A TEMPORAL DEEP CONVOLUTIONAL NEURAL NETWORK MODEL ON SENTINEL-1 IMAGE TIME SERIES FOR PIXEL-WISE FLOOD CLASSIFICATION

Konstantinos Vlachos, Anastasia Moumtzidou, Ilias Gialampoukidis, Stefanos Vrochidis, Ioannis Kompatsiaris

Information Technologies Institute (ITI)
Centre for Research and Technology Hellas (CERTH)
6th km Charilaou - Thermi Rd, 57001, Thermi, Thessaloniki, Greece

ABSTRACT

Accurate and timely flood mapping is important in emergency management during and after extreme flood events which can be greatly served by Synthetic Aperture Radar (SAR). This work is focused on open-land regions where a custom annotation based on expert knowledge and NDFI is used. Research on SAR flood detection is mostly based on histogram thresholding that has low time complexity and seems ideal for emergency response, although human intervention is needed. Machine learning methods have fewer errors and minimize the need for human intervention but their computational complexity is higher. This work aims to provide a lightweight convolutional neural network baseline for pixel-wise time series classification in flood monitoring on SAR satellite image time series. Quantitative and qualitative evaluation of results indicate that the approach is promising.

Index Terms— Sentinel-1, SAR, flood detection, time series classification, deep learning, CNN

1. INTRODUCTION

Extreme geohazard phenomena are more and more frequent due to climate change. One of them are extreme flood events that may also be enhanced because of human intervention in the environment. Therefore the societal, economical and environmental impact is significant, and for this reason rapid and accurate flood mapping is equally important in emergency response management and long-term planning.

One of the most useful means for flood mapping is the use of Big satellite Data and, in particular, Synthetic Aperture Radar (SAR) has been proved to show quite a lot of advantages compared to multispectral/optical sensors. It provides usable data regardless of solar illumination and weather conditions which is a reason for being an excellent tool for flood monitoring. However, SAR data do not come without a price because radar backscatter is affected by many factors which makes interpretation non-intuitive and difficult especially to non-experts. Some factors are related to the atmospheric conditions and the orbital geometry of the satellite during the acquisition, while others depend on the Earth's surface morphology. Regarding surface morphology three separable effects can be distinguished a) topography b) surface dielectric properties and c) surface roughness. For the flood detection problem, surface roughness is the crucial factor that mainly determines the backscatter amplitude.

The reasoning behind SAR flood detection is that water acts as a specular reflector which results in very weak backscatter. As a result, the backscatter of a given region becomes lower after flooding. However, this reasoning may change due to alteration of various surface characteristics that are interconnected with surface roughness. According to [1] those are land cover classes, weather conditions, and geometry of the target itself as well as relative to the sensor Line of Sight. Concerning land cover classes, the backscatter becomes higher after an event in urban environments and canopies. This is owned to double bounce effects of the radar beam between water and buildings/canopies [1, 2]. In addition, similar counter-intuitive effect holds for targets such as roads which have backscatter similar to water, thus making it unfeasible to detect a change due to flooding. Last but not least, weather conditions such as strong winds result in a weaker than expected decrease of backscatter since they increase the roughness of the water surface. In addition, the aforementioned effects depend on SAR polarization channels [3], resulting in different channels revealing different kind of information. For this reason, using all available channels seems to have an assisting effect in flood detection [4].

Flood detection using SAR data can be categorized according to the number of dates used [1], i.e. single date, dual date and multiple dates. The number of dates also defines the range of techniques that can be used. For example, a time series analysis can only be applied on the third case, while semantic segmentation can theoretically be applied in all cases.

Core flood detection approaches regardless of number of images are thresholding, change detection [5], change detection and thresholding [6], supervised [7, 8], semi-supervised [9, 10] and unsupervised [2] image classification based on

This work has received funding from the European Union's Horizon2020 research and innovation project aqua3S under grant agreement No 832876.

classical machine and deep learning. Multi-temporal pixel-wise time series approaches such as [11, 12], despite being lightweight, do not exploit the temporal autocorrelation, as they treat the time series using statistical analysis without any learning procedure. On the other hand, sophisticated deep learning methods may be more accurate but they have a high computational complexity. Our approach aims to tackle this gap by exploiting the temporal information using a 1D-temporal Convolutional Neural Network (CNN) [13].

