Computer Vision Application for Effective Flood Forecasting in Public Health Systems

Uruj Jaleel[1], R Lalmawipuii[2]
[1]*Associate Professor, Department of CS & IT, Kalinga University, Raipur, India*
[2]*Research Scholar, Department of CS & IT, Kalinga University, Raipur, India*

KEYWORDS

Flood, Public Health, Greylag goose driven redefined Long Short-Term Memory (GG-RLSTM)

ABSTRACT

Among the natural disasters, floods are considered as one of the most devastating since they can damage infrastructures, force people leave their homes, and negatively impact the public health. The training of the model seems to be computationally intensive and could prove an interruption where computational resources are a limited commodity. In this paper, we present the Greylag goose driven redefined Long Short-Term Memory (GG-RLSTM) model to be implemented by public health systems in effective flood forecasting. Thus, adapting behavioural patterns in Greylag geese, GG-RLSTM enhances the theoretical structure of LSTM and increases its capability to capture complex relations in the flood processes. From the Kaggle dataset, we gathered meteorological evidence, and geographical features. The results from the experiments indicate that effectiveness of the proposed GG-RLSTM model is higher than the other conventional methods in terms of accuracy (89%), precision (88%), recall (87%), and f1-score (90%). Due to the effectiveness of the model and its applicability in various situations, public health systems can likely adopt it and commence preventive flood measures.

1. Introduction

Minimizing the risks as well as increasing the preparedness for natural disasters are two concepts that make flood forecasting important for the protective health of the public health systems [3]. With the situation in the field of public health described in detail, floods present significant threats to populous regions since they are normally unpredictable and disastrous [1]. Hurricanes can be very destructive and can affect the facility through flooding in several direct ways, and also by interrupting fundamental services and posing severe threats to the health of the community. One of the immediate concerns is that the movement of populations, which can lead to crowded confinement and increased spread of communicable diseases, is another one [2]. These are the same as above only with the perspective of injury and death due to drowning or physical harm. In flood-affected areas, drinking water sources also get contaminated and this consequence in diseases such as cholera and typhoid fever [15]. To minimize such health concerns, preventive measures must be taken and this must therefore be preceded by timely and accurate flood predictions [4]. Because of the increase in the methods of predicting floods, experts can do it better using hydrological models, flood occurrences, and meteorological factors [12]. The authorities of public health systems can stock sufficient quantities of medical supplies, mobilize rescue groups, and evacuate the susceptible population when the intensity and severity of floods can be predicted [5]. Therefore, in this paper, we introduce the Greylag goose driven redefined LSTM model or the GG-RLSTM to be adopted by the public health systems when carrying out effective flood forecasting [6]. The related works are described in the second part. The approach is shown in the third part. The outcome is shown in Part four and the fifth part presents the conclusion.

Related works

In an attempt to overcome these problems, Anbarasan et al. [13] proposed the use of Big Data (BD), Convolutional Deep Neural Networks (CDNN) and Internet of Things (IoT) for the identification of flood catastrophes. Many opportunities have been provided by the recent advancements in IoT and BD devices for disaster management platforms and disaster-related authority emergency participants, public health, police, and firefighters to obtain outstanding support and enhanced perception for prompt and dependable decisions [9]. To develop a food forecasting approach, they employed the multilayer perceptron of Dtissibe et al. [7] in the investigation with discharged as the input-output parameter. The efficacy of the concept with a strong forecasting power was demonstrated by the outcomes of extensive experiments conducted on the constructed model. Kan et al. [8] proposed a unique hybrid machine learning (HML) hydrological system for flood forecasting purposes in the investigation by combining the Artificial Neural Network (ANN) with the K-nearest neighbor approach. The HML hydrological algorithm's practical applications demonstrated its good performance and consistent stability, opening up the possibility of additional uses in flood forecasting issues. Xu et al. [14] presented a prediction model that mimics the hourly rainfall runoff association using the architecture of a temporal convolutional network (TCN). It was demonstrated that the TCN was a useful technique for hydrological forecasting and had a quicker rate of convergence. The LSTM method, Bayesian optimizing, and transfer learning technique were all integrated in the investigation of Zhou et al. [10] Deep learning technology-based data-driven flood forecasting strategy. The outcomes unequivocally demonstrated that the model could, given a variety of hyetograph inputs, reliably construct time series flood maps as well as maximum water depths at substantially lower computing costs [16].

