ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume X-4/W1-2022
GeoSpatial Conference 2022 – Joint 6th SMPR and 4th GIResearch Conferences, 19–22 February 2023, Tehran, Iran (virtual)

MAPPING OF URBAN FLOOD INUNDATION USING 3D DIGITAL SURFACE MODEL AND SENTINEL-1 IMAGES

M. Sharif [1], S. Heidari [2] *, S. M. Hosseini [3]

[1]Master of Science Remote Sensing and GIS, Faculty of Geography, University of Tehran. morteza.sharif@ut.ac.ir
[2]PhD Student of Climatology, Department of Physical Geography, Faculty of Geography, University of Tehran. Heidari.s@ut.ac.ir
[3]Physical Geography Department, University of Tehran, P.O. Box 14155-6465, Tehran, Iran. smhosseini@ut.ac.ir

Commission IV, WG IV/3

KEYWORD: Remote Sensing, Convolutional Neural Network, Synthetic Aperture Radar, Flood Inundation Map.

ABSTRACT:

Flooding in urban areas poses serious risks to citizens, infrastructures, and transportation. Precise and real-time delineation of the inundated areas is crucial for a better understanding of the extent of damage and high-risk areas and people evacuation actions. It also increases citizens' awareness that living in areas with high flood risk. Yazd city is characterized by low rainfall (<70 mm/yr) and the desert climate is considered the study area of this research. This city encountered a flash flood event that was generated by severe rainfall with a depth of 75 mm in 3hr (i.e., the intensity of 25 mm/hr) on July 29, 2022. Many strategic infrastructures of this city especially the railway station were flooded, which caused heavy casualties and financial losses. This study aims to monitor the flood inundated areas of Yazd city due to this flood event using remote sensing. In this research, the Sentinel-1 polarimetric radar images and the 3D model of the Yazd city surf ace were used to delineate the flooded areas. The field information of the flooded areas and the available Sentinel-1 images during or near the occurrence time of maximum flood extension were adopted. The Convolutional Neural Network (CNN) model in combination with the 3D model of the studied area was used to identify the flooded pixels in the city of Yazd. The results showed that the adopted 3D model and CNN algorithm indicated a good ability to identify flooded areas with an accuracy of 88% and a kappa coefficient of 0.83.

1. Introduction

Flood is a main natural disaster that possess a serious risk to the urban environment. (Peng et al., 2019). Real-time monitoring of urban inundated areas is crucial for mitigation of flood risk and damage and planning of people evacuation strategy, especially in developing countries which are characterized by low investment in hydraulic and conveyance structures (Bayik et al., 2018). However, the increasing development of advanced technologies such as satellite images, the development of algorithms, and environmental risk monitoring software can overcome these challenges to some degree (Li et al., 2018). While, the remotely-sensed derivative of flood inundation area by Synthetic Aperture Radar (SAR) has been widely used, but, this issue in urban areas is ongoing challenging (Li et al., 2019).

In particular, radar satellite images have the potential to monitor flood-inundated areas compared to optical images (Li et al., 2018; Surampudi & Yarrakula, 2020; Tanim et al., 2022). Polarimetric radar images from Sentinel-1, an active remote sensing satellite that produces intensities of backscattered signals from the Earth's surface (DeVries et al., 2020). This data is freely available to public users. The possibility of inverse scattering intensity identification is of individual characteristics of Sentinel-1 images (Carreño Conde & De Mata Muñoz, 2019). However, the scattering of radar pulses from land surface phenomena in distinguishing between pixels in flooded areas and other phenomena such as vegetation that is hidden under water during flooding (Mason et al., 2021), or sandy areas where the intensity of scattering of radar pulses is similar to water bodies is of challenges associated with this process (Martinis et al., 2018). As a result, it increases the error in detecting inundated areas. Using 3D models can overcome these challenges to some extent.

In several studies, using the integration of remote sensing images, including radar, optical and LiDAR images, to increase the accuracy of the detection of earth surface phenomena, such as urban surfaces, vegetation, or other effects on the earth's surface, has been done (Al-Najjar et al., 2019; DeVries et al., 2020; Jahan et al., 2018). These extra topographies can improve the classification accuracy for specific features (Al-Najjar et al., 2019). For instance, datasets synthesized from RGB (Red, Green and Blue) images obtained from UAVs or other sources together with digital surface models (DSM) provided a more holistic representation for the production of accurate maps (Jahan et al., 2018). Considering DSMs as additional features were proven to increase the classification for image segmentation (Marcos et al., 2018).

