





Article

Damage Attention-Aware Dense Layered Framework for Surface Crack Classification

Molaka Maruthi ¹, Munisamy Shyamala Devi ^{2,3}, Young Choi ^{4,*} and Chang-Yong Yi ^{5,*}

¹ School of Architecture, Civil, Environment and Energy Engineering, Kyungpook National University, 80 Daehak-ro, Buk-gu, Daegu 41566, Republic of Korea; maruthi0396@knu.ac.kr

² Department of Robot and Smart System Engineering, Kyungpook National University, 80 Daehak-ro, Buk-gu, Daegu 41566, Republic of Korea; drshyamala@knu.ac.kr

³ National Satellite Information Research Institute, Kyungpook National University, 80 Daehak-ro, Buk-gu, Daegu 41566, Republic of Korea

⁴ Earth Turbine, 36, Dongdeok-ro 40-gil, Jung-gu, Daegu 41905, Republic of Korea

⁵ Intelligent Construction Automation Center, Kyungpook National University, 80 Daehak-ro, Buk-gu, Daegu 41566, Republic of Korea

* Correspondence: youngch5@naver.com (Y.C.); cyyi@knu.ac.kr (C.-Y.Y.)

Abstract

Accurate surface defect classification is a critical requirement in structural health monitoring and infrastructure inspection, where defects, including cracks, spalling, delamination and noncrack regions, often appear with low-contrast and complex background textures. Motivated by the need for a robust and discriminative framework that can enhance defect visibility and focus learning on damage-critical regions, this research proposes a novel damage-aware DenseNet-201 (DA-DenseNet-201) model for surface defect classification. As a critical novelty, a damage-aware adaptive contrast-limited adaptive histogram equalisation (DAC) filtering strategy is introduced as a preprocessing stage. The proposed DAC filter dynamically adjusts contrast enhancement parameters based on damage indicators, selectively amplifying crack edges and defect textures while preserving healthy surface regions and suppressing noise. Building on this method, enhanced images are processed using a pretrained DenseNet-201 backbone, retaining the benefits of dense feature propagation and efficient gradient flow. To strengthen the discriminative learning of DA-DenseNet-201 further, an attention refinement block is integrated into the network, combining channel attention to emphasise defect-relevant feature responses and spatial attention to localise damage regions accurately. In addition, a multiscale feature fusion mechanism aggregates feature maps from multiple dense blocks to capture fine-grained crack patterns, texture-level degradation and high-level semantic damage information. Extensive experiments conducted on surface defect datasets demonstrate its effectiveness, achieving a superior classification accuracy of 98.93%, along with notable improvements in sensitivity, specificity and the intersection over union compared with state-of-the-art models. These results confirm that the proposed DA-DenseNet-201 provides a reliable and high-performance solution for automated surface defect classification.



Academic Editor: Giuseppina Uva

Received: 26 March 2026

Revised: 22 May 2026

Accepted: 5 June 2026

Published: 9 June 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article

distributed under the terms and

conditions of the [Creative Commons](https://creativecommons.org/licenses/by/4.0/)

[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: accuracy; attention; classification; contrast-limited adaptive histogram equalisation (CLAHE); damage-aware adaptive contrast-limited adaptive histogram equalization (DAC); deep learning; feature extraction; filtering; multiscale feature fusion; structural health monitoring

1. Introduction

Surface condition assessment is a critical aspect of structural health monitoring and infrastructure maintenance, particularly for concrete and masonry systems exposed to ageing, repeated loading, and harsh environmental conditions [1,2]. Visible surface defects, such as cracks, delamination and spalling, often act as early warning indicators of deeper deterioration mechanisms, including corrosion, fatigue, moisture ingress and progressive material loss [1,3]. In practical inspection workflows, surface images are widely collected using handheld cameras, mobile devices and uncrewed aerial vehicle platforms because they are inexpensive and fast to acquire and do not require contact [4,5]. Despite these advantages, manual visual inspection remains time-consuming, subjective and strongly dependent on inspector experience, lighting conditions, viewing angles and background clutter, potentially leading to inconsistent assessments and missed defects [1,6]. The above limitations motivated the development of automated and reliable image-based defect recognition systems to support rapid inspection and consistent decision-making. Among surface-level damage, cracks are prevalent and critical indicators of structural deterioration in civil infrastructure, mechanical components and industrial assets [7]. If not detected and addressed in a timely manner, such surface-level cracks can propagate into severe structural failures, resulting in safety hazards, economic loss and increased maintenance costs. Consequently, accurate, reliable surface crack detection is vital in ensuring the structural integrity, serviceability and long-term sustainability of engineered systems [8]. Traditionally, surface crack inspection has heavily relied on manual visual assessment performed by trained inspectors. Although this approach remains widely practised, it is time-consuming, labour-intensive and subject to human error and subjectivity. Moreover, manual inspection becomes increasingly impractical for large-scale infrastructure, such as highways, bridges, tunnels, and high-rise buildings [9]. In addition to manual visual inspection, both destructive and non-destructive diagnostic methods have been extensively investigated for structural damage assessment in concrete infrastructures. Destructive testing approaches, such as core extraction, compression testing and laboratory-based material characterization, provide direct evaluation of material strength and internal deterioration; however, these methods are often time-consuming, costly and may further damage the inspected structure. To overcome these limitations, several non-destructive testing (NDT) techniques, including ultrasonic testing, acoustic emission, infrared thermography, radiographic testing, vibration-based monitoring and ground penetrating radar, have been widely adopted for crack detection and subsurface damage evaluation [10,11]. These methods enable safer and more practical structural inspection without damaging the structural components. Nevertheless, conventional NDT techniques may still suffer from limitations related to environmental noise, complex signal interpretation, operator dependency, limited spatial localization capability and reduced performance under complex real-world conditions. Consequently, recent studies have increasingly focused on integrating automated image processing, machine vision and deep learning techniques for reliable and efficient structural damage assessment [12,13]. To address these limitations, automated crack detection and classification techniques based on image processing and computer vision have been extensively explored over the past two decades [14].

Early methods predominantly relied on handcrafted features, thresholding techniques, edge detection and morphological operations. Although these approaches have demonstrated promising results under controlled conditions, their performance often significantly degrades in real-world scenarios due to variations in lighting, surface texture, background noise and crack morphology [15]. Recently, deep learning (DL) models have demonstrated strong performance in vision-based defect analysis due to their ability to learn hierarchical feature representations directly from raw images. Convolutional neural networks (CNNs)

have been widely adopted for crack detection and surface defect classification because they can capture discriminative patterns, including edges, texture irregularities and local structural discontinuities [16]. Recent studies have further extended CNN and Transformer-based architectures for automatic post-fire reinforced concrete damage detection and assessment, demonstrating improved capability in identifying multiple fire-induced damage categories and severity levels [17]. Recent advancements have also explored AI-driven sensing frameworks integrating distributed fiber optic sensing and TCN-Transformer architectures for monitoring damage evolution and structural behavior under complex environmental conditions, demonstrating the growing potential of intelligent sensing-assisted structural health monitoring systems [18]. Nevertheless, robust surface defect classification remains challenging in real-world scenarios [19]. Surface cracks and deterioration patterns often exhibit low contrast, varying widths, irregular geometry and strong background interference from stains, shadows, rough textures, surface markings or illumination changes. These factors can degrade feature learning when models are trained solely on raw images, leading to confusion between defect and nondefect regions and reduced generalisation across diverse surface conditions [20]. Furthermore, many conventional single-stream CNN architectures treat all spatial regions with equal importance, despite defect regions typically occupying a small portion of the image. Without mechanisms to emphasise damage-relevant regions and capture multiscale characteristics, models may overfit to background textures rather than learning true defect signatures [21].

To address these challenges, this study proposes a damage attention-aware dense layered network 201 (DA-DenseNet-201) for multiclass surface defect classification. Although DenseNet-201 is a powerful deep feature extractor, its conventional architecture is not optimised for surface damage inspection, where cracks, spalling and delamination often appear as subtle, low-contrast patterns embedded in complex backgrounds. The absence of damage-aware preprocessing, explicit attention mechanisms and multiscale feature integration limits its ability to focus on defect-critical regions and discriminate between visually similar surface conditions. To overcome these shortcomings and improve the robustness, sensitivity and interpretability in surface crack classification, the proposed DA-DenseNet-201 is designed with targeted architectural enhancements for damage-centric feature learning with the following contributions.

- i. A novel damage-aware adaptive contrast-limited adaptive histogram equalisation (CLAHE) filtering mechanism (DAC) is introduced before feature extraction to enhance crack edges, spalling and delamination boundaries selectively while preserving background integrity.
- ii. To guide the network toward defect-relevant information, an attention refinement block is incorporated into the DenseNet-201 backbone. This block combines channel attention, emphasising what damage-related features are important, and spatial attention, indicating where the damage regions are located. The attention-guided feature refinement enables the model to suppress irrelevant background responses and concentrate on critical surface defect regions.
- iii. A dedicated multiscale feature fusion strategy is introduced by aggregating feature maps from dense Blocks 2, 3 and 4, enabling the model to learn fine-scale crack patterns, texture-level surface degradation and high-level semantic damage cues simultaneously.
- iv. A lightweight, regularised classification head is designed by extending DenseNet-201 using batch normalisation (BN), multiple fully connected layers and dropout regularisation. This lightweight head improves feature abstraction, reduces overfitting and enhances generalisation performance, especially under limited or imbalanced training data scenarios.

This article is structured as follows: Section 1 outlines the critical contributions and process of the research. Next, Section 2 reviews related studies and examines prior approaches to surface defect classification using DL models. Then, Section 3 details the research method for the proposed DA-DenseNet-201 model. Section 4 presents the implementation setup, results and performance analysis of DA-DenseNet-201 compared with the baseline methods. Finally, Section 5 concludes the work and suggests promising directions for future research and model enhancements.

2. Background Study

2.1. Surface Crack Models Based on Image Processing

Surface crack models based on image processing have been widely explored as early approaches to automated crack identification due to their simplicity, low computational cost and independence from training data. Classical edge-based operators, such as Sobel, Canny and Laplacian of Gaussian, have been extensively applied to enhance crack boundaries by exploiting intensity discontinuities, although their detection performance is highly sensitive to illumination variations, surface stains and background textures [22]. Threshold-based segmentation techniques, particularly Otsu's method, have demonstrated reasonable effectiveness in extracting crack regions and measuring crack geometry under controlled laboratory and field conditions; however, their robustness significantly deteriorates when applied to complex surfaces with nonuniform lighting environments [23]. Morphological operations, including dilation, erosion, opening and closing, are commonly incorporated to improve crack continuity and suppress noise, yet these operations remain strongly parameter-dependent and require careful tuning for various surface characteristics [24]. Region-based approaches, such as region growing and watershed algorithms, have also been adopted to delineate irregular and branching crack patterns, but they are prone to over-segmentation in the presence of shadows, stains and heterogeneous backgrounds [25]. Texture-based descriptors, including local binary patterns and a grey-level cooccurrence matrix, have further been applied to characterise crack-related texture variations on rough concrete surfaces; nevertheless, their discriminative capability weakens when crack widths are extremely small or when background textures resemble crack features [26]. Overall, comprehensive assessments consistently indicate that, although traditional image-processing approaches are easy to implement and computationally efficient, their limited robustness, poor generalisability, and sensitivity to environmental variability restrict their practical applicability in real-world structural health monitoring systems [27].

2.2. Machine Learning-Based Surface Crack Models

Machine learning (ML)-based crack detection methods have been increasingly adopted to overcome the limitations of purely image-processing-based approaches by enabling data-driven learning of crack characteristics under complex surface and environmental conditions. Traditional ML classifiers, such as support vector machines, k-nearest neighbours, and artificial neural networks, have been employed using handcrafted features extracted from crack images, demonstrating improved robustness against noise and background interference compared to rule-based methods, although their performance depends on feature quality and dataset representativeness [28]. Ensemble learning techniques, including random forest and gradient boosting models, have enhanced crack classification accuracy by capturing nonlinear relationships and reducing overfitting in heterogeneous surface conditions [29]. Several studies have demonstrated that ML-based surface crack identification benefits from integrating texture, geometric and statistical features, enabling reliable discrimination between crack and noncrack regions across materials [30]. To improve detection stability, hybrid frameworks combining classical image preprocessing with

ML classifiers have been proposed, where image enhancement and filtering are applied prior to classification to reduce noise sensitivity and improve feature separability [27]. Recent research has also explored ML-based crack severity assessment by mapping extracted features to crack width or damage levels, providing quantitative indicators for structural condition evaluation [25]. Comparative studies indicate that ML approaches offer a favourable trade-off between accuracy and computational efficiency, making them suitable for practical deployment when training data are limited and hardware resources are constrained [26]. Despite these advantages, existing ML-based crack detection models still require careful feature engineering and hyperparameter tuning, and their generalisability remains challenged when applied to unseen imaging environments [31].

2.3. Deep Learning-Based Fracture Classification Models

Recent advances in DL have significantly enhanced fracture and crack detection across civil, mechanical and medical engineering applications by enabling automatic feature extraction from raw data, overcoming the limitations of traditional image-processing methods that rely on handcrafted features. Early studies have demonstrated the feasibility of CNN-based image classification for concrete crack detection, where one study [32] employed transfer learning with VGG-16 to achieve robust performance under heterogeneous surface and lighting conditions. Building on this, an image-based DL framework [33] was designed for concrete crack classification and quantification, achieving high accuracy in estimating the crack length, width and orientation. Moreover, YOLOv5-based object detection models for pavement crack detection [34] have been explored, highlighting their effectiveness in real-time and large-scale roadway inspections. To address multiscale and internal damage detection, another study [35] integrated wavelet-based multiresolution analysis with CNNs for ultrasonic concrete crack monitoring, enabling early-stage crack identification. Beyond concrete and pavement structures, a DL-based segmentation framework for masonry crack detection [36] enabled real-life crack length measurement, and a semantic segmentation approach for macroscale fracture surface analysis [37] achieved crack-size measurements comparable to expert manual assessments, reducing subjectivity. Expanding DL-based fracture analysis to nonvisual domains, a DL-based acoustic emission clustering framework was proposed to evaluate fatigue cracks in welded bridge joints under operational noise [38]. The applicability of object detection models, such as SSD and YOLOv8, was assessed for automatic midfacial fracture detection from computed tomography (CT) images in medical diagnostics [39]. A comprehensive review [40] confirmed the increasing dominance of DL-based crack detection approaches, identifying semantic segmentation as a promising direction for precise damage characterization.

More recent studies have highlighted the diversification and specialization of DL-based crack and fracture classification across complex structures, materials and operating conditions. For example, one study [41] developed a CNN-based computational framework integrating vision and frequency-response data to identify crack locations and characteristics in engineering structures. Moreover, another study [42] proposed a large-scale CNN model trained on heterogeneous concrete surface images, demonstrating high robustness under varying illumination and surface conditions. In bridge inspection applications, a hierarchical DL framework was introduced that can classify surface information at structural, component and defect levels [43], significantly improving inspection efficiency and interpretability. For pavement engineering, a study [44] proposed a DL fusion model combining SSD and U-Net architectures to perform crack classification, segmentation and geometric parameter estimation. A comprehensive synthesis [45] emphasized the increasing dominance of CNN-based and encoder-decoder-based architectures in automated crack detection and segmentation tasks. An enhanced YOLOv8-based algorithm

was developed targeting architectural heritage structures to detect cracks in fair-faced walls [46], improving accuracy under complex environmental conditions. To address computational efficiency and model interpretability, a one-dimensional data DL framework coupled with explainable artificial intelligence was introduced for real-time concrete crack classification on mobile platforms [47]. Beyond civil infrastructure, an automatic image labelling and DCNN-based framework was designed for multimode damage quantification in 2.5-dimensional woven composites using CT images [48], and the effectiveness of DL was demonstrated in identifying soil types based on desiccation crack patterns under noisy conditions [49]. Finally, pretrained CNN-based crack detection frameworks have consistently outperformed traditional image-processing methods, reinforcing the paradigm shift toward scalable, data-driven and objective fracture assessment systems [50]. Collectively, these studies reflect a clear paradigm shift toward automated, data-driven fracture assessment frameworks that apply advanced DL architectures for accurate, scalable and objective structural integrity evaluation.

2.4. Deep Learning Approaches for Post-Fire Concrete Damage Assessment

Recently, deep learning techniques have demonstrated significant potential for post-fire concrete damage detection and assessment [17] proposed a feature fusion framework integrating mobilenetV3 and swin-transformer for automatic classification of post-fire reinforced concrete (RC) damage, achieving reliable performance for multiple damage categories and severity levels. Ref. [51] developed an enhanced YOLOv5s-D model for multicategory fire damage detection in post-fire RC structural components, including soot, cracks, concrete spalling, and rebar exposure, while also enabling real-time mobile deployment. Similarly, ref. [52] introduced a hybrid CNN-LSTM architecture for autonomous detection of concrete damage under elevated temperature conditions, demonstrating improved capability in identifying fire-induced cracks and surface deterioration. Furthermore, recent review studies by [53,54] highlighted the growing importance of artificial intelligence, computer vision, and advanced post-fire assessment techniques for evaluating the residual structural integrity and damage conditions of concrete structures. Although these studies demonstrate the effectiveness of deep learning techniques for post-fire structural assessment, challenges still remain regarding complex background conditions, damage localization accuracy, computational efficiency, and the generalization capability of models under diverse environmental conditions. Therefore, the development of robust and adaptive deep learning frameworks for reliable crack classification and damage assessment remains an active and important research direction. These studies collectively indicate that deep learning-based analysis has emerged as a promising and efficient approach for rapid, automated, and reliable post-fire structural damage assessment.

2.5. Research Gap and Motivation

Despite substantial progress in applying DL techniques to surface defect and crack classification, several critical limitations remain unaddressed. Most contemporary approaches rely on generic CNN architectures designed for natural image recognition, which have been fine-tuned for defect detection tasks. Such models often fail to account for the unique characteristics of surface damage imagery, including uneven illumination and low contrast between defects and the background. Furthermore, conventional preprocessing typically employs fixed or global contrast enhancement methods, inadequately adapting to localized damage severity, which could suppress subtle defects or amplify noise. Recent developments have further demonstrated the importance of attention-guided learning, adaptive preprocessing, and multiscale feature integration in improving defect classification performance under challenging environmental conditions. Several studies have shown that attention-based CNN

architectures can effectively localize defect-sensitive regions and suppress irrelevant background information, thereby improving classification robustness for fine crack structures and heterogeneous surfaces. Similarly, multiscale feature learning frameworks have gained considerable importance for capturing both local crack textures and high-level semantic damage representations across varying defect sizes and structural conditions. In addition, adaptive image enhancement and contrast-aware preprocessing techniques have been increasingly explored to improve visibility of low-contrast defects under uneven illumination and noisy backgrounds. Despite these advancements, existing studies still largely rely on independent preprocessing and feature extraction stages without jointly optimizing damage-aware enhancement and deep hierarchical feature learning within a unified framework. Therefore, the proposed DA-DenseNet-201 framework aims to bridge this gap by integrating DAC filtering, attention refinement, and multiscale feature fusion into a unified deep learning architecture for robust and generalized surface defect classification.

The motivation for proposing DA-DenseNet-201 arises from the need for a damage-aware, feature-adaptive framework that integrates contrast enhancement with deep hierarchical feature learning. By incorporating a DAC filtering mechanism, the proposed approach dynamically enhances defect-relevant regions while preserving structural continuity, enabling more discriminative feature extraction at early network stages. Coupled with a densely connected architecture, DA-DenseNet-201 facilitates efficient feature reuse, strengthens gradient propagation and captures multiscale damage representations without excessive parameter growth. This approach addresses the identified research gap by aligning preprocessing and network design with the physical and visual characteristics of surface defects, improving detection reliability, generalizability and interpretability in automated structural health monitoring applications. Moreover, recent studies on advanced deep learning, multimodal sensing, and 3D damage localization have shown promising capabilities for automated structural damage assessment. Approaches based on UAV inspection, LiDAR-camera fusion, edge AI systems, and multimodal learning frameworks have improved crack localization, post-disaster reconnaissance, and real-time structural monitoring. However, several limitations still remain regarding computational complexity, real-time deployment, robustness under noisy environments, and accurate localization of fine-scale defects in complex structural scenes. In addition, many existing frameworks require expensive hardware, large annotated datasets, and complex reconstruction processes, limiting their practical applicability for lightweight and adaptive structural health monitoring systems. These limitations highlight the need for efficient, robust, and damage-aware frameworks capable of reliable crack detection and structural assessment under diverse real-world conditions. Table 1 summarizes the inferences from surface crack classification, including limitations.

