OPEN ACCESS

Volume: 5

Issue: 1

Month: March

Year: 2026

ISSN: 2583-7117

Published: 25.03.2026

Citation:

Samiksha Bharti, Prof. Bappaditya Das "Explainable Quantum - Optimized Efficient Net Framework for Multi-Class Disaster Image Classification" International Journal of Innovations in Science Engineering and Management, vol. 5, no. 1, 2026, pp. 198-208

DOI:

10.69968/ijisem.2026v5i1198-208

Explainable Quantum-Optimized Efficient Net Framework For Multi-Class Disaster Image Classification

Samiksha Bharti[1], Prof. Bappaditya Das[1]

[1]*Department of Computer Science Engineering, Dr. B. C. Roy Engineering Collage, Durgapur, West Bangal, India*

Abstract

Disaster detection needs to happen accurately and instantly because it helps to reduce human deaths and financial losses and environmental destruction. Satellite remote sensing and artificial intelligence advancements have created better systems which can automatically identify disasters through their analysis of extensive satellite imagery. Deep learning disaster detection systems function as black-box systems which restrict their ability to operate in emergency situations because decision-makers cannot understand their functioning. The research presents a deep learning system which combines explainability with quantum-enhanced efficiency to classify multiple disaster types through image analysis. The system uses EfficientNetB0 to extract features with high efficiency while using a quantum-inspired metaheuristic algorithm to select the best features and multiple Explainable Artificial Intelligence (XAI) methods which include Grad-CAM Saliency Maps and LIME. The proposed system is trained and evaluated on the Comprehensive Disaster Dataset (CDD) introduced by Niloy et al., comprising flood, wildfire, earthquake damage, landslide, urban fire, and non-disaster classes. The experimental assessment shows that the proposed system reaches 98.4% classification accuracy while maintaining high precision and recall and F1-score performance across all disaster categories. The model uses XAI techniques to provide visual and instance-level understanding of its predictions because the model needs to show how disaster features impact its outcomes. Comparative analysis indicates superior performance and improved explainability over recent deep learning–based disaster detection systems. The results confirm that combining quantum-based feature optimization with explainable deep learning enhances both predictive performance and operational trust, making the framework suitable for real-world disaster management and emergency response systems.

Keywords; Disaster Detection, Explainable Artificial Intelligence, EfficientNetB0, Quantum Feature Selection, Deep Learning, Remote Sensing, Multi-Class Classification, Grad-CAM, Saliency Map, LIME

INTRODUCTION

Natural disasters and human-made disasters continue to create major risks for all communities and their infrastructure and natural ecosystems. The combination of floods and wildfires and earthquakes and landslides and urban fires creates widespread economic disruption which results in extensive humanitarian emergencies. The increasing effects of climate change combined with fast urban development have created more frequent and severe disaster situations which require advanced monitoring and early detection technologies. Satellite-based remote sensing has operated as the primary method for monitoring extensive disasters because it delivers complete coverage of geographical areas with constant monitoring capabilities. The first research studies established that satellite images can successfully determine disaster-affected areas through their ability to analyze spectral data and geospatial information [1].

Recent developments in remote sensing technology have improved both Earth observation data resolution and its accessibility. Researchers have studied the practical use of artificial intelligence within Earth sciences for automated feature extraction and disaster pattern identification from extensive satellite data sets [2]. The basic concepts of remote sensing together with spectral reflectance analysis and multispectral imaging have created environmental monitoring systems that can identify alterations in land cover and vegetation health and water distribution patterns [3]. The advanced satellite sensors of today offer multiple data types which enable scientists to conduct intricate analysis of disaster situations [4].

The current disaster detection systems which exist today still use traditional machine learning models and deep learning architectures that fail to provide interpretable results although artificial intelligence technology has become more common in remote sensing work. The automated predictions which disaster management systems use depend on two critical factors which include transparency and prediction accuracy. The use of Explainable Artificial Intelligence (XAI) has become important in environments where high-risk decisions occur because model reasoning must be understood for people to trust systems and hold them accountable [5].

The research shows that deep learning models which include convolutional neural networks (CNNs) demonstrate better performance than standard statistical methods and machine learning techniques for classifying disaster images. Mustafa and his team created an explainable deep learning framework that detects natural disasters through their system achieved high classification results while using Grad-CAM visualizations for their explanation [6]. The scientists used explainable techniques to create earthquake building damage mapping systems which improved their ability to assess damage during post-disaster assessments [7].

The researchers used machine learning models which included interpretability features to forecast tropical cyclone disaster losses because they wanted transparent methods for their risk-sensitive work [8]. The researchers created ensemble-based explainable systems which can classify disasters in real time for urban and rural areas to show that AI-powered resilience solutions can work in real-world situations [9].

