RESEARCH REPORT DOCUMENTATION PAGE

1. Report No.	2. Report Date	3. Contract No.	4. Project No.
UND 20-01	October 2021	91200653	

5. Title and Subtitle	6. Report Type	7. Project No.
Advanced Evaluation Methods for Concrete Bridge Decks: Data Acquisition, Validation and Annotation	Work Plan ☐ Construction ☐ Evaluation ☐ **Final** ☒	8. Project No. 9. Project No. 10. Project No.

11. Author(s)/Principle Investigator(s)
Sattar Dorafshan, PhD.

12. Performing Organization Name and Address	13. Sponsoring Agency Name and Address
NDDOT M+R ☐ NDDOT OTHER* ☐ NDSU ☐ UND ☒ UGPTI ☐ OTHER* ☐ *see supplementary notes University of North Dakota College of Engineering and Mines 241 Centennial Drive Grand Forks, ND 58202	North Dakota DOT Materials and Research Division 300 Airport Road Bismarck ND 58504-6005

14. Supplementary Notes

15. Abstract

Objective

The objective of this research was to collect nondestructive evaluation data from five bridge decks in the county of Grand forks, ND, and to perform initial analysis for their evaluation of use and outline parameters for their future use.

Scope

Bridge decks experience faster deterioration than other bridge components and they cannot be fully evaluated through visual inspections. Non-destructive evaluation (NDE) methods can provide vital information about bridge deck condition. North Dakota Department of Transportation has scheduled five bridges in Grand Forks District for deck repair. The principal investigator proposes collecting NDE data in contact and non-contact (unamend aerial systems) manners from these bridge before, during, and after repair. The NDE data will be analyzed to evaluate the decks and then compared with ground truth to identify their shortcomings and potentials in the fields. In addition, accurate annotated NDE datasets will be formed for the first time that can effectively be used to develop robust artificial intelligent models, capable of NDE data analysis for bridge deck evaluation without relying on experts' opinion.

Summary

The research team reviewed recent advanced NDE methodologies including, but not limited to, resolution, environmental, sensor, and platform requirements for equipment, and specifications. The research team also performed a thorough literature review. All NDE data were collected before the bridges were repaired. Impact Echo is commonly used to detect shallow delamination while GPR can detect the location of bars and their level of corrosion. IRT images were collected using Unmanned Aerial Systems (UAS) to detect delamination. The bridge decks in this study were prepared for ground truth data collection by scarifying delaminated portions of the deck identified by chain-dragging and other visual investigation methods. A delamination survey map locating the areas of delamination and their classes/level of deck removal, produced by NDDOT, served as the ground truth for annotating IRT, IE and GPR field tests data in this study. The maps were aligned with the ground truth maps using image registration, affine transformation, image binarization, morphological operations, connected components and region props techniques to execute a semi-automatic pixel – wise annotation. The research team created the annotated dataset that can be used for training Artificial Intelligence models for defect detection, validating NDE for bridge evaluation, data fusion for improved defect detection, development of new data quality assessment methods based on probability of defect detection etc.

16. Key Words	17. Distribution Statement	18. No. of Pages
Non Destructive Testing, Impact Echo, Infrared Imagining, Bridge Deck, UAS, Artificial Intelligence, Transportation Infrastructure	No restrictions. This document is available to the public from: North Dakota Department of Transportation Materials and Research Division: 300 Airport Road Bismarck ND 58504-6005	103 19. File type/Size PDF/1.6 MB

ADVANCED EVALUATION METHODS FOR CONCRETE BRIDGE DECKS: DATA ACQUISITION, VALIDATION AND ANNOTATION.

North Dakota Department of Transportation

Final Report

Prepared for:

North Dakota Department of Transportation

Prepared by:

Eberechi Ichi (Graduate Research Assistant).

Amrita Das (Graduate Research Assistant).

Sattar Dorafshan, PhD.

Department of Civil and Environmental Engineering

University of North Dakota,

Grand Forks, North Dakota, USA.

October 2021.

Acknowledgement

The authors would like to acknowledge the North Dakota Department of Transportation (NDDOT) for funding this research project. The authors also thank NDDOT project supervisory team and staff for their invaluable assistance and incessant flow of communication and information that was instrumental in ensuring we collected the data promptly and accurately.

Our acknowledgement goes to BDI and SkySkopes for the rapid response to our call for data collection. Several research personnel and students helped to complete this project at the University of North Dakota, including, Bruce Dockter and Tala Talaei Khoei.

Disclaimer

The contents of this report reflect the work of the authors, who are responsible for the facts and the accuracy of the information presented.

Table of Contents

List of Tables

List of Figures

EXECUTIVE SUMMARY

Distress and deterioration of bridges and large infrastructures is becoming a growing concern in North Dakota, the United States, and the world at large. Engineers and researchers are faced with the challenges of developing and deploying viable techniques from conceptual phase to construction and maintenance phases of bridges to ensure that they are safe, stable and durable over their service years. Although destructive tests on core extraction may provide detailed information of the concrete condition, they are labor intensive, time consuming, alter the structural condition and cannot be applied ubiquitously over large areas of the bridge deck. Traditional NDE methods such as, chain dragging, rely on the experience and subjective judgement of skilled inspectors to differentiate the defected sub-surface or surface regions from the sound parts of the deck. Other concerns are difficulties in accessing some parts of the bridge out of reach by physical inspection and safety issues during inspection due to heights or open traffic. Also, infrared thermography (IRT) which has been adopted in bridge projects and studies has not been well validated over the entire area of bridge decks which has resulted in varying digress of accuracy in delamination detection.

These drawbacks have led to adopting advanced non-destructive evaluation (NDE) techniques and remote structural evaluation methods with reliable annotated dataset in detecting, assessing, and monitoring structural sub-surface defects in bridge decks. Recent advances in artificial intelligence can address these limitations; however, development and adoption of the proper artificial intelligence model for proper processing and interpretation of NDE results requires reliable annotated datasets. There exist datasets for surface defects; but due to lack of "ground truth," datasets for subsurface defects have been limited to laboratory specimens and destructive methods which cannot mimic reality. In this study, we have developed a reliable annotated dataset designated for subsurface defects for Infrared Thermography (IRT), Ground Penetrating Radar (GPR), and Impact Echo (IE). The data were collected from five in-service Grand forks, North Dakota (ND) bridges that were scheduled for repair and maintenance.

The research team reviewed recent advanced NDE methodologies including, but not limited to, resolution, environmental, sensor, and platform requirements. We have also examined the previous peer-reviewed conferences and journal papers, as well as state and federal DOT published research reports, focusing on testing standards such as ASTM. Also, we reviewed previous bridge deck evaluation studies using contact and non-contact NDE to determine the best practices for NDE data collection. The procedures, equipment, specifications, settings, and environmental conditions under which the data for the bridges were collected are presented in this report.

All NDE data were collected before the bridges were repaired. IE is commonly used to detect shallow delamination while GPR can detect the location of bars and their level of corrosion. IRT images were collected using Unmanned Aerial Systems (UAS) to detect delamination. We carried out a quality

assessment of the image and signal datasets. Value checks, null value, duplicate values, descriptive statistical, and correlation relationships were considered for the signal dataset, i.e., IE and GPR. The entire, 2,275 IE signals and 655,818 GPR signals were satisfactory without any known error or inconsistencies. The degradation of the images at the acquisition and transmission stages of the data collection affects the results of the image processing and evaluation. The quality of the images was evaluated objectively by computing the amount of degradation (in form of distortions, blurriness, noises, etc.) based on a benchmarked numerical distribution of pixel intensities. A numerical score ranging from poor to excellent quality was assigned to each image. The image quality assessment (IQA) metrics adopted were Perception based image quality evaluator (Piqe), Naturalness image quality evaluator (Niqe), and Blindness/reference-less image spatial quality evaluator (Brisque). The assessment results of the raw image data show good-to-excellent quality for future processing and use.

The bridge decks in this study were prepared for ground truth data collection by scarifying delaminated portions of the deck identified by chain-dragging and other visual investigation methods. A delamination survey map locating the areas of delamination and their classes/level of deck removal, produced by NDDOT, served as the ground truth for annotating IRT, IE and GPR field tests data in this study. IRT images of each bridge deck were stitched together to create an Orthomosaic stitched map. The maps were aligned with the ground truth maps using image registration, affine transformation, image binarization, morphological operations, connected components and region props techniques to execute a semi-automatic pixel –wise annotation. These pixels within the labelled regions are on the stitched images and have been annotated as Green (255) for a class 2 removal and Red (255) for a class 3 removal and others as sound.

Similarly, the collected IE and GPR signals data were also annotated and validated with the ground truth delamination survey layout. The location of each IE and GPR signal was located and established on the delamination map layout which has the different classes of delamination. The classes of delamination (class 2 for shallow delamination, class 3 for deeper delamination and class 1 for sound portions) were therefore assigned to each of the 2,275 IE and signal and 655,818 GPR signals.

The annotated dataset can be found in Appendix A of this report that can be used for training Artificial Intelligence models for defect detection, validating NDE for bridge evaluation, data fusion for improved defect detection, development of new data quality assessment methods based on probability of defect detection etc.

List of Abbreviations and Symbols

Abbreviation	Meaning
ASTM	American Society for Testing and Materials
AI	Artificial Intelligence
DL	Deep Learning
DCNN	Deep Convolutional Neural Network
EM	Electromagnetic
FLIR	Forward-looking infrared
FHWA	Federal Highway Administration
GPR	Ground penetrating radar
IE	Impact echo
IQA	Image Quality Assessment
IRT	Infrared tomography
ML	Machine Learning
ND	North Dakota
NDE	Non-destructive evaluation
SHM	Structural Health Monitoring
UAS	Unmanned Aerial Systems
UAV	Unmanned Aerial Vehicle
UND	University of North Dakota

1. INTRODUCTION

Bridge decks and other structural elements are subjected to continuous surface and sub-surface deterioration because of exposure to environmental and service loads. The bridges in North Dakota (ND) are not only subjected to the usual loads but are also exposed to additional harsh frost action as well as de-icing chemicals and salts to keep the roads and bridges serviceable during winter months. These actions consequently lead to deterioration and defects of the bridge components. According to a 2019 report of the North Dakota section, ASCE journals, out of the 4,377 existing bridges, 86% are in fair or good condition while 14% are classified as structurally deficient (ASCE Report, 2019). The maintenance of in-service bridges sometimes means performing major repair on the decks, which are more susceptible to deterioration than other parts of bridges. A section of existing bridge in ND (Forest River North Bound) scheduled for deck repair is shown in Figure 1.

Figure 1. Picture showing a section of the Forest River North & South-Bounds approach and concrete bridge deck captured using a UAS during one of the flight operations in 2020 summer.

Bridge deck defects, such as cracks, delamination, spalling etc. were commonly detected using traditional inspection methods (visual, destructive, chain dragging, and sounding). Recent studies are focused on adopting advanced Non-destructive Evaluation (NDE) techniques, image-based structural evaluation methods, computer vision techniques, Artificial Intelligence (AI) (Dorafshan et al. 2018), Structural Health Monitoring (SHM), Unmanned Aerial Systems (UAS) (Dorafshan et al. 2021), etc. for full bridge deck evaluation (Rolander et al. 2001; Ahmadvand et al. 2021). Combing NDE and AI for bridge deck evaluation has created more reliable results and are more suitable for the future of bridge asset management due to their data-driven nature (Dorafshan and Azari 2020a & b).

In the traditional destructive testing, cored samples of the bridge deck are collected by coring through the bridge, thereby compromising the integrity of the bridge. These samples collected from localized portions of the bridge decks are used as a basis for evaluating the condition of the entire bridge. Results from such evaluation if not well interpreted by an expert may present false recommendations and impending challenges that may be detrimental.

Similarly, other contact methods of evaluation such as traditional chain drag, GPR, IE, Ultrasonic Pulse Velocity etc. would disrupt traffic flow and in other instances, closures of traffic lane to carry out inspections on the bridge decks. In the same manner, human inspections of bridges are usually laborious, costly, and time-consuming (Dorafshan and Maguire 2018). In addition, interpretation of field data inspection results can be cumbersome with likelihood of errors and in-consistencies. Other techniques adopted in the evaluation of bridges are SHM, an emerging field in NDE and finite element modeling, which is computational in nature and most times requires field testing or data for validation (Ahmed et al. 2020).

A combination of technologically advanced tools such as infrared cameras mounted on UAS enable professionals and inspectors to carry out bridge condition assessment and NDE of the bridge decks without disrupting traffic flow. It also becomes possible to inspect other parts of the bridge that might not be accessible by inspectors. Inspections and Non-destructive Evaluation of bridges are carried out by non-contact means usually at elevated altitudes rapidly and more accurately than human methods, by collecting high-resolution thermal images for further processing, helping to reduce operational cost, and minimizing safety risks.

UAS's when paired with variety of related functions, capabilities, and expertise such as video recording, photography, or multi spectral imaging, can safely, economically, and efficiently carry out a broad variety of surveying and inspection services. UAS applications and capabilities is known to provide experts across allied fields with a unique aerial view which provides access to remote and inaccessible zones without compromising the safety of the pilot. The role and capabilities of UAS in bridge assessment and other allied infrastructural inspections is growing tremendously.

Computer vision application which uses AI to "see" and interpret visuals is also gaining more attention in evaluation, testing, monitoring and maintenance of bridges. A major area of application of Computer vision is in cameras and image processing. Images and visuals are therefore becoming major sources of input data from UAS flights that engineers, researchers, and image processing experts adopt in bridge assessment and evaluation. In some drone flights, several thousands of high-resolution images are captured. This large number of images cannot be effectively evaluated and processed with the human eyes and perception. Majority of AI-based bridge inspections need annotated datasets for input into training, validation, and testing of these models which in turn will help in solving computer vision and prediction problems with a higher accuracy and reliability. The

generated IRT, IE and GPR datasets, validated by an innovative ground truth in this study will make available reliable dataset for further innovative studies and research in automated advanced NDE techniques using AI applications.

This report covers the operations, tasks and activities carried out by our team for the NDE of five concrete deck bridges; Forest River North and South Bound, Park River North and South Bound and Park River Median Bridges.; from the planning stage, review of NDE data collection methodologies, data acquisition, data quality, data annotation and labeling, future works, recommendations, and conclusion. This study has been reported and discussed under the following major sub-sections; Introduction (the current section), Review of NDE methodologies, Data Collection, Data Quality, Data Annotation, and Conclusion.

2. REVIEW OF NDE METHODOLOGIES.

2.1 NDE Methods for Bridges Decks.

Evaluating and ascertaining the degree of deterioration of bridge decks can be one of the most challenging problems researchers, inspectors, and agencies encounter. Traditionally, hammer sounding, or chain dragging with visual inspections are implemented by qualified engineers and inspectors. These methods require substantial field labor, experience, and lane closures for concrete bridge deck inspection. The use of advanced, sophisticated, reliable, and accurate NDE techniques such as GPR, IE, IRT, among others, have been developed to implement inspections efficiently and effectively. The emergence of UAS for bridge evaluations has increased field activities using non-contact NDE. Generally, NDE methods for inspecting and evaluating concrete structures have been classified based on the governing property changes and principles, such as Sonic/Ultrasonic/Acoustic emission, Electromagnetic methods, Electrical methods, Thermal methods, Radiography Vibration analysis, and other methods. The methods investigated in this study are IRT imagery using UAS as a non-contact method, IE and GPR as contact NDE methods. A brief review of the investigated NDE methods is needed to summarize the requirements for data collection phase. Therefore, the research team has reviewed past studies to determine requirements for data acquisition, resolution, procedures, environmental factors, and instrumentation.

2.1.1 Impact-Echo (IE)

Impact-Echo (IE) uses elastic waves to identify and characterize delamination in concrete structures. This method uses the transient vibration response of a plate-like structure subjected to a mechanical impact. IE is commonly implemented in deck evaluations by conducting point testing on a grid with pre-selected spacing. The transient time response of a solid structure is measured with an accelerometer placed on the surface close to an impact source. In an air-coupled IE testing system, the accelerometer is replaced with a contactless sensor such as a microphone; however, the impactor still needs contact to generate seismic waves (Choi et al., 2018). The impactor could theoretically be replaced by a speaker to generate high-frequency waves through the air and through to the inspected media to be fully non-contact; however, fully air-coupled IE has not been implemented in practice and is still in an exploratory phase. A typical contact IE instrument includes an impact source, a nearby receiver, and a data acquisition system. Figure 2 shows personnel carrying out an IE test in this study.

