Application-Oriented Classification And Segmentation In Remote Sensing: A Satellite Image Analysis Approach

Adithi Ram
Department of AIML
N.M.A.M Institute of Technology,
Nitte Deemed to be University
adithiram123@gmail.com

Aditi S Rao
Department of AIML
N.M.A.M Institute of Technology,
Nitte Deemed to be University
aditisrao2@gmail.com

Mr. Mahesh B L
Department of AIML
N.M.A.M Institute of Technology,
Nitte Deemed to be University
mahesh.bl@nitte.edu.in

***Abstract*—This research paper presents an advanced approach to semantic segmentation in aerial imagery, focusing on pixel-level classification of satellite images to identify features like buildings, roads, vegetation, and water bodies. By leveraging cutting-edge segmentation techniques, we aim to boost the accuracy and reliability of these classifications. Employing U-Net and DeepLabV3+ architectures, our model enhances the precision and efficiency of detecting and categorizing complex land cover types within diverse urban and natural landscapes. These encoder-decoder frameworks are particularly effective at capturing nuanced spatial details, supporting high-fidelity feature extraction across various terrain types. Extending this framework, we introduce a specialized neural network for post-hurricane damage assessment, tailored to recognize disaster-specific cues such as structural damage and scattered debris. Trained on a broad, heterogeneous dataset, this model offers swift, resource-efficient mapping of affected areas, providing crucial insights for immediate disaster response. Our results highlight the potential of deep learning to deliver accurate, real-time damage assessment, ultimately facilitating more focused and efficient recovery efforts in post-disaster contexts.**

Keywords—Semantic segmentation, Aerial Imagery, Cutting-edge segmentation techniques, Encoder-decoder frameworks , Spatial details, Neural Network, Disaster-specific, Deep learning.

I. Introduction

Aerial imagery segmentation serves as a cornerstone of geospatial analysis, enabling precise, pixel-level classification of satellite or aerial images into distinct categories such as infrastructure, vegetation, water bodies, and open land. This granular mapping capability is critical in addressing modern challenges across environmental conservation, urban planning, and disaster management. With urbanization intensifying, environmental concerns rising, and disaster response demands growing, accurate aerial image segmentation has become indispensable for extracting actionable insights from vast, complex landscapes. As geospatial applications expand, the role of segmentation in producing high-fidelity, interpretable data from aerial perspectives continues to gain prominence.

This research leverages the latest in deep learning advancements, particularly the U-Net and DeepLabV3+ architectures, to enhance segmentation accuracy and model efficiency. U-Net's encoder-decoder architecture captures fine-grained spatial details, enabling the model to distinguish complex features in varied land cover types, even within dense urban and heterogeneous natural environments. DeepLabV3+, with its atrous convolution and spatial pyramid pooling layers, augments this capability by extracting broader contextual details, making it highly effective for capturing intricate terrain nuances and improving the model's adaptability to diverse datasets. By integrating these complementary architectures, our approach establishes a robust, versatile framework for high-resolution aerial segmentation.

In the context of post-hurricane damage assessment, our approach starts with a thorough exploratory analysis of disaster datasets, aiming to capture specific patterns of destruction and debris distribution unique to hurricane-affected landscapes. This groundwork enables us to develop a customized convolutional neural network model, finely tuned to identify disaster-specific indicators such as structural collapse, scattered debris, and visible signs of environmental impact. Processing aerial imagery taken immediately after a hurricane, this model offers a rapid, accurate assessment of affected areas, making it an invaluable asset for first responders and disaster management teams in delivering high-precision situational insights.

This paper fundamentally enhances landscape analysis capabilities, fostering data-driven decision-making across various sectors, from environmental sustainability and infrastructure resilience to emergency response. By swiftly generating detailed damage maps from a single aerial flyover, our model provides real-time insights that expedite recovery efforts. Offering an efficient and precise method for gauging disaster impact, it enables agencies to direct resources more effectively, ensuring that recovery actions are targeted and streamlined. This approach not only bolsters immediate situational awareness in the wake of a disaster but also promotes long-term resilience, equipping planners and policymakers with data-rich insights to shape sustainable environmental management and urban development practices.

