ISSLIDE: A new InSAR dataset for Slow SLIding area DEtection with machine learning

Antoine Bralet, Emmanuel Trouvé, Jocelyn Chanussot, Abdourrahmane Atto

ISSLIDE: A new InSAR dataset for Slow SLIding area DEtection with machine learning

Antoine BRALET † Emmanuel TROUVÉ † Jocelyn CHANUSSOT ⋆ Abdourrahmane M. ATTO †

†LISTIC, Université Savoie Mont Blanc, 74940 Annecy, France
⋆Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, Grenoble, 38000, France

***Abstract*—Due to the high data demand of machine learning algorithms, multiple datasets are emerging in remote sensing. But these datasets are costly and time consuming to annotate especially for change detection or natural phenomena monitoring. In particular, early warning systems on slow-moving disasters are lacking of training datasets as they require both geomorphological and SAR interferometry expertise. In this paper, (i) we propose a novel InSAR dataset for Slow SLIding area DEtection (ISSLIDE) with machine learning algorithms. The latter consists of manually annotated patches of generated interferograms over slow moving areas. (ii) We implement the segmentation of ISSLIDE interferograms with classical deep learning approaches. FCN, DeepLabV3 and U-Net-like architectures are explored to serve as baseline for future works. To the best of our knowledge, this is the first dataset adapted to machine learning and targeting slow sliding area detection.**

***Index Terms*—Dataset, InSAR, Deep learning, ISSLIDE, Ground motions**

I. Introduction

Deep learning approaches are increasingly used for Earth observation and monitoring due to their strong ability in terms of performances, reliability and inference computation time. These approaches are especially applied on satellite images to benefit from their wide spatial coverage, their high revisit frequency and their multiple available sensors (*e.g.* SAR, multispectral or hyperspectral). Thus, to leverage both remote sensing characteristics and deep learning efficiency, more and more multimodal datasets are created [1]–[5] to ensure relevant trainings. If getting access to remote sensing images is straightforward with the Copernicus program, annotating this amount of data is not only costly and time consuming but requires also strong expert knowledge. This is the reason why land-cover classification and change detection datasets are often based respectively on pre-existing maps [6]–[8] or inventories [4], [5], [9] for labelling. Early warning systems are more challenging and, in particular, detecting ground deformations or motions requires both geomorphological expertise and radar interferometry knowledge. Some studies are emerging to predict volcanoes unresting periods and prevent catastrophic tolls by using interferometric SAR (InSAR) [10]. But to the best of our knowledge, there is no dataset targeting ground motion detection for deep approaches. Our contribution summarized in Figure 1 is two-fold:

1) We release a **novel InSAR dataset for Slow SLIding area DEtection (ISSLIDE) with machine learning**. A total 200 French Alps ground motions are manually annotated ending up with 13,230 samples of interferograms ready for deep learning training purpose. The dataset is available on IEEE Dataport (https://ieee-dataport.org/documents/isslide-insar-dataset-slow-sliding-area-detection-machine-learning).
2) We propose a **first deep segmentation study applied on ISSLIDE samples**. Different configurations of FCN [11], DeepLabV3 [12] and U-Net-like models [13] are compared to establish primary baseline results.

The remaining of the paper reviews related works in Section II before introducing the ISSLIDE dataset in Section III. Section IV details deep segmentation baselines and their performances. Section V concludes and offers perspectives.

II. Related Works

In remote sensing field, applications such as super-resolution [14], despeckling of radar images [15] or modality translation [16] do not need any annotations. But other tasks as land-cover mapping [17] require specific labels often based on pre-existing maps or inventories. Authors of BigEarthNet-MM [1] extracted 590,326 pairs of Sentinel-1 and 2 patches and used the CORINE land-cover map [6] for labelling. SEN12MS [2] is composed of 180,662 Sentinel image couples annotated with the MODIS [7] land-cover map. Authors of SpaceNet-6 [3] computed labelling with the 3DBAG [8] land-cover map on 3,401 radar-optical couples respectively from Capella Space & Metasensing and Maxar Worldview-2.