Some past studies have laid out certain data acquisition procedures. For instance, Gucunski et al. (2008) performed IE on an existing bridge with a 2 ft x 2 ft to 3 ft x 3 ft grid. The sensor box was placed at the test point, a series of impacts with a duration of 50 μs were applied, and the acceleration histories were recorded. Traffic-induced vibrations did not impair this testing because of the high-frequency range used—typically between 2 and 30 kHz. Zhang et al. (2016) studied concrete

condition assessments using the IE with grid lines spaced at 4 inch and data acquisition frequency of 500 kHz. Azari et al. (2019) conducted a study to evaluate IE method for concrete bridge decks without and with overlay systems with testing grid of 4 inch x 4 inch and sampling frequency of 200 kHz for 10 ms of IE data acquisition at any point of the gird.

Figure 2.Crew members carrying out IE with the standard set-up device.

2.1.2 Ground-Penetrating Radar (GPR) Method

GPR uses radar pulses to image the subsurface by utilizing two types of antennas: air and ground-coupled. GPR can provide a qualitative assessment of bridge decks and is used as an assessment tool to detect suspected delamination or corrosive deterioration in an existing deck. This method can also be used as a quality assurance tool for new construction or rehabilitation. The GPR method transmits electromagnetic radio waves, with frequencies ranging from 10 MHz to 2.5 GHz, through the concrete. The reflected electromagnetic (EM) waves are then recorded when the reflector's dielectric constants, such as rebar or delamination, are different from the concrete. The GPR data usually consists of changes in reflection strength and arrival time of specific reflections, source wave distortion, and signal attenuation.

Some considerations regarding GPR data collections can be found in the past studies. Despite a slightly higher frequency and the ability to be mounted on a vehicle easily, Hing et al. (2006) and Halabe et al. (2007) stated that GPR systems with a 2 GHz air-coupled antenna would not provide the

ability for the detection of small defects, such as 0.6 ft in width and length, due to the large footprint of the antennas, approximately 1.8x0.6 ft^2. In another study, a bridge deck was evaluated using GPR with 1 ft. spacing path (Abouhamad et al. 2017). Finding smaller defects were studied using data collected from a 1.5 GHz ground-coupled antenna-based GPR system (Hing et al. 2010). Air-coupled GPR can be used to scan bridges at highway speeds, unlike ground-coupled GPR. Rhee et al. (2019) conducted a GPR survey on a pilot bridge, creating a case study to examine the Korean Expressway using a 4-channel, 1 GHz air-coupled GPR antennas of GSSI, which was installed at the rear of a vehicle. The GPR signals were acquired with a sampling rate of 40 scans/ft in the longitudinal direction at a speed of 50–62 mph to avoid disturbing the traffic flow. This sampling rate is much lower than the widely-used rate for ground-coupled GPR, generally conducted at a walking speed of about 167–667 scans/ft. The assessment of the concrete bridge deck and asphalt overlay was determined through field observations. The best time for a GPR survey was deemed to be after a day of rain, when water entrapment in the interface was maximized. Water trapped in the voids and delamination were detected easier than the empty void itself, which provided a more accurate evaluation of the inspected decks. Weather conditions for the studied bridge location from the Korea Meteorological Administration for 30 days prior to the survey date (September 12, 2007-October 10, 2007) for the location of the pilot bridge is hazy having a 76% R.H of air and an average temperature range of 46-75^{o}F. GPR surveys of wet conditions likely identify deteriorated concrete more clearly than when dry. The weather of the pilot bridge location 30 days before the survey was recorded as humid with frequent rainy days. The temperature immediately before the GPR survey was in the range of 50 °F to 57 °F; therefore, the water contained in the asphalt and concrete pore systems was not frozen, indicating suitable survey conditions (Rhee and Choi 2018).

2.1.3 Infrared thermography (IRT) Method

Infrared thermography (IRT) is commonly used for detecting subsurface flaws in bridge decks, such as internal cracks and delamination, in a non-contact manner. IRT performance and accuracy is seen to be different from one study to another. This is due to significant changes in IRT sensor reading caused by varying factors such as change in ambient weather conditions (wind speed, air temperature, air humidity), sensor types and specifications, condition and type of concrete deck, altitude and speed of UAS flight, sizes and depth of delamination, presence or absence of sub-surface delamination, and scenario of deck whether laboratory or in-service. However, the overwhelming majority of past studies indicate that presence of the defects such as delamination would alter IRT reading both in theory and practice. The level of accuracies achieved were heavily affected by the data collection scenario and condition. Therefore, in this study, the research team have conducted a review to identify best practices for IRT data collection and implement them to the highest possible degree.

IRT is a technology to detect thermal radiation emitted from materials with a temperature greater than absolute zero in Kelvin. The rate at which this energy is emitted is a function of the material's temperature and emissivity. A material's emissivity defines the correlation between the true kinetic temperature and the object's radiant temperature (Robert 1982). Hing et al (2010) investigated IRT for detection of water and air-filled voids in concrete specimens. The deck was exposed for 24 hours on a sunny day with ambient temperature highs of around 70°F. Temperature variances were found to be higher between 9 a.m. and 9 p.m. because of solar radiation when compared to night-time. It was noted that the best time to conduct an IRT test under solar heating would be between 12 p.m. and 5 p.m. in Fall, where it was primarily sunny with air temperatures greater than 21°C. In summer, when temperatures were typically around ~86°F, it was possible to see large thermal differences as early as 10:30 or 11 a.m. Solar radiation is a practical heating source for deck testing using the IRT technique since the heating is uniform over the entire deck area, and the process of IRT data acquisition can proceed rapidly.

The nightly cooling ambient condition works better than the daytime heating condition for IRT data collection, as mentioned by Kee et al. (2012); however, Washer et al. (2013) noted that high daily variations in ambient temperature are needed to allow for the detection of subsurface damage, or delamination, in concrete through IRT. This change, approximately 46°F, in ambient temperature is desirable for field inspection. He also concluded that a high rate of change in the ambient temperature improves the detection of subsurface damage in concrete; therefore, a rate of change of at least 37°F/hour for daytime inspections is desirable. When comparing the temperature range settings of 36°F and 50°F to produce the thermal image, large settings result in images that do not indicate the presence of temperature variations associated with delamination. Kee et al. (2012) further noted that the successful use of the IR thermography depends on appropriate weather conditions and on being able to collect data at the right point in time of the day/night cycle.

Gucunski et al. (2013), Oh et al. (2013), Vaghefi et al. (2013), and Kee et al. (2012) evaluated the effectiveness of NDE technologies for the detection of bridge deck deterioration when examining a bridge in service. These studies revealed that IRT was faster than the other NDE methods. Air coupled IE and IRT technology studies showed accurate detection of delamination by comparing the outcome with eight core samples collected from the bridge deck to the IE and IRT heat maps (Oh et al. 2013). Vaghefi et al. (2013) compared IRT and chain drag at the locations of 10 core samplings; in this case, IR showed 40% of core samples (4 out of 10 cores) while chain drag performed 80% (8 out of 10 cores) of accuracy for defect detection. On the other hand, in a comparative study at another bridge, the air-coupled (Prototype A) IE and IRT correctly detected 100% , while chain dragging and air-coupled (Prototype E) IE detected 75% of core samples showing delamination.

The IE equipment used were from different sources with several important differences that affected the overall performance of the equipment (Oh et al. 2013). Table 1 shows performance characteristics of NDT methods adopted in their study. Gucunski et al. (2013) mentioned that IE method is the most accurate technology for delamination detection, followed by IRT and GPR; however, IRT results were less consistent between different specimens and studies. He also mentioned that IRT has a good potential for subsurface damage detection, although it was not the most accurate option among other NDE techniques.

In more general terms, the accuracy of chain drag is higher than that of IRT, it is invariably subjected to human judgment and can only be carried out while vehicle lanes are closed. Recent studies are carried out using IRT techniques, image processing, CNN and other techniques which objectively and autonomously detect defects without necessarily any lane closure. Models are developed to considerably increase the accuracy of delamination detection to almost equal or surpass that of chain dragging or any chosen reliable ground truth.

Vaghefi et al. (2013) collected thermal IRT images over the Mannsiding road bridge deck in the State of Michigan, in August of 2011. The ambient temperature and humidity at the beginning of data collection were 35.8°C and 35.3%, respectively. These temperatures were recorded into the thermal IR data collection software to calibrate the temperature measurements. Sun direction was also considered during the data collection procedure to avoid shadows on the thermal IR images. The Thermal IR camera, a forward looking infrared (FLIR) SC640 with a 24° × 19° lens, was used for data collection. This equipment indicated that the camera needed to be located 26 ft above the concrete bridge deck to cover the width of an entire lane in a single image. During field deployment, the thermal IR camera was installed on a cart 6.2 ft above the deck surface, and thermal IR images were collected by pulling the cart over the bridge in 12 passes at a walking speed of approximately 1.3mph, creating a field of view for each image of approximately 2.6×2 ft.

Table 1. Performance characteristics of each NDT method to detect delamination in bridge decks. (Oh et al. 2013).

Considerations	**Air-coupled IE (Prototype A)**	**Air-coupled IE (Prototype B)**	**IRT**	**Chain drag**
Capital cost	low to medium 1.5	medium 1.0	Medium to high 0.5	low 2.0
Operational cost (film, labor, materials, etc.)	Low 0	Low 0	Very Low 2	Medium 1
Time of operation	1.5hr 1	2.5hr 0	30mins 2	1.5hr 1
Time of analysis	10mins 0	10mins 0	0-10mins 1	10mins 2
Relative sensitivity to ambient noise and/or environmental conditions	Low 2	Low 2	High 0	Medium 1
Operator skill	Medium	Medium	Medium	Low

	1	1	1	2
Weight and size	Bulk but moveable 0	Bulk but moveable 0	Portable 2	Portable 1
Potential use without lane closure	Good 2	Good 2	Good 2	Medium 1
Surface operation	Not needed 2	Not needed 2	Recommended 0	Not needed 2
Delamination detection accuracy	100% (8/8) 2	75% (6/8) 0	100% (7/7) 2	75% (6/8) 0
Objectivity and repeatability	High 2	High 2	Medium 1	Low 0
Evaluation point average	1.23	0.91	1.23	1.18

There have been studies to explore the viability of carrying out IRT survey on bridge decks without lane closures. Although ASTM standard recommends less than 10 mph for IRT data collection, IRT studies have been conducted at higher speeds by mounting NDE testing devises on carts, vehicles, or autonomous platforms for data collection and processing. This is primarily targeted to avoid lane closures, increase the speed data collection and evaluation, reduce safety concerns of personnel (Hiasa et al 2014, La et al 2014, ASTM 2014).

Farrag et al. (2015) investigated mix-variation effects on defect-detection abilities using IRT. In preparing the slabs for testing, each of them was marked and divided into 16 squares, at 1ft x 1ft each using chalk lines after three days of hardening and curing with damp burlap. The study investigated the effect of different mixes of normal weight concrete, high strength concrete, lightweight concrete, and self-compacting concrete on the results of IRT. Earlier studies have researched the best time for data collection and have concluded this time to be during either early heating or late cooling during the day. For this experiment, the later cooling period of between 4 and 7 p.m. in the United Arab Emirates region was used to image specimens. During data collection, environmental conditions, which include temperature, humidity, and wind speed were recorded for all events shown in Table 2.

Hiasa et al. (2016) compared three infrared cameras with different specifications at different times and data collection speeds to explore several factors affecting the utilization of IRT regarding subsurface damage detection in concrete structures. The results exhibit how IRT can detect up to a 1 inch delamination at any time and observed that nighttime would be the most suitable time frame, with fewer false detections and interferences from “noise” created by sunlight. A higher resolution of 640 × 480 pixels was preferable than a lower resolution of 320 × 240 pixels for efficient bridge deck inspection. Table 3 lists the cameras used and the primary specifications.

Table 2. Environmental conditions during IRT inspection of concrete specimens. (Farrag et al. 2015)

Parameter	December	January	February	March	April	March
Temperature (^{O}C)	24	26	28	33	37	39
Wind speed(km/hr)	6	10	19	12	12	15
Relative Humidity (%)	47	50	55	56	53	58

Table 3. Comparison of IRT sensors (Hiasa et al., 2016)

Camera Type	T420	T650sc	SC5600
Detector Type	Uncooled micro bolometer	Uncooled micro bolometer	InSb
Thermal sensitivity (NETD)	<0.045^{O} at 86^{O}F	<0.02^{O} at 86^{O}F	<0.02^{O} at 77^{O}F
Accuracy	±36^{O}F or ±2%	±34^{O}F or ±1%	±34^{O}F or ±1%
Resolution	320x240 pixels	640x480 pixels	640x512 pixels
Spectral range	25-43µinch	25-47µinch	8.3-17 µinch
Frame rate	60Hz	30Hz	100Hz
Field of view	25^{o}x19^{o}	25^{o}x19^{o}	25^{o}x19^{o}
Integration time/Time constant (Electronic shutter speed)	40 inch-sec	27 inch-sec	33-66 inch-sec

Tran et al. (2017) studied the effects of ambient temperature and relative humidity on subsurface defect detection in concrete structures. They assessed the influence of environmental conditions on the subsurface delamination detection of a concrete specimen with active thermal imaging. The temperatures ranged from 60.4 ºF to 71.4 ºF, and the corresponding relative humidity varied from 71% to 35% for the samples tested. The study results revealed that the existence of a highly humid atmosphere might lead to an increase in the absolute contrast, especially for shallow delamination at 4 inch deep. For deeper delamination, the absolute contrast shows similar values for the two test cases at 71% and 58%, despite the difference in relative humidity of up to 13%. The experiment results show that the absolute contrast between defective and non-defective areas increase as ambient temperatures increase and decrease as ambient temperatures decrease. The contrast increases at a faster rate with broad and shallow delamination than with smaller and deeper delamination.

Although there is a good amount of research on IRT up to the present, each study has been conducted under different conditions and mostly on laboratory specimens, making it difficult to draw generalized conclusions. Therefore, there are still several uncertainties regarding the accuracy and reliability of IRT for application to bridge inspection when compared to other tests such as sounding test (Catbas et al. 2015). Nevertheless, the possibilities of IRT technique in bridge evaluation leaves no doubt that if well explored and implemented in future bridge studies and projects, will help to accrue greater benefits in terms of overall cost reduction, time, safety, effective structural health monitoring (SHM) and longevity of bridge infrastructures.

2.2 Data collection standards

2.2.1 Standard Test Method for Measuring the P-Wave Speed and the Thickness of Concrete Plates Using the IE Method'(ASTM C1383).

Impact on the surface of the concrete generates stress waves, of which the P-wave is of primary importance. The P-wave propagates into the plate and is reflected from the opposite surface. Multiple reflections of the P-wave between the plate surfaces give rise to a transient thickness resonance with a frequency related to the plate thickness. A receiving transducer, located adjacent to the impact point, records the surface displacement caused by the arrival of the reflected waves. The output of the transducer is captured as a time domain waveform. The recorded waveform is transformed into the frequency domain using a Fourier transform technique and an amplitude spectrum is obtained. The thickness resonance produces one dominant peak in the spectrum, which can be readily identified. The frequency value of this peak is used in conjunction with the apparent P-wave speed to calculate the thickness of the plate by using the equation:

$$T = \frac{C_{p,plate}}{2f}$$

T = Thickness of Plate, m

f = frequency of the p-wave thickness mode of the plate obtained from the amplitude spectrum, Hz.

According to ASTM C1383 (2000), the impactor shall be spherical or spherically tipped to deliver sufficient energy to a solid plate so that a well-defined amplitude spectrum is obtained with a single predominant peak. Hardened steel balls with diameters ranging from 0.32 to 63 inch in diameter and attached to steel spring rods and spherically tipped mechanical impactors are used as suitable impactors for typical highway pavements. The smaller the ball, the shorter will be the contact time at a given impact point. The typical sampling frequency is usually between 500 kHz and 250 kHz. The typical number of data points in the recorded waveform is 1024 or 2048. The surface of the bridge deck to be evaluated should be cleared of dirt and debris from the surface and for extremely rough surface that is difficult to achieve good contact between the transducer tip and the concrete, the surface should be groun so that good contact is achieved.

2.2.2 Standard Guide for Using the Surface GPR Method for Subsurface Investigation. (ASTM D6432 – 11)

GPR is most often employed as a technique that uses high-frequency EM waves, from 10 to 3000 MHz, to acquire subsurface information. It detects changes in EM properties such as dielectric permittivity, conductivity, and magnetic permeability. Data are acquired using antennas placed on the ground surface. The transmitting antenna radiates EM waves that propagate in the subsurface and

reflect from boundaries where there are property contrasts. The receiving GPR antenna records the reflected waves over a selected time range. The depths to the reflecting interfaces are calculated from the arrival times in the GPR data if the EM propagation velocity in the subsurface can be estimated or measured.

