HBRP PUBLICATION

Recent Innovations in Wireless Network Security
Volume 7 Issue 3 Sep-Dec, 2025
e-ISSN: 3048-8354

Glacier Lake Outburst Floods Prediction Using XGBoost

[1]Prasanth K. Baby*, [2]Aljo Davis P., [3]Ameya Jojo, [4]Annliya Baiju
[1-4] Dept. of Computer Science and Engineering, Christ College of Engineering, Thrissur, India

Corresponding Author
E-Mail Id: *prasanthkbaby@cce.edu.in*

ABSTRACT

The Glacial Lake Outburst Floods (GLOFs) prediction using XGBoost is an advanced early detection platform aimed at reducing the risks associated with GLOFs, which are becoming more frequent due to climate change. Glacial lake outburst floods (GLOFs) occur when a glacial lake overflows its natural boundaries, releasing large amounts of water that can cause significant damage downstream. While global temperatures rise, the area of these lakes is increasing, which, in turn, increases the chances of such events happening and poses harm to human settlements and infrastructure, as well as ecosystems? It does so by monitoring the size of the lakes, water levels, temperatures, and the state of natural dam structures. By combining real-time data with algorithms, the system can forecast potential damage. With this, authorities can take measures in advance and mitigate the damage. In addition, the system provides valuable information for forecasting, easterly infrastructure services, and resource saving. This assists governments and responders in flood control planning and realization of social measures. By continuous monitoring and studying the situation, the GLOF warning system increases disaster preparedness and shifts the focus to protecting people and minimizing flood consequences.

Keywords: *GOLF, XGBoost*

1. INTRODUCTION

Glacial Lake Outburst Floods (GLOFs) represent a significant environmental danger, happening when water is abruptly released from glacial lakes, resulting in devastating floods in areas downstream. The occurrence of these events is increasing as a result of climate change, which speeds up glacier melting and creates unstable lakes. The conventional approaches taken in trying to estimate floods appear very much to be lacking in their accuracy as well as the timely provisions of warnings. This gives credence to the development of better approaches for early detection and risk mitigation.

Attempts made to improve remote sensing, along with the strengthening of algorithms and the advanced data analysis techniques have dramatically improved the management and especially the anticipation of disasters like flooding. By using these technologies, an intelligent system can be developed that would continuously monitor the condition of the glacial lake and detect all of the early signals of an ever-impending outburst of the lake. This clockwork-style monitoring enables the authorities to take proactive measures, and thereby reduce risks to the communities and physical assets.

The GLOF warning system is an automated system which merges satellite images and data analysis to assess the existing conditions of the environment and issue a crisis forecast concerning the possibility of GLOF occurring. It is also capable of monitoring certain parameters such as the area of the lake along with water level of the lake, temperature changes, and the stability of the dams. Employing advanced algorithms to this data enables the system to recognize certain patterns that are indicative of potentially dangerous flooding, and the respective authorities are notified in a timely fashion.

.

The system is systematic which enables it to accurately detect and forecast GLOFs.

1. **Environmental Data Acquisition and Processing:** Study how GLOF events can be triggered through environmental shifts via space, satellite images and historical data analysis
2. **Predictive Modeling:** Extreme Gradient Boosting (XGBoost), Support Vector Machine (SVM), and Random Forest maching learning methods come in to play for the identification of high risk glacial lakes and flooding prediction.
3. **Automated Alert System:** GLOF warning systems aids in mitigation of impending glacial lake outburst floods using real time data monitoring through predictive analytics on social media.

Combining real-time data processing, predictive analytics, and automated alert systems, the GLOF Warning System enhances disaster preparedness, reduces economic losses, and strengthens community resilience in areas prone to glacial lake outburst floods. Integration of advanced AI systems, disaster response teams can be alerted instantly for swift action. This plan enables fewer disasters, stronger community resilience, and lesser economic damage in the target areas. Climate mindful disaster mitigation has never been this easier.

2. MOTIVATION

The rising occurrence of Glacial Lake Outburst Floods (GLOFs) presents a significant environmental and socioeconomic threat due to climate change. As global temperatures increase, glaciers melt more rapidly, forming larger glacial lakes often restrained by unstable natural dams. When these lakes burst, they can cause devastating floods that put nearby towns, buildings, and nature at risk. This problem is serious, but the old ways of keeping an eye on things—using past records and occasional on-site checks—often don’t give timely and correct warnings. These old methods can’t predict things in real-time making it hard for officials to plan ahead for disasters. So, we need better computer methods to improve how well and how we can forecast GLOFs.

