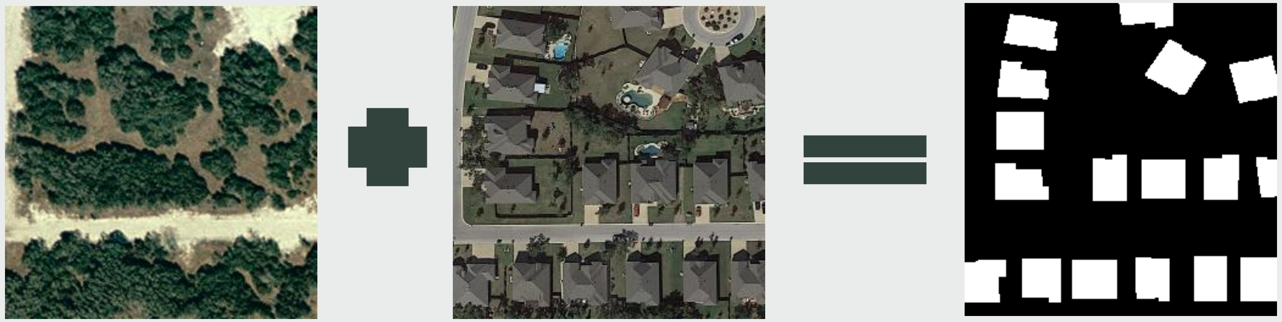
AYANet: A Gabor Wavelet-based and CNN-based Double Encoder for Building Change Detection in Remote Sensing

Priscilla Indira Osa[1,2], Josiane Zerubia[2], Zoltan Kato[3,4]

(1)*University of Genoa, Italy* ***(2)****Inria, Université Côte d'Azur, France* ***(3)****University of Szeged, Hungary* ***(4)****J. Selye University, Slovakia*

Introduction

Building Change Detection (BCD)?



t_0 t_1 Change map

Importance

Urban expansion
Deforestation
Natural disaster
…

Challenges

NOT ONLY	Differentiate changed and unchanged pixels of buildings
BUT ALSO	Ignore changes in other objects e.g., roads, vegetation, etc. Ignore unimportant changes mainly from sensing conditions

Related Works

Non-deep methods

Support Vector Machine
Random Forest
Etc.

Deep methods

Strategy-wise

- 3D-CNN to consider time axis : *AFCF3D-Net[1]*
- Various attention mechanisms : *DMI-Net[2], DUNE-CD[3], FHD[4], TINYCD[5]*
- Considering geometric of the object : *GVA-CD[6]*

Etc.

Architecture-wise

- Pure CNN : *STANet[7]*
- Pure Transformer : *ChangeFormer[8]*
- Hybrid CNN + Transformer : *BIT[9]*

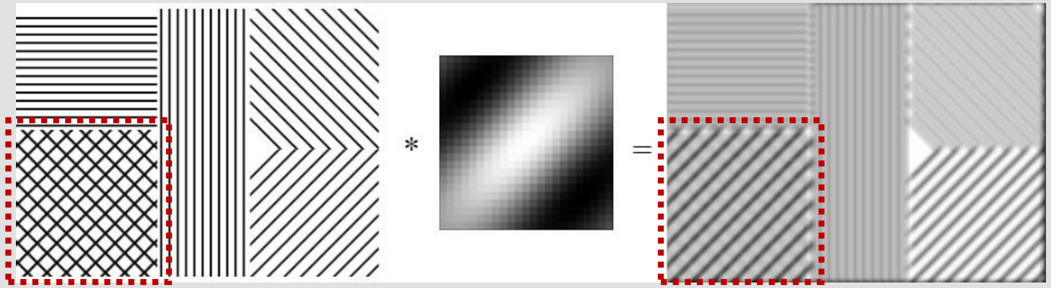
Methodology

The idea

Building's characteristics
in remote sensing imagery

Repetitive texture

Gabor Filter



Proposal

AYANet

1. Gabor encoder : explicitly extract texture features
2. CNN encoder : extract high-level features
3. Feature Conjunction Module to combine features from both encoders

Methodology

Gabor Filter

$$G(u,v) = \frac{\|\boldsymbol{k}_{u,v}\|^2}{\sigma^2} e^{-\left(\frac{\|\boldsymbol{k}_{u,v}\|^2 \|\boldsymbol{z}\|^2}{2\sigma^2}\right)} \left[e^{i\boldsymbol{k}_{u,v}\boldsymbol{z}} - e^{\frac{-\sigma^2}{2}} \right]$$

