AGPE The Royal Gondwana Research Journal of History, Science, Economic, Political and Social Science
ISSN (P): 2582:8800 | A Peer reviewed | Open Accsess & Indexed
Volume 02 | Issue 02| Oct-Dec 2021| *www.agpegondwanajournal.co.in* | Page No. 52-59

FLOOD DETECTION SYSTEM USING SENTINEL - 1 IMAGES AND EXTREME LEARNING MACHINE CLASSIFIER

Prakash S.[1], Joseph S.[2], Singh R.[3]*, Preethi P.[4] and Sreejith S.[5]

[1]Swetha Prakash, LBS Institute of Technology for Women, Trivandrum, Kerala, India.
[2] Shweta Joseph, LBS Institute of Technology for Women, Trivandrum, Kerala, India.
[3]Ritu Singh, LBS Institute of Technology for Women, Trivandrum, Kerala, India.
[4]Preethi P., LBS Institute of Technology for Women, Trivandrum, Kerala, India.
[5]Sreejith S., LBS Institute of Technology for Women, Trivandrum, Kerala, India.

Abstract:

Floods are one of the most common natural disasters, affecting millions of people worldwide. Floods occur when streams burst their banks, generally as a result of significant rainfall, and inundate areas that are not normally flooded. In August 2018, Kerala witnessed the devastating effect of floods which resulted in 13.86% of the land being inundated. Satellite imagery is one of the most effective techniques for assessing the extent of flood-affected areas with high spatial resolution. The approach of monitoring floods by using Sentinel-1 satellite data from Google Earth Engine (GEE) is presented. Using the satellite images, we have created a dataset after applying preprocessing techniques like resizing and thresholding. We have used Otsu thresholding in this work due to its ability to easily distinguish water and non-water pixels. An Extreme Learning Machine (ELM) model is proposed to identify the flood-prone regions in the chosen study area. We have compared our model with existing classifiers such as Decision Tree and Support Vector Machine and found our model performs better with a good consistency and accuracy score of 0.8787. These systems can be used for better preparedness and aid in monitoring the change or reconstructing the progress of a past flood.

Keywords: Extreme Learning Machine, Flood detection, Sentinel-1, Google Earth Engine

CORRESPONDING AUTHOR:	RESEARCH ARTICLE
Ritu Singh Research Scholar LBS Institute of Technology for Women, Trivandrum, Kerala, India. Email: ritutweets46@gmail.com	

Introduction:

Floods are one of the world's most prevalent and destructive natural disasters. The progress of flood detection models has contributed to risk reduction, economic crisis minimization, loss of human life, and property damage reduction related to floods. Machine learning (ML) techniques have greatly aided the progress of autonomous flood prediction and detection systems, resulting in improved performance and cost-effective solutions. One of the most cost-effective technologies for analysing inundated areas and drawing judgments about the level of damage in a given location is satellite-based imageries. They are very helpful in observing remote areas and are usually the only tool in places where rain gauge equipment is not installed to take real-time ground data. They have good spatial resolution and accuracy. Remote sensing is a dependable method of providing synoptic coverage over a large area at a low cost. It also overcomes the ground stations' inability to register data during a severe hydrological event. The readily available satellite data include Landsat 8, MODIS, Sentinel-1, Sentinel-2, etc. SAR data can be collected in all-weather conditions at both day and night time. Distinguishing water and non-water pixels is much easier in SAR images, which is why they are generally used in flood detection problems. Furthermore, multi-date imagery can serve in tracking change or recreating the path of a previous flood. Analysis of past flood occurrences is important to understand the flood-prone areas and what needs to be done to dampen the effects and to ensure safety to the people living near these areas.

1.1 Literature Survey

SAR data is most suitable for flood detection purposes as it can easily distinguish between water and non-water areas. (Anusha et al., 2020)[1].

There is a lack of availability of public datasets in the context of floods. Bonafilia et al. (2020)[3] have introduced Sen1Floods11, a surface water dataset that includes raw Sentinel-1 footage from 11 flood incidents across six continents.

