AdaptFormer: An Adaptive Hierarchical Semantic Approach for Change Detection on Remote Sensing Images

edges and shapes, proving instrumental for the identification of minor changes. However, this stage is limited in its ability to unravel more intricate details. Advancing to stage 2, the model deepens its semantic exploration to intermediate levels, thereby refining its detection capabilities to encompass moderate changes through the discernment of more complex shapes and patterns, albeit with remaining challenges in capturing the finest nuances. The culmination occurs in stage 3, where an intensive dive into deep semantic realms enables the model to grasp comprehensive contextual relationships and substantial structural shifts, thus extending its detection acumen to substantial changes. This graduated approach aligns closely with GT data, indicating minimal discrepancies and highlighting the model's adaptability and scalability. The framework effectively addresses the diverse requirements of CD in RS imagery, accommodating changes across a wide range of magnitudes.

3) Error Maps: We employ error maps as a visual technique to rigorously assess the effectiveness of CD on RS images, highlighting discrepancies between predicted and true values. Fig. 8 elucidates the confidence visualization results for various CD models when applied to the LEVIR-CD dataset. Primarily, the majority of the figures—columns 1 to 6—display error analysis from several mainstream models on their respective test images, whereas the concluding column distinctively represents the outcomes of our AdaptFormer approach. A unique measurement system was employed wherein the differences between the model outputs and the GT were visualized on a scale from 0 to 1. A shade closer to blue (indicating a value nearer to 0) epitomizes high confidence in detection, while a hue leaning toward red (signifying a value approaching 1) designates lesser assurance.

In this visualization, AdaptFormer's adeptness is consistently evident across various test images. Particularly notable is its proficiency in small object detection, where the near absence of the red hue in the first row suggests its enhanced capability to identify scattered minor entities. For medium-sized objects, many contemporary models manifest continuous red zones, indicating lapses in their detection confidence. In stark contrast, AdaptFormer's results, especially in the fourth row, underscore its superiority by almost flawlessly identifying these areas. This prowess extends to large object detection as well, as observed in the fifth row, where the dearth of red regions in our method's visualization stands testament to its exceptional confidence and accuracy in recognizing substantial object changes.

V. Conclusion

This study presents AdaptFormer, a groundbreaking solution to CD in RS imagery. Distinctly adaptive, AdaptFormer systematically interprets hierarchical semantics, tailoring its operations across three depth levels: simple techniques for shallow semantics, spatial data assimilation for medium details, and cascaded depthwise attention for in-depth insights. Our experimental evaluations, particularly on the LEVIR-CD and DSIFN-CD datasets, showcase AdaptFormer's superior accuracy and performance over other models, underscore its potential in applications from urban development to environmental surveillance. In essence, AdaptFormer emerges as a benchmark in CD, ushering in new avenues for future research and development in the domain. In future work, we aim to enhance the computational efficiency of the AdaptFormer model to better support real-time analysis, while maintaining its accuracy and effectiveness in CD tasks.

References

[1] A. Singh, "Review article digital change detection techniques using remotely-sensed data," *Int. J. Remote Sens.*, vol. 10, no. 6, pp. 989–1003, Jun. 1989.

[2] Z. Lv et al., "Land cover change detection with heterogeneous remote sensing images: Review, progress, and perspective," *Proc. IEEE*, vol. 110, no. 12, pp. 1976–1991, Dec. 2022.

[3] L. Ding, H. Guo, S. Liu, L. Mou, J. Zhang, and L. Bruzzone, "Bi-temporal semantic reasoning for the semantic change detection in HR remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5620014.

[4] H. Chen, W. Li, S. Chen, and Z. Shi, "Semantic-aware dense representation learning for remote sensing image change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5630018.

[5] Y. Liang, C. Zhang, and M. Han, "RaSRNet: An end-to-end relation-aware semantic reasoning network for change detection in optical remote sensing images," *IEEE Trans. Instrum. Meas.*

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

[7] J. Z. Xu, W. Lu, Z. Li, P. Khaitan, and V. Zaytseva, "Building damage detection in satellite imagery using convolutional neural networks," 2019, *arXiv:1910.06444*.

