International Journal Of Innovative Trends In Multi Disciplinary Scientific Research
ISSN (Online): 3139 – 2687
(www.ijtmsr.co.in) Volume 9, Issue 6, 2026

Dual-Network Deep Learning with Grad-CAM for Post-Flood Aerial Image Classification

Punyamanthula Bhagya Madhuri
Assistant Professor, Department of CSE (AI&ML)
NRI Institute of Technology
Agiripalli, A.P, India
bhagyamadhuripunyamanthula@gmail.com

Dangudubiyyam Yagnesh Lakshman Sai
Student, Department of CSE (AI&ML)
NRI Institute of Technology
Agiripalli, A.P, India

Kanukurthi Pujitha
Student, Department of CSE(AI&ML)
NRI Institute of Technology
Agiripalli, A.P, India

Jannu Varun Kumar
Student, Department of CSE (AI&ML)
NRI Institute of Technology
Agiripalli, A.P, India

Puvvaladasu Arun Kumar
Student, Department of CSE (AI&ML)
NRI Institute of Technology
Agiripalli, A.P, India

Abstract-Flood is one of the most destructive forms of natural disasters that usually cause a rampage destruction of human life and infrastructure. The use of aerial imagery is one of the vital instruments in measuring the damage caused by floods, since it allows identifying the territories and the radius of destruction. This paper gives novelty in classifying aerial images after floods with the help of latest deep learning models, i.e., MobileNet and DenseNet, to be able to classify different regions inside the images. The suggested models are trying to define the regions in six different forms, namely, buildings, flooded zones, forests, mountains, seas, and streets. In order to promote transparency and trustworthiness of the models, we also make use of Explainable Artificial Intelligence (XAI) models, namely, Class Activation Maps (CAMs) produced by Grad-CAM, which can be used to interpret the model decision-making process better. The results prove that the deep learning models proved efficient to identify the flooded areas among other environmental characteristics and to be able to assess the disaster and recovery process after floods. Furthermore, the models provide an increased degree of trust on their prediction, due to the interpretability included, with the help of Grad-CAM, essential when disaster response and recovery planning should be carried out in real-time. Effectiveness of the suggested strategy has been measured through different indicators, and that also shows its potential in the development of AI-based solutions in disaster management and post-disaster recovery. This study can be added to the increasing body of literature that tries to leverage AI technologies to enhance the recovery process of floods.

Keywords: Aerial Image Classification, Post-Disaster Geospatial Analysis, MobileNet, DenseNet, Explainable Artificial Intelligence (XAI), Grad-CAM.

I. INTRODUCTION

Flooding is a catastrophic natural disaster that severely damages infrastructure, the environment, and human life. Conventional ground-based damage assessment methods are time-consuming, resource-intensive, and might not be able to survey large areas quickly. In order to assess the post-flood conditions, aerial data (or UAVs and satellites) has proven to be a crucial tool. These images provide a complete picture of the destruction, allowing for quicker and more accurate assessments. However, successfully classifying and interpreting aerial imagery and the massive amounts of data it produces is a challenge.

.

MobileNet and DenseNet are the most widely used convolutional neural networks (CNNs) that have demonstrated remarkable efficacy in image classification, including disaster management. Because these models can automatically extract hierarchical features from raw image data, they have been very successful in classifying the flood-affected areas.

Deep learning models are regarded as "black boxes" despite their accuracy, which undermines trust, especially during crucial humanitarian response. The decision-making process can be made more transparent rather than using Explainable Artificial Intelligence (XAI) like Grad-CAM (Gradient-weighted Class Activation Mapping). It generates heatmaps, which are significant regions of an image that facilitate the interpretation of the model's associations.

In order to classify post-flood aerial photos, this paper proposes using MobileNet, DenseNet, and Grad-CAM. Its goal is to create an automated system that can recognize buildings, streets, forests, and other areas impacted by flooding in order to evaluate the damage and facilitate recovery. Explainable AI is incorporated to improve the system's dependability and transparency, which helps with disaster recovery decision-making.

The structure of the paper is as follows: The review of disaster management and earlier research on aerial image classification is provided in Section II. Section III discusses the data steps and preprocessing. The methodology, including XAI and deep learning models, is presented in Section IV. The experiment's setups and results are in Section V, and Section VI will discuss them. The paper's conclusion and suggestions for future research are found in Section VII.

A. *Contributions of the Paper*

The following is a summary of this study's primary contributions. In order to identify buildings, flooded areas,

forests, mountains, seas, and streets using deep learning techniques, a six-class post-flood aerial image classification framework is first proposed. Second, the accuracy and computational efficiency of a deep CNN (DenseNet) and a lightweight CNN (MobileNet) are compared. Third, Grad-CAM is used to integrate Explainable Artificial Intelligence, which improves decision-making transparency and trust by making model predictions visually interpretable. Lastly, the suggested system's efficacy for post-flood disaster assessment is demonstrated by its up to 96% classification accuracy on the validation dataset.

