DEEP LEARNING-BASED IDENTIFICATION OF FLOOD-AFFECTED REGIONS

[1]Dr.Bhargavi Peddi Reddy[2]Banka Dharma Rao

[1]Department of Computer Science and Engineering [2]Department of Computer Science and Engineering

[1]Vasavi College of Engineering, Hyderabad, India.[2]Vasavi College of Engineering, Hyderabad, India

-- ****--

Abstract-- In recent years, numerous countries have continued to experience the impact of floods and related disasters.Individuals residing in low-lying areas and those living near water bodies such as lakes, dams, rivers, and otherreservoirs are particularly vulnerable to flooding each year. This vulnerability is largely attributable to inadequatePlanning in the construction of buildings and other public infrastructure, such as proper sewage and drainage systems. While there are various factors that contribute to the occurrence of floods, these two issues are among the most significant, as floods often result from heavy rainfall. Moreover, in the current context, even moderate rainfall can lead to flooding due to the lack of adequate space for rainwater to drain or reach coastal areas. This study aims to identify regions that may suffer damage from flooding.Additionally, it provides information to assess whether a specific area is at risk of flood-related damage.

***Keywords*:Examination of inundated regions, Flood prediction, Flood risk, Deep Learning**

I. INTRODUCTION

In the contemporary world, nearly every nation is confronted with various forms of natural disasters, including landslides, earthquakes, hurricanes, tsunamis, lightning storms, tornadoes, and volcanic eruptions. These natural calamities, inflict, significant damage on populations globally. Among the most perilous disasters encountered worldwide is river flooding. Each year, millions of communities across the numerous countries suffer extensive damage due to floods. Regions such as India, Bangladesh, China, Thailand, Indonesia, and Iraq are particularly susceptible to river flooding. A flood is characterized as an overflow of rainwater onto dry or unmoistened ground, inundating areas that are typically arid. Flooding occurs because of excessive rainfall in rivers or dams, it begins to overflow into adjacent areas, leading to flooding. Floods are categorized into four main types: Flash floods, Coastal floods, Urban floods, River floods, and Ponding, which represent the five primary categories of flooding. Flash floods occur rapidly in low-lying areas and are typically triggered by intense rainfall associated with severe storms, typhoons, cyclones or the melting of ice or snow from glaciers or icebergs.

Coastal floods, another form of flooding, usually take place in coastal regions as sea levels rise,particularly during heavy rainfall, leading to flooding events. Urban floods occur in metropolitan areas when rivers exceed their banks or when hurricanes generate storm surges along the coastline, exacerbated by excessive rainfall that has no drainage outlet. River floods arise from a sudden increase in water levels in rivers, dams, lakes or other water bodies during periods of heavy precipitation. Ponding refers to flooding that happens in low-lying areas where rainwater accumulates in regions that typically retain water in the ground, channels or ponds, which are displaced or elevated when additional rainfall enters the water management system. Flooding occurs when water levels reach critical heights. Floods can inflict significant damage on individuals residing in lower-lying areas, such as those near flowing rivers, lakes, or dams, which often experience severe destruction during flooding events. Some studies indicate that altering land levels or obstructing river paths or lakes that have been dry for extended periods to build industries or residential complexes may adversely impact local populations during floods. These factors contribute to the occurrence of floods in both developed and developing nations. It is challenging to analysed flood-prone areas as they change annually.

II. RELATED WORK

Employing data from ATER and Radiometer, Mohammad Sohadi accurately determines the present water flow path and locations of water buildup. ArcGIS uses ATER and DEMs as data sources for hydrological feature analysis. Bands A and B are combined into one image by image processing. The diversity of land uses, including urban, agricultural, and rocky regions, makes land cover analysis a challenging task. In order to find possible land regions, supervised learning uses picture categorization. Rahman notes that while evaluating pertinent data from conventional datasets, the LSTM model concurrently discards irrelevant information using the ignore gate. Using these models in several areas, the CNN model may detect patterns in social events. For the purpose of rainfall level forecasting, the LSTM model makes use of long-term dependencies. For forecasts, the LSTM has worked well. According to Pinar Ozturk, precipitation could be the cause of the flooding. As the temperature rises, the likelihood of precipitation grows and it falls more quickly. Clouds are more likely to form when temperatures rise, which eventually leads to precipitation. The ratio of the air's moisture content to its capacity at a specific temperature is used to determine the humidity. When precipitation from the sky seeps into the ground and becomes assimilated by the soil, this phenomenon is called soil conversion. In complex flood functional processes, where datasets are insufficient or missing data, the ANNS approach is heavily used in machine learning, according to Perna Jain, because of its high processing capabilities and remarkable accuracy. However, when it comes to ANNs, generalization is still very challenging [4]. There are two types of flood damage, according to Ramli Adnan: training and testing. If we want to know how well a wavelet transform-based flood prediction system works, we need to run two tests on the model's output: one for correlation and one for bias [6].

III. METHODOLOGIES

We use USGS satellite pictures consistently while preserving the original data to identify the areas inundated for this research. Accurate data about the environment is obtained by using the.tiff satellite photos. This project's overarching goal is to map out the allocated territory's heights, low-lying areas, and water reservoirs. Since the majority of riverine floods strike this area, causing immense damage and suffering for the locals, this is of utmost importance. Figure 1 depicts the capabilities of the suggested system. The input images that make up the original image showcase twelve different bandwidths that were used for processing. We create NDVI, VARI, and NDMI images to find out how much water is there, if there are any hills or other elevated features, and how dense the vegetation is in the specified regions. The principal component analysis (PCA) method makes it easier to train low-dimensional images. Differentiating between swampy lowlands and dry soil is accomplished using the K-Means algorithm. In the end, the judgments taken determine the correctness.

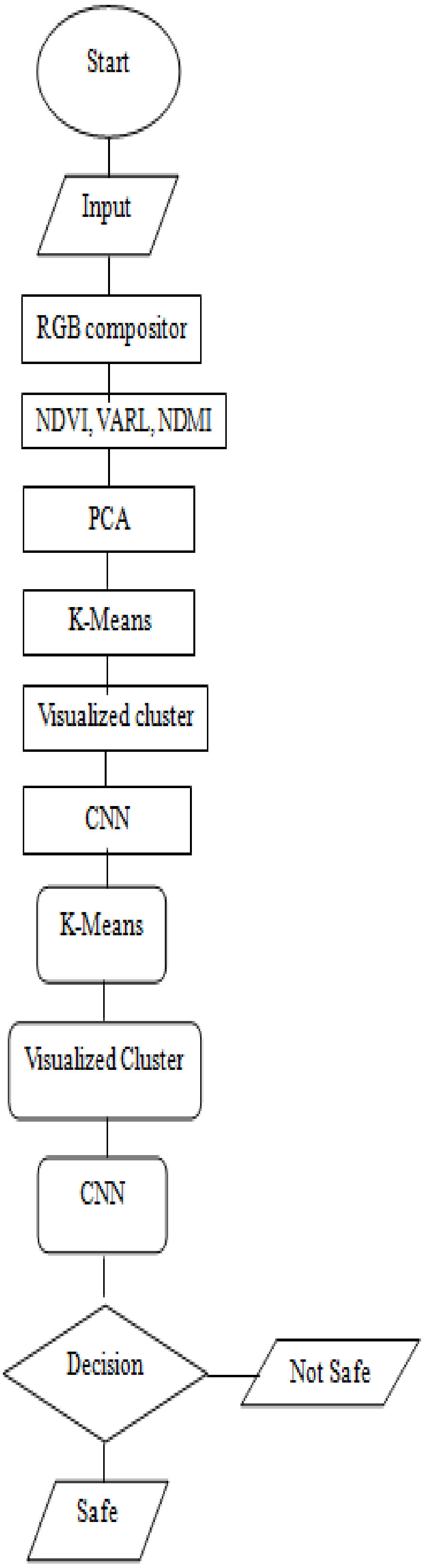

Figure.1. Flow diagram of the proposed system

In the study, we utilize several Python packages, includingRasterio, Geopandas, Matplotlib, and Pandas, which play a significant role in reading and collecting information from satellite images across the various bands(band A and band B). Figure 2 illustrates the combined input satellite image of the city in India, featuring four distinct bands.

Figure.2. Source of Combined I/P images

For processing satellite photos rasterically and for reading and writing geographic raster formats in Python, there is a specific library called Rasterio. Raster, which is based on GDAL, uses these images to extract data from satellite images. In this sense, Geopandas is another important Python library for working with geographical data and easing Raster-related procedures. We can check if the satellite image contains land and water once we load the required Python packages. We separate land from water using the Normalized Difference Vegetation Index (NDVI) picture. The Normalized Difference Vegetation Index (NDVI) allows for the separation of land cover from water settings by using normalized differences in imagery to detect vegetative areas. Still, the NDVI can't tell the difference between low-lying areas or river channels and high-lying terrains on its own. Since the VARI formula is intended to highlight crops in significant color regions while minimizing flash differences and weather effects, we use it to select high-priority zones. We can distinguish between different types of land cover (e.g., low-lying and elevated places) by looking at the effects of the atmosphere. Pictures are used in the Visible Atmospheric Resistance Index (VARI) chart.

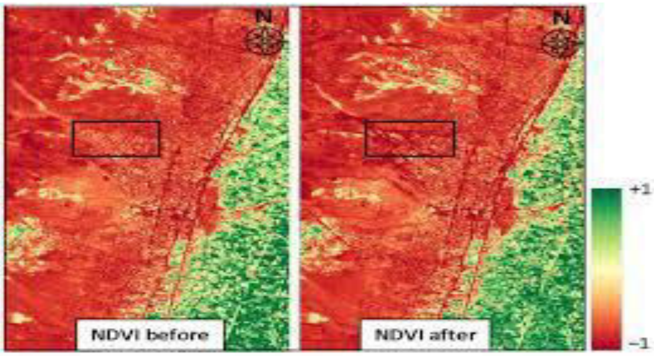

Figure. 3(a). NDVI 3(b). VARI image

To gain a more detailed understanding of the water resource area, we utilize the NDMI, which is typically employed to assess crop water levels. By applying the NDMI, we can effectively pinpoint moisture-rich areas such as dams, rivers, seas, and other water reservoirs. The NDMI (Normalized Difference Moisture Index) is capable of accurately predicting moisture areas through the normalization difference derived from image A and image B. Following the application of the NDMI, we proceeded to implement the PCA Algorithm. Figure.4 illustrates the NDMI (Normalized Difference Moisture Index) alongside images to provide a detailed identification of the water resource area.

Figure.4. NDMI Image

Using mathematical techniques to manage the dimensions of the dataset, the principal component analysis (PCA) approach is used for major data compression. Now that everything is processed, the principal component analysis shows the band with water resource zones, highlands, and lowlands depicted in great detail. Figure 5 shows the final pictures that came out of the PCA method. Our methodology has been followed in refining this technology so that it can detect changes in satellite photos.

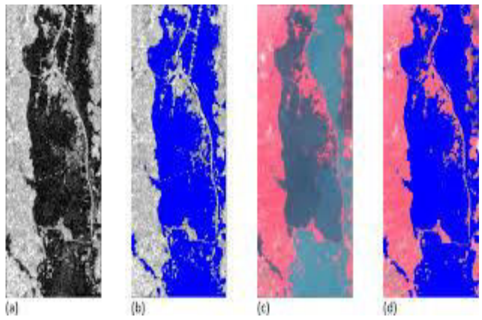

Figure.5. Output of PCA images

After principal component analysis (PCA) is run, a number of detailed insights are derived from the PCA data. In this study, we used the K-Means approach to find the cluster values from the VARI data. The cluster values were then displayed visually by means of the visual cluster. By comparing the RGB image with the cluster photos, this visualization makes it possible. The results of the K-Means algorithm are shown in Figure 6. We can see areas of stagnant water in green and areas of barren land in red. With the use of principal component analysis (PCA), we can generate k-means values, and the visible cluster displays the cluster values.

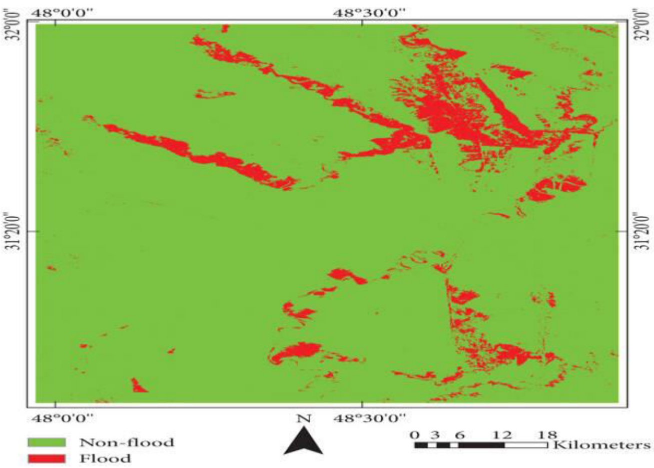

Figure.6. Output of CNN Algorithm

Thevisualized output is subsequently in the CNN. This CNN is trained using multiple images of land cover areas, along with images of water resource areas, to determine whether a specific region is likely to experience severe damage due to flooding or not.

IV. RESULT AND DISCUSSION

This project seeks to determine the specific requirements and regions that may be susceptible to damage from flooding. It collects fundamental information regarding these areas. Flooding typically occurs in low-lying areas, along rivers, near water resources such as dams, lakes, and dry lakes. These locations frequently experience damage due to floods. Additionally, this area may include damage during periods of heavy or moderate rainfall. By compiling this information, the project focuses on three key objectives: identifying low-lying areas, identifying elevated areas, and identifying water resources. These results are processed using a CNN algorithm, which is then compared with CNN-Trained models (e.g., river areas and land areas), ultimately producing a final output.

Table.1. Classification Accuracy

S.NO	Surface	Accuracy
1	Land	80%
2	River	85%

Table.1 illustrates the classification accuracy concerning land surface and water resource areas. The land surface was located in proximity to the river or other dry water reservoirs, indicating the rivers path that has been traversed over the course of a year. This straightforward information will assist individuals without being impacted by flooding. This project will utilize only satellite images from various band sets, ranging from image A and image B, both maintaining the same dimensions.

Table.2 indicates the comparison between the K-Means and CNN algorithms, focusing on precision, recall, and accuracy.

Algorith	Precisio	Recall	Accuracy
K-Mean	75.7	73.8	74.7
CNN	86.5	82.8	84.6

Table.2. Classification Accuracy

V. CONCLUSION

Nearly 85 percent of individuals lack prior knowledge regarding the suitability of a location, whether it has been affected by flooding or if it is situated at a low elevation near any water sources. This fundamental information is often overlooked by those considering moving to or purchasing land for their dream home. This project aims to elucidate the characteristics of the land offers valuable data to individuals who are unfamiliar with the area where they intend to build their dream residence or seek a safe living environment. It will assist those who require additional information about the specific location. With this knowledge, they can determine whether the area is at risk of flooding. Furthermore, this project includes a feature that assesses whether a given area is safe from flood hazards. The application of SAR data alongside deep learning has revolutionized flood plain mapping, resulting in increased speed, accuracy, and reliability. In contrast to conventional techniques that depend on optical imagery and ground surveys, SAR-based deep learning models are capable of detecting floods in real-time, even in overcast or nighttimeconditions.

VI. REFERENCES

[1]. Mohammad, (2019). The application process of satellite image and GIS-spatial modelling for flash floods. Accepted on 15May 2019 Journal of African Earth Sciences 158.

[2]. Rahman. LSTM model, CNN model used for to observe the flood and rainwater level. Accepted on May 2020.

[3]. Pinar Ozturk, Amir Mosavi, d Kwok-wing Chau (2018). Flood Prediction Using ML,a literature review on floods analysis. Accepted on 5 October 2018.

[4]. Perna Jain,Sresta Satish, and Aishwarya Krishna (2020). Flood prediction based on deep learning. IWA Publish 2020 Journal of Water and Climate Change. 2021 2nd International Conference on Intelligent Engineering and Management (CIIEM).

[5]. Ramli Adnan, Fazlina Ahmat Ruslan, Khadijah Haron, Abd Manan Samad(2017). AMISO, ARMAX model of the flood prediction system.2017 IEEE 13th International Colloquium on Signal Processing & its Applications (CSPA) 10.1109/CSPA.2017.8064947.

[6]. Yuan. J, Zhang Shaozhong (2010). Flux and level using a wavelet based on neural network flood model.2010 Third International Symposium on Knowledge Acquisition and Modelling.10.1109/KAM.2010.5646325.

[7]. Riza Budi Prasetya, Arna Fariza, Jauari Akhmad Nur Hasim, Achmad Basuki (2015) International Conference on Data and Software Engineering.
(IcoDSE).10.1109/CSPA.2017.8064941.

[8].Kei Hiroi, Nobuo Kawaguchi (2016). Flood Eye Real-time flash flood prediction system for urban complex water flow. 2016 IEEE SENSORS. 10.1109/ICSENS.2016.7808626.

[9]. Chenming Li, Jianhua Peng, Huibin Wang; Simon X. Yang. A flood prediction system for based on streaming big data processing. 2017 IEEE International conference on information and Automation.10.1109/ICInfA.2017.8079030.

[10]. Swapnil Bande; Virendra V. Shete. Smart flood disaster prediction system using IoT& neural networks. Swapnil Bande; Virendra V. Shete. 10.1109/Smart Tech Con.2017.8358367.

[11]. Mohammad Musiyam, Jumadi, Yunus Aris Wibow (2020). Analysis of Flood Affected Areas. Accepted: 16 July 2020.