[8] P. de Bem, O. de Carvalho Junior, R. F. Guimar aes, and R. T. Gomes, "Change detection of deforestation in the Brazilian Amazon using Landsat data and convolutional neural networks," *Remote Sens.*, vol. 12, no. 6, p. 901, Mar. 2020.

[9] L. Bruzzone and D. F. Prieto, "Automatic analysis of the difference image for unsupervised change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 3, pp. 1171–1182, May 2000.

[10] Y. Pang et al., "Improved crop row detection with deep neural network for early-season maize stand count in UAV imagery," *Comput. Electron. Agricult.*, vol. 178, Nov. 2020, Art. no. 105766.

[11] M. Liu, Z. Chai, H. Deng, and R. Liu, "A CNN-transformer network with multiscale context aggregation for fine-grained cropland change detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 4297–4306, 2022.

[12] G. Pei and L. Zhang, "Feature hierarchical differentiation for remote sensing image change detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[13] Y. Feng, J. Jiang, H. Xu, and J. Zheng, "Change detection on remote sensing images using dual-branch multilevel intertemporal network," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 4401015.

[14] T. Lei et al., "Ultralightweight spatial–spectral feature cooperation network for change detection in remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023.

[15] W. G. C. Bandara and V. M. Patel, "A transformer-based Siamese network for change detection," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2022, pp. 207–210.

[16] Y. Pang et al., "Slim UNETR: Scale hybrid transformers to efficient 3D medical image segmentation under limited computational resources," *IEEE Trans. Med. Imag.*, vol. 43, no. 3, pp. 994–1005, Mar. 2024.

[17] Y. Zhang, Y. Yuan, Y. Feng, and X. Lu, "Hierarchical and robust convolutional neural network for very high-resolution remote sensing object detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 5535–5548, Aug. 2019.

[18] H. Sun, S. Li, X. Zheng, and X. Lu, "Remote sensing scene classification by gated bidirectional network," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 1, pp. 82–96, Jan. 2020.

[19] Z. Chen et al., "A new approach for detecting urban centers and their spatial structure with nighttime light remote sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 11, pp. 6305–6319, Nov. 2017.

[20] X. Zhang, S. Cheng, L. Wang, and H. Li, "Asymmetric cross-attention hierarchical network based on CNN and transformer for bitemporal remote sensing images change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023.