Data gathering

Kaggle dataset https://www.kaggle.com/datasets/rajanand/rainfall-in-india has been acquired. The dataset includes monthly rainfall information for 36 Indian meteorological subdivisions.

Greylag goose driven redefined LSTM (GG-RLSTM)

Greylag Goose Optimization (GGO)

The GGO technique initially creates a set of randomly initialized people, each of whom stands for a possible fix for the current issue. The population is referred to as $W_j (j = 1,2, \dots, m)$ in the GGO structure, where m is the size of a 'gaggle'. To assess these people's quality, an impartial function E_m, is determined for every W_j. The value O stands for the best possible solution. The GGO algorithm splits people into two groups: exploiting agents m_2 and exploring m_1 using a dynamic grouping technique. The efficiency of the current finest solution is used to iteratively change the distribution of solutions among these categories and the procedures for overseeing these groups responsible for exploring and exploiting. In the beginning, GGO splits people equally between exploring and exploitation. The exploiting category m_2 grows larger and the exploring category m_1 gets smaller as the iterations go on. But if the impartial parameter value of the optimum solution stays constant for three sequential iterations, the algorithm expands the composition of the exploring category m_1 to avoid becoming trapped in local optima.

Exploring

Exploring intriguing areas in the search space (SS) and directing the technique away from less-than-ideal solutions and finally toward the globally optimal solution are two important tasks that the exploring phase performs. Individual geese scout the area around their present location in quest of better spots during the exploring period. Throughout this procedure, possible local solutions are evaluated iteratively to determine which one produces the highest fitness value. This is achieved via the GGO method using the following equations, which are given as:

$$W(s+1) = W^*(s) - B.|D.W^*(s) - W(s)| \tag{1}$$

The search agent's (SAs) current location is represented by $W(s)$ in this case. $W^*(s)$ represents the prey's position, while $W(s+1)$ represents the SAs next updated location. Vectors B and D are provided as:

$$B = 2 \propto q_1 - b \tag{2}$$

$$D = 2q_2 \tag{3}$$

In this case, matrices B and D contain random values within the interval [0, 1], and parameter b drops linearly from 2 to 0 throughout iterations. Three randomly chosen SAs additionally traverse the SS known as W_{o1} W_{o2} W_{o3}, enhancing the algorithm's exploring capability and ensuring that it is not solely dependent on leader location. Next, any SAs position update for $B > 1$ is provided as:

$$W(s+1) = \omega_1 * W_{O1} + y * \omega_2 * (W_{O2} - W_{O3}) + (1-y) * \omega_3 * (W - W_{O1}) \tag{4}$$

ω_1,ω_2 and ω_3 are located between [0-2]. The exponential decline in y is given as:

$$y = 1 - \left(\frac{s}{S}\right)^2 \tag{5}$$

The location update equation is provided as follows for the variable $B < 1$ and the declining value of b:

$$W(s+1) = \omega_4 * |W^*(s) - W(s).e^{bl.\cos(2\pi k)} + [2\omega_1(q_4 + q_5)] * W^*(s) \tag{6}$$

In this case, the parameters a are all constants, ω_4 ranges within [0, 2], while q_4 *and* q_5 both fall between [0, 1]. The random parameter k has values in [-1, 1].

Exploiting

The GGO algorithm recognizes the individual with the optimum fitness after each phase as the executive. To direct its exploiting activities, the GGO uses two unique methodologies, which are explained below. Moving regarding the optimum solution: To find the optimal solution, apply the preceding equation. The three optimal SAs W_{t1} W_{t2} *and* W_{t3}, direct additional random SAs W to modify their locations approaching the optimal location of prey. In mathematics, it is defined as:

$$W_1 = W_{t1} - B_1.|D_1.W_{t1} - W| \tag{7}$$

$$W_2 = W_{t2} - B_2.|D_1.W_{t2} - W| \tag{8}$$

$$W_3 = W_{t3} - B_3.|D_1.W_{t3} - W| \tag{9}$$

The population's revised location is provided as

$$W(s+1) = \frac{W_1 + W_2 + W_3}{3} \tag{10}$$

B and D are calculated using Equations 2 and 3.

