

A Review on Flood Prediction Algorithms and A Deep Neural Network Model for Estimation of Flood Occurrence

Tabassum Farhana Ullah ^{1,*}, O.S. Gnana Prakasi ¹, P. Kanmani ¹

¹ Department of Computer Science and Engineering, CHRIST (Deemed to be University), Bangalore, India.

*Corresponding author E-Mail ID: tabassum.farhana@mtech.christuniversity.in

DOI: <https://doi.org/10.34256/irjmt2052>

Abstract

Flood occurs as often as possible happens due to many environmental changes in our planet in the present years. The occurrence and damages caused by flood is very high. Major cause of flood is due to heavy rainfall which in turn increases the water level of the rivers and other water bodies. The various factors that play a major role in the occurrence of rainfall are rise in temperature, humidity level, dew point, pressure in and around the area of concern, wind speed, etc. In order to reduce the number of victims due to flood it is necessary to have a system to predict flood occurrence. In this paper, we classify and analyzed the various prediction algorithms which show usage of Deep Neural Network produces better results. In addition, a design model has been proposed to predict the flood by training the Deep Neural Network with the above-mentioned factors.

Keywords: *Precipitation, Multilayer Perception Layer (MLP) architecture, Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM)network, Gated Recurrent Units (GRU).*

1. Introduction

Flood visits most part of the country, nearly on a yearly premise. The floods may occur due to heavy rainfall and the overflow of the rivers in those rainy regions. The districts of those regions under the flood encounter precipitation, year after year. The precipitation contributes a great deal to the flooding. Areas situated in a plain, this does not improve the situation much. When it rains in the upper area, the lower fields experience the anger of the waterway. But due to recurrent and merciless occurrence of flood every year the people of those regions, the flora and fauna faces lots of loss and endangerment. To control the flooding, the experts can manufacture dykes on either side of the channel to confine the flood of water. Without such measures, the flooding will keep on unleashing ruin and decimation. The intermittent decimation can be ceased just if measures are taken to spare individuals and property.

In other to control this damage, it is necessary to propose a prediction algorithm to know about the occurrence of flood in the area beforehand so that proper measures can be taken to prevent damages and preserve the nature. However, the prediction of flood lead time and occurrence location is fundamentally complex due to the dynamic nature of climate condition. Therefore, nowadays major flood prediction models are mainly data-specific and involve various simplified assumptions. The neural networks play an important role in flood prediction and the accuracy could be increased by applying suitable deep learning algorithms. This paper helps to analyze few prediction models and propose a model based on Recurrent Neural Network with LSTM and GRU.

The rest of the paper has been organized as follows. The section II shows various classification of the prediction algorithm and few methods in each classification. Section III shows the proposed model and its working functions, followed by section IV that concludes the paper.

2. Existing Techniques

In this, various existing techniques for flood prediction are discussed. Fig.1 shows few classifications of flood prediction algorithms.

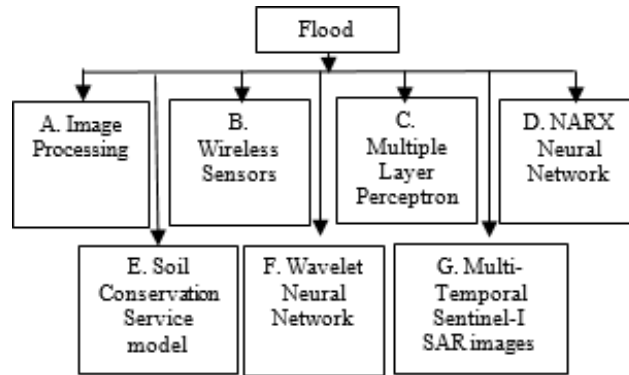


Fig 1. Various flood prediction techniques used during the survey

A. Digital Image Processing

Image Processing is one of most prominent method to predict the floods with the help of digital image made up of finite number of components known as pixels [1].

Complex water flow caused by Local Rainfall was predicted by [2] infrared images captured at low cost. Linear Regression algorithm was applied to predict the flood. This is also integrated with sensors to read the data and to improve the accuracy [3]. Using the digital images obtained, the land area and the water level are extracted. Estimate the difference between the land area and the water-level for the selected region. If the difference between the two areas is close to zero it is highly possible that the area could be flooded. If the difference is negative then flood has already occurred [4-7].

[8] used image segmentation with semi-supervised learning algorithm to handle flood mapping. The algorithm helps to divide the satellite image into patches with graph-based approach to detect the water areas. Monitoring and prediction of flood will be efficient when we made the observation from the space. But the usage of optical image may be failed in the presence of cloud. The image capture from Synthetic Aperture Radar (SAR) [16] provide valuable information as they are independent of weather conditions.

