Transportation Consortium of South-Central States

Tran-SET

Solving Emerging Transportation Resiliency, Sustainability, and Economic Challenges through the Use of Innovative Materials and Construction Methods: From Research to Implementation

Performance monitoring leveraging advanced AI technique with CNN

Project No. 21CUTA01

Lead University: The University of Texas at Arlington

Final Report

August 2022

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

Acknowledgements

The authors acknowledge the Tran-SET for providing the financial support of this study. The authors also would like to thank Project Review Committee, Do Soo Moon , Austin Mckeon, and Trent Ballard for their support and direction of the project.

TECHNICAL DOCUMENTATION PAGE

1. Project No.	2. Government Accession No.	3. Recipient's Catalog No.
21CUTA01		

4. Title and Subtitle	5. Report Date
Performance monitoring leveraging advanced AI technique with CNN	Aug. 2022
	6. Performing Organization Code
7. Author(s)	**8. Performing Organization Report No.**
PI: Dr. Suyun Ham (ORCID: 0000-0001-6375-211X) Co-PI: Stefan Romanoschi GRA: Yin Chao Wu GRA: Dafnik Saril Kumar David GRA: Sanggoo Kang	
9. Performing Organization Name and Address	**10. Work Unit No. (TRAIS)**
Transportation Consortium of South-Central States (Tran-SET) University Transportation Center for Region 6 3319 Patrick F. Taylor Hall, Louisiana State University, Baton Rouge, LA 70803	**11. Contract or Grant No.**
12. Sponsoring Agency Name and Address	**13. Type of Report and Period Covered**
United States of America Department of Transportation Research and Innovative Technology Administration	Final Research Report Aug. 2021 – Aug. 2022
	14. Sponsoring Agency Code

15. Supplementary Notes

Report uploaded and accessible at Tran-SET's website (http://transet.lsu.edu/).

16. Abstract

The main goal of this project is to study and develop a reliable nondestructive testing (NDT)-based structural performance prediction model framework leveraging the advanced machine learning convolutional neural network (CNN) technique and rapid crack evaluation system. There are two steps of application CNN technique in this project: 1) the first step is to identify delamination, noise, and the unexpected signal produced by the existing damage identification algorithm to improve the accuracy of NDT results. The input image or training data of NDT data for CNN is comprehensively studied with several features, such as the duration of the signal, the starting time of the signal, the resolution of images, and the number of images. 2) The second step is to study damage prediction with four different stress levels. The FE model is used to simulate structural performance with different delamination conditions. Moreover, except for field test results, the artificial delamination model is created. We performed numerous finite element (FE) simulation to create inputs for CNN for damage detection. The result shows improved NDT results, and CNN can achieve structural performance prediction. We performed six tasks based on these objectives: Task 1. literature review; Task 2. Collect data from bridges; Task 3. perform filed test NDT results; Task 4. Develop FE model based on field test results; Task 5. development of a machine learning model for damage prediction.

17. Key Words	18. Distribution Statement
NDT, Bridge deck evaluation, delamination, ML, CNN, structural performance prediction, FE,	No restrictions. This document is available through the National Technical Information Service, Springfield, VA 22161.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. 47	22. Price
Unclassified	Unclassified		

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized.**

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
		NOTE: volumes greater than 1000 L shall be shown in m^3		
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
^oF	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	^oC
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

APPROXIMATE CONVERSIONS FROM SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact degrees)		
^oC	Celsius	1.8C+32	Fahrenheit	^oF
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS, ABBREVIATIONS, AND SYMBOLS

NDT	nondestructive testing
DL	deep learning
FE	finite element
ANN	artificial neural network
CNN	convolution neural network
MEMS	micro-electromechanical system
ACE	automated crack evaluation
ML	machine learning
AI	artificial intelligence
SHM	structural health monitoring
TRB	transportation research board
GPR	ground-penetrating radar
KCE	kinematic contact enforcement method
Acc.	accuracy
CT.	computation time

EXECUTIVE SUMMARY

The overall goal of this study is to develop a framework integrating infrastructure performance evaluation leveraging an advanced evaluation system, so-called automatic crack evaluation system (ACES), and advanced machine learning (ML) techniques (e.g., convolutional neural network (CNN)). Ultimately, the results of developed framework 1) enable reliable traffic disruption-free assessment, 2) provide structural performance data incorporating with the damage, and 3) help accurate prediction of structural damage with proper damage classification. The proper maintenance and operation of deteriorating infrastructures require timely detection, precise diagnosis, and accurate estimation of possible structural performance degradation induced by various damages. Many advanced bridge assessment techniques have been recently practiced to evaluate and monitor concrete structures while requiring the evaluation of numerous measurement data for more accurate interpretation and assessment. Current technologies have been developed to prevent time-consuming and labor-intensive field tests. However, most high-speed techniques still present practical challenges, such as 1) limitations in accuracy, sensitivity, and coverage by focusing on indirect response and external surface conditions, and 2) not considering structural performances that are not readily available for engineers, decision-makers, and stakeholders. In detail, many field data present a challenge due to unexpected signals such as noise, which may greatly affect signal analysis results, leading to over- or under-estimation of structural damage prediction.

The primary objectives of this study are 1) to improve nondestructive testing (NDT) systems by using ACES with the high-speed traffic disruption-free damage detecting system and state-of-art signal processing algorithms to enhance damage recognition capability and speed in <u>field environments</u>, 2) to perform finite element (FE) modeling for an efficient structural performance incorporating ACES data; and 3) to develop a CNN framework that provides a quick decision of its structure performance to make reliable asset management decisions. This study presents a hybrid model featuring field assessments with NDT and deep learning (DL) to improve the accuracy of damage detection and predict the structural behavior of a bridge deck. Seismic wave that is obtained from sensors in the field is reflected from internal objects such as a crack to evaluate different damages. Since various uncertainty factors may affect collected data, such as field NDT results on surface conditions, road slope, and bridge material, a more comprehensive and in-depth study should be performed on the analysis of sensing signals. A series of procedures for damage prediction is devised and applied based on a comparative assessment. The steps are described below: 1) The NDT field test results are produced, 2) an in-depth study of the relationship between the DL model and NDT results with several parameter studies to improve NDT results, 3) develop FE model based on a field test result to obtain structural performance, 4) FE model maps to develop DL model for structural performance identification and prediction model. The results indicate that an image-based CNN can improve NDT results by using signals to identify delamination and noise or insignificant signals. Also, using the FE model and NDT results, CNN can identify and predict different kinds of structural levels. Ultrimaty, for roadway safety and bridge deck service life, these results significantly contribute to maintaining the bridge deck in the early-stage deterioration to ensure the infrastructure is operating safely and efficiently.

1. INTRODUCTION

The maintenance of transportation infrastructures (e.g., bridges) has become a major concern for safety and economic loss due to many factors. In particular, infrastructure deterioration (e.g., external and internal damages) can significantly impact service life of bridges and other infrastructures and often require extensive repairs or replacements. Thus, reliable inspections and monitors of bridge conditions are required to improve the service life to ensure roadway safety and provide proper time for appropriate preservation and rehabilitation treatments for asset owners or transportation agencies.

The most serious problem in bridge monitoring is internal damages like vertical cracks or delamination, which crack horizontally, mainly caused by corroded steel reinforcements. Generally, reinforcement corrosion in bridge decks occurs due to environmental conditions such as migrated moisture and chemicals (e.g., chloride ions). The formation of corrosion products causes volume expansion of the reinforcement—eventually, delamination forms due to this mechanism. In addition, the produced delamination also causes vertical cracks to extend the delamination to the surface. Increased levels and a number of these vertical crack damages even accelerate the corrosion process. As a result, further degradation (e.g., potholes) can be caused by the negative interactions among delamination, vertical cracks, and reinforcement corrosion with external environmental factors such as traffic load and freeze-thaw cycle).

Proper maintenance and inspection operation for deteriorated infrastructure require timely detection, precise diagnosis, and accurate estimation of possible structural performance degradation induced by various damages. Many advanced bridge assessment techniques have been recently practiced to evaluate and monitor bridge structures. In addition serveral technologies have been developed to prevent time-consuming and labor-intensive field tests. However, these high-speed inspection techniques still present practical challenges such as 1) limitations in accuracy, sensitivity, and coverage by focusing on visual inspection and external surface conditions, 2) not considering structural performances that are not readily available for engineers, decision-makers, and stakeholders.

To improve the current system, a rapid damage inspection system, was developed without interrupting traffic using noncontact microelectromechanical systems (MEMSs) and multichannel acoustic scanning called automatic crack evaluation system (ACES). The noncontact manner in the system enables faster, easier, and more accurate evaluations for improving timely maintenance. Finite element (FE) analysis was also performed to simulate damaged structures incorporating the damage index map obtained from ACES. However, there still challenge to simulate and evaluate damage for each bridge, while ACES will provide quick and real-time internal damage. In addition, many field data present challenges due to unexpected signals such as noise, which may greatly affect signal analysis results, leading to over- or underestimation of structural damage prediction. Thus, advanced deep learning (DL) techniques introduced in this study can help identify structural damages and process and compile raw sensing data and ensuing damage map results to improve accuracy. The NDT field test results are used to reference to design FE model, which used as training input for DL. Results so far show that the proposed DL model trained with field test and simulation data significantly enables improved assessment results for bridge damage identification and prediction.

2. OBJECTIVES

Overall goal of this study is to develop a framework integrating infrastructure performance evaluation leveraging an advanced evaluation system and advanced ML techniques (e.g., convolutional neural network (CNN)), which ultimately enable reliable traffic disruption-free assessment, provide structural performance data incorporating with the damage, and help accurate prediction of structural damage with proper damage classification.

To achieve the goall, the primary objectives of the proposal are i) to improve NDT systems by using ACES with the high-speed reference-free damage detecting system and state-of-art signal processing algorithms to enhance damage recognition capability and speed, and ii) to perform FE modeling for an efficient structural performance incorporating ACES data; and 3) to develop a CNN framework that provides a quick decision of its structure performance to make reliable asset management decisions.

3. LITERATURE REVIEW

The National Bridge Inspection program regulations require states to inspect highway bridges persistently on a reoccurring timetable that can vary depending on the type of infrastructure. Bridges are critical components of transportation infrastructure. Bridge decks, in particular, are the most susceptible components in a bridge to traffic safety and material deterioration due to direct exposure to traffic and deteriorating factors (e.g., temperature, moisture, deicing agents). Their service life is shorter than other components. Thus, monitoring the degree of deterioration of bridge decks is important for determining appropriate maintenance and rehabilitation strategies.

Structural health monitoring (SHM) and NDT techniques have been widely studied and utilized in infrastructure evaluation for maintenance over the past decades. Traditional SHM and NDT techniques are time-consuming, labor-intensive, and error-prone to realize in some situations[1]. Advanced sensing has gained more attention in addressing these limitations: advanced sensing for surface condition assessments (e.g., smartphones, automated vehicles) [2], acoustic-ultrasound[3], electrochemical sensors[4], and fiber Bragg grating sensors [5]. There are also current NDT techniques focusing on crack evaluation: a damage monitoring system with GIS and acceleration[6], crack development studdies[7], [8], and in-situ monitoring[9], [10]. Previous attempts have persistent issues in identifying the damage, especially “internal” deteriorations (e.g., delamination, cracks) requiring lane closure due to their slow speed or stationary measurement on a bridge (e.g., chain drag, impact-echo, contact 3D tomography). There are also advanced studies for high-speed bridge deck scanning for crack detection by several researchers using infrared thermography[11], ground-penetrating radar (GPR), light detection and ranging (LiDAR), and mechanical wave-based evaluation[11]. In spite of these efforts, several high-speed systems (e.g., thermography, GPR, LiDAR) provide indirect crack information or superficial information and weather-induced variations. In addition, most “high-speed” mechanical wave-based approaches provide limited internal damage identification or present relatively slow inspection requiring lane closure. For performing accurate structural assessments in a real-time manner, rapid traffic disruption-free damage inspection without lane closure so-called automated crack evaluation system (ACES) has been developed[11]. The noncontact manner in the system enables faster, easier, and more accurate evaluations for timely maintenance. The rapidly obtained mechanical waves propagate through infrastructure elements (e.g., bridge decks) to provide a 2-D or 3-D damage image similar to an MRI, to show a hidden damage map.

However, there are a couple of questions and challenges as following for improving infrastructure maintenance: 1) can we evaluate both potential driving safety (e.g., potholes, cracks) and structural performance using the data from ACES? 2) how to evaluate the bridge capacity and how to compare it with a hidden damage map? 3) what are the appropriate and quick approaches to evaluate its structural performance, aiding FE modeling results? To addresses challenges considering these questions, advanced machine-learning (ML) techniques (e.g., convolutional neural network (CNN), artificial neural networks (ANN), which ultimately enable reliable traffic disruption-free assessment, provide structural performance data incorporating with the damage, and help accurate prediction of structural damage with proper damage classification. The accuracy of structural damage evaluated by NDT needs to be improved by incorporating advanced technology.

To achieve the goal of developing a framework, three objectives are to 1) to enhance inspection system and data using ML and ACES for high-speed reference-free damage detecting system and state-of-art signal processing algorithms, 2) to perform FE modeling for an efficient structural performance incorporating field data; and 3) to develop an ANN and CNN framework that provides a quick decision of its structure performance to make reliable asset management decisions.

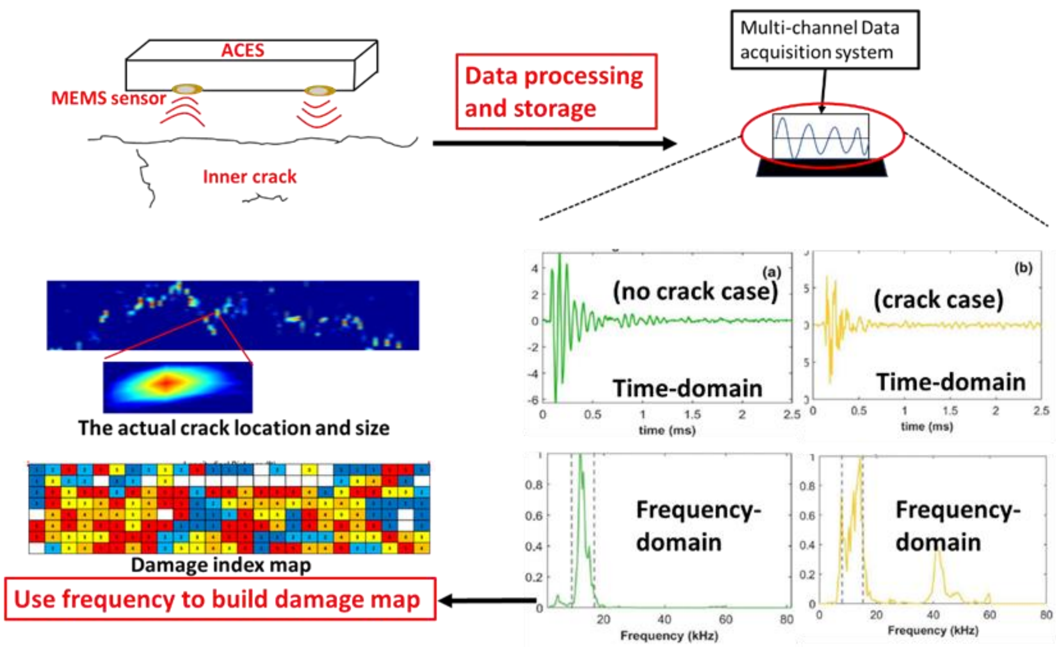


Figure 1. The concept of collecting and post-processing field data obtained from ACES

In recent years, the influence of information technology (**IT**) has grown tremendously with regard to different aspects of today's society. One of the most well-known IT is machine learning. Machine learning (ML) in the field of computer science of using statistical techniques to enable computers to act and make data-driven decisions and progressively learn and improve over time without being explicitly programmed. The ML is the dicision maker to find the optimal output, which has the minimum data loss value(e.g., mean square error). There are several important parameter affect ML performance and effiecny, which including loss function, active function, optimizor and learning rate. Jaocha et al. studied about the loss function for deep neural networks especially for classification. They investigate how the specific loss function affect ML models and their restuls[12]. For active function, Nwankpa et al. compared the commonly used active function and their trends in different purpose for deep learning to provide the advise of active functions[13]. The learning rate is control the effiency to find the minimum loss value. If learning rate is too long, it is never find the global minimum value. However, if learning rate is too short, the computation time will be increased more. Thus, to find the reasonable learning rate is important. Smith provides new approach to set the learning rate, which practically eliminates the need to experimentally find the best values and schedule for the global learning rates[14]. Besides the conventional ML model, there are many more advanced ML as known as deep learning (DL) which applied with different algorithms. The most two common DL model is ANN and CNN. ANN already applied in many field, for example, Mirhosseini et al. develop and predict rainfull intensity model for future climate change[15] ; Neves et al. studied the damage detection for strcutrual healthy monitoring applied with ANN [16]. Jerome et al. deeply studied about how to improve the effiency of computation time and accuracy with new algorithms[17]. There are more recently developed DL model, for example, the generative adversarial network, which has two training structures in one model. The first is generator and the other is discriminator. Unlike ANN or CNN, GAN can create fake sample itself and use these fake sample mixed with real sample to train the model. GAN commonly applied on face identification in real time[18]. Overall, the ML techniques already developed and applied in different area, and it still growing with engineers efforts, therefore, people start to believe ML can make better decision regarding data analysis.