The application of deep learning techniques (e.g., convolutional neural network, CNN) in satellite image classification has led to

* Corresponding author

https://doi.org/10.5194/isprs-annals-X-4-W1-2022-715-2023

a significant increase in accuracy compared to machine learning techniques (Al-Najjar et al., 2019; Sameen et al., 2018; Wang & Wang, 2019). This study presents a method of combining the CNN model and DSM images to extract flooded areas. Therefore, in this study, an attempt was made using the DSM of the studied area to reduce this error. The polarimetric radar images during or near the occurrence of a flood on July 29, 2022, in Yazd city was processed. Then, by adding the three-dimensional layer of the area and the difference in the digital model of the area, it was introduced to the CNN algorithm using eCognition Developer software. Finally, the accuracy of the flood inundated map was assessed with field surveys of the inundated points.

2. METHODOLOGY

2.2 Case study

The area studied in this research is Yazd city. This city, with an area of about 2397 square kilometers, is located in 54°19' to 54°24' east longitude and 31°49' to 31°55' north latitude (Figure 1a). Figure 1-b shows the image of the false color combination of polarizations of the Sentinel-1 sensor for Yazd city. Yazd city is located between the two mountain ranges of Shirkoh and Kharanagh at an altitude of about 1200 meters above sea level (Figure 1c). The average annual rainfall in this province is about 69.5 mm, the minimum average temperature is -20 and the maximum is +47 degrees Celsius. Also, the maximum relative humidity of 23% in January shows that this province has had relatively high relative humidity in recent years. Yazd has a hot, dry, and desert climate. The fluctuation of temperature in summer and winter and even at night and day are high and variable. Therefore, it has two long hot seasons (from March to October) and a short cold one (from November to February) (https://data.irimo.ir).

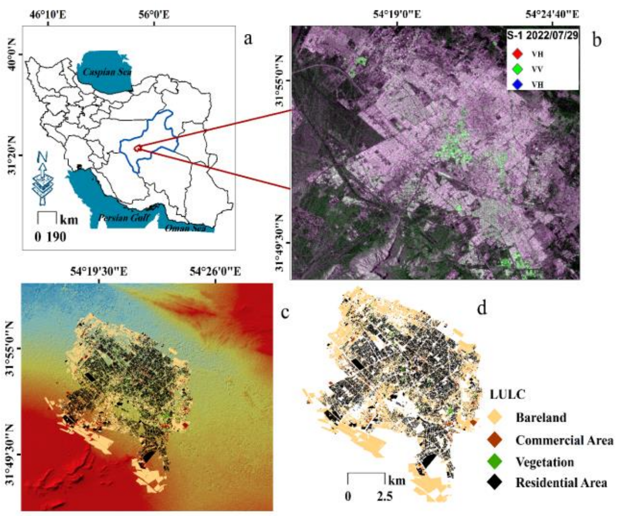

Figure 1. Location and geographic context of the Yazd city in central Iran (a), Sentinel1 image RGB: VH, VV, VH (b), SRTM 1sec of Yazd city (c), and (d) land use land cover (LULC) map of Yazd city.

2.3 Dataset

In this study, three polarimetric radar images, Sentinel-1 images before, after, and during a flood event in the city of Yazd, were pre-processed by removing thermal and border noise of the images, radiometric calibration, and correction from the ground surface with the help of the digital model of the area, and conversion to decibel units. Sentinel-1 C band images (Table 1) are accessed from the Alaska archive. Sentinel-1, C band Interferometric Wide (IW) swath Ground Range Detected (GRD) datasets (Table 1) are chosen for mapping flooding in Yazd city. The image preprocessing is done in a SNAP workflow that consists of seven major steps.

