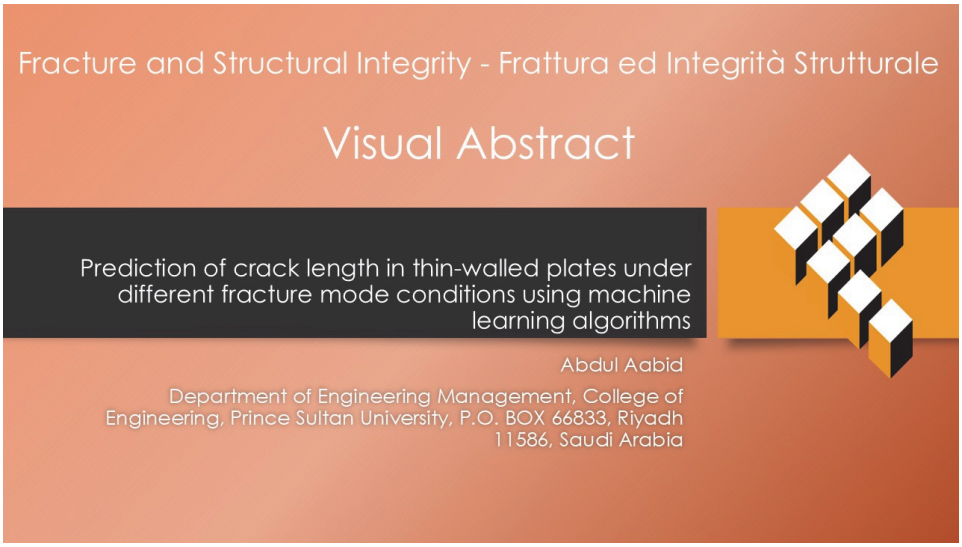
Prediction of crack length in thin-walled plates under different fracture mode conditions using machine learning algorithms

Abdul Aabid
Department of Engineering Management, College of Engineering, Prince Sultan University, P.O. BOX 66833, Riyadh 11586, Saudi Arabia
aaabid@psu.edu.sa, https://orcid.org/0000-0002-4355-9803



Citation: Aabid, A., Prediction of crack length in thin-walled plates under different fracture mode conditions using machine learning algorithms, Fracture and Structural Integrity, 75 (2026) 55-75.

Received: 29.05.2025
Accepted: 11.10.2025
Published: 16.10.2025
Issue: 01.2026

KEYWORDS. SIF, Fracture modes, Thin-walled plates, Machine learning, damages/cracks.

INTRODUCTION

Material structures can fail due to mechanical and thermal loads. The crack can be initiated through the thickness or only on the surface of the structures. When it comes to the thin structures, the crack can initiate through the thickness, while in the thick structures, the surface cracks can occur in many cases. Early researchers have conducted a vast number of studies on damaged structures considering the fundamentals of fracture mechanics (FM). On the other hand, the researcher used a different approach to show the results of damaged structures [1]. This can be characterized by different scenarios for thick and thin structures. The damage or crack can be initiated in three modes that have Mode I, II, and III. These modes of propagation can be predicted through the fracture parameters such as stress intensity, stress concentration, J-integral evaluation, or fracture toughness. The FM has been classified into two major

categories: Linear Elastic Fracture Mechanics (LEFM) and Elastic-Plastic Fracture Mechanics (EPFM), and this research considered the thin plate under three modes within the LEFM study.

Early studies demonstrated the Stress Intensity Factor (SIF) for the semi-elliptical surface cracks in tension plates using the finite element (FE) method, then compared with the Society of Experimental Stress Analysis, and also compared to photoelastic K-measurements [2]. Semi-elliptical surface cracks were found important in early research, and the researcher explored with another approach SIF and weight functions for longitudinal semi-elliptical surface cracks in thin pipes and demonstrated using the three-dimensional FE method [3].

Cracks have always shown an important study as the focus on safety; therefore, the study of SIF of an arbitrarily located circumferential crack in a thin-walled cylinder with axisymmetrically loaded ends has been analyzed through the numerical and analytical methods, and results showed that the SIF increased when the cylinder length decreased and when the crack is located near the cylinder edge [4]. Also, a dynamic SIF for a longitudinal semi-elliptical crack in a thick-walled cylinder has been calculated [5]. The T-plate weld found a critical object that creates the crack; therefore, this was further explored by considering the effects of residual stress and determining the SIF [6].

Load always matters on the structure, hence buckling of cracked thin plates under tension or compression studies has been found in the literature [7], which defines the crack propagation. Using the FE method, SIF was calculated for semi-elliptical surface cracks in pressure vessels, focusing primarily on Mode I [8] and Mode II [9] propagation. Next, the SIF was computed using a hybrid method coupling with the point weight-function method in cracked plates under bending in static and fatigue loading conditions [10].

FE simulations have been conducted for Mode I and II conditions in sharp notched plates under in-plane shear and bending loads; the results confirmed the accuracy and efficiency of the FE method to compute the SIFs for notch problems [11]. A new FE formulation to simulate embedded strong discontinuity for the study of the fracture process in brittle or quasi-brittle solids was presented, and the crack path prediction under three-point bending loading conditions [12].

The SIF has also been calculated in the joint system of two dissimilar materials or shapes. During the joint process, the crack can be initiated, and Qian, [13] investigated such crack formations in which he simulated the model of V-joints circular hollow section with a rack-plate chord and determined the SIF. In some cases, researchers extracted the work of mixed-mode effect in a crack-front field in ductile thin plates that effects of T-stress using the FE method [14]. A SIF has been determined at the tips of multi-site cracks in the unstiffened aluminium panel using the extended FE method, and this predicts the crack propagation under tensile load [15]. The extended FE method is also found in determining the SIF in thin-walled plates [16].

In contrast, machine learning (ML) has gained popularity in many engineering disciplines and has been extensively used for optimizing, predicting, and analyzing results [17]. Related to the current work, a researcher explores the SIF results using an artificial neural network approach from acoustic emission measurements [18], and this provides an opportunity to explore the ML technique for SIF prediction. Same as the previous study, predicted the SIF in pavement cracking with the same neural networks method using semi-analytical FE results [19]. Also, a fatigue crack repair has been optimized with this ML approach for a cracked plate [20]. A prediction method through deep learning in coal rock for SIF Mode I crack [21] and mixed-mode SIF of a crack in composites has been investigated [22].

Recent work by Yao et al. [23] integrates SVR with FE simulations to accurately predict mode I–II crack growth paths. Jan et al. [24] transfer-learning study demonstrates Random Forest's strength in fatigue life prediction for welded steel joints. Omar et al. [25] benchmarked SVR, Random Forest, and Gradient Boosting Regressor, showing GBR's superior accuracy in crack propagation across composite, metal, and polymer materials. Zhao et al. [26] further show how deep-learning ensembles like attention-residual networks can improve fatigue life prediction, underscoring the relevance of modern boosting techniques.

While many existing studies apply ML to predict SIF values from acoustic emission, image data, or simulation outputs (e.g., [18], [19], [21]), the present study inverts the problem by using theoretical SIF values to predict crack length directly. Moreover, unlike works focusing solely on Mode I or using experimental AE signals, this research explores Modes I, II, and III, purely from analytical formulations, offering a generalizable, low-cost alternative for early-stage crack length estimation.

Based on the current literature studies, it has been found that damage/cracks in any kind of structure can occur while having any type of load. Therefore, this has also been important to investigate the crack or damage propagation, and hence this study explores the prediction of crack propagation in a thin-walled plate. As per the existing research, only an SIF has been predicted; there is no study proposed or investigated to predict the crack length using the SIF data. Current work uses an advanced ML algorithm that can predict the damage propagation. Initially, the data was obtained from Tada's analytical method, and further empirical relations were used to define Modes.

FRACTURE MECHANICS

Over the last five decades, FM studies have been extensively studied for thin-walled plates. In the early years, the research was done through mathematical modelling, and the fracture parameters were calculated. Current work focused on the fracture parameter known as the SIF for a thin-walled plate has been considered. This parameter is considered the most predominant parameter in FM studies. On the other hand, the crack in a thin-walled plate can be propagated in three scenarios called Mode I (opening), Mode II (shearing), and Mode III (tearing). Considering all these three modes, one of the most common modes is called Mode I, or called opening mode, because this mode has a high frequency in the propagation of cracks with external loads such as mechanical or environmental loads.
The determination of SIF for the current work has been calculated using Tada's analytical formula, which was derived for thin-walled plates. The dimension and material properties have been chosen for the application of aerospace engineering, and these prototype model evaluations have been extensively done in the existing work. However, this study extracted the fundamental information about the determination of SIF to be assessed in this current work.

Theoretical calculation of SIF

LEFM studies the crack propagation in materials assuming linear elasticity and small-scale yielding near the crack tip. According to LEFM, the stress field near the crack tip in polar coordinates (r, θ), where r is the distance from the crack tip and θ is the angle from the crack line, is expressed as:

$$\sigma_{i,j}(r,\theta)=\frac{K}{\sqrt{\pi a}}f_{i,j}(\theta)+higher\ order\ term \tag{1}$$

where, $f_{i,j}(\theta)$ is the angular distribution function (mode-dependent).
The Stress Intensity Factor K characterizes the intensity of the stress field near the crack tip and is the key parameter in LEFM. Each fracture mode has its own SIF. For Mode I, this can be determined for cracked plate dimensions under uniform uniaxial load, where 'σ' represents the applied load and 'a' represents crack length, and this is represented by:

$$K_I=\sigma\sqrt{\pi a} \tag{2}$$

The above expression (Eqn. 2) is for the infinite plate, whereas for a finite plate, the geometrical fracture needs to be included, and it is expressed as:

$$K_I=\sigma\sqrt{\pi a}\,Y\left(\frac{a}{W}\right) \tag{3}$$

where $Y\left(\frac{a}{W}\right)$ represents the geometrical factor, and this factor will depend on the crack location, such as the edge or the center cracked plate. As this focused on the edge-cracked plate, the geometrical factor can be written as:

$$Y\left(\frac{a}{W}\right)=1.122-0.231\left(\frac{a}{W}\right)+10.550\left(\frac{a}{W}\right)^2-21.710\left(\frac{a}{W}\right)^3+30.382\left(\frac{a}{W}\right)^4 \tag{4}$$

The relation of this geometrical factor has been expressed by Tada [27], and this has an accuracy better than 0.5% for $(a/W)\le 0.6$. Finally, the Mode I SIF for an edge-cracked plate known plate under uniform uniaxial load can be expressed as [27]:

$$K_I=\sigma\sqrt{\pi a}\,1.122-0.231\left(\frac{a}{W}\right)+10.550\left(\frac{a}{W}\right)^2-21.710\left(\frac{a}{W}\right)^3+30.382\left(\frac{a}{W}\right)^4 \tag{5}$$

Similarly, for Mode II and Mode III, the SIF expressions are illustrated below [28]:

$$K_{II} = \tau_s \sqrt{\pi a}\, 1.122 - 0.561\left(\frac{a}{W}\right) + 4.284\left(\frac{a}{W}\right)^2 - 12.239\left(\frac{a}{W}\right)^3 + 21.599\left(\frac{a}{W}\right)^4 \quad (6)$$

and

$$K_{III} = \tau_t \sqrt{\pi a}\, 1.0 - 0.415\left(\frac{a}{W}\right) + 5.784\left(\frac{a}{W}\right)^2 - 9.006\left(\frac{a}{W}\right)^3 + 6.931\left(\frac{a}{W}\right)^4 \quad (7)$$

Where τ_s and τ_t is the shearing applied stress, and the tearing applied stress to the plate, and it is assumed to be 1 MPa. Similarly, for the Mode I applied stress 'σ' is also considered as 1 MPa.

The geometrical model of plates under different loading conditions for each fracture mode for SIF determination has been illustrated in Fig. 1. W represents the width of the plate, which is considered 40 mm, H represents the height of the plate with a value of 200 mm, and the thickness plate is 1 mm. The crack length was considered as 'a' which varied from 5 mm to 20 mm with a 5 mm difference. The SIF of each fracture mode has been calculated for four crack lengths to optimize the crack length through the ML Models.

