Efficient Flood Detection through Hybrid Machine Learning and Metaheuristic Methods using Sentinel-1

Behnam Ebadati [1], Mohammad Alikhani [2], Fahimeh Youssefi [3], Saied Pirasteh[3]

[1] Dept. of Geomatic Engineering, South Tehran Branch, Islamic Azad University, Tehran, Iran - st_b_ebadati@azad.ac.ir
[2] Geodesy and Geomatics Engineering Faculty, K. N. Toosi University of Technology, Tehran, Iran - m.alikhani1@email.kntu.ac.ir
[3] Institute of Artificial Intelligence, Shaoxing University, 508 West Huancheng Road, Yuecheng District, Shaoxing, Zhejiang Province, Postal Code 312000, China - youssefi@usx.edu.cn, sapirasteh1@usx.edu.cn

Keywords: Flood Detection, Machine Learning, VGG-16, Hyperparameter Optimization, Metaheuristic Algorithms, Sentinel-1.

Abstract

Floods are considered among the most destructive natural disasters, requiring precise and timely management. Remote sensing, utilizing diverse satellite imagery data, enables effective monitoring and assessment of flood impacts. In this context, machine learning and deep learning methods, as effective and scalable approaches, can significantly enhance the accuracy of flood detection and management by analyzing remote sensing data, thereby playing a crucial role in mitigating flood-related risks. In this study, to flood detection using Sentinel-1 SAR data, machine learning algorithms, including Random Forest (RF) and Histogram-based Gradient Boosting Decision Tree, were employed, along with two metaheuristic algorithms, Harris Hawks Optimization (HHO) and Ant Colony Optimization (ACO), for hyperparameter optimization. Additionally, to enhance the models' ability to detect flooded pixels and improve overall performance accurately, a pre-trained VGG-16 Neural Network was used as a deep feature extractor. Finally, four ensemble flood detection models—RF-HHO, RF-ACO, HGBDT-HHO, and HGBDT-ACO—were implemented, and their performance was evaluated and compared based on statistical metrics. Based on the obtained results, all four ensemble flood detection models demonstrated excellent performance in the validation and testing phases. The overall accuracy of these models reached over 95% in the validation phase and exceeded 97% in the testing phase. However, the HGBDT-ACO model achieved the highest accuracy and the lowest error rate in detecting flood pixels, making it the best-performing model in this study. Generally, HGBDT models showed a relative advantage over RF models, as they required significantly less time for training while achieving comparable results. Therefore, they were efficient and performed better in terms of computational complexity.

1. Introduction

Recently, the frequency of natural disasters worldwide has significantly increased (Munawar et al., 2022). Among these, floods stand out as one of the most prevalent water-related calamities, directly or indirectly affecting approximately 23% of the global population, equivalent to 1.8 billion individuals (Amitrano et al., 2024). The United Nations Office for Disaster Risk Reduction (UNDRR) reports that the occurrence of flood events has seen a remarkable rise across the globe over the past two decades (Sadiq et al., 2023). Climate change, intense rainfall, snowmelt, glacier retreat, and dam breaches are among the primary triggers for flooding (Jeyaseelan, 2004). Additionally, rapid urbanization and increased human activities pose an even more significant threat of floods to human communities (Alidoost and Arefi, 2018). Floods can severely damage agricultural lands, residential areas, and transportation infrastructures such as roads and railways (Lamovec et al., 2013).

Despite the seemingly insurmountable challenge of preventing floods, effective management can reduce the risks and mitigate the damages they cause (Rahman and Di, 2017). Access to essential flood-related information is crucial for achieving such a goal. Satellite imagery, particularly through remote sensing (RS) methods, is essential for gathering and providing accurate data, for flood mapping and monitoring. (Amitrano et al., 2018; Giustarini et al., 2016; Klemas, 2014; Martinis et al., 2015). The continuous advancements in optical and radar satellite sensor technologies in recent years have made remote sensing a more cost-effective and rapid method compared to traditional approaches for flood detection or monitoring (Islam et al., 2020). Radar sensors such as ALOS PALSAR, TerraSAR-X, RADARSAT-1, RADARSAT-2, ENVISAT, and Sentinel-1, along with optical sensors like Landsat and Sentinel-2, have become prominent in this regard (Anusha and Bharathi, 2020). Data from Sentinel missions, such as Sentinel-1 and Sentinel-2, are accessible at no cost to the public through the European Space Agency (ESA), and the United States Geological Survey (USGS) provides Landsat data (Nyamekye et al., 2021; Solórzano et al., 2021). Flood detection and Flood mapping using optical imagery presents challenges due to its reliance on sunlight for imaging and high sensitivity to adverse weather conditions. Conversely, radar sensors, with their ability to acquire Synthetic Aperture Radar (SAR) images day and night and in adverse weather conditions, along with dense cloud cover, are recognized as more effective tools for flood detection (Cao et al., 2019). Nevertheless, multi-temporal optical imagery can sometimes mitigate the impact of clouds, thereby enhancing the performance of optical sensors (Priyatna et al., 2023; Uddin et al., 2019).

Flood detection can be achieved through various methods utilizing remote sensing data. For example, floods and other surface water can be detected through multiple spectral indices (Albertini et al., 2022) and automatic thresholding techniques (Moharrami et al., 2021; Tran et al., 2022) such as Otsu (Otsu, 1979). However, these methods are often time-consuming and require specialized knowledge and skills. Moreover, they are typically tailored to specific case studies and may not apply to different scenarios with varying scales and conditions (Tanim et al., 2022). Machine learning-based methods offer a promising alternative to traditional flood detection methods using remote sensing data. Various machine learning algorithms, including Random Forests (RF), Convolutional Neural Networks (CNN), and Support Vector Machines (SVM), are employed for flood detection and other surface water extraction tasks (Vongkusolkit et al., 2023). A significant obstacle in deep learning is the necessity of training models on vast amounts of data. This requirement not only extends the duration of the training

process but also amplifies the computational resources needed, making the process more demanding and time-consuming. Several approaches address these limitations; one is using transfer learning techniques, such as employing pre-trained models, which can significantly reduce the time and computational resources required for training models (Tulasi et al., 2019).

