Disaster Detection from SAR Images with Different **Off-Nadir Angles Using Unsupervised Image Translation**

Jian Song^{1,2}, Bruno Adriano² and Naoto Yokoya^{1,2,*}

¹The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8561, Japan ²RIKEN AIP, 1-4-1 Nihonbashi, Chuo-ku, Tokyo 103-0027, Japan

Abstract

Synthetic aperture radar (SAR) images observed at different off-nadir angles have different intensities, and change detection methods using difference images do not work well. This problem hinders emergency response when there is no archive data with a consistent off-nadir angle as emergency SAR observation. In this paper, we investigate unsupervised image translation methods based on generative adversarial networks and autoencoders to detect flood and landslide areas using SAR images observed at different off-nadir angles. Comprehensive experiments of disaster detection using ALOS-2 PALSAR-2 images for three floods and two landslides show that the developed methods can significantly improve the accuracy of disaster detection using pre- and post-disaster images observed at different off-nadir angles.

Keywords

Autoencoders, disaster detection, generative adversarial networks, image translation, off-nadir angle, synthetic aperture radar

1. Introduction

Synthetic aperture radar (SAR) can acquire images in bad weather conditions and at night, and thus it is effective in assessing the situation during emergency response. In particular, floods and landslides are often caused by heavy rainfall, and since the weather conditions are often bad, SAR images are more suitable than optical images for rapid detection of disaster areas in a wide area.

The straightforward method for detecting flood and landslide areas from SAR is image analysis [1], such as thresholding the difference between pre- and postdisaster images. This method assumes that the two images are observed at the same off-nadir angle. Recently, with the success of advanced machine learning algorithms, researchers have developed techniques using pairs of SAR data and deep convolutional neural networks [2, 3]. Although previous methods have excellent performance for detecting land changes, they still follow the primary assumption of image analysis: having a set of pre- and post-disaster images with almost similar acquisition conditions (e.g., off-nadir angle). Attempts to perform change detection analysis using SAR images with different acquisition conditions have been focused mainly on urban environments or forest areas by exploit-

*Corresponding author.

(B. Adriano); yokoya@k.u-tokyo.ac.jp (N. Yokoya)

https://brunoadriano.com/ (B. Adriano);

ing the double-bounce features of SAR images [4, 5, 6]. However, considering that floods primarily occur on plain surfaces, where the SAR's double-bounce is almost negligible, we can not apply the previous techniques directly to assess flood extent assessment.

In this paper, we study disaster area detection using SAR images with different off-nadir angles by unsupervised image translation. Two methods are developed, one that accounts for changes in the ground surface and one that does not. We conduct experiments on several datasets of flood and landslide disasters and investigate the characteristics of these methods with respect to the magnitude of the difference in off-nadir angles.

2. Image Translation for Change Detection

To cope with nonlinear differences between two period images caused by differences in off-nadir angles, we investigate unsupervised image translation based on convolutional neural networks (CNNs). Encoder-decoder neural networks are applied to translate one image into an image obtained at the same off-nadir angle as the other image. The following subsections describe two methods investigated in this paper: 1) image translation based on conditional generative adversarial networks (cGAN) [7, 8] that does not consider the presence or absence of change, and 2) image translation based on code-aligned autoencoders [9] that takes into account surface change.

CDCEO 2022: 2nd Workshop on Complex Data Challenges in Earth Observation, July 25, 2022, Vienna, Austria

Song@ms.k.u-tokyo.ac.jp (J. Song); bruno.adriano@riken.jp

https://naotoyokoya.com/ (N. Yokoya)

D 0000-0002-0877-7063 (J. Song); 0000-0002-4318-4319

⁽B. Adriano); 0000-0002-7321-4590 (N. Yokoya) • 2022 copyright for this paper by its authors. Use permitted under Creative Commons License Arithmitun 40 International (CC BV 4.0). CEUR Workshop Proceedings (CEUR-WS.org)



Figure 1: cGAN-based image translation without consideration of change.



Figure 2: Code-aligned autoencoders (CAA).