B. *Scope Of The Study*

Using MobileNet and DenseNet, this project aims to develop a deep learning-based system for automated post-flood aerial imagery classification. The system uses high-resolution aerial photos to identify six key land-type categories: buildings, flooded areas, forests, mountains, seas, and streets. In order to give visual support for model predictions, the study also uses Explainable AI techniques, particularly Grad-CAM. The scope is restricted to the examination of the provided dataset and the creation of a classification framework that facilitates quick post-flood damage assessment and helps disaster response teams organize recovery efforts.

C. *Problem statement*

Due to the extensive geographic coverage and the pressing need for decision-making, post-flood damage assessment is difficult. Inspection of impacted areas by hand is labor-intensive, slow, and inefficient when applied to large areas. It is still challenging to differentiate flooded areas from other visually similar land types, even though aerial imagery provides a more comprehensive view of the disaster. Although deep learning models can automate this process, they frequently act as opaque, unintelligible black boxes. An automated, precise, and comprehensible system that can categorize aerial photos taken after floods and offer clear visual explanations is therefore required. In order to solve this issue, this study uses Grad-CAM in conjunction with MobileNet and DenseNet models to provide disaster management professionals with outputs that are both interpretable and have high classification accuracy.

II. LITERATURE SURVEY

A. *Aerial Image Classification for Post-Flood Assessment*

Objective: The goal is to effectively assess disasters by classifying post-flood aerial photos into meaningful land categories, such as streets, buildings, forests, and flooded areas.

Key Studies:

Using MobileNet, DenseNet, and Grad-CAM, Manaf et al. (2025) proposed a robust deep learning system to classify post-flood scenes with enhanced interpretability.

Hussain et al. (2024) used deep learning on UAV imagery to quickly detect floods, demonstrating the importance of drone-based systems for real-time disaster assessment.

By using ensemble learning models to distinguish between flooded and non-flooded areas, Yasi et al. (2024) showed increased accuracy and robustness in noisy environments.

Key Takeaway:

While integrating XAI improves transparency and trust in disaster analysis systems, deep learning models greatly

improve the identification of flood-impacted areas from aerial images.

B. *Deep Learning Models (MobileNet & DenseNet) in Remote Sensing*

Objective: The goal is to classify land-cover and disaster-related images with high accuracy using lightweight and dense CNN architectures.

Key Studies:

Bi et al. (2019) showed how DenseNet's dense connectivity and effective feature reuse in aerial imagery tasks contributed to its excellent performance in scene classification.

In order to validate CNN's effectiveness in aerial classification, Kyrkou & Theocharides (2021) introduced EmergencyNet, which uses atrous convolution for real-time drone-based emergency monitoring.

Lei et al. (2023) used deep learning to detect multiple objects in drone images in real-time, facilitating quick post-disaster evaluation.

Key Takeaway:

Large-scale post-flood image classification and real-time disaster response can benefit from models like MobileNet and DenseNet, which offer high accuracy with efficient computation.

C. *Explainable Artificial Intelligence (XAI) for Disaster Interpretation*

Objective: To highlight significant areas in aerial photos used for classification in order to increase the transparency of deep learning predictions.

Key Studies:

Kadiyala & Woo (2021): Applied Grad-CAM to expose decision-critical image regions for flood prediction tasks, improving interpretability during emergencies.

Manaf et al. (2025): Integrated Grad-CAM into MobileNet and DenseNet models to visualize flood damage cues such as submerged structures and water spread patterns.

Key Takeaway:

By providing visual justification for classification decisions, XAI techniques such as Grad-CAM boost confidence in AI-driven flood analysis systems.

D. *Drone and Satellite-Based Flood Monitoring Systems*

Objective: To improve post-disaster evaluation by employing satellite imagery and high-resolution UAVs for extensive surveillance.

Key Studies:

Duarte et al. (2018): Leveraged satellite and airborne imagery to classify building damage, demonstrating the value of high-resolution remote sensing in post-event evaluation.

Whitehurst et al. (2022): Used drone data to quantify flood damage and support restoration planning with detailed spatial insights.

Zhang & Zhu (2023): Reviewed UAV platforms and imaging sensors, emphasizing their role in flood detection, mapping, and environmental monitoring.

Key Takeaway:

The accuracy and speed of flood damage assessment are significantly improved by the high-resolution data provided by drone and satellite imagery.

E. *AI Techniques for Natural Disaster Management*

Objective: To automate disaster detection and recovery planning using artificial intelligence techniques, such as CNNs, ensemble models, and hybrid frameworks.

Key Studies:

Akhyar et al. (2024): Reviewed deep AI techniques used across natural disaster management, highlighting classifications, early warnings, and post-disaster analytics.