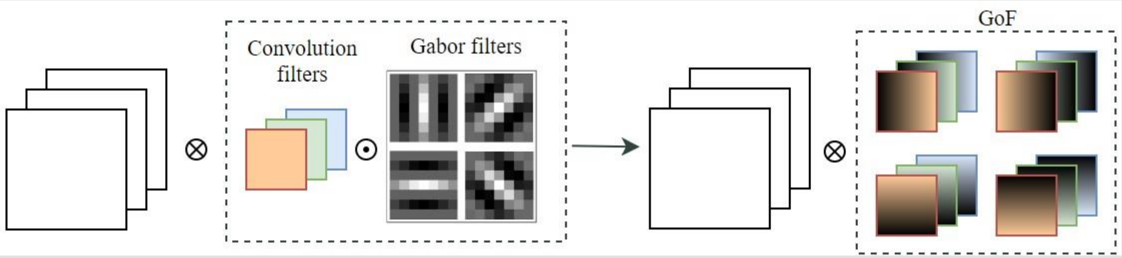
$$\boldsymbol{z} = (x,y), \boldsymbol{k}_{u,v} = \begin{pmatrix} k_v \cos k_u \\ k_v \sin k_u \end{pmatrix}, \boxed{k_v = \frac{\pi/2}{\sqrt{2}^{v-1}}}, \boxed{k_u = u\frac{\pi}{U}}, \sigma = 2\pi$$

frequency orientation

Integrated to CNN

Gabor Orientation Filters[10]

Gabor filters modulated in learnable CNN filters



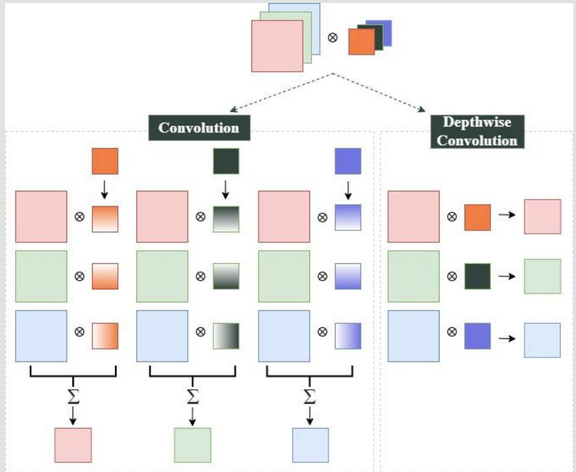
Change the standard convolution to have less parameter

Depthwise Convolution[11]

$$Conv_{k,l,n} = \sum_{i,j,m} \mathrm{K}_{i,j,m,n} \mathrm{F}_{k+i-1,l+j-1,m}$$

$$\downarrow$$

$$DConv_{k,l,m} = \sum_{i,j} \mathrm{K}_{i,j,m} \mathrm{F}_{k+i-1,l+j-1,m}$$



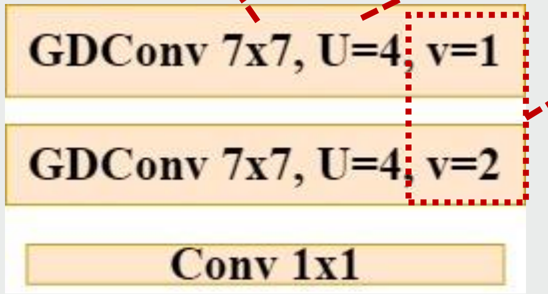
The building block of the Gabor encoder **(GDConv)**

Methodology

Gabor encoder

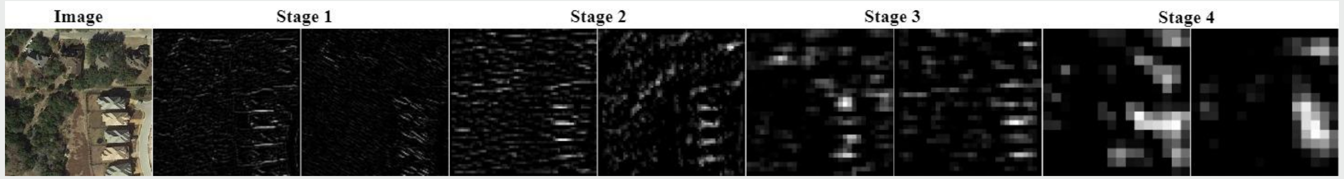
Bigger filter size = wider receptive field.
Higher computational cost is alleviated by the depthwise convolution