Another typical application on remote sensing images is the detection and monitoring of landscape changes. But tasks as flood [18], forest fire [18] and landslide detection [19] or the evaluation of their damage on buildings [20] require more expertise in the annotation process. The xBD dataset [21] contains 700,000 annotated buildings based on their post-disaster destruction level. The Landslide4Sense [19] competition provides 3,799 annotated optical patches for post-event landslide detection. The dataset from [9] contains 1,918 landslides from the United States Geological Survey to implement their detection on optical images. Authors of SEN12-FLOOD [4] gathered Sentinel-1 and 2 images as well as labels provided by Copernicus Emergency Management Service to create a flood dedicated multimodal dataset with 412 time series. The same labelling source is used in MMFlood [5] for flood delineation based on 1,748 triplets of Sentinel-1 images, Digital Elevation Model (DEM) and hydrography maps.

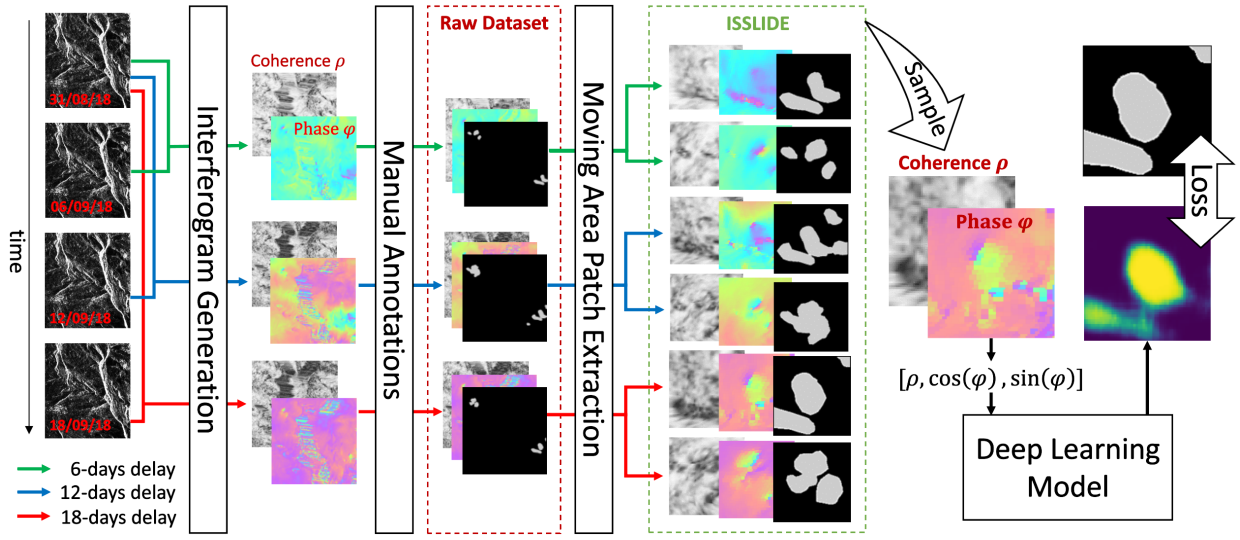


Fig. 1: Flowchart of the generation and use of ISSLIDE. A collection of successive radar images is used to generate 6-, 12- and 18-days delay interferograms. These interferograms are manually annotated to release the *raw dataset*. Every move is extracted within a patch to create the ISSLIDE dataset. Delay-independent samples are then selected to train the neural network inputting the coherence (ρ), the cosine and the sine of the phase difference (φ). The output probability map is then compared to the manual annotation to optimise the network.

