The impact of different fusion modules on change detection with deep learning

Vasil Dakov [1] Dessislava Petrova-Antonova [2] Jan van Gemert [1]

[1]Delft University of Technology [2]GATE Institute, Sofia University "St. Kliment Ohridski"

Introduction - What is change detection?

Change detection - detecting differences images $\mathcal{X}_1, \mathcal{X}_2$ of the same area taken at different times $\mathcal{T}_1, \mathcal{T}_2$.

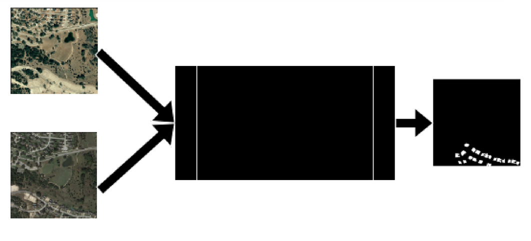


Figure 1. A blackbox change detection pipeline.

Facilitates **analyzing population trends** [1], **automatic map updates** [2], **illegal building detection** [3], etc.

Background - Encoder-Decoders and Fusion

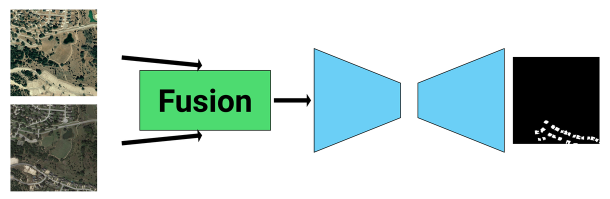


Figure 2. Early Fusion

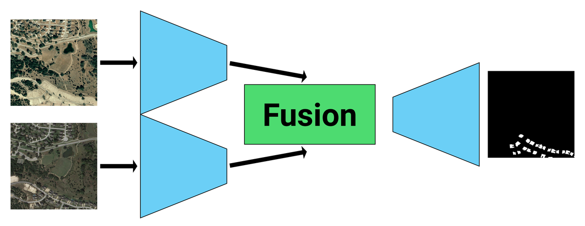


Figure 3. Middle Fusion

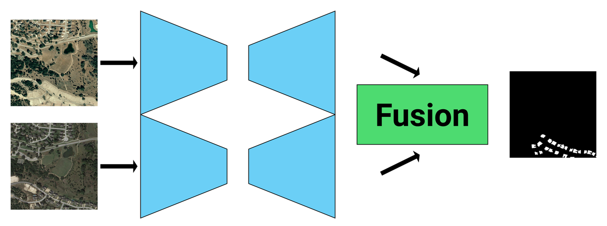


Figure 4. Late Fusion

- Dominant architecture for change detection - **encoder-decoder** [4]
- Fully Convolutional Networks - FCN, U-Net [5], [6]
- Shared trait - **fusion** – a point in the network where $\mathcal{X}_1, \mathcal{X}_2$ switch from being examined independently to being correlated together [4] [7].

Motivation and Research Hypothesis

Hypothesis: Changing the location fusion module impacts performance on changes that differ in size, spread and number. Inspired by the theory of CNNs.