The following briefly reviews several studies on flood detection and surface water identification using various remote sensing techniques. In a study, four methods were employed for flood detection in urban areas: automatic thresholding using the OTSU algorithm, change detection, supervised classification using the Random Forest algorithm, and unsupervised classification using the K-means algorithm. Various Sentinel-1 polarization combinations were used as inputs for all four methods. Additionally, to enhance the accuracy of the results, the zero-depth method was applied with the SRTM DEM. The findings indicated that all four methods provided satisfactory accuracy and performance. (Tazmul Islam and Meng, 2022). In another study, Sentinel-1 mission data from 5,296 tiles were used to train several deep learning-based models for flood detection in the Yangtze River Basin. The results of this study demonstrated that neural network models exhibited significantly higher accuracy and efficiency than thresholding methods. Additionally, the impact of VV and VH polarizations was evaluated, revealing that VH polarization was more effective than VV polarization in flood detection (Wu et al., 2023). In a study, neural network algorithms were employed for permanent and temporary water mapping during flood events. This study utilized fusion techniques to leverage the advantages of Sentinel-2 multispectral and Sentinel-1 SAR images for water detection and differentiation from non-water areas. According to the results, the models performed well, achieving an overall accuracy of over 92%, and demonstrated significant generalizability (Bai et al., 2021). In a study, Sentinel-1 and Sentinel-2 missions data were used simultaneously to monitor floods in northeastern Bangladesh and assess the resulting damages, aiming to maximize the advantages of these datasets. Two classification methods, Random Forest and Maximum Likelihood, were applied to flood detection and evaluate the associated damages. The results indicated that the Random Forest algorithm performed better in flood detection, achieving a classification accuracy of 90%, while the Maximum Likelihood method reported a classification accuracy of 74% (Billah et al., 2023). In a study, the capabilities and performance of four pre-trained models, including ResNet18, ResNet50, EfficientNet, and VGG16, were evaluated for flood detection using Sentinel-2 MSI data. The results of this study indicated that pre-trained models with different convolutional neural network architectures can exhibit highly significant performance in flood detection using remote sensing data and be effective in this regard (Jain et al., 2020). In another study, the VGG-16 model was utilized alongside two machine learning models, Random Forest, and Gradient Boosting, to develop several advanced flood detection models using Sentinel-1 data. The results of this study demonstrated that the models performed very well, with the VGG-16 model playing a significant role in enhancing the accuracy and performance of the models in detecting flood-affected pixels by extracting deep features (Ebadati et al., 2023).

Given the increasing importance of floods and the need for more effective management of this natural disaster, remote sensing data, particularly SAR images, has significantly improved in recent years with advanced machine learning and deep learning methods. The primary goal of this study is to design and develop advanced models for flood detection using Sentinel-1 data. Additionally, this research provides a comprehensive comparison of these models in terms of accuracy, performance, and computational efficiency during the training process, aiming to offer optimized solutions for flood detection with minimal human intervention. To accomplish this objective, a hybrid approach has been adopted, integrating deep learning methods, machine learning techniques, and metaheuristic algorithms to develop ensemble flood detection models.

2. Methodology and Materials

The first step toward achieving this goal involves feature extraction employing the VGG-16 model through a transfer learning approach. In the next step, Random Forest (RF) and Histogram-based Gradient Boosting Decision Tree (HGBDT) algorithms are used to train the classification models. Two metaheuristic algorithms, Harris Hawks Optimization (HHO) and Ant Colony Optimization (ACO), are utilized for hyperparameter optimization. Therefore, ensemble models are implemented using two machine learning-based classifiers and two metaheuristic algorithms, referred to in this study as RF-HHO, RF-ACO, HGBDT-HHO, and HGBDT-ACO. Figure 1 presents a flowchart of the steps taken in this study.

2.1 Dataset

This study used the "Sen1Floods11" dataset to train and develop flood detection models. This dataset consists of 4,831 raw Sentinel-1 Imagery data, along with classified images of floods and permanent water bodies used as ground truth labels. The data in this dataset are all provided as georeferenced image patches in 512×512 pixels size. The Sen1Floods11 covers 2018-2019 flood events across six continents, encompassing an area of approximately 12,400 km² in various geographical regions (Bonafilia et al., 2020). Table 1 shows an overview of Sentinel-1 SAR data acquired during flood events in multiple locations worldwide. These data are available in the Sen1Floods11 as 512x512 image patches in Geotiff file format. This dataset was generated by "Cloud to Street," a Public Benefit Corporation with access provided via their GitHub repository at https://github.com/cloudtostreet/Sen1Floods11.

Region	Acq. date	Orbit	Polarization
Bolivia	2018-02-15	Descending	VV/VH
Paraguay	2018-10-31	Ascending	VV/VH
Ghana	2018-09-18	Ascending	VV/VH
Somalia	2018-05-07	Ascending	VV/VH
Nigeria	2018-09-21	Ascending	VV/VH
Pakistan	2017-06-28	Descending	VV/VH
Cambodia	2018-08-05	Ascending	VV/VH
India	2016-08-12	Descending	VV/VH
Sri Lanka	2017-05-30	Descending	VV/VH
Spain	2019-09-17	Descending	VV/VH
USA	2019-05-22	Ascending	VV/VH

Table 1. Flood events data overview.

2.2 Data Preparation

Data cleaning is a crucial step to perform before model training. In this study, the data underwent thorough and repeated reviews.