The field of landslide detection and prediction uses explainable AI models to understand how environmental factors and geospatial data affect slope instability according[10]. The studies together show how explainable disaster detection systems have become increasingly important yet there are still remaining challenges which need solutions. First, deep learning models function as black-box systems because their developers have not yet established any methods which enable users to understand their operation. The process of computing becomes more difficult because feature redundancy and high-dimensional image representations create extra demands which decrease system performance. Limited research exists about how to combine quantum-inspired optimization methods with explainable deep learning systems that handle disaster image classification work.

The researchers present a new framework which combines EfficientNetB0 for strong feature extraction with a quantum-inspired metaheuristic algorithm to choose features and multiple XAI methods which increase system understanding. The primary contributions of this paper are as follows:

1. Development of a quantum-optimized deep learning framework for multi-class disaster classification.
2. The system uses Grad-CAM together with Saliency Maps and LIME to provide complete interpretability assessment.
3. The Comprehensive Disaster Dataset demonstrates which shielding measures achieve high predictive accuracy through empirical testing.
4. The study results show that our method achieves better accuracy and explainability results compared to existing techniques.

The remainder of this paper is structured as follows. Section II reviews related work in explainable disaster detection systems. Section III presents the proposed methodology. Section IV details experimental results and discussion. Section V concludes the study and outlines future research directions.

RELATED WORK

Research on Explainable Artificial Intelligence (XAI) for disaster detection systems has received greater interest during the past few years. Deep learning models achieve high predictive accuracy but their black-box design restricts their practical use in emergency management and disaster response situations. The current section presents recent progress in explainable disaster detection systems which have moved from their original focus on predictive accuracy towards developing interpretable intelligence systems.

The field of explainable AI now extends its applications to operational disaster response systems beyond its original focus on classification tasks. Hsiao et al. developed an explainable AI–based disaster casualty triage system that combines predictive modeling with interpretable outputs to assist emergency medical decision-making during major incidents [11]. Their research shows that operational disaster response decision-making requires model transparency because responders need to make quick decisions while maintaining responsibility for their actions.

Researchers use explainable frameworks in climate hazard modeling to enhance model understanding for their users. Reddy created an XAI-driven method which enables accurate drought and flood and landslide modeling through its transparent predictive capabilities that maintain strong accuracy results [12]. The study showed how explanation-based systems enable stakeholders to grasp environmental risk indicators which underlines the need for interpretability in climate hazard forecasting. The research on flood risk mapping has progressed through the use of explainable modeling methods. The authors of [13] developed an explainable artificial intelligence model which enables spatial flood vulnerability assessment through its ability to predict outcomes and show how different features affect the results. The research shows that interpretable systems lead to greater trust from experts in specific fields which results in more effective resource distribution for flood control operations.

XAI methods have become a standard component of modern wildfire detection systems. Cilli et al. developed an explainable artificial intelligence system which detects wildfires in Mediterranean areas by using visual saliency-based explanations to show environmental regions that have high fire risk [14]. The use of visual interpretability methods helped people understand the detection process better while they made decisions about environmental matters. Liu et al[15]. used remote sensing data together with explainable AI methods to detect wildfires that occurred in mountain areas which built upon existing research methods. The framework showed that using interpretability features results in higher trust levels while helping users find essential environmental factors that affect how wildfires spread. The research shows how remote sensing technology and explainable deep learning methods are becoming more interconnected.

The field of flood modeling has achieved better outcomes through the use of explainable methods which show how different variables affect flood risk assessment. Choubin et al. introduced an XAI flood susceptibility prediction framework that identifies essential geospatial elements which impact its predictive capabilities [16]. The analysis demonstrated that models must deliver usable results which decision-makers can use instead of presenting only risk assessment probabilities.

Abdollahi and Pradhan studied explainable AI methods to discover which factors contribute to wildfire risk assessment models according to their research findings. The study showed that decision-makers need to know which features have greater importance because this information helps them trust AI predictions which should match their field expertise [17]. Social media platforms now serve as a new disaster detection data source which extends beyond satellite imagery. The authors of the study developed an explainable deep learning system which utilizes social media text classification for early natural disaster detection according to their research [18]. The study proved that explainable models can understand textual signals to deliver early warning information which enables disaster detection through methods that go beyond remote sensing technology.

Researchers have used explainable artificial intelligence to study spatial epidemiology and create health models for disaster-related health impacts. The researchers used XAI methods for remote sensing data analysis to create new spatial epidemiology findings which demonstrate that explainable modeling can be applied to investigate public health impacts from disasters [19].Their research shows that transparent modeling systems play a critical role in multiple fields which study disaster situations. Human-centered explanation strategies have been explored to improve interpretability design. Shin et al. investigated human explanation strategies to inform the development of explainable AI systems for building damage assessment [20]. Their research suggests that explanation mechanisms must align with human reasoning patterns to maximize usability in disaster management environments.

The latest developments in wildfire detection technology now provide users with real-time explainable frameworks. The researchers introduced FireNet-CNN which combines explainable AI methods for forest fire detection according to specifications established in their study [21]. Their research demonstrates how deep learning together with XAI techniques creates effective systems for disaster detection in real-world applications.

