Improving change detection performance with generative-adversarial augmentation of dataset

V.A. Knyaz[1,2], V.V. Kniaz[1,2] , S.Yu. Zheltov[2]

[1] Moscow Institute of Physics and Technology (MIPT), Moscow, Russia - (kniaz.va, kniaz.vv)@mipt.ru
[2] State Research Institute of Aviation Systems (GosNIIAS), Moscow, Russia - zhl@gosniias.ru

Keywords: Change detection, unmanned aerial vehicle, maps updating, neural networks, generative adversarial learning.

Abstract

The relevance of maps is a prerequisite for most geospatial applications such as navigation, urban planning, cadastre updating, etc. The main source of information used for maps updating is remote sensing, and the progress in sensors and methods of data analysis allows automatically retrieving the changes in observed scene from multi-time image series. Nowadays unmanned aerial vehicles (UAV) became a readily available and power mean for acquiring aerial images of a given territory. But solving change detection task using UAV-acquired images has some additional specificity, caused by additional disturbing specifics. This study addresses the problem of improving change detection performance in UAV-acquired imagery. The proposed approach firstly provide accurate image registration based on orthophoto generation, and then uses special technique for augmentation of data, that allows to improve the performance of the network model for change detection. The evaluation of the developed framework on the collected UAV-acquired multi-temporary image dataset has demonstrated change detection improving.

1. Introduction

The relevance of maps is a prerequisite for most geospatial applications such as navigation, urban planning, cadastre updating, etc. Satellite and aerial imagery is a rich source of information for retrieving the changes automatically, but multi-time image series usually contain a lot of differences, and not all of them should not be taken in account in specific task of change detection. For example, if the goal is to detect changes in the road network and the appearance of buildings, changes in vegetation or the distribution of cars on the road should be omitted.

With the progress in unmanned aerial vehicles and sensors, the UAV-based surveys become the powerful mean for fast acquisition of high resolution imagery. However, in comparison with satellite images, UAV imagery has additional disturbing specifics due to the different view angles and scales of the images, and diversity in illumination conditions.

Machine learning methods now demonstrate the state-of-the-art performance for a great range of data analysis problems, and for task of change detection in particular. But the performance of a network model significantly depends not only on adequate architecture and loss function, but also on relevant and diverse training dataset. The main issues that restrict their performance are the lack of variability and and classes imbalance in training dataset.

The presented study proposes the framework for semantic change detection using the images acquired from unmanned aerial vehicle.

The main contributions of the study are the following:

- the framework for UAV images processing aimed at change detection;
- generative adversarial network for creating augmented images with realistic changes based on real images;
- the representative dataset containing real and augmented images for efficient change detection network model training;
- evaluation of the proposed framework on created dataset and baselines.

2. Related work

The problem of change detection is addressed by scientific community for a several decades. The first studies on automatic change detection in digital images can be found in 1980-ths (Singh, 1989). The advent of digital images created the basis for the development of automatic methods for solving change detection problem. Before the period of the emergence and development of deep machine learning methods, point classification methods (Cao et al., 2014, Bovolo et al., 2008, Li et al., 2015, Hao et al., 2020, Zhou et al., 2016, Bruzzone and Fernández-Prieto, 2000, Touati et al., 2020) have made significant progress in solving the problem of detecting changes.

The methods for automatic change detection can be divided into hand-crafted techniques, that intensively developed till 2010-ths, and deep-learning-based ones, that now are the dominating in this scientific area. The hand-crafted methods, in their turn, could be subdivided into algebra-based (Lu Corresponding author et al., 2004), statistics-based (Liu et al., 2019) and transformation-based groups (Kuncheva and Faithfull, 2012).

The core of most methods of this class is the detection of modified pixels and their subsequent classification to create a map of changes. To solve the problem of detecting changes, various machine learning methods are used, such as random forests (Seo et al., 2018) and decision trees (Xie et al., 2016), support vector machine (Huo et al., 2016, Bovolo et al., 2008), Markov random fields (Bruzzone and Fernández-Prieto, 2000, Touati et al., 2020) and conditional random fields (Li et al., 2015, Hao et al., 2020, Zhou et al., 2016).