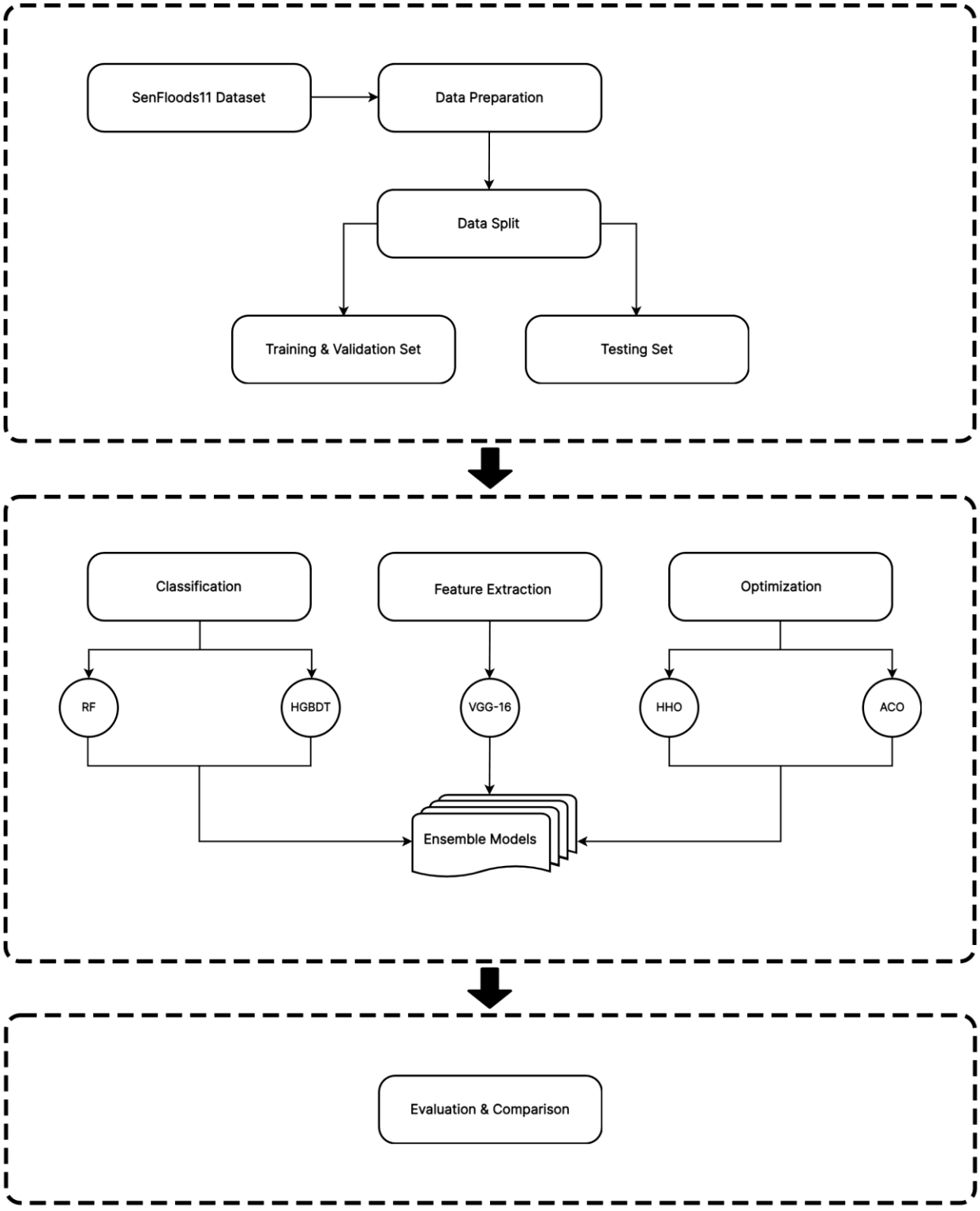

Figure 1. Methodology flowchart.

This process prevented the inclusion of corrupted, low-quality data or unreliable ground truth labels, which could cause errors in model training. Figure 2 shows examples of data removed from the dataset because it contains null pixels.

In binary classifications, such as flood detection by classifying flooded and non-flooded pixels, the presence of these null pixels can introduce errors in the classification results or decrease the accuracy.

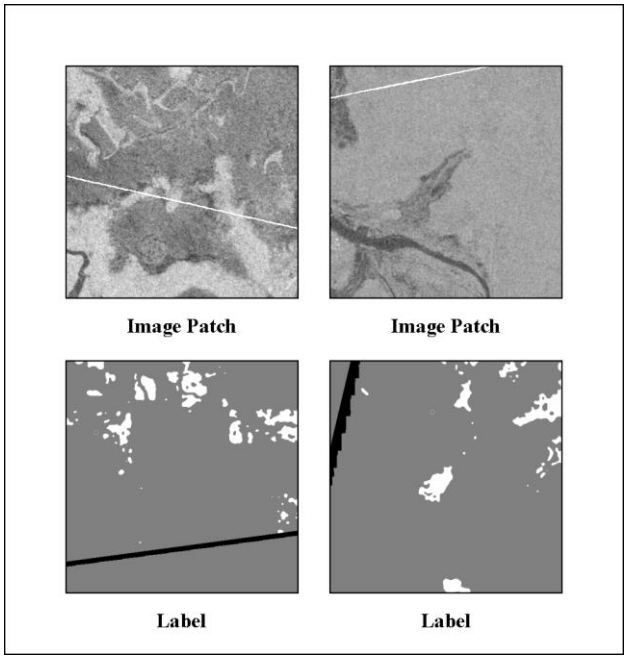

Figure 2. Examples of corrupted data and labels (contain null pixels).

Due to hardware limitations and to avoid excessive costs and time, not all cleaned data from Dataset Sen1Floods11 were used. Instead, the best samples were selected as the training, validation, and testing sets. This approach aimed to overcome limitations while ensuring high-quality model training. After thorough final reviews and verification of data quality and ground truth labels, 287 image patches with their corresponding labels were chosen for the training and validation set. Additionally, 60 new image patches were chosen for the testing set to ensure the models' effectiveness with entirely new and unseen data. Table 2 shows the selected image patches overview and final for each count geographical region for the models' training, validation, and testing.

