Vol. 08, No. 2 (2026) 999-1016, doi: 10.24874/PES08.02A.014

Proceedings on Engineering Sciences

www.pesjournal.net

A HYBRID QUANTUM–CLASSICAL LEARNING FOR FINE-GRAINED POST-DISASTER BUILDING DETECTION REMOTE SENSING IMAGERY

Ankita Dixit[1]
Indu Chawla

Received 09.03.2026
Revised 21.04.2026
Accepted 13.05.2026

Keywords:

Natural disaster, Remote sensing images, Building damage assessment, Attention mechanism, U-Net, Quantum computing

A B S T R A C T

Rapid and accurate building damage assessment is essential for effective emergency response and recovery. This study proposes a Dual-Path Quantum–Classical U-Net (DQC-U-Net) for fine-grained damage assessment and structured mapping. The framework integrates building segmentation and damage classification within a unified two-stage architecture. It leverages quantum-enhanced feature modeling with U-Net to capture both global dependencies and local spatial details using six- and eight-qubit configurations. In the first stage, building footprints are extracted from pre-disaster imagery. In the second stage, building masks guide multi-class damage classification into no, minor, major, and destroyed categories. Comparative experiments with U-Net++, DeepLabv3+, HRNet, and U-NeXt demonstrate consistent performance improvements across evaluation metrics.

© 2026 Published by Faculty of Engineering

1. INTRODUCTION

Natural disasters such as earthquakes, floods, hurricanes, wildfires, and tsunamis cause extensive damage to buildings and infrastructure worldwide, resulting in significant human casualties and economic losses. Rapid and accurate assessment of building damage is essential for effective emergency response, resource allocation, loss estimation, and post-disaster recovery (Wu et al., 2016). Consequently, reliable damage information is a critical requirement for disaster management authorities. Recent natural disasters have caused extensive infrastructure damage across different regions. This includes multi-billion-dollar losses from floods and extreme events in China (Reuters, 2026), severe monsoon-induced destruction of roads, bridges, and public assets in Himachal Pradesh, India (Economic Times, 2025), and persistent large-scale impacts across Africa, where natural disasters result in average annual infrastructure losses of approximately USD 12.7 billion, affecting transport, energy, and water systems. Conventional building damage assessment methods rely primarily on field surveys and manual interpretation of satellite imagery. Field-based surveys require trained personnel to physically access affected areas. As disaster-affected regions can span thousands of square kilometres, manual inspection is labor-intensive, costly and often unsafe when transportation networks are disrupted. In many post-disaster environments, such as those impacted by floods, earthquakes, or landslides, on-ground assessments can be hazardous**,** exposing survey teams to unstable structures, debris, contaminated water, and secondary hazards.

[1] Corresponding author: Ankita Dixit
Email: ankitadixit177@gmail.com

The challenges are further exacerbated during large-scale disasters, where emergency responders operate under severe time constraints. Rapid situational awareness is critical in the first few hours to days following a disaster, yet traditional assessment methods often fail to deliver actionable insights within this narrow time window. Delays in damage assessment can result in inefficient allocation of rescue resources, slowed relief operations, and increased humanitarian and economic losses. Consequently, the growing frequency and intensity of natural disasters underscore the urgent need for automated, scalable, and reliable damage assessment approaches that can operate effectively under time-critical and resource-constrained conditions.

Remote sensing imagery, particularly optical satellite data, has emerged as a reliable approach due to its wide-area coverage, frequent revisit times, and cost efficiency (Dong & Shan, 2013). Advances in satellite imaging technology enable rapid acquisition of post-disaster imagery, facilitating timely damage analysis. However, accurate interpretation of satellite images remains challenging due to complex post-disaster conditions such as debris, shadows, water coverage, smoke, and varying illumination, which can degrade model performance (Ghaffarian & Kerle, 2019).

Early automated approaches focused on damage detection using only post-disaster imagery. While computationally simple, these methods often struggle to distinguish actual structural damage from environmental artifacts. To improve reliability, many recent studies incorporate both pre-disaster and post-disaster images. By analysing temporal changes, these approaches more effectively capture damage-related variations. Nevertheless, building damage assessment remains more complex than traditional change detection, as it requires precise building localization and multi-class damage classification, including no damage, minor damage, major damage, and destruction.

Recent advances in deep learning have greatly improved automated building damage assessment using remote sensing imagery, with U-Net–based encoder–decoder architectures demonstrating strong performance in pixel-level segmentation tasks (Ronneberger et al., 2015). It has been proposed by the authors that deep neural networks can extract fine-grained spatial features from high-resolution satellite imagery for accurate damage mapping (Xia et al., 2023). The work (Gerard et al., 2024) demonstrated strong performance on the xBD dataset using carefully designed yet simple CNN architectures. However, another work (Su et al. 2020) highlighted persistent challenges such as class imbalance and limited generalization across disaster types, indicating remaining gaps in current deep learning–based approaches. Despite these advances, existing approaches often struggle to jointly capture global structural dependencies and fine-grained spatial details, particularly under complex post-disaster conditions, while also being affected by background interference and class ambiguity.

To address these limitations, this study proposes a hybrid framework that integrates quantum and classical learning paradigms for building damage assessment. The primary contributions of this work are summarized as follows:

- A hybrid quantum–classical U-Net framework is proposed for automated building damage assessment using pre- and post-disaster satellite imagery.
- A two-stage architecture is introduced that separates building localization from damage classification, reducing background interference and improving prediction consistency.
- A dual-path design is developed to combine quantum-derived global feature representations with convolutional local features, enabling improved modelling of structural dependencies.
- An attention-based fusion mechanism and building-aware masking strategy are incorporated to enhance feature integration and focus on structurally relevant regions.
- Experimental results on the xBD dataset demonstrate consistent performance improvements over baseline methods, indicating the effectiveness of the proposed approach for fine-grained damage assessment.

The remainder of this paper is organized as follows. Section II reviews related work using remote sensing imagery. Section III describes the datasets and preprocessing procedures. Section IV presents the proposed methodology, while Section V provides implementation details. Section VI outlines the evaluation metrics. Section VII presents the experimental results and discusses the findings. Section VIII describes the transfer learning, and Section IX concludes the paper.

2. RELATED WORK

This section reviews recent methods for building damage assessment using remote sensing satellite imagery. Due to the severe impact of natural disasters on human life and urban infrastructure, damage assessment has remained an active research topic. Over the past two decades, satellite imagery has been increasingly adopted for this task because of its wide spatial coverage and rapid data acquisition capabilities (Gupta et al., 2019).

Early automated approaches primarily relied on traditional machine learning techniques. For example, texture and spectral features extracted from pre- and post-disaster Sentinel-1 and Sentinel-2 images were used to generate building damage maps and classified using a Random Forest model (Rathje et al., 2005). Another study employed an ANFIS-based decision-making system that combined geometric features, bi-temporal satellite images, and building vector data to detect damaged buildings following an earthquake event (Kakooei & Baleghi, 2023).

In recent years, deep learning–based methods have been widely explored due to their strong capability to learn rich and discriminative image representations (Putri et al., 2022). A deep object-based semantic change detection framework was proposed to assess building damage caused by both natural and human-induced disasters (Janalipour & Mohammadzadeh, 2015). A two-stage U-Net architecture enhanced with a Squeeze-and-Excitation module was introduced to further improve damage assessment accuracy (Shen et al., 2021). In addition, attention-based U-Net variants were developed to enhance building damage localization and classification performance (Zheng et al., 2021). Several state-of-the-art semantic segmentation networks are used as baseline models in this study. UNet++ enhances the U-Net architecture through nested dense skip connections to improve multi-scale feature fusion (Deng & Wang, 2022). DeepLabv3+ employs an encoder–decoder framework with atrous separable convolutions and ASPP to capture contextual information at multiple scales (Zhou et al., 2018).HRNet maintains high-resolution representations via parallel multi-resolution branches, enabling precise spatial localization (Chen et al.,2018), while U-NeXt integrates MLP-based token mixing into a lightweight U-Net–style design for efficient segmentation (Wang et al., 2019).

Several studies focused on modelling correlations between pre- and post-disaster images. A two-stage convolutional neural network was proposed in (Valanarasu & Patel, 2022), where a U-Net is first used for building localization, followed by a dual-branch multiscale U-Net with cross-directional attention for damage assessment. This approach achieved state-of-the-art performance on the xBD dataset. Transformer-based frameworks have also been explored, integrating attention mechanisms and auxiliary information to capture long-range dependencies and complex damage patterns for improved damage classification (Qing et al., 2022).

To better handle scale variation and contextual information, a Selective-Kernel U-Net was introduced to adaptively capture multi-scale convolutional features, improving sensitivity to varying building sizes and damage patterns (Xiao & Mostafavi, 2025). A local–global collaborative attention-based network was proposed to jointly model fine-grained details and global contextual relationships for flood-damaged building extraction (Ahmadi et al., 2024). Furthermore, a Swin Transformer-based framework with a cross-fusion pyramid was developed to enhance multi-scale feature integration for flood-related building damage classification (Jiang et al., 2024). Finally, an incremental learning framework was proposed to enable rapid damaged-building identification by transferring knowledge from historical disaster datasets, improving generalization across disaster scenarios with limited labeled data (Wu et al., 2025). An automated large-scale building damage assessment framework integrating satellite imagery, GIS data, and deep learning models was also introduced to support post-disaster mapping and decision-making (Ge et al., 2023).