Twelve parameters were considered for susceptibility assessment- slope, elevation, aspect, curvature, Stream Power Index, Topographic Wetness Index, lithology, rainfall, Normalized Difference Vegetation Index, river density, proximity to the river, and soil type. (Mohammadi et al., 2020)[2].

Cao et al. (2019)[4] introduced an unsupervised large-scale flood detection using an iteratively multi-scale chessboard segmentation-based tiles selection approach. For tiles with a bimodal Gaussian distribution, this method was used, and a non-parametric histogram-based thresholding procedure was used to locate water areas.

Images are filtered to remove speckle noise and further subjected to binarization. For the extraction of water bodies from optical remote sensing data, a spectral water index-based method called Modified Normalized Differential Water Index (MNDWI) is used (Vishnu, C. L., et al., 2019)[6].

Islam et al. (2021)[9] applied and evaluated two new hybrid ensemble models, Dagging and Random Subspace (RS) coupled with Artificial Neural Network (ANN), Random Forest (RF), and

Support Vector Machine (SVM) for modelling flood susceptibility maps. Dagging was found to be one of the most capable methods for flood-susceptible modeling.

Anupam et al. (2020)[10] used the mean daily gauge heights, mean daily rainfall, and the mean daily river discharge values as input for an extreme learning machine (ELM) regression model. To obtain the maximum coefficient of determination using the particle swarm optimization algorithm (PSO), the number of units in the ELM was optimized and a hybrid ELM-PSO model was created..

Sentinel-1 images, remote sensing data, field surveys, aerial photographs, and Google Earth was used to map flooded areas. A flood inventory was assembled based on flood events in 2008, 2012, 2016, and 2017 (Shahabi et al., 2020)[5].

2 Materials and Methods

2.1 Dataset

ALOS DSM: GLOBAL 30m, a global digital surface model (DSM) dataset with approximately 30 meters horizontal resolution is provided by JAXA Earth Observation Research Center. The S1 Ground Range Detected (GRD) scenes are included in the Sentinel-1 SAR GRD: C-band Synthetic Aperture Radar Ground Range Detected dataset. Thermal noise removal, radiometric calibration, and terrain correction steps by Sentinel-1 Toolbox were applied to pre-process each scene. NASA GES DISC at NASA Goddard Space Flight Center provides the GPM: Global Precipitation Measurement (GPM) v6 dataset. Every three hours, it delivers rainfall estimates and observations of rain and snow around the world. This dataset was used to retrieve the precipitation values of the study area during the image acquisition time period. JRC Global Surface Water Mapping Layers, v1.3 dataset provides statistics on the extent and change of those water surfaces along with maps of the location. WWF HydroSHEDS Flow Accumulation dataset offers geo-referenced datasets (vector and raster) for regional and global-scale applications. It is based on elevation data provided by NASA's Shuttle Radar Topography Mission(SRTM) in 2000 and contains hydrographic information such as flow accumulations. The Sentinel-1 image dataset of the area of interest is taken and the ALOS DSM dataset is used to identify the permanent water bodies in the area of interest and these images form the dataset that is used along with the precipitation values obtained from the GPM dataset in the area of interest. The JRC Global Surface Water Mapping Layers and WWF HydroSHEDS dataset are used to calculate flood extent in the study area on the basis of which flood or non-flood image is decided. We have used 108 images that were downloaded from Google Earth Engine, then preprocessed and then converted to CSV file which is the dataset we have used.

2.1.1 Study Area

The study area chosen is the area around the Mettur Stanley Reservoir constructed across the Kaveri River in Tamil Nadu. Sentinel-1 satellite images of the same are taken. We have acquired multi-temporal images between the period 2016-2020. A sample non-flood image of our created dataset is shown below in Fig 1.