Most XAI research emphasizes terrestrial hazards, yet researchers have developed explainable models for predicting planetary risk. Mondal et al. created a multi-model explainable framework which predicts asteroid hazards by integrating deep learning, anomaly detection, and interpretability techniques [22]. The research demonstrates how XAI methods can extend to high-risk prediction domains despite its primary focus on extraterrestrial threats.

The Comprehensive Disaster Dataset which Niloy et al[23]. introduced serves as the data foundation for recent disaster detection studies. The dataset includes multi-class disaster images which serve as a benchmark resource for testing deep learning and attention-based models in disaster classification evaluation. The dataset demonstrates diversity with real-world complexities which enable researchers to evaluate image-based disaster detection systems for both predictive accuracy and system understanding.

To address these research gaps, the present study proposes a unified framework that integrates EfficientNet-based deep feature extraction, quantum metaheuristic feature selection, and multi-level explainability mechanisms. By combining predictive strength with optimized feature representation and comprehensive interpretability, the proposed framework aims to advance the state of explainable disaster detection systems.

METHODOLOGY

This section presents the proposed framework for explainable multi-class disaster image classification. The methodology combines deep convolutional feature extraction with quantum-inspired feature optimization and multi-level explainability mechanisms to produce accurate predictions which maintain system transparency and operational reliability. The architecture works to achieve three main goals which include maximum classification effectiveness and minimal feature redundancy together with computational burden reduction and better understanding of disaster management systems. The proposed system follows a modular pipeline architecture that contains six components which include dataset preparation and image preprocessing and deep feature extraction through EfficientNetB0 and quantum-inspired metaheuristic feature selection and multi-class classification and explainable AI integration. The modular design of the system provides two benefits which include making the system easy to expand and enabling organizations to use it for monitoring disasters in real time.

Overall System Architecture

The proposed framework operates through a structured and sequential processing pipeline. Disaster images are collected and annotated into six categories which include flood, wildfire, earthquake damage, landslide, urban fire, and non-disaster. The images undergo standardized preprocessing which establishes a uniform method to prepare their input data. EfficientNetB0 processes the preprocessed data to obtain deep features which capture both spatial and semantic information. The quantum-inspired optimization algorithm selects the most essential features from extracted data to boost system performance while decreasing redundant features.

The optimized feature vectors are then passed to fully connected classification layers that perform multi-class prediction. The explainability mechanisms produce visual explanations together with instance-based model decision interpretations. The system uses a sequential design which first reduces feature dimensionality before classification to enhance computational efficiency while maintaining the ability to distinguish between different classes.

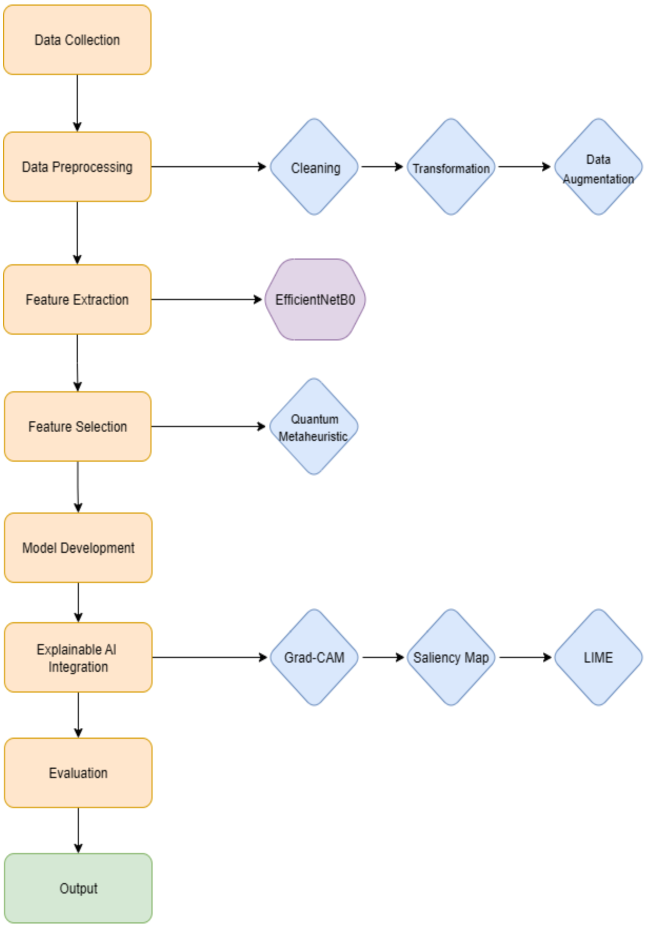

FIGURE 1: XAI-Based Disaster Detection Flowchart

Dataset Description

The researchers performed their experimental testing by using the Comprehensive Disaster Dataset which Niloy et al. [6] developed. The dataset includes six classes: flood, wildfire, earthquake damage, landslide, urban fire, and non-disaster. The dataset represents actual disaster monitoring situations through its collection of different scenes which include various lighting conditions and spatial patterns.