Boucher (2017): Studied the effect of satellite imagery resolution on land-use classification accuracy, emphasizing the importance of high-quality spatial data for flood mapping.

Mukherjee (2022): Explored AI-driven Earth-observation systems to map and detect floods, identifying future improvements for large-scale disaster predictions.

Key Takeaway:

The automation, precision, and efficiency of disaster detection systems are greatly improved by contemporary AI techniques, facilitating quick decision-making in post-flood situations.

Research Gap:

Even though previous research has shown that deep learning is useful for detecting floods from aerial photos, there are still a number of drawbacks. The majority of earlier studies concentrate on the binary classification of flooded versus non-flooded areas, which limits their applicability in thorough post-disaster evaluation. Additionally, many approaches lack explainability, which makes deep learning models operate as "black boxes" and undermines confidence in critical disaster response scenarios. There hasn't been much focus on comparing lightweight models with deeper architectures to achieve a balance between computational efficiency and accuracy. To address these gaps, this study proposes a framework that combines Grad-CAM with deep learning models to enable interpretable decision-making.

III. RELATED WORK

Deep learning's capacity to extract intricate spatial features and automate extensive analysis has made it a popular method for post-flood damage assessment using aerial imagery. While more recent research has moved toward multi-class classification and explainability-based decision support, earlier studies concentrated on binary flood detection.

The usefulness of UAV imagery in conjunction with CNN models for quick post-flood assessment was shown by Hussain et al. [1], allowing for real-time aerial data monitoring. The robustness of flood classification under noisy and heterogeneous conditions was enhanced by ensemble-based methods like those by Yasi et al. [2]. The ability to differentiate visually similar land regions was further improved by texture-enhanced CNN models, such as the binary pattern encoded framework put forth by Anwer et al. [3].

The necessity of interpretable AI systems in disaster response has been highlighted by recent research. In order to visually support predictions, Manaf et al. [4] combined Grad-CAM with MobileNet and DenseNet, emphasizing submerged structures and water-spread patterns in post-flood imagery. Studies by Duarte et al. [5] and Whitehurst et al. [6] demonstrated that drone and satellite imagery greatly increase the accuracy of large-area damage mapping. High-resolution remote sensing has also been extensively investigated. The expanding role of AI-supported Earth-observation systems for large-scale flood detection and monitoring was confirmed by more comprehensive reviews by Mukherjee [7] and Zhang & Zhu [8].

Overall, previous studies show that integrating XAI techniques improves trust, transparency, and usability in disaster management workflows, and that deep learning combined with high-resolution aerial imagery offers powerful capabilities for flood classification.

IV. PROPOSED SYSTEM

An automated framework for post-flood aerial image classification is offered by the suggested system. Aerial image acquisition is the first step in the workflow, which is followed by simple preprocessing to standardize the input data. Two deep learning models, MobileNet and DenseNet, are then used to analyze the processed images and classify each one into one of six land categories. Grad-CAM is used to produce visual explanations that highlight the regions influencing each prediction in order to improve interpretability. Lastly, efficient post-flood damage assessment and disaster response planning are supported by the categorized outputs and related visualizations.

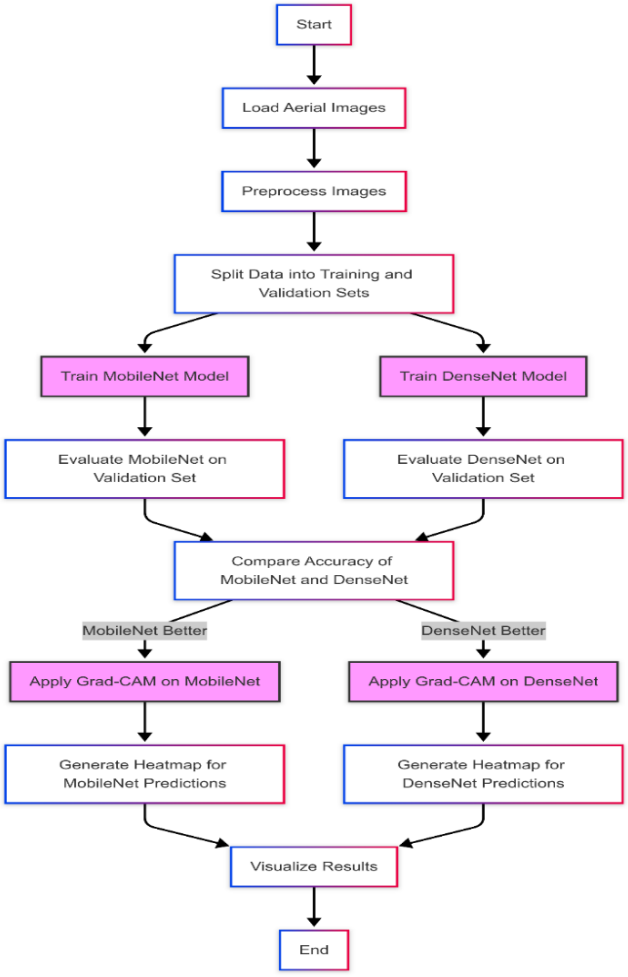


Figure 1 illustrates the overall workflow of the proposed post-flood aerial image classification system.