2.1. Conditional Generative Adversarial Networks

Figure 1 shows an overview of the cGAN-based image translation method that does not consider the presence or absence of change. The developed method consists of a generator with an encoder-decoder structure (i.e., U-Net [10]) to translate low-pass filtered image 1 into low-pass filtered image 2 during the training phase, and a discriminator to determine whether the transformed image is the true low-pass filtered image 2. The generator learns to trick the discriminator and the discriminator learns to detect that the translated image is false. In the inference stage, we discard the discriminator and apply the generator to the original image 1 to convert to image 2 with the same off-nadir angle. The method assumes that the low-pass filtered image pair contains relatively slight changes in the ground surface. Thus, differences in images due to differences in off-nadir angles are somewhat easier to transform using a U-Net than differences in images due to changes in the ground surface; therefore, the generator only achieves the translation corresponding to the former. The difference image between translated image 1 and image 2 is subjected to thresholding to detect the change area.

2.2. Code-Aligned Autoencoders

We extend the image translation-based change detection method proposed by Luppino, et al. [9] that takes changes into account. Figure 2 shows an overview of the method. We train two different encoder-decoders for each disaster image pair by switching decoders. One is used to transform the pre-disaster image $X \in \mathbb{R}^{h \times w}$ to the post-disaster image $Y \in \mathbb{R}^{h \times w}$ with the same off-nadir angle, and the other is used to transform the post-disaster $Y \in \mathbb{R}^{h \times w}$ image to the pre-disaster image $X \in \mathbb{R}^{h \times w}$ with the same offnadir angle. This dual encoder-decoder structure helps to increase the robustness of the final change detection. *h* and *w* are the height and width of the image used when training the CNNs. Each CNN consists of an encoder (E_X , $E_{\rm Y}$) that extracts semantic information about the ground surface and a decoder (D_X, D_Y) that restores the style of the image (including the effect of off-nadir angles). The features (Z_X, Z_Y) extracted for each image are passed to the other decoder to transform the image style. To achieve unsupervised learning, the final loss function is a combination of the four loss functions described below. The differences from [9] are that the loss function is changed from the L2 norm to the L1 norm and that image modulation based on the Fourier transform is utilized for preprocessing.

Reconstruction loss: The solid black line process in Figure 2 indicates the image reconstruction; the reconstruction loss is the *L*1 norm of the difference between the inputs and outputs so that each encoder-decoder network acts as an autoencoder that outputs the same image as the one used for input:

$$L_r = \mathbb{E}_X[\|\tilde{X} - X\|_1] + \mathbb{E}_Y[\|\tilde{Y} - Y\|_1], \quad (1)$$

where $\tilde{X} = D_X(E_X(X))$ and $\tilde{Y} = D_Y(E_Y(Y))$.

Weighted translation loss: The following weighted translation loss is considered so that the result of image translation by switching decoders matches the other image in areas where there is no change in the ground surface:

$$L_t = \mathbb{E}_X[\|M \odot (\hat{X} - X)\|_1] + \mathbb{E}_Y[\|M \odot (\hat{Y} - Y)\|_1], \quad (2)$$

where $\hat{X} = D_X(E_Y(Y))$, $\hat{Y} = D_Y(E_X(X))$, *M* is a weight matrix that approaches 0 for pixels with ground surface changes and 1 for pixels without changes, and \odot is an element-wise product.

Cycle-consistency loss: Two iterations of an image translation should match the original image, which can be implemented by the following cycle consistency loss:

$$L_{c} = \mathbb{E}_{X}[\|\bar{X} - X\|_{1}] + \mathbb{E}_{Y}[\|\bar{Y} - Y\|_{1}], \qquad (3)$$

where $\bar{X} = D_X(E_Y(\hat{Y}))$ and $\bar{Y} = D_Y(E_X(\hat{X}))$.