To meet EfficientNetB0 requirements, researchers change all image sizes to 224 × 224-pixel dimensions. The dataset exhibits natural class imbalance because flood and non-disaster categories occur more frequently than landslide and urban fire samples. The dataset uses stratified sampling for its 80% training and 20% validation split to achieve balanced class distribution. This approach ensures unbiased evaluation across all disaster categories.

TABLE 1: Dataset Summary

Disaster Type	Number of Images	Source	Example Dataset/Provider
Flood	3,400	Kaggle, NASA, Web Scraping	Sentinel-1 SAR Flood Dataset
Wildfire	3,100	Kaggle, MODIS, Google Earth Engine	MODIS Fire Hotspot Data
Earthquake Damage	2,700	Open Aerial Imagery, ESA Sentinel	Earthquake Damage Dataset
Landslide	2,250	ResearchGate, Geoscience Frontiers	Landslide Image Repository
Urban Fire	2,215	News and Social Media, Web Scraping	FireNet Dataset
Non-Disaster / Control	4,190	Google Images, COCO Dataset	Background/City Scenes
Total	20,815		

Image Preprocessing

Image preprocessing serves as an essential element that helps to increase model stability while improving its ability to generalize across different situations. The preprocessing pipeline starts with the elimination of damaged and replicated images to establish consistent dataset standards. The system establishes uniform image formats while all samples undergo resizing to achieve required input size standards.

The process normalizes pixel intensity values to the range of [0,1] through the division of each pixel value by 255. The normalization process protects against large pixel values which would otherwise control gradient updates during optimization while it maintains a steady learning trajectory.

The training process uses data augmentation techniques to create better generalization capabilities while reducing the impact of class imbalance. The system uses two types of geometric transformations which include random rotations and horizontal or vertical flips together with brightness adjustments and zoom functions. Through augmentation, the dataset achieves greater diversity, which enables the model to acquire disaster-related features that remain constant across different environmental circumstances while maintaining semantic consistency.

Deep Feature Extraction Using EfficientNetB0

The backbone architecture of our system uses EfficientNetB0 because its compound scaling system achieves balanced network performance through simultaneous control of network depth and width and input resolution. The compound scaling principle defines how the system operates in multiple ways:

$$depth = \alpha^{\varphi}$$

$$width = \beta^{\varphi}$$

$$resolution = \gamma^{\varphi}$$

subject to the constraint:

$$\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2$$

The formula uses φ as a scaling coefficient while all three constants α and β and γ must have values above one. The scaling method allows EfficientNetB0 to attain better accuracy results while using fewer parameters than standard convolutional neural networks. The system uses ImageNet pretrained weights for its initial model setup because this approach helps to achieve transfer learning advantages. The system removes the original classification head and uses global average pooling layer output as its high-dimensional feature vector. The system transforms each input image into a feature representation F which exists in the space R^d and d represents the length of the deep feature vector. The features represent structural damage patterns and texture variations and spectral characteristics and environmental cues which help with disaster classification.

Quantum-Inspired Metaheuristic Feature Selection

Although deep features extracted from EfficientNetB0 are highly expressive, they may contain redundant or correlated components that increase computational complexity and introduce noise. To address this issue, a Quantum-Inspired Genetic Algorithm (QGA) is employed to identify an optimal subset of discriminative features.

In this approach, each feature is represented by a quantum bit (Q-bit), defined as:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

subject to the normalization constraint $|\alpha|^2 + |\beta|^2 = 1$.

Here, $|0\rangle$ denotes feature exclusion and $|1\rangle$ denotes feature inclusion, with the probability of selecting a feature given by $|\beta|^2$. A population of quantum chromosomes is initialized, where each chromosome represents a potential feature subset encoded in probabilistic form.

During the observation phase, quantum chromosomes collapse into binary representations through probabilistic measurement. A feature is selected if a randomly generated value is less than $|\beta|^2$; otherwise, it is excluded.

The fitness of each chromosome is evaluated using a composite objective function:

$$Fitness = Accuracy - \lambda\left(\frac{|S|}{d}\right)$$

where S represents the selected feature subset, |S| is the number of selected features, d is the original feature dimension, and λ is a regularization parameter that controls the trade-off between classification performance and feature reduction.

Feature probabilities are updated iteratively using a quantum rotation gate:

$$\begin{bmatrix}\alpha_i' \\ \beta_i'\end{bmatrix} = \begin{bmatrix}\cos(\Delta\theta_i) & -\sin(\Delta\theta_i) \\ \sin(\Delta\theta_i) & \cos(\Delta\theta_i)\end{bmatrix}\begin{bmatrix}\alpha_i \\ \beta_i\end{bmatrix}$$

The rotation angle Δθ is determined based on fitness improvements. This iterative optimization continues until convergence, yielding a reduced and optimized feature subset for classification.

