FLOOD DETECTION IN TIME SERIES OF OPTICAL AND SAR IMAGES

C. Rambour[1, 2,][*] N. Audebert[1], E. Koeniguer[2], B. Le Saux[2], M. Crucianu[1], M. Datcu[1]

[1] CEDRIC (EA4629), Conservatoire National des Arts et Métiers, HESAM Université, 75003 Paris

[2] ONERA / DTIS, Université Paris-Saclay, F-91123 Palaiseau, France

ABSTRACT:

These last decades, Earth Observation brought a number of new perspectives from geosciences to human activity monitoring. As more data became available, Artificial Intelligence (AI) techniques led to very successful results for understanding remote sensing data. Moreover, various acquisition techniques such as Synthetic Aperture Radar (SAR) can also be used for problems that could not be tackled only through optical images. This is the case for weather-related disasters such as floods or hurricanes, which are generally associated with large clouds cover. Yet, machine learning on SAR data is still considered challenging due to the lack of available labeled data. To help the community go forward, we introduce a new dataset composed of co-registered optical and SAR images time series for the detection of flood events and new neural network approaches to leverage these two modalities.
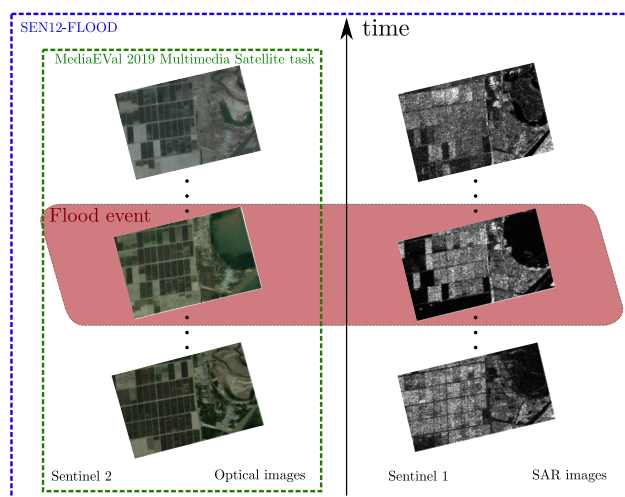

Figure 1. The SEN12-FLOOD dataset is composed of SAR and optical time series in which a flood event may occur.

1. MOTIVATION AND SIGNIFICANCE OF THE TOPIC

Recently, new datasets dealing with natural disaster detection were proposed. Most of them are constituted of RGB or multispectral images. These datasets are built from acquisitions performed by sensors with a high resolution, either temporal (*e.g.* Landsat, Sentinel 2) as in the MediaEval 2019 Multimedia Satellite task (Bischke et al., 2019) or spatial (*e.g.* Quickbird, WorldView) as in XView2 (Gupta et al., 2019). AI and, in particular, deep learning techniques proved to be efficient in retrieving semantic land cover information and specific behaviors from such datasets (Zhu et al., 2017, Audebert et al., 2019). However, the analysis of optical images time series may be impossible when looking at areas where the cloud cover is important. Even when the Earth's surface is visible, wetlands and floods are very difficult to characterize visually. SAR images offer an alternative as they can be acquired without the sun's illumination and independently from the cloud cover. In particular, new satellites such as Sentinel 1 provide an extensive

amount of data with a high time-frequency (an image every 6 days), allowing to monitor large areas of the Earth. So, there is a need for new machine learning approaches for disaster monitoring that leverage both passive optical and active radar imaging modalities.

Detecting floods and measuring their extent on the basis of satellite images is a core topic in remote sensing for disaster management, especially as floods can develop slowly or sometimes very quickly. Several previous works have investigated detecting flooding events from satellite imagery, either multispectral – Landsat/IKONOS (Gläßer, Reinartz, 2005), MODIS (Brakenridge, Anderson, 2006) – or SAR (Nico et al., 2000). As recent work (Gómez-Chova et al., 2015) has shown, multimodal machine learning can leverage the complementary information from multiple sensors to improve the accuracy of the models. For example, in MediaEval 2017, (Bischke et al., 2017) learned deep models to perform flood detection in natural images using ancillary data from social networks. Our work digs in the same direction: we aim to provide strategies to perform multimodal flood detection to leverage as many remote sensing data as possible.

We present in this paper the new SEN12-FLOOD dataset composed of both Sentinel 1 and Sentinel 2 images to foster the development of new flood detection techniques (Fig. 1). Then, we propose a first baseline based on off-the-shelf deep networks for multimodal time series analysis to classify the images in the dataset.