These sudden natural disasters are especially studied for post-event policies, but early warning systems and monitoring of slow-moving phenomena are of crucial importance for disaster forecasting. Their objective is to localize deformations which are invisible by offset tracking from standard satellite resolutions (*e.g.* 10m for Sentinel constellations). This issue is tackled by studying the phase of radar images through InSAR-based approaches [22]. If authors of [23] propose to compute interferograms on large zones with advanced algorithms, they do not provide any annotations. In contrary, authors of [24] make an inventory of slow-moving areas as rock glaciers and landslides without providing a dataset adapted to machine learning trainings. Hephaestus [10] is one of the few annotated InSAR datasets for volcanoes unresting period detection. It is composed of 216,106 samples among which only 2,247 contains volcanic deformations. The InSAR dataset proposed in this work extends deformation studies to ground motions as slow landslides and rock glaciers.

III. ISSLIDE Dataset

A. Interferogram Generation

Interferograms are generated based on Copernicus Sentinel-1 SAR images provided by the ESA. Before 2021, this two-satellites constellation released a new image of a same zone every six days over Europe. To benefit from this high revisit frequency, SAR images are selected every six days from the 2nd of July 2018 to the 24th of October 2018. Summer season is favoured to keep a high coherence between radar images and avoid non-coherent snowing periods. Images are Level-1 Single Look Complex products acquired on descending pass in interferometric wide swath mode. Only the VV polarization is used to ensure a better coherence between the acquisitions.

Interferograms are generated with the Sentinel Application Platform (SNAP)[1] and based on their standard "Command-Line InSAR Tutorial". The following process does not require any additional information as atmospherical predictions or meteorological assessments: the freely available SRTM [25] 1sec HGT DEM is the only requirement at this point.

The generation procedure is divided is 8 successive steps:

- Application of σ_0-calibration to the VV band;
- Orbital distortion corrections;
- Co-registration of the images with respect to the DEM;
- Interferogram generation;
- Topographic phase correction with respect to the DEM;
- Goldstein Filtering for interferogram smoothing;
- Multilooking of size 8×2 (range×azimuth);
- Orthorectification with the WGS84 projection map.

These steps result in the coherence and phase difference maps provided in the ISSLIDE dataset.

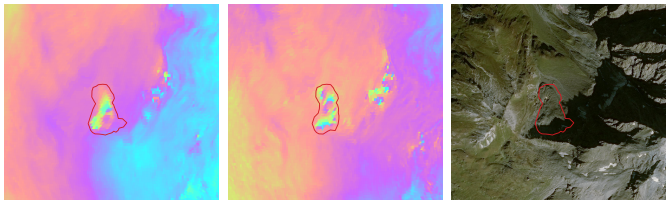
B. Annotation Generation

Annotation process is set up in collaboration with geomorphologists to establish robust identification strategies. Every following steps are computed with the QGIS software[2] on each generated interferogram.

Manual annotations are based on the research of phase difference odd patterns. Two of them are targeted: bubbles and fringes as illustrated in Figure 2. *Bubbles* (Figure 2a) are the most common patterns characterizing slow-moving phenomena as landslides and rock glaciers. Compared to the stable background phase difference, they show significant phase changes slowly attenuated when reaching their borders. This is revealing a fast moving area triggering aside earth to follow. The second pattern of interest is *fringes* (Figure 2b). They appear when continuous displacements occur and when a 2π delay is created between the stable background and the moving area. It creates a cyclic pattern ranging from $-\pi$ to π.

[1]https://step.esa.int/main/toolboxes/snap/, visited on the 6th of July 2023.

[2]https://www.qgis.org/fr/site/index.html, visited on the 6th of July 2023.



(a) Bubble pattern (b) Fringe pattern (c) Optical image

Fig. 2: Example of moving areas patterns in the phase difference map of a same zone at (a) 6 days and (b) 18 days of delay.

Fringes appear if the phenomenon is fast enough compared to the interferogram delay. This is illustrated in Figure 2 where a bubble pattern in a 6-days interferogram (Figure 2a) changes into a fringe pattern in the 18-days interferogram (Figure 2b).