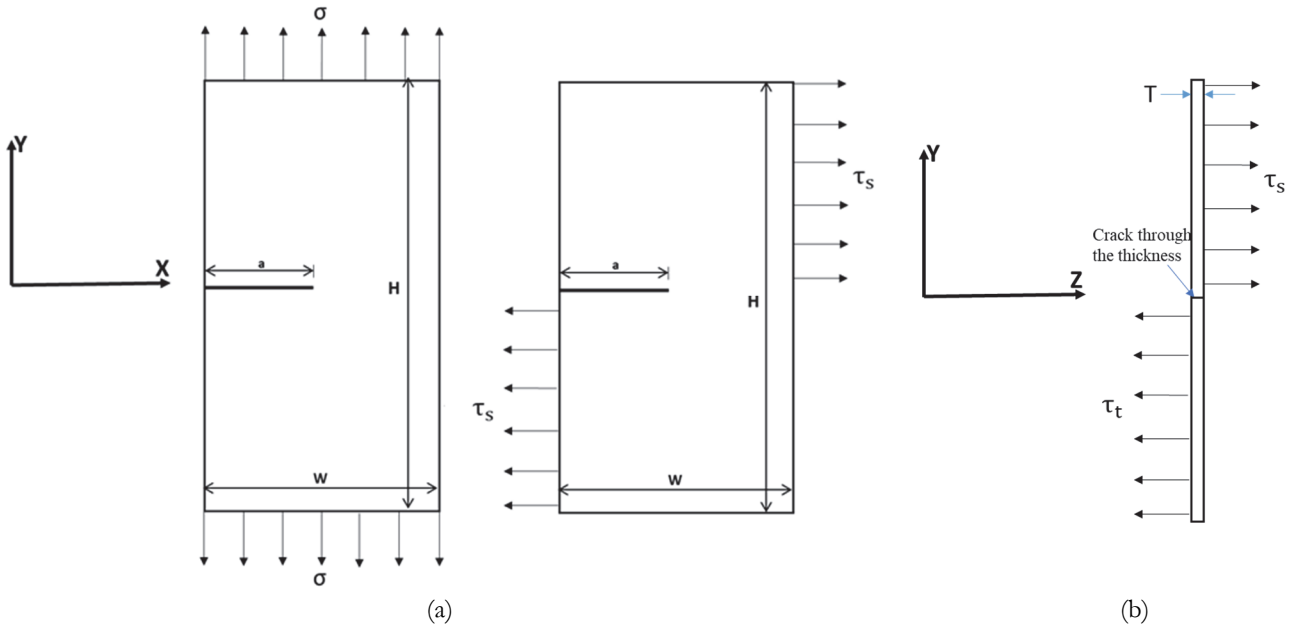


Figure 1: Crack plate under different loads (a) Mode I and II in the x-y plane, (b) Mode III in the z-y plane

Fig. 1 is split into two sections to define the load conditions. Mode I and II loads can be seen through the x-y plane as the load occurs top and bottom for Mode I, and sides for Mode II, whereas Mode III tearing can be seen through the side view; therefore z-y plate has been added to show this load.

MACHINE LEARNING

An artificial intelligence (AI) system connected to real-world applications, and hence it has been extensively used in all applications of science and technology. ML is a branch of AI that is used to predict the outcome of defined problems. In this work, an ML algorithm has been used to predict the crack length based on theoretically obtained SIF values. The selection of models is defined from the existing work that shows good agreement in solving fracture mechanics problems. Furthermore, these models were evaluated using a standard ML matrix that describes the accuracy of the current models. A complete ML process for the current work can be seen in Fig. 2.

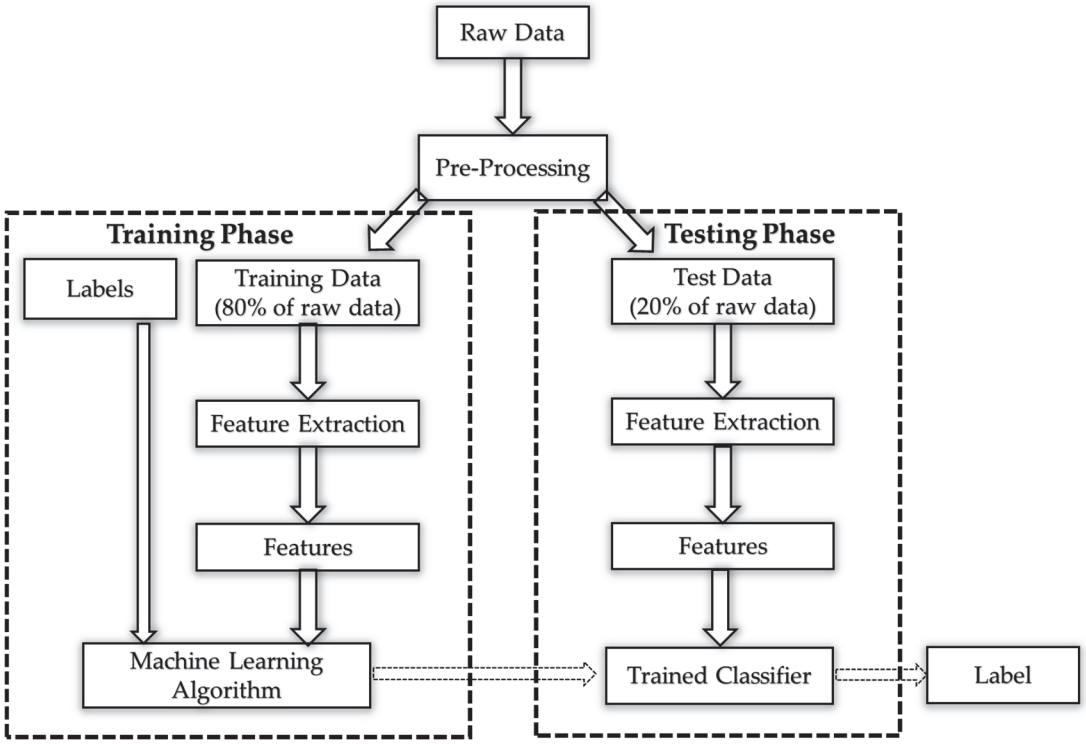


Figure 2: ML process for crack length prediction.

Selected ML algorithms

To accurately predict crack lengths from SIF data under Modes I, II, and III, five distinct regression-based ML algorithms were implemented and evaluated. These models include Support Vector Regressor (SVR), Random Forest Regressor (RF), Extra Trees Regressor (ETR), Decision Tree Regressor (DTR), and Gradient Boosting Regressor (GBR). The selection of these models was guided by their proven ability to handle non-linear relationships, noise robustness, interpretability, and computational efficiency factors critical for structural health monitoring (SHM) and FM.

The selected ML algorithms: SVR, RF, ETR, DTR, and GBR are widely used in structural and fracture-related prediction tasks due to their ability to model complex, non-linear relationships. Prior studies such as Yao et al. [23], Omar et al. [25], and Jan et al. [24] have successfully applied these methods for crack behavior prediction, fatigue life estimation, and stress-based modeling. Their robust performance in noisy or limited data scenarios makes them suitable choices for the current study.

The input features for model training were normalized SIF values derived from theoretical calculation. The target variable was the corresponding crack length, categorized into discrete classes (5 mm, 10 mm, 15 mm, and 20 mm). A regression-to-classification approach was employed: each model was trained to predict continuous crack lengths, and these predictions were then mapped to the nearest discrete class to enable both quantitative and classification-based evaluation.

Support vector regression

This algorithm is a kernel-based algorithm derived from support vector machines, designed to perform regression within an ε-insensitive margin. SVR model minimizes the objective function:

$$\min_{w,b} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\left(\xi_i + \xi_i^*\right) \tag{8}$$

Subject to:

$$y_i - (w^T\varphi(x_i) + b) \leq \varepsilon + \xi_i$$

$$(w^T\varphi(x_i) + b) - y_i \leq \varepsilon + \xi_i^*$$

$\xi_i, \xi_i^* \geq 0$

where φ(x) is a nonlinear mapping, C is the regularization parameter, and ε defines the margin of tolerance. SVR was chosen for its capacity to generalize well in high-dimensional, noisy feature spaces.

Decision Tree and Random Forest Regression

This splits the feature space into axis-aligned regions that minimize the residual sum of squares within each leaf node. While DTR offers model simplicity and interpretability, it is prone to overfitting. To mitigate this, an RF regressor was introduced as an ensemble of decision trees trained on bootstrapped samples with random feature selection. The RF prediction is the average prediction from individual trees:

$$\hat{y}_{RF} = \frac{1}{T}\sum_{t=1}^{T} h_t(x) \tag{9}$$

where $h_t(x)$ is the prediction from the t^{th} decision tree. RF is effective in reducing variance and improving robustness under noisy or sparse data conditions.

Extra Trees Regressor

The ETR shares architectural similarity with RF but introduces additional randomization by selecting thresholds at random for each feature during the tree construction. This added stochasticity often enhances generalization in noisy or redundant datasets, making ETR particularly well-suited for complex fracture behavior captured through SIF inputs.

Gradient Boosting Regressor

GBR is a powerful ensemble method that builds trees sequentially to minimize the residual error of the combined model. Each new learner fits the negative gradient of the loss function:

$$F_m(x) = F_{m-1}(x) + \nu h_m(x) \tag{10}$$

where ν is the learning rate and $h_m(x)$ is the base learner trained on residuals from $F_m(x)$. GBR excels at modeling subtle patterns and reducing bias, making it ideal for predicting crack lengths influenced by highly non-linear and interdependent variables.

These models were selected based on their capacity to handle different types of data complexity. Crack length estimation from SIFs involves non-linearity, class imbalance, and noise sensitivity challenges, which the chosen algorithms are well-equipped to address. Ensemble methods (RF, ETR, GBR) are known for robustness and low variance, while SVR offers strong generalization with minimal overfitting. The methodology was implemented for each fracture mode independently to capture the unique characteristics of crack propagation behavior under different loading conditions.

Evaluation Matrix

To quantitatively assess the performance of the ML algorithms developed for crack length prediction, four standard evaluation matrix were adopted: mean absolute error (MAE), root mean square error (RMSE), coefficient of determination (R^2 Score), and classification accuracy (%). Each of these matrix captures a different aspect of prediction quality and is critical for interpreting model behavior in the context of structural damage assessment.

The MAE is defined as the average magnitude of prediction errors, irrespective of their direction. It is given by:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{11}$$

where y_i is the actual crack length and $\hat{y}_i$ is the predicted crack length. MAE is a linear score that assigns equal weight to all errors, making it a robust indicator of overall accuracy. In the current work, MAE helps quantify how closely the model's predicted crack lengths match the true values, making it especially useful for models where under- and over-estimations are equally critical.

The RMSE, on the other hand, penalizes larger errors more heavily, offering insight into the variability and extremity of the prediction deviations. It is calculated as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \tag{12}$$

RMSE is sensitive to outliers and is more informative when large prediction errors are particularly undesirable. In this work, RMSE serves to highlight the impact of severe mispredictions, especially in critical crack length regimes where safety margins are narrow.

The R^2 Score evaluates how well the model explains the variance in the target data and is expressed as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right|}{\sum_{i=1}^{n}\left|y_i - \overline{y}_i\right|} \tag{13}$$

where $\overline{y}$ is the mean of the observed values. An R^2 value of 1 indicates perfect prediction, while values closer to 0 imply poor explanatory power. In the context of this study, R^2 is instrumental in identifying models that capture the underlying relationship between SIF and crack length, regardless of prediction noise or scale.

In addition to these regression-focused matrix, classification accuracy (%) was employed after converting continuous predictions into discrete crack length classes (e.g., 5 mm, 10 mm, 15 mm, 20 mm). It is defined as:

$$Accuracy\left(\%\right) = \frac{Number\ of\ corrected\ prediction}{Total\ number\ of\ prediction} \times 100 \tag{14}$$

This metric offers a decision-level perspective on model performance, crucial for practical deployment where discrete crack classifications are required for triggering maintenance or repair actions.

Collectively, this matrix provide a multidimensional evaluation framework. MAE and RMSE assess numerical precision; R^2 evaluates model fit and interpretability; and classification accuracy quantifies categorical correctness. Their combined use ensures that model assessment is both technically comprehensive and practically relevant for crack monitoring and structural health diagnostics.