V. METHODOLOGY

A. *Loading Dataset*

In order to train and be exponentially validated, the system will first upload a dataset of labeled aerial photos. Buildings, flooded areas, forests, mountains, seas, and streets are among the categories into which the images are classified. The training data and the validation of the model's performance on the previously unidentified data comprise the two primary subsets of the data. The sample size consists of 900 images set aside for validation and 3600 images for training. In order to assess the model's capacity for generalization, this division ensures that it is trained on one set of data and tested on another. Each class is given a label so that they can be efficiently categorized:

- Building: 0
- Flooded: 1
- Forest: 2
- Mountains: 3
- Sea: 4
- Street: 5

During the classification process, this class to index mapping helps the model accurately label the target. The dataset is preprocessed to make it compatible with DenseNet and MobileNet models after it has been loaded. The computational efficiency and speed of model training are significantly increased by using a CUDA-enabled GPU.

Table 1 Dataset Distribution

Dataset Information	**Training Set**	**Validation Set**
Total Images	3600	900
building	600	150
flooded	600	150
forest	600	150
mountains	600	150
sea	600	150
street	600	150

B. *Model Training and Classification*

The system trains two deep learning models—MobileNet and DenseNet—to classify post-flood aerial images into six categories. Both models learn to identify buildings, flooded areas, forests, mountains, seas, and streets from the processed dataset. Grad-CAM is integrated to generate heatmaps that highlight the regions influencing each prediction, providing clear interpretability and helping users understand the model's decision-making process.

Model performance is evaluated using accuracy, precision, recall, F1-score, and confusion matrices, offering a comprehensive assessment of classification quality. The final comparison between MobileNet and DenseNet considers accuracy, computational efficiency, and robustness to determine the most suitable architecture for post-flood damage assessment.

Model Architecture Choice

Because of its lightweight architecture, MobileNet is chosen for real-time or resource-constrained applications, like drone-based flood assessment. It maintains competitive accuracy while drastically lowering computational costs through the use of depthwise separable convolutions.

Dense connectivity, which allows each layer to receive input from all preceding layers, is what makes DenseNet so effective at extracting features. This effective feature reuse enhances performance on complex aerial imagery and improves gradient flow, especially when identifying subtle land type differences in post-flood scenes.

Training Process and Data Augmentation:

The models are trained on a dataset of 3600 labeled images and evaluated on 900 validation images across six classes: buildings, flooded areas, forests, mountains, seas, and streets. Data augmentation techniques—includingrandom rotations, flips, zooms, and shifts—are applied to increase dataset variability and reduce overfitting.

Training is performed using the Adam optimizer and categorical cross-entropy loss, with model performance monitored after each epoch. Validation accuracy and loss trends are used to assess the learning stability and generalization of MobileNet and DenseNet throughout the training process.

C. Mobilenet

MobileNet is an efficient deep learning network designed specifically for mobile and embedded devices that require high-speed image classification with minimal computational capacity. Its lightweight architecture and efficient parameter usage make it ideal for applications where speed and computational efficiency are essential. This reduction in complexity makes MobileNet capable of processing more images in less time, making it well suited for real-time tasks such as flood-damage assessment using aerial imagery.

MobileNet Architecture

MobileNet can be subdivided into two main architectural components: depthwise convolution and pointwise convolution.

Depthwise Convolution

A conventional convolution applies a filter across all input channels, but MobileNet replaces this with depthwise convolution, where a single filter is applied to each input channel independently. This drastically reduces computation. Mathematically, depthwise convolution is expressed as:

where Kd is the depth-wise kernel running on channel k and X is the input feature map. A standard convolution has the

$$Yd(i,j,k) = m,n\sum X(i+m, j+n, k) \cdot Kd(m,n,k) \quad (1)$$

following computational cost:

$$C_{standard} = D_K^2 M\ N\ D_F^2$$

while depthwise convolution reduces this to:

$$C_{depthwise} = D_K^2 M D_F^2$$

showing a significant reduction in operations.

