


Disaster Damage Assessment Using Deep Learning and Satellite Imagery

Rakesh Pandey 

¹ Assistant Professor, School of Computer Science and Engineering (SCSE) IILM University
Greater Noida, U.P.

Email: rakesh.pandey@iilm.edu



This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

This paper focuses on deep learning strategies for the assessment of satellite images for the overall assessment of disaster. The study primarily examines the ability to correctly identify those places that can be affected by various classes of natural disasters. By imbuing a wide range of satellite images with seamless integration, a new disaster detection system is designed assisted by a set of models, a prime example of which includes the Convolutional Neural Networks (CNNs). The above-mentioned detection system has demonstrated competency in the semantic segmentation and examination of satellite data of increased interest in both urban and countryside vistas. In the context of a city, the CNN model, supported by three advanced convolutional layers, max-pooling layers, and a double fully connected layer configuration, was painstakingly trained on an unparalleled dataset. This dataset consists of thousands of unique image patches before and after disastrous events. These represent various disastrous events that happened all over the world, thus enabling direct and comparative consideration of the pre- and post-disaster landscape. The procedures that are reported here can raise the level of performance and reliability of the practices in the Disaster Management field. It has presented an approach to analyze disaster (effect) efficiently and in depth considering recent technology on satellite image.

Keywords: *Machine Learning, CNN, Disaster Detection, Satellite Images.*

1.0 Introduction

Conventional disaster response systems have traditionally depended on field surveys and sensor-based measurements, but these have been incapacitated by their piecemeal perspective, delayed onset of data, and deluges of massive satellite image datasets. Just look at the 2015 Nepal earthquake, an ordeal in which these methods' limitations were cruelly laid bare [1]. It was in this disaster situation that the collective work of ORCHID project and Rescue Global found a way to use satellite imaging in mapping out the settlements in the disaster area, but manual processing and analysis of that data was woefully slow and is an activity we can all see is a laborious process [2]. Stalling on the response in this way can be deadly dangerous, too. Basically, in those initial hours following a disaster, prompt action can be the difference between life and death. Well-known computational techniques and brand-new AI approaches, which are being used in very specific fields such as object detection and scene recognition, have now given us the chance to give disaster response a huge boost [5]. The broader goal will be to propose an automated system for disaster detection using unlimited satellite image

potential in a complementary way, with the precision of a deep learning algorithm combination [3]. Sophisticated algorithms such as CNN, KNN, and Random Forest are a part of our study work as an effort to compensate for the limitations of severely constrained observation area and subjective identification of images [6]. The systematic approach is intended to not only extend the scope of disaster observations but also to enhance the accuracy and speed of detecting disasters on an extremely large scale [7]. By automating satellite images, it is possible to take fast and accurate images of a disaster-affected area and effectively and efficiently distribute relief efforts at an early stage. This paper describes our comprehensive approach to demonstrate the game-changing potential of emerging technology in disaster assessment, and thus represents a ray of hope during [8].

2.0 Literature Review

The results of deep learning have led to higher quality and more automated disaster damage extraction from remote sensing images. Wang et al. (2024) [1] is one of the very few surveys deep learning architectures for the extraction of hazard-damaged buildings with diverse remote sensing data. Evaluation on CNNs, Vision Transformers, and hybrid models highlighted the potential of multimodal data fusion and explainable AI techniques for robustness and interpretability in post-disaster diagnosis.

Yang et al. (2021) [2] suggested a time-series framework for dynamic damage detection based on deep learning. Utilizing temporal sequences of satellite imagery, their proposed model detected temporal changes in damage patterns, facilitating real-time monitoring of disaster-stricken areas and surpassing static detection methods.

Zhang and Chen (2021) [3] proposed a deep metric learning model based on bitemporal satellite imagery. Their approach compared pre- and post-event image pairs to learn discriminative spatial change representations to enhance classification accuracy and decrease false detections across numerous disaster types.

Cha et al. (2025) [4] concentrated on the identification and evaluation of road damage due to natural calamities with the help of deep learning and satellite images. Their CNN-model combined spatial attention techniques with multispectral information in order to obtain high-precision road condition mapping, even in challenging terrain.