Meanwhile, the problem of detecting changes in time images of the same scene is complicated by a number of factors, such as differences in time of day and shooting season, differences in illumination and shadows, the presence of some interfering

https://doi.org/10.5194/isprs-archives-XLVIII-3-2024-265-2024

(a) Image acquired by SONY DSC-RX1R color camera (b) Image acquired by Phase One iXU150 color camera (c) Image acquired by Phase One IQ180 color camera

Figure 1. Images taken during the surveying

factors and noise. Another issue, that should be taken into account when developing change detection algorithms, is the requirements of application for the change detection procedure. In other words, change detection algorithms should be task-oriented, and should retrieve those differences, that are required by the application.

The progress in deep learning methods make it possible to solve the problem of semantic change detection, focused on retrieving application-oriented differences in image set. The number of deep learning methods for change detection is growing from year to year and now not complete list of these methods include more than forty network models. Transformer network models coupled with UNet-based models now demonstrate the state-of-the-art performance in terms of F_1 score and Overall Accuracy metrics (Parelius, 2023). Recently, the impressive progress in deep learning (and particularly, in deep convolutional neural networks), provided superiority of neural network models, that shows significantly better performance in solving the problem of detecting changes compared to traditional methods.

This fact is confirmed by a large number of works (Zhan et al., 2017, Huang et al., 2021, Liu et al., 2018, Gong et al., 2016, Hou et al., 2020, Khurana and Saxena, 2020, Kerner et al., 2019, Zhao et al., 2020, Lin et al., 2020), and it is mainly explained by the exceptional capabilities of deep neural networks in data representation and the ability to model the nonlinear characteristics of the processes under study, which determines the undeniable advantages of neural network models to achieve the best performance in this area.

For deep learning methods, as for any data-driven techniques, the very important role plays training dataset. Among the most frequently used in task of change detection LEVIR-CD (Chen and Shi, 2020) and CDD dataset (Lebedev et al., 2018) can be noted.

3. Materials and Methods

3.1 Change detection problem

In general, change detection is the task of finding differences in the same scene using remote sensing data obtained at different points in time. And the trivial solution for this problem is finding different elements of the data (e.g., pixels in case of images) related to the same object or area of a scene.

But for most applications, the change detection problem is task-specific, and aims to identify the specific differences identified by the task. For the task of updating the map, it is important to find changes in road and river routes, agricultural and urban areas, etc. For natural disaster monitoring, it is necessary to identify damaged facilities and assess losses.

The general technology for solving the problem of detecting changes includes solving several subtasks, such as:

3.1.1 Preliminary image processing. At this stage, the tasks of noise removal, correction of geometric and radiometric distortions and image quality improvement are being solved, providing the possibility of comparison for further analysis. Also, the tasks of images aligning is to be solved to provide accurate matching and correct comparison of the corresponding pixels in the images. This phase includes the estimate of the quality of image preprocessing for further reasoning on reliability of detected changes.

3.1.2 Choice of change retrieving technique. The approach for change detection significantly depend on the specific of the application. It is determined by the specifics of the semantic problem being solved, and are based on the analysis of the image content. Currently, the best results in solving the problems of detecting changes are demonstrated by methods based on machine learning.

3.1.3 Post-processing analysis. At this stage, results of change detection are analysing for eliminating outliers, false positive detections and making decision on the quality of the resulting change map, taking into account of the estimates of image preprocessing obtained at the first stage.

3.2 Evaluation Metrics

A measure of the correctness of detecting changes in an image, the Jacquard coefficient (or Intersection over Union, IoU, the ratio of the intersection of sets to their union) expresses the ratio of the intersection areas S_i and combining S_o sets of pixels of the reference markup M_{label} and the results of detecting changes M_{det}:

$$IoU = \frac{S_i}{So} \tag{1}$$

$$S_i = M_{label} \cap M_{det}. \tag{2}$$

$$S_o = M_{label} \cup M_{det} \tag{3}$$

With

(a) Tile from SONY DSC-RX1R image (b) Tile from Phase One iXU150 image (c) Tile from Phase One IQ180 image

Figure 2. Tiles for the same place produced from multi-temporary images of Figure 1.