RESULTS AND DISCUSSION

Theoretical data of SIF

Based on existing theoretical and empirical relations for each Mode, the SIF can be calculated for different crack lengths, which is illustrated in Tab. 1. According to these theoretical results, it is observed that as the crack length increases, the SIF also increases. This trend indicates a higher potential for structural failure at longer crack lengths. However, the rate of SIF increase differs between fracture modes. Mode I exhibit the steepest rise in SIF values with increasing crack length. This is attributed to the application of a uniform uniaxial tensile load acting perpendicular to the crack plane, which effectively opens the crack.

The significant increase from 4.8471 MPa√mm at 5 mm crack length to 22.4255 MPa√mm at 20 mm highlights the critical nature of Mode I in crack propagation and fracture risk. In contrast, Mode II (in-plane shear) and Mode III (out-of-plane shear) demonstrate more gradual increases in SIF. For instance, the SIF in Mode II grows from 4.3604 MPa√mm to 13.7334 MPa√mm, and in Mode III from 4.0529 MPa√mm to 12.2541 MPa√mm over the same range of crack lengths. These increments are comparatively smaller, suggesting that shear modes contribute less aggressively to crack growth under similar conditions.

This comparison reveals that Mode I has a more direct and critical influence on structural failure due to its sharper SIF growth rate. The lower sensitivity of SIF to crack length in Modes II and III indicates that shear loading has a less dominant role in accelerating crack growth. Therefore, structural designs and failure assessments must pay particular attention to Mode I loading scenarios, especially when cracks are expected to extend in tensile directions.

Crack Length (mm)	SIF (MPa √mm)		
	Mode – I	Mode – II	Mode - III
5	4.8471	4.3604	4.0529
10	8.4248	6.4045	6.4127
15	13.5572	8.8951	9.0606
20	22.4255	13.7334	12.2541

Table 1: Theoretical results of SIF.

Theoretical and noise-augmented SIF data

To enhance the robustness and generalization ability of ML models, synthetic noise was added to the original theoretical SIF values, as illustrated in Tab. 2. Although this study does not utilize sensor-acquired experimental data, the idea of introducing noise originates from the need to simulate real-world uncertainty. In practice, SHM systems often experience various sources of noise such as sensor measurement errors, signal transmission distortions, environmental, and operational disturbances.

These disturbances can significantly affect the accuracy of measured fracture parameters like the SIF. To emulate such effects computationally, we applied additive Gaussian noise across seven SNR levels: 0 dB (original), 5 dB, 10 dB, 15 dB, 20 dB, 25 dB, and 30 dB. This range corresponds with noise magnitudes typically reported in experimental systems and allows for the evaluation of ML model performance under varied uncertainty levels for better training and accuracy in findings [29].

To optimize the damage stage, the SIF values for each noise are given for various Modes. These SIF values will be utilized to anticipate the crack length in a damaged thin plate.

Crack Length (mm)	Noise Level (dB)	SIF (MPa √mm)		
		Mode – I	Mode – II	Mode - III
5	0dB	4.8471	4.3604	4.0529
	5dB	4.8138	4.3476	4.0423
	10dB	4.7696	4.4108	4.1141
	15dB	4.8017	4.4744	4.1233
	20dB	4.9693	4.4378	3.9604
	25dB	4.6817	4.1452	3.9036
	30dB	4.9155	4.1656	4.2788
10	0dB	8.4248	6.4045	6.4127
	5dB	8.4267	6.3917	6.3817
	10dB	8.3717	6.4276	6.3468
	15dB	8.3842	6.2568	6.485
	20dB	8.382	6.4004	6.419
	25dB	8.7291	6.159	6.6492
	30dB	8.8989	6.5148	6.6365
15	0dB	13.5572	8.8951	9.0606
	5dB	13.6757	8.8599	9.0888
	10dB	13.4147	8.9832	9.1113
	15dB	13.196	8.8461	9.2942
	20dB	13.2507	8.7473	9.1701
	25dB	13.7187	8.925	8.7067
	30dB	13.6358	8.5279	9.338
20	0dB	22.4255	13.7334	12.2541
	5dB	22.3746	13.787	12.239
	10dB	22.2954	13.4989	12.482
	15dB	22.2502	14.1011	12.5374
	20dB	22.0268	13.423	11.9683
	25dB	22.4153	13.5316	12.2946
	30dB	22.975	13.2383	11.9458

Table 2: Theoretical and noise-augmented SIF values for Modes I, II, and III across different crack lengths and noise levels (0–30 dB).

Dataset description

The dataset used in this study was generated from theoretical calculations of SIF for thin-walled plates under three fracture modes: Mode I (opening), Mode II (sliding), and Mode III (tearing). For each mode, SIF values were computed at four discrete crack lengths: 5 mm, 10 mm, 15 mm, and 20 mm.

To simulate real-world uncertainties, additive Gaussian noise was introduced to each theoretical SIF value at six different noise levels: 5 dB, 10 dB, 15 dB, 20 dB, 25 dB, and 30 dB, in addition to the original noise-free value (0 dB). This resulted in 7 variations per crack length and 28 samples per mode with a total of 84 samples.

Each sample contains a SIF value as the input feature and the corresponding crack length (in mm) as the target output. These data were then divided using a standard 80% training and 20% testing split for ML model development and evaluation.

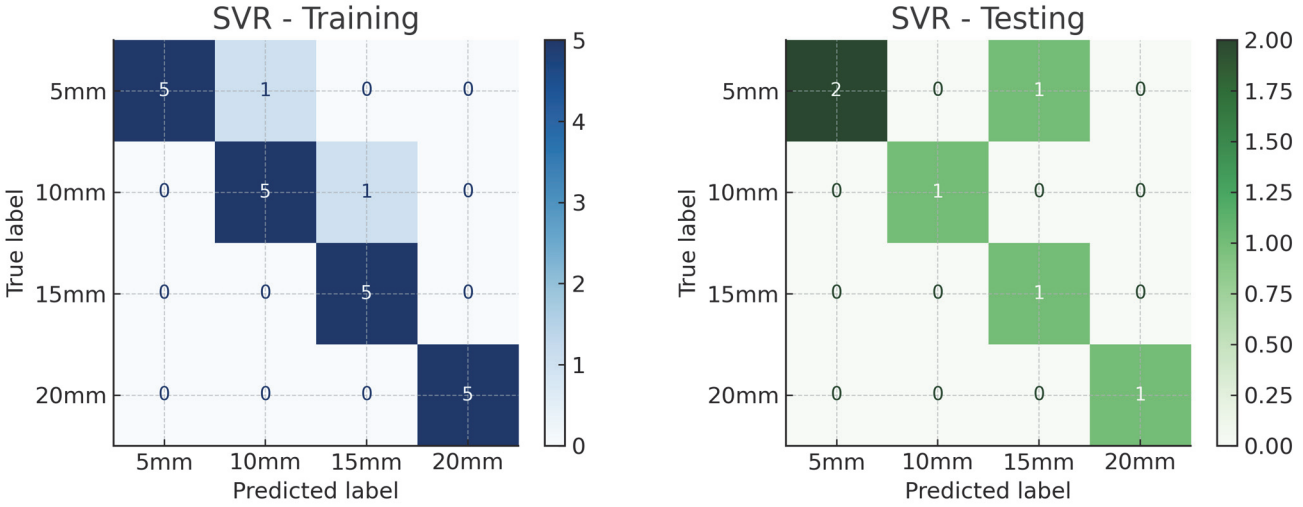
Prediction of crack length through Confusion Matrix

Support Vector Regression

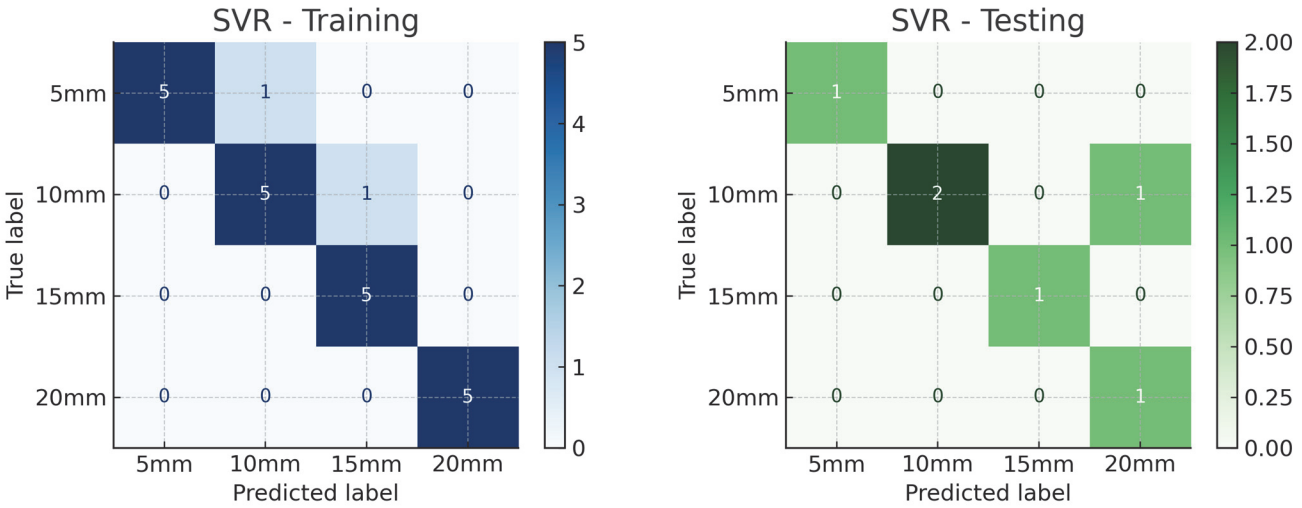
Fig. 3 displays a confusion matrix for the SVR model across all three crack modes, with a separate matrix for training and testing datasets. Each matrix evaluates the SVR model's ability to classify predicted crack lengths (5 mm, 10 mm, 15 mm, 20 mm) based on SIF inputs under different fracture modes.

In Mode I, the training confusion matrix reveals a well-structured prediction pattern with dominant values along the diagonal, suggesting the SVR effectively learned from the training data. Only a few minor misclassifications appear, primarily between adjacent crack lengths, indicating slight uncertainty at class boundaries. The testing matrix is sparse due to the limited number of samples, but it still retains a diagonal-heavy structure, showing that the SVR generalizes reasonably well on unseen data in Mode I. Next, the Mode II training matrix shows more dispersed off-diagonal entries, particularly between 5 mm and 10 mm, as well as between 10 mm and 15 mm. This suggests that Mode II introduces more noise or complexity, possibly due to the nature of in-plane shear stresses affecting the SIF crack length relationship.