In this project, two deep learning techniques are used, data-based ANN, and image-based CNN. The difference between these two is ANN uses data, while CNN uses image information known as a pixel value. Khan et al. [19]use STFT to convert time signals to spectrum imaging, which can present the frequency within a certain time frame, and uses the input for CNN to identify delamination [19]. [20]Willard et al. provide the approach to analyze spectrograms, which is used for CNN by pixel value to identify signal differences [20].[21] Ahmadvand et al. used structural responses obtained from the GPR as input images to identify defects [21]. The research results showed that CNN can identify signals correctly with high accuracy. For ANN, [16][22]Dworakowski et al. use ANN to identify damage in aircraft [22] and [23] Güemes et al. identify structural damage based on the SHM system [23]. Thus, ANN and CNN have already proved the ability to identify damage or signals. In this project, they are used to 1) identify delamination signal, non-delamination signal, and noise signal to improve NDT results; 2) identify structural performance simulated from FE models to predict four concentrated stress level.

4. METHODOLOGY

In this section, the research tasks will be explained sequentially. Task 2 and 3 provide the evaluation principle of bridge inspection using ACES and a detailed procedure of field signal data enhancement using CNN. In Task 4, the FE modeling and analysis of structural performance are established based on the field test results. Depending on the damage index of inspection results, the FE model is designed with different damage levels. In Task 5, the machine learning framework is developed and discussed. The field test data (from Task 2) and FEM data (from Task 4) are the main resources as input or training data to develop an ML model for damage prediction. All results are shown in Chapter 5, which include: 1) the NDT evaluation of bridge inspection and signal enhancement, 2) The FE model development with actual defect, 3) the in-depth study of the damage prediction model developed with field test and FE data, and 4) the structural performance prediction with four different levels. According Figure 2, the CNN procedure to improve delamination map. First is a field test to obtain data; secondly, try to find the relationship between input signal images and CNN identification accuracy to develop the optimal CNN model. Thus, the delamination map are improved by CNN classification. The CNN procedure for structural performance identification and prediction(concentrated stress value). After obtaining improved delamination maps, 110 FE models are created, which include low damage index and artificial delamination maps. Four different levels of stress concentration are calculated from the post-processed map and used for labeling. The final step is to use labeled FE model images with CNN to predict stress distribution results from the damage map in Figure 3.

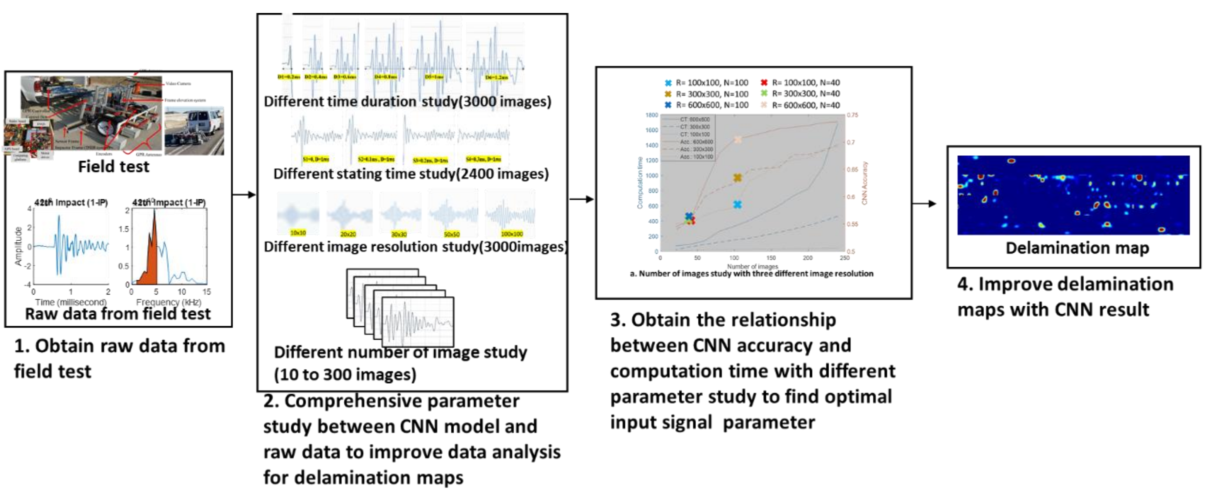


Figure 2. Overview of Objective 1 to enhance NDT inspection.

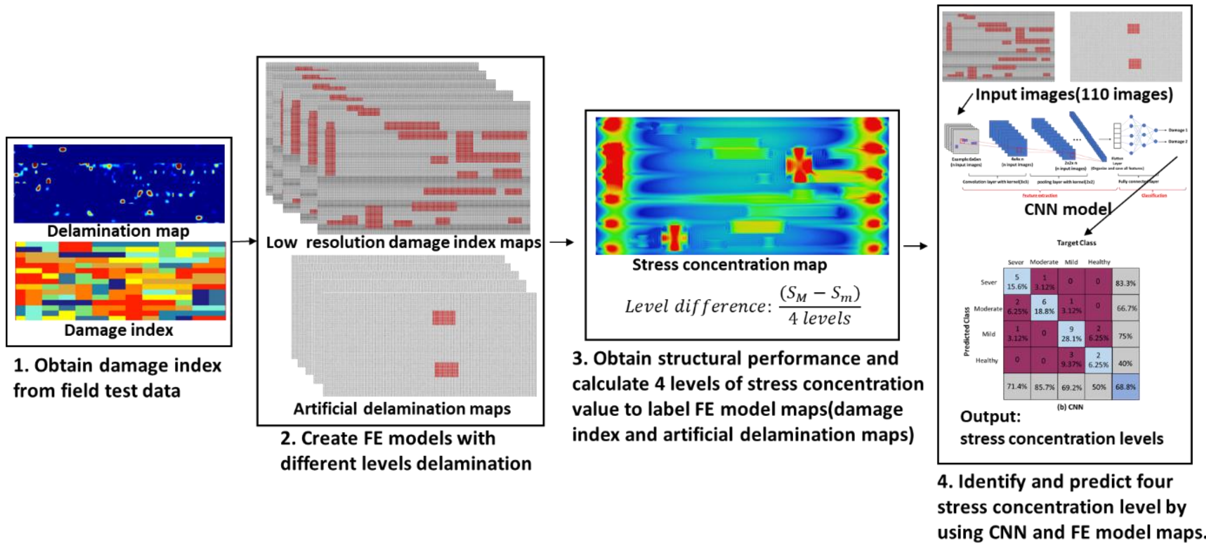


Figure 3. Overview of Objectives 2 and 3 for performing FE structure performance and prediction.

4.1 Task2: Select the bridges to deploy ACES along critical Texas corridors

During the early stage of this project, the research team evaluated the Texas highway bridge deck for identifying internal damages. Two bridges are C is significantly higher than on other bridges. The location of bridge 1 as shown in Figure 4. The scanning length of the first bridge referred to as Bridge 1 is 200 ft., and the bridge deck depth is 6.75 inches, covered with hot mix asphalt, as shown in Figure 55. The length of a second bridge, referred to as Bridge 2, is 312 ft, and there are two bridge deck depths, which are 6.75 inches and 7.25 inches, as shown in Figure 66. The scanned width of ACES is 6.6 feet, as shown in Figure 77; thus, we have three scanning lanes on Bridge 1 and four scanning lanes on Bridge 2. To ensure the correctness and decrease the uncertainty (e.g., the vehicle speed and moving direction) of inspection results, all lanes are scanned twice. After finishing one lane scan, the inspection team goes back to the bridge and scan for 10 minutes. The inspection results provide information about the severity of damage by performing the quantified damage evaluation proposed in the data analysis. The fundamental background and detailed data analysis process of NDT are discussed in Task 3.

Figure 4. Inspection area and the sectional view of slabs of the first bridge.

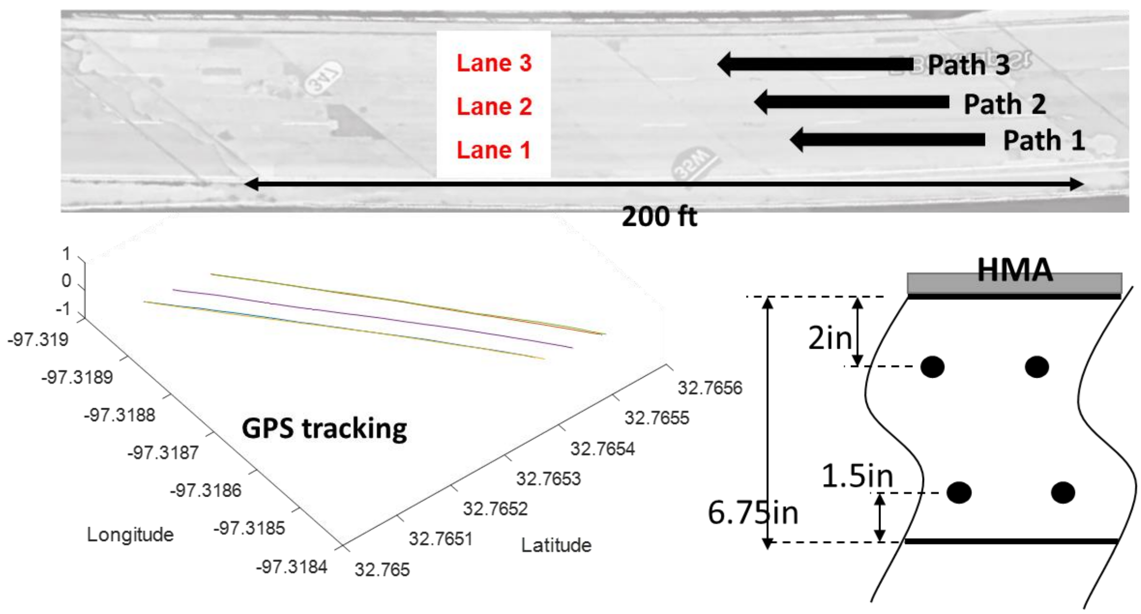


Figure 5. The information on the first bridge. Scanning length:200 ft. and the bridge deck is 6.75 inches.

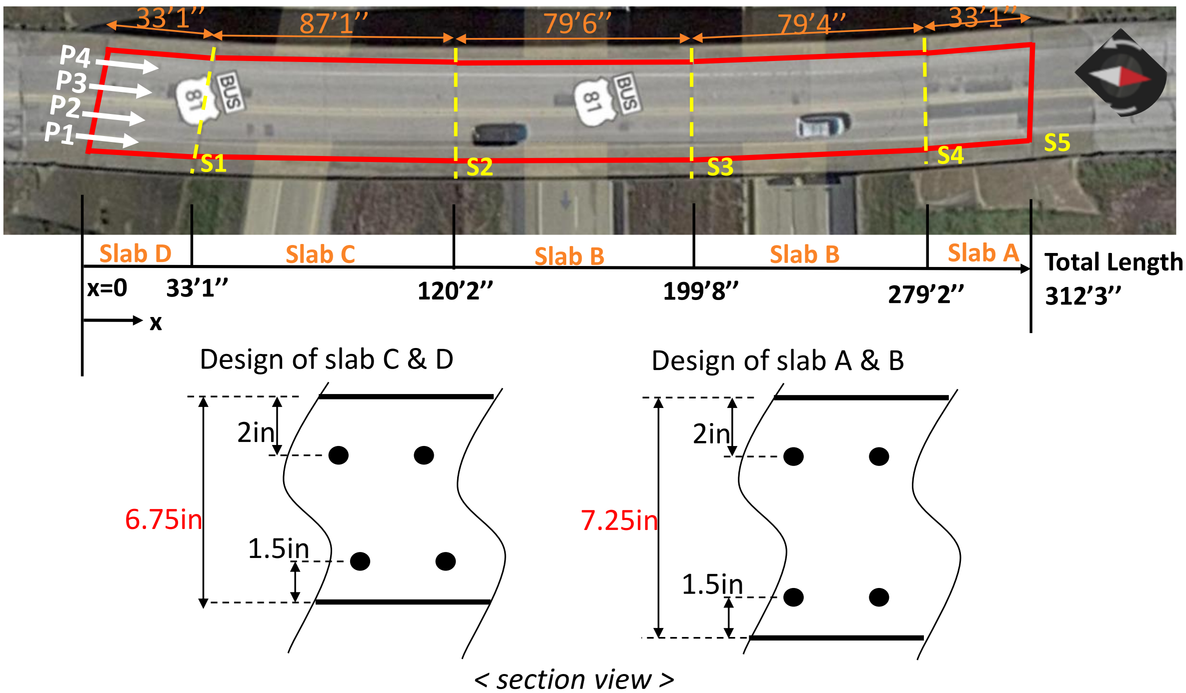


Figure 6. The information of the second scanned bridge. Length is 312 ft. with two types of deck (depth:6.75 in and 7.25 in).

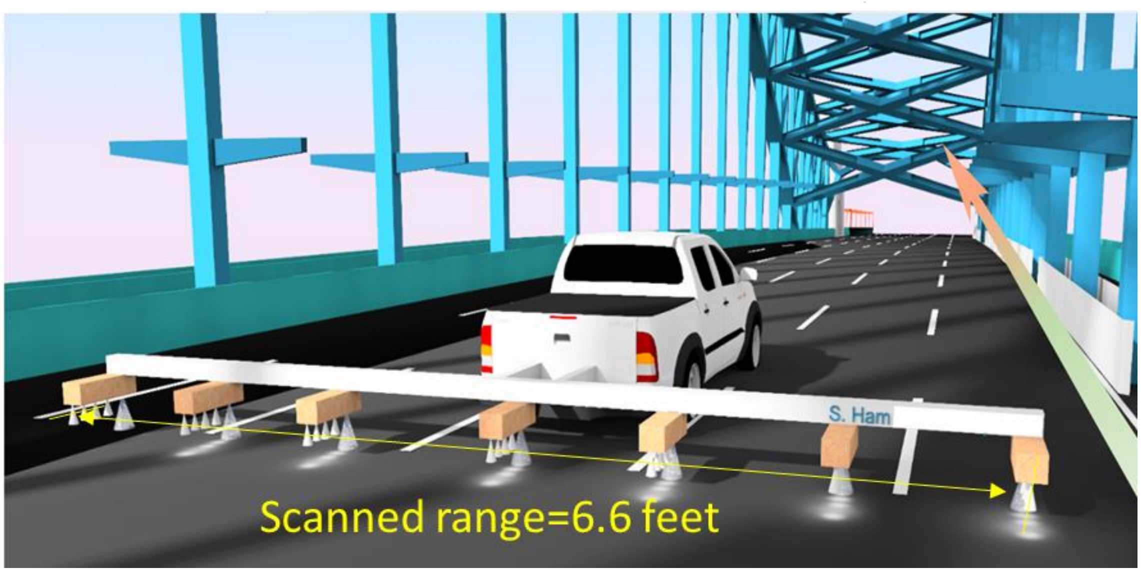


Figure 7. The ACES scanned range is 6.6 feet, which can cover the bridge lane by lane.

3.2 Task3: Obtain Field Data and Perform Data Analysis

The method called the impact-echo principle is the main technique was used in this project. There are wave generators, transducers, noncontact sensors and data acquisition. Based on the laboratory test system, this task focus on further improvement of signal processing obtained from ACES using advanced ML such as CNN. ACES is an automatic rapid damage inspection system that will be unitized. Previously, the research team developed an ACES based on the high-speed reference-free acoustic crack measurement method, which detects early-stage internal damages in real time. ACES is also composed of a novel acoustic scanning system of rapid stress wave excitation sources, air-coupled sensors, a GPS positioning system, and an advanced post-processing system. The ACES system will be used to image internal damage information on transportation pavements.