2. PRESENTATION OF THE SEN12-FLOOD DATASET

The city-centered satellite sequences provided by the Media-Eval 2019 Multimedia Satellite task (Bischke et al., 2019) give access to series of multispectral Sentinel 2 images. The observed areas correspond to African, Iranian, and Australian cities and their surroundings with or without a flood event occurring during the time series. These images are composed of 12 bands, 10m ground-sampling distance and are provided with Level 2A atmospheric correction. Here, we propose a new dataset corresponding to the Sentinel 1 sequences for the same areas and periods. However, since SAR is independent of

---

* Corresponding author

Figure 2. Map of the main areas contained in the dataset. Areas in red correspond to sequences in the training set and areas in blue correspond to sequences in the validation set. Most of the scenes correspond to South East African areas while the rest of the dataset is obtained from West African, Iranian or Australian locations. The behavior of the flood may differ greatly from one area to another: while open water areas appear clearly in SAR images, flooded vegetation or soaked ground areas are harder to discriminate from dry areas.

cloud cover, more SAR images are retrieved for the same time period, leading to a higher sampling rate. This SAR dataset is composed of roughly two times more images than the optical one. To leverage both SAR and optical modalities, we merge the MediaEval dataset and our own in the new SEN12-FLOOD dataset.

Each image has a binary label specifying whether a flood event is visible or not in the observed area. The labels have been provided by the original MediaEval 2019 dataset and were obtained from the Copernicus Emergency Management Service[1]. The Sentinel 1 images were downloaded from the Scientific ESA hub website[2]. The data were acquired in Interferometric Wide Swath (IW) mode at polarization VV and VH. The SAR images are delivered in Ground Range Detected High Resolution (GRDH) products with a resolution of $10 \times 10 \ m$. Pre-processing – including radiometric calibration (Miranda, 2015) as well as Range Doppler Terrain Correction using the shuttle radar topographic mission digital elevation model – was applied to the SAR images thanks to the SNAP ESA software (Brockmann Consult, C-S, 2019). The dataset is composed of 412 time series with 4 to 20 optical images and 10 to 58 SAR images in each sequence. On average, there are 9 optical and 14 SAR images per sequence. The period of acquisition goes from December 2018 to May 2019. A flood event is occuring in 40% of the optical Sentinel 2 images and in 47% of the SAR Sentinel 1 images. As in the MediaEval dataset, once a flood occurred in a sequence, all the subsequent images are labeled as flooded which corresponds to the hypothesis that the surface still presents characteristic modifications after the event.

---

[1] https://emergency.copernicus.eu/
[2] https://scihub.copernicus.eu/

3. BENEFIT OF MULTISPECTRAL AND SAR DATASET

This dataset has been constituted to train a new architecture of neural networks for dual-mode and multi-temporal flood classification. We provide an in-depth study of the various components of the model. Indeed, our goal is to assess the relevance of each modality and the contribution of temporal analysis.

First, SAR images are expected to help the ground classification generally conducted on multispectral data. For example, the normalized water difference index (Gao, 1996) is widely used to detect the presence of water bodies. However, depending on the sensor, this index may suffer from one drawback: bands associated with the near-infrared and short-waved infrared can present a loss of resolution compared to the RGB ones. On the other hand, SAR images are more sensitive to the geometrical distribution of the backscattering elements. For instance, smooth, plane surfaces such as roads or open water areas behave as mirrors and backscatter most of the transmitted wave in the specular direction from the sensor. These surfaces produce typical dark areas in the resulting SAR images, allowing to identify these classes quickly. Moreover, polarization is also affected by the presence of water, and statistical approaches combining the VV and VH bands have shown promising performances (Cazals et al., 2016).

Finally, the time consistency is essential to distinguish floods from permanent elements like water bodies. So, multitemporal analysis is the key for the detection of abnormal events such as natural disasters and even their prediction ahead of time. This is becoming more and more necessary to avoid potential harms.

4. FLOOD DETECTION

In order to show the interest of this dataset, we used the state-of-the-art ResNet-50 (He et al., 2015) network for the detection