Table 1. Inferences from the literature review.

	Representative Studies	Inference and Advantages	Limitations
ML models	<ul style="list-style-type: none"> • SVM, random forest, KNN, decision trees [55] • KNN [56] • random forest and gradient boosting [57,58] 	<ul style="list-style-type: none"> • Classify crack vs. noncrack regions based on handcrafted features. • Lower computational cost than DL. • Handle high-dimensional feature spaces with appropriate feature selection. 	<ul style="list-style-type: none"> • Strong dependence on handcrafted feature quality and selection. • Performance degrades under complex backgrounds. • Limited ability to capture spatial context.

Table 1. Cont.

Representative Studies	Inference and Advantages	Limitations	
IP models	<ul style="list-style-type: none"> • Edge detection and morphological operations [59,60] • Thresholding-based methods [61] • Texture descriptors, such as LBP and GLCM [62] 	<ul style="list-style-type: none"> • Detect cracks by identifying intensity gradients, texture variations and edge discontinuities. • Simple, fast and computationally efficient. • Effective in controlled environments with uniform lighting and smooth surfaces. • Do not require labelled training datasets. 	<ul style="list-style-type: none"> • Highly sensitive to noise, stains, shadows and illumination variation. • Ineffective for blurred, discontinuous or low-contrast cracks. • Limited robustness in complex real-world scenarios.
DL models	<ul style="list-style-type: none"> • CNN- and FCN-based methods [63,64] • Deep fully CNN and deep segmentation networks [65–67] 	<ul style="list-style-type: none"> • Automatically learn hierarchical spatial features from raw images. • Enable accurate crack classification, localisation and pixelwise segmentation. • Robust to variations in crack width, orientation, geometry and background complexity. • Achieve high accuracy in real-world conditions with sufficient training data. 	<ul style="list-style-type: none"> • Large labelled datasets and extensive training time are required. • High computational and memory demands during training and inference. • Performance may deteriorate for micro-crack detection and severe class imbalance.
Advanced DL, 3D damage localization models	<ul style="list-style-type: none"> • Knowledge-driven 3D damage mapping and multimodal sensing [68] • YOLO-CLIP multimodal crack detection [69] • Edge AI-based intelligent crack detection [70] • Deep learning post-disaster reconnaissance [71] • Knowledge-based post-disaster road damage detection [72] • 3D vision technologies for crack detection robots [73] 	<ul style="list-style-type: none"> • Provides accurate 3D damage localization and mapping for fire-damaged RC structures. • Combines image and text-based learning to improve crack detection and segmentation. Performs well under blurry crack and complex background conditions. • Enables real-time crack detection using edge devices and embedded AI systems. Reduces inspection time and supports autonomous monitoring. • Detects multiple structural damages such as cracks spalling and exposed rebars using faster R-CNN. Useful for UAV and robotic inspection after disasters. • Detects damaged roads from high-resolution remote sensing images without requiring pre-disaster data. Useful for rapid disaster assessment. • Uses LiDAR and camera fusion for high-precision 3D crack localization and measurement. Achieves submillimeter-level crack measurement accuracy. 	<ul style="list-style-type: none"> • Requires multiple sensing devices and complex 3D reconstruction processing. Performance may reduce in highly noisy or complex environments. • Training process is more complex and requires multimodal supervision data. Mainly focused on pavement crack applications. • Detection performance depends on hardware capability and computational resources. Complex crack patterns may still affect accuracy. • Limited 3D spatial localization capability. Requires large annotated datasets for training. • Depends on rule-based feature extraction and may not perform well for detailed crack localization. • Requires expensive hardware, sensor calibration and higher computational cost for practical deployment.

3. Materials and Methods

3.1. Research Methodology

The proposed DA-DenseNet-201 model was designed to classify four surface crack class types. Figure 1 presents the overall workflow of DA-DenseNet-201. The proposed DADenset-201 begins with systematic data organization to ensure robust learning across diverse surface damage conditions in Stage 1. Extensive data augmentation is applied in Stage 2 to enhance generalizability and mitigate overfitting caused by limited samples. Spatial transformations (e.g., horizontal and vertical flipping, rotation, translation, scaling and zooming) are employed to simulate diverse camera viewpoints and surface orientations encountered during field inspections. Therefore, the augmented dataset considerably enhances robustness to geometric distortions and environmental variations. In Stage 3, the augmented images are subjected to the DAC filtering process. Unlike conventional CLAHE, the proposed method adaptively adjusts contrast enhancement parameters based on local damage-sensitive regions, selectively emphasizing crack edges and defect boundaries without amplifying the background noise.

The histogram analysis confirms improved intensity redistribution, enhancing local contrast and superior edge visibility, which are critical for accurate surface damage recognition. The enhanced images are input into the proposed DA-DenseNet-201 architecture for deep feature extraction and classification in Stage 4. The network employs a DenseNet-201 backbone, with an initial convolution and pooling layer to capture low-level texture features, followed by four dense blocks interleaved with transition layers, progressively extracting hierarchical representations, ranging from fine patterns to high-level semantic damage features. An attention refinement block is integrated after dense feature extraction to improve discriminative learning further, comprising channel attention to identify which damage features are important and spatial attention to localize where the damage occurred. Although the proposed attention refinement block employs a sequential combination of channel and spatial attention mechanisms, the design was intentionally selected to maintain a balance between discriminative feature refinement and computational efficiency. Advanced attention architectures, including Transformer-based self-attention, non-local attention, and cross-scale attention frameworks, generally introduce substantial parameter complexity and require significantly larger training datasets for stable optimisation. Considering the relatively limited structural defect dataset and the objective of achieving efficient surface defect classification, the lightweight cascade attention strategy was adopted in the proposed DA-DenseNet-201 framework. The channel attention stage focuses on identifying defect-sensitive feature channels by modelling inter-channel dependencies, whereas the spatial attention stage enhances localisation of crack boundaries, spalling regions, and delamination structures by emphasising spatially important regions. This sequential refinement process effectively suppresses irrelevant background responses while strengthening defect-discriminative representations. Moreover, the ablation analysis demonstrated that integrating the proposed attention refinement block with DAC filtering and multiscale feature fusion consistently improved classification accuracy, sensitivity, and IoU performance compared with baseline DenseNet-201 configurations.

A multiscale feature fusion module aggregates features from multiple dense blocks to capture fine-grained cracks, mid-level textures and severe damage semantics. Finally, a lightweight classification head with BN, dropout regularisation and a fully connected layer (FCL) produces reliable surface damage predictions. In the final stage, the DA-DenseNet-201 model is evaluated using quantitative performance metrics and comparative analysis. The receiver operating characteristic (ROC) and precision–recall curves are plotted against multiple DL baselines to assess discriminability. The proposed model consistently achieves superior performance, demonstrating higher values for the area under the curve (AUC) and

improved classification accuracy. The final output categorises surface conditions into four classes: crack, no crack, spalling and delamination, providing an effective and automated solution for surface damage inspection and structural health monitoring.

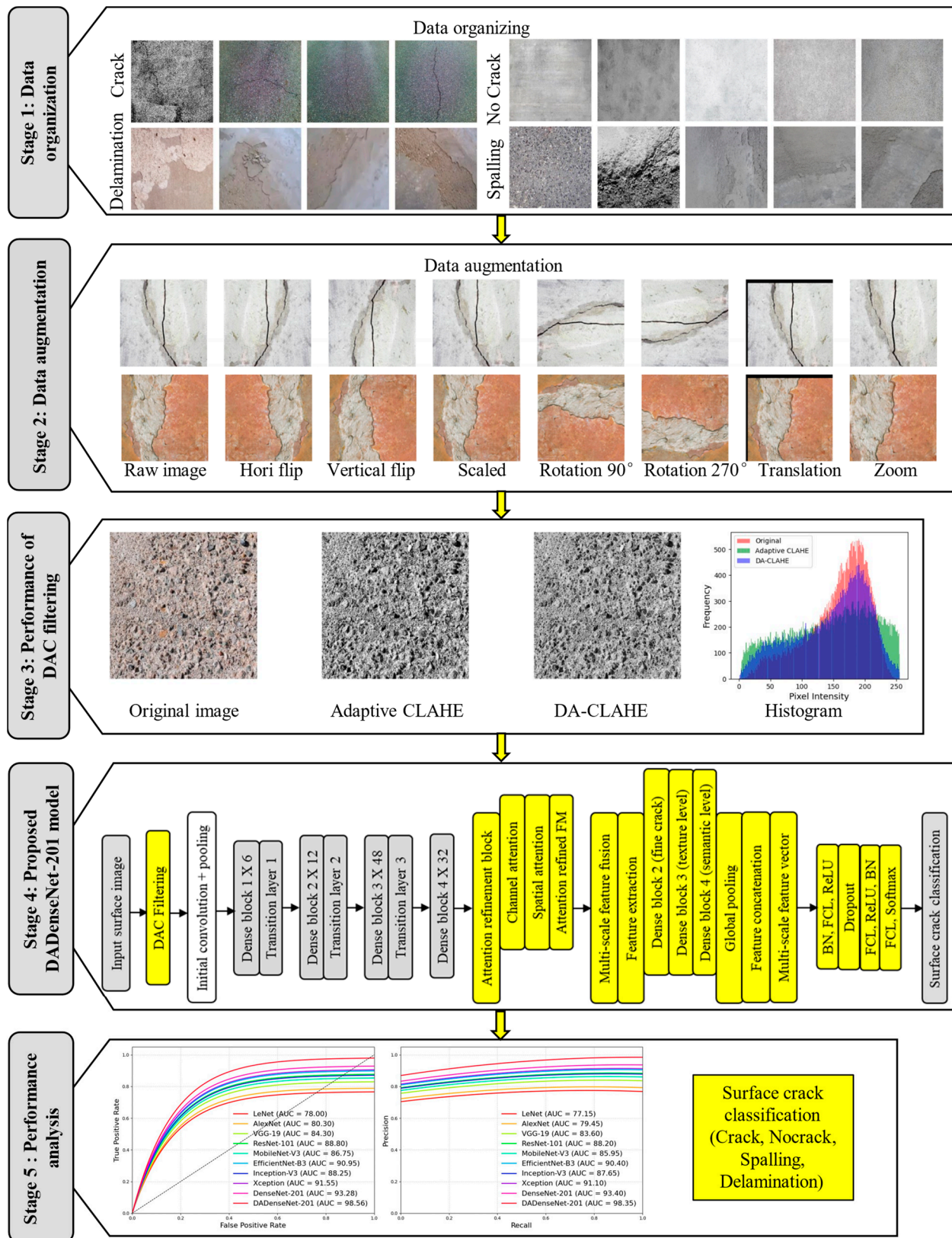


Figure 1. Proposed DA-DenseNet-201 model research method.

3.2. Architectural Innovation of DA-DenseNet-201

Figure 2 illustrates the architectural distinction between the conventional DenseNet-201 and the proposed DA-DenseNet-201 designed for surface crack classification. In the existing DenseNet-201 framework (Figure 2a), the input surface image is processed via an initial convolution and pooling layer, followed by four dense blocks comprising 6, 12, 48 and 32 densely connected convolutional layers. Transition layers are inserted between successive dense blocks for feature map (FM) compression and spatial down-sampling. The dense connectivity pattern promotes feature reuse and alleviates vanishing gradient problems. However, the architecture primarily relies on global average pooling (GAP) and a single FCL followed by a SoftMax classifier. This standard process lacks explicit mechanisms to enhance damage-specific contrast, focus attention on critical defect regions, or integrate multiscale structural information, which are crucial for accurate surface damage interpretation under complex background conditions.

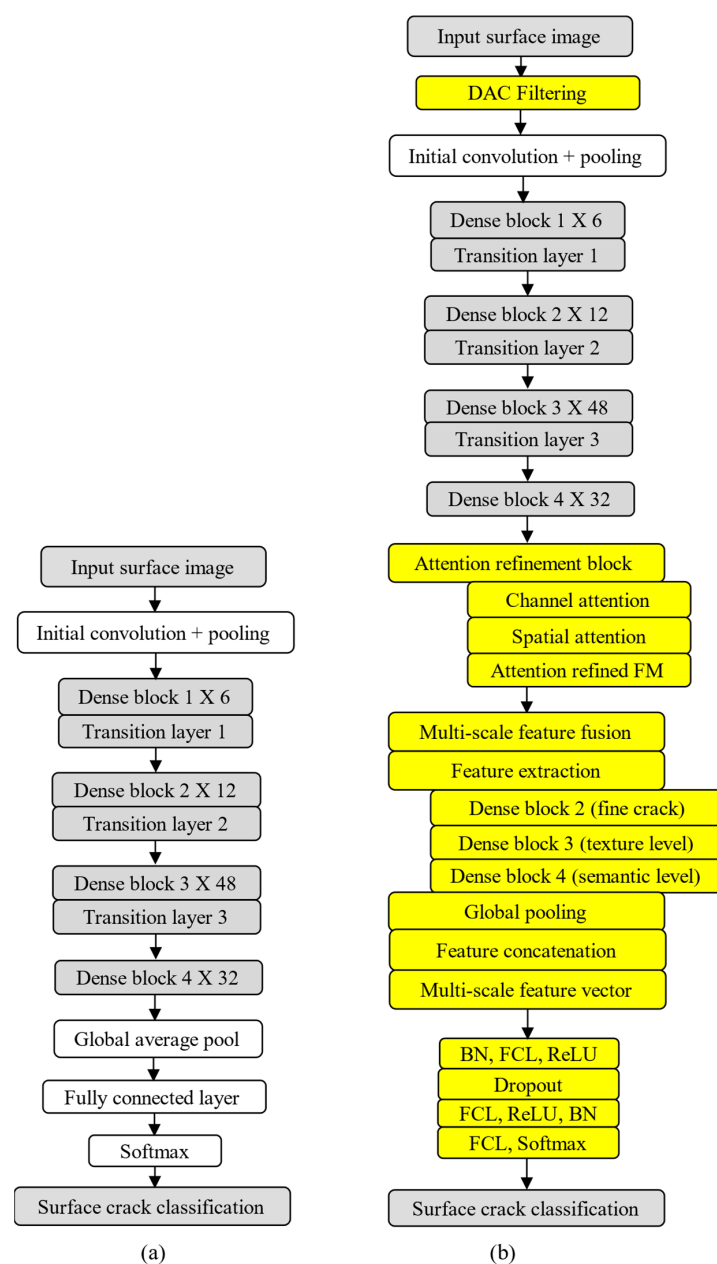


Figure 2. Frameworks: (a) Existing DenseNet-201. (b) Proposed DA-DenseNet-201.

In contrast, the proposed DA-DenseNet-201 (Figure 2b) introduces several task-specific enhancements to address these limitations. Before feature extraction, the proposed DAC filtering is applied to the input surface image to amplify crack edges and defect textures selectively while suppressing irrelevant background noise. The enhanced image is passed through the same DenseNet-201 backbone to retain the benefits of dense feature propagation. An attention refinement block is integrated to improve discriminability after dense feature extraction, comprising channel attention to emphasize damage-relevant feature channels and spatial attention to localize crack and defect regions accurately.

Moreover, the multiscale feature fusion strategy aggregates FMs from dense Blocks 2, 3 and 4, enabling learning of fine-scale crack patterns, texture-level surface degradation and high-level semantic damage representations. These fused features are globally pooled and passed through a lightweight classification head incorporating BN, dropout and multiple FCLs to improve generalization. Collectively, these enhancements allow DA-DenseNet-201 to achieve superior sensitivity, robustness and classification accuracy for surface crack detection compared with the conventional DenseNet-201 architecture.

3.3. Dataset Collection and Augmentation

The dataset used to evaluate the proposed DA-DenseNet-201 model comprises four surface damage categories (i.e., crack, delamination, spalling and no crack), with an equal number of samples maintained across all classes to ensure class balance. The images were collected through on-site inspections of reinforced concrete structures, including residential and commercial buildings, located in Daegu, South Korea. The dataset was acquired under diverse lighting conditions and surface textures to reflect realistic field environments often encountered during structural inspections. Figure 3 depicts the four surface damage classes, and Table 2 details the dataset distribution.

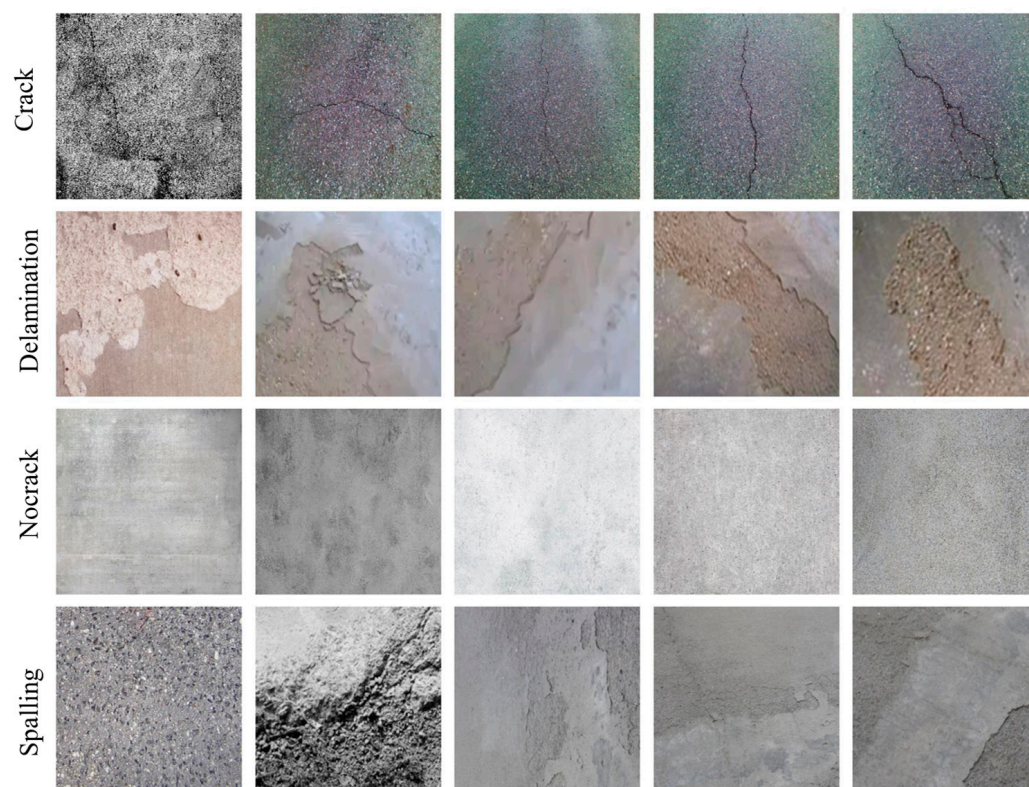


Figure 3. Surface crack classes.

Table 2. Dataset Distribution of the DA-DenseNet-201 Model.

Defect Class	Data Distribution						
	Original	Testing	Actual	Augment	Total	Training	Validation
Crack	250	50	200	1400	1600	1280	320
Delamination	250	50	200	1400	1600	1280	320
Spalling	250	50	200	1400	1600	1280	320
No crack	250	50	200	1400	1600	1280	320
Total	1000	200	800	5600	6400	5120	1280

Initially, 1000 original surface images were collected, comprising 250 images per defect class. From this original dataset, 20% of the images were strictly reserved for the testing set and were isolated at the beginning of the experiment. These testing samples were not involved in data augmentation or model training processes, preserving their authenticity and ensuring an unbiased and realistic evaluation of the generalisability of the model. In addition, class-balanced splitting was maintained across training, validation, and testing subsets to minimize sampling bias and ensure consistent class representation during model evaluation. The remaining 80% of the images were designated as the dataset for training and validation. To enhance the robustness of the model against variations in crack orientation and surface appearance, data augmentation techniques were applied to the dataset. Figure 4 presents the representative augmentation results.

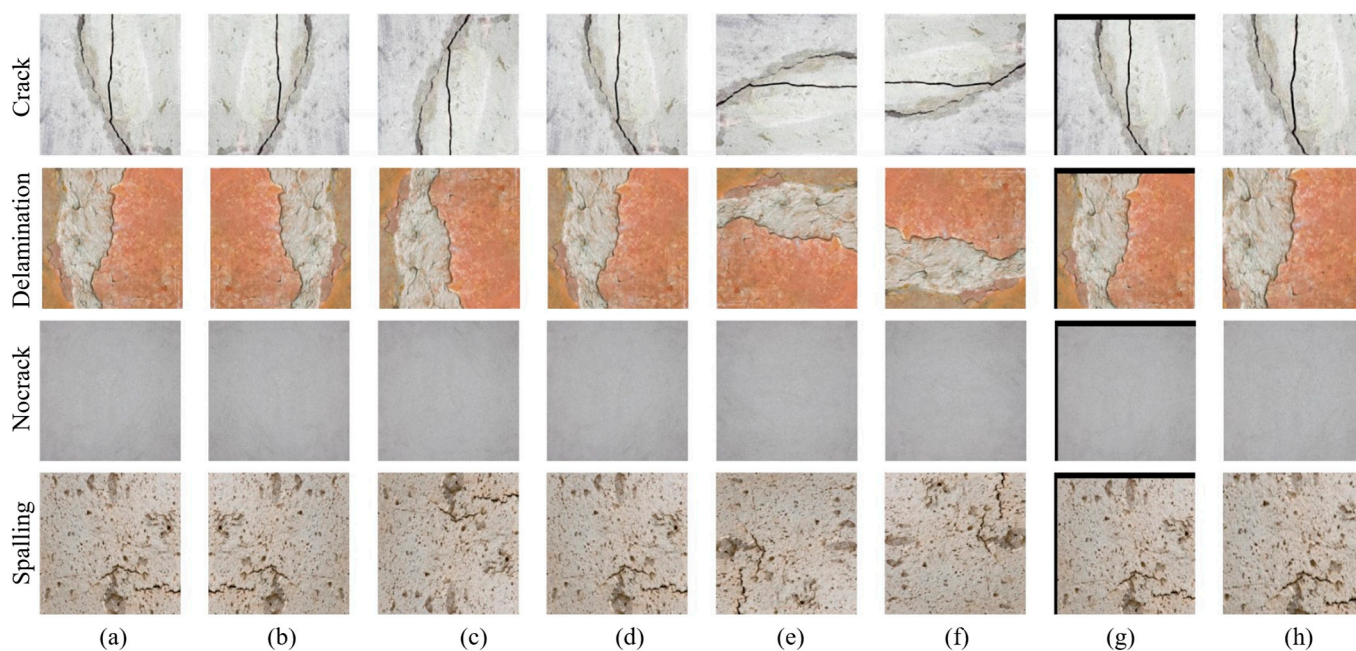


Figure 4. Augmentation of DA-DenseNet-201: (a) Original image, (b) horizontal flip, (c) vertical flip, (d) scaled, (e) rotation 90°, (f) rotation 270°, (g) translation, and (h) zoom.

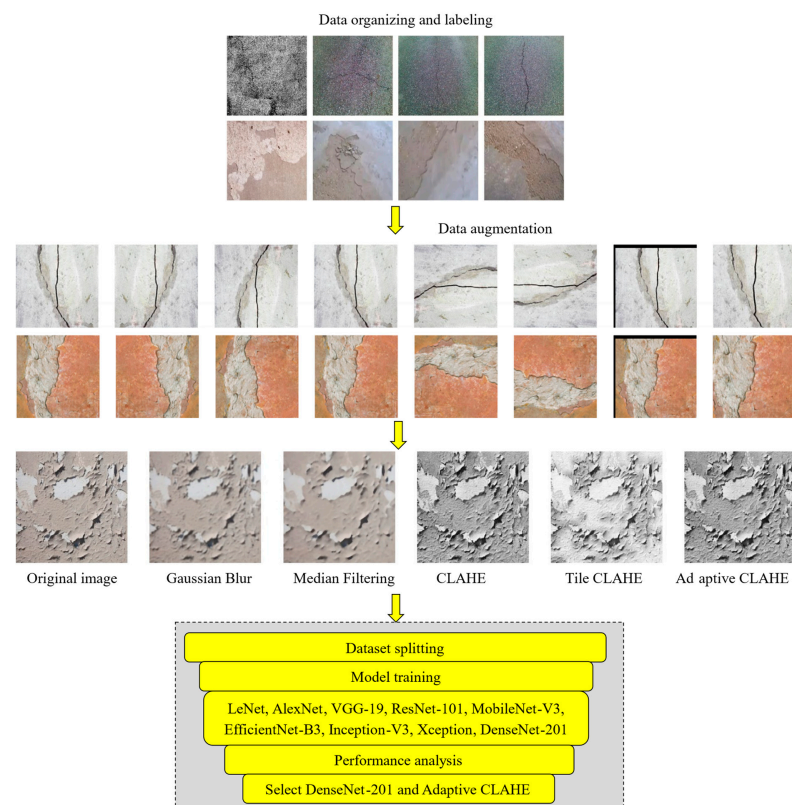
Table 3 lists the data augmentation techniques applied to each class. Through augmentation, an additional 1400 images were generated per class, resulting in 1600 images per class and an overall augmented dataset of 6400 images. Then, the augmented dataset was divided into training and validation subsets using an 80:20 split, yielding 1280 training images and 320 validation images per class, for 5120 training and 1280 validation images. By excluding the testing data from augmentation and model optimisation, the proposed method avoids data leakage and ensures a fair and reliable performance assessment.

Table 3. Augmentation Inferences of the DA-DenseNet-201 Model.

Augmentation Technique	Purpose	Expected Influence on Fracture Detection Accuracy
Horizontal flip	Mirrors the surface to simulate cracks and patterns across structural orientations.	Improves robustness against directional bias irrespective of orientation.
Vertical flip	Flips images top-to-bottom to expose the model to inverted surface conditions.	Reduces sensitivity to the acquisition viewpoint and improves generalisation.
Scaling	Resizes surface damage regions to simulate variations in camera distance.	Enhances scale invariance, allowing the accurate detection of fine cracks.
Rotation (90°, 270°)	Introduces orthogonal rotational variations common in field-acquired surface images.	Prevents misclassification due to rotated crack patterns and improves orientation-independent feature learning.
Translation	Shifts images spatially to simulate off-centre damage regions.	Improves localisation and reduces dependence on damage position.
Zoom	Simulates closer and farther inspection of surface defects by cropping and resizing.	Enhances detection of subtle cracks and edges while preserving the global context.

3.4. Selection Rationale of DenseNet-201

The selection of DenseNet-201 as the backbone architecture for the proposed DA-DenseNet-201 model followed a systematic and progressive evaluation process (Figure 5). Initially, surface defect images were organized and labelled to ensure dataset consistency and class balance.

**Figure 5.** Selection procedure for DenseNet-201.

Extensive data augmentation strategies were applied to enhance data diversity and mitigate overfitting, generating multiple variations of surface conditions, crack orientations and texture patterns. Several filtering techniques, including Gaussian blur, median filtering, CLAHE, tile-based CLAHE and adaptive CLAHE, were employed to analyze their influence on improving surface defect visibility. Following preprocessing, the enhanced datasets were subjected to dataset splitting for training, validation and testing to ensure a fair and unbiased evaluation. Multiple DL models, including LeNet, AlexNet, VGG-19, ResNet-101, MobileNet-V3, EfficientNet-B3, Inception-V3, Xception and DenseNet-201, were trained and evaluated for performance analysis. Among all evaluated combinations, DenseNet-201 consistently demonstrated superior performance, particularly when trained on adaptive CLAHE. Based on this empirical evidence, DenseNet-201 with adaptive CLAHE was selected as the optimal baseline configuration, forming the foundation for the proposed DA-DenseNet-201 model.

3.5. DAC Filtering

The proposed DAC filtering technique is designed to enhance surface damage characteristics selectively while preserving the visual integrity of undamaged regions. Initially, the input surface image I_{RGB} is acquired and resized to a fixed square of $N \times N$, ensuring uniformity for processing and compatibility with DL architectures. The resized image is converted into a grayscale representation I_g , as intensity-based transformations are more effective for highlighting surface irregularities, such as cracks, spalling and delamination. Canny edge detection is applied to the grayscale image to localise damage-prone regions, yielding an edge map $E(x, y)$ that captures crack boundaries and structural discontinuities. To quantitatively guide the adaptive enhancement process, the proposed DAC filtering framework employs the Laplacian variance σ_L^2 as a texture irregularity metric for determining the CLAHE tile grid size and clip limit. The parameter selection is guided using the quantitative texture irregularity metric obtained from the Laplacian variance rather than subjective manual tuning alone. In parallel, the local texture irregularity is quantified by computing the Laplacian variance σ_L^2 , defined in Equation (1):

$$\sigma_L^2 = \text{Var}(\nabla^2 I_g), \quad (1)$$

where I_g represents the grayscale image, ∇^2 denotes the Laplacian operator used to capture second-order intensity variations, and $\text{Var}(\cdot)$ indicates the statistical variance of the Laplacian response. Higher values of σ_L^2 correspond to abrupt intensity changes and complex surface textures typically associated with cracks, spalling, and delamination, whereas lower values indicate smoother and relatively undamaged regions. Based on the computed texture variance, the CLAHE tile grid size T is adaptively determined, where T denotes the local contextual region size used for contrast enhancement. Larger tile sizes are assigned to low-variance regions to avoid unnecessary contrast amplification, whereas moderate and smaller tile sizes are employed for medium- and high-variance regions to emphasise fine crack patterns and localised defects. Adaptive CLAHE is further applied using a dynamically adjusted clip limit C , where C controls the maximum contrast amplification to suppress excessive noise enhancement. The normalized edge map $\hat{E}(x, y)$ is then fused with the CLAHE-enhanced image I_{CLAHE} to focus enhancement on damage-relevant regions. The fusion process is expressed in Equation (2):

$$I_{DAC}(x, y) = \alpha \cdot I_{CLAHE}(x, y) \cdot \hat{E}(x, y) + (1 - \alpha) \cdot I_g(x, y). \quad (2)$$

where $I_{DAC}(x, y)$ denotes the final damage-aware contrast (DAC) enhanced image at pixel location (x, y) , $I_{CLAHE}(x, y)$ represents the image enhanced using contrast-limited adaptive histogram equalization (CLAHE), $\hat{E}(x, y)$ is the normalized edge map obtained from canny

edge detection, and $I_g(x, y)$ denoted the grayscale input image. The parameter $\alpha \in [0, 1]$ is a weighting coefficient that controls the contribution of the damage-aware enhancement. Higher values of α emphasise crack boundaries and local defects, whereas lower values preserve the appearance of undamaged surface regions. The adaptive thresholds and enhancement parameters used in the proposed DAC filtering framework were experimentally selected based on validation performance and texture sensitivity analysis. Specifically, the lower and upper texture variance thresholds were fixed as $T_{low} = 45$ and $T_{high} = 120$, respectively. For low-variance regions, a larger CLAHE contextual region of 16×16 with a clip limit of 1.5 was used to avoid excessive enhancement in smooth background areas. Medium-variance regions employed a 8×8 tile size with a clip limit of 2.5 to moderately enhance surface texture details. High-variance regions corresponding to dense crack and defect structures used a smaller 4×4 tile size with a clip limit of 4.0 to strongly emphasize localized structural irregularities.

The edge-aware fusion coefficient was fixed as $\alpha = 0.7$, allowing stronger enhancement contribution from defect-sensitive regions while preserving the overall surface appearance. In addition, the background preservation threshold was set as $\varepsilon = 0.15$, ensuring that low-edge regions retained the original grayscale intensity to suppress unnecessary contrast amplification in healthy surface areas. These parameter settings provided stable enhancement behaviour across crack, spalling, delamination, and no-crack surface categories. Finally, the resulting I_{DAC} image improves the visibility of structural damage while preventing excessive enhancement of healthy surface areas, as summarised in Algorithm 1.

Algorithm 1: DAC Filter

Input: Red, Green, Blue (RGB) surface damage image I

Output: Damage-aware adaptive CLAHE-enhanced image I_{DA}

- [1] Read input image I
 - [2] Define the parameters
 - [3] $T_{low} = 45$
 - [4] $T_{high} = 120$
 - [5] $large_tile = 16 \times 16$
 - [6] $medium_tile = 8 \times 8$
 - [7] $small_tile = 4 \times 4$
 - [8] $low_clip = 1.5$
 - [9] $medium_clip = 2.5$
 - [10] $high_clip = 4.0$
 - [11] $\alpha = 0.7$
 - [12] $\varepsilon = 0.15$
 - [13] Resize I to a fixed size ($H \times W$)
 - [14] Convert RGB image I to grayscale G
 - [15] Perform edge detection on G .
 - [16] $E = \text{Canny}(G)$
 - [17] Compute texture irregularity
 - [18] $L = \text{Laplacian}(G)$
 - [19] $V = \text{Variance}(L)$
 - [20] Determine CLAHE parameters based on V
 - [21] if $V < T_{low}$ then
 - [22] $tile_size = large_tile$
 - [23] $clip_limit = low_clip$
-

Algorithm 1: *Cont.*

```

[24]     else if  $V \geq T_{low}$  and  $V < T_{high}$  then
[25]          $tile\_size = medium\_tile$ 
[26]          $clip\_limit = medium\_clip$ 
[27]     else
[28]          $tile\_size = small\_tile$ 
[29]          $clip\_limit = high\_clip$ 
[30]     end if
[31] Apply adaptive CLAHE.
[32]  $C = CLAHE(G, tile\_size, clip\_limit)$ 
[33] Normalise edge map.
[34]  $E\_norm = Normalise(E)$ 
[35] Fuse edge-aware enhancement
[36]      $I_{DA} = (1 - \alpha) \times C + \alpha \times (C \times E\_norm)$ 
[37] Preserve background regions
[38]     For each pixel  $p$  in  $I_{DA}$  do
[39]         if  $E\_norm(p) < \epsilon$  then
[40]              $I_{DA}(p) = G(p)$ 
[41]         end if
[42]     end for
[43] Return  $I_{DA}$ 

```

3.6. Attention and Multiscale Feature Fusion in DA-DenseNet-201

The proposed DA-DenseNet-201 framework integrates damage-aware image enhancement with an attention-guided DL feature strategy to achieve robust, accurate surface defect classification. The workflow in Figure 6 acquires raw surface images, which are processed using DAC filtering. This preprocessing stage selectively enhances crack edges, spalling boundaries and delamination regions by combining adaptive contrast enhancement with edge-guided fusion, enhancing the visibility of damage-relevant structures while preserving the background texture consistency. The DAC-enhanced image is forwarded to the DenseNet-201 backbone, where an initial convolutional and pooling layer extracts low-level structural features and reduces spatial redundancy. In the DenseNet-201 architecture, feature extraction proceeds through four densely connected blocks interleaved with transition layers. Each dense block includes a sequence of BN, rectified linear unit (ReLU) activation, and convolutional operations, enabling efficient feature reuse and strong gradient propagation. The transition layers employ 1×1 convolutions followed by average pooling to compress feature dimensions and control model complexity. As the network depth increases, dense Blocks 2, 3 and 4 progressively capture fine crack patterns, texture-level damage characteristics and high-level semantic representations of severe surface defects, respectively. An attention refinement block is introduced after the dense feature extraction stage to enhance discriminability. This block incorporates channel and spatial attention mechanisms. Channel attention uses GAP followed by the FCL and sigmoid activation to learn channel-wise importance weights, identifying damage-related features. Spatial attention applies channel-wise pooling and convolutional operations to generate a spatial attention map that highlights the locations of damaged regions. The combined attention process refines the FM via elementwise multiplication, suppressing irrelevant background information and amplifying salient defect regions. The refined FM undergoes multiscale feature fusion, where the outputs from dense Blocks 2, 3 and 4 are individually processed using 1×1 convolutions, BN, ReLU activation and GAP.

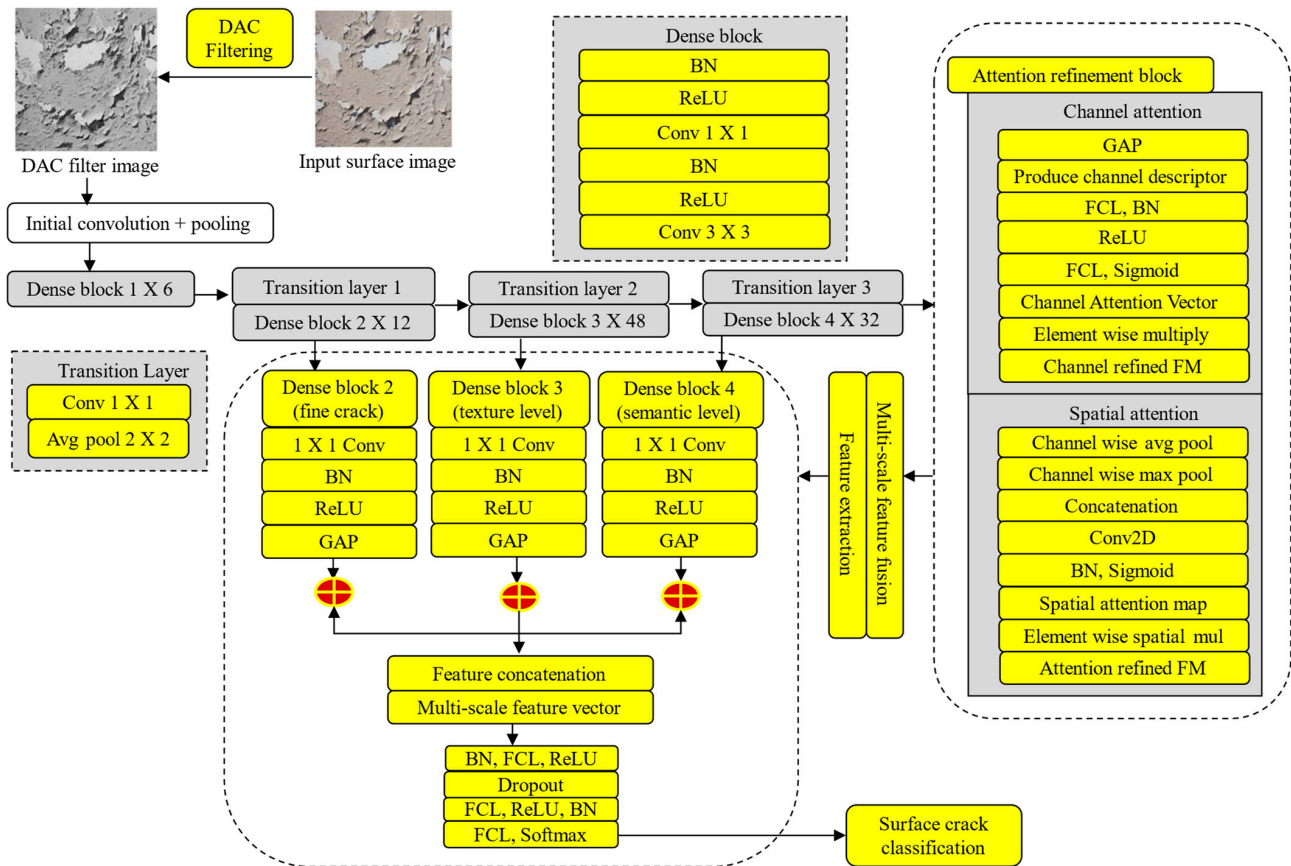


Figure 6. Architecture of the proposed DA-DenseNet-201 model.

These pooled representations are concatenated to form a comprehensive multiscale feature vector. This fusion strategy enables the model to capture subtle crack details and the broader damage context simultaneously, improving robustness across diverse surface defect types. Finally, the multiscale feature vector is passed to a lightweight classification head comprising BN and FCLs with ReLU activation and dropout for regularisation. A final FCL with softmax activation produces class probabilities corresponding to the crack, no-crack, spalling, and delamination categories using Algorithm 2 for the proposed DA-DenseNet-201.

In addition, the dense connectivity mechanism within DenseNet-201 enables each convolutional layer to receive feature information from all preceding layers through direct feature concatenation, thereby improving feature reuse and reducing information loss during deep propagation. This dense feature transmission is particularly beneficial for surface defect classification because low-level crack edge information extracted in earlier layers can continuously contribute to deeper semantic feature learning stages. Furthermore, the proposed attention refinement block operates adaptively on the extracted FM by assigning higher importance to structurally damaged regions while suppressing redundant texture responses from healthy background surfaces. The channel attention mechanism recalibrates inter-channel feature importance using global contextual information, whereas the spatial attention mechanism localizes damage-sensitive regions through spatial dependency analysis. The multiscale feature fusion stage further strengthens the robustness of the framework by integrating shallow, intermediate, and deep semantic representations obtained from Dense Blocks 2, 3, and 4. This hierarchical fusion enables simultaneous representation of narrow crack boundaries, texture discontinuities, and severe surface degradations within a unified feature space. In the final classification stage, the lightweight classification head progressively transforms the fused multiscale feature vector into highly discriminative defect representations through nonlinear dense layers and dropout-based regularization.

The SoftMax activation function finally computes normalized class probabilities for Crack, Delamination, Spalling, and No-crack categories, enabling stable multiclass surface defect classification under varying texture and illumination conditions.

The proposed DA-DenseNet-201 framework performs multiclass surface defect classification by learning discriminative structural and texture representations from surface images. Surface scratches and structural cracks are not treated as identical patterns because the proposed framework learns variations in edge continuity, texture irregularity, defect intensity, and semantic surface characteristics through hierarchical feature extraction. The proposed DAC filtering enhances damage-sensitive edge regions, while the attention refinement block emphasizes relevant defect features and suppresses irrelevant background responses. Furthermore, the multi-scale feature fusion strategy captures both fine crack-level patterns and deeper semantic damage characteristics from different DenseNet-201 stages. Finally, the lightweight classification head with SoftMax activation performs the final defect category prediction based on the refined feature representations. In addition, crack width information is indirectly learned through the hierarchical feature extraction process, where earlier dense blocks capture narrow and fine crack structures, while deeper layers represent wider and severe defect patterns with stronger semantic responses.

Although Dense Block 1 preserves very low-level edge and gradient information, its feature representations mainly contain generic structural transitions and noise-sensitive texture responses that are less discriminative for robust surface defect classification. Experimental analysis showed that directly incorporating Dense Block 1 into the multiscale fusion stage increased redundant edge responses and sensitivity to illumination variations, particularly for rough surface backgrounds. Therefore, the proposed framework selectively aggregates features from Dense Blocks 2, 3, and 4, which provide a more effective hierarchical balance between fine crack representation, texture-level degradation analysis, and high-level semantic defect understanding. Nevertheless, Dense Block 1 remains an integral part of the DenseNet-201 backbone and contributes indirectly through dense feature propagation and gradient reuse across subsequent layers, as shown in Algorithm 2.

Algorithm 2: DA-DenseNet-201

Input: Surface images

Output: Classified surface defect labels

- [1] Damage Class = 4
- [2] Class = {crack, no crack, delamination, spalling}

Step A: Data Collection and Preprocessing

- [3] Organise and label the surface damage images according to the four defect classes.
- [4] Resize to 224×224 each input image and normalise pixel intensities to $[0,1]$.
- [5] Apply DAC filtering to enhance local contrast and emphasise fine crack structures and surface irregularities.
- [6] Generate the DAC filter image as the final input to the deep network.

Step B: DenseNet-201 Backbone Feature Extraction

- [7] Apply the initial 7×7 convolution, BN, ReLU activation, and max pooling.
 - [8] Pass FMs sequentially through four dense blocks and three transition layers.
 - [9] For each dense block:
 - [10] Perform BN, ReLU, 1×1 convolution, BN, ReLU, and 3×3 convolution.
 - [11] Concatenate FMs from all preceding layers to promote feature reuse.
 - [12] For each transition layer:
 - [13] Perform 1×1 convolution and average pooling 2×2 .
-

Algorithm 2: *Cont.*

Step C: Attention Refinement Block

- [14] Extract deep semantic FMs from dense Block 4.
- [15] Apply channel attention:
- [16] Perform GAP.
- [17] Pass through two FCLs with BN and ReLU.
- [18] Apply sigmoid activation to generate channel attention weights.
- [19] Multiply channel attention weights by the FM to emphasise damage-relevant channels.
- [20] Apply spatial attention:
- [21] Perform channelwise average pooling and max pooling.
- [22] Concatenate pooled maps.
- [23] Apply 7×7 convolution, BN and sigmoid activation.
- [24] Multiply the spatial attention map with channel features to obtain the attention-refined FM.

Step D: Multiscale feature fusion

- [25] Extract FMs from:
- [26] Fine crack patterns from dense Block 2.
- [27] Texture-level damage from dense Block 3.
- [28] Semantic damage severity from dense Block 4.
- [29] For each extracted feature set:
- [30] Apply 1×1 convolution for channel alignment.
- [31] Perform BN and ReLU activation.
- [32] Apply GAP to generate scale-specific feature vectors.
- [33] Concatenate all pooled feature vectors to form a multiscale refined feature vector.

Step E: Lightweight Classification Head

- [34] Apply BN to the fused feature vector.
 - [35] Feed features into an FCL for high-level abstraction.
 - [36] Apply ReLU activation followed by dropout to prevent overfitting.
 - [37] Pass features through a second FCL for refinement.
 - [38] Apply the SoftMax activation to produce class probability scores.
 - [39] Damage classification output.
 - [40] Predict the surface damage category: crack, no crack, delamination or spalling.
 - [41] Display the final surface damage classification result.
-

4. Results and Discussion

4.1. Implementation Setup

The proposed DA-DenseNet-201 model was implemented using the TensorFlow Keras DL framework, applying transfer learning to achieve stable and efficient training. The proposed DA-DenseNet-201 model was implemented using Python 3.10 with its high-level Keras API, enabling the modular construction of the DenseNet-201 backbone, multiscale feature extraction layers and the custom refinement head. For numerical computation and data manipulation, NumPy 1.26 was extensively employed, and Pandas 2.0 facilitated dataset handling and experimental result logging. Images were preprocessed using various operations, including resizing, normalisation and batch generation using OpenCV 2.0 and TensorFlow 2.0. Scikit-learn supported model training, evaluation and performance visualization. Table 4 lists the base model, which was selected by applying all filtered surface images with DL models for assessment.

Table 4. Performance of DL Models for Model Selection and Filtering.

Model	Accuracy (%)					
	Raw Images	Gaussian Blur	Median Filtering	CLAHE	Tile-Based CLAHE	Adaptive CLAHE
LeNet	72.00	73.10	73.85	74.60	75.20	76.00
AlexNet	75.40	76.55	77.30	78.40	79.10	80.20
VGG-19	78.90	80.10	81.25	82.60	83.40	84.30
ResNet-101	82.40	83.95	85.10	86.50	87.35	88.60
MobileNet-V3	80.30	81.60	82.55	83.90	84.70	85.80
EfficientNet-B3	84.60	86.10	87.40	88.95	89.80	90.70
Inception-V3	83.20	84.60	85.75	87.10	87.95	88.90
Xception	83.90	85.40	86.60	88.20	89.10	90.10
DenseNet-201	85.30	87.10	88.55	90.10	90.85	91.25

The performance analysis in Table 4 demonstrates that progressive image enhancement techniques consistently improve classification accuracy across all DL models. Among the evaluated methods, adaptive CLAHE yields the highest performance gains, indicating its superior ability. Notably, DenseNet-201 achieves the highest accuracy of 91.25% under adaptive CLAHE, outperforming all other architectures and enhancement strategies. Table 5 validates the robustness of the adaptive CLAHE by evaluating multiple performance metrics beyond accuracy. These results indicate that DenseNet-201 improves correct classification rates and maintains a strong balance between damage detection and background discrimination. Overall, this performance justifies the adoption of DenseNet-201 with adaptive CLAHE as the core refinement model in the proposed framework.

Table 5. Performance Metrics of DL with Adaptive CLAHE Images.

Model	Adaptive CLAHE Surface Images (%)						
	Accuracy	Sensitivity	Specificity	Precision	Recall	F1-Score	IoU
LeNet	76.00	73.40	78.10	74.20	73.40	73.80	58.50
AlexNet	80.20	77.60	82.40	78.10	77.60	77.85	63.60
VGG-19	84.30	82.10	86.50	83.00	82.10	82.55	69.30
ResNet-101	88.60	86.40	90.20	87.10	86.40	86.75	76.40
MobileNet-V3	85.80	83.30	87.90	84.10	83.30	83.70	71.80
EfficientNet-B3	90.70	88.90	92.10	89.40	88.90	89.15	79.50
Inception-V3	88.90	86.80	90.40	87.30	86.80	87.05	76.80
Xception	90.10	88.20	91.70	88.80	88.20	88.50	78.90
DenseNet-201	91.25	89.60	92.80	90.10	89.60	89.85	81.20

Table 6 summarizes the hyperparameter configuration for implementing the proposed DA-DenseNet-201 model. The table presents the architectural choices, training parameters and regularization strategies adopted to ensure stable learning and high classification performance. DenseNet-201 was selected as the backbone network due to its deep feature extraction and efficient feature reuse. The top classification layers were removed to allow task-specific learning on surface damage images resized to a uniform spatial resolution. The internal configuration of the DenseNet-201 backbone, including the number of dense blocks, transition layers, growth rate and activation and normalization strategies, supports hierarchical feature learning and stable gradient propagation. Multiscale feature extraction was performed to enhance discriminative representation, using intermediate dense connectivity layers and the final DenseNet output, followed by GAP and feature concatenation. A lightweight classification head with FCLs and dropout regularization was employed to

refine the fused features and reduce overfitting. Additionally, training-related hyperparameters (e.g., the loss function, optimizer, learning rate and data augmentation strategy) are presented for reproducibility and clarity. During training, the DenseNet-201 backbone pretrained on ImageNet was initially frozen to perform transfer learning-based feature extraction, while the proposed classification head was optimized using the Adam optimizer with a batch size of 32 over 150 epochs. Early stopping and ReduceLROnPlateau scheduling were employed to stabilize convergence and prevent overfitting. A fixed random seed of 42 was used to ensure reproducibility across all experimental runs.

Table 6. Hyperparameter Configuration of DA-DenseNet-201.

Model	Hyperparameter	Configuration	Description
Base network	Backbone architecture	DenseNet-201	Pretrained CNN for feature extraction
	Pretrained weights	ImageNet	Enables transfer learning and faster convergence
	Include the top layer	False	Removes ImageNet classifier head
	Input image size	$224 \times 224 \times 3$	RGB surface damage images
DenseNet-201 backbone	Total layers	201	Deep feature extraction capacity
	Growth rate	32	Feature map growth per dense layer
	Dense blocks	4	Progressive hierarchical feature learning
	Transition layers	3	Feature compression and down-sampling
	Activation function	ReLU	Nonlinear transformation
	Normalisation	BN	Stabilises and accelerates training
Multiscale feature extraction	Feature layers	conv3_block12_concat	Fine-grained crack-level features
		conv4_block24_concat	Texture-level features
		DenseNet output	Semantic-level features
	Pooling method	Global average pooling	Converts feature maps to vectors
Feature fusion	Concatenation	Combines multiscale descriptors	
Fusion normalisation	Batch normalisation	Balances fused features	
Classification head	Dense Layer 1	512 neurons, ReLU	High-level feature refinement
	Dropout Rate 1	0.5	Prevents overfitting
	Dense Layer 2	256 neurons, ReLU	Compact discriminative learning
	Dropout Rate 2	0.4	Regularisation
	Output layer	Class = 4, softmax	Multiclass damage classification
Training configuration	Loss function	Categorical cross-entropy	Multiclass classification objective
	Optimiser	Adam	Adaptive learning rate optimisation
	Learning rate	0.001	Stable convergence
	Batch size	32	Stable mini-batch optimisation
	Training epochs	150	Sufficient convergence for DA-DenseNet-201
	Frozen layer policy	DenseNet-201 backbone frozen initially	Transfer learning-based feature extraction
	Early stopping	Patience = 15	Prevents overfitting
	Learning-rate scheduling	ReduceLROnPlateau	Adaptive learning-rate reduction
	LR reduction factor	0.5	Gradual convergence improvement
	Minimum learning rate	1.00×10^{-6}	Prevents excessive LR decay
Random seed	42	Ensures reproducibility	
Regularisation	Dropout	0.5, 0.4	Reduces model overfitting
	Data augmentation	Applied to training data only	Improves generalisation

4.2. Inferences and Performance Analysis of DAC Images

Figure 7 illustrates the comparative effects of filtering and enhancement techniques on surface damage images. The original images (Figure 7a) display low contrast and subtle damage patterns that are challenging to distinguish. Gaussian blur and median filtering (Figure 7b,c) reduce noise but smooth critical edge information, leading to partial loss of fine crack details. Standard CLAHE (Figure 7d) improves global contrast but over-enhances

certain regions, amplifying background texture and noise. Tile-based CLAHE (Figure 7e) offers better local contrast control but does not enhance tiles uniformly. In contrast, adaptive CLAHE (Figure 7f) produces the most balanced output by selectively enhancing damage regions while preserving the background homogeneity.

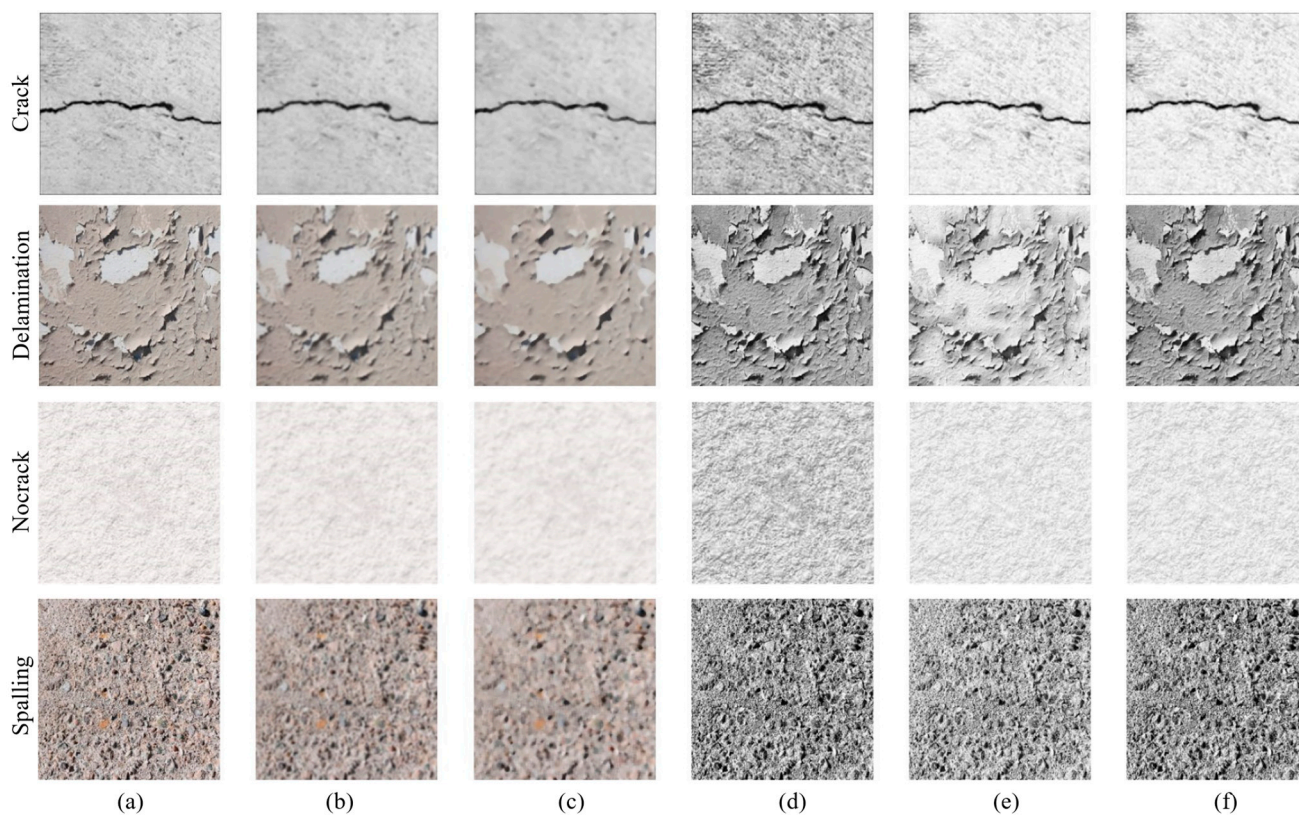


Figure 7. Filtering results: (a) Original image, (b) Gaussian blur, (c) median filtering, (d) CLAHE, (e) tile-based CLAHE, and (f) adaptive CLAHE.

Figure 8 qualitatively compares filtering images along with their corresponding histogram analyses. The original image presents a poor contrast distribution with limited intensity spread, resulting in weak visibility of damage features. Adaptive CLAHE improves local contrast and redistributes pixel intensities more effectively, uniformly enhancing background textures. In contrast, the proposed DAC produces a more discriminative enhancement, where crack edges, spalling boundaries and delamination contours are distinctly highlighted while maintaining background stability. The DAC histogram demonstrates a well-balanced intensity distribution with controlled stretching, indicating effective contrast enhancement without over-amplifying noise, confirming that DAC filtering is more damage-sensitive and visually robust than conventional adaptive CLAHE. Table 7 quantitatively evaluates filtering techniques using the structural similarity index measure (SSIM) and sharpness, entropy, contrast and edge density metrics. Traditional smoothing filters (e.g., Gaussian blur and median filtering) reduce sharpness, contrast and edge density, indicating the loss of structural damage information. The CLAHE and tile-based CLAHE methods substantially improve sharpness, entropy and contrast, reflecting enhanced texture and edge visibility. The SSIM values indicate a moderate deviation from the original structural content due to uniform enhancement. Adaptive CLAHE improves all metrics, achieving a better balance between enhancement and structural preservation. The proposed DAC filter consistently outperforms all other methods, reaching the highest sharpness (0.75), entropy (6.48), contrast (0.72) and edge density (0.57) while maintaining

a high SSIM value (0.96), demonstrating that the DAC filter enhances damage-relevant features, preserves structural integrity and minimises background distortion.

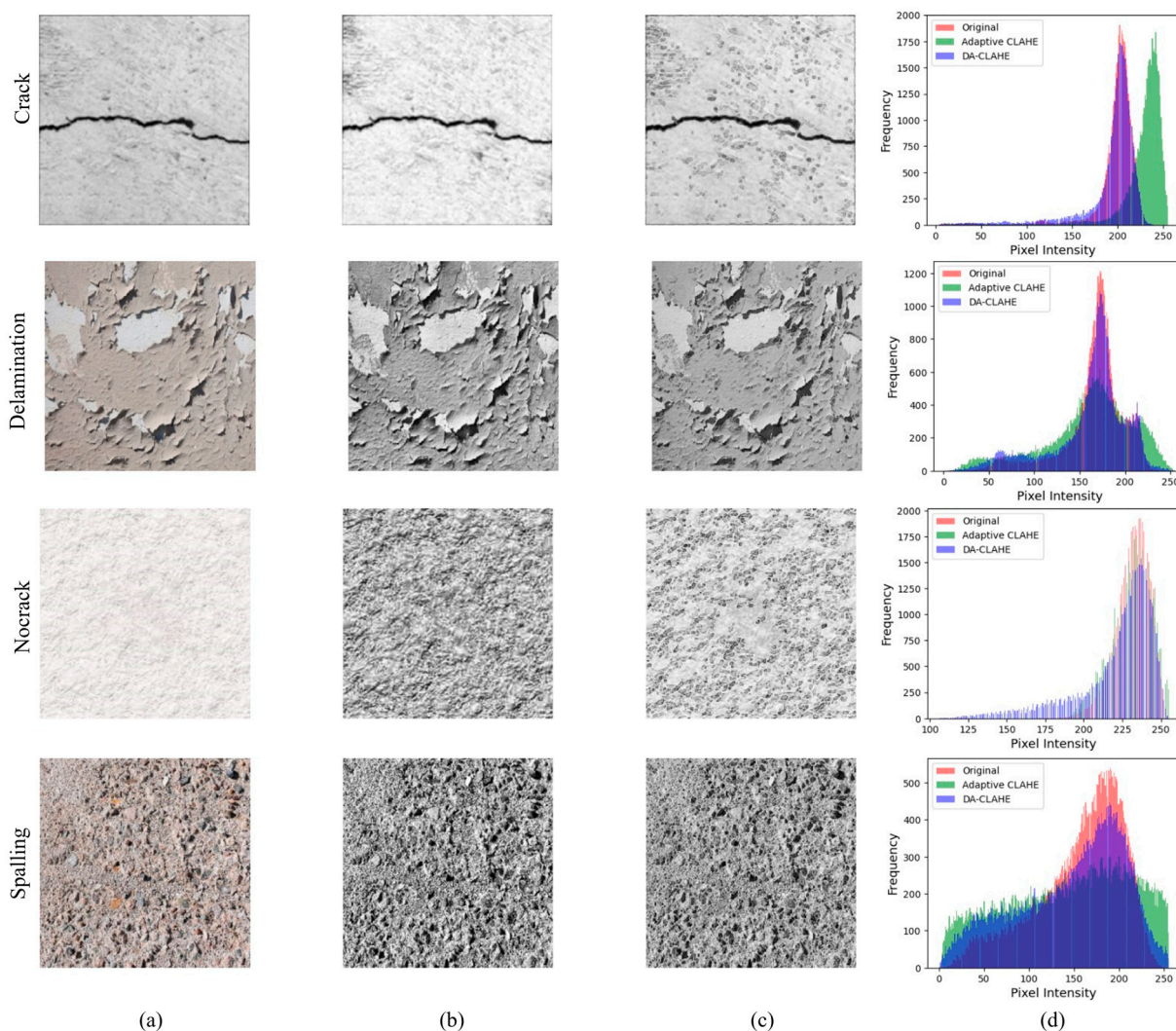


Figure 8. Performance of DAC image: (a) Original image, (b) adaptive CLAHE, (c) DA-CLAHE, and (d) histogram analysis.

Table 7. Performance metrics of DAC images.

Filter Image	Sharpness	Entropy	SSIM	Contrast	Edge Density
Original image	0.41	5.12	1.000	0.38	0.21
Gaussian blur	0.28	4.60	0.75	0.25	0.14
Median filtering	0.33	4.85	0.78	0.30	0.18
CLAHE	0.56	5.78	0.81	0.54	0.39
Tile-based CLAHE	0.62	5.95	0.82	0.59	0.44
Adaptive CLAHE	0.68	6.12	0.89	0.65	0.49
Proposed DAC filter	0.75	6.48	0.96	0.72	0.57

4.3. Performance Analysis of DA-DenseNet-201 Model

The proposed DA-DenseNet-201 model applies DAC filtering to surface damage images to enhance crack-related structures while suppressing the background noise. Figure 9 illustrates the learning behaviour of the model via the training and validation accuracy and loss curves over 150 training epochs.

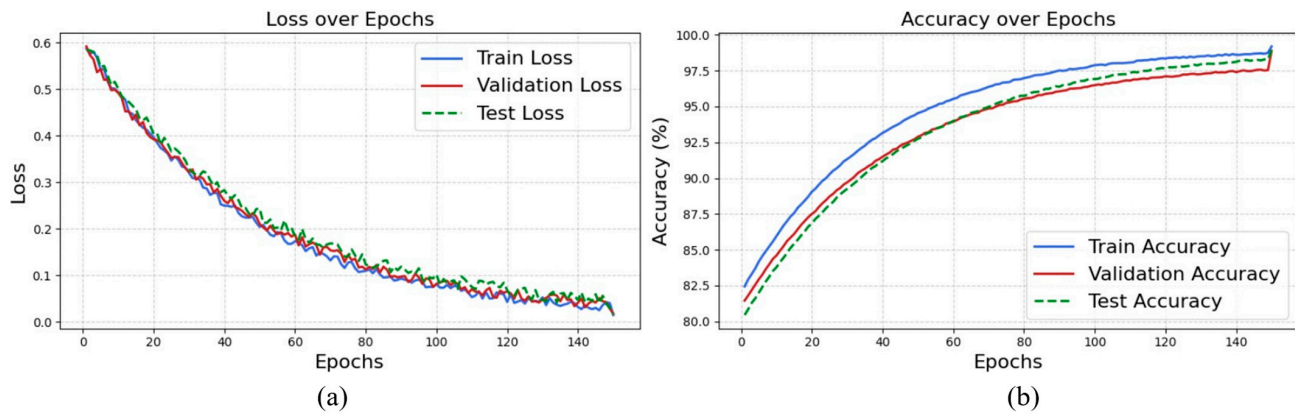


Figure 9. DA-DenseNet-201 model performance: (a) Loss and (b) accuracy curves.

Loss curves (Figure 9a) decrease consistently to very low values, and the validation loss drops sharply in the initial epochs and stabilises afterwards with minimal oscillations. This trend indicates effective optimisation, stable convergence and the absence of significant overfitting. The accuracy curves (Figure 9b) rapidly increase during the early epochs, indicating efficient feature learning facilitated by damage-aware preprocessing and multiscale dense feature extraction. As training progresses, training and validation accuracy values improve steadily and converge at 98.93%, reflecting strong generalisability. The performance curves confirm that the proposed DA-DenseNet-201 achieves robust, reliable learning, supported by DAC filtering, attention refinement and multiscale feature fusion, resulting in highly accurate surface damage classification.

To comprehensively evaluate the effectiveness of the proposed DA-DenseNet-201 framework, several widely adopted DL architectures were considered for comparative analysis, including LeNet, AlexNet, VGG-19, ResNet-101, MobileNet-V3, EfficientNet-B3, Inception-V3, Xception, and DenseNet-201. LeNet represents a shallow CNN architecture with low computational complexity, suitable for basic image feature extraction but limited in learning complex structural patterns. AlexNet introduced deeper convolutional representations and ReLU-based nonlinear learning, improving large-scale image classification performance. VGG-19 employs sequential small-sized convolution filters that enhance hierarchical feature extraction at the cost of increased computational complexity. ResNet-101 introduces residual skip connections that alleviate vanishing gradient issues and enable stable deep feature learning. MobileNet-V3 is a lightweight architecture optimized for computational efficiency and mobile deployment using depthwise separable convolutions and attention mechanisms. EfficientNet-B3 improves feature representation through compound scaling of network depth, width, and resolution. Inception-V3 utilizes multiscale convolutional kernels within parallel branches to capture diverse spatial patterns, whereas Xception extends this concept using depthwise separable convolutions for improved feature learning efficiency. DenseNet-201 employs dense feature connectivity, enabling efficient feature reuse, strong gradient propagation, and improved multiscale representation learning. Comparative analysis indicates that deeper architectures with efficient feature reuse and multiscale representation capability generally achieve superior performance for surface defect classification due to their ability to capture subtle crack patterns, texture irregularities, and semantic structural damage characteristics. Based on these advantages, DenseNet-201 was selected as the baseline backbone architecture for the proposed DA-DenseNet-201 framework, which was further enhanced using DAC filtering, attention refinement, and multiscale feature fusion mechanisms.

Table 8 presents a comprehensive comparison of DL models evaluated on DAC-filtered surface images. The proposed DA-DenseNet-201 significantly outperforms all baseline

architectures across the performance metrics. In particular, DA-DenseNet-201 achieves the highest accuracy of 98.93%, along with superior sensitivity (98.20%) and specificity (99.40%), indicating its strong ability to identify damaged and undamaged surface regions correctly. The consistently high precision (98.50%), recall (98.20%), and F1-score (98.35%) metrics confirm the robustness and balanced classification performance of the model. The proposed DA-DenseNet-201 displays substantial improvement, highlighting the effectiveness of integrating DAC filtering, attention refinement and multiscale feature fusion.

Table 8. Performance Metrics for DAC Images with DA-DenseNet-201.

Model	DAC Surface Images						
	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	Recall (%)	F1-Score (%)	IoU (%)
LeNet	79.20	76.80	81.40	77.50	76.80	77.15	62.80
AlexNet	81.50	79.10	83.60	79.80	79.10	79.45	65.90
VGG-19	85.40	83.20	87.10	84.00	83.20	83.60	70.80
ResNet-101	89.70	87.90	91.20	88.50	87.90	88.20	77.40
MobileNet-V3	87.90	85.60	89.80	86.30	85.60	85.95	73.60
EfficientNet-B3	91.80	90.10	93.20	90.70	90.10	90.40	80.90
Inception-V3	89.20	87.30	90.80	88.00	87.30	87.65	76.90
Xception	92.30	90.80	93.80	91.40	90.80	91.10	82.40
DenseNet-201	93.45	93.10	93.60	93.70	93.10	93.40	85.60
DA-DenseNet-201	98.93	98.20	99.40	98.50	98.20	98.35	94.10

4.4. Confusion Matrix of DA-DenseNet-201

The confusion matrix of the proposed DA-DenseNet-201 provides a detailed class-wise evaluation of the classification performance on DAC-enhanced surface damage images. The training confusion matrix in Figure 10 demonstrates that the proposed DA-DenseNet-201 model achieved nearly perfect learning across all four defect categories. Out of 1280 training samples per class, the model correctly classified 1271 crack, 1265 spalling, 1273 delamination, and 1272 no-crack images, highlighting its strong discriminability. Only a minimal number of samples were misclassified, such as crack images confused with other defects. These minimal off-diagonal errors indicate that the model learned class-specific and interclass distinguishing features, even for visually similar damage patterns.

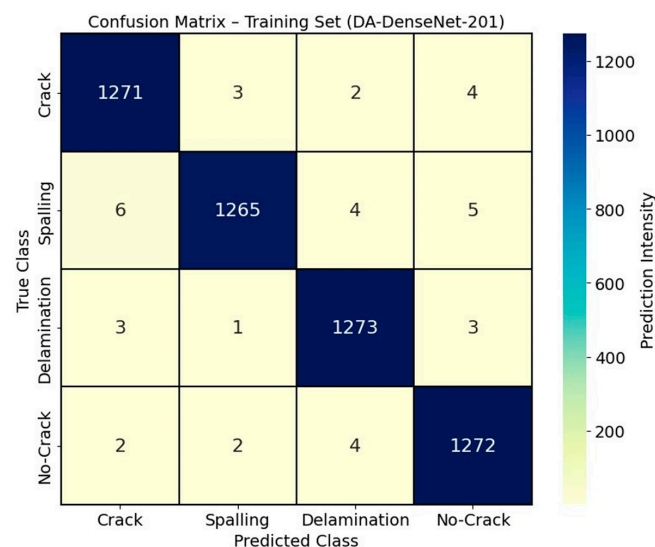


Figure 10. Confusion matrix for the training dataset.

The results confirm the robustness of the DAC filtering and multiscale feature fusion strategy, ensuring highly reliable feature learning and stable convergence of the DA-DenseNet-201 model during training. Figure 11 presents the confusion matrices for the validation and testing phases, demonstrating the strong generalisability of the proposed DA-DenseNet-201 model.

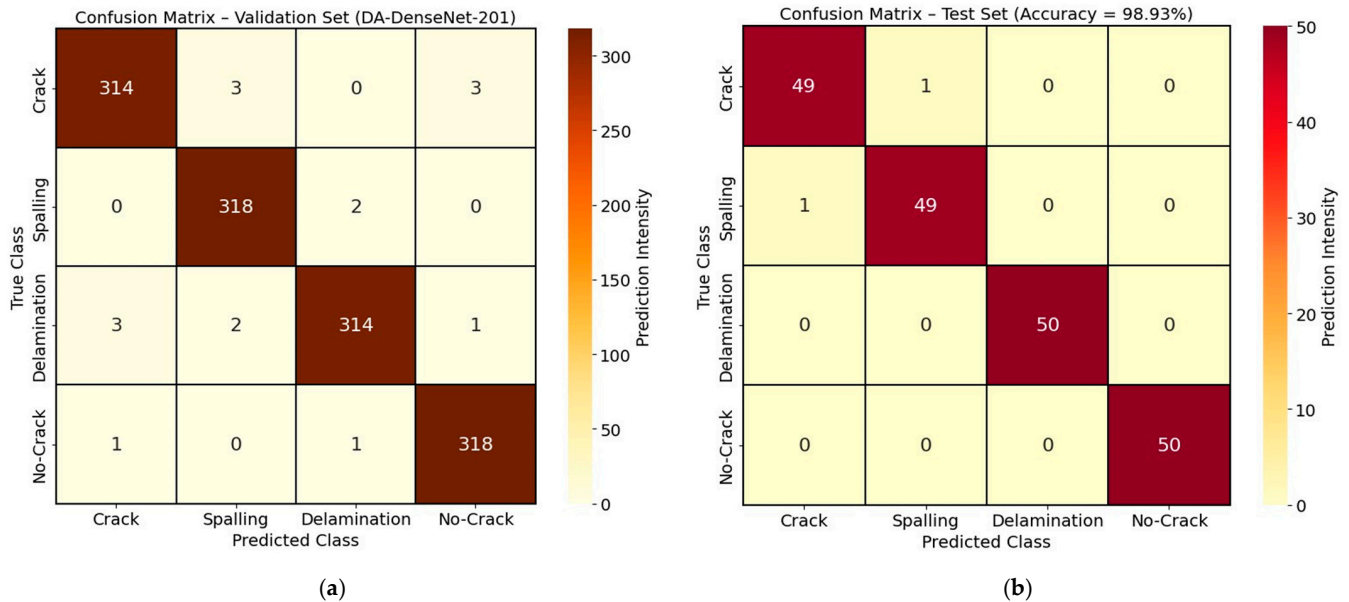


Figure 11. Confusion matrix: (a) Validation and (b) Testing dataset.

In the validation set (Figure 11a), out of 320 samples per class, the model correctly identified 314 crack, 318 spalling, 314 delamination and 318 no-crack images. Only a minimal number of misclassifications were observed, primarily between visually similar classes. The sparsity of off-diagonal entries confirms that the model maintains high class separability on unseen validation data. The testing confusion matrix (Figure 11b) further validates the robustness and reliability of the proposed framework. The model correctly predicted 49 out of 50 crack images and 49 out of 50 spalling images and had 100% accuracy for the delamination and no-crack classes, with only two misclassifications occurring between the crack and spalling categories.

These results lead to an overall testing accuracy of 98.93%, demonstrating that the DA-DenseNet-201 model transfers learned damage-aware features to completely unseen data. The ROC and precision–recall curves in Figure 12 reveal that the proposed DA-DenseNet-201 model consistently outperformed all benchmark CNN architectures in terms of the discriminability and reliability of the positive predictions. The DA-DenseNet-201 model attained the highest AUC values in both analyses, whereas traditional models (e.g., LeNet and AlexNet) exhibited noticeably lower AUC scores below 81% on both curves. These results confirm that DA-DenseNet-201 achieves superior trade-offs between the true- and false-positive rates and maintains high precision across a wide range of recall levels, making it highly suitable for accurate concrete surface defect detection.

4.5. Feature Map Inferences of DA-DenseNet-201

This work performs an FM analysis at various network stages to gain deeper insight into the internal learning behaviour of the proposed DA-DenseNet-201 model. The FM provides a visual interpretation of how the model progressively transforms raw surface images into discriminative representations by capturing hierarchical information at multiple depths. Figure 13 depicts the progressive feature learning behaviour of the proposed

DA-DenseNet-201 model by visualising the FM extracted at the transition layer. At the initial convolutional stage (Figure 13a), the model primarily captures low-level structural cues (e.g., edges, intensity gradients and basic texture variations), which are crucial for identifying early crack patterns and surface discontinuities. These shallow features offer a foundational representation of damage contours and background structures. Following the first transition layer (Figure 13b), the FMs become more refined and organised, highlighting localised textural irregularities and crack-like regions while suppressing redundant background information. This stage enhances mid-level features related to surface roughness and damage orientation. After the second transition layer (Figure 13c), the network learns more abstract and semantically meaningful representations, where damage regions appear more prominent and spatially consistent, indicating the effective aggregation of multiscale contextual information. Finally, at the third transition layer (Figure 13d), the FM exhibits highly discriminative patterns that emphasise class-specific damage characteristics while considerably reducing noise and irrelevant background details.

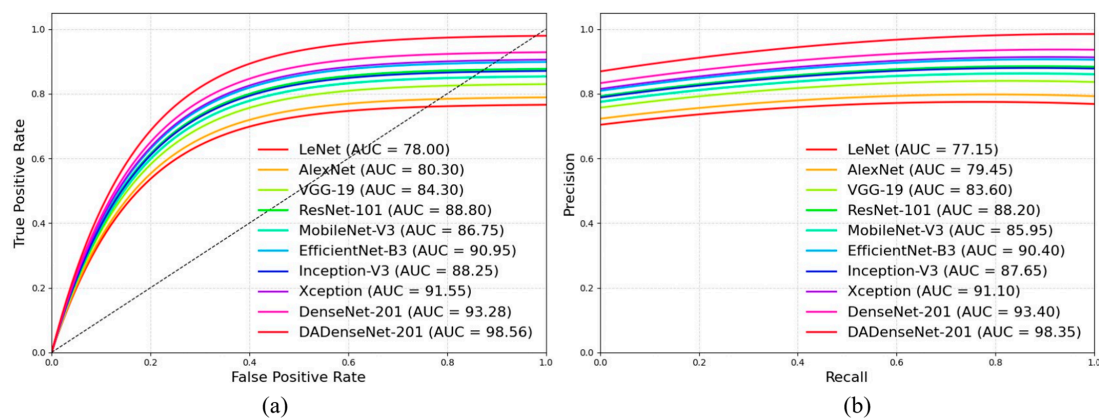


Figure 12. DA-DenseNet-201 Curve analysis: (a) ROC and (b) Precision–recall curves.

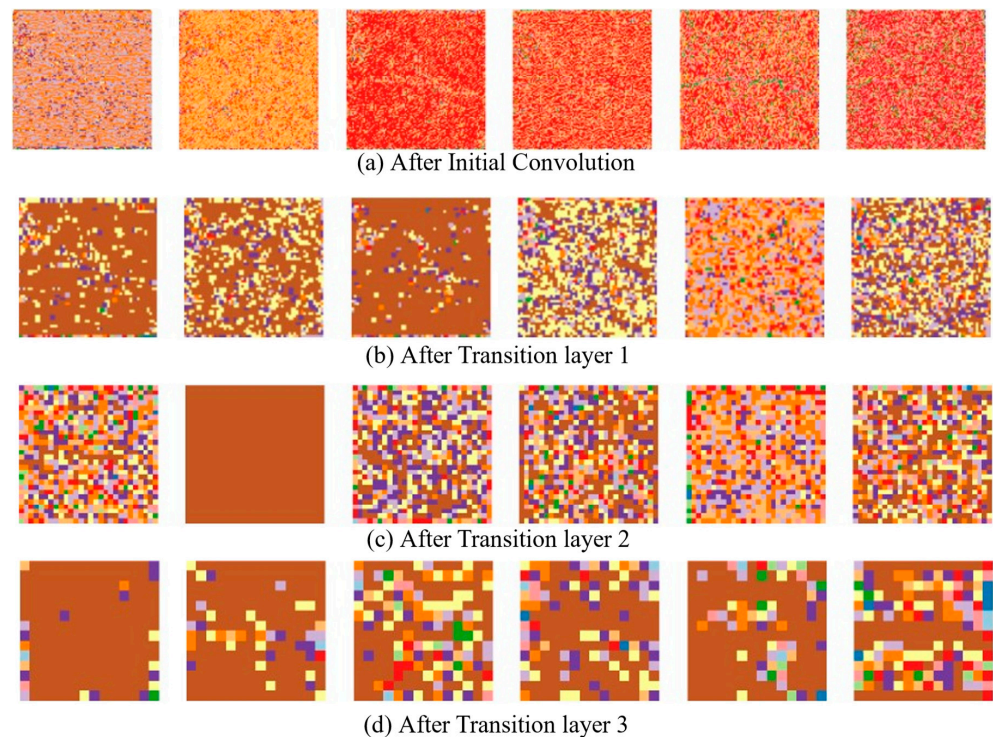


Figure 13. FM of initial stage and transition layer of DA-DenseNet-201.

Figure 14 illustrates the progressive evolution of FMs at successive dense blocks, highlighting how discriminative representations are hierarchically learned across the network depth. After dense Block 1 (Figure 14a), the FMs predominantly capture low-level structural cues (e.g., fine edges, micro-cracks and local intensity variations introduced by the DAC-enhanced input), indicating a strong sensitivity to basic damage contours. Following dense Block 2 (Figure 14b), the activations become more organised and texture-aware, emphasising crack connectivity, spalling boundaries and surface roughness patterns while suppressing irrelevant background responses. In dense Block 3 (Figure 14c), the FMs reveal more abstract and semantically meaningful representations, where damage regions are distinctly localised, and interclass variations (e.g., crack versus delamination textures) are more clearly separated. Finally, after dense Block 4 (Figure 14d), the deepest FM exhibits highly condensed and discriminative semantic responses, highlighting severe damage regions and global structural patterns while minimising noise and redundant information.

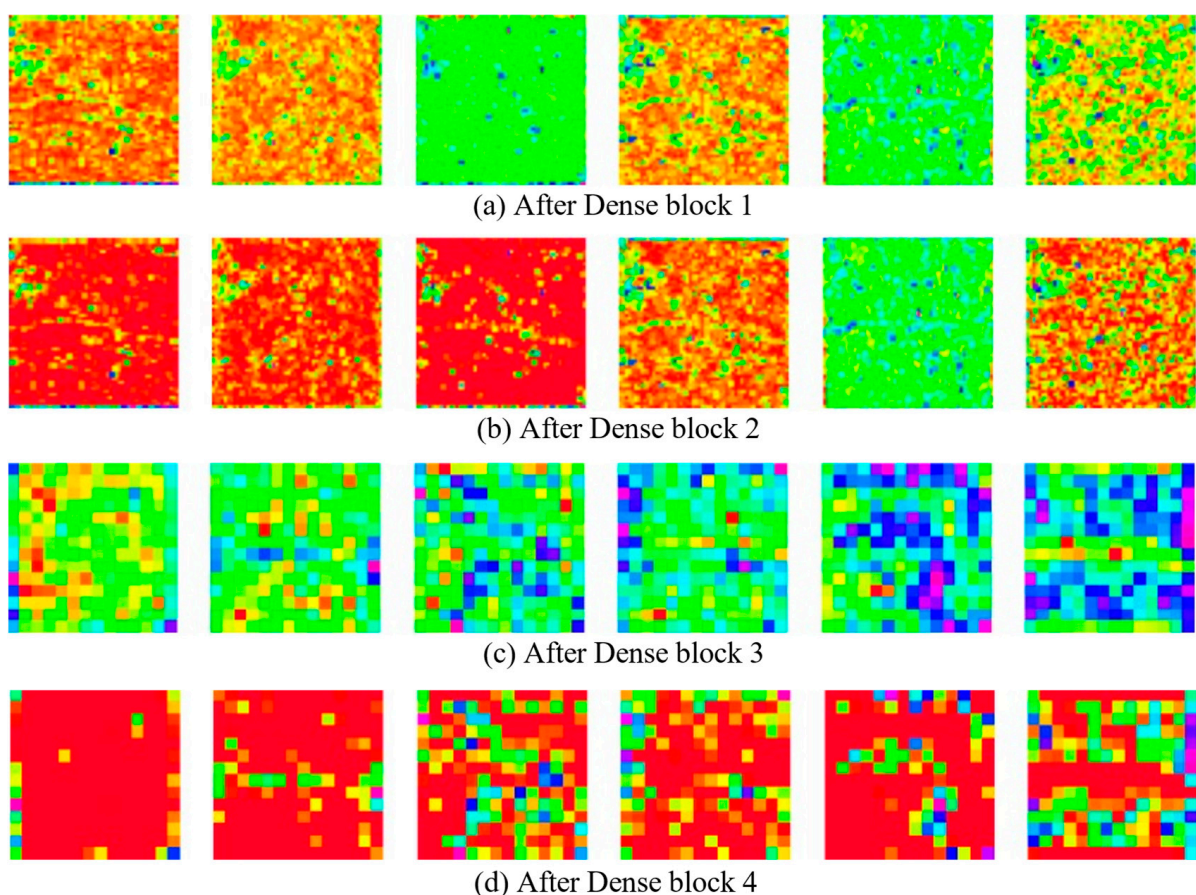


Figure 14. Feature maps of the dense block layer of DA-DenseNet-201.

Figure 15 illustrates the FM inference of the attention refinement block of DA-DenseNet-201, illustrating how attention mechanisms enhance damage-relevant representations at various stages. In Figure 15a, the channel attention vector highlights the relative importance of individual feature channels, assigning higher weights to channels that encode fracture-related patterns (e.g., edges, discontinuities and intensity variations). Following this, Figure 15b depicts the channel-refined FMs, where damage-sensitive features are amplified, and background noise is noticeably reduced.

Figure 15c presents the spatial attention map, emphasising the spatial locations corresponding to fracture regions, contours and critical structural disruptions. Finally, Figure 15d illustrates attention-refined FMs, obtained by sequentially applying channel and spatial

attention, resulting in highly localised and discriminative activations concentrated around fracture regions.

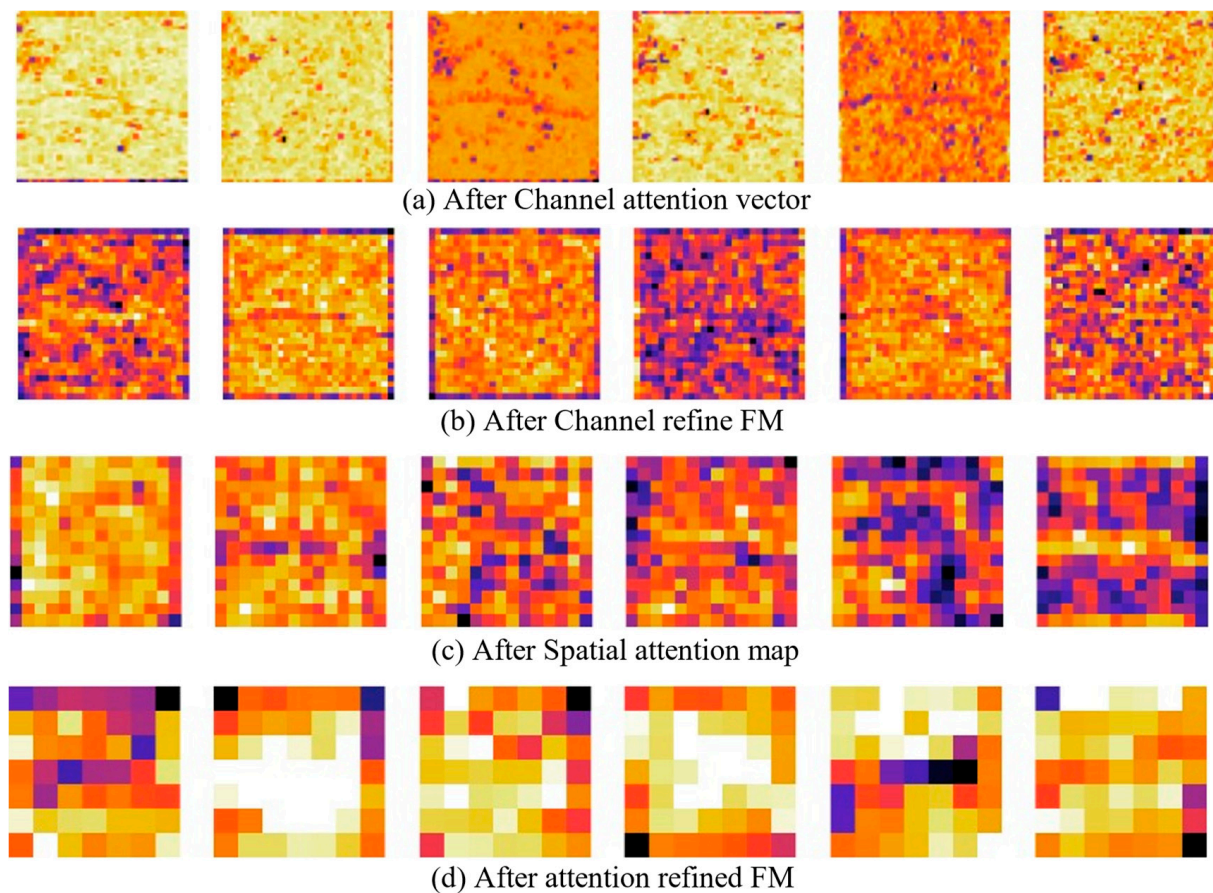


Figure 15. Feature maps of the attention refinement block of DA-DenseNet-201.

Figure 16 illustrates the FM inference of the multiscale feature fusion module in the proposed DA-DenseNet-201, revealing how information from various network depths is integrated to represent surface damage robustly. In Figure 16a, the FMs extracted from dense Block 2 capture primarily fine-scale crack patterns, including thin fissures and micro-level discontinuities, owing to their higher spatial resolution and sensitivity to low-level edge and textural cues. Figure 16b presents the FMs from dense Block 3, emphasising texture-level damage characteristics, including roughness variations, partial spalling and intermediate crack propagation patterns. These maps reflect a balance between spatial details and semantic abstractions, enabling the effective discrimination of moderate surface defects. In Figure 16c, the FMs from dense Block 4 represent high-level semantic information, capturing severe damage, global structural irregularities and class-specific contextual cues with strong discriminative power. Finally, Figure 16d reveals the multiscale fused feature vector, formed by globally pooling and concatenating features from dense Blocks 2, 3 and 4. This fusion integrates fine, intermediate and semantic representations into a unified descriptor, preserving complementary information across scales. The resulting multiscale feature representation substantially enhances the ability of the model to distinguish between crack, delamination, spalling and no-crack classes accurately, contributing to the superior performance of the proposed DA-DenseNet-201 model.

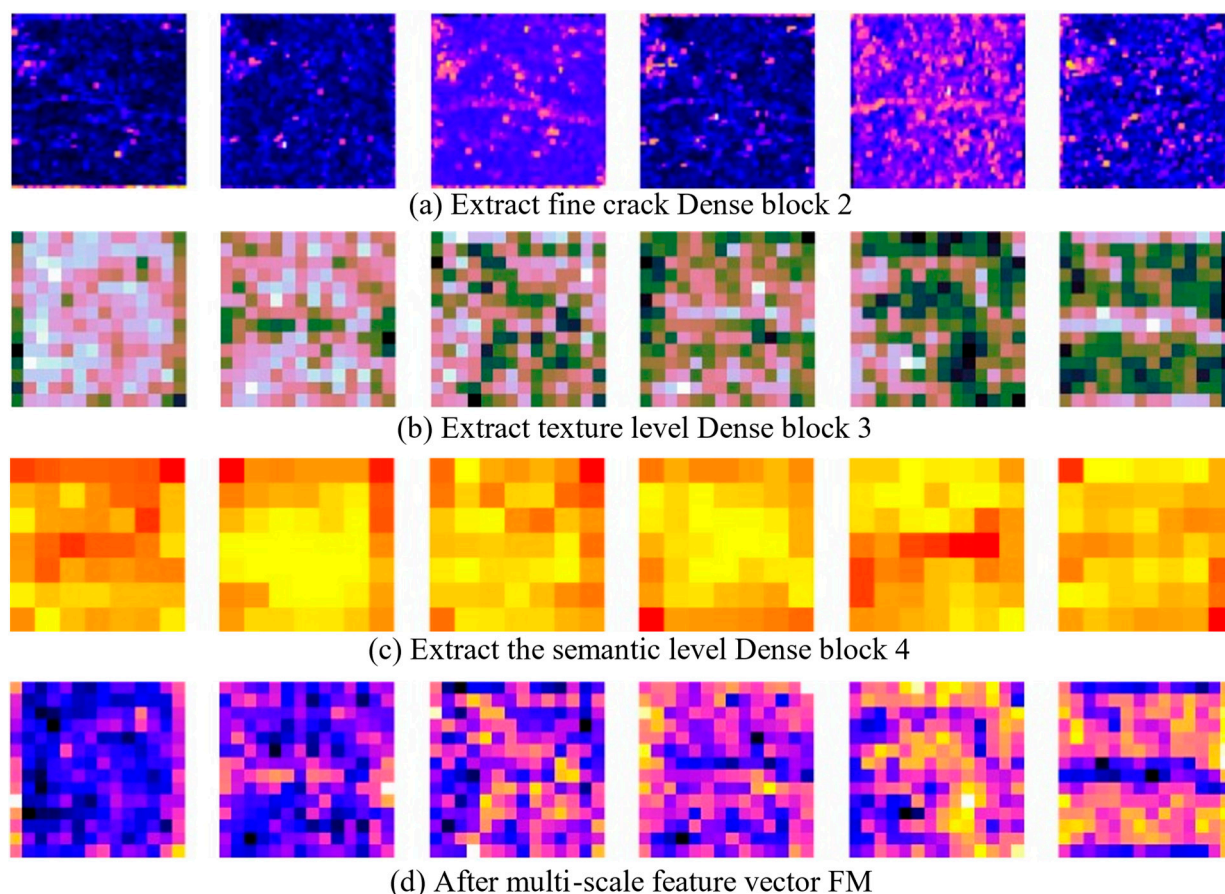


Figure 16. Feature maps of multiscale feature fusion for DA-DenseNet-201.

5. DA-DenseNet-201 Generalization Performance

To further validate the robustness and cross-domain generalization capability of the proposed DA-DenseNet-201, additional experiments were conducted using two publicly available surface defect datasets, namely the Concrete Structural Defect Imaging Dataset [74] and the NEU Surface Defect Dataset [75]. These datasets contain surface defect images acquired under different environmental conditions, texture distributions, illumination variations, and structural backgrounds, thereby providing a more challenging evaluation scenario compared to the original dataset collected from Daegu, South Korea. The purpose of this analysis is to examine whether the proposed DAC filtering, attention refinement, and multi-scale feature fusion mechanisms can maintain stable performance across heterogeneous datasets. For both datasets, the same preprocessing pipeline, DAC enhancement strategy, augmentation protocol, and training configuration employed in the original DA-DenseNet-201 framework were preserved to ensure fair comparison and consistent evaluation. Similar to the primary dataset preparation strategy, 20% of the original images were reserved exclusively for testing and were not included in augmentation or training processes. This ensures unbiased evaluation and prevents data leakage during generalization analysis. Table 9 presents the dataset distribution details for the external datasets used in this study. The Concrete Structural Defect Imaging Dataset contains 1400 original images distributed across crack, delamination, spalling, and no-crack categories, while the NEU Surface Defect Dataset contains 960 original images distributed across the same defect classes. After augmentation, the datasets were expanded to improve intra-class diversity and strengthen model generalization.

Table 9. Dataset Distribution for DA-DenseNet-201 Generalization.

Defect Class	Concrete Structural Defect Imaging Dataset Distribution						
	Original	Testing	Actual	Augment	Total	Training	Validation
Crack	350	70	280	1960	2240	1792	448
Delamination	350	70	280	1960	2240	1792	448
Spalling	350	70	280	1960	2240	1792	448
No crack	350	70	280	1960	2240	1792	448
Total	1400	280	1120	7840	8960	7168	1792
Defect class	NEU surface defect dataset distribution						
	Original	Testing	Actual	Augment	Total	Training	Validation
Crack	240	48	192	1344	1536	1229	307
Delamination	240	48	192	1344	1536	1229	307
Spalling	240	48	192	1344	1536	1229	307
No crack	240	48	192	1344	1536	1229	307
Total	960	192	768	5376	6144	4915	1229

To evaluate the effectiveness of the proposed framework under cross-dataset conditions, comparative experiments were conducted using conventional CNN architectures and advanced deep learning models. The obtained performance metrics are summarized in Table 10. The results clearly demonstrate that the proposed DA-DenseNet-201 consistently outperforms baseline models across both external datasets, indicating strong adaptability to varying surface textures and defect distributions. The cross-dataset experimental results confirm that the proposed DA-DenseNet-201 maintains highly stable and superior performance even when evaluated on external datasets collected under different acquisition conditions. In particular, the proposed framework achieves 97.85% accuracy on the Concrete Structural Defect Imaging Dataset and 96.94% accuracy on the NEU Surface Defect Dataset, outperforming conventional DenseNet-201 and other state-of-the-art architectures. The consistent performance improvement demonstrates that DAC filtering effectively enhances defect-sensitive structures, while attention refinement and multi-scale feature fusion significantly improve feature generalization across heterogeneous surface conditions. These findings validate the robustness, scalability, and practical applicability of the proposed DA-DenseNet-201 for real-world structural defect inspection scenarios.

Table 10. Performance Metrics for DA-DenseNet-201 Generalization.

Model	Concrete Structural Defect Imaging Dataset DAC Images						
	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	Recall (%)	F1-Score (%)	IoU (%)
LeNet	77.80	75.40	80.10	76.20	75.40	75.80	60.50
AlexNet	80.60	78.30	82.70	79.10	78.30	78.70	65.10
VGG-19	84.10	81.90	86.20	82.70	81.90	82.30	69.90
ResNet-101	88.20	86.10	90.00	86.90	86.10	86.50	76.20
MobileNet-V3	86.40	84.00	88.10	84.70	84.00	84.35	72.50
EfficientNet-B3	90.60	88.70	92.10	89.30	88.70	89.00	80.30
Inception-V3	88.50	86.40	90.20	87.00	86.40	86.70	76.60
Xception	91.20	89.60	92.80	90.20	89.60	89.90	81.90
DenseNet-201	93.10	91.70	94.10	92.20	91.70	91.95	84.80
DA-DenseNet-201	97.85	97.10	98.50	97.40	97.10	97.25	92.90
Model	NEU surface defect dataset DAC images						
	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	Recall (%)	F1-Score (%)	IoU (%)
LeNet	76.40	73.80	78.60	74.50	73.80	74.15	58.60
AlexNet	79.30	76.60	81.20	77.30	76.60	76.95	63.20
VGG-19	82.90	80.50	84.60	81.20	80.50	80.85	68.20
ResNet-101	87.10	85.00	88.70	85.70	85.00	85.35	74.90
MobileNet-V3	85.20	82.70	86.90	83.40	82.70	83.05	71.20
EfficientNet-B3	89.30	87.30	90.80	87.90	87.30	87.60	78.80
Inception-V3	87.40	85.30	88.90	85.80	85.30	85.55	75.20
Xception	90.10	88.40	91.60	88.90	88.40	88.65	80.40
DenseNet-201	92.20	90.60	93.40	91.10	90.60	90.85	83.30
DA-DenseNet-201	96.94	96.20	97.70	96.50	96.20	96.35	90.80

6. Ablation Study on DA-DenseNet-201

To investigate the individual contribution of each proposed component in the DA-DenseNet-201 framework, an ablation study was conducted by progressively integrating enhancement modules into the baseline DenseNet-201 architecture. The analysis evaluates the effectiveness of Adaptive CLAHE preprocessing, DAC filtering, attention refinement, and multi-scale feature fusion for surface defect classification. All experiments were performed under identical training configurations and dataset settings to ensure fair comparison. Initially, the conventional DenseNet-201 model was evaluated using raw surface images. Subsequently, Adaptive CLAHE enhancement was introduced to improve local contrast and defect visibility. The proposed DAC filtering was then incorporated to selectively strengthen crack-sensitive regions while suppressing unnecessary background enhancement. Further experiments integrated the attention refinement block and multi-scale feature fusion independently to analyse their individual impact on discriminative feature learning. Finally, the complete DA-DenseNet-201 framework combining all proposed modules was evaluated.

The obtained results presented in Table 11 demonstrate a consistent improvement in performance with each additional enhancement module. While Adaptive CLAHE improves the baseline DenseNet-201 performance through contrast enhancement, DAC filtering provides a larger improvement by preserving structural defect boundaries and texture irregularities. The attention refinement block further improves feature localization by emphasizing defect-relevant spatial and channel information. Similarly, multi-scale feature fusion enhances classification robustness by integrating fine crack patterns, texture-level degradation, and semantic-level defect representations from different DenseNet stages. The complete DA-DenseNet-201 achieves the highest classification performance with 98.93% accuracy and 94.10% IoU, confirming the complementary effectiveness of all proposed modules.

Table 11. Performance Metrics for the ablation study.