B. Wireless Sensors

Wireless Sensor Network has a wide range of applications including weather prediction and a major impact in predicting the flood [5]. Wireless Sensor Network (WSN) is a component of information retrieval that provides more optimal results in obtaining data time series.

[5] predicts the water elevation level in data time series controlled by remote site and control center. Water level sensor system with real-time river-level monitoring helps in flood prediction. The water level sensor system defined by [2] uses Linear Regression Approach for data assimilation to the collected information from the sensor. This helps to improve the accuracy in predicting the flood.

C. Multiple Layer Perceptron

Multi-Layer Perceptron is a feed-forward artificial neural network. The network is integrated with numerous nodes through linking weights. The main purpose of the MLP network is to map the group of inputs into a set of desired output [6]. This MLP also plays a major role in flood prediction by processing the collected information.

[5] used Multilayer Perceptron a deep learning algorithm in time series data with neural network. Produce better results in water elevation on downstream canal.

[6] The activation functions such as Sigmoid or logistic, tanH, Rectified Linear Unit (ReLU) are used in hidden layers to fetch the occurrence of flood in the future is predicted. The following equations helps to define the activation function as defined in [6] is given in equation 1 and 2.

$$H_j = f\left(\sum_{i=1}^n (w_{ji}x_i + b_i)\right) \quad (1)$$

$$y = f\left(\sum_{i=1}^m (w_{kj}H_i + b_0)\right) \quad (2)$$

D. NARX Neural Network

Nonlinear dynamic systems are modelled using dynamical neural architecture (NARX) [7]. It has been proven that utilizing NARX systems, instead of traditional recurrent systems helps in improving the prediction with no computational loss and that they are at any rate equal to Turing machines [9] helps to improve the prediction by incorporating machine learning algorithms to train the dataset generated. [10] shows gradient descent learning can be more successful in NARX systems than in other intermittent structures with "hidden states". [7] proposed an architectural recurrent neural network with embedded memory for flood prediction. This uses non-linear autoregressive model to model the input and the analysis is done with the chaotic time series prediction. [11] proposed flood water level modelling to overcome nonlinearity problem in flood prediction. The methods check the preciseness of the true model and the prediction model to predict the error. The best fit for this model is 87.0799% and RMSE is 0.0806m.

E. Soil Conservation Service Model

[13] The traditional Soil Conservation Service model is low efficient for medium based flood forecast. To overcome this streaming big data processing is integrated with Soil Conservation Service in which basins are divided into sub-basins and are executed in parallel to predict the flood. The results show that this overall computation time is reduced by 73%. [12] an analytical model using Soil Conversation serviced is proposed for flood prediction. [14] proposed a survey of different soil conservation methods and its various applications and the various improvement to be done in the model for flood prediction.

[15] shows an improvement in SCS model to adapt runoff calculation in large and small basins. [16] uses remote sensing and geographic information system to extract and analyse the spatial and attribute data. Soil Conservation Service curve number method (SCC-CN) is applied to calculate the runoff coefficients of different underlying surfaces characteristics. Then the amount of rainwater collected in the basin is estimated and it achieves good results.

F. Wavelet Neural Network

Wavelet neural network theory works on reconstruction theory of wavelet function. The combination between wavelet transformation and neural network is built by wavelet decomposition of consistent approaching. This helps in predicting the flood. [15] used wavelet

neural networks in which water flux and water level are used as an input for which the time frequency feature was analysed. The accuracy of the result is increased by combing the low and high frequencies of signal which are decomposed into large scales.

[17] used scale vector, transfer vector, weight vector for the wavelet network. In which the weights and displacement and scale vectors are updated based on the activation function the main advantage of the prediction is the non-linear time series and non-linear function approach.

G. Multi Temporal Sentinel-1SAR images

[17] used temporal and spatial availability of data in which binarization and image differences are considered for pre-processing which is mapped with class of type water, urban, forest, agriculture and bare land respectively. Results show that it produces a fast and more appropriate result than existing methods.