Exploiting the division around the optimal response:

The algorithm prioritizes identifying solutions located near the current optimum solution. This is based on the assumption that further improvements may be found in this area. A subset of individuals, designated as W_e, focuses on this localized search. The GGO accomplishes this process utilizing the following equation.

$$W(s+1) = W(s) + C(1-y) * \omega * (W - W_e) \tag{11}$$

RLSTM

Although LSTM performs well when processing data patterns with evenly dispersed information, it fails to function effectively when dealing with sequences having unequal information transfer among phases. It improves the LSTM input gate, adds an Exponential Linear Unit (ELU) activation section, eliminates elements, and streamlines operation. The control signal determines what data from the current stage can be transferred, and the output gates (OGs) finish the next hidden step. The model can arbitrarily forget gates (FGs) or learn new data by adjusting FGs and OGs. By purposefully erasing data, the FG effectively reduces overfitting.

$$e_s = \sigma(X_e.[g_{s-1}, w_s] + a_e) \tag{12}$$

$$D_s = e_s * D_{s-1} + \tanh(X_D.[g_{s-1}, w_s] + a_D) \tag{13}$$

$$p_s = elu\left(X_p.[g_{s-1}, w_s] + a_p\right) \tag{14}$$

$$g_s = p_s * \tanh(D_s) \tag{15}$$

The RLSTM approach loads and transmits data based on variations in moment steps, selecting and discarding fresh or prior cell status data. It is possible to effectively eliminate the issue of data dispersion generated by too long a duration by transferring the output data to the storage unit at a specific point in duration and by utilizing lengthier setting parameters.

GG-RLSTM

Public health management during flood disasters is made feasible via the flexibility of GG-RLSTM, which issues forecast results in a timely and accurate manner. They contain equally valuable real-time adaptation to the dynamic environmental status. GG-RLSTM reinterprets LSTM, which provides a new way of handling the complexity that flood prediction involves in the contexts of public health systems by utilizing avian behavior. Due to the inclusion of biological adaptability and computer modeling, GG-RLSTM serves as a way to progress early warning systems, facilitate more proactive catastrophe prevention, and reduce the health hazards linked with flooding.

2. Results and discussion

The system features an Intel i5 13th Gen and runs on Windows 10 with 16 GB RAM. Python 3.11 was used to execute the GG-RLSTM strategy and evaluate the method's efficiency. Our proposed GG-RLSTM strategy is evaluated with the existing methods such as Decision Tree Classifier (DTC), K-Nearest Neighbours (KNN), and Support Vector Classifier (SVC) [11]. Accuracy measures evaluate how well flood events are predicted, which is important for organizing public health responses and preparations. A comparison of accuracy is presented in Figure 1. Our suggested GG-RLSTM approach performed (89%), in contrast to (84%), (84%), and (79%) of the current techniques such as SVC, KNN, and DTC. The outcomes demonstrate that the suggested strategy outperforms existing methods.

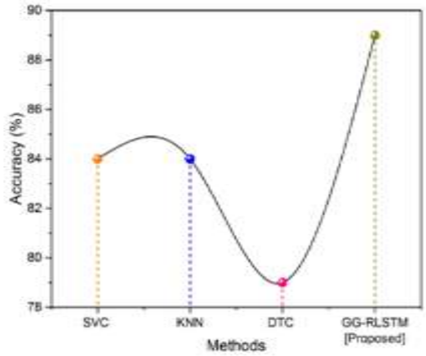

Figure. 1 Result of accuracy

Precision guarantees effective early warnings, which are essential for prompt public health actions and the distribution of resources in the case of flooding. A comparison of precision is presented in Figure 2. Our suggested GG-RLSTM approach performed (88%), in contrast to (83%), (83%), and (80%) of the current techniques such as SVC, KNN, and DTC. The outcomes demonstrate that the suggested strategy outperforms existing methods.