4 orientations (horizontal, vertical, diagonal)

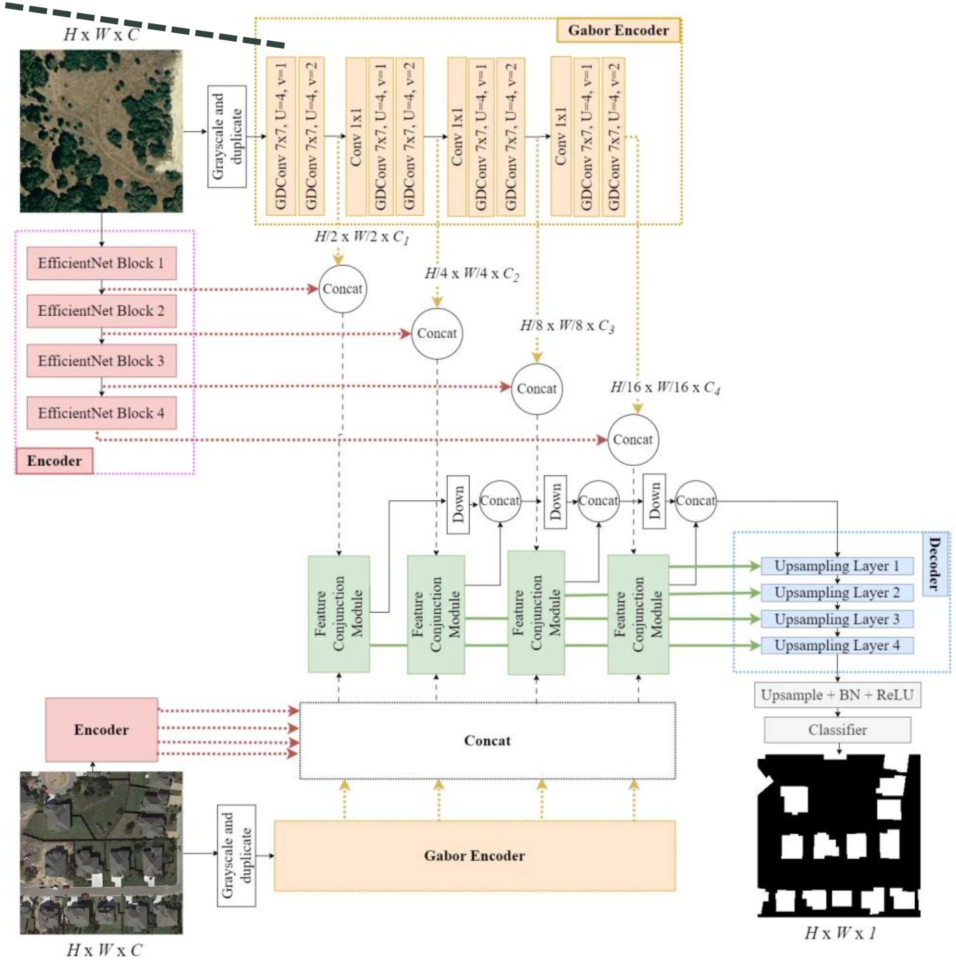


2 scales which correspond to 2 frequencies

Stacked GDConv at different stage = applying Gabor filters at the resized image

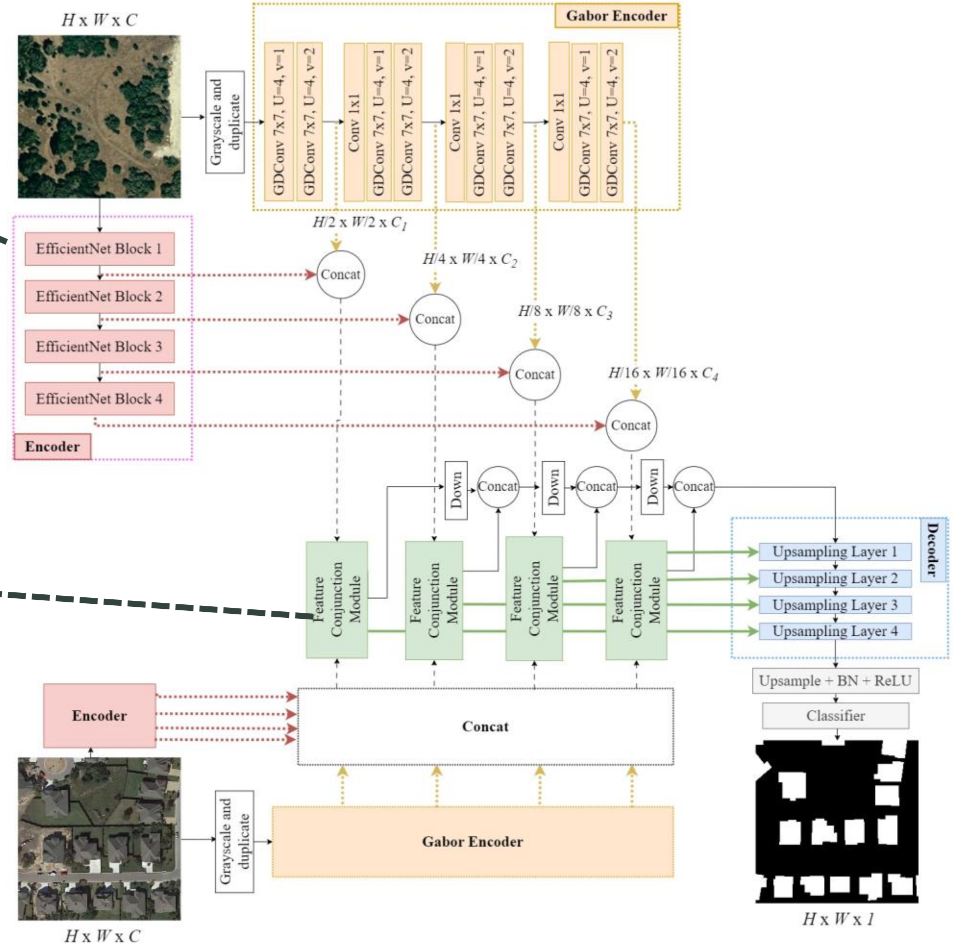


Overall Architecture



Methodology

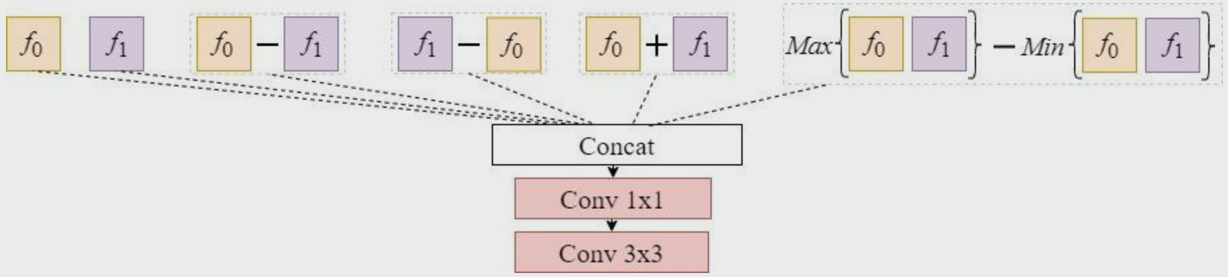
Overall Architecture



CNN-based encoder

EfficienNet-B7[12]-based encoder learns other high-level features

Feature Conjunction Module



Experimental Validation

LEVIR-CD

Model	Precision	Recall	F1-score	IoU
AFCF3D-Net	91.35%	90.17%	90.76%	83.08%
BIT	89.24%	89.37%	89.31%	80.68%
ChangeFormer	92.05%	88.80%	90.40%	82.48%
DMI-Net	92.52%	89.95%	90.71%	82.99%
DUNE-CD	92.27%	88.83%	90.52%	82.68%
FHD	92.61%	89.61%	91.09%	83.63%
GVA-CD	92.63%	87.88%	90.31%	82.51%
MSFCTNet	92.06%	90.00%	91.02%	83.52%
STANet-PAM	83.81%	**91.00%**	87.26%	77.40%
TINYCD	**92.68%**	89.47%	91.05%	83.57%
AYANet	92.60%	90.25%	**91.41%**	**84.17%**