II. Literature review

Bilel Benjdira et al. effectively addressed the issue of applying a model to new geographical areas not represented in its training dataset through the innovative use of a Generative Adversarial Network (GAN) architecture. Their approach consisted of two distinct GAN networks: the first network converted a selected image from the target domain into a corresponding semantic label. The second network then transformed this generated semantic label into an image that aligned with the source domain while preserving the semantic integrity of the original target image. Utilizing the ISPRS semantic segmentation dataset, their methodology achieved a notable 24% improvement in accuracy when transitioning from the Potsdam domain to the Vaihingen domain. [1]

In 2023, Behera et al. proposed a unique multi-scale Convolution Neural Network (CNN) architecture that incorporates superpixels to enhance classification accuracy in complex urban aerial images. Their innovative two-tier segmentation framework begins by generating superpixel images through linear iterative clustering, which captures essential contextual information. In the second tier, a multi-scale CNN analyzes these superpixels to extract invariant features for each pixel. Tested on the National Institute of Technology Rourkela drone dataset and the urban drone dataset, their method demonstrated superior performance, achieving higher accuracy than existing approaches in challenging urban environments. [2]

A recent study by D. Jozi et al. introduced an innovative image-processing approach for assessing post-disaster building damage through UAV imagery. This method integrated texture-based features, like dissimilarity and homogeneity, with edge-based characteristics derived from Canny edge detection. Unique indices were proposed to find the structural irregularities by analyzing the uniformity of edge angle distributions and entropy. Using these features, a Naïve Bayesian classifier effectively distinguished damaged from undamaged buildings, achieving 89.3% accuracy on real-world disaster images. This research highlighted the potential of UAV-based image analysis for fast, reliable damage assessment in post-disaster scenarios. [3]

Thakkar et al. proposes an innovative segmentation approach using a single-day image dataset from a village in Gujarat, obtained via Google Earth’s Sentinel satellite imagery, to classify two key crops: wheat and Ricinus. The study evaluates various segmentation architectures, including machine learning, U-Net and deep learning, optimizing performance through different learning rate strategies. The U-Net architecture achieved a validation accuracy of 94.6% at 500m altitude, while the ResUNet with a cyclic learning rate for 1000m altitude images surpassed traditional methods, reaching 98.5% accuracy. These findings highlight the superior performance of U-Net and its variants with cyclic learning rates over conventional machine learning and deep learning techniques in agricultural segmentation tasks. [4]

In 2021, Zhang et al. proposed an innovative atrous spatial pyramid pooling U-Net (ASPP-U-Net) to enhance the speed and accuracy of land-cover classification for medium-resolution remote-sensing images. This advanced model is specifically designed to leverage the unique characteristics of medium-resolution imagery,

outperforming conventional techniques such as SVM, patch-based CNN, and traditional U-Net in terms of both classification precision and inference efficiency. Their results suggest that ASPP-U-Net is a highly effective method for mapping land cover in medium-resolution images. However, the authors note that the potential for further improvements in classification accuracy may be hindered by the presence of inaccurately labelled reference data. [5]

The article by C. Kyrkou and T. Theocharides, explored improvements in aerial image classification for emergency response using UAVs. It introduced a specialized Aerial Image Database for Emergency Response applications and performed a detailed comparison of existing methods. From the analysis, a lightweight convolutional neural network called EmergencyNet was developed which used atrous convolutions to efficiently process multi-resolution features. This architecture was designed to run effectively on low-power embedded platforms, achieving up to 20 times better performance than existing models. Importantly, it maintained minimal memory use and had less than 1% accuracy loss compared to state-of-the-art methods. [6]

This study presented a machine learning approach for autonomous drones to identify and prioritize flood-affected areas using image classification. The system combined Inception v3 and DenseNet CNNs with a sorting algorithm to guide relief efforts, prioritizing the most affected areas first. Inception V3 performed better, achieving 83% accuracy compared to DenseNet’s 81% on a custom flood severity dataset. This approach demonstrated CNNs' potential in supporting autonomous decision-making for robotic disaster response. [7]

This study investigated various deep learning techniques combined with geographical hash codes. In particular, the geohash method are used to encode the geographic coordinates of satellite photos into a string of binary codes. To improve the deep neural network's semantic segmentation capabilities for satellite images, the binary codes of the geographic coordinates are then input into the network using three distinct techniques. The efficacy of incorporating geographic locations into neural networks is demonstrated through experiments conducted on three datasets. [8]

The article suggests a unique multichannel water body detection network that takes advantage of multispectral imagery. To overcome the aforementioned upsampling problem, the model employs Sentinel-2 RGB, NIR, and SWIR bands and designs a multichannel fusion module to handle the various image resolutions. To maintain the saliency of the high-dimensional representations, a new Enhanced Atrous Spatial Pyramid Pooling module has also been used to extract multireceptive feature representations and Space-to-Depth/Depth-to-Space operations in place of max pooling and upsampling. [9]