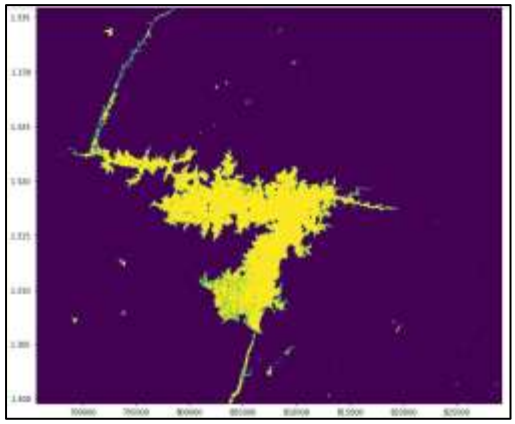

Fig 1: Sample non-flood image from the dataset

2.2 Methodology

2.2.1 Extreme Learning Machine (ELM)

G. Huang [11] invented the extreme learning machine (ELM) algorithm in 2006. It is basically a feedforward neural network. These neural networks were designed for feedforward neural networks with a single hidden layer (SLFNs). The input weights are picked at random, while the output weights of SLFNs are calculated analytically. It's popular for its quick and efficient learning pace, rapid convergence, and ease of use. They can have a single layer or numerous layers of hidden nodes, with no requirement to tune the parameters of the hidden nodes. Classification, regression, clustering, and feature learning can be implemented with ELM classifiers. ELM classifiers have better generalisation performance and can learn considerably faster than backpropagation networks. We propose to train our dataset with an ELM classifier and classify whether the Sentinel-1 image is flood or non-flood (label).

2.2.2 Proposed Work

Figure 2 depicts the methods employed in our suggested research.

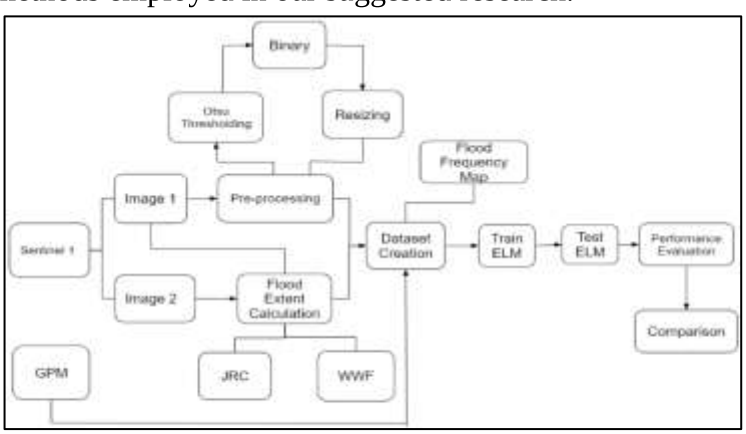

Fig 2: Proposed Model

The steps that used to complete the model are as follows:

I. Image Acquisition:

Sentinel-1 satellite images of the area of interest are taken. The transmitter-receiver polarization of these images is filtered out. VV(vertical transmit and Vertical receive) polarization is used for Otsu thresholding in assigning water and non-water pixels.VH(Vertical transmit and Horizontal receive) polarization is used for flood extent calculation. Along with the current image, an image after a period of 10 days is retrieved. Applying JRC Global surface water mapping and WWF HydroSHEDS flow accumulation layers, a water mask is formed. By dividing these two images, flood extent is computed in hectares. This value is used to assign class labels (i.e. flood and non-flood) in our dataset.

II. Image preprocessing:

Initial image preprocessing is done in Google Earth Engine Code Editor. The steps followed here are:

a) Otsu Thresholding

Otsu's thresholding method is a binarization algorithm commonly used to perform automatic image thresholding in computer vision and image processing. The algorithm determines this single intensity threshold automatically. Otsu's thresholding method comes under global thresholding wherein a single threshold is used for the whole image. In our proposed model, our aim is to find the threshold value to assign the image pixel value either 1(indicating water pixel) or 0(indicating non-water pixel). We first process the input image to obtain the distribution of pixels in the image as a histogram, then use the otsu technique to compute the threshold value T, and finally set the pixel intensity value to 1 for pixels with a value more than T and 0 for pixels with a value less than T.After applying Otsu thresholding to our dataset, we get binary images as output.