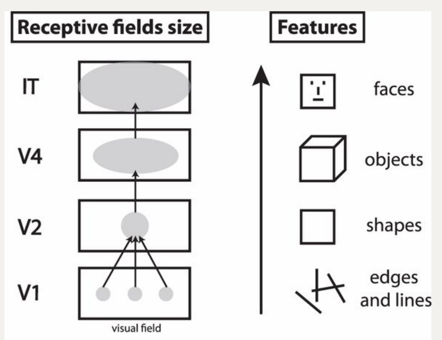
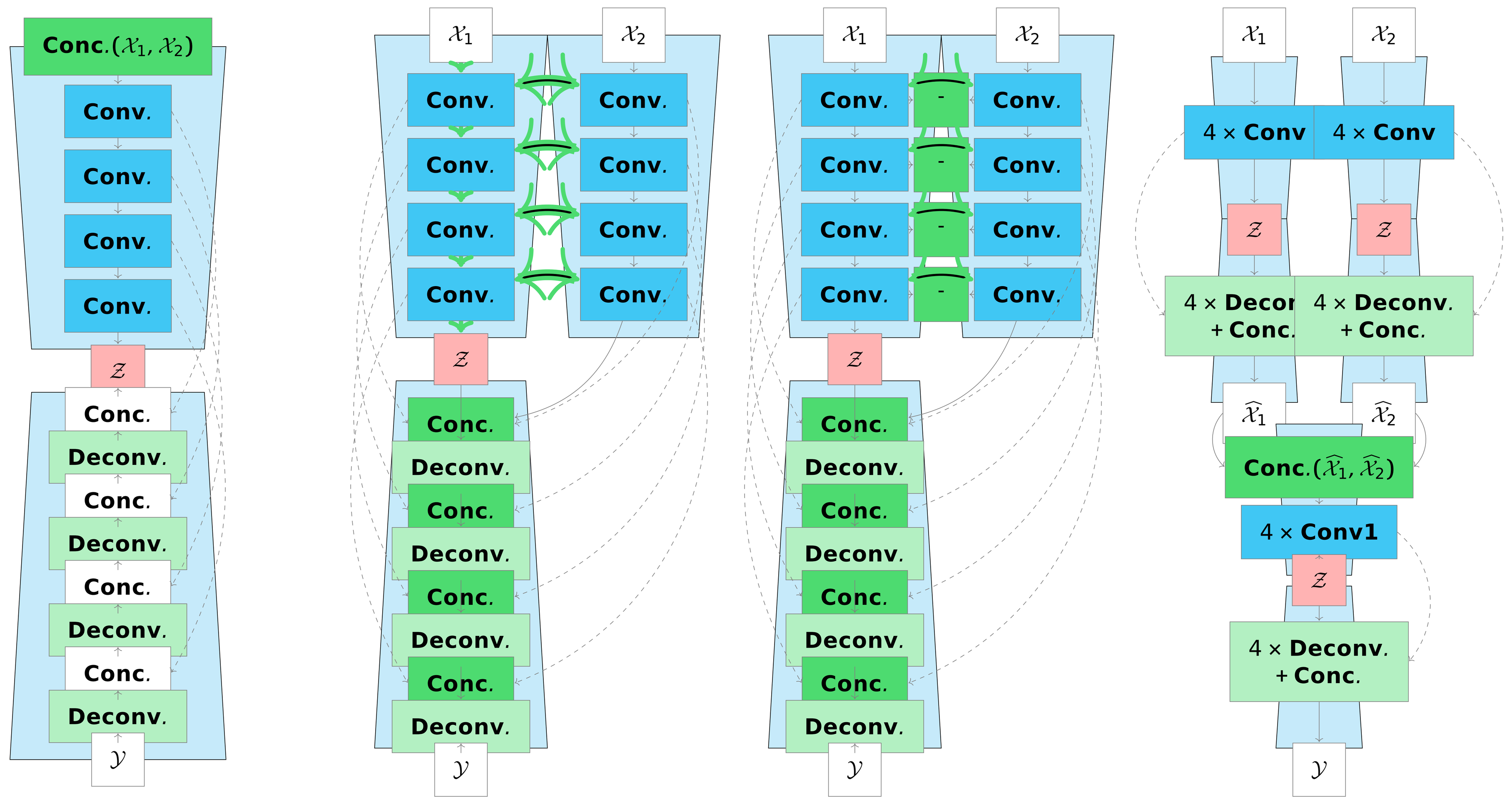


Figure 5. The increasing size of the receptive field at different visual cortex layers. Adapted from [8].

Aim: Change detection analysts can better understand what model to employ, based on the data in their area.

Methodology - Experimental Setup



(a) FC-EF (b) FC-Siam-Conc (c) FC-Siam-Diff (d) FC-LF

Figure 6. A showcase of all fusion architectures used. Point of fusion is in dark green. Figure adapted from [9].

Methodology - Datasets

- Two remote sensing datasatets - **LEVIR-CD** and **HiUCD** and a fully-controlled morphological dataset, **CSCD**.

