A CHANGE DETECTION REALITY CHECK

Isaac Corley
University of Texas at San Antonio
San Antonio, TX, USA
isaac.corley@utsa.edu

Caleb Robinson & Anthony Ortiz
Microsoft AI for Good Research Lab
Redmond, WA, USA
{caleb.robinson,anthony.ortiz}@microsoft.com

ABSTRACT

In recent years, there has been an explosion of proposed change detection deep learning architectures in the remote sensing literature. These approaches claim to offer state-of-the-art performance on different standard benchmark datasets. However, has the field truly made significant progress? In this paper we perform experiments which conclude a simple U-Net segmentation baseline without training tricks or complicated architectural changes is still a top performer for the task of change detection. All experiments are openly available at github.com/isaaccorley/a-change-detection-reality-check.

1 INTRODUCTION

The task of change detection from remotely sensed imagery is a canonical and important problem that allows us to analyze how our planet changes over time. It is crucial that change detection methods are highly accurate given their primary applications for building damage assessment disaster response (Gupta et al., 2019a; Sublime & Kalinicheva, 2019) and monitoring of the environment (Cambrin et al., 2024; Watch, 2002). The machine learning and remote sensing fields are currently experiencing their ImageNet (Deng et al., 2009) benchmarking era with new proposed model architectures claiming incremental performance gains on standard benchmark datasets at a rapid rate. For example, the Changer paper (Fang et al., 2023) – first preprint released on Sep. 2022 – is cited by 56 papers, the majority of them proposing a new architecture for change detection while also not releasing any open-source code or model checkpoints (Ren et al., 2024; Zhang et al., 2024b; Zhou et al., 2024; Li & Wu; Liu et al., 2024; Huang et al., 2024; Lin et al., 2024; Wang et al., 2024; Lu et al., 2024; Tan et al., 2024; Xu et al., 2024; Yan et al., 2024; Peng et al., 2023; Wan et al., 2023; Zhao et al., 2023c; Liu et al., 2023c; Quan et al., 2023; Wang et al., 2023; Liu et al., 2023b; Fazry et al., 2023). Many of these proposed methods contain complicated architectural layers and modules specifically designed to better handle the bi-temporal image change detection task (Bai et al., 2023; Zhang et al., 2024a; Yuan et al., 2024; Li et al., 2024; Pang et al., 2024; Zheng et al., 2024; Zhao et al., 2023a; Li et al., 2023; Yan et al., 2023; Chen et al., 2023; Liu et al., 2023a; Zhao et al., 2023b). Notably many of these papers are peer-reviewed and published at reputable conferences and journals. However, this begs the question if these new methods are actually improvements over generic segmentation architectures or have the benchmarks just been poorly executed? In this paper, we seek to answer this question.

Background When proposing a new model architecture it is common to perform a comparison to prior works through benchmark experiments to show statistically significant improvement. However over time, comparisons become less fair due to differences in training methodologies that are misconstrued as achieving state-of-the-art performance. Put simply, baseline experiments are often underpowered. This has been shown to be prevalent in many areas of research such as deep metric learning (Musgrave et al., 2020), unsupervised domain adaption (Musgrave et al., 2021), image classification (Bello et al., 2021), deep reinforcement learning (Henderson et al., 2018), point cloud classification (Uy et al., 2019), and video recognition (Du et al., 2021). Furthermore, these analyses commonly conclude that simple baselines outperform complicated and task-specific architectures. Most recently, Gerard et al. (2024) discovered that a generic architecture like U-Net, without training or evaluation tricks, is still competitive on the xBD dataset from the xView2 challenge (Gupta et al., 2019b).

When analyzing relevant change detection papers and source code for benchmarking we find that it is common for authors to just compare to metrics reported in prior literature rather than re-running experiments with the prior methods in a consistent training setup. Further, we find that recent work often changes multiple aspects of the experimental setup, including the training routine (optimization methods, learning rate schedule, etc.) and loss functions, beyond just the proposed model architecture. This is problematic as any observed improvements on the benchmark dataset could be due to any combination of the new architecture, training routine, or loss function. This has the potential to result in unfair comparisons, especially if the improvement in quantitative performance is very small.