This work focuses on open-land flooded areas using a temporal convolutional neural network framework on VV and VH Sentinel-1 polarization channels for flood event classification in satellite image time series. To our knowledge, the proposed approach is the first one that relies on pixel-wise SAR time series classification using deep CNN for flood detection. Thus, our main contribution is that we provide a pure time series classification CNN-based baseline (inspired from [14, 13]) that exploits the autocorrelation of pixel-wise Sentinel-1 SAR time series of VV and VH channels in a supervised deep learning setting and at the same time is lightweight, flexible and tuned for rapid flood mapping.

2. METHODS & MATERIALS

2.1. Dataset

The area of interest (AOI) is the wider Monfalcone region (N. Italy) and is defined by (12.801928, 46.171155) and (13.733412, 45.659611) geographical coordinates. A time series of Sentinel-1 A/B IW Level-1 GRD, path 95, descending node images was acquired from Alaska Satellite Facility Data Search Vertex. The downloaded data represent the year 2019 which results in 61 images with a constant sampling period of 6 days, and the flood event happened between the 15th and the 27th of November with a peak on the 21st. Some distinctions that need to be made at this point are a) the complete time series was used for the multi-temporal speckle filtering b) data from August to December 3rd were used for exploration and c) data from August to November 21st were used for the annotation and modelling.

2.2. Pre-processing

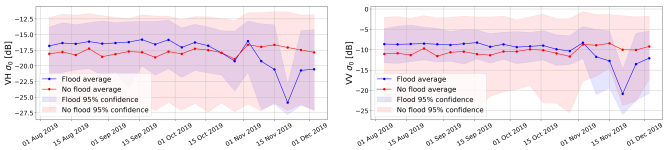
The image pre-processing steps have been carried out in SNAP Toolbox. These are: a) image subset; b) orbit file application; c) thermal noise removal; d) calibration to σ_o; e) range-Doppler terrain correction with SRTM 3-sec as DEM and WGS84/UTM zone 33N as reprojection CRS using nearest-neighbour interpolation; f) co-registration based on geo-location using nearest-neighbour interpolation and stack creation; g) multi-temporal speckle filtering (7x7 Lee filter); h) linear to dB scale conversion. The dataset used for annotation and modelling is composed of 19 images with the last one being November 21st.

2.3. Reference data

A known issue in flood detection is the scarcity of reference data due to difficulties in generating/acquiring them precisely and on time. Two curated datasets have been released recently for that purpose [15, 16] but they are inflexible when one desires to apply custom preprocessing and specific timestep length, among others. In addition, the reference data source used widely in the remote sensing community is the Copernicus Emergency Management Service[1]. However, these come with their own limitations such as restricted number of events and geographical coverage, since their generation depends on regional authorities' request for service activation. We note that the service is widely based on optical data which usually come with cloud coverage during an event.

Taking all these into account, we create reference data by combining expert knowledge from the local water authority and feature engineering. Hydrology experts of the AOI from Alto Adriatico Water Authority[2] (AAWA) delineated the regions with high probability of flood occurrence. Following, the Normalized Difference Flood Index (NDFI) [11] was calculated on the pre-processed image stack (in linear scale). This is an index-based thresholding approach that aims to depict open-land temporary waters by using a quasi-fixed threshold value. For the threshold value after trial-n-error we used the suggested 0.7 above which floods are denoted. Finally, 32,707 pixels that fall in AAWA delineation and have NDFI over 0.7 were annotated as flood, while 34,053 pixels that do not satisfy that condition were annotated as non-flood constituting a balanced training dataset.

We then produce a time series cumulative statistics graph for VV and VH Fig.1 aiming to validate the assumption that time series of flooded pixels are characterized by a drop in backscatter. In particular, a temporal pattern with small dispersion is clear with a sudden drop on November 21st which denotes the peak of the flood. On the other hand, pixels annotated as non-flooded show a relatively stable average signal with large dispersion which is owed to the variety of different sub-classes that it depicts such as permanent water, urban environments etc.



(a) (b)

Fig. 1. Sub-figures show the pixel-wise time series for VV and VH that were annotated as flood/non-flood.

