Landslide Mapping in Multi Spectral Satellite Imagery using Deep Learning

submitted in partial fulfilment of the requirements
for the degree of

BACHELOR OF TECHNOLOGY

Adnan Abbas Rizvi

Supervisor
Dr. Avadh Bihari Narayan

INDIAN INSTITUTE OF TECHNOLOGY TIRUPATI

December 2024

DECLARATION

Table of Content   Page No.

ABSTRACT..5

ABBREVIATIONS...6

1. **CHAPTER 1**...7

   INTRODUCTION..7

2. **CHAPTER 2**...9

   LITERATURE REVIEW..9

3. **CHAPTER 3**...12

   METHODOLOGY...12

4. **CHAPTER 4**...15

   OUTPUT & RESULTS..15

5. **CHAPTER 5**...19

   APPLICATIONS..19

6. **CHAPTER 6**...23

   REFERENCES...23

ABSTRACT

Landslides are a critical natural hazard, causing widespread damage to infrastructure, ecosystems, and human lives, particularly in mountainous and hilly regions. Traditional methods of landslide detection, which often involve on-ground surveys and manual assessments, are not only labor-intensive but also constrained by time and accessibility to remote and hazardous areas. In response to these challenges, this project leverages advanced deep learning techniques, specifically the U-Net architecture, to develop an automated solution for landslide detection and mapping using multispectral satellite imagery and digital elevation models (DEMs).

The primary datasets utilized include Sentinel-2 multispectral imagery and DEMs from USGS, which together provide crucial indicators such as vegetation loss (NDVI), terrain slope, elevation, and exposed soil. The U-Net model, a specialized convolutional neural network for image segmentation tasks, was trained on the Landslide4Sense dataset, consisting of 3799 training images and 120 test images, with pixel-level labels indicating landslide-prone areas. Rigorous preprocessing steps, including feature extraction and data augmentation, were undertaken to ensure high-quality input for the model.

The developed model demonstrates exceptional performance with an accuracy of 99.15%, precision of 86.58%, recall of 74.84%, and an F1-score of 80.29%. These metrics underline the model's capability to accurately identify landslide-prone regions while maintaining robustness across diverse terrains and image resolutions. The use of skip connections in the U-Net architecture ensures the retention of spatial details, enabling precise segmentation critical for practical applications.

By automating landslide mapping, this project offers a scalable and cost-effective solution for real-time monitoring and early warning systems, significantly reducing the reliance on manual intervention. The results also highlight the potential for deploying this approach in disaster management, environmental monitoring, and urban planning. Future work will focus on enhancing real-time processing capabilities and incorporating additional environmental factors such as rainfall and soil type to improve the system's predictive power.

This project marks a significant step forward in leveraging remote sensing and artificial intelligence for disaster risk reduction, showcasing the transformative impact of modern technology in addressing one of the most pressing environmental challenges.

ABBREVIATIONS

- **DEM –** Digital Elevation Model
- **AI –** Artificial Intelligence
- **USGS –** United States Geographic Survey
- **RGB –** Red, Green, Blue bands
- **NDVI –** Normalized Differenced Vegetation Index
- **ML** – Machine Learning
- **SNAP –** Sentinel Application Platform
- **GDAL –** Geospatial Data Analysis Library
- **GIS –** Geographic Information System

CHAPTER 1

INTRODUCTION

1.1 Background

Landslides are a significant geological hazard that pose a grave threat to human life, infrastructure, and ecosystems. They are primarily triggered by factors such as heavy rainfall, earthquakes, deforestation, and human interventions like mining and construction. Landslides lead to widespread damage, including destruction of property, disruption of transportation networks, and loss of agricultural land, particularly in hilly and mountainous regions. Identifying landslide-prone areas and detecting their occurrence is crucial for disaster management and risk reduction.

Traditional methods for landslide detection, such as field surveys and manual mapping, require significant time, expertise, and resources. Moreover, these methods are often constrained by the inaccessibility of remote or hazardous terrains during post-disaster scenarios. This limitation underscores the need for innovative, automated approaches to landslide detection that can provide timely and accurate results over large areas.

With advancements in remote sensing technologies and the availability of high-resolution satellite imagery, significant progress has been made in environmental monitoring and disaster management. Satellite-based solutions, especially those utilizing multispectral and elevation data, have emerged as powerful tools for understanding terrain characteristics and environmental changes. When combined with artificial intelligence (AI) and machine learning (ML) algorithms, these data sources have the potential to revolutionize landslide detection.