In this paper we revisit bi-temporal change detection benchmarks with simple baselines to get an overview on progress. Specifically, we experiment with a simple semantic segmentation U-Net architecture, and siamese network variants (Koch et al., 2015), to explore how these baselines perform against the latest state-of-the-art change detection methods. We find that this baseline architecture, from 2015, is a top performer on change detection benchmark datasets.

2 Methods

State-of-the-Art Models In our experiments, in addition to compiling results from many prior change detection works, we retrain and evaluate the following state-of-the-art methods for change detection: BIT (Chen et al., 2021b) is a transformer-based siamese network architecture which uses a shared convolutional backbone to extract image features and transformer encoder decoder networks to perform change detection. ChangeFormer (Bandara & Patel, 2022b) is an end-to-end transformer-based siamese network architecture for change detection. TinyCD (Codegoni et al., 2023) is a change detection architecture which uses an EfficientNet (Tan & Le, 2019) backbone to extract convolutional features to feed to a custom attention-based decoder network.

Our Baseline Models (Daudt et al., 2018) proposed three fully-convolutional (FC) architectures for change detection. The first is Early Fusion (FC-EF) which is an encoder-decoder style architecture with the change detection image pair concatenated as input. The other methods contain shared siamese encoders which either concatenate intermediate feature maps (FC-Siam-Conc) or take the difference (FC-Siam-Diff). While these networks are similar to the U-Net architecture (Ronneberger et al., 2015), the original implementation is customized and hardcoded to be lightweight in parameter count and thus are unable to take advantage of pretrained encoder backbones like ResNet (He et al., 2016). These methods are commonly seen in change detection benchmarks being compared to methods which utilize pretrained ImageNet backbones.

In our experiments, we generalize these implementations to use the standard U-Net framework which is able to take advantage of numerous ImageNet pretrained backbones. To avoid confusion with the original implementations of (Daudt et al., 2018), we refer to these architectures in our experiments as simply `U-Net`, `U-Net SiamConc`, and `U-Net SiamDiff`. Specifically, we utilize the implementations of these models in the TorchGeo (Stewart et al., 2022) library for reproducibility. We perform benchmarks using the ResNet-50 and EfficientNet-B4 (Tan & Le, 2019) backbones.

2.1 Datasets

For comparisons of change detection architectures we utilize the following benchmark datasets:

LEVIR-CD (Chen & Shi, 2020) A binary change detection dataset containing 637 high resolution (0.5m) 1024×1024 image pairs extracted from Google Earth. We utilize the splits provided with the dataset and, following other change detection papers, we convert the images to non-overlapping 256×256 patches.

WHU-CD (Ji et al., 2018) A binary change detection dataset containing one pair of high-resolution (0.075m) aerial images of size 32507×15354. We utilize the train and test splits provided with the dataset and convert the images to non-overlapping 256×256 patches. During training runs we randomly split a 10% holdout set from the train set to use as a validation set.

Table 1: Comparison of state-of-the-art and change detection architectures to a U-Net baseline on the LEVIR-CD dataset. We report the test set precision, recall, and F1 metrics of the positive change class. For the baseline experiments we perform 10 runs while varying random the seed and report metrics from the highest performing run. All other metrics are taken from their respective papers. The top performing methods are highlighted in bold. Gray rows indicate our baseline U-Net and siamese encoder variants.

Model	Backbone	Precision	Recall	F1
FC-EF (Daudt et al., 2018)	-	86.91	80.17	83.40
FC-Siam-Conc (Daudt et al., 2018)	-	91.99	76.77	83.69
FC-Siam-Diff (Daudt et al., 2018)	-	89.53	83.31	86.31
DTCDSCN (Liu et al., 2020)	SE-Resnet34	88.53	86.83	87.67
STANet (Chen & Shi, 2020)	ResNet-18	83.81	**91.00**	87.26
CDNet (Chen et al., 2021a)	ResNet-18	91.60	86.50	89.00
BIT (Chen et al., 2021b)	ResNet-18	89.24	89.37	89.31
ChangeFormer (Bandara & Patel, 2022b)	MiT-b1	92.59	89.68	91.11
Tiny-CD (Codegoni et al., 2023)	EfficientNet-b4	92.68	89.47	91.05
ChangerVanilla (Fang et al., 2023)	ResNet-18	92.66	89.60	91.10
ChangerEx (Fang et al., 2023)	ResNet-18	92.97	90.61	**91.77**
`U-Net` (Ronneberger et al., 2015)	EfficientNet-b4	92.69	87.16	89.25
`U-Net` (Ronneberger et al., 2015)	ResNet-50	91.97	89.78	90.38
`U-Net SiamConc`	ResNet-50	92.87	89.48	90.41
`U-Net SiamDiff`	ResNet-50	**93.21**	89.50	90.46

