

ACCEPTED MANUSCRIPT

Lamb wave damage severity estimation using ensemble-based machine learning method with separate model network

To cite this article before publication: Syed Haider Rizvi *et al* 2021 *Smart Mater. Struct.* in press <https://doi.org/10.1088/1361-665X/ac2e1a>

Manuscript version: Accepted Manuscript

Accepted Manuscript is “the version of the article accepted for publication including all changes made as a result of the peer review process, and which may also include the addition to the article by IOP Publishing of a header, an article ID, a cover sheet and/or an ‘Accepted Manuscript’ watermark, but excluding any other editing, typesetting or other changes made by IOP Publishing and/or its licensors”

This Accepted Manuscript is © 2021 IOP Publishing Ltd.

During the embargo period (the 12 month period from the publication of the Version of Record of this article), the Accepted Manuscript is fully protected by copyright and cannot be reused or reposted elsewhere.

As the Version of Record of this article is going to be / has been published on a subscription basis, this Accepted Manuscript is available for reuse under a CC BY-NC-ND 3.0 licence after the 12 month embargo period.

After the embargo period, everyone is permitted to use copy and redistribute this article for non-commercial purposes only, provided that they adhere to all the terms of the licence <https://creativecommons.org/licenses/by-nc-nd/3.0>

Although reasonable endeavours have been taken to obtain all necessary permissions from third parties to include their copyrighted content within this article, their full citation and copyright line may not be present in this Accepted Manuscript version. Before using any content from this article, please refer to the Version of Record on IOPscience once published for full citation and copyright details, as permissions will likely be required. All third party content is fully copyright protected, unless specifically stated otherwise in the figure caption in the Version of Record.

View the [article online](#) for updates and enhancements.

Lamb wave damage severity estimation using ensemble-based machine learning method with separate model network

Syed Haider M. Rizvi^{1*}, and Muntazir Abbas^{1,2}

¹ Corresponding author, department of Engineering Sciences, PN Engineering College, National University of Science & Technology, Pakistan
² SWEE, Cranfield University, College Road, Cranfield, Bedfordshire, United Kingdom

1. Abstract:

Lamb wave-based damage estimation have great potential for structural health monitoring. However, designing a generalizable model that predicts accurate and reliable damage quantification result is still a practice challenge due to complex behavior of waves with different damage severities. In the recent years, machine learning algorithms have been proven to be an efficient tool to analyze damage-modulated Lamb wave signals. In this study, ensemble-based machine learning algorithms are employed to develop a generalizable crack quantification model for thin metallic plates. For this, the scattering of Lamb wave signals due to different configuration of crack dimension and orientation is extensively studied. Various finite element simulations signals, representing distinct crack severities in term of crack length, penetration and orientation are acquired. Realizing that both temporal and spectral information of signal is extremely important to damage quantification, three time-frequency based damage sensitive indices namely energy concentration, time-frequency flux and coefficient of energy variance are proposed. These damage features are extracted by employing smoothed-pseudo Wigner-Ville distribution. After that data augmentation technique based on the spline-based interpolation is applied to enhance the size of the dataset. Eventually, these fully developed damage dataset is deployed to train ensemble-based models. Here we propose separate model network (SMN), in which different models are trained and then link together to predict new and unseen datasets. The performance of the proposed model is demonstrated by two cases: first simulated data incorporated with high artificial noises are employed to test the model and in the second scenario, experimental data in raw form are used. Results indicate that the proposed model has the potential to develop a general model that yields reliable answer for crack quantification.

Keyword: machine learning; ensemble-based machine learning algorithm; bagging random forest algorithm; structural health monitoring; guided waves; Lamb waves; structural damage estimation.

2. Introduction:

Damage or corrosion-induced material degradation is the major cause of the catastrophic failure of civil and mechanical infrastructures [1, 2]. The application of structural health monitoring (SHM) is therefore an area of growing interest in the engineering community. For SHM, using ultrasonic guided waves is widely acknowledged as a reliable and effective approach for damage identification in recent years. Its robust and fast-growing technique requires affordable instrumentation, which makes it more attractive nowadays. Lamb waves, a type of guided wave travel with little attenuation, cover a large area with minimum tool dependencies [3, 4] and hence is widely used for the inspections of large structures like long pipelines and rails with minimum use of transducers. The waves are highly sensitive to abnormalities, and inhomogeneity near its propagation path [4]. Recent development in sensor technology, diagnostics, informatics, signal processing, Engineering mechanics and material sciences, SHM based upon the guided waves have been the subject of intensive research.

Several studies proposed different techniques, like tomography and probabilistic algorithm for structural damage identification and localization in both homogenous and composite materials. The technique is based on the

spatially distributed actuator array. When guided waves are excited in a structure through an actuator, wave propagates and interacts with structural defects. The variation in wavefield is recorded by sensors placed at different locations on the surface of the material. These damage-modulated signals are compared with the baseline signal to detect, localize, and quantify the defect [5, 6]. Much research has been carried out associated with this physics-based signal processing methods [5][7][8]. The major shortcoming of physics-based method is its reliance primarily on physical laws, which govern the structural behavior, to extract the meaningful data about the structural damages [9]. However, due to the complex underlying mechanism of wave scattering with structural defects, it could be highly non-trivial to set up a physics-based general model. For instance, the behavior of Lamb wave is almost stochastic, with different severities of the damage and the only possible way of dealing is to consider them in a case-by-case manner [10]. Another possible way is to employ machine learning algorithms and artificial intelligence methods that use extensive data mining methods as pre-processing and information fusion to form a general model.

In the last two decades, researchers have extensively employed Machine Learning (ML) algorithms to develop a broad range of GW based SHM techniques with high accuracy. Su and Ye used artificial neural network (ANN) to perform quantitative identification of delamination in composite structures [11]. Agarwal and Mitra compared ANN and Support Vector Machine (SVM) algorithm by detecting damage in a metallic plate. They used Matching Pursuit (MP) to denoise and increase the sparsity of datasets[12]. Similarly, Hossein et al. used multiclass SVM to classify and estimate severity of cracks [13]. Yang exploited Bayesian learning method to accurately predict the length of the crack [14] while Cofre-Martel et al. proposed deep neural network-based damage localization and quantification method using transmissibility functions [15]. Recently, Sun et al. proposed a damage quantification model using Genetic Algorithm (GA) and least square SVM [16]. Recently, Rai and Mitra used 1-D multi-headed convolution neural network to detect damages in metallic plates. The model is trained by wide range of database obtained by changing parameters like scanning length and frequency and adding different level of white noises. The classification model achieves 100% accuracy in identifying damages for unseen data [17]. Similarly, Ewald et al. utilized deep neural network for Lamb wave signal classification and the proposed scheme achieved high accuracy in classifying pristine and different damage cases [18]. However, the generalization to different damage severities were not investigated in these research works. One recently published research by Zhang *et al.*, tried to bridge this gap, used diverse dataset of 17 damage state in terms of damage type, size, and orientation to train SVM classification model. The model achieved great accuracy when considering single feature to classify. The paper briefly addressed the mixed data types and achieved maximum accuracy of 89.54% on it [10].

In this study, a Lamb wave-based damage quantification model is presented by using ensemble-based machine learning algorithm for metallic plate. A comprehensive learning framework in which different damage severities like length, penetration, and orientation are simultaneously considered, to obtain a model generality that quantifies the structural damage. In this work, we propose separate model network (SMN) that links two separate ensemble-based models to predict the required label. Overall, the evaluation of accuracy and performance of proposed scheme is done by three generalization levels: 1) the model is tested by the simulation result incorporated with high white noises, 2) testing by the experimental data with same damage severities and orientations used in model training, 3) testing by experimental data with different damage severities and orientations used in model training. The only shortcoming of the proposed ML scheme is the necessity of baseline signal to compare with damage signal. However, this problem might be resolved by adopting time reversal method [19–23]. However, in this study, baseline signals are obtained separately by Finite Element Method (FEM). The layout of this paper is as follow: the methodology is presented in section 3, including the detail of machine learning method, mode selection, feature extraction, data augmentation techniques and SMN. Section 4 describes the

numerical and experimental setups. Result and discussion on the performance and accuracy of the proposed method are presented in section 5. Conclusion is drawn in section 6.

3. Methodology:

The objective of this paper is to present a reliable method to find the severity of a crack by exploiting a powerful machine learning algorithm. Although much research on the damage assessment has been reported recently, [14, 16, 24, 25]. These studies estimated temporal or spectral damage features by targeting only a single dimension of the crack, while other dimensions are kept constant. In a realistic situation, damage features are heavily influenced by other dimensions of the crack. Similarly, in most literature, [14, 16, 26, 27] the arrangement of actuator and sensor is usually perpendicular to the line of the crack. However, crack initiation and its propagation are stochastic and do not always favor the sensors' arrangement. Considering a fixed crack orientation may introduce uncertainties in damage assessment. Therefore, crack orientation is an important factor that needs to be considered in damage assessment.

In this paper, intensive studies have been carried out to find the relationship between scattering of damage signals and different severities of the crack, i.e. length, depth/penetration, and orientation. For this purpose, simulated data is obtained by a pitch-catch configuration, in which a single actuator is placed on one side of the cracks while different sensors aligned in a line on the other side. Each sensor makes a different angle of incidence, which is equivalent to a different orientation of the crack. Figure 1 explains the geometrical equivalency of the two scenarios. Moreover, multi-modal and dispersion are inherent properties of lamb waves that depend on the excitation, frequency, and thickness of the material under investigation. This dispersion behavior leads to a complex waveform structure. Therefore, in order to extract the damage information, the dispersion phenomenon must be reduced. From Figure 2, it is observed that in the low-frequency region, the S_0 mode shows significantly non-dispersive behavior, propagating at high speed and very low attenuation rate. In this study, we will consider S_0 mode, generated at 200 kHz frequency on 1.6mm-thick aluminum plates.

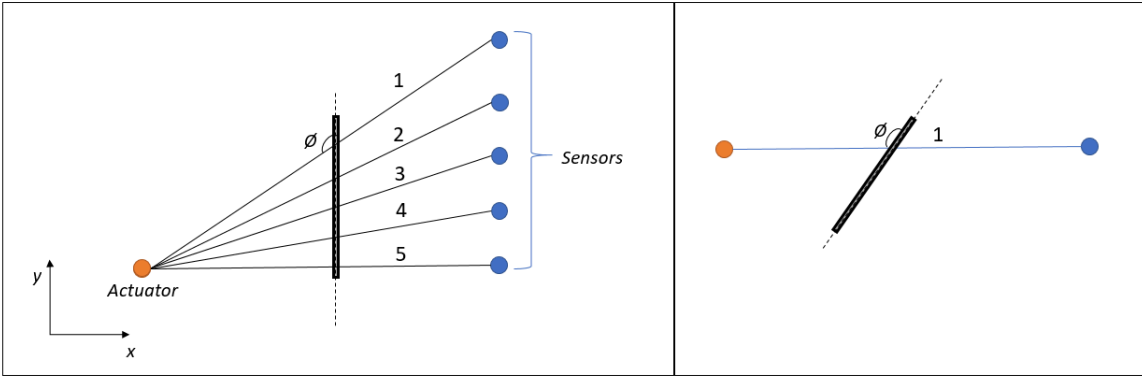


Figure 1: Schematic representation of crack orientation

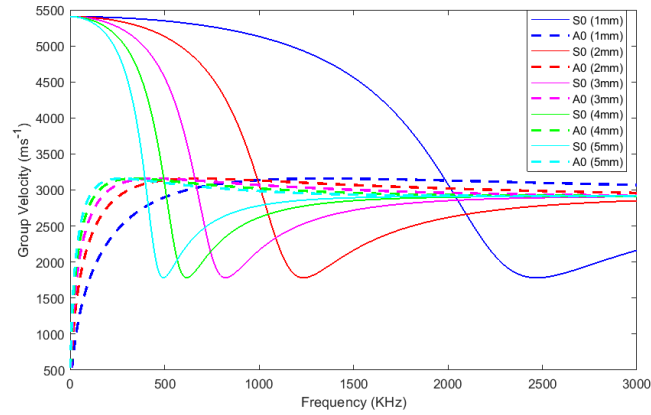


Figure 2: Dispersion curve of S_0 and A_0 mode Lamb wave at various material thickness

3.1 Machine Learning Approach

The ensemble method is a set of machine learning techniques that create multiple models, and their decisions are combined in this way to improve the performance of the overall system. The intuitive concept of ensemble learning is based on the belief that no machine learning approach can claim to be consistently superior to any other strategy and the only appropriate combination of several single methods may improve the performance (accuracy, reliability, and comprehensibility) of the final classifier. Empirical studies have shown that ensemble-based machine learning schemes have much better accuracy than the individual base learner, also referred to as weak learner. However, the performance and effectiveness of the ensemble method are highly dependent on the reliability and diversity of the base learner [28]. In terms of base learner, ensemble learning is classified into two types,

Homogenous Ensemble: It consists of the same sets of base learning algorithms. Popular method is Bagging (Bootstrap aggregating) in which each base learner is trained by a subset of the initial training sets. Bootstrap sampling is used to sample the training sets of each learner i.e., randomly selecting N (size of dataset) out of N observation with replacement. The sampling method also omits some observation, on average 37%, for each decision tree which is “out-of-bag” observation [29]. Figure 3 represents the bagging algorithm. Bagging mitigates variance in high variance low bias datasets and is particularly effective with unstable base-learner [30]. The targeted output is estimated by voting or averaging results of base-learners.

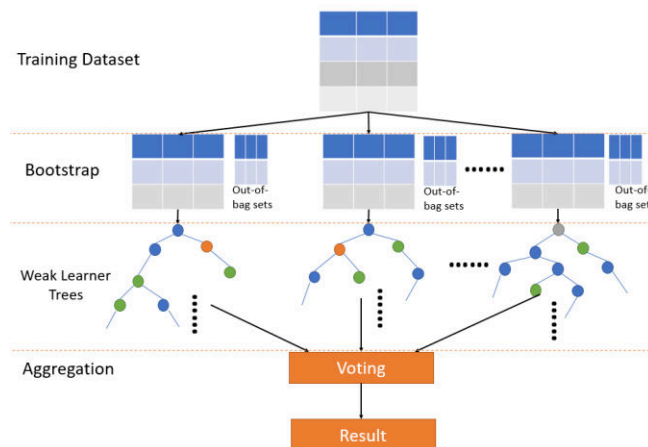


Figure 3: Schematic representation of ensemble based random forest model

Heterogeneous Ensemble: It is formed by utilizing different learning algorithms. The better predictive accuracy is proportional to different model characteristics, which is the basic idea behind heterogenous ensemble. Combining different learning algorithm improves results as compared to a single base learner [31]. Since each algorithm has an entirely different prediction approach and therefore has different result accuracy. Here, soft voting or averaging may lead to inaccurate results because it requires each learner to contribute equally. In this situation, accuracies generated by different learners are carefully calibrated by assigning appropriate weights so that each learner contributes to results in accordance with their estimated performance [30].

Following are the features and limitations of ensemble learning:

1. Size and nature of the dataset: ensemble learning works efficiently even when just small amount of data relative to space of hypothesis is available for training. Combining different learners significantly reduces the risk of selecting incorrect hypothesis [32]. Similarly, for highly skewed data, ensemble learning significantly mitigates the effects of such class-imbalance problem, especially bagging in which each learner is trained using balanced subsample of the data [33]
2. Computational advantage: singleton learner performs local search and may stuck in local optima. Combining several learners can solve this problem [32].
3. Representation: individual learner may have limited space to search for optimal hypothesis. Combining multiple learners extend the search space and generate better predicted results [32].
4. Computational cost: it increases the computational cost and complexities as multiple models need to be trained [34].
5. Interpretation: as compared to singleton model, sometimes it is very difficult to interpret the decision of ensemble model and extract meaningful result [35] .

3.2 Damage Feature Extraction:

It is a technique of identifying the damage-modulated properties and parameters in a received signal [36]. Precise measurement of features plays a vital role in the damage quantification process. Typically, separate temporal or spectral domains of a signal are exploited to find the damage-modulated features for SHM applications. However, lamb wave signals have non-stationary attributes whose spectra vary with time, so separate domain features may lack sufficient discriminating information about the acquired signal. In this study, time-frequency distribution (TFD) is employed to measure features of the arrival signals. TFDs represent the energy of the signal in joint time-frequency domain, which is useful for analyzing the signal variation in both axes. TFD provides additional information about the non-stationary signal that cannot be directly obtained by a separate domain [37]. Therefore, TFD is the powerful analysis tool for non-stationary signals. [36].

In the literatures, there are different time-frequency (TF) analysis methods are used like Short-Term Fourier Transform (STFT) [38], Wavelet Transform (WT) [39] and Wigner-Ville distribution (WVD) [6]. The most widely used TF analysis methods belong to Cohen class like Wigner-Ville distribution which is bilinear TF distribution. The WVD represents the energy distribution in both the time (t) and frequency domain (ω). Mathematically, WVD is the Fourier Transform of the instantaneous auto-correlation sequence with time lag (τ). It is defined as,

$$W_x(t, \omega) = \int_{-\infty}^{+\infty} z\left(t + \frac{\tau}{2}\right) z^*\left(t - \frac{\tau}{2}\right) e^{-i\omega\tau} d\tau \quad (1)$$

Where, $Z(*)$ is the analytical signal of the original signal obtained by Hilbert transform. The WVD contains interference or cross-terms that often provide misleading information and hence make the interpretation more complicated. Therefore, to reduce these cross-terms, the smoothed pseudo-Wigner-Ville distribution (SPWVD) is

employed. It uses an independent window to sharpen the distribution. Let g and H are two uniform positive windows, the smoothed pseudo-WVD is given as,

$$W_x(t, \omega) = \int_{-\infty}^{+\infty} g(t)H(\omega)z\left(t + \frac{\tau}{2}\right)z^*\left(t - \frac{\tau}{2}\right)e^{-i\omega\tau}d\tau \quad (2)$$

SPWVD has two inherent properties; first is the anti-noise property as it can be observed as filtered energy passing through extremely narrow passband so the effect of background noises and interference in Lamb wave signal can be reduced [6]; the second property is that, unlike the linear TF distributions which only express the approximate energy distribution in the TF domain, SPWVD represents the true time-frequency energy distribution of the signal which can precisely estimate the damage features [37]. In this paper, three damage features are extracted by joint time-frequency SPWVD to estimate the variation of the signals to quantify crack severity. These are,

- Time-frequency flux (TFF): It is the rate of change of energy content of a signal in time-frequency domain. It is expressed as,

$$TFF(x) = \sum_{t=1}^{T-l} \sum_{\omega=1}^{F-q} |W(t+l, \omega+q) - W(t, \omega)| \quad (3)$$

Where $W(*)$ represents the TFD of size $N \times M$ while, l and q are the predefined values which are correlated with the rate of change of the signal energy. The value can be any integer value from 0 to $T-1$ and 0 to $F-1$ which represents the size of the window over which the rate of change in TFR is determined. Small value tends to consider and magnify small variation in TFR, while large value may neglect the meaningful changes [40]. Since in this study, all the features are compared with baseline signals, therefore finding a small variation in TFR is desirable. Therefore, in this paper, l and q are taken as 1 to estimate TFF. It can compute small energy variation of the signal take place either in temporal or spectral axis [41].

- Energy concentration (EC); it characterizes the energy distribution over the time-frequency plane. Its mathematical expression is expressed as

$$EC(x) = \left(\sum_{t=1}^T \sum_{\omega=1}^F |W(t, \omega)| \right)^2 \quad (4)$$

- Coefficient of variation (CV): it is defined as ratio of standard deviation to the mean. It is an important tool to measure of the relative dispersion of data around its mean ($\mu_{(t,\omega)}$).

$$CV(x) = \frac{\frac{1}{TF} \sqrt{\sum_{t=1}^T \sum_{\omega=1}^F |W(t, \omega) - \mu_{(t,\omega)}|}}{\mu_{(t,\omega)}} \quad (5)$$

3.3 Data augmentation technique:

To improve the performance of machine learning algorithms and to make the model more stable for multiple unseen predictions, a large amount of training datasets is required, which is not always available. Therefore, data augmentation is a powerful scheme to virtually enlarge the training datasets. This scheme is popular in the field of computer vision, in which image data is used to train the model. However, unlike image data, time-series data is vulnerable to label changes, especially in regression cases and therefore image data augmentations methodology cannot be applied to time-series data. To cope with this problem, recently some researchers have proposed

different data enhancement schemes for time-series data [42–44]. Cheolhwan recently proposed a data augmentation scheme based on interpolation, which is more efficient against the impairment of trend information of time-series data [45]. In this paper, we use the same time-series data augmentation technique in which a virtual sensor is considered between the two true sensors paired. This is shown in Figure 4. Same technique is applied for different damage sizes. Damages indices associated with a virtual sensor and damage size are estimated by the spline-based linear interpolation method. It is important to mention that the intuition of data augmentation is to diversify the input data by estimating the unknown values that lie between the known observations.

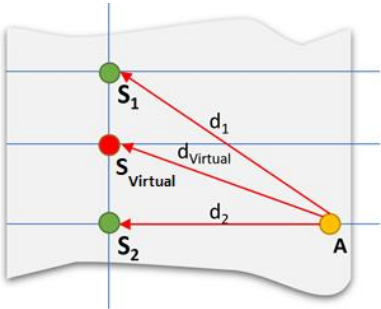


Figure 4: Data augmentation technique: where A and S_1 , S_2 represent the actuator and real sensors respectively, while S_{Virtual} is an assumed sensor. Here, all d s show the Euclidian distances.

3.4 Separate Model Network (SMN):

Since different damage attributes are simultaneously considered in estimating damage indices, therefore, it is likely that data distribution involves significantly ambiguous or overlapping structures and It would be difficult for a single regression model to predict accurate results. To alleviate this problem, we propose a separate model network (SMN), in which the entire dataset is decomposed into various sub-classes or labels and then multiple baseclassifiers are simultaneously employed to train an ensemble model. Here, labels are set according to the penetration levels of crack, therefore instead of clustering, supervised ML algorithms are used to classify the datasets. The output of the base-classifier will become the valuable input for the meta-learners along with the original dataset, that would help to improve the efficiency and reliability of the final result. Moreover, here ensemble learning also helps to mitigate the effect of imbalanced datasets that has a strong correlation with overlapping data [46].

To implement SMN, first, an ensemble model with five strong base learners/classifiers namely Random Forest (RF)/Decision Tree (DT), Support Vector Machine (SVM), Artificial Neural Network (ANN), K-Nearest Neighbors (KNN), and Gaussian Naïve Bayes (GNB), are trained on all available data with label 1 to 4 representing damage penetration level. We call them damage class. Since each learner has a different learning approach, therefore, they predict output with different accuracy. This prediction accuracy plays a vital role in the consensus process and so consensus result may be significantly affected by the low accuracy-based classifier. To overcome this problem and enhance the consensus performance, most researchers use majority voting and weight voting schemes. The majority voting is based on the implicit belief that all classifiers have the same accuracy and impact on the result. This scheme is applicable in the case of the bagging and boosting ensemble approach. In heterogeneous ensembles, instead of voting, weights must be assigned to all individual learners. Weights are usually calculated by a performance evaluation scheme, in which validation accuracy of all base learners is determined and based on them, weights are assigned to each learner. K-fold cross-validation (CV) is a powerful technique used as a fitness function that allows the model to be tested K times by the means of resampling and helps to protect the model against

overfitting. Therefore, in this paper, we use K-fold CV to estimate the accuracy of the model and based on them weights of learners are computed. These weights are then used to predict the result of the damage class. Figure 5 represents the architecture of all five learners. After that, meta-learners/ secondary-level ensemble-based regression learner is trained to estimate the crack size. Here we use the term meta-learner because crack size is predicted by taking the output of first-level learners as input.

Moreover, it is important to note that here the sensitive signal features are exploited to predict the length and penetration level of the crack. These features are extracted by changing crack dimensions (length and penetration) and orientations not in a continuous manner, but a discrete one. Therefore, it is likely that the model may slightly deviate from true values and generate a bit higher error value. To estimate the performance of the proposed algorithm, here we use three parameters; prediction accuracy (PA), root mean square error (RMSE), and mean absolute error (MAE).

$$PA = 1 - \frac{\sum |actual\ response - predicted\ response|}{actual\ response} \quad (6)$$

$$RMSE = \sqrt{\frac{\sum (actual\ response - predicted\ response)^2}{N}} \quad (7)$$

$$MAE = \frac{\sum |actual\ response - predicted\ response|}{N} \quad (8)$$

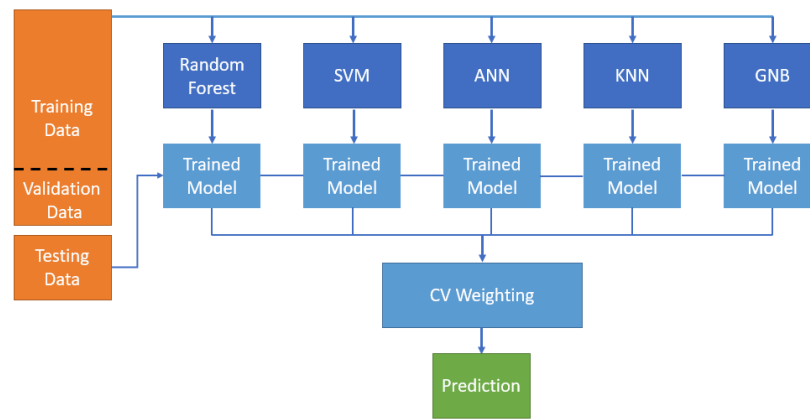


Figure 5: Architecture of all five-basic classifier.

PA simply refers to how the closeness or conformity of a predicted value to actual value. On the other hand, RMSE and MAE summarize the mean difference between actual and predicted values. RMSE is highly sensitive to outliers and their probability of occurrence due to quadratic scoring while MAE is a linear score which weight all the differences equally [47]. Therefore, both parameters give useful information about the model prediction. Figure 6 summarizes the steps of SMN algorithm while Figure 7 represents the overall methodology of the crack estimation method used in this paper.

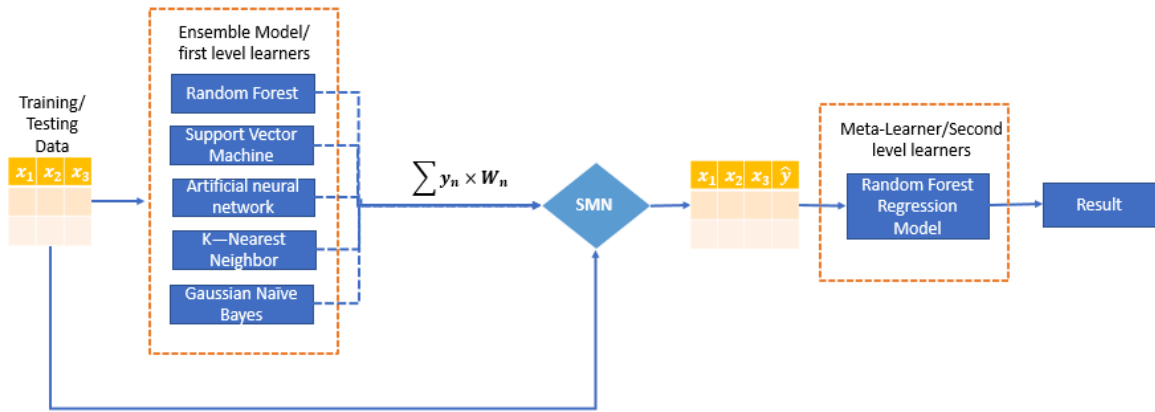


Figure 6: Schematic diagram of Separate Model Network

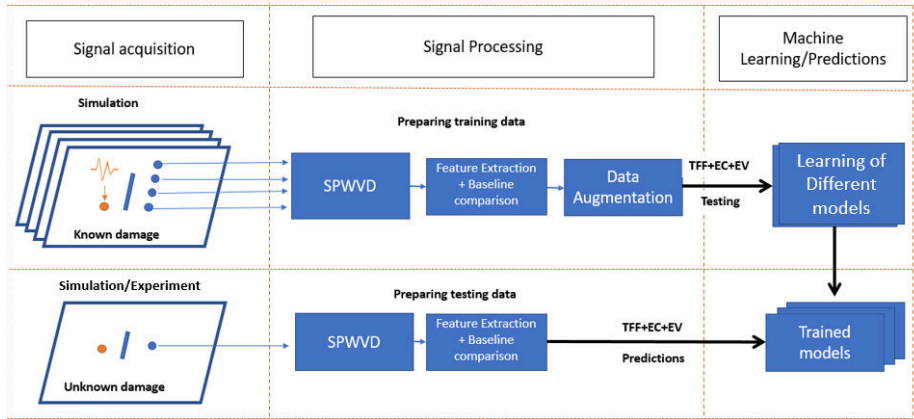


Figure 7: Overall framework of damage quantification using machine learning model.

4. Preparing data for model training, validation, and testing

4.1 Training and validation datasets:

To obtain the training and validation datasets, numerical simulation on 1.6mm-thick aluminum alloy with rectangular cracks of different lengths, depths, and orientations is performed. Mechanical properties of the alloy are given in Table 1. The lamb wave is excited by a five-cycle Gaussian-modulated tone burst with a central frequency of 200kHz. These waves, after encountering cracks, are received at different points depicting different excitation angles. Four different types of damage classes are designed, including mild damage, medium damage, high damage, and severe damage. These types are categorized based on the depth/penetration of cracks. This is summarized in Table 2. We consider crack depths from 0.3mm to 1.3mm and 1.6mm (through-thickness case). Beyond these range not only surge the computational cost but also increase the complexities of signal interpretation. However, features between 1.3mm and 1.6mm are computed by spline interpolation technique. All the damages are located at the center of the plate, while the minimum distance between the actuator and sensor is 150mm. This ensures the complete separation of S_0 mode from the other wave modes. In addition, each damage class has different crack sizes, ranging from 5mm to 70mm. The orientation of the cracks is kept between 0° and 12° while the width of the crack is fixed to 2mm for all cases.

Table 1: Mechanical properties of Aluminum alloy (6061-T6)

Density (kgm^{-3})	Poisson's ratio	Young's Modulus (GPa)	Yield Stress (MPa)
2780	0.33	72.4	276

The numerical simulation is performed by using a commercial software Abaqus/CAE Explicit module. The accuracy of numerical results can be improved by selecting the smallest step size of the solution, [48]. High time step generates high-frequency components that may not be resolved accurately. Therefore, Moser [49] determined that the ratio between the integration time step and the maximum frequency of the exciting pulse is a critical value and must be evaluated carefully. Moser suggested that the integration time step must be minimum, i.e. at least 20 points per cycle at the maximum frequency. Therefore, the step size of $0.1\mu\text{s}$ is selected. On the other hand, mesh density is also crucial for good spatial resolution. In this study, 1mm as the approximate global size of an element with linear hexahedral element of type C3D8R is used. It ensures that there are approximately ten nodes across the shortest wavelength. However, near the crack area minimum 0.25mm is taken as the size of the element. Moreover, free-free boundary condition is considered during simulation. Figure 8 represents the finite element non-uniform mesh and cross-section of the crack geometry.

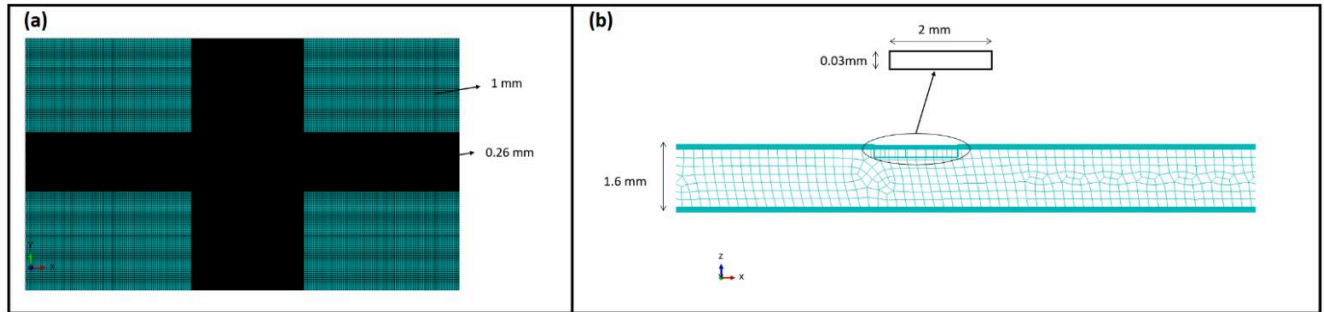


Figure 8: Finite Element Mesh: a. Non-uniform Meshes, b. Wireframe cross-sectional view of the crack.

Table 2 Damage cases based on crack depth

Damage Type	depth of the crack (d)	Maximum percentage of crack penetration	Number of samples	
Mild damage	0<d≤0.5mm	31.3%	FEM =	96
			Augmentation =	237
			Total =	333
Medium damage	0.5<d≤1mm	62.5%	FEM =	192
			Augmentation =	480
			Total =	672
High damage	1<d≤1.2mm	75%	FEM =	96
			Augmentation =	237
			Total =	333
Severe damage	1.2<d≤1.6mm	100%	FEM =	96
			Augmentation =	237
			Total =	333
			Overall =	1677

4.2 Testing data:

Trained models are tested by datasets obtained by simulation and experiments. The simulated dataset is prepared by adding white Gaussian noise with a very low SNR value to signals. These signals are obtained by the same numerical simulation setup discussed in the previous section. For the experimental dataset, experiments are performed on a rectangular aluminum plate to generate and detect the Lamb wave. Plates are made of aluminum alloy of grade 6061-T6. The mechanical properties and dimensions of all plates are the same as used in numerical analysis. Small circular piezoelectric wafer active transducers (10mm dia. and 1mm-thick) are used to generate and detect Lamb waves signals. Cyanoacrylate adhesive is used to fix the PWAS on the plate surface, while a rectangular crack is introduced by the pantograph machine to obtain high accuracy. The complete model is given in Figure 9. The distance between sensors is 150mm while the crack is placed between the exciter and sensor. A Gaussian-modulated pulse with a central frequency of 200kHz frequency is transmitted to a GW Instek AFG-2225 function generator to trigger the exciting PWAS. The Lamb wave generated and propagated along the boundaries of the plate. When the wave encountered the crack present on its path, some parts of the wave reflected while the rest continued to travel and captured by the sensing PWAS. The transmitted signal is obtained by a GW Instek GDS-1052U digital oscilloscope, which is sent to the computer for signal analysis.

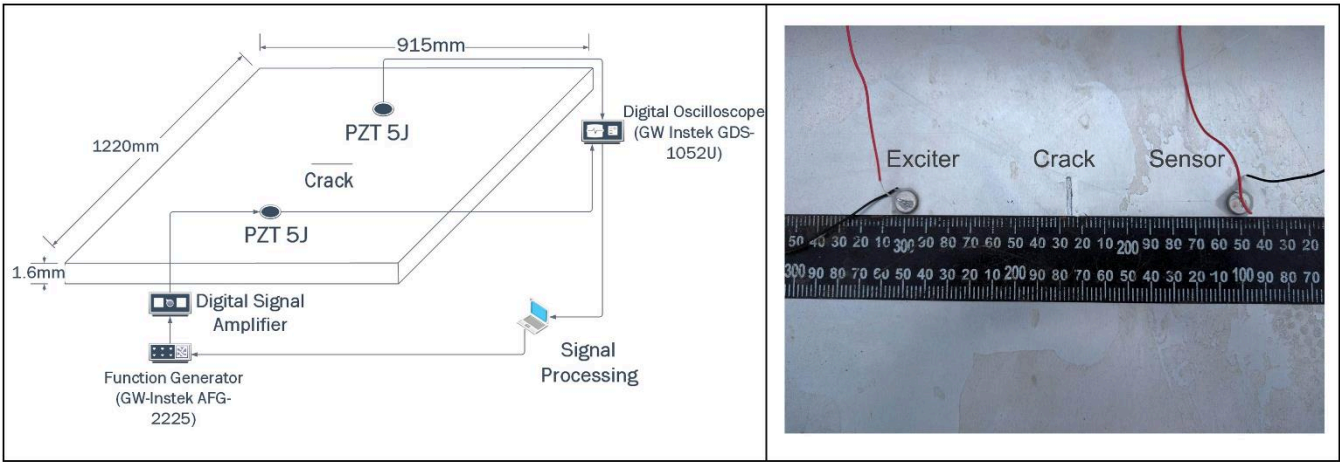


Figure 9: Experimental Setup for wave generation using pitch-catch configuration.

5. Results and Discussion:

5.1 Model training:

The time-domain signals obtained by the numerical simulations are first converted into the time-frequency scale by using SPWVD. After that, three damage-sensitive features from time-frequency distribution namely time-frequency flux, energy concentration, and coefficient of energy variation are estimated for all damage sets. Figure 10(a) represents the wave propagation in metallic structure. In the figure actuator and sensors are represented in circle for the illustrative purpose. In reality, point-force method is used to excite and sense the wave. In Figure 10(b), SPWVD of the damage signal is presented. After that, the indices are then normalized by dividing them with baseline (non-damage samples). To avoid model overfitting and to make it more robust and stable, a data augmentation scheme is applied using the spline-based linear interpolation method. It enhances the training datasets and overall, 1677 x 3 samples are collected as training. Figure 11 represents the three damage indices estimated by signals received at different severities of the cracks. It shows that the crack penetration significantly changes the values of the damage indices and therefore must be considered for structural damage estimation.

After that, these datasets are labeled according to the scheme discussed in section 4.1 (Table 2) and then fed for model training using the heterogeneous ensemble classification learning model. In the first scenario, five models; RF, SVM, ANN, KNN, and GNB are trained to classify the damage in four different classes. In order to estimate the performance of the classifier, 10-fold cross-validation has been carried out. Figure 12 is a confusion matrix that represents the validation accuracy of all five learners. The diagonal values represent the percentage of true positive rate (TPR), while the non-diagonals are the false-negative rate (FNR). Table 3 represents the configuration parameters of all five-classifier estimated by the Bayesian optimization with some constraints. These classifiers are trained by MATLAB® classification learner app [50]. Table 4 represents the 10-fold cross-validation accuracy of all classifiers, which shows that ANN outperforms with 97.61% accuracy. After that, these crossvalidation accuracies are used to assign weights to each classifier for future prediction.

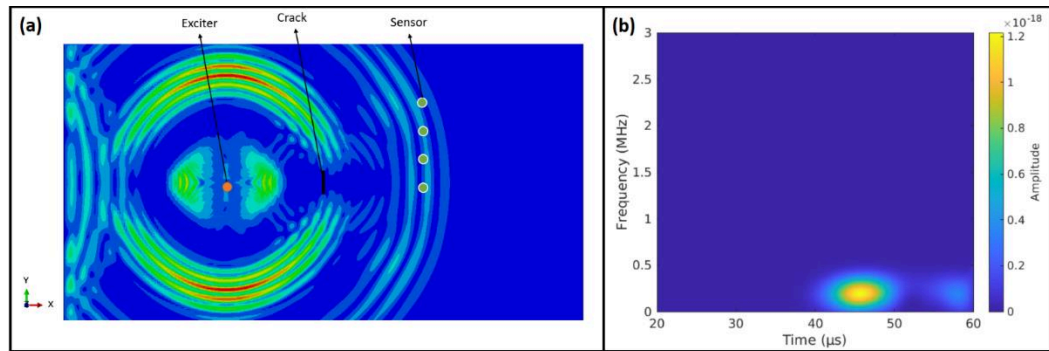


Figure 10: a. FEM model representing wave propagation, b. SPWVD of a simulated signal.

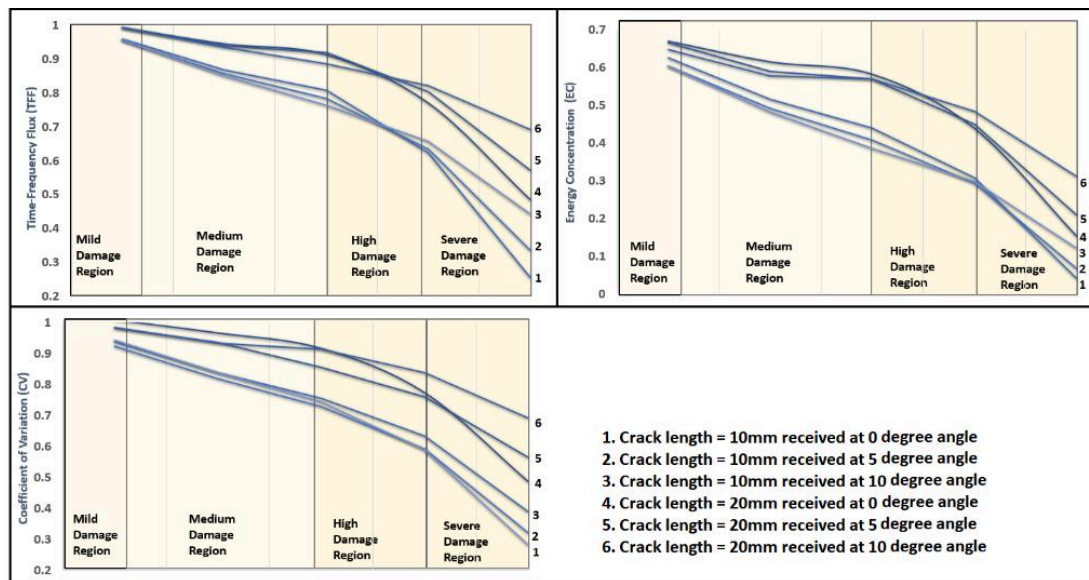


Figure 11: Three damage indices in four different damage regions

Now, the next step is to train the regression model by using three predictors along with their respective labels used in the first-level learning model. Bagging is employed with a random forest sampling technique that combines the prediction of several base learners (decision tree) to form a final result. Here, the tuning parameters for the learning process are estimated by using the Bayesian optimization algorithm. Figure 13 represents the hyperparameter optimization of the bagging ensemble model, in which a maximum of 30 iterations is employed to compute the

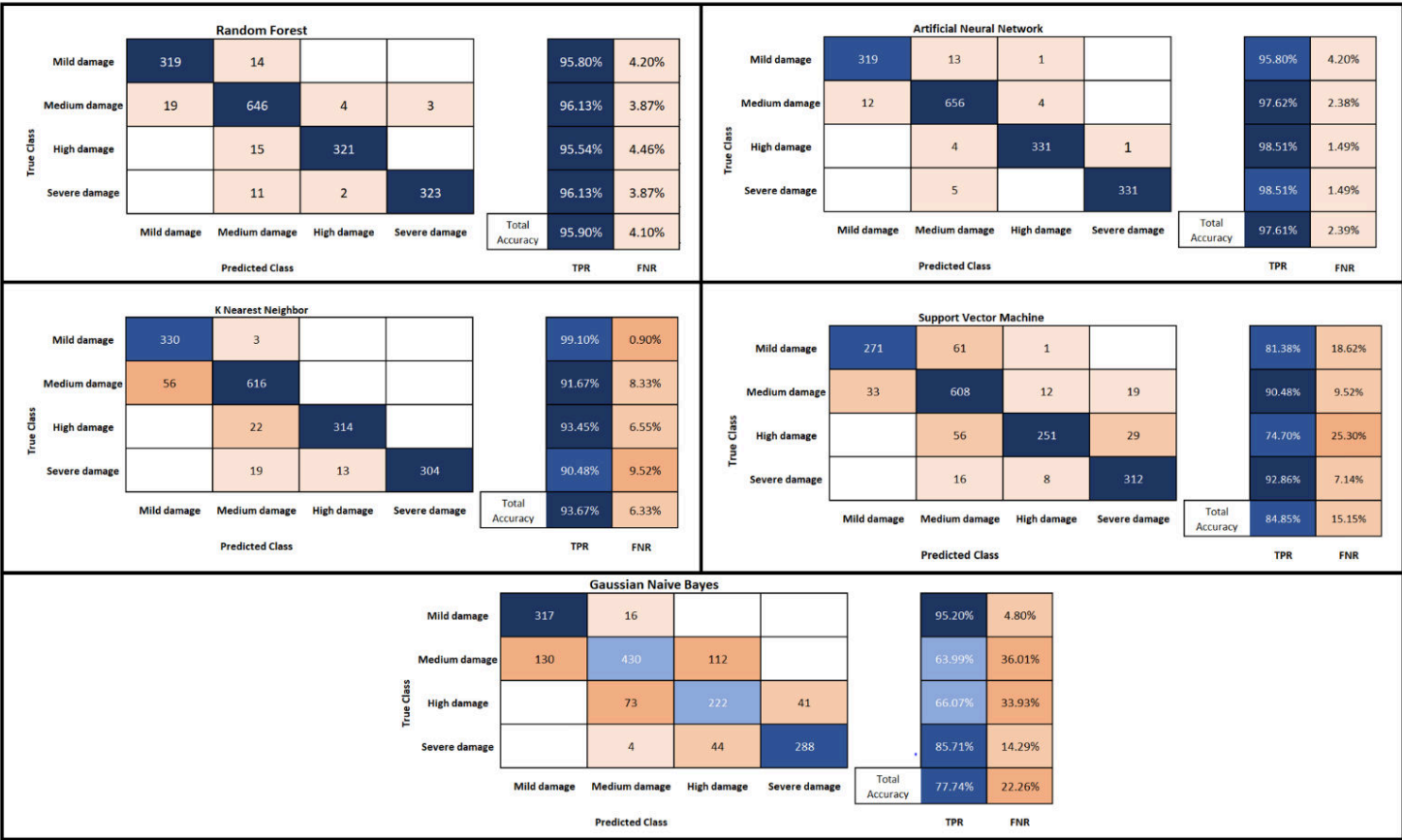


Figure 12: Confusion matrix that represents the validation accuracy of all five classifiers

tuning parameters. Figure 14 represents the prediction results using the validation dataset while Table 5 summarizes the performance of the trained model with and without SMN. RMSE and MAE are the important parameters that measure how precisely the model predicts the response. The values are in an acceptable range for such a volatile set of data.

Table 3: Configuration of all five-learners

Classifier	Configurations	Values
Random Forest	Method	Bagging*
	Learner Type	Decision Tree*
	Number of Learners	30
	Loss function	Classiferror (Misclassified rate in decimal)
	Maximum Number of Split	1676
Artificial Neural Network	Network Type	Feedforward
	Number of Hidden Layers	1
	Size of the hidden layer/neurons	25
	Activation function	Rectified Linear Unit (ReLU)
	Output layer activation function	Softmax
	Maximum Iteration	1000*
K-Nearest Neighbor	Validation patience	6*
	Loss function	Classiferror (Misclassified rate in decimal)

	Method	Kd tree
	Number of neighbors	10*
	Size of bucket	50
	Distance Metric	Euclidian
	Distance weight	Equal
	Standardize data	True
Support Vector Machine	Kernel	Linear*
	Box constraint level	1
	Method	One-vs-One
	Loss function	Hinge
	Standardize data	True
Gaussian Naïve Bayes	Kernel type	Gaussian
	Loss Function	ClassifError

(*) = constraint/unoptimized attribute; All classifiers' learning rate is equal to 1.

Table 4: K-fold cross validation accuracy of all classifiers

Classifier	K-fold CV
Random Forest	0.959
SVM	0.848
ANN	0.976
KNN	0.936
GNB	0.777

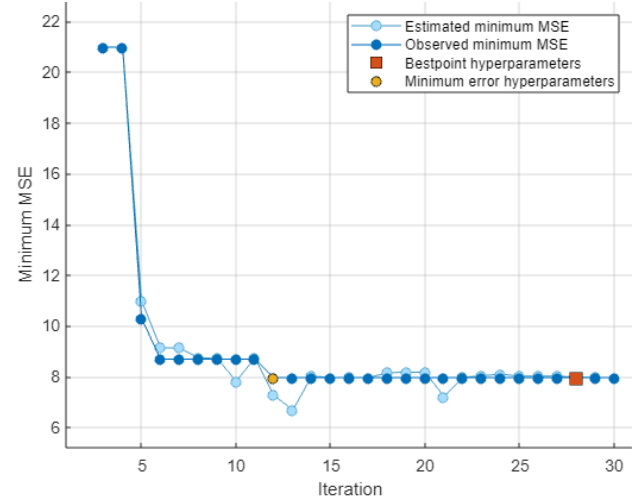


Figure 13: Bagging Model tuning using Bayesian optimization.

Table 5: Statistical summary of ensemble regression model

Learning Algorithm	R-square	RMSE	MAE	Out-of-bag Loss (RMSE)
Ensemble with SMN (Bagging)	0.97	2.8309	1.7365	2.764
Ensemble without SMN	0.73	7.8621	5.2572	7.775

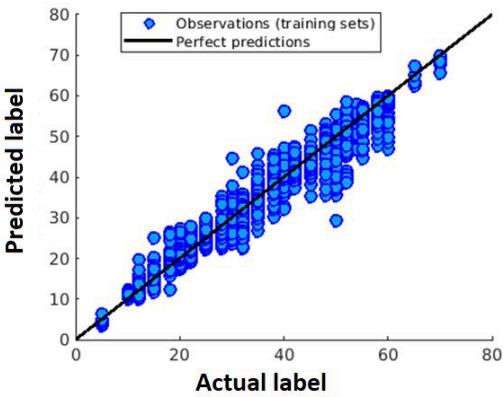


Figure 14: Actual vs predicted label (damage size) of ensemble regression model (Meta-learner) using SMN.

5.2 Model testing:

To check the validity of the trained models, two cases are considered. First, simulated signals with low SNR values are exploited to test the model. These simulated data were never used during model training or validation. However, all the damage attributes (damage size, penetration level, and orientation) are the same as used in model training. In the second scenario, experimental data are fed into the model to check its credibility.

- Simulated results:

To imitate the environmental noises, which is sometimes inevitable during the experimental measurement, different SNR values ranging from 5dB to 20 dB are achieved by adding white gaussian noises to the simulated signal. Due to the very low SNR values, it would be a great challenge for the model to predict the class and size of the crack. Figure 15 shows simulated signals at the aforementioned SNR values.

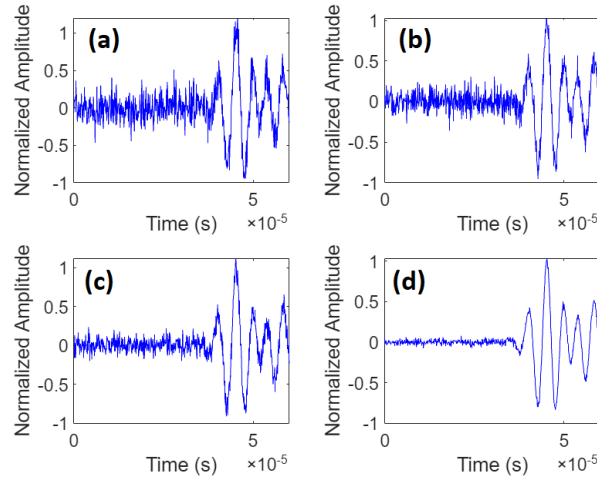


Figure 15: Damage signals with different SNR values a. 5dB, b. 8dB, c. 10db, d. 20dB

The damage features are extracted and then fed into the trained models. Figure 16 presents the model prediction results of all four regions using the proposed random forest ensemble model and SMN. The straight line shows the median prediction, while hollow dots represent the prediction results of the test data. Table 6 summarizes the overall prediction efficiencies of four regions. It is important to mention here that some data of different cracks deviate slightly from the general trend and yield inaccurate results when dealing with such low SNR values. For instance, some values of 70mm cracks in mild damage and 50mm in severe damage region are outliers that generate higher error values. The average RMSE is found to be 4.51mm, while the MAE value is 3.9mm. However, these errors significantly improve when the SNR value is increased. For example, between 21db to 30db, RMSE and MAE will become 2.54mm and 2.24mm, respectively.

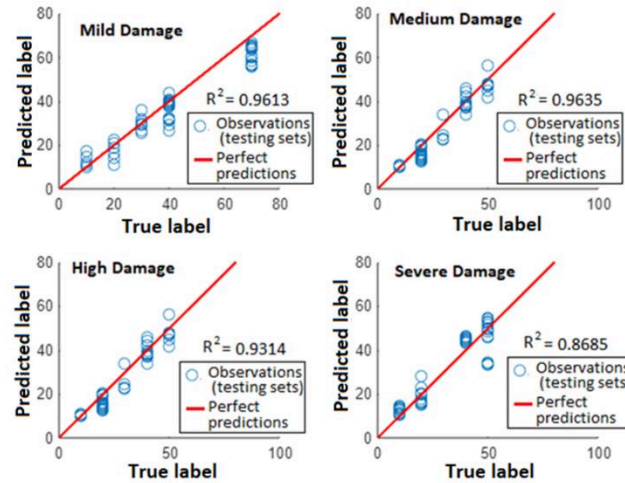


Figure 16: Actual vs predicted label using simulated testing datasets.

• Experiment Results:

Experiments are carried out on different plates with different crack dimensions. Experiments are repeated several times to have sufficient datasets for analysis and error estimation. Figure 17 represents the received damage signals in the raw form and their comparison with their respective pristine signals. Without using any type of filter,

these signals are directly used to extract damage features. Table 7 summarizes the prediction results of these datasets using the proposed methodology. The average accuracy of all the cases is greater than 90% and RMSE and MAE are nearly equivalent to that of model validation. Moreover, it is important to mention here that the last two cases have distinct damage attributes, and the algorithm predicts results with high accuracy that indicates the model generalization.

Table 6: Average prediction accuracy using simulated testing datasets

Regions	Prediction accuracy (5db to 20db)
Mild Damage	0.909
Medium Damage	0.897
High Damage	0.960
Severe Damage	0.869
Total	0.908

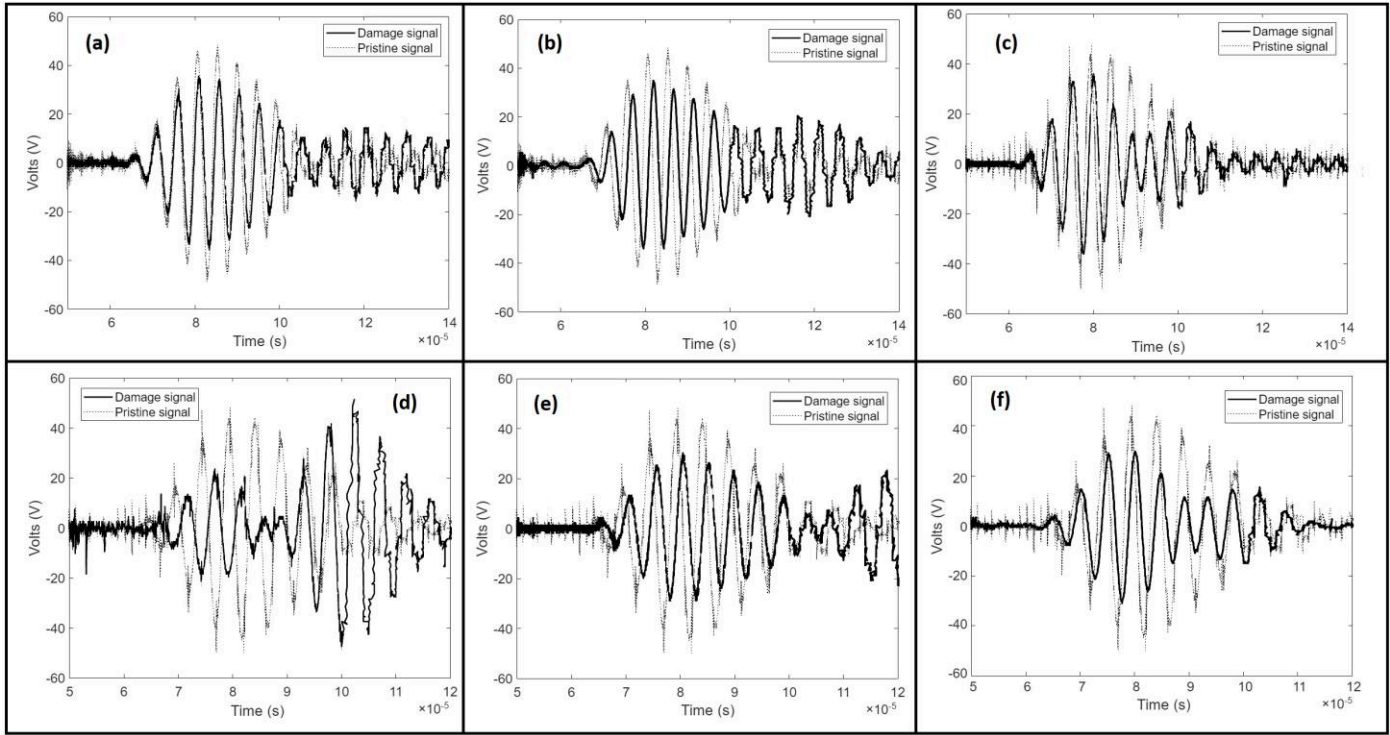


Figure 17: Experimental raw signals with crack dimension: **a.** 10 x 2 x 0.8mm, **b.** 40 x 2 x 0.8mm, **c.** 70 x 2 x 0.8mm, **d.** 70 x 2 x 1.4mm, **e.** 70 x 2 x 0.8mm (45°), **f.** 70 x 3 x 0.8mm

Table 7: Prediction summary of experimental raw signals

Dimension ($L \times W \times D$) Mm	Orientation (degree)	Number of experiments	Average prediction (mm)	Prediction Accuracy	RMSE (mm)	MAE (mm)
10 x 2 x 0.8	5	10	10.97	0.90	1.045	0.975
70 x 2 x 0.8	0	10	68.11	0.97	4.433	1.973

$40 \times 2 \times 0.8$	0	10	42.18	0.94	2.183	2.180
$70 \times 2 \times 1.4$	0	10	63.95	0.91	6.018	6.07
$70 \times 2 \times 0.8$	45	10	68.16	0.97	2.186	1.82
$70 \times 3 \times 1$	0	10	65.95	0.94	4.210	4.050
			Average:	0.938	3.345	2.844

6. Conclusion

This paper presents a machine learning framework for damage quantification model using ultrasonic Lamb waves and a separate model network. The underlying mechanism and behavior of Lamb waves with different severities of the damage; crack size, penetration level, and orientations are studied using numerical simulation based on FEM. Three time-frequency-based, damage-sensitive features are extracted from the first-wave packets of S_0 mode. These data are used to train two ensemble-based learners; one is the classification model and another one is the regression model. These models are linked together by using the separate model network. The overall proposed method is validated using low SNR simulated data and raw experimental data. The results indicate that the proposed machine learning scheme is efficient and reliable for crack severity estimation. From result analysis following conclusion are drawn,

- Time-frequency-based damage features exhibit reliable damage sensitivity as they contain the information in time and frequency scales. Moreover, smoothed pseudo-Wigner-Ville distribution has the true and excellent time-frequency energy distribution property, which is beneficial to accurately estimate the signal features.
- Since ensemble learning requires a large amount of dataset to protect the model from overfitting. Proposed spline-based interpolation is a robust and simple technique to augment the learning datasets that shows satisfactory results. However, a more efficient, and reliable method is required for time-series data, particularly for the regression model that considerably lacks in the literature.
- Ensemble-based learning, along with a separate model network, have great potential to predict the severity of the damage. The proposed framework shows a test accuracy of 90.8% and 93.8% for simulated and experimental datasets, respectively.
- High-accuracy results of two experimental data obtained by distinct damage-severity features that were excluded during the learning process prove the model generalization.
- Since the model requires baseline measurement, which is the only shortcoming of the proposed method.
- Although the proposed method generates acceptable results, there is still room to improve. The accuracy of the method can be further enhanced if a large number of reliable training datasets is used instead of synthetic data obtained by the aggressive augmentation process. Moreover, there is a future research scope in the quantification of cracks, corrosion, and dis-bonds in advanced metallic and composite materials.

7. Acknowledgement:

The authors are thankful to Dr. Sohaib Zia Khan (Department of Mechanical Engineering, Islamic University of Madinah, Kingdom of Saudi Arabia), Muhammad Tariq (Faculty of Engineering and Applied Science, University of Ontario Institute of Technology, Oshawa, Canada) and undergraduate students Adil Zafar and Muhammad Omer Humayun at Mechanical department, PNEC for providing experimental assistance.

8. Reference:

[1] Cho S, Yun C-B, Lynch JP, et al. Smart wireless sensor technology for structural health monitoring of civil structures. *Steel Structures* 2008; 8: 267–275.

[2] Abbas M, Simms N, Lao L, et al. A dimensional metrology-based approach for corrosion measurement of ship grade steels exposed to various marine environmental conditions. *Corrosion Engineering, Science and Technology* 2021; 1–13.

[3] Khan SZ, Khan MA, Tariq M, et al. Response of Gaussian-modulated guided wave in aluminum: An analytical, numerical, and experimental study. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* 2017; 231: 3057–3065.

[4] Giurgiutiu V. *Structural Health Monitoring with Piezoelectric Wafer Active Sensors: with Piezoelectric Wafer Active Sensors*. Elsevier, 2007.

[5] Michaels JE. Detection, localization and characterization of damage in plates with an in situ array of spatially distributed ultrasonic sensors. *Smart Materials and Structures* 2008; 17: 35035.

[6] Dai D, He Q. Structure damage localization with ultrasonic guided waves based on a time–frequency method. *Signal Processing* 2014; 96: 21–28.

[7] Hu N, Shimomukai T, Fukunaga H, et al. Damage identification of metallic structures using A0 mode of Lamb waves. *Structural Health Monitoring* 2008; 7: 271–285.

[8] Yu L, Tian Z, Leckey CAC. Crack imaging and quantification in aluminum plates with guided wave wavenumber analysis methods. *Ultrasonics* 2015; 62: 203–212.

[9] Yuan F-G, Zargar SA, Chen Q, et al. Machine learning for structural health monitoring: challenges and opportunities. 2020. Epub ahead of print 2020. DOI: 10.1117/12.2561610.

[10] Zhang Z, Pan H, Wang X, et al. Machine Learning-Enriched Lamb Wave Approaches for Automated Damage Detection. *Sensors* 2020; 20: 1790.

[11] Su Z, Ye L. Lamb wave-based quantitative identification of delamination in CF/EP composite structures using artificial neural algorithm. *Composite Structures* 2004; 66: 627–637.

[12] Agarwal S, Mitra M. Lamb wave based automatic damage detection using matching pursuit and machine learning. *Smart materials and structures* 2014; 23: 85012.

[13] HosseinAbadi HZ, Amirfattahi R, Nazari B, et al. GUW-based structural damage detection using WPT statistical features and multiclass SVM. *Applied acoustics* 2014; 86: 59–70.

[14] Yang J, He J, Guan X, et al. A probabilistic crack size quantification method using in-situ Lamb wave test and Bayesian updating. *Mechanical Systems and Signal Processing* 2016; 78: 118–133.

[15] Cofre-Martel S, Kobrich P, Lopez Droguett E, et al. Deep convolutional neural network-based structural damage localization and quantification using transmissibility data. *Shock and Vibration*; 2019.

- [16] Sun F, Wang N, He J, et al. Lamb wave damage quantification using GA-based LS-SVM. *Materials* 2017; 10: 648.
- [17] Rai A, Mitra M. Lamb wave based damage detection in metallic plates using multi-headed 1-dimensional convolutional neural network. *Smart Materials and Structures* 2021; 30: 35010.
- [18] Ewald V, Groves RM, Benedictus R. DeepSHM: a deep learning approach for structural health monitoring based on guided Lamb wave technique. 2019. Epub ahead of print 2019. DOI: 10.1117/12.2506794.
- [19] Falcetelli F, Venturini N, Romero MB, et al. Broadband signal reconstruction for SHM: An experimental and numerical time reversal methodology. *Journal of Intelligent Material Systems and Structures*; 32. Epub ahead of print 2021. DOI: 10.1177/1045389X20972474.
- [20] Gangadharan R, Murthy CRL, Gopalakrishnan S, et al. Time reversal technique for health monitoring of metallic structure using Lamb waves. *Ultrasonics*; 49. Epub ahead of print 2009. DOI: 10.1016/j.ultras.2009.05.002.
- [21] Ing RK, Fink M. Time-reversed lamb waves. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*; 45. Epub ahead of print 1998. DOI: 10.1109/58.710586.
- [22] Kannusamy M, Kapuria S, Sasmal S. Accurate baseline-free damage localization in plates using refined Lamb wave time-reversal method. *Smart Materials and Structures*; 29. Epub ahead of print 2020. DOI: 10.1088/1361-665X/ab8028.
- [23] Xu B, Giurgiutiu V. Single mode tuning effects on lamb wave time reversal with piezoelectric wafer active sensors for structural health monitoring. *Journal of Nondestructive Evaluation*; 26. Epub ahead of print 2007. DOI: 10.1007/s10921-007-0027-8.
- [24] He J, Ran Y, Yang J, et al. A novel crack size quantification method based on lamb wave simulation. In: *2016 Prognostics and System Health Management Conference (PHM-Chengdu)*. IEEE, 2016, pp. 1–6.
- [25] Yu L, Leckey CAC. Lamb wave–based quantitative crack detection using a focusing array algorithm. *Journal of Intelligent Material Systems and Structures* 2013; 24: 1138–1152.
- [26] Wang J, Shen Y, Rao D, et al. Physical-virtual time reversing of nonlinear Lamb waves for fatigue crack detection and quantification. *Mechanical Systems and Signal Processing* 2021; 160: 107921.
- [27] He J, Ran Y, Liu B, et al. A fatigue crack size evaluation method based on lamb wave simulation and limited experimental data. *Sensors* 2017; 17: 2097.
- [28] Tan AC, Gilbert D. Ensemble machine learning on gene expression data for cancer classification.
- [29] The Math Works, Inc., <https://www.mathworks.com/help/stats/ensemble-algorithms.html> (2021).
- [30] Zhou ZH. *Ensemble methods: Foundations and algorithms*. 1st Editio. New York: CRC press, 2012. Epub ahead of print 2012. DOI: 10.1201/b12207.
- [31] Jiang W, Chen Z, Xiang Y, et al. SSEM: A novel self-adaptive stacking ensemble model for classification. *IEEE Access* 2019; 7: 120337–120349.

[32] Sagi O, Rokach L. Ensemble learning: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 2018; 8: e1249.

[33] Nikulin V, McLachlan GJ, Ng SK. Ensemble approach for the classification of imbalanced data. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2009. Epub ahead of print 2009. DOI: 10.1007/978-3-642-10439-8_30.

[34] Ni S, Wu T-Y, Chang T-H, et al. Advances in Smart Vehicular Technology, Transportation, Communication and Applications.

[35] el Gayar N, Kittler J, Roli F. *Multiple Classifier Systems: 9th International Workshop, MCS 2010, Cairo, Egypt, April 7-9, 2010, Proceedings*. Springer, 2010.

[36] Su Z, Ye L. *Identification of damage using Lamb waves: from fundamentals to applications*. Springer Science & Business Media, 2009.

[37] Rizvi SH, Abbas M. An advanced Wigner-Ville time-frequency analysis of lamb waves signals based upon AR model for efficient damage inspection. *Measurement Science and Technology* 2021; 32: 095601.

[38] Ihn J-B, Chang F-K. Detection and monitoring of hidden fatigue crack growth using a built-in piezoelectric sensor/actuator network: I. Diagnostics. *Smart materials and structures* 2004; 13: 609.

[39] Kim YY, Kim E-H. Effectiveness of the continuous wavelet transform in the analysis of some dispersive elastic waves. *The Journal of the Acoustical Society of America* 2001; 110: 86–94.

[40] Biswal M, Brahma SM, Cao H. Supervisory Protection and Automated Event Diagnosis Using PMU Data. *IEEE Transactions on Power Delivery*; 31. Epub ahead of print 2016. DOI: 10.1109/TPWRD.2016.2520958.

[41] Boashash B. *Time-frequency signal analysis and processing: a comprehensive reference*. Academic Press, 2015.

[42] Jaitly N, Hinton GE. Vocal tract length perturbation (VTLP) improves speech recognition. In: *Proc. ICML Workshop on Deep Learning for Audio, Speech and Language*. 2013.

[43] Cui X, Goel V, Kingsbury B. Data augmentation for deep neural network acoustic modeling. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 2015; 23: 1469–1477.

[44] Ko T, Peddinti V, Povey D, et al. Audio augmentation for speech recognition. In: *Sixteenth Annual Conference of the International Speech Communication Association*. 2015.

[45] Oh C, Han S, Jeong J. Time-Series Data Augmentation based on Interpolation. *Procedia Computer Science* 2020; 175: 64–71.

[46] Sáez JA, Galar M, Krawczyk B. Addressing the overlapping data problem in classification using the one-vs-one decomposition strategy. *IEEE Access* 2019; 7: 83396–83411.

[47] Chai T, Draxler RR. Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*; 7. Epub ahead of print 2014. DOI: 10.5194/gmd-7-1247-2014.

- 1
2
3 [48] Luangvilai K. Attenuation of ultrasonic Lamb waves with applications to material characterization and
4 condition monitoring.
5
6 [49] Moser F, Jacobs LJ, Qu J. Modeling elastic wave propagation in waveguides with the finite element
7 method. *Ndt & E International* 1999; 32: 225–234.
8
9 [50] The Math Works, Inc., <https://www.mathworks.com/help/stats/classificationlearner-app.html> (2021,
10 accessed September 2, 2021).
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



DSPACE

<https://dspace.org/>

Lamb wave damage severity estimation using ensemble-based machine learning method with separate model network

Rizvi, Syed Haider; Abbas, Muntazir

2021-10-08

IOP Publishing

<https://doi.org/10.1088/1361-665X/ac2e1a>

Rizvi SH, Abbas M. (2021) Lamb wave damage severity estimation using ensemble-based machine learning method with separate model network. Smart Materials and Structures, Volume 30, Issue 11, November 2021, Article number 115016