The test area is covered by 25 Interferometric Wide Swath Sentinel-1 data sets (IW, with a spatial resolution of 5×20 m) during the period 20 July 2022 to 08 august 2022 (as shown in Figure 1b), which are used for generating the Sand Exclusion Layer (SEL). The considered site covers an area of ~2397 km² (2612×2203 pixels). Because the data have been acquired in single polarization VH and VV, with two VH and VV, polarized data are used to have a longer and more consistent time series. These images were prepared between 2022/07/20-2022/07/28 before the flood, 2022/07/29 during the flood, and 2022/07/29-2022/08/09 after the flood in Yazd city. Sentinel-1A satellite was launched on April 2014 and Sentinel-1B on April 25, 2016 (Sharif et al., 2022). These images have a temporal resolution of once every six days for each area of the earth's surface. This temporal capability is very useful for monitoring natural disasters with a spatial resolution of about 10 meters.

Image Acquisition Parameters	Before image	During image	After image
Product type	Sentinel-1A GRD	Sentinel-1B GRD	Sentinel-1B GRD
Time	2022/07/24-02:22	2022/07/29-14:26	2022/08/09-14:27
Acquisition mode	IW	IW	IW
Pass	Descending	Ascending	Ascending
Relative orbit	488	130	130
Absolute orbit	44236	44302	44477

Table 1. Sentinel-1 datasets used in the present study.

2.4 Methodology

Figure 2 presents the step-by-step methodology adopted in this study. First, the raw images of the Sentinel-1 sensor were collected for the study area in the period before and after the flood, as well as the nearest days to the occurrence of the flood (July 29, 2022, to July 30, 2022). Then the steps required for pre-processing include removing thermal noise, removing image noise to reduce the amount of noise, audiometric corrections, correction using a digital model of the earth's surface (SRTM 1sec), applying a median filter to reduce image speckle, and finally converting into decibel (dB) units. The pre-processing has been done using the ESA SNAP application platform software of the European Space Agency (Figure 2).

The goal of image processing using SNAP software is to preserve consistent image features of model datasets from multiple images. Figure 3 shows the steps of image preprocessing and Sentinel-1 image processing from the raw data set to the generation of the scattering intensity value.

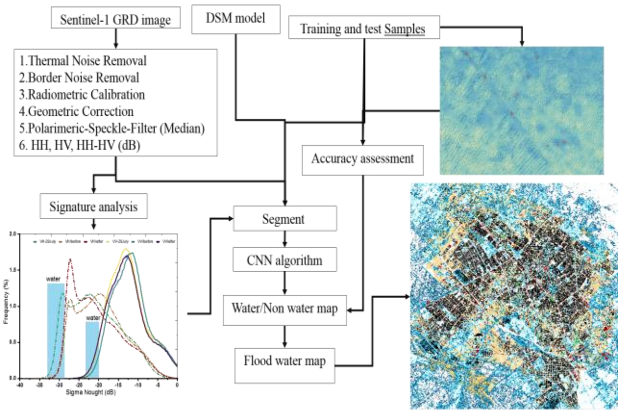

Figure 2. Schematic flowchart of the methodology adopted in this study to process the Digital Surface Model (DSM) and Sentinel-1 image in Convolutional Neural Network (CNN) algorithm.

2.5 AFTER THE PROCESSING STEPS, IT WAS ANALYZED ACCORDING TO THE FLOODED AREAS FOR THE PIXELS BELONGING TO WATER AND OTHER NON-WATER FEATURES IN BOTH POLARIZATIONS FOR THE PERIODS BEFORE, AFTER, AND DURING THE FLOOD TO IDENTIFY THE THRESHOLD OF PIXELS RELATED TO WATER. IN THE NEXT STEP, BY ADDING A THREE-DIMENSIONAL LAYER Evaluation Metrics

In this study, the criteria of overall accuracy (*OA*), and the Kappa index (*K*) were used to evaluate the flood inundation area detected by the developed approach in this study. The *OA* measures the percentage of total classified pixels that are truly labeled into the specific land cover classes. The *OA* was computed by dividing the total correctly classified pixels by the total number of pixels (*N*) in the error matrix, as is shown in Equation (1).

(DSM) of the ground surface with polarization images for three periods before, during, and after the flood, the flooded areas were classified using the CNN algorithm. As a result, the final flood map in Yazd city was prepared by removing the overlapping areas of flooded areas from each period. To evaluate the results, according to the available images of the flood that were prepared from the website of the Yazd municipality, training points were prepared for validation. This effort was visually captured and converted into training points in vector format.