To avoid incorrect annotations, phase difference images are masked with a layover map. Each suspected pattern is then confirmed or refuted based on three major criteria. First, the coherence map is used to ensure phase difference reliability. Indeed, the ratio between the targeted displacement range (cm) and the image pixel resolution (m) prevents from coherence losses between the two acquisition dates. Therefore, sufficiently high coherence is expected at least in the surroundings to validate the presence of a ground motion. The magenta circled area in Figure 3 illustrates a typical non coherent suspicious pattern to avoid. The second indicator is the repetitiveness in time of the movement. Indeed, targeted phenomena are both rock glaciers and landslides. The first ones are moving due to the melting of the underground ice which results in a continuously moving zone. On the other hand, landslides are expected to move for several weeks between two possible deactivation periods. Thus, both phenomena should have visible patterns on successive interferograms. The last criterion is based on an optical assessment from SPOT6-7 images in 2018. This final verification removes nonsense areas (*e.g.* forests or lakes) and identify optical patterns such as downhill earth accumulation (see Figure 2c) or fallen rocks.

C. Final Dataset

Three specific French Alps regions are selected to be manually annotated, namely the Vanoise mountains, the north-east of the Écrin Park and the Queyras Park. Additional moving areas are identified outside of these areas to be used as a more variable source of data for the training. The dataset is composed of 200 different phenomena identified on 19, 18 and 17 interferograms with a respective delay of 6, 12 and 18 days. Every single interferogram is manually annotated so that each phenomenon has a distinct segmentation on each interferogram. A total of 13,230 images surrounding moving areas are extracted from the interferograms with their respective segmentation. 8,701 images show actual moving periods while the remaining 4,529 are provided for multi-temporal studies. The dataset is released under different configurations.

1) Raw dataset: The raw dataset contains the 54 interferograms used for the annotation process with the 54 shapefiles identifying moving areas within them. A given shapefile contains not only the annotated polygons in its corresponding interferogram but also a confidence index, the number of repetitions among the X-days interferograms as well as a movement class: *Landslides*, *Rock Glaciers* or *Unknown phenomena*. The raw dataset is released for future users to process their own extraction, use available attributes for training or enrich the dataset with more annotations.

2) Ready to be used dataset: The second dataset is already cropped around the annotated moves, merging all classes into a single *Moving Area* class for detection purpose. Based on a global annotation map - including all annotations at all dates for every delays - we used a distance transformation to extract moving areas coordinates. Each interferogram and its respective annotation map is then cropped around these coordinates to extract patches of size 100×100px. For multitemporal monitoring approaches, samples of interferograms with movements showing a deactivation period are kept but associated with a "no-move" annotation maps (only 0 values).

IV. Learning on ISSLIDE

A. Baseline models

To study the performances of approaches from the literature, we used the FCN [11] and DeepLabV3 [12] architectures with a ResNet50 [26] backbone already implemented in Pytorch libraries. These models are both trained from scratch (Model$_S$) and with available PASCAL VOC dataset pre-trained weights (Model$_P$).

Three shallower networks are also implemented to explore their relevancy in comparison with the deep ones from the literature. These U-Net-like [13] architectures differ in their encoding strategy - classical convolutional blocks [13], residual blocks [26] or separable convolutional blocks [27]. They are respectively named as UNet, ResUNet and SepUNet.

B. Experimental setup

For generalizability purpose, the whole dataset is split in training and evaluation sets. The first one contains the patches extracted on a random 80% selection of the interferograms in every zone except the Queyras Park which is reserved for the evaluation set. The latter is thus composed of the samples from the 20% remaining interferograms. For this single interferogram study, only images with visible moving areas are kept. Resulting datasets contain respectively 4,821 and 572 images. Data augmentation is applied for phase shifting invariance by introducing a random additional offset to the phase difference map.