2.2.3 Standard Test Method for Evaluating Asphalt-covered concrete Bridge Deck Using GPR. (ASTM D6087 - 08).

In the case of asphalt-overlaid bridge decks with rebar cover less than 50 mm, data collected with the air-launched horn antenna system must be obtained with two antennas positioned in-line in the longitudinal direction, and with radiating perpendicular polarizations due to the possible interference between the asphalt concrete interface reflection and the rebar reflections. In the case of concrete surface bridge decks without overlay or with Portland cement concrete overlay, a single antenna per lane position will suffice regardless of the rebar cover.

For the air-launched antenna system, the GPR moves in a longitudinal direction parallel to the bridge deck's centerline, with an antenna mounted to maintain a manufacturer-recommended distance from the bridge deck surface. Using a longitudinal distance between the GPR and the surface of less than or equal to 0.5 ft. with a transverse distance of less than 1m between GPR inspection passes is recommended. For a ground-coupled antenna system, GPR inspection passes should be either parallel or perpendicular to the direction of traffic, depending on the direction of the top layer of reinforcement. The pass direction should be chosen so that the antenna crosses over the top layer of reinforcement at an angle nearest to 90^0. Using a longitudinal distance between the GPR and the surface of less than 0.5 ft. and a transverse distance of less than 2 ft between GPR inspection passes is recommended.

2.2.4 Standard Test Method for Detecting Delamination in Bridge Decks using IRT. (ASTM D4788 – 03)

From the ASTM D4788 standard manual, for the delamination to be identified by an infrared imaging scanner, there must be a temperature difference between the delaminated or debonded area and the adjacent solid concrete of at least 33°F. A minimum of three hours of direct sunlight is sufficient to create a temperature difference of 33°F.

While data can be collected at any time of the year, temperature difference will be smaller during the winter months and may inhibit bridge deck testing. Testing should not be attempted when ambient air temperatures are less than 32ºF, as ice in the delamination will cause false readings. An ambient temperature of 68ºF with four hours of sun, and winds less than 15 mph, will allow accurate data collection on Portland cement concrete surfaces. On asphalt-covered decks, an ambient air

temperature increment of 68ºF with at least six hours of sun, and winds less than 15 mph, are necessary for the use of this procedure during the same time period.

The NDE evaluation using IRT should be completed on the entire bridge deck by making one pass per lane. When shoulder areas are wider than 3 ft, a separate pass along the shoulder should be completed. Data should not be collected at speeds greater than 10 mph, and field confirmation of the existence of a delaminated or debonded area is required.

2.3 FHWA Research and Technology NDE Web Manual.

Federal highway administration (FHWA) has published some guidelines for NDE data collection and analysis on their NDE manual website (FHWA, 2021). The following recommendations for IE, GPR, and IRT are adapted from the FHWA NDE web manual.

2.3.1 IE

The deck surface to be surveyed needs to be free of debris before IE testing. The survey should be conducted on the densest grid possible: a 2 ft × 2 ft. This data collection grid and coordinate system for bridge decks should be to locate test points on the deck grid and should be marked using washable paint or chalk. Data collection can be conducted using various devices, each providing a different level of accuracy, speed, and automation.

The equipment used for IE testing consists of an impactor and a sensor recording the response of the deck or other structure. The impactor should be selected carefully to provide enough energy in the frequency range of interest, and some are equipped with a sensor to measure the characteristics of the impact. The wavelength frequency requirements are shorter than the defect's lateral size, and shorter than twice the depth of the defect to be detected. The maximum useful frequency is calculated as the inverse of the impact duration, or contact time, typically 20 to 100 μs. Many IE systems utilize steel balls of different sizes as impact sources, where the center and maximum frequencies are inversely proportional to the ball diameter. The deck should be clear of any debris while the data collection is carried out along any test line. The operator should place the impact device and sensor in direct contact with the sampled concrete surface, avoiding hitting or resting on exposed aggregate, edges, or depressions within surface voids. Otherwise, erroneous signals may be generated, collected, and stored. According to the FHWA Long-Term Bridge Performance Program, during field operations, operators should ensure that the data collected in the time and frequency domains are consistent with seemingly expected signals from the structure being surveyed. For instance, it is usually expected that a frequency corresponding to the deck thickness dominates the frequency response of an intact portion of a deck while other dominant frequency signals will likely include very low resonant frequencies consistent with flexural oscillations, corresponding to shallow delamination—those which should be audible during hammer sounding or the chain drag test. Also, there should be high-frequency resonant

responses from deeper, less extensive delamination or from incipient delamination, either of which fully or partially blocks incident waveforms from reaching the bottom of the deck. There should be no chain drag, hammer sounding, coring, impact or hammer drilling, or similar operations producing high-frequency vibration in the proximity (within 100 ft) of the IE equipment during the data collection. Operation of such equipment may produce frequency responses within the range of interest, which decrease the signal-to-noise ratio and makes it either difficult or impossible to collect and interpret data.

2.3.2 GPR

An experienced and qualified operator should carry out data collection with GPR. The following list comprises the general steps required to achieve quality results on bridge decks:

1. Turn on the GPR system and allow it to stabilize, or warm up, while adjusting the settings and calibrating the survey wheel/distance measuring instrument.
2. Calibrate the survey wheel/distance measurement instrument on a fixed distance over the deck's surface, roughly one-third to one-half of the deck length.
3. Set up the signal position for the top of the scan, gain, vertical high-pass (HP), and low-pass (LP) filters based on the manufacturer's recommendations.
4. Set up the GPR sampling window for a sufficient time range (8 to 12 ns for a ground-coupled antenna, and up to 40 ns for an air-coupled antenna), vertical sampling (512 samples/scan, but ranging from 256 to 1024), scan rate (100 to 400 scans/second), and a spatial scan density of 267-667 scans/ft for a ground-coupled antenna, and approximately 133 scans/ft for an air-coupled antenna.
5. Collect data and perform a quality control review of onscreen output sent to permanent storage (hard drive, flash drive, memory stick, or other media). A ground coupled GPR setup is shown in Figure 3.

Simple data collection on a bridge deck can be accomplished using a hand-pushed or hand-pulled cart with an encoder-based wheel or vehicle-mounted air-coupled antenna, following a series of lines spaced about 2 ft apart. Smaller carts allow access into tighter confines, such as where the deck meets the curb, parapet, jersey barriers, or other limiting factors. The GPR survey lines should be perpendicular to the direction of top reinforcement. If the survey is conducted parallel to the reinforcement orientation, it should be conducted with two antennas of dual polarization. For concrete girders such as precast prestressed, post-tensioned, or cast-in-place, GPR field data are collected at a spacing equal to or less than that of the rebar within the girder. A 1ft × 1ft or more closely spaced grid is used to map voids and delamination, and to image, locate, and map

reinforcements within the girders and the layout of any embedded conduits, post-tensioned cables, or prestressing strands. Position and depth will be accurate if the measurements are appropriately performed.

Figure 3. Data Collection on a Bridge Deck Using Ground-Coupled GPR

2.3.3 IRT

Depending on the heating source, IRT surveys are described as passive or active. Passive IRT is more commonly utilized since natural heating sources, such as the sun, directly expose structures, such as bridge decks, to solar heating. Depending on the ambient weather conditions, the appropriate time for IRT survey recommended by FHWA web manual is between five (5) and nine (9) hours after sunrise or after sunset, since previous experiments have shown that the highest temperature contrasts are achieved during those periods. Most studies have been carried out during the day heating for the passive IRT. Other environmental factors can affect the effectiveness of an IRT survey, such as cloud cover or precipitation. Modern cameras can simultaneously take IR and high-resolution video images. Figure 4 shows the raw output from a UAS equipped with IRT sensors.

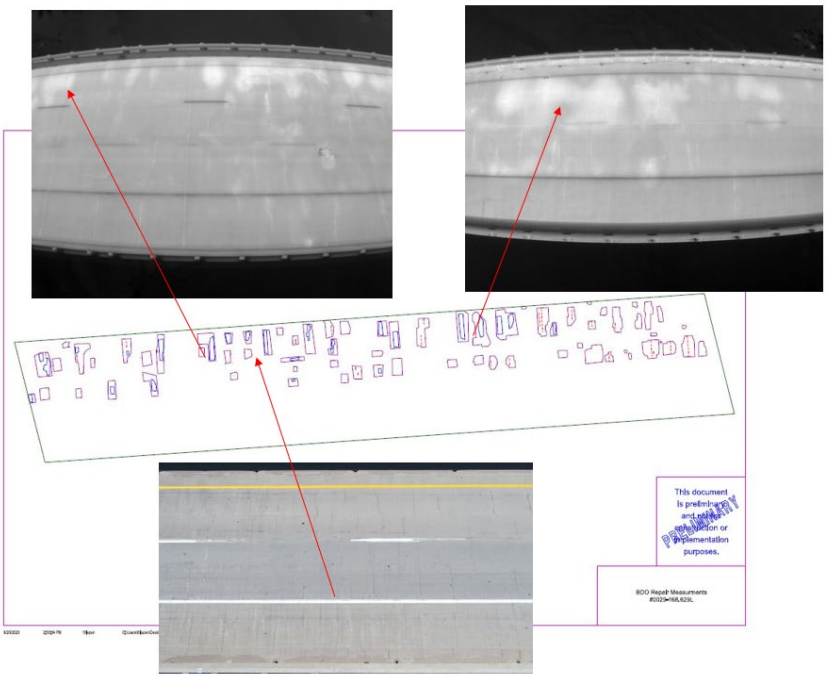


Figure 4. UAS output for deck inspection (IRT)

2.3.4 Chain dragging and Ground Truth.

A comparison of different NDE methods for detecting bridge deck delamination damage requires reliable ground truth information. Chain dragging and hammer sounding are mainly used to detect moderate to severe delamination in concrete structures. The resulting data acquired is commonly used as ground truth data when destructive methods are not available. In some cases, core samples are also collected over selected portions of the bridge deck. The locations are therefore overlain with the chain drag survey layout as ground truth data. Scott et al (2003) compared chain dragging to IE and GPR for bridge deck assessment by collecting ten core samples. Gucunski et al. (2010) carried out a similar study by validation IE and GPR data by comparing them to 4 cores and chain dragging. The positions of the cores were marked by white circles or plus signs, where the circle indicates a delaminated and deteriorated core, and the plus sign indicates a good condition.).

For chain dragging, the operator drags the chain on the deck while listening to the sound that the chain makes (ASTM 2007). Delaminated areas will be detected based on the audible response of the deck. A clear ringing sound represents a sound deck while a hollow sound represents a delaminated deck. The operator marks the locations that are perceived as delaminated. Figure 5 shows a chain drag used for site inspection of delaminated sub-surfaces.

Figure 5. Chain dragging tool for inspecting sub-surface delamination on bridge decks.

Chain dragging gives a general indication of the delamination location. To get more accurate information about the size and shape of the delamination, hammer sounding is conducted by tapping the concrete surface with an ordinary masonry hammer and listening for the response. Like the chain dragging, a clear ringing sound represents a sound deck while a mute/hollow sound represents the delaminated deck. According to FHWA Hammer Sound & Chain Drag recommendations, the operator marks the detected delaminated areas during the survey and the marked areas can be mapped on a grid or placed in drawing/mapping software for output without necessarily requiring any further processing. Figure 6 shows markings for concrete removal after chain drag task on site.

Figure 6. Forest River Bridge deck markings after chain dragging for removal of delaminated sub-surface.

Some of the advantages in adopting chain dragging are rapid and well-established field collection method, an ability to identify severe to moderate delamination by a skilled technician, field method and operation is simple, mapping out the data is not difficult and limited training required.

Some limitations in its adoption are it is labor intensive, can become expensive due to traffic disruption and control, cannot pick up the onset of delamination, results are subjective and vary from one technician to another, must be performed when traffic noise is minimal, generally ineffective for delamination detection on decks with asphalt overlays and chain drag is limited to horizontal surfaces, but hammer sounding can be used for a wider range of structures.

2.4 Evaluation of NDE data Acquisition for Investigated Bridges

The reliability and effectiveness of NDE data collection depends on different factors and conditions. Several standardized codes and past publications have been reviewed to provide a guide for performing an effective and result oriented NDE solution when evaluating five concrete bridge decks in North Dakota. Table 4 shows the summary of conditions for three NDE methods in the past studies and in the standard manuals: IE, GPR, and IRT.

In relation to the review performed on surface conditions, NDE was carried out while the deck was clear of debris and surface water to ensure reliable data. For the data review during collection, conditions such as ambient temperature and average temperature change, humidity, wind speed, and other factors were recorded at the site closest to the climatological substation. These factors were documented to ascertain the variability of the ambient conditions, and whether the prevailing site conditions were suitable for the evaluation. The ambient conditions during which the data for this evaluation were collected falls within the reviewed parameters and therefore provides a reliable dataset for further study, analysis, and reference.

As seen from the review, there are different equipment types and models that can be used for performing IE, GPR, and IRT evaluations. Some of these studies presented the camera types, sensor specifications and settings, models, resolutions, fields of view, altitudes, flight durations, scan rates, and sampling frequencies. The equipment used, along with the settings and specifications for evaluating the bridges under study in North Dakota has been provided in Table 4. As seen, the data acquisition in this study matched the successful past implementation of IE, GPR, and IRT. When possible, data were collected with the same settings and reference points throughout the NDE assessment for consistency and variability reduction. The NDE data collection for the Forest and Park river bridges investigated the detection of subsurface delamination defects and has been carried out within these bounds, as discussed in the reviewed literature and standard manuals. These testing conditions have been presented herein for additional information, reference, further validation, and research purposes.

Table 4. Summary of testing conditions using NDE in the past studies.

Method	Data collection summary	Reference
IE	**Grid:** 2 ft × 2 ft spacing, 3 ft x 3 ft Contact time: 20 to 100 µs. **Equipment used:** Impact-Echo Instruments, LLC, Ithaca, New York, USA. **Sampling frequency:** 200, 500 kHz **Surface condition:** Concrete, polished carefully with sandpaper, and then rinsed with water.	FHWA, Gucunski et al. (2008), ASTM C1383 Azari et al. (2019), Zhang et al. (2016), Gucunski et al. (2008).
GPR	**Grid arrangement:** 1ft x 1ft. survey lines should be perpendicular to the direction of top reinforcement. **Surface Type and condition: A**ir-dried surface of the asphalt concrete layer **Weather:** RH 76%, 50 °F -57°F **Antenna Type and Frequency:** For Ground coupled, 1.5GHz, GPR inspection passes should either be parallel or perpendicular to the direction of the traffic. **Air-launched antenna system:** GPR inspection passes in a longitudinal direction parallel to the bridge deck's centerline. **Scanning rate:**12 scans/m in the longitudinal direction at a speed of 50–62 mph **GPR sampling window:** 8 -12 ns for a ground-coupled, and up to 40 ns for an air-coupled antenna.	Abouhamad et al. 2017, FHWA, Farrag et al. (2015). Rhee et al. (2019) Abouhamad et al. (2017), Hing et al. (2010), Nectaria et al. (2017), ASTM D6087 - 08 Rhee et al. (2019) FHWA,
IRT	**Camera Spec.:** Thermal IR camera FLIR SC640 with a 24° × 19° lens, 640 × 480 pixels resolution **Weather**: **<u>Air temperature</u>** greater than 70°F during Fall weather and 86°F at hours with the largest temperature difference (20°F) from the sunrise or sunset temperature (for day heating and night cooling). 60 °F - 71°F, 75°F -102°F. **<u>Wind speed</u>:** 3.7-12 mph. **<u>Relative Humidity</u>:** 47%-58%, 71% - 58%. **Altitude:** 6.6-26.6 ft. above **Overlapping rates:** 60% -65% for end lap and 30% ±15% for side lap. **Angles and offset distances and number of images taken:** Angles ranging from 0 to 45°, offset distances of 66.6–133 ft, and a total of 295 images taken in 3 flights.	Washer et al. (2013), FLIR systems (2017), Vaghef et al. (2013), Hiasa et al. (2016). Hing et al. (2010), Farrag et al. (2015), Tran et al. (2017), Vaghefi et a.l (2013) Vaghefi et al. (2013), Paine (2003), Siyuan et al. (2019)

3. FIELD DATA COLLECTION

3.1 Background and Description of Bridges in ND

According to the 2019 ASCE report card for ND's infrastructure report card, 4,377 bridges are currently existing in North Dakota. The report further states that 14% are structurally deficient, compared with 8.9% nationally, which are deemed to require maintenance, rehabilitation, or replacement. It is also noteworthy that 43% of North Dakota bridges are over 50 years in age, meaning they are past their initial design life. These bridges are continually subjected to heavy and light vehicular loads in addition to the usual environmental load exposures due to the harsh weather exposures; cold and snowy conditions with heavy freeze/thaw cycles increase bridge deterioration. Also, the use of de-icing chemicals contributes to the deterioration of the bridge's structural elements. (ASCE, 2019).