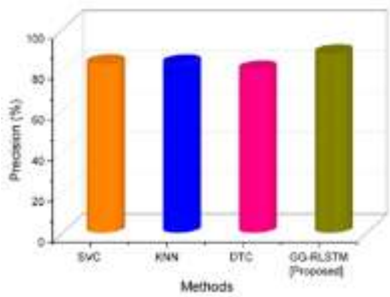

Figure 2 Result of precision

Recall assesses the model's capacity to include all pertinent flood cases, which is important for prompting public health initiatives and mitigation plans. A comparison of recall is presented in Figure 3. Our suggested GG-RLSTM approach performed (87%), in contrast to (84%), (84%), and (79%) of the current techniques such as SVC, KNN, and DTC. The outcomes demonstrate that the suggested strategy outperforms existing methods.

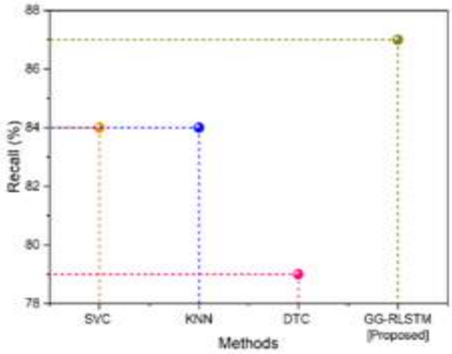

Figure 3. Result of recall

F1-score shows how accurately forecasts match actual flood events, which is important for allocating resources during emergencies and for prompt public health interventions. A comparison of the f1-score is presented in Figure 4. Our suggested GG-RLSTM approach performed (90%), in contrast to (82%), (82%), and (79%) of the current techniques such as SVM, KNN, and RF. The outcomes demonstrate that the suggested strategy outperforms existing methods.

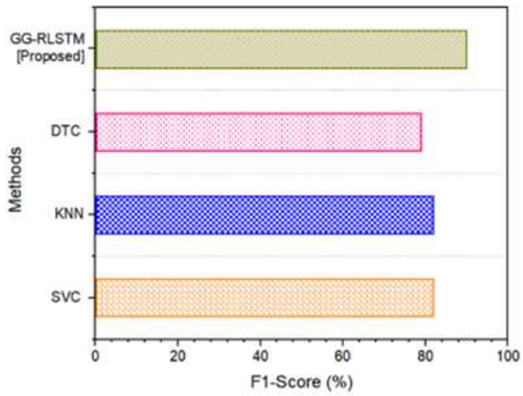

Figure 4. Result of f1-score

3. Conclusion and future scope

This research proves that the GG-RLSTM model is a useful tool for enhancing the ability of public health organizations to offer accurate flood predictions by using data of meteorological conditions, and geography obtained from Kaggle source, based on our evaluation, GG-RLSTM exhibits higher accuracy (89%), precision (88%), recall (87%), and f1-score (90%) than the conventional approaches. The results of this study support the proposed model and posit that public health systems might apply those methods to prevent floods. This would mark a significant enhancement of disaster preparedness and response undertakings. Some of the implementation challenges might also include dependability in different situations concerning floods and expansibility to other areas. Moreover, the extension of the applicability of these tools in other public health systems around the world as well as the development of superior scalability and interpretability might enhance early warning systems and evacuation strategies.

Reference

[1] S. Nevo, E. Morin, A. Gerzi Rosenthal, A. Metzger, C. Barshai, D. Weitzner, D. Voloshin, F. Kratzert, G. Elidan, G. Dror, and G. Begelman, “Flood forecasting with machine learning models in an operational framework”, *Hydrology and Earth System Sciences, 26(15),* pp.4013-4032, 2022. https://doi.org/10.5194/hess-26-4013-2022

[2] S. Puttinaovarat, and P. Horkaew, "|Flood forecasting system based on integrated big and crowdsource data by using machine learning techniques", *IEEE Access, 8*, pp.5885-5905, 2020. https://doi.org/10.1109/ACCESS.2019.2963819