M_{label} - the set of pixels of the reference markup,
M_{det} - sets of pixels of the change detection results.

The following criteria are also considered to assess the quality of detection:

precision

$$P = \frac{N_{tr}}{N_r} \tag{4}$$

recall:

$$R = \frac{N_{tr}}{N} \tag{5}$$

N_{tr} is a true positive number (the number of correctly detected modified pixels);
N_r is the number of all (true positive + false positive) pixels attributed to the changes.
N is the total number (true positive + false negative) of pixels in the image.

The metric F_1 is the average harmonic value of precision p_i and recall r:

$$F_1 = \frac{2}{\frac{1}{P} + \frac{1}{R}} = 2\frac{P \times R}{P + R}, \tag{6}$$

The Overall Accuracy represents proportion of correctly classified samples. The Overall Accuracy is calculated by dividing the number of correctly classified samples by the total number of samples.

$$A_{overall} = \frac{N_{tr}}{N}, \tag{7}$$

3.3 Study materials

In our study we address the problem of change detection in multi-temporary images acquired from UAV platform with various cameras in different survey conditions. Examples of images used in the study are shown in Figure 1.

The Figure 1 demonstrates the differences in images, that should be firstly eliminated by photogrammetric image transformation into orthophoto. It creates conditions for efficient image processing by machine learning.

The developed framework for retrieving semantic change detection from UAV-based imagery exploits two-stage image processing. The first stage is photogrammetric image processing for obtaining true orthophoto (Figure 3) of given area at different times. At the second stage image tiles from the orthophoto are processed by change detection network model, trained on adversarially augmented dataset.

Algorithm 1: Orthophoto generation

Input:
orthophoto region $S : \{X_1, Y_1, X_2, Y_2\}$
scale m and resolution $\delta x, \delta y$ of the orthophoto

Output:
orthophoto O_S of the region S

Procedure `Orthophoto generation():`
- $\Delta X = m \cdot \delta x$
- $\Delta Y = m \cdot \delta y$
- $X_c = X_1$
- **for** $i = 1, \ldots N$ **do**
  - $X_c = X_c + \Delta X$
  - $Y_c = Y_1$
  - **for** $j = 1, \ldots M$ **do**
    - $Y_c = Y_c + \Delta Y$
    - `/* Determining image point, corresponding to 3D point (using co-linearity equations and camera orientation) */`
    - (x_i, y_j) = FindImagePoint(X_c, Y_c)
    - `/* Assigning to the point of the orthophoto the color of the corresponding image point */`
    - Color($S(i, j)$) = Color(x_i, y_j)
- **return** O_S;

Figure 3. Algorithm for orthophoto generation

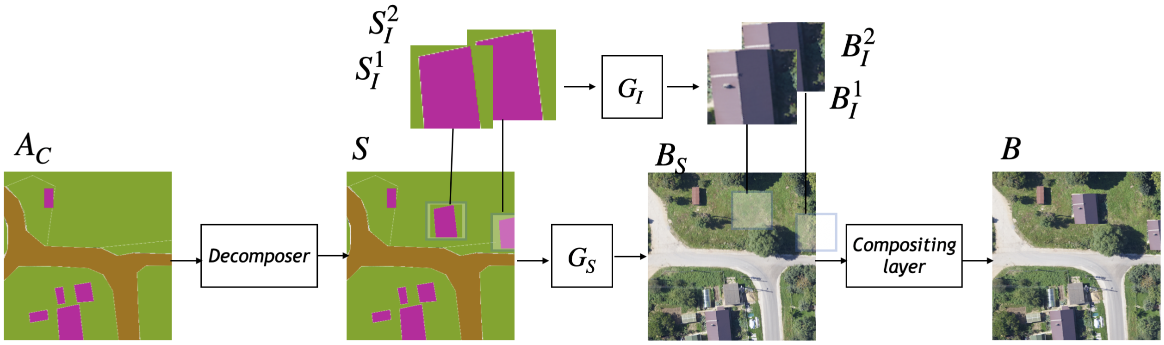


Figure 4. Annotation decomposition into background labeling S and instance labeling S_I. We use two generators G_S and G_I to generate background B_S and objects B_I from the annotations. The final output image B is generated by our compositing layer

For satellite images in many cases the distortions, caused by central projection imaging and by Earth surface relief, are negligible. So, in this case, image preliminary processing requires only transform to one scale and to one point of view.