3 EXPERIMENTS

3.1 BASELINE TRAINING DETAILS

Throughout our baseline training experiments we use the same hyperparameter setup from the BIT (Chen et al., 2021b) source which consists of batch size of 8, 200 epochs, stochastic gradient descent (SGD) optimizer with an initial learning rate $\gamma = 0.01$, momentum $\mu = 0.9$, weight decay $\alpha = 5E - 4$, and a linear decaying scheduler. We train each model by optimizing a cross entropy loss on a multiclass segmentation output where the number of classes is 2 for binary change detection. We select the checkpoint with the top performance based on the validation loss and do not perform any early stopping. We note that this setup results in a number of iterations which allows each model to converge.

During training we use augmentations consisting of random horizontal and vertical flips with probability $p = 0.5$ and random resize crop with scale in the range $[0.8, 1.0]$, aspect ratio of 1, and $p = 1$. For image normalization we simply rescale the images to the range $[-1, 1]$ following the BIT (Chen et al., 2021b) source.

3.2 UPDATED WHU-CD BENCHMARKS

Many papers utilize a preprocessed and randomly split version of the WHU-CD dataset created in (Bandara & Patel, 2022a) for change detection benchmarks. However these sets have been known to introduce data leakage (Bandara, 2023) – $\approx 85\%$ of the test set is included in the train set due to a bug in the preprocessing scripts – which makes them impossible to use to benchmark methods. We perform a new benchmark using the original train and test set splits from the WHU-CD dataset and retrain the BIT, ChangeFormer, and TinyCD models for comparison. We use the same experimental settings described in Section 3.1.

We find that performance of some models, particularly the transformer based methods BIT and ChangeFormer, can vary significantly over different random seeds, therefore we train each model on each dataset for runs with 10 different seeds. For comparisons we select the results from the highest performing run as well as report the average and standard deviation over runs for transparency. Any additional parameters can be found in our open source implementation.

Table 2: Experimental results on the WHU-CD dataset. We retrain several state-of-the-art methods using the original dataset's train/test splits instead of the commonly used randomly split preprocessed version created in Bandara & Patel (2022a). We find that these state-of-the-art methods are outperformed by a U-Net baseline. We report the test set precision, recall, F1, and IoU metrics of the positive change class. For each run we select the model checkpoint with the lowest validation set loss. We provide metrics averaged over 10 runs with varying random seed as well as the best seed. Gray rows indicate our baseline U-Net and siamese encoder variants.

Model	Backbone	F1	Pre.	Rec.	IoU
Averaged Over 10 Seeds					
ChangeFormer	MiT-b1	75.65 ± 1.58	77.06 ± 3.22	74.67 ± 1.97	61.60 ± 2.05
TinyCD	EfficientNet-b4	78.53 ± 1.28	80.15 ± 2.49	77.56 ± 2.13	65.52 ± 1.72
BIT	ResNet-18	72.67 ± 2.69	70.30 ± 6.36	76.84 ± 4.53	58.06 ± 3.24
`U-Net`	ResNet-50	81.85 ± 1.32	83.72 ± 2.65	80.39 ± 2.32	69.96 ± 1.83
`U-Net SiamConc`	ResNet-50	81.33 ± 1.08	79.30 ± 2.78	**84.19 ± 1.55**	69.40 ± 1.52
`U-Net SiamDiff`	ResNet-50	**82.02 ± 1.48**	**83.82 ± 3.80**	80.92 ± 2.51	**70.29 ± 2.00**
Best Seed					
ChangeFormer	MiT-b1	77.75	82.60	78.57	64.22
TinyCD	EfficientNet-b4	78.53	80.15	77.56	65.52
BIT	ResNet-18	77.68	78.58	82.13	64.34
`U-Net`	ResNet-50	**84.17**	**88.65**	83.08	**73.23**
`U-Net SiamConc`	ResNet-50	82.75	83.69	**86.56**	71.15
`U-Net SiamDiff`	ResNet-50	84.01	88.56	85.63	73.02

4 DISCUSSION

Results In Table 1 we provide experimental results for our baseline on the LEVIR-CD dataset. We compare to prior work by compiling results from their respective papers. While the compiled results represent the best reported metrics for each previous method, it still is evident that a simple U-Net baseline is a top performer. Table 2 provides additional updated experimental results for the WHU-CD dataset and compares several state-of-the-art change detection architectures. Again, we can see that the baselines outperform all other methods.