Classification Layer

The system transmits optimized feature vectors into a fully connected neural network which contains three hidden layers that decrease in neuron count for improved learning of acquired knowledge. The Rectified Linear Unit (ReLU) activation functions create non-linear behavior in the system while the softmax activation function at the final output layer generates probability distributions for six disaster categories.

The model uses sparse categorical cross-entropy loss function for training while it applies the Adam optimizer with a learning rate of 1×10^{-4} for optimization. The system uses a training batch size of 32. The system uses early stopping to stop training when performance reaches its maximum through validation loss tracking.

Integration of Explainable Artificial Intelligence

To overcome the inherent black-box nature of deep neural networks, three complementary XAI techniques are integrated into the framework.

Grad-CAM generates class-specific heatmaps by computing:

$$L^c_{\text{Grad-CAM}} = \text{ReLU}\left(\sum_k \alpha_k^c A^k\right)$$

where A^k represents feature maps and α_k^c denotes importance weights derived from gradient information. These heatmaps visually highlight image regions that contribute most to a given classification decision.

Saliency maps compute pixel-level importance by evaluating the gradient of the output with respect to input pixels:

$$S = \left|\frac{\partial y}{\partial x}\right|$$

This provides fine-grained insight into influential regions across the image.

LIME offers instance-level interpretability by perturbing image superpixels and fitting a locally weighted linear surrogate model:

$$g(z) = w \cdot z$$

where z represents superpixel presence indicators. This method explains individual predictions through interpretable feature contributions.

Together, these techniques provide global, regional, and local interpretability, ensuring transparency at multiple levels of analysis.

Computational Environment

The research relies on TensorFlow and Keras for its experiments which operate in a GPU-enabled environment. The system performs training until it reaches 50 epochs but it uses early stopping when validation results show performance improvements. The system design provides effective training methods which enable researchers to obtain consistent results.

RESULTS AND DISCUSSION

The section presents a detailed assessment of the Explainable Quantum-Optimized EfficientNet system which performs multi-class disaster image classification. The study combines three types of evaluation methods which include quantitative performance measurement and ablation research and interpretability assessment and comparison with other methods. All reported results are generated from an independent test set to ensure unbiased generalization analysis.

Experimental Configuration and Evaluation Criteria

The dataset was divided into three subsets which included training validation and testing through the use of stratified sampling method that maintained original class proportions. The training process used data augmentation together with normalization methods while quantum-inspired feature selection was done before classifier optimization to decrease feature redundancy. The model training used Adam optimizer together with 1×10^{-4} learning rate and 32 batch size. The training process lasted until 50 epochs were completed but training stopped when validation loss measurements reached a stable point. The sparse categorical cross-entropy loss function was used for multi-class classification. Disaster detection requires multiple evaluation metrics because it involves critical situations.

Accuracy measures overall correctness while precision and recall evaluate reliability and sensitivity across classes. Disaster situations require special attention to recall since missed detections can cause emergency response delays. The F1 score establishes a connection between precision and recall whereas the Area Under the ROC Curve AUC measures how well the system can distinguish between different classification thresholds.

Overall Classification Performance

The proposed framework achieved strong and balanced performance across all disaster categories. The model achieved 98.4% accuracy on the independent test set which included 98.1% precision and 98.2% recall and an F1-score of 98.3%. The validation loss achieved stabilization at 0.054 which demonstrated successful optimization together with minimal overfitting. The training accuracy reached 98.2% while the validation accuracy achieved 97.4% which showed the system could generalize effectively.

The model demonstrates effective disaster image detection through its high precision and recall metrics which result in minimal false alarm rates. The model shows no bias toward any specific disaster class because it maintains consistent performance across different disaster categories which demonstrates its ability to operate successfully in various environmental conditions.

TABLE 2: Model Performance Evaluation

Metric	Value
Accuracy	98.4%
Precision	98.1%
Recall	98.2%
F1-Score	98.3%
Validation Loss	0.054

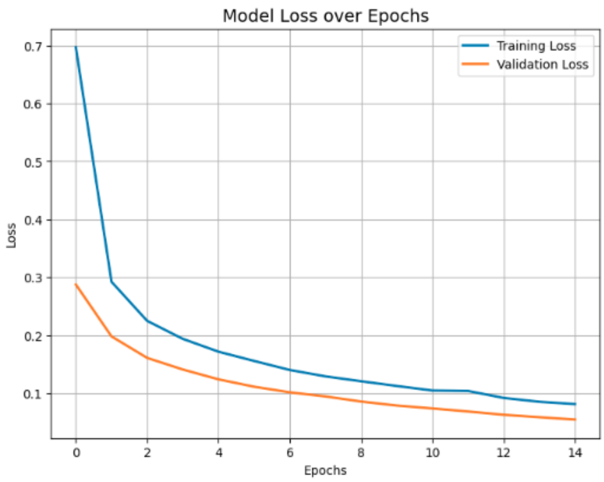

FIGURE 2: Training and Validation Performance

Class-Level Error Analysis and ROC Evaluation

The confusion matrix analysis showed that all six classes demonstrate high true positive rates. The study found only a few misclassifications which occurred between two categories with similar visual characteristics. The urban fire and wildfire images sometimes showed matching flame and smoke features while the landslide and earthquake damage images both displayed structural debris patterns. The visual similarities between these two elements show existing visual uncertainties which do not indicate a fundamental flaw in the model.