The specificity of UAV-aquired imagery comparing with satellite images is more variability in scale, imaging conditions and imaging distortion. For providing accurate aligning of multitemporary and multi-scale images, we produce orthophoto images of the given scale for the area of interest.

For generation digital elevation model from the set of UAV-acquired images we have used the Structure-from-Motion and Multi View stereo techniques (Knyaz and Zheltov, 2017).

For orthophoto generation from a set of UAV-acquired images the standard procedure is used (Algorithm 1, Figure 3).

The collection and annotating of a training dataset containing examples of changes remains a significant issue. At the same time, the quality of the existing deep learning algorithms largely depends on the correspondence of the examples from the dataset to the images on which the algorithm is planned to be applied.

3.4 Dataset augmentation technique

To solve the problem of dataset preparation, the method for adversarial augmentation of the training dataset using generative-adversarial neural networks is developed. The main idea of the method is to train the generator network to create a modified version of the image based on an existing image, but containing artificial realistic changes. This allows, on the one hand, to use unlabeled images of any area as a training samples, having accurate labels automatically from image composing technique.

On the other hand, the competition of the detector and the generator networks improves the quality of detecting changes (Kniaz et al., 2021). With an increase in the number of training epochs, the nature of changes becomes more difficult to detect (due to the competitive loss function), that, in its own turn, forces the detector network to resolve more and more complicated cases.

3.4.1 Framework Overview. The goal of the developed network, termed as Instance-Quantized Generative Adversarial Network (IQ-GAN) (Kniaz et al., 2023), is training a mapping

$$G:\ (A_C, A_I, N) \rightarrow B \tag{8}$$

from semantic labeling A_C, instance labeling A_I, and 3D noise N to a color image B. The original semantic labeling A_C of the image is decomposed into a set instance semantic labels $A_C = \{A_C^1, A_C^2, \ldots, A_C^M\}$, where M is the number of object instances in the image. The main assumption of this approach is that the target image $\hat{B}$ can be represented as an alpha weighted sum of the background scene $\hat{B}_S$ and a set of object instances images $\hat{B}_1^A, B_2^A, \ldots, B_M^A$. Then, the resulting mapping G can be achieved by performing two mappings: a scene mapping G_S

$$G_S:\ (A_C, N) \rightarrow B \tag{9}$$

and instance mapping G_I

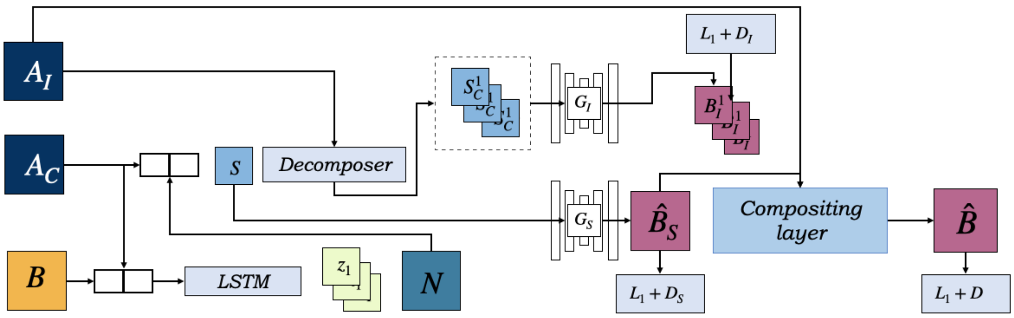
$$G_I:\ (A_C^i, N^i) \rightarrow B_I^i \tag{10}$$

In the developed framework there are five domains in consideration:

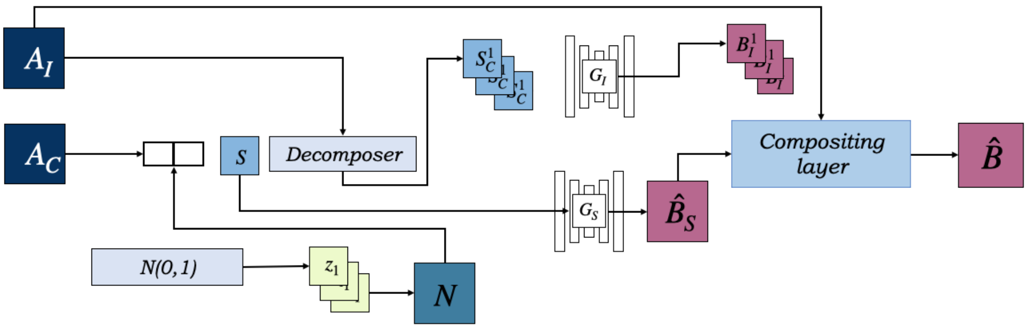
- the semantic label domain $\mathcal{A}_\mathcal{C} \in \{0, 1, \ldots, N\}^{W\times H}$, where N is the number of object classes,
- an instance labeling domain $\mathcal{A}_\mathcal{I} \in \{0, 1, \ldots, M\}^{W\times H}$,
- the real image domain $\mathcal{B} \in \mathbb{R}^{W\times H\times 3}$,
- the instance image domain $\mathcal{B}_\mathcal{I} \in \mathbb{R}^{W\times H\times 4}$, in which each pixel represents R, G, B components and an α mask,
- the semantic 3D noise domain $\mathcal{Z} \in \mathbb{R}^{W\times H\times D}$, where D is a hyper-parameter defining the number of noise channels.

The concatenation of the semantic labeling A_C and semantic noise Z is termed as $S = A_C \oplus Z$, where $\oplus$ is concatenation along channel axis.

The task of translation semantic labelling A_C into a color image B is multimodal (i.e. multiple images may correspond to a single semantic labeling A_C). The 3D noise (Schönfeld et al., 2021) serves as a starting point for developed multimodal sampling technique. The original 3D noise (Schönfeld et al., 2021) was used to train a stochastic model, and the developed model is to be deterministic. Specifically, the aim is to control individually the style of the whole scene and the style of each



(a) Training phase



(b) Inference phase

Figure 5. Overview of the IQ-GAN framework.

object instance. To obtain this aim, the 3D noise (Schönfeld et al., 2021) is extended to the semantic 3D noise.

We sample N random noise vectors z^j for each semantic region in the image and fill each region j in our semantic 3D noise tensor with the corresponding value z^j. The 3D semantic noise allows to solve two tasks. Firstly, it provides consistency between scene generator G_S and instance generator G_I, allowing instance generator to match the style of the instance and the whole scene. Secondly, it allows us to train an additional image semantic 3D noise encoder that estimates z^j conditioned by the distribution of values of the semantic region j in the ground truth color image B. This allows to carry out the transfer of instance level style from an arbitrary input image to the synthesized image $\hat{B}$.

The resulting pipeline of the developed framework is shown in Figure 5. Here six models are trained simultaneously: two generators G_S and G_I, semantic 3D noise encoder E_N, and three adversarial discriminators. Scene discriminator D_S distinguishes real images and background scene images $\hat{B}_S$ before instance compositing. Instance discriminator D_I evaluates instance synthetic images $\hat{B}_I$ produced by generator G_I and real instance images B_I produced be scene decomposer. Composite discriminator D aims to distinguish synthetic images $\hat{B}$ produced after instance compositing and real images B. Semantic noise encoder E_N generates 3D noise vector conditioned by the semantic labeling A_C and real images B.

3.4.2 Framework training and evaluation. We created the special UAV-CD dataset for change detection task using the imagery acquired during UAV surveys of the same area in different time periods. The collected images were process using the described above technique to create the set of one-scale and aligned samples. The resulting dataset contains 60k annotated images.