To address these problems, the GLOF Warning System uses satellite pictures along with smart computer programs like Extreme Gradient Boosting (XGBoost), Support Vector Machine (SVM), and Random Forest. By always checking lake size, water levels, temperature changes, and how stable the dams are, the system can figure out the risks and send out early warnings. Using smart prediction tools, it gives disaster management teams real-time alerts helping them to act to lower flood risks. Setting up a system that can predict GLOFs on a large scale , and based on data is key to getting better at preparing for disasters cutting down on money lost, and making communities in areas at risk from glacier dangers more resilient. This system combines machine learning to assess risks and real-time tracking of the environment. It offers a way to get ahead of disasters and adapt as needed, which helps protect people and important buildings in danger zones more

3. LITERATURE SURVEY

Climate change has caused more Glacial Lake Outburst Floods (GLOFs) leading to a lot of research to watch, forecast, and lessen their effects. Many studies have looked into how climate change plays a role using satellites to keep an eye on things applying remote sensing tech, and putting machine learning to work to make GLOF warning systems better.

The study by Immerzeel et al. (2010) outlines the manner in which the rising temperature influences the glacial lakes, particularly in the Himalayas, thereby leading to a higher possibility of GLOFs with an increase in melting of glaciers. [1]. Similarly, Johns-Putra (2016) discusses broader social and environmental effects of climate change, such as higher likelihood of natural hazards like GLOFs [2]. Meanwhile, Anderegg and Goldsmith (2014) analyze public awareness and policy changes regarding climate change, highlighting their indirect impact on disaster management and preparedness [3]. Rankl et al. (2014) use remote sensing techniques in glacier fluctuation and glacial lake expansion analysis and obtain vital information for the assessment of GLOF hazard [4]. Yde and Knudsen (2007) also give a historical overview of glacier fluctuations and their contribution to triggering GLOFs [5]. Chen et al. (2021) apply SAR observations to track inter-seasonal changes of glacial lakes, which can be useful inputs to flood hazard mapping [6]. Likewise, Li et al. (2024) introduce a new satellite-based method for detection and classification of dangerous glacial lakes in mountainous areas [7].

To have better forecasting of GLOF, Zhu et al. (2023) propose an AI approach to observing temporal alterations of glaciers for upgrading early warning systems [8]. Patel et al. (2024) suggest an automated approach for detecting land cover changes, aiding in the identification of GLOF hazards. [9].

Along with improving detection methods, Ahmed et al. (2023) integrate machine learning with historical data to boost the accuracy of glacial lake monitoring and risk evaluation. [10]. Wang et al. (2022) explore changes in glacial lake water storage and provide insights into their stability and likelihood of outbursts. [11].

Luo et al. (2021) discuss the effect of climate change on glacial lakes and how they form with time [12]. Johansson and Brown (2013) use machine learning-based classification methods to examine and classify glacial lakes in terms of their flood risk [13].

In addition, Zhu et al. (2024) also suggest a multi-year glacial lake analysis technique, which presents a long-term perspective of their origin and accompanying flood risk [14].Lastly, Rastner et al. (2014) contrast different image classification techniques for accurate identification and tracking of glacial lakes [15].

Together, these studies improve GLOF early warning systems by combining remote sensing, satellite data, and machine learning models to improve risk assessment, detection, and mitigation. The results of these studies form a solid basis for the formulation of more efficient disaster preparedness and response mechanisms in high-risk glacial areas.

4. PROPOSED MODEL

This project aims to create an advanced Glacial Lake Outburst Flood (GLOF) Warning System by utilizing satellite imagery, machine learning, and predictive modeling to detect potential flood risks in glacial areas. The system consistently tracks key environmental factors like lake expansion, water levels, temperature, and ground stability to deliver early warnings and mitigate disasters.