Regarding the siamese variants of our U-Net baseline (`U-Net SiamDiff` and `U-Net SiamConc`), our experimental results indicate that processing each image in the bi-temporal pair with a shared encoder results in reasonable gains and is a promising approach. Methods such as Changer (Fang et al., 2023) also conclude that additional feature interaction is key for achieving performance improvements for change detection.

Limitations Due to the fast pace of the field of change detection, with new architectures being proposed almost weekly, it is difficult to benchmark each new method. Furthermore, we do not compare to methods which propose a technique, e.g. a loss function, which is dependent on the architecture or methods, or to methods which do not publish open source code to reproduce the experimental results.

We note that there exists several other common change detection datasets such as DSIFN-CD (Zhang et al., 2020), S2Looking (Shen et al., 2021), SECOND (Yang et al., 2020), LEVIR-CD+ (Shen et al., 2021), and xBD (Gupta et al., 2019b). While we do not benchmark against these, we leave this for future work.

5 CONCLUSION AND NEXT STEPS

In this paper we analyzed whether the field of change detection has actually made significant improvements on benchmark datasets in recent years. We conclude that many claimed improvements are questionable by demonstrating that a simple baseline of U-Net is still a top-performing method. To be clear, this is not an issue unique to change detection; other machine learning fields such as language modeling are finding it crucial to standardize fair benchmarking (Srivastava et al., 2022) when new methods are rapidly proposed. To mitigate this, we recommend utilizing and contributing proposed models to libraries and projects such as OpenCD (Li, 2022), GEO-Bench (Lacoste et al., 2023), and TorchGeo (Stewart et al., 2022) which standardize datasets and trainers for reliable benchmarking of remote sensing tasks. As this field has important downstream applications, we hope our results motivate the community to perform more reliable benchmarks of performance so that realistic advancements in change detection can be achieved.

ACKNOWLEDGMENTS

We thank Jonathan Lwowski, Conor Wallace, Robin Cole, Sebastian Gerard, and Juan M. Lavista Ferres for their valuable feedback.

REFERENCES

Ting Bai, Le Wang, Dameng Yin, Kaimin Sun, Yepei Chen, Wenzhuo Li, and Deren Li. Deep learning for change detection in remote sensing: a review. *Geo-spatial Information Science*, 26 (3):262–288, 2023.

Chaminda Bandara. Changeformer_cd source code data leakage issue 1. https://github.com/wgcban/ChangeFormer/issues/52, 2023.

Wele Gedara Chaminda Bandara and Vishal M Patel. Revisiting consistency regularization for semi-supervised change detection in remote sensing images. *arXiv preprint arXiv:2204.08454*, 2022a.

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. IEEE, 2022b.

Irwan Bello, William Fedus, Xianzhi Du, Ekin Dogus Cubuk, Aravind Srinivas, Tsung-Yi Lin, Jonathon Shlens, and Barret Zoph. Revisiting resnets: Improved training and scaling strategies. *Advances in Neural Information Processing Systems*, 34:22614–22627, 2021.

Daniele Rege Cambrin, Luca Colomba, and Paolo Garza. Cabuar: California burned areas dataset for delineation. *arXiv preprint arXiv:2401.11519*, 2024.

Hao Chen and Zhenwei Shi. A spatial-temporal attention-based method and a new dataset for remote sensing image change detection. *Remote Sensing*, 12(10):1662, 2020.

Hao Chen, Wenyuan Li, and Zhenwei Shi. Adversarial instance augmentation for building change detection in remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1–16, 2021a.

Hao Chen, Zipeng Qi, and Zhenwei Shi. Remote sensing image change detection with transformers. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–14, 2021b.