2.6 Fusion in-decoder CNN Model

In this study, a Fusion in-Decoder networks using VV, VH, and DSM images was assumed and applied to detect the flood inundation classes. The proposed method is used from image patches with the size of N×N×3 pixels or N×N×3 pixels as input for each Encoder. The size of the dilation filter was defined as 5×5, which was the same as the kernel size in CNN model (Figure 3). Two following techniques were used to improve the network performance (Fathi & Shah-Hosseini, 2021; Garbin et al., 2020):

1) Batch Normalization, is applied to keep the distribution of the input values of each layer and increase the speed of learning. and

2) Dropout, is applied to reduce over-fitting and create different architectures by using removing neurons randomly in the last layer of each.

The patch-level analysis is usually used with deep-learning methods (e.g., CNN) in order to overcome challenges posed by speckle noise and segmentation optimization. These problems arise from pixel-level and object-level feature extraction (Garbin et al., 2020; Sameen et al., 2018). In a patch-level analysis, images are divided into a grid of N×N and then each patch is separately analyzed. The size of the image patch used to train the CNN was determined based on the spatial resolution of the Sentinel-1 image and the expected size of the objects in the scene.

Parameters input into the CNN model and architectures of Fusion in-Decoder are shown in Table 2 and Figure 3.

Parameters	Value
Image	2 image VV&VH and DSM model
Filter Size	5*5
Pooling Size	5*5
Convolution kernel Size	5
Batch Size	32
Steps per epoch	12
Optimizer	RMSprob (lr=0.001, rho=0.8, decay=0.8)
Activation function	ReLU/sigmoid (last layer)

Table 2. Parameters input into the CNN models.

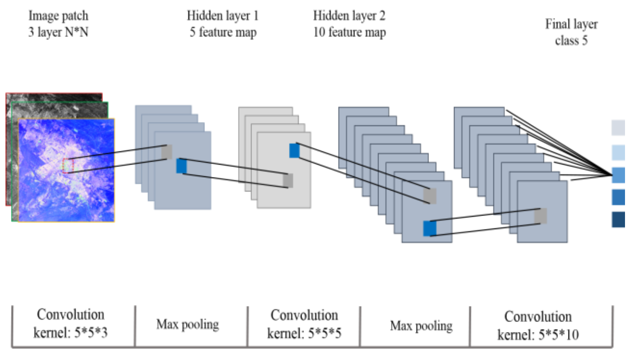

Figure 3. Architectures and parameter of the convolution CNN model in this study.

$$OA = \frac{\sum Dij}{N}, \qquad (1)$$

The K statistics is a discrete multivariate criterion used to assess the accuracy (Congalton, 1991; Elijah & Jensen, 1996). A Kappa analysis yields a K statistic, which is a quantitative metric of the level of agreement in correctly classified pixels (2):

$$K = \frac{N \sum_{i,j=1}^{m} Dij - \sum_{i,j=1}^{m} Ri.Cj}{N^2 - \sum_{i,j=1}^{m} - Ri.Cj}, \qquad (2)$$

Where *Dij*= the total number of correctly classified pixels,
i, *j*= row and column,
Ri= the total number of pixels in a row,
m= is the number of classes, and
Dij= the number of correctly classified pixels in a row and column
Ri = the total number of pixels in a row,
Cj= the total number of pixels in a column, and

N= the total number of pixels.

3. RESULTS AND DISCUSSION

The open water surfaces lead to low backscattering coefficients due to the specular reflection of the SAR signal on such features and the high permittivity of water (Martinis et al., 2018). This can also be seen in Figure 4, which shows the spatial distribution of water backscattering in both VV and VH polarization within the validation areas. In both polarizations, the classes have a great overlapping, which leads to misclassifications of the water areas. However, the distributions show a small shift between VV and VH polarization. While in VV polarization, in all three time periods before, after, and during the considered flood event, the mean backscattering is almost the same (water surfaces in VH= -30.34 dB for 29 to 30 July) and the standard deviation is similar (water surfaces in VH= 1.6).
Pixel counts appear to be slightly better distinguishable in VH polarized data (water surfaces in VV= -20.22 to -25 dB VV) (Figure 4). The difference in VH polarization redistribution during the flood compared to after and before the flood occurrence time is about 3.4 dB. However, this difference in polarization VV has a lot of overlap.
The optimal threshold value used to differentiate the flooded and non-flooded area was found as -30 dB when applying the Histogram thresholding method to the VH polarization dataset. The flood extent was increased (non-flood affected area was detected as a flood) when the threshold value was less than -30 dB. Whereas the actual flood-affected area was missed when applying the threshold value on the histogram of signal backscatter intensity greater than -24 dB. Hence, determining an optimal threshold value is crucial to overcome both over and under estimation in flood inundation mapping. Due to the low viewing angle of the VV polarization, the optimal threshold was found at -23dB. However, increased false flood alarms were obtained when applying threshold values less than the optimal (-23 dB).