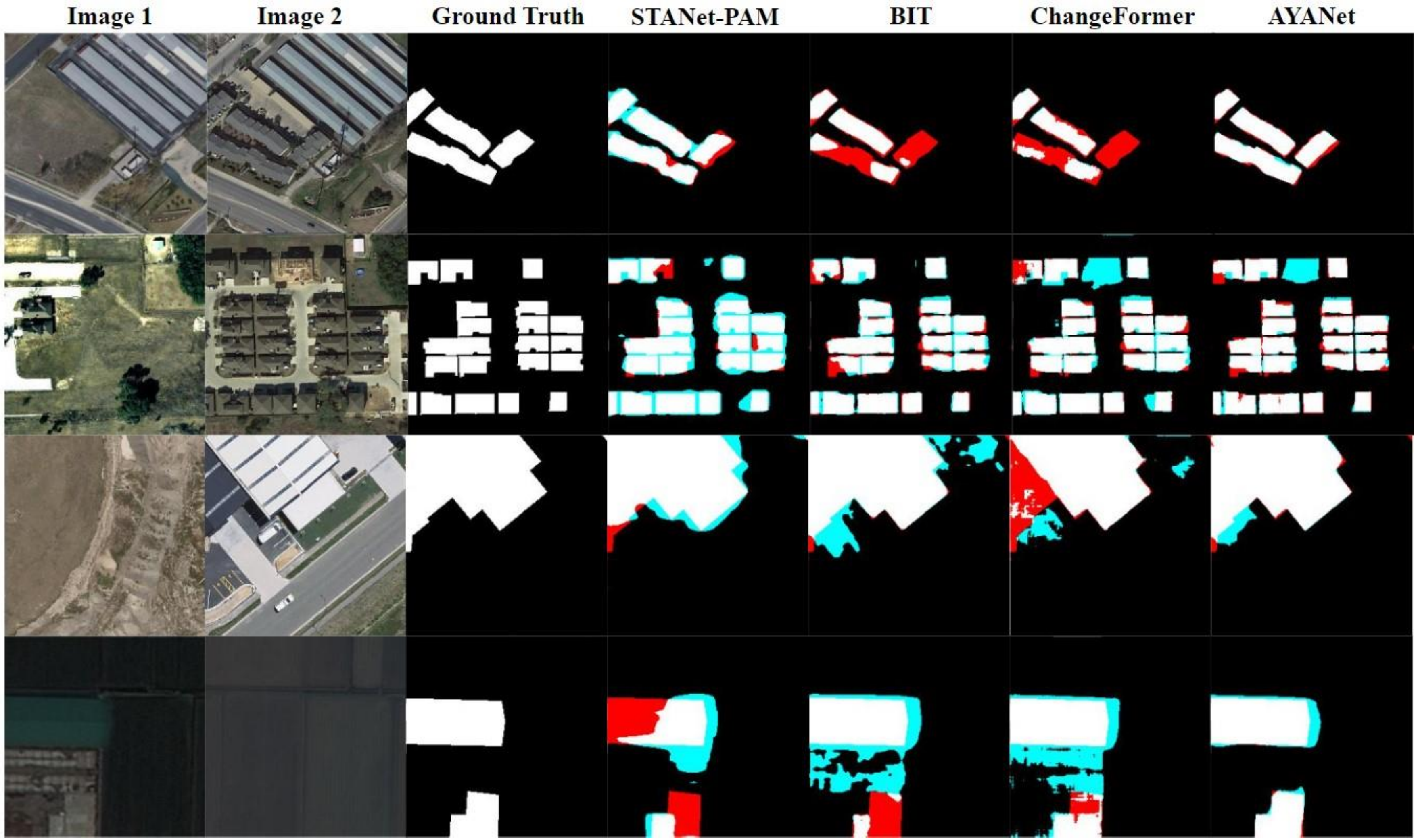
WHU-CD

Model	Precision	Recall	F1-score	IoU
BIT	87.65%	90.91%	89.25%	80.59%
ChangeFormer	94.15%	85.52%	89.63%	81.20%
STANet-PAM	70.65%	**93.54%**	80.50%	67.37%
AYANet	**95.56%**	92.89%	**94.21%**	**89.05%**

S2Looking

Model	Precision	Recall	F1-score	IoU
BIT	**73.99%**	52.73%	61.58%	44.49%
ChangeFormer	68.04%	57.03%	62.05%	44.98%
STANet-PAM	36.30%	**61.84%**	45.74%	29.65%
AYANet	69.37%	58.70%	**63.59%**	**46.62%**