Region	Training and Validation Set	Testing Set
Bolivia	18	6
Colombia	20	0
Paraguay	34	4
Ghana	19	5
Somalia	14	0
Nigeria	27	3
Pakistan	11	4
Cambodia	34	10
India	27	6
Sri Lanka	24	6
Spain	27	11
USA	32	5
SUM	287	60

Table 2. Geographic distribution of selected data.

2.3 Feature Extraction Using VGG-16

In this study, A transfer learning approach utilizing the VGG-16 deep neural network model was employed as an auxiliary tool for feature extraction. VGG-16, introduced in 2015, is a pre-trained Convolutional Neural Network (CNN) model renowned for its effectiveness in various image-related tasks such as machine vision. (Simonyan and Zisserman, 2015). VGG-16 has been trained based on the ImageNet database by the Visual Geometry Group (VGG). (Deng et al., 2009; Theckedath and Sedamkar, 2020). The ImageNet database contains millions of images belonging to various classes. The VGG16 network architecture employs 3x3 convolutional layers stacked for depth, initially processing 224x224 RGB images. It comprises 13 convolutional layers with subsequent 2x2 max-pooling layers. The network ends with three fully connected layers of varying depths and configurations. (Tune et al., 2021). Due to its extensive training regimen, VGG-16 showcases exceptional accuracy, even when confronted with limited image datasets (Theckedath and Sedamkar, 2020).

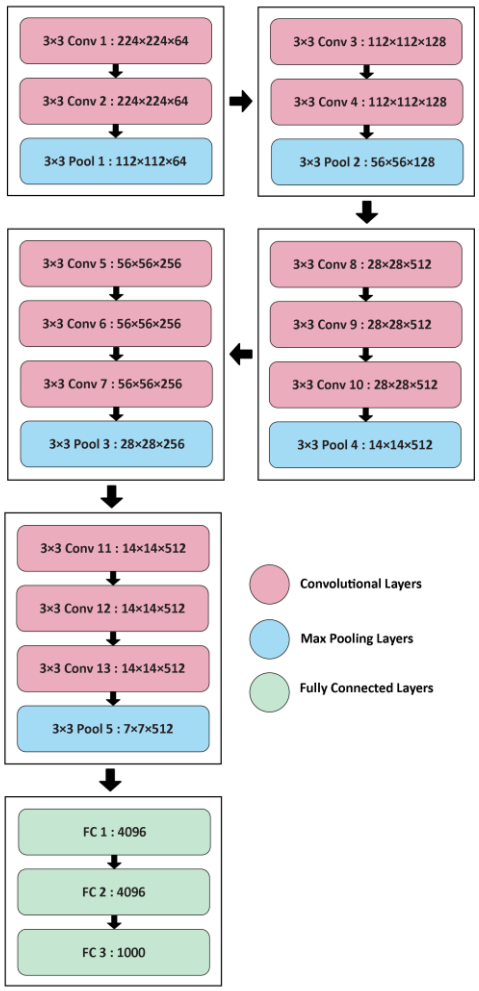

Figure 3. VGG-16 network architecture

2.4 Machine Learning Classifiers

For training flood detection models using features extracted by the VGG-16 model, classifiers are needed to achieve the highest accuracy rate and lowest error, overcome challenges such as overfitting, and exhibit acceptable computational performance. This issue is not limited to flood detection but is also crucial for all classification tasks in other studies. This study employed Random Forest (RF) (Breiman, 2001) and Histogram-based Gradient Boosting Decision Tree (HGBDT) (Guryanov, 2019) that was developed to enhance the Gradient Boosting Decision Tree (GBDT) (Friedman, 2002) to train classifier models based

on features extracted from 287 images using the VGG-16 model. 70% of the training and validation set was used exclusively for model training, while the remaining 30% was allocated for validation.

2.5 Hyperparameter Optimization

Hyperparameters refer to parameters used either for configuring a machine learning model, such as the C parameter and kernel type for SVM models, or the learning rate for CNN models (Diaz et al., 2017). Hyperparameter optimization is crucial for improving machine learning classification models' performance and accuracy rate. Recently, this task has been pursued more vigorously by employing metaheuristic algorithms (Morales-Hernández et al., 2023; Yang and Shami, 2020; Elshawi et al., 2019) In this study, the Harris Hawks Optimization (HHO) and Ant Colony Optimization (ACO) algorithms were employed to optimize and tune the hyperparameters of the RF and HGBDT classifiers. The HHO (Heidari et al., 2019) and ACO (Dorigo et al., 1996) algorithms, representing the new and traditional generations of population-based metaheuristic algorithms, have been widely used by researchers for optimization tasks in various studies. These algorithms repeatedly search for the best hyperparameters to train models with the highest possible performance. These algorithms in this study aim to identify the optimal configurations that enhance classifier models performance. To ensure identical conditions and create a fair benchmark for comparing the final performance of the models, a limited and equal search space was defined to find the best hyperparameters for each model. In other words, for both algorithms, similar parameter values common to each algorithm, such as the number of Hawks, Ants, and iterations, were set to be limited and uniform. This was done to ensure equal conditions for hyperparameter search while also preventing significant time loss due to extended iterations.

2.6 Models Evaluation Methods

In this study, common statistical metrics were employed to evaluate and compare the performance of the ensemble flood detection model. The statistical metrics used include overall accuracy, Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). Additionally, the kappa coefficient and Receiver Operating Characteristic (ROC) curve were utilized to provide a more detailed assessment of the models' performance. The Kappa coefficient, or Cohen's Kappa, measures the reliability of categorical data by assessing agreement between two raters or classifications beyond chance. It ranges from -1 (disagreement) to 1 (strong agreement) and 0, indicating no agreement beyond chance (Cohen, 1960), while The ROC curve offers a visual representation of the trade-off between the true positive rate (TPR) and the false positive rate (FPR) for both training and test datasets (Termeh et al., 2018). By analyzing this curve and calculating the Area Under the Curve (AUC), the model's overall performance can be effectively assessed, reflecting the balance between the true positive and false positive rates. (Luu et al., 2021).