After that, the created UAV-CD dataset was augmented using the developed augmentation technique (Section 3.4.1). The overall number of images and annotations in augmented UAV-CDa dataset is 80k.

The developed IQ-GAN framework was trained on three image datasets: *ADE20k* (Zhou et al., 2017), *Cityscapes* (Cordts et al., 2016), *MHPv2* (Zhao et al., 2018). The PyTorch library (Paszke et al., 2017) was used for training. The training was performed using two NVIDIA 2090 Ti GPUs and took 176 hours. For network optimization, we use minibatch SGD with an Adam solver. We set learning rate to 0.0002 with momentum parameters $\beta_1 = 0.5$, $\beta_2 = 0.999$ similar to (Isola et al., 2017).

We evaluated performance of four state-of-the-art change detection network models (FC-EF (Caye Daudt et al., 2018), SNUNet (Fang et al., 2022), BiT (Chen et al., 2022), and ChangeFormer (Bandara and Patel, 2022) on the initial UAV-CD dataset and on the augmented UAV-CDa dataset.

F_1, Precision, Recall, and Intersection over Union (IoU) in terms of change class were used to evaluate the CD performance of the models. The results of evaluation is presented in Table 1. Table 1 demonstrates that the proposed framework allows to improve change detection performance comparing with training on non-augmented dataset.

Conclusion

The two-staged framework for retrieving semantic change detection from UAV-based imagery is developed. It firstly uses photogrammetric image processing for obtaining true orthophoto, and then solves change detection problem by deep learning technique. To improve the performance of change detection for network models special technique for augmentation of dataset is

Method	Backbone	UAV-CD					UAV-CDa				
		IoU	P	R	F_1	A_o	IoU	P	R	F_1	A_o
ConvNet-based											
FC-EF	-	78.44	82.18	97.81	89.32	89.41	79.67	84.02	98.47	90.67	89.41
SNUNet	-	88.23	94.12	96.54	95.31	95.15	89.94	95.26	97.81	96.52	89.41
Transformer-based											
BiT	(ResNet18)	84.77	96.17	92.74	94.42	94.63	86.01	97.58	93.08	95.28	95.41
ChangeFormer	(MiT-B1)	82.93	94.53	92.92	93.72	93.52	83.76	95.32	94.28	94.80	94.98

Table 1. Performance evaluation for proposed approach

developed, that allows to create the diverse and relevant dataset by adding realistic images.

The evaluation of the developed framework for four state-of-the-art change detection network models on the collected UAV-acquired multi-temporary image dataset has demonstrated that it provides change detection improving in terms F_1 score and Overall Accuracy.

Acknowledgements

The research was carried out at the expense of a grant from the Russian Science Foundation No. 24-21-00269,
https://rscf.ru/project/24-21-00269/

References

Bandara, W. G. C., Patel, V. M., 2022. A transformer-based siamese network for change detection. *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*, 207–210.

Bovolo, F., Bruzzone, L., Marconcini, M., 2008. A Novel Approach to Unsupervised Change Detection Based on a Semisupervised SVM and a Similarity Measure. *IEEE Trans. Geosci. Remote. Sens.*, 46(7), 2070–2082.

Bruzzone, L., Fernández-Prieto, D., 2000. Automatic analysis of the difference image for unsupervised change detection. *IEEE Trans. Geosci. Remote. Sens.*, 38(3), 1171–1182.

Cao, G., Li, Y., Liu, Y., Shang, Y., 2014. Automatic change detection in high-resolution remote-sensing images by means of level set evolution and support vector machine classification. *International Journal of Remote Sensing*, 35(16), 6255-6270.

Caye Daudt, R., Le Saux, B., Boulch, A., 2018. Fully convolutional siamese networks for change detection. *Proceedings - International Conference on Image Processing, ICIP*, 4063–4067.

Chen, H., Qi, Z., Shi, Z., 2022. Remote Sensing Image Change Detection With Transformers. *IEEE Trans. Geosci. Remote Sens.*, 60, 1-14.