These 4,377 bridges are either owned by the North Dakota Department of Transportation (NDDOT), a local jurisdiction (county/township), or the Federal government (including US Park Service, Bureau of Indian Affairs, US Forest Service, etc.). Figure 7 shows the current ownership of the ND bridge; 982 NDDOT-owned bridges, 3,339 county/township-owned bridges, and 56 federal bridges-owned.

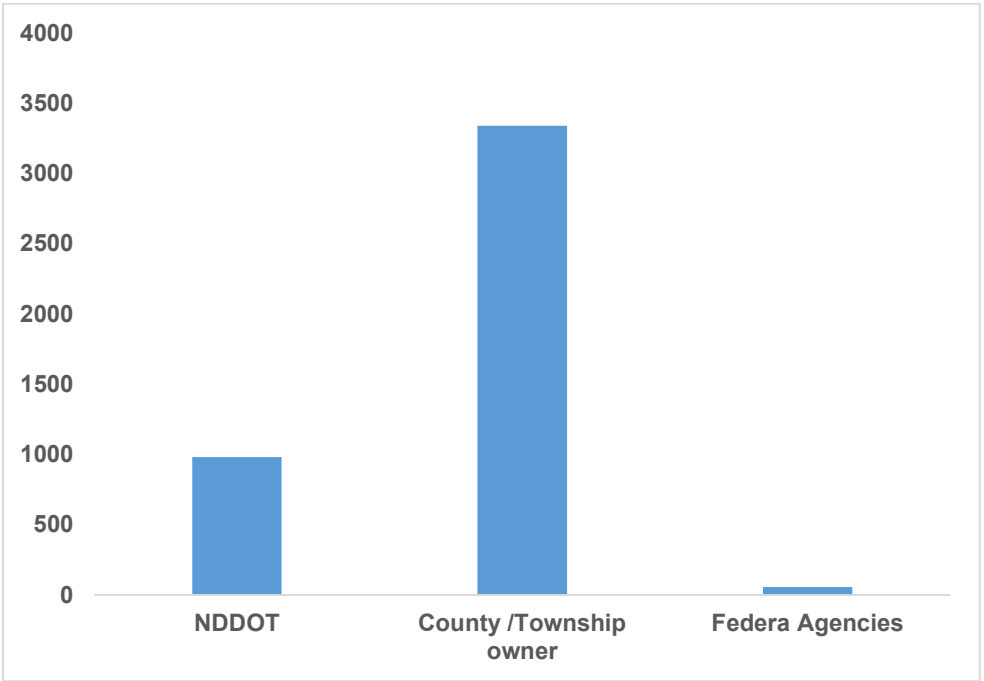


Figure 7. Bridge ownership/maintenance in ND

Most structurally deficient bridges are on locally controlled roads while only 20 of the structurally deficient bridges are on the state system managed by NDDOT. Figures 8,9, and 10 show a summary of the ND bridge structural condition, their rating, and their age.

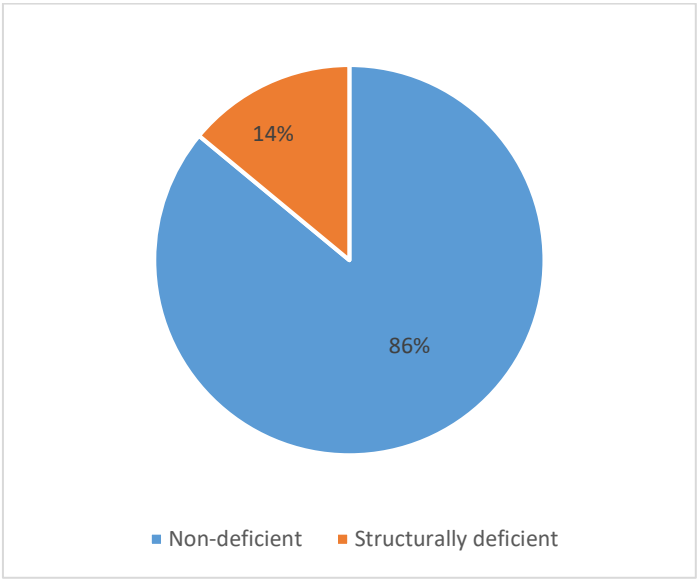


Figure 8. Bridge condition in North Dakota (adapted from ASCE, 2019)

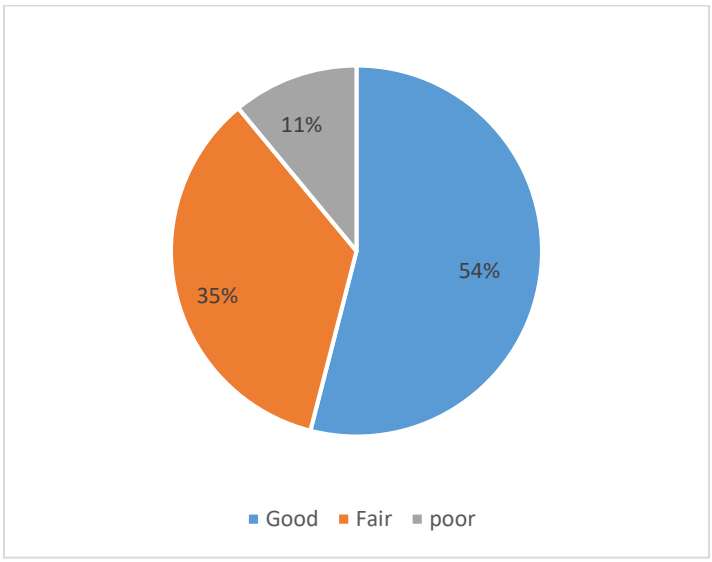


Figure 9. Bridge condition to FAST Act condition-based performance management system (adapted from ASCE, 2019)

The average age of all North Dakota, NDDOT and Local bridges is almost 47 years, 44 years and 48 years respectively with 43% (1,883), 46% (455) and 44% (1,463) of all bridges in North Dakota over 50 years old. It was reported that nationally, 39% of bridges are over 50 years old.

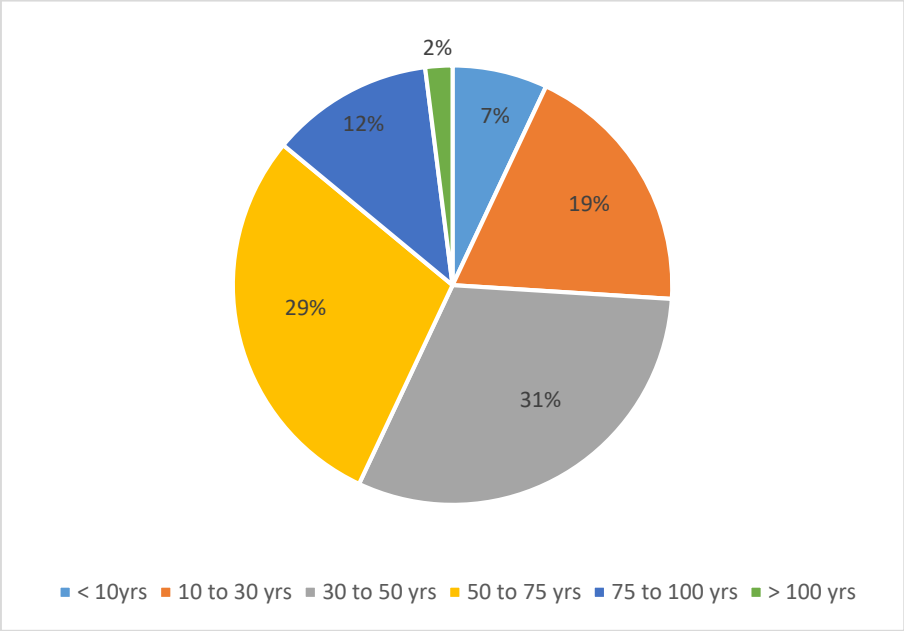


Figure 10. Bridge ages in North Dakota in 2018 (adapted from ASCE, 2019)

The earliest age of the bridges under investigation as at the time of study was 47 years continuously subjected to high vehicular traffic, thermal, freeze-thaw cycles, environmental loads, etc. This has resulted in sub-surface delamination and corrosion of embedded reinforcements. The delaminated condition of one of the bridges under investigation is shown in the Figure 11. The regions marked on the bridge show the actual level of deck removal during bridge repair. About 35% of the deck area would require shallow class 2 sub-surface repair while about 23% would require a deeper sub-surface class 3 repair.

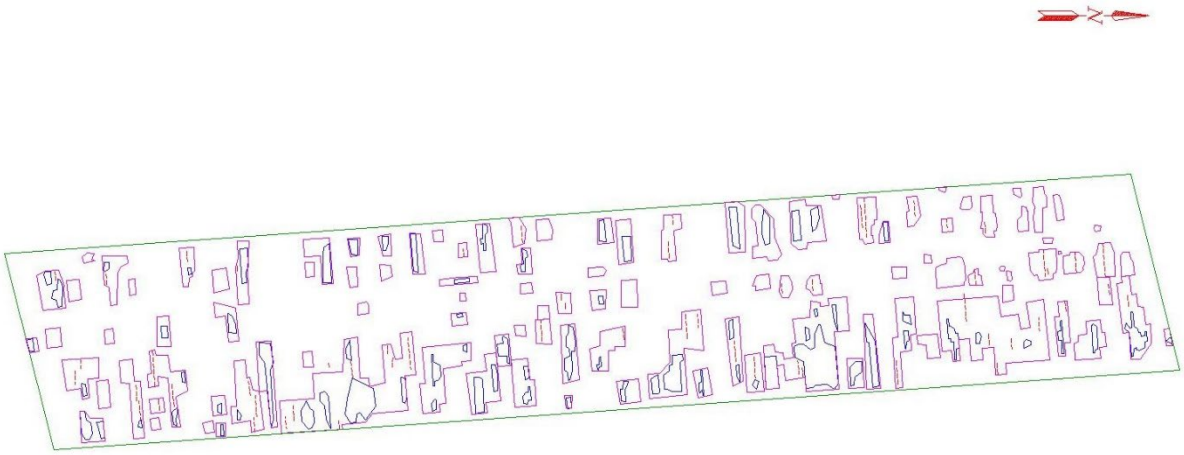


Figure 11. Bridge condition of Forest River South Bound.

3.1.1 Description of the Investigated Bridges

NDDOT identified five bridges in the Walsh county, Grand Forks District with significant amounts of delamination present in the concrete decks for repair. These bridges are located on interstate highway 29 as passing over Forest and Park rivers as shown on google earth map in Figures 12 and 13.

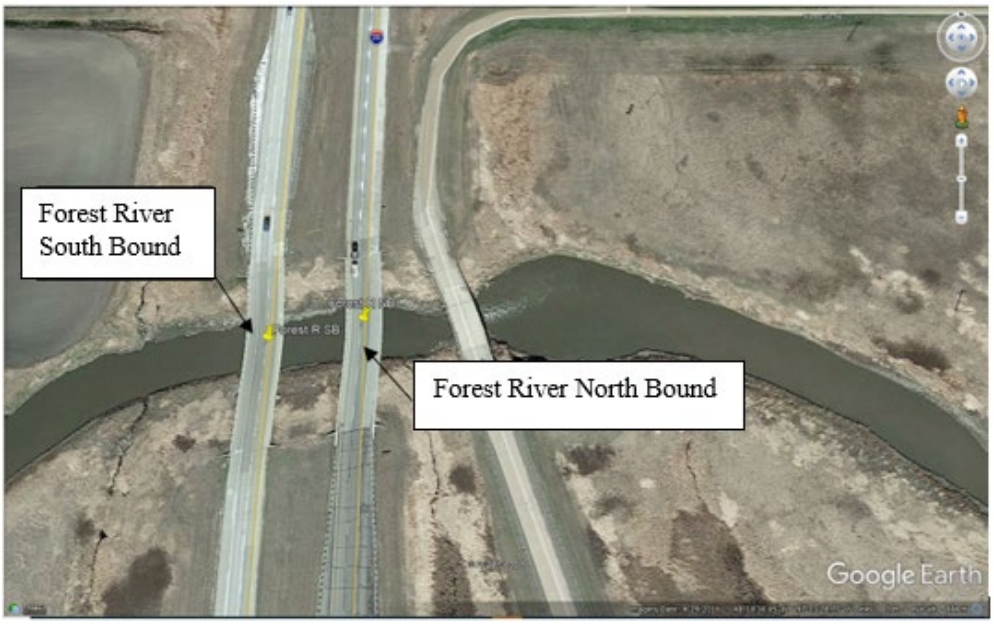


Figure 12. Forest River North and South Bound (Source: Google earth map).

Figure 13. Park River North-Bound (L), Park River Median (Middle), Park River South Bound (R).

A summary information of the bridges is listed in Table 5. The bridges were built between the years 1971 and 1973 making the earliest bridge to be 47 years and the oldest 49 years as at the time of evaluation. They are built to carry the interstate 29 highway except for the Park-River median which is designed for a service road function. The bridge length ranges from 210 ft for the Forest River

bridges to 464.9 ft for the Park River (L) bridge. Table 6 shows weather conditions and the hours of solar exposure from the time of sunrise for rounds of image data collection.

Table 5.Summary description of inspected bridges

Bridge ID	Structure Number	Features intersected	Facilities carried by structure	Length (ft)	width curb-to-curb (ft)	Location of bridges
FR-SB	0029168629 L	Forest River	Interstate 29	210	37.1	$48^{0}18^{I}30.85^{II}$ N $97^{0}11^{I}18.77^{II}$ W
FR-NB	0029168632 R	Forest River	Interstate 29	210	37.1	$48^{0}18^{I}30.85^{II}$ N $97^{0}11^{I}18.77^{II}$ W
PR-NB	0029179087 L	Forest River	Interstate 29	464.9	37.1	$48^{0}27^{I}33.05^{II}$ N $97^{0}11^{I}32.02^{II}$ W
PR-MD	0029179123 M	Park River	Service Road (Park)	365.2	24	$48^{0}27^{I}33.05^{II}$ N $97^{0}11^{I}32.02^{II}$ W
PR-SB	0029179147 R	Park River	Interstate 29	395	48.9	$48^{0}27^{I}33.13^{II}$ N $97^{0}11^{I}31.89^{II}$ W

Table 6. Summary of weather conditions for NDE data collection

First round of data collection					
Bridge ID	**Time collected (7th July, 2020)**	**Ambient and sunrise Tempt. (^{O}F)**	**Sun light exposure (hour)**	**Humidity (%)**	**wind speed (mph)**
FR_SB	9:55-10:25am	78.8	3.0	47.0	6.5
FR_NB	10:26-10:44am	80.0	3.5	44.0	8.0
PR_NB	11:36-11:55am	81.0	4.5	47.0	8.0
PR_MD	12:09-12:32pm	82.0	5.0	44.5	9.0
PR_SB	12:34-12:55pm	82.0	5.5	45.0	10.0
Second round of data collection					
Bridge ID	**Time collected (29th July-27th Aug. 2020)**	**Temperature (^{O}F)**	sun light exposure (hour))	**Humidity (%)**	**wind speed (mph)**
FR_SB	7:25-7:46am	61.0	2.0	83	11.0
FR_NB	7:14-7:36am	61.5	2.0	90	1.5
PR_NB	7:39-8:16am	60.1	2.0	80	1.0
PR_MD	5:59-6:25am	52.5	2.0	96.5	5.5
PR_SB	7:14-7:57am	61.0	2.0	87	4.0
Third round of data collection					
Bridge ID	**Time collected (16th Oct. 2020)**	**Temperature (^{O}F)**	sun light exposure (hour))	**Humidity (%)**	**wind speed (mph)**
FR_SB	8:59-9:06am	27.0	1.0	61	7.0
FR_NB	9:19-9:26am	27.5	1.5	58	8.0
PR_NB	9:47-9:57am	32.0	2.0	49	14.0
PR_MD	7:56-8:11am	24.0	1.0	70	7.0
PR_SB	8:20-8:34am	26.0	1.5	64.5	7.0

The concrete bridge decks are supported on steel beam girders. Our evaluation did not include the steel girder supports, vertical supporting piers and the sub-structure but was limited to the concrete bridge deck. The bridge decks are reinforced with bars ranging from 0.625-0.750 inches diameters in both directions, top and bottom. The approach portion is made of asphalt pavement while the concrete bridge deck is overlaid with Asphalt. The bridge decks had a topping cover of about 2 inches to the top of the top bars embedded in the concrete deck and an overall avearge deck thickness of about 8 inches. The bridges have concrete barrier rails on both sides shown in Figure 14. There are expansion joints (Figure 15) at intervals to control thermal cracks caused by temperature changes.