3. Results

3.1 Extracted Features

To prevent computational complexity and reduce the risk of overfitting, only the first two layers out of the 16 layers of the VGG-16 network were selected to search for and extract deep pixel features in 64 channels. Figure 4 shows a sample of the data in these 64 channels. In total, 14,714,688 parameters are involved across all layers of VGG-16, which are reduced to 38,720 parameters by selecting only the first two layers. With the reduction in parameters, the final number of extracted features was significantly decreased, resulting in models being trained with fewer and more manageable features.

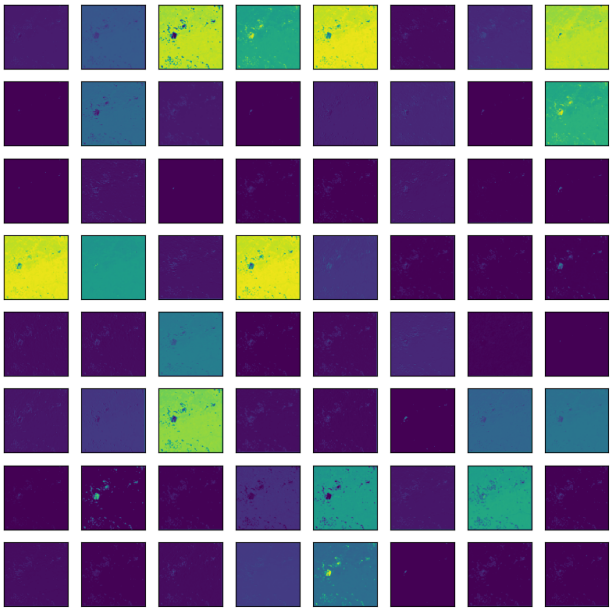

Figure 4. Sample data representation in 64 distinct channels.

3.2 Models Evaluation Results

In the first stage of comparing trained models, the computational speed, or in other words, the required time for training these models, is examined. Under the identical conditions set for all four models in this study, HGBDT models demonstrated a remarkable advantage in terms of speed and efficiency over RF models.

Model	Hyperparameter		Time Spent (Sec)
RF-HHO	n of trees	50	38,449.6567
	max depth	7	
RF-ACO	n of trees	67	36,528.4194
	max depth	6	
HGBT-HHO	max iteration	50	842.733
	max depth	8	
	learning rate	0.038	
HGBDT-ACO	max iteration	56	843.375
	max depth	6	
	learning rate	0.037	

Table 3. Best hyperparameters and training duration.

The HGBDT-HHO and HGBDT-ACO Models proved significantly more efficient in training speed than RF-HHO and RF-ACO and had substantially lower computational complexity; therefore, they required nearly 40 times less training time while maintaining overall accuracy levels close to those of RF models. Table 3 shows the best hyperparameters and the time spent for trained models. For further emphasis, it

should be noted that the time spent for each model, as presented in Table 3, is the total time required for training the models and the iterations needed for hyperparameter optimization using the HHO and ACO. The preliminary evaluation of the ensemble flood detection models was conducted using 30% of the training and validation dataset to provide an overview of these models' predictive capabilities and classification accuracy. Table 4 presents the values for Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and overall accuracy in this evaluation phase. The RF-HHO, RF-ACO, HGBDT-HHO, and HGBDT-ACO models recorded RMSE values of 0.1968, 0.2155, 0.2119, and 0.1739, respectively, along with overall accuracies of 96.1280%, 95.3535%, 95.5233%, and 96.9740%.

Model	MSE	RMSE	Accuracy (%)
RF-HHO	0.0378	0.1968	96.1280
RF-ACO	0.0465	0.2155	95.3535
HGBDT-HHO	0.0448	0.2119	95.5233
HGBDT-ACO	0.0302	0.1739	96.9740

Table 4. Initial evaluation using validation data.

Subsequently, after the initial evaluation, the models were tested with 60 new image patches outside the training and validation set. This was done to evaluate the stability of the ensemble models' performance when confronted with new data and to demonstrate their generalizability for flood detection using unknown data. Table 5 shows the evaluation results of each ensemble flood detection model in the testing phase., underscoring their improved performance and reliability. In the secondary evaluation, the RF-HHO, RF-ACO, HGBDT-HHO, and HGBDT-ACO models demonstrated enhanced performance, achieving RMSE values of 0.1531, 0.1607, 0.1613, and 0.1490, respectively. Their overall accuracies were 97.6552%, 97.4164%, 97.3951%, and 97.7789%. These results indicate that the models achieved higher accuracy and showed reduced RMSE compared to the initial evaluation

Model	MSE	RMSE	Accuracy (%)
RF-HHO	0.0234	0.1531	97.6552
RF-ACO	0.0258	0.1607	97.4164
HGBDT-HHO	0.0260	0.1613	97.3951
HGBDT-ACO	0.0222	0.1490	97.7789

Table 5. Secondary evaluation using testing data

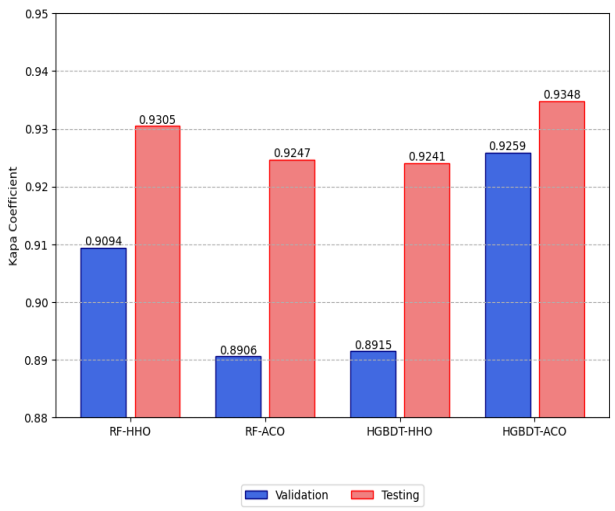

Figure 5. Ensemble models’ Kappa coefficient values.

Figure 5 shows the model's Kappa coefficient during each validation and testing phase, showcasing the models' performance and agreement levels. HGBDT-ACO model achieved the highest Kappa coefficients in both the validation and testing phases, making it the best performer according to this statistical metric among this study's ensemble flood detection models. The Kappa coefficient values for RF-HHO, RF-ACO, HGBDT-HHO, and HGBDT-ACO models were 0.9094, 0.8906, 0.8915, and 0.9259 in the validation phase and 0.9305, 0.9247, 0.9241, and 0.9348 in the testing phase, respectively.

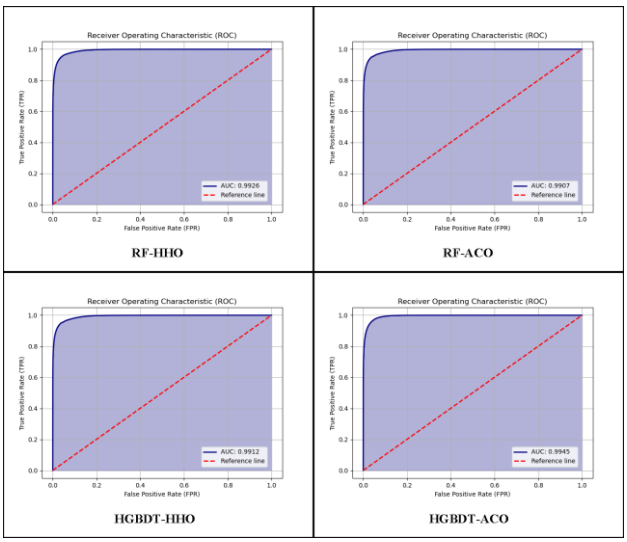

Figure 6. Ensemble models’ ROC Curves.

Figure 6 shows the ROC curves for each ensemble flood detection model based on the training and validation set data. The AUC values for RF-HHO, RF-ACO, HGBDT-HHO, and HGBDT-ACO models were 0.9926, 0.9907, 0.9912, and 0.9945, respectively. The HGBDT-ACO model, with the highest AUC value, demonstrated a better balance between the true positive rate (TPR) and the false positive rate (FPR) than the other models. Therefore, based on the ROC curve analysis, this model was the best-performing ensemble flood detection model. Figure 7 shows the classification output of a new image patch without any filtering by each of the four ensemble flood detection models. This figure shows how each model detected the flooded pixels within the image patch. Alongside these classification outputs, the ground truth label of the image patch is provided to facilitate comparison between the models' results and the actual data. This label, shown as a reference in Figure 7, provides a benchmark for assessing the models' performance.

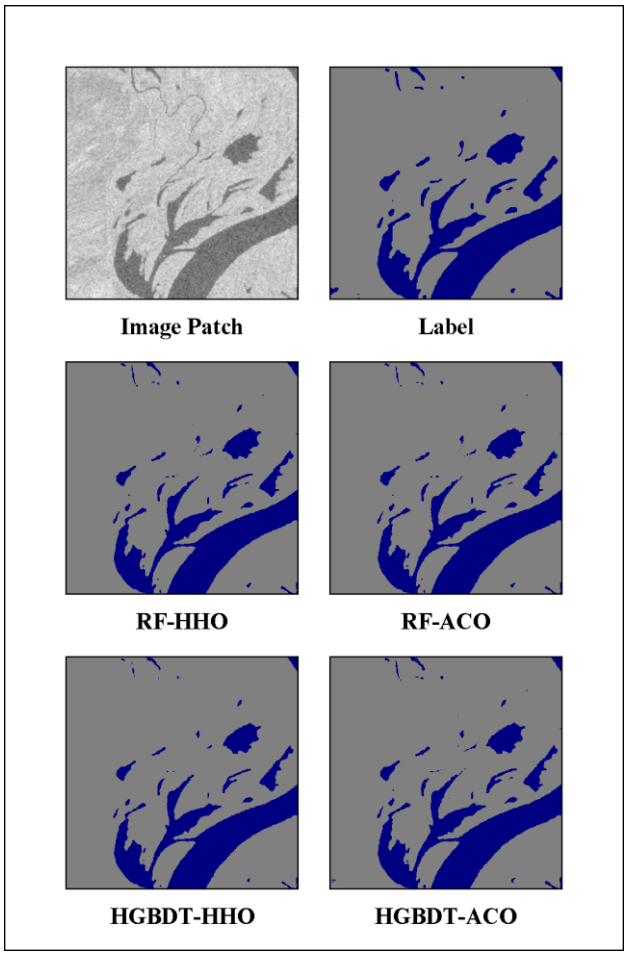

Figure 7. Comparison of models in classifying flooded and non-flooded pixels.

4. Discussion and Conclusion

This study adopted a two-step assessment to ensure greater confidence in the models' accuracy and generalizability. These indices are critical metrics for evaluating the performance of ensemble flood detection models. According to Tables 4 and 5, the HGBDT-ACO model demonstrated the lowest error rate and the highest overall accuracy compared to the other models in the validation and testing phases. Therefore, the HGBDT-ACO model was the best in this study, and the RF-HHO model ranked second. Based on these results, all four models exhibited acceptable and closely aligned performance and accuracy. This indicates the high quality and efficiency of ensemble models for flood detection purposes using Sentinel-1 imagery data.

The results of this study demonstrated that ensemble flood detection models, RF-HHO, RF-ACO, HGBDT-HHO, and HGBDT-ACO models achieved high accuracy, exceeding 95%, in detecting flood pixels from Sentinel-1 images. A simple visual comparison between the classification results of these models and the ground truth labels, as shown in Figure 7, clearly demonstrates the models' satisfactory performance in identifying flood pixels. Additionally, these models maintained their accuracy and low error rates and retained their stability and generalizability when faced with new data from the testing set in the secondary evaluation. This demonstrates that the approach of this study in implementing and enhancing the performance of flood detection models can effectively perform flood detection in various case studies, even with multispectral data or other classifiers. Although all four models achieved high accuracy, the overall performance of the HGBDT-HHO and HGBDT-ACO models was particularly noteworthy. These two models provided excellent performance and accuracy and completed the hyperparameter optimization and training processes in significantly less time than the RF models. One of the main reasons the HGBDT algorithm was used alongside the RF algorithm is their structural similarity and some shared hyperparameters. For instance, both algorithms utilize multiple decision trees, which leads to several common hyperparameters, such as the number of trees and maximum depth. However, in HGBDT, the number of trees is referred to as the maximum number of iterations. Despite these structural similarities, the two algorithms differ significantly in their operation and execution speed. The HGBDT models trained approximately 40 times faster than the RF models under the same conditions. This highlights the HGBDT's ability to handle large datasets much more efficiently and sometimes more accurately than well-known algorithms like RF in various machine learning applications. Additionally, the study underscored the importance of transfer learning techniques with pre-trained neural network models in improving feature extraction and overall model performance. Specifically, using the VGG-16 neural network model with machine learning methods and metaheuristic algorithms significantly contributed to developing advanced flood detection models. This approach can be seen as an effective strategy for enhancing the detection of complex patterns and accurately classifying surface features.

Ultimately, this study demonstrated that using optimized ensemble models could significantly enhance the accuracy and speed of flood detection. If the criteria for selecting the best models are logical performance and acceptable accuracy in the shortest time, the HGBDT-based models had an absolute advantage over the RF-based models. HGBDT models achieved higher efficiency by training a sequence of decision trees sequentially, continuously improving the performance by aggregating the results of weaker models. On the other hand, RF models, which operate based on a combination of independent decision trees, require more time. The findings highlight the potential for using HGBDT models in flood detection tasks, especially when combined with metaheuristic optimization techniques like HHO and ACO. Furthermore, implementing transfer learning with pre-trained models such as VGG-16 proved highly effective in the feature extraction automatic process and the overall performance of the machine learning models used in this study. Pre-trained neural network models like VGG-16 can significantly enhance the overall performance of machine learning models by extracting deep features.

Future studies should consider leveraging advanced pre-trained neural network models, such as VGG-19, and innovative machine learning methods and metaheuristic algorithms combinations. This approach can significantly enhance outcomes in flood detection, forest monitoring, agriculture, water resource management, and other machine learning applications using various remote sensing data.

References

Albertini, C., Gioia, A., Iacobellis, V., Manfreda, S. 2022. Detection of surface water and floods with multispectral satellites. *Remote Sensing*, 14(23), 6005.

Alidoost, F., Arefi, H. 2017. Application of deep learning for emergency response and disaster management. In *Proceedings*

of the AGSE Eighth International Summer School and Conference, pp. 11-17.

Amitrano, D., Di Martino, G., Di Simone, A., Imperatore, P. 2024. Flood detection with SAR: A review of techniques and datasets. *Remote Sensing*, 16(4), 656.

Amitrano, D., Di Martino, G., Iodice, A., Riccio, D., Ruello, G. 2018. Unsupervised rapid flood mapping using Sentinel-1 GRD SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 56(6), 3290-3299.

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.

Bai, Y., Wu, W., Yang, Z., Yu, J., Zhao, B., Liu, X., Koshimura, S. 2021. Enhancement of detecting permanent water and temporary water in flood disasters by fusing sentinel-1 and sentinel-2 imagery using deep learning algorithms: Demonstration of sen1floods11 benchmark datasets. *Remote Sensing*, 13(11).

Billah, M., Islam, A. S., Mamoon, W. B., Rahman, M. R. 2023. Random forest classifications for landuse mapping to assess rapid flood damage using Sentinel-1 and Sentinel-2 data. *Remote Sensing Applications: Society and Environment*, 30, 100947.

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.

Breiman, L. 2001. Random forests. *Machine learning*, 45, 5-32.

Cao, H., Zhang, H., Wang, C., Zhang, B. 2019. Operational flood detection using Sentinel-1 SAR data over large areas. *Water*, 11(4), 786.

Cohen, J. 1960. A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1), 37-46.

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., Fei-Fei, L. 2009. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pp. 248-255.

Diaz, G. I., Fokoue-Nkoutche, A., Nannicini, G., Samulowitz, H. 2017. An effective algorithm for hyperparameter optimization of neural networks. *IBM Journal of Research and Development*, 61(4/5), 9-1.

Dorigo, M., Maniezzo, V., Colorni, A. 1996. Ant system: optimization by a colony of cooperating agents. *IEEE transactions on systems, man, and cybernetics, part b, cybernetics*, 26(1), 29-41.

Ebadati, B., Mohammadi Ahoei, M. A., Attarzadeh, R. 2023. A Comparative Study of Flood Detection Models with Optimized Machine Learning Methods Using Sentinel-1 Data. *Earth Observation and Geomatics Engineering*, 7(1).

Elshawi, R., Maher, M., Sakr, S. 2019. Automated machine learning: State-of-the-art and open challenges. *arXiv preprint arXiv:1906.02287*.

Friedman, J. H. 2002. Stochastic gradient boosting. *Computational statistics and data analysis*, 38(4), 367-378.

Giustarini, L., Hostache, R., Kavetski, D., Chini, M., Corato, G., Schlaffer, S., Matgen, P. 2016. Probabilistic flood mapping using synthetic aperture radar data. *IEEE Transactions on Geoscience and Remote Sensing*, 54(12), 6958-6969.

Guryanov, A. 2019. Histogram-based algorithm for building gradient boosting ensembles of piecewise linear decision trees. In *Analysis of Images, Social Networks and Texts*. Springer International Publishing.

Heidari, A. A., Mirjalili, S., Faris, H., Aljarah, I., Mafarja, M., Chen, H. 2019. Harris hawks optimization: Algorithm and applications. *Future generation computer systems*, 97, 849-872.

Islam, K. A., Uddin, M. S., Kwan, C., Li, J. 2020. Flood detection using multi-modal and multi-temporal images: A comparative study. *Remote Sensing*, 12(15), 2455.

Islam, M. T., Meng, Q. 2022. An exploratory study of Sentinel-1 SAR for rapid urban flood mapping on Google Earth Engine. *International Journal of Applied Earth Observation and Geoinformation*, 113, 103002.

Jain, P., Schoen-Phelan, B., Ross, R. 2020. Tri-Band Assessment of Multi-Spectral Satellite Data for Flood Detection. In *MACLEAN@ PKDD/ECML*.

Jeyaseelan, A. T. 2003. Droughts and floods assessment and monitoring using remote sensing and GIS. *Satellite remote sensing and GIS applications in agricultural meteorology*, 291.

Klemas, V. 2015. Remote sensing of floods and flood-prone areas: An overview. *Journal of Coastal Research*, 31(4), 1005-1013.

Krishna, S. T., Kalluri, H. K. 2019. Deep learning and transfer learning approaches for image classification. *International Journal of Recent Technology and Engineering*, IJRTE, 7(5S4), 427-432.

Kulis, B., Saenko, K., Darrell, T. 2011. What you saw is not what you get: Domain adaptation using asymmetric kernel transforms. In *CVPR 2011*, pp. 1785-1792.

Lamovec, P., Matjaž, M., Krištof, O. 2013. Detection of flooded areas using machine learning techniques: Case study of the Ljubljana moor floods in 2010. *Disaster Advances*, 6(7), 4-11.

Li, W., Duan, L., Xu, D., Tsang, I. W. 2013. Learning with augmented features for supervised and semi-supervised heterogeneous domain adaptation. *IEEE Transactions on Pattern analysis and machine intelligence*, 36(6), 1134-1148.

Luu, C., Bui, Q. D., Costache, R., Nguyen, L. T., Nguyen, T. T., Van Phong, T., Pham, B. T. 2021. Flood-prone area mapping using machine learning techniques: a case study of Quang Binh province, Vietnam. *Natural Hazards*, 108(3), 3229-3251.

Martinis, S., Kersten, J., Twele, A. 2015. A fully automated TerraSAR-X based flood service. *ISPRS Journal of Photogrammetry and Remote Sensing*, 104, 203-212.

Moharrami, M., Javanbakht, M., Attarchi, S. 2021. Automatic flood detection using sentinel-1 images on the google earth engine. *Environmental monitoring and assessment*, 193(5), 248.

Morales-Hernández, A., Van Nieuwenhuyse, I., Rojas Gonzalez, S. 2023. A survey on multi-objective hyperparameter optimization algorithms for machine learning. *Artificial Intelligence Review*, 56(8), 8043-8093.

Munawar, H. S., Hammad, A. W., Waller, S. T. 2022. Remote sensing methods for flood prediction: A review. *Sensors*, 22(3), 960.

Nyamekye, C., Ghansah, B., Agyapong, E., Kwofie, S. 2021. Mapping changes in artisanal and small-scale mining (ASM) landscape using machine and deep learning algorithms.-a proxy evaluation of the 2017 ban on ASM in Ghana. *Environmental Challenges*, 3, 100053.

Otsu, N. 1975. A threshold selection method from gray-level histograms. *Automatica*, 11(285-296), 23-27.

Priyatna, M., Wijaya, S. K., Khomarudin, M. R., Yulianto, F., Nugroho, G., Afgatiani, P. M., Hussein, M. A. 2023. The use of multi-sensor satellite imagery to analyze flood events and land cover changes using change detection and machine learning techniques in the Barito watershed. *Journal of Degraded and Mining Lands Management*, 10(2).

Rahman, M. S., Di, L. 2017. The state of the art of spaceborne remote sensing in flood management. *Natural Hazards*, 85, 1223-1248.

Sadiq, R., Imran, M., Ofli, F. 2023. Remote Sensing for Flood Mapping and Monitoring. *International Handbook of Disaster Research*, 1-19.

Simonyan, K., Zisserman, A. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Solórzano, J. V., Mas, J. F., Gao, Y., Gallardo-Cruz, J. A. 2021. Land use land cover classification with U-net: Advantages of combining sentinel-1 and sentinel-2 imagery. *Remote Sensing*, 13(18), 3600.

Tammina, S. 2019. Transfer learning using vgg-16 with deep convolutional neural network for classifying images. *International Journal of Scientific and Research Publications*, 9(10), 143-150.

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.

Termeh, S. V. R., Kornejady, A., Pourghasemi, H. R., Keesstra, S. 2018. Flood susceptibility mapping using novel ensembles of adaptive neuro fuzzy inference system and metaheuristic algorithms. *Science of the Total Environment*, 615, 438-451.

Theckedath, D., Sedamkar, R. R. 2020. Detecting affect states using VGG16, ResNet50 and SE-ResNet50 networks. *SN Computer Science*, 1(2), 79.

Tran, K. H., Menenti, M., Jia, L. 2022. Surface water mapping and flood monitoring in the Mekong delta using sentinel-1 SAR time series and Otsu threshold. *Remote Sensing*. 14, 2022.

Tun, N. L., Gavrilov, A., Tun, N. M., Aung, H. 2021. Remote sensing data classification using a hybrid pre-trained VGG16 CNN-SVM classifier. In *2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering*, pp. 2171-2175.

Uddin, K., Matin, M. A., Meyer, F. J. 2019. Operational flood mapping using multi-temporal Sentinel-1 SAR images: A case study from Bangladesh. *Remote Sensing*, 11(13), 1581.

Vongkusolkit, J., Peng, B., Wu, M., Huang, Q., Andresen, C. G. 2023. Near real-time flood mapping with weakly supervised machine learning. *Remote Sensing*, 15(13), 3263.

Wu, X., Zhang, Z., Xiong, S., Zhang, W., Tang, J., Li, Z., Li, R. 2023. A Near-Real-Time Flood Detection Method Based on Deep Learning and SAR Images. *Remote Sensing*, 15(8), 2046.

Yang, L., Shami, A. 2020. On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*, 415, 295-316.