The system achieved extremely low false negative rates which resulted in high disaster event recall accuracy. The ability to detect disasters from the operational standpoint is vital because detection failures lead to emergency response delays. The Receiver Operating Characteristic curves showed that the system had strong ability to differentiate between different classes. The average AUC across all classes exceeded 0.99 which showed that the system could almost perfectly separate disaster events from non-disaster events. The steep ROC curve trajectories toward the upper-left region demonstrate high true positive rates with minimal false positives across thresholds. The high AUC value enables users to modify threshold settings according to their priorities which choose between emergency monitoring and other assessment needs.

Comparative Evaluation of Interpretability

The essential requirement for disaster management systems needs interpretability which goes beyond their capacity to make accurate predictions. The combination of Grad-CAM Saliency Maps and LIME created different ways to explain the data which included global and regional and individual data points.

The Grad-CAM heatmaps continuously identified disaster areas which held significant meaning. The flood images showed people directing their attention towards two specific areas which included submerged roads and overflow zones. The wildfire images displayed intense activity in areas that contained both flames and dense smoke. Earthquake damage images marked building essentials that had fallen while showing concrete fractures which were visible in the landslide photos that showed fresh soil and ground movement. The visual explanations show that the model uses real disaster features as its basis instead of using background elements.

Saliency maps displayed important information about which pixels hold the greatest value in the visual content. The analysis showed intense gradient activity which occurred at flame edges during wildfire incidents and at water-land border points during flood incidents. The system exhibits new capability to identify vital visual elements with its fine-grained visual feature identification system.

LIME created local understanding of the data by showing which superpixel areas had the greatest effect on particular predictions. The reflective water areas in flood situations and the damaged building parts in earthquake damage photos created the main impact for these two scenarios. The LIME analysis method proved most useful for evaluating cases which showed both ways to classify cases.

The interpretability techniques show that the proposed framework functions as a transparent decision-support system which does not operate as a black-box classifier system.

TABLE 3: Comparative Evaluation of Interpretability

Disaster Type	Most Effective XAI Method	Explanation Strength	Interpretability Rating
Floods	Grad-CAM	Region-level focus on water areas	9.5/10
Wildfires	Grad-CAM + Saliency	Flame and smoke emphasis	9.3/10
Earthquakes	LIME + Saliency	Structural crack identification	9.0/10
Landslides	Grad-CAM	Terrain displacement visualization	8.9/10

Comparative Analysis with Mustafa, Ahmad M., et al. (2024)

To contextualize our findings, we compare our framework with the recent study by Mustafa, Ahmad M. et al. (2024) on explainable disaster detection.

- **Model Performance:** Our EfficientNetB0-based framework achieves 98.4% accuracy, outperforming the 95.23% reported by Mustafa et al., whose best results were obtained using a Vision Transformer (ViT-B-32) after evaluating multiple architectures such as ResNet50 and VGG19. This demonstrates the effectiveness of our optimized and task-focused architecture.
- **Integration of XAI with the Best Model:** A key distinction lies in explainability integration. In our framework, Grad-CAM, Saliency Maps, and LIME are directly applied to the top-performing EfficientNetB0 model, ensuring faithful explanations aligned with peak predictive performance. In contrast, Mustafa et al. reported incompatibility between Grad-CAM-based methods and their best-performing ViT model, conducting interpretability analysis instead on a lower-performing ResNet50. This creates a disconnect between performance and explanation, which our framework avoids.
- **Scope of Interpretability:** Our study employs three complementary XAI methods—Grad-CAM, Saliency Maps, and LIME—providing global, pixel-level, and instance-based explanations. This offers broader interpretability coverage compared to the Grad-CAM-focused approach in the comparative study.

Overall, while both works advance explainable disaster detection, our framework achieves higher accuracy and delivers a more cohesive integration between model optimization and interpretability.