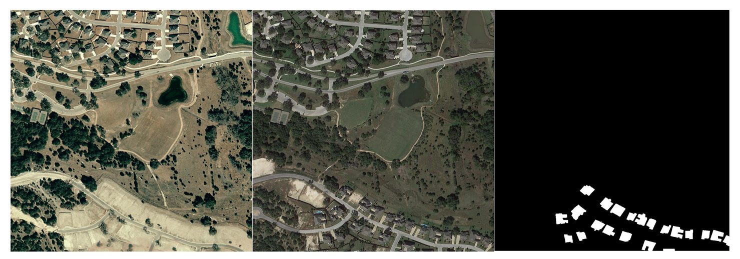
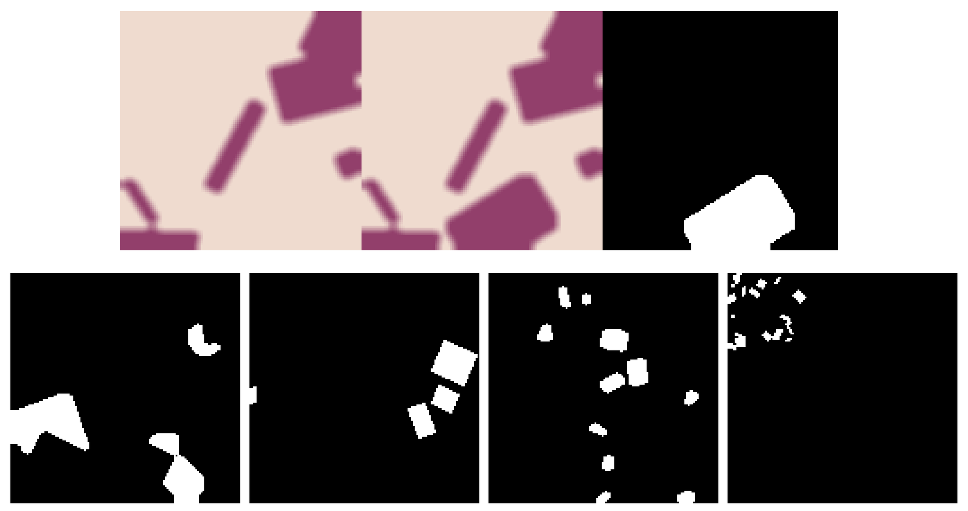


Figure 7. LEVIR-CD.

Figure 8. HiUCD.



(a) LCU. (b) LCNU. (c) SCU. (d) SCNU.

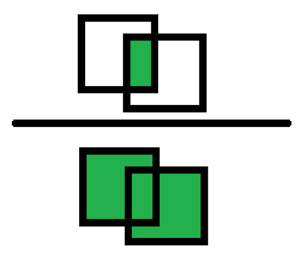
Figure 9. CSCD's labels (top) and categories (bottom).

Methodology - Training and Evaluation

- Training until early stopping.
- Categorical Evaluation.
- Contour counting.

$$\mu_{\text{diff}} = \frac{1}{N}\sum_{i=1}^{N} \left| |\mathcal{Y}_i| - |\hat{\mathcal{Y}}_i| \right|$$

IoU Threshold ≥ 0.5



Results and Discussion

Architecture	Acc.	Prec.	Rec.	F-1	μ_{diff}
FC-EF	70.31	0.70	0.96	0.81	97.55
FC-Siam-Conc	82.03	0.83	0.96	0.89	80.21
FC-Siam-Diff	39.06	0.36	0.87	0.51	276.5
FC-LF	81.24	0.81	0.99	0.89	53.6

Table 1. Evaluation on LEVIR-CD per architecture.

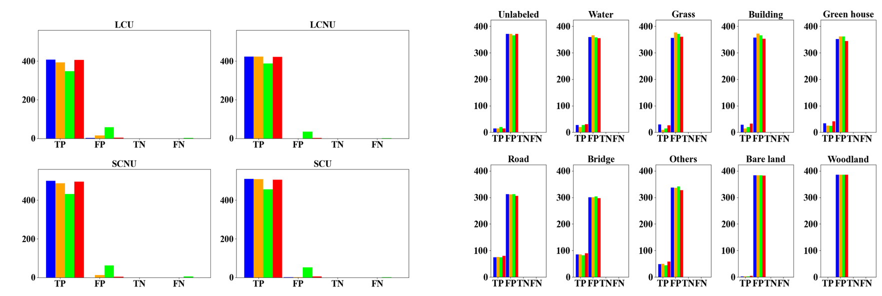
Architecture	Acc.	Prec.	Rec.	F-1	μ_{diff}
FC-EF	4.92	0.04	1	0.09	726
FC-Siam-Conc	7.51	0.07	1	0.13	776.9
FC-Siam-Diff	15.54	0.15	0.96	0.26	1336.6
FC-LF	17.09	0.17	0.85	0.29	975.5

Table 2. Evaluation on HiUCD per architecture.