Chen, H., Shi, Z., 2020. A Spatial-Temporal Attention-Based Method and a New Dataset for Remote Sensing Image Change Detection. *Remote Sensing*, 12(10). https://www.mdpi.com/2072-4292/12/10/1662.

Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., Schiele, B., 2016. The Cityscapes Dataset for Semantic Urban Scene Understanding. *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Fang, S., Li, K., Shao, J., Li, Z., 2022. SNUNet-CD: A Densely Connected Siamese Network for Change Detection of VHR Images. *IEEE Geosci. Remote Sens. Lett.*, 19, 1-5.

Gong, M., Zhao, J., Liu, J., Miao, Q., Jiao, L., 2016. Change Detection in Synthetic Aperture Radar Images Based on Deep Neural Networks. *IEEE Trans. Neural Networks Learn. Syst.*, 27(1), 125–138.

Hao, M., Zhou, M., Jin, J., Shi, W., 2020. An Advanced Superpixel-Based Markov Random Field Model for Unsupervised Change Detection. *IEEE Geosci. Remote. Sens. Lett.*, 17(8), 1401–1405.

Hou, B., Liu, Q., Wang, H., Wang, Y., 2020. From W-Net to CDGAN: Bitemporal Change Detection via Deep Learning Techniques. *IEEE Trans. Geosci. Remote. Sens.*, 58(3), 1790–1802.

Huang, R., Zhou, M., Xing, Y., Zou, Y., Fan, W., 2021. Change detection with various combinations of fluid pyramid integration networks. *Neurocomputing*, 437, 84–94.

Huo, C., Chen, K., Ding, K., Zhou, Z., Pan, C., 2016. Learning Relationship for Very High Resolution Image Change Detection. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.*, 9(8), 3384–3394.

Isola, P., Zhu, J.-Y., Zhou, T., Efros, A. A., 2017. Image-to-Image Translation with Conditional Adversarial Networks. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 5967–5976.

Kerner, H. R., Wagstaff, K. L., Bue, B. D., Gray, P. C., III, J. F. B., Amor, H. B., 2019. Toward Generalized Change Detection on Planetary Surfaces With Convolutional Autoencoders and Transfer Learning. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.*, 12(10), 3900–3918.

Khurana, M., Saxena, V., 2020. A Unified Approach to Change Detection Using an Adaptive Ensemble of Extreme Learning Machines. *IEEE Geosci. Remote. Sens. Lett.*, 17(5), 794–798.

Kniaz, V., Knyaz, V., Moshkantsev, P., 2023. IQ-GAN: Instance-Quantized Image Synthesis. *Studies in Computational Intelligence*, 1064, 277–291. https://doi.org/10.1007/978-3-031-19032-2_30.

Kniaz, V. V., Knyaz, V. A., Mizginov, V., Papazyan, A., Fomin, N., Grodzitsky, L., 2021. Adversarial dataset augmentation using reinforcement learning and 3d modeling. B. Kryzhanovsky, W. Dunin-Barkowski, V. Redko, Y. Tiumentsev (eds), *Advances in Neural Computation, Machine Learning, and Cognitive Research IV*, Springer International Publishing, Cham, 316–329.

Knyaz, V. A., Zheltov, S. Y., 2017. Accuracy evaluation of structure from motion surface 3D reconstruction. *Proc.SPIE*, 10332, 10332 - 10332 - 10. http://dx.doi.org/10.1117/12.2272021.

Kuncheva, L. I., Faithfull, W. J., 2012. PCA feature extraction for change detection in multidimensional unlabelled streaming data. *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*, 1140–1143.

Lebedev, M. A., Vizilter, Y. V., Vygolov, O. V., Knyaz, V. A., Rubis, A. Y., 2018. CHANGE DETECTION IN REMOTE SENSING IMAGES USING CONDITIONAL ADVERSARIAL NETWORKS. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2, 565–571. https://isprs-archives.copernicus.org/articles/XLII-2/565/2018/.

Li, H., Li, M., Zhang, P., Song, W., An, L., Wu, Y., 2015. SAR Image Change Detection Based on Hybrid Conditional Random Field. *IEEE Geosci. Remote. Sens. Lett.*, 12(4), 910–914.