Hao Chen, Haotian Zhang, Keyan Chen, Chenyao Zhou, Song Chen, Zhengxia Zhou, and Zhenwei Shi. Remote sensing image change detection towards continuous bitemporal resolution differences. *arXiv preprint arXiv:2305.14722*, 2023.

Andrea Codegoni, Gabriele Lombardi, and Alessandro Ferrari. Tinycd: a (not so) deep learning model for change detection. *Neural Computing and Applications*, 35(11):8471–8486, 2023.

Rodrigo Caye Daudt, Bertr Le Saux, and Alexandre Boulch. Fully convolutional siamese networks for change detection. In *2018 25th IEEE International Conference on Image Processing (ICIP)*, pp. 4063–4067. IEEE, 2018.

Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pp. 248–255. Ieee, 2009.

Xianzhi Du, Yeqing Li, Yin Cui, Rui Qian, Jing Li, and Irwan Bello. Revisiting 3d resnets for video recognition. *arXiv preprint arXiv:2109.01696*, 2021.

Sheng Fang, Kaiyu Li, and Zhe Li. Changer: Feature interaction is what you need for change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 2023.

Lhuqita Fazry, Mgs M Luthfi Ramadhan, and Wisnu Jatmiko. Change detection of high-resolution remote sensing images through adaptive focal modulation on hierarchical feature maps. *IEEE Access*, 2023.

Sebastian Gerard, Paul Borne-Pons, and Josephine Sullivan. A simple, strong baseline for building damage detection on the xbd dataset. *arXiv preprint arXiv:2401.17271*, 2024.

Ritwik Gupta, Bryce Goodman, Nirav Patel, Ricky Hosfelt, Sandra Sajeev, Eric Heim, Jigar Doshi, Keane Lucas, Howie Choset, and Matthew Gaston. Creating xbd: A dataset for assessing building damage from satellite imagery. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, pp. 10–17, 2019a.

Ritwik Gupta, Richard Hosfelt, Sandra Sajeev, Nirav Patel, Bryce Goodman, Jigar Doshi, Eric Heim, Howie Choset, and Matthew Gaston. xbd: A dataset for assessing building damage from satellite imagery. *arXiv preprint arXiv:1911.09296*, 2019b.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

Peter Henderson, Riashat Islam, Philip Bachman, Joelle Pineau, Doina Precup, and David Meger. Deep reinforcement learning that matters. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32, 2018.

Bin Huang, Yichen Xu, and Feng Zhang. Remote sensing image change detection based on adjacent-level feature fusion and dense skip connections. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2024.

Shunping Ji, Shiqing Wei, and Meng Lu. Fully convolutional networks for multisource building extraction from an open aerial and satellite imagery data set. *IEEE Transactions on geoscience and remote sensing*, 57(1):574–586, 2018.

Gregory Koch, Richard Zemel, Ruslan Salakhutdinov, et al. Siamese neural networks for one-shot image recognition. In *ICML deep learning workshop*, volume 2. Lille, 2015.

Alexandre Lacoste, Nils Lehmann, Pau Rodriguez, Evan David Sherwin, Hannah Kerner, Björn Lütjens, Jeremy Andrew Irvin, David Dao, Hamed Alemohammad, Alexandre Drouin, et al. Geo-bench: Toward foundation models for earth monitoring. *arXiv preprint arXiv:2306.03831*, 2023.

Jialu Li and Chen Wu. Using difference features effectively: A multi-task network for exploring change areas and change moments in time series remote sensing images. *Available at SSRN 4779358*.

Kaiyu Li. Opencd. https://github.com/likyoo/open-cd, 2022.

Zhe Li, Xiaoxin Wang, Sheng Fang, Jianli Zhao, Shuqi Yang, and Wen Li. A decoder-focused multi-task network for semantic change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 2024.

Zhenglai Li, Chang Tang, Xinwang Liu, Changdong Li, Xianju Li, and Wei Zhang. Ms-former: Memory-supported transformer for weakly supervised change detection with patch-level annotations. *arXiv preprint arXiv:2311.09726*, 2023.

Hui Lin, Renlong Hang, Shanmin Wang, and Qingshan Liu. Diformer: A difference transformer network for remote sensing change detection. *IEEE Geoscience and Remote Sensing Letters*, 2024.

