Deep Learning-Based Change Detection of Forest, Water, Barren Land, and Human Activity (Buildings, Roads, Agriculture)

The objective of this project is to perform change detection and semantic segmentation of forest, water, barren land, and human activity (Buildings, Roads, Agriculture) to simplify the process of monitoring changes on the Earth's surface, with potential uses in climate analysis, ecosystem assessment, and disaster response. Currently this is done by using deeply supervised image fusion network (DSIFN) [1] and using a semantic segmentation model such as Deeplabv3 [2] model to create segmentation of the area of interest from satellite images. Code at https://github.com/luhouyang/Deep_Learning-Based_Change_Detection_of_Urban_Landscape.git

Datasets Used

LoveDA (8.91 GB): https://github.com/Junjue-Wang/LoveDA [3]

The dataset consists of 3 splits and a further split of Rural and Urban. This amounts to Test (Rural: 976, Urban: 820), Train (Rural: 1366, Urban: 1156), Validation (Rural: 992, Urban: 677) instances of 1024x1024 high spatial resolution (HSR) satellite images. The segment labels were merged for buildings and roads, producing 7 labels for forest, water, barren land, impervious surfaces (buildings and roads), agriculture, background, and no-data.

Change Detection Dataset (4.57 GB):
https://drive.google.com/file/d/1GX656JqqOyBi_Ef0w65kDGVto-nHrNs9 [4]
The dataset consists of 3 splits for satellite images, each containing subdirectories 'A', 'B', 'OUT' for post-change image, pre-change image and ground truth change mask. This amounts to Test (3000), Train (10000), Validation (2998) instances of 244x244 satellite images.

Unexplored Datasets Due to Limited Hardware Capability

DynamicEarthNet: Daily Multi-Spectral Satellite Dataset for Semantic Change Segmentation (524 GB; Categories: Impervious surfaces (man-made surfaces), water, soil, agriculture, wetlands, snow & ice, and forest & other vegetation): https://mediatum.ub.tum.de/1650201 [5].

Results

The deeply supervised image fusion network (DSIFN) [1] model trained on the Change Detection Dataset for 20 epochs achieved a validation F1 score of 65.466%. Trained model weights: Google_Drive

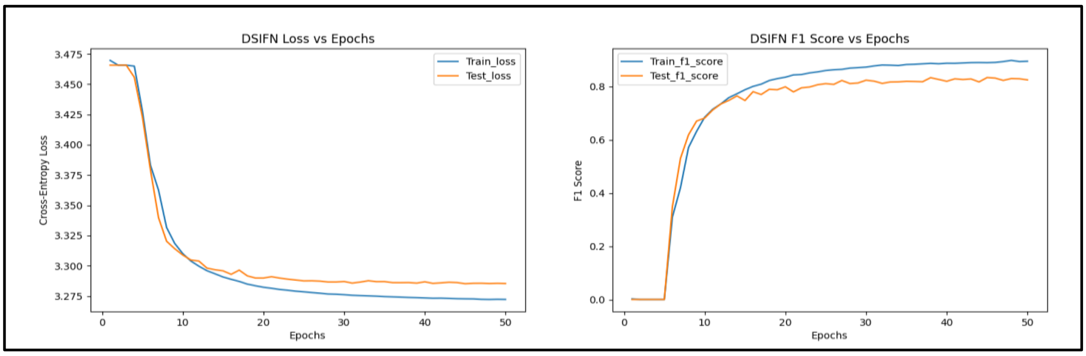

Figure 1: DSIFN Loss vs Epochs and F1 Score vs Epochs

Deeplabv3 [2] trained on the Urban Train Set of LoveDA dataset for 20 epochs achieved lowest test loss of 0.428 and highest test accuracy of 95.82%. Below shows the semantic segmentation of sample 3539 from the validation set. Trained model weights: Google_Drive

