



Towards in-situ failure assessment: Deep learning on DIC results for laminated composites

Amir Mohammad Mirzaei

Faculty of Engineering and Applied Science, Cranfield University, Cranfield, MK43 0AL, UK

ARTICLE INFO

Keywords:

Laminated composites
Notch
Digital image correlation
Machine learning
Artificial neural networks

ABSTRACT

Predicting fracture load in laminated composites with stress raisers is challenging due to complex failure mechanisms such as delamination, fibre breakage, and matrix cracking, which are heavily influenced by fibre orientation, layup sequence, and notch geometry. This study aims to address this by developing a novel deep learning framework that leverages solely experimental strain field data from Digital Image Correlation (DIC) for accurate, in-situ predictions—bypassing the need for finite element simulations or empirical calibrations. Two alternative architectures are explored: a multi-layer perceptron (MLP) that processes numerical values of maximum principal strain from a targeted rectangular region ahead of the notch, enhanced by advanced feature selection (mutual information, Lasso, and SHAP) to focus on critical data points; and a convolutional neural network (CNN) trained on full-field strain images, bolstered by data augmentation to handle variability and prevent overfitting. Validated across 116 quasi-static tests encompassing 31 distinct configurations—including six layups (quasi-isotropic to highly anisotropic) with four off-axis angles for open-hole specimens, and one cross-ply layup with four off-axis and four on-axis notch orientations for U-notched specimens—the MLP and CNN achieve coefficients of determination (R^2) of 0.86 and 0.82, respectively. The framework captures a broad spectrum of damage modes and responses, from brittle fibre-dominated fracture to ductile delamination-driven failure, and due to its computational efficiency and reliance only on DIC measurements, the approach enables practical in-situ fracture load estimation.

1. Introduction

The anisotropic behaviour of laminated composites poses significant challenges in predicting fracture loads for notched structures, where intricate interactions among failure mechanisms—such as fibre breakage, matrix cracking, and delamination—often prevent straightforward analysis through conventional models, particularly under the influence of variables like fibre orientation, stacking sequence, and notch shape [1,2]. Classical approaches such as the point stress/average stress criteria by Whitney and Nuismer [3] were among the first to relate stress distributions around notches to laminate failure strengths. These methods introduced a characteristic length over which the stress is averaged to account for damage tolerance, but the optimal parameters often depend on empirical calibration. Another early contribution to understanding notch sensitivity in composites was made by Pipes et al. [4]. They proposed a notch strength model that accounted for the influence of notch size and material anisotropy on the failure of laminated composites, providing a complementary perspective to stress-based criteria. In contrast, modern full-field techniques like digital image

correlation (DIC) [5] have revolutionized damage visualization, allowing precise tracking of strain patterns and evolving degradation in composites [6]. Initial experiments underscored how pre-catastrophic damage, including matrix cracks and delaminations, reshapes strain fields and diminishes residual load capacity in notched laminates [7]. For instance, Ambu et al. [8] leveraged DIC to trace progressive failure in notched carbon/epoxy and carbon/PEEK systems, revealing how varied mechanisms—fibre rupture, delamination, and beyond—trigger unique strain redistributions that directly impact strength. Such DIC insights have further illuminated the dual-edged role of subcritical damage in modulating stress raisers, either amplifying or mitigating them to influence ultimate laminate endurance, positioning DIC as an instrument for dissecting failure dynamics and their effects on structural integrity [9]. With advances in computational power, researchers have increasingly turned to machine learning (ML) to use strain field data—both from experiments (e.g., DIC) and high-fidelity simulations—in order to predict fracture loads more accurately across different notch geometries, off-axis fibre orientations, and laminate layups [10].

Given the complexity of composite failure and the richness of DIC

E-mail address: amir.mirzaei@cranfield.ac.uk.

<https://doi.org/10.1016/j.compositesb.2025.113241>

Received 20 August 2025; Received in revised form 11 November 2025; Accepted 27 November 2025

Available online 29 November 2025

1359-8368/© 2025 The Author. Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).

data, machine learning offers a promising approach to model and predict fracture loads more accurately. Azeem and Pinho [11] introduced a machine-learning-enhanced characteristic length method for open-hole tension laminates. In their approach, multiple regression models were trained to directly predict the ultimate failure load of an open-hole composite specimen. By learning from a database of finite element simulations and experiments, their ML models could capture the influence of variables like hole size and plate width that go beyond the ideal infinite-width assumption of analytical solutions. Balducci et al. [12] redefined open-hole failure prediction as a classification problem, learning the failure envelope (the set of stress combinations leading to failure) for a composite plate with a hole. They trained Support Vector Machine classifiers to distinguish safe vs. failure loading states and showed that their best SVM (with a tuned radial-basis kernel) achieved over 90 % accuracy in classifying failure vs. non-failure cases. This demonstrates that even “traditional” ML algorithms (SVMs, decision trees, etc.), when provided with sufficient training data from simulations or experiments, can learn complex failure criteria in composites. Beyond open-hole tension, researchers have applied ML to other loading and geometric scenarios. Li et al. [13] proposed a high-throughput simulation and ML framework to predict the allowable compression load of notched laminates under compression. They generated a large dataset of finite element analyses (FEA) for various notch sizes, laminate layups, and loading conditions, and then trained regression models to map these input parameters to the laminate’s critical buckling or failure load. This approach effectively produced a surrogate model that can instantly predict the notched compression strength given a set of design parameters, eliminating the need for exhaustive incremental testing. Similarly, a deep transfer learning approach by Li et al. [14] constructed an entire allowable load space for notched composites under different design configurations. In that study, an ensemble of deep neural networks was first trained on a broad range of notch geometries, materials (fibre and matrix properties), stacking sequences (including various off-axis ply angles and ply counts), and loading types (tension, compression, biaxial). This ensemble DNN could predict the failure load for a given set of design parameters within the training distribution.

While the above approaches often use summary features (like geometric parameters or aggregate stress values), an alternative paradigm is to employ the full-field strain data from experiments or high-fidelity simulations as direct input to machine learning models. The idea is that the strain field reflects both stress concentrations and damage states, thus containing rich information for predicting the load-bearing capacity. Nastos et al. [15] demonstrated this concept in a non-destructive strength prediction method for composites. In their work, a quasi-isotropic carbon/epoxy laminate with a central hole was loaded to only 20 % of its expected failure load, and the surface strain field was measured with DIC. These moderate-load strain maps – well before any visible damage or cracking – were then fed into a deep learning model (a convolutional neural network) that had been trained on thousands of random FEA simulations of the same laminate under random strength properties. The CNN learned to correlate the pattern of strains at 10–20 % load with the eventual ultimate tensile strength of the specimen. When tested on actual specimens (loaded to 10–20 % and then unloaded), the model predicted failure loads with good accuracy and low error variability. Another use of full-field strain data is in damage identification and early warning of failure. Rather than directly outputting a numeric failure load, some studies train ML models to detect and locate damage from strain patterns, which can indirectly predict failure. Wang et al. [16] developed a CNN-based semantic segmentation model to analyze DIC strain maps of CFRP laminates with open-holes. Using a DeepLabv3+ architecture [17] (with a ResNet-50 backbone [18]), they trained the network to perform pixel-level classification of the strain field – essentially labelling regions as damaged (plasticity, fibre break, delamination) or healthy based on the strain distribution. The training data were generated via high-fidelity finite element simulations of tension tests, ensuring the network saw a wide

variety of damage patterns. When applied to real-time DIC data from composite coupons, the model achieved a high segmentation accuracy (mean IoU $\approx$ 0.92) in identifying damage zones. This automated interpretation of strain contours means that the onset of critical damage can be detected online, without relying on a human inspector.

Following the reviewed studies, this paper builds on previous investigations by the author to employ machine learning models trained only on field distributions to predict the failure of notched components. In Ref. [19], stress, strain, and energy fields were utilised to estimate the fatigue life of notched components. The framework was validated across diverse loading conditions, notch geometries, and materials, including steel, additively manufactured PLA, and laminated composites [20]. The study demonstrated that all field distributions effectively predicted fatigue life, with stress fields generally providing the most stable and accurate predictions across all cases. This framework was further extended in Ref. [21] to predict fracture load (defined as maximum load based on load-displacement curve) and crack initiation angle under mixed-mode in-plane loading conditions in cracked bodies. The model was validated using five distinct specimen configurations, each subjected to at least six different mode mixity conditions. Results showed that stress field data, when combined with a multi-layer perceptron (MLP), outperformed conventional models such as Generalized Maximum Tangential Stress (MTS) or Theory of Critical Distances (TCD) in prediction accuracy.

In this paper, the approach is extended to include not only anisotropic behaviour but also the direct use of DIC results. This research addresses a critical gap by exclusively utilizing experimental strain field data from Digital Image Correlation (DIC) to train deep learning models for predicting fracture loads across diverse layup configurations and notch geometries, eliminating the need for additional finite element analyses or post-processing. To utilize DIC results, two distinct approaches are employed: first, the use of numerical values of maximum principal strain at representative nodes ahead of the notch; second, images of the principal strain field distribution around the notch obtained from DIC. Studies have shown that the principal strain distribution is sensitive to changes in layup configuration and notch geometry, making it a robust metric for fracture prediction [6,9,16]. For the first approach, an MLP model with stepwise feature selection algorithms is used to handle sparse to dense sets of representative nodes in DIC data. To accommodate different notch geometries, a novel approach is introduced, where nodes are considered within a rectangular region ahead of the notch. For the second approach, a Convolutional Neural Network (CNN) based on the DenseNet-121 architecture is employed. Since both approaches are data-driven and typically require extensive data, the models in this paper are trained with a reasonable number of experimental tests by employing techniques such as feature selection for the MLP and data augmentation for the CNN, enhancing efficiency and applicability across various cases. To this end, no additional finite element analyses are used to generate data, and only strain fields calculated from DIC in the experimental campaign are employed. In Section 2, the experimental campaign is presented. In Section 3, the deep learning algorithms are introduced. In Section 4, the results are presented, and the robustness of the framework across varied input data is demonstrated and discussed.

2. Experimental tests

This section presents the experimental data used to validate the proposed framework. In this study, two different geometries are considered: the open-hole tests, whose results were presented by Mitrou et al. [22], and the U-notch specimens, which were manufactured and tested by the same group but are presented in this paper for the first time. The specimens utilised in the experimental campaign were composed of carbon fibre thin-ply carbon/epoxy laminates produced from North Thin-Ply Technology T800/736 LT unidirectional prepreg, with a nominal ply thickness of 0.05 mm and ply properties of $E_1 = 143.83$ GPa, $E_2 = 7.85$ GPa, $G_{12} = 3.80$ GPa and $\nu_{12} = 0.38$. Six 32-ply

laminates (≈ 1.6 mm thick overall) were tested covering a wide range from quasi-isotropic to highly anisotropic behaviour as defined in Ref. [22]: QI [90/-45/0/45]_{4s}, SOFT [90/ ± 45 /90/ ± 45 /90/0]_{2s}, CP30 [90/0/ ± 30 /0/90/ ± 30]_{2s}, CP1560 [90/ ± 60 /0/ ± 15 /90/0]_{2s}, CP1575 [90/ ± 75 /0/ ± 15 /0/90]_{2s} and cross-ply CP [90/0]_{gs}. All specimens were 20 mm wide and 170 mm long overall, with the effective gauge length of 120 mm. The open-hole specimens had a central hole with a diameter of 5 mm, while the U-notched specimens had a notch 5 mm deep and 1 mm wide, created using a 0.5 mm radius drill bit. It is noted that the U-notched specimens were only manufactured from a CP layup. Off-axis loading was introduced by cutting the coupons in the parent plate at 0°, 15°, 45° and 60° relative to the fibre-0° direction, see Fig. 1a). As stated in Ref. [22], each laminate and angle combination had a number of samples that ranged from 3 to 5, giving 109 overall tests for the open-hole configuration. For the U-notched samples, apart from the off-axis testing similar to the open-hole testing, angled notched configurations were introduced, meaning that the off-axis angle was fixed and equal to 0°, while the notch direction was oriented at 15°, 45°, 60° and 90° with respect to the fibre-0° (loading) direction, which is referred to as the on-axis notch test in this study (Fig. 1b)). For the U-notched specimens only one sample per configuration was tested yielding 7 samples. Notably, the 0° off-axis configuration corresponds to the 90° on-axis case. In total, 116 mechanical tests were conducted across 31 distinct configurations. Configuration names follow the pattern FAMILY θ_{off} , where FAMILY denotes the laminate class (e.g., CP = cross-ply; QI = quasi-isotropic) and θ_{off} is the off-axis angle in degrees. The suffix “_U” indicates a U-notched specimen, and the suffix “_UO” denotes on-axis loading for a U-notched specimen (i.e., CP θ_{on} _UO). Replicate counts per configuration were: 5 tests for {CP1575_00, CP1575_45, CP1575_60, CP30_00, CP30_15, CP30_45, CP30_60, QI_60, CP1560_00, CP1560_15, CP1560_45, CP1560_60, SOFT_00, SOFT_60}; 4 tests for {CP1575_15, QI_15, QI_45, SOFT_15, SOFT_45, CP_00, CP_15, CP_45, CP_60}; 3 tests for {QI_00}; and 1 test for each U-notched case {CP_00_U, CP_15_U, CP_15_UO, CP_45_U, CP_45_UO, CP_60_U, CP_60_UO}. A comprehensive overview of all tested specimen configurations, including detailed geometrical parameters, layup sequences, and loading conditions, is provided in Table 1.

To obtain strain fields based on DIC, full-field kinematics of the specimens were monitored with a single-camera 2D digital image correlation setup. Specifically, the system utilised a Canon EOS 80D camera fitted with a Canon EF-S 18–135 mm f/3.5–5.6 IS STM lens, recording at 60 fps with a 1080p Full-HD resolution. Images were processed to extract the strain field in VIC-2D software using a 15-pixel subset, a 7-

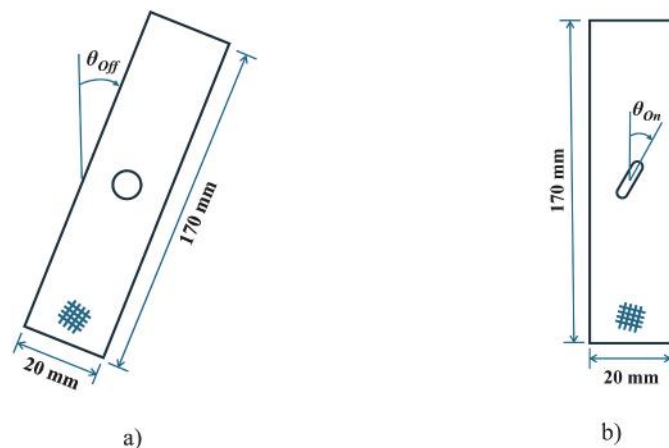


Fig. 1. Schematic of loading configurations for notched composite specimens. a) Off-axis loading with angles $\theta_{off} = 0^\circ, 15^\circ, 45^\circ, 60^\circ$ defined relative to the fibre-0° direction [22]. b) On-axis loading for U-notched cross-ply (CP) specimens, with angles $\theta_{on} = 15^\circ, 45^\circ, 60^\circ, 90^\circ$ relative to the loading direction (fibre-0°).

Table 1

Summary of specimen geometries, notch dimensions, layup sequences, loading configurations, and the corresponding number of tests for each configuration.

Notch Geometry	Notch Dimensions (mm)	Layup	Loading Configuration	Number of Tests
Open-hole	Hole radius: 2.5	QI [90/-45/0/45] _{4s} , SOFT [90/ ± 45 /90/ ± 45 /90/0] _{2s} , CP30 [90/0/ ± 30 /0/90/ ± 30] _{2s} , CP1560 [90/ ± 60 /0/ ± 15 /90/0] _{2s} , CP1575 [90/ ± 75 /0/ ± 15 /0/90] _{2s} , CP [90/0] _{gs}	4 off-axis: $\theta_{off} = 0^\circ, 15^\circ, 45^\circ, 60^\circ$	109
U-notch	Depth: 5 Width: 1 Root radius: 0.5	CP [90/0] _{gs}	4 off-axis: $\theta_{off} = 0^\circ, 15^\circ, 45^\circ, 60^\circ$ +4 on-axis (3 unique): $\theta_{on} = 15^\circ, 45^\circ, 60^\circ, 90^\circ$	7

pixel step, and the fill-boundary option to obtain data close to the notch boundary. In the Appendix, the strain fields obtained using DIC for some arbitrary cases are presented and compared with the finite element results (for validation purposes only).

In order to emphasise and demonstrate the different mechanical responses between configurations, the results of mechanical testing are presented as force–displacement curves for eight different configurations in Fig. 2. In this figure, the subfigure labels indicate the layup configuration, with different colours representing various off-axis angles—except for the U-notched specimens tested under the on-axis configuration (CP_UO), where the angles correspond to the notch direction relative to the fibre-0° (loading) direction.

From Fig. 2, it can be seen that the configurations exhibited a variety of failure behaviours, ranging from brittle to ductile, with a wide range of failure loads—from approximately 8 kN to 26 kN—and displacements at failure from less than 1.5 mm to more than 17 mm. This confirms the diversity and complexity of the experimental data available for validating the model.

To give a clear sense of specimen-to-specimen variability, the coefficients of variation (CV) in ultimate load across configurations are: the largest scatter for CP1575_00 (8.8 %), followed by SOFT_60 (7.1 %) and CP1560_15 (6.9 %). In contrast, the smallest scatter is observed for CP1575_60 (1.3 %), followed by CP1575_45 and CP_60 (both 1.8 %), then CP_45 (2.0 %), and CP30_00 and CP1560_60 (both 2.2 %).

The diversity in observed failure modes is evident from the actual appearance of the failed samples and failure surfaces. Fig. 3 demonstrates representative failure patterns from six different notched specimens, highlighting some distinct failure modes and underlying mechanisms associated with the failure of the tested laminates. As reported in Ref. [22], the failure patterns observed varied widely—from brittle net-section, through-thickness cracks characteristic of thin-ply laminates, to more complex modes involving delamination, matrix cracking, and fibre pull-out also observed in thin-ply systems. The authors directly attributed these differences to the level of anisotropy in the laminates. They observed that specimens with lower anisotropy tended to fail in a brittle manner, regardless of the loading angle. In contrast, highly anisotropic laminates, especially under off-axis loading, exhibited more complex failure patterns involving extensive subcritical damage before final fracture.

Overall, greater anisotropy in the laminate resulted in more

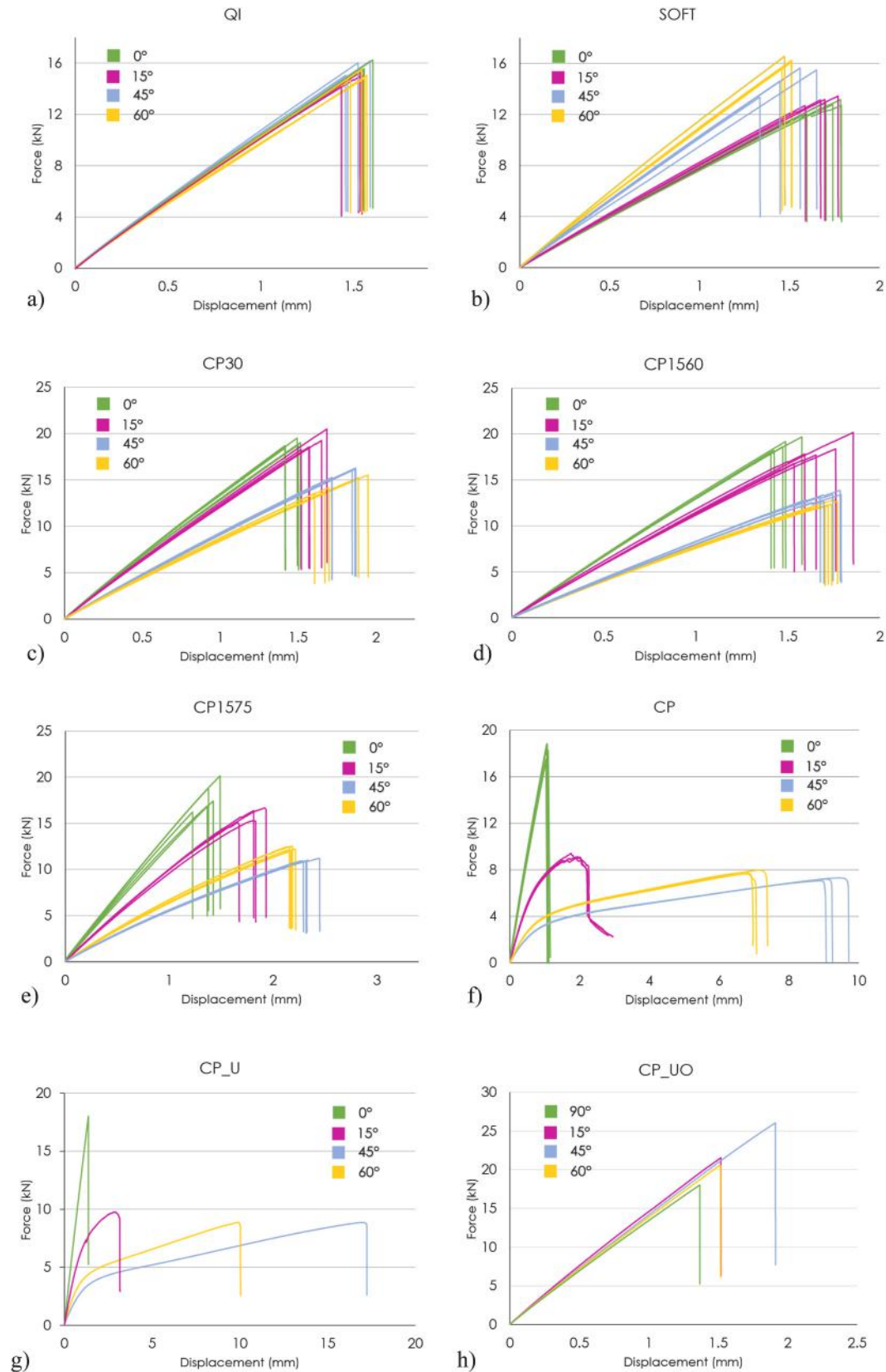


Fig. 2. Force-displacement curves for notched laminated composites under quasi-static loading [22]. Subfigures (a–h) represent configurations: a) QI, b) SOFT, c) CP30, d) CP1560, e) CP1575, f) CP with open-hole, g) CP with U-notch (CP_U) under off-axis loading (0°, 15°, 45°, 60°), and h) CP with U-notch under on-axis loading (CP_UO, notch angles 15°, 45°, 60°, 90°). Each colour denotes a specific off-axis or notch angle. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

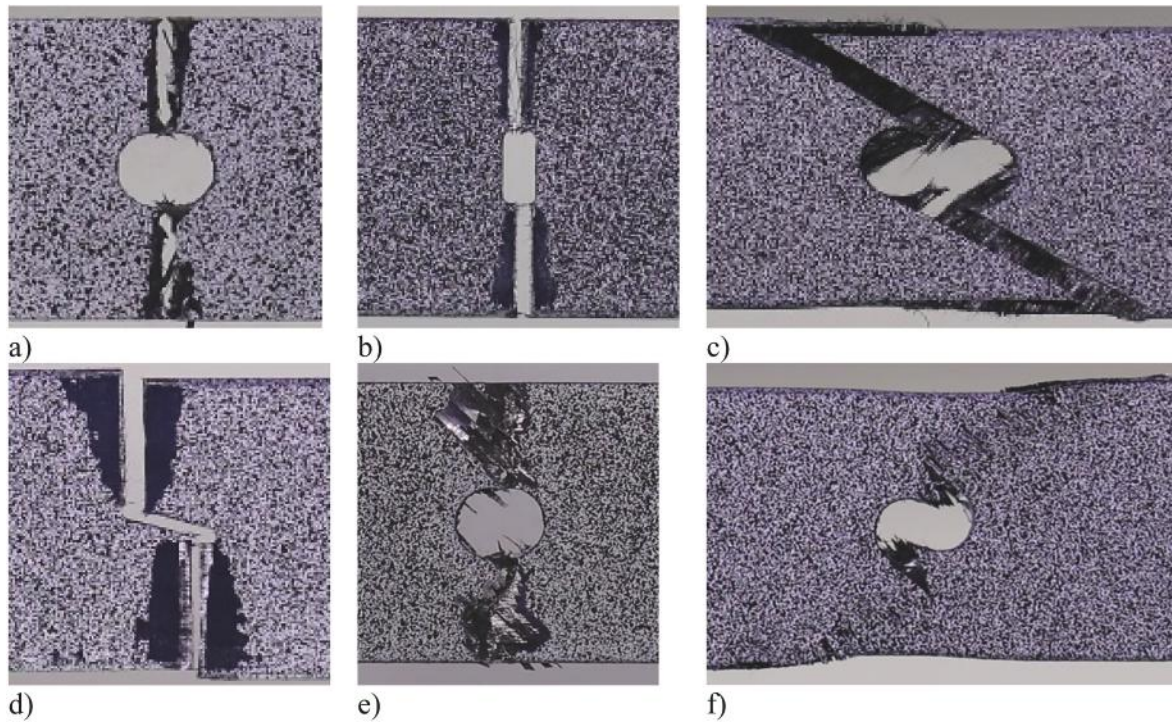


Fig. 3. Representative post-failure fracture patterns in notched laminated composites, captured via optical photography [22]. Specimens include: a) SOFT_00 and b) CP_00_U (brittle, fibre-dominated net section failure), c) CP_60 (matrix-dominated shear-out with interlaminar shear), d) CP_15_UO (fibre-dominated net section failure along a path perpendicular to the load application direction), e) QI_45 (brittle, fibre-dominated net section failure), and f) CP_15 (localized band of rotated fibres accompanied by matrix damage).

pronounced changes in failure mode across different off-axis loading angles. The U-notched specimens exhibited similar failure mechanisms to the open-hole specimens; however, these failure processes were confined to a more localized region near the notch, Fig. 4. This increased localization is likely attributed to the higher stress concentration associated with the smaller notch tip radius. It is noted that the angle of the notch with respect to load direction, inducing a mixed mode state, appears to have minimal effect on the failure pattern since the crack still propagated from the strong concentration points in a net-section through-thickness failure form along a path perpendicular to the load application direction. For more details and a more extensive discussion on the experimental findings, refer to Ref. [22].

3. Machine learning tools

3.1. Utilisation of DIC results in deep learning for fracture load prediction

The results obtained from DIC are utilised as inputs for machine learning (ML) algorithms to predict fracture loads. As mentioned earlier,

here, the fracture load is defined as the maximum load that the specimen can withstand. Two distinct ML models were employed: the MLP and the CNN. The MLP model processes numerical values derived from DIC analysis, whereas the CNN model directly utilizes images of the strain field obtained from DIC. Specifically, the maximum principal strain field was selected as the indicator of failure at a predetermined load level across all specimens. Since the failure load is unknown, reference measures like 20 % of the failure load, used in previous studies [15], cannot be employed to obtain a reference load to read the strain field. The chosen load level must be sufficiently high that near-notch strain gradients are resolvable with adequate strain signal to noise ratio for the selected virtual strain gauge (effective spatial averaging length), yet low enough to stay below the minimum observed failure load so the model can be applied to every specimen. When no prior strength data of the sample are available, a simple physics-based rule of thumb calculations is useful. For example, in our case, the laminate most likely to fail first is CP [90/0]_{ss} cut at 45°, which behaves as a through-thickness [± 45] laminate in axial tension. In this configuration, the response is shear-dominated (fibres are not aligned with the load), and first-ply



Fig. 4. Representative post-failure fracture patterns in the U-notched laminated composites. Specimens include: a) CP_00_U (brittle, fibre-dominated net section failure), b) CP_45_U (fibre-dominated net section failure along a path perpendicular to the load application direction).

failure is governed primarily by the in-plane shear strength rather than the fibre tensile strength. For an unnotched coupon, the axial capacity therefore scales with the ply shear strength and section size; introducing an open hole/notch reduces this capacity by a conservative notch-efficiency factor (<1). Under standard coupon geometries, even a conservative notch factor leaves a reasonable margin between this weakest-case and a reference load such as 7 kN. Therefore, 7 kN is a scientifically sound choice: it ensures well-resolved strain fields, while remaining below any possible failure load across all layups and off-axis angles. The impact of varying this reference load level is further investigated in Section 4.2.

For future studies or tests on other laminated or notched configurations, an effective procedure to select a single reference load is: (1) Identify the weakest expected configuration in the planned test matrix (considering material properties, layup anisotropy, notch severity, and loading mode). (2) Estimate a conservative lower bound on its failure load (from prior data, simple net-section or shear-dominated screening estimates, or preliminary tests). (3) Choose a campaign-wide reference load below this bound while ensuring that the maximum-principal-strain (ϵ_1) field is clearly resolvable for the chosen virtual strain-gauge settings—i.e., sufficient strain signal-to-noise ratio (SNR) at the selected subset/step and strain-calculation window. (4) Validate the choice with a brief pilot (fixed region of interest, fixed mm/px sampling, constant subset/step/strain window), confirming stable correlation quality and spatially coherent near-notch gradients without signs of damage. This procedure ensures reproducible, comparable results while remaining flexible enough to apply across different laminate families and notch geometries.

For the MLP approach, after computing the strain field at the selected load level, numerical values surrounding various notches were extracted. To address the variability in notch geometries, a simplified strategy was implemented: a rectangular region was defined ahead of the notch at a consistent relative distance from the notch tip, with fixed relative nodal positions. Key considerations include determining the rectangle's dimensions, the number of nodes it should contain (discretized), and its relative position to the notches. Previous studies by the author [19,21] suggest using as wide a region as possible; however, this is not mandatory, as evidence indicates that while nodes at greater distances contribute to predictions, their influence is less significant than that of nodes closer to the notch, where failure initiates and stresses are higher due to the presence of a notch. This will be further explored in Section 4.4. Regarding the number of discretisation nodes, in simple terms, a relatively fine mesh, similar to that used in a proper finite element analysis of the specimen, is recommended to capture critical information around the notches without excessive computational cost. In practice, a simple script can be used to extract the strain values at predefined nodes. If a given node is not present in the DIC analysis, the script can interpolate the strain value at that location. The effect of density of nodes in the region will be examined in Section 4.3. The rectangle should ideally be positioned as close as possible to the notch tip; however, due to reduced accuracy of DIC calculations (very) close to the notch boundary, proximity is limited. This limitation will be studied in Section 4.4. In this investigation, the rectangle spans from 0.1 mm away from the notch tip to 1 mm from the specimen edge, with the specimen edge located 7.5 mm from the notch tip (6.4 mm length). The longer side of the rectangle is twice the length of the shorter side and is symmetrically aligned with the notch bisector line (see Fig. 5). This area was discretized into a grid of 200 nodes (10 along the shorter side and 20 along the longer side), to consider the largest practically valid neighbourhood around the notch while avoiding boundary artifacts. The strain values at these nodes were flattened into a vector, serving as the sole feature for the MLP model. Note that for on-axis cases, the rectangle maintains the same relative position to the notch bisector line. The input dataset included this vector, specimen identifiers (used only for categorization to iterate over different specimens), and the corresponding failure loads. It is worth emphasising that these identifiers are not included in the predictive

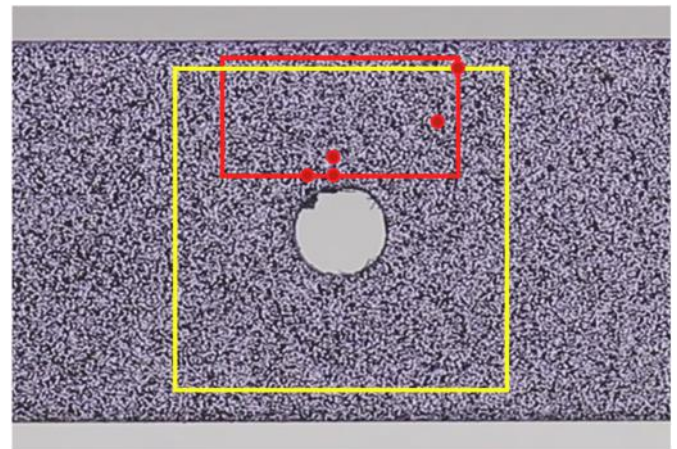


Fig. 5. Definition of analysis windows and sampling points overlaid on the reference image of an open-hole laminate. The red rectangle shows the window used for the MLP; the red dots indicate the selected sampling nodes at which principal strain is read. The yellow square marks the notch-centred region of interest used to crop the DIC-derived principal strain field for the CNN. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

input vector or used in model calculations and are only used to group test data.

For the CNN methodology [23,24], the strain field was extracted from a centred square region at the notch centre, measuring 16 mm per side. To ensure data quality, a margin of 2 mm from the specimen's edge was excluded to eliminate the low accuracy and noise typically observed in DIC measurements near boundaries (see Fig. 5). These strain field maps were saved as images with a resolution of 224×224 pixels, a standard size in CNN due to its balance of feature capture and computational efficiency. To enhance prediction efficiency, the original three-channel (RGB) colour images were converted to single-channel grayscale images. This transformation is appropriate since the first principal strain field is represented by a scalar quantity at each pixel and can be effectively represented with a single channel, significantly reducing model complexity and data size. Prior studies confirm that reducing image channels lowers parameter count and computation time without compromising geometric information critical for model performance [25]. It is crucial that the strain range (minimum and maximum) remains consistent across all specimens (so colour intensities are directly comparable), a condition achievable with many DIC software tools. The input data for the CNN analysis comprised these grayscale images and a file detailing the failure loads for each specimen.

To illustrate the actual DIC outputs at the fixed reference load, Fig. 6 shows representative maximum principal strain maps measured by 2D-DIC at 7 kN for six notched laminate configurations (the same configurations as in Fig. 3). Each panel includes its own colour bar with panel-specific strain limits; the minimum and maximum differ across panels to highlight the diversity of strain magnitudes.

3.2. Multi-layer perceptron (MLP)

This study develops a comprehensive and robust machine learning approach using an MLP neural network [26,27] to accurately predict fracture loads in composite materials. The method integrates advanced feature selection, cross-validation, hyperparameter tuning, and thorough training practices to ensure high accuracy predictions.

The input dataset comprises numerical features extracted from experimental setups and the related failure loads for each specimen. Given the complexity and noise in high-dimensional datasets, a multi-step feature selection process is implemented. This begins with Mutual Information (MI) analysis [28], which evaluates the dependency

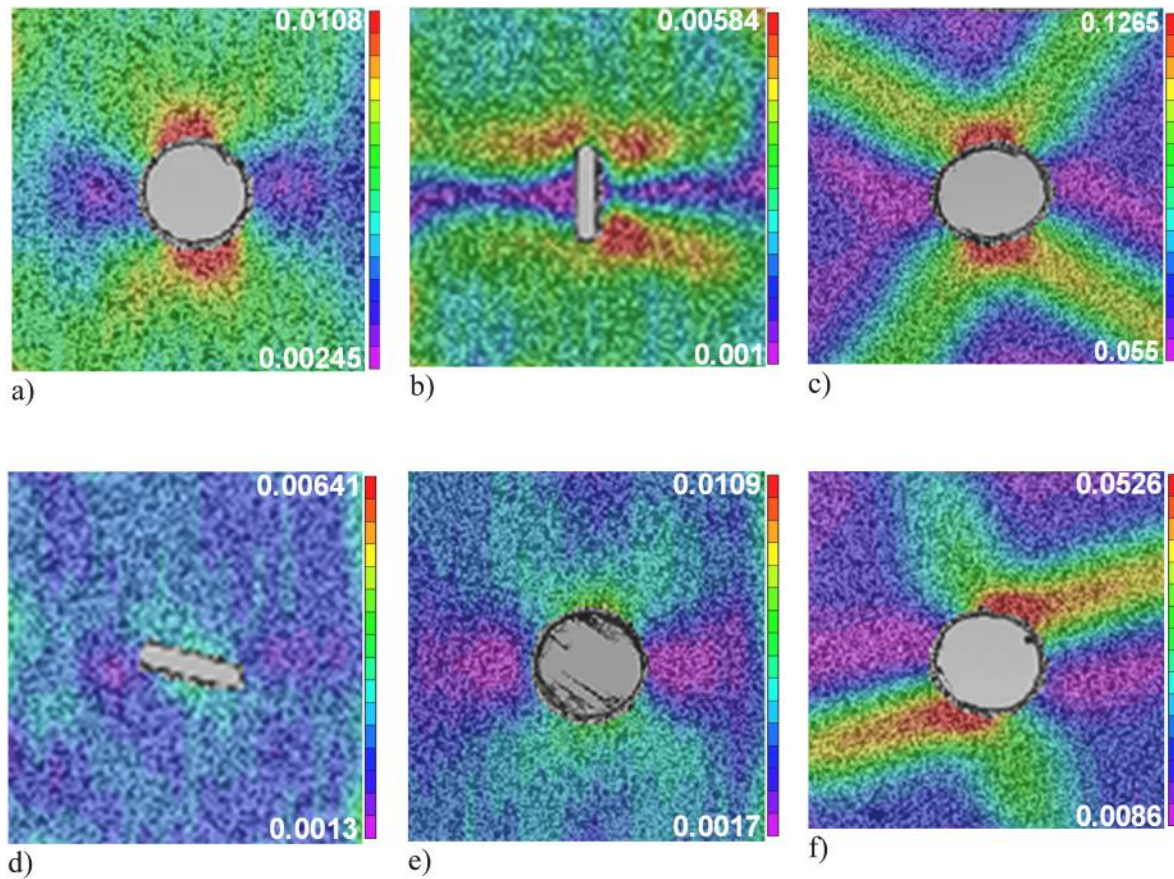


Fig. 6. Representative maximum principal strain maps, measured by 2D-DIC at the reference load (7 kN) for six notched laminate configurations in Fig. 3. Specimens include: a) SOFT_00 and b) CP_00_U, c) CP_60, d) CP_15_UO, e) QI_45, and f) CP_15. Each panel uses its own colour scale; the maximum (top) and minimum (bottom) values are printed adjacent to the colour bar. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

between individual features and target variables to effectively identify the most relevant predictors. Mutual Information between a feature X and the target variable Y is defined as:

$$MI(X; Y) = \sum_{x \in X} \sum_{y \in Y} p(x, y) \log \left(\frac{p(x, y)}{p(x)p(y)} \right) \quad (1)$$

where $p(x, y)$ is the joint probability distribution, and $p(x)$ and $p(y)$ are the marginal probabilities. This step significantly reduces dimensionality by retaining approximately 90 % of the most informative variables. The threshold of retaining 90 % of the most informative variables after MI analysis was chosen to balance dimensionality reduction with information preservation. This cutoff was determined based on preliminary tests showing that lower thresholds led to a decrease in prediction accuracy, while higher thresholds increased computational cost with negligible gains. Following MI, the dataset undergoes further refinement using Lasso regression (Least Absolute Shrinkage and Selection Operator) [29] combined with cross-validation. Lasso regression minimizes the following objective function:

$$\min_{\beta} \left(\frac{1}{2n} \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^p |\beta_j| \right) \quad (2)$$

where n is the number of samples, p is the number of features, y_i is the target variable, x_{ij} are the feature values, β_j are the coefficients, and λ is the regularisation parameter. Lasso penalizes less relevant features by shrinking their coefficients towards zero, resulting in a sparse yet highly informative subset. The final stage of feature selection employs a Random Forest (RF) model [30,31] in conjunction with SHAP (SHapley

Additive exPlanations) analysis [32,33]. After trial and error across preliminary tests, the final model employed the top 5 features to achieve the best combination of efficiency and accuracy.

SHAP values quantify the importance of each feature by calculating the average marginal contribution of a feature across all possible subsets of features:

$$\phi_j = \sum_{S \subseteq N \setminus j} \frac{|S|!(p - |S| - 1)!}{p!} (f(S \cup j) - f(S)) \quad (3)$$

where N is the set of all features, S is a subset of features, p is the total number of features, and $f(S)$ is the model output with features in S . Because univariate MI and linear Lasso can under-rank purely interaction-driven features, the RF + SHAP stage recovers non-linear and interaction effects before final selection. This procedure yields a concise and reliable set of features tailored for accurate prediction.

To thoroughly evaluate the generalization capabilities of the model and prevent data leakage, the pipeline utilizes a two-layer cross-validation approach. An outer Leave-One-Group-Out (LOGO) cross-validation strategy is combined with an inner group k-Fold cross-validation [34] for hyperparameter optimization. The outer cross-validation loop was implemented with 31 folds, corresponding exactly to the 31 unique experimental configurations (combinations of layup, notch geometry, and loading angle), so that all specimens and repeats of that configuration are held out together, never mixing training and test data from the same geometry. Each outer fold isolates all specimens from a single configuration as the test set, ensuring that no data from the same physical scenario is seen during training or feature selection. Within each fold, scaling and feature selection are performed

exclusively on the training subset to strictly prevent any possibility of data leakage. While this rigorous per-fold selection ensures no data leakage, a post-hoc analysis confirmed the stability of the process, revealing that a core subset of nodes—predominantly located in the physical region of highest stress concentration ahead of the notch—was consistently identified as most informative across the majority of cross-validation folds.

Within each outer fold, the inner loop implements group k-Fold splitting to accurately tune hyperparameters, ensuring results accurately reflect model performance rather than sample dependencies. Hyperparameters including hidden layer sizes, dropout rates and learning rates are optimised in Optuna [35] using the Tree-structured Parzen Estimator (TPE) sampler. Optuna efficiently navigates the hyperparameter space through sequential model-based optimization, systematically searching for optimal values. Specifically, hidden layer sizes range from 32 to 128 neurons, dropout rates vary between 0 and 0.2, and learning rates span a logarithmic scale from 0.001 to 0.05. This strategic exploration provides comprehensive yet computationally efficient hyperparameter tuning.

The neural network architecture selected is a three-layer MLP [26, 27] enhanced by batch normalization (BN) to stabilise and accelerate the training process. Hyperbolic tangent (Tanh) activation functions are utilised within each hidden layer, interspersed with dropout layers to prevent overfitting and enhance generalization. The employed MLP architecture can be represented as:

$$\begin{aligned}
 &\text{Input : } \mathbf{x} \in \mathbb{R}^d \\
 &\text{HiddenLayers : For } i = 1, 2, 3 : \\
 &\quad \mathbf{z}_i = \mathbf{W}_i \mathbf{h}_{i-1} + \mathbf{b}_i \quad \text{with } \mathbf{h}_0 = \mathbf{x} \\
 &\quad \mathbf{z}'_i = \text{BN}(\mathbf{z}_i) = \gamma_i \frac{\mathbf{z}_i - \mu_{\mathbf{z}_i}}{\sqrt{\sigma_{\mathbf{z}_i}^2 + \epsilon}} + \beta_i \\
 &\quad \mathbf{h}_i = \tanh(\mathbf{z}'_i) \\
 &\quad \mathbf{h}'_i = \text{Dropout}(\mathbf{h}_i, p_i) \\
 &\text{OutputLayer : } \hat{\mathbf{y}} = \mathbf{w}_4^\top \mathbf{h}_3 + b_4
 \end{aligned} \tag{4}$$

where $\mathbf{W}_i$ and $\mathbf{b}_i$ are the weights and biases, p_i is the dropout probability, and $\hat{\mathbf{y}}$ is the predicted fracture load. Training employs the Adam optimizer for efficient gradient descent. The Adam optimizer updates the parameters using:

$$\begin{aligned}
 \mathbf{m}_t &= \beta_1 \mathbf{m}_{t-1} + (1 - \beta_1) \mathbf{g}_t, \\
 \mathbf{v}_t &= \beta_2 \mathbf{v}_{t-1} + (1 - \beta_2) \mathbf{g}_t^2, \\
 \hat{\mathbf{m}}_t &= \frac{\mathbf{m}_t}{1 - \beta_1^t}, \quad \hat{\mathbf{v}}_t = \frac{\mathbf{v}_t}{1 - \beta_2^t} \\
 \theta_t &= \theta_{t-1} - \eta \frac{\hat{\mathbf{m}}_t}{\sqrt{\hat{\mathbf{v}}_t} + \epsilon}
 \end{aligned} \tag{5}$$

where $\mathbf{g}_t$ is the gradient at time t , η is the learning rate, β_1 and β_2 are exponential decay rates, and ϵ is a small constant. Training efficiency is optimised further using batch-wise gradient updates and an early stopping mechanism based on validation loss performance. Additionally, an adaptive learning rate scheduler (ReduceLROnPlateau) [36] dynamically adjusts learning rates when training stagnation is detected, ensuring efficient convergence towards optimal model weights without unnecessary computational overhead.

Following hyperparameter optimization and training within each fold, comprehensive evaluation metrics—including Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R^2)—are calculated. These metrics are defined as:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \tag{6}$$

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \tag{7}$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \tag{8}$$

where y_i is the true value, $\hat{y}_i$ is the predicted value, $\bar{y}$ is the mean of the true values, and n is the number of samples.

These metrics provide detailed insights into predictive accuracy, error magnitude, and overall reliability. Results from all outer cross-validation folds are combined to deliver an extensive and holistic assessment of the model's predictive performance. Fig. 7 demonstrates the summary of the employed MLP algorithm for failure load prediction.

3.3. Convolutional neural network (CNN)

This study employs a CNN [23,24], specifically a modified DenseNet-121 [37] architecture, to predict fracture loads in notched laminated composites using strain field images. The developed method integrates image preprocessing, data augmentation, and training practices. These components are carefully designed to achieve high predictive accuracy, prevent data leakage, and ensure the generalizability and reliability of predictions. Given the sensitivity of fracture load predictions to strain variations near notches, precise and consistent image preprocessing is crucial. Fig. A.1 (in Appendix) illustrates representative examples of strain field maps used as input for the CNN. The preprocessing pipeline features two primary enhancements: automatic notch masking via in-painting [38] and Contrast Limited Adaptive Histogram Equalization (CLAHE, clip limit = 2.0, tile grid = 8 × 8) [39]. Notch masking involves automatically detecting the notch region (black pixels) using Otsu thresholding (3 iterations with a 3 × 3 elliptical kernel) [40]. Otsu thresholding maximizes the inter-class variance:

$$\sigma_b^2(t) = w_0(t)w_1(t)(\mu_0(t) - \mu_1(t))^2 \tag{9}$$

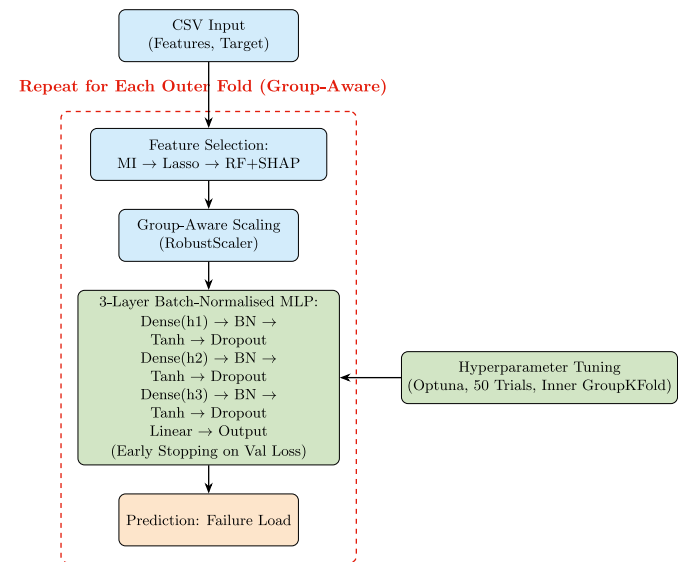


Fig. 7. Schematic representation of the 3-layer batch-normalised MLP used for predicting fracture load from full-field numerical features derived from strain field data. The pipeline includes outer-loop cross-validation with group-wise holdout, followed by per-fold feature selection (mutual information, Lasso, and SHAP), robust group-aware scaling, and hyperparameter tuning via Optuna. The MLP architecture consists of three hidden layers, each followed by batch normalization, Tanh activation, and dropout. The final output layer predicts fracture load for each specimen.

where $w_0(t)$ and $w_1(t)$ are the probabilities of the two classes separated by threshold t , and $\mu_0(t)$ and $\mu_1(t)$ are their means. Morphological dilation [41] is then applied to refine the detected region. Dilation is defined as:

$$A \oplus B = z|(B)_z \cap A \neq \emptyset \quad (10)$$

This region is replaced with local mean grayscale values calculated from adjacent pixels, improving the signal-to-noise ratio by clearly revealing critical strain gradients around the notch. This in-painting strategy was specifically chosen over masking with a constant value to preserve the local continuity of the strain field data and prevent the introduction of artificial high-frequency edges at the mask boundary, which could be misinterpreted as physical features by the CNN's convolutional filters. Following notch masking, CLAHE is applied to enhance the contrast of subtle strain differences by equalising local histograms within small image regions. This technique effectively highlighted key features necessary for the neural network's learning process. The chosen DenseNet-121 CNN architecture is pretrained on ImageNet to ensure robust initial feature extraction capabilities. Given the single-channel principal-strain inputs, the ImageNet-pretrained DenseNet-121 was adapted by replacing the RGB stem (conv0) with a 1-channel convolution. To preserve the pretrained representation, all remaining feature layers were frozen, and their BatchNorm layers were held in evaluation mode; only the new 1-channel stem and the regression head were optimised. The final fully connected layer output a single continuous value (fracture load), and the optimizer was applied solely to the trainable stem and head.

To enhance model robustness and avoid overfitting, the training dataset undergoes data augmentation. This includes random horizontal and vertical flips, slight rotations ($\pm 10^\circ$), and minor brightness and contrast adjustments [42]. It is worth emphasising that these augmentations do not alter the underlying physics: they preserve the spatial distribution and magnitude of the principal-strain field around the notch (i.e., the salient gradient/intensity cues); this is consistent with Whitney–Nuismer model that failure is governed by the local field in a neighbourhood of the notch [3]. Each transform is applied on-the-fly with the default 0.5 probability in torchvision.transforms, yielding an average four-fold increase in distinct training samples per epoch. The modest rotation span and conservative colour jitter widen the effective dataset while preserving the local strain-gradient patterns essential for accurate fracture-load regression. Small jitter helps robustness to minor DIC exposure/lighting changes. These augmentation techniques diversify the training data, allowing the CNN to learn generalizable features rather than memorizing specific pixel configurations. Note that all augmentations are applied only to the inner training subset within each outer fold; inner-validation and outer-test images are evaluated without augmentation.

The model optimization process involves a thorough two-stage training approach within each LOGO cross-validation fold, resulting in a total of 31 folds for 31 distinct test configurations. Initially, each training fold is internally partitioned into an 80 % training and 20 % validation subset, ensuring complete isolation of the test specimen to strictly prevent data leakage. It is worth emphasising that hyperparameter tuning is conducted exclusively on this internal subset, utilizing five chosen trials exploring critical parameters: learning rate (logarithmically sampled between 10^{-5} and 10^{-3}), and weight decay (logarithmically sampled between 10^{-6} and 10^{-3}). The AdamW optimizer is used, and batch size is fixed at 8 to stabilise training on the relatively small dataset, with each trial trained for up to 50 epochs on the internal training set and evaluated on the internal validation set. This strategy is adopted to ensure a balance between accuracy and computational cost, as CNN-based analyses are considerably more computationally intensive than the MLP.

With optimised hyperparameters determined, the final CNN for each fold is trained on the complete internal dataset (combined training and

validation subsets) for an additional 20 epochs. During this 20-epoch refit, no validation-driven decisions are made: early stopping and any validation-keyed schedulers are disabled. A fixed cosine-annealing schedule is used, and only the final checkpoint is evaluated on the outer-fold test set. The training utilizes the Smooth $L1$ loss function [43] together with a very light target shrinkage ($\alpha = 0.01$) applied during training and validation to stabilise optimization in small batches; test uses the original targets (no shrinkage). The Smooth $L1$ loss is defined as:

$$\text{Smooth } L1(x) = \begin{cases} 0.5x^2 & |x| < 1 \\ |x| - 0.5 & \text{otherwise} \end{cases} \quad (11)$$

where $x = y - \hat{y}$, with y the true label and $\hat{y}$ the predicted value. During training we replace each target by a lightly shrunk value toward the mini-batch mean:

$$y'_i = (1 - \alpha) y_i + \alpha \bar{y}_B, \bar{y}_B = \frac{1}{|B|} \sum_{j \in B} y_j \quad (12)$$

where α is small shrinkage coefficient (set to 0.01), B denotes the current training mini-batch and $\bar{y}_B$ is the mini-batch mean target. This preserves the expected label mean and reduces batch-to-batch variance; no shrinkage is applied in validation or testing. A small ablation showed that ($\alpha = 0.01$) lowered error slightly without altering conclusions, so we retain it.

Performance evaluation adhered strictly to the LOGO scheme, where the model predicts fracture loads for each unseen test configuration (31 in total) individually. Reproducibility and robustness are ensured by consistently setting model initialization seeds and data shuffling generators throughout all stages. Training dynamics are monitored using cosine annealing learning rate scheduling [44], facilitating stable convergence to optimal model parameters without manual intervention. Cosine annealing adjusts the learning rate as:

$$\eta_t = \eta_{\min} + \frac{1}{2} (\eta_{\max} - \eta_{\min}) \left(1 + \cos\left(\frac{t}{T}\pi\right) \right) \quad (13)$$

where η_t is the learning rate at epoch t , $\eta_{\min}$ and $\eta_{\max}$ are the minimum and maximum learning rates, and T is the total number of epochs. We decay η from its initial value to 1×10^{-6} using $T_{\max} = 50$ in the search phase and $T_{\max} = 20$ during the final 20-epoch fine-tune.

A schematic view of the developed CNN structure is shown in Fig. 8.