[1]https://emergency.copernicus.eu
[2]https://distrettoalpiorientali.it

A/A	l_r	Loss	OA [%]	F1-score [%]
Train	0.01	8.78e-03	99.7398	99.7349
Test		9.02e-03	99.7303	99.7243
Train	0.1	7.87e-04	99.9771	99.9769
Test		8.04e-04	99.9735	99.9725
Train	1.	**4.13e-06**	**100**	**100.0**
Test		**4.35e-05**	**99.9984**	**99.9983**

Table 1. Loss, Overall Accuracy (OA) and F1-score on train/test sets for different learning rate (l_r) initialization values. Metrics are averaged over 30 model runs.

2.4. Dataset preparation

The per-pixel annotated time series were extracted from the 19-image stack. As a result, the model input data shape is (samples, timesteps, channels) where samples denote the number of the time series (66,760), timesteps the sequence length (19) and channels the two SAR polarizations (2). They were z-score standardized using the mean and standard deviation across all training samples, timesteps and channels at once. This scaling approach was used because it keeps the relative distances between input channels stable, which is needed when applying convolutions across time and channel dimension. Finally, the labels were one-hot encoded which resulted in a (samples, number of classes) vector where the two number of classes denotes flood/non-flood.

2.5. Architecture

The network is composed of one batch normalization, two convolutional (8 units), one fully-connected (4 units) and a softmax layer (2 units). Further settings used are ReLU activation function, categorical cross-entropy loss, and Adadelta optimizer[3] with 0.1 initial learning rate. Regarding convolutional layers kernels size was set to 5 and stride to 1. The modified settings from [14] are a result of experimentation. In particular, a technique to handle learning rate decay and noisy gradient information, such as Adadelta, was imperative since simple use of Adam with various learning rates made the learning process unstable.

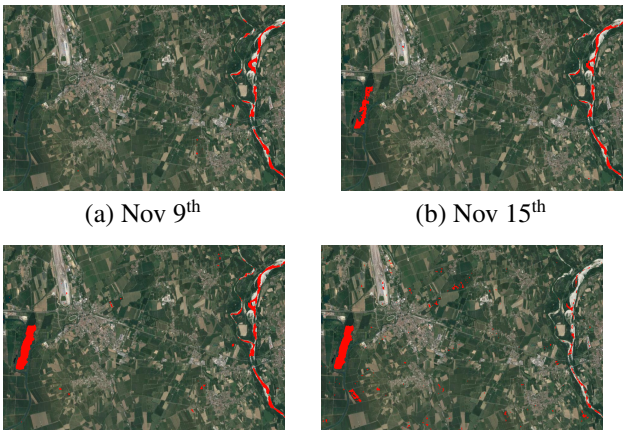
The basic reasoning behind the network is to learn the temporal pattern and in particular the presence of a local minimum value that will occur at the last 1 timestep of the sequence. This also means that the network should not learn that a sequence represents a flood event assuming there is global or local minimum value(s) in a timestep prior to the last one. Therefore, for flood detection on a given date the user should use a fixed-length time series as input where the last date should be the one s/he wants to predict on. Based on the architecture and the training setting the model is capable of predicting an ongoing flood event.

[3]https://arxiv.org/abs/1212.5701

2.6. Experiments

The experimentation is based on a 50/50 and 80/20 train/test and train/validation split, respectively. Acceptable model fitting was achieved using an early stopping mechanism with a patience of 10 epochs according to validation set loss. Training epochs were set to 60 with a batch size of 32. The experiments were ran on a CPU machine using Tensorflow[4] and Keras[5].

After deciding on the network baseline architecture we experimented with different learning rates producing relevant error metrics Tab. 1. Adadelta comes with a suggested initial value equal to 1, however 0.1 showed the best bias-variance tradeoff.



(a) Nov 9th (b) Nov 15th

(c) Nov 21st (d) Nov 27th

Fig. 2. Figure shows prediction of the CNN model for four dates. The temporal CNN model was trained with November 21st as the last date.

3. RESULTS AND DISCUSSION

After training the model the results were evaluated quantitatively and qualitatively. For the former case, average error metrics such as Loss, Overall Accuracy and F1-score were calculated for train and test sets over 30 model runs. For the latter case, flood predictions from the best model were used on time series with a temporal lag relative to the time series used for training. For instance, in Figure 2d a time series that begins on August 11th and ends on November 27th (total 19 images) was used to predict for November 27th. Same reasoning was applied for the rest of the dates.

