Natural Disaster Damage Identification Using Satellite Images
A Deep Learning Approach to Satellite Imagery Analysis

Prepared by:

Esraa Bashar Aletaywi (320200603020)

Reema Abdelrazeq (320200603030)

Supervised by:

Dr. Mohammad Almseiden

TABLE OF CONTENTS

01 INTRODUCTION

Abstract 1
Introduction 1
Problem Statement 2
Objectives 2

02 METHODOLOGY

Data Collection 3
Data Preprocessing Model 4
Architecture 4
Training and Evaluation 5

03 CHALLENGES

Unbalanced data 6
Image Quality and Variability 6
Computational Resources 6
Data collection 6

04 RESULTS

Performance Metrics 7
Class specific Performance 8
Confusion Matrix Analysis 9

05 CONCLUSION

Summary 10
Future Work 10
References 11

Abstract

Accurate and timely identification of damage after natural disasters, such as earthquakes, landslides, and hurricanes, is crucial for effective disaster management. This project explores the application of deep learning techniques to classify post-disaster damage using image data. The proposed model leverages convolutional neural networks (CNNs) to achieve high accuracy and robustness in classification tasks, addressing challenges such as data imbalance and varying image quality.

Introduction

Disasters such as earthquakes, hurricanes and droughts – kill approximately 40,000 to 50,000 people per year. This is the average over the last few decades. While that's a relatively small fraction of all deaths globally, disasters can have much larger impacts on specific populations. Single extreme events can kill tens to hundreds of thousands of people. In the 20th century, more than a million deaths per year were not uncommon.

Disasters have other large impacts, too. Millions of people are displaced some left homeless by them each year. And the economic costs of extreme events can be severe, and hard to recover from. This is particularly true in lower-income countries.

In the aftermath of these catastrophic disasters, having the ability to quickly and precisely estimate damage is essential for coordinating successful disaster response and management operations. Satellite imagery, with its extensive coverage and real-time availability, is a vital tool in this process. Recent advances in deep learning, especially Convolutional Neural Networks (CNNs), have significantly increased our ability to analyze and understand satellite imagery.

CNNs specialize in extracting complex features from image data, allowing for accurate classification and inspection of disaster-related damage. This project investigates the use of this deep learning architecture to classify post-disaster damage using satellite images. Using CNNs, the proposed model seeks to achieve high levels of accuracy in damage classification. It gets over significant challenges, including data imbalance and varying image quality. This approach converts satellite images into valuable information, hence improving disaster response operations .

Problem Statement

Traditional approaches to identify disaster damage are mostly labor-intensive and manual, which increases the risk of human error and frequently delays responses. Since imagery taken by satellite offers a rich supply of information, its complexity demands the use of sophisticated analytical methods. This project's main goal is to provide an automatic, precise, and effective way to categorize satellite images of disaster-affected areas using convolutional neural networks (CNNs) to improve the speed, accuracy, and reliability of damage assessments.

Objectives

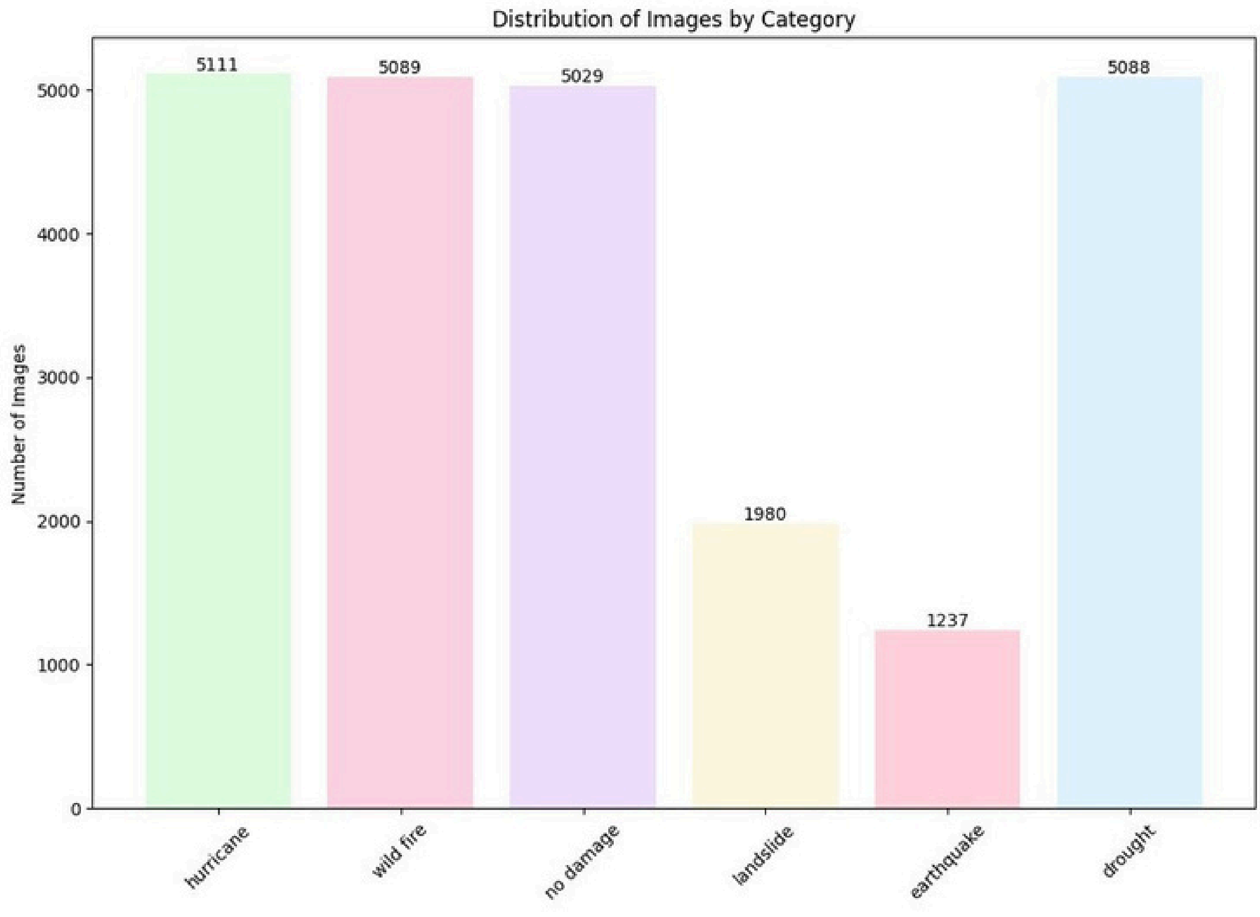
- designing and training a CNN model capable of categorizing satellite images into 5 post disaster categories.
- Address Data Imbalance: Implement techniques to manage and mitigate the effects of data imbalance in the training datasets.
- Achieve High Accuracy: Ensure the model maintains high accuracy across diverse and unpredictable disaster scenarios.

Methodology

Data Collection

The foundation of any deep learning model is the quality and quantity of the data it is trained on. For this project, we collected an extensive dataset of satellite images depicting various natural disasters, including hurricanes, landslides, earthquakes, droughts, wildfires, and non-disaster images. The data was sourced from a combination of publicly available satellite imagery repositories, government databases, and crowd-sourced platforms that provide high-resolution images which are listed in References.

The provided chart shows the image distribution over classes.



FIG(1) .Image Distribution Over Classes

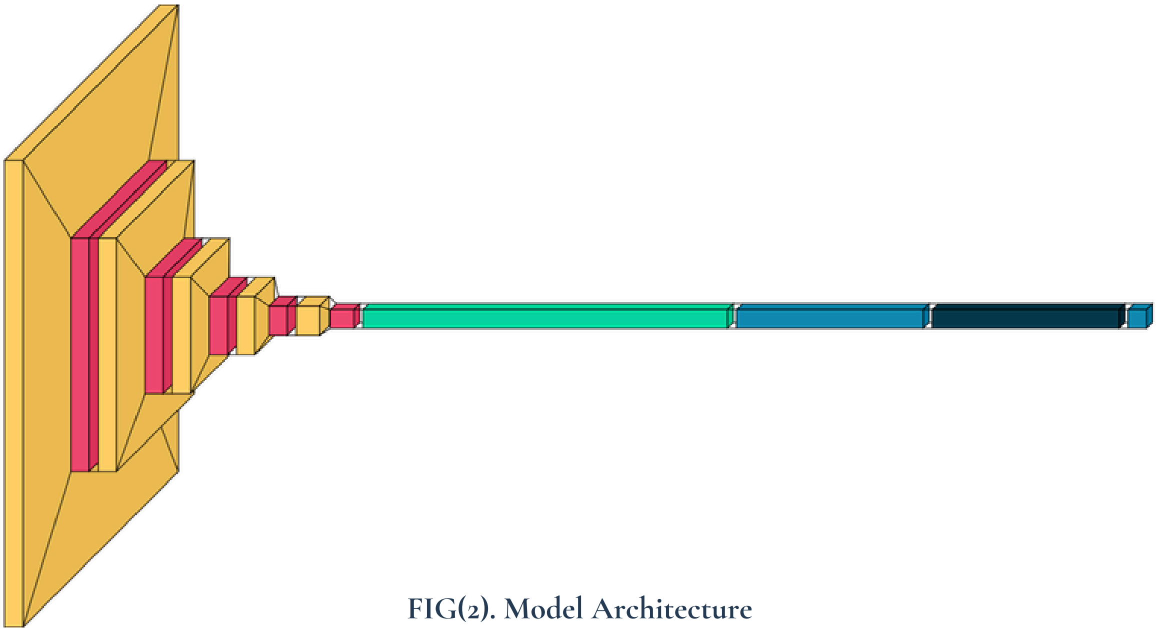
Data Preprocessing

The dataset consists of satellite images representing various damages from natural disasters. Extensive preprocessing steps, including splitting data using stratify by categories to handle imbalance, image augmentation, normalization, resizing, and computing class weights, were employed to enhance the dataset's quality and address the issue of class imbalance.

Model Architecture

1. Input Layer:
   - 32 Convolutional Filters (Size: 3x3): Processes the input image of size 200x200 with 3 color channels (RGB). Each filter extracts unique features, producing a feature map of the same size due to 'same' padding.
   - Activation Function: ReLU (Rectified Linear Unit), which introduces non-linearity, helping the model capture complex patterns in the data.
2. MaxPooling Layer:
   - Function: Reduces the size of feature maps by half, simplifying the network by reducing the number of parameters, which helps in minimizing overfitting and reducing computational load. Applied after every convolutional layer.
3. Additional Convolutional Layers:
   - Increasing Filters: The network depth increases with layers using 64, two layers with 128, and 256 filters, enhancing the model's ability to extract more complex and high-level features.
   - MaxPooling: each convolutional layer is followed by a max pooling layer to continuously reduce data dimensionality, focusing on the most essential features.

4. Flatten Layer:
   - Function: Converts the 3D output of the final pooling layer into a 1D vector, necessary to transition from convolutional layers to fully connected layers.
5. Dense Layers:
   - First Dense Layer (2048 units): A fully connected layer that interprets the features extracted and flattened previously, utilizing ReLU activation for non-linearity.
   - Dropout (0.5): Randomly ignores 50% of the nodes during training to prevent overfitting.
   - Output Layer (6 units): Uses softmax activation to classify the input into one of six categories: earthquake, hurricane, landslide, wildfire, and drought. Softmax ensures the output is a probability distribution across these classes.



FIG(2). Model Architecture

Training and Evaluation

The model was trained using the Adam optimizer, with careful tuning of learning rates and other hyperparameters to achieve optimal convergence. The evaluation process involved a thorough analysis of the model's performance across several metrics, with particular attention paid to its ability to accurately classify images from underrepresented classes.

Challenges

Unbalanced Data

The notable class imbalance in the dataset was one of the main issues that had to be dealt with. The underrepresentation of some disaster categories, such landslides and earthquakes, made it more difficult for the model to learn and generalize across all classes. To reduce this problem, strategies like stratify split, data augmentation, and computing class weights were used.

Image Quality and Variability

The quality and variability of satellite images posed another significant challenge. Differences in resolution, lighting conditions, and angles of capture introduced noise and variability into the dataset. Preprocessing techniques, such as image normalization and augmentation, were crucial in addressing these issues and improving the model's robustness.

Computational Resources

Training deep neural networks requires high computational power. High-performance GPUs are essential for handling the intensive computations involved in (CNN) training. Due to the unavailability of such GPUs on our local devices, we utilized cloud-based GPU resources to develop and train our model.

Data collection

Acquiring relevant and high-quality satellite imagery for a range of natural disasters, such as hurricanes, landslides, earthquakes, and droughts, was one of the primary challenges. Due to the variety and complexity of the needed photos, data had to be sourced from several platforms, each with different licensing and access requirements. Additional challenges were brought up by inherent variances in image quality and resolution. Careful preparation and teamwork were necessary to guarantee the accuracy and representativeness of the data that was gathered. In order to manage limited datasets and guarantee their suitability for model training, extensive preprocessing was also necessary.

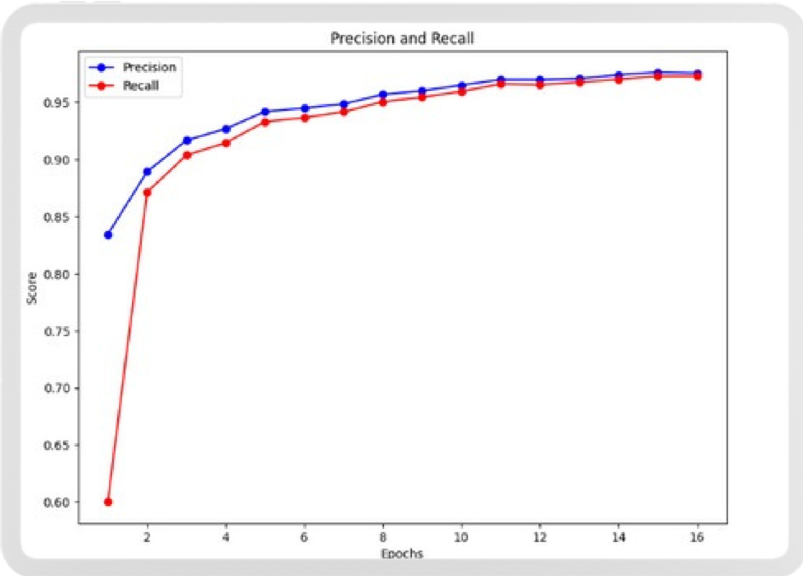
Results

Performance Metrics

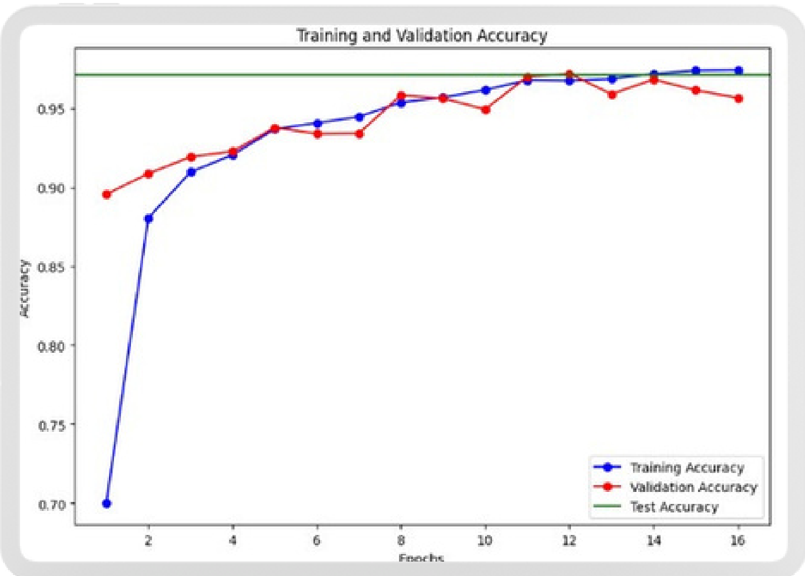
The model's performance was evaluated using a comprehensive suite of metrics. The final results demonstrated a high degree of accuracy and precision, particularly in identifying disaster classes with sufficient representation in the training data. The precision, recall, and F1-scores provided additional insights into the model's capability to handle imbalanced datasets. The evaluation metrics are as follows:

- Training Performance:
  - Accuracy: 0.9818
  - Loss: 0.0746
  - Precision: 0.9819
  - Recall: 0.9814

- Test Performance:
  - Test Loss: 0.1308
  - Test Accuracy: 0.9708



FIG(3) Training Precision and Recall



FIG(4) Training and Validation Accuracy

Class-specific Performance:

- Drought:
  - Precision: 0.9589
  - Recall: 0.9928
  - F1-score: 0.9755
- Earthquake:
  - Precision: 0.9002
  - Recall: 0.9704
  - F1-score: 0.9340
- Hurricane:
  - Precision: 0.9854
  - Recall: 0.9707
  - F1-score: 0.9780
- Landslide:
  - Precision: 0.9983
  - Recall: 0.9916
  - F1-score: 0.9949
- No Damage:
  - Precision: 0.9654
  - Recall: 0.9417
  - F1-score: 0.9534
- Wild Fire:
  - Precision: 0.9821
  - Recall: 0.9699
  - F1-score: 0.9759

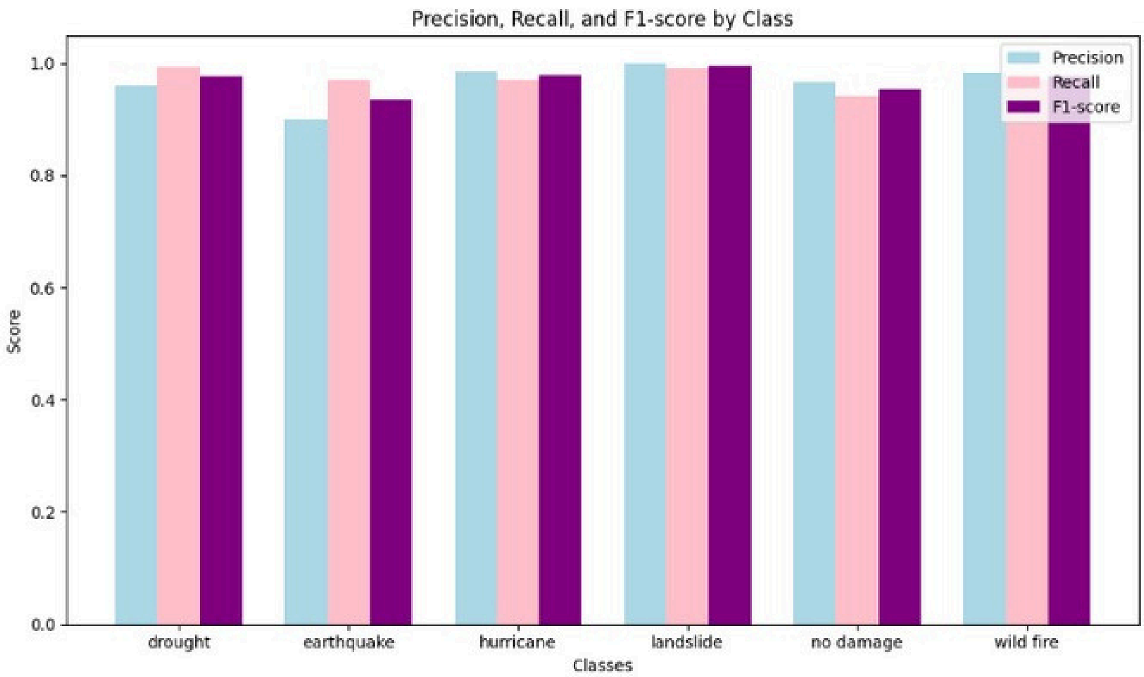


FIG (5) Recall, Precision and F1 Score Class wise

Confusion Matrix Analysis

To assess our model's ability to classify natural disasters from satellite images, we analyzed its performance using a confusion matrix. The table below summarizes the correct classifications and common errors for each type of disaster.

Disaster Type	Correctly Classified	Common Misclassifications
Drought	1516	Wild Fire: 6, No Damage: 4, Earthquake: 1
Earthquake	361	No Damage: 11
Hurricane	1489	No Damage: 10, Drought: 27
Landslide	589	No Damage: 6
No Damage	1421	Wild Fire: 37, Drought: 31
Wild Fire	1481	No Damage: 28, Drought: 16

FIG (6) Confusion Matrix Result Analysis Table

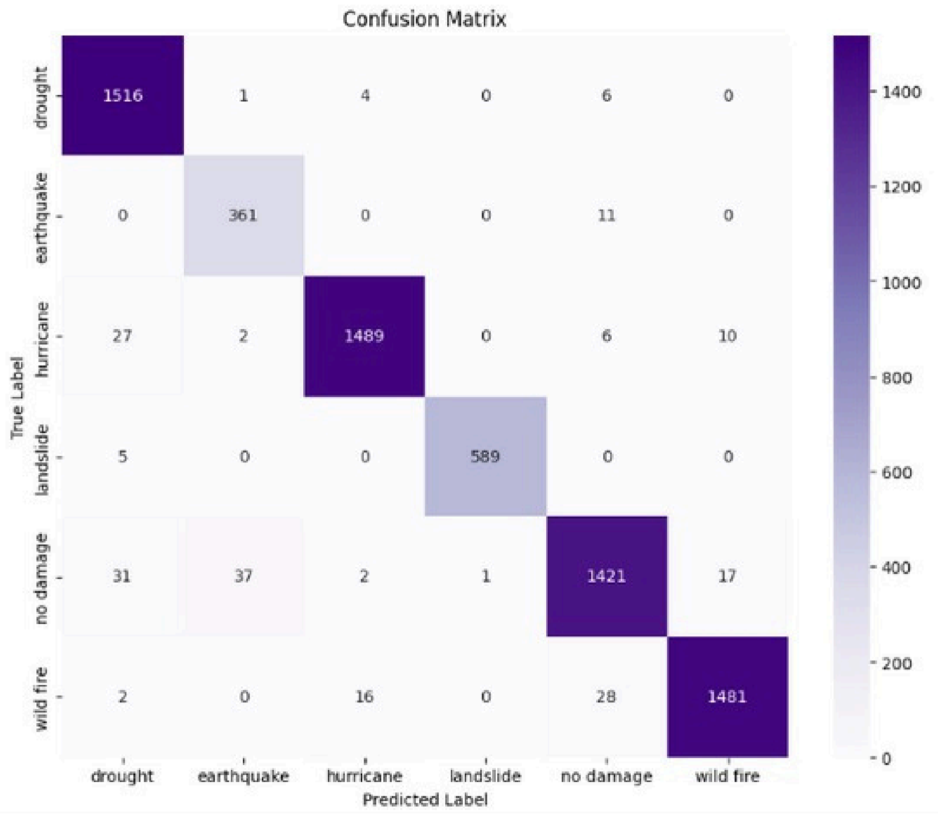


FIG (7) Confusion Matrix

Conclusion

Summary

To sum everything up, the model succeeds effectively when it comes to identifying different types of natural disasters damages from satellite photos. It accomplishes excellent recall, precision, and F1-scores across a variety of disaster classes. Nevertheless, it has some difficulties classifying scenarios correctly, especially when separating disaster-related scenarios from scenarios with no damage. Despite these difficulties, the model's overall classification of post-disaster damage accuracy and reliability indicates its potential value in disaster management and response operations.

Future Work

Future efforts will focus on expanding the dataset to include additional classes and exploring advanced techniques alongside CNNs for disaster classification. This will involve integrating multispectral and temporal data for richer analysis and leveraging more sophisticated architectures and ensemble methods. Real-world applications include incorporating the model into disaster management platforms for real-time damage assessment, enhancing decision-making for emergency responders, and providing rapid situational awareness. Additionally, validating the model across diverse datasets, collaborating with domain experts, and integrating auxiliary data sources such as social media and sensor networks will ensure its robustness and practical effectiveness in real-world disaster scenarios. These improvements aim to enhance the model's accuracy, interpretability, and utility in managing and responding to natural disasters.

References

WildFire:

- Data Source: Forest Fires - Open Government Portal, Canada.
- Original License: Creative Commons 4.0 Attribution (CC-BY) license – Quebec.
- Number of Images: 5,089

Hurricane:

- Data Source: Detecting Damaged Buildings Post-Hurricane
- Citation: http://dx.doi.org/10.21227/sdad-1e56
- Original Paper: https://arxiv.org/abs/1807.01688
- Number of Images: 5,111

Landslide:

- Original Dataset: Created by IARAI for landslide4sense 2022. The original dataset contained many images without annotations and was difficult to read.
- Current Version: Simplified and cleaned to remove zero annotation masks.
- Number of Images: 1,980

Earthquake:

- Data Source: Satellite images and UAV images from the HGM KURE application and screenshots from Google Maps. Dataset Composition: 1237 images of damaged buildings and 1498 images of undamaged buildings, totaling 2735 images. Citation: Tasci, B., Acharya, M. R., Baygin, M., Dogan, S., Tuncer, T., & Belhaouari, S. B. (2023).Inception and concatenation residual block-based deep learning network for damaged building detection using remote sensing images. International Journal of Applied Earth Observation and Geoinformation, 123, 103483. Number of Images: 1,237

References

Drought:

- Data Source: developers.google.com/earth-engine/datasets
- Citation: National Drought Mitigation Center; U.S. Department of Agriculture; National Oceanic and Atmospheric Administration (2023). United States Drought Monitor. University of Nebraska-Lincoln. https://droughtmonitor.unl.edu/. Accessed 2023-09-17
- Number of Images: 5,088

No Disaster:

- Data Source: A mix from the earthquake dataset labeled undamaged buildings, the TDS satellite dataset with 1,000 images each of normal bridges, rivers, and roads(kaggle.com/datasets/dipensaini/sateelite-images-for-computer-vision-tasks), and 600 images from the wildfire dataset labeled no fire.
- Number of Images: 5,029