Figure 14. Shows Forest River bridge steel girders and concrete barrier rails.

Figure 15. Crew preparing Park river bridge expansion joint for repair.

3.2 Description of NDE data collection Techniques.

Visual inspections are the oldest and most frequent type of bridge inspection. Visual inspections can involve walking on the deck, using binoculars to see a point of interest, or using either scaffolding or an Under-Bridge Inspection Truck for regions that are difficult to access.

Physical inspections are recommended when visual inspections are not sufficient for rating a certain region due to uncertainty of defect presence or measurement requirements of a member. The most common practice for physical inspections of bridge decks uses a sounding hammer and chain dragging to locate delaminated regions by comparing the resonating sounds of the defected and undamaged areas (Ryan et al., 2012).

3.2.1 Advanced NDE inspections (NDE)

Practitioners and researchers recognized the shortcomings of visual and physical inspections and therefore suggest more accurate and advanced bridge evaluation methods (Rens et al., 1997). Some of these limitations are in-situ structural characteristic determinations, accurate evaluation of the current serviceability level, economic efficiency, and degree of dependency on inspector skill or experience. To address these limitations, NDE methods may be applied for bridge inspections. In a recent survey by Lee et al (2014), it was observed that out of thirty (30) states with their own bridge inspection manuals, only eight (8) of them addressed using NDE methods in 2014. The most practiced NDE method for concrete bridge inspection was GPR, which was used at least once by 77.5% of surveyed state DOTs, while half of the surveyed states used acoustic methods during their inspections.

3.2.2 Unmanned Aerial Systems (UASs) and their applications.

A UAS is a combination of an Unmanned Aerial Vehicle (UAV), either fixed wing or multi-copter aircraft, the payload (what it is carrying), and the ground control system. UASs are generally defined as any aircraft or aerial device which can fly without an onboard human pilot. Figure 16 shows a drone flight operation executed in this study.

Figure 16. UAS flight operation by crew during bridge inspection.

The type and number of sensors mounted on a UAS depend on the mission requirements. In most cases, the sensors on a UAS must be non-contact, significantly limiting the possible NDE techniques. The most popular sensors for evaluating structures are thermal cameras. There is also a suite of sensors available that are necessary to perform autopilot functions.

3.3 NDE data collection for the investigated bridges

The entire data collection was carried out during 2020 summer between June and September. This study focused on a rare opportunity to collect a significant amount of NDE techniques of in-service bridges identified as having considerable delamination. Ambient weather conditions such as temperature, humidity, and wind speed, as well as equipment type and settings, camera specifications, models, altitude, angle of observations for the contact or non-contact methods are among several factors that were duly considered and recorded during data collection. The task carried out during data collection will be categorized and discussed under the following: Chain drag & ground truth, GPR and IE (Contact methods), and IRT (non-contact methods) data collection techniques.

Classes of removal

NDDOT adopted the following definitions of the class of removal for bridge deck repairs (also shown in Figure 17).

Class 1: Removal of 1.5" of the concrete deck by machine scarification . This does not include any asphalt overlay that is removed from the surface of the deck.

Class 2: Removal of delaminated portions of the bridge deck from the top of class 1 to the top mat using a 30 lbs pneumatic hammer.

Class 2A: is any removals around the top bar of the top mat using a 15 lbs pneumatic hammer with a minimum of 3/4" clearance around the bars.

Class 3: Removal of delaminated portions of the bridge deck that goes down to the top of the top bars of the bottom mat using a 15 lbs pneumatic hammer.

Class 4: removal of delaminated concrete portions beyond the top bars of the bottom mat down to the bottom of the deck. It should be noted that there was no class 4 removal in this study.

Figure 17 defines the different classes of removal and Figure 20 shows a typical ground truth generated based on the class of removal for each bridge deck.

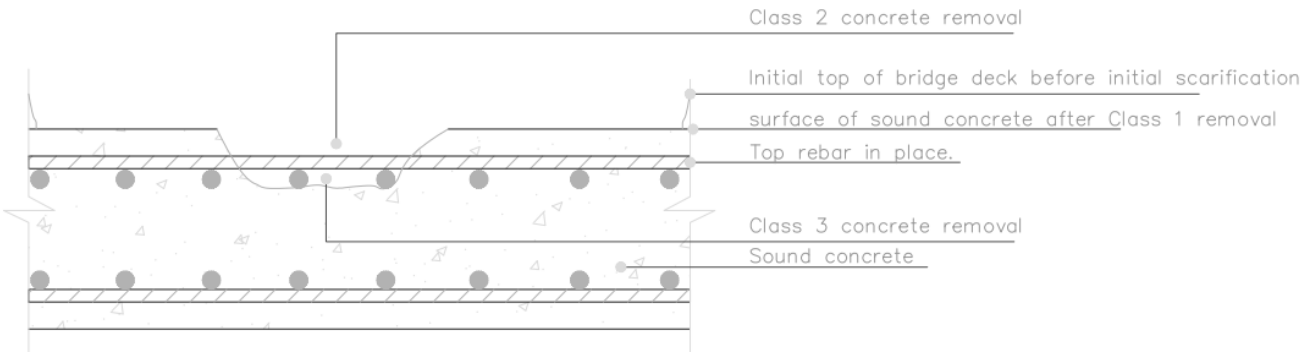


Figure 17. Typical section showing classes of removal

Since the emphasis of this study is to create a dataset for AI applications, the research team have defined the following classes to categorize NDE data collected in this project:

Class 1: Sounds portions of the bridge deck after removal of 1.5" of the concrete deck.

Class 2: Shallow delamination after removal of delaminated portions of the bridge deck from the top of class 1 to the top mat. This also contains class 2A exposed top bars having showing corrosion.

Class 3: Deeper delamination after removal of delaminated portions of the bridge deck that goes down to the top of the top bars of the bottom mat.

3.4 Chain dragging and Ground Truth.

The chain dragging field operation was carried out to determine the sub-surface delamination. The chain dragging was preceded by foremostly removing the entire asphalt surface overlay of the investigated bridge. Afterwards, the class 1 removal was done with the scarification machine. The top of the concrete deck was carefully milled-off up to 1.5 inches by machine scarification. Furthermore, the entire exposed concrete deck was inspected by chain dragging to determine the sub-surface shallow delamination just beneath the exposed concrete deck. The portions deemed to be delaminated

due to the judgment of the sounds heard by the inspectors were marked for removal. This portion of the deck scarified was named class 2. The class 2 delaminated regions were therefore carefully removed by using a 30lbs pneumatic hammer to the top of the reinforcing bars. The exposed reinforcing bars were inspected for visible rebar corrosion. The top of the bars exposed were inspected for further removal of delaminated concrete surrounding the bar. Class 2A removal was thereafter carried out by removing delaminated concrete portions around the top bars of the top mat with a minimum 0.75 inches clearance around the bar. Lastly, further chain dragging was carried out within the class 3 regions. Areas with delamination were noted and class 3 removal was carried out with a 15lbs pneumatic hammer. The delaminated regions were removed up to the top of the bottom bar of the top mat. The field inspection revealed no further delamination beyond class 3 and so there were no class 4 removal. Figure 18 and 19 show class 2 markings of delaminated surface for bridge deck under study.

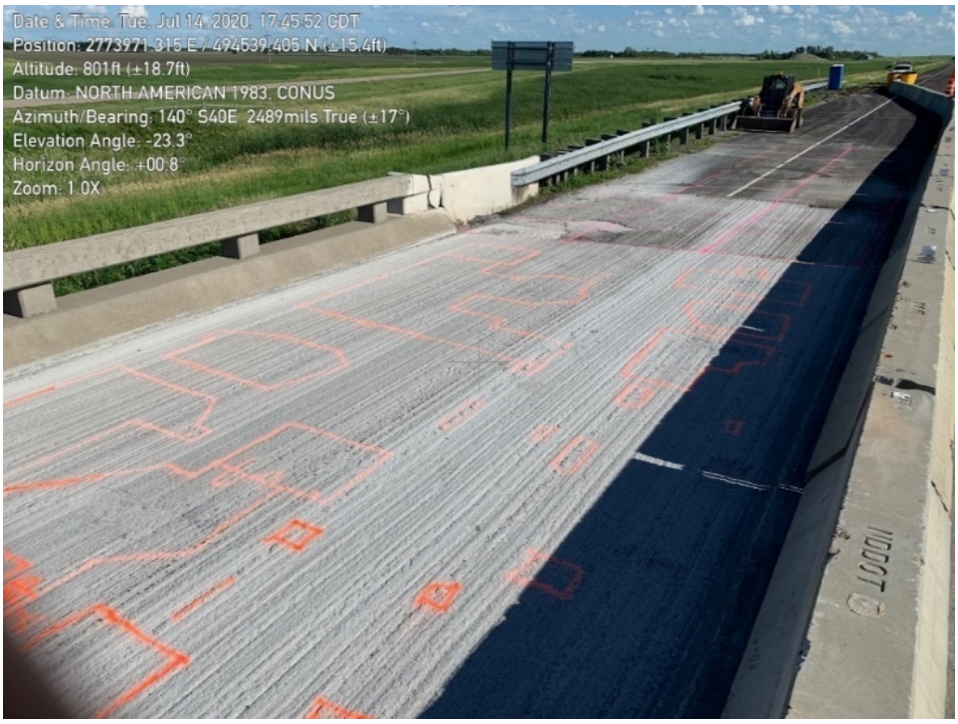


Figure 18.Forest River Bridge deck markings prepared for class 2 removal after chain drag and topping removal

Figure 19.Park River Bridge deck markings prepared for class 2 removal after chain drag and class one removal.

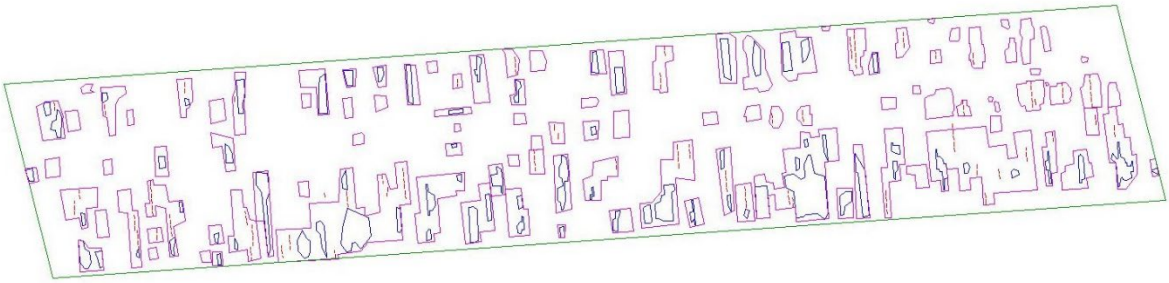


Figure 20.Showing classes of delamination.

3.4.1 Additional ground truth information.

During our assessment, it was deemed fit to gather additional ground truth information for the bridges under investigation. After the bridge decks have been properly scarified in preparation for repair, and prior to concrete pour, the research team captured images of the removed patches and collected the following information.

- Classes of depth of deck exposure (Classes 1, 2, 2A and 3).

- Approximate lateral dimension of the patches.
- Possible Section loss of exposed reinforcement (classified as very mild, mild, severe, and very severe).
- A rough layout sketch of the exposed portions of the bridge deck.

The patches were numbered, and it was ensured that the numbers on the ground matches the numbers of the patches sketched by the personnel. Figure 21(a & b) shows different degree of delamination for classes 2 and 3.

(a) (b)

Figure 21.Image showing portion of (a) class 2 and (b) class 3 removal for Park River south Bound.

The images captured on site for the bridge under investigation shows the condition of the exposed reinforcement bars. The approximate sizes of the patches, rebar diameter, classes of concrete removal, approximate section loss (%) were recorded for the images taken on site. A sample of this ground truth collected is shown in Table 6.

Table 7. Forest River South Bound additional ground truth.

Patch Identity	Image of delamination removal	Classes or removal	Approximate depth of removal (inches)	Approximate Section loss
FR-SB01		2	3	0
FR-SB02		2 and 2A	3	0
FR-SB03		2&3	4	0
FR-SB04		2&3	4	0

3.5 GPR and IE data collection (Contact methods).

IE and GPR tests were performed by BDI. GPR testing were performed across all five bridge decks to a depth of about 8inches while collecting all relevant rebar and concrete activity for delamination

mapping. Testing was performed such that longitudinal scans were taken every 2 ft. along the width of the bridge deck and transverse scans are taken every 10 ft. along the length of the bridge deck.

IE testing was performed across all five bridge decks at three locations per bridge for a total of 15 test locations. Testing at each location was completed on a 1ft x1 ft grid with overall dimensions of 10 ft x10 ft (Figures 22-26). Darker parts in the IE grids shown in these figures were designated as delamination by a trained IE operator. As a reference, chain drag was performed again as discussed earlier at each test location to map out the features that were indicated on the bridge deck layout. The following information were delivered to the research team:

- Narrative report documenting inspection work completed and any key observations on-site.
- Plan view mapping of each bridge with NDE results overlaid.
- Access to raw data gathered using each NDE method, including:
- IE data files (lvm files) and relevant notes
- Ground Penetrating Radar data files (DZT files) and relevant notes.
- Any relevant field photos