As illustrated in Figure 1, all networks input the coherence map, the cosine and the sine of the phase difference map. They are optimized with a binary cross entropy loss and an Adam optimizer on mini-batches of size 8. The three shallow networks are trained with a learning rate of 1.10^{-5} for 300 epochs. To avoid overfitting with FCN [11] and DeepLabV3 [12], a learning rate of 1.10^{-7} is used for 600 epochs.

In Section IV-C, five metrics are used to evaluate the performances of the networks. The Hausdorff Distance (HD) is used as a distance metric, the Dice and the Intersection over Union (IoU) scores as ensemble metrics and the F1-score (F1-S) and the Area Under the Receiver Operating Characteristic Curve (AUC) as classification metrics. Metrics are calculated

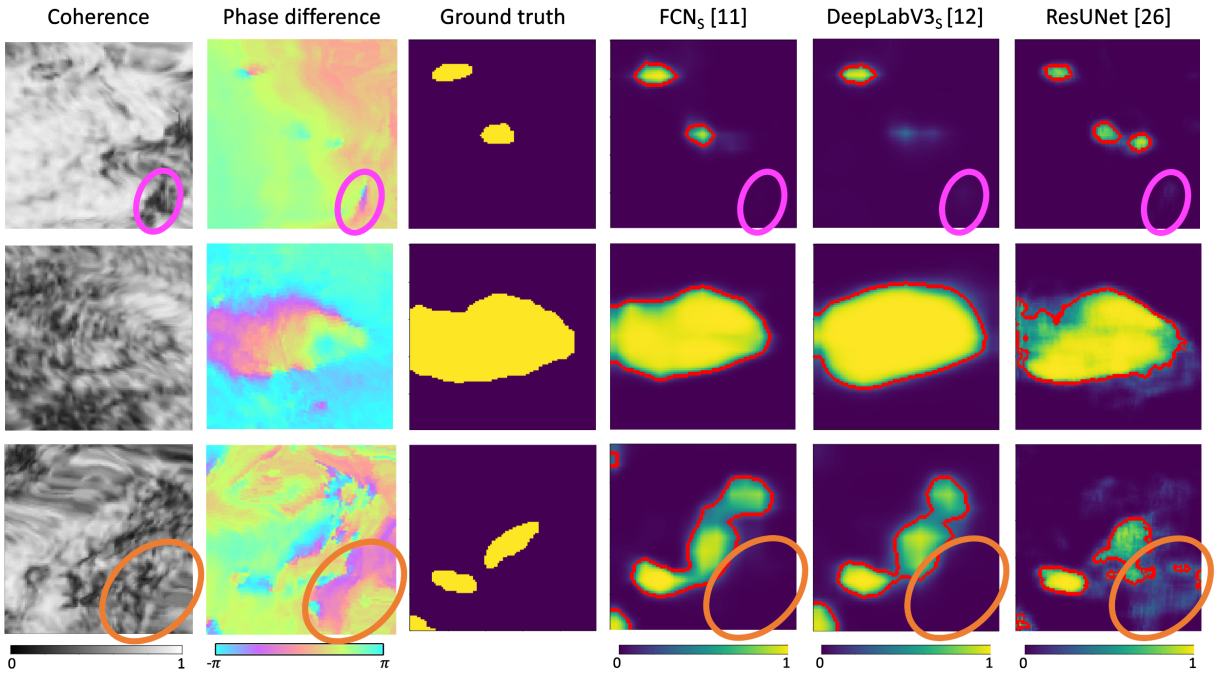


Fig. 3: Qualitative comparison of the predictions of FCN$_S$ [11], DeepLabV3$_S$ [12] and ResUNet [26] which shown the best quantitative results. Red borders denote the segmentation used to calculate Table I results. Magenta circles illustrate a suspicious pattern in the phase difference map but with a low coherence identifying an unreliable pattern. Orange circles identify an area with strong layover distortions.