Code correlation loss: All three losses above represent the conditions that the translated image must satisfy as loss functions, but no explicit regularization is given as to what features the encoder should extract. We consider the following loss function so that the information about the change in correlation between all pixels in each input image is reflected in the features to be extracted by the encoder:

$$L_{z} = \mathbb{E}_{X,Y}[\|R - S\|_{1}], \tag{4}$$

where the correlation matrix $S \in \mathbb{R}^{n \times n}$ $(n = h \times w)$ is defined by

$$S_{i,j} = 1 - \frac{1}{\sqrt{n}} \|A_i^X - A_j^Y\|_2 \quad i, j \in \{1, ..., n\},$$

$$A_i^X = [A_{i,1}^X, ..., A_{i,n}^X], \quad A_j^Y = [A_{j,1}^Y, ..., A_{j,n}^Y], \quad (5)$$

$$A_{i,j}^X = \exp(-(X_i - X_j)^2), \quad A_{i,j}^Y = \exp(-(Y_i - Y_j)^2).$$

The correlation matrix $R \in \mathbb{R}^{n \times n}$ $(n = h \times w)$ is defined by

$$R_{i,j} = \frac{(Z_i^X)^T Z_j^Y + Z_{max}}{2Z_{max}} \quad i, j \in \{1, ..., n\},$$

$$Z_{max} = \max_{i \in \{1, ..., n\}} \{ \| Z_i^X \|, \| Z_i^Y \| \}.$$
(6)

Correlations between pixels of input images are assumed to reflect information about the ground surface rather than observational conditions such as off-nadir angles, because they are not affected by changes in the overall brightness of the image. So with L_z we can restrict the features to have a similar correlation to force the encoder to output the features we want.

The final loss function is defined as the linear sum of the above four loss functions as follows: $L = \lambda_r L_r + \lambda_t L_t + \lambda_c L_c + \lambda_z L_z$. The matrix *M* is initialized to a 0 matrix in the initial stage of training, i.e., only L_r , L_c and L_z are involved in the optimization. Along with updating the parameters of the CNNs, the weight matrix *M*, which represents the presence or absence of change, is also updated according to the following equation:

$$M = 1 - \delta,$$

$$\delta = \frac{d(\hat{X}, X) + d(\hat{Y}, Y)}{2},$$
(7)

where d denotes the calculation of the difference map by setting the threshold value. Partially cropped patches of the image to be analyzed are used to train the CNN models. After training, image translation is applied to the entire image, and Otsu's method [11] is applied to the difference image for final change detection.

2.3. Image Modulation Based on Fourier Transform

The success of the CAA algorithm depends on whether or not it can proceed from a reasonable weight matrix M in the initial stage of learning. If the differences in off-nadir angles are large, there is a problem of not converging to a good local optimal solution. Therefore, as a pre-processing step, we use as the input image for the



Figure 3: Image modulation based on Fourier transform.



Figure 4: Example of Fourier transform-based image modulation for Kinugawa (9.12) data. (Best viewed with zoom in.)

CAA algorithm an image in which the low-frequency component of one image (pre-disaster image) is made closer to the low-frequency component of the other image (post-disaster image), thereby reducing the difference between the two images in terms of overall intensity.

Figure 3 shows an overview of image modulation based on the Fourier transform. First, a two-dimensional Fourier transform is applied to each image to separate the amplitude and phase information. Next, the amplitude of the low-frequency component of *image 1* is replaced by the amplitude of the low-frequency component of *image* 2. The simplest method for replacing the amplitude is to use a two-dimensional rectangular function to replace components below a certain frequency, but in this paper, a linear summation using a two-dimensional Gaussian function was employed for more natural image modulation. The standard deviation (σ) of the Gaussian function is a parameter that adjusts the degree to which highfrequency components are modulated. Finally, the modulated image is obtained by applying a two-dimensional inverse Fourier transform to the phase of *image 1* and the amplitude modulated by the linear sum. Figure 4 shows an example of image modulation based on the Fourier transform. It can be seen that after modulation, the brightness of the entire image approaches that of image 2, but the detailed pattern of the image is inherited from *image* 1.

Table 1	
Observation conditions and image size of ALOS-2 PALSAR-2 images for seven d	latasets.

Disaster Dataset name	Acquisition date	Orbit direction	Observation direction	Off-nadir angle	lmage size pixels (km)
Sep. 2015 Kanto/Tohoku heavy rainfall <i>Kinugawa (9.12)</i>	2015.09.12 2014.10.06 2014.09.13	Left Left Left	Descending Descending Descending	35.4 25.6 35.4	1504×3656 (7.5×18.3)
Sep. 2015 Kanto/Tohoku heavy rainfall <i>Kinugawa (9.13)</i>	2015.09.13 2014.09.13	Left Left	Descending Descending	52.1 35.4	1490×3479 (7.5×17.4)
Sep. 2015 Kanto/Tohoku heavy rainfall <i>Kinugawa (9.16)</i>	2015.09.16 2019.06.29	Left Left	Descending Descending	13.9 35.4	1490×3479 (7.5×17.4)
July 2018 heavy rainfall <i>Kurashiki</i>	2018.07.07 2018.06.14 2018.04.04	Right Right Right	Ascending Ascending Ascending	38.7 45.1 38.7	3081×2983 (21.9×21.2)
July 2020 heavy rainfall Saga	2020.07.04 2020.06.08 2016.04.16	Left Right Left	Descending Descending Descending	50.5 32.8 50.5	8224×2720 (16.4×5.4)
July 2018 heavy rainfall Hiroshima	2018.07.08 2018.06.20 2015.04.19	Right Right Right	Descending Descending Descending	48 35.4 48	3117×2603 (7.8×6.5)
2018 Hokkaido Eastern Iburi earthquake <i>Hokkaido</i>	2018.09.06 2017.10.28 2018.08.23	Left Left Left	Ascending Ascending Ascending	38.2 42.7 38.2	6134×5904 (12.2×11.8)

3. Experiments

3.1. Data

Experiments on disaster area detection are conducted using ALOS-2 PALSAR-2 imagery of three floods (the September 2015 Kanto/Tohoku heavy rainfall, the July 2018 heavy rainfall, and the July 2020 heavy rainfall) and two landslide disasters (the July 2018 heavy rainfall and the 2018 Hokkaido Eastern Iburi earthquake) in Japan. Binary polygon data of inundation areas and landslide disaster areas published by the Geospatial Information Authority of Japan (GSI) are used as the ground truth data. For the September 2015 Kanto/Tohoku heavy rainfall, there are images observed on three different days after the disaster and the corresponding ground truth data, so they are treated as three independent experiments. Therefore, there are seven experimental datasets in total.

Table 1 shows the observation conditions and image sizes for the seven datasets composed of 19 ALOS-2 PALSAR-2 images used in this study. For those datasets having 3 images, we have a pair of images with the same off-nadir angle before and after the disaster. The results of change detection using the same off-nadir angle images are used as reference information to evaluate the accuracy of the developed methods. In all experiments, images with speckle noise reduced by bilateral filtering were used as input.

3.2. Evaluation Method

The F-score is used as a metric to evaluate the accuracy of change detection. The effectiveness of the investigated methods is verified by comparing the following five methods.

- Thresholding of difference images of the same off-nadir angle images; it will be referred to as *Reference*
- Thresholding of difference images of different offnadir angle images
- cGAN-based image translation
- CAA-based image translation
- CAA-based image translation using image modulation based on Fourier transforms as preprocessing (denoted as CAA-FT)

In order to evaluate the effectiveness of image translation for change detection, the change detection method in this study was unified into the simplest thresholding method. Thresholding is applied to the case of reduced pixel values in the case of floods, and to both increased and decreased pixel values in the case of landslides.

3.3. Results

Table 2 shows the quantitative evaluation results of disaster area detection using F-scores for the seven experiments. Figures 5 and 6 show the visual results of flood

Quantitative evaluation of disaster area detection (F-score)). Bold type indicates the best performance of the method using
images with different off-nadir angles.	

Angle diff.	Kinugawa (9.12) 9.8	Kinugawa (9.13) 16.7	Kinugawa (9.16) 21.5	Kurashiki 6.4	Saga 17.7	Hiroshima 12.6	Hokkaido 4.5
Reference	0.6551	_	_	0.4216	0.2065	0.0849	0.2967
Threshold	0.5187	0.1322	0.0560	0.3410	0.0798	0.0825	0.2861
cGAN	0.6446	0.3843	0.1331	0.4046	0.1341	0.0620	0.1962
CAA	0.5739	0.4672	0.4087	0.3783	0.3037	0.1535	0.2010
CAA-FT	0.3936	0.5142	0.3767	0.3272	0.3463	0.1646	0.2659



Figure 5: Flood detection results in Kinugawa (2015.09.12) during the September 2015 Kanto/Tohoku heavy rainfall.

detection in the experiments using the Kinugawa (9.12) and Saga datasets, respectively. Compared to thresholding the difference images of different off-nadir angle images, the investigated methods provide superior detection results in the six experiments. More detailed results and discussion on flood detection and landslide detection are presented below.

3.3.1. Flood detection

For flood detection, the F-scores of around 0.35 to 0.65 were achieved, which can be considered to be of practical accuracy considering that the simplest change detection method of thresholding is used. For the Kurashiki and Kinugawa (9.12) datasets, the F-scores are comparable to



Figure 6: Flood detection results in Saga during the July 2020 heavy rainfall.

Table 2

that of using the same off-nadir angle images, indicating that our objective is achieved. For the Saga dataset, the Fscores of the investigated methods exceeded the reference value, which may be due to the fact that the seasonal difference for the same off-nadir angle images is larger than that for different off-nadir angles. A comparison of the performance among the methods under investigation in this paper reveals the following trends.

- When the difference in off-nadir angles is about 5° to 10°, as in Kurashiki and Kinugawa (9.12), the cGAN-based image translation that does not consider the presence of change is effective.
- When the difference of off-nadir angles is about 10°to 25°, as in Kinugawa (9.13), Kinugawa (9.16), and Saga, the CAA-based image translation focusing on the unchanged area is effective.
- If the difference in off-nadir angles is not large, image modulation based on the Fourier transform is not necessary for CAA.

The above trends are consistent with the conditions and design principles assumed by each method, suggesting that reasonable experimental results were obtained. The cGAN-based method assumes that image differences due to differences in off-nadir angles are relatively easier to transform using a CNN than image differences due to changes in the ground surface, and this condition is considered to be satisfied when the differences in offnadir angles are relatively small. If this assumption is not satisfied, i.e., the differences in images due to differences in off-nadir angles are relatively large, then the CAAbased methods, which perform image translation while taking into account the presence or absence of changes in the ground surface, is effective. Furthermore, it was shown experimentally that the pre-processing based on the Fourier transform contributed to the improvement of accuracy by reducing the difficulty of the problem.

3.3.2. Landslide detection

In the experiment for detecting landslides using the Hiroshima dataset, the CAA-based method showed a clear improvement in accuracy. In the Hokkaido experiment, the thresholding results for the difference images of different off-nadir angle images were almost the same as the reference value using the same off-nadir angle images, and no further improvement could be obtained by the investigated methods. This may be due to the fact that the difference in off-nadir angles is so small (4.5°) that its effect on the observed image is negligible. Although the presented methods did not improve the accuracy of change detection, this is a reasonable result.

In the experiment for detecting landslides, the F-score was generally low. This is because the detection of landslide areas is more difficult than the detection of inundation areas, and the simple thresholding process cannot

Table 3

Off-nadir angle conditions and the effectiveness of the methods under investigation. One, two, or three • signs mean low, mideum, and high.

Angle diff.	cGAN	CAA	CAA-FT
5-10°	•••	••	•
10-25°	•	••	•••

achieve sufficient detection accuracy even when using the same off-nadir angle images. In [12], F-scores of 0.33 and 0.58 were achieved for Hiroshima and Hokkaido, respectively, using an algorithm specialized for landslides detection. In this paper, the simplest thresholding was used as a basis to evaluate the effectiveness of change detection based on image translation. However, it is expected that the combination of landslide detection algorithms that take land cover and topography into account and the image translation methods investigated in this paper will improve the detection accuracy of landslide areas to a practical level.