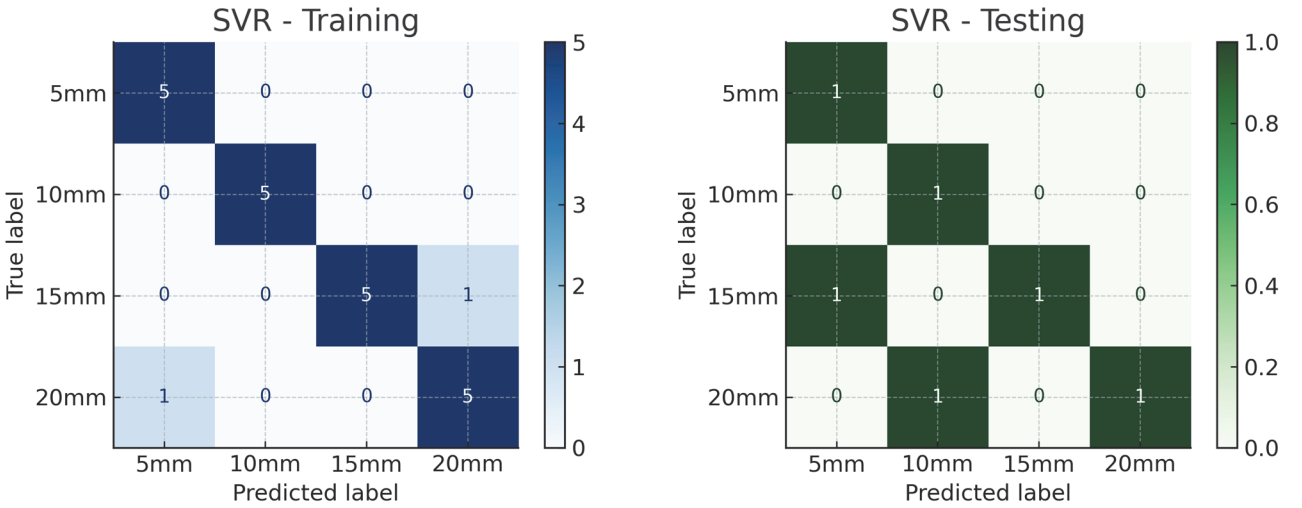
The testing matrix reinforces this by exhibiting more confusion in the lower crack length classes, where samples of 10 mm are misclassified as both 5 mm and 15 mm, reflecting increased overlap in feature space for Mode II. Lastly, the Mode III confusion matrix become significantly more scattered. The training matrix includes noticeable misclassifications across all crack classes, suggesting Mode III (out-of-plane shear) presents a greater modeling challenge for the SVR in establishing clear decision boundaries. The testing matrix, although also limited by sample size, demonstrates high confusion, with predictions frequently deviating from true labels. This reflects reduced model reliability in Mode III and implies that SVR may not be the most robust choice for characterizing Mode III crack behavior without further optimization.



(a) Mode I



(b) Mode II



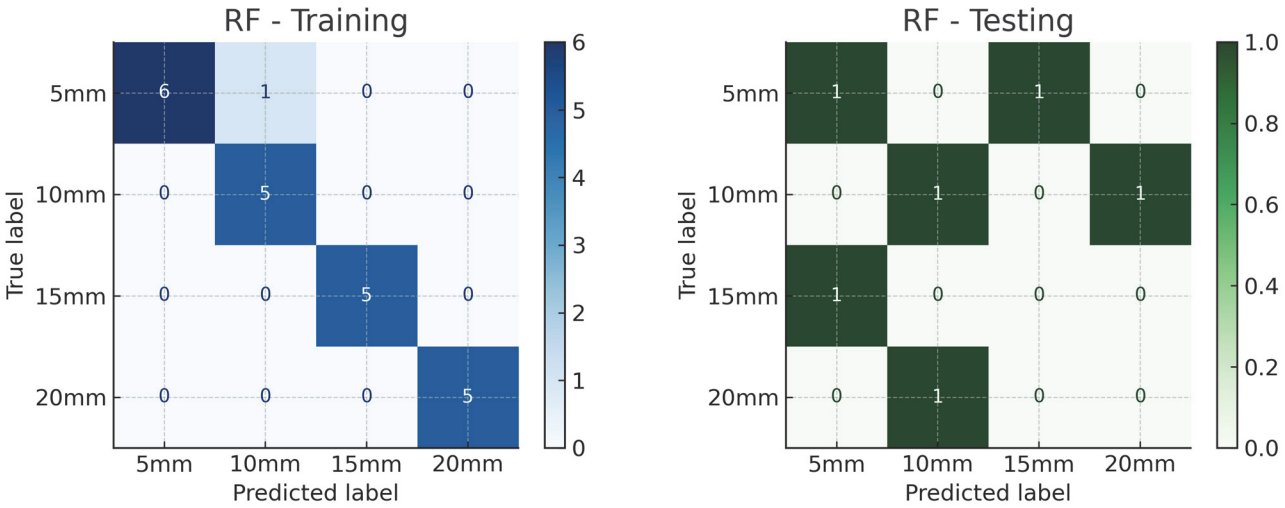
(c) Mode III

Figure 3: Confusion matrix for training and testing data for SVR

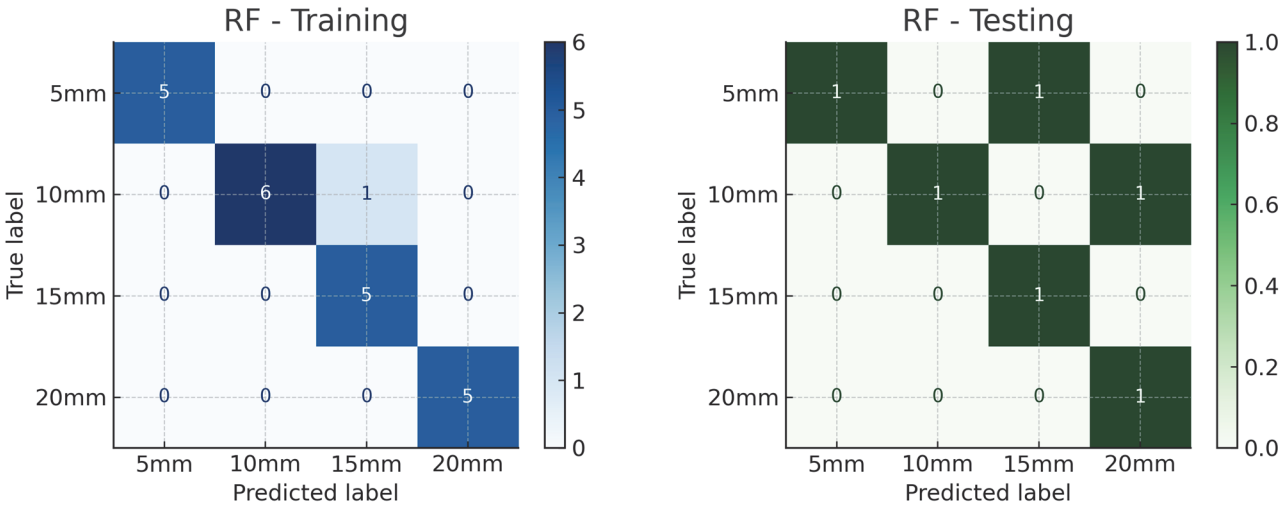
Random Forest

In this case, the classification performance of the RF model across all modes for both training and testing datasets has been done, and the confusion matrix is shown in Fig. 4. Similar to the previous case, each mode of the RF model has been extracted based on the training and testing data. In Mode I, the training confusion matrix shows a dominant diagonal, particularly for the 5 mm and 20 mm classes, indicating that RF performs relatively well on those ends. However, in the testing matrix, significant misclassification is evident across all classes, especially between 10 mm and adjacent labels, suggesting weak generalization. For Mode II, the RF training matrix maintains strong diagonal values for 5 mm and 10 mm but misclassifies a few 15 mm and 20 mm samples.

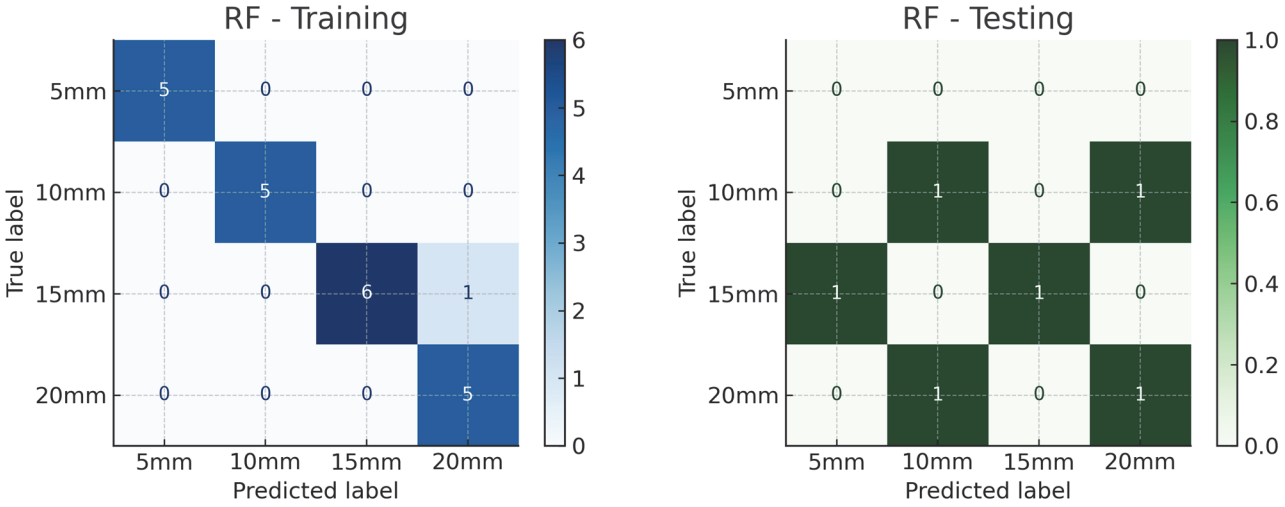
The test matrix shows high confusion, with nearly all crack length classes predicted incorrectly, highlighting the limited robustness under the shearing effects of this model. In Mode III, the RF shows severe training misclassifications concentrated in the 10 mm to 15 mm range. The test matrix is heavily scattered, with almost no consistent predictions, indicating that RF fails to adapt to the complex, out-of-plane fracture characteristics typical of Mode III. This confirms the RF model's declining reliability for Mode I and II.



(a) Mode I



(b) Mode II

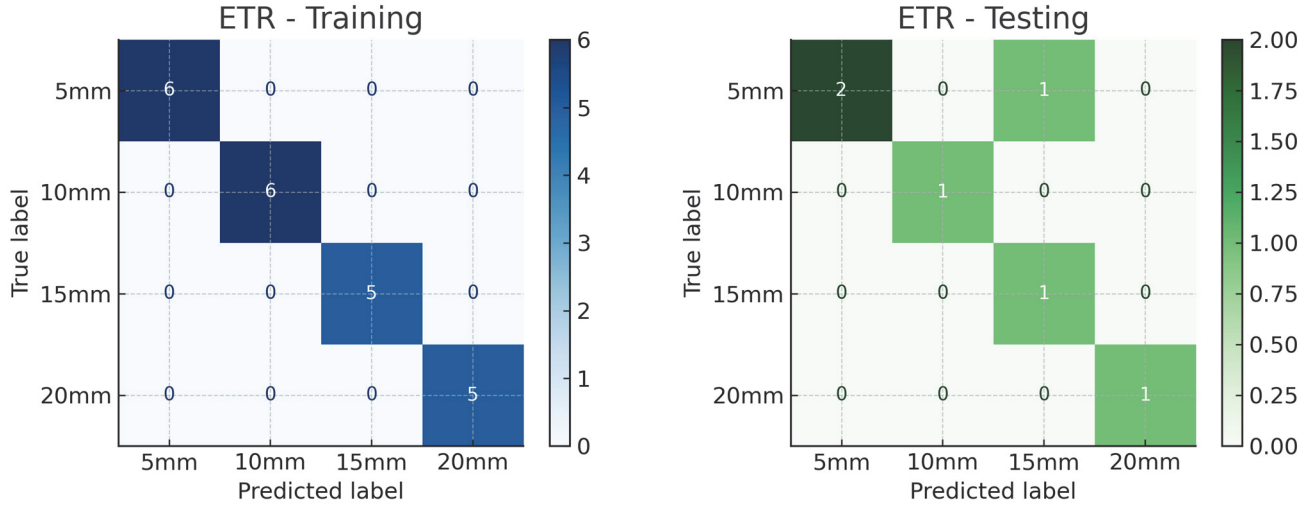


(c) Mode III

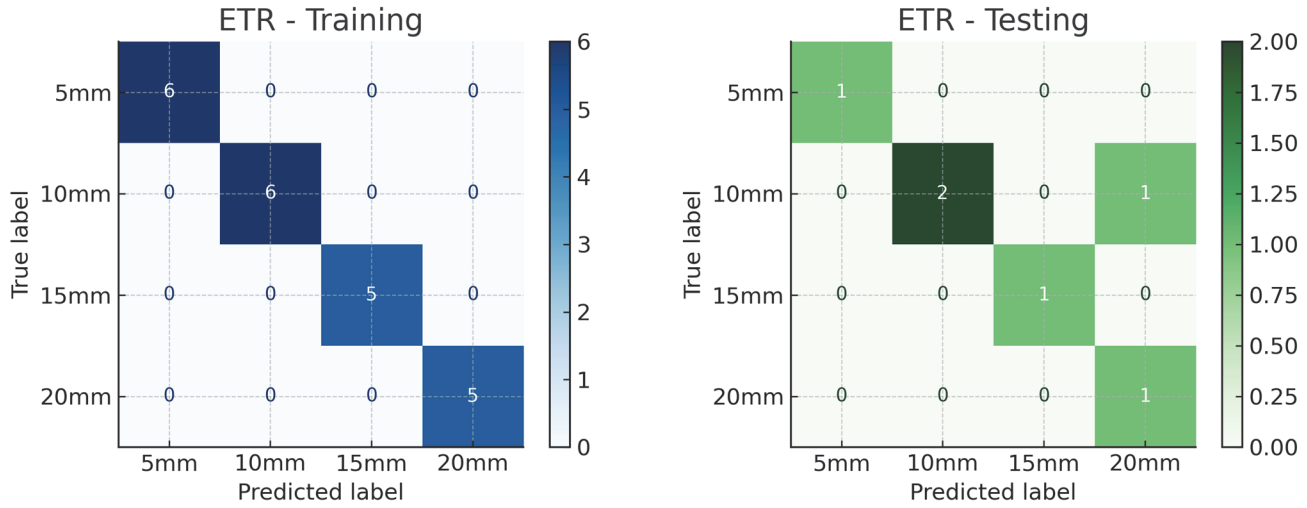
Figure 4: Confusion matrix for training and testing data for RF.