The design of the ACES aims to perform rapid traffic disruption-free bridge inspection with enhanced scanning qualities leveraging the integrated scanning platform, advanced impacting system, and a multichannel acoustic sensing unit [11]. The system comprises 22 noncontact MEMS sensors, GPS, GPR antennas, data acquisition, and a control panel, as shown in Figure 8. The MEMS is used to obtain mechanical waves reflected from bridge interior damage. The signal is transferred to the frequency domain to identify the healthy condition. The peak (or energy) appears in the range between 1kHz-5kHz, presenting delamination, while the range above 5kHz presents healthy conditions. The two examples of delamination and non-delamination field test data as shown in Figure 9. The following delamination map presents the total energy in the frequency domain between 1-5kHz; the higher value of energy means the damage severity is higher in the location.

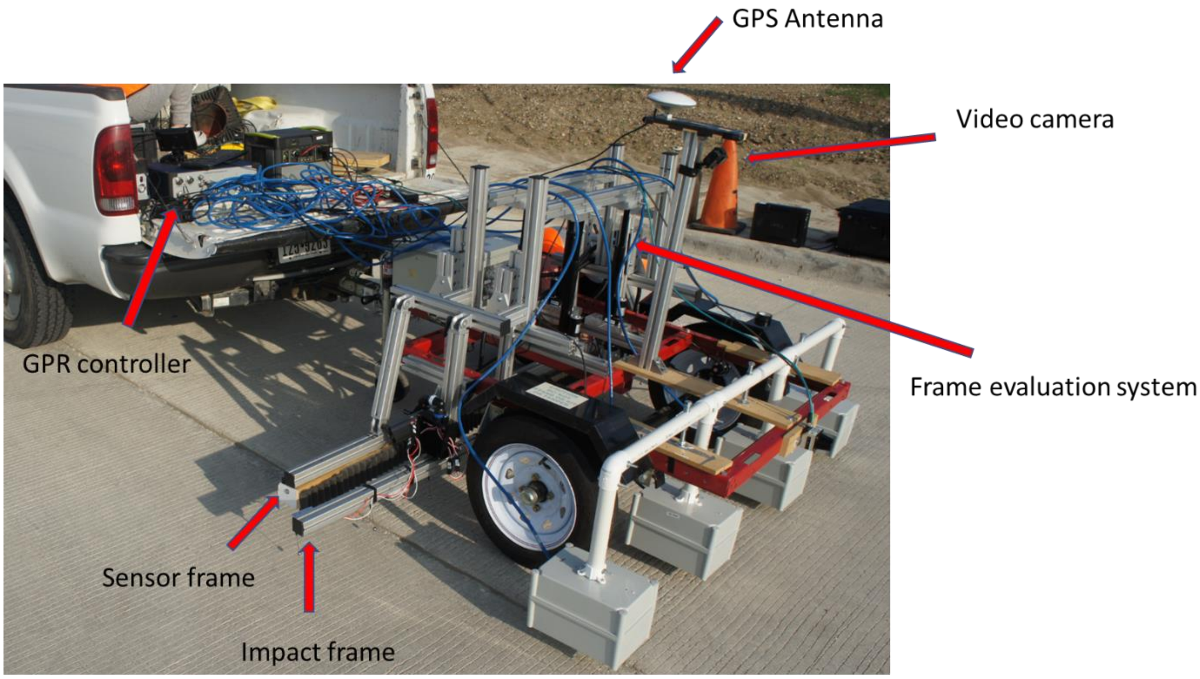
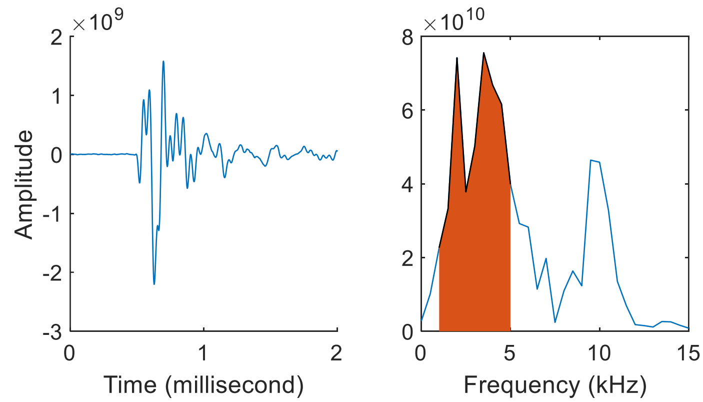
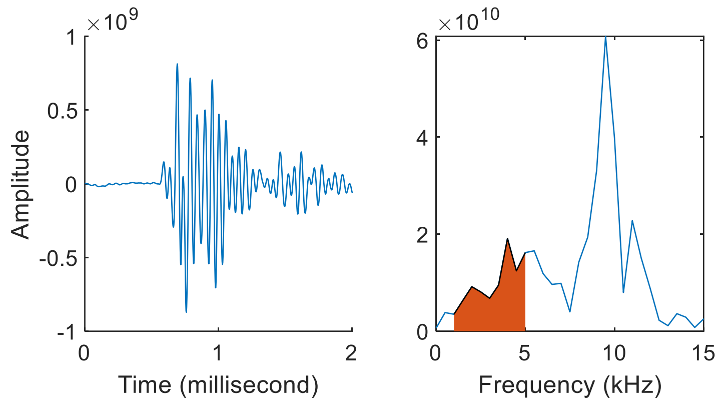


Figure 8. Photo of ACES. The system comprises 22 MEMS, GPS antenna, GPR antennas, data acquisition and a control panel.



(a)



(b)

Figure 9. The examples of signal analysis from field test data on (a) delamination and (b) non-delamination.

4.3 Task4: Perform FE Modeling and Analysis for Structural Performance Evaluation

After processing the delamination map and damage index from Task3, the FE model is designed based on delamination map results. In this study, ABAQUS/EXPLICIT is used to simulate a 3-D FE model representing concrete structures wherein both cases, the same isotropic, elastic plate with Young's modulus (E), Poisson's ratio, and the structure's thickness. The model element type is a four-node plane stress element defined as C3D8R. The mesh size for the plain solid 3-D simulation is 25 mm. Typical values of material properties are assumed for concrete (2400 kg/m3, E = 30 GPa, and t = 0.2). A computer workstation with 16 GB RAM and eight CPUs with a clock speed of 1.90 GHz and a 250 GB hard drive is used to carry out the computations. The kinematic contact enforcement method (KCE) simulates various delamination conditions by giving interaction boundary conditions to simulate mechanical waves reflected by damage. The main result of the FE model is the stress changing at a time when an external force is applied. The stress change can help us to understand the structural performance with certain damages. Besides, this result is one of the significant CNN inputs to identify damage severity.

4.4 Task5: Development of a machine learning framework

In Task 5, there are two parts needed to be solved. The first part is related to Objective 1 which is to improve the NDT accuracy because the existing data processing algorithm is not fit for all bridges. Sometimes noise or insignificant signals are chosen as the correct signal to result in the wrong evaluation. Thus, image-based ML CNN is used to identify whether it is a correct signal or noise. Also, the data-based ML ANN is used to compare with CNN. The second part is related to Objective 3 in which once the NDT delamination results are improved, the FE model and damage index will be training data for the advanced MI model to predict structure performance levels from the different severity of damage case. Besides, the relationship between CNN accuracy and computation time was also studied. The goal is to find the optimal NDT results input for CNN to identify damage severity.

4.4.1 Artificial neural networks (ANN) study

For the training data of ML, in this ANN model, the maximum value and energy area between 1kHz-6kHz and 7 kHz -12kHz are used as data features, which means the value can be used to represent a significant characteristic of signal:

$$\boldsymbol{max}\left(\boldsymbol{Amp}(\boldsymbol{t})_{\boldsymbol{x},\boldsymbol{y}}\right) \qquad \textbf{Eq. 1}$$

$$\boldsymbol{max}\left(\boldsymbol{Amp}(\boldsymbol{f})_{\boldsymbol{x},\boldsymbol{y}}\right) \qquad \textbf{Eq.2}$$

$$\sum_{1kHz}^{6kHz}\boldsymbol{E}(\boldsymbol{f}) \;\; \boldsymbol{and} \;\; \sum_{7kHz}^{12kHz}\boldsymbol{E}(\boldsymbol{f}) \qquad \textbf{Eq. 3}$$

The road surface condition is related to the amplitude value in Eq.1, if the surface condition is bad, the amplitude is lower; the maximum frequency and frequency energy are used to present whether there is interior damage or not. The delamination happened if the peak and energy were high at 1-6kHz.

4.4.2 Convolutional neural network (CNN) study

For the CNN model, the input data always important, which includes time-domain signal, frequency domain signal, and short-time Fourier transformation (STFT). There are four parametric studies for signal, 1) different signal duration (D) study, 2) different starting time (S) study, 3) different image resolution study (R), and 4) a different number of image studies (N). The input efficiency is used to present the relationship between accuracy and computation time(CP):

$$\boldsymbol{Input}\ \textbf{efficiency} = \frac{\boldsymbol{Normalized\ Acc.}}{\boldsymbol{Normalized\ CT.}} \qquad \textbf{Eq. 4}$$

An example of input studies is shown in Figure 10. There are 10 different duration (D) from 0.2 milliseconds to 2 milliseconds in the signal duration study. The signal starts at a zero-crossing point which is the time right before obtained structural signal. For different starting times (S) study, there are 8 different starting times. The first case starts from the zero-crossing point with three different durations(D=0.3ms, 0.6ms, and 1ms). For later cases, the starting time is delayed 0.1ms compared with the previous case. For example, the second case starting time is 0.1ms late from zero-crossing; the third case is 0.2ms later from zero-crossing as shown in Figure 11. An example of different resolutions is shown in Figure 12. There are 10 different resolutions that start from 10x10 to 500x500. The image resolution affects accuracy and computation time dramatically; thus, how to find the balance to save computation time and obtain good accuracy is the main goal of this study. After the time-domain signal study is completed with the optimal time-signal parameter setting, the frequency-domain signal is combined with the time-signal to become a new input image, as shown in Figure 13.

The number of image used in the parametric study mentioned above, duration(D) study use 3000 images (300images*10cases); stating time (S) study use 2400 images (300images*8cases); image resolution (R) study use 3000images (300images*10cases), and combined study each also use 3000images (300images*10cases). And 70 % of data is used for training a CNN model, and the other 30% is used for testing accuracy.

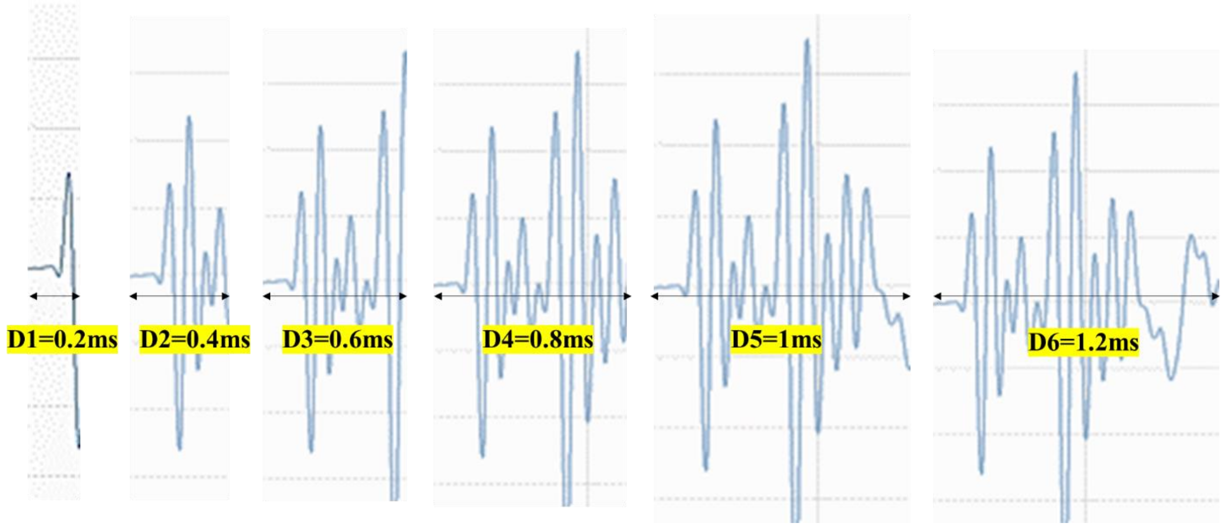


Figure 10. The example signal of different time duration studies. There are 10 different duration (D) from 0.2 milliseconds to 2 milliseconds in the signal duration study. The signal starts at the zero-crossing point, the time right before the obtained structural signal.

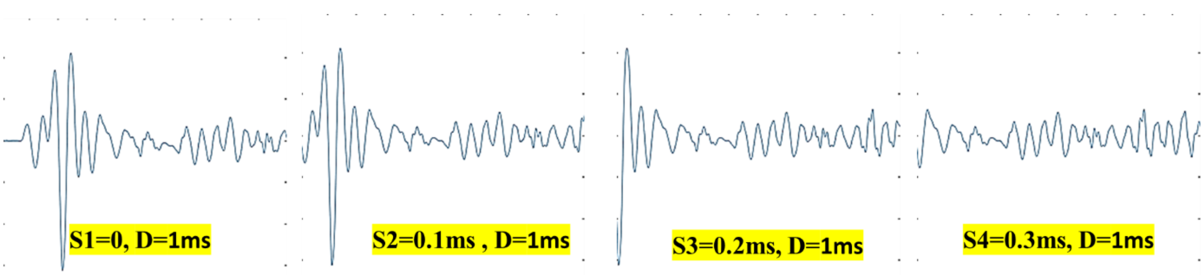


Figure 11. The example signal of different starting time study. There are 8 different starting time, the first case starts from a zero-crossing point, and the second case starting time is 0.1ms late from zero-crossing.

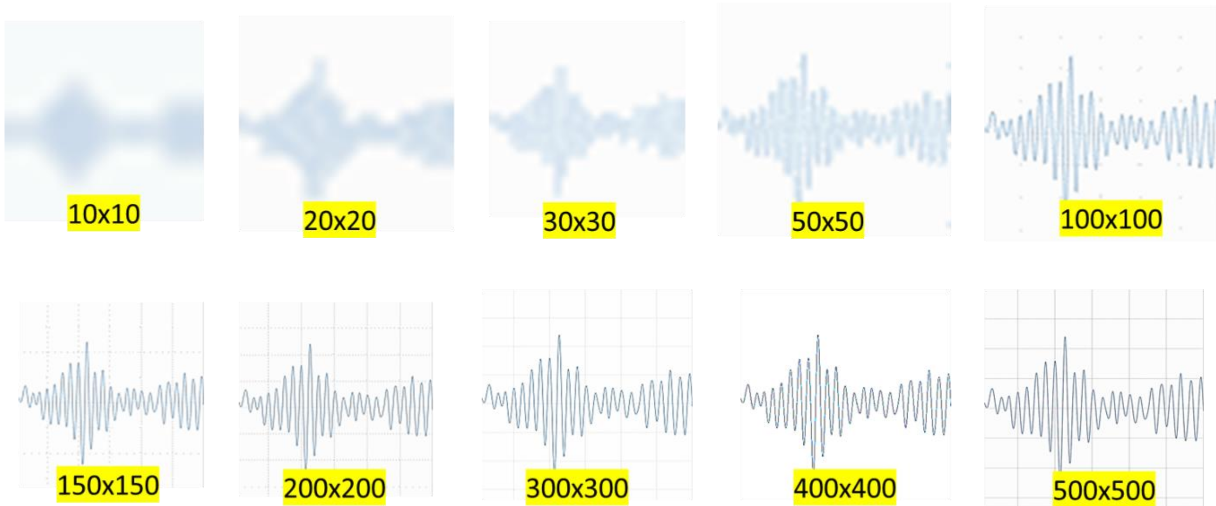


Figure 12. The example of different image resolutions indicates that finding the balance is to save computation time and to obtain good accuracy.