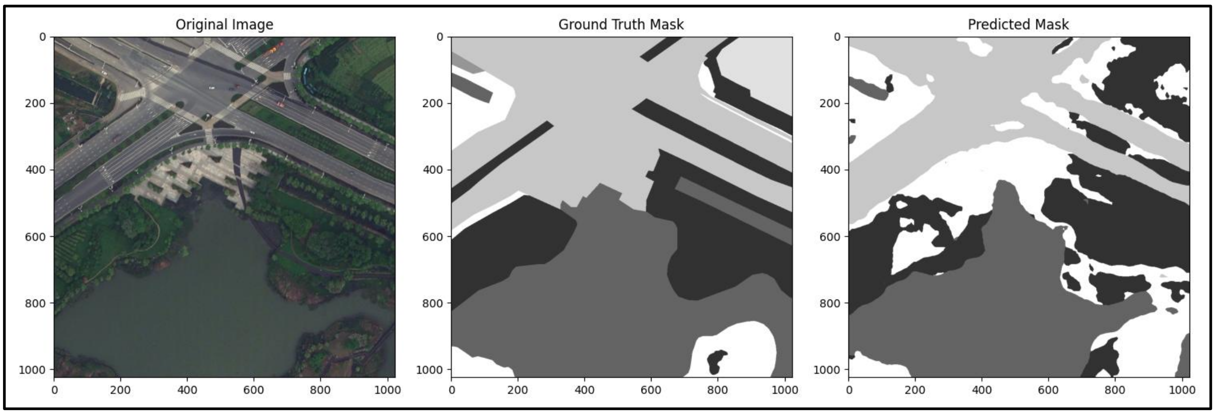

Figure 2: Deeplabv3 semantic segmentation

Below is the semantic change map generated from a 1600x1600 cutout area from CDD's satellite image.

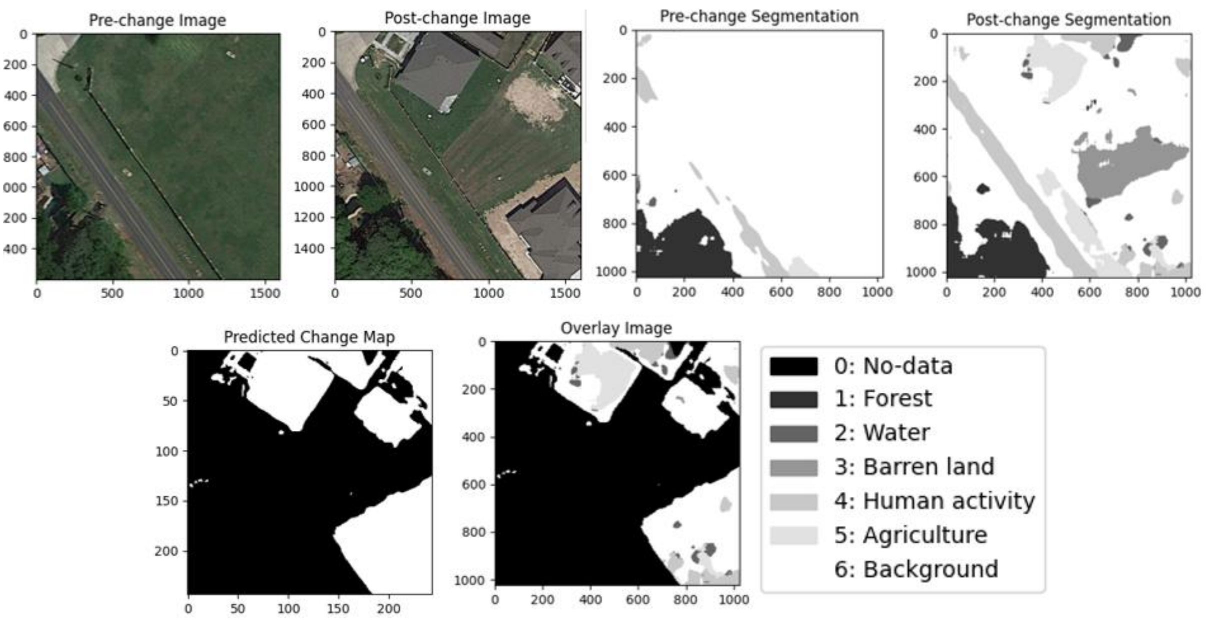

Figure 3: Pre- Post-change image, Pre- Postchange semantic segmentation, Change map and Semantic change map

Brief Discussion

1. Change Detection - DSIFN

The DSIFN model, trained on the Change Detection Dataset (CDD), achieved an F1 score of 65.466%. These relatively low metrics can be attributed to the CDD's lack of geographic diversity, as it contains samples from only one location. To address this issue transformations such as flip, re-scale, crop, Gaussian blur, and colour jittering could be applied to the data. Furthermore, a more diverse dataset, such as DynamicEarthNet, which spans 75 locations could improve the model's ability to generalize across different geographical regions.