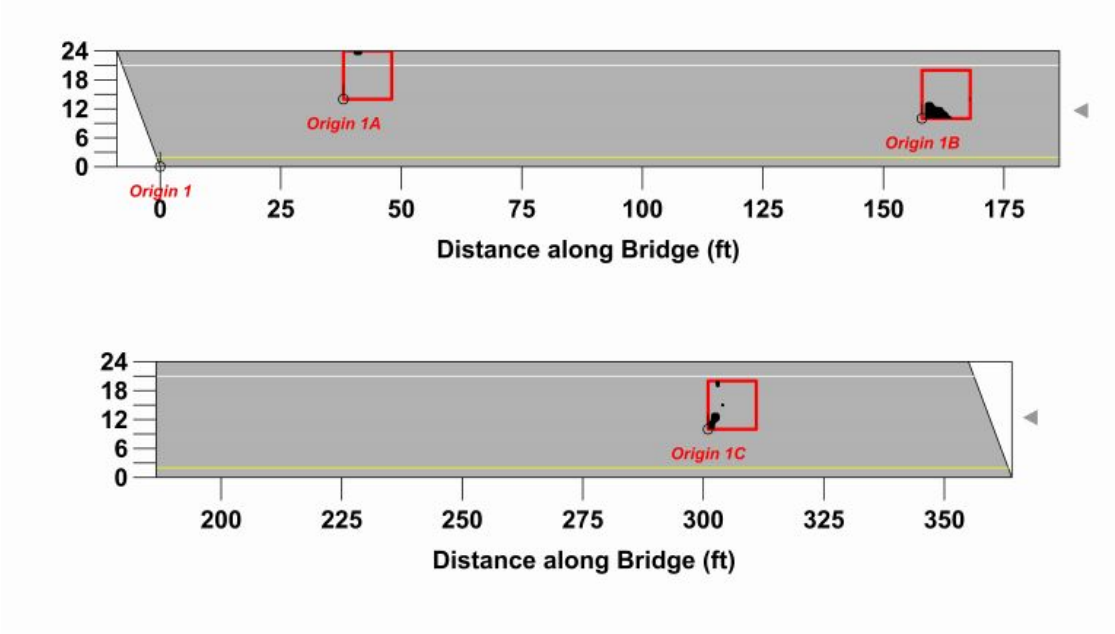


Figure 22.Park River median bridge test locations for IE test

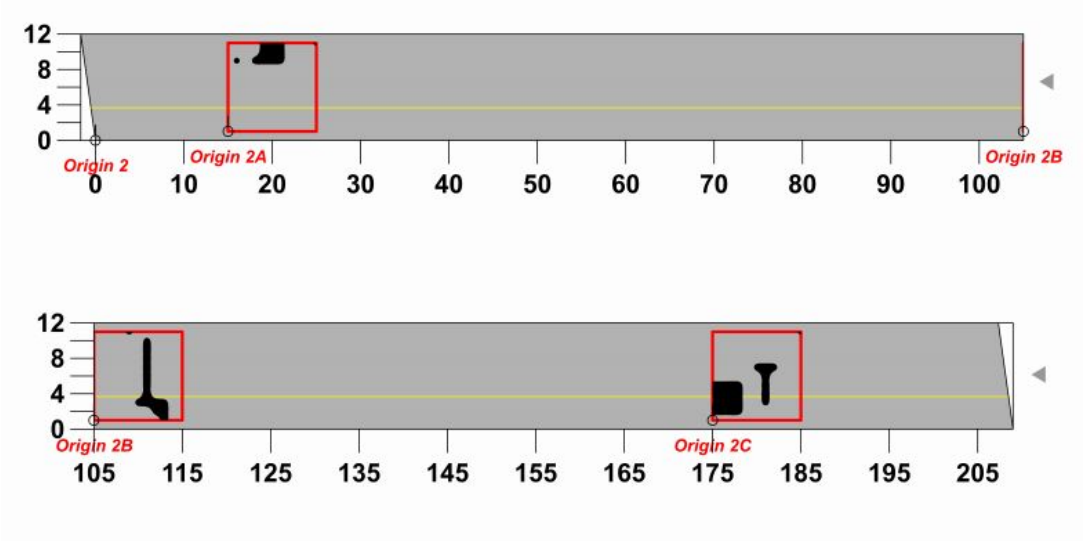


Figure 23.Forest River NB bridge test locations for IE test

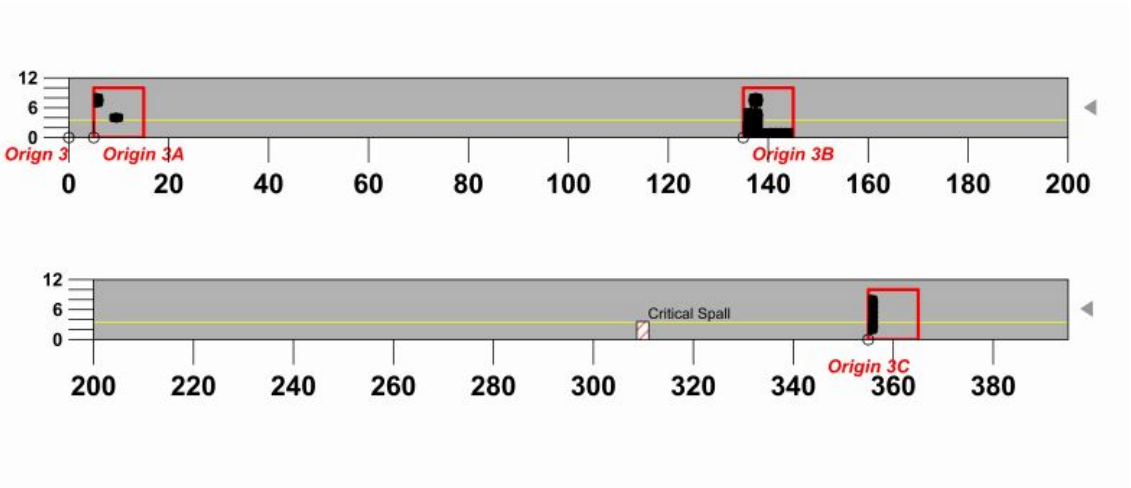


Figure 24.Park River median bridge test locations for IE test

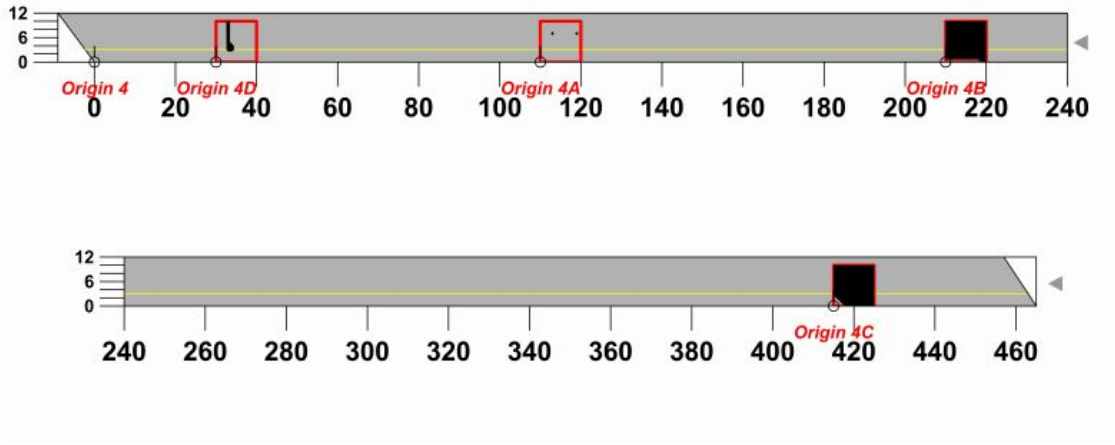


Figure 25.Park River NB bridge test locations for IE test

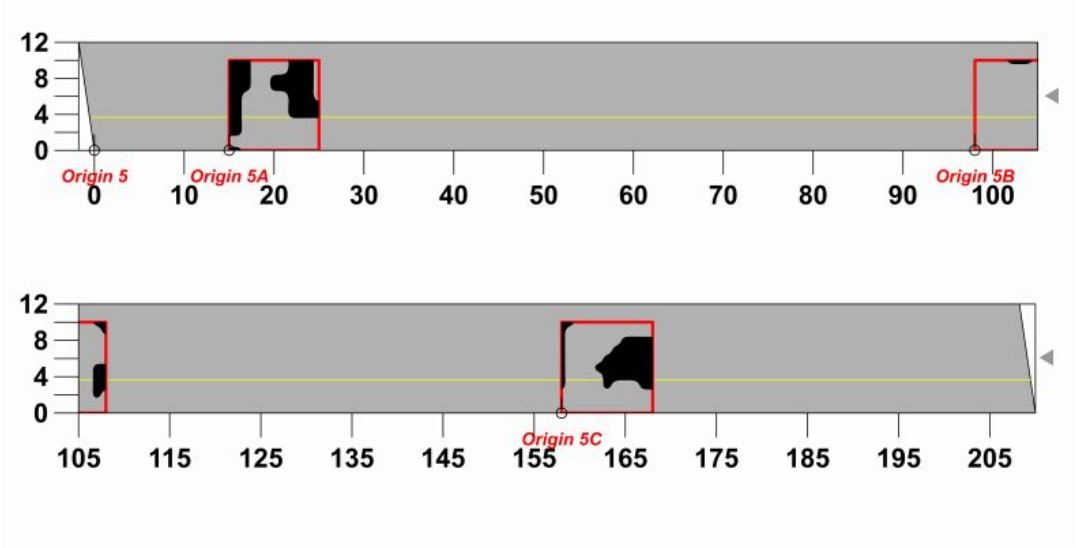


Figure 26.Forest River SB bridge test locations for IE test

3.5.1 Grid layout and hammer sounding

Upon arriving on-site, our team defined the best coordinate system as appropriate for each bridge deck. A fixed origin point was selected from which a convenient test grid was laid out that encompasses the bridge deck's full area. It was critical that the selected origin square off the corner of the bridge such that the skew of a given bridge is captured properly. This test grid was used to provide contextual positions for every measurement taken. This grid was recorded in the field sketch of the bridge with clear indications of the X-Y orientation, North Arrow, relevant landmarks such as corners and bridge joints, etc. After laying out the test grids for IE and setting up the coordinate system, BDI team performed a combination of visual inspections with chain drag sounding at each of the IE test areas. Moving in a systematic manner through the test areas, the team documented any clear signs of spalling concrete (broken up concrete) per the test grids while also performing sounding of the concrete. Areas that indicate the presence of delamination (a void in the concrete that will sound hollow relative to intact concrete) was recorded in the notes as well.

3.5.2 GPR test procedure

For the GPR testing, the field team evaluated the entire testing area using both a 2600MHz antenna. The 2600MHz antenna was set to a depth of 18in for evaluating the rebar, and the sampling rate should have made to be sufficient to resolve the rebar spacing. The following general procedure was followed:

For each type of antenna being used in the inspection, the following steps were completed:

- Set up the GPR hardware and software, using the appropriate settings for the selected antenna.

- At the start of a new project, calibrate the encoder wheel to ensure proper distance measurement.
- For this project, the goal lies in identifying the steel reinforcement as well as any subsurface anomalies, so reviewing the data as it is taken is critical to ensuring all relevant features are being captured.
- It is useful to take some test scans on the concrete slab before formal testing to ensure the data quality is sufficient for the settings.
- After setting everything up, choose to begin gathering either all of the X-Direction scans or Y-Directions scans and then get positioned at the origin in the direction of choice.
- Starting with the antenna centered on the gridline at which the scan is to be taken, set the GPR system to begin collecting data once you start moving the antenna and encoder wheel.
- Moving along the scan line at a steady pace that captures every scan (moving too fast will skip scans), look for anomalies in the concrete or subsurface that are worth further investigation.
- Upon reaching the end of the scan line, save the file if you believe the scan data to have been good or restart the scan.
- Record the file number in the notes table, making sure to include the appropriate location information. In addition, make note of any anomalies observed during data collection in the comments for that scan.
- Continue this process until you have taken scans in both the X-Direction and Y-Direction along the test surface in accordance with your diagram and field notes. Best Practices for GPR Inspection
- Make sure to account for the antenna wheel in your positioning. If the concrete slab has a wall and makes it impossible to center the antenna on the true start of the scan line because of the encoder wheel, then make sure you document the actual placement of the antenna on the scan line, not the theoretical placement.
- Since the offset will be fixed, you can make this adjustment on the backend after completing testing (i.e. it might be confusing to go from 0 ft, 5 ft, 10 ft, to 2 ft, 7 ft, 12 ft… after you have laid out your test grid).

GPR data collection was performed using the following equipment (also shown in Figure 27):

- GSSI SIR-3000 and 2600 MHz Cart Antennae Packing List:

- GSSI SIR-3000 Data Acquisition System.
- GPR Antenna – appropriate frequency for use in application, 2600MHz Antenna (Survey cart & 11-Pin black cable) and 19-pin blue cable.
- 2x Inspired Energy Lithium-Ion Batteries with Inspired Energy Battery Charger

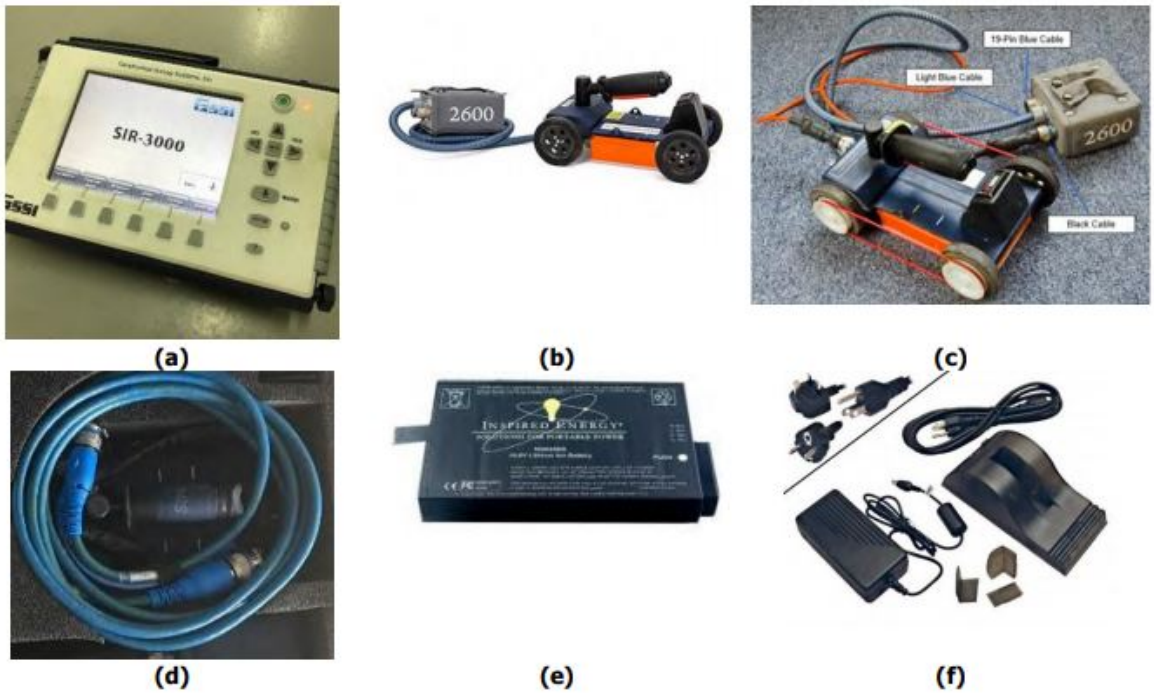


Figure 27.GSSI GPR Equipment – (a) SIR-3000, (b) 2600MHz antenna, (c) 11-pin black cable, (d) 19-pin blue cable, (e) Lithium-Ion Battery, (f) Battery charger

3.5.3 IE test procedure

Each of the selected locations on a given bridge was assigned a unique local origin ID to connect to global coordinate system used to cover the bridge itself. It was ensured this has been completed in advance of testing so the test grid itself can be effectively locked down in space. A total of five individual files was taken at each point of the 1 ft x 1 ft test grid with dimensions of 10 ft x 10 ft for an overall total of 500 files per test grid. This amount of data will yield a sufficient sample of the behavior at each location being evaluated. It was ensured that each set of files was appropriately labelled to make sure they are appropriately connected to the right location, and updated the notes with the file numbers, positions, and any comments. In particular, files that are directly over areas where the sounding identified an issue were labelled (delamination). Finally, since testing was done on locations with known defects, files were saved when a consistent signal was observed, even if the quality of the signal itself was poor or irregular relative to intact concrete. Figure 28 and 29 show personnel carrying our IE test and the IE equipment set up, respectively.

(a) (b)

Figure 28.BDI crew member (a) preparing deck surface points for IE test (b) marking carrying out IE test on Forest River NB bridge deck.

Figure 29.IE test equipment (1) NI-USB-4431 USB DAQ System. (2) Laptop w/ LabView software. (3) USB-A to USB-B Cable. (4) Accelerometer w/ mounting base. (5) Accelerometer BNC Cable. (6) Impact Hammer

3.6 IRT

The UAS flights were conducted by SkySkopes at an ambient temperature range of 68 - 82° F, wind speed range of 5-8 kmph and at an average altitude of 60ft. IE and GPR data were collected on some portions of the bridge deck for eventual comparison with the outcome of the delamination survey and to eventually provide more information on the IRT data collected to evaluate and identify defected sub-surface portions of the bridge deck. Figure 30 shows sample images of IRT collected during investigation.

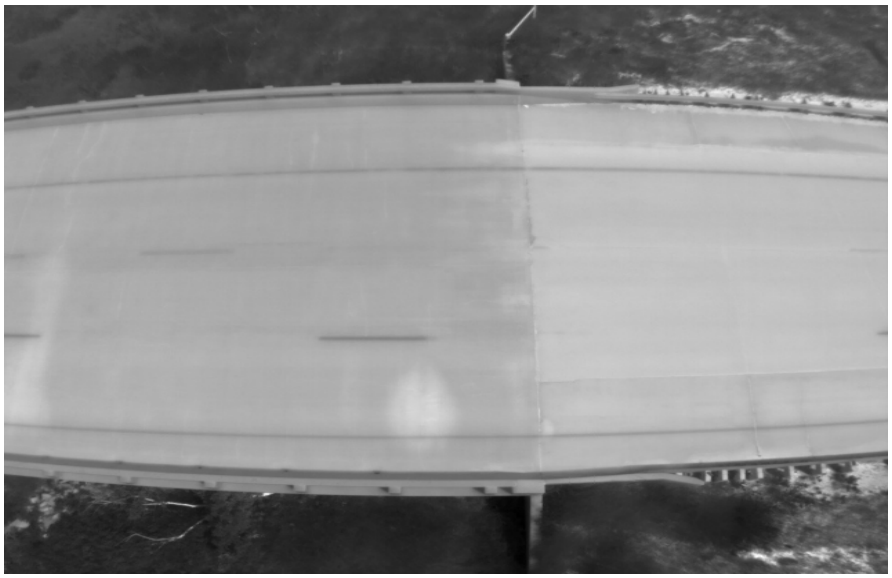


Figure 30.IRT image sample for Park river North Bound 1st flight operation

In collecting the IRT data, the UAS was employed with IRT camera sensors. The name and specifications of the UAS used for the planned flights are highlighted below. All Three flights were planned with same UAS and sensor specifications as shown in Table 7 and 8. A tabulated summary detailing the IRT, GPR and IE data collected are shown in Table 9.

Table 8. UAS specifications

UAS specification	Description
Name of Aircraft	DJI Matrice 210.
Takeoff Weight Maximum	13.5 lb .
Maximum Flight Altitude	8200 ft (above sea level)
Maximum Flight Time	32 minutes (approx.) with TB55 batteries and no payload.
Maximum Tilt Angle P-mode (GPS)	35° (Forward Vision System enabled: 25°).
Hovering Accuracy (GPS Mode)	Vertical: ±1.64 ft (±0.33 ft, with Downward Vision System enabled).
Horizontal:	±4.92ft (±0.98 ft, with Downward Vision System enabled).