Despite recent advances in deep learning–based building damage assessment, most existing methods rely solely on classical models, which often struggle to capture global structural correlations and subtle damage patterns in complex post-disaster scenes. Transformer-based approaches improve global context modeling but introduce high computational complexity and data demands. Recent studies in quantum machine learning indicate that quantum-enhanced models can efficiently represent complex feature relationships through superposition and entanglement (Braik & Koliou, 2024; Schuld et al., 2015; Biamonte et al., 2017; Havlíček et al., 2019). Motivated by this, we propose a hybrid quantum–classical framework that integrates a quantum encoder–bottleneck module with a convolutional segmentation network to enhance global context modeling and improve building damage classification accuracy.

3. PRELIMINARIES

This section introduces the core principles of the proposed dual U-Net framework and the quantum computing components incorporated within it.

3.1 U-Net

U-Net is a fully convolutional neural network architecture specifically designed for image segmentation tasks (Ronneberger et al., 2015). It follows an encoder–decoder structure, where the encoder progressively captures high-level semantic features through down sampling, and the decoder reconstructs spatial resolution via up sampling operations. Skip connections between corresponding encoder and decoder layers enable the fusion of contextual and fine-grained spatial information, improving localization accuracy. Due to its ability to perform precise segmentation with limited training data, U-Net has been widely adopted in medical imaging, such as cell segmentation, organ delineation, and lesion detection in microscopy, CT, and MRI images (Iglovikov & Shvets, 2018; Milletari et al., 2016; He et al., 2016). Beyond healthcare, U-Net has been widely adopted in remote sensing for land-cover mapping, change detection, building footprint extraction, and disaster damage assessment using satellite and aerial imagery (Daudt et al., 2018).

3.2 QUANTUM COMPUTING

Quantum computing provides an alternative computational framework in which information is processed using quantum states rather than classical binary variables. In hybrid quantum–classical learning, quantum components are embedded within conventional deep learning architectures to enhance representation capability without replacing classical processing entirely.

The fundamental unit of quantum information is the qubit, which can exist in a superposition of basis states, enabling compact and expressive feature encoding. In the proposed framework, the quantum module is designed to function as a global feature extractor that complements the local feature extraction of convolutional neural networks. Classical image features are mapped to quantum states through parameterized encoding, allowing spatial information to be represented in a high-dimensional Hilbert space.

The quantum layer is implemented as a variational quantum circuit composed of basic rotational gates, including R_x (θ), R_y(θ) and R_z(θ), which perform rotations of the qubit state around the corresponding axes of the Bloch sphere. These gates introduce trainable parameters into the quantum circuit, enabling adaptive feature transformation during end-to-end learning. To capture higher-order correlations across encoded features, multi-qubit entangling operations are employed. In addition to controlled-NOT (CNOT) operations, the circuit incorporates the controlled-controlled-NOT (CCNOT) gate, which acts on three qubits and enables conditional interactions based on multiple qubits. Following quantum state evolution, measurement operations are applied to extract expectation values from the qubits. These measurements produce a compact classical feature vector that encodes global structural information from the input. This quantum-enhanced representation is particularly advantageous where subtle structural variations and contextual dependencies must be captured across large-scale, heterogeneous disaster scenes.

Within the Dual-Path Quantum–Classical U-Net architecture, these quantum-derived features are fused with convolutional features extracted by the U-Net encoder. While the CNN branch captures fine-grained spatial and textural details, the quantum branch models global contextual dependencies. This complementary integration enhances feature representation and improves the model's ability to distinguish subtle damage patterns in complex post-disaster environments.

4. PROPOSED METHODOLOGY

The proposed methodology addresses building damage assessment through a two-stage framework that consists of building segmentation and damage classification. A detailed description is mentioned in sections 4.1 and 4.2. Each of these modules is explained in the following sections. The figure illustrates the overall architecture of the proposed Dual-Path Quantum–Classical U-Net (DQC-U-Net) framework for accurate building localization and multi-class damage assessment using pre- and post-disaster satellite images. A hybrid quantum-classical neural network is implemented using TFQ. The architecture consists of quantum blocks and convolutional blocks, as shown in Figure 1. The design leverages quantum gates for state preparation and entanglement, followed by measurement to extract classically interpretable outputs. The input image patch is sized 256×256.

4.1 Quantum-based building localization

The objective of building localization is to identify building and non-building region**s** using only the pre-disaster images. This ensures reliable localization without interference from post-disaster distortions such as debris, collapse, or shadows. A linear projection layer is applied to pre disaster images to transform each image patch into a fixed-dimensional feature embedding suitable for quantum encoding. The image embeddings are first passed through a Quantum Encoder, where image patches are encoded into quantum states using angle encoding. Then, the encoder performs initial feature extraction, followed by quantum entanglement operations using CNOT and a variational quantum layer, enabling the modeling of global correlations and structural relationships within the scene. Quantum measurement is then applied to collapse the quantum states into compact classical feature representations. These features are processed through a Quantum Bottleneck, which acts as a compression layer that retains the most informative global features. After quantum measurement, the obtained expectation values are transformed into classical feature vectors. Since the measured quantum outputs are low-dimensional representations, a fully connected projection layer followed by a reshape operation is applied to reconstruct spatial feature maps compatible with the decoder architecture. These reconstructed feature maps are then forwarded to the decoder for building footprint localization and dense feature reconstruction. The output is forwarded to a decoder, which progressively up samples the features to produce a binary building mask indicating building and non-building regions. This mask serves as a spatial prior for the subsequent damage assessment stage.

Algorithm: DQC-U-Net for Post-Disaster Building Damage Assessment

Step 1: The pre-disaster image X_{pre} is partitioned into fixed-size image patches.
to control input dimensionality and enable efficient processing.
Step 2: Each patch is transformed into a compact global feature representation using a quantum encoder:
a. The classical pixel values are encoded into a quantum state using angle encoding(R_y) summarize patch-level structural.
b. Perform entanglement using the CNOT gate that models long-range structural relationships across the scene.
Step 3: The encoded global features are passed through a bottleneck layer.
Step 4: A decoder progressively up samples the features to generate a binary building footprint mask.

Step 5: Building mask obtained, which identifies building and non-building regions.
Step 6:Again, both pre- and post-disaster images are processed in parallel.
Step 7: Apply:
a. Local spatial features using a conventional U-Net encoder.
b. Global structural features using the quantum encoder, using phase rotation with amplitude rotation, and then entangling using the CCNOT gate, followed by measurement.
Step 8: Global (structural) and local (textural) features are merged using a cross-attention-based fusion mechanism, allowing the model to focus on damage-relevant regions.
Step 9: The fused features are restricted to building areas using the building mask obtained in Step 5, reducing interference from background objects such as vegetation and roads.
Step 10: A decoder reconstructs the fused features to produce a pixel-wise damage map, assigning each building pixel to one of four damage levels.

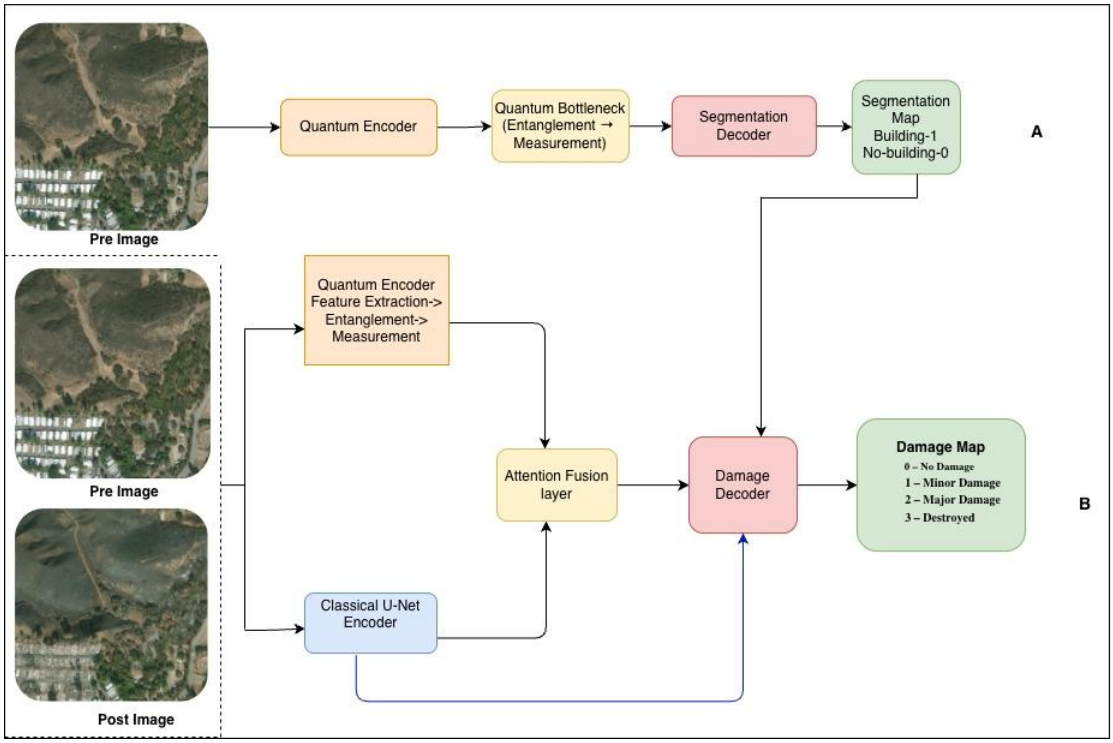


Figure 1. Proposed DQC-U-Net architecture for post-disaster building damage assessment. The framework consists of a quantum-based building localization stage using pre-disaster images, followed by a dual-path quantum–classical damage assessment stage that fuses global quantum features and local U-Net features to produce four-class building damage maps

4.2 Dual path quantum classical damage assessment

In the second stage, both pre- and post-disaster images are used to assess damage severity at the building level. This stage adopts a dual-path that integrates quantum and classical feature representations. The pre- and post-disaster images are processed by a quantum encoder, extracting global correlation-aware features through entanglement and measurement. In parallel, the pre- and post-disaster images are processed by a classical U-Net encoder that serves as a convolutional feature extractor, capturing spatial, textual, and semantic details from the input remote sensing image before fusion with quantum features. The quantum block operations enable the network to capture both global structural dependencies and local spatial details for accurate building damage assessment. The extracted features are integrated through an attention-based fusion module, which selectively emphasizes salient spatial and contextual information relevant to damage assessment. The attention-refined fused feature maps are then directly forwarded to the decoder, where hierarchical feature reconstruction and upsampling are performed. The final output is a four-class damage classification map, where each pixel within the building footprint is assigned one of the following labels: 0, No Damage; 1, Minor Damage; 2, Major Damage; and 3, Destroyed.

Separating building localization from damage assessment and integrating quantum-based global feature modeling with classical spatial feature extraction reduces background confusion and false positives caused by vegetation, shadows, and surrounding objects. The inclusion of quantum entanglement–driven features enhances the model's awareness of global structural damage patterns that are difficult to capture using local convolutional operations alone. At the same time,

classical CNN features preserve fine spatial details, enabling more accurate and fine-grained damage classification across multiple severity levels. Owing to the compact and information-rich nature of quantum representations, the overall architecture maintains computational efficiency while delivering robust and reliable performance in complex post-disaster scenarios.

5. IMPLEMENTATION

The proposed approach is a two-stage process: building footprint extraction and damage assessment. The first instance takes input from the pre-disaster image only, which is encoded using a quantum encoder, followed by entanglement and variational layers that capture global structural features, which are compressed via a quantum bottleneck and decoded to generate a binary building mask. In the next stage, pre- and post-disaster images are processed through a quantum encoder and a U-net encoder in parallel. The quantum mask and classical mask are fused using cross attention and decoded with the decoder. This enables accurate four-class building damage classification while reducing background confusion and false positives. The detailed framework is discussed below.

5.1 Dataset description

The xBD dataset is the largest publicly available optical satellite dataset for building segmentation and post-disaster damage assessment. It consists of high-resolution (~0.3–0.8 m) pre- and post-disaster optical images collected from multiple disaster types, including earthquakes, floods, wildfires, hurricanes, and tsunamis. The pre-disaster images capture the scene before the event, while post-disaster images show the same area after damage has occurred. The mask files contain pixel-level building footprint annotations, distinguishing building and non-building regions. The label files are stored in JSON format, where each building polygon is annotated with a unique ID and an associated damage class.

Buildings are categorized into four damage levels: no damage, minor damage, major damage, and destroyed, as shown in Table 1. Overall, the dataset comprises 22,068 images, covering an area of approximately 45,361.79 km² and featuring 850,736 annotated buildings. For efficient training, each image was divided into 16 non-overlapping patches of 256 × 256 pixels. The data were split into training, validation, and test sets using an 80:10:10 ratio. Data augmentation techniques, including flipping, rotation, and contrast enhancement, were applied during training to improve model robustness and reduce overfitting. In this study, we sampled 8,000 image pairs from the xBD dataset, consisting of paired pre- and post-disaster high-resolution optical satellite images annotated with building footprints and damage labels.

5.2 Experimental setup

The model was implemented using TensorFlow and TensorFlow Quantum (TFQ). Training was performed end-to-end using the Adam optimizer with a learning rate of 1×10^{-4}, batch size of 16, and 100 epochs. A dropout rate of 0.3 was applied for regularization. Binary cross-entropy and categorical cross-entropy losses were used for building segmentation and damage classification, respectively, while transposed convolution (UpConv) layers were employed in the decoder for up-sampling.

Table 1. Building Damage Classification Criteria

Damage Level	Damage Category	Structural Description
0	No Damage	Buildings remain structurally intact with no visible evidence of water intrusion, roof or wall damage, surface cracking, or fire-related effects.(Green)
1	Minor Damage	Buildings exhibit limited localized damage, such as partial roof element loss, minor façade cracking, superficial burn marks, or surrounding water without structural deformation.(Yellow)
2	Major Damage	Buildings show significant structural degradation, including partial wall or roof collapse, severe deformation caused by debris or volcanic flow, or prolonged exposure to surrounding water, mud, or sediment.(Orange)
3	Destroyed	Buildings are completely or near-completely destroyed, characterized by total structural collapse, extensive charring, full submergence under water or mud, or physical absence of the structure.(Red)

5.3 Quantum-based building localization

The first step in the proposed building damage assessment framework is the accurate localization of building footprints. Since post-disaster images often contain severe distortions such as debris, collapsed structures, occlusions, and shadows that can degrade localization performance, only pre-disaster images are utilized in this stage. The first stage of the proposed framework focuses on accurate building footprint localization using only pre-disaster imagery, thereby avoiding the adverse effects of post-disaster distortions such as debris, collapsed structures, occlusions, and shadows. Given a pre-disaster image X_{pre}

$$X_{pre} \in R^{H*W*C} \quad (1)$$

where H, W, and C represent the height, width, and number of spectral channels, respectively. The image is partitioned into non-overlapping patches $\{x_i\}_N^{i=1}$. Before quantum encoding, pixel intensities are normalized as six

and eight qubits to [0,1] and mapped to quantum rotation angles respectively as:

$$x_{ik} = \frac{p_{ik}}{\max(p)}, k = 1 \ldots\ldots 6 \quad (2)$$

$$x_{ik} = \frac{p_{ik}}{\max(p)}, k = 1 \ldots\ldots 8 \quad (3)$$

Each patch is encoded into a 6-qubit and 8-qubits quantum encoder, where each qubit is initialized in the ground state $|0\rangle = |0\rangle^{\otimes 6}$ and$|0\rangle = |0\rangle^{\otimes 8}$.

$$\theta_i = \pi \cdot x_{ik} \quad (4)$$

Angle encoding is performed using R_y rotations:

$$K= \\ |\psi_i\rangle = \otimes R_y(\theta_{ik})|0\rangle \quad (5)$$

To capture long-range structural dependencies that are difficult to model using local convolutional operations alone, the encoded state is passed through a quantum bottleneck implemented as a variational entanglement layer. Entanglement is then introduced using a sequence of CNOT gates between neighboring qubits,

$$\phi_i\rangle = (\prod_{k=1}^{5} CNOT_{(q_k, q_{k+1})})|\Psi_i\rangle \quad (6)$$

$$|\phi_i\rangle = (\prod_{k=1}^{7} CNOT_{(q_k, q_{k+1})})|\Psi_i\rangle, \quad (7)$$

optionally combined with trainable single-qubit rotations to enhance representational capacity. This entangled quantum representation enables efficient modeling of global structural correlations across spatial regions of the image. Quantum measurement is then performed in the Pauli-(Z) basis to collapse the entangled state into a compact classical representation for six and eight qubits, respectively, as:

$$f_q = [\langle\phi_i|Z_1|\phi_i\rangle, \ldots\ldots, \langle\phi_i|Z_6|\phi_i\rangle] \quad (8)$$

$$f_q = [\langle\phi_i|Z_1|\phi_i\rangle, \ldots\ldots, \langle\phi_i|Z_8|\phi_i\rangle] \quad (9)$$

resulting in a 6-dimensional and 8-dimensional quantum feature vector that retains the most informative global context. The measured quantum features are subsequently fed into a classical decoder. The decoder output produces a pixel-level binary map.

$$M \in \{0,1\}^{H*W} \quad (10)$$

where 1 denotes building pixels, and 0 denotes non-building regions. This mask effectively suppresses background elements such as vegetation, roads, water bodies, and shadows, enabling precise building footprint extraction. This building mask effectively suppresses background elements such as vegetation, roads, water bodies, and shadows, thereby providing accurate building localization and a reliable spatial prior for subsequent damage classification.

5.4 Dual path quantum classical damage assessment

In the second stage, both pre-disaster and post-disaster images are used to assess the severity of building damage. Let the paired inputs be denoted as:

$$X_{pre}, X_{post} \in R^{H \times W \times C} \quad (11)$$

where H, W, and C represent the height, width, and number of spectral channels, respectively. The image is partitioned into non-overlapping patches x_i^{pre}, x_i^{post}, which are processed by a quantum classical dual-path encoder. Followed by normalizing on each image patch, phase rotation is applied first, followed by amplitude rotation, to encode classical pixel intensities into quantum states. The circuit diagram is shown in Figure 2.

For each patch x_i, pixel intensities p_{ik} are normalized to the range [0,1]. All qubits are initialized in the computational basis state and mapped to quantum rotation angles $\theta_{ik} = \pi x_{ik}$. Phase encoding is first applied using parameterized R_z rotations for six and eight qubits, followed by amplitude encoding via R_y rotations:

$$U_{phase}(x_i) = \prod_{k=1}^{6} R_z(\theta_{ik}) \quad (12)$$

$$U_{phase}(x_i) = \prod_{k=1}^{8} R_z(\theta_{ik}) \quad (13)$$

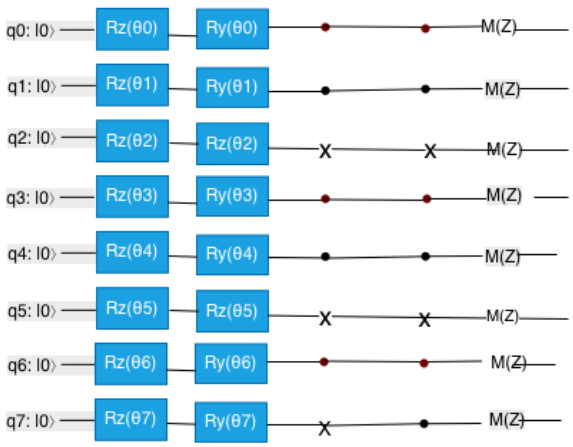


Figure 2. Structure of quantum circuits: Each qubit undergoes a sequence of parameterized rotations using Rz and Ry gates. CCNOTs gate introduces entanglement between the qubits, followed by measurement in the computational (Z) basis to generate quantum-enhanced feature representations

$$U_{amp}(x_i) = \prod_{k=1}^{6} R_y(\theta_{ik}) \quad (14)$$

$$U_{amp}(x_i) = \prod_{k=1}^{8} R_y(\theta_{ik}) \quad (15)$$

The resulting quantum-encoded state is expressed as:

$$|\psi_i\rangle = U_{amp}(x_i)\, U_{phase}(x_i)\, |0\rangle \quad (16)$$

where n is the number of qubits, $U_{phase}(\cdot)$ represents the phase rotation operator, and $U_{amp}(\cdot)$denotes the amplitude rotation operator.

This phase–amplitude encoding jointly preserves intensity magnitude and relative structural information within each image patch. This enables the quantum encoder to generate compact and expressive feature representations suitable for subsequent entanglement and correlation modeling. After phase–amplitude encoding, multi-qubit entanglement is introduced to capture higher-order correlations among spatial features within each image patch. Toffoli (CCNOT) gates are employed to model conditional dependencies between groups of qubits, enabling the representation of complex structural relationships that are difficult to capture using pairwise entanglement alone.

$$U_{ent} = \prod_{(a,b,c)\in T} \mathrm{CCNOT}\ a, b \rightarrow c \quad (17)$$

where T denotes the set of selected qubit triplets from the six and eight-qubit register. The fully entangled state is given by:

$$|\phi_i\rangle = U_{ent|}\psi_i\rangle \quad (18)$$

This configuration enables the quantum circuit to encode nonlinear, global interactions among encoded image features. By leveraging CCNOT-based entanglement, the quantum encoder captures higher-order structural dependencies, such as joint roof-wall context relationships. This enables better discrimination between similar damage classes, like major damage and destroyed. Following entanglement, the quantum state for six qubits and eight qubits is measured to obtain a compact classical representation:

$$z_i=[\langle\phi_i|Z_1|\phi_i\rangle,\langle\phi_i|Z_2|\phi_i\rangle,\ldots,\langle\phi_i|Z_6|\phi_i\rangle] \quad (19)$$

$$z_i=[\langle\phi_i|Z_1|\phi_i\rangle,\langle\phi_i|Z_2|\phi_i\rangle,\ldots,\langle\phi_i|Z_8|\phi_i\rangle] \quad (20)$$

At each encoder stage, convolutional blocks composed of 3×3 convolutions, batch normalization, and ReLU activations are applied, followed by 2×2 max-pooling to reduce spatial resolution while increasing feature depth. After the L down-sampling stages, the encoder produces the deepest feature map:

$$f_c \in R^{H'\times W'\times d}, H' = \frac{H}{2^L}, W' = \frac{W}{2^L} \quad (21)$$

where d denotes the number of channels. This representation captures rich semantic and contextual information, serving as the classical feature input for quantum–classical fusion. Skip connections connect encoder features to the corresponding decoder layers. They preserve fine spatial details and reduce information loss during downsampling. This improves boundary accuracy and provides reliable classical features for quantum–classical fusion.

Further, the outputs from the quantum and classical encoders are merged in a fusion layer, where global quantum-derived features and local CNN features complement each other. F_q and F_c denote the quantum and classical features, respectively. These quantum embeddings serve as queries within a cross-attention fusion block, enabling them to guide the selection and refinement of classical U-Net features. Unlike the CNN encoder that focuses on local textures and spatial continuity, the quantum representation captures long-range correlations and structural relationships, providing complementary information for damage assessment. The query, key, and value are computed as

$$Q=W_qF_q,\ K=W_kF_c,\ V=W_vF_c \quad (22)$$

The cross-attention weights are calculated as

$$Attn = \frac{QK^T}{\sqrt{d_k}} \quad (23)$$

The attended features are obtained by

$$F_{att}= \mathrm{Attn}\odot V \quad (24)$$

Finally, the fused representation is formed using residual addition and channel-wise concatenation:

$$F_{fused}=\mathrm{Concat}(F_q,\ F_{att}+F_q) \quad (25)$$

This fusion allows the model to reason about global structural consistency and local damage cues jointly. The fused features are then combined with the building mask obtained from Stage 1, resulting in a fusion output constrained to building regions, effectively suppressing background elements such as vegetation, water bodies, and roads.

$$F_{bm} = F_{fused} \odot M. \quad (26)$$

The fused building-aware features are passed to a decoder, which gradually up samples the feature maps to the original image resolution.

$$F^{(l-1)} = \sigma(\mathrm{Conv}(\mathrm{UpConv}(F^{(l)}))) \quad (27)$$

where UpConv(·) denotes transposed convolution and σ(·) is the ReLU activation. Skip connections from the U-Net encoder are incorporated as:

$$F^{(l-1)} = Concat(F^{(l-1)}, F_c^{(l-1)}) \quad (28)$$

preserving spatial details and improving boundary reconstruction. The final prediction is generated using a 1×1 convolution:

$$Y = Conv_{1*1}(F^{(0)}) \quad (29)$$

producing a building-aware damage map at full resolution. This high-resolution damage map enables precise localization of damage severity at the building level, facilitating reliable post-disaster assessment and decision support. By integrating quantum-extracted features with classical processing, the proposed framework enhances feature discrimination and supports accurate damage inference. The quantum layer captures global structural dependencies, while the attention mechanism selectively emphasizes damage-relevant regions. Together, these components improve both classification accuracy and model interpretability.

6. EVALUATION METRICS

To predict the performance and effectiveness, several criteria are commonly used. The evaluated metrics include:

Accuracy: This metric represents the overall correctness by calculating the proportion of true classified images. Accuracy is calculated as:

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \quad (30)$$

Precision: Precision quantifies the ratio of accurately predicted positive pixels to the total number of pixels predicted as positive. It is calculated as:

$$\text{Precision} = TP / (TP + FP). \quad (31)$$

Recall: It offers an overall assessment of the model's performance across all categories; however, it may present a distorted view in cases of imbalanced datasets.

$$\text{Recall} = TP / (TP + FN). \quad (32)$$

F1-Score: The F1-score represents the harmonic mean of precision and recall.

$$\text{F1-Score} = 2 \times (\text{Precision} \times \text{Recall}). \quad (33)$$

Intersection over Union: IoU is a key evaluation metric used primarily in object detection and image segmentation tasks, including flood mapping and damage detection in disaster analysis.

$$\text{IoU} = TP / (TP + FP + FN). \quad (34)$$

7. RESULTS AND DISCUSSION

To evaluate the performance of the proposed model, extensive experiments were conducted on the xBD dataset. This study aims to develop a robust framework for addressing bi-temporal remote sensing challenges in post-disaster building damage assessment using high-resolution satellite imagery.

7.1 Results

To assess the reliability of the proposed model, its segmentation performance was evaluated. Non-building regions were assigned a value of zero, while building regions were masked and labeled as one. We have evaluated the performance of DQC-U-Net on the xBD dataset using Precision, Recall, F1-score, and IoU metrics. Table 2 presents the quantitative comparison of building localization. Further, compares the building localization performance across five models: U-Net++ (Zhou et al., 2018), DeepLabv3+ (Chen et al., 2018), HRNet (Sun et al., 2019), U-NeXt (Valanarasu & Patel, 2022), and the proposed DQC-U-Net.

Table 2. Quantitative Comparison of Building Localization Performance

Models	F1-score	Precision	Recall	IoU
U-Net++ (Zhou et al., 2018)	0.8243	0.8189	0.8298	0.7041
DeepLabv3 + (Chen et a., 2018)	0.8225	0.8228	0.8223	0.7091
HRNet (Sun et al., 2019)	0.8464	0.8476	0.8452	0.7233
U-NeXt (Valanarasu & Patel, 2022)	0.8539	0.8391	0.8692	0.7276
DQC-U-Net(six qubits)	0.8641	0.8531	0.8755	0.761
DQC-U-Net(eight qubits)	0.8964	0.859	0.8801	0.7691

Among conventional architectures, U-NeXt achieves the strongest performance, with an F1-score of 0.8539 and an IoU of 0.7276, followed by HRNet, whereas U-Net++ and DeepLabv3+ achieve comparatively lower scores. The proposed DQC-U-Net with 6 qubits surpasses all classical baselines, achieving an F1-score of 0.8641 and IoU of 0.7610, corresponding to absolute gains of +1.02% in F1 and +3.34% in IoU over U-NeXt. The improvement in IoU is particularly notable, indicating enhanced spatial overlap and boundary consistency in segmented building regions.

Further increasing the qubits from six to eight yields additional improvements, with F1-score rising to 0.8694 and IoU to 0.7691. This reflects incremental gains of +0.53% in F1 and +0.81% in IoU compared to the six-qubit configuration. Precision and recall also increase simultaneously with 0.8590 and 0.8801, respectively, demonstrating that the performance enhancement is balanced and not driven by a trade-off between false positives and false negatives. The monotonic improvement from classical models to six-qubit and then eight-qubit configurations confirms that controlled

expansion of the quantum state space strengthens feature representation and inter-feature correlation modeling.

Figure 3 shows a consistent upward shift across all four metrics when transitioning from classical networks to the quantum-enhanced variants. The eight-qubit configuration attains the highest values in every metric, while the relative separation in IoU is more pronounced than in F1-score, further emphasizing improved segmentation quality at object boundaries. The proposed quantum-assisted architecture delivers systematic and statistically meaningful performance improvements, validating its effectiveness for high-resolution remote sensing building segmentation. Furthermore, Table 3 and Figure 4 show the multi-class building damage classification performance of different models. Among the classical architectures, U-NeXt achieves the best baseline performance with an F1-score of 0.7611 and an IoU of 0.6544. HRNet follows with an F1-score of 0.7533 and an IoU of 0.6389. DeepLabv3+ and U-Net++ obtain lower scores across all metrics. This indicates the limitations of purely convolution-based models in capturing complex multi-class damage patterns.

Table 3. Quantitative Comparison of Multi-Class Building Damage Classification Performance.

Models	F1-score	Precision	Recall	IoU
U-Net++	0.7259	0.7321	0.7036	0.6125
DeepLabv3+	0.7301	0.7465	0.7169	0.6242
HRNet	0.7533	0.7641	0.7339	0.6389
U-NeXt	0.7611	0.7801	0.7542	0.6544
DQC-U-Net(six qubits)	0.7752	0.7921	0.7643	0.6743
DQC-U-Net(eight qubits)	0.7813	0.8023	.7685	0.6834

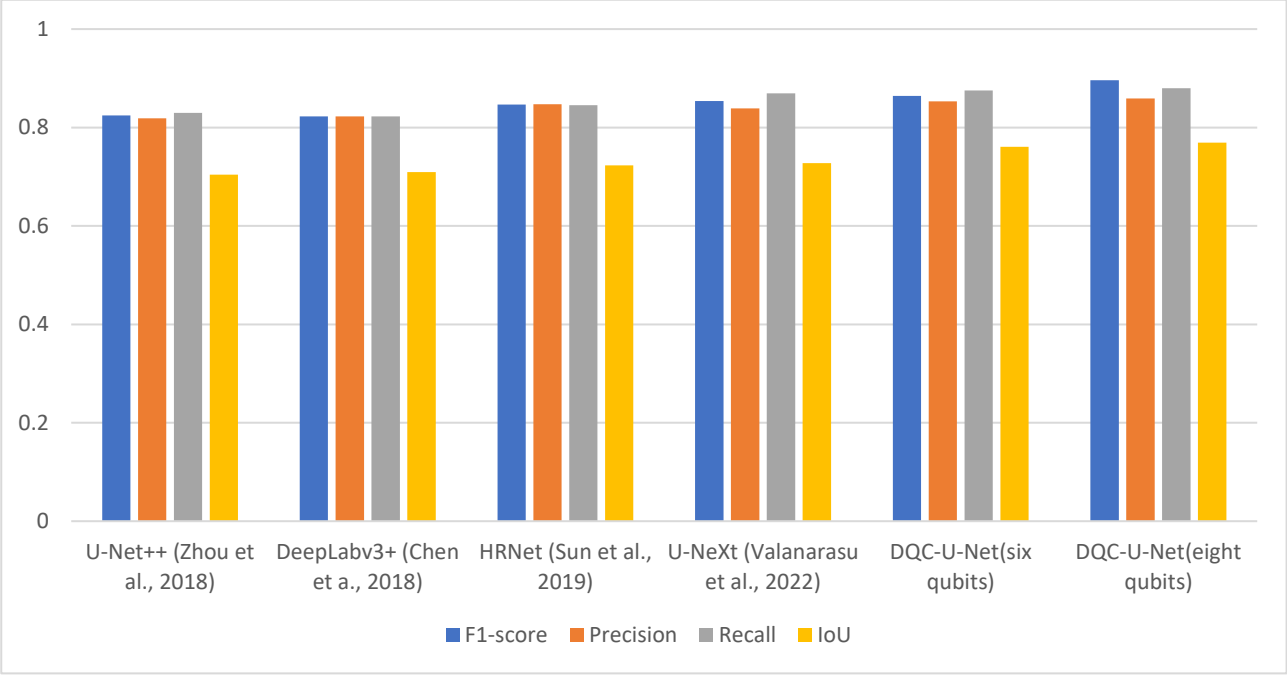


Figure 3. Building segmentation performance comparison of U-Net++, DeepLabv3+, HRNet, U-NeXt, and the proposed DQC-U-Net using F1-score, Precision, Recall, and IoU metrics.

Increasing the qubit count from six to eight provides further gains. F1-score rises from 0.7752 to 0.7813 (+0.61%). The IoU improves by +1.13%. Precision increases from 0.7921 to 0.8023, and recall also shows improvement. These results indicate stronger class discrimination and better boundary delineation. Although the gain from six to eight qubits is moderate, the improvement is consistent across all metrics. The upward trend confirms that increasing quantum dimensionality enhances representational capacity without destabilizing training. Overall, the results validate the effectiveness of the proposed quantum-assisted framework for fine-grained multi- class building damage assessment.

The next stage of the evaluation assessed multi-class building damage classification performance using the proposed quantum and U-Net framework with satellite imagery. Table 4 summarizes the quantitative results across four damage levels using six and eight-qubit configurations. The model achieves the strongest performance in the no damage and destroyed categories, with an F1 score of 0.928 for six qubits to 0.934 for eight qubits. This indicates reliable identification of intact buildings with minimal false positives and false negatives.

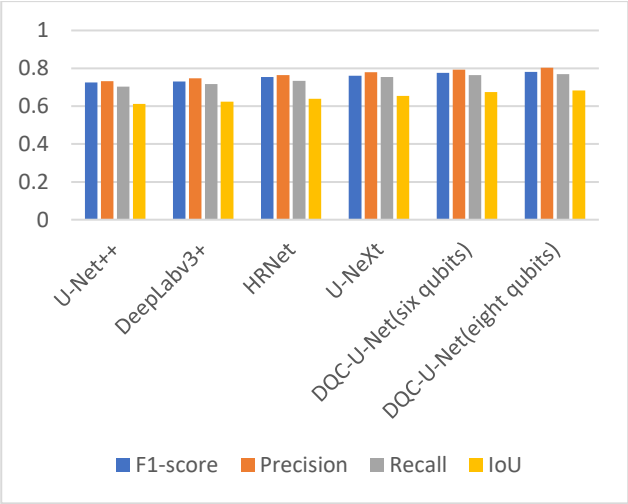


Figure 4. Multi-class building damage classification performance comparison across different models

This strong performance can be attributed to the distinct visual characteristics of undamaged structures, which are easier to separate from damaged classes.

For the Destroyed class, the model also provides effective recognition of severely damaged buildings, which typically exhibit prominent structural collapse patterns and debris signatures. The balanced precision and recall values suggest that the model can accurately detect destroyed buildings while maintaining a low misclassification rate. The F1-score improves from 0.872 to 0.881 for six and eight qubits, respectively. The minor damage class remains the most challenging. The F1-score improves from 0.619 to 0.634. Precision and recall increase slightly but remain lower than in other classes.

This reflects the difficulty of identifying subtle structural changes in satellite imagery. Overall, eight qubits consistently improve performance across all classes. The gains are moderate but stable. The model performs best for extreme damage levels and shows improvement for intermediate categories. For major damage, the F1-score rises from 0.826 to 0.839 when using eight qubits. Recall improves from 0.892 to 0.905. This shows effective detection of heavily damaged buildings, which is important.

Table 4. Distribution of buildings across damage classes (no damage, minor, major, destroyed) for different disaster types.

Damage Level	No. of qubits	F1-score	Precision	Recall
No damage	Six	0.928	0.946	0.911
	Eight	0.934	0.95	0.918
Minor damage	Six	0.619	0.593	0.647
	Eight	0.634	0.609	0.661
Major damage	Six	0.826	0.77	0.892
	Eight	0.839	0.782	0.905
Destroyed	Six	0.872	0.893	0.852
	Eight	0.881	0.901	0.861

In contrast, the minor damage class exhibits the lowest performance. F1-score improves from 0.619 to 0.634 for six and eight qubits. This reflects the inherent difficulty in distinguishing minor damage from both undamaged and moderately damaged structures, as subtle façade-level or roof-level changes are often visually ambiguous in satellite imagery. Performance of eight qubits consistently improves across all classes. The gains are moderate but stable. Overall, the class-wise results highlight the model's strong capability in identifying extreme damage levels, while revealing the remaining challenges in accurately discriminating intermediate damage categories.

Figure 5 illustrates the disaster-wise performance of the proposed model across different hazard types. Overall, the model demonstrates stable and consistently high performance across all disaster categories, indicating strong generalization capability to diverse damage patterns and disaster mechanisms. For earthquake and hurricane events, the model achieves the highest scores across all metrics, reflecting its effectiveness in capturing large-scale structural collapse, roof failure, and debris distribution that are characteristic of these disasters. Tsunami and tornado scenarios also exhibit strong performance, with high F1-score and IoU value, suggesting reliable detection of both structural displacement and severe inundation-induced damage.

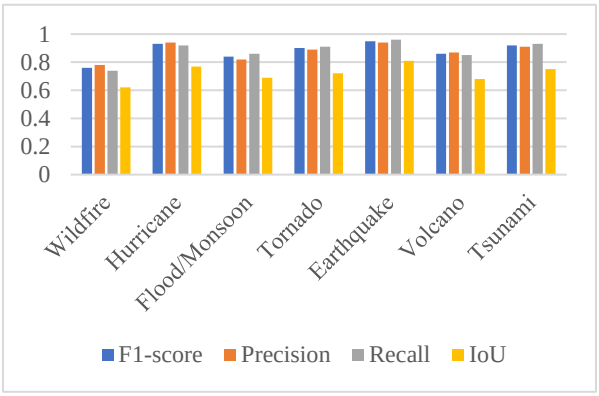


Figure 5. Performance of the proposed model across different disaster types in terms of F1-score, precision, recall, and Intersection over Union (IoU)

The lower IoU values are observed for wildfire and flood events. This can be attributed to the more diffuse and fewer geometric damage patterns associated with these disasters, where buildings may remain structurally intact

while exhibiting partial roof damage, smoke occlusion, waterlogging, or sediment coverage. Such conditions introduce ambiguity in building boundary delineation and damage severity estimation, leading to reduced overlap accuracy. Despite these challenges, the performance gap across disaster types indicates that the model effectively captures disaster-invariant structural cues while adapting to disaster-specific damage characteristics. This robustness across heterogeneous disaster scenarios highlights the suitability of the proposed approach for large-scale, multi-hazard post-disaster building damage assessment.

7.2 Ablation study

To systematically evaluate the contribution of each component, an ablation study was performed by incrementally integrating the proposed modules into a baseline U-Net, as presented in Table 5. The baseline model (V1) achieves an F1-score of 0.7192 and an IoU of 0.6110, serving as the reference point. Introducing the quantum encoder (V2) improves performance to an F1-score of 0.7335 and IoU of 0.6350. This gain reflects the encoder's role as a nonlinear feature projection module that captures global structural relationships, enabling more expressive representations than purely convolutional features. Further improvement is observed with the addition of the dual-path architecture (V3), which raises the F1-score to 0.7558 and IoU to 0.6592. By jointly leveraging quantum-derived global features and CNN-based local spatial features, the model benefits from complementary representations, enhancing its ability to distinguish between damage categories. Incorporating the attention-based fusion mechanism (V4) yields additional gains, achieving an F1-score of 0.7712 and an IoU of 0.6780. The attention module facilitates selective feature integration by emphasizing damage-relevant regions and suppressing irrelevant background information, thereby improving localization and classification accuracy.

Table 5. Ablation study evaluating the contribution of key components in the proposed DQC-U-Net framework for building damage classification

Model Variant	Quantum Encoder	Dual Path	Attention	Building Mask	F1-score	IoU
V1: U-Net (baseline)	×	×	×	×	.7192	0.6110
V2:Quantum encoder	✓	×	×	×	.7335	.6350
V3: Dual Path	✓	✓	×	×	.7558	.6592
V4:Attention	✓	✓	✓	×	.7712	.6780
V5: DQC-U-Net	✓	✓	✓	✓	.7813	.6834

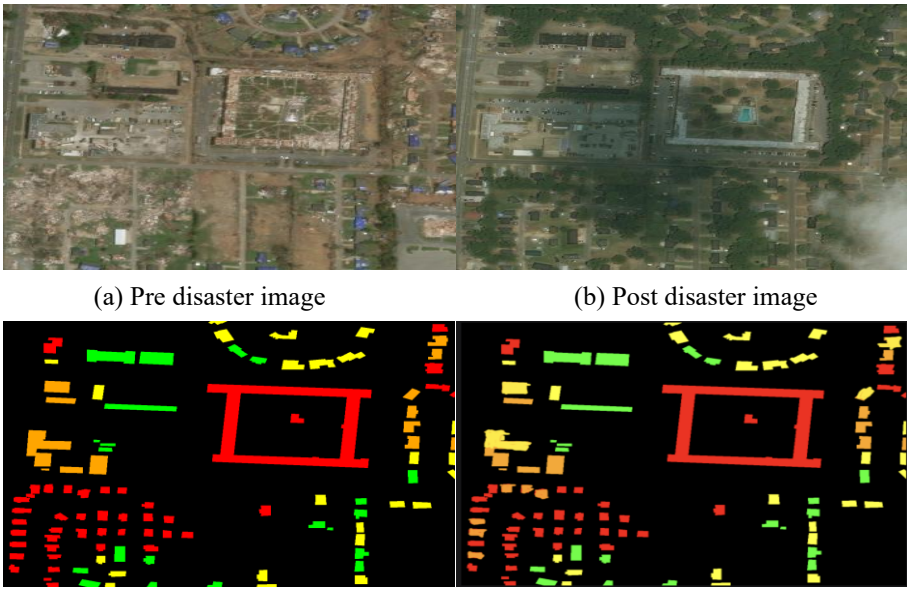
Finally, the complete DQC-U-Net model (V5), which includes the building-aware masking strategy, attains the best performance with an F1-score of 0.7813 and IoU of 0.6834. The masking operation constrains predictions to building regions, effectively reducing background interference from non-structural elements such as vegetation and roads, and resulting in more precise damage estimation. Overall, the consistent improvement across all configurations demonstrates that each component contributes positively to the framework. The combination of quantum-based global representation, dual-path feature extraction, attention-guided fusion, and spatial masking leads to a robust and accurate solution for fine-grained building damage assessment.

7.3 Discussion

Dual-Path Quantum–Classical U-Net (DQC-U-Net) framework for building damage assessment supports disaster response operations. The proposed framework integrates quantum neural networks with a classical U-Net to improve damage mapping accuracy from high-resolution satellite imagery using six and eight qubits. DQC-U-Net follows a two-stage workflow aligned with operational needs. In the first stage, quantum-enhanced building segmentation accurately localizes building footprints using pre-disaster imagery. This ensures stable asset identification despite post-disaster debris, shadows, and occlusions. In the second stage, a dual-path quantum-classical module analyses pre- and post-disaster images to classify buildings into four damage levels: no damage, minor damage, major damage, and destroyed. To further observe the results of the proposed work for eight qubits, two sample images in the verification set are shown, as shown in Figures 6 and 7. The pre-disaster image depicts an area with relatively low spatial resolution and sparsely distributed buildings. This makes it difficult to accurately delineate building footprints. The post-disaster image exhibits substantial visual degradation caused by debris, partial collapses, shadows, and occlusions, further obscuring structural boundaries and complicating damage assessment. The ground-truth mask provides fine-grained multi-class damage annotations, highlighting the challenges associated with distinguishing varying damage severities at the building level.

In Figure 6, the model successfully captures the large building footprints and the surrounding severely affected areas. Minor discrepancies are observed at building boundaries and among smaller structures, where subtle damage patterns are more difficult to distinguish. These differences primarily occur in regions with dense building distributions or visually ambiguous damage signatures. Overall, the predicted damage map closely

aligns with the ground truth, particularly for severely damaged and destroyed buildings, achieving reliable localization and accurate damage severity classification. Minor discrepancies are observed at building boundaries and among smaller structures, where subtle damage patterns are more difficult to distinguish. These differences primarily occur in regions with dense building distributions or visually ambiguous damage signatures.

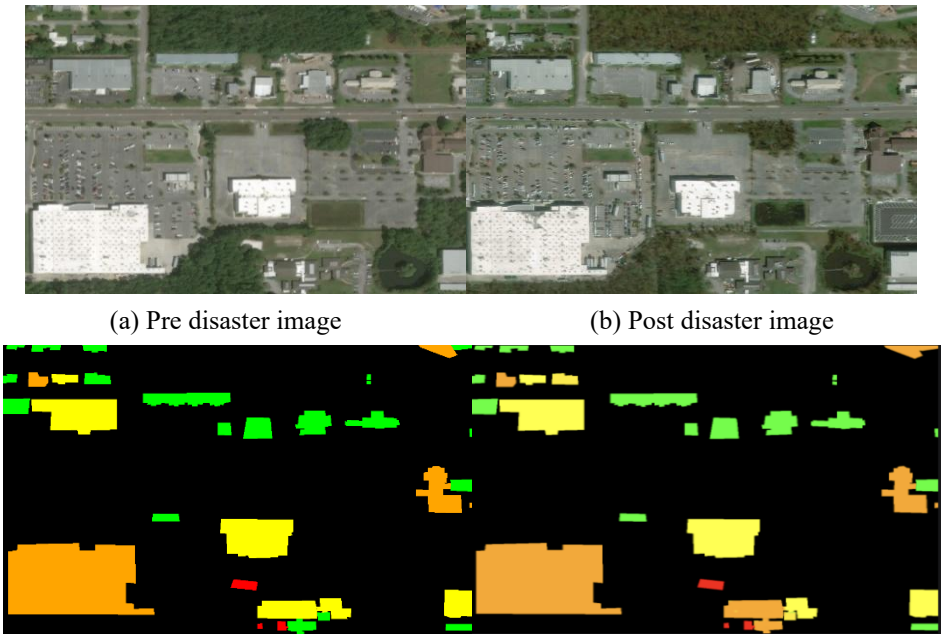


(a) Pre disaster image (b) Post disaster image

(c) Mask (d) Predicted output

Figure 6. Qualitative comparisons of chosen samples are shown: (a) Pre-event image, (b) Post-event image, (c) Mask, (d) Predicted output of the proposed model

In Figure 7, the predicted output demonstrates that discrepancies between the mask and prediction are evident for minor-damaged buildings. Several small structures labeled as minor damage in the ground truth are incorrectly assigned to the major damage class. These errors arise because minor damage often manifests as subtle roof deformation or localized texture change, which is difficult to distinguish from normal appearance variations in optical imagery.



(a) Pre disaster image (b) Post disaster image

(c) Mask (d) Predicted Output

Figure 7. Qualitative comparisons of chosen samples are shown: (a) Pre-event image, (b) Post-event image, (c) Mask, (d) Predicted output of the proposed model

Overall, the predicted damage map preserves the spatial organization and captures the dominant damage patterns observed in the ground-truth mask. The results confirm that the proposed model performs reliably for large and severely affected buildings while maintaining reasonable consistency for minor damage categories, demonstrating its robustness and applicability for post-disaster building damage assessment in complex urban environments. Nevertheless, the overall visual consistency between the ground truth mask and the predicted output demonstrates the model's strong capability in localizing buildings and accurately classifying damage severity, validating its effectiveness for post-disaster building damage assessment from high-resolution remote sensing imagery.

The quantum pathway captures global structural relationships, while the classical pathway preserves local spatial details. The compact quantum representations reduce computational overhead, thereby demonstrating the robustness and effectiveness of the proposed approach. Experimental results demonstrate consistent improvements in F1-score and IoU across multiple disaster scenarios. These findings highlight the potential of DQC-U-Net as a reliable and scalable tool for post-disaster assessment of building damage.

7.4 Limitation

Despite its strong overall performance, the proposed DQC-U-Net framework exhibits several limitations. First, the integration of quantum-inspired feature modeling components increases computational complexity, which may restrict scalability when processing large volumes of high-resolution satellite imagery. Second, the proposed two-stage architecture is sensitive to inaccuracies in the initial building footprint extraction stage. Although DQC-U-Net demonstrates reliable performance in identifying minor and severe damage, particularly for clearly distinguishable damage patterns, its accuracy degrades in scenarios where building boundaries are ambiguous or partially occluded. In such cases, small or dense areas may lead to undetected damaged structures. Finally, building footprint annotations derived from large-scale datasets may contain spatial misalignments and labelling inaccuracies. This introduces noise during training and affects boundary precision and damage-level discrimination.

8. TRANSFER LEARNING

To rigorously evaluate cross-event generalization, a strict zero-shot transfer-learning protocol is adopted. Specifically, the model parameters corresponding to the highest validation F1-score obtained during multi-event training are frozen and directly applied to five completely unseen disaster events, without any fine-tuning or domain adaptation. The held-out events include the 2018 Sulawesi earthquake and tsunami, 2017 Puebla earthquake, 2018 Woolsey Fire, 2013 Moore tornado, and 2017 Nepal floods. These events are strictly excluded from the training stage to ensure a genuine cross-domain evaluation scenario that realistically simulates deployment on newly occurring disasters. Building segmentation performance is quantified using F1-score and Intersection-over-Union (IoU), which jointly measure classification balance and spatial overlap consistency. Under this zero-shot transfer setting, DQC-U-Net achieves consistently strong segmentation performance across heterogeneous hazard types. The quantitative results are summarized in Table 6.

Table 6. Transfer learning performance on unseen disaster events.

Disaster Event	F1-score	IoU
Sulawesi Tsunami	0.83	0.71
Puebla Earthquake	0.81	0.69
Woolsey Fire	0.79	0.66
Moore Tornado	0.80	0.68
Nepal Flood	0.78	0.65
Average	**0.80**	**0.68**

The proposed framework achieves an average F1 Score of 0.80 and an IoU of 0.68, with event-wise F1 scores ranging from 0.78 to 0.83. The narrow performance variation (approximately 0.05) across events indicates stable behavior under substantial domain shift. Relatively higher accuracy is observed in earthquake and tsunami scenarios, where structural collapse typically generates pronounced geometric discontinuities and fragmented rooftop patterns that are effectively captured by the learned representations. Wildfire and tornado scenes exhibit comparable robustness, suggesting that the model successfully generalizes to irregular debris distributions and boundary distortions induced by wind and fire dynamics. Slightly reduced performance is observed in the flood scenario (F1 = 0.78, IoU = 0.65), which can be attributed to water-induced occlusion, partial rooftop submergence, and increased spectral ambiguity. These factors degrade edge IoU consistently remains above 0.65, confirming reliable spatial alignment between predicted and ground-truth building masks. The stable IoU trend across all events further indicates that DQC-U-Net preserves structural coherence despite substantial radiometric and contextual variability. Beyond binary building segmentation, multi-class damage classification is also evaluated under the same transfer protocol. The average class-wise F1-scores across the five unseen events are reported in Table 7.

Table 7. Class-wise damage classification performance under transfer learning.

Damage Level	Average F1-score
No Damage	0.88
Minor Damage	0.76
Major Damage	0.72
Destroyed	0.81
Macro Avg.	**0.79**

The highest transfer performance is achieved for the no-damage class, with F1 0.88, indicating that intact structural patterns are consistently preserved across domains. The destroyed category also demonstrates strong robustness, with an F1 of 0.81, particularly in earthquake- and tsunami-dominated scenes where collapse signatures produce distinctive debris fields and roof disintegration patterns. In contrast, performance decreases for minor are 0.76 and major 0.72 damage classes, reflecting the intrinsic ambiguity between intermediate degradation levels. The categories often differ only in subtle roof deformation, partial structural compromise, or localized texture variation, making them more susceptible to domain shift. The performance degradation is most evident in flood imagery, where water coverage reduces visible structural cues, and in wildfire scenes where burn marks obscure rooftop textures. Despite these challenges, the macro-average F1-score of 0.79 under zero-shot transfer demonstrates that DQC-U-Net retains substantial discriminative capability across all four damage categories without event-specific adaptation.

The relatively modest gap between destroyed and major classes further suggests that the model captures hierarchical structural degradation patterns rather than hazard-specific artefacts. Overall, the transfer learning results confirm that DQC-U-Net maintains stable segmentation and damage classification performance when deployed on previously unseen disasters. The low inter-event variance and consistent multi-class robustness indicate that the learned representations encode transferable structural priors, thereby enhancing resilience under cross-hazard and cross-geographical domain shifts. These findings substantiate the suitability of the proposed framework for real-world disaster response scenarios where rapid adaptation to newly emerging events is essential.

9. CONCLUSION

This study presented a Dual-Path Quantum Classical U-Net (DQC-U-Net) for fine-grained building segmentation and damage assessment using satellite imagery. By integrating quantum-enhanced global feature modeling with classical convolutional neural networks, the proposed approach effectively captures both long-range structural dependencies and local spatial details, addressing key challenges in post-disaster building damage mapping. The two-stage assessment strategy separates building localization from damage classification, contributing to improved consistency by restricting damage prediction to relevant structural regions. The dual-path architecture further enables complementary feature learning through the combination of quantum-derived global representations and CNN-based local features.

Comprehensive experiments and comparisons with U-Net models demonstrated the effectiveness, robustness, and generalization capability of the proposed approach across diverse disaster scenarios. The strong performance of DQC-U-Net highlights the potential of quantum–classical hybrid learning to enhance disaster damage assessment from remote sensing imagery. Furthermore, the framework's ability to generalize to unseen geographic regions indicates its suitability for scalable, real-world post-disaster response applications. Future work will focus on exploring more advanced quantum encoding strategies, extending the framework to multi-temporal and multi-modal data sources, and improving computational efficiency for large-scale deployment. Overall, this study provides a promising direction for integrating quantum-inspired techniques into operational disaster damage assessment systems.

Acknowledgement: The authors would like to thank the institutions that provided the datasets and computational resources used in this research. The authors also appreciate the valuable feedback that helped improve the quality of this work.