Pointwise Convolution

After the depthwise convolution, MobileNet applies a pointwise (1×1) convolution to combine the outputs of all channels and learn cross-channel interactions. It is defined as:

$$Yp(i,j,n) = k\sum Yd(i,j,k) \cdot Kp(1,1,k,n) \quad (2)$$

The total computational cost of MobileNet's depthwise separable convolution becomes:

$$C_{DSC} = C_{depthwise} + C_{pointwise} = D_K^2 M D_F^2 + M N D_F^2$$

which is nearly 8–9× more efficient than standard convolution. This architectural design enhances MobileNet's ability to deliver fast and accurate aerial image classification, even on low-resource devices

Model Training

In order to train MobileNet for post-flood aerial image classification, pre-processed images labeled into six classes—building, flooded area, forest, mountains, sea, and street—are first fed into the model. To enhance generalization and avoid overfitting, these images are resized, normalized, and enhanced by arbitrary rotations, flips, zooms, and shifts. By exposing the model to a greater range of real-world variations, augmentation ensures robustness when classifying various flood scenarios. Beginning with basic edges and textures and working its way up to more intricate shapes and structural patterns pertinent to the identification of post-flood damage, MobileNet learns hierarchical visual features from the input images.

The multi-class categorical cross-entropy loss function, which is ideal for classification across several categories, and the Adam optimizer are used to train the model.

$$\mathcal{L} = -\sum_{c=1}^{6} y_c \, log(\hat{y}_c) \quad (3)$$

Where $\boldsymbol{yc}$ represents the true label and $\boldsymbol{y^c}$ is the predicted probability for class $\boldsymbol{c}$. Adam updates the model weights using the moment estimates

$$m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t, v_t = \beta_2 v_{t-1} + (1-\beta_2) g_t^2, \quad (4)$$

and the parameter update rule

$$\theta_{t+1} = \theta_t - \alpha \frac{m_t}{\sqrt{v_t} + \epsilon} \quad (5)$$

Training is performed for several epochs, allowing the model to refine its parameters and achieve strong generalization on unseen validation samples.

Feature Extraction and Classification

Because MobileNet uses depthwise separable convolutions, which allow the model to capture rich spatial patterns while retaining computational efficiency, it is an effective feature extractor. In order to distinguish between buildings, floodwater areas, forest vegetation patterns, rocky mountain textures, and the smooth surfaces of streets and seas, MobileNet learns discriminative features as aerial images flow through the network. The fully connected layers that carry out the final classification receive these extracted features.

Class probabilities are calculated by the output layer using the softmax activation function, which is defined as

$$\hat{y}_c = \frac{e^{z_c}}{\sum_{k=1}^{6} e^{z_k}} \quad (6)$$

where the logit for class c is represented by z_c.Accurate classification of aerial imagery and efficient interpretation of post-flood damage are made possible by the model's selection of the class with the highest probability.

Interpretability using Grad-CAM

MobileNet incorporates Grad-CAM (Gradient-weighted Class Activation Mapping) to provide interpretability for its predictions. Grad-CAM highlights the regions of an image that contribute most strongly to the model's output, enabling

disaster management teams to understand how and why the model classified a region as flooded or non-flooded. Grad-CAM computes importance weights using

$$\alpha_k^c = \frac{1}{z} \sum_{i,j} \frac{\partial y^c}{\partial A_{ij}^k} \quad (7)$$

and generates the final heatmap using

$$L_{GradCAM}^c = \mathrm{ReLU}\left(\sum_k \alpha_k^c A^k \right) \quad (8)$$

These visual explanations provide transparency and increase trust in automated flood assessment systems by revealing the exact image regions influencing the classification.

The performance of MobileNet is compared with DenseNet to examine the balance between classification accuracy, computational efficiency, and interpretability. This comparison ensures the selection of the most suitable model for post-flood aerial image analysis, taking into account the trade-offs between speed and predictive performance.

D. DenseNet

The dense block, which is the central part of DenseNet, is made up of several convolutional layers, each of which receives as input the feature maps of all layers that came before it rather than just the output of the layer that came right before it. This creates a highly connected structure in which the features learned earlier in the network are directly reused in later layers, eliminating redundant feature extraction and significantly improving the richness and efficiency of

representation learning. Mathematically, the output of the l-th layer in a dense block is expressed as

$$xl = Hl([x0, x1, x2, \dots, xl - 1]) \tag{9}$$

where $[x_0, x_1, \dots]$denotes the concatenation of feature maps from all preceding layers and H_l represents the composite transformation consisting of Batch Normalization (BN), ReLU activation, and convolution. Following each dense block, a transition layer is applied to control the size of the feature maps and maintain computational efficiency. This layer uses a 1×1convolution to reduce dimensionality and a 2×2average or max pooling operation to downsample the spatial resolution. These transitions prevent excessive growth of feature maps and maintain manageable computational cost.

Model Architecture:

The DenseNet architecture is made up of several dense blocks with transition layers in between. The number of dense layers within each block varies among different configurations, such as DenseNet-121, DenseNet-169, and DenseNet-201, but the basic design principle is the same: each layer receives feature maps from all previous layers, allowing for stronger gradient flow and feature reuse. DenseNet creates increasingly richer representations by concatenating feature maps across layers, which improves its capacity to identify intricate patterns in images. DenseNet is particularly useful for difficult tasks like identifying minute variations in post-flood aerial photos because of its highly interconnected architecture, which enhances classification performance.