1.2 Problem Statement

Despite the availability of high-quality satellite data, the manual interpretation and analysis of such data remain labor-intensive and prone to human error. Moreover, many existing models for landslide detection struggle to achieve the precision and scalability required for operational deployment in real-world scenarios. These challenges necessitate a robust, automated solution capable of processing vast quantities of remote sensing data to provide accurate, timely, and actionable insights.

1.3 **Objectives**

The primary objective of this project is to develop a deep learning-based framework to automate landslide detection and mapping using multispectral satellite imagery and digital elevation models (DEMs). The specific goals include:

1. Utilizing Sentinel-2 satellite data and DEMs to extract key features indicative of landslides, such as vegetation loss, slope, elevation, and exposed soil.
2. Designing a U-Net model, a state-of-the-art convolutional neural network (CNN) architecture for image segmentation, to detect landslide-prone areas at the pixel level.
3. Validating the model's performance on the Landslide4Sense dataset and assessing its generalizability to unseen geographic locations.
4. Providing a scalable and cost-effective solution for real-time landslide detection and early warning systems.

1.4 **Significance of the Study**

This study holds immense importance in advancing the state of landslide detection and monitoring. By leveraging multispectral satellite imagery and machine learning, it provides:

- Faster response times: Automated detection enables authorities to respond promptly to landslide events, minimizing casualties and damage.
- Scalability: The model can analyze vast geographic regions, making it suitable for large-scale disaster management.
- Cost-efficiency: Reducing the need for ground surveys lowers operational costs.
- Accuracy: The U-Net model's pixel-level segmentation ensures high precision in identifying landslide-prone regions, aiding planners and policymakers in risk assessment.

1.5 **Scope of the Project**

This project focuses on training and evaluating a U-Net-based landslide detection model using Sentinel-2 multispectral imagery and DEM data. The study emphasizes pixel-level accuracy and explores metrics such as precision, recall, F1-score, and accuracy to assess model performance. Case studies, including tests on Tirupati's satellite data, illustrate the model's practical applications.

By automating landslide detection, this project not only contributes to disaster risk management but also establishes a framework that can be extended to other remote sensing applications, such as flood mapping, deforestation monitoring, and urban planning.

CHAPTER 2
LITERATURE REVIEW

2.1 Introduction to Landslide Detection

Landslides are a major natural hazard that require precise and timely identification to mitigate their devastating effects. Over the years, researchers have explored various methodologies for landslide detection, ranging from manual surveys to advanced computational techniques. The introduction of remote sensing technologies and machine learning has transformed landslide detection by enabling large-scale, automated analyses of environmental and terrain changes. This section explores the evolution of landslide detection techniques, their limitations, and how deep learning is addressing these gaps**.**

2.2 Traditional Methods of Landslide Detection

Traditional landslide detection primarily relied on field surveys, geotechnical investigations, and photogrammetry. While these methods provided accurate results for small-scale studies, they suffered from several limitations:

1. Time and Resource Intensive: Field surveys require significant manpower and time, making them impractical for real-time monitoring.
2. Limited Accessibility: Remote and hazardous areas often remain unexplored due to logistical challenges.
3. Inconsistent Accuracy: Manual interpretations are prone to human error and subjectivity.

Photogrammetry and aerial surveys improved the scope of detection, but their dependency on good weather conditions and limited spatial coverage remained significant challenges.

2.3 Advances in Remote Sensing for Landslide Detection

The development of satellite technology marked a turning point in landslide detection. Multispectral and hyperspectral imaging, along with Synthetic Aperture Radar (SAR), allowed for the monitoring of terrain and environmental changes over large areas. Key datasets such as Sentinel-2 (multispectral) and DEMs became widely used due to their ability to capture crucial indicators like vegetation loss, soil disturbance, and terrain slope.

Several remote sensing indices have proven instrumental in landslide detection:

1. NDVI (Normalized Difference Vegetation Index): Detects changes in vegetation cover, often indicative of landslide events.
2. Slope and Elevation: Derived from DEMs, these parameters highlight terrain instability.
3. Soil Reflectance and Moisture Content: Indicators of landslide-prone areas.