[21] H. Yao, R. Qin, and X. Chen, "Unmanned aerial vehicle for remote sensing applications—A review," *Remote Sens.*, vol. 11, no. 12, p. 1443, Jun. 2019.
[22] H. Yin et al., "Attention-guided Siamese networks for change detection in high resolution remote sensing images," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 117, Mar. 2023, Art. no. 103206.
[23] Z. Wang et al., "Toward learning joint inference tasks for IASS-MTS using dual attention memory with stochastic generative imputation," *IEEE Trans. Neural Netw. Learn. Syst*.
[24] R. S. Lunetta, J. F. Knight, J. Ediriwickrema, J. G. Lyon, and L. D. Worthy, "Land-cover change detection using multi-temporal MODIS NDVI data," in *Geospatial Information Handbook for Water Resources and Watershed Management*, vol. 2. Boca Raton, FL, USA: CRC Press, 2022, pp. 65–88.
[25] S. I. Toure, D. A. Stow, H.-C. Shih, J. Weeks, and D. Lopez-Carr, "Land cover and land use change analysis using multi-spatial resolution data and object-based image analysis," *Remote Sens. Environ.*, vol. 210, pp. 259–268, Jun. 2018.
[26] S. Schlaffer, P. Matgen, M. Hollaus, and W. Wagner, "Flood detection from multi-temporal SAR data using harmonic analysis and change detection," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 38, pp. 15–24, Jun. 2015.
[27] T. Peterson, C. Folland, G. Gruza, W. Hogg, A. Mokssit, and N. Plummer, *Report on the Activities of the Working Group on Climate Change Detection and Related Rapporteurs*. World Meteorological Organization Geneva, 2001.
[28] K. Rokni, A. Ahmad, A. Selamat, and S. Hazini, "Water feature extraction and change detection using multitemporal Landsat imagery," *Remote Sens.*, vol. 6, no. 5, pp. 4173–4189, 2014.
[29] M. Hussain, D. Chen, A. Cheng, H. Wei, and D. Stanley, "Change detection from remotely sensed images: From pixel-based to object-based approaches," *ISPRS J. Photogramm. Remote Sens.*, vol. 80, pp. 91–106, Jun. 2013.
[30] R. C. Daudt, B. Le Saux, and A. Boulch, "Fully convolutional Siamese networks for change detection," in *Proc. 25th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2018, pp. 4063–4067.
[31] Y. Liu, C. Pang, Z. Zhan, X. Zhang, and X. Yang, "Building change detection for remote sensing images using a dual-task constrained deep Siamese convolutional network model," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 5, pp. 811–815, May 2021.
[32] J. Chen et al., "DASNet: Dual attentive fully convolutional Siamese networks for change detection in high-resolution satellite images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 1194–1206, 2020.
[33] H. Chen, Z. Qi, and Z. Shi, "Remote sensing image change detection with transformers," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–14, 2021.
[34] S. Fang, K. Li, and Z. Li, "Changer: Feature interaction is what you need for change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023.
[35] D. Hendrycks and K. Gimpel, "Gaussian error linear units (GELUs)," 2016, *arXiv:1606.08415*.
[36] A. F. Agarap, "Deep learning using rectified linear units (ReLU)," 2018, *arXiv:1803.08375*.
[37] P.-T. de Boer, D. P. Kroese, S. Mannor, and R. Y. Rubinstein, "A tutorial on the cross-entropy method," *Ann. Oper. Res.*, vol. 134, no. 1, pp. 19–67, Feb. 2005.
[38] C. Zhang et al., "A deeply supervised image fusion network for change detection in high resolution bi-temporal remote sensing images," *ISPRS J. Photogramm. Remote Sens.*, vol. 166, pp. 183–200, Aug. 2020.
[39] Q. Shi, M. Liu, S. Li, X. Liu, F. Wang, and L. Zhang, "A deeply supervised attention metric-based network and an open aerial image dataset for remote sensing change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022.
[40] N. Chinchor and B. Sundheim, "MUC-5 evaluation metrics," in *Proc. 5th Conf. Message Understand.*, 1993, pp. 1–10.
[41] X. Song, Z. Hua, and J. Li, "Remote sensing image change detection transformer network based on dual-feature mixed attention," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022.
[42] J. Huang, Q. Shen, M. Wang, and M. Yang, "Multiple attention Siamese network for high-resolution image change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2021.
[43] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Mari, J. L. Rojo-Alvarez, and M. Martinez-Ramon, "Kernel-based framework for multitemporal and multisource remote sensing data classification and change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 6, pp. 1822–1835, Jun. 2008.
[44] H. Chen, F. Pu, R. Yang, R. Tang, and X. Xu, "RDP-net: Region detail preserving network for change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5635010.
[45] D. M. W. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation," 2020, *arXiv:2010.16061*.
[46] S. H. Khan, X. He, F. Porikli, and M. Bennamoun, "Forest change detection in incomplete satellite images with deep neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 9, pp. 5407–5423, Sep. 2017.
[47] S. Fang, K. Li, J. Shao, and Z. Li, "SNUNet-CD: A densely connected Siamese network for change detection of VHR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2021.
[48] M. Lin, G. Yang, and H. Zhang, "Transition is a process: Pair-to-video change detection networks for very high resolution remote sensing images," *IEEE Trans. Image Process.*, vol. 32, pp. 57–71, 2022.