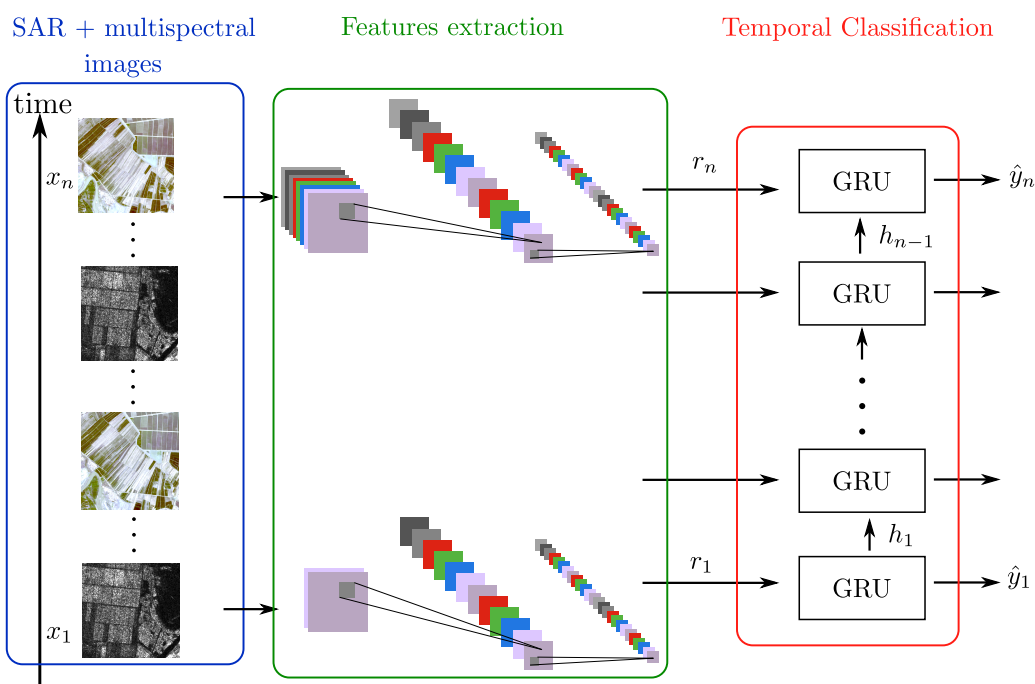Figure 3. The proposed pipeline for flood detection in sequences of remote sensing images.

Model	Data	Accuracy
Resnet-50	RGB	0.885
	Multispectral	0.793
	SAR	0.753
Resnet-50 + GRU	RGB	0.930
	SAR	0.875
	SAR+RGB	**0.957**

Table 1. **First row :** accuracy achieved by a state of the art network (Resnet-50) for a flood detection task on each image. **Second row :** accuracy given by a recurrent network on the sequence of image features.

of flooding events in each image. This network was designed for RGB images classification and the first convolutional layer was modified to take into account the correct number of bands for multispectral (12 bands) and SAR (2 bands) data. This implies that the network had to be retrained from scratch for these two configurations while on the RGB images the pretrained network could be used explaining part of the accuracy gap in the baseline. For the SAR and multispectral configuration, the models were trained for 400 epochs using Stochastic Gradient Descent with a learning rate of 7.10^{-6}. The ResNet-50 network may not be optimal when considering a large number of channels, which may also explain the lower accuracy on multispectral data. When no temporal information is considered, only the spatial context gives insight on the presence or absence of flood phenomena and optical images seem to be more suited for this task than SAR images.

To take into account the temporal dimension, we extracted the features of each image with respect to the trained Resnet-50 and applied a Gated Recurrent Unit (GRU) (Cho et al., 2014) on each sequence. For the multimodal classification, we simply concatenate the features from SAR and RGB ResNets and feed those to the GRU layer. The final result is a sequence of binary labels

giving the result of the flood detection task for every frame. The pipeline of the model is illustrated in Fig. 3.

5. DISCUSSION

The accuracy scores obtained with the ResNets on each image as well as the proposed network on each sequence are presented in Table 1 where the metric was computed for each image in the dataset. The experiments performed on the SEN12-FLOOD dataset give several insights on the interest to use optical and SAR images for flood detection. On the first hand, state of the art techniques perform well on both optical and SAR images and manage to retrieve most of the flood phenomena. On the second hand, it appears clearly that the temporal dimension is a key to characterize these events. Indeed, considering the temporal dimension leads to a significant error reduction using optical as well as SAR images. Moreover, SAR and optical modalities appear to be complementary for the detection of flood phenomena as the best accuracy on the dataset is achieved by using both kinds of data.

To go further, several points may be explored. First, architectures able to handle multispectral data (Sumbul et al., 2019) should allow to retrieve more information from all the Sentinel 2 images and thus achieve a better classification. Second, flood detection on SAR images may seem harder than on optical images. However, it appears that on areas where the water is clearly visible to the sensor, the detection is close or even better than with optical images, whereas when it is occluded by vegetation, optical images are more useful. In the results, areas in Zimbabwe where the flood mostly consisted in soaked ground and vegetation, were often miss-classified using the SAR images sequences but not with the optical ones. On the contrary, areas in Iran where the flood events consisted in large open water areas, the accuracy may go down to 0.5 using optical images while the detection is almost perfect using SAR images. This may speak in favor of a specific weighting of the input data depending on the observed area configuration, as illustrated in Fig. 2.