Lin, Y., Li, S., Fang, L., Ghamisi, P., 2020. Multispectral Change Detection With Bilinear Convolutional Neural Networks. *IEEE Geosci. Remote. Sens. Lett.*, 17(10), 1757–1761.

Liu, J., Gong, M., Qin, A. K., Zhang, P., 2018. A Deep Convolutional Coupling Network for Change Detection Based on Heterogeneous Optical and Radar Images. *IEEE Trans. Neural Networks Learn. Syst.*, 29(3), 545–559.

Liu, S., Marinelli, D., Bruzzone, L., Bovolo, F., 2019. A Review of Change Detection in Multitemporal Hyperspectral Images: Current Techniques, Applications, and Challenges. *IEEE Geoscience and Remote Sensing Magazine*, 7(2), 140-158.

Lu Corresponding author, D., Mausel, P., Brondízio, E., Moran, E., 2004. Change detection techniques. *International Journal of Remote Sensing*, 25(12), 2365–2401. https://doi.org/10.1080/0143116031000139863.

Parelius, E. J., 2023. A Review of Deep-Learning Methods for Change Detection in Multispectral Remote Sensing Images. *Remote Sensing*, 15(8). https://www.mdpi.com/2072-4292/15/8/2092.

Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L., Lerer, A., 2017. Automatic Differentiation in PyTorch. *arXiv preprint arXiv:1710.10903*.

Schönfeld, E., Sushko, V., Zhang, D., Gall, J., Schiele, B., Khoreva, A., 2021. You only need adversarial supervision for semantic image synthesis. *International Conference on Learning Representations*.

Seo, D. K., Kim, Y., Eo, Y. D., Lee, M. H., Park, W. Y., 2018. Fusion of SAR and Multispectral Images Using Random Forest Regression for Change Detection. *ISPRS Int. J. Geo Inf.*, 7(10), 401.

Singh, A., 1989. Review Articlel: Digital change detection techniques using remotely-sensed data. *Int. J. Remote Sens.*, 10(6), 989–1003.

Touati, R., Mignotte, M., Dahmane, M., 2020. Multimodal Change Detection in Remote Sensing Images Using an Unsupervised Pixel Pairwise-Based Markov Random Field Model. *IEEE Trans. Image Process.*, 29, 757–767.

Xie, Z., Wang, M., Han, Y., Yang, D., 2016. Hierarchical decision tree for change detection using high resolution remote sensing images. Y. Xie, A. Zhang, H. Liu, L. Feng (eds), *Geoinformatics in Sustainable Ecosystem and Society*, Communications in Computer and Information Science, 980, Springer, 176–184.

Zhan, Y., Fu, K., Yan, M., Sun, X., Wang, H., Qiu, X., 2017. Change Detection Based on Deep Siamese Convolutional Network for Optical Aerial Images. *IEEE Geosci. Remote. Sens. Lett.*, 14(10), 1845–1849.

Zhao, J., Li, J., Cheng, Y., Sim, T., Yan, S., Feng, J., 2018. Understanding humans in crowded scenes: Deep nested adversarial learning and A new benchmark for multi-human parsing. S. Boll, K. M. Lee, J. Luo, W. Zhu, H. Byun, C. W. Chen, R. Lienhart, T. Mei (eds), *2018 ACM Multimedia Conference on Multimedia Conference, MM 2018, Seoul, Republic of Korea, October 22-26, 2018*, ACM, 792–800.

Zhao, W., Mou, L., Chen, J., Bo, Y., Emery, W. J., 2020. Incorporating Metric Learning and Adversarial Network for Seasonal Invariant Change Detection. *IEEE Trans. Geosci. Remote. Sens.*, 58(4), 2720–2731.

Zhou, B., Zhao, H., Puig, X., Fidler, S., Barriuso, A., Torralba, A., 2017. Scene parsing through ade20k dataset. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 5122–5130.

Zhou, L., Cao, G., Li, Y., Shang, Y., 2016. Change Detection Based on Conditional Random Field With Region Connection Constraints in High-Resolution Remote Sensing Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.*, 9(8), 3478–3488.