To justify the selection of DenseNet-121 for the CNN architecture, a comparative evaluation was conducted against other well-known models, including ResNet-50 and EfficientNet-B0, during the model selection phase. These models were chosen as candidates due to their established performance in image-based tasks and availability of pre-trained weights on ImageNet, which facilitate robust feature extraction. However, ultimately, DenseNet-121 provided slightly more accurate predictions. The superior performance of DenseNet-121 might be attributed to its dense connectivity pattern, which promotes feature reuse and enhances the capture of complex strain gradients around notches. Note that because we use only the maximum principal strain component, the MLP operates on a low-dimensional tabular input (after the three-step per-fold feature selection, reduces to five features), which reduces model complexity and the data needed for effective learning and generalisation. Also, the CNN ingests a single-channel image but still observes a high-dimensional field, so explicit regularisation (dropout, weight decay, early stopping) and physics-admissible augmentation remain essential to prevent overfitting.

All experiments were conducted in Python, with PyTorch as the core deep-learning framework. The CNN uses torchvision's DenseNet-121 pretrained on ImageNet (we modified the first convolution to accept single-channel strain maps and replaced the classifier with a single-output regression head). Image ingestion and preprocessing for the CNN (notch in-painting and CLAHE) were implemented with OpenCV and Pillow. The MLP pipeline employed scikit-learn for data scaling,

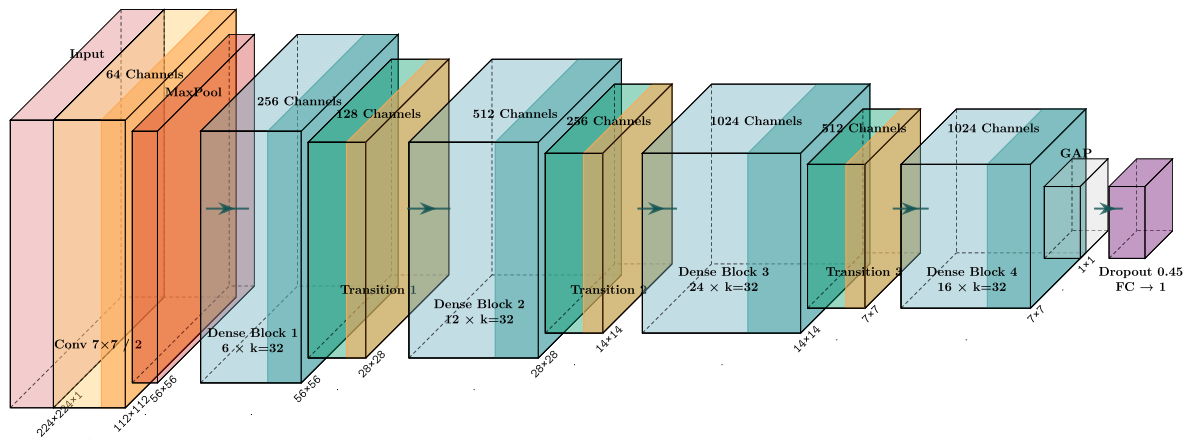


Fig. 8. Schematic representation of the DenseNet-121-based CNN used for fracture load prediction from 224×224 grayscale strain field images. The network includes four Dense Blocks interleaved with Transition Layers. The final layers comprise a Global Average Pooling (GAP) and a fully connected regression output. The number of output channels and spatial dimensions are indicated for each stage.

group-aware splitting, and evaluation metrics; Optuna was used for MLP hyperparameter optimization, and SHAP provided post-hoc feature attribution for tabular inputs. Figures were generated with Matplotlib. All scripts execute unchanged on CPU or CUDA-equipped GPUs when a compatible GPU is available.

4. Results and discussion

4.1. Prediction results

In this section, the fracture load predictions by the proposed framework for the notched laminated composites are presented. In order to compare the different cases and clearly demonstrate the accuracy of the framework, the analysis utilizes three key performance metrics: R^2 , MAE, and MAPE. MAE reports the average magnitude of residuals in the same units as the response variable. R^2 expresses the proportion of observed variance explained by the model, while MAPE normalises residuals by the true values. It is worth emphasising that predictions are obtained using Leave-One-Group-Out (LOGO), such that all replicates of the held-out configuration are unseen during training, feature selection, and scaling. Replicates are treated as independent samples; no averaging or smoothing is applied to preserve real-world variability. This deliberate, data-driven choice allows the model to learn from specimen-to-specimen differences and, in turn, improves training robustness. To understand the accuracy of predictions across different specimens and loading conditions, the predictions by the MLP model are visualized in Fig. 9, while the predictions by the CNN model are shown in Fig. 10. Replicate specimens within a group share the same marker style; each point denotes a single specimen. The solid black line represents the ideal prediction, serving as a benchmark for comparison. Furthermore, $\pm 20\%$ scatter bands, defined as 0.8 and 1.2 times the ideal prediction, are depicted using red dashed lines to provide a clearer understanding of the model's performance and its deviation from perfect agreement with the experimental data.

Considering Figs. 9 and 10, both models demonstrated excellent alignment with experimental data for fracture load predictions, with almost all predictions falling within the $\pm 20\%$ scatter bands, except for the specimen with a 15° U-notch. Focusing on Fig. 2 h), the failure loads for CP_15_UO and CP_60_UO are nearly identical, suggesting a possible issue with experimental testing which is also suggested by both models. The clustering of predictions for all different specimens with different notches, anisotropy levels, and off-axis angles around the ideal line suggests that the proposed framework effectively captures the mechanics driving fracture loads based on the principal strain field. In this case, the MLP achieved a slightly higher accuracy of $R^2 = 0.86$ compared

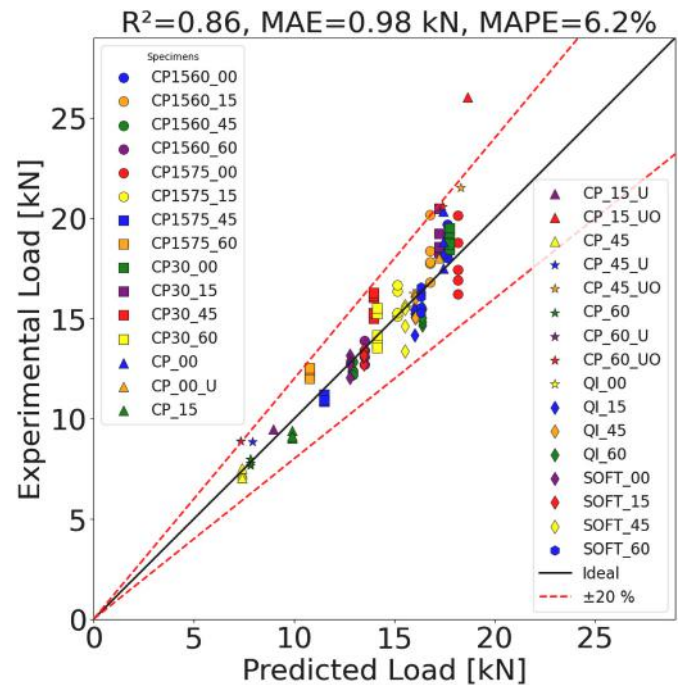


Fig. 9. Predictions of fracture load for different specimens using the maximum principal strain field as input features in the MLP model. The solid black line represents the ideal prediction, while the red dashed lines indicate the $\pm 20\%$ scatter bands. The experimental data were obtained from Ref. [22]. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

to the CNN's $R^2 = 0.82$ for predicting fracture load using DIC data. Although CNNs are generally superior for spatial data due to their ability to detect local patterns, several factors likely contributed to the MLP's better performance. The main reason may be the limited size of the dataset. The employed CNN model contains around eight million parameters and thus needs heavy data augmentation, transfer learning, or weight decay to prevent overfitting. An MLP with three hidden layers and carefully tuned regularisation has a far lower number of parameters, so it can reach its bias-variance sweet spot sooner. On the other hand, MLP was fed by data only in the action zone. This focus reduces noise and maintains model accuracy. However, the CNN takes in the whole strain field image—everything from the critical notch area to far-off regions. It is worth noting that the results of the CNN model would

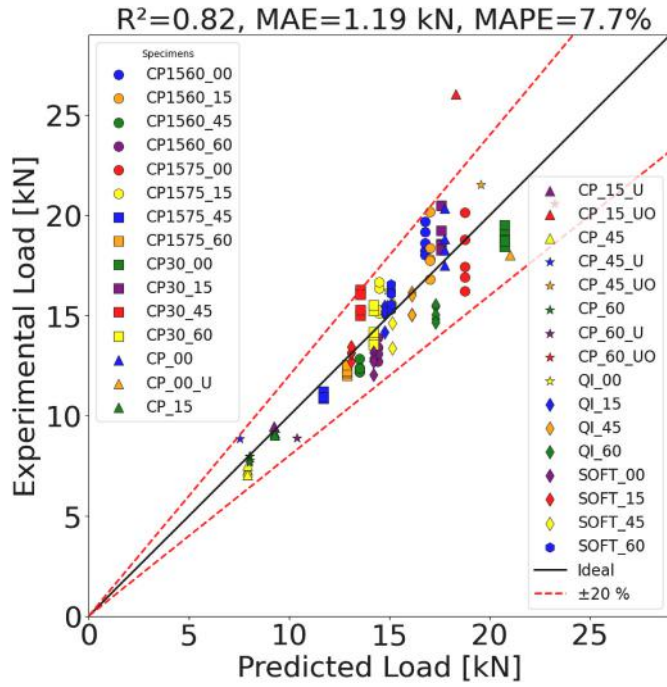


Fig. 10. Predictions of fracture load for different specimens using the maximum principal strain field as input features in the CNN model. The solid black line represents the ideal prediction, while the red dashed lines indicate the $\pm 20\%$ scatter bands. The experimental data were obtained from Ref. [22]. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

decrease to $R^2 = 0.57$ without augmentation (see section 3.3), but still most data are within $\pm 20\%$ scatter bands. Overall, considering the complexity of the problem, it can be argued that both models provide remarkable predictive accuracy in an efficient way, without the need for any simulations, or further analyses.

4.2. Effect of load level selection

The load level at which strain field data are collected is a key factor in predicting fracture loads in this framework. Using DIC, the strain field must be captured at a load that provides clear, reliable data without causing any specimen to fail. Ideally, the reference load should be high enough to provide a well-resolved, spatially coherent strain field, yet low enough to remain below a conservative lower bound of the failure load across all configurations. Achieving this balance is essential for the deep learning models, MLP and CNN, to perform well. The main analysis used a load level of 7 kN, selected to be high enough to be able to represent the strain field correctly and exceed DIC noise thresholds while remaining below the minimum observed fracture load (~ 8 kN). At 5 kN, the same model parameters were used to test the same conditions under increased noise and a different load level, with the results shown in Figs. 11 and 12. Note that the mean value of all failure loads is 14.7 kN, with 7 kN representing less than half of this value, and the maximum failure load is around 26 kN.

For the MLP model, Fig. 11 showed that predictions at 5 kN closely match those at 7 kN. Most data points stay near the ideal prediction line, which suggests the MLP handles lower load levels well, likely because it uses specific numerical features focused on strain near the notch, which remain clear even with some noise. The CNN model, however, performs differently. As seen in Fig. 12, its predictions at 5 kN scatter more widely around the ideal line compared to 7 kN. This drop in accuracy points to the CNN's sensitivity to noise in the strain field images at lower loads. Since the CNN processes full images (including strain patterns and minor details like speckle or lighting flaws), a lower load weakens the signal-

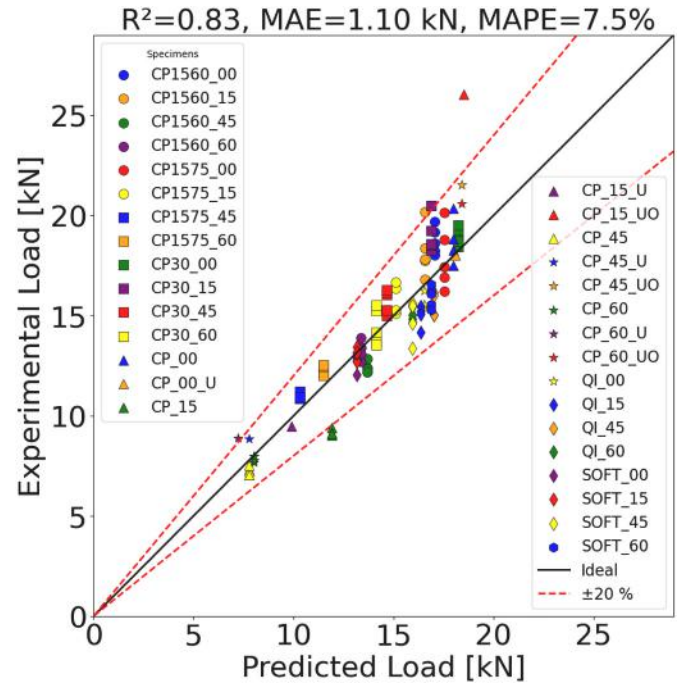


Fig. 11. Predictions of fracture load for different specimens using the maximum principal strain field as input features in the MLP model evaluated at a low load level of 5 kN. The solid black line represents the ideal prediction, while the red dashed lines indicate the $\pm 20\%$ scatter bands. The experimental data were obtained from Ref. [22]. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

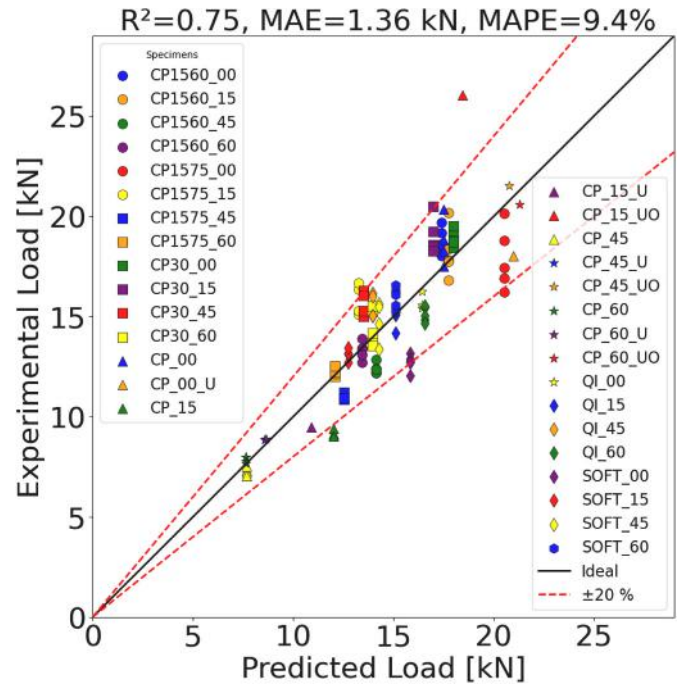


Fig. 12. Predictions of fracture load for different specimens using the maximum principal strain field as input features in the CNN model evaluated at a low load level of 5 kN. The solid black line represents the ideal prediction, while the red dashed lines indicate the $\pm 20\%$ scatter bands. The experimental data were obtained from Ref. [22]. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

to-noise ratio, making it harder for the model to focus on the most important features. It is important to note that, for the 5 kN evaluation, both models were run with exactly the same settings as those used at 7 kN. Specifically, re-tuning the CNN to the 5 kN input distribution could reasonably improve its predictive accuracy.

4.3. Effect of number of data points

For the MLP model, an important consideration in implementing the framework is selection of representative nodes for the field distribution. The baseline analysis used a grid of 200 nodes (10 along the shorter side and 20 along the longer side of the rectangle) to extract strain values. To test the model's adaptability, two additional configurations were analysed: Case 1 with a high-density grid of 800 nodes (20×40), and Case 2 with a low-density grid of 50 nodes (5×10). These results are presented in Fig. 13 a) and Fig. 13 b) respectively. By comparing these setups, the analysis highlights how parameter selection influences the prediction accuracy of the proposed approach.

Fig. 13 showed that for both high- and low-density configurations, predictions remained robust for most specimens. However, slightly larger deviations were observed for CP_15 and CP_45, which fall further outside the $\pm 20\%$ bands compared to the baseline. These findings demonstrate the flexibility of the feature selection algorithm in handling different node density levels, from 50 to 800 nodes, without significant loss of performance. The feature selection process plays an important role by identifying the most informative nodes and enhancing the efficiency of the framework. For practical applications, a moderate node density considering a typical finite element mesh for the problem (e.g., 200 nodes) offers an optimal balance between capturing essential strain data and minimizing computational cost.

4.4. Effect of region of interest positioning

As mentioned, the MLP model relies on strain values extracted from a rectangular region positioned close to the notch. However, placing the sampling zone too close to the notch boundary risks unreliable

measurements due to DIC limitations, while positioning it too far may miss critical strain gradients. This section investigates how the distance of the rectangular region from the notch tip as well as its size affects the MLP's predictive performance, offering insights into optimizing this critical parameter for robust and practical applications.

The baseline configuration positioned the nearest edge of the sampling zone 0.1 mm from the notch tip and 1 mm from the specimen edge, yielding a rectangle with dimensions of 6.4×12.8 mm. To investigate the impact of proximity, two alternative configurations are tested: Case 1, with the rectangle starting 0.5 mm from the notch tip (sampling zone size: 6×12 mm), and Case 2, with the rectangle starting 0.3 mm from the notch tip (sampling zone size: 6.2×12.4 mm). Note that in the baseline configuration, the close proximity increases the likelihood of missing strain data due to DIC inaccuracies. To reduce this, a simple nearest-neighbour interpolation algorithm can be used to estimate missing values. Results for these configurations are shown in Fig. 14, with subfigures a) and b) representing the 0.5 mm and 0.3 mm distances, respectively.

In Fig. 14 a), the 0.5 mm configuration produced predictions that closely follow the ideal line for most specimens, comparable to the baseline's performance (0.1 mm). Data points cluster tightly within the $\pm 20\%$ scatter bands, indicating that the MLP effectively captures essential strain patterns even at a slightly greater distance and smaller rectangle. Fig. 14 b) presents the closer distance to the notch tip, where predictions remain robust, with most points within the $\pm 20\%$ bands. The MLP's stable performance despite this challenge highlights the strength of the feature selection pipeline—utilizing Mutual Information, Lasso regression, and SHAP-based ranking—to prioritize high-impact strain values.

4.5. Effect of DIC processing parameters

In the employed DIC software (VIC-2D), two key processing parameters—subset (correlation window) size and step (grid spacing), govern the trade-off between precision, spatial smoothing, and sampling density, and therefore can influence the prediction quality. Larger subsets

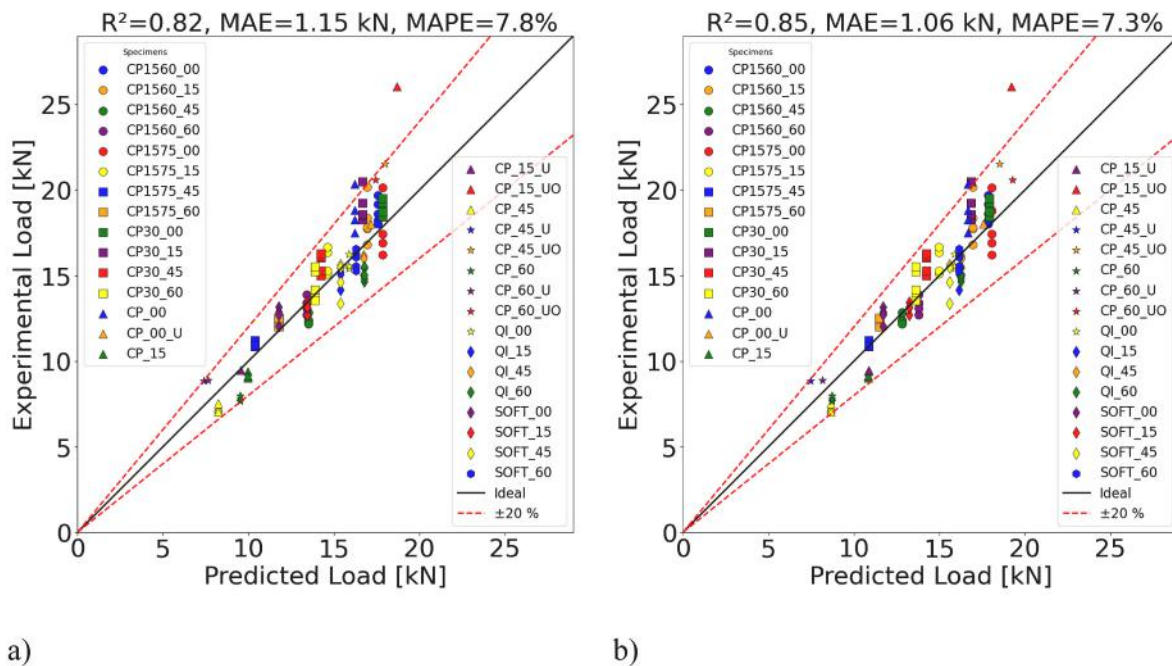


Fig. 13. Predictions of fracture load for different specimens using the maximum principal strain field as input features. The results are calculated based on discretizing the region of interest using a) 800 nodes, b) 50 nodes. The solid black line represents the ideal prediction, while the red dashed lines indicate $\pm 20\%$ scatter bands. The experimental data were obtained from Ref. [22]. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

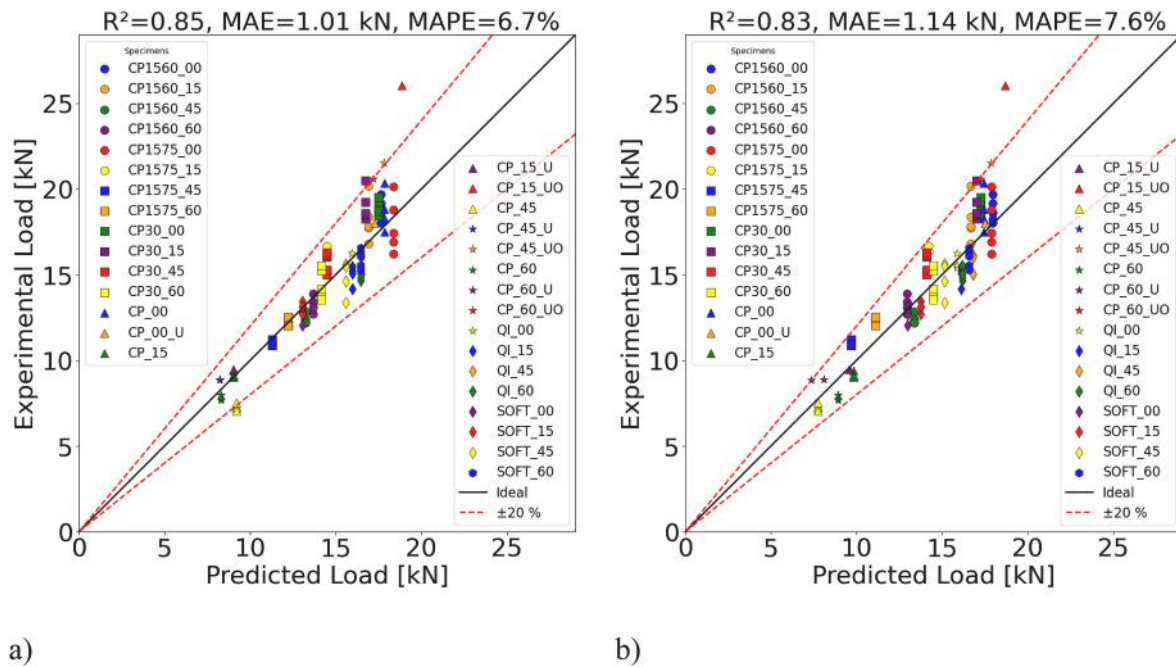


Fig. 14. Predictions of fracture load for different specimens using the maximum principal strain field as input features in the MLP model. The results are calculated for: a) 0.5 mm, and b) 0.3 mm distance between the region of interest and the notch tip. The solid black line represents the ideal prediction, while the red dashed lines indicate $\pm 20\%$ scatter bands. The experimental data were obtained from Ref. [22]. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

reduce random noise but increase the effective virtual strain-gauge (VSG) length (the effective spatial averaging length used to compute strain from the displacement field), which smooths strain gradients near the notches. The step controls spatial sampling density and overlap between adjacent subsets; reducing the step increases the number of evaluated points but, by itself, does not lower the per-point measurement uncertainty; per-point uncertainty is primarily set by subset size. A

practical and widely used choice is to keep $\text{step} \approx (1/3 - 1/2) \times \text{subset}$ to ensure smooth, well-sampled fields without redundant computations [45,46]. As stated, all primary analyses used a 15-pixel subset with a 7-pixel step.

To examine sensitivity, we reprocessed the dataset with a 21-pixel subset and an 8-pixel step, which increases smoothing while maintaining a typical overlap ratio. This combination was selected to test

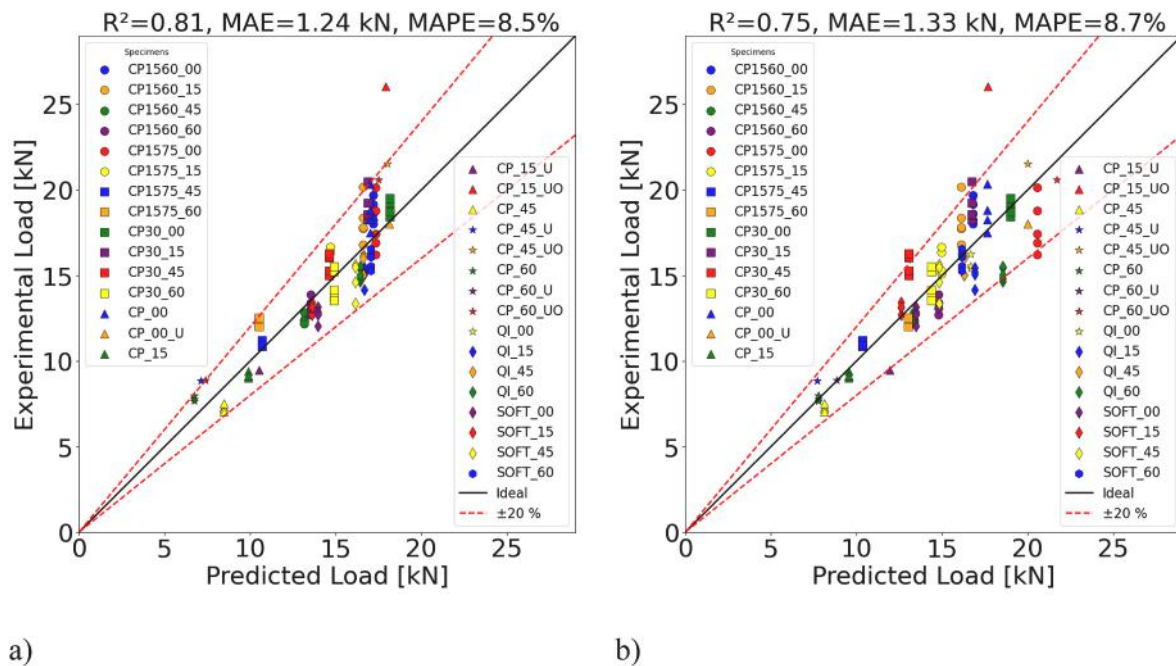


Fig. 15. Predictions of fracture load for different specimens using the maximum principal strain field as input: (a) MLP model, (b) CNN model. The results are calculated using Subset = 21 px and Step = 8 px. The solid black line represents the ideal prediction, while the red dashed lines indicate $\pm 20\%$ scatter bands. The experimental data were obtained from Ref. [22]. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

whether additional smoothing alters the strain features learned by the models. On the lower side, a 13-pixel subset was the smallest viable value (given the available image resolution and speckle sampling); using subsets <13 px led to local loss of correlation (small regions where the strain field could not be computed), and those runs were therefore not considered. For several checked cases, the strain fields obtained with 13 px and 15 px subsets differed only slightly and did not change the model metrics, supporting the robustness of the baseline settings. The prediction results for 21 px/8 px are presented in Fig. 15.