The preprocessing technique done in Google Colab's python notebook is:

b) Resizing

Machine learning models train faster on smaller images. So images will be resized to a consistent width by height ratio. Here, we have downsized the images in the dataset to a dimension of 512 X 512-pixel format.

III. Image processing

Training and Testing the ELM: Here, we train the ELM classifier to predict whether an area in the area of interest is flooded or non-flood using the created dataset. For training and testing, the dataset is separated into training and testing sets.The accuracy and efficiency of the machine learning model are determined by the accuracy after testing.

For the purpose of comparison, we employed three classifiers: ELM classifier, SVM classifier, and Decision Tree.

IV. Image postprocessing

In this step, we are mainly analysing all the images in the dataset and generating a flood frequency map by factorising the water pixels. Generated flood frequency map provides valuable insights regarding the areas susceptible to flooding based on varying frequencies. This information can facilitate effective crisis management of floods.

3 Experimental Results and Discussions

We used 3 different classifiers to compare their performance namely ELM classifier, SVM classifier, and Decision Tree Classifier. We divided the data into two sets: training and testing. The classifiers were implemented using different train test split ratios- 7:3, 8:2, and 9:1. From Table 1 we can understand that the ELM classifier is the most consistent performer in terms of classification using different train-test split ratios.

Table 1. Accuracy Comparison of Classifiers

Classifiers	**Test Accuracy Score**		
	Split Ratio 7:3	**Split Ratio 8:2**	**Split Ratio 9:1**
Extreme Learning Machine(ELM)	0.8787	0.86	0.909
Support Vector Machine(SVM)	0.69	0.69	0.909
Decision Tree(DT)	0.81	0.909	0.72

3.1 ROC Curve for the ELM Classifier:

The ROC curve (receiver operating characteristic curve) for our proposed model (ELM classifier) is shown in fig 3. The ROC curve is obtained by graphing the points (TPR, FPR), i.e. it is a plot of true positive rate (y-axis) vs false positive rate (x-axis) (x-axis). The closer the ROC curve is to the ROC space coordinates (0, 1), the greater the classifier's accuracy. The value of the area under the ROC curve (AUC) for a perfect classifier is 1.0. AUC is also a measure of the performance of a classifier. Here, we get AUC=0.88.

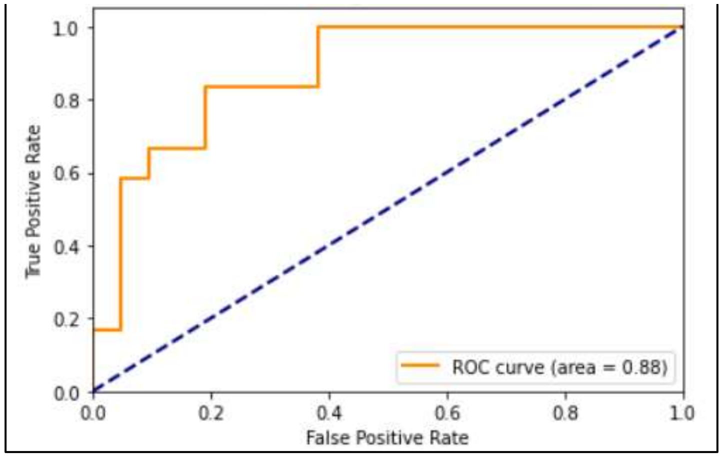

Fig 3 ROC curve for the proposed model

3.2 Flood Frequency Map:

The flood frequency map is plotted by calculating the frequent water pixels and colors were assigned based on the ratio of pixel occurrence. The red-colored area in the flood frequency map indicates the areas most susceptible to flood as the frequency of these pixels being water pixels is high. The other colors to the pixels are also assigned similarly.