Finally, other fusion strategies could be considered in order to give more freedom to the network to learn specific behaviors associated to each kind of input data. For example, attentional models could be used to focus more specifically on one sensor or the other based on image characteristics such as cloud cover, noise or environmental properties.

6. CONCLUSION

In this paper we presented a new dataset composed of optical and SAR images for the detection of flood events in time series. We also proposed a baseline for multitemporal classification of floods based on spatial-temporal porcessing by residual and GRU networks. Our experiments show the interest to consider both of these modalities for this task. Future work may include the search for better fusion strategies as well as the efficient processing of multispectral data. The SEN12-FLOOD dataset as well as the code of the proposed approach can be downloaded at `https://clmrmb.github.io/SEN12-FLOOD`.

7. ACKNOWLEDGMENT

REFERENCES

Audebert, N., Le Saux, B., Lefevre, S., 2019. Deep Learning for Classification of Hyperspectral Data: A Comparative Review. *IEEE Geoscience and Remote Sensing Magazine*, 7(2), 159-173.

Bischke, B., Bhardwaj, P., Gautam, A., Helber, P., Borth, D., Dengel, A., 2017. Detection of flooding events in social multimedia and satellite imagery using deep neural networks. *Working Notes Proceedings of the MediaEval 2017. MediaEval Benchmark, September 13-15, Dublin, Ireland*, MediaEval.

Bischke, B., Helber, P., Schulze, C., Srinivasan, V., Dengel, A., Borth, D., 2019. The Multimedia Satellite Task at MediaEval 2019. *In Proc. of the MediaEval 2019 Workshop*.

Brakenridge, R., Anderson, E., 2006. Modis-based flood detection, mapping and measurement: The potential for operational hydrological applications. J. Marsalek, G. Stancalie, G. Balint (eds), *Transboundary Floods: Reducing Risks Through Flood Management*, Springer Netherlands, Dordrecht, 1–12.

Brockmann Consult, A. S. C., C-S, 2019. SNAP Software.

Cazals, C., Rapinel, S., Frison, P.-L., Bonis, A., Mercier, G., Mallet, C., Corgne, S., Rudant, J.-P., 2016. Mapping and Characterization of Hydrological Dynamics in a Coastal Marsh Using High Temporal Resolution Sentinel-1A Images. *Remote Sensing*, 8(7). https://www.mdpi.com/2072-4292/8/7/570.

Cho, K., van Merrienboer, B., Çaglar Gülçehre, Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y., 2014. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. *ArXiv*, abs/1406.1078.

Gao, B., 1996. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58(3), 257 - 266.

Gläßer, C., Reinartz, P., 2005. Multitemporal and Multispectral Remote Sensing Approach for Flood Detection in the Elbe-Mulde Region 2002. *Acta hydrochimica et hydrobiologica*, 33(5), 395-403. https://onlinelibrary.wiley.com/doi/abs/10.1002/aheh.200400599.

Gupta, R., Hosfelt, R., Sajeev, S., Patel, N., Goodman, B., Doshi, J., Heim, E., Choset, H., Gaston, M., 2019. xBD: A Dataset for Assessing Building Damage from Satellite Imagery.

Gómez-Chova, L., Tuia, D., Moser, G., Camps-Valls, G., 2015. Multimodal Classification of Remote Sensing Images: A Review and Future Directions. *Proceedings of the IEEE*, 103(9), 1560-1584.

He, K., Zhang, X., Ren, S., Sun, J., 2015. Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770-778.

Miranda, N.; Meadows, P., 2015. Radiometric Calibration of S-1 Level-1 Products Generated by the S-1 IPF, ESA-EOPG-CSCOP-TN-0002.

Nico, G., Pappalepore, M., Pasquariello, G., Refice, A., Samarelli, S., 2000. Comparison of SAR amplitude vs. coherence flood detection methods - a GIS application. *International Journal of Remote Sensing*, 21(8), 1619-1631. https://doi.org/10.1080/014311600209931.

Sumbul, G., Charfuelan, M., Demir, B., Markl, V., 2019. Bigearthnet: A Large-Scale Benchmark Archive for Remote Sensing Image Understanding. *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*. http://dx.doi.org/10.1109/IGARSS.2019.8900532.

Zhu, X. X., Tuia, D., Mou, L., Xia, G., Zhang, L., Xu, F., Fraundorfer, F., 2017. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geoscience and Remote Sensing Magazine*, 5(4), 8-36.