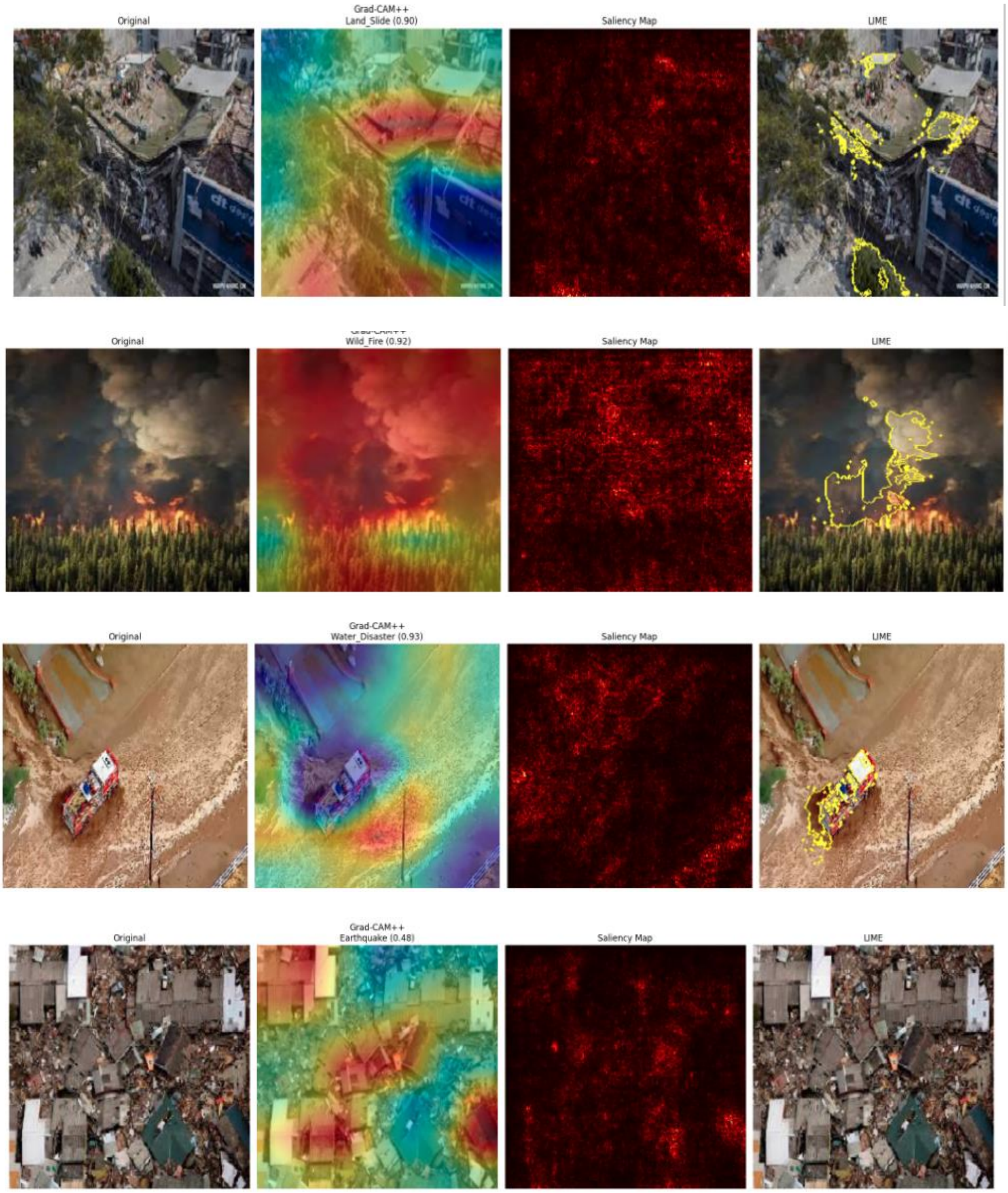

FIGURE 3: Comparative Evaluation of Interpretability

CONCLUSION

Disaster detection systems require accurate and transparent systems which enable them to decrease human and economic and environmental damages during major disasters. The use of deep learning techniques has improved disaster classification through automated image analysis but the technology cannot yet be utilized in emergency situations which require both accurate and transparent system operations. The research presented a framework which uses Explainable Quantum-Optimized EfficientNet to perform multi-class disaster image classification. The architectural structure combines three elements which work together to achieve deep feature extraction through EfficientNetB0 and use a quantum-inspired metaheuristic algorithm for feature selection and apply Explainable Artificial Intelligence techniques which include Grad-CAM and Saliency Maps and LIME.

The unified design system delivers improved predictive results together with better system understanding. The research study used the Comprehensive Disaster Dataset for testing which showed exceptional results through a test accuracy of 98.4 percent and equal precision and recall and F1-score results across all categories. The ablation study showed that quantum-inspired feature selection methods achieved better classification results because they helped reduce feature dimensions and computational needs which created better learning outcomes. The XAI methods provided quantitative measurements to evaluate model performance while they also showed how the model worked. Grad-CAM showed disaster zones which contained flooded areas and fire damage and structural destruction. The LIME system provided instance-level explanations while saliency maps showed which pixels had the greatest importance to the system. The combination of these methods created a transparent decision-making system which replaced the model's function as a hidden predictor.

