COMPUTER VISION – BITS F459

Class Project: Semantic Segmentation of Flooded regions with UNet++ and SAR Preprocessing

MALAY MISHRA – 2022A7PS0116P

TABLE OF CONTENTS

1. Introduction to the Problem and Model Overview
2. Problem
3. Dataset Description
4. Pre-processing
5. **Model Description**
6. Results and Analysis
7. Conclusions
8. References

1. Introduction to the Problem and Model Overview

Floods have devastating impacts on infrastructure and human life. Detecting flood-affected regions accurately using satellite imagery is crucial for timely disaster response. This project focuses on pixel-wise segmentation of flood regions from synthetic aperture radar (SAR) satellite images using deep learning. We implement a UNet++ model with 6-channel input (stacked pre- and post-flood images), trained with Focal Tversky Loss to handle class imbalance.

2. Problem

The goal is to build a semantic segmentation model that can accurately predict flood-affected areas from a pair of SAR images taken before and after a flooding event. The model outputs a binary segmentation mask identifying flooded regions.

---

3. Dataset Description

The dataset consists of SAR images collected from various Indian flood events. Each sample includes:

- **Before Image**: A 3-channel SAR image before the flood
- **After Image**: A 3-channel SAR image after the flood
- **Mask**: A binary image marking flooded pixels

All images are preprocessed to a uniform resolution of 256×256. The final dataset is saved in a compressed.npz format containing before, after and masked arrays.

4. Preprocessing

SAR data is prone to speckle noise. To address this:

- **Lee Filtering** is applied to smooth noise while preserving edges.
- **Normalization** scales image intensity between 0 and 1.
- **Resizing** ensures uniform input dimensions.
- **Dilation** enhances flood boundaries in ground-truth masks.
- The before and after images are stacked along the channel dimension to create 6-channel inputs.

All-zero masks are filtered or flagged as warnings to avoid skewed learning.

---

5. Model Description

The model used for this flood segmentation task is **UNet++**, an advanced variant of the original UNet architecture designed for semantic segmentation. It is particularly well-suited for medical and remote sensing applications due to its improved ability to capture multiscale contextual information and reduce semantic gaps between encoder and decoder feature maps.

□ *UNet++ Architecture Overview*

UNet++ is built upon a U-shaped encoder–decoder architecture with the following key enhancements:

- **Nested and dense skip connections**: Unlike standard UNet, which connects encoder and decoder blocks at the same level, UNet++ introduces a nested structure of convolutional blocks connected with intermediate skip paths. This allows better feature fusion and reduces the semantic gap between encoder and decoder.
- **Deep Supervision** (optionally enabled): Intermediate outputs from the decoder can be supervised, improving gradient flow during training. In our implementation, this was kept in mind but not explicitly used.

Input and Output

- **Input**: 6-channel tensors created by stacking 3-channel "before" and "after" SAR images.
- **Output**: 1-channel segmentation mask with sigmoid activation for binary classification.

Layer Details

- The encoder consists of 5 stages of convolutional blocks with down sampling via max-pooling.
- The decoder is constructed from convolutional blocks with transposed convolutions (up sampling) and nested connections.
- Each block consists of two Conv-Batch Norm-ReLU operations.
- Final layer: A 1×1 convolution followed by a sigmoid activation for mask prediction.

Model Configuration

- **Input channels**: 6
- **Output channels**: 1
- **Base feature maps**: 64 → 128 → 256 → 512 → 1024
- **Optimizer**: Adam (lr=1e-4, weight_decay=1e-5)
- **Loss Function**: Focal Tversky Loss

- **Metric**: Dice Score

 Loss Function – Focal Tversky Loss

Designed to address the severe class imbalance often found in medical and disaster segmentation:

- **Tversky Index** is a generalization of Dice that penalizes false positives and false negatives asymmetrically.
- **Focal extension** emphasizes learning hard examples by raising (1 - Tversky Index) to a power γ.

This helped the model better capture minority class pixels (i.e., flooded regions), which are sparse compared to the background.

 Why UNet++?

- Handles multiscale feature aggregation effectively.
- Improves boundary delineation, crucial for flood edges.
- Reduces semantic gap via intermediate skip fusion layers.

6. Results and Analysis

The model was trained using:

- **UNet++** architecture with skip-connections and deep supervision
- **Focal Tversky Loss**, suitable for imbalanced segmentation
- **Dice Score** as the evaluation metric

Best Validation Dice Score Achieved: **0.5589** (Epoch 30)

Training stabilized after a few epochs but showed fluctuations in validation performance, likely due to class imbalance and noisy data.

I also trained on a **tiny subset (10 samples)** to ensure overfitting capability, which is a basic sanity check.

Visual Results

Visualization confirms the model is able to localize major flood regions, although finer boundaries could still be improved.

7. Conclusions

- UNet++ with 6-channel SAR inputs effectively segments flood regions.
- Focal Tversky Loss improved performance on sparse foreground masks.
- Preprocessing, especially speckle noise filtering and mask enhancement, was essential.
- Model generalization could benefit from more data or pre-trained encoders.

8. References

1. Zhou et al., *UNet++: A Nested U-Net Architecture for Medical Image Segmentation*, DLMIA 2018.
2. Salehi et al., *Tversky Loss Function for Image Segmentation Using 3D Fully Convolutional Deep Networks*, MLMI 2017.
3. Sentinel-1 SAR Datasets – European Space Agency.
4. [Preprocessing and Modeling Code] – Provided in project files.
5. **Donato Amitrano, Gerardo Di Martino, Alessio Di Simone, Pasquale Imperatore**,

Flood Detection with SAR: A Review of Techniques and Datasets,
Remote Sensing 2024, 16(4), 656.
https://doi.org/10.3390/rs16040656