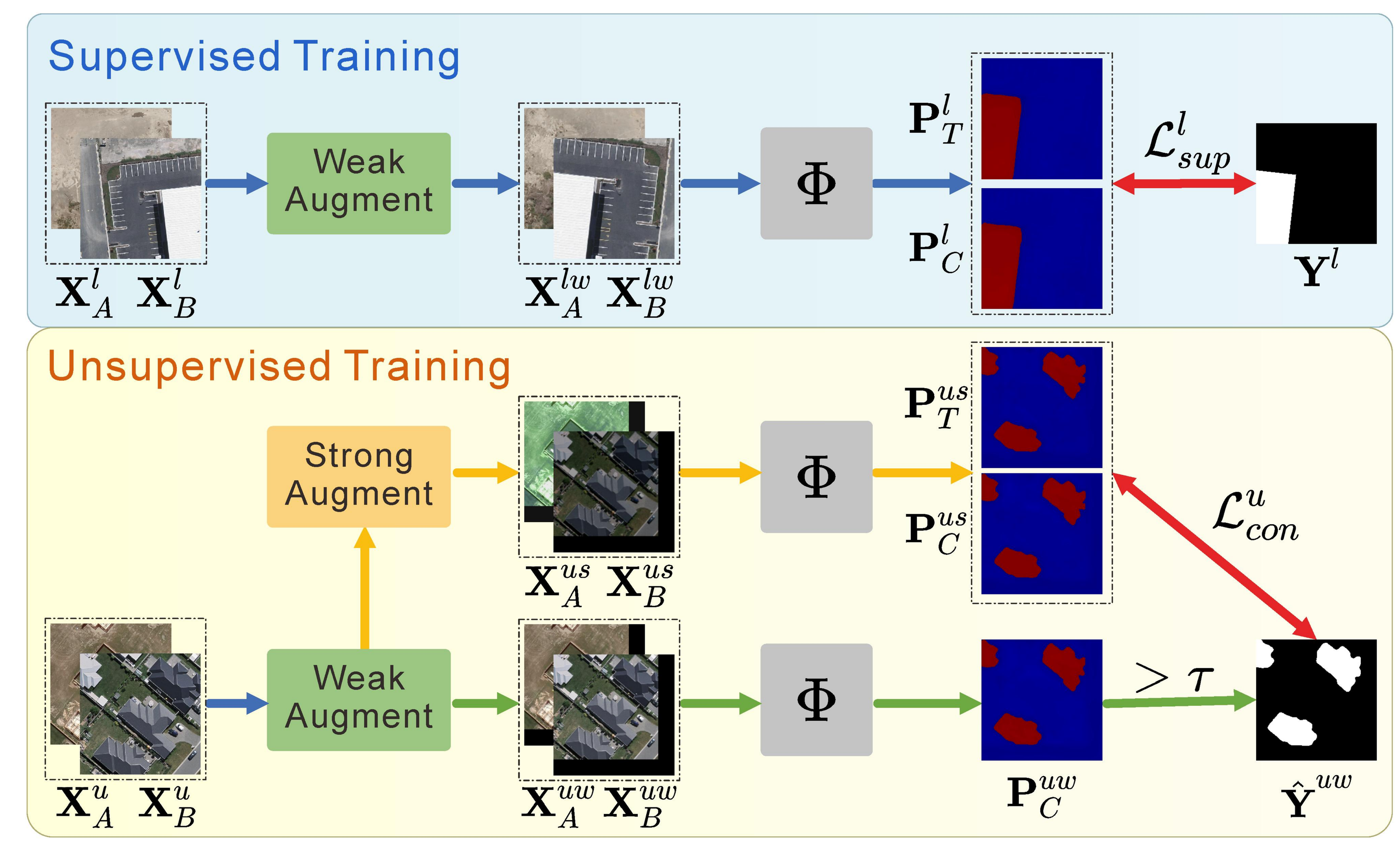
Method	WHU-CD							
	5%		10%		20%		40%	
	IoU	OA	IoU	OA	IoU	OA	IoU	OA
AdvEnt	57.7	97.87	60.5	97.79	69.5	98.50	76.0	98.91
s4GAN	57.3	97.94	58.0	97.81	67.0	98.41	74.3	98.85
SemiCDNet	56.2	97.78	60.3	98.02	69.1	98.47	70.5	98.59
SemiCD	65.8	98.37	68.0	98.45	74.6	98.83	78.0	99.01
RC-CD	57.7	97.94	65.4	98.45	74.3	98.89	77.6	99.02
SemiPTCD	74.1	98.85	74.2	98.86	76.9	98.95	80.8	99.17
UniMatch	78.7	99.11	79.6	99.11	81.2	99.18	83.7	99.29
Ours	**81.0**	**99.20**	**81.1**	**99.18**	**83.6**	**99.29**	**86.5**	**99.43**

Dataset-02: LEVIR-CD

Method	LEVIR-CD							
	5%		10%		20%		40%	
	IoU	OA	IoU	OA	IoU	OA	IoU	OA
AdvEnt	67.1	98.15	70.8	98.38	74.3	98.59	75.9	98.67
s4GAN	66.6	98.16	72.2	98.48	75.1	98.63	76.2	98.68
SemiCDNet	67.4	98.11	71.5	98.42	74.9	98.58	75.5	98.63
SemiCD	74.2	98.59	77.1	98.74	77.9	98.79	79.0	98.84
RC-CD	67.9	98.09	72.3	98.40	75.6	98.60	77.2	98.70
SemiPTCD	71.2	98.39	75.9	98.65	76.6	98.65	77.2	98.74
UniMatch	82.1	99.03	82.8	99.07	82.9	99.07	83.0	99.08
Ours	**82.6**	**99.05**	**83.2**	**99.08**	**83.2**	**99.09**	**83.9**	**99.12**