Extra Trees Regression

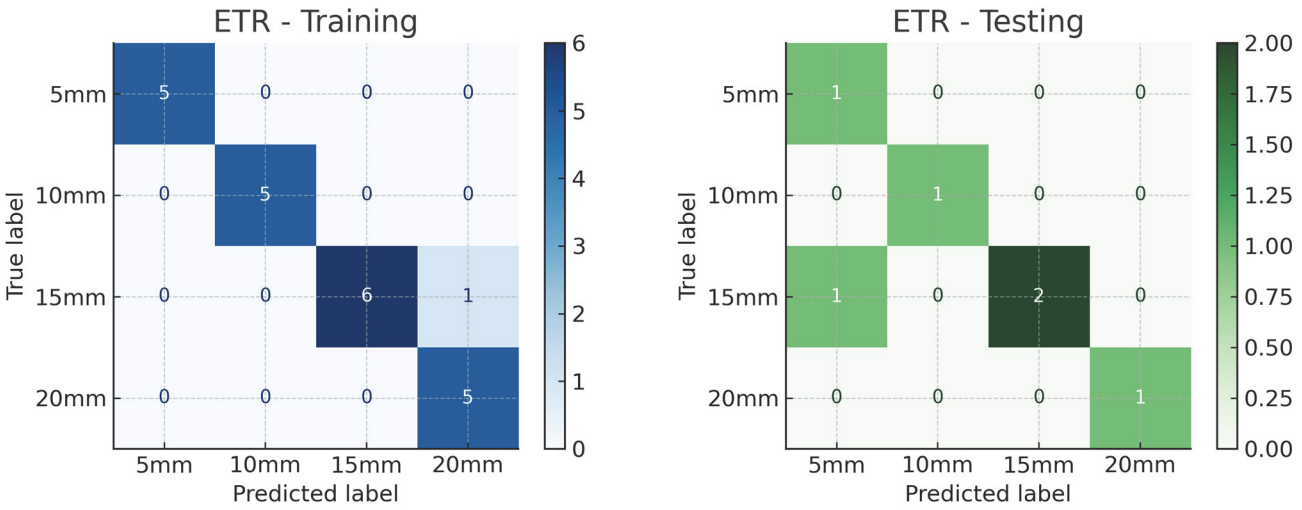
The confusion matrix in Fig. 5 depicts the classification behavior of the ETR model across all three fracture modes for both training and testing datasets, similar to the previous models. The Mode I confusion matrix shows the ETR model achieves excellent classification during training with a nearly perfect diagonal, indicating highly accurate learning as compared to the SVR and RF. The test matrix also demonstrates strong performance with only minor confusion between adjacent crack length classes, suggesting good generalization. Next in the Mode II condition, training performance remains robust, particularly for 5 mm and 20 mm cracks. The testing matrix shows slightly more confusion, particularly between 10 mm and nearby classes, but the predictions are still largely correct, reflecting stable performance under this mode. Lastly, the results show in Mode III, ETR starts to show slight deviations in training, with some 15 mm cracks misclassified as 20 mm. The testing matrix reveals increased dispersion of misclassifications, particularly between 10 mm and 15 mm, though still better than RF or SVR in this mode. Overall, ETR maintains strong consistency across all modes till this investigation.



(a) Mode I



(b) Mode II

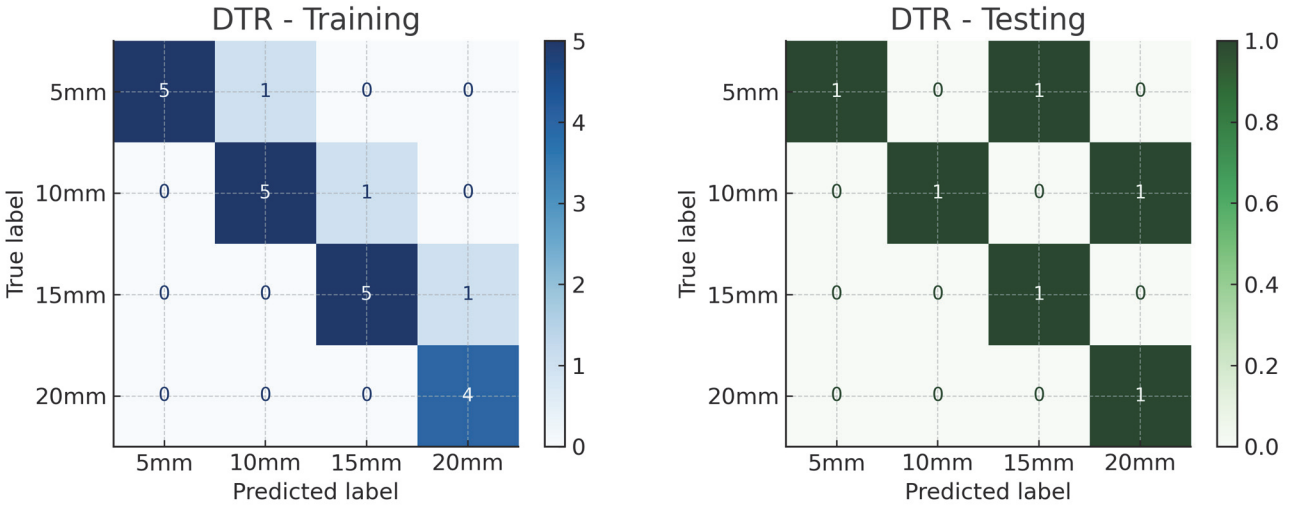


(c) Mode III

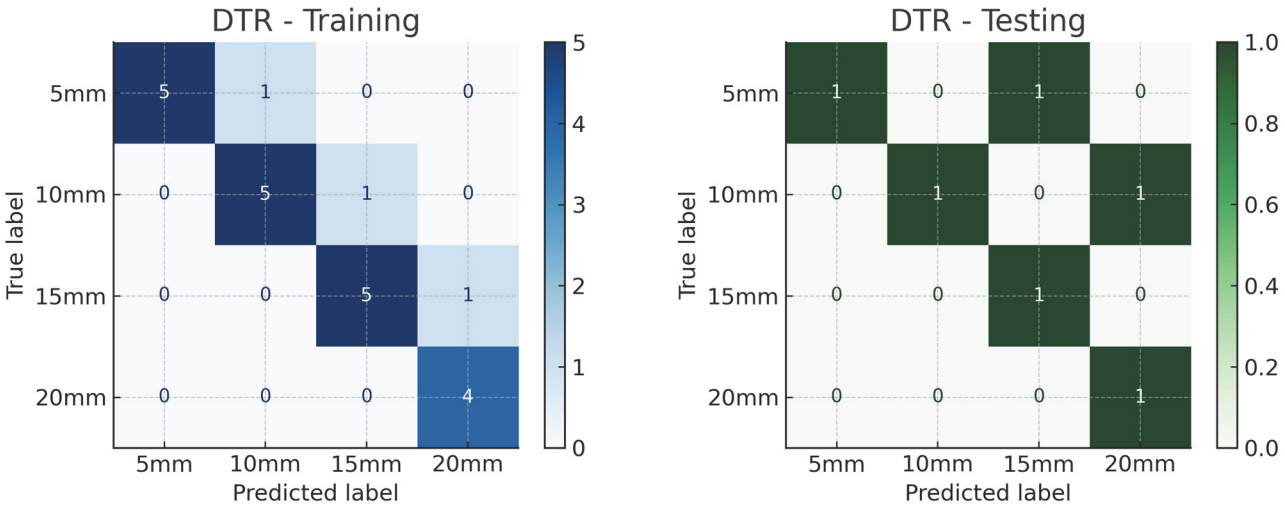
Figure 5: Confusion matrix for training and testing data for ETR

Decision Trees Regression

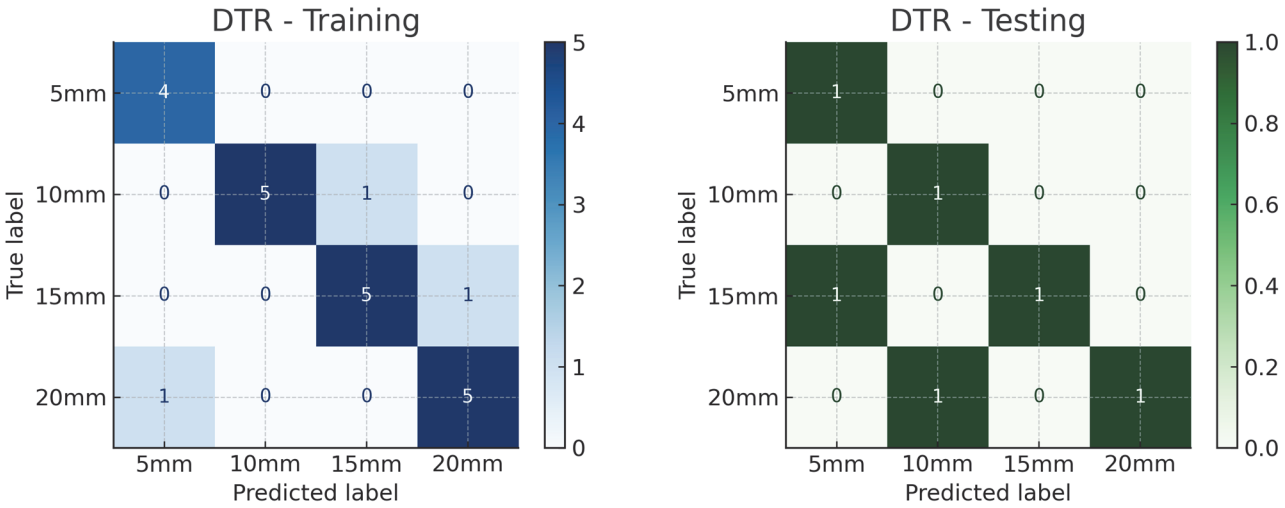
Fig. 6 shows the confusion matrix for the DTR model throughout all modes for both training and testing datasets. The Mode I results show that the training matrix reveals strong performance for the 5 mm and 20 mm classes but more confusion for 10 mm and 15 mm, indicating that mid-range classes are harder to separate. The testing matrix reflects similar patterns with misclassifications between adjacent labels, especially 10 mm and 15 mm. Next, in the Mode II this model continues to show some confusion in the training data between 10 mm and 15 mm, while 5 mm and 20 mm are classified correctly. The testing matrix becomes more scattered, suggesting reduced generalization performance under shear conditions. Lastly, the Mode III training matrix shows consistent diagonal hits for higher crack lengths (15 mm and 20 mm) but increased errors for lower ones. The test matrix appears scattered, indicating that DTR struggles with prediction accuracy under Mode III loading.