Data Collection and Feature Selection

The model processes historical records of GLOF using remote sensing and satellite geospatial data. The training parameters were based on selected features, including location related attributes such as Lat-Lake, Lon-Lake, and Elev-Lake, event properties like Driver-Lake, Driver-GLOF, and Mechanism, impact-related factors including Impact Type and Repeat, and geographical information such as Region-RGI and Region-HiMAP. These features were used to train the model, ensuring it captures the key factors influencing GLOF events. In light of the dataset structure of the provided image, the most prominent features selected for model training are as follows:

Selected Characteristics the GLOF Prediction System takes into account multiple factors to enhance the accuracy of flood prediction. Location-based features like latitude, longitude, and glacial lake elevation determine its geographical position and outburst energy potential. Event-based features like the primary driver, specific triggering mechanism, and failure mode describe how the past GLOFs have occurred. Impact based features examine the type and frequency of past floods to determine the high-risk areas. Geographical features, such as regional classifications from the Randolph Glacier Inventory (RGI) and High Mountain Asia database (HiMAP), help categorize influencing factors. The system uses Impact as the target variable to train machine learning models. To ensure accurate predictions, the data undergo preprocessing steps like handling missing values, normalization, and feature engineering before applying XGBoost, SVM, and Random Forest algorithms for analysis. One of the major challenges we faced was acquiring a reliable and comprehensive dataset for training the GLOF prediction model. Due to the limited availability of historical data on glacial lake outburst floods, we had to rely on the data we could obtain through remote sensing and satellite imagery. Consequently, the model's accuracy and predictions are inherently tied to the quality and quantity of this data. Any gaps or inaccuracies in the dataset may affect the model's performance.

The synthetic dataset used in this study was developed by integrating three primary datasets to enhance the analysis of glacial changes and associated hazards in High Mountain Asia. The first source is the GLOF database of High Mountain Asia, which provides extensive data on glacial lake outburst floods across the region. The second source is the dataset compiled by Mo et al. (2024), which examines regional disparities in glacier dynamics and their responses to climate change. The third source is the Dataset of Glacier Changes in the Northern Hemisphere, created by Ren et al. (2022), which reconstructs glacier fluctuations over the past 2000 years using a combination of temperature, precipitation, and ice core data. By synthesizing these datasets, the study constructs a more comprehensive dataset that facilitates improved modeling and understanding of glacier-related processes in High Mountain Asia.

Machine Learning for GLOF Prediction

To enhance accuracy in flood forecasting, the system utilizes machine learning models trained on past GLOF events and environmental data. The process starts with feature extraction, identifying key risk factors like changes in lake size, rainfall patterns, and terrain stability. Three models — XGBoost, Support Vector Machine (SVM), and Random Forest — are trained on historical data to predict GLOF probabilities. After evaluation, XGBoost was selected as the final model because of its excellent predictive performance and computational efficiency.

A major advantage is its ability to manage imbalanced datasets, which is essential given the infrequency of GLOF events. In contrast to SVM, which needs extensive parameter tuning and has difficulty with large datasets, XGBoost can effectively handle large-scale environmental data while still achieving high accuracy. Moreover, XGBoost uses regularization techniques to reduce overfitting, leading to more stable and generalizable predictions. Another important benefit is its ability to rank feature importance, offering insights into the key factors that influence GLOF occurrences, such as lake expansion, temperature changes, and terrain stability.

Model	Accuracy	Macro Avg F1-Score	Weighted Avg F1-Score
SVM	0.949	0.95	0.95
Random Forest	0.949	0.95	0.95
XGBoost	0.987	0.99	0.99

Fig. 1: *Comparative Performance of Machine Learning Models for GLOF Prediction.*

XGBoost is highly efficient in processing large-scale environmental data,to its ability to perform parallel computations. It also automatically manages missing values and offers extensive hyperparameter tuning options, allowing it to adapt well across different datasets. Unlike Random Forest, which constructs trees separately, XGBoost refines its predictions step by step, continuously learning and correcting errors for better accuracy. Additionally, its strong resistance to data noise and capability to handle imbalanced datasets make it a more reliable choice compared to SVM and Random Forest. The trained model then classifies risk, classifying lakes as low, moderate, or high risk based on the probability of an outburst. Through the application of machine learning techniques, the system maximizes its predictions in real time, allowing for more accurate early warning.

Automated Warning System

When the system detects a potential GLOF event, an automated alert system is triggered to inform relevant agencies and communities. These include:

Real-Time Notifications: Sending email, SMS, and siren alerts to high-risk groups.

Interactive Dashboard: Presenting hazard maps and risk zones on a GIS platform for informed decision-making.

Emergency Coordination Response: Alerting disaster response teams to enable swift action and mitigation.

By integrating real-time monitoring, predictive analysis, and automated alerts, the system supports timely intervention, reducing infrastructure damage, economic loss, and the risk to human lives.