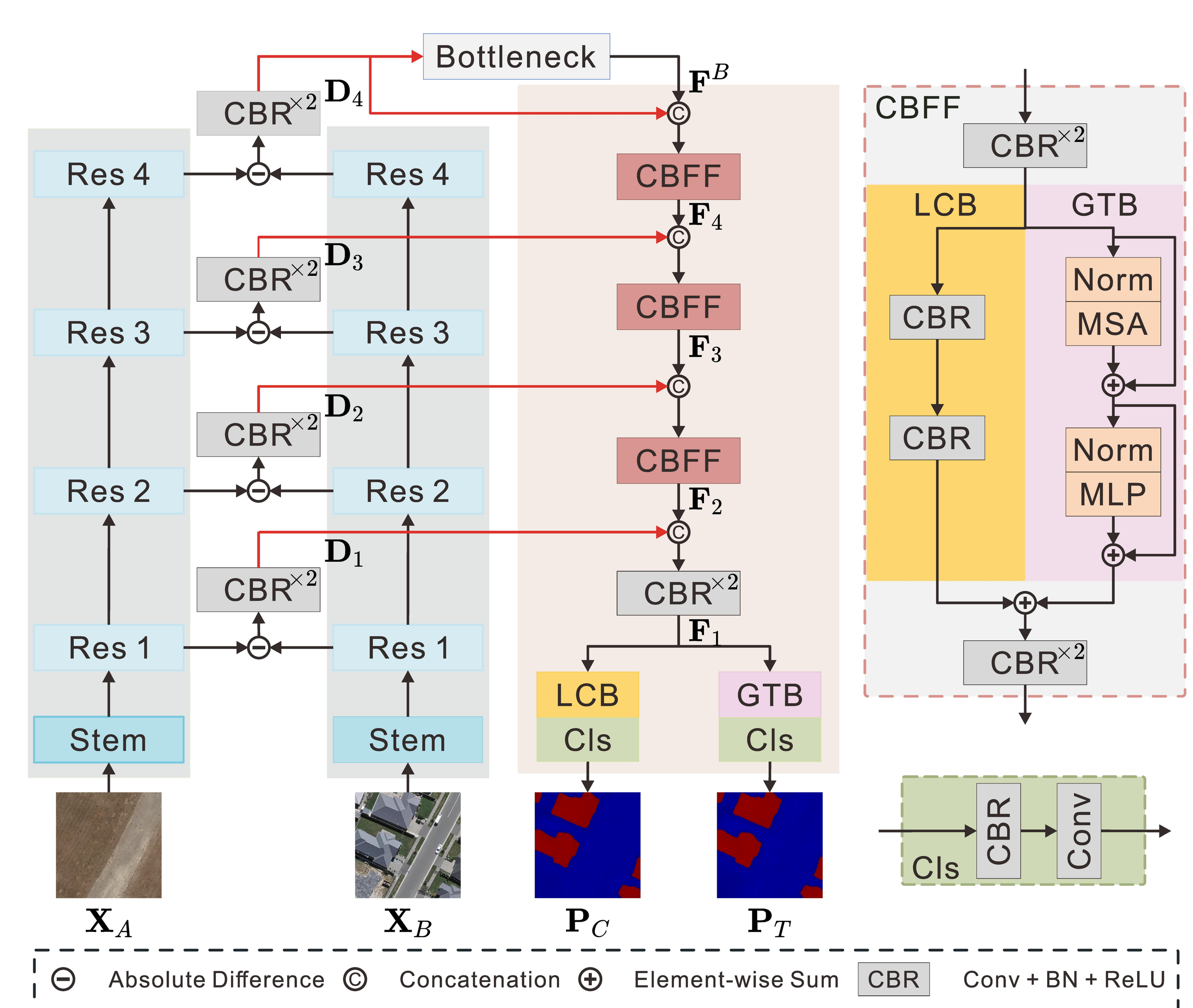
FRAMEWORK



- In the supervised training phase, we use labeled training data to train the change detection network Φ.
- In the unsupervised training phase, we employ strong-to-weak consistency regularization, utilizing the change map generated from weakly augmented input to create pseudo-labels.

Change Detection Network

- Firstly, we apply Siamese ResNet-50 to extract basic features, and calculate the difference features.
- Secondly, to extract richer feature information, Atrous Spatial Pyramid Pooling (ASPP) is used in the Bottleneck.
- Finally, we propose the Cross Branch Feature Fusion (CBFF) decoder, incorporating a Local Convolutional Branch (LCB) and a Global Transformer Branch (GTB), to generate accuate change maps.



CONCLUSION

- We introduce a new decoder, Cross Branch Feature Fusion (CBFF), which consists of two branches: a local convolutional branch and a global transformer branch.
- Using CBFF, we have built a SSCD model based on a strong-to-weak consistency strategy.
- Experiments on two benchmark datasets demonstrate that our method outperforms seven state-of-the-art SSCD methods.

Contact e-mail: 2021052074@cauc.edu.cn

The 2024 IEEE International Conference on Acoustics, Speech, and Signal Processing.

ICASSP 2024