2. Semantic Segmentation - Deeplabv3

The results of the fine-tuned Deeplabv3 model on LoveDA dataset shows promising results on the validation set, as shown in Figure 2. However, it is unable to handle novel data from other geographical locations (i.e. different house architecture and snow coverage in CDD). The model also struggles with objects of different scales, since the input size for the DSIFN model and Deeplabv3 model are 244x244 and 1024x1024 respectively. Generalization could be improved with a more diverse dataset. Detecting objects at different scales could be helped by implementing Shunted Self-Attention (SSA) which selectively merges tokens to represent objects of different sizes, while keeping tokens for fine-grained features [6].

3. Semantic Change Detection - DSIFN + Deeplabv3

As seen in the segmentation results, the semantic segmentation map and change map do not overlap fully, producing semantic change maps that are unable to fully capture the change over time. This is caused by the separation of the semantic segmentation generation process and the change detection process.

4. Future Direction

The above-mentioned issues could be resolved by unifying the semantic segmentation and change detection process. Further exploration can be done by combining two successful techniques into a new model - Shunted Self-Attention (SSA) [6] which is currently used by Yuan et al. [7] in bi-temporal satellite image semantic change detection (SCD) and Temporal Attention [8] which was recently used by Vincent et al. [9] for satellite image time series semantic change detection (SITS-SCD). The proposed model would be able to output bi-temporal or multi-temporal SCDs. Source code for SSA is available at https://github.com/OliverRensu/Shunted-Transformer [6]. Source code for Temporal Attention is available at https://github.com/VSainteuf/utae-paps [8]. Source code for SITS-SCD is available at https://github.com/ElliotVincent/SitsSCD [9].

References

[1] C. Zhang et al., "A deeply supervised image fusion network for change detection in high resolution bi-temporal remote sensing images," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 166, pp. 183–200, Aug. 2020, doi: https://doi.org/10.1016/j.isprsjprs.2020.06.003.

[2] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking Atrous Convolution for Semantic Image Segmentation," arxiv.org, Jun. 2017, Available: https://arxiv.org/abs/1706.05587

[3] J. Wang, Z. Zheng, A. Ma, X. Lu, and Y. Zhong, "LoveDA: A Remote Sensing Land-Cover Dataset for Domain Adaptive Semantic Segmentation," Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks, Oct. 2021, Available: https://www.researchgate.net/publication/355390292_LoveDA_A_Remote_Sensing_Land-Cover_Dataset_for_Domain_Adaptive_Semantic_Segmentation

[4] M. A. Lebedev, Y. V. Vizilter, O. V. Vygolov, V. A. Knyaz, and A. Y. Rubis, "CHANGE DETECTION IN REMOTE SENSING IMAGES USING CONDITIONAL ADVERSARIAL NETWORKS," The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XLII–2, pp. 565–571, May 2018, doi: https://doi.org/10.5194/isprs-archives-xlii-2-565-2018.

[5] A. Toker et al., "DynamicEarthNet: Daily Multi-Spectral Satellite Dataset for Semantic Change Segmentation," arXiv.org, 2022. https://arxiv.org/abs/2203.12560

[6] S. Ren, D. Zhou, S. He, J. Feng, and X. Wang, "Shunted Self-Attention via Multi-Scale Token Aggregation," arXiv.org, 2021. https://arxiv.org/abs/2111.15193

[7] P. Yuan, Q. Zhao, X. Zhao, X. Wang, X. Long, and Y. Zheng, "A transformer-based Siamese network and an open optical dataset for semantic change detection of remote sensing images," International Journal of Digital Earth, vol. 15, no. 1, pp. 1506–1525, Sep. 2022, doi: https://doi.org/10.1080/17538947.2022.2111470.

[8] V. Sainte and L. Landrieu, "Panoptic Segmentation of Satellite Image Time Series with Convolutional Temporal Attention Networks," arXiv.org, 2021. https://arxiv.org/abs/2107.07933.

[9] E. Vincent, J. Ponce, and M. Aubry, "Satellite Image Time Series Semantic Change Detection: Novel Architecture and Analysis of Domain Shift," arXiv.org, 2024. https://arxiv.org/abs/2407.07616.