References:

Ahmadi, S. A., Mohammadzadeh, A., Yokoya, N., & Ghorbanian, A. (2024). BD-SKUNet: Selective-kernel U-Nets for building damage assessment in high-resolution satellite images. *Remote Sensing*, *16*(1), 182. https://doi.org/10.3390/rs16010182

Biamonte, J., Wittek, P., Pancotti, N., Rebentrost, P., Wiebe, N., & Lloyd, S. (2017). Quantum machine learning. *Nature*, *549*, 195–202. https://doi.org/10.1038/nature23474

Braik, A. M., & Koliou, M. (2024). Automated building damage assessment and large-scale mapping by integrating satellite imagery, GIS, and deep learning. *Computer-Aided Civil and Infrastructure Engineering*, *39*, 2389–2404. https://doi.org/10.1111/mice.13197

Chen, L., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. In *Lecture notes in computer science* (pp. 833–851). https://doi.org/10.1007/978-3-030-01234-2_49

Daudt, R. C., Le Saux, B., & Boulch, A. (2018). Fully convolutional siamese networks for change detection. In *2018 25th IEEE International Conference on Image Processing (ICIP)* (pp. 4063–4067). IEEE. https://doi.org/10.1109/ICIP.2018.8451652

Deng, L., & Wang, Y. (2022). Post-disaster building damage assessment based on an improved U-Net. *Scientific Reports*, *12*, 15862. https://doi.org/10.1038/s41598-022-20114-w

Dong, L., & Shan, J. (2013). A comprehensive review of earthquake-induced building damage detection with remote sensing techniques. *ISPRS Journal of Photogrammetry and Remote Sensing*, 84, 85–99. https://doi.org/10.1016/j.isprsjprs.2013.06.011

Economic Times. (2025). Himachal Pradesh monsoon toll reaches 92 lives; infrastructure damage estimated. Retrieved from https://economictimes.indiatimes.com

Ge, J., Tang, H., Yang, N., & Hu, Y. (2023). Rapid identification of damaged buildings using incremental learning with transferred data from historical natural disaster cases. *ISPRS Journal of Photogrammetry and Remote Sensing*, 195, 105–128. https://doi.org/10.1016/j.isprsjprs.2022.11.010

Gerard, S., Borne-Pons, P., & Sullivan, J. (2024). A simple, strong baseline for building damage detection on the xBD dataset. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2401.17271

Ghaffarian, S., & Kerle, N. (2019). TOWARDS POST-DISASTER DEBRIS IDENTIFICATION FOR PRECISE DAMAGE AND RECOVERY ASSESSMENTS FROM UAV AND SATELLITE IMAGES. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. XLII-2/W13. 297-302. https://doi.org/10.5194/isprs-archives-XLII-2-W13-297-2019.

Gupta, R., Hosfelt, R., Sajeev, S., Patel, N., Goodman, B., Doshi, J., . . . Gaston, M. (2019). xBD: A Dataset for Assessing Building Damage from Satellite Imagery. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1911.09296

Havlíček, V., Córcoles, A. D., Temme, K., Harrow, A. W., Kandala, A., Chow, J. M., & Gambetta, J. M. (2019). Supervised learning with quantum-enhanced feature spaces. *Nature*, *567*(7747), 209–212. https://doi.org/10.1038/s41586-019-0980-2

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 770–778). https://doi.org/10.1109/CVPR.2016.90

Iglovikov, V., & Shvets, A. (2018). TernausNet: U-Net with VGG11 Encoder Pre-Trained on ImageNet for Image Segmentation. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1801.05746

Janalipour, M., & Mohammadzadeh, A. (2015). Building damage detection using Object-Based Image Analysis and ANFIS from High-Resolution Image (Case study: BAM Earthquake, Iran). *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *9*(5), 1937–1945. https://doi.org/10.1109/jstars.2015.2458582

Jiang, Z., Wu, Y., Huang, L., & Gu, M. (2024). FDB-Net: Fusion double branch network combining CNN and transformer for medical image segmentation. *Journal of X-Ray Science and Technology*, *32*(4), 931–951. https://doi.org/10.3233/xst-230413

Kakooei, M., & Baleghi, Y. (2023). Fusion of vertical and oblique images using intra-cluster classification for building damage assessment. *Computers & Electrical Engineering*, *105*, 108536, ISSN 0045-7906. https://doi.org/10.1016/j.compeleceng.2022.108536.

Milletari, F., Navab, N., & Ahmadi, S.-A. (2016). V-net: Fully convolutional neural networks for volumetric medical image segmentation. *2016 Fourth International Conference on 3D Vision (3DV)*, 565–571. Stanford, CA, USA: IEEE. https://doi.org/10.1109/3DV.2016.79

Putri, A. F. S., Widyatmanti, W., & Umarhadi, D. A. (2022). Sentinel-1 and Sentinel-2 data fusion to distinguish building damage level of the 2018 Lombok Earthquake. *Remote Sensing Applications: Society and Environment*, *26*, 100724. ISSN 2352-9385. https://doi.org/10.1016/j.rsase.2022.100724.

Qing, Y., Ming, D., Wen, Q., Weng, Q., Xu, L., Chen, Y., Zhang, Y., & Zeng, B. (2022). Operational earthquake-induced building damage assessment using CNN-based change detection. I*nternational Journal of Applied Earth Observation and Geoinformation*, 112, 102899, ISSN 1569-8432. https://doi.org/10.1016/j.jag.2022.102899.

Rathje, E. M., Crawford, M., Woo, K., & Neuenschwander, A. (2005). Damage patterns from satellite images of the 2003 Bam, Iran earthquake. *Earthquake Spectra*, 21, 295–307. https://doi.org/10.1193/1.2101047

Reuters. (2026). China's natural disasters led to $35 billion losses. Retrieved from https://www.reuters.com

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-NET: Convolutional Networks for Biomedical Image Segmentation. In *Lecture notes in computer science* (pp. 234–241). https://doi.org/10.1007/978-3-319-24574-4_28

Schuld, M., Sinayskiy, I., & Petruccione, F. (2015). An introduction to quantum machine learning. *Contemporary Physics*, 56(2), 172–185. https://doi.org/10.1080/00107514.2014.964942

Shen, Y., Zhu, S., Yang, T., Chen, C., Pan, D., Chen, J., Xiao, L., & Du, Q. (2021). BDANet: Multiscale CNN with cross-directional attention for Building Damage Assessment From Satellite Images. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1-14. https://doi.org/10.1109/TGRS.2021.3080580.

Su, J., Bai, Y., Wang, X., Lu, D., Zhao, B., Yang, H., Mas, E., & Koshimura, S. (2020). Technical Solution Discussion for Key Challenges of Operational Convolutional Neural Network-Based Building-Damage Assessment from Satellite Imagery: Perspective from Benchmark xBD Dataset. *Remote Sensing*, *12*(22), 3808. https://doi.org/10.3390/rs12223808

Sun, K., Xiao, B., Liu, D., & Wang, J. (2019). Deep High-Resolution Representation learning for human pose estimation. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1902.09212

Valanarasu, J. M. J., & Patel, V. M. (2022). UNeXt: MLP-based rapid medical image segmentation network. In L. Wang, Q. Dou, P. T. Fletcher, S. Speidel, & S. Li (Eds.), *Medical image computing and computer assisted intervention – MICCAI 2022* (Vol. 13435, pp. 23–33). Springer. https://doi.org/10.1007/978-3-031-16443-9_3

Wang, J., Sun, K., Cheng, T., Jiang, B., Deng, C., Zhao, Y., . . . Xiao, B. (2019). Deep High-Resolution Representation learning for visual recognition. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1908.07919

Wu, C., Wang, Z., Xia, J., & Zhang, F. (2025). CFP-SwinT: building damage classification during floods using cross fusion pyramid Swin-T. *International Journal of Digital Earth*, *18*(1). https://doi.org/10.1080/17538947.2025.2523490

Wu, F., Gong, L., Wang, C., Zhang, H., Zhang, B., & Xie, L. (2016). Signature analysis of building damage with TerraSAR-X new staring spotlight mode data. *IEEE Geoscience and Remote Sensing Letters*, *13*(11), 1696-1700. https://doi.org/10.1109/LGRS.2016.2604841

Xia, H., Wu, J., Yao, J., Zhu, H., Gong, A., Yang, J., ... & Mo, F. (2023). A deep learning application for building damage assessment using ultra-high-resolution remote sensing imagery in Turkey earthquake. *International Journal of Disaster Risk Science*, *14*(6), 947-962. https://doi.org/10.1007/s13753-023-00526-6

Xiao, Y., & Mostafavi, A. (2025). DamageCAT: A deep learning transformer framework for typology-based post-disaster building damage categorization, *International Journal of Disaster Risk Reduction*, Volume 128, 105704, ISSN 2212-4209. https://doi.org/10.1016/j.ijdrr.2025.105704

Zheng, Z., Zhong, Y., Wang, J., Ma, A., & Zhang, L. (2021). Building damage assessment for rapid disaster response with a deep object-based semantic change detection framework: From natural disasters to man-made disasters. *Remote Sensing of Environment*, *265*, Article 112636. https://doi.org/10.1016/j.rse.2021.112636

Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., & Liang, J. (2018). UNET++: a nested U-Net architecture for medical image segmentation. *Lecture Notes in Computer Science*, *11045*, 3–11. https://doi.org/10.1007/978-3-030-00889-5_1

Ankita Dixit
Jaypee Institute of Information Technology
Noida,
India
ankitadixit177@gmail.com
ORCID 0009-0006-3651-7580

Indu Chawla
Jaypee Institute of Information Technology
Noida
India
indu.chawla@mail.jiit.ac.in
ORCID 0000-0002-5333-9185