Experimental Validation



Experimental Validation

Ablation Study

Encoder	Precision	Recall	F1-score	IoU
Gabor	90.51%	87.38%	88.92%	80.05%
EfficientNet	92.15%	90.35%	91.24%	83.90%
AYANet	**92.60%**	90.25%	**91.41%**	**84.17%**

Conclusion

1. Challenge in BCD includes differentiating between changes in buildings and changes in other objects
2. Utilizing the buildings' distinctive features in RS imagery (repetitive texture) → Gabor filters
3. AYANet: a BCD model with double encoders
    a. Gabor encoder: Gabor filters modulated in CNN to extract buildings' features
    b. CNN-based encoder: extracting other high-level features
4. AYANet showed promising results when compared with SOTA on 3 benchmark BCD datasets

References

[1] Ye, Y., Wang, M., Zhou, L., Lei, G., Fan, J., Qin, Y.: Adjacent-level feature cross-fusion with 3-D CNN for remote sensing image change detection. IEEE Transactions on Geoscience and Remote Sensing 61, 1–14 (2023). https://doi.org/10.1109/TGRS.2023.3305499

[2] Feng, Y., Jiang, J., Xu, H., Zheng, J.: Change detection on remote sensing images using dual-branch multilevel intertemporal network. IEEE Transactions on Geoscience and Remote Sensing 61, 1–15 (2023). https://doi.org/10.1109/TGRS.2023.3241257

[3] Adil, E., Yang, X., Huang, P., Liu, X., Tan, W., Yang, J.: Cascaded U-Net with training wheel attention module for change detection in satellite images. Remote Sensing 14(24) (2022). https://doi.org/10.3390/rs14246361, https://www.mdpi.com/2072-4292/14/24/6361

[4] Pei, G., Zhang, L.: Feature hierarchical differentiation for remote sensing image change detection. IEEE Geoscience and Remote Sensing Letters 19, 1–5 (2022). https://doi.org/10.1109/LGRS.2022.3193502

[5] Codegoni, A., Lombardi, G., Ferrari, A.: TINYCD: a (not so) deep learning model for change detection. Neural Computing and Applications 35, 8471–8486 (2023). https://doi.org/10.1007/s00521-022-08122-3, https://doi.org/10.1007/s00521-022-08122-3

[6] Huo, S., Zhou, Y., Zhang, L., Feng, Y., Xiang, W., Kung, S.Y.: Geometric variation adaptive network for remote sensing image change detection. IEEE Transactions on Geoscience and Remote Sensing 62, 1–14 (2024). https://doi.org/10.1109/TGRS.2024.3363431

[7] Chen, H., Shi, Z.: A spatial-temporal attention-based method and a new dataset for remote sensing image change detection. Remote Sensing 12, 1662 (2020). https://doi.org/https://doi.org/10.3390/rs12101662

[8] Bandara, W.G.C., Patel, V.M.: A Transformer-based Siamese network for change detection. In: IEEE International Geoscience and Remote Sensing Symposium IGARSS. pp. 207–210 (2022). https://doi.org/10.1109/IGARSS46834.2022.9883686

[9] Du, P., Liu, S., Gamba, P., Tan, K., Xia, J.: Fusion of difference images for change detection over urban areas. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 5(4), 1076–1086 (2012). https://doi.org/10.1109/JSTARS.2012.2200879

[10] Luan, S., Chen, C., Zhang, B., Han, J., Liu, J.: Gabor convolutional networks. IEEE Transactions on Image Processing 27(9), 4357–4366 (2018). https://doi.org/10.1109/TIP.2018.2835143

[11] Chollet, F.: Xception: Deep learning with depthwise separable convolutions. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) pp. 1800–1807 (2016), https://api.semanticscholar.org/CorpusID:2375110

[12] Tan, M., Le, Q.: EfficientNet: Rethinking model scaling for convolutional neural networks. In: Chaudhuri, K., Salakhutdinov, R. (eds.) Proceedings of the 36th International Conference on Machine Learning. Proceedings of Machine Learning Research, vol. 97, pp. 6105–6114. PMLR (09–15 Jun 2019), https://proceedings.mlr.press/v97/tan19a.html

Thank you

Paper available at

Codes available at *https://github.com/Ayana-Inria/AYANet*