Table 1. Machine Learning Algorithms and Their Gaps

Method	Explanation	Gaps found
Decision Tree	A tree structure where the nodes represent questions and the leaves represent the corresponding category of documents. After having created the tree, a new document can easily be categorized by putting it in the root node of tree and let it run through the query structure until it reaches a certain leaf	The high classification error rate while training set is small in comparison with the number of classes
k-Nearest Neighbour	The categorization itself is usually performed by comparing the category frequencies of the k nearest documents (neighbours)	The main disadvantage of the KNN algorithm is that it is a lazy learner, i.e., it does not learn anything from the training data and simply uses the training data itself for classification
SVM Classifiers	SVM classifiers attempt to partition the dataspace with the use of linear or non-linear delineations between the different classes	Limitations are: speed and size, both in training and testing
Neural Network Classifiers	CNN are basically just several layers of convolutions with different functions	Whole network is disturbed if any of the layer's functionality goes wrong
Bayesian Classifiers	There are two groups of Bayesian approaches in document categorization: Naïve and non-naïve Bayesian approaches. The naïve part of the former is the assumption of word independence, meaning that the word order is irrelevant and consequently that the presence of one word does not affect the presence or absence of another one	A disadvantage of Bayesian approaches in general is that they can only process binary feature vectors
Other Classifiers	Other classification techniques include random forest, recurrent neural networks, pattern-based classification	Every method has its own advantage and disadvantages, few of the disadvantages, few of the disadvantages of various techniques have been discussed

[16] applies backscatter temporal variability on the SAR image to classify open water body and the land. They done the analysis in the heritage site Kaziranga National Park, Natural Wonder of the World, Assam during the monsoon period of 2017 with six SAR images captured during the period of 21st march to 26th July 2017. This helps to predict the flood in the region. [19] proposed image fusion technique, a combination of two statistical threshold approaches to increase the accuracy of the prediction for optical and SAR images.

3. Novel Approach for Flood Prediction

From the above survey, it has been identified that Neural network plays an important role in predicting the flood. So, in this, a model for flood prediction based on Recurrent Neural Network is proposed. Two main issues faced by existing RNN are how to get the knowledge of the previous steps and vanishing gradient are solved in this model.

First problem can be solved by the introduction of the hidden layers [18] which is responsible for remembering the information about a sequence. For each input, the hidden layer uses the same set of parameters to perform the same task. Downside of RNN is that it has to put up with the vanishing gradient problem [19] with Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU).

This LSTM and GRU have an impact in analyzing the data using deep learning algorithms. The combination of these two methodologies will improve in the accuracy of the prediction and the occurrence of flood in the near future. The flow diagram of the model is shown in the Fig. 2.

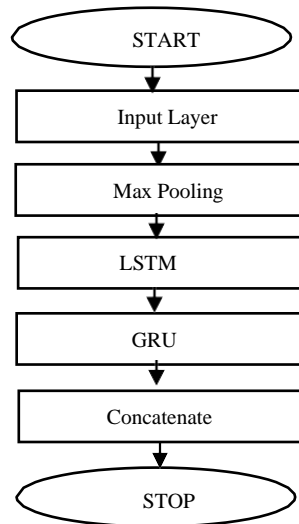


Fig 2. Flow diagram of proposed methodology for predicting flood in Assam using Neural Network (LSTM & GRU)

Input layer will have the parameters as pressure, precipitation amount, dew point, wind speed which are the factors considered for the proposed model. The next step is Max pooling, a sample-based discretization process that down sample an input representation by reducing its dimensionality with the features contained in the sub-regions. Next to overcome the vanishing gradient problem of recurrent neural network, Long Short-Term Memory (LSTM) is used. LSTM contains four gates namely the input gate, forget gate, cell gate and the output gate along with pointwise multiplication and pointwise addition. The trained data will be then fed to the Gated Recurrent Unit (GRU). The GRU has two gates namely reset gate and the update gate along with pointwise multiplication and pointwise addition.

All these are concatenated and formulated into a single output, which gives us the desired output for predicting the occurrence of flood using the above-mentioned parameters as inputs. The conventional method uses sensors in major rivers to detect the water level of the rivers. These data are sent to a remote server where using data processing the occurrence of flood is predicted. Various other factors that causes flood are not taken into consideration. However, with this model the various other factors that lead to the occurrence of flood are taken into consideration which improves the accuracy for predicting flood over the conventional method.

4. Conclusion

The classification of the prediction algorithms is discussed and we modeled a prediction system based on neural network using a combination of LSTM and GRU to predict the occurrence of flood. In future, the model can be trained with the factors like temperature, pressure, precipitation amount, dew point, wind speed, etc. which plays an important role in the occurrence of flood. These factors will be used to train and test the neural network model and finally predict the occurrence of flood in the near future.