A framework utilizing segmentation neural networks is suggested to identify affected areas and accessible roadways in order to offer fast and actionable information for disaster response. The performance of well-known segmentation models has been evaluated, and the efficacy of pretraining with ImageNet for the job of aerial image segmentation has been examined. According to experimental findings, pretraining on ImageNet typically enhances several models' segmentation ability. Training is done using open data from

OpenStreetMap (OSM), eliminating the need for laborious hand annotation. In order to detect changes brought on by a natural disaster and to update road network data that is available via OSM, the method additionally uses graph theory. [10]

III. System design

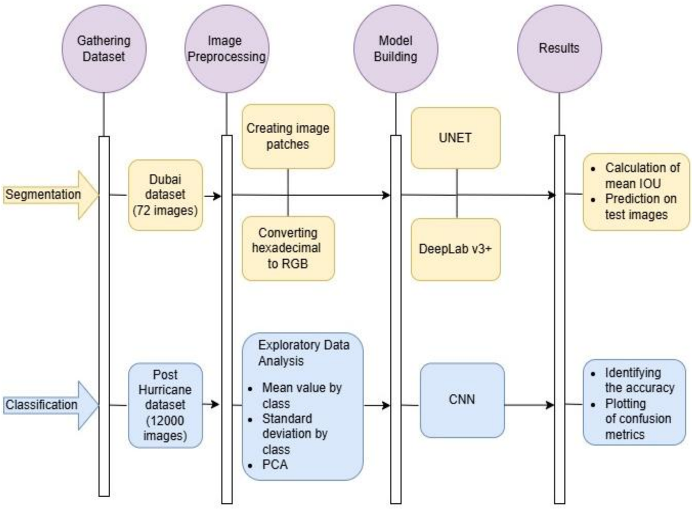

Fig. 1. Illustrates the sequence diagram of the system

Fig. 1. illustrates the process of image segmentation, starting with the Dubai dataset as input. The images are prepared by converting to RGB and dividing them into smaller patches for easier processing. Two powerful models, UNET and DeepLab V3+ are then employed to identify and separate objects within the images. The models' performance is evaluated using the mean Intersection over Union (IoU) score, which measures segmentation accuracy.

The next task focuses on classifying images from the post-hurricane dataset into predefined classes. This involves preprocessing the images, conducting exploratory data analysis to understand data distribution, and training a Convolutional Neural Network model for classification. The model's effectiveness is assessed using accuracy and confusion matrices.

IV. Methodology

In this paper, we design U-Net and DeepLabV3+ models to segment essential categories in aerial imagery including buildings, roads, vegetation and water bodies, and implement a CNN-based architecture tailored to detect disaster-specific indicators. The framework for building our segmentation and damage assessment system follows a structured process that can be divided into two stages: Segmentation Phase and Classification Phase.

A. Segmentation phase

U-Net and DeepLab V3+ are leading models for image segmentation, each with a distinct approach. U-Net's U-shaped architecture combines an encoder-decoder framework to capture fine details and maintain spatial accuracy at the pixel level. DeepLabV3+, part of Google's DeepLab series, enhances segmentation by applying Atrous Spatial Pyramid Pooling (ASPP) and dilated convolutions, allowing it to capture features across multiple scales for effective segmentation of objects in varied sizes.

This study employs a satellite imagery dataset of Dubai, provided by Humans in the Loop and captured by MBRSC satellites. Available under the CC0 1.0 license, this dataset was carefully annotated at the pixel level by Roia Foundation trainees in Syria. It includes 72 unique images divided across eight large tiles, each further split into nine smaller patches, covering six classes: Building, Land, Road, Vegetation, Water, and Unlabeled. For preprocessing, images are divided into 256 x 256 patches using Patchify, generating 145 patches per image and a total of 1,305 patches across all images and masks. Hexadecimal color codes are also converted to RGB by interpreting each pair of hex digits for precise color accuracy.

After data preprocessing, the U-Net and DeepLabv3+ models are constructed using multiple layers of neural networks for both the encoder and decoder structures. The models are subsequently trained and tested on satellite imagery from the dataset. For performance evaluation, we utilized accuracy and the Jaccard Coefficient (also known as Intersection over Union, IoU) as key metrics for analyzing the results of the project. Both the models gave an accuracy of 86% on the test data and the mean IoU obtained was 60% which means that, on average, 60% of the pixels in the predicted segmentation overlap with the pixels in the ground truth segmentation for each class. Higher IoU values indicate better performance, with 100% being a perfect overlap.