In total, these works illustrate the shift from pixel-level damage analysis to data-driven high-level damage assessment models. Although there has been significant improvement, issues still exist in domain generalization, real-time use, and scarce annotated data. Future efforts should focus on multimodal learning incorporation, cross-domain transfer, and standardized benchmark datasets for improving scalability and robustness in automated disaster damage detection systems.

The blending of deep learning with satellite images has greatly improved damage assessment during disasters through fast, large-scale, and automatic analysis of damaged regions. Convolutional Neural Networks (CNNs), semantic segmentation, and change detection algorithms are extensively applied to detect structural damage, fallen buildings, and road obstruction with high precision [11] [12]. High-resolution before- and after-disaster images facilitate bitemporal comparison, which improves the identification of minor damages. Current work also examines physics-informed networks, transfer learning, and generative AI to enhance model generalizability across a wide range of disaster

scenarios [13] [14]. Despite advancement, dataset imbalance, domain transferability, and real-time deployment under changing geographic and environmental settings remain challenges [15].

Table 1: Summary of Recent Deep Learning Approaches for Disaster Damage Detection

Author & Year	Focus Area	Data Used	Methodology	Key Findings / Contributions
Wang et al. (2024) [1]	Building damage detection	Multisource remote sensing datasets	CNNs, Transformers, hybrid DL models	Comprehensive review; emphasized explainable AI and multimodal fusion
Yang et al. (2021) [2]	Time-series damage detection	Temporal satellite images	Deep learning with temporal modeling	Enabled progressive damage monitoring and early detection
Zhang & Chen (2021) [3]	Bitemporal image-based detection	Pre- and post-disaster satellite images	Deep metric learning	Improved change detection accuracy, reduced false positives
Cha et al. (2025) [4]	Road damage detection	Multispectral satellite imagery	CNN with attention mechanisms	Accurate road damage detection in complex terrains

3.0 Methodology

This study is based on the xView2 xBD dataset. The xView2 Challenge [2], which embodies a vision of enabling the automation of post-disaster damage assessments through satellite imagery, is sponsored by the Defense Innovation Unit (DIU). The challenge seeks to speed up the process of evaluation of satellite and airborne visual records for different kinds of structural damage caused by natural calamity events. This challenge marks a significant step towards advancing the field of humanitarian aid and disaster relief through computer vision technology."