Wei Liu, Yiyuan Lin, Weijia Liu, Yongtao Yu, and Jonathan Li. An attention-based multiscale transformer network for remote sensing image change detection. *ISPRS Journal of Photogrammetry and Remote Sensing*, 202:599–609, 2023a.

Xu Liu, Yu Liu, Licheng Jiao, Lingling Li, Fang Liu, and Dan Zhang. Swin resnetswin transformers for change detection in remote sensing images. In *IGARSS 2023-2023 IEEE International Geoscience and Remote Sensing Symposium*, pp. 6660–6663. IEEE, 2023b.

Xu Liu, Yu Liu, Licheng Jiao, Lingling Li, Fang Liu, Shuyuan Yang, and Biao Hou. Mutsimnet: Mutually reinforcing similarity learning for rs image change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 2024.

Yi Liu, Chao Pang, Zongqian Zhan, Xiaomeng Zhang, and Xue Yang. Building change detection for remote sensing images using a dual-task constrained deep siamese convolutional network model. *IEEE Geoscience and Remote Sensing Letters*, 18(5):811–815, 2020.

Yikun Liu, Mingsong Li, Tao Xiao, Yuwen Huang, and Gongping Yang. Feature alignment and refinement for remote sensing images change detection. *International Journal of Remote Sensing*, 44(24):7827–7856, 2023c.

Wen Lu, Lu Wei, and Minh Nguyen. Bi-temporal attention transformer for building change detection and building damage assessment. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2024.

Kevin Musgrave, Serge Belongie, and Ser-Nam Lim. A metric learning reality check. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXV 16*, pp. 681–699. Springer, 2020.

Kevin Musgrave, Serge Belongie, and Ser-Nam Lim. Unsupervised domain adaptation: A reality check. *arXiv preprint arXiv:2111.15672*, 2021.

Chao Pang, Xingxing Weng, Jiang Wu, Qiang Wang, and Gui-Song Xia. Hicd: Change detection in quality-varied images via hierarchical correlation distillation. *IEEE Transactions on Geoscience and Remote Sensing*, 62:1–16, 2024.

Wenguang Peng, Wenzhong Shi, Min Zhang, and Lukang Wang. Fda-ffnet: A feature-distance attention based change detection network for remote sensing image. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2023.

Yujun Quan, Anzhu Yu, Wenyue Guo, Xuanbei Lu, Bingchun Jiang, Shulei Zheng, and Peipei He. Unified building change detection pre-training method with masked semantic annotations. *International Journal of Applied Earth Observation and Geoinformation*, 120:103346, 2023.

Wuxu Ren, Zhongchen Wang, Min Xia, and Haifeng Lin. Mfinet: Multi-scale feature interaction network for change detection of high-resolution remote sensing images. *Remote Sensing*, 16(7): 1269, 2024.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, pp. 234–241. Springer, 2015.

Li Shen, Yao Lu, Hao Chen, Hao Wei, Donghai Xie, Jiabao Yue, Rui Chen, Shouye Lv, and Bitao Jiang. S2looking: A satellite side-looking dataset for building change detection. *Remote Sensing*, 13(24):5094, 2021.

Aarohi Srivastava, Abhinav Rastogi, Abhishek Rao, Abu Awal Md Shoeb, Abubakar Abid, Adam Fisch, Adam R Brown, Adam Santoro, Aditya Gupta, Adrià Garriga-Alonso, et al. Beyond the imitation game: Quantifying and extrapolating the capabilities of language models. *arXiv preprint arXiv:2206.04615*, 2022.

Adam J Stewart, Caleb Robinson, Isaac A Corley, Anthony Ortiz, Juan M Lavista Ferres, and Arindam Banerjee. Torchgeo: deep learning with geospatial data. In *Proceedings of the 30th international conference on advances in geographic information systems*, pp. 1–12, 2022.

Jérémie Sublime and Ekaterina Kalinicheva. Automatic post-disaster damage mapping using deep-learning techniques for change detection: Case study of the tohoku tsunami. *Remote Sensing*, 11 (9):1123, 2019.

Mingxing Tan and Quoc Le. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, pp. 6105–6114. PMLR, 2019.