4. Conclusion and Future Outlook

We studied the potential of image translation via unsupervised deep learning to detect flood and landslide areas using SAR images with different off-nadir angles. We investigated a cGAN-based method that does not consider changes and a CAA-based method that focuses on unchanged areas of the ground surface during image translation. Experiments on seven cases of flood and landslides disasters demonstrate the effectiveness of the development methods.

Table 3 summarizes the relationship between off-nadir angle conditions and the effectiveness of the methods under investigation for flood detection. When the difference in off-nadir angle is about 5° to 10°, cGAN-based image translation without considering change is effective, and when the difference in off-nadir angle is about 10° to 25°, CAA-based image translation focusing unchanged areas can significantly improve the accuracy of flood detection. Cases with larger differences in off-nadir angles (greater than 25°) are not included in the experiments and require further verification.

The overall accuracy in detecting landslide areas was low because the simplest method of change detection, thresholding of difference images, was used. Combining landslide detection algorithms that take land cover and topography into account with image translation is a future challenge.

Acknowledgments

The authors would like to thank the Japan Aerospace Exploration Agency for providing ALOS-2 PALSAR-2 data.

References

- [1] L. Landuyt, A. Van Wesemael, G. J. Schumann, R. Hostache, N. E. Verhoest, F. M. Van Coillie, Flood Mapping Based on Synthetic Aperture Radar: An Assessment of Established Approaches, IEEE Transactions on Geoscience and Remote Sensing 57 (2019) 722–739. doi:10.1109/TGRS.2018.2860054.
- [2] Y. Li, S. Martinis, M. Wieland, Urban flood mapping with an active self-learning convolutional neural network based on TerraSAR-X intensity and interferometric coherence, ISPRS Journal of Photogrammetry and Remote Sensing 152 (2019) 178–191. doi:10.1016/j.isprs.jprs.2019.04.014.
- [3] X. Li, Z. Du, Y. Huang, Z. Tan, A deep translation (gan) based change detection network for optical and sar remote sensing images, ISPRS Journal of Photogrammetry and Remote Sensing 179 (2021) 14–34. URL: https://www.sciencedirect.com/ science/article/pii/S0924271621001842. doi:https: //doi.org/10.1016/j.isprsjprs.2021.07.007.
- [4] P. Ge, H. Gokon, , K. Meguro, Building damage assessment using intensity sar data with different incidence angles and longtime interval, Journal of Disaster Research 14 (2019) 456–465. doi:10.20965/ jdr.2019.p0456.
- [5] J. Tao, S. Auer, P. Reinartz, R. Bamler, Object-based change detection for individual buildings in sar images captured with different incidence angles, 2013 IEEE International Geoscience and Remote Sensing Symposium - IGARSS (2013) 1238–1241.
- [6] M. W. Lang, P. A. Townsend, E. S. Kasischke, Influence of incidence angle on detecting flooded forests using c-hh synthetic aperture radar data, Remote Sensing of Environment 112 (2008) 3898–3907. URL: https://www.sciencedirect.com/ science/article/pii/S0034425708002228. doi:https: //doi.org/10.1016/j.rse.2008.06.013.
- [7] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, Advances in neural information processing systems 27 (2014).
- [8] P. Isola, J.-Y. Zhu, T. Zhou, A. A. Efros, Imageto-image translation with conditional adversarial networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 1125–1134.
- [9] L. T. Luppino, M. A. Hansen, M. Kampffmeyer, F. M.

Bianchi, G. Moser, R. Jenssen, S. N. Anfinsen, Codealigned autoencoders for unsupervised change detection in multimodal remote sensing images, IEEE Transactions on Neural Networks and Learning Systems (2022).

- [10] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical image computing and computer-assisted intervention, Springer, 2015, pp. 234–241.
- [11] N. Otsu, A threshold selection method from graylevel histograms, IEEE transactions on systems, man, and cybernetics 9 (1979) 62–66.
- [12] B. Adriano, N. Yokoya, H. Miura, M. Matsuoka, S. Koshimura, A semiautomatic pixel-object method for detecting landslides using multitemporal alos-2 intensity images, Remote Sensing 12 (2020) 561.