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	Recall (%)	F1-Score (%)	IoU (%)
DenseNet-201 (Raw images)	85.30	83.20	86.90	84.10	83.20	83.65	71.40
DenseNet-201 + Adaptive CLAHE	91.25	89.60	92.80	90.10	89.60	89.85	81.20
DenseNet-201 + DAC	94.10	92.80	95.20	93.20	92.80	93.00	85.90
DenseNet-201 + DAC + Attention Refinement Block	96.40	95.30	97.10	95.80	95.30	95.55	90.10
DenseNet-201 + DAC + Multi-Scale Feature Fusion	97.20	96.10	98.00	96.50	96.10	96.30	91.80
DA-DenseNet-201	98.93	98.20	99.40	98.50	98.20	98.35	94.10

7. Computational Complexity and Edge Inference Analysis

To evaluate the practical deployment capability of the proposed DA-DenseNet-201 framework, computational complexity and inference efficiency analysis were performed and compared with conventional DL architectures. The evaluation considered the total number of trainable parameters, floating-point operations (FLOPs), model size, and average inference time per image. The experiments were conducted using an NVIDIA RTX-series GPU with a batch size of 1 under identical testing conditions. Table 12 summarizes the computational performance comparison.

Table 12. Computational Complexity and Inference Performance Analysis.

Model	Parameters (M)	FLOPs (G)	Model Size (MB)	Inference Time (ms/Image)	Accuracy (%)
LeNet	0.06	0.01	2.1	3.2	79.2
AlexNet	61	0.72	233	6.8	81.5
VGG-19	143.7	19.6	548	18.4	85.4
ResNet-101	44.5	7.8	170	14.6	89.7
MobileNet-V3	5.4	0.23	21	5.1	87.9
EfficientNet-B3	12	1.8	47	8.4	91.8
Inception-V3	23.8	5.7	92	11.2	89.2
Xception	22.9	8.4	88	12.5	92.3
DenseNet-201	20.2	4.3	77	10.8	93.45
DA-DenseNet-201	23.6	5.1	89	12.1	98.93

The computational analysis indicates that the proposed DA-DenseNet-201 introduces only moderate computational overhead compared with conventional DenseNet-201 despite incorporating DAC filtering, attention refinement, and multiscale feature fusion mechanisms. Although lightweight architectures such as MobileNet-V3 exhibit lower inference time and parameter complexity, their classification performance remains comparatively lower. In contrast, the proposed framework achieves substantially higher classification accuracy while maintaining practical inference efficiency suitable for edge-assisted structural inspection systems.

8. Dataset Distortion Diversity Analysis of DA-DenseNet-201

To evaluate the robustness capability of the proposed DA-DenseNet-201 framework under practical infrastructure inspection conditions, a distortion diversity analysis was performed by introducing multiple noise, blur, and brightness variations into the original surface defect dataset. Real-world infrastructure inspection images are frequently affected by sensor noise, motion-induced blur, illumination inconsistency, and environmental interference, which significantly degrade crack visibility and defect discriminability. Therefore, the proposed robustness evaluation investigates the ability of DA-DenseNet-201 to maintain stable classification performance under adverse imaging conditions. Initially, four defect categories, namely crack, delamination, spalling, and no-crack, were subjected to multiple image degradation operations, including Gaussian noise, salt-and-pepper noise, speckle noise, Poisson noise, Gaussian blur, motion blur, median blur, defocus blur, low-brightness variation, high-brightness variation, uneven illumination, contrast reduction, and gamma adjustment.

Figure 17 presents the noise diversity dataset generated for robustness analysis of the proposed DA-DenseNet-201 framework. Gaussian noise introduces random intensity fluctuations that simulate sensor acquisition disturbances, whereas salt-and-pepper noise generates sparse impulsive corruption resembling transmission artifacts and environmental contamination. Speckle noise produces multiplicative granular distortion affecting texture continuity, while Poisson noise simulates photon-dependent acquisition irregularities under low-light imaging conditions.

Figure 18 illustrates the blur diversity dataset used to evaluate the robustness of DA-DenseNet-201 against motion and focus-related degradation conditions. Gaussian blur smooths fine structural details and reduces crack-edge sharpness, while motion blur simulates camera displacement during image acquisition. Median blur suppresses high-frequency texture information and partially removes narrow crack continuity, whereas defocus blur replicates out-of-focus inspection conditions. These blur variations significantly challenge defect visibility and edge localization capability, providing a comprehensive

robustness evaluation environment for assessing multiscale feature learning and attention-guided defect enhancement. Figure 19 presents the brightness diversity dataset generated using multiple illumination variations. Low-brightness images simulate shadowed or poorly illuminated inspection conditions, whereas high-brightness images represent over-exposed outdoor environments. Uneven illumination introduces non-uniform lighting distribution commonly observed in practical infrastructure inspections. Contrast reduction decreases defect-background separability, while gamma adjustment modifies nonlinear brightness perception across surface regions. These illumination distortions create substantial appearance variability and demonstrate the capability of the proposed DAC filtering and attention refinement mechanisms to preserve damage-sensitive structural information under challenging lighting conditions.

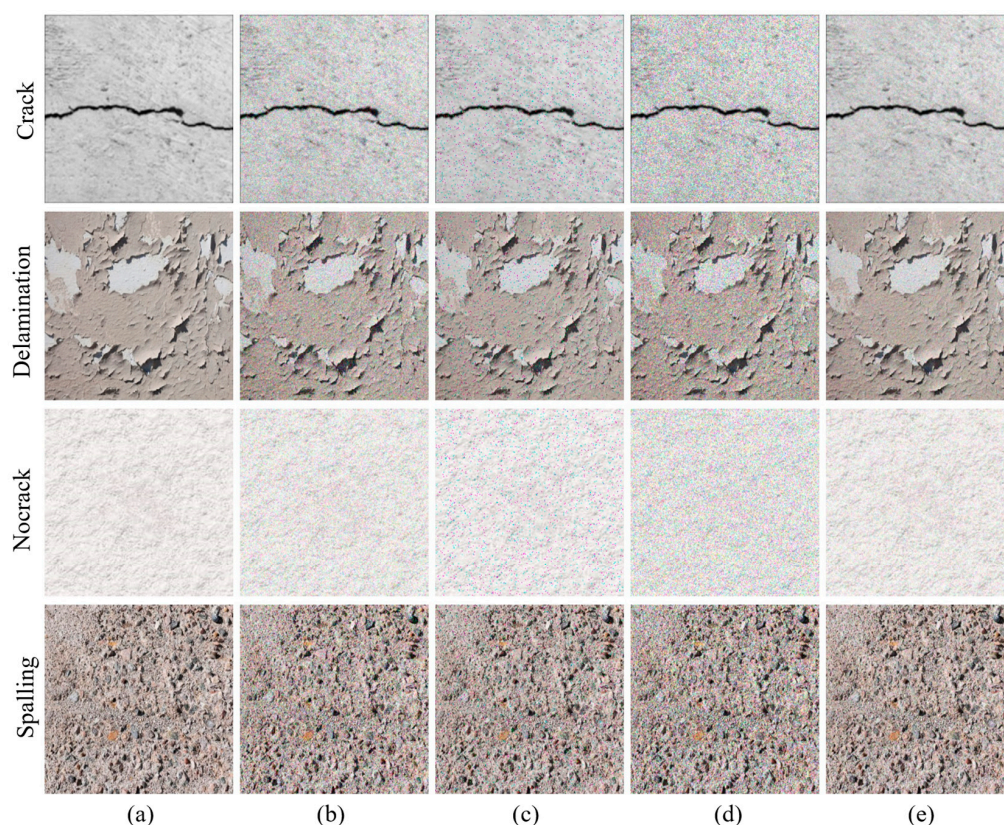


Figure 17. Noise diversity dataset for DA-DenseNet-201. (a) Original image, (b) Gaussian noise, (c) Salt and pepper, (d) Speckle noise, and (e) Poisson noise.

The overall distortion diversity dataset generation is summarized in Table 13. The distortion diversity dataset significantly increases the variability of imaging conditions by incorporating multiple environmental degradations into each defect category.

Table 13. Dataset count for distortion diversity.

Defect	Original	Noise	Blur	Brightness	Actual
Crack	250	1000	1000	1250	3500
Delamination	250	1000	1000	1250	3500
Spalling	250	1000	1000	1250	3500
No crack	250	1000	1000	1250	3500
Total	1000	4000	4000	5000	14,000

The generated distortion diversity dataset was subsequently divided into training, validation, and testing subsets for robust model evaluation, as summarized in Table 14. The dataset distribution maintains balanced defect representation across all subsets to ensure unbiased robustness evaluation under multiple distortion conditions. The inclusion of diverse degradation scenarios enables the proposed DA-DenseNet-201 framework to learn distortion-invariant damage representations while preserving defect-sensitive structural characteristics.

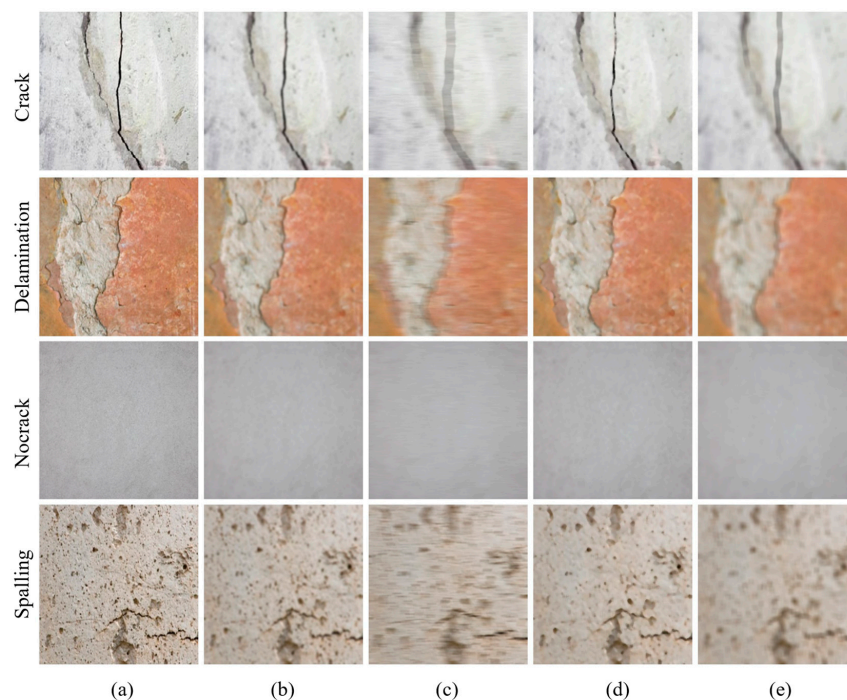


Figure 18. Blur diversity dataset for DA-DenseNet-201. (a) Original image, (b) Gaussian blur, (c) Motion blur, (d) Median blur, and (e) Defocus blur.

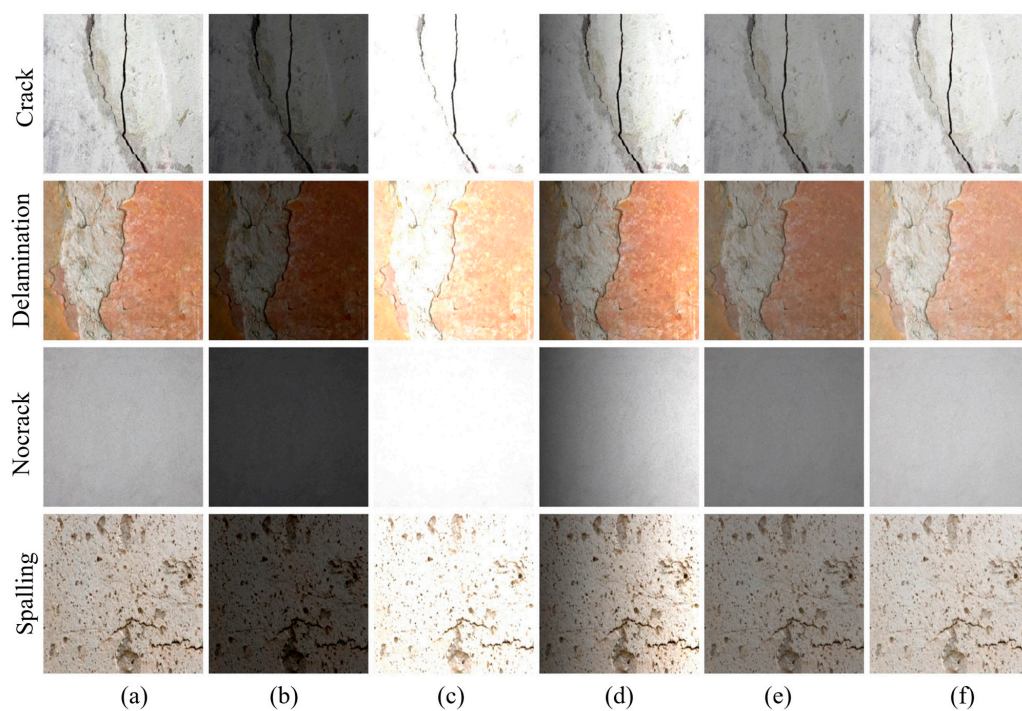


Figure 19. Brightness diversity dataset for DA-DenseNet-201. (a) Original image, (b) Low bright, (c) High bright, (d) Uneven bright, (e) Contrast, and (f) Gamma adjustment.

Table 14. Dataset Distribution for DA-DenseNet-201 distortion diversity.

Defect Class	Actual	Testing	Actual	Training	Validation
Crack	3500	700	2800	2240	560
Delamination	3500	700	2800	2240	560
Spalling	3500	700	2800	2240	560
No crack	3500	700	2800	2240	560
Total	14,000	2800	11,200	8960	2240

The performance analysis under distortion diversity conditions, as shown in Table 15, demonstrates that the proposed DA-DenseNet-201 framework consistently outperforms with 98.94% accuracy against conventional DL architectures across all evaluation metrics. Although all baseline models experience performance degradation under severe noise, blur, and illumination variations, the proposed framework maintains highly stable classification performance owing to the integration of DAC filtering, attention refinement, and multi-scale feature fusion. The DAC filtering mechanism effectively preserves defect-sensitive contrast under varying brightness and noise conditions, while the attention refinement block suppresses distortion-induced background interference. Furthermore, multiscale feature fusion enables robust extraction of fine crack structures and high-level semantic defect characteristics even under degraded imaging environments. The obtained results confirm the strong robustness and practical deployment capability of DA-DenseNet-201 for real-world infrastructure surface defect inspection applications.

Table 15. Performance Metrics for DA-DenseNet-201 under the distortion diversity dataset.

Model	Distortion Diversity Dataset DAC Images						
	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	Recall (%)	F1-Score (%)	IoU (%)
LeNet	74.60	72.10	76.80	73.00	72.10	72.55	57.40
AlexNet	77.90	75.80	79.50	76.20	75.80	76.00	61.80
VGG-19	82.50	80.60	84.30	81.10	80.60	80.85	67.90
ResNet-101	87.10	85.40	88.80	86.20	85.40	85.80	74.60
MobileNet-V3	85.60	83.20	87.10	84.00	83.20	83.60	71.50
EfficientNet-B3	90.20	88.60	91.80	89.10	88.60	88.85	79.20
Inception-V3	87.80	85.90	89.40	86.50	85.90	86.20	75.30
Xception	91.10	89.70	92.60	90.20	89.70	89.95	80.80
DenseNet-201	92.60	91.80	93.20	92.10	91.80	91.95	84.10
DA-DenseNet-201	98.94	98.10	98.30	98.40	989.10	98.25	93.20

9. Conclusions

This research was motivated by the increasing demand for accurate, reliable automated surface defect classification in civil infrastructure inspections, where various defects, including cracks, spalling, and delamination, often exhibit low-contrast, irregular textures and complex background interference. Conventional DL models, including standard DenseNet architectures, lack damage-aware preprocessing and mechanisms to emphasise defect-critical regions, limiting their robustness under real-world surface conditions. To address these challenges, this work proposes a novel DA-DenseNet-201 framework that integrates damage-sensitive enhancement, attention-guided feature refinement and multiscale representation learning into a unified architecture. The proposed DA-DenseNet-201 model employs DAC filtering, which dynamically enhances contrast based on local damage indicators, selectively amplifying defect edges and textures while preserving the background integrity. Then, the enhanced images are processed using a pretrained DenseNet-201 backbone to exploit dense feature reuse and deep semantic learning. An attention refinement block, comprising channel and spatial attention, is incorporated to emphasise damage-relevant features and localise defect regions more precisely. Furthermore, multiscale feature fusion aggregates FMs from multiple dense blocks, enabling the model to learn fine crack

patterns, texture-level degradation and high-level semantic damage characteristics. A lightweight and regularised classification head employing BN and dropout ensures stable convergence and improved generalisability. The extensive experimental evaluation demonstrates that the proposed DA-DenseNet-201 model significantly outperformed conventional CNN and DL architectures. The model achieved a maximum classification accuracy of 98.93%, along with superior sensitivity, specificity, F1-score and IoU values, confirming the effectiveness of damage-aware preprocessing, attention-based refinement and multiscale learning. The confusion matrix analysis validated the robustness and reliability of the proposed framework across the training, validation and testing datasets.

From a broader SHM perspective, the proposed DA-DenseNet-201 framework demonstrates the potential of specimen-invariant DL-based surface inspection by learning generalized defect representations across varying surface textures, illumination conditions and structural materials. The integration of DAC filtering improves defect visibility under heterogeneous imaging environments, enabling the framework to reduce sensitivity to local brightness variations and background interference commonly encountered in practical infrastructure monitoring systems. Unlike conventional enhancement techniques that uniformly amplify image contrast, the proposed DAC strategy selectively enhances damage-sensitive regions through adaptive tile selection and edge-guided fusion, thereby improving the preservation of structural defect morphology while minimizing unnecessary enhancement of non-damaged regions.

Moreover, the multiscale feature fusion mechanism enables simultaneous extraction of localized crack edges, intermediate texture degradations and high-level semantic defect patterns from multiple dense blocks. This hierarchical representation strategy is particularly beneficial for complex SHM scenarios where defect characteristics vary significantly in scale, orientation and severity. The integration of attention-guided refinement further strengthens defect discrimination by dynamically emphasizing structurally relevant feature regions. Collectively, these mechanisms demonstrate that the proposed DA-DenseNet-201 framework not only improves surface defect classification accuracy but also contributes toward more reliable and adaptive automated SHM systems suitable for practical real-world infrastructure inspection applications.

Despite its robust performance, the proposed DA-DenseNet-201 framework still exhibits several limitations. The integration of DAC filtering, attention refinement, and multiscale fusion increases computational overhead and inference time, which may affect real-time deployment in resource-constrained construction monitoring environments. The DAC filtering approach relies on heuristic threshold selection for edge density and texture variance that may require tuning for different surface defect types and imaging conditions. Although the attention mechanism improves focus on damage-relevant regions, the current framework primarily performs surface defect classification rather than precise defect localization and three-dimensional structural assessment. Furthermore, the DA-DenseNet-201 model independently processes images and does not explicitly exploit temporal or contextual information from sequential inspections. The model performance under highly complex real-world conditions, including severe illumination variation, occlusion, motion blur, and heterogeneous background interference, may require further investigation using larger and more diverse datasets. Future enhancements will focus on lightweight model compression techniques, adaptive learnable enhancement modules, multimodal sensing integration, Transformer-assisted localization frameworks, and temporal modeling approaches to improve computational efficiency, localization capability, and real-time deployment performance for large-scale structural health monitoring applications. In addition, cross-dataset validation experiments conducted using publicly available infrastructure defect datasets confirmed the generalization capability of the proposed DA-DenseNet-201

framework under varying imaging conditions, defect scales, and heterogeneous surface textures. Although the current testing setup employed balanced class distributions for unbiased evaluation, real-world infrastructure inspections often involve highly imbalanced defect occurrences and complex environmental conditions. Therefore, future research will focus on large-scale real-world infrastructure datasets, imbalance-aware learning strategies, and domain adaptation frameworks to further improve deployment-level robustness and practical generalization performance.

Author Contributions: M.M.: Conceptualization, Data curation, Formal analysis, Writing—original draft, Methodology; M.S.D.: Software, Validation, Visualization, Writing—review and editing; Y.C.: Formal analysis, Project administration, Resources; C.-Y.Y.: Project administration, Funding, Supervision, Resources. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2025-00558871). This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2024-00336025).