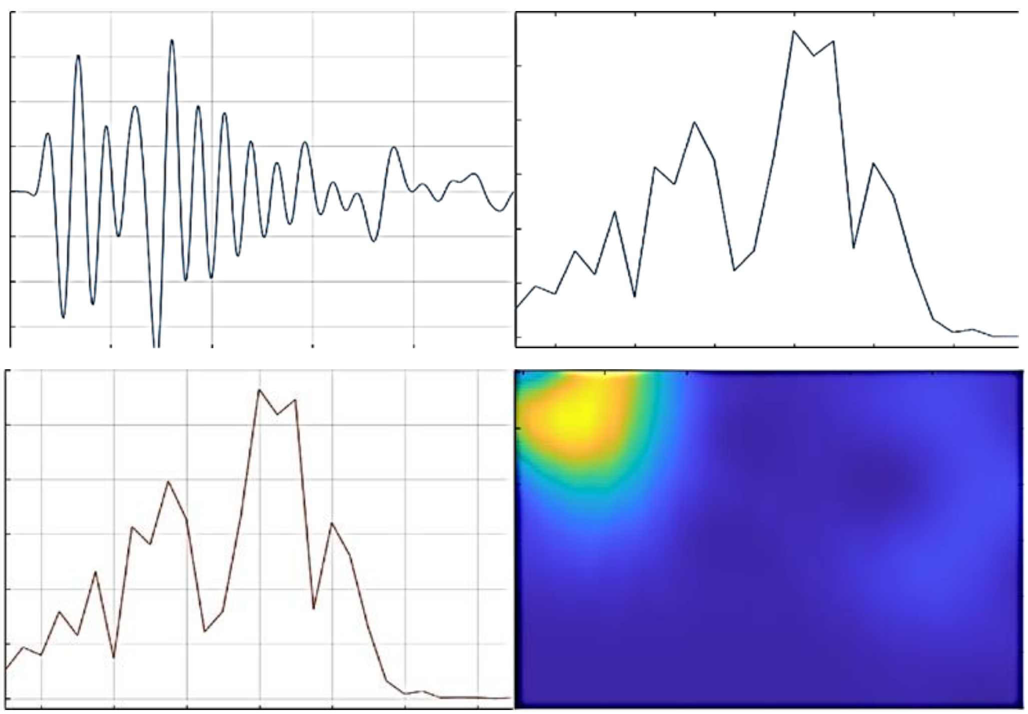


Figure 13. The example of the combined input image for CNN training. Top: time-domain signal + frequency-domain signal (T+F); Bottom: frequency-domain signal + STFT(F+STFT)

5. RESULT, ANALYSIS, AND FINDINGS

5.1 Preliminary study of the CNN model

These results and findings are for Task 1 and Task 5. For a preliminary study of the CNN model, there are three different input models: time-domain signal, frequency-domain signal, and STFT. Note that the preliminary test input hasn't gone through in-depth study yet. In this stage, only for testing whether CNN can work properly with these input signals. The example of the CNN framework is shown in Figure 14. There are several sets of convolution layer and polling layer, which convolution layer is used to calculate with certain algorithm(e.g., averaging value in 5x5 pixels) to find a significant value(as known as a feature) from pixel number. And the pooling layer is used to further pool out the most important feature from the convolution layer. The tail part of CNN is connected with a multilayer perceptron(MLP), which uses features extracted from the pooling layer to find the regression relationship. The preliminary results of the convolution layer as shown in Figure 15. The left matrix (6x6) is manually created by the research team to present pixel numbers in an image. The kernel (3x3) with the average algorithm is used to find the average with those 9-pixel numbers covered by the kernel. After the kernel scans the whole image, the original image (6x6) becomes a convolution feature image (4x4) which represents the significant character of the image. This convolution feature is sent to the pooling layer to extract features further. The pooling layer results are shown in Figure 16; in this case, the kernel(2x2) is using a maximum algorithm that only pools out the maximum feature from the covered pixel to make a new pooling feature map.

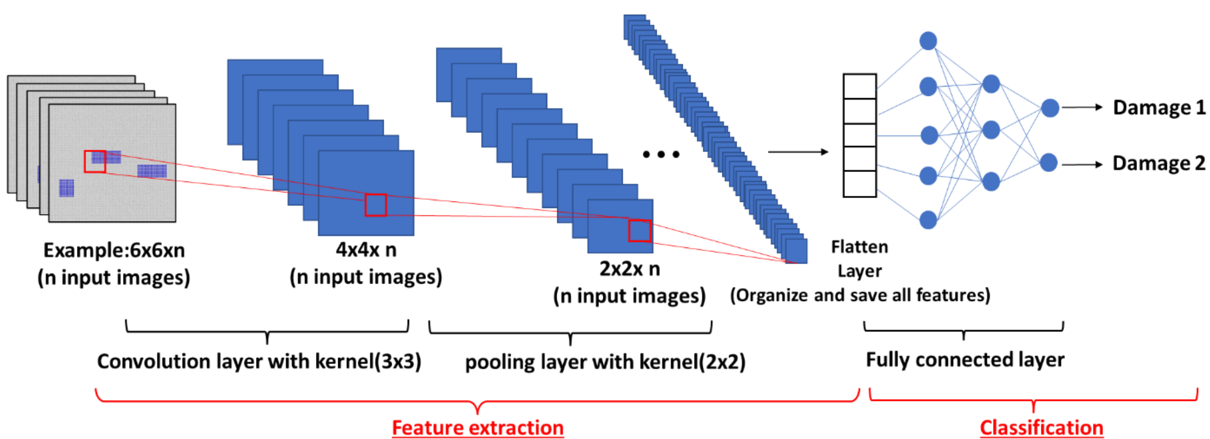


Figure 14. The framework of CNN.

There are several sets of convolution layer and polling layer, which convolution layer is used to calculate with certain algorithm (e.g., averaging value in 5x5 pixels) to find a significant value (as known as a feature) from pixel number. And the pooling layer is used to further pool out the most important feature from the convolution layer. The tail part of CNN is connected with a multilayer perceptron(fully connected layer), which is the layer that uses and classifies features.

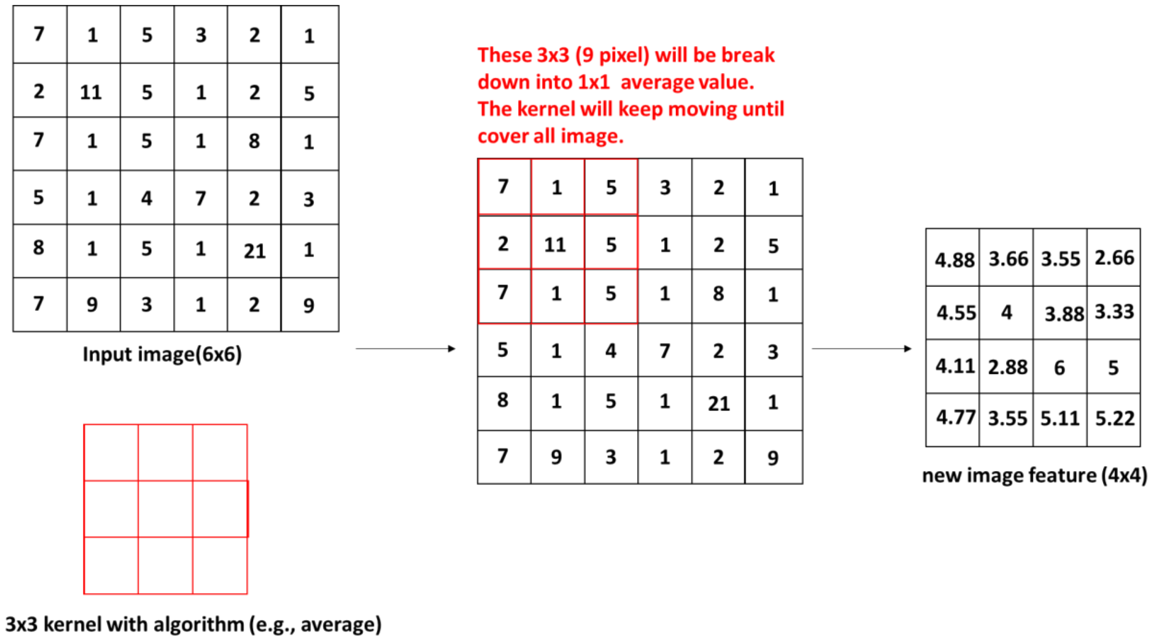


Figure 15. The preliminary results of convolution layer.

In Figure 15, the left matrix (6x6) is manually created by the research team to present pixel numbers in an image. The kernel(3x3) with the average algorithm is used to find the average with those 9-pixel numbers covered by the kernel. After the kernel scans the whole image, the original image(6x6) becomes a convolution feature image (4x4) representing the image's significant character.

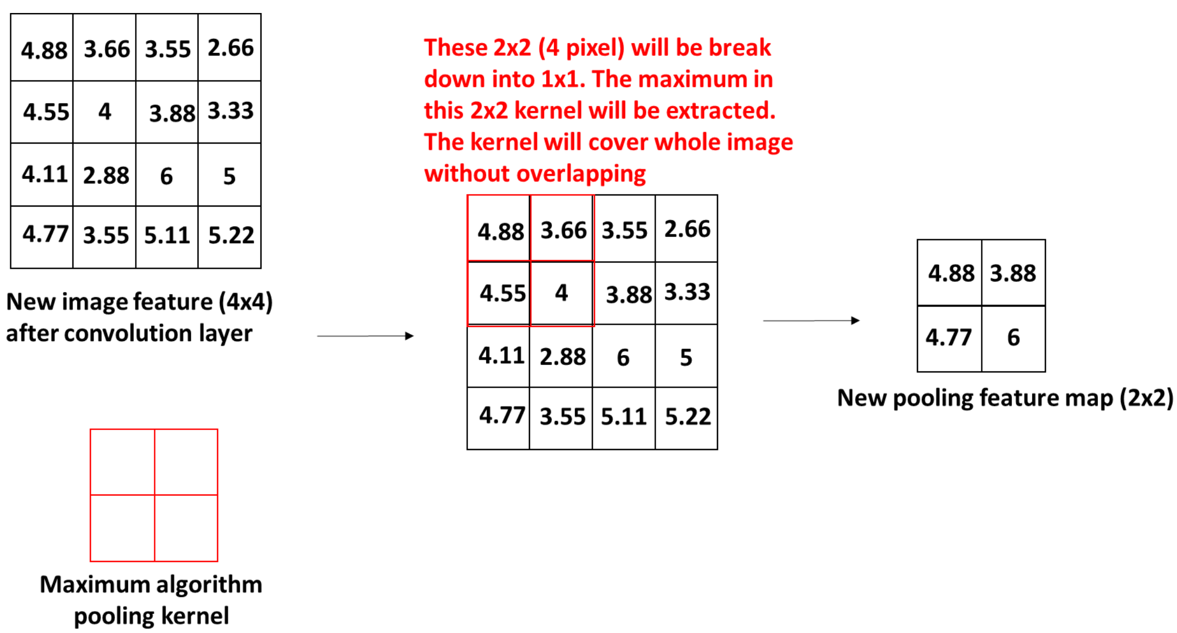


Figure 16. The kernel(2x2) uses a maximum algorithm that only pools out the maximum feature from the covered pixel to make a new pooling feature map.

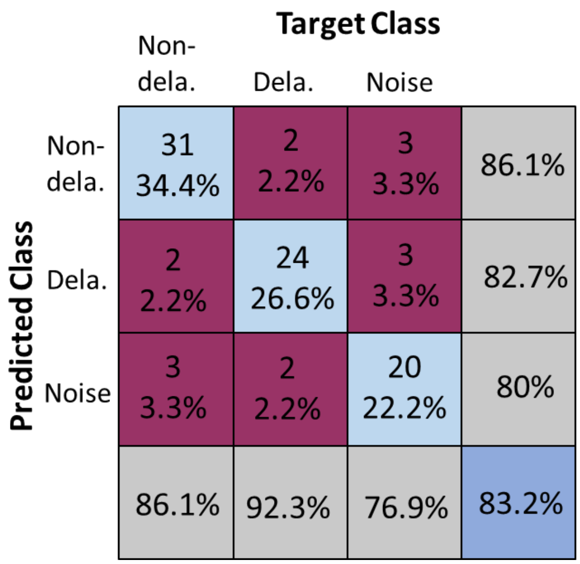
The CNN signal identification accuracy is shown in Figure 17. It should be reminded that there are 210 images for training and 90 images for testing. For e.g., in Figure 17(a), 33.3% is calculated from the correct predicted sample divided by the total testing image (30/90=33.3%); The number of 85.7% on the bottom is calculated from the correct predicted class sample divided by the total class sample (30/35=85.7%); The number represents 73.2% is the accuracy of CNN model, which is a summation of corrected prediction of each class (33.3%+23.3%+16.6%=73.2%). Among three different inputs, STFT has higher prediction accuracy, and it has the highest accuracy for predicting noise class.

Predicted Class \ Target Class	Non-dela.	Dela.	Noise	
Non-dela.	30 33.3%	1 1.1%	5 5.5%	83.3%
Dela.	2 2.2%	21 23.3%	8 8.8%	67.7%
Noise	3 3.3%	5 5.5%	15 16.6%	65.2%
	85.7%	77.7%	53.5%	73.2%

(a) Time-domain signal result

Predicted Class \ Target Class	Non-dela.	Dela.	Noise	
Non-dela.	25 27.7%	1 1.1%	6 6.6%	78.1%
Dela.	1 1.1%	20 22.2%	5 5.5%	76.9%
Noise	4 4.4%	3 3.3%	25 27.7%	78.1%
	83.3%	83.3%	69.4%	77.6%

(b) Frequency-domain signal result

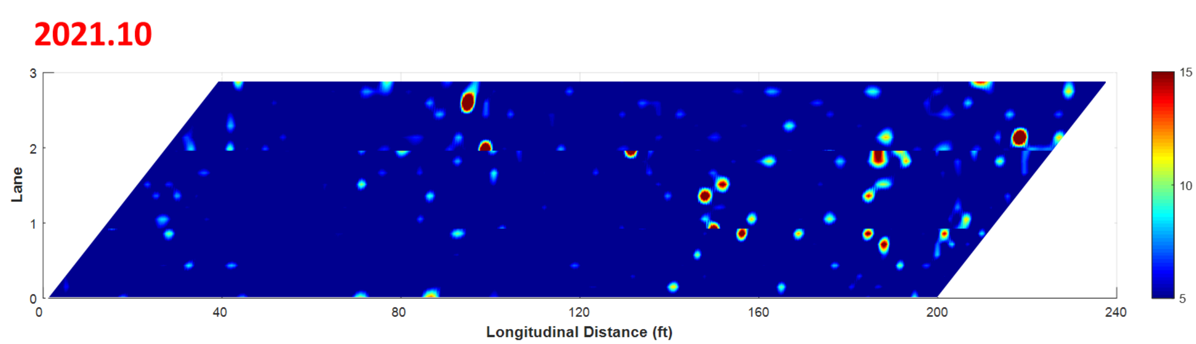


(c) STFT result

Figure 17. The preliminary test of CNN with different inputs.

5.2 Field test results

Tasks 2 and 3 results are presented in this section. The NDT results for Bridges 1 and 2 are provided. The inspection team scanned Bridge 1 three times in 2021 and 2022; thus, the changing of damaged delamination area can be easy to point out. Figure **18**18 shows a delamination map of three different scanning times of bridge 1(First scanning time: 2021.10; second scanning time: 2022.1; and third scanning time: 2022.3). Red area means high energy between 1kHz-6kHz represents delamination; the yellow and light blue also presents delamination but has smaller size, and dark blue is good condition area. The delamination is growing faster between January 2022 to March, while delamination is growing slower between 2021.10 to 2022.1. The delamination map of bridge 2 is shown in Figure 19. As shown in Figure 20, the damage index is calculated from the total value within a 5 ft. length from the delamination map.



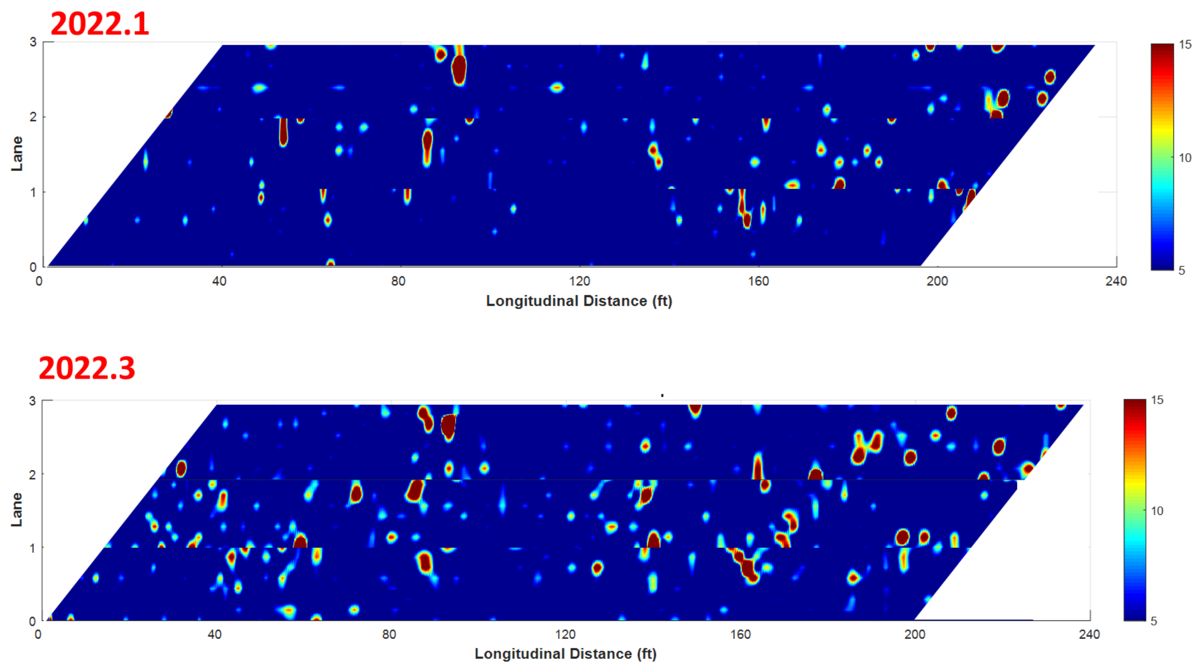


Figure 18. The delamination map of bridge 1 with three different scanning times.