Using the baseline (subset = 15 px, step = 7 px), the models achieved $R^2 = 0.86$ (MLP) and $R^2 = 0.82$ (CNN). When increasing smoothing via (subset = 21 px, step = 8 px), performance decreased to $R^2 = 0.81$ (MLP) and $R^2 = 0.75$ (CNN)—absolute drops of 0.05 and 0.07, corresponding to $\approx 5.8\%$ and $\approx 8.5\%$ relative reductions, respectively. As mentioned, enlarging the subset increases the effective virtual strain-gauge length; it reduces strain noise but also smooths the near-notch strain gradients that carry predictive information for failure-load estimation. The small step increase (7 $\rightarrow$ 8 px) slightly reduces sampling density without lowering per-point noise; therefore, the observed degradation is dominated by the larger subset. The CNN is more sensitive to this smoothing because its convolutional filters rely on high-frequency spatial content around the notch; the MLP, which aggregates a limited set of node values within a fixed physical region of interest, shows a smaller but still noticeable reduction. Together with our lower-bound trials (subsets <13 px led to local correlation loss in high-gradient/low-texture zones), these results indicate that the baseline 15/7 setting provides a robust bias–variance trade-off. Importantly, despite this degradation, both settings still deliver high predictive performance ($R^2 \approx 0.81$ for MLP; 0.75 for CNN), demonstrating the framework’s robustness to moderate DIC-processing variations.

To enhance the discussion, the effect of DIC camera settings on prediction accuracy can be summarised as follows. Camera and optics choices in single-camera 2D-DIC determine the object-space sampling ($\text{mm}\cdot\text{px}^{-1}$) and, together with the processing parameters, define the virtual strain-gauge length of the reported strain field; both directly influence the information available to the ML models. A wider field-of-view (FOV) at fixed sensor resolution coarsens the $\text{mm}\cdot\text{px}^{-1}$ scale and, unless the subset (correlation window) and step (grid spacing) are re-tuned, effectively enlarges the VSG, which can weaken notch-proximal gradients that the models rely on. The recommended practice is to select magnification/FOV so that speckle features are $\sim 3\text{--}5$ px in diameter and each subset contains $\geq 3 \times 3$ speckles [45,46]. Lens selection and calibration are equally critical: standard ‘entocentric’ photographic lenses are sensitive to out-of-plane motion and perspective, and uncorrected radial/tangential distortion introduces spatially structured bias. It is recommended to use a calibrated lens–camera system, i.e., perform full geometric calibration and apply the correction so that reprojection (distortion) residuals are sub-pixel (typically $\leq 0.1\text{--}0.2$ px). Maintain adequate depth of field so the entire region of interest—and any small out-of-plane motion during loading—remains sharply focused, and lock focus and zoom after calibration to preserve the calibrated geometry [45]. Moreover, changes in brightness over time introduce random error into the measured displacements. Using zero-mean normalised cross-correlation (ZNCC) or zero-mean normalised sum of squared differences (ZNSSD) makes the correlation effectively insensitive to linear intensity changes (offset and gain) [47]. Finally, the imaged speckle field should be sharp, high-contrast, and well-sampled so that each subset contains a sufficiently distinctive intensity pattern; otherwise, correlation precision degrades and a larger subset is required, reducing spatial resolution and effectively increasing the virtual strain-gauge window. Low-contrast or undersampled/blurred speckles also promote interpolation-related errors (e.g., pixel-locking) [45–47]. Collectively, these practices minimise camera-induced domain shift and stabilise the input distribution for both the MLP (node samples) and the CNN (full fields), thereby supporting robust and accurate predictions.

4.6. Discussion on feature selection algorithm

In this study, the MLP predicts fracture loads based on strain values sampled from a grid of nodes near the notch. However, not all nodes contribute equally to the prediction; including too many nodes can introduce redundant information, increase computational cost, and potentially reduce model accuracy. Feature selection solves this by identifying the most critical nodes, ensuring the MLP focuses on the strain values that matter most. Fig. 16 provides insight into how features interact and contribute to the model’s predictive performance. It should be noted that the results in this section are based on features from all groups, not one-leave-out algorithm. Also, in Fig. 16 b), the origin is at the centre of notches.

Fig. 16a) shows the SHAP summary for the Random Forest model used in the feature-selection stage. Each point corresponds to one specimen; the abscissa is the SHAP value (impact on the predicted fracture load relative to the model baseline), and colour encodes the scaled principal strain value at that node (blue = low, red = high). Positive SHAP values mean that the node’s strain increases the predicted fracture load for that specimen; negative values mean the opposite. As can be seen, the two highest-ranked nodes are located in the near-notch high-gradient zone, consistent with notch-controlled behaviour. For the most influential node (−1.68, 4.02), lower principal strain at the reference load tends to push the prediction upward—consistent with a locally stiffer response being associated with higher capacity—whereas a far-field node (e.g., 6.40, 8.29) shows the opposite trend, indicating that more compliant global responses are associated with lower predicted capacity.

Following the analysis of the framework’s behaviour under varying data selection strategies, the study acknowledges inherent limitations. The small dataset size limits both models, particularly the CNN’s ability to fully exploit its spatial learning capabilities. Also, CNNs are computationally more expensive; this is the reason why further parameter tuning or analysis of the parameters’ effect was not performed for the CNN model.

It can be argued that the proposed framework is material-agnostic in principle because it learns a mapping from strain field to failure load without embedding a material model (no constitutive law, failure criterion, or finite-element simulation is employed at inference). It can therefore be extended to other composite families such as glass-fibre (GFRP), aramid/Kevlar (AFRP), and fibre–metal laminates (FMLs, e.g., GLARE). It is worth emphasising that, in practice, this mapping must be calibrated for each material system. The following remarks summarise widely reported trends across these materials. Under uniaxial tension, CFRP generally fails in a fibre-dominated manner, with matrix cracking, ply splitting and interlaminar delamination that develop and co-evolve before fibre rupture [48]. In GFRP, the lower longitudinal modulus and lower in-plane shear modulus (higher shear compliance) typically produce more spatially diffuse notch-root strain fields, with damage dominated by matrix cracking, fibre–matrix debonding and shear-driven failure [49]. Aramid/Kevlar laminates exhibit high tensile toughness but comparatively poor compressive strength due to microbuckling/kink-banding, and characteristically show fibrillation and fibre pull-out in tension [50]. In fibre–metal laminates (e.g., GLARE), metal-layer plasticity interacts with interfacial debonding and interlaminar delamination in the composite plies, yielding hybrid strain fields that reflect both plastic zones in the metal and interlaminar shear in the composite [51]. It is worth emphasising that the dataset used here already spans a broad range of failure behaviours; moreover, the same framework has been applied by the author to predict failure in PLA, PMMA and steel [19,21], indicating the applicability of the method across material classes.

Moreover, computational efficiency is a critical advantage of the proposed framework for in-situ failure assessment. DIC processing in VIC-2D required from seconds to a few minutes per test on a standard workstation (Core i7-class CPU, 16 GB RAM; 12 GB VRAM). For the

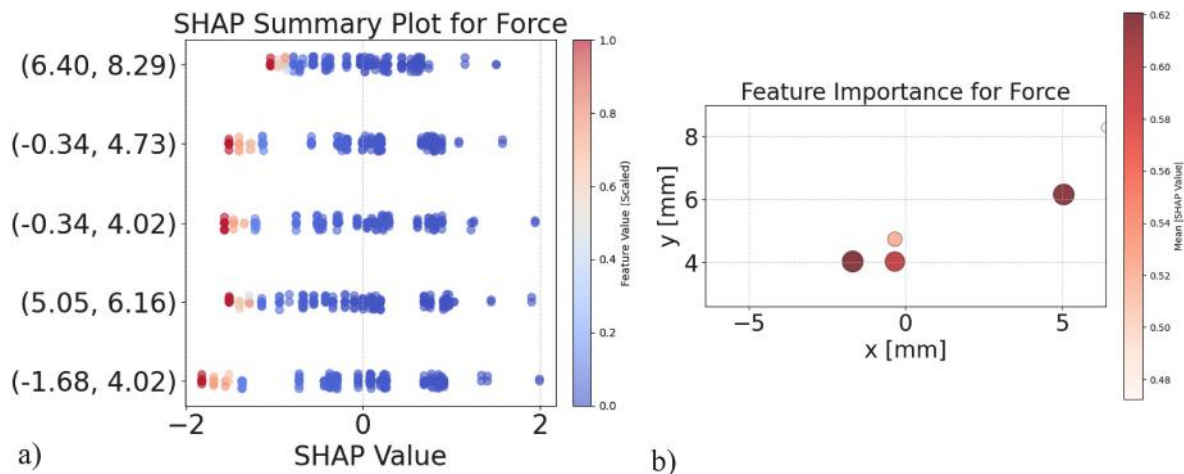


Fig. 16. a) SHAP summary plot for the top 5 features presenting them in Cartesian coordinates. b) Feature importance based on SHAP value of top 5 features considering their locations. For the purpose of feature interpretation, this analysis was performed on a model trained on the full dataset.

learning stage per held-out group, final model training took ~ 0.8 s for the MLP and ~ 26 s for the CNN (DenseNet-121) on the same workstation. Post-training, inference is near-real-time: the MLP evaluates a single sample in ~ 0.004 s, and the CNN processes a 224×224 strain image in ~ 0.01 – 0.07 s on GPU (a few tenths of a second on CPU for small batches). Compared to FE-based progressive damage simulations (hours per case), analytical models like Whitney-Nuismer (seconds but several empirical calibrations and possibly lower accuracy), our simulation-free, localized DIC approach is orders of magnitude faster, ideal for in-situ failure assessment in complex structures. A key efficiency benefit is that the workflow needs only the local strain field in a small region around the notch. It does not require building a full finite-element model of the entire structure or reconstructing detailed boundary conditions and load paths (e.g., grips, contacts). By reducing the problem to the notch-centred region of interest that governs failure, the approach removes the dominant cost drivers in FE pipelines, thereby enabling fast predictions from a single DIC frame.

Looking forward, several avenues can enhance this framework. Expanding the dataset with additional test cases, particularly for U-notch configurations and anisotropic layups, could improve the CNN's performance and generalization. Extending the framework to other loading conditions, such as compression or fatigue [52], would broaden its applicability in composite design and testing. On the other hand, the proposed framework estimates fracture loads from single-frame DIC maximum-principal-strain fields without explicit layup inputs, which supports broad applicability across configurations for a known material system. Another extension is to predict the lamina-level failure sequence. To infer it, two additions are required (i) encode ply architecture (e.g., orientations/stacking): so the model can condition on laminate design, and (ii) use temporal DIC (multi-frame strain maps over load) to capture the progression of failure. Sequence-aware models—such as temporal CNNs, ConvLSTMs or Transformers—can exploit these time series, while theory-guided ML can regularise learning and multi-task outputs for onset loads and event ordering. Related strategies have been validated for progressive damage identification in composites [53–55].

From a practical point of view, it can be argued that since the framework was able to successfully predict failure in notched laminated composites for a variety of anisotropy levels, loading conditions, and notch geometries using DIC, the previous problems solved by the same framework, like finite fatigue life of notched components across different materials [19], as well as fracture load and fracture angle predictions in mixed mode loading conditions [21] can be tackled by the framework using DIC and in-situ online approaches.

5. Conclusions

This study presented a robust deep learning framework for in-situ prediction of fracture loads in notched laminated composites using only maximum principal strain field data acquired by DIC. The study employed two different deep learning approaches: an MLP that processes numerical strain values sampled from a grid of nodes near the notch, and a CNN that operates directly on full-field strain images extracted from DIC measurements to capture the complex interplay of notch geometry, fibre orientation, and layup. The experimental dataset used was notably diverse, comprising 116 tests across 31 distinct configurations, including six laminate layups, four off-axis angles, and two notch geometries. This encompassed a wide range of anisotropy levels, both brittle and ductile behaviours, and diverse failure modes such as fibre-dominated fracture and delamination. The results demonstrated exceptional predictive accuracy, with the MLP achieving an $R^2 = 0.86$ and the CNN $R^2 = 0.82$, both exhibiting tight clustering of predictions within $\pm 20\%$ scatter bands. These results confirmed the framework's ability to model the intricate mechanics driving fracture in notched laminated composites.

The superior performance of the MLP highlighted the effectiveness of its targeted feature selection process, which combines Mutual Information, Lasso regression, and SHAP-based ranking to prioritize the most valuable information around the notch tip. This targeted, numerically-driven approach both enhanced accuracy and maintained computational efficiency, making the framework especially suited for real-time or resource-limited monitoring. In contrast, the CNN utilizes the detailed information in full-field DIC images to identify complex patterns, though its predictions are more sensitive to the selected load level. Sensitivity analyses further established that the MLP's predictive accuracy remains stable across a wide range of numbers of input nodes and locations of region of interest, and that it is robust to moderate variation in DIC load levels, which is an important practical advantage. Furthermore, sensitivity analyses of the DIC processing parameters indicate that increasing the effective virtual strain-gauge length—i.e., applying stronger smoothing by using larger correlation subsets—reduces the sharpness of near-notch strain gradients and produces a modest decrease in predictive accuracy. Nevertheless, both models maintained high overall performance, demonstrating robustness to practical levels of smoothing. A key strength of this framework lies in its simplicity and practicality, as it requires no additional finite element analysis—which avoids full-geometry modelling and detailed boundary condition specification—and allows for the direct use of full-field strain measurements. Once trained, both the MLP and CNN models demonstrate low computational times while relying solely on near-field strain distributions

around the notch. This streamlined approach is particularly valuable for in-situ structural health monitoring.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix

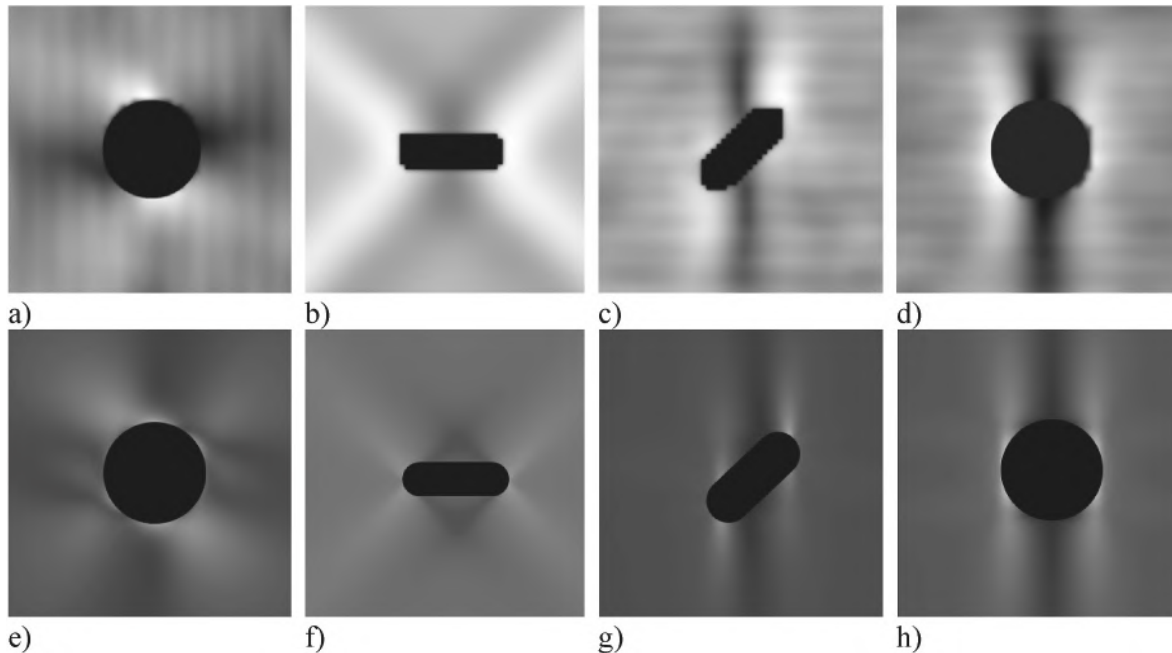


Fig. A.1. Maximum principal strain distributions for four composite specimens, showing experimental DIC measurements alongside finite-element validations. (a–d) DIC-measured strain fields for (a) CP1560_15, (b) CP_45_U, (c) CP_45_UO and (d) CP_00. (e–h) Corresponding FE predicted strain fields for (e) CP1560_15, (f) CP_45_U, (g) CP_45_UO and (h) CP_00. The FE contours are provided solely as a benchmark against the experimental DIC fields; no FE data were employed at any stage of the machine-learning model's training, validation, or inference.

Data availability

Data will be made available on request.

References

- [1] Barbero EJ. Introduction to composite materials design. CRC press; 2017.
- [2] Jones RM. Mechanics of composite materials. CRC press; 2018.
- [3] Whitney JM, Nuismer RJ. Stress fracture criteria for laminated composites containing stress concentrations. *J Compos Mater* 1974;8:253–65.
- [4] Pipes RB, Wetherhold RC, Gillespie Jr JW. Notched strength of composite materials. *J Compos Mater* 1979;13:148–60.
- [5] Sutton MA, Wolters WJ, Peters WH, Ranson WF, McNeill SR. Determination of displacements using an improved digital correlation method. *Image Vis Comput* 1983;1:133–9.
- [6] Esmaili A, Mohammadi B, Yousefi A. Fracture parameters and crack initiation assessment employing mixed-mode I/II fracture criterion in laminate composites using digital image correlation method. *Theor Appl Fract Mech* 2023;127:104095.
- [7] Wisnom MR, Hallett SR. The role of delamination in strength, failure mechanism and hole size effect in open hole tensile tests on quasi-isotropic laminates. *Compos Part A Appl Sci Manuf* 2009;40:335–42.
- [8] Ambu R, Aymerich F, Bertolino F. Investigation of the effect of damage on the strength of notched composite laminates by digital image correlation. *J Strain Anal Eng Des* 2005;40:451–61.
- [9] Caminero MA, Lopez-Pedrosa M, Pinna C, Soutis C. Damage monitoring and analysis of composite laminates with an open hole and adhesively bonded repairs using digital image correlation. *Composites Part B* 2013;53:76–91.
- [10] Khalid S, Yazdani MH, Azad MM, Elahi MU, Raouf I, Kim HS. Advancements in physics-informed neural networks for laminated composites: a comprehensive review. *Mathematics* 2025;13.
- [11] Azeem OAI, Pinho ST. A machine learning enhanced characteristic length method for failure prediction of open hole tension composites. *Compos Part C Open Access* 2024;15:100524.
- [12] Balducci GT, Chen B, Möller M, Gerritsma M, De Breuker R. Predicting open-hole laminates failure using support vector machines with classical and quantum kernels. *ArXiv Prepr ArXiv2405* 2024:02903.
- [13] Li Z, Qin H, Wang Q, Jia L, Zhang G, Li Y, Liu Y. Prediction of allowable compression load for notched composite laminates combining FEA simulation and machine learning. *Compos Struct* 2024;340:118188.
- [14] Li Y, Qin H, Tan VBC, Jia L, Liu Y. A deep transfer learning approach to construct the allowable load space of notched composite laminates. *Compos Sci Technol* 2024;247:110432.
- [15] Nastos C, Kominos P, Zarouchas D. Non-destructive strength prediction of composite laminates utilizing deep learning and the stochastic finite element methods. *Compos Struct* 2023;311:116815.
- [16] Wang Y, Luo Q, Xie H, Li Q, Sun G. Digital image correlation (DIC) based damage detection for CFRP laminates by using machine learning based image semantic segmentation. *Int J Mech Sci* 2022;230:107529.
- [17] Chen L-C, Papandreou G, Schroff F, Adam H. Rethinking atrous convolution for semantic image segmentation. <https://arxiv.org/abs/1706.05587>; 2017.
- [18] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *Proc. IEEE conf. Comput. Vis. Pattern recognit.*; 2016.
- [19] Mirzaei AM. Stress, strain, or energy? Which one is superior predictor of fatigue life in notched components? A novel machine learning-based framework. *Eng Fract Mech* 2024;309:110401.

- [20] Mirzaei AM, Mirzaei AH, Shokrieh MM, Sapora A, Cornetti P. Fatigue life assessment of notched laminated composites: experiments and modelling by finite fracture mechanics. *Compos Sci Technol* 2024;246:110376.
- [21] Mirzaei AM. Stress, strain, or displacement? A novel machine learning based framework to predict mixed mode I/II fracture load and initiation angle. *Eng Fract Mech* 2025;111349.
- [22] Mitrou A, Arteiro A, Reinoso J, Camanho PP. Effect of the level of anisotropy on the macroscopic failure of notched thin-ply laminates. *Compos Struct* 2024;348:118407.
- [23] Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *Proc IEEE* 1998;86:2278–324.
- [24] LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, Jackel LD. Backpropagation applied to handwritten zip code recognition. *Neural Comput* 1989;1:541–51.
- [25] Hirra I, Ahmad M, Hussain A, Ashraf MU, Saeed IA, Qadri SF, Alghamdi AM, Alfakheh AS. Breast cancer classification from histopathological images using patch-based deep learning modeling. *IEEE Access* 2021;9:24273–87.
- [26] Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol Rev* 1958;65:386–408.
- [27] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature* 1986;323:533–6.
- [28] Shannon CE. A mathematical theory of communication. *Bell Syst Tech J* 1948;27:379–423.
- [29] Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc Ser B* 1996;58:267–88.
- [30] Ho TK. Random decision forests. In: *Proc. 3rd int. Conf. Doc. Anal. Recognit.*, vol. 1; 1995. p. 278–82.
- [31] Breiman L. Random forests. *Mach Learn* 2001;45:5–32.
- [32] Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: *Proc. 22nd ACM SIGKDD int. Conf. Knowl. Discov. Data min.* New York, NY, USA: Association for Computing Machinery; 2016. p. 785–94.
- [33] Lundberg S. A unified approach to interpreting model predictions. *ArXiv Prepr ArXiv1705.2017:07874*.
- [34] Stone M. Cross-validatory choice and assessment of statistical predictions. *J R Stat Soc Ser B* 1974;36:111–33.
- [35] Akiba T, Sano S, Yanase T, Ohta T, Koyama M. Optuna: a next-generation hyperparameter optimization framework. In: *Proc. 25th ACM SIGKDD int. Conf. Knowl. Discov. Data Min.*; 2019. p. 2623–31.
- [36] Al-Kababji A, Bensaali F, Dakua SP. Scheduling techniques for liver segmentation: Reducelronplateau vs onecyclelr. In: *Int. Conf. Intell. Syst. Pattern recognit*; 2022. p. 204–12.
- [37] Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: *2017 IEEE conf. Comput. Vis. Pattern recognit.*; 2017. p. 2261–9.
- [38] Bertalmio M, Sapiro G, Caselles V, Ballester C. Image inpainting. In: *Proc. 27th annu. Conf. Comput. Graph. Interact. Tech.* USA: ACM Press/Addison-Wesley Publishing Co.; 2000. p. 417–24.
- [39] Pizer SM, Johnston RE, Ericksen JP, Yankaskas BC, Muller KE. Contrast-limited adaptive histogram equalization: speed and effectiveness. In: [1990] *proc. First conf. Vis. Biomed. Comput.*; 1990. p. 337–45.
- [40] Otsu N. A threshold selection method from gray-level histograms. *IEEE Trans Syst Man Cybern* 1979;9:62–6.
- [41] Serra JP, Serra J. *Image analysis and mathematical morphology*. Academic Press; 1982. <https://books.google.co.uk/books?id=BpdTAAAYAAJ>.
- [42] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. In: Pereira F, Burges CJ, Bottou L, Weinberger KQ, editors. *Adv. Neural inf. Process. Syst.* Curran Associates, Inc.; 2012.
- [43] Girshick R. Fast R-CNN. In: *2015 IEEE int. Conf. Comput. Vis.*; 2015. p. 1440–8.
- [44] Loshchilov I, Hutter F. SGDR: stochastic gradient descent with warm restarts. <https://arxiv.org/abs/1608.03983>; 2017.
- [45] Schreier H, Orteu J-J, Sutton MA. *Image Correlation for Shape, Motion and Deformation Measurements: Basic Concepts, Theory and Applications*. 2009.
- [46] International Digital Image Correlation Society. *A good practices guide for digital image correlation*. 2025.
- [47] Pan B, Qian K, Xie H, Asundi A. Two-dimensional digital image correlation for in-plane displacement and strain measurement: a review. *Meas Sci Technol* 2009;20:62001.
- [48] Liu Y, Xie J, Chen X, Zhang J, Wan M, Sun Z, Yang C. Characterization of progressive damage behaviour and failure mechanism of carbon fiber reinforced composite laminates. *Sci Rep* 2025;15:13791.
- [49] Kulpa M, Wiater A, Rajchel M, Siwowski T. Comparison of material properties of multilayered laminates determined by testing and micromechanics. *Mater (Basel Switzerland)* 2021;14.
- [50] Dai Y, Yuan Y, Luo L, Liu X. A facile strategy for fabricating aramid fiber with simultaneously high compressive strength and high interfacial shear strength through cross-linking promoted by oxygen. *Compos Part A Appl Sci Manuf* 2018;113:233–41.
- [51] Khan SU, Alderliesten RC, Benedictus R. Delamination in fiber metal laminates (GLARE) during fatigue crack growth under variable amplitude loading. *Int J Fatig* 2011;33:1292–303.
- [52] Mirzaei AH, Haghi P. In-situ fatigue prognosis in laminated composites: a machine learning approach using self-temperature rise data. *Composites Part B* 2025;301:112537.
- [53] Zobeiry N, Reiner J, Vaziri R. Theory-guided machine learning for damage characterization of composites. *Compos Struct* 2020;246:112407.
- [54] Zhang C, Li Y, Jiang B, Wang R, Liu Y, Jia L. Mechanical properties prediction of composite laminate with FEA and machine learning coupled method. *Compos Struct* 2022;299:116086.
- [55] Loh JYY, Yeoh KM, Raju K, Pham VNH, Tan VBC, Tay TE. A review of machine learning for progressive damage modelling of fiber-reinforced composites. *Appl Compos Mater* 2024;31:1795–832.



DSPACE

<https://dspace.org/>

Towards in-situ failure assessment: deep learning on DIC results for laminated composites

Mirzaei, Amir Mohammad

2026-03-01

Elsevier

<https://doi.org/10.1016/j.compositesb.2025.113241>

Mirzaei AM. (2025) Towards in-situ failure assessment: deep learning on DIC results for laminated composites. Composites Part B: Engineering, Volume 312, March 2026, Article number 113241