2024 Girl Hackathon Ideathon Round: Solution Submission

Project Name:**Semantic Segmentation of Flood Events using U-Net for Real-Time Flood Hazard Assessment**

Pariticipant Name : Swastika Satya

Participant Email ID : swastika22102@iiitnr.edu.in

Participant GOC ID : 954790390322

ReadMe File Links (Eg: Github) : swastika2108/Google-Girl-Hackathon (github.com)

Brief summary

The problem aims to design a **deep learning model for accurate identification of flooded regions** from satellite imagery for disaster response.The model should be able to differentiate between flooded and non-flooded pixels with a high degree of accuracy.

The solution provided utilizes Convolutional neural networks trained on **pairs of water-bodies and their masks** that leverages remote sensing data for robust water segmentation,ultimately aiding flood response teams with precise location for targeted rescue efforts.It uses the **U-Net Architecture** into the model pipeline, ensuring efficient training and evaluation process.

Problem Statement

Why?
Surviving the brutal floods of Cyclone HudHud in 2014 left me stranded for two days without water or food, compounded by the anguish of not knowing where my parents were. In the midst of this disaster, surrounded by the sounds of the storm, I then decided to use technology to assist others enduring similar crises. This experience inspired me to launch a project aimed at using technology to provide aid and relief to those affected by natural disasters.

What Am I Doing?
Rightly identifying flooded regions for rescue during disaster is actually a race with time. Every wasted second searching for all those in need is surely a lost life or a life opportunity for rescue.I am developing an AI-based solution which can pinpoint location with accuracy and help the rescue team to automate their tasks to reach the needy as fast as possible.

Whom?
This project would be for the countless individuals like me, especially in the coastal regions, whose lives and livelihoods are constantly threatened by floods or cyclones.It serves a dual purpose; firstly, it acts as a tool for rescue teams to target their efforts effectively with great speed and precision, which actually mean the difference between life and death. Secondly, this model would benefit the government and disaster management agencies by automating flood detection.

Design Idea and approach

- Technologies I will use:
    - Python as a programming language and their libraries-Tensorflow and Keras.
    - Cloud Based-GPU for accelerated machine learning training and inference.
    - Pandas and Numpy for data manipulation and preprocessing.
    - Google Earth-Engine(GEEE) API:Provides real time and historical satellite images.
- New components
    - Data Preprocessing-python script takes real-time satellite images using the API or the historical dataset, processes the data via Image Augmentation and feeds it to the model.
    - Model-This trained on water bodies images and mask calculates the differences and predicts for the flood affected regions.
- Dominant Scaling Parameters
    - Data Size : Dominant scaling parameters in satellite image processing encompass the variability in size influenced by both image resolution and spectral bands. Additionally, employing **superpixel techniques** proves beneficial in reducing computational demand by aggregating pixels into larger, more manageable units while preserving spatial context.
    - Applying **Sampling Techniques**: Refers to selecting a subset of data points from the larger image dataset for analysis It reduces computational complexities while maintaining the integrity of the analysis.
- RollOut Strategy:
    - Model Size and Packaging: Machine learning model can be wrapped within a **FastAPI** framework and containerized using Docker. FastAPI offers a lightweight and efficient solution for building APIs, while Docker provides a standardized environment for packaging and deploying applications.
    - Cloud Services Leverage:Implementing A/B Testing or Canary Deployment for machine learning models can be seamlessly achieved by leveraging cloud services. These deployment strategies allow for the simultaneous testing of multiple model versions in real-world scenarios, enabling effective comparison of performance metrics and user feedback. With cloud platforms offering scalable infrastructure and built-in services for deploying and managing applications, setting up A/B testing or Canary Deployment requires minimal additional setup.

- Data Privacy and Security:
    - Implement Access Control: Establish robust access control mechanisms to safeguard sensitive data, restricting access only to authorized personnel or systems.
    - Anonymize Data: Whenever feasible, anonymize data to remove or alter personally identifiable information (PII), reducing the risk of privacy breaches and unauthorized access.
    - Obtain User Consent: Prioritize obtaining user consent for processing identifiable information, ensuring transparency and accountability in data handling practices.
    - Comply with DPD Act: Ensure compliance with the Digital Personal Data Protection Act, 2023 (DPD Act), especially when dealing with satellite imagery containing identifiable information. Adhering to relevant regulations protects individual privacy rights and mitigates legal risks associated with data processing.

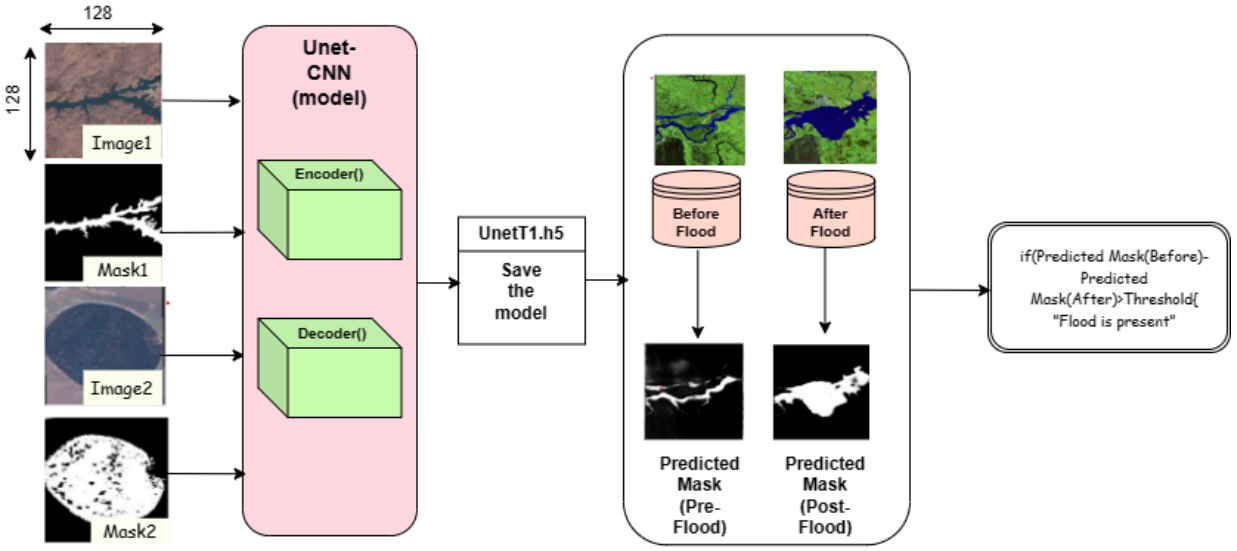

Fig 1 : Proposed Methodology

The approach used to improve the algorithm.

1. Evaluate Baseline Performance - Before making refinements, establish a baseline performance using current metrics. This helps in measuring the impact of changes.
2. Data Augmentation - Implement more sophisticated data augmentation techniques to increase the diversity of the training set without collecting new data.
3. Feature Selection - Analyze the importance of different input features. Remove redundant or less informative features that do not contribute to model performance
4. Advanced Architectures - Explore more advanced neural network architectures that are specifically designed for image segmentation tasks, such as attention U-Net or PSPNet.

Impact

Every year, coastal cyclones inflict devastating losses. The project uses AI and satellite data to revolutionize disaster management in these vulnerable regions. By analyzing real-time satellite imagery, our AI model can predict and track coastal threats, enabling more targeted resource deployment. This project is built on a strong foundation. Collaborations with ISRO and leading space agencies provide access to cutting-edge data and invaluable scientific expertise. Partnerships with the Indian Meteorological Department and disaster management bodies ensure seamless integration of our AI model into existing frameworks.

The impact will be far-reaching. Communities will benefit from significantly reduced loss of life and property damage. Better resource allocation will save lives during emergencies. *Beyond immediate response, the project's data will inform long-term strategies for land-use planning, infrastructure development that can withstand future threats, and climate adaptation.* This project promises a future where coastal communities can thrive despite the challenges of natural disasters.

Feasibility

The plan would be as follows:-

The initial data access would be to utilize Kaggle datasets as the initial source of data collection. Pre-collected datasets available the platform to start the project. We can then establish collaboration with ISRO to access historical archives and obtain live satell data.

ISRO's expertise and resources will enrich the dataset and facilitate access to real-time information crucial for disaster monitoring and response.Also,we can collaborate with other space agencies such as NASA and ESA to expand the dataset's

breadth and diversity. Accessing data from multiple sources enhances the dataset's robustness and allows for better analysis of disaster-prone regions.

We can prioritize data sampling in coastal and river regions, known for their susceptibility to floods and cyclones.Concentrating on these regions heightens the dataset's relevance to disaster management, reducing resource complexity, and enhancing utility.Developing a predictive model to estimate the frequency of satellite revisits over specific areas.Optimizing satellite revisit schedules enhances workflow efficiency, ensuring timely data acquisition for disaster monitoring and analysis purposes.

Partnerships with Key Institutions : Forge partnerships with key institutions such as the **Indian Meteorological Department (IMD)** and government organizations like **Indian Space Research Organization** in India. Collaborating with these entities provides access to domain expertise, valuable resources, and authoritative data sources, strengthening the project's credibility and impact in disaster management and risk assessment endeavors. This model is not just narrowed with floods but also can be used for other disasters.

Use of AI

The core of the solution lies in the **deep learning model**. Deep learning models use complex neural networks with multiple layers,which enable the model to extract the features from the satellite images and do an accurate flood segmentation.

Alternatives considered

1)The **X-StreamView Drone** is equipped with advanced imaging technology, ensuring it captures high-quality photographs and videos in a variety of settingsHowever, it's important to note that while the X-StreamView Drone excels in standard weather conditions, its performance is significantly hindered during extreme weather events such as typhoons and cyclones,where satellite possess an advantage.

2**)DeepLab** is a semantic segmentation model known for its effectiveness in capturing fine details and handling object boundaries. It utilizes dilated convolutions to increase the receptive field without significantly increasing parameters. However, one disadvantage of DeepLab is its higher computational complexity compared to U-Net, resulting in longer training times and higher resource requirements.

References and appendices

Kaggle Dataset Link:->Satellite Images of Water Bodies (kaggle.com)

Research Papers:->

[1]Y. Chen, H. Li, Q. Yuan, Z. Wang, C. Hu and W. Ke, "Underwater Image Enhancement based on Improved Water-Net," 2022 IEEE International Conference on Cyborg and Bionic Systems (CBS), Wuhan, China, 2023.

[2]G. Ling, F. Suo, Z. Lin, Y. Li and J. Xiang, "Real-time Water Area Segmentation for USV using Enhanced U-Net," 2020 Chinese Automation Congress (CAC), Shanghai, China, 2020.

[3]F. Ye, R. Zhang, X. Xu, K. Wu, P. Zheng and D. Li, "Water Body Segmentation of SAR Images Based on Sar Image Reconstruction and An Improved UNet," in IEEE Geoscience and Remote Sensing Letters.