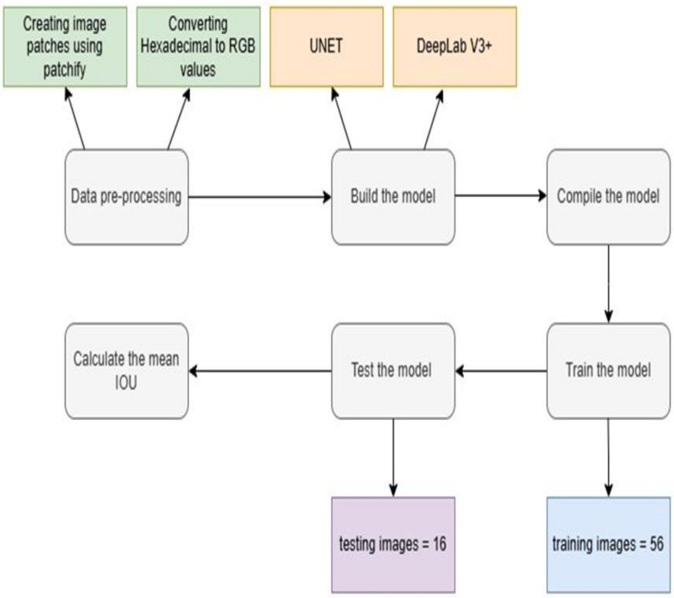

Fig. 2. Illustrates the workflow of the models

B. CLASSIFICATION PHASE

This project makes use of a labeled dataset from Quoc Dung Cao and Youngjun Choe at the University of Washington's Disaster Data Science Lab, consisting of aerial images of buildings in Houston and surrounding areas after Hurricane Harvey in August 2017. Each image is annotated to indicate whether the building was damaged. The dataset is divided into training, validation, and test sets, with 10,000 images in the training set (5,000 damaged, 5,000 undamaged) and 2,000 images in both the validation and test sets (1,000 of each label). The filenames of the images contain latitude and longitude coordinates, which allow for plotting the buildings' locations in relation to each other.

Prior to training our neural network, we performed an extensive data exploration to uncover patterns within the dataset. We started by visualizing example images from each class to establish a common visual reference. Following this, we analyzed the dataset's characteristics by generating summary images to visualize metrics such as mean, standard deviation, and contrast across the classes.

Fig. 3. Represents sample images from the dataset

Fig. 3. shows a comparison of three randomly selected images from both damaged and undamaged areas highlights several key patterns. Many of the damaged area images feature floodwaters surrounding the structures, resulting in noticeable changes in ground texture and color, though this is not consistent across all cases. Some images of damaged areas also display small objects scattered on the ground, a pattern that can also be seen in some undamaged images. Images of undamaged structures are more likely to show clear ground or pools of bluish water, free of floodwaters. In many instances, distinguishing between damaged and undamaged areas proves challenging for the human eye.

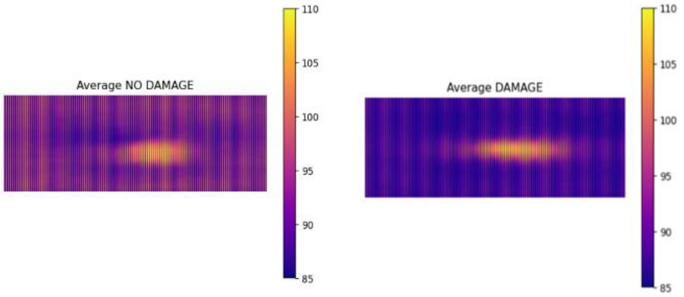

Fig. 4. Represents mean value by class of damage and no damage images

Fig. 4. depicts the mean pixel values for each class, damage and no damage, revealing distinct spatial patterns across images. Structures generally appear centrally located in both categories. However in damage images, the surrounding pixel values are lower compared to those in no damage images. This difference may be attributed to floodwaters around damaged structures, which introduce darker hues, unlike the brighter tones of dry, undisturbed ground in no damage images.

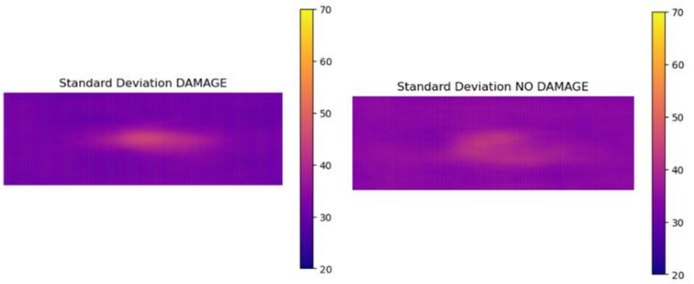

Fig. 5. Represents standard deviation value by class of damage and no damage images

Images in the Fig. 5. are comparable to those shown earlier, but they represent each pixel's standard deviation by class rather than the mean pixel value. Notably, the no-damage images exhibit a slightly higher standard deviation along the edges. This increased variability may result from the visible ground surrounding the structures, adding more fluctuation between images within the no-damage category.

To further understand the data, we applied principal component analysis (PCA) to identify the underlying patterns and relationships. The resulting eigen images are not immediately recognizable to the human eye, especially compared to those based on more regular patterns like faces. Instead, the eigen images here suggest basic geometric shapes with right angles, emphasizing only the primary patterns. We retained 70% of the variation in the data to focus on the most significant components.

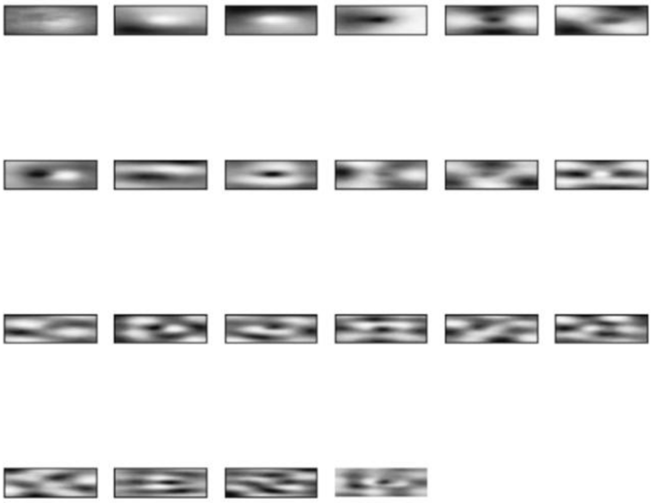

Fig. 6. Eigen images depicting structures with damage

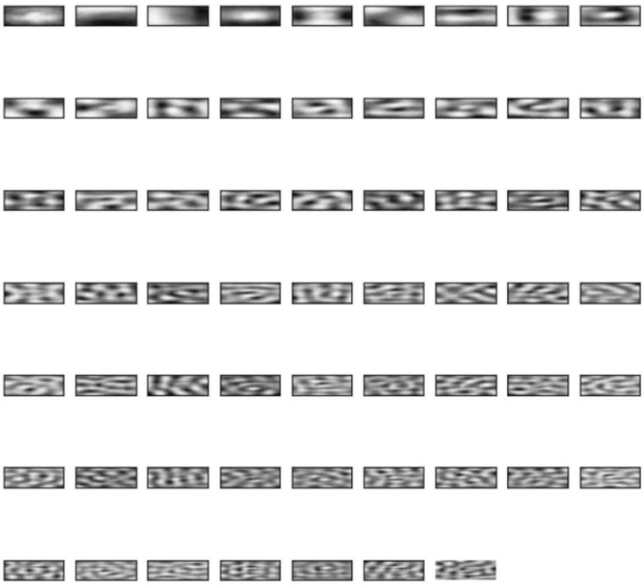

Fig. 7. Eigen images depicting structures with no damage

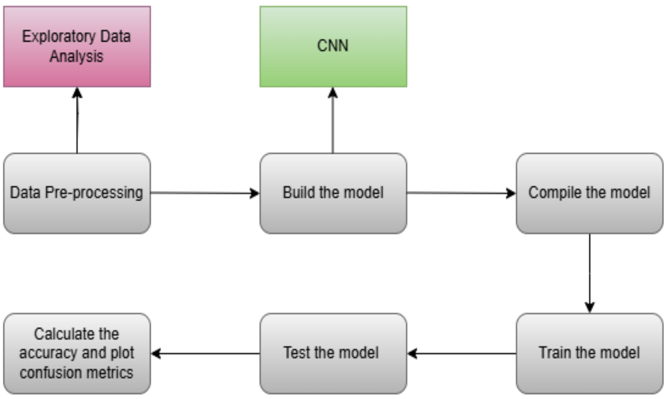

Fig. 8. Illustrates the workflow of classification model

Fig. 5.8 outlines the steps involved in building and evaluating a Convolutional Neural Network (CNN) for image classification. Following an initial Exploratory Data Analysis, the CNN model is created and compiled, specifying the optimizer, loss function, and evaluation metrics. The model is trained on the dataset, then tested to gauge its performance. Finally, the accuracy and confusion matrix are calculated and displayed to assess the model's classification performance. This structured approach ensures a comprehensive process for training and evaluating the CNN model. The model achieved an accuracy of 97% on the test data, with accuracy and the confusion matrix used as the key evaluation metrics.

V. Result Analysis

The color palette is shown in Fig. 9. with six colors and their corresponding RGB values. Each color is named with labels like Building, Land, Road and many more, suggesting that this palette might be used to visually represent different land cover types or objects in a geographic dataset. This could be useful for creating color-coded maps or visualizations in fields like remote sensing, GIS, or urban planning.

Name	R	G	B	Color
Building	60	16	152	
Land	132	41	246	
Road	110	193	228	
Vegetation	254	221	58	
Water	226	169	41	
Unlabeled	155	155	155	

Fig. 9. Shows the color palette for the geographical data

The following images present the results of the U-Net model for segmentation phase.

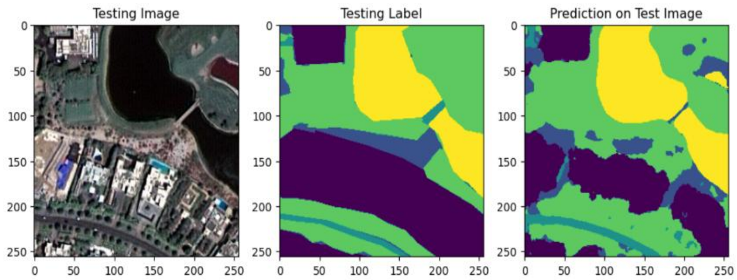

Fig. 10. Illustrates the results of U-Net segmentation model

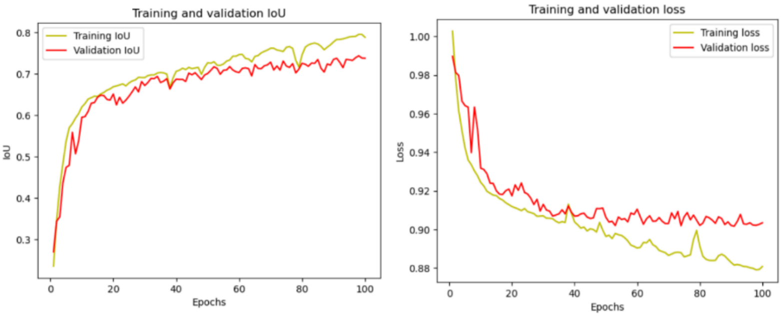

Fig. 11. Illustrates the graphs for IoU and loss metrics of U-Net segmentation model

Epochs	Loss	Accuracy	Val_loss	Val_accuracy
5	0.9454	0.7561	0.9641	0.6880
25	0.9092	0.8462	0.9183	0.8182
50	0.8970	0.8685	0.9063	0.8525
75	0.8870	0.8902	0.9065	0.8592
100	0.8827	0.9026	0.9034	0.8686

Table. 1. Illustrates U-Net's model performance over time

The following images present the results of the DeepLab V3+ model for segmentation phase

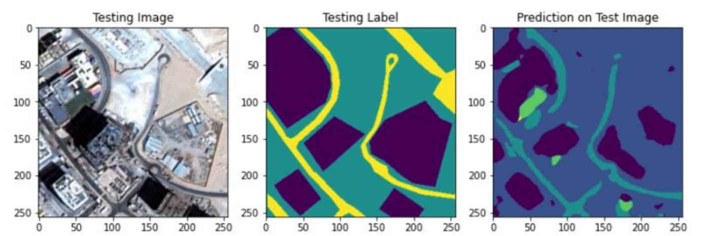

Fig. 12. Illustrates the results of DeepLab V3+ segmentation model

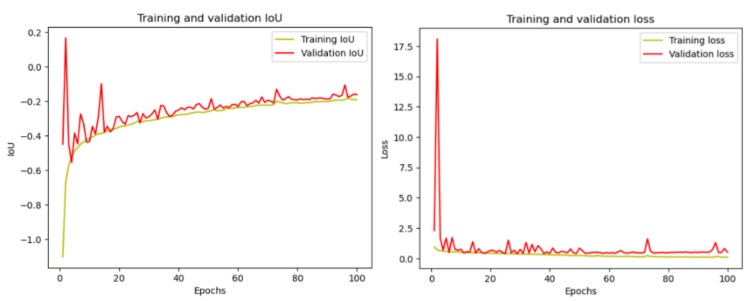

Fig. 13. Illustrates the graphs for IoU and loss metrics of DeepLab V3+ segmentation model

Epochs	Loss	Accuracy	Val_loss	Val_accuracy
5	0.5618	0.8146	0.1688	0.2813
25	0.3993	0.8588	0.4133	0.8580
50	0.2213	0.9198	0.8549	0.6801
75	0.1683	0.9364	0.4506	0.8688
100	0.1040	0.9593	0.5218	0.8638

Table. 2. Illustrates DeepLab V3+ model performance over time

The following images present the results of the CNN model for classification phase.