[3] S. Neelima, Manoj Govindaraj, Dr.K. Subramani, Ahmed ALkhayyat, & Dr. Chippy Mohan. (2024). Factors Influencing Data Utilization and Performance of Health Management Information Systems: A Case Study. *Indian Journal of Information Sources and Services, 14*(2), 146–152. https://doi.org/10.51983/ijiss-2024.14.2.21

[4] S. Miau, and W.H. Hung, "River flooding forecasting and anomaly detection based on deep learning", *Ieee Access, 8*, pp.198384-198402, 2020. https://doi.org/10.1109/ACCESS.2020.3034875

[5] F.T. Zahura, J.L. Goodall, J.M. Sadler, Y. Shen, M.M. Morsy, and M. Behl, "Training machine learning surrogate models from a high-fidelity physics-based model: Application for real-time street-scale flood prediction in an urban coastal community", *Water Resources Research, 56(10)*, p.e2019WR027038, 2020. https://doi.org/10.1029/2019WR027038

[6] Alamer, L., & Shadadi, E. (2023). DDoS Attack Detection using Long-short Term Memory with Bacterial Colony Optimization on IoT Environment. *Journal of Internet Services and Information Security*, *13*(1), 44-53.

[7] F.Y. Dtissibe, A.A.A. Ari, C. Titouna, O. Thiare, and A.M. Gueroui, "Flood forecasting based on an artificial neural network scheme", *Natural Hazards, 104,* pp.1211-1237, 2020. https://doi.org/10.1007/s11069-020-04211-5

[8] G. Kan, K. Liang, H. Yu, B. Sun, L. Ding, J. Li, X. He, and C. Shen, "Hybrid machine learning hydrological model for flood forecast purpose", *Open Geosciences, 12(1),* pp.813-820, 2020. https://doi.org/10.1515/geo-2020-0166

[9] Uchida, N., Takahata, K., Shibata, Y., & Shiratori, N. (2012). Never Die Network Based on Cognitive Wireless Network and Satellite System for Large Scale Disaster. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 3*(3), 74-93.

[10] Q. Zhou, S. Teng, Z. Situ, X. Liao, J. Feng, G. Chen, J. Zhang, and Z. Lu, "A deep-learning-technique-based data-driven model for accurate and rapid flood predictions in temporal and spatial dimensions", *Hydrology and Earth System Sciences, 27(9),* pp.1791-1808, 2023. https://doi.org/10.5194/hess-27-1791-2023

[11] M.M.A. Syeed, M. Farzana, I., Namir, I. Ishrar, M.H. Nushra, and T. Rahman, "Flood prediction using machine learning models", *In 2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA),* pp. 1-6, 2022.

[12] Bobir, A.O., Askariy, M., Otabek, Y.Y., Nodir, R.K., Rakhima, A., Zukhra, Z.Y., Sherzod, A.A. (2024). Utilizing Deep Learning and the Internet of Things to Monitor the Health of Aquatic Ecosystems to Conserve Biodiversity. *Natural and Engineering Sciences*, *9*(1), 72-83.

[13] M. Anbarasan, B. Muthu, C.B. Sivaparthipan, R. Sundarasekar, S. Kadry, S. Krishnamoorthy, and A.A. Dasel, "Detection of flood disaster system based on IoT, big data and convolutional deep neural network", *Computer Communications, 150,* pp.150-157, 2020. https://doi.org/10.1016/j.comcom.2019.11.022

[14] Y. Xu, C. Hu, Q. Wu, Z. Li, S. Jian, and Y. Chen, "Application of temporal convolutional network for flood forecasting", *Hydrology Research, 52(6),* pp.1455-1468, 2021. https://doi.org/10.2166/nh.2021.021

[15] D.T. Nguyen, and S.T. Chen, "Real-time probabilistic flood forecasting using multiple machine learning methods", *Water, 12(3),* p.787, 2020. https://doi.org/10.3390/w12030787

[16] Ana, L. (2023). GIS Analysis of the Vulnerability of Flash Floods in the Porečka River Basin (Serbia). *Archives for Technical Sciences*, *1*(28), 57-68.