Operating Temperature:	-4 to 113°F
Dimensions Unfolded:	34.9 x 34.6 x 16.1inches

Table 9. Camera sensors and specifications

X5s 15mm lens sensor	**FLIR XT V2 thermal sensor**
F-stop: f/8 Exposure Time: 1/800 seconds. Focal Length: 0.6 inches ISO Speed: ISO-400	Full Frame Rates: 640 × 512 ： 30 Hz (NTSC) 25 Hz (PAL). 336 × 256 ： 30 Hz (NTSC) 25 Hz (PAL). Spectral Band: 300 - 540 μ-inches. Pixel Pitch: 680 μ-inches. Digital Zoom: 2x, 4x Thermal Imager: Uncooled VOx Microbolometer.

According to review of past studies, an appropriate timing for IRT survey should be when the highest temperature difference will be recorded for the day from the sunrise time for passive IRT day heating. Depending on the prevailing weather conditions for the site, this time may fall between five and nine hours after sunrise. For our study, the sunrise varied between 5:30-7am on the days of investigation. Therefore, depending on the weather forecast and projection, appropriate timing for investigation was selected to be between 10:30am-12noon for the first flight to investigate the sub-surface delamination prior to removal of defected portions. The second and third flights were collected when the decks have been exposed and repaired respectively for the purpose of investigating further the effect of temperature on the quality of IRT data. The temperature difference from the time of sunrise for the first flight on 7th of July was recorded to be about 18°F. The summary of the data collected for the NDE are shown in Table 10. In addition, the summary of the data collected for the study is shown in Table 11.

Table 10. Summary of NDE data collection.

Method	Data collection summary
IRT/ (UAS)	Flight 1 Locations: • Forest River Left (South Bound) / 48.308472, -97.189608. • Forest River Right (North Bound) / 48.308492, -97.189158. • Park River Right (North Bound) / 48.460053, -97.191763. • Park River Median / 48.459933, -97.194187. • Park River Left (South Bound) / 48.459628, -97.195598 Time and Flight Durations: • 10minutes and 16 seconds (min) to 23 minutes and 12 seconds (max). • {10:53pm (earliest) - 4:53pm (latest)}. Altitude: • 50.6 – 58.0 ft AGL. Ambient temp, change (Δt) in temp. from sunrise and wind speed: • 83.0°F, Δt =18°F, 3.0 – 5.0mph. Flight 2 Locations: • Forest River (South Bound) / 48.308472, -97.189608 Date: 08/01/2020. • Park River Right (North Bound) / 48.460053, -97.191763 Date: 08/03/2020 Time and Flight Duration: • 18 minutes 22 seconds(min) - 23 minutes 47 seconds (max). • {9:23am (earliest)-10:17am (latest)}. Altitude: 61.0- 71.0 ft AGL. Ambient tempt, change (Δt) in temp. from sunrise and wind speed: 68-70° F, Avg. Δt =13°F, 6.0-8.0 mph,
IE	Equipment and specifications: NI-USB-4431 USB DAQ System Laptop w/ LabView software. PCB 353B15 Accelerometer. Samples/Scan: 204800. Sampling Rate: 102.4 kHz., Grid: 1 ft x 1 ft test grid.
GPR	Grid arrangement: 2 ft x 10 ft spacing grid. Equipment: GSSI SIR-3000 Data Acquisition System. GPR Antenna Frequency: 2600MHz. Horizontal Parameters: Scans/Foot – 60. Vertical Parameters: Samples/Scan:512 Bits/Sample: 16-bit. Dielectric Constant: 6.25. Range (ns): 12.00. Depth to Direct Coupling (ns): 2.1 Surface Type and condition: concrete surface cleared of debris.

Table 11. Summary of Data collected for bridges.

Stages of Data Collection	Data Types and formats	FR- NB	FR- SB	PR-NB	PR-SB	PR-MD	Number of files
Images (Round 1)	Thermal Image (JPEG):	122	66	76	95	121	480
Images (Round 2)	Thermal Image (JPEG):	76	84	48	152	100	460
Images (Round 3)	Thermal Image (JPEG):	19	16	24	31	34	124
GPR	Shown in field Note (csv &DZT):	50-77	181-209	78-127	128-180	1-49	1-209
	Downloaded (csv,DZT,DZX):	28	29	50	53	49	209
IE	Shown in field Note (lvm):	415	415	440	542	466	2275
	Downloaded (lvm):	415	415	440	542	466	2275

4. DATA QUALITY ASSESSMENT

4.1 Background and Brief Review of Data Quality.

According to the Federal Enterprise Architecture Data Reference Model, a datum is a value, or set of values, representing a specific concept or concepts. (Data Reference Model, 2005). Invariably, data becomes ‘information when processed, analyzed, and possibly combined with other data in order to extract meaning, results and recommendations. Data can be represented in forms including written, digital, textual, numerical, graphical, audio, images, and video, but broadly categorized into quantitative and categorical datatypes. Figure 31 shows the broad classification of datatypes and their sub-division.

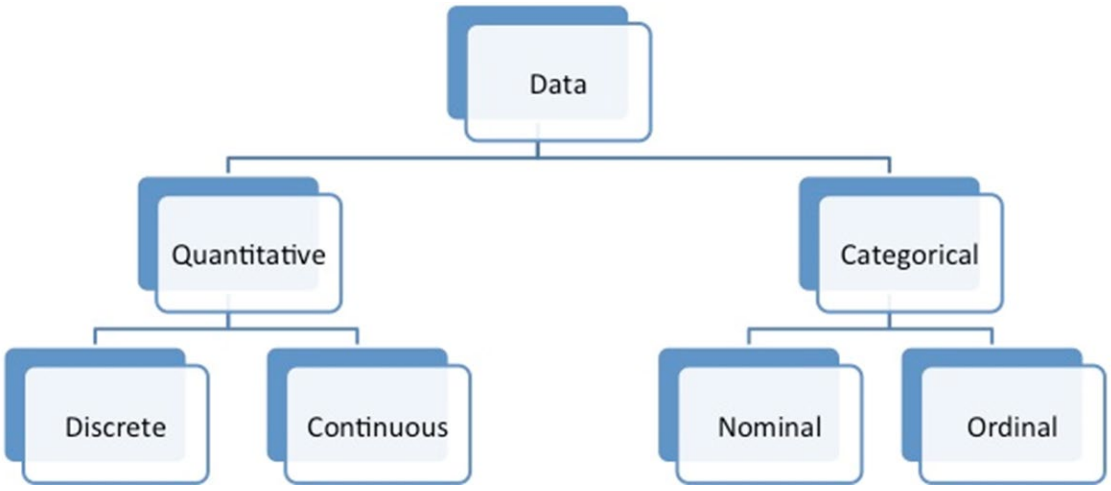


Figure 31.Broad classification of datatype

Acquisition of NDE data in evaluating Civil engineering infrastructures and ensuring the reliability and quality of these data is key and central to the entire NDE evaluation process. The quality of a dataset to a large extent determines the reliability and dependability of the analysis and end results of the study or evaluation.

Data quality and especially the assessment of data quality have been intensively discussed in research and practice alike. There are several definitions for data quality control. According to ISO 9000:2015, data quality can be defined as the degree to which a set of characteristics of data fulfills requirements. Juran et al. (1999) stated that “Data are of high quality if they are fit for their intended uses in operations, decision making and planning.” This implies that data quality means that data are relevant to their intended uses and are of sufficient detail and quantity, with a high degree of accuracy and completeness, consistent with other sources, and presented in appropriate ways. Gassman et al. (1999) asserted that data quality assurance is the process of profiling the data to discover inconsistencies, inaccuracy, incompleteness, and other anomalies in the data, as well as performing data cleansing activities, data aggregation, data transfer to improve big data quality. In general, big data quality assurance refers to the study and application of various assurance processes, methods, standards,

criteria, and systems to ensure the quality of big data in terms of a set of quality parameters. In general, it can be said that the quality of information or data is the degree to which features of this data product meet the requirements (Hinrichs 2002). The quality of data is also defined as ‘‘multi-dimensional measure of the suitability of data to fulfill the purpose bound in its acquisition / generation. This suitability may change over time as needs change’’ (Wu¨rthele 2003). Data quality thus means the ‘‘degree of usability of information for the respective application’’ (Krcmar 2015). This underlines the subjective requirements for data quality in respective institutions and illustrates a possible dynamic data quality process. The definition makes it clear that ‘‘the quality of data depends on the time of the consideration and on the level of claims placed at the time on the data’’ (Apel et al. 2015). Heinrich et al (2018) stated that to support an economically and quality-oriented data management for a reliable result and decision making under uncertainty, it is vital to assess the data quality level of collected data by means of well-founded parameters and metrics. They further stated that, if these metrics are not adequately defined and established, they can lead to wrong interpretations, decisions and losses. Almost for all applications high data quality is an essential consideration; however, only a few have invested the time and resources to maintain and improve the data quality. In most cases, a poor data basis in individual departments is either reluctantly tolerated, or in the worst case not even perceived (Apel et al. 2015).

Objective assessments of data can be task-independent or task-dependent. As the name implies, task-independent metrics reflect states of the data without the contextual knowledge of its application and can be applied to any dataset regardless of the tasks at hand. Conversely, task dependent metrics are developed based on specific organization’s business operations, company and government rules and regulations, and constraints provided by the database administrator.

Data quality metrics are vital in evaluating data quality dimensions. They are quality measures with which a quantitative statement is possible. These quality metrics form the operational basis for determining data quality (Gebauer and Windheuser 2015). A metric is understood to mean methods and systems which, as a result, provide quantifiable values and key figures. In order to make the quality of data measurable, certain characteristics (quality dimensions) are required, which must be assigned to the data (Apel et al. 2015). Several data quality dimensions are highlighted in Table 12.

Table 12.Data quality dimensions (Pipino et al. 2002))

	Quality Metric	**Data quality dimensions**
1	Accessibility	The extent to which data is available, or easily and quickly retrievable
2	Appropriate amount of data	The extent to which the volume of data is appropriate for the task at hand.
3	Completeness	The extent to which data is not missing and is of sufficient breadth and depth for the task at hand.
4	Data accuracy	This measures the closeness of results of observations to the true values or values accepted as being true. This quality indicator usually is used to measure collected sensor data by comparing multiple sources. Data accuracy could be evaluated in two different aspects.
5	Concise representation	The extent to which data is compactly represented.
6	Consistent representation	The extent to which data is presented in the same format
7	Ease of manipulation	The extent to which data is easy to manipulate and apply to different task.
8	Free of error/Data correctness	This data quality factor is useful to evaluate the correctness and reliability of collected (or acquired) big data sets in term of data types, formats, profiles, and soon.
9	Interpretability	The extent to which data is in appropriate languages, symbols, and units and the definitions are clear.
10	Objectivity	The extent to which data is unbiased, unprejudiced, and impartial.
11	Relevancy	The extent to which data is applicable and helpful for the task at hand
12	Reputation	The extent to which data is highly regarded in term of its source or content.
13	Security	The extent to which access to data is restricted appropriately to maintain its security.
14	Data currency and timeliness	Currency measures the extent to which the data is sufficiently up to date for the task at hand, and whether it is correct despite the possibility of modifications or changes that impact time and date values.

		Timeliness refers to the time expectation for the accessibility of data. Timeliness can be measured as the time between when data is expected and when it is readily available for use
15	Understandability	The extent to which data is easily comprehended.
16	Value-added	The extent to which data is beneficial and provides advantages from its use.

Wang et al. (1996) and Krcmar (2015) highlighted the major dimensions of data quality to be Completeness, Correctness/free of error, Timeliness and Consistency. According to Wang et al. (1996), the data quality requirements can be divided into the following four categories shown in Table 13. Defining data quality requirements therefore entails:

- Procedures in defining the requirements for data quality in a particular environment/context
- Professional or institution responsible for making the data quality requirements.
- Procedures in making the requirements measurable and control measures.

Table 13. Categories and definitions of data quality (Wang and Strong 1996)

Accuracy	The extent to which data values are in conformance with the actual or true values
Relevancy	The extent to which data are applicable (pertinent) to the task of the data user
Representation	The extent to which data are presented in an intelligible and clear manner
Accessibility	The extent to which data are available or obtainable

4.2 NDE data quality assessment.

The data collected for the NDE evaluation of the ND bridge deck studies are broadly classified into time series data and images data. Data quality control is often ignored bridge evaluations.

4.2.1 Data Quality for Time series.

IE and GPR data are usually large and therefore ensuring the quality of these data is critical. Data quality control for these data were performed to make sure they are fit, reliable and ready to be used for further analysis in generating AI models which rely on the quality of the data.

Arbesser et al. (2016) recommends the following for data quality control of large time series:

Missing: data values being NULL.

Anomaly (data values that are not impossible but improbable given the data context):

Zero at daytime: power production being zero at daytime.

Non-zero at night: non-zero power production at night

Univariate outliers: meteorological data values which are outliers with respect to their univariate normal distribution – Non-zero duplicates: repeating identical sensor data values other than zero (indication for malfunctioning sensor).

Constraint Violation (semantically impossible data values or relations between values):

Boundary: e.g., negative values for power production

Time holes: time steps of more than one hour

Time duplicates: duplicate time stamps.

Shrivastava et al. (2019) designed an automated data quality advisor for data quality checks and validation. They developed a validator library for a time series data acquired from real world sensors used in the industry. Problems such as missing data, wrong data, and duplicates may prevent an application of analytical methods or may cause unusable or misleading results and interpretation. As summary of the validators adopted by Shrivastava et al. (2019) in shown in Table 13.

Table 14.Validator library overview (Adapted from Shrivastava et al. (2019)

Category	Validator
General validators	
Value checks	• Null missing • Infinity values in data • Columns with constant values
Uniqueness	• Number of zeros in columns • Column-wise unique values • Column-wise repeating values • Column-wise non-repeating values
Duplicate values	• Duplicate rows in all columns • Duplicate rows in specified columns • Duplicate column names • Duplicate columns with different names but same value.
Statistical	• Low variance • Statistical summary
Correlation	• Pearson correlation • Spearman correlation • Distance correlation • MIC correlation
Time series Validators	
Timestamp based	• Time column regular sampling • Time column frequency

4.2.2 Image Data Quality Evaluation

Most objective quality measures for images are full reference since they require a perfect quality original as a basis for comparison. Nevertheless, quality often depends on characteristics which are difficult to infer, even given an undistorted reference. For instance, visual attention and gaze direction are known to significantly influence subjective quality (Moorthy et al. 2009). Objective quality measures are convenient because they do not have the costs associated with human subjects. However, each objective measure is designed to estimate specific quality aspects and is tuned to map to a unique dataset. Thus, an objective measure will only produce predictable results for the environment, error conditions and impairments it was developed for. This sensitivity is especially severe for no-reference measures, which are typically developed to detect one specific impairment. (Ribeiro et al. 2011).

Images and multimedia data can be evaluated either by subjective or objective methods. Subjective quality metrics are considered to give the most reliable results, but are expensive, time-consuming, and impractical for real-time implementation and system integration. Objective Image Quality Evaluation (IQE) methods can be classified by whether a reference image, representing the original signal, exists. When such a reference is accessible, the evaluation is known as full-reference (FR) image quality assessment. Another IQE approach is the reduced-reference (RR) quality assessment, which assumes that partial information about the reference signal is available and used for quality evaluation. In cases where a reference image cannot be obtained (this project), a no-reference, or blind, objective image quality metric is needed to measure the perceptual image quality. Such a measure should be capable of identifying the differences in distorted images; correlate with human perception, reliably benchmark image processing algorithms and assist in selecting the optimal operating parameters, have low computational complexity and be implementable in real time (Yang et al. 2015).

Most existing no-reference image quality metrics were developed for measuring the grayscale image quality of JPEG-2000 coded images, where the pre-dominant distortions are due to blurring and ringing. The widely used quality metrics for grayscale images are contrast or edge sharpness (Agaian et al. 2000 & 2007; Panetta et al. 2008; Marilina et al. 2003 and Caviedes et al. 2004). Measuring the perceived quality of a color image is extremely difficult because human vision is highly nonlinear for different colors. Most proposed color image quality metrics for atmospheric images are based on modifications of grayscale image quality measures. Some methods apply grayscale measures on color images by converting the color image into a grayscale image or by measuring the quality in each color component individually and then combining the measure values with different weights (Wang et al. 2004; Gao et al. 2012). However, the color-to-grayscale conversion procedure is one that involves

losing some useful information. In this research, only greyscale image quality control methods are used. Table 15 shows some image data quality metrics.