From Figures 2a-2c and cross-comparison with VV and VH False Color Composites (R: pre-event, G: post-event, B: post-event) for the three dates we have noticed that the model is accurate. On the contrary, the model predicts a false positive ongoing flood at the left red region for Figure 2d (Nov

[4]https://www.tensorflow.org

[5]https://keras.io

27^{th}) where the backscatter value starts to increase as seen in Figure 1. Although, the results are accurate on dates prior to the event, i.e., flood is not detected. A possible explanation of this is that the model needs more variability in the training set, which in turn will make it able to consider edge cases such as the one described.

Last but not least, the wide majority of model predictions for November 21^{st} coincide with the thresholded NDFI pixels that were not used for train/test. Additionally, it also detects floods on regions that were edge cases for NDFI. In other words, the result of a smaller NDFI threshold coincides with the model output (i.e., more floods were depicted) but the former is more noisy. This leads to the conclusion that the model has the potential to detect floods as good as NDFI does with less noise and no human intervention, which is an important advantage in rapid flood mapping.

4. CONCLUSIONS

In this work we presented a time series classification for flood detection using a deep CNN approach. The main advantage of this model is that it offers a lightweight way to indicate ongoing flood events in open land without human intervention. In addition it has high predictive power during a flood event occurrence, and its lightweight training makes it easy for quantitative error assessment. In the future we plan to explore additional features to be used as input in the temporal neural network architecture for other disasters beyond floods.

5. REFERENCES

[1] M. Chini et al., "Sar-based flood mapping, where we are and future challenges," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 2021, pp. 884–886.

[2] X. Jiang et al., "Rapid and large-scale mapping of flood inundation via integrating spaceborne synthetic aperture radar imagery with unsupervised deep learning," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 178, pp. 36–50, 2021.

[3] J. Liang and D. Liu, "A local thresholding approach to flood water delineation using sentinel-1 sar imagery," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 159, pp. 53–62, 2020.

[4] S. Plank et al., "Mapping of flooded vegetation by means of polarimetric sentinel-1 and alos-2/palsar-2 imagery," *International Journal of Remote Sensing*, vol. 38, no. 13, pp. 3831–3850, 2017.

[5] W. Liu et al., "Inundation assessment of the 2019 typhoon hagibis in japan using multi-temporal sentinel-1 intensity images," *Remote Sensing*, vol. 13, no. 4, 2021.

[6] P. Chakma and A. Akter, "Flood mapping in the coastal region of bangladesh using sentinel-1 sar images: A case study of super cyclone amphan," *Journal of the Civil Engineering Forum*, vol. 7, no. 3, pp. 267–278, 2021.

[7] A. Nallapareddy and B. Balakrishnan, "Automatic flood detection in multi-temporal sentinel-1 synthetic aperture radar imagery using ann algorithms," *INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL*, vol. 15, no. 3, 2020.

[8] R. M. de la Cruz et al., "Near-realtime flood detection from multi-temporal sentinel radar images using artificial intelligence," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLIII-B3-2020, pp. 1663–1670, 2020.

[9] K. A. Islam et al., "Flood detection using multi-modal and multi-temporal images: A comparative study," *Remote Sensing*, vol. 12, no. 15, 2020.

[10] P. Sayak and G. Siddha, "Flood segmentation on sentinel-1 sar imagery with semi-supervised learning," 2021.

[11] F. Cian et al., "Normalized difference flood index for rapid flood mapping: Taking advantage of eo big data," *Remote Sensing of Environment*, vol. 209, pp. 712–730, 2018.

[12] K. Karamvasis and V. Karathanassi, "Flompy: An open-source toolbox for floodwater mapping using sentinel-1 intensity time series," *Water*, vol. 13, no. 21, 2021.

[13] Zhiguang Wang, Weizhong Yan, and Tim Oates, "Time series classification from scratch with deep neural networks: A strong baseline," in *2017 International joint conference on neural networks (IJCNN)*. IEEE, 2017, pp. 1578–1585.

[14] C. Pelletier et al., "Temporal convolutional neural network for the classification of satellite image time series," *Remote Sensing*, vol. 11, no. 5, 2019.

[15] D. Bonafilia et al., "Sen1floods11: A georeferenced dataset to train and test deep learning flood algorithms for sentinel-1," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2020.

[16] C. Rambour et al., "Flood detection in time series of optical and sar images," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLIII-B2-2020, pp. 1343–1346, 2020.