5. IMPLEMENTATION DETAILS

The GLOF Forecasting System that integrates and operationalizes historical records of GLOF events, satellite images, and the recent hydrological data has been developed and operationalized for the system in use. The dataset contains historical information concerning the number of outbursts of a glacier, glacial retreat rates, and temporal climatic data available from the monitoring regions of the world. The deployment processentails data analysis and machine learning algorithms formulation, along with risk visualization for improved prediction and timely alarms.

Feature Processing and Engineering

To ensure the model predictions, features are consistent and accurate; the data are subjected to multiple pre-processing steps. Satellite images of glaciers are taken with standardized settings; therefore, all images have the same resolution and their location on the globe is consistent. Satellite images containing the glacial lakes are segmented, and the resulting glacial lake outlines are

Recent Innovations in Wireless Network Security
Volume 7 Issue 3 Sep-Dec, 2025
e-ISSN: 3048-8354

used to monitor the growth of the lake over time. Climatic variables are adjusted and transformed to account for changes in temperature, precipitation, and other environmental factors in normalization, which would increase the performance of models using such variables.

Machine Learning Deployment

The system uses Pythonbased libraries like Scikit-learn, OpenCV, and XGBoost to train models and manage real-time data processing. OpenCV is employed to process satellite images to acquire lake boundaries and track lake area changes over time.

Three machine learning models, namely XGBoost, SVM, and Random Forest, are employed for prediction modeling to train on historical GLOF data and classify lakes into different risk stages. The last model is selected to be XGBoost on the basis of model performance since it can deal with imbalanced data and improve prediction accuracy. Risk Visualization and Decision Support to help the decision-makers determine flood-risk zones, the system employs an interactive GIS-based dashboard for the plotting of risky zones. The officials can monitor the glacial lake behavior, detect potential risk, and prepare ahead in the event of disasters through the application. The risk categorization system classifies the lakes as low, medium, or high in terms of risk and prioritizes intervention in the highest-risk zones.

Early Warning and Risk Mitigation

When a GLOF event is detected, the system automatically triggers an early warning system to alert relevant authorities and at-risk communities. Real-time notifications are sent via SMS, email, and sirens to prompt immediate action. The system also mobilizes disaster response teams to carry out emergency mitigation strategies, reducing risks to human settlements and infrastructure. By combining remote sensing, machine learning models, and automated alert systems, the GLOF Warning System offers a data-driven, scalable solution for early hazard detection, disaster response planning, and community protection. With real-time environmental monitoring and predictive analysis, the system enhances disaster resilience, lowers economic impacts, and ensures safety in vulnerable glacial regions.

6. CONCLUSION

The proposed Glacial Lake Outburst Flood (GLOF) Warning System introduces the possibility of early effect detection and risk mitigation by combining satellite imagery, machine learning models, and real-time environmental monitoring. The system is working continuously to monitor the most critical factors affecting the lakes, including lake growth, variations in water levels, variations in temperature, and stability of the terrain so that it can provide better predictions.

To achieve a high level of accuracy in predictions, the system initially trained three different models: XGBoost, support vector machine (SVM), and random forest. Due to its incapability to effectively handle imbalanced datasets, XGBoost was selected as the final model for this system due to its successful improvement of prediction accuracy. The trained model has the capability to classify lakes into low, moderate, and high-risk zones, building up a state of readiness for the possible outburst floods.

This system incorporates the automated early warning system that sends out real-time alerts through SMS, emails, and sirens to disaster management authorities and at-risk communities. By ingesting historical data alongside real-time environmental monitoring, the GLOF nasalities answer the unpredictable behavior of dynamics in glacial lakes and

Recent Innovations in Wireless Network Security
Volume 7 Issue 3 Sep-Dec, 2025
e-ISSN: 3048-8354

the stability of moraine dams and hence enhance disaster preparedness.
GLOF risk management presents one such challenge related to the indeterminate nature of the evolution of the glacial lakes, which in turn makes early detection the most important. This system uses a data-driven and machine-learning-based approach to risk assessment to counter such challenges and assure precise forecasting and proactive action against possible disaster eminence.

The system will help reduce the loss of life, damage to vital infrastructure, and consequent economic losses that impact GLOFs by working alongside governmental agencies, environmental researchers, and emergency responders. The development and implementation will ensure enhanced climate revisioning, better water management, and improved disaster preparedness in fourth-gone glacial areas.