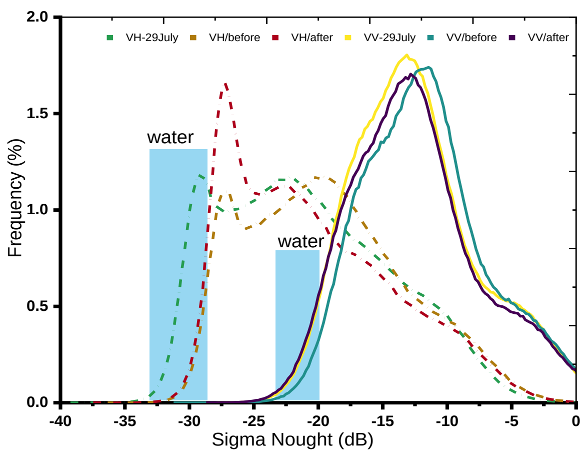

Figure 4. Histograms for water backscattering for VV and VH polarized Sentinel-1 data sets within the validation area in Yazd city (during 29 to 30 July 2022).

Figure 5 visually shows the results of Sentinel-1 image processing techniques for the identification of flood inundated areas after applying the CNN algorithm. According to the results shown in Figure 5, the proposed method showed good reliability in the classification of flooded areas in Yazd city, which is representative of a dry and desert area. The results showed user accuracy (UA) for the water class at 85%, overall accuracy (OA) at 88%, and a kappa coefficient of 0.83. In this algorithm, it was identified according to the redistribution range of five classes, which is considered to be about 5 cm for every 1 dB. This was achieved by experimental processing after trial and error. The weaker the power of the radar pulses, the smoother the water surface and the deeper the water layer. Moreover, the adopted methodology in this study makes it possible to estimate the depth of water in the flooded areas as shown in Figure 5. However, more studies are needed to attain robust and precise results, but the systematic methodology adopted in this study indicated that the use of the three-dimensional earth layer (DSM) has great potential as a promising tool in real-time monitoring of flood inundation areas, especially in urban cities with a dry and desert climate.

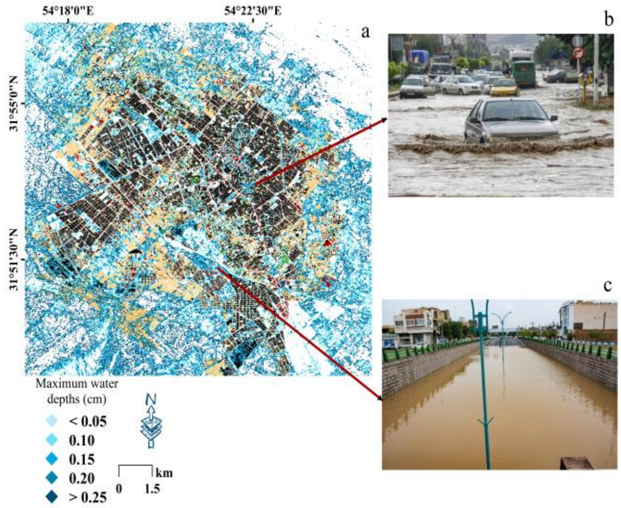

Figure 5. Flood inundation map and depth of water in Yazd city due to event storm July 28-30, 2022 produced by image processing and CNN Classifier (a), images of two flooded places in Yazd city: main street (b) , and railway station (c).