While remote sensing offered broader coverage, traditional analysis methods often failed to harness the full potential of such data due to their inability to manage and interpret large datasets effectively.

2.4 Emergence of Machine Learning in Landslide Detection

Machine learning (ML) emerged as a promising solution for automating the interpretation of remote sensing data. Early applications of ML, including decision trees, support vector machines (SVM), and random forests, were applied to classify landslide-prone areas. These models utilized terrain and spectral features extracted from satellite data to predict landslide occurrences. However, several limitations hindered their effectiveness:

1. Feature Engineering: Classical ML models rely on manual feature extraction, which is both time-consuming and prone to bias.
2. Scalability Issues: Traditional ML models often struggled to handle the vast and complex datasets generated by modern satellite sensors.

2.5 Deep Learning for Image Segmentation and Landslide Detection

The advent of deep learning revolutionized landslide detection, particularly through its ability to process high-dimensional datasets and perform pixel-level image segmentation. Convolutional Neural Networks (CNNs) became the foundation of this transformation due to their capacity to extract hierarchical features from input images.

2.5.1 U-Net Architecture

U-Net, introduced by Ronneberger et al. (2015), was initially developed for biomedical image segmentation but quickly gained popularity in geospatial and remote sensing applications. Its U-shaped architecture, combining encoder and decoder paths with skip connections, enables the model to preserve spatial details while extracting features at multiple scales. This architecture has been applied in:

- Landcover Classification: Segmentation of vegetation, water bodies, and urban areas.
- Disaster Management: Flood and landslide mapping through pixel-level segmentation.
- Medical Imaging: Tumor and organ segmentation in CT and MRI scans.

The U-Net model's strength lies in its ability to generalize across diverse datasets, making it a prime choice for landslide detection.

2.5.2 Applications of Deep Learning in Remote Sensing

Several studies have highlighted the effectiveness of deep learning for landslide mapping:

- Ghorbanzadeh et al. (2019): Demonstrated the use of CNNs for landslide susceptibility mapping, achieving higher accuracy compared to traditional methods.
- Hong et al. (2020): Used U-Net for detecting landslide-prone areas with multispectral data, achieving precise segmentation results.
- Pham et al. (2021): Applied deep learning models to Sentinel-2 data for landslide detection, emphasizing the importance of integrating NDVI and DEM features.

2.6 Challenges in Existing Research

Despite the success of deep learning in landslide detection, challenges remain:

1. Data Availability: The lack of labeled datasets, particularly for remote regions, limits the training and validation of deep learning models.
2. Computational Complexity: High-resolution data and deep learning models require significant computational resources.
3. Generalization: Models trained on specific geographic regions may not perform well when applied to different terrains.

This project addresses these challenges by using the Landslide4Sense dataset, combining Sentinel-2 and DEM data, and optimizing the U-Net architecture for better generalization and computational efficiency.

CHAPTER 3

METHODOLOGY

The methodology for landslide mapping in multispectral satellite images using deep learning involves a systematic approach that combines data preprocessing, feature extraction, model development, training, and evaluation. This section elaborates on each step in detail.

3.1 Data Collection

The datasets used for this project were chosen for their relevance to landslide detection and availability.

1. **Sentinel-2 Multispectral Data**
   - Provided by the Copernicus Open Access Hub, Sentinel-2 data includes multispectral images across various bands (visible, near-infrared, and shortwave infrared).
   - RGB bands (Red, Green, Blue) were used to analyze exposed soil, while vegetation health was measured using the Normalized Difference Vegetation Index (NDVI).
2. **Digital Elevation Model (DEM)**
   - DEM data, sourced from the United States Geological Survey (USGS), provides information on elevation and slope. These are critical indicators of terrain instability.
3. **Landslide4Sense Dataset**
   - The dataset contains 3799 training images and 120 test images, each labeled with binary masks indicating landslide-prone areas.

3.2 Data Preprocessing

Preprocessing the data was essential to ensure compatibility with the U-Net architecture and to improve model performance. The following steps were performed:

1. **Feature Extraction**
   - The features extracted include:
     - RGB bands for soil and terrain analysis.
     - NDVI to evaluate vegetation health.
     - Elevation and slope derived from DEM data to assess terrain characteristics.
   - The extracted features were combined into a six-channel image with a shape of 128×128×6128 \times 128 \times 6128×128×6.
2. **Resampling and Normalization**
   - Images were resampled to a resolution of 128×128128 \times 128128×128 pixels to standardize the input size.
   - Normalization of pixel values was performed to scale data between 0 and 1, ensuring uniformity.
3. **Data Augmentation**

- Techniques such as flipping, rotation, cropping, and scaling were applied to enhance dataset diversity and reduce overfitting.