	F1-S ↑	AUC ↑	Dice ↑	IoU ↑	HD ↓
FCN$_S$ [11]	**0.690**	**0.980**	**0.645**	**0.503**	**21.695**
FCN$_P$ [11]	0.677	0.976	0.628	0.486	22.951
DeepLabV3$_S$ [12]	0.673	0.978	0.630	0.486	<u>22.352</u>
DeepLabV3$_P$ [12]	0.666	0.976	0.620	0.480	22.887
UNet [13]	0.675	0.977	0.636	0.497	25.912
ResUNet [26]	<u>0.681</u>	<u>0.979</u>	<u>0.639</u>	<u>0.500</u>	25.921
SepUNet [27]	0.651	0.972	0.606	0.466	28.026

TABLE I: Quantitative comparison of the performances of the networks trained on ISSLIDE. Bold and underline results show respectively best and second-best performances.

based on the binarized output probability map, the threshold of which maximizes the F1-Score.

C. Results

Following discussions are based on quantitative and qualitative results from Table I and Figure 3. For readability reasons, we select the best configuration of each architecture to appear in Figure 3 based on their quantitative performances.

The first major conclusion emerging from these results is the poor transferability of the pre-trained weights to our detection task. Indeed, performances of FCN$_P$ (resp. DeepLabV3$_P$) are dropping of 0.013 (resp. 0.007) points in F1-Score, 0.017 (resp. 0.01) points in Dice Score, 0.17 (resp. 0.006) points in IoU and 1.256 (resp. 0.535) points in Hausdorff distance. This significant gap illustrates the crucial need of interferogram datasets as their patterns, distributions and contents are completely different from those of classical image processing datasets making transfer learning inefficient.

In terms of architecture, FCN show actual benefits compared to DeepLabV3. Indeed, FCN$_S$ increases the performances of DeepLabV3$_S$ by 0.013 points in F1-Score, 0.015 points in Dice, 0.017 points in IoU and 0.657 in Hausdorff Distance. This gain in metrics can be explained by the use of dilated convolutions in DeepLabV3 against classical (local) convolutions in FCN. Dilated convolutions are extracting characteristics in a larger field of view assuming that multiple non-neighbouring pixels can serve for the segmentation of the targeted object. Nevertheless, the resolution of Sentinel-1 images is 10m$\times$10m meaning that two pixels in the same patch of 100px$\times$100px can be located at several kilometers the one from the other. Thus exploring so distanced pixels to segment an actual move covering some square meters is helpless and lead to coarse segmentations as in the last row of Figure 3.

On the other hand, the comparison of the three shallower networks show that ResUNet reached the best results. If ResUNet obviously overperform UNet as stated in [26], SepUNet reaches unexpected low results. This may be due to two reasons. First, in order to keep a similar amount of parameters in the networks, we used only 2 kernels for each channel in the depthwise convolutions. This might be too restrictive and introduces a lack of variability in terms of spatial extracted patterns. Then, separable convolutions are especially used to compute convolutions on channels which are not or weakly correlated. But these experiments input the coherence map and the cosine and sine of the phase difference map, *i.e.* the interferogram in its whole. Thus, each channel is extremely correlated with the others which may explain the drop in performance of SepUNet.

The most striking point based on Table I is that the second best performances are reached by the small ResUNet network. Indeed, despite its shallower architecture, ResUNet metrics are lower than FCN$_S$ ones by only 0.009 points in F1-Score,

0.006 points in Dice and 0.003 points in IoU. The most significant drop lies in Hausdorff Distance as it is 4.226 points greater. The latter can be explained by the use of skip connections between the encoder and the decoder of ResUNet introducing more high resolution details which disturbs the final decision as visible on the last row of Figure 3 within the orange circled layover area. The use of skip connections is still beneficial to extract more detailed segmentations. As illustrated by the landslide on the second row, the segmentation by ResUNet has more nuanced and precise borders than the coarse segmentation of FCN_S and DeepLabV3_S.