Objectives

• Develop an early warning system for predicting GLOFs

• Improve flood prediction accuracy through machine learning techniques

7. ACKNOWLEDGMENT

We want to thank everyone who gave us the opportunity to finish this in the sincerest way possible. We would also want to express our gratitude to Mr. Prasanth K Baby, our project coordinator for the Final year, whose contribution of energizing suggestions and support helped us to plan our project, particularly in drafting this paper. Additionally, we would like to express our gratitude to the department head, Dr. Vince Paul, who invested her full effort in guiding the team in achieving the goal. We must be grateful for the advice provided by other departmental faculty members, especially in regard to our paper presentation, which has enhanced our presentation skills as a result of their comments and suggestions.

REFERENCES

1. Immerzeel, W. W., van Beek, L. P. H., & Bierkens, M. F. P. (2010). Climate change will affect the Asian water towers. *Science*. https://doi.org/10.1126/science.1183188
2. Johns-Putra, A. (2016). Climate change in literature and literary studies: From cli-fi, climate change theater and ecopoetry to ecocriticism and climate change criticism. *Wiley Interdisciplinary Reviews: Climate Change*. https://doi.org/10.1002/wcc.385
3. Anderegg, W. R. L., & Goldsmith, G. R. (2014). Public interest in climate change over the past decade and the effects of the 'climategate' media event. *Environmental Research Letters*, *9*(5), 054005. https://doi.org/10.1088/1748-9326/9/5/054005
4. Rankl, M., Kienholz, C., & Braun, M. (2014). Glacier changes in the Karakoram region mapped by multimission satellite imagery. *The Cryosphere*, *8*, 977–989. https://doi.org/10.5194/tc-8-977-2014
5. Yde, J. C., & Knudsen, N. T. (2007). 20th-century glacier fluctuations on Disko Island (Qeqertarsuaq), Greenland. *Annals of Glaciology*. https://doi.org/10.3189/172756407782871422
6. Chen, L., Tan, T., Zhang, L., & Huang, H. (2021). Seasonal changes of glacier lakes in Tibetan Plateau revealed by multipolarization SAR data. *IEEE Transactions on Geoscience and Remote Sensing*. https://doi.org/10.1109/TGRS.2021.9506350
7. Li, Y., Zhang, C., Zhang, K., & Liu, W. (2024). A glacial lake mapping framework in high mountain areas: A case study of the southeastern Tibetan Plateau. *IEEE Transactions on*

Geoscience and Remote Sensing. https://doi.org/10.1109/TGRS.2024.10036278

8. Zhu, X., Wang, M., Li, J., & Wang, T. (2023). A progressive learning strategy for large-scale glacier mapping. *IEEE Transactions on Geoscience and Remote Sensing*. https://doi.org/10.1109/TGRS.2023.9985678
9. Patel, R., Bhattacharya, S., & Kumar, V. (2024). Application of inverse mapping for automated determination of normalized indices useful for land surface classification. *IEEE Transactions on Geoscience and Remote Sensing*. https://doi.org/10.1109/TGRS.2024.10073902
10. Ahmed, S., Singh, B., & Verma, M. K. (2023). Mapping glacial lakes using historically guided segmentation models. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. https://doi.org/10.1109/JSTARS.2023.9924600
11. Wang, H., Huang, Z., Wen, Z., & Luo, Z. (2022). Lake water storage changes in Northeast Hoh Xil observed by Cryosat-2 and Landsat-5/7/8: Impact of the outburst of Zhuonai Lake in 2011. *IEEE Geoscience and Remote Sensing Letters*. https://doi.org/10.1109/LGRS.2022.3164735
12. Luo, R., Yuan, Q., Yue, L., & Shi, X. (2021). Monitoring recent lake variations under climate change around the Altai Mountains using multimission satellite data. *IEEE Transactions on Geoscience and Remote Sensing*. https://doi.org/10.1109/TGRS.2021.9492724
13. Johansson, A. M., & Brown, I. A. (2013). Adaptive classification of supraglacial lakes on the West Greenland Ice Sheet. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. https://doi.org/10.1109/JSTARS.2012.2233722
14. Zhu, T., Cui, X., Zhang, Y., Lu, K., & Yang, Y. (2024). A conditional random fields-based identification for small lakes over multiple melt seasons using Sentinel-2 imagery in the Larsemann Hills, East Antarctica. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. https://doi.org/10.1109/JSTARS.2024.3391881
15. Rastner, P., Bolch, T., Notarnicola, C., & Paul, F. (2014). A comparison of pixel- and object-based glacier classification with optical satellite images. *IEEE Transactions on Geoscience and Remote Sensing*. https://doi.org/10.1109/TGRS.2014.6917362