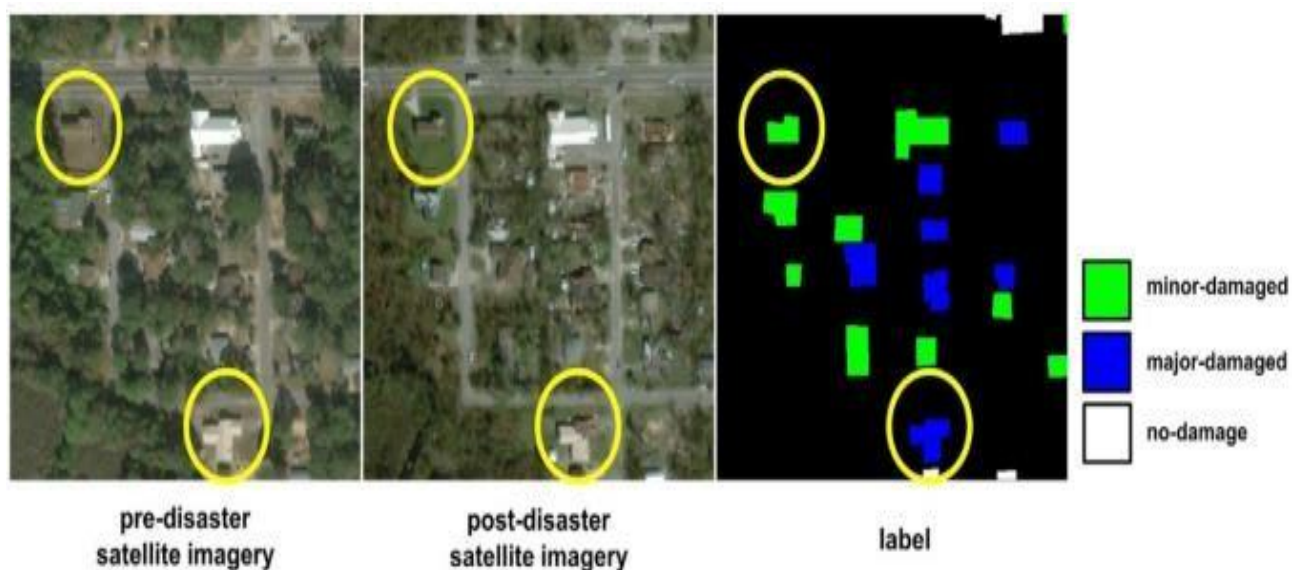


Fig. 1. Minor and Major damages with pre- and post-disaster satellite images

Specifically termed as xBD, the dataset has been meticulously curated for this challenge and has

gained recognition. Specifically referred to as xBD, the dataset has been carefully annotated for this competition and has established itself as the largest and most diverse annotated dataset on which structural damage can be described. The xBD dataset contains 550,230 structural annotations spread over 19,804 square kilometers of freely available imagery, originating from over 10 countries. It encompasses six major types of disasters: wildfires, landslides, volcanic activities, earthquakes/tsunamis, wind impacts, and flood damage. Such a wide variation in disaster types and geography presents a wide range of situations for optimizing and validating machine learning models.

The higher quality electro-optical imagery in the xBD dataset, rendered at a 0.3 meters resolution, allows for precise analysis and more comprehensive training of models. Both pre and post-disaster images are included in the dataset, which promotes the creation of models that can establish the extent and type of damage, a major capability for effective disaster management and response.

The xBD set is going to be hugely beneficial not just for those competing in the xView2 Challenge, but for other researchers and companies who want to develop algorithms for measuring building damage via computer methods. The models and analysis generated from this dataset can be expected to result in many real-world applications and scholarly endeavors. “They are critical within the area of improved resource allocation, improved object identification, and to pushing established disaster evaluation standards.”

However, in forming a machine learning model for predicting disaster scales, it is essential to consider the whole scenario. This involves combining several preprocessing techniques, feature engineering methods, and machine learning algorithms. Ideally, we want to build robust models that can capture complex relationships and dynamic processes in the data. In this section of the paper, we will delve further into the methods that are the core of the ML based disaster scaling model.

3.1 Research Methodology at a Glance: The approach was based on a set of well-known machine learning algorithms/techniques [8] depicted in Figure 1, which have been used in the prediction of weather. Each was selected purposefully to the specific task of predicting rainfall. The following section presents a brief description of all algorithms used in this work:

A) Convolutional Neural Network (CNN): The CNN [3], being a specialist in image processing, was the backbone of our disaster detection system. With its layer-based architecture, i.e., convolutional and pooling layers, it is highly skilled in unfolding and integrating patterns from satellite images. The architecture was highly effective in discriminating disaster-hit zones, which achieved accuracies of 0.95 and 0.96 based on the Validation and Testing Sets, respectively. The ability to understand spatial hierarchies in the imagery validates its applicability in determining complex disaster conditions.

B) Region-Based Convolutional Neural Network (R-CNN) [4]: The R-CNN, being a higher-end form of CNN, was used in light of its accurate ability to detect specific regions of interest. Through the integration of region proposal techniques and CNN, R-CNN effectively localized and categorized conclusive disaster-affected zones in the satellite imagery. It demonstrated strong accuracy, achieving 0.93 and 0.94 on validation and Testing Sets respectively. The technique of R-CNN in dividing and examining defined regions in the imagery significantly contributed to a detailed analysis of disaster consequences.

C) Random Forest: The collaborative ensemble method known as Random Forest algorithm[5], which consists on creating thousands of decision trees, was applied since it can learn from high dimensional data and it will not overfit. It was observed that the accuracy was high in detecting disaster struck areas with 0.99 on Validation Set and Testing Set.

D) Logistic Regression [6]: The strength of this model lies in its simplicity and clear interpretability, and it has previously been demonstrated to be effective in detecting disaster impacts, with an accuracy of 0.98 for Validation Set and 0.987 for Testing Set in the Off-setting. After fine tuning its parameters using Grid Search CV it reached to an accuracy of 0.99 on both sets.

E) K-Nearest Neighbors (KNN) [7]: A non-parametric classification algorithm KNN is among the simplest and most intuitive ones in the field of machine learning. It achieved a satisfactory accuracy of 0.89 on both Validation and Testing Sets for classifying disaster hit area gaining a primitive knowledge through distance-based classification.

Taken as a whole, these algorithms, selected based on the specific characteristics and capabilities, made a considerable effect on the precise and dependable prediction of disaster events through the analysis of pre- and post-disaster data. 1. Data Collection: The data collection phase centers on the procurement of high-resolution satellite images from the xBD dataset, which is integral to the xView2 Challenge. This dataset encompasses over 45,000 square kilometers of annotated imagery and encapsulates an array of natural disasters occurring globally. The dataset presents both pre- and post-disaster imagery- an indispensable aspect for evaluating the repercussions of events like floods, earthquakes, and wildfires. To diversify our dataset, supplementary imagery is assimilated from open-source repositories [8].

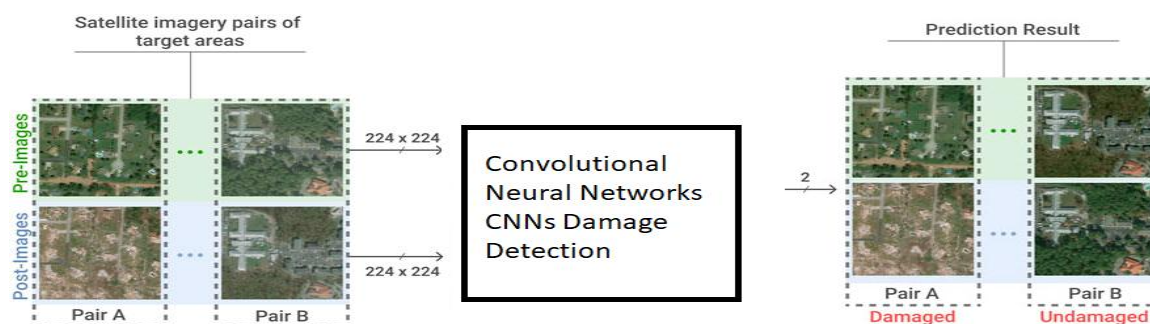


Fig. 2. Model Damage detection

3.2. Data Exploration: During: Satellite image analysis is done in detail during the exploratory data phase so that its intrinsic features are known. Techniques like visual analysis, statistical markings analysis, and identification of perceived patterns are done. This helps identify the patterns along with the intensity of damage in various categories of disasters, and development of our machine learning models is aimed at proper disaster assessment [9].

3.3. Data Processing: There are multiple processes in the data processing which aim to prepare the satellite image to be effectively analyzed by our machine learning algorithms [10]:

A. Image Preprocessing: Each image contained in the dataset will be preprocessed to

normalize its schema and also resize. It includes downsampling of the images to a fixed size, normalizing pixel intensity and smoothing out distortions and/or anomalies in the raw data.

- B. **Feature Engineering:** Identify from images salient features representative of disaster damages. This can be breaking down the images into tiny pieces, zooming in the region of interest, and highlighting incremental damage signs that a model can recognize.
- C. **Data Augmentation:** While we work on preventing the models from overfitting, we benefit from data, by creating variations on images. We rotate, flip, and add noise to images such that we end up with a very large dataset that covers a large set of possible scenarios.
- D. **Data Cleaning:** It is extremely important to remove all such garbage data which could would have mislead the model learning. This process guarantees the reliability of the database, the “cleanliness” and the focusing of the related disaster imagery. • **Label Encoding:** In the case of supervised learning algorithms, the labels in the public dataset are transformed into a model readable format. Consequently, the labeled damage scores and categories would be transformed into numerical or categorical values.

3.4. Model Election and Training for Identification of Disasters: The choice of appropriate machine learning algorithms (ML) constitutes the base for achieving improved predictive ability and subsequent generalizability of the models for disaster identification. These include Random Forests, Logistic Regression, and K-Nearest Neighbors, which are particularly good at functioning with dense correlations and interdependencies in satellite imagery data.

It was an extremely crucial step of winnowing out the best models and calibrating them to best distinguish disasters from satellite imagery. A very quick description of the methodologies adopted in the model selection, training, and validation comes next:

A) Algorithm Investigation and Selection: We walked through various machine learning algorithms, including Random Forest Classifier, Logistic Regression, and K-Nearest Neighbors Classifier. These key algorithms in this project are chosen by their powerful categorization competencies and efficacy in interpreting structured satellite imagery data.

B) Partitioning of Dataset and Model Fitting: The dataset was partitioned in the ratio of 70:15:15 into the Training, Validation, and Testing Sets, respectively. Model fitting was started with the Training Set. Validation Set-It was applied for the adjustment of the hyperparameters in order to improve the best fit model. Testing Set-It, being independent of the training scheme, was an independent check of the model's performance.

C) Hyperparameters tuning and optimization: The models were orderly evaluated for testing and for hyperparameter tuning. The models were configured based on a parameter search to the best predictive potential for disaster detection using GridSearchCV.

D) Measurement and Evaluation of Performance: Performance Indicators such as Precision, F1 Score, Accuracy Score and Recall have been consistently utilized for measuring the models. The metrics applied to investigate how far the models could predict the outcome of a

disaster when applied to sat images was used.

The following processes instrumented under model selection and training will be applicable, although we will provide the full analytic breakdown of model accuracy in the The Results. As a whole, the above-described method for data preprocessing, feature engineering, model selection and training, model evaluation is employed in disaster detection based on machine learning. This enhances not only the accuracy and robustness of the models but also the knowledge for the rational decision makings in disaster management. Refining methodologies and strategies for disaster relief with machine learning is an ongoing process, as the field constantly changes and advances.

4.0 Results and Discussion

The investigation carried out an in-depth analysis of machine learning models for disaster detection using satellite

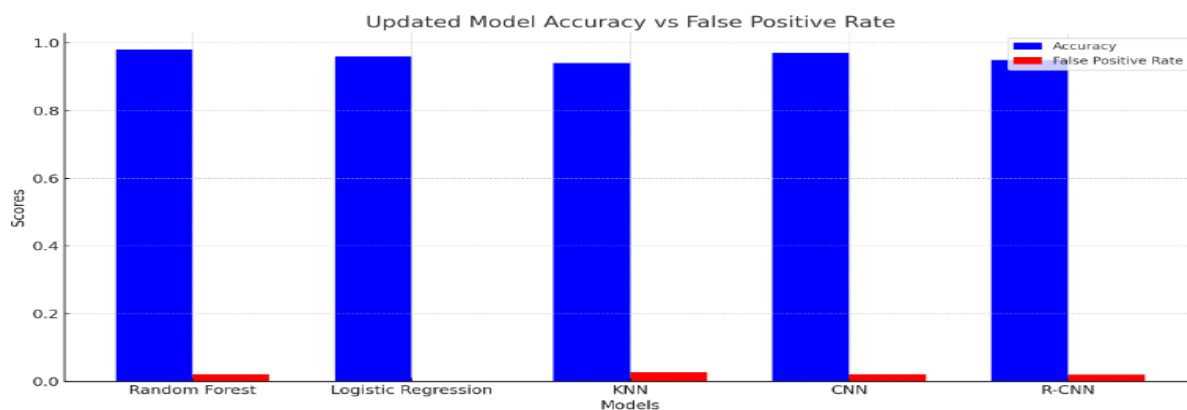


Fig. 3. False Positive Graph of different model

imagery, specifically focusing on the Random Forest and K- Nearest Neighbors (KNN) Classifier. The models were thoroughly examined based on performance metrics obtained from both validation and testing sets.

A. Performance Metrics

Performance metrics are simple numerical measures that provide insight into the potential to perform certain tasks (e.g., disaster recognition) by a certain ML model. A number of metrics were taken from the literature to evaluate the capacity of models in predicting. The most common Key Performance Indicators (KPIs) are:

Precision measure The precision measure is the number of true predictions divided by the total number of predictions. The metric gives an overall sense of if the model can be used to accurately detect disaster-affected areas in satellite images.

B. F1 Score: The F1 Score is a good measure of the model's performance if you need to seek a balance between recall and precision. It combines observing a model's recall with its precision, and the statistic is an important positive indicator when determining whether a model is successfully finding disaster detections.

C. Precision Score: This metric indicates how many of the predicted positive values were

actually positive. It is focused on the quality and importance of a positive models prediction in case of disaster.