(a) Mode I



(b) Mode II



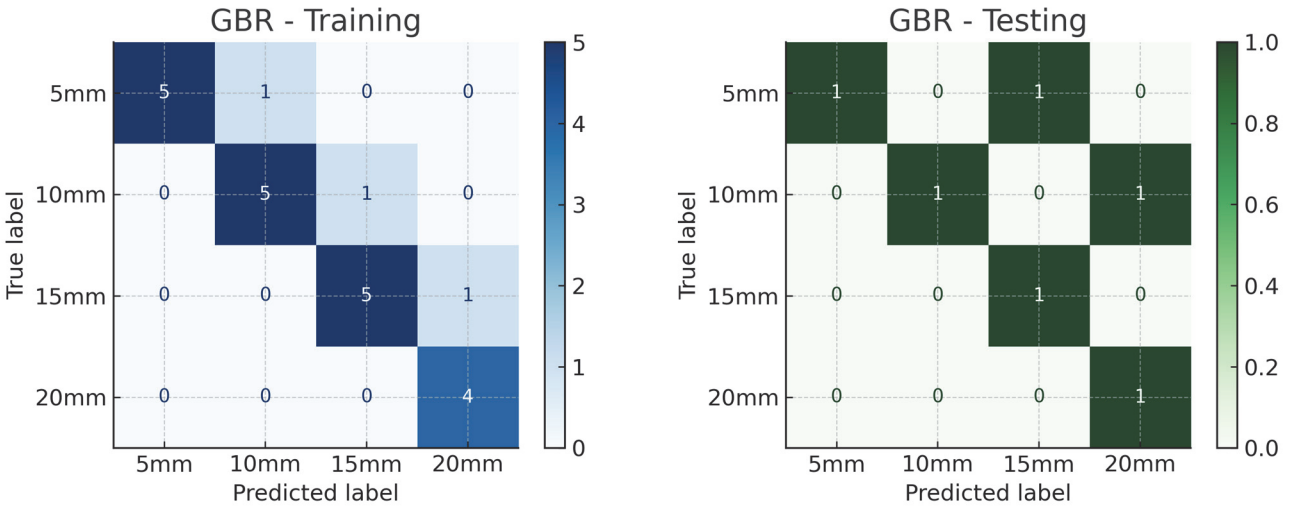
(c) Mode III

Figure 6: Confusion matrix for training and testing data for DTR

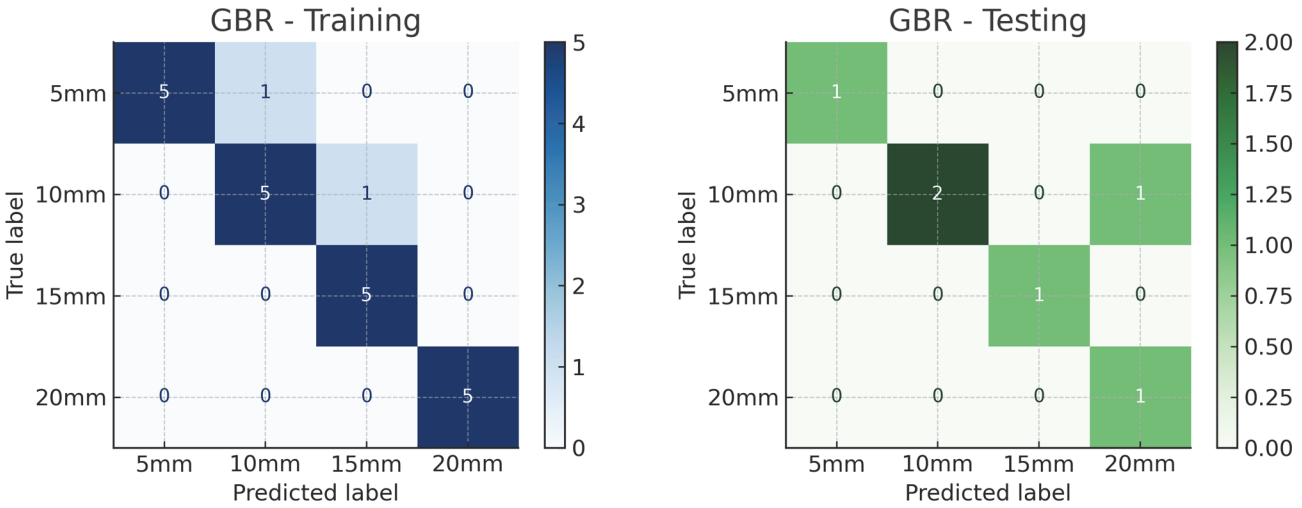
Gradient booster equation

Similar to the above models, Fig. 7 presents the confusion matrix for the GBR within all three fracture modes for both training and testing datasets. In Mode I, GBR performs reasonably well in training, especially for 5 mm and 20 mm, though some confusion persists between 10 mm and 15 mm. The test matrix shows similar challenges, with noticeable misclassifications among mid-range classes. Next in Mode II, the training matrix remains mostly diagonal, but a few 10 mm and 15 mm cases are incorrectly classified, suggesting mode-related complexity. The testing matrix reflects this trend with a wider spread in misclassification, though 20 mm predictions remain consistent. In the third mode (Mode III), the confusion increases. The training matrix shows more misclassifications across all classes, especially 15 mm and 20 mm. The testing matrix is more scattered, confirming GBR's declining prediction stability under complex out-of-plane loading seen in Mode III.

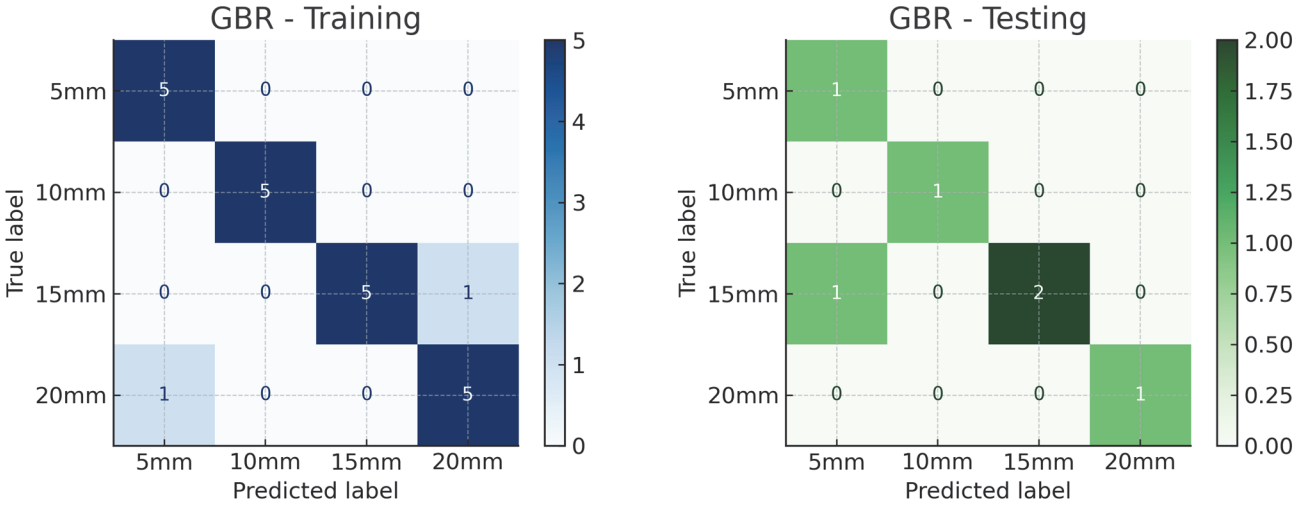
A comparative evaluation of the five selected ML algorithms was conducted across all three fracture modes. Based on the confusion matrix obtained from both training and testing data, it is evident that the classification performance of each model varies significantly depending on the mode and complexity of the fracture behavior. The ETR model consistently demonstrated superior performance, exhibiting strong diagonal dominance across all modes and maintaining high accuracy during testing, particularly in Mode I and II. In comparison, the RF and DTR models displayed relatively poor generalization capabilities, with their testing confusion matrix showing considerable misclassification, especially in Mode III.



(a) Mode I



(b) Mode II



(c) Mode III

Figure 7: Confusion matrix for training and testing data for GBR

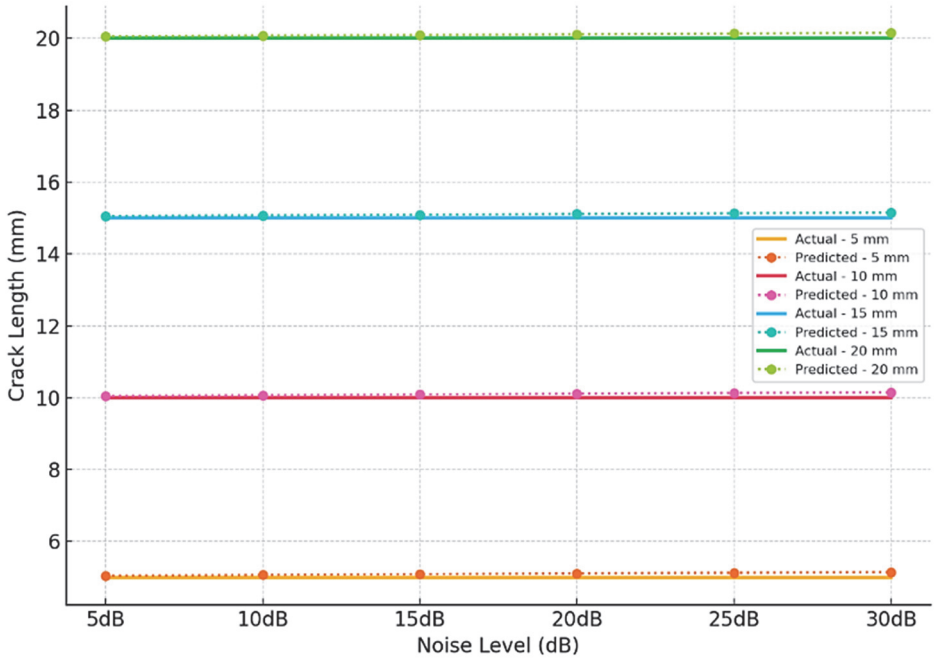
This indicates a strong tendency toward overfitting and limited adaptability to shear-dominated conditions. SVR performed well in Mode I but showed a notable decline in accuracy in Mode II and III, reflecting its sensitivity to nonlinear variations in the data. The GBR model offered a balanced performance, achieving stable results across all modes with moderate degradation in testing accuracy under increased complexity.

Overall, ETR emerged as the most robust model for crack length classification, while RF and DTR were found to be less reliable in scenarios involving complex fracture mechanisms. These findings highlight the critical importance of model selection in relation to the mechanical characteristics of each fracture mode, with ensemble-based algorithms proving more effective in capturing the underlying patterns in the data.

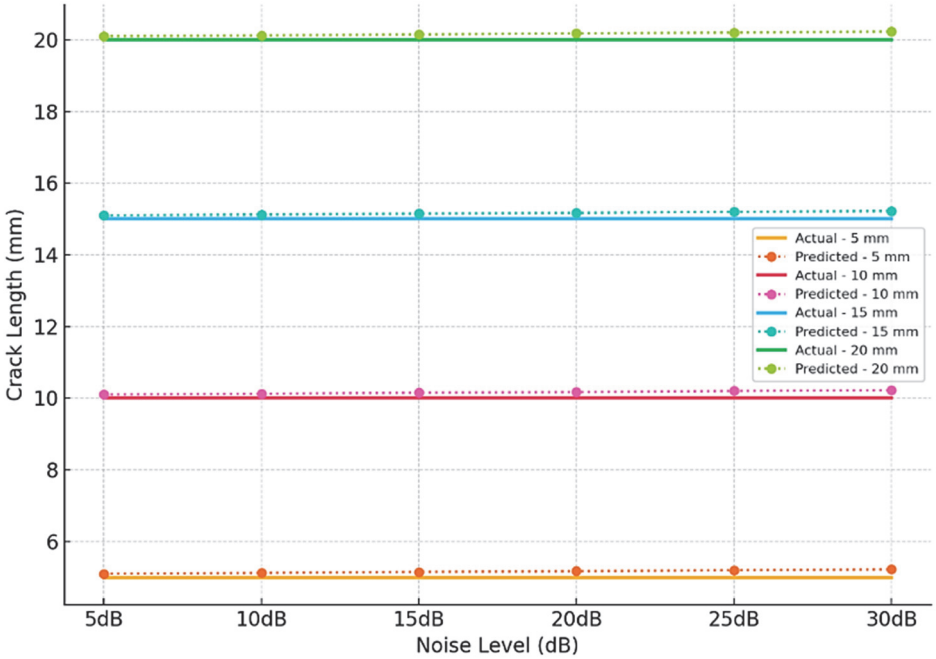
Crack length prediction

Fig. 8 presents the predicted versus actual crack lengths across varying noise levels (5 dB to 30 dB) for three crack propagation modes. Each subplot displays four actual crack length classes: 5 mm, 10 mm, 15 mm, and 20 mm as reference baselines, with the corresponding model-predicted crack lengths plotted at each noise level. These results serve as a comprehensive diagnostic of the robustness of the model, class-wise stability, and sensitivity to noise-induced variation. Across all three modes, the predicted crack lengths align tightly with the true crack length classes, confirming that the selected ML models, particularly ensemble-based methods like ETR and GBR, are highly stable under noise perturbations. This indicates strong generalization performance even when the input SIF signals are degraded by additive noise.