Data Availability Statement: The datasets used in this study are publicly available and can be accessed from the following sources: Concrete Structural Defect Imaging Dataset [74]—A publicly available dataset containing surface defect categories including crack, delamination, spalling, and no-crack images. Available at: <https://www.kaggle.com/datasets/programmer3/concrete-structural-defect-imaging-dataset> (accessed on 1 November 2025). NEU Surface Defect Database [75]—An open-source industrial surface defect dataset used for additional generalization analysis and robustness evaluation. Available at: <https://www.kaggle.com/datasets/kaustubhdikshit/neu-surface-defect-database> (accessed on 1 November 2025). The custom dataset collected through on-site inspections of reinforced concrete structures, including residential and commercial buildings, located in Daegu, South Korea for experimental analysis for this research will be made available on reasonable request for research purpose.

Conflicts of Interest: Author Young Choi was employed by the company Earth Turbine. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Fan, C.; Ding, Y.; Liu, X.; Yang, K. A review of crack research in concrete structures based on data-driven and intelligent algorithms. *Structures* **2025**, *75*, 108800. [CrossRef]
2. Zhang, J.; Zheng, Z.; Ling, T. Transformer in civil engineering defect detection: A survey. *J. Traffic Transp. Eng. (Engl. Ed.)* **2025**, *12*, 1330–1359. [CrossRef]
3. Chen, Y.; Li, H.; Zhu, H.; Ren, T.; Cao, Z. Concrete Bridge Crack Detection Using Unmanned Aerial Vehicles and Image Segmentation. *Infrastructures* **2025**, *10*, 161. [CrossRef]
4. Song, F.; Liu, B.; Yuan, G. Pixel-Level Crack Identification for Bridge Concrete Structures Using Unmanned Aerial Vehicle Photography and Deep Learning. *Struct. Control Health Monit.* **2024**, *2024*, 1299095. [CrossRef]
5. Zhang, J.; Qian, S.; Tan, C. Automated bridge surface crack detection and segmentation using computer vision-based deep learning model. *Eng. Appl. Artif. Intell.* **2022**, *115*, 105225. [CrossRef]
6. Khan, S.; Jan, A.; Seo, S. Structural crack detection using deep learning: An in-depth review. *Korean J. Remote Sens.* **2023**, *39*, 371–393. [CrossRef]
7. Xu, Z.; Qian, S.; Ran, X.; Zhou, J. Application of deep convolution neural network in crack identification. *Appl. Artif. Intell.* **2022**, *36*, 2014188. [CrossRef]
8. Mayya, A.M.; Alkayem, N.F. Enhance the Concrete Crack Classification Based on a Novel Multi-Stage YOLOV10-ViT Framework. *Sensors* **2024**, *24*, 8095. [CrossRef]
9. Ye, G.; Dai, W.; Tao, J.; Qu, J.; Zhu, L.; Jin, Q. An improved transformer-based concrete crack classification method. *Sci. Rep.* **2024**, *14*, 6226. [CrossRef]
10. Hota, J.; Schabowicz, K. State-of-the-art non-destructive methods for diagnostic testing of building structures—Anticipated development trends. *Arch. Civ. Mech. Eng.* **2010**, *10*, 5–18. [CrossRef]

11. Dwivedi, S.K.; Vishwakarma, M.; Soni, A. Advances and researches on non destructive testing: A review. *Mater. Today Proc.* **2018**, *5*, 3690–3698. [[CrossRef](#)]
12. Zhao, Z. Review of non-destructive testing methods for defect detection of ceramics. *Ceram. Int.* **2021**, *47*, 4389–4397. [[CrossRef](#)]
13. Razi, P.; Esmaeel, R.A.; Taheri, F. Application of a robust vibration-based non-destructive method for detection of fatigue cracks in structures. *Smart Mater. Struct.* **2011**, *20*, 115017. [[CrossRef](#)]
14. Fu, R.; Xu, H.; Wang, Z.; Shen, L.; Cao, M.; Liu, T.; Novak, D. Enhanced intelligent identification of concrete cracks using multi-layered image preprocessing-aided convolutional neural networks. *Sensors* **2020**, *20*, 2021. [[CrossRef](#)]
15. Lv, Z.; Tian, J.; Zhu, Y.; Li, Y. Automatic crack detection of dam concrete structures based on deep learning. *Comput. Concr.* **2023**, *32*, 615. [[CrossRef](#)]
16. Liu, D.; Xu, M.; Li, Z.; He, Y.; Zheng, L.; Xue, P.; Wu, X. A multi-scale residual encoding network for concrete crack segmentation. *J. Intell. Fuzzy Syst.* **2024**, *46*, 1379–1392. [[CrossRef](#)]
17. Liu, C.; Tian, L.; Wang, P.; Wang, X.; Miao, J. Automatic damage detection and evaluation of post-fire reinforced concrete structures using deep learning. *Autom. Constr.* **2025**, *177*, 106376. [[CrossRef](#)]
18. Zhang, X.; Wen, K.; Liu, L.; Cao, J.; Zheng, Z.; Chen, Z.; Xiang, P. Distributed Fiber Optic Sensing Integrating TCN-Transformer for Damage Evolution Analysis of Asphalt Concrete under Freeze–Thaw Cycles. *J. Transp. Eng. Part B Pavements* **2026**, *152*, 04026018. [[CrossRef](#)]
19. Wen, X.; Li, S.; Yu, H.; He, Y. Multi-scale context feature and cross-attention network-enabled system and software-based for pavement crack detection. *Eng. Appl. Artif. Intell.* **2024**, *127*, 107328. [[CrossRef](#)]
20. Yang, L.; Bai, S.; Liu, Y.; Yu, H. Multi-scale triple-attention network for pixelwise crack segmentation. *Autom. Constr.* **2023**, *150*, 104853. [[CrossRef](#)]
21. Liu, G.; Wu, X.; Dai, F.; Liu, G.; Li, L.; Huang, B. Crack-MsCGA: A Deep Learning Network with Multi-Scale Attention for Pavement Crack Detection. *Sensors* **2025**, *25*, 2446. [[CrossRef](#)]
22. Azouz, Z.; Honarvar Shakibaei Asli, B.; Khan, M. Evolution of crack analysis in structures using image processing technique: A review. *Electronics* **2023**, *12*, 3862. [[CrossRef](#)]
23. Pham, M.V.; Ha, Y.S.; Kim, Y.T. Automatic detection and measurement of ground crack propagation using deep learning networks and an image processing technique. *Measurement* **2023**, *215*, 112832. [[CrossRef](#)]
24. Kirthiga, R.; Elavenil, S. A survey on crack detection in concrete surface using image processing and machine learning. *J. Build. Pathol. Rehabil.* **2024**, *9*, 15. [[CrossRef](#)]
25. Dwivedi, D.; Babu, K.V.S.M.; Yemula, P.K.; Chakraborty, P.; Pal, M. Identification of surface defects on solar pv panels and wind turbine blades using attention based deep learning model. *Eng. Appl. Artif. Intell.* **2024**, *131*, 107836. [[CrossRef](#)]
26. Wu, Z.; Tang, Y.; Hong, B.; Liang, B.; Liu, Y. Enhanced precision in dam crack width measurement: Leveraging advanced lightweight network identification for pixel-level accuracy. *Int. J. Intell. Syst.* **2023**, *2023*, 9940881. [[CrossRef](#)]
27. Tumrate, C.S.; Saini, D.K.; Gupta, P.; Mishra, D. Evolutionary computation modelling for structural health monitoring of critical infrastructure. *Arch. Comput. Methods Eng.* **2023**, *30*, 1479–1493. [[CrossRef](#)]
28. Dai, Q.; Ishfaq, M.; Khan, S.U.R.; Luo, Y.L.; Lei, Y.; Zhang, B.; Zhou, W. Image classification for sub-surface crack identification in concrete dam based on borehole CCTV images using deep dense hybrid model. *Stoch. Environ. Res. Risk Assess.* **2024**, *39*, 4637–4654. [[CrossRef](#)]
29. Nguyen, L.N.; Le, T.H.; Nguyen, L.Q.; Tran, V.Q. Machine learning approaches for predicting Cracking Tolerance Index (CTIndex) of asphalt concrete containing reclaimed asphalt pavement. *PLoS ONE* **2023**, *18*, e0287255. [[CrossRef](#)]
30. Srivastava, V.; Basu, B.; Prabhu, N. Application of Machine Learning (ML)-based multi-classifications to identify corrosion fatigue cracking phenomena in Naval steel weldments. *Mater. Today Commun.* **2024**, *39*, 108591. [[CrossRef](#)]
31. Xu, L.; Yang, J.; Ge, M.; Su, Z. Three-dimensional fatigue crack quantification using densely connected convolutional network-assisted ultrasonic guided waves. *Int. J. Fatigue* **2024**, *180*, 108094. [[CrossRef](#)]
32. Silva, W.R.L.D.; Lucena, D.S.D. Concrete cracks detection based on deep learning image classification. *Proceedings* **2018**, *2*, 489. [[CrossRef](#)]
33. Flah, M.; Suleiman, A.R.; Nehdi, M.L. Classification and quantification of cracks in concrete structures using deep learning image-based techniques. *Cem. Concr. Compos.* **2020**, *114*, 103781. [[CrossRef](#)]
34. Hu, G.X.; Hu, B.L.; Yang, Z.; Huang, L.; Li, P. Pavement crack detection method based on deep learning models. *Wirel. Commun. Mob. Comput.* **2021**, *2021*, 5573590. [[CrossRef](#)]
35. Arbaoui, A.; Ouahabi, A.; Jacques, S.; Hamiane, M. Concrete cracks detection and monitoring using deep learning-based multiresolution analysis. *Electronics* **2021**, *10*, 1772. [[CrossRef](#)]
36. Dang, L.M.; Wang, H.; Li, Y.; Nguyen, L.Q.; Nguyen, T.N.; Song, H.K.; Moon, H. Deep learning-based masonry crack segmentation and real-life crack length measurement. *Constr. Build. Mater.* **2022**, *359*, 129438. [[CrossRef](#)]
37. Rosenberger, J.; Tlatlik, J.; Münstermann, S. Deep learning based initial crack size measurements utilizing macroscale fracture surface segmentation. *Eng. Fract. Mech.* **2023**, *293*, 109686. [[CrossRef](#)]

38. Li, D.; Chen, Q.; Wang, H.; Shen, P.; Li, Z.; He, W. Deep learning-based acoustic emission data clustering for crack evaluation of welded joints in field bridges. *Autom. Constr.* **2024**, *165*, 105540. [[CrossRef](#)]
39. Morita, D.; Kawarazaki, A.; Soufi, M.; Otake, Y.; Sato, Y.; Numajiri, T. Automatic detection of midfacial fractures in facial bone CT images using deep learning-based object detection models. *J. Stomatol. Oral Maxillofac. Surg.* **2024**, *125*, 101914. [[CrossRef](#)]
40. Hamishebahr, Y.; Guan, H.; So, S.; Jo, J. A comprehensive review of deep learning-based crack detection approaches. *Appl. Sci.* **2022**, *12*, 1374. [[CrossRef](#)]
41. Satpathy, R.P.K.; Kumar, K.; Hirwani, C.K.; Kumar, V.; Kumar, E.K.; Panda, S.K. Computational deep learning algorithm (vision/frequency response)-based damage detection in engineering structure. *Acta Mech.* **2023**, *234*, 5919–5935. [[CrossRef](#)]
42. Le, T.T.; Nguyen, V.H.; Le, M.V. Development of deep learning model for the recognition of cracks on concrete surfaces. *Appl. Comput. Intell. Soft Comput.* **2021**, *2021*, 8858545. [[CrossRef](#)]
43. Zhang, H.; Shen, Z.; Lin, Z.; Quan, L.; Sun, L. Deep learning-based automatic classification of three-level surface information in bridge inspection. *Comput.-Aided Civ. Infrastruct. Eng.* **2024**, *39*, 1431–1451. [[CrossRef](#)]
44. Feng, X.; Xiao, L.; Li, W.; Pei, L.; Sun, Z.; Ma, Z.; Shen, H.; Ju, H. Pavement crack detection and segmentation method based on improved deep learning fusion model. *Math. Probl. Eng.* **2020**, *2020*, 8515213. [[CrossRef](#)]
45. Pandey, V.; Mishra, S.S. A review of image-based deep learning methods for crack detection. *Multimed. Tools Appl.* **2025**, *84*, 35469–35511. [[CrossRef](#)]
46. Li, Q.; Zhang, G.; Yang, P. CL-YOLOv8: Crack detection algorithm for fair-faced walls based on deep learning. *Appl. Sci.* **2024**, *14*, 9421. [[CrossRef](#)]
47. Geetha, G.K.; Sim, S.H. Fast identification of concrete cracks using 1D deep learning and explainable artificial intelligence-based analysis. *Autom. Constr.* **2022**, *143*, 104572. [[CrossRef](#)]
48. Zheng, J.; Qian, K.; Liu, X.; Pang, Z.; Yang, Z.; Sun, J.; Zhang, D. An improved automatic image labeling and classification algorithm for multi-mode damage quantification of 2.5 D woven composites based on deep learning strategy. *Compos. Sci. Technol.* **2025**, *259*, 110932. [[CrossRef](#)]
49. Daimari, E.; Ratna, S.; Mouli, P.C.; Madhurima, V. A Comprehensive study on the different types of soil desiccation cracks and their implications for soil identification using deep learning techniques. *Eur. Phys. J. E* **2024**, *47*, 57. [[CrossRef](#)]
50. Golding, V.P.; Gharineiat, Z.; Munawar, H.S.; Ullah, F. Crack detection in concrete structures using deep learning. *Sustainability* **2022**, *14*, 8117. [[CrossRef](#)]
51. Wang, P.; Liu, C.; Wang, X.; Tian, L.; Miao, J.; Liu, Y. Multicategory fire damage detection of post-fire reinforced concrete structural components. *Comput.-Aided Civ. Infrastruct. Eng.* **2025**, *40*, 91–112. [[CrossRef](#)]
52. Ibrahim, E.A.; Goff, D.; Keyvanfar, A.; Jonaidi, M. Assessing post-fire damage in concrete structures: A comprehensive review. *Buildings* **2025**, *15*, 485. [[CrossRef](#)]
53. Jovanović, B.; Caspeepele, R.; Lombaert, G.; Reynders, E.; Van Coile, R. State-of-the-art review on the post-fire assessment of concrete structures. *Struct. Concr.* **2023**, *24*, 5370–5387. [[CrossRef](#)]
54. Andrushia, A.D.; Anand, N.; Neebha, T.M.; Naser, M.Z.; Lubloy, E. Autonomous detection of concrete damage under fire conditions. *Autom. Constr.* **2022**, *140*, 104364. [[CrossRef](#)]
55. Choudhary, G.K.; Dey, S. Crack detection in concrete surfaces using image processing, fuzzy logic, and neural networks. In *2012 IEEE Fifth International Conference on Advanced Computational Intelligence (ICACI)*; IEEE: Piscataway, NJ, USA, 2012; pp. 404–411. [[CrossRef](#)]
56. Lins, R.G.; Givigi, S.N. Automatic crack detection and measurement based on image analysis. *IEEE Trans. Instrum. Meas.* **2016**, *65*, 583–590. [[CrossRef](#)]
57. Akagic, A.; Buza, E.; Omanovic, S.; Karabegovic, A. Pavement crack detection using Otsu thresholding for image segmentation. In *2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*; IEEE: Piscataway, NJ, USA, 2018; pp. 1092–1097. [[CrossRef](#)]
58. Safaei, N.; Smadi, O.; Masoud, A.; Safaei, B. An automatic image processing algorithm based on crack pixel density for pavement crack detection and classification. *Int. J. Pavement Res. Technol.* **2022**, *15*, 159–172. [[CrossRef](#)]
59. Hsieh, Y.A.; Tsai, Y. Automated asphalt pavement raveling detection and classification using convolutional neural network and macrotexture analysis. *Transp. Res. Rec.* **2021**, *2675*, 984–994. [[CrossRef](#)]
60. Taj, M.N.A.B.G.; Alruwais, N.; Alshahrani, H.M.; Vijayalakshmi, J.; Shanmugapriya, N.; Jayaprakash, S. Precision crack analysis in concrete structures using CNN, SVM, and KNN: A machine learning approach. *Matéria* **2024**, *29*, e20240551. [[CrossRef](#)]
61. Fakhri, D.; Khodayari, A.; Mahmoodzadeh, A.; Hosseini, M.; Ibrahim, H.H.; Mohammed, A.H. Prediction of Mixed-mode I and II effective fracture toughness of several types of concrete using the extreme gradient boosting method and metaheuristic optimization algorithms. *Eng. Fract. Mech.* **2022**, *276*, 108916. [[CrossRef](#)]
62. Mir, B.A.; Sasaki, T.; Nakao, K.; Nagae, K.; Nakada, K.; Mitani, M.; Tsukada, T.; Osada, N.; Terabayashi, K.; Jindai, M. Machine learning-based evaluation of the damage caused by cracks on concrete structures. *Precis. Eng.* **2022**, *76*, 314–327. [[CrossRef](#)]

63. Cha, Y.J.; Choi, W.; Büyüköztürk, O. Deep learning-based crack damage detection using convolutional neural networks. *Comput.-Aided Civ. Infrastruct. Eng.* **2017**, *32*, 361–378. [[CrossRef](#)]
64. Liu, J.; Yang, X.; Lau, S.; Wang, X.; Luo, S.; Lee, V.C.S.; Ding, L. Automated pavement crack detection and segmentation based on two-step convolutional neural network. *Comput.-Aided Civ. Infrastruct. Eng.* **2020**, *35*, 1291–1305. [[CrossRef](#)]
65. Dung, C.V. Autonomous concrete crack detection using deep fully convolutional neural network. *Autom. Constr.* **2019**, *99*, 52–58. [[CrossRef](#)]
66. Yang, L.; Huang, H.; Kong, S.; Liu, Y. A deep segmentation network for crack detection with progressive and hierarchical context fusion. *J. Build. Eng.* **2023**, *75*, 106886. [[CrossRef](#)]
67. Ali, R.; Chuah, J.H.; Talip, M.S.A.; Mokhtar, N.; Shoab, M.A. Structural crack detection using deep convolutional neural networks. *Autom. Constr.* **2022**, *133*, 103989. [[CrossRef](#)]
68. Liu, C.; Tian, L.; Wang, P.; Yu, Q.Q.; Zhong, X.; Miao, J. Knowledge-driven 3D damage mapping and decision support for fire-damaged reinforced concrete structures using enhanced deep learning and multi-modal sensing. *Adv. Eng. Inform.* **2025**, *68*, 103715. [[CrossRef](#)]
69. Zhao, L.; Xu, T. Multimodal perception-guided intelligent pavement crack detection approach: A novel combining vision and semantic supervision framework. *Measurement* **2025**, *258*, 119333. [[CrossRef](#)]
70. Rizia, M.M.; Reyes-Munoz, J.A.; Ortega, A.G.; Choudhuri, A.; Flores-Abad, A. Intelligent Crack Detection in Infrastructure Using Computer Vision at the Edge. *Expert Syst.* **2025**, *42*, e13784. [[CrossRef](#)]
71. Ghosh Mondal, T.; Jahanshahi, M.R.; Wu, R.T.; Wu, Z.Y. Deep learning-based multi-class damage detection for autonomous post-disaster reconnaissance. *Struct. Control Health Monit.* **2020**, *27*, e2507. [[CrossRef](#)]
72. Wang, J.; Qin, Q.; Zhao, J.; Ye, X.; Feng, X.; Qin, X.; Yang, X. Knowledge-based detection and assessment of damaged roads using post-disaster high-resolution remote sensing image. *Remote Sens.* **2015**, *7*, 4948–4967. [[CrossRef](#)]
73. Hu, K.; Chen, Z.; Kang, H.; Tang, Y. 3D vision technologies for a self-developed structural external crack damage recognition robot. *Autom. Constr.* **2024**, *159*, 105262. [[CrossRef](#)]
74. Concrete Structural Defect Imaging Dataset: High-Quality Multi-Defect Concrete Images for Detection and Rehabilitation. Kaggle Dataset. 2025. Available online: <https://www.kaggle.com/datasets/programmer3/concrete-structural-defect-imaging-dataset> (accessed on 1 November 2025).
75. NEU Surface Defect Database. Kaggle Dataset. 2025. Available online: <https://www.kaggle.com/datasets/kaustubhdikshit/neu-surface-defect-database> (accessed on 1 November 2025).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.