Finally, the first row of Figure 3 shows that the three networks are mimicking the manual annotation process. As expected, bubbles and fringes are correctly identified by the networks on the phase difference map and the coherence map is used to discriminate reliable patterns. This is illustrated by the magenta circle identifying a low coherent odd phase pattern. All networks identify the zone as non-moving despite its similarities in phase with the two northern moves. This behaviour helps to limit the number of false positives in the predictions and focus on actual moving areas as illustrated by the high AUC score of every networks in Table I.

V. Conclusion

In this work we first propose a new InSAR dataset for Slow SLIding areas DEtection (ISSLIDE) using machine learning. The latter contains 200 moves manually annotated on 54 generated interferograms following geomorphologist criteria. We then establish baseline results for the deep segmentation of these moves on a single interferogram. They end up with three major conclusions: transfer learning is inefficient on interferograms, shallow networks can be as performing as deep ones and all networks can mimic the annotation process.

The ISSLIDE dataset is released to encourage for more studies on Earth deformations. In particular, a multi-temporal approach would certainly reinforce the segmentation process as it mimics geomorphologist annotation strategies. On the other hand, ISSLIDE can be used to follow the evolution of the moves in time by stacking interferograms with different delays but with the same starting or ending date.

Acknowledgment: This work is supported by the Region Auvergne-Rhône-Alpes (France) through the project IATOAURA. Special thanks to Diego Cusicanqui and Xavier Bodin from ISTerre and EDYTEM for their collaboration in the annotation routine set up and their geomorphological expertise.

References

[1] G. Sumbul, A. de Wall, T. Kreuziger, F. Marcelino, et al., "BigEarthNet-MM: A Large-Scale, Multimodal, Multilabel Benchmark Archive for Remote Sensing Image Classification and Retrieval [Software and Data Sets]," *IEEE Geosci. Remote Sens. Mag.*, vol. 9, no. 3, pp. 174–180, Sept. 2021.

[2] M. Schmitt, L. H. Hughes, C. Qiu, and X. X. Zhu, "SEN12MS – A Curated Dataset of Georeferenced Multi-Spectral Sentinel-1/2 Imagery for Deep Learning and Data Fusion," June 2019, arXiv:1906.07789 [cs].

[3] J. Shermeyer, D. Hogan, J. Brown, A. Van Etten, et al., "SpaceNet 6: Multi-Sensor All Weather Mapping Dataset," in *2020 IEEE/CVF Proc. IEEE Comput. Soc. Conf. Comput. Vis.*, June 2020, pp. 196–197.

[4] C. Rambour, N. Audebert, E. Koeniguer, B. Le Saux, et al., "Flood Detection in Time Series of Optical and SAR Images," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. 43, no. B2, pp. 1343–1346, Aug. 2020.

[5] F. Montello, E. Arnaudo, and C. Rossi, "MMFlood: A Multimodal Dataset for Flood Delineation from Satellite Imagery," *IEEE Access*, vol. 10, pp. 96774–96787, 2022.

[6] J. Feranec, T. Soukup, G. Hazeu, and G. Jaffrain, *European Landscape Dynamics: CORINE Land Cover Data*, CRC Press, Aug. 2016.

[7] D. Sulla-Menashe, J. M. Gray, S. P. Abercrombie, and M. A. Friedl, "Hierarchical Mapping of Annual Global Land Cover 2001 to Present: the MODIS Collection 6 Land Cover Product," *Remote Sens. Environ.*, vol. 222, pp. 183–194, Mar. 2019.

[8] B. Dukai, H. Ledoux, and J. E. Stoter, "A Multi-Height LOD1 Model of All Buildings in the Netherlands," *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. IV-4-W8, pp. 51–57, Sept. 2019.

[9] S. Nagendra, D. Kifer, B. B. Mirus, T. Pei, K. Lawson, et al., "Constructing a Large-scale Landslide Database Across Heterogeneous Environments Using Task-Specific Model Updates," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 15, pp. 4349–4370, 2022.