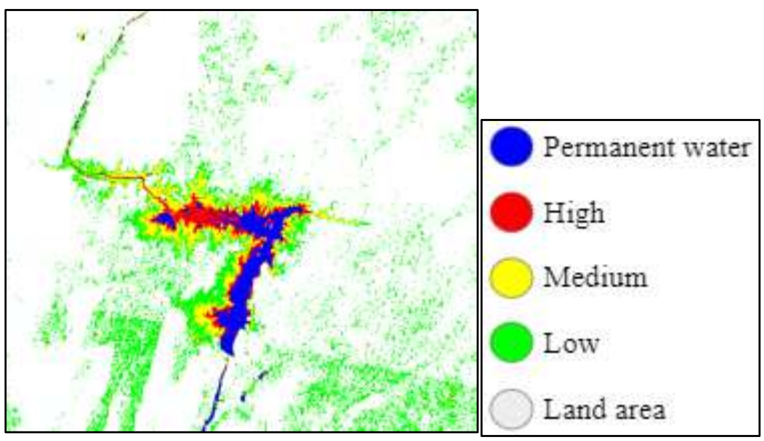

Fig 4 Flood Frequency Map

4 Conclusion:

In this project, we have developed an ELM-based classifier for flood detection in Sentinel-1 satellite images. Experiments carried out using the created dataset show that the proposed model was performing well in comparison to other existing machine learning classifiers. An accuracy score of 0.8787 (test-size: 30%) was obtained when the collected images were used for training. The proposed ELM-based model proves to be suitable for automatic flood detection problems, making use of readily available satellite imagery and classifying by training the ELM classifier. Obtained results will aid in effective crisis management of floods in the future.

References:

1. Anusha, N., & Bharathi, B. (2020). Flood detection and flood mapping using multi-temporal synthetic aperture radar and optical data. The Egyptian Journal of Remote Sensing and Space Science, 23(2), 207-219.
2. Mohammadi, A., Kamran, K. V., Karimzadeh, S., Shahabi, H., & Al-Ansari, N. (2020). Flood detection and susceptibility mapping using sentinel-1 time series, alternating decision trees, and bag-adtree models. Complexity, 2020.
3. Bonafilia, D., Tellman, B., Anderson, T., & Issenberg, E. (2020). 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 Workshops (pp. 210-211).
4. Cao, H., Zhang, H., Wang, C., & Zhang, B. (2019). Operational flood detection using Sentinel-1 SAR data over large areas. Water, 11(4), 786.
5. Shahabi, H., Shirzadi, A., Ghaderi, K., Omidvar, E., Al-Ansari, N., Clague, J. J., ... & Ahmad, A. (2020). Flood detection and susceptibility mapping using sentinel-1 remote sensing data and a machine learning approach: Hybrid intelligence of bagging ensemble based on k-nearest neighbor classifier. Remote Sensing, 12(2), 266.
6. Vishnu, C. L., Sajinkumar, K. S., Oommen, T., Coffman, R. A., Thrivikramji, K. P., Rani, V. R., & Keerthy, S. (2019). Satellite-based assessment of the August 2018 flood in parts of Kerala, India. Geomatics, Natural Hazards and Risk, 10(1), 758-767.
7.Miau, S., & Hung, W. H. (2020). River flooding forecasting and anomaly detection based on deep learning. IEEE Access, 8, 198384-198402.
8. Thakur, P. K., Dhote, P. R., Ranjan, R., Singh, S., Sharma, V., Srivastav, V., ... & Chauhan, P. Integrated Hydrological Study of August 2018 Kerala Floods using Remote Sensing and Hydrological Modelling.
9. Islam, A. R. M. T., Talukdar, S., Mahato, S., Kundu, S., Eibek, K. U., Pham, Q. B., ... & Linh, N. T. T. (2021). Flood susceptibility modelling using advanced ensemble machine learning models. Geoscience Frontiers, 12(3), 101075.
10. Anupam, S., & Pani, P. (2020). Flood forecasting using a hybrid extreme learning machine-particle swarm optimization algorithm (ELM-PSO) model. Modeling Earth Systems and Environment, 6(1), 341-347.
11. Huang, G. B., Zhu, Q. Y., & Siew, C. K. (2004, July). Extreme learning machine: a new learning scheme of feedforward neural networks. In 2004 IEEE international joint conference on neural networks (IEEE Cat. No. 04CH37541) (Vol. 2, pp. 985-990). Ieee.

■ ■ ■