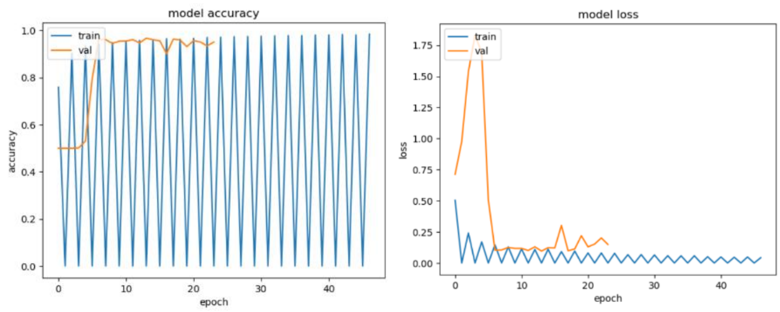

Fig. 14. Illustrates the graphs for accuracy and loss metrics of CNN model

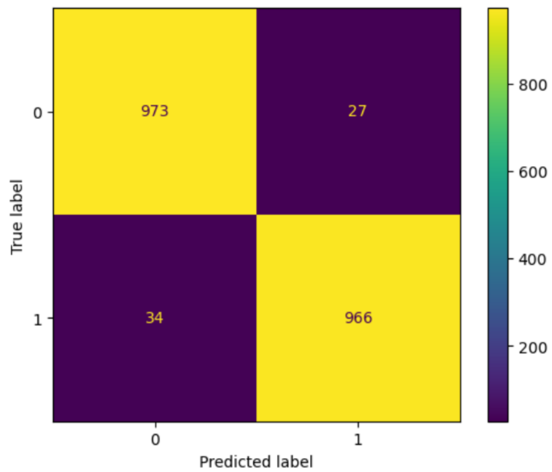

Fig. 15. Represents the confusion matrix

This confusion matrix provides an overview of the model's performance in classifying images as either damage (1) or no damage (0). The model correctly identified no damage in 973 cases and damage in 966 cases, showing a strong ability to make accurate predictions. However, there were some misclassifications: 27 cases were incorrectly predicted as damage when there was actually no damage (false positives), and 34 cases were incorrectly predicted as no damage when there was actually damage (false negatives).

V. Conclusion

The research emphasizes the vital role of satellite imagery segmentation in advancing both land cover analysis and disaster resilience. Through precise classification of diverse land features including buildings, vegetation, water bodies, and infrastructure, our model generates high-resolution insights that are invaluable for environmental monitoring, urban development, and resource management. By providing stakeholders with an accurate and detailed view of land cover types, our approach promotes informed, data-driven decisions that support sustainable development and foster responsible interaction with natural landscapes.

In addition, our work demonstrates the power of applying these segmentation techniques to post-disaster scenarios, where the model accurately detects disaster-specific features such as structural damage, debris fields, and flood zones. This rapid identification and mapping of affected areas provide an essential resource for disaster response teams, enabling faster, more efficient resource allocation and streamlined emergency operations. The detailed, high-precision assessments generated from post-disaster aerial images enhance both the speed and accuracy of response, ensuring that relief efforts are focused where they are needed most.

This integrated framework not only addresses the immediate requirements following a disaster but also lays the groundwork for building long-term resilience by informing sustainable urban and environmental planning. By seamlessly transitioning from routine land cover analysis to high-stakes disaster assessment, our approach offers a versatile tool that adapts to a range of geospatial challenges. As climate change intensifies the frequency and scale of environmental disruptions, our project provides a forward-looking solution that enhances preparedness and promotes resilient, sustainable communities—positioning it as an essential resource in navigating complex environmental challenges in the years ahead.

Future directions for research in satellite imagery segmentation for land cover and disaster assessment could include integrating diverse data sources, such as LiDAR and ground surveys, to improve segmentation accuracy and model robustness. The development of real-time processing algorithms would allow for immediate insights during disaster scenarios, facilitating quicker, more effective responses and better resource management. Further advancements in deep learning could refine classification results, enhancing the precision of feature detection. Implementing automated change detection systems would support ongoing monitoring of land cover and provide early warnings for areas vulnerable to disasters. In addition, developing intuitive, user-friendly platforms would enable urban planners and disaster responders to easily utilize the insights for better decision-making. Field validation studies would be essential for confirming model accuracy in real-world conditions, while long-term tracking of climate change impacts could inform proactive policy and help shape adaptive strategies for future environmental challenges.

References

[1]. B. Benjdira, A. Ammar, A. Koubaa, and K. Ouni, “Data-Efficient Domain Adaptation for Semantic Segmentation of Aerial Imagery Using Generative Adversarial Networks,” Appl. Sci., vol. 10, no. 3, p. 1092, 2020, doi: 10.3390/app10031092.

[2]. T. K. Behera, S. Bakshi, M. Nappi and P. K. Sa, "Superpixel-Based Multiscale CNN Approach Toward Multiclass Object Segmentation From UAV-Captured Aerial Images," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 16, pp. 1771-1784, 2023, doi: 10.1109/JSTARS.2023.3239119.

[3]. D. Jozi, N. Shirzad-Ghaleroudkhani, G. Luhadia, S. Abtahi, and M. Gül, "Rapid post-disaster assessment of residential buildings using Unmanned Aerial Vehicles," *International Journal of Disaster Risk Reduction*, vol. 111, p. 104707, 2024, doi: 10.1016/j.ijdrr.2024.104707.

[4]. M. Thakkar, R. Vanzara, and A. Patel, “Semantic segmentation approaches for crop classification with multi-altitude Google Earth imagery”, J Integr Sci Technol, vol. 12, no. 6, p. 832, Jun. 2024, doi: 10.62110/sciencein.jist.2024.v12.832.

[5]. W. Zhang, P. Tang and L. Zhao, "Fast and accurate land-cover classification on medium-resolution remote-sensing images using segmentation models", Int. J. Remote Sens., vol. 42, no. 9, pp. 3277-3301, May 2021.

[6]. C. Kyrkou and T. Theocharides, "EmergencyNet: Efficient Aerial Image Classification for Drone-Based Emergency Monitoring Using Atrous Convolutional Feature Fusion," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 13, pp. 1687-1699, 2020, doi: 10.1109/JSTARS.2020.2969809.

[7]. Md A. Islam, S. I. Rashid, N. U. I. Hossain, R. Fleming, and A. Sokolov, "An integrated convolutional neural network and sorting algorithm for image classification for efficient flood disaster management," Decision Analytics Journal, vol. 7, p. 100225, 2023. doi: 10.1016/j.dajour.2023.100225.

[8]. Yang, N., & Tang, H. (2021). Semantic Segmentation of Satellite Images: A Deep Learning Approach Integrated with Geospatial Hash Codes. Remote Sensing, 13(14), 2723. DOI: 10.3390/rs13142723

[9]. K. Yuan, X. Zhuang, G. Schaefer, J. Feng, L. Guan, and H. Fang, "Deep-Learning-Based Multispectral Satellite Image Segmentation for Water Body Detection," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 7422-7434, 2021. DOI: 10.1109/JSTARS.2021.3098678.

[10]. Ananya Gupta, Simon Watson and Hujun Yin "Deep learning-based aerial image segmentation with open data for disaster impact assessment", ScienceDirect, Volume 439,pp.22-33, doi: 10.1016/j.neucom.2020.02.139.

[11]. I. Ahmed, M. Ahmad, G. Jeon and A. Chehri, "An Internet of Things and AI-Powered Framework for Long-Term Flood Risk Evaluation," in IEEE Internet of Things Journal, vol. 11, no. 3, pp. 3812-3819, 1 Feb.1, 2024, doi: 10.1109/JIOT.2023.3308564.

[12]. A. A. Deshmukh, S. D. B. Sonar, R. V. Ingole, R. Agrawal, C. Dhule and N. C. Morris, "Satellite Image Segmentation for Forest Fire Risk Detection using Gaussian Mixture Models," 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 2023, pp. 806-811, doi: 10.1109/ICAAIC56838.2023.10140399.

[13]. L. He, J. Shan and D. Aliaga, "Generative Building Feature Estimation From Satellite Images," in IEEE Transactions on Geoscience and Remote Sensing, vol. 61, pp. 1-13, 2023, Art no. 4700613, doi: 10.1109/TGRS.2023.3242284.

[14]. K. Kaku, "Satellite remote sensing for disaster management support: A holistic and staged approach based on case studies in Sentinel Asia," International Journal of Disaster Risk Reduction, vol. 33, pp. 417-432, 2019, doi: 10.1016/j.ijdrr.2018.09.015.