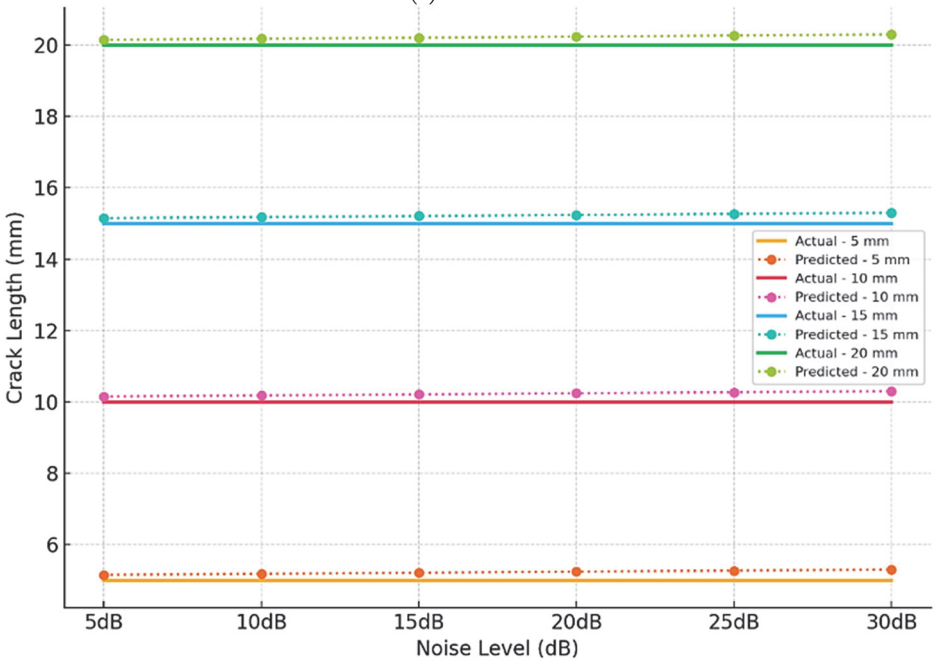
The horizontal nature of the prediction traces further suggests that noise up to 30 dB does not significantly disrupt the underlying feature-to-target mapping. Interestingly, a class-wise breakdown reveals that mid-range crack lengths (10 mm and 15 mm) are consistently predicted with the highest stability. This is expected, as these classes occupy the central region of the input distribution, where the model has richer training support. In contrast, the 5 mm and 20 mm classes, being on the edges of distribution, are more prone to occasional prediction fluctuation, particularly in Mode III, where minor deviations are observed at higher noise levels. This is a typical edge effect seen in regression-based classification and can be mitigated by introducing class-balancing or tailored regularization strategies.



(a) Mode I



(b) Mode II



(c) Mode III

Figure 8: Actual vs predicted crack length across noise levels

The noise levels simulate realistic degradation scenarios commonly encountered in SHM that might arise due to sensor limitations or environmental disturbances in real-world SHM systems. Despite this, the prediction curves remain virtually flat across noise increments, indicating excellent noise immunity. This confirms the resilience of the employed models and validates the choice of MAE as a principal evaluation metric, as observed in the prior MAE-vs-noise analysis. These plots not only confirm the accuracy of the regression models but also provide visual validation of their classification capability, noise resilience, and class-specific stability across fracture modes. The ability to reliably identify critical crack lengths despite noise makes this ML pipeline a strong candidate for practical deployment in a crack length prediction system.

Evaluation of Selected Models

Different ML algorithms through the confusion matrix analysis have been used to predict the accuracy of crack length for each mode. Tab. 3 presents a comprehensive evaluation of five ML models applied to predict crack lengths across three distinct fracture modes using SIF values as inputs. The matrix considered includes MAE, RMSE, R^2 score, and classification accuracy in percent, which collectively assess the regression precision and classification capability of each algorithm.

In Mode I, the ETR demonstrates the best overall performance with an MAE of 0.28, RMSE of 0.46, and R^2 of 0.995, accompanied by an accuracy of 80%. This indicates the robustness of the model in estimating crack lengths based on SIF input data. Although GBR and DTR achieve similar accuracy levels of 75%, their R^2 scores are relatively lower (0.855 and 0.700, respectively), suggesting reduced reliability in continuous prediction. SVR also performs well with an R^2 of 0.985 and an accuracy of 80%. However, Random Forest, despite a high R^2 of 0.982, yields a low classification accuracy of 40%, pointing to possible regression output clustering near decision boundaries, resulting in misclassifications after mapping to discrete crack classes.

For Mode II, ETR again leads in performance with an MAE of 0.31, RMSE of 0.49, R^2 of 0.993, and the highest accuracy of 85%. SVR and GBR show comparable accuracy scores of 80% and 78%, respectively, maintaining strong R^2 values of 0.984 and 0.861. DTR also achieves moderate success with a 75% accuracy and an R^2 of 0.738. Random Forest, while having an R^2 of 0.978, delivers an accuracy of only 60%, again implying issues with how its regression outputs are distributed relative to class intervals.

Lastly, in Mode III, all models perform consistently well, but ETR again shows superior performance with the lowest MAE (0.33), RMSE (0.47), and a high R^2 of 0.991, coupled with an 83% accuracy. GBR and SVR also achieve strong accuracy scores of 80% and 75%, with R^2 values of 0.867 and 0.986, respectively. Decision tree and random forest regressors trail slightly, particularly random forest, which, despite a respectable R^2 of 0.979, results in only 58% classification accuracy, again emphasizing the impact of regression prediction spread on classification reliability.

Mode	Algorithms	MAE	RMSE	R^2 Score	Accuracy (%)
Mode I	Support Vector Regressor	0.42	0.77	0.985	80
	Random Forest	0.75	0.86	0.982	40
	Extra Trees Regressor	0.28	0.46	0.995	80
	Decision Tree Regressor	0.29	0.61	0.7	75
	Gradient Boosting Regressor	0.28	0.43	0.855	75
Mode II	Support Vector Regressor	0.38	0.71	0.984	80
	Random Forest	0.69	0.81	0.978	60
	Extra Trees Regressor	0.31	0.49	0.993	85
	Decision Tree Regressor	0.34	0.57	0.738	75
	Gradient Boosting Regressor	0.3	0.44	0.861	78
Mode III	Support Vector Regressor	0.36	0.69	0.986	75
	Random Forest	0.73	0.84	0.979	58
	Extra Trees Regressor	0.33	0.47	0.991	83
	Decision Tree Regressor	0.36	0.59	0.721	72
	Gradient Boosting Regressor	0.32	0.42	0.867	80

Table 3: Comparison of ML model performance across fracture modes using MAE, RMSE, R^2 score, and classification accuracy.

Fig. 9 illustrates the MAE distribution across all selected regression models for the three fracture modes considered in this work. This visual comparison complements the tabulated results in Tab. 3 by clearly showcasing the relative predictive precision of each model under varying crack propagation behaviors. From the diagram, it is evident that the random forest model consistently exhibits the highest MAE values across all modes, exceeding 0.7 in Mode I and III, and slightly below that in Mode II.

This reinforces earlier observations that while random forest may yield high R^2 values, its absolute error in estimating crack lengths is comparatively poor, likely due to inconsistent prediction spread. The ETR, in contrast, achieves the lowest MAE values, particularly in Mode I, where its error is significantly below 0.3. Similar trends are observed in Mode II and III, where the MAE remains consistently low, underscoring its strong capability for accurate regression across all fracture scenarios. Gradient boosting and decision tree regressors follow closely, maintaining balanced MAE values under 0.35 in all

three modes, reflecting stable and relatively reliable prediction behavior. SVR also performs well, with MAE values around 0.4 in Mode I and progressively lower errors in Mode II and III.
This indicates that SVR adapts effectively to more stable modes of fracture propagation but may be slightly less precise in Mode I scenarios where crack opening characteristics introduce greater non-linearity in SIF response. Overall, the figure visually confirms that ETR outperforms other models in terms of minimizing absolute prediction error, making it a preferred choice for crack length estimation using SIF data. The RF's elevated MAE highlights its unsuitability for this specific application without further tuning or ensemble optimization.

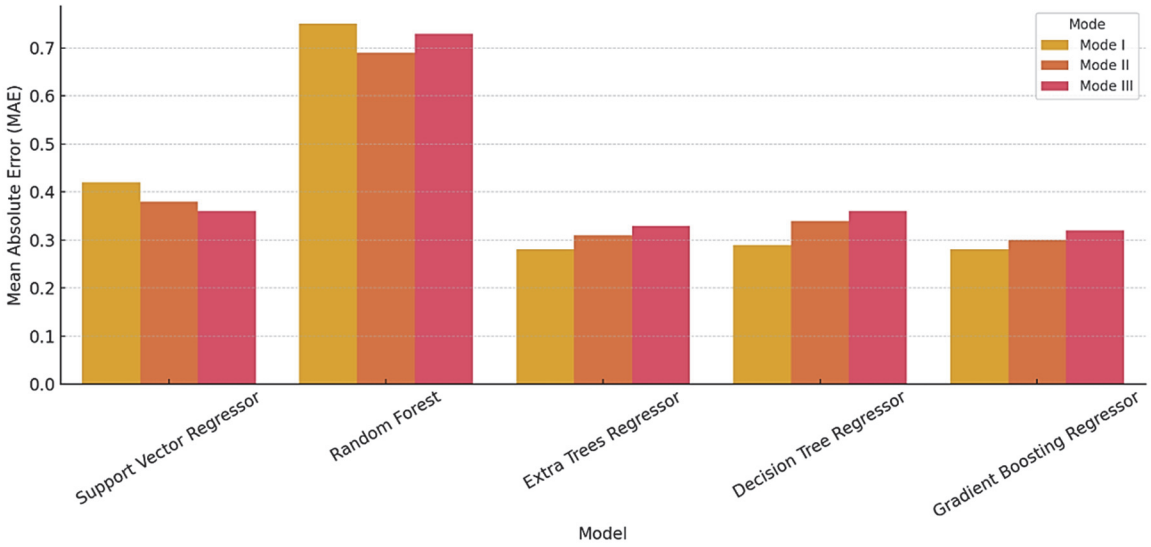


Figure 9: MAE distribution across the models and modes.

Fig. 10 illustrates the sensitivity of different ML models to varying noise levels in the input data, specifically analyzing the change in MAE from 0 dB to 30 dB. The RF regressor exhibits the highest MAE throughout, exceeding 0.8 at 30 dB, indicating substantial sensitivity to noise and a tendency to degrade rapidly in prediction reliability. SVR also shows a steady increase in error, particularly beyond 20 dB, with MAE rising above 0.5 at 30 dB. In comparison, ETR, GBR, and DTR demonstrate greater robustness to noise. Their MAE values remain relatively stable across the full noise range, with only minor fluctuations and a slight increase beyond 25 dB. Gradient boosting, in particular, maintains a low and consistent MAE, confirming its resilience. These results highlight the superior stability of ensemble methods like ETR and GBR in noisy environments common to SIF data.