Yonghui Tan, Xiaolong Li, Yishu Chen, and Jinquan Ai. Bd-msa: Body decouple vhr remote sensing image change detection method guided by multi-scale feature information aggregation. *arXiv preprint arXiv:2401.04330*, 2024.

Mikaela Angelina Uy, Quang-Hieu Pham, Binh-Son Hua, Thanh Nguyen, and Sai-Kit Yeung. Revisiting point cloud classification: A new benchmark dataset and classification model on real-world data. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 1588–1597, 2019.

Ling Wan, Ye Tian, Wenchao Kang, and Lei Ma. Cldrnet: A difference refinement network based on category context learning for remote sensing image change detection. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2023.

Renfang Wang, Zijian Yang, Hong Qiu, Xiufeng Liu, and Dun Wu. Spatial and channel exchange based on efficientnet for detecting changes of remote sensing images. In *2023 26th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, pp. 1595–1600. IEEE, 2023.

Yukun Wang, Mengmeng Wang, Zhonghu Hao, Qiang Wang, Qianwen Wang, and Yuanxin Ye. Msgfnet: Multi-scale gated fusion network for remote sensing image change detection. *Remote Sensing*, 16(3):572, 2024.

Global Forest Watch. Global forest watch. *World Resources Institute, Washington, DC Available from http://www. globalforestwatch. org (accessed March 2002)*, 2002.

Zhongrong Xu, Chengkun Zhang, Jun Qi, Xilai Li, Bin Yao, and Lu Wang. A dual-difference change detection network for detecting building changes on high-resolution remote sensing images. *Geocarto International*, 39(1):2322080, 2024.

Wenkai Yan, Yikun Liu, Mingsong Li, Ruifan Zhang, and Gongping Yang. Multilevel feature aggregation and enhancement network for remote sensing change detection. *Journal of Applied Remote Sensing*, 18(1):016513–016513, 2024.

Yinglong Yan, Jun Yue, Jiaxing Lin, Zhengyang Guo, Yi Fang, Zhenhao Li, Weiying Xie, and Leyuan Fang. When vectorization meets change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 2023.

Kunping Yang, Gui-Song Xia, Zicheng Liu, Bo Du, Wen Yang, Marcello Pelillo, and Liangpei Zhang. Semantic change detection with asymmetric siamese networks. *arXiv preprint arXiv:2010.05687*, 2020.

Shiying Yuan, Ruofei Zhong, Cankun Yang, Qingyang Li, and YaXin Dong. Dynamically updated semi-supervised change detection network combining cross-supervision and screening algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, 2024.

Chenxiao Zhang, Peng Yue, Deodato Tapete, Liangcun Jiang, Boyi Shangguan, Li Huang, and Guangchao Liu. A deeply supervised image fusion network for change detection in high resolution bi-temporal remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 166:183–200, 2020.

Haotian Zhang, Hao Chen, Chenyao Zhou, Keyan Chen, Chenyang Liu, Zhengxia Zou, and Zhenwei Shi. Bifa: Remote sensing image change detection with bitemporal feature alignment. *IEEE Transactions on Geoscience and Remote Sensing*, 2024a.

Xingpeng Zhang, Yuru Li, Qiuli Wang, and Sijing Wu. Mfi-cd: a lightweight siamese network with multidimensional feature interaction for change detection. *International Journal of Remote Sensing*, 45(8):2548–2566, 2024b.

Xiaoyang Zhao, Keyun Zhao, Siyao Li, Chuanming Song, and Xianghai Wang. Gampf: A full-scale gated message passing framework based on collaborative estimation for vhr remote sensing image change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 2023a.

Xiaoyang Zhao, Keyun Zhao, Siyao Li, and Xianghai Wang. Gesanet: Geospatial-awareness network for vhr remote sensing image change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 2023b.

Yang Zhao, Yuxiang Zhang, Yanni Dong, and Bo Du. Adapting vision transformer for efficient change detection. *arXiv preprint arXiv:2312.04869*, 2023c.

Dalong Zheng, Zebin Wu, Jia Liu, Chih-Cheng Hung, and Zhihui Wei. Detail enhanced change detection in vhr images using a self-supervised multi-scale hybrid network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2024.

Meng Zhou, Weixian Qian, and Kan Ren. Multistage interaction network for remote sensing change detection. *Remote Sensing*, 16(6):1077, 2024.