4. **Mask Preparation**
    - Binary masks were generated to indicate landslide-prone regions, with pixel values of 1 representing landslide areas and 0 for non-landslide areas.

3.3 Feature Selection

Key features selected for this project were chosen based on their correlation with landslide indicators:

- **Exposed Soil**: Captured using RGB bands.
- **Vegetation Health**: Computed using NDVI.
- **Slope and Elevation**: Extracted from DEM data to evaluate terrain stability.

These features were instrumental in identifying critical patterns associated with landslide-prone areas.

3.4 Model Architecture

The U-Net architecture was employed for its efficacy in image segmentation tasks.

1. **Encoder (Contraction Path)**
    - The encoder extracts hierarchical features using convolutional layers.
    - Max pooling layers reduce spatial dimensions while retaining key features.
2. **Decoder (Expansion Path)**
    - The decoder reconstructs spatial details using transposed convolutional layers.
    - Skip connections link encoder and decoder layers to preserve spatial information, ensuring precise segmentation.
3. **Output Layer**
    - The final layer uses a sigmoid activation function to generate a binary segmentation mask, where each pixel value indicates the probability of landslide presence.
4. **Model Input and Output**
    - Input shape: 128×128×6128 \times 128 \times 6128×128×6 (six feature channels: RGB, NDVI, slope, elevation).
    - Output shape: 128×128128 \times 128128×128, representing the segmentation map.

3.5 Training Strategy

1. **Loss Function**
    - **Binary Cross-Entropy (BCE)** was used to calculate the error between predicted and actual values for binary classification tasks.
2. **Optimizer**
    - The **Adam optimizer** was employed for efficient gradient-based optimization, balancing convergence speed and model performance.
3. **Regularization**

- Dropout layers were used to prevent overfitting by randomly deactivating a fraction of neurons during training.

4. **Hyperparameter Tuning**
    - Parameters such as learning rate, batch size, and dropout rate were fine-tuned using grid search to achieve optimal performance.
5. **Training Procedure**
    - The model was trained for 100 epochs on an NVIDIA GPU, with early stopping implemented to prevent overfitting.

3.6 Model Evaluation

1. **Performance Metrics**
    - Accuracy: Measures the proportion of correctly classified pixels.
    - Precision: Indicates the percentage of true positive predictions.
    - Recall: Reflects the model's ability to identify all landslide-prone pixels.
    - F1-Score: The harmonic mean of precision and recall for a balanced evaluation.
    - Confusion Matrix: Analyzes true positives, true negatives, false positives, and false negatives.
2. **Test Dataset Evaluation**
    - The model was validated on unseen test images from the Landslide4Sense dataset.
    - Predictions were compared against ground truth masks to calculate the performance metrics.

3.7 Satellite Data Preparation

For practical application, real-world satellite data from Sentinel-2 and DEM were preprocessed using the Sentinel Application Platform (SNAP) and GDAL libraries. The preprocessing included:

- Mosaicking: Combining multiple satellite tiles for larger coverage.
- Clipping: Extracting regions of interest for targeted analysis.
- Slope Calculation: Deriving terrain slope from DEM data.

CHAPTER 4

OUTPUT AND RESULTS

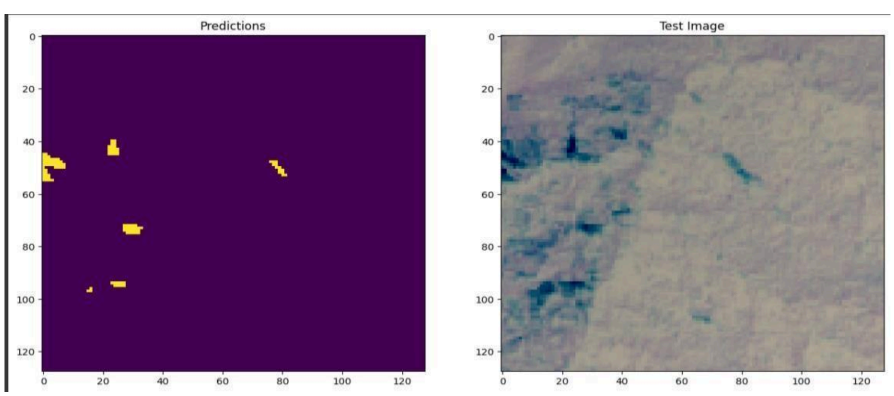
4.1 Model Performance

The model achieved the following metrics:

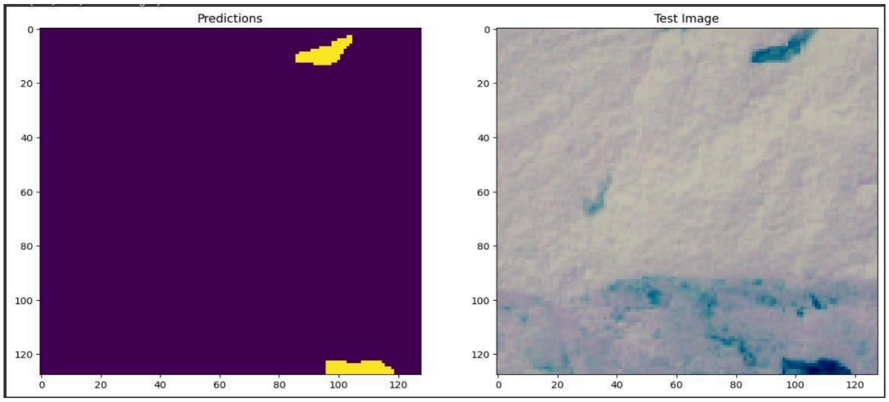
- **Accuracy**: 99.15%
- **Precision**: 86.58%
- **Recall**: 74.84%
- **F1-Score**: 80.29%
- **Confusion Matrix**: [12,126,55472,602 33,451

  72,602 215956]