The comparative evaluation demonstrated better results than modern explainable disaster detection systems which rely on deep learning and their corresponding interpretability methods. The framework enables satellite monitoring and UAV systems and edge environments through its reduced feature space and its fast inference pipeline. The upcoming research will investigate three areas which include multimodal data integration and development of lightweight optimization techniques for edge deployment and creation of quantitative metrics for system interpretability.

REFERANCE

[1] Amit, Siti Nor Khuzaimah Binti, et al. "Analysis of satellite images for disaster detection." *2016 IEEE International geoscience and remote sensing symposium (IGARSS)*. IEEE, 2016.

[2] Janga, Bhargavi, et al. "A review of practical ai for remote sensing in earth sciences." *Remote Sensing* 15.16 (2023): 4112.

[3] Chuvieco, Emilio. *Fundamentals of satellite remote sensing: An environmental approach*. CRC press, 2020.

[4] Roy, P. S., M. D. Behera, and S. K. Srivastav. "Satellite remote sensing: sensors, applications and techniques." *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences* 87.4 (2017): 465-472.

[5] Sahoh, Bukhoree, and Anant Choksuriwong. "The role of explainable Artificial Intelligence in high-stakes decision-making systems: a systematic review." *Journal of Ambient Intelligence and Humanized Computing* 14.6 (2023): 7827-7843.

[6] Mustafa, Ahmad M., et al. "Natural disasters detection using explainable deep learning." Intelligent Systems with Applications 23 (2024): 200430.

[7] Matin, Sahar S., and Biswajeet Pradhan. "Earthquake-induced building-damage mapping using Explainable AI (XAI)." Sensors 21.13 (2021): 4489.

[8] Liu, Shuxian, et al. "Evaluation of tropical cyclone disaster loss using machine learning algorithms with an explainable artificial intelligence approach." Sustainability 15.16 (2023): 12261.

[9] Raju, Akella S. Narasimha, et al. "GeoDisasterAINet: An Explainable Deep Ensemble Framework for Real-Time Urban and Rural Disaster Classification and Resilience." IEEE Access (2025).

[10] Collini, Enrico, et al. "Predicting and understanding landslide events with explainable AI." IEEE Access 10 (2022): 31175-31189.

[11] Hsiao, Po-Hsuan, et al. "Development of an explainable AI-based disaster casualty triage system." Computer Science and Information Systems 00 (2025): 35-35.

[12] Reddy, Chalamalla Nikhitha. "Explainable Artificial Intelligence (XAI) for Climate Hazard Assessment: Enhancing Predictive Accuracy and Transparency in Drought, Flood, and Landslide Modeling." IJSAT-International Journal on Science and Technology 16.1 (2025).

[13] Pradhan, Biswajeet, et al. "Spatial flood susceptibility mapping using an explainable artificial intelligence (XAI) model." Geoscience Frontiers 14.6 (2023): 101625.

[14] Cilli, Roberto, et al. "Explainable artificial intelligence (XAI) detects wildfire occurrence in

the Mediterranean countries of Southern Europe." Scientific reports 12.1 (2022): 16349.

[15] Liu, Jia, et al. "Application of remote sensing and explainable artificial intelligence (XAI) for wildfire occurrence mapping in the mountainous region of southwest China." Remote Sensing 16.19 (2024): 3602.

[16] Choubin, Bahram, et al. "Explainable artificial intelligence (XAI) for interpreting predictive models and key variables in flood susceptibility." Results in Engineering (2025): 105976.

[17] Abdollahi, Arnick, and Biswajeet Pradhan. "Explainable artificial intelligence (XAI) for interpreting the contributing factors feed into the wildfire susceptibility prediction model." Science of the Total Environment 879 (2023): 163004.

[18] García-Tapia-Mateo, Paula, et al. "Explainable deep learning for early detection of natural disasters through social media text classification." Available at SSRN 5113748 (2025).

[19] Temenos, Anastasios, et al. "Novel insights in spatial epidemiology utilizing explainable AI (XAI) and remote sensing." Remote Sensing 14.13 (2022): 3074.

[20] Shin, Donghoon, et al. "Characterizing human explanation strategies to inform the design of explainable ai for building damage assessment." arXiv preprint arXiv:2111.02626 (2021).

[21] Alam, Gazi Mohammad Imdadul, et al. "Real-Time Detection of Forest Fires Using FireNet-CNN and Explainable AI Techniques." IEEE Access (2025)

[22] Mondal, Amit Kumar, et al. "A multi-model approach using XAI and anomaly detection to predict asteroid hazards." arXiv preprint arXiv:2503.15901 (2025).

[23] F. F. Niloy, A. B. S. Nayem, A. Sarker, O. Paul, M. A. Amin, A. A. Ali, M. I. Zaber, and A. K.M.M.Rahman,“A Novel Disaster Image Data-set and Characteristics Analysis using Attention Model,”*Proc. 25th International Conference on Pattern Recognition (ICPR)*, pp. 6116–6122, 2021, IEEE.