Model Training

For post-flood aerial image classification, DenseNet is trained on preprocessed images that contain complex structures such as submerged regions, buildings, vegetation patches, and street layouts. DenseNet's densely connected architecture allows it to identify subtle visual patterns present in these images by reusing features generated throughout the network. The model is optimized using the Adam optimizer, while the multi-class categorical cross-entropy loss function guides the learning process. Training is performed across several epochs, and a validation set is used to continuously monitor model performance, ensuring that DenseNet generalizes well and does not overfit to the training data.

Characteristics Extraction and Classification

Because of the dense connectivity between layers—each layer receives feature maps from all previous layers—DenseNet performs exceptionally well at hierarchical feature extraction. This makes it possible for the network to record both high-level structures like building outlines, water boundaries, and road markings as well as low-level patterns like edges and textures. A global average pooling layer reduces the spatial dimensions while maintaining the learned semantic information after feature extraction within the dense blocks. A fully connected layer receives the pooled feature vector and uses the softmax classifier to produce a probability distribution across the six classes. The definition of the softmax function is

$$\hat{y}_c = \frac{e^{z_c}}{\sum_{k=1}^{6} e^{z_k}} \tag{10}$$

with the predicted category corresponding to the class of highest probability. Through this process, DenseNet effectively distinguishes between different land types in post-flood imagery.

Interpretability using Grad-CAM

To make its classification results easier to understand, DenseNet incorporates Grad-CAM (Gradient-weighted Class Activation Mapping). Grad-CAM generates a heatmap that identifies the areas of the input image that had the biggest impact on the model's choice. The weights of importance are calculated as

$$\alpha_k^c = \frac{1}{Z} \sum_{i,j} \frac{\partial y^c}{\partial A_{ij}^k} \tag{11}$$

and the final Grad-CAM heatmap is generated using

$$L_{GradCAM}^c = \mathrm{ReLU}\left(\sum_k \alpha_k^c A^k \right) \tag{12}$$

These heatmaps enable disaster management teams to understand why specific regions were classified as flooded or non-flooded, improving transparency and trust in the model's predictions. Such interpretability is essential in critical decision-making scenarios during post-flood response operations.

Model Evaluation

Accuracy, precision, recall, F1-score, and confusion matrix analysis are used to assess DenseNet's classification performance. These metrics offer a thorough evaluation of the model's capacity to differentiate between the various land categories found in aerial flood imagery. Robust learning and dependable classification performance are indicated by strong evaluation across these metrics..

In addition to standalone evaluation, DenseNet's results are compared to those of MobileNet to determine the optimal balance between classification accuracy, computational efficiency, and interpretability. This comparison aids in determining which model is better suited for implementation in real-time post-flood assessment applications, where processing speed and predictive performance are crucial.

VI. Discussion and Results

Using a validation dataset of 900 post-flood aerial photos, the accuracy, precision, recall, and F1-score of the suggested MobileNet and DenseNet models were assessed. In all six land-cover categories, both models showed excellent classification capabilities; however, DenseNet's overall performance was slightly better. Table II provides a summary of the two models' quantitative comparison.

Table II. Model Performance Comparison

Model	Accuracy	Precision	Recall	F1-Score
DenseNet	0.96	0.96	0.96	0.96
MobileNet	0.95	0.95	0.95	0.95

A. *MobileNet*

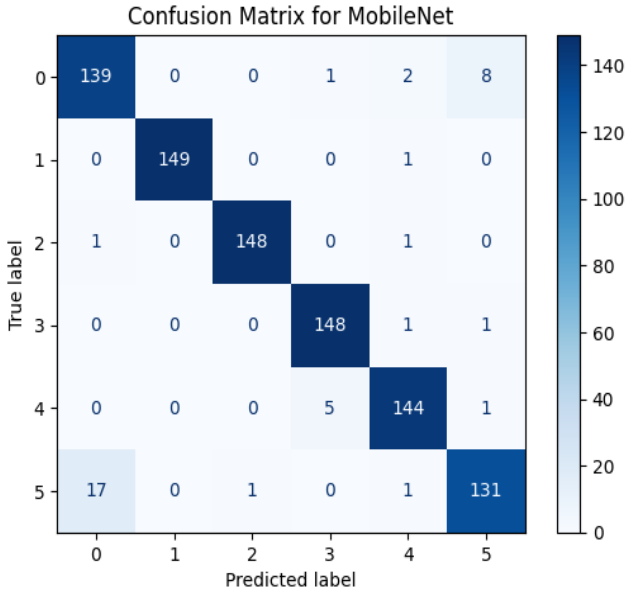


Figure 2. Confusion matrix for MobileNet showing class-wise classification performance on post-flood aerial images.

The MobileNet model's confusion matrix, which shows class-wise prediction performance for each of the six land categories, is shown in Figure 2. Strong correct-class predictions are indicated by the model's high diagonal values. MobileNet's efficacy in differentiating various terrain patterns found in post-flood aerial imagery is demonstrated by the high accuracy with which flooded areas, forests, mountains, and sea classes are identified.

According to the classification report, the overall accuracy is 0.95. The most consistently detected categories were those that represented flooded areas (class 1) and forests (class 2), which attained almost perfect precision and recall. Due to visual similarities between nearby built-up areas and streets, the street class (class 5) performs slightly worse, with precision and recall around 0.87. In spite of this, the weighted average metrics verify that MobileNet performs dependably and has good generalization for post-flood damage assessment.

B. *DenseNet*

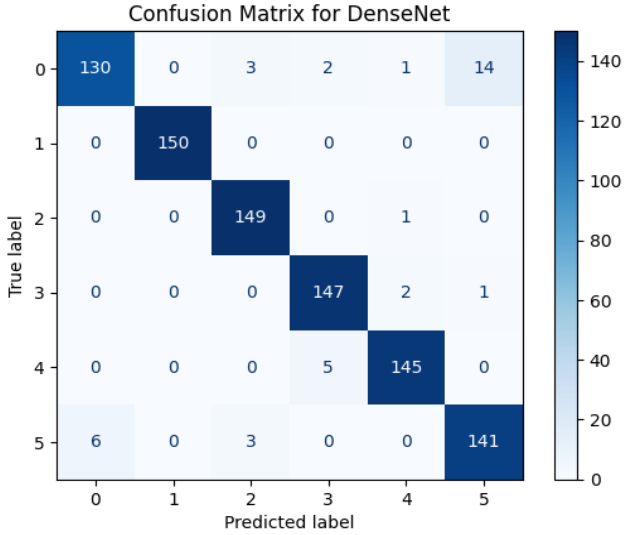


Figure 3. Confusion matrix for DenseNet illustrating high classification accuracy across all six land categories.

The DenseNet model's confusion matrix is shown in Figure 3. With an accuracy of 0.96 and consistently high precision and recall across all classes, DenseNet performs exceptionally well overall. Although there are a few minor misclassifications between buildings (class 0) and streets (class 5), which are visually similar in aerial imagery, the overall performance is not greatly impacted by these mistakes.

DenseNet successfully identifies every flooded region with a recall of 1.00 for the flooded class (class 1). Classes 2, 3, and 4, which correspond to forests, mountains, and the sea, also show high precision and recall values, demonstrating DenseNet's strong feature extraction and discrimination capabilities. The overall results validate DenseNet's suitability for precise multi-class post-flood aerial image classification, despite minor performance reductions for buildings and streets.

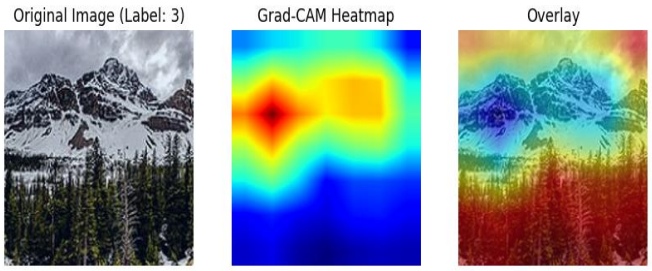


Figure 4.1. Grad-CAM visualization highlighting important regions used by the model for mountain class prediction.

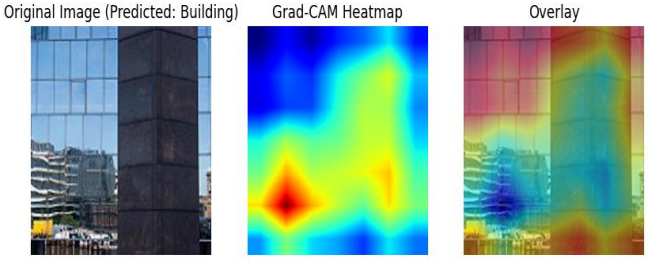


Figure 4.2. Overlay of Grad-CAM heatmap on the original aerial image illustrating model attention regions.

The use of Grad-CAM to illustrate the deep learning models' decision-making process is shown in Figures 4.1 and 4.2. This example examines an aerial photo that is categorized as a mountainous area (class 3). The Grad-CAM heatmap displays the regions of the image that were most crucial to the model's prediction; warmer colors denote greater significance.

It is evident from the heatmap overlay on the original image that the model considers pertinent terrain features when making decisions. The classification is based on significant spatial patterns rather than unimportant background data, as this visual explanation demonstrates. Such interpretability supports the model's applicability in actual disaster management and post-flood assessment scenarios and increases confidence in its predictions.