The Sentinel-1 data are acquired at 6 days intervals in Yazd city. Hence, not all flood events can be captured by satellite image. However, the flood maps generated from the satellite image proved to be useful information when long-time series multi- temporal images were analyzed as applied in this research.
Future studies could monitor flooding in Yazd city per 6 days interval using the Sentinel-1 images to provide useful information for flood risk assessment and management. The flood extent detected from the Sentinel-1 image can be useful for validating hydrodynamic flood models of the study area.
Detailed characteristics and areas exposed and vulnerable to flooding by combining remote sensing satellite images and field surveys are very valuable for flood risk assessment (Risi et al., 2020; Bekele et al., 2022). This information can reduce costs incurred by cities due to free access to remote sensing images. Therefore, it is suggested to investigate the combination of optical and radar images in future studies.
These results suggest that the combined model of CNN and DSM is a promising tool to classify approximately all of the classes with relatively high accuracy. However, more studies are needed to achieve better results. Also, using different remote sensing images can improve these results. Remotely-sensed images can provide real-time information on flood extension areas which is vital in flood risk mitigation. It also caused flooding on the city roads and damage to the city's historical buildings, blocking of traffic routes, and damage to the city's infrastructure facilities. The developed systematic approach is a

promising tool for real-time detection of urban flood extended areas, which can be repeatable for other cities.

4. CONCLUSION

In this study, the data of polarizations (VH, VV) Sentinel-1 were evaluated for mapping a recent flood event that occurred in Yazd city. In general, the method used in this study can help rapidly monitor flood-prone areas using the Sentinel-1 dataset. The following key results have been obtained based on the analysis of the research results:

The accuracy of generated flood maps using Sentinel-1 images is significantly affected by image polarization, the method of flood detection. Wet soil and low depths of flood show more errors in flood inundation detection. The final accuracy of the flood map was significantly higher using VH polarization data instead of VV one. However, these results can be different in areas with characteristics of surface phenomena as well as the type of soil in the area (for example, sandy and clay soils that have similar redistribution of radar pulses to flooded areas).

To observe the flooded areas of the city, the histogram threshold method was used in each polarization. The results of classification in each polarization alone do not show good results in some parts. Therefore, to reduce the final classification error, according to the threshold of the pixels representing water cover, the classification obtained from both polarizations was combined. Therefore, the use of both S-1 polarization to monitor the flood at the Yazd city level will lead to better results.

Overall, the multi-temporal analysis of Sentinel-1 images provides useful information about the extent of flooding. The surrounding area, due to being a desert, is mainly sandy and clay soil, which causes errors in the final results for areas outside the city. However, Sentinel-1 images showed acceptable results for flood-affected areas in the city.

In this study, the single threshold limit of Sentinel-1 radar sensor polarizations was evaluated to distinguish between flooded and non-flooded areas. But in future studies, this threshold can be defined more precisely according to the type of phenomena on the ground surface and the time conditions. Because checking the histogram of each polarization according to local conditions, soil characteristics and topographic conditions improve the results.

The processing of remote sensing images is very important in the mapping urban flooded areas. To achieve more accurate results in future studies, it is suggested to use high spatial resolution images of Sentinel-2 as optical images along with radar images. Optical images can improve the results according to the spectral characteristics.

REFERENCES

Al-Najjar, H. A. H., Kalantar, B., Pradhan, B., Saeidi, V., Halin, A. A., Ueda, N., & Mansor, S., 2019. Land Cover Classification from fused DSM and UAV Images Using Convolutional Neural Networks. Remote Sensing, 11(12), 1461. https://doi.org/10.3390/rs11121461.

Bayik, C., Abdikan, S., Ozbulak, G., Alasag, T., Aydemir, S., & Sanli, F. B., 2018. Exploiting multi-temporal Sentinel-1 SAR data for flood extend mapping. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci, 42(3), W4.

Bekele, T. W., Haile, A. T., Trigg, M. A., & Walsh, C. L., 2022. Evaluating a new method of remote sensing for flood mapping in the urban and peri-urban areas: Applied to Addis Ababa and the Akaki catchment in Ethiopia. Natural Hazards Research, 2(2), 97–110. https://doi.org/10.1016/j.nhres.2022.03.001

Carreño Conde, F., & De Mata Muñoz, M., 2019. Flood Monitoring Based on the Study of Sentinel-1 SAR Images: The Ebro River Case Study. Water, 11(12), 2454. https://doi.org/10.3390/w11122454.

Congalton, R. G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment, 37(1), 35–46.

DeVries, B., Huang, C., Armston, J., Huang, W., Jones, J. W., & Lang, M. W., 2020. Rapid and robust monitoring of flood events using Sentinel-1 and Landsat data on the Google Earth Engine. Remote Sensing of Environment, 240, 111664. https://doi.org/https://doi.org/10.1016/j.rse.2020.111664.

Elijah, W. R., & Jensen, J. R., 1996. Remote sensing of mangrove wetlands: relating canopy spectra to site-specific data. Remote Sensing, 62(939).

Fathi, M., & Shah-Hosseini, R., 2021. Automatic rice fields mapping by fusion of in-decoder CNN and data augmentation techniques on Landsat-8 multi-temporal images. Earth Observation and Geomatics Engineering, 5(2), 74–95.

Garbin, C., Zhu, X., & Marques, O., 2020. Dropout vs. batch normalization: an empirical study of their impact to deep learning. Multimedia Tools and Applications, 79(19–20), 12777–12815. https://doi.org/10.1007/s11042-019-08453-9.

Jahan, F., Zhou, J., Awrangjeb, M., & Gao, Y., 2018. Fusion of Hyperspectral and LiDAR Data Using Discriminant Correlation Analysis for Land Cover Classification. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11(10), 3905–3917. https://doi.org/10.1109/JSTARS.2018.2868142.

Li, Y., Martinis, S., Plank, S., & Ludwig, R., 2018. An automatic change detection approach for rapid flood mapping in Sentinel-1 SAR data. International Journal of Applied Earth Observation and Geoinformation, 73, 123–135. https://doi.org/https://doi.org/10.1016/j.jag.2018.05.023.

Li, Y., Martinis, S., & Wieland, M., 2019. 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, 178-191.

Marcos, D., Volpi, M., Kellenberger, B., & Tuia, D., 2018. Land cover mapping at very high resolution with rotation equivariant CNNs: Towards small yet accurate models. ISPRS Journal of Photogrammetry and Remote Sensing, 145, 96–107. https://doi.org/10.1016/j.isprsjprs.2018.01.021.

Martinis, S., Plank, S., & Ćwik, K., 2018. The Use of Sentinel-1 Time-Series Data to Improve Flood Monitoring in Arid Areas. Remote Sensing, 10(4), 583. https://doi.org/10.3390/rs10040583.

Mason, D. C., Dance, S. L., & Cloke, H. L., 2021. Floodwater detection in urban areas using Sentinel-1 and WorldDEM data. Journal of Applied Remote Sensing, 15(03). https://doi.org/10.1117/1.JRS.15.032003.

Risi, R. De, Jalayer, F., Paola, F. De, Carozza, S., 2020. From flood risk mapping toward reducing vulnerability: the case of Addis Ababa. Nat. Hazards 100 (1), 387–415. https://doi.org/10.1007/s11069-019-03817-8.

Sameen, M. I., Pradhan, B., & Aziz, O. S., 2018. Classification of Very High Resolution Aerial Photos Using Spectral-Spatial Convolutional Neural Networks. Journal of Sensors, 2018, 1–12. https://doi.org/10.1155/2018/7195432.

Sharif, M., Attarchi, S., & A.Kakroodi, A., 2022. Investigating the phenology changes of three plant species in different ecosystems using radar and optical data. Physical Geography Research Quarterly, 54(1), 111–133. https://doi.org/10.22059/jphgr.2022.334134.1007658.

Surampudi, S., & Yarrakula, K., 2020. Mapping and assessing spatial extent of floods from multitemporal synthetic aperture radar images: a case study on Brahmaputra River in Assam State, India. Environmental Science and Pollution Research, 27(2), 1521–1532. https://doi.org/10.1007/s11356-019-06849-6.
Peng, B., Meng, Z., Huang, Q., & Wang, C., 2019. Patch Similarity Convolutional Neural Network for Urban Flood Extent Mapping Using Bi-Temporal Satellite Multispectral Imagery. Remote Sensing, 11(21), 2492. https://doi.org/10.3390/rs11212492.

Tanim, A. H., McRae, C. B., Tavakol-Davani, H., & Goharian, E., 2022. Flood Detection in Urban Areas Using Satellite Imagery and Machine Learning. Water, 14(7), 1140. https://doi.org/10.3390/w14071140.

Wang, Y., & Wang, Z., 2019. A survey of recent work on fine-grained image classification techniques. Journal of Visual Communication and Image Representation, 59, 210–214. https://doi.org/10.1016/j.jvcir.2018.12.049.