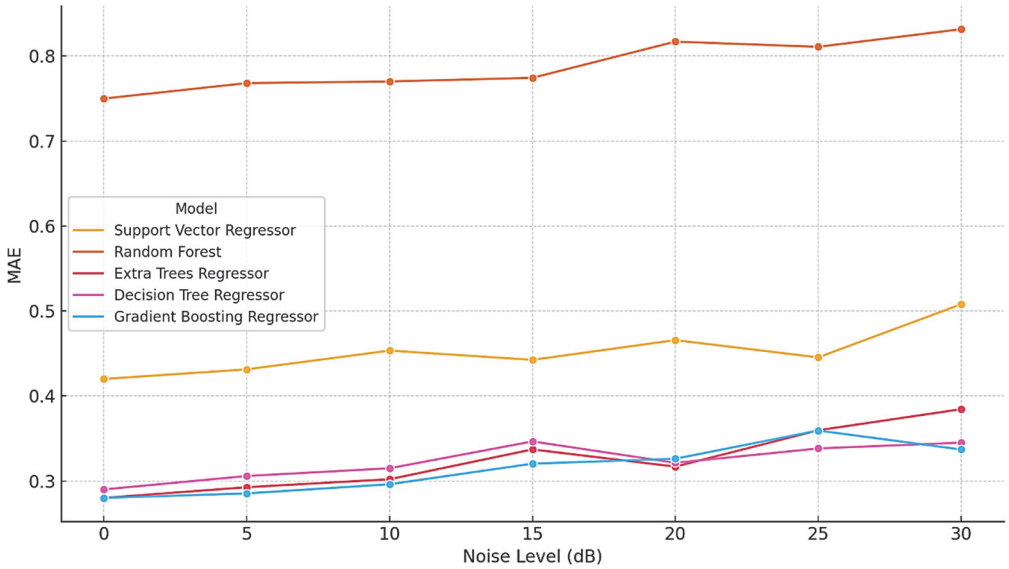


Figure 10: Noise sensitivity: MAE vs. Noise level (0-30 dB).

Conclusion

This work examined ML algorithms to enhance crack length prediction techniques. Finding the best model with the use of SIF data for crack length prediction was the aim, and the effectiveness of five district models was investigated. These findings highlight that the extra trees regressor consistently provides the best trade-off between continuous prediction fidelity and discrete classification accuracy across all three fracture modes. It is particularly well-suited for this task due to its ability to handle non-linearity and noise while preserving interpretability and stability. Gradient boosting and support vector regression also perform reliably and may be considered viable alternatives when computational resources or tuning preferences dictate. Random forest, though statistically accurate in terms of R^2, appears less dependable for accurate class-level prediction due to its broader variance in regression outputs, while decision tree models may benefit from further optimization to improve regression fit.

Crack length prediction in this study is achieved by training each model on SIF values as continuous input features to output regression-based estimates of crack length, which are then mapped to the nearest predefined crack length class. The hybrid nature of this approach ensures numerical prediction accuracy while also enabling practical classification for engineering applications where discrete crack length identification is required for structural assessment and maintenance planning. Furthermore, the study emphasizes how important ML algorithms are for creating novel ideas, especially when considering different fracture modes. Finally, the current work concludes notable improvements in crack length prediction using SIF data and specific ML algorithms under different fracture modes.

A limitation of this study is the exclusive use of theoretical SIF data without experimental or simulation-based validation. While this approach provides a computationally efficient proof of concept, it limits direct applicability. In future work, the framework could be expanded by incorporating experimental measurements or FE-generated SIF data for training and validation. This would allow benchmarking the performance of the model under real-world conditions and enhancing its generalizability across more complex structural configurations.

Acknowledgement

This research is supported by the Structures and Materials (S&M) Research Lab of Prince Sultan University.

References

[1] Zerbst, U., Heinimann, M., Donne, C.D., Steglich, D. (2009). Fracture and damage mechanics modelling of thin-walled structures - An overview, Eng. Fract. Mech., 76(1), pp. 5–43, DOI: https://doi.org/10.1016/j.engfracmech.2007.10.005.

[2] Hosseini, A., Mahmoud, M.A. (1985). Evaluation of stress intensity factor and fatigue growth of surface cracks in tension plates, Eng. Fract. Mech., 22(6), pp. 957–974, DOI: https://doi.org/10.1016/0013-7944(85)90036-0.

[3] Wang, X., Lambert, S.B. (1996). Stress intensity factors and weight functions for longitudinal semi-elliptical surface cracks in thin pipes, Int. J. Press. Vessel. Pip., 65(1), pp. 75–87, DOI: https://doi.org/10.1016/0308-0161(94)00160-K.

[4] Meshii, T., Watanabe, K. (1999). Stress intensity factor of an arbitrarily located circumferential crack in a thin-walled cylinder with axisymmetrically loaded ends, Eng. Fract. Mech., 62(4–5), pp. 371–382, DOI: https://doi.org/10.1016/s0013-7944(98)00104-0.

[5] Rasiulis, K., Gurkšnys, K. (2010). Analyses of the stress intensity of the cylindrical tank wall at the place of the geometrical defect, J. Civ. Eng. Manag., 16(2), pp. 209–215, DOI: https://doi.org/10.3846/jcem.2010.23.

[6] O'Dowd, N.P., Nikbin, K.M., Wimpory, R.C., Lee, H.Y., Biglari, F.R. (2004). Stress intensity factors due to residual stresses in T-plate welds, Am. Soc. Mech. Eng. Press. Vessel. Pip. Div. PVP, 479, pp. 139–146, DOI: https://doi.org/10.1115/PVP2004-2658.

[7] Brighenti, R. (2005). Buckling of cracked thin-plates under tension or compression, Thin-Walled Struct., 43(2), pp. 209–224, DOI: https://doi.org/10.1016/j.tws.2004.07.006.

[8] Diamantoudis, A.T., Labeas, G.N. (2005). Stress intensity factors of semi-elliptical surface cracks in pressure vessels by global-local finite element methodology, Eng. Fract. Mech., 72(9), pp. 1299–1312,

DOI: https://doi.org/10.1016/j.engfracmech.2004.10.004.
[9] Subbaiah, A., Bollineni, R. (2020). Stress Intensity Factor of Inclined Internal Edge Crack in Cylindrical Pressure Vessel, J. Fail. Anal. Prev., 20(5), pp. 1524–1533, DOI: https://doi.org/10.1007/s11668-020-00948-0.
[10] Hachi, B.K., Rechak, S., Haboussi, M., Taghite, M., Maurice, G. (2007). Computation of stress intensity factor in cracked plates under bending in static and fatigue by a hybrid method, Int. J. Fatigue, 29(9–11), pp. 1904–1912, DOI: https://doi.org/10.1016/j.ijfatigue.2007.04.005.
[11] Treifi, M., Oyadiji, S.O., Tsang, D.K.L. (2009). Computation of the stress intensity factors of sharp notched plates by the fractal-like finite element method, Int. J. Numer. Methods Eng., 77(February), pp. 558–580, DOI: https://doi.org/10.1002/nme.
[12] Brighenti, R. (2008). A new discontinuous FE formulation for crack path prediction in brittle solids, Int. J. Solids Struct., 45(25–26), pp. 6501–6517, DOI: https://doi.org/10.1016/j.ijsolstr.2008.08.008.
[13] Qian, X. (2009). Stress-intensity factors for circular hollow section V-joints with a rack-plate chord, Fatigue Fract. Eng. Mater. Struct., 32(1), pp. 61–79, DOI: https://doi.org/10.1111/j.1460-2695.2008.01321.x.
[14] Ding, P., Wang, X. (2017). Three-dimensional mixed-mode (I and II) crack-front fields in ductile thin plates — effects of T-stress, Fatigue Fract. Eng. Mater. Struct., 40(3), pp. 349–363, DOI: https://doi.org/10.1111/ffe.12498.
[15] Aldarwish, M., Grbović, A., Kastratović, G., Sedmak, A., Lazić, M. (2018). Stress intensity factors evaluation at tips of multi-site cracks in unstiffened 2024-t3 aluminium panel using XFEM, Teh. Vjesn., 25(6), pp. 1616–1622, DOI: https://doi.org/10.17559/TV-20170309133824.
[16] Yuan, H., Yang, W., Zhang, L., Hong, T. (2023). Model Development of Stress Intensity Factor on 7057T6 Aluminum Alloy Using Extended Finite Element Method, Coatings, 13(3), DOI: https://doi.org/10.3390/coatings13030581.
[17] Anjum, A., Hrairi, M., Aabid, A., Yatim, N., Ali, M. (2025). Integrating AI and statistical methods for enhancing civil structures: current trends, practical issues and future direction, Frat. Ed Integrita Strutt., 19(71), pp. 164–181, DOI: https://doi.org/10.3221/IGF-ESIS.71.12.
[18] Kim, K.B., Yoon, D.J., Jeong, J.C., Lee, S.S. (2004). Determining the stress intensity factor of a material with an artificial neural network from acoustic emission measurements, NDT E Int., 37(6), pp. 423–429, DOI: https://doi.org/10.1016/j.ndteint.2003.08.007.
[19] Wu, Z., Hu, S., Zhou, F. (2014). Prediction of stress intensity factors in pavement cracking with neural networks based on semi-analytical FEA, Expert Syst. Appl., 41(4 PART 1), pp. 1021–1030, DOI: https://doi.org/10.1016/j.eswa.2013.07.063.
[20] Remadi, A., Ayeb, M., Bahloul, A., Bouraoui, C. (2025). Optimization of Fatigue Crack Repair Parameters for Al7075-T6, J. Mater. Eng. Perform., DOI: https://doi.org/10.1007/s11665-025-10769-7.
[21] Xia, B., Ma, Z., Hu, H., Li, Y., Zhao, W. (2022). A prediction method of stress intensity factor for mode-I crack in coal rock based on deep learning, Theor. Appl. Fract. Mech., 122, pp. 103645, DOI: https://doi.org/10.1016/J.TAFMEC.2022.103645.
[22] Zhang, X., Zhao, T., Liu, Y., Chen, Q., Wang, Z., Wang, Z. (2023). A data-driven model for predicting the mixed-mode stress intensity factors of a crack in composites, Eng. Fract. Mech., 288, pp. 109385, DOI: https://doi.org/10.1016/J.ENGFRACMECH.2023.109385.
[23] Yao, J., Xiang, J. (2024). Support vector regression-assisted finite element method for mode I-II fatigue crack growth path prediction, Theor. Appl. Fract. Mech., 131, pp. 104336, DOI: https://doi.org/https://doi.org/10.1016/j.tafmec.2024.104336.
[24] Schubnell, J., Fliegener, S., Rosenberger, J., Feth, S., Braun, M., Beiler, M., Baumgartner, J. (2025). Data-driven fatigue assessment of welded steel joints based on transfer learning, Weld. World, DOI: https://doi.org/10.1007/s40194-025-01967-x.
[25] Omar, I., Khan, M., Starr, A. (2023). Comparative Analysis of Machine Learning Models for Predicting Crack Propagation under Coupled Load and Temperature, Appl. Sci., 13(12), DOI: https://doi.org/10.3390/app13127212.
[26] Zhao, C., Wang, J., He, F., Bai, X., Shi, H., Li, J., Huang, X. (2025). A fatigue life prediction method based on multi-signal fusion deep attention residual convolutional neural network, Appl. Acoust., 235, pp. 110646, DOI: https://doi.org/https://doi.org/10.1016/j.apacoust.2025.110646.
[27] Tada, H., Paris, P.C., Irwin, G.R. (2000). The Stress Analysis of Cracks Handbook, Third Edition, DOI: https://doi.org/10.1115/1.801535.
[28] Rooke, D.P., Cartwright, D.J. (1976). Compendium of Stress Intensity Factors, London.
[29] Anjum, A., Hrairi, M., Aabid, A., Yatim, N., Ali, M. (2024). Damage detection in concrete structures with impedance data and machine learning, Bull. Polish Acad. Sci. Tech. Sci., 72(3), pp. 1–11, DOI: https://doi.org/10.24425/bpasts.2024.149178.