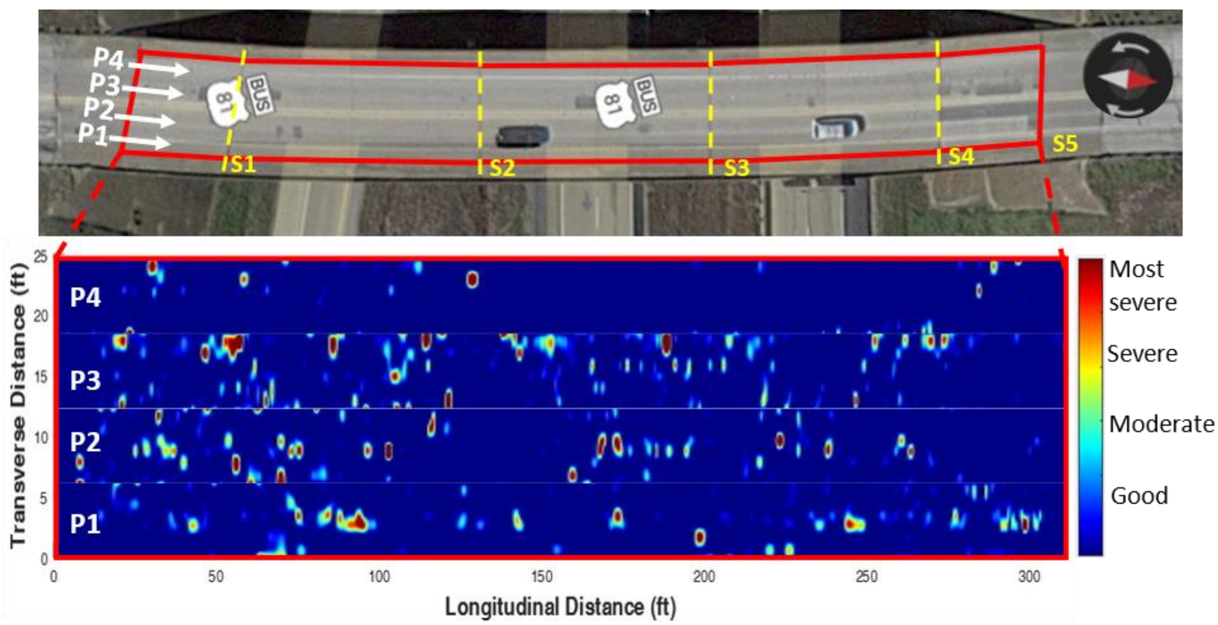


Figure 19. The delamination map of bridge 2.

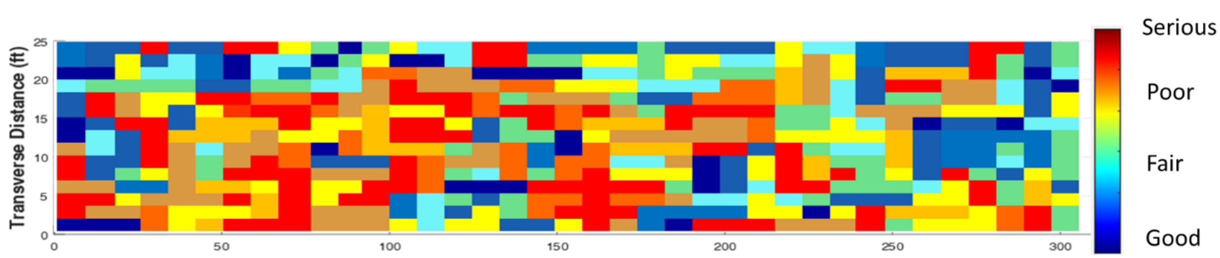


Figure 20. The total energy within a 5 ft. area are calculated for the damage index map

5.4 In-depth Parameter Study of machine learning input with NDT signal

These results are for Task 3 and the first step of Task 5. Because of the various environmental condition and field test uncertainty, the existing signal processing algorithm is not perfectly fitting with different bridges. To improve NDT results (e.g., delamination map), CNN is used to identify three different types of signals: delamination signal, non-delamination signal, and noise(or insignificant signal). Based on this classification, we have four types of in-depth signal study: different signal duration (D), different signal starting time (S), different image resolutions (R), and a different number of images (N). By using these parameters, the relationship between CNN accuracy and computation time can be found. It can provide the optimal input for the CNN model to identify signals.

The results of different signal duration (D) are shown in Figure 21. D=1ms has higher prediction accuracy as, within 1ms, the important impact signal reflected structural interior damage. On the other hand, D=2ms has lower accuracy as it covers too much insignificant signal(e.g., tail resonance signal or noise).

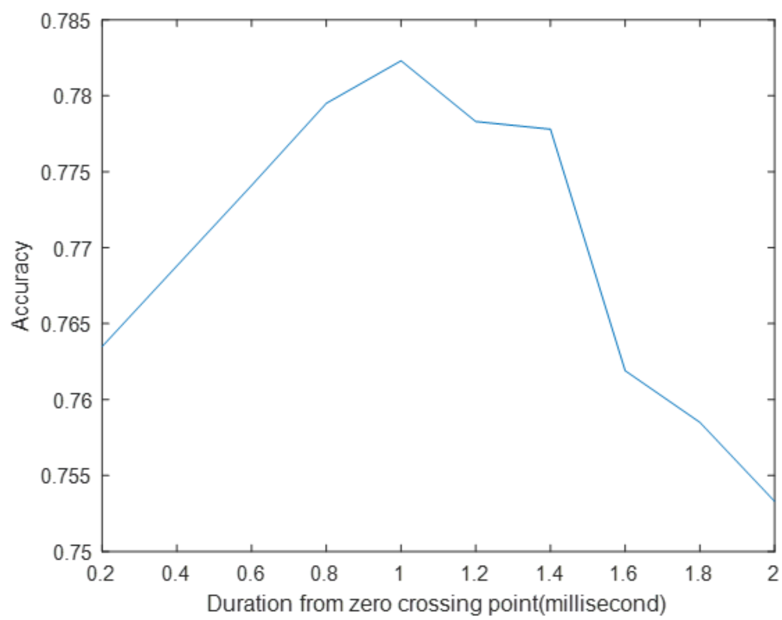


Figure 21. The result of different signal duration studies.

The results of different signal starting times (S) with three different D are shown in Figure 22. Six results all show a similar pattern: the highest accuracy appears when S=50(0.1ms) or 100(0.2ms) because the front part surface wave, which dramatically affects signal behavior, is ignored. Besides, the significant signal always happens right after the surface wave. On the other hand, S=250(0.5ms) gives lower accuracy as the signal only covers noise as a target.

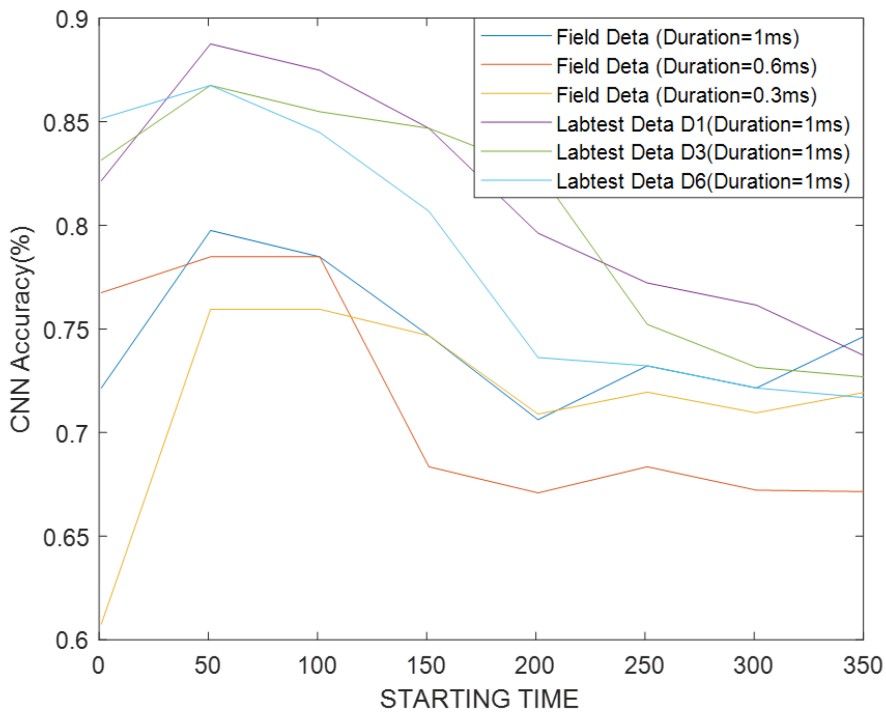
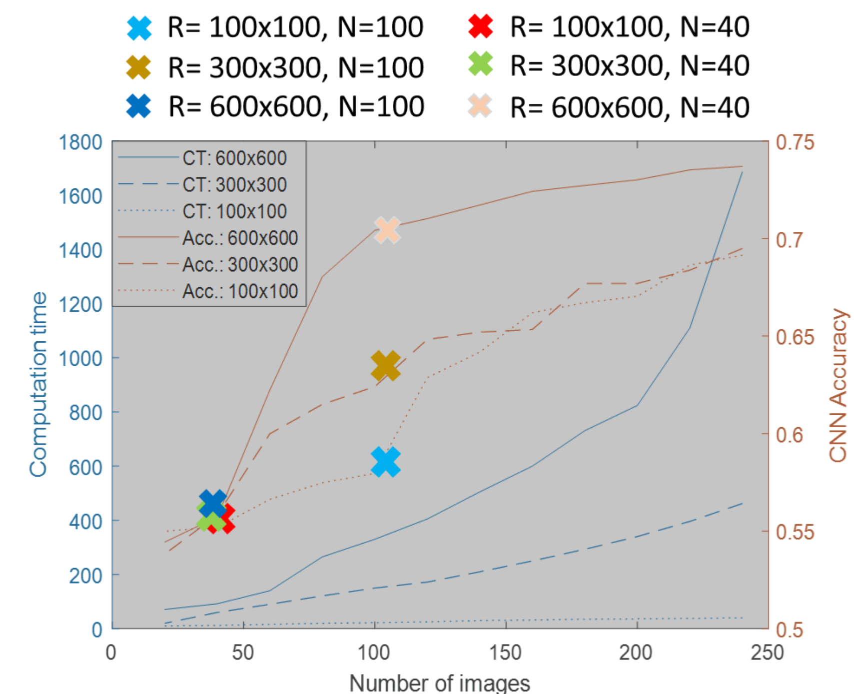
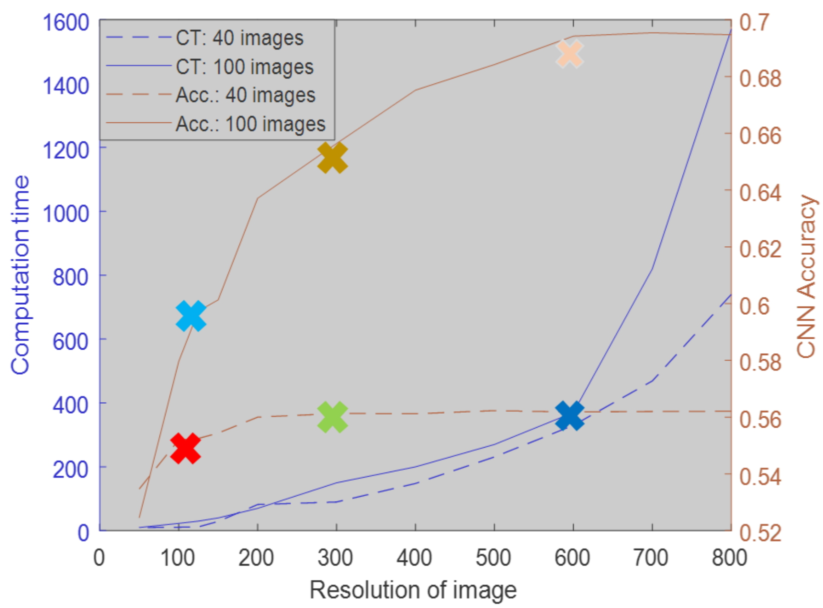


Figure 22. The results of different signal starting times (S) with three different duration and laboratory test delamination.

The results of image resolution and number of an image are shown in Figure 2323. Figure 23(a) shows that the accuracy (Acc.) reaches the threshold when N=100, while computation time (CT.) suddenly increases after N=200. Figure 23(b) shows the accuracy reaches the threshold when R=600 and computation time suddenly increases after R=600. This information helps to understand the relationship between the CNN model and input data. Finding the threshold of accuracy and computation time is important as it represents the efficacy of the CNN model. For e.g., in Figure 23 (b), although the accuracy slowly increased when R increased from 300 to 600, the computation time increased dramatically. Similar pattern can be obtained from Figure 23(a) when N increased from 100 to 200, the accuracy increased slowly, but computation time increased significantly.



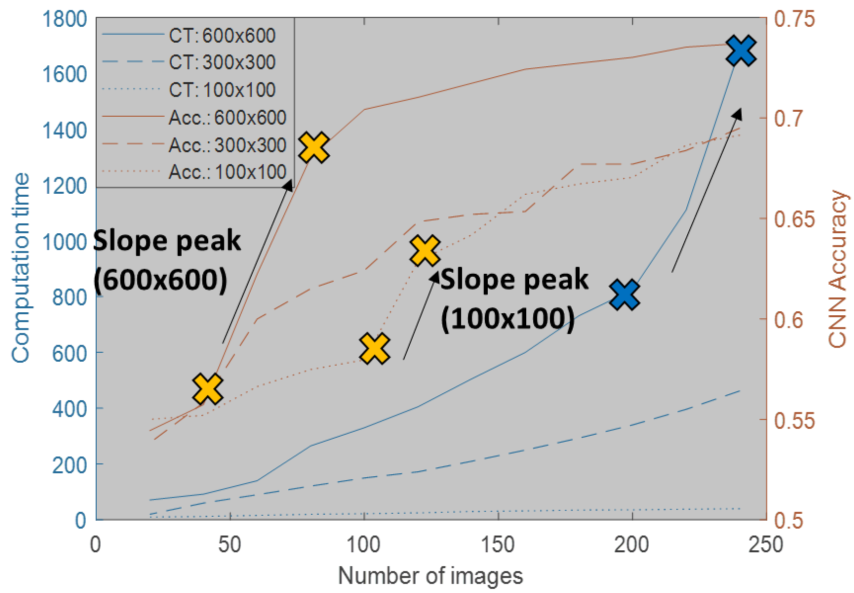
a. Number of images study with three different image resolution



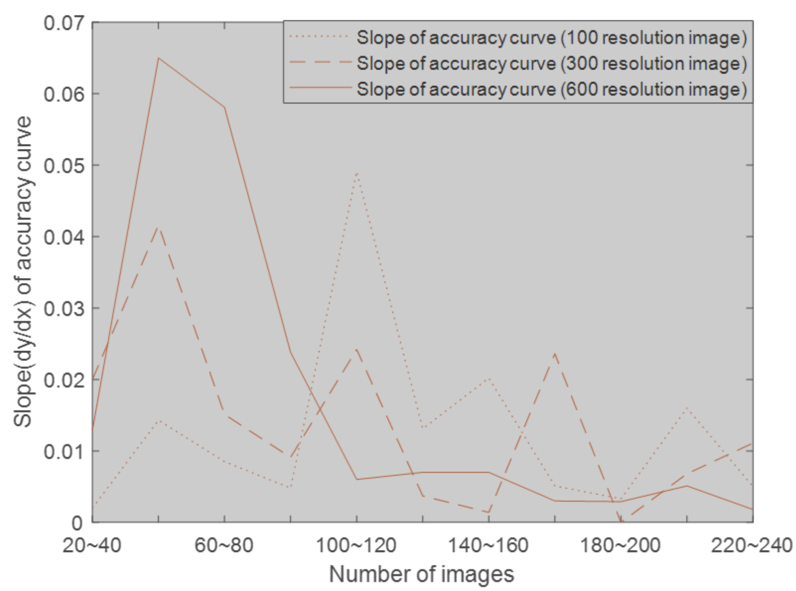
b. Same image resolution with different number of images

Figure 23. The result of image resolution(R) and a number of images(N) study.

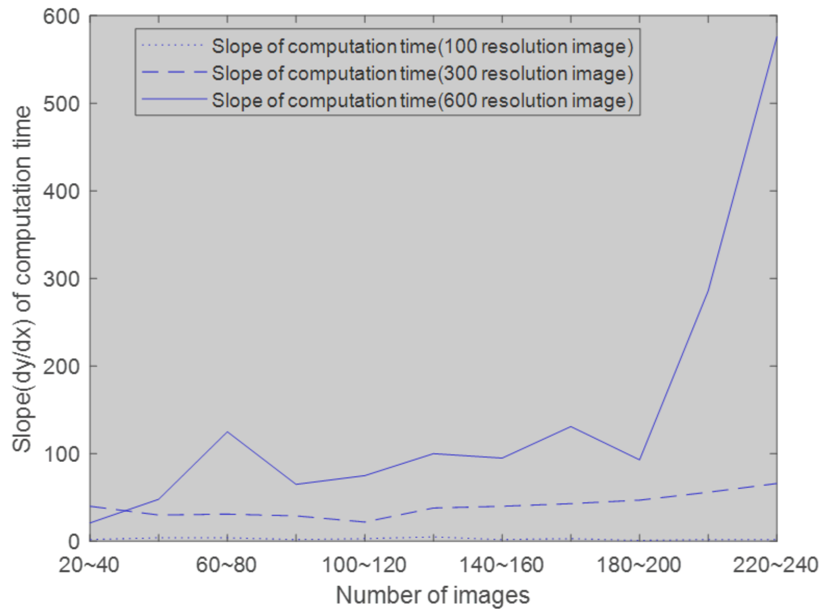
Further slope analysis of a number of images with three image resolutions is shown in Figure 24. According to Figure 24(b), the first peak occurs when N=40~60 and when N=100~120, which means the accuracy increased dramatically between these cases (marked yellow in Figure 24a). Only the R=600 case didn't have a second peak because it had already reached a threshold which means accuracy won't increase dramatically if using more N. According to Figure 24(c), when R=100x100 and 300x300 with N=120~140 images, the slope of computation time shows the highest value, which means the computation time suddenly increased. The same situation can be found in (blue mark in Figure 24a). Besides, If N> 140 images, the computation time increases exponentially.



a. Number of images study with three different image resolution



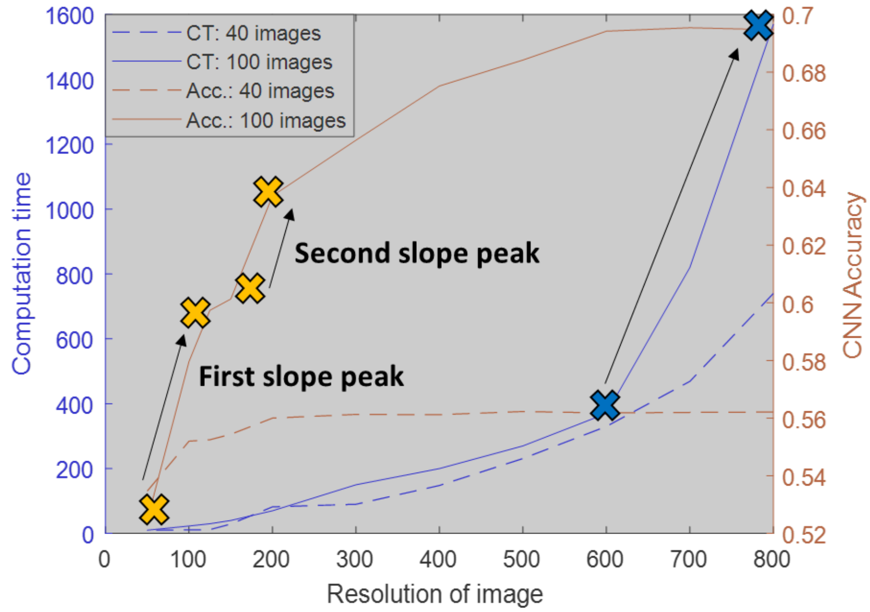
b. Slope of accuracy curve (Number of images study)



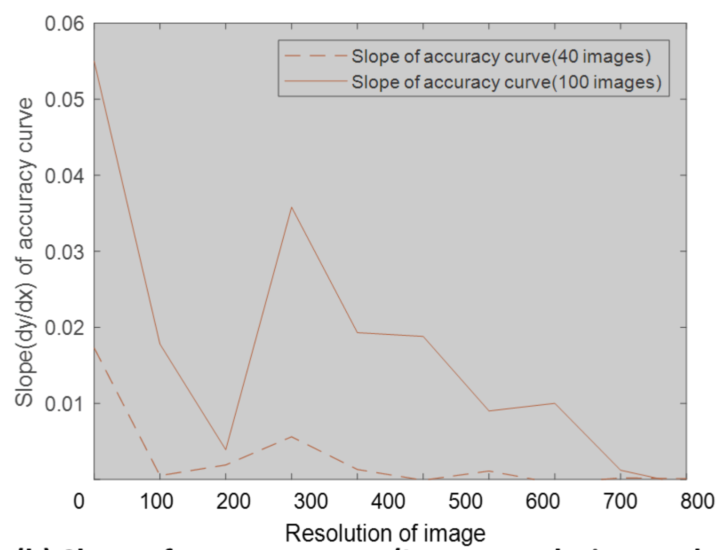
c. Slope of computation time (Number of images study)

Figure 24. The slope analysis of a number of image studies (N) with three different image resolutions.

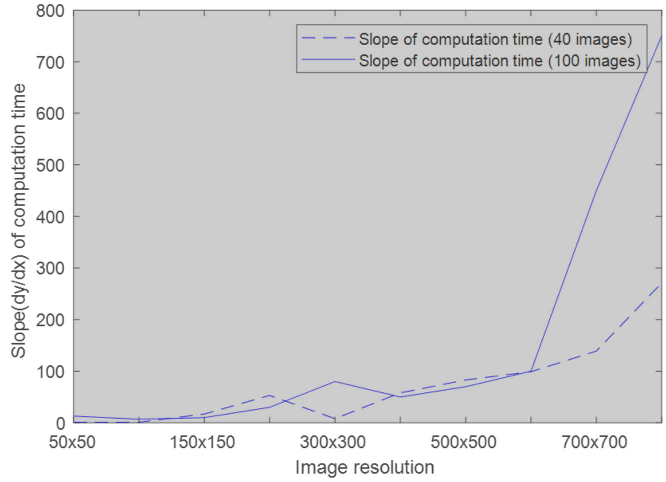
The slope analysis of image resolution is shown in Figure 2525. According to Figure 25(b), when image resolution around R= 200x200 shows the peak, the CNN accuracy suddenly increases (marked yellow in Figure 25a). If R>200x200 images, the slope is decreased, which means the accuracy has already reached the threshold when R=200x200. According to Figure 25 (c), the slope exponentially increased after R=600x600 with the N=100 image case. For N=40 image, the computation time slope also started to increase when R=700x700, which means the CNN model efficiency is getting lower when using R≥600x600 images.



(a)Image resolution study with two different image data size



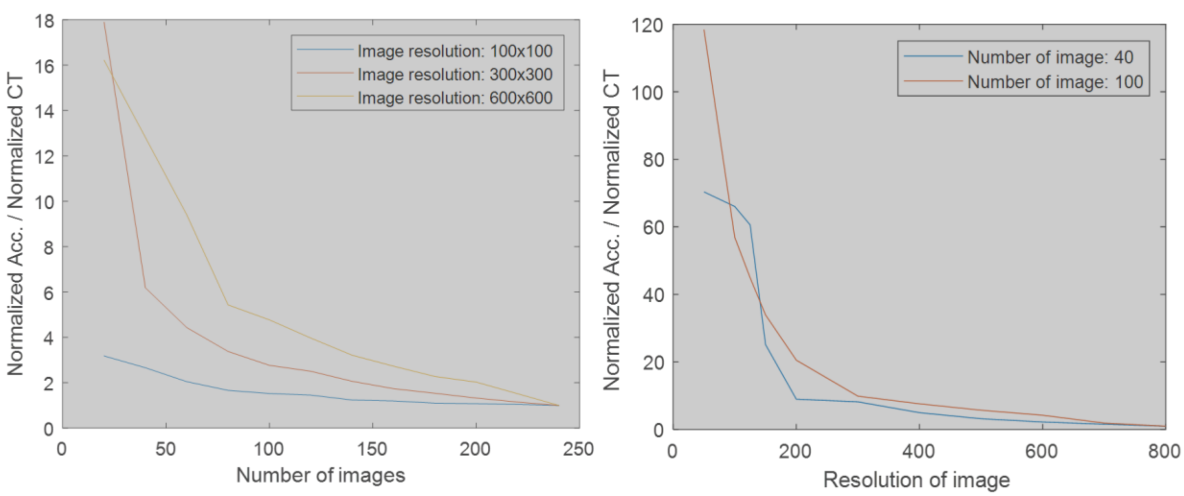
(b) Slope of accuracy curve (Image resolution study)



(c) Slope of computation time curve (Image resolution study)

Figure 25. The slope analysis of image resolution study.

Based on Figure 24 and 25, the CNN efficiency is shown in Figure 26. The higher value represents the good efficiency of the CNN model. According to Figure 26(a), the efficiency is higher when the number of images N≤ 100; however, there is still a need to consider the accuracy since accuracy is lower when N is smaller. Accuracy reaches the threshold when N≅100 (Figure 24 a). According to Figure 26 (b) efficiency of resolution study shows there is higher efficiency when R ≤ 200. The accuracy threshold of image resolution is 200x200 (Figure 25 a). After an in-depth study of an input signal, the optimized input parameter is found. Signal duration, D is 1ms, starting at 0.1ms after the zero-crossing point with image resolution R=200x200 and N=100 images as our input setting. By using these parameters, the NDT results(e.g., delamination and damage index) are checked and improved with the CNN model.



(a) Efficiency of N study

(b) Efficiency of R study

Figure 26. The result of CNN efficiency.

The higher CNN efficiency value represents the good efficiency of the CNN model. However, there is still a need to consider the accuracy value present in Figure 34 and Figure 35.

The combined input results as shown in Figure 27. There are three different combined siganl:1) C1:frequency and STFT; 2) C2:time-signal and frequency and 3) C3:time signal, frequency, and STFT. ANN results are also shown and compared. C1 gives the highest accuracy, while C2 has lower accuracy. Compared to C2 and C3, the time signal makes accuracy decrease because the noise may affect the signal dramatically. This pattern fits with the time-signal starting time study (Figure 22) when S=50 (0.1ms) and 100 (0.2ms) have the highest CNN accuracy. For ANN, when S=200(0.4ms), accuracy suddenly decreased because the signal started to be affected by tail part noise, which is similar to Figure 22.

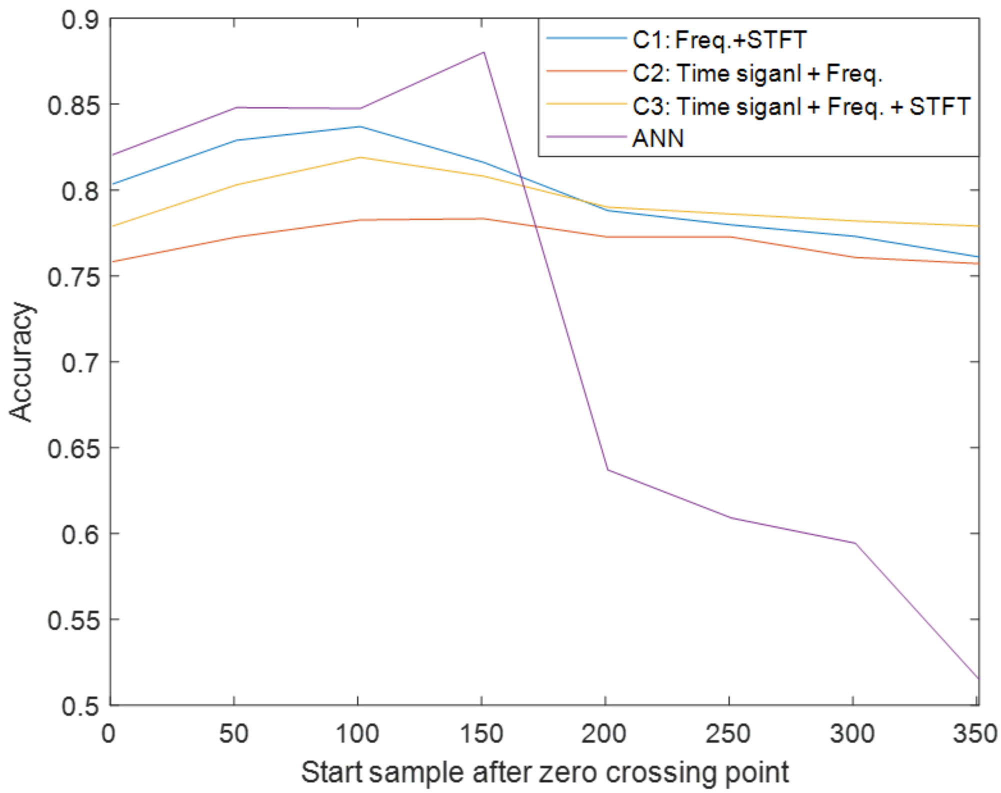


Figure 27. The CNN results with the combined input image.

5.3 FE Modeling result for Structural Performance

This result for Task 4. The FE model result is mainly used to calculate structural performance value(e.g., stress) to further calculate damage levels for damage prediction. The other purpose of the FE model is to understand the relationship between delamination and structural performance, especially in nearby girder areas. The delamination FE model is designed by the delamination inspection result, as shown in Figure 28. The artificial delamination is created in between the 1st and 2nd layers, which is a 2-inch depth from the top surface. The contact constraint enforcement method (called contact method) of Abaqus/Standard is considered to simulate delamination areas by using a stiff approximation of hard contact (penalty method), which is the condition that the upper element cannot penetrate the lower element. The spring-damper-mass system, which consists of discrete elements interconnected with spring and damper, is adopted to define the surface behavior. The penalty stiffness and damping coefficients are provided to distribute the load to the lower element. The contour plot from these simulation results is shown in Figure 29 and

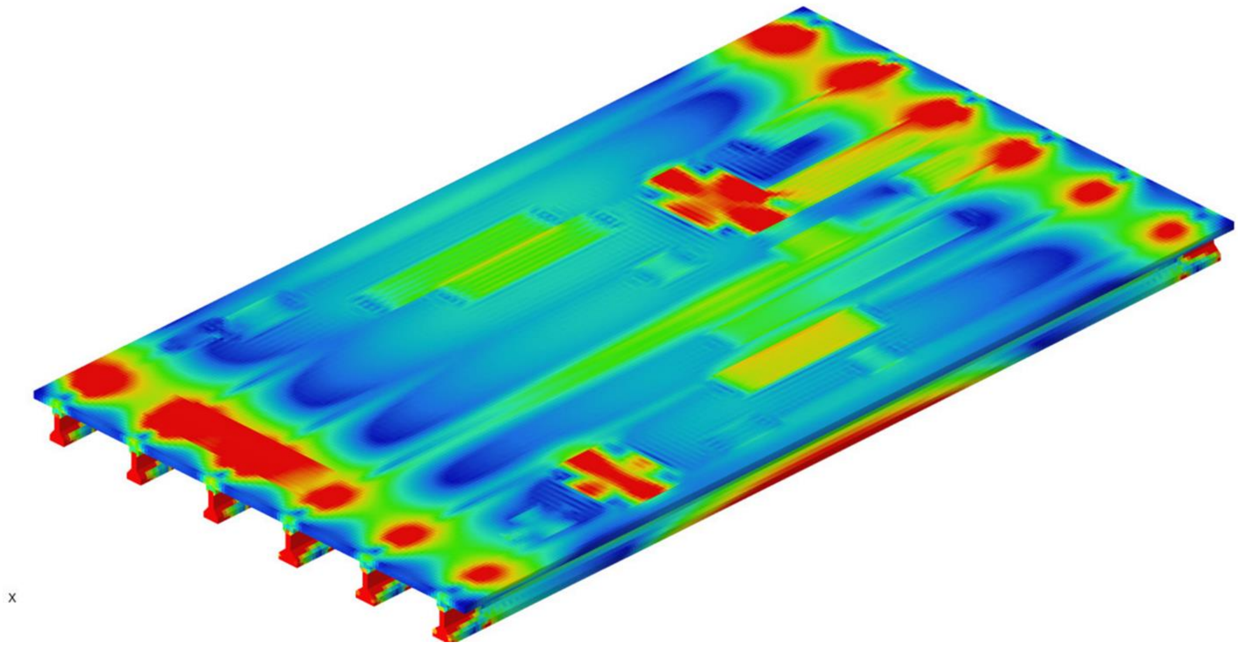


Figure 30 with a color index. The results of the FE model with delamination show that the maximum principal stress (MPS) occurs at the bottom surface. The maximum principal stress distribution is one of the most important structural performance analysis methods related to concrete crack behavior. The delamination model shows a different load distribution than the control model on the artificial delamination region. The highlighted delamination shape on the top surface indicates the delamination damage affects the load distribution. From these simulation results, 28 elements are chosen from each model to compare the MPS at the elements. The detail of the elements is described in Figure 29.

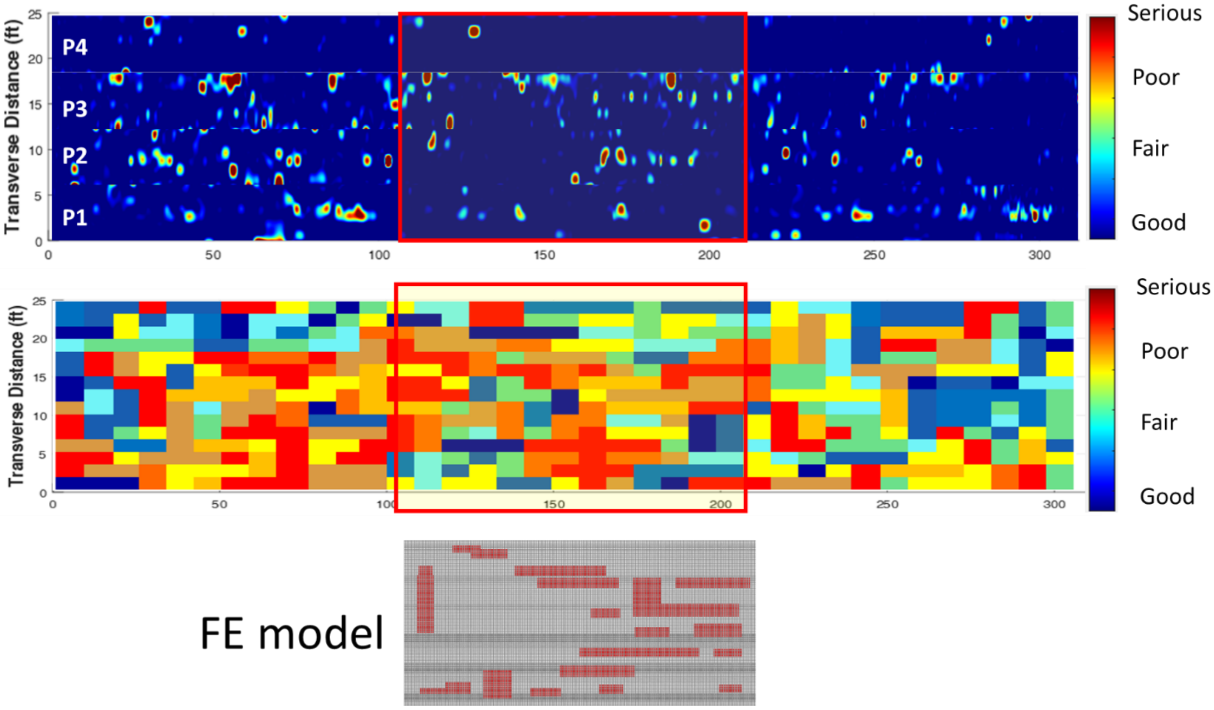


Figure 28. Delamination FE models are designed by field test results

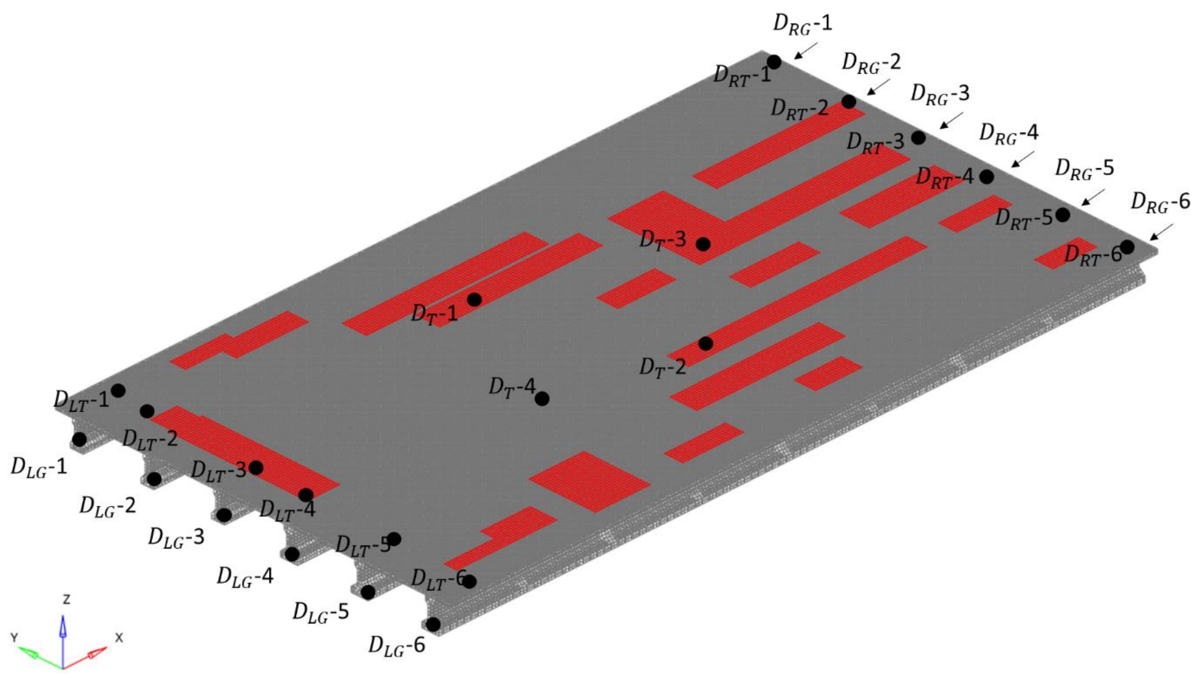


Figure 29. The picking element locations: left side top layer element (D_{LT}-1-6), right side top layer element(D_{RT} -1-6), left side girder element (D_{LG}-1-6), right side girder element (D_{RG}-1-6), and element

on the delamination(D_T-1-4).

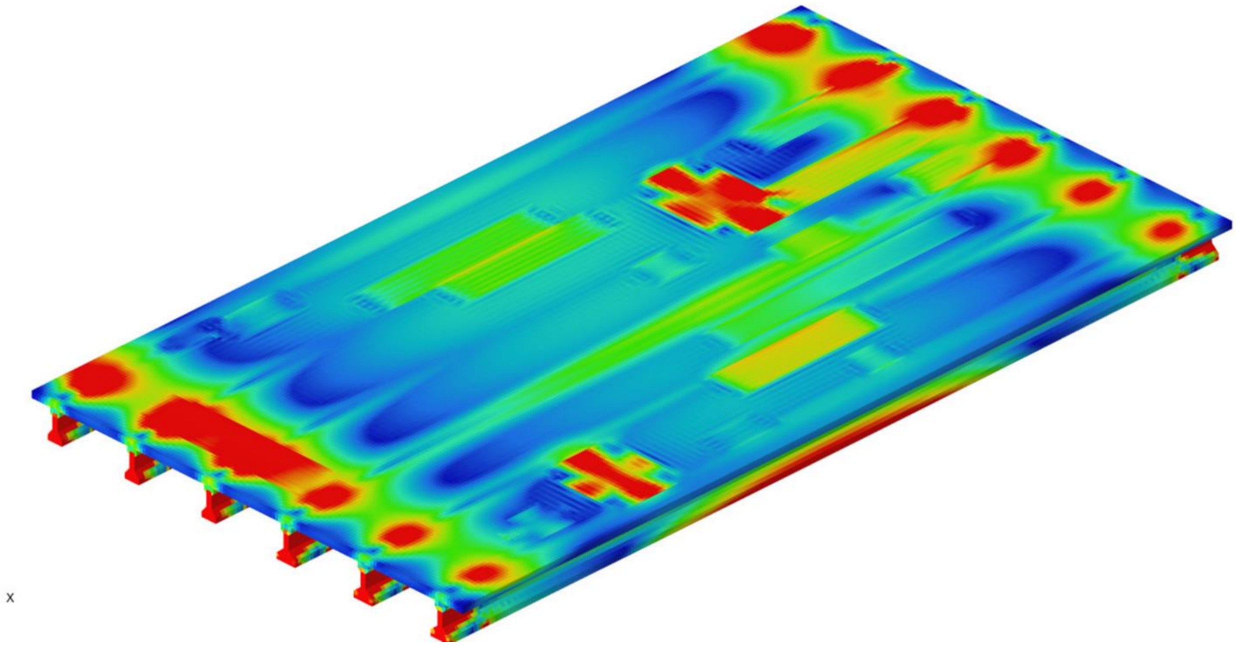


Figure 30. The stress map of the delamination model

The calculated stresses at the elements are described in **Error! Reference source not found.**. Most of the selected elements show more increased stress in the delamination model than in the non-delamination model. Elements DT-1, DT-2, DT-3, and DLT-4, which are on delamination centers, show the dramatically increased MPS. In addition, the relatively higher stress increment at DLG-3 (10 %) indicates the delamination on the girder affects the load distribution to the girder from a slab, and it is critical for structural performance.

Table 1. Maximum Principal Stress of the selected elements

	Maximum Principal Stress (N/m^2, Pa)		Increase (%)
	Delamination	Non-Delamination	

D_{LT}-1	125332	120202	4.26
D_{LT}-2	129105	123730	4.34
D_{LT}-3	115990	61675	88.06
D_{LT}-4	97494	93752	3.99
D_{LT}-5	90105	84650	6.44
D_{LT}-6	73725	72406	1.82
D_{RT}-1	125332	119692	4.71
D_{RT}-2	130587	123983	5.32
D_{RT}-3	118397	114960	2.98
D_{RT}-4	110144	105873	4.03
D_{RT}-5	101460	95937	5.75
D_{RT}-6	84029	79047	6.30
D_{LG}-1	265772	247988	7.17
D_{LG}-2	283465	266830	6.23
D_{LG}-3	281978	254880	10.63
D_{LG}-4	290788	289176	0.55
D_{LG}-5	280997	273320	2.80
D_{LG}-6	236347	233441	1.24
D_{RG}-1	265772	249579	6.48
D_{RG}-2	283465	269328	5.24
D_{RG}-3	281978	269250	4.72
D_{RG}-4	290788	291020	-0.07
D_{RG}-5	227549	225355	-7.90
D_{RG}-6	236347	231913	1.91
D_{T}-1	83167	334	24800.29
D_{T}-2	91266	7911	1053.65
D_{T}-3	72345	32117	125.25
D_{T}-4	295	300	-1.66

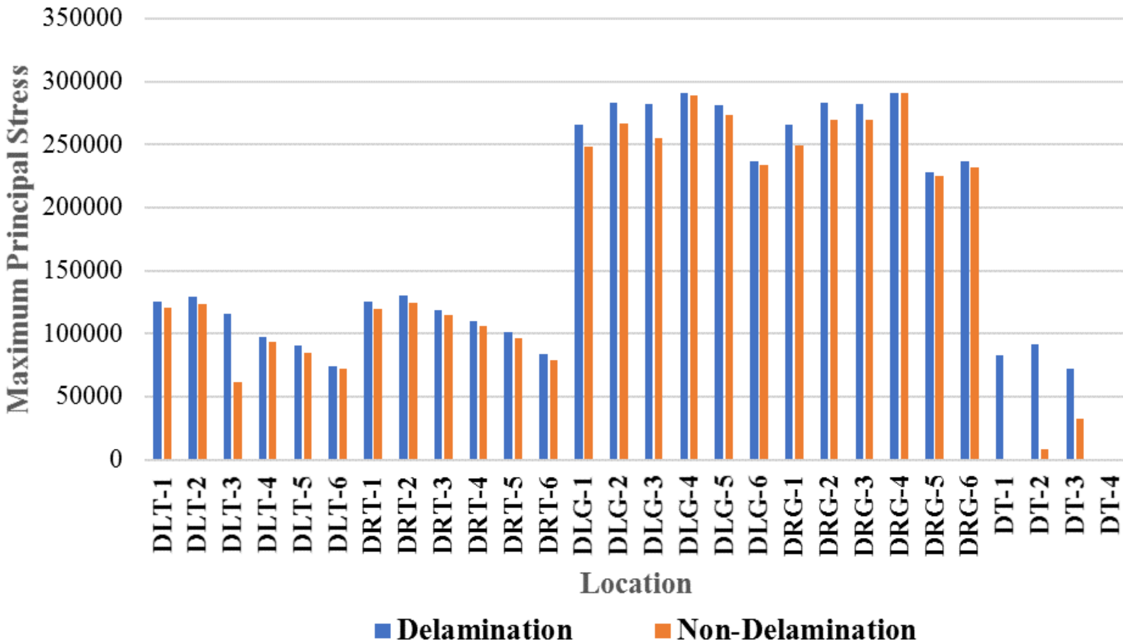


Figure 31. MPS bar chart of the selected elements in the delamination simulation

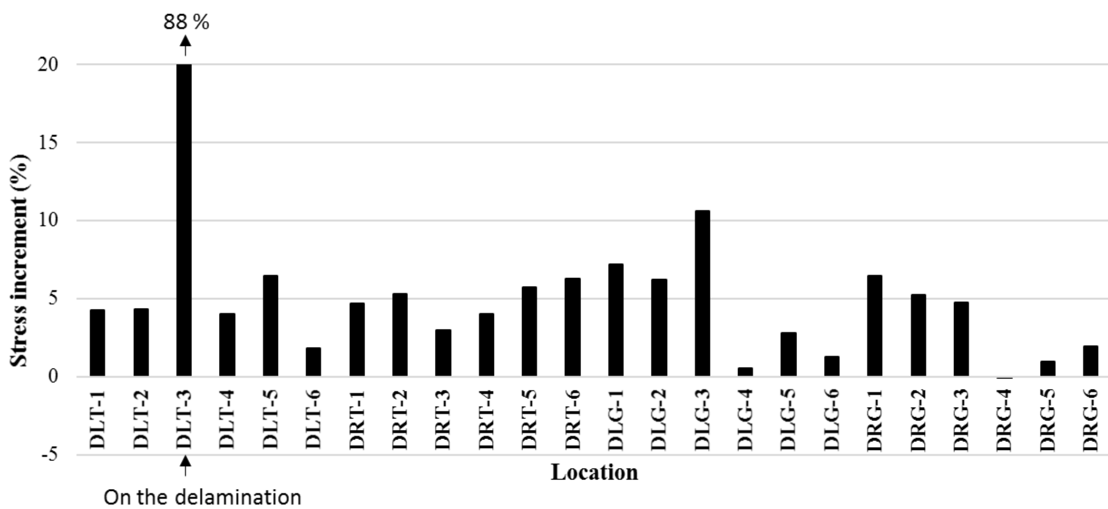


Figure 32. MPS increment bar chart of the selected elements in the delamination simulation

5.5 Structural Performance Prediction Result

These results are related to Task 5. The structural performance prediction model has four different damage levels: severe stress concentration, moderate stress concentration, mild stress concentration, and healthy. According to Figure 33, the delamination FE model is designed and represents to different damage level by delamination size, element interaction scale value of different types of delamination, and the number of delamination shown in Table 2. The 110 FE models are created with these random parameters. Generally, if the model has more delamination with a bigger delamination size and lower scale value, then this model represents a severe damage condition because the stress value will concentrate around delamination. The damage level classification is labeled by FE simulation stress analysis. The four damage levels are calculated from $\boldsymbol{Damage\ level\ difference} = \frac{(S_M - S_m)}{\boldsymbol{4\ levels}}$ Eq 5, the summation maximum stress case minus minimum stress case then divided by 4 levels. Thus, the damage level of all 106 models can be defined as shown in **Error! Reference source not found.**: Healthy ≤ 2.996; 2.996 < Mild damage ≤ 3.119; 3.119 < Moderate damage ≤ 3.242; 3.242021 <Serious≤ 3.365 as shown in Table 3.