[10] N. I. Bountos, I. Papoutsis, D. Michail, A. Karavias, et al., "Hephaestus: A Large Scale Multitask Dataset Towards InSAR Understanding," *IEEE Proc. Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 1453–1462, Apr. 2022.

[11] J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," *IEEE Proc. Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 3431–3440, 2015.

[12] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking Atrous Convolution for Semantic Image Segmentation," Dec. 2017, arXiv:1706.05587 [cs].

[13] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Med. Image Comput. Comput. Assist. Interv. – MICCAI 2015*. 2015, pp. 234–241, Springer International Publishing.

[14] X. Wang, J. Yi, J. Guo, Y. Song, et al., "A Review of Image Super-Resolution Approaches Based on Deep Learning and Applications in Remote Sensing," *Remote Sens.*, vol. 14, no. 21, pp. 5423, Jan. 2022.

[15] E. Dalsasso, L. Denis, and F. Tupin, "SAR2SAR: A Semi-Supervised Despeckling Algorithm for SAR Images," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 14, pp. 4321–4329, 2021.

[16] A. Bralet, A. M. Atto, J. Chanussot, and E. Trouvé, "Deep Learning of Radiometrical and Geometrical Sar Distorsions for Image Modality translations," in *2022 IEEE Int. Conf. Signal Image Processing (ICIP)*, Oct. 2022, pp. 1766–1770.

[17] V. S. F. Garnot, L. Landrieu, and N. Chehata, "Multi-Modal Temporal Attention Models for Crop Mapping from Satellite Time Series," *ISPRS J. Photogramm. Remote Sens.*, vol. 187, pp. 294–305, Dec. 2021.

[18] L. T. Luppino, M. A. Hansen, M. Kampffmeyer, F. M. Bianchi, et al., "Code-Aligned Autoencoders for Unsupervised Change Detection in Multimodal Remote Sensing Images," *IEEE Trans. Neural Netw. Learn*, pp. 1–13, 2022.

[19] O. Ghorbanzadeh, Y. Xu, H. Zhao, J. Wang, et al., "The Outcome of the 2022 Landslide4Sense Competition: Advanced Landslide Detection from Multi-Source Satellite Imagery," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 15, pp. 9927–9942, Sept. 2022.

[20] B. Adriano, N. Yokoya, J. Xia, H. Miura, et al., "Learning from Multimodal and Multitemporal Earth Observation Data for Building Damage Mapping," *ISPRS J. Photogramm. Remote Sens.*, vol. 175, pp. 132–143, Sept. 2020.

[21] R. Gupta, B. Goodman, N. Patel, R. Hosfelt, et al., "Creating xBD: A Dataset for Assessing Building Damage from Satellite Imagery," *IEEE Proc. Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 10–17, 2019.

[22] Y. Zhao, G. Feng, Y. Wang, X. Wang, et al., "A New Algorithm for Intelligent Detection of Geohazards Incorporating Attention Mechanism," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 113, pp. 102988, Sept. 2022.

[23] F. Thollard, D. Clesse, M.-P. Doin, J. Donadieu, et al., "FLATSIM: The ForM@Ter LArge-Scale Multi-Temporal Sentinel-1 InterferoMetry Service," *Remote Sens.*, vol. 13, no. 18, pp. 3734, Jan. 2021.

[24] A. Bertone, C. Barboux, X. Bodin, T. Bolch, et al., "Incorporating InSAR Kinematics into Rock Glacier Inventories: Insights from 11 Regions Worldwide," *The Cryosphere*, vol. 16, no. 7, pp. 2769–2792, July 2022.

[25] T. G. Farr, P. A. Rosen, E. Caro, R. Crippen, et al., "The Shuttle Radar Topography Mission," *Rev. geophys.*, vol. 45, no. 2, 2007.

[26] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *IEEE Proc. Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770–778.

[27] F. Chollet, "Xception: Deep Learning With Depthwise Separable Convolutions," in *IEEE Proc. Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 1251–1258.