References

- [1] R. C. Gonzalez, R. E. Woods, (1987) Digital Image Processing 2nd edition, *Prentice Hall*, 1.
- [2] K. Hiroi, and N. Kawaguchi, (2016) Flood Eye- Real-time flash flood prediction system for urban complex water flow, *IEEE Sensors*, IEEE, 1-3.
- [3] J. Liang, P. Jacobs, S. Parthasarathy, (2016) Human-guided Flood Mapping on Satellite Images, *KDD 2016 Workshop on Interactive Data Exploration and Analytics*, 76-85.
- [4] S A. Krishnan, S. Ganesh, S. S. Satheendran, A. Varghese, Flood Mapping and Damage Assessment of Kaziranga National Park, Assam using Multi-Temporal Sentinel-1 Synthetic Aperture Radar Images, *Indian Journal of Scientific Research*, 18 (2018) 1-8,
- [5] H. Karl, A. Willig, (2007) Protocols and Architectures for Wireless Sensor Networks, 1st edition, *John Wiley & Sons*, 1-507.
- [6] I. R. Widiyari, L. E. Nugroho, W. Widyawan, Deep Learning Multilayer Perceptron (MLP) for Flood Prediction Model Using Wireless Sensor Network Based Hydrology Time Series Data Mining, *In 2017 International Conference on Innovative and Creative Information Technology (ICITech)* IEEE, 1-5.
- [7] N. H. B. Harun, W. L. Woo, S. S. Dlay, (2010) Performance of Keystroke Biometrics Authentication System Using Multilayer Perceptron Neural Network (MLP NN), *In International Conference on Computer and Communication Engineering (ICCCE'10)*, IEEE, 1-6.
- [8] S. M. Faradonbeh, F. Safi Esfahani (2019) A review on Neural Turing Machine, *arXiv*,
- [9] S. Mahmoud, A. Lotfi, C. Langensiepen, Behavioural Pattern Identification and Prediction in Intelligent Environments, *Applied Soft Computing Journal*, 13 (2013) 1813-1822.
- [10] E. Diaconescu, The Use of NARX Neural Networks to Predict Chaotic Time Series, *WSEAS Transactions on Computer Research*, 3 (2008) 182-191,
- [11] F. A. Ruslan, Z. M. Zain, R. Adnan, Flood Water Level Modeling and Prediction Using Narx Neural Network: Case Study at Kelang River, *In 2014 IEEE 10th International Colloquium on Signal Processing and its Applications*, IEEE, 204-207.

- [12] M. Huang, J. Gallic hand, Z. Wang, M. Goulet, A Modification to the Soil Conservation Service Curve Number Method for Steep Slopes in The Loess Plateau of China, *Hydrological Processes*, 20 (2006) 579-589.
- [13] Y. Wu, Y. Tao, F. Liu, Flood Zoning Calculation in Zhejiang Province Based on Regional Regression Analytical Method, *Journal of Hohai University, Nature Sciences*, pp. 39-43, 2015.
- [14] P. K. Singh, P. K. Bhunya, S. K. Mishra, U. C. Chaube, A Sediment Graph Model Based on SCS-CN Method, *Journal of Hydrology*, 349 (2008) 244-255.
- [15] Z. Shaozhong, Y. Juqin, (2010) Flux and Level Prediction based on A Wavelet Neural Network Flood Model, *In 2010 Third International Symposium on Knowledge Acquisition and Modeling*, IEEE. 67-70.
- [16] B. Tavus, S. Kocaman, C. Gokceoglu, H. A. Nefeslioglu, Considerations on the use of Sentinel-1 Data in Flood Mapping in Urban Areas: Ankara (Turkey) 2018 Floods, *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, XLII-5 (2018) 575-581.
- [17] Y-J Kwak, R. Pelich, J. Park, W. Takeuchi, (2018) Improved Flood Mapping Based on the Fusion of Multiple Satellite Data Sources and In-Situ Data, *In 2018 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, 3521-3523,
- [18] T. Mikolov, G. Zweig, (2012) Context Dependent Recurrent Neural Network Language Model, *In 2012 IEEE Spoken Language Technology Workshop (SLT)*, IEEE, 234-239.
- [19] H. Sak, A. Senior, F. Beaufays, (2014) Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modeling, *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, 338-342,

Funding

This study was not funded by any grant

Conflict of interest

None of the authors have any conflicts of interest to declare.

About The License

The text of this article is licensed under a Creative Commons Attribution 4.0 International License

Cite this Article

Tabassum Farhana Ullah, O.S. Gnana Prakasi, P. Kanmani, A Review on Flood Prediction Algorithms and A Deep Neural Network Model for Estimation of Flood Occurrence, *International Research Journal of Multidisciplinary Technovation*, Vol 2, Iss 5 (2020) 8-14.

DOI: <https://doi.org/10.34256/irjmt2052>