Architecture	Acc.	Prec.	Rec.	F-1	μ_{diff}
FC-EF	98.97	0.98	1	0.99	10.08
FC-Siam-Conc	91.79	0.91	0.99	0.95	3.7
FC-Siam-Diff	67.91	0.76	0.80	0.78	5.34
FC-LF	97.31	0.97	0.99	0.98	6.71

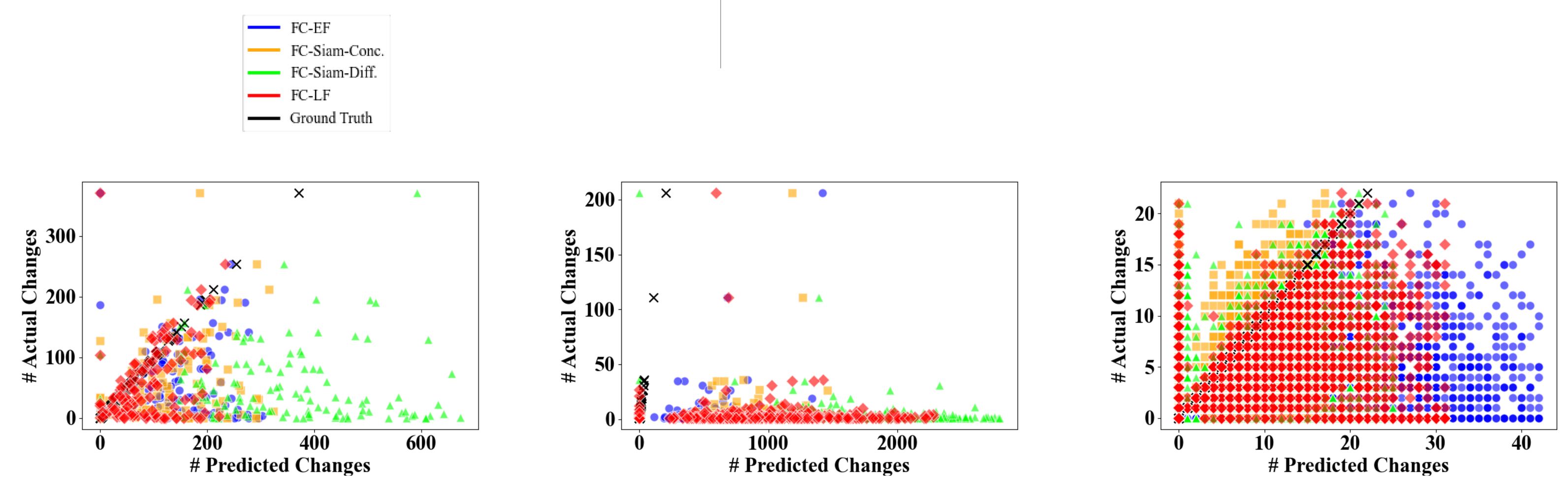
Table 3. Evaluation on CSCD per architecture.

- Inconclusive in accuracy
- Inconclusive categorically
- Overestimation
- Early fusion consistently overestimate most



(a) CSCD. (b) HiUCD.

Figure 10. Histograms with categorical evaluation results of dat



(a) LEVIR. (b) HiUCD. (c) CSCD.

Figure 11. Number of actual vs. predicted changes in different datasets.

Conclusion and Future Work

- Architecture isolation
- Class imbalance - loss
- Labels / CSCD Generation
- Extension to other networks

Per dataset, changing fusion impacts performance, but the only conclusively impacted metric is μ_{diff}.

References

[1] I. S. Serasinghe Pathiranage, L. N. Kantakumar, and S. Sundaramoorthy, "Remote Sensing Data and SLEUTH Urban Growth Model: As Decision Support Tools for Urban Planning," *Chinese Geographical Science*, vol. 28, pp. 274–286, Apr. 2018.

[2] F. Dahle, K. Arroyo Ohori, G. Agugiaro, and S. Briels, "Automatic change detection of digital maps using aerial images and point clouds," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLIII-B2-2021, pp. 457–464, 2021.

[3] S. Xu, G. Vosselman, and S. Oude Elberink, "Detection and classification of changes in buildings from airborne laser scanning data," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. II-5/W2, pp. 343–348, 2013.

[4] G. Cheng, Y. Huang, X. Li, S. Lyu, Z. Xu, Q. Zhao, and S. Xiang, "Change Detection Methods for Remote Sensing in the Last Decade: A Comprehensive Review," May 2023. arXiv:2305.05813 [cs, eess].

[5] E. Shelhamer, J. Long, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," May 2016. arXiv:1605.06211 [cs] version: 1.

[6] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," May 2015. arXiv:1505.04597 [cs].

[7] 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, p. 1688, Jan. 2020.

Number: 10 Publisher: Multidisciplinary Digital Publishing Institute.

[8] M. H. Herzog and A. M. Clarke, "Why vision is not both hierarchical and feedforward," *Frontiers in Computational Neuroscience*, vol. 8, Oct. 2014.

[9] R. Caye Daudt, B. Le Saux, and A. Boulch, "Fully Convolutional Siamese Networks for Change Detection," in *2018 25th IEEE International Conference on Image Processing (ICIP)*, pp. 4063–4067, Oct. 2018.