70% of 110 models are used to train the model and 30% for test model accuracy. The input for damage prediction is FE delamination model; the output is the classification and prediction of level(Figure 3). The advantage of this approach is, that it does not require simulation results, but just using a model can be done performance prediction. The efficiency is better since the simulation time is eliminated. Moreover, the data-based ANN was also applied for damage prediction to compare with CNN results. The used features in ANN are the location of delamination, size of delamination, element interaction scale value of different types of delamination, and the number of delamination. The biggest difference between CNN and ANN is CNN extracts image features in the convolution and pooling layer, and in ANN, a user needs to provide features themselves.

Figure 33. The example of artificial delamination FE model with different parameters

Table 2. The design parameter for artificial delamination FE model

Delamination size	Number of delamination	Contact scale value
10x10	1	0.1
20x20	2	0.5
25x25	3	0.8
15x20		

$$\boldsymbol{Damage\ level\ difference} = \frac{(S_M - S_m)}{4\ levels} \qquad \textbf{Eq 5}$$

Table 3. The four stress concentration level calculated from the FE model

Healthy	Mild concentration	Moderate concentration	Serious concentration
≤ 2.996	2.996 < x ≤ 3.119	3.119 < x ≤ 3.242	3.242 < x

The prediction results of CNN and ANN are shown in **Error! Reference source not found.**. ANN prediction accuracy is 78.1%, and CNN is 68.8%. Take Figure 34 (a) as an example; the target class means "real classification" of severity level, and the predicted class is the "answer" from CNN. There are 11 samples of moderate sample target class; however, ANN gives two wrong samples as an answer of severe class, while the other 9 samples are in the correct prediction class. The accuracy percentage of 28.1%, colored in a light blue box, means the correct CNN answer sample divided by the total test sample (9/32*100%=28.1%). The 60% of the gray box means the prediction accuracy only in the healthy class(3/5*100%=60%). The total ML accuracy is a summation of accuracy in each class (9.37%+28.1%+21.8%+18.7%=78.1%). Compared with ANN, CNN has lower accuracy, which means the input image requires more detail of delamination. For example, more fine resolution of delamination size, more colors to represent damages, and more field test delamination maps. These studies should be considered as a future work plan.

For the further study of the ability to identify different levels of CNN and ANN prediction, the input samples are fixed in both models (severe:20, moderate:26, mild: 22, healthy: 10 samples). Once they use the same sample to train the model, it is clear to demonstrate which model is better for predicting which type of structural performance. The results, as shown in Figure 35, for the ANN model has 77.8% accuracy in predicting severe levels, 78.5% in predicting moderate levels, and 50%,33.3% for mild and healthy levels. Compared with ANN, CNN has a higher ability to detect severe levels, which has 88.8% accuracy, although CNN has a lower prediction for moderate levels. For mild and healthy cases, CNN has 33.3% accuracy for both.

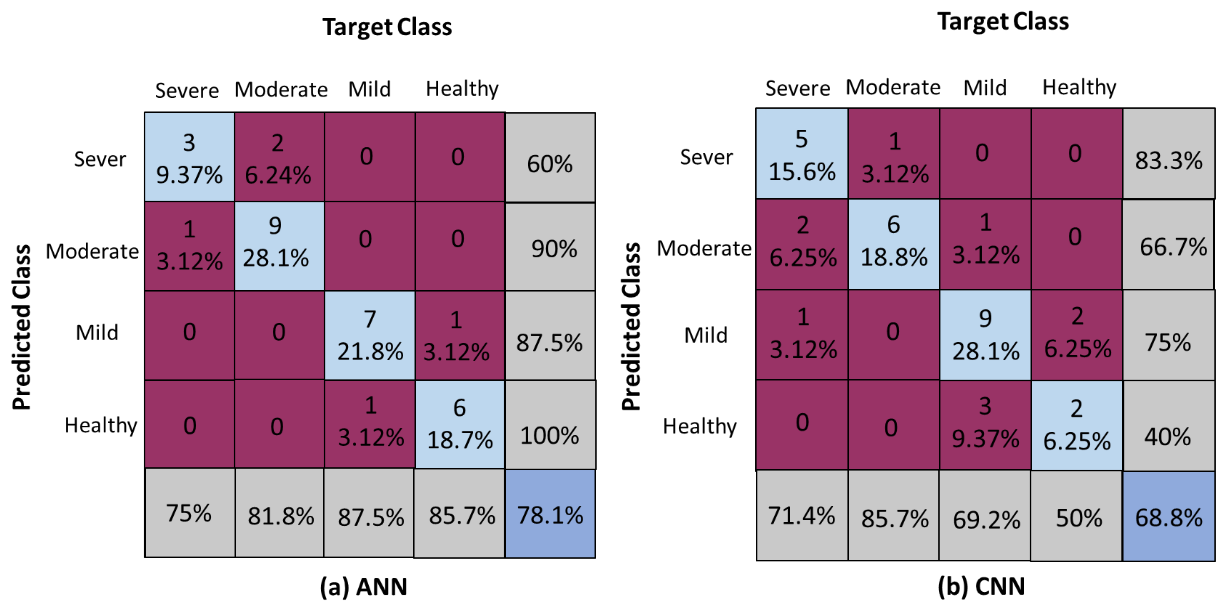


Figure 34. The damage prediction accuracy of (a) ANN, (b) CNN.

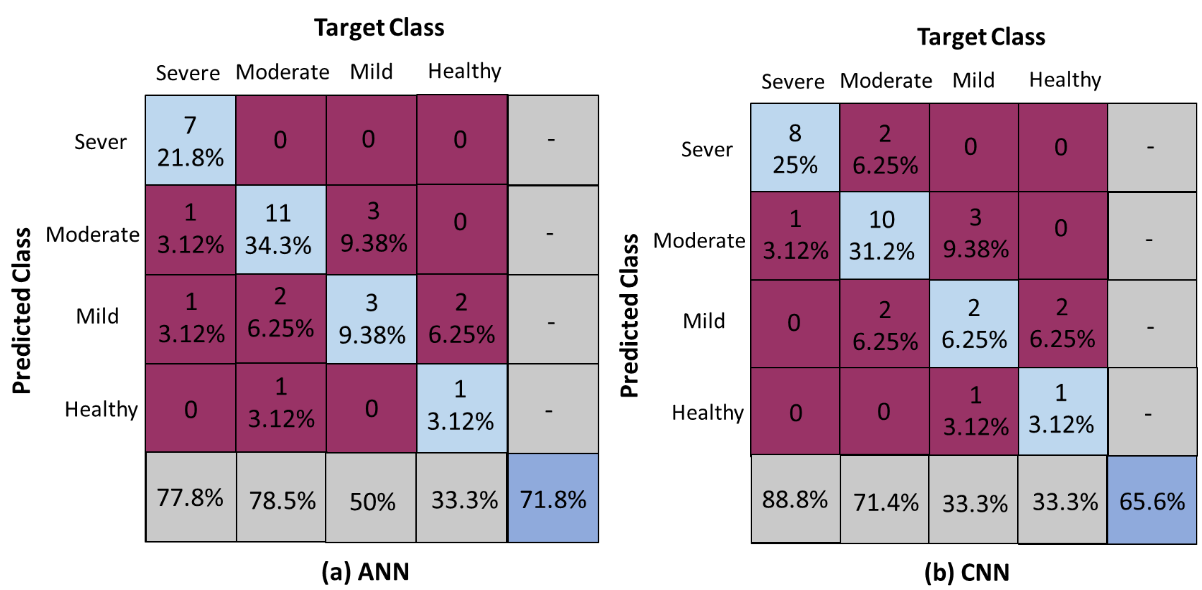


Figure 35. The accuracy in predicting different levels of structural performance.

6. CONCLUSIONS

In this study, the NDT filed test result is improved by using CNN with different types of signal data to reduce noise and insignificant signal to decrease the error to identify delamination from the existing algorithm. The in-depth study of field test data with CNN provides an idea of how to train or input data of NDT signal to obtain a proper prediction model. The relationship between accuracy based on CNN process and field scanned signal is found. For time-domain signal, the optimal design for the CNN model is signal duration D=1ms, starting time S=0.1ms after a zero-crossing point, with image resolution R=200x200 and number of images N=100. The CNN model efficiency is also calculated from accuracy and computation time, which shows when $N \leq 100$ and $R \leq 200$ have higher efficiency. Moreover, using FE structural performance and NDT damage index can identify and predict the different levels of structural performance, with 93.7% accuracy in the ANN model. <u>The study demonstrates the framework of NDT using AI technology to improve 1) data post-processing procedure and also 2) provide the idea of damage prediction with an FE model delamination map</u>. In the future, once the bigger database is built with field test data and an FE model, the ML approach developed in the project can be further improved and applied.

- Firstly, NDT results can be improved by using signal images applied with CNN.
- Secondly, an in-depth CNN study of the NDT time-domain signal provides an optimal image parameter for damage identification.
- Thirdly, FE model 2-dimension top-view, which is designed by field test delamination map and artificial damage maps are input for CNN can be used to predict structural performance.
- For different levels of prediction, ANN has 78.1% accuracy, while CNN has 68.8% accuracy.
- For severe level prediction, ANN has 77.8% accuracy, and CNN has 88.8% accuracy; for moderate level, ANN has 78.5%, and CNN has 71.4% accuracy. CNN is better at predicting severe levels, while ANN has a better ability to predict moderate levels.

This study demonstrates a framework of improvement of NDT data analysis with ML and also provides an idea for damage prediction with NDT results which is a more efficient way compared with other numerical analysis approaches. However, based on the research procedure and results, there are some future work needs to be considered: 1) keep adding and improving the database of FE model images; 2) improve FE model delamination resolution with smaller mesh size; 3) keep collecting field test data for the training ML model. By these three main tasks, the prediction accuracy could be improved.

REFERENCE

[1] C. Chen, S. Zhang, G. Zhang, S. M. Bogus, and V. Valentin, "Discovering temporal and spatial patterns and characteristics of pavement distress condition data on major corridors in New Mexico," *J Transp Geogr*, vol. 38, pp. 148–158, 2014, doi: 10.1016/j.jtrangeo.2014.06.005.

[2] C. Ho, M. Snyder, and D. Zhang, "Application of Vehicle-Based Sensing Technology in Monitoring Vibration Response of Pavement Conditions," *Journal of Transportation Engineering, Part B: Pavements*, vol. 146, no. 3, p. 04020053, 2020, doi: 10.1061/jpeodx.0000205.

[3] W. G. Buttlar and M. S. Islam, "Integration of Smart-Phone-Based Pavement Roughness Data Collection Tool with Asset Management System," *Region V Regional University Transportation Center Final Repor*, no. 098IY04, 2014, [Online]. Available: https://www.purdue.edu/discoverypark/nextrans/assets/pdfs/098IY04 Integration of Smartphone-Based-Pavement Roughness data collection tool with asset management system.pdf

[4] R. B. Figueira, "Electrochemical sensors for monitoring the corrosion conditions of reinforced concrete structures: A review," *Applied Sciences (Switzerland)*, vol. 7, no. 11, 2017, doi: 10.3390/app7111157.

[5] E. Oromiehie, B. G. Prusty, P. Compston, and G. Rajan, "Characterization of process-induced defects in automated fiber placement manufacturing of composites using fiber Bragg grating sensors," *Struct Health Monit*, vol. 17, no. 1, pp. 108–117, 2018, doi: 10.1177/1475921716685935.

[6] C. Ho, C. Lai, and A. Almonnieay, "Using geographic information systems and smartphone-based vibration data to support decision making on pavement rehabilitation," *Communications in Computer and Information Science*, vol. 610, no. 2, pp. 475–485, 2016, doi: 10.1007/978-3-319-40596-4_40.

[7] B. Mi, T. E. Michaels, and J. E. Michaels, "In-situ ultrasonic monitoring of crack growth under static and dynamic loading conditions," *Nondestructive Evaluation and Health Monitoring of Aerospace Materials, Composites, and Civil Infrastructure IV*, vol. 5767, no. March 2015, p. 1, 2005, doi: 10.1117/12.598968.

[8] B. X. Yu and X. Yu, "Vibration-based system for pavement condition evaluation," *Applications of Advanced Technology in Transportation - Proceedings of the Ninth International Conference on Applications of Advanced Technology in Transportation*, pp. 183–189, 2006, doi: 10.1061/40799(213)31.

[9] R. O. Guldiken, O. Onen, M. Gul, and F. N. Catbas, "A structural health monitoring system with ultrasonic MEMS transducers," *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2011*, vol. 7981, no. April 2011, p. 79810F, 2011, doi: 10.1117/12.881036.

[10] S. Vanlanduit, E. Parloo, and P. Guillaume, "An on-line combined linear-nonlinear fatigue crack detection technique," *NDT and E International*, vol. 37, no. 1, pp. 41–45, 2004, doi: 10.1016/S0963-8695(03)00094-X.

[11] S. Kang, Y. C. Wu, D. S. David, and S. Ham, "Rapid damage assessment of concrete bridge deck leveraging an automated double-sided bounce system," *Autom Constr*, vol. 138, Jun. 2022, doi: 10.1016/j.autcon.2022.104244.

[12] K. Janocha and W. M. Czarnecki, "On Loss Functions for Deep Neural Networks in Classification," Feb. 2017, [Online]. Available: http://arxiv.org/abs/1702.05659

[13] C. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, "Activation Functions: Comparison of trends in Practice and Research for Deep Learning," Nov. 2018, [Online]. Available: http://arxiv.org/abs/1811.03378

[14] L. N. Smith, "Cyclical Learning Rates for Training Neural Networks," Jun. 2015, [Online]. Available: http://arxiv.org/abs/1506.01186

[15] G. Mirhosseini, P. Srivastava, and X. Fang, "Developing Rainfall Intensity-Duration-Frequency Curves for Alabama under Future Climate Scenarios Using Artificial Neural Networks," *J Hydrol Eng*, vol. 19, no. 11, Nov. 2014, doi: 10.1061/(asce)he.1943-5584.0000962.

[16] A. C. Neves, I. González, J. Leander, and R. Karoumi, "Structural health monitoring of bridges: a model-free ANN-based approach to damage detection," *J Civ Struct Health Monit*, vol. 7, no. 5, pp. 689–702, 2017, doi: 10.1007/s13349-017-0252-5.

[17] J. Friedman, R. Tibshirani, and T. Hastie, "Additive logistic regression: a statistical view of boosting (With discussion and a rejoinder by the authors)," *The Annals of Statistics*, vol. 28, no. 2, pp. 337–407, 2000, doi: 10.1214/aos/1016120463.

[18] I. J. Goodfellow *et al.*, "Generative Adversarial Networks," Jun. 2014, [Online]. Available: http://arxiv.org/abs/1406.2661

[19] A. Khan, D. K. Ko, S. C. Lim, and H. S. Kim, "Structural vibration-based classification and prediction of delamination in smart composite laminates using deep learning neural network," *Compos B Eng*, vol. 161, no. December 2018, pp. 586–594, 2019, doi: 10.1016/j.compositesb.2018.12.118.

[20] J. Willard, X. Jia, S. Xu, M. Steinbach, and V. Kumar, "Integrating Physics-Based Modeling with Machine Learning: A Survey," vol. 1, no. 1, pp. 1–34, 2020.

[21] M. Ahmadvand, S. Dorafshan, H. Azari, and S. Shams, "1D-CNNs for autonomous defect detection in bridge decks using ground penetrating radar," no. March, p. 18, 2021, doi: 10.1117/12.2580575.

[22] Z. Dworakowski, K. Dragan, and T. Stepinski, "Artificial neural network ensembles for fatigue damage detection in aircraft," *J Intell Mater Syst Struct*, vol. 28, no. 7, pp. 851–861, 2017, doi: 10.1177/1045389X16657428.

[23] A. Güemes *et al.*, "Damage index for a RPA structure by a fiber-optic based SHM system," *9th European Workshop on Structural Health Monitoring, EWSHM 2018*, pp. 1–11, 2018.