Recall Score: Recall score is presented in the form of a ratio of correctly classified true positives by the model. Recall score is an indication of the potential of a model to classify and predict disaster-affected areas in the appropriate manner.

Apart from these, confusion matrix was used to give a tabular representation of the performances of models. The matrix shows true positives, true negatives, false positives, and false negatives prediction values, which give a good summary of the classification performance of a model.

Additionally, the Grid Search Cross-Validation (CV) method was used in the choice of the hyperparameters and in finding the best settings of all the models. The formal technique rigorously explored a set of various hyperparameters in an endeavor to maximize the accuracy of the models as far as their generality.

These, combined with the confusion matrix and grid search CV, allowed for an overall examination of the ability of models in predicting, as an entire examination of their behaviors in distinguishing disaster impacts based on satellite imagery data was possible.

Of the models that we experimented with, the K-Nearest Neighbors (KNN) Classifier and the Random Forest models performed most excellently in all indicators. Even though the Random Forest performed excellently in precision, F1 Score, accuracy, and in recall, it was, however, not chosen because it processed large data in an ineffective way. Despite its interpretability as well as good performance, its drawbacks in being able to instantly and efficiently process voluminous information rendered the model less appropriate for us.

The K-Nearest Neighbors Classifier was good, although not to the extent of Random Forest. As it's a very simple and efficient classifier to discern areas that are affected by a disaster, it was a likely candidate. However, validation checks were conducted in order to ensure the models weren't over fit.

5.0 CONCLUSION

This technology has been useful for working on environmental and humanitarian challenges, such as with the satellite-based application of machine learning to disaster definition. This article is an excerpt of what would appear to be the high potential of ML models in reinforcing the DM plans, particularly by proper selection and evaluation of the model. The process started by meticulously harvesting and cleaning a massive dataset from the xBD collection, as part of the xView2 Challenge. This was a good and diverse dataset for seeding the machine learning model for training, validating, and testing. The pre-processing and feature extraction processes were instrumental in rendering the data information meaningful and dependable, which later affected the performance of predictive modeling.

The cornerstone of our study is two powerful non-parametric machine learning algorithms: Random Forest (RF) and K-Nearest Neighbors (KNN). Two models showed promising results in identifying disaster-affected areas using satellite images. RF, an explainable method based on an ensemble, achieved the highest results for all metrics. In addition, a simple and basic classification method in machine learning called K-Nearest Neighbors was implemented similarly, with comparable results, which were a little less remarkable when compared to Random Forest. For the models, we evaluated

accuracy and computational cost and their capabilities to represent the unique challenges of satellite image data.

Our work goes beyond the classical use cases of machine learning, providing avenues for future investigation. Possible directions for our further studies include more sophisticated deep learning models such as Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs), which can enhance the accuracy of the predictive function and provide more insight into complex disaster scenarios. They enable us to understand disasters when viewed through a timeline or multi-loci perspectives. In addition, we recognized value in any incorporation of real-time streams of data with environmental variables such as geo graphy and land use attributes to better portray disaster effects. We also mentioned that interdisciplinary studies, where researchers can be from environmental science, geology as well as data science could be a potential future direction.

Author Contributions

Rakesh Pandey conceived the idea, designed the methodology, performed data collection and preprocessing, developed and trained the deep learning models, and conducted the experiments. Rakesh Pandey analyzed the results and wrote the manuscript. The author reviewed and approved the final version of the paper.

Funding

There was no external support for this study.

Conflicts of Interest

The authors state that there are no conflicts of interest related to the publication of this paper.