Teng Huang received the Ph.D. degree in computer science from Beihang University, Beijing, China, in 2019.

Currently, he is working at the Institute of Artificial Intelligence and Blockchain, Guangzhou University, Guangzhou, China. His research interests include computer vision, RS, and blockchain.

Yile Hong received the B.S. degree from Huizhou University, Huizhou, China, in 2022. He is currently pursuing the master's degree with Guangzhou University, Guangzhou, China.

His research areas of interest span computer vision, remote sensing and smart contract.

Yan Pang (Member, IEEE) received the Ph.D. degree from the University of Colorado, Denver, CO, USA, in 2021.

From April 2021 to May 2022, he was a Machine Learning Scientist at Moffett AI, Los Altos, CA, USA. From August 2018 to May 2021, he was an Instructor with the Department of Electrical Engineering, University of Colorado Denver, and the Department of Electrical Engineering Technology, Metropolitan State University of Denver, Denver. His research interests span machine learning, computer vision, and efficient deep learning.

Jiaming Liang received the B.S. degree from Foshan University, Foshan, China, in 2021. He is currently pursuing the master's degree with Guangzhou University, Guangzhou, China.

His research areas of interest span computer vision, remote sensing, and medical image analysis. As a master's student, he focuses on advancing the field through innovative research and applications, leveraging his background in electronic technology and a keen interest in the computer vision domain.

Jie Hong received the B.S. degree from Nanjing Institute of Technology, Nanjing, China, in 2021. He is currently pursuing the master's degree with Guangzhou University, Guangzhou, China.

His research areas of interest span computer vision, remote sensing, and medical image analysis.

Lin Huang is a Professor with the Metropolitan State University of Denver, Denver, CO, USA. Her research interests include the areas of biometrics, pattern recognition, signal processing, computer vision, machine learning, embedded system design, and VLSI.

Dr. Huang has been an Active Member of International Conference on Machine Learning and Computing (ICMLC) as a Conference Chair/Keynote Speaker/Session Chair, since 2011. She is a Program Chair of the International Conference on Signal Processing (ICOSP 2023 & 2024). She has been an Editor Board Member of several international journals, for example, she has been the Editor-in-Chief of the *International Journal of Machine Learning and Computing* (IJMLC), since 2012. And she has been reviewing papers on regular basis for some conferences and journals since 2008.

Yuan Zhang received the B.S., M.S., and Ph.D. degrees from Beijing University of Aeronautics and Astronautics, Beijing, China, in 2004, 2008, and 2017, respectively.

He was a joint Ph.D. Student with German Aerospace Center, Cologne, Germany, and the Technical University of Munich, Munich, Germany, from 2015 to 2016. He is an Associate Professor with the North China University of Technology, Beijing, China. His main research direction is synthetic aperture radar (SAR) imaging.

Yan Jia received the double M.S. degree in telecommunications engineering and computer application technology from Politecnico di Torino, Turin, Italy, and Henan Polytechnic University, Jiaozuo, China, in 2013, and the Ph.D. degree in electronics engineering from Politecnico di Torino in 2017.

Now she is working with Nanjing University of Posts and Telecommunications. In 2013, she was with the Department of Electronics and Telecommunications, Politecnico di Torino, where she performed research on GNSS system construction and GNSS antenna analysis. In 2014, she worked with the SMAT project, mainly focusing on the retrieval of soil moisture and vegetation biomass content by GNSS-R. Her research interests include microwave remote sensing, soil moisture retrieval, and Global Navigation Satellite System Reflectometry (GNSS-R) applications to land remote sensing and antenna design.

Patrizia Savi (Senior Member, IEEE) received the Laurea degree in electronic engineering from the Politecnico di Torino, Turin, Italy, in 1985.

In 1986, she was a Consultant with Alenia, Caselle Torinese, Italy, where she conducted research on the analysis and design of dielectric radomes. From 1987 to 1998, she was a Researcher with Italian National Research Council, Rome, Italy.