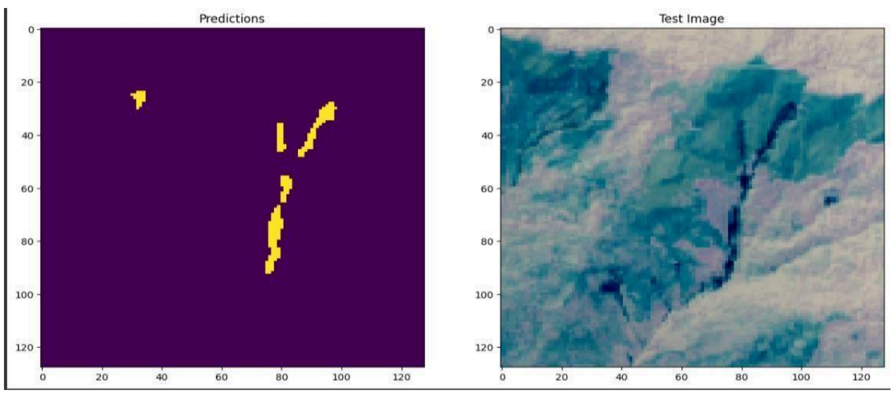
4.2 Visual Results



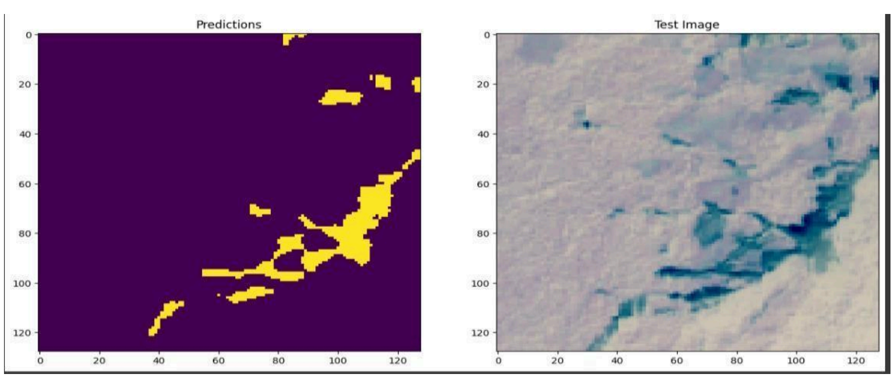
Prediction Actual Image



Prediction Actual Image



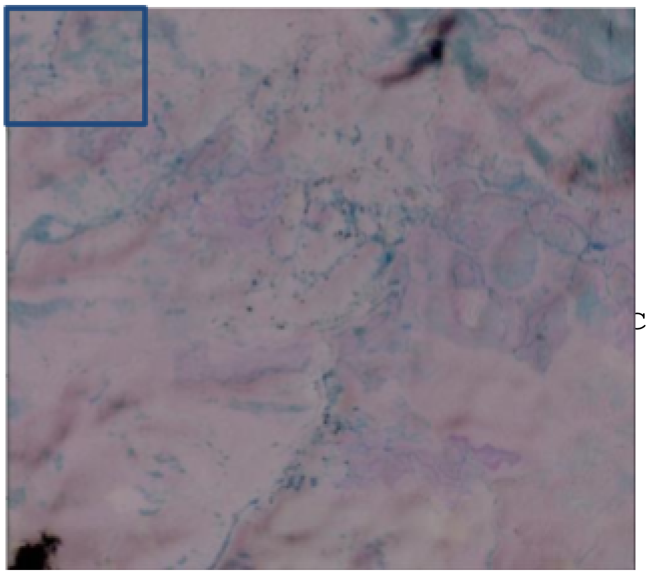
Predictions Actual Images

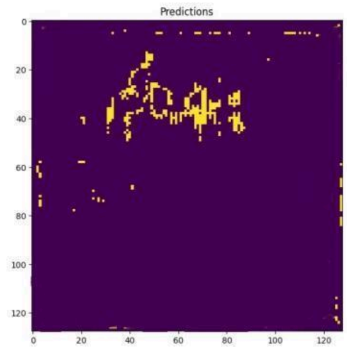


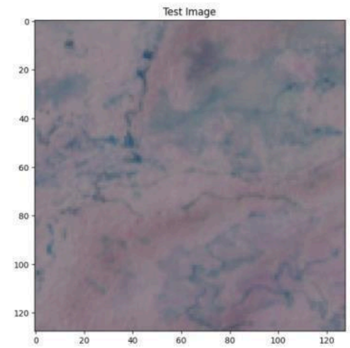
Predictions Actual Images

4.3 Case Study: Tirupati

128 x 128



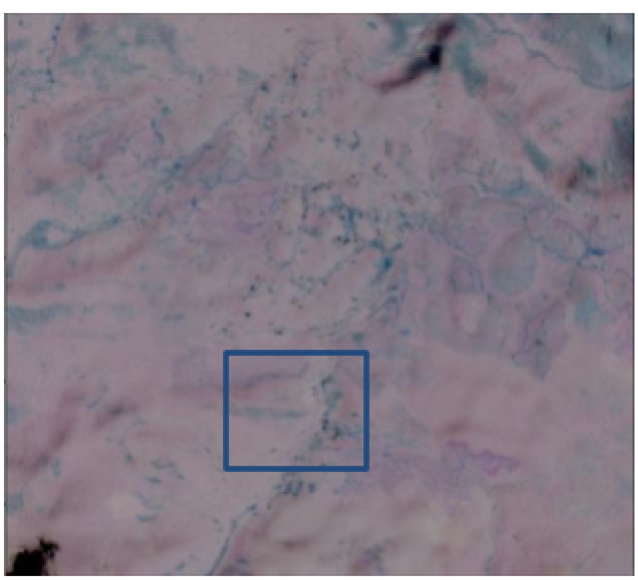




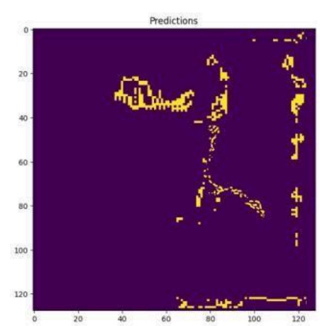
Actual Image

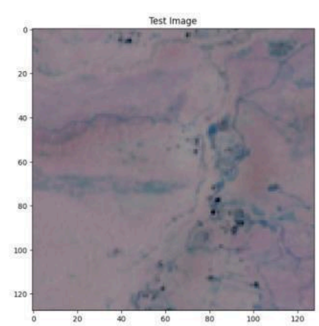
ction

460 x 567



128 x 128

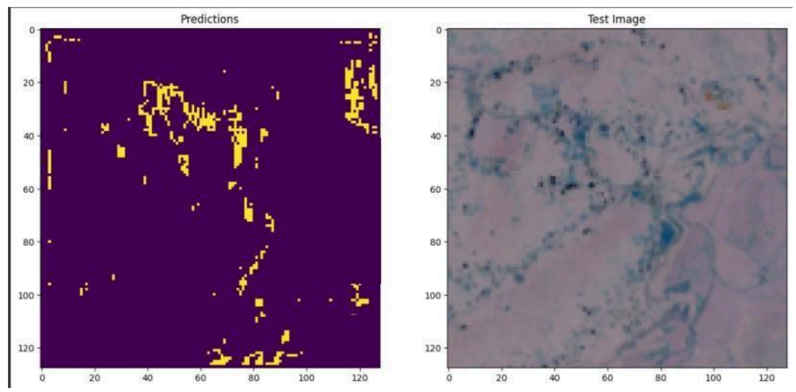


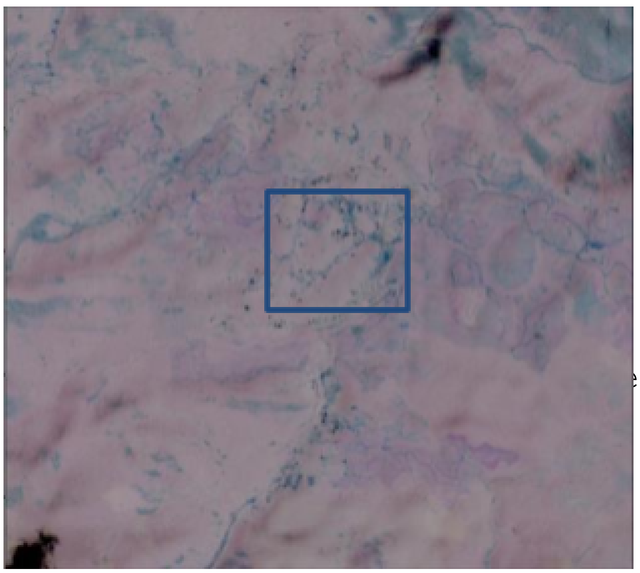


Prediction Actual Image

460 x 567

128 x 128





ediction

Actual Image

460 x 567

CHAPTER 5

APPLICATION

The U-Net-based landslide detection system developed in this project has wide-ranging applications in multiple domains. The integration of deep learning with remote sensing technologies enables innovative solutions for disaster management, environmental monitoring, and urban planning. Below are the detailed applications of this system:

6.1 Disaster Management and Mitigation

The primary application of this project is in disaster risk reduction and response.

1. **Early Warning Systems**
    - **Real-time Monitoring**: By continuously analyzing satellite data, the model can identify potential landslide-prone areas before they occur.
    - **Proactive Evacuation**: Early warnings allow authorities to evacuate residents and reduce casualties.
2. **Post-Disaster Assessment**
    - **Rapid Damage Mapping**: The model can quickly identify landslide-affected regions, enabling efficient allocation of rescue and relief resources.
    - **Infrastructure Restoration Planning**: By mapping damaged areas, the model aids in prioritizing repair work and rebuilding infrastructure.
3. **Policy and Decision-Making Support**
    - **Risk Zoning**: Using historical data, the model can help create hazard maps, highlighting high-risk areas for better urban and rural planning.
    - **Budget Allocation**: Governments can use the system to allocate funds more effectively for mitigation measures in vulnerable regions.

6.2 Environmental Monitoring

The model's capability to detect changes in vegetation, soil, and terrain can be extended to various environmental applications:

1. **Deforestation and Land Degradation Monitoring**
    - The model can be adapted to detect areas of vegetation loss, often associated with illegal logging, mining, or natural degradation.

2. **Soil Erosion Studies**
   - By analyzing changes in soil exposure, the model can help identify areas prone to erosion, aiding in the design of soil conservation measures.
3. **Climate Change Impact Analysis**
   - Landslides are closely linked to changing climatic patterns, such as increased rainfall intensity. This model can help study the impact of climate change on terrain stability.

6.3 Urban Planning and Infrastructure Development

The system is a valuable tool for planning and maintaining sustainable infrastructure in regions prone to landslides.

1. **Infrastructure Development**
   - **Road and Railway Planning**: The model can identify safer routes by avoiding landslide-prone areas during the construction of highways, railways, and tunnels.
   - **Hydropower Projects**: Assists in site selection for dams and reservoirs by analyzing terrain stability.
2. **Urban Expansion Management**
   - The model can help urban planners assess the safety of expanding residential or commercial zones into hilly areas.
3. **Maintenance of Existing Infrastructure**
   - Continuous monitoring of infrastructure in vulnerable areas can help schedule preventive maintenance and avoid catastrophic failures.

6.4 Integration with Existing Systems

This project can be seamlessly integrated with other geospatial and AI systems for broader applications.

1. **Geographic Information Systems (GIS)**
   - The outputs of the U-Net model can be overlaid on GIS platforms for enhanced visualization and analysis.
   - Integration with GIS enables the creation of comprehensive hazard maps that combine landslide risk with other factors like flood zones and earthquake susceptibility.
2. **Internet of Things (IoT)**

    - Coupled with IoT-enabled ground sensors, the model can improve landslide detection by combining satellite data with real-time ground-based observations.
3. **Time-Series Analysis**
    - By analyzing temporal satellite data, the system can detect gradual changes in terrain stability, providing insights into slow-moving landslides.

6.5 Research and Academic Applications

The system also has significant potential for research and education:

1. **Geoenvironmental Studies**
    - Researchers can use the model to study landslide causes, patterns, and relationships with environmental factors like rainfall and deforestation.
2. **Educational Tools**
    - Universities and training institutions can use the model as a teaching tool to demonstrate the application of deep learning in geospatial analysis.
3. **Extending to Other Natural Hazards**
    - The model architecture and methodology can be adapted for detecting other hazards such as floods, earthquakes, and wildfires.

6.6 Commercial Applications

The model can be commercialized to offer services to governments, private industries, and NGOs:

1. **Consulting Services**
    - Offer geospatial analysis for urban planners, civil engineers, and disaster management agencies.
2. **Real Estate and Insurance**
    - Assist real estate developers in identifying safer construction sites.
    - Help insurance companies assess risks for properties in landslide-prone areas.
3. **Agricultural Planning**
    - Aid farmers and agricultural planners in identifying safer areas for cultivation, particularly in hilly regions.

6.7 Global Applications

The scalability of the system allows it to be deployed in regions worldwide, especially in areas prone to landslides:

1. **Himalayan and Alpine Regions**: Frequent landslides due to high elevation and heavy rainfall.
2. **South-East Asia**: Monsoon-triggered landslides.
3. **Andean and Rocky Mountains**: Landslides triggered by seismic activities and deforestation.

CHAPTER 6

REFERENCES

 Ronnburger, Olaf, et al. (2015). "U-Net: Convolutional Networks for Biomedical Image Segmentation."

- Available at: arXiv U-Net Paper

 Sentinel-2 Satellite Data

- Provided by the Copernicus Open Access Hub.
- Available at: Sentinel Hub

 Digital Elevation Models (DEM) from USGS

- Acquired from the Earth Explorer portal of the United States Geological Survey.
- Available at: USGS Earth Explorer

 Landslide4Sense Dataset

- Dataset used for training and evaluation.
- Available at: Kaggle - Landslide4Sense

 Hong, H., et al. (2020). "A hybrid deep learning model for landslide susceptibility mapping using sentinel-2 data."

- Available at: ScienceDirect

 Ghorbanzadeh, O., et al. (2019). "Evaluation of Different Machine Learning Methods and Deep Learning Convolutional Neural Networks for Landslide Detection."

- Available at: Springer Link

 GDAL Documentation

- Guide for geospatial data preprocessing.
- Available at: GDAL Official Site

 TensorFlow and Keras Documentation

- Deep learning framework used for model implementation.
- Available at: TensorFlow | Keras

 Avtar, R., et al. (2019). "Monitoring landslides using remote sensing data: A review."

- Available at: ResearchGate

 SNAP (Sentinel Application Platform) Software

- For preprocessing Sentinel-2 imagery.
- Available at: SNAP Software

 Pham, B. T., et al. (2021). "Deep Learning for Landslide Susceptibility Mapping: A Case Study of Sentinel-2 Multispectral Imagery."

- Available at: Elsevier