VII. CONCLUSION

To sum up, MobileNet and DenseNet performed well and consistently in all six land-type categories of post-flood aerial image classification, attaining high recall, accuracy, and precision. Both models successfully identified important areas like buildings, transportation networks, forests, and flooded areas, demonstrating their applicability for extensive post-disaster assessment. By emphasizing the areas of the image that most affected model predictions, the Grad-CAM integration improved transparency and confidence in the decision-making process by offering insightful visual explanations. For disaster response teams working under time constraints, these comprehensible insights are especially helpful. Overall, by facilitating quick, dependable, and comprehensible post-flood damage analysis, the suggested dual-network framework demonstrates great promise for real-time, AI-assisted disaster management, ultimately supporting more effective recovery and planning efforts.

Table III. Base Paper vs Proposed System Comparison

Method	Algorithm	Accuracy	Precision	Recall	F1-Score
Base Paper	CNN (Traditional CNN)	0.90	0.89	0.88	0.88
Proposed System	Mobilenet, DenseNet (with Grad-CAM)	0.96	0.96	0.96	0.96

Table III compares the outcomes of the algorithms used in the base paper and the proposed system. It also shows how the application of MobileNet, DenseNet, and Grad-CAM has enhanced accuracy, precision, recall, and F1-score to make them more comprehensible.

VIII. FUTURE ENHANCEMENT

The current system's ability to classify aerial photos following a flood shows promise for further development. The first step in enhancing the model's robustness and generalization is to include a range of floods in the dataset, such as images captured at different times of day and in different weather condition. Similarly, to improve the classification accuracy, more advanced deep learning techniques like transformer-based architectures or hybrid models would be used. Additional suggestions for improvement include using drones or unmanned aerial vehicles (UAVs) to analyze flood footage in real time and provide a dynamic perspective of the floods. Enhancing the Grad-CAM approach to offer additional insights and integrate multi-moderate data, such as satellite imagery, in addition to sensor data, could be the second refining step. This could result in more coordinated efforts for post-flood damage assessment and recovery planning. Ultimately, on-site decision-making would become considerably quicker and carried out in real time once the model performance has been optimized to operate on edge devices or in low-resource settings.

IX. REFERENCES

[1] A. Hussain, G. Latif, J. Alghazo, and E. Kim, "Flood detection using deep learning methods from visual images," AIP Conference Proceedings, vol. 3034, no. 1, Mar. 2024, doi: 10.1063/5.0194669.

[2] Q. Bi, K. Qin, Z. Li, H. Zhang, and K. Xu, "Multiple instance dense connected convolution neural network for aerial image scene classification," IEEE Transactions on Image Processing, vol. 29, pp. 4911–4926, Aug. 2019, doi: 10.1109/TIP.2020.2975718.

[3] E. Yasi, T. U. Shakib, N. Sharmin, and T. H. Rizu, "Flood and non-flood image classification using deep ensemble learning," Water Resources Management, vol. 38, no. 13, pp. 5161–5178, Oct. 2024, doi: 10.1007/s11269-024-03906-9.

[4] A. Manaf, N. Mughal, K. R. Talpur, B. A. Talpur, G. Mujtaba, and S. R. Talpur, "Aerial image classification in post-flood scenarios using robust deep learning and explainable artificial intelligence," IEEE Access, vol. 13, pp. 35973–35984, 2025, doi: 10.1109/ACCESS.2025.3543078.

[5] D. Duarte, F. Nex, N. Kerle, and G. Vosselman, "Satellite image classification of building damages using airborne and satellite image samples in a deep learning approach," ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. IV-2, no. 2, pp. 89–96, May 2018, doi: 10.5194/isprs-annals-IV-2-89-2018.

[6] C. Kyrkou and T. Theocharides, "EmergencyNet: Efficient aerial image classification for drone-based emergency monitoring using atrous convolutional feature fusion," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 13, pp. 1687–1699, Apr. 2021, doi: 10.1109/JSTARS.2020.2969809.

[7] A. Akhyar et al., "Deep artificial intelligence applications for natural disaster management systems: A methodological review," Ecological Indicators, vol. 163, p. 112067, Jun. 2024, doi: 10.1016/j.ecolind.2024.112067.

[8] J. R. B. Fisher, E. A. Acosta, P. J. Dennedy-Frank, T. Kroeger, and T. M. Boucher, "Impact of satellite imagery spatial resolution on land-use classification accuracy and modeled water quality," Remote Sensing in Ecology and Conservation, vol. 4, no. 2, pp. 137–149, Jun. 2018, doi: 10.1002/rse2.61.

[9] S. Hossain and D. J. Lee, "Deep learning-based real-time multiple-object detection and tracking from aerial imagery via a flying robot with GPU-based embedded devices," Sensors, vol. 19, no. 15, p. 3371, Jul. 2019, doi: 10.3390/s19153371.

[10] D. Whitehurst et al.,"Post-flood analysis for damage and restoration assessment using drone imagery," Remote Sensing, vol. 14, no. 19, p. 4952, Oct. 2022, doi: 10.3390/rs14194952