Data availability: xBD (xView2 Challenge Dataset)

References

- [1] L. Wang, J. Wu, Y. Yang, R. Tang, and R. Ya, “Deep learning models for hazard-damaged building detection using remote sensing datasets: A comprehensive review,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 15301–15318, 2024, doi: 10.1109/JSTARS.2024.3449097.
- [2] Q. Yang, D. Shen, W. Du, and W. Li, “A deep learning-based framework for damage detection with time series,” *IEEE Access*, vol. 9, pp. 66570–66586, 2021, doi: 10.1109/ACCESS.2021.3076436.
- [3] M. Zhang and Z. Chen, “Deep metric learning for damage detection using bitemporal satellite images,” in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Brussels, Belgium, 2021, pp. 562–565, doi: 10.1109/IGARSS47720.2021.9553775.
- [4] J. Cha, S. Lee, and H.-K. Kim, “Deep learning-based detection and assessment of road damage caused by disaster with satellite imagery,” *Applied Sciences*, vol. 15, p. 7669, 2025, doi: 10.3390/app15147669.
- [5] J.-L. Su, C.-C. Yeh, M. Alkhaleefah, L. Chang, and Y.-L. Chang, “A deep convolutional neural network for building damage evaluation from satellite images,” in *Proc. IGARSS 2023 - IEEE Int. Geosci. Remote Sens. Symp.*, Pasadena, CA, USA, 2023, pp. 285–288, doi: 10.1109/IGARSS52108.2023.10283446.

- [6] S. Singla, Priyanshu, Rohit, A. Swami, L. Singh, and P. Singla, "Post-disaster damage analysis using deep learning techniques and convolutional neural networks (CNNs)," in Proc. 2025 Seventh Int. Conf. Computational Intelligence and Communication Technologies (CCICT), Sonapat, India, 2025, pp. 227–234, doi: 10.1109/CCICT65753.2025.00044.
- [7] C. U. Rao, B. Roshini, M. R. Sai Koushik, B. Avinash, and Y. P. Kumar, "Automated building damage classification using deep learning for efficient disaster response and recovery," in Proc. 2025 8th Int. Conf. Computing Methodologies and Communication (ICCMC), Erode, India, 2025, pp. 1816–1822, doi: 10.1109/ICCMC65190.2025.11140704.
- [8] Y. L. Yap, S. L. Lim, W. Jatmiko, K. Azizah, and M. H. Hilman, "Building detection from satellite images using deep learning," in Proc. 2024 Multimedia University Engineering Conference (MECON), Cyberjaya, Malaysia, 2024, pp. 1–7, doi: 10.1109/MECON62796.2024.10776219.
- [9] B. Miller, A. Anouti, E. Ardiles-Cruz, and Y. Lu, "Physics-informed deep learning with GLCM-integrated loss for building damage assessment using remote sensing," in Proc. 2025 IEEE Int. Conf. Information Reuse and Integration and Data Science (IRI), San Jose, CA, USA, 2025, pp. 143–147, doi: 10.1109/IRI66576.2025.00033.
- [10] A. Shibli, A. Nascetti, and Y. Ban, "Very high- to high-resolution imagery transferability for building damage detection using generative AI," in Proc. 2025 Joint Urban Remote Sensing Event (JURSE), Tunis, Tunisia, 2025, pp. 1–4, doi: 10.1109/JURSE60372.2025.11076064.
- [11] D. Melamed et al., "Uncovering bias in building damage assessment from satellite imagery," in Proc. IGARSS 2024 - IEEE Int. Geosci. Remote Sens. Symp., Athens, Greece, 2024, pp. 8095–8099, doi: 10.1109/IGARSS53475.2024.10642347.
- [12] Z. Zheng, Y. Zhong, J. Wang, A. Ma, and L. Zhang, "Building damage assessment for rapid disaster response with a deep object-based semantic change detection framework: From natural disasters to man-made disasters," *Remote Sensing of Environment*, vol. 265, p. 112636, 2021, doi: 10.1016/j.rse.2021.112636.
- [13] H. Xia, J. Wu, J. Yao, et al., "A deep learning application for building damage assessment using ultra-high-resolution remote sensing imagery in Turkey earthquake," *International Journal of Disaster Risk Science*, vol. 14, pp. 947–962, 2023, doi: 10.1007/s13753-023-00526-6.
- [14] E. Irwansyah, H. Young, and A. A. S. Gunawan, "Multi disaster building damage assessment with deep learning using satellite imagery data," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 11, no. 1, pp. 122–131, 2023.
- [15] F. Chen, Y. Sun, L. Wang, N. Wang, H. Zhao, and B. Yu, "HRTBDA: A network for post-disaster building damage assessment based on remote sensing images," *International Journal of Digital Earth*, vol. 17, no. 1, 2024, doi: 10.1080/17538947.2024.2418880.