Table 15. Data Quality Dimensions and matrix

Quality	References
• Distortion • Entropy • Brightness • Sharpness • Contrast or Colorfulness • Color Quality Enhancement (CQE), Color Root Mean Enhancement (CRME), • Color Image Quality Index (CIQI)	Ribeiro et al (2011), Wang et al (1999), Kim et al (2008). Chen et al (2003), Wang et al (2005), Ooi et al (2009). Bringier et al (2006), Maalouf et al (2003). Bringier et al (2006), Hasler et al (2003) Panetta et al.(2013) Yang et al. (2015).

4.3 NDE data quality of the investigated bridge

The quality of the collected data is discussed under the following sections;

- Image data quality and
- Signals data quality for Impact Echo and GPR data

4.3.1 Image data quality for ND bridges.

Prior to proceeding to actual annotation of the images, our team conducted an image quality assessment (IQA) on the raw thermal image data collected during the UAS flight operations. A total of 1,064 IRT images were collected during the UAS flight. The thermal images were taken with a FLIR camera, a resolution of 640x512, focal length 0.354 inches. Objective full reference, Non-reference and reduced reference quality assessment techniques are one of the methods adopted in evaluating IQA of image dataset. Objective assessment was adopted in place of subjective methods due to high cost and time implications.

4.3.2 Non-reference Image Quality Assessment

In this study, Non-Reference (NR) objective algorithms were adopted to assess the quality of the image dataset collected before proceeding to full annotation. Image quality metrics such as Perception Based image quality evaluator (Piqe), Naturalness image quality evaluator (Niqe), Blindness/Reference less Image Spatial Quality Evaluator (Brisque), were adopted in classifying the image quality.

The preliminary work in the area of no-reference quality assessment started by modeling blur, noise or compression artifacts to address the blurriness, graininess, blockiness and ringing in images Ferzli et al (2009). However, it is of great interest to develop dual approaches which model the statistics of natural images. In this way, it may be possible to avoid detailed modeling of distortions, making possible approaches that are independent from the distortion images are afflicted with. Mittal et al (2011). The approach adopted in their study, Blind/Referenceless Image Spatial Quality Evaluator, was built on a foundation of models of statistical regularities observed in natural scenes. Mathematical representation of this approach is shown in Equation 1, 2 & 3.

Given an image, I;

$$\hat{I}(i,j) = \frac{I(i,j) - \mu(i,j)}{\sigma(i,j) + C} \tag{1}$$

$$\mu(i,j) = \sum_{k=K}^{K} \sum_{l=L}^{L} \boldsymbol{w}_{k,l} \boldsymbol{Y}_{k,l}(\boldsymbol{i},\boldsymbol{j}) \tag{2}$$

$$\sigma(i,j) = \sqrt{\sum_{k=K}^{K} \sum_{l=L}^{L} \boldsymbol{w}_{k,l} (\boldsymbol{Y}_{k,l}(\boldsymbol{i},\boldsymbol{j}) - \mu(i,j)2} \tag{3}$$

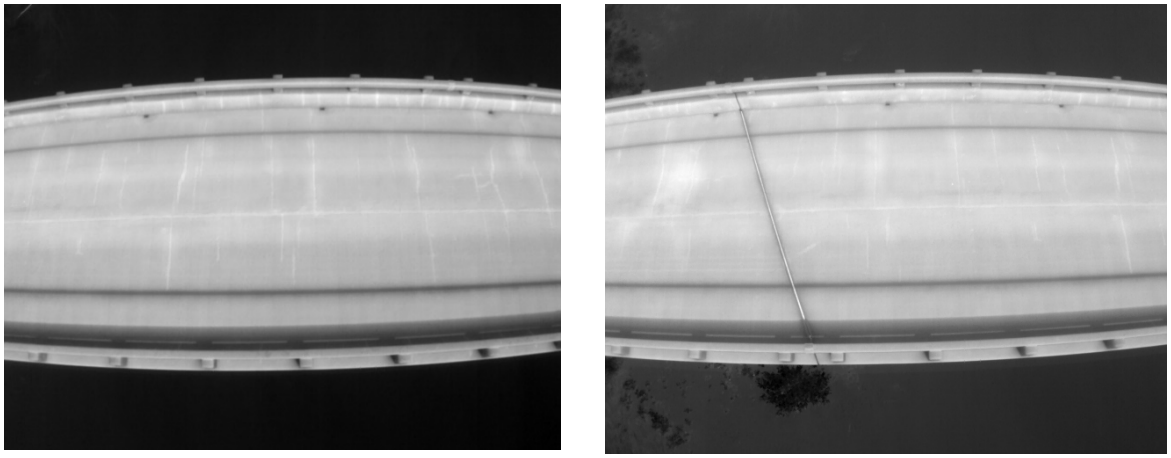
$i \in 1,2.....M$, $j \in 1,2.....N$ are spatial indices, M, N are the image height and width respectively, C=1 is a constant that is used to prevent instabilities from occurring when the denominator tends to zero and W= $\boldsymbol{w}_{k,l}/K = -K,....K, l = -K,.....L\}$ is a circularly-symmetric Gaussian filter sampled out to 3 standard deviations. In our implementation, K = L = 3.

Compared with the illumination index, the factors affecting image quality are much more complex. For example, illumination, noise and contrast will also affect the image quality. The naturalness index considered in this work assumes that a high-quality fundus image should look natural as well. Thus, we further evaluate the naturalness index for those images with good illumination. In recent years, many existing NR-IQA metrics were built based upon the natural scene statistics (NSS) to train a regression model. Mittal et al (2012) presented in their studies of NIQE index as the quality of the distorted image expressed as the distance between the quality aware Natural Scene Statistics (NSS) feature model and the multivariate Gaussian (MVG) fit to the features extracted from the distorted image:

$$D(\boldsymbol{v1}, \boldsymbol{v2}, \sum\boldsymbol{1}, \sum\boldsymbol{2}) = \sqrt{(\boldsymbol{v1} - \boldsymbol{v2})^T \left(\frac{\sum\boldsymbol{1} + \sum\boldsymbol{2}}{2}\right)^{-1} (\boldsymbol{v1} - \boldsymbol{v2})} \quad (4)$$

where $\boldsymbol{v1}, \boldsymbol{v2}, \sum\boldsymbol{1}, \boldsymbol{and} \sum\boldsymbol{2}$ are the mean vectors and covariance matrices of the natural MVG model and the distorted image's MVG model.

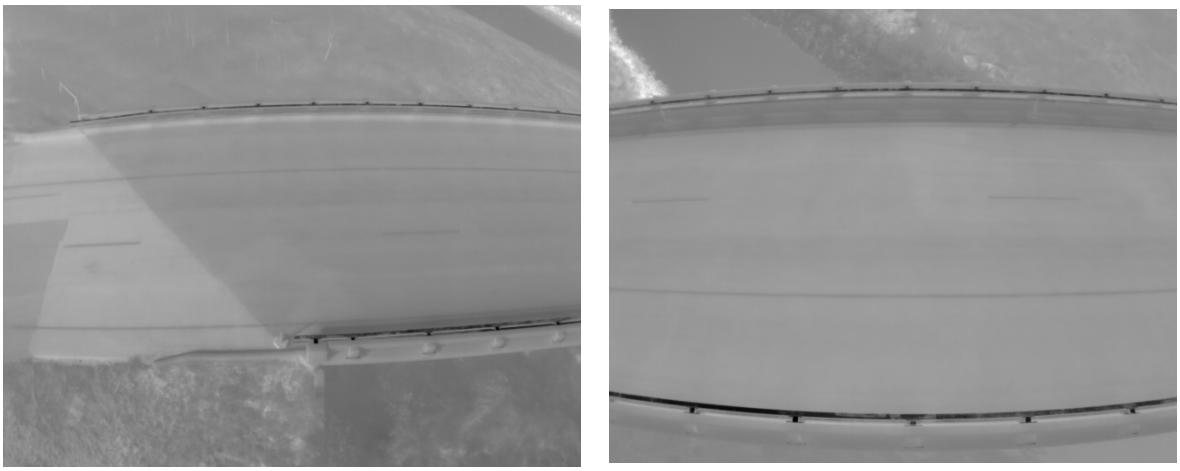
In evaluating the quality of images collected in this study, the thermal images had 3% and 18% of the images within the poor-bad category for the Piqe and Niqe matrix and 100% of the entire thermal images were within the good-excellent category for Brisque. The initial results from the IQA in addition with human perception shows that the images are satisfactory for use. Adopting an accurate and efficient image quality assessment (IQA) technique and algorithm that can out-rightly estimate the inherent quality of an image content under various kinds of distortions is exceedingly important and vital to achieving an overall reliable bridge assessment exercise and results. IRT images with excellent Brisque scores are shown in Figure 32. Brisque compare an image to a default model computed from images of natural scenes with similar distortions. A smaller score indicates better perceptual quality. The metric scores for Brisque are; Excellent [0,20], Good [21,35], Fair [36,50], Poor [51,80], Bad [81,100] (Taylor et al. (2020))



(a) (b)

Figure 32.(a) and (b) Samples of IRT images having Excellent/Good Brisque score.

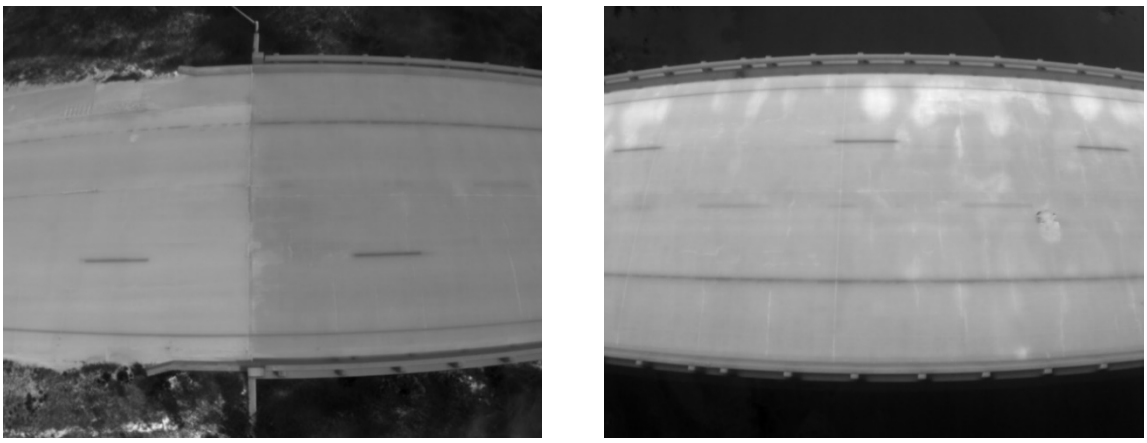
Pique calculates the no-reference image quality score for image A using a perception-based image quality evaluator. Excellent [0,20], Good [21,35], Fair [36,50], Poor [51,80], Bad [81,100] (Taylor et al (2020)). Examples of IRT images with fair and poor Pique scores are shown in Figure 33.



(a) (b)

Figure 33. IRT images having (a) fair and (b) poor Pique.

Niqe calculates the no-reference image quality score for an image using the Naturalness Image Quality Evaluator (NIQE). Niqe compares an image to a default model computed from images of natural scenes. A smaller score indicates better perceptual quality, ranging from 1 (very unnatural) to 7 (very natural). Sample IRT images and their Niqe-score are shown in Figure 34.



(a) (b)

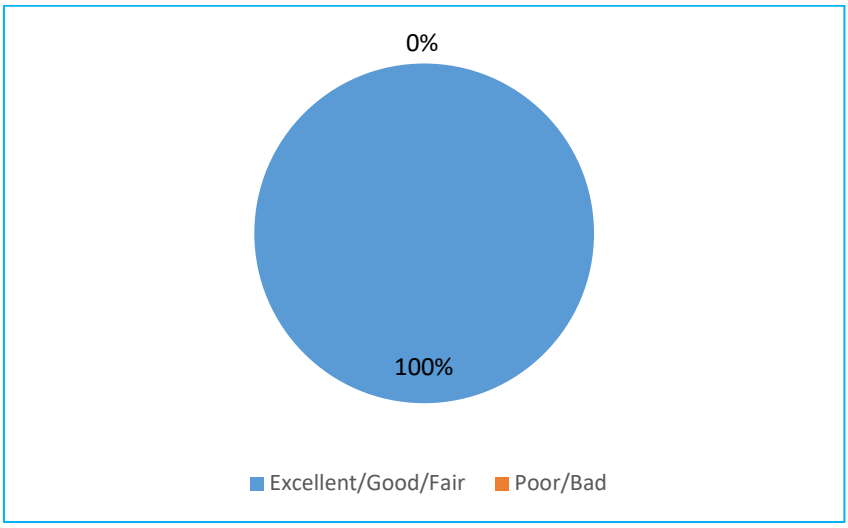
Figure 34.Sample of IRT images having (a) 18% of the images having poorly natural/unnaturalness & (b) 82% of images having Excellent/good naturalness.

Results of IQA carried out on the IRT images for methods considered, shows that high quality images were collected during the flights operations in relation to Niqe, Piqe and Brisque. Only 3% of the entire IRT images shows poor Piqe value while the remaining showed an outstanding high value of 97% as deemed good/excellent. For the Niqe assessment, 18% of the IRT images were categorized as poorly natural/unnatural while the remaining 82% had excellent/good/fair natural results. The Brisque results showed all the IRT images are categorized to be excellent/good/fair. These results satisfied the

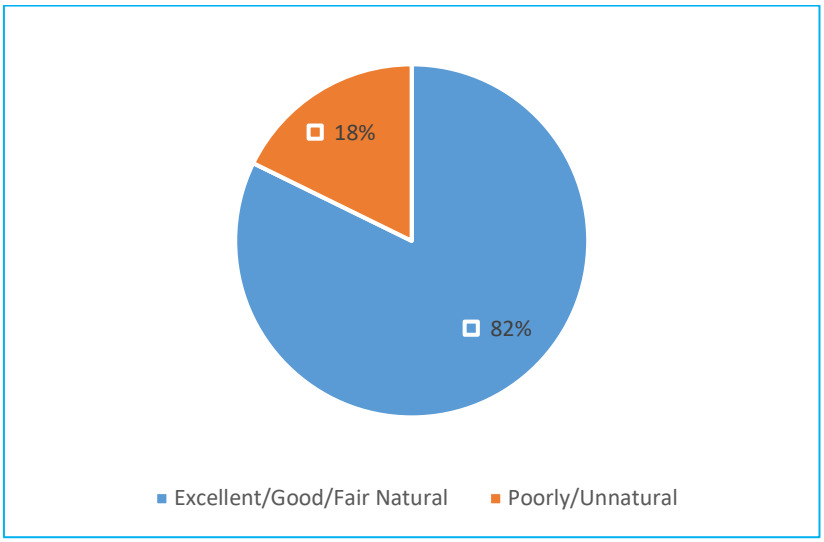
quality requirements based on the metrics highlighted in Table 15. The IQA results for IRT images are shown in Figures 35 (a) through (c).

Table 16. Summary of IQA assessment matrices for Non-Reference

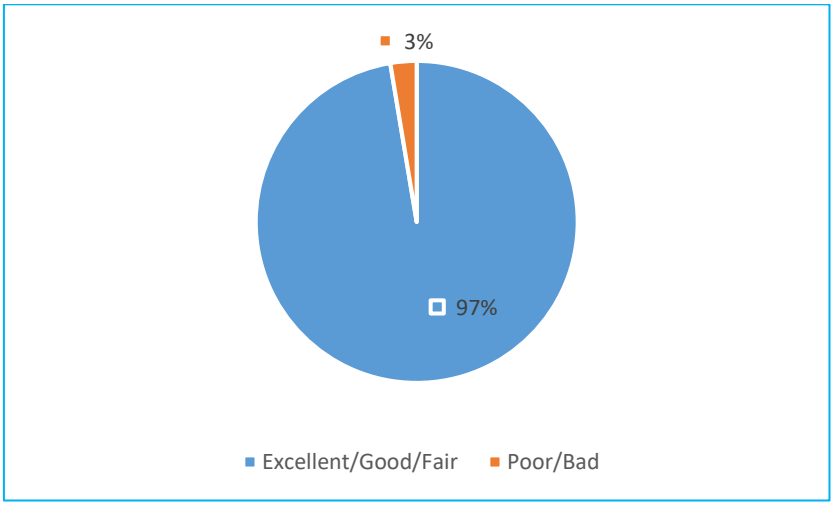
Variable	Total Count	Comments	Reference
Piqe_score	1064	Excellent [0,20], Good [21,35], Fair [36,50], Poor [51,80], Bad [81,100]	Taylor et al (2020).
Niqe_score	1064	Lower values of score reflect better perceptual quality of image. 1 (very unnatural) to 7 (very natural).	Halonen, et al (2011).
Brisque_score	1064	Excellent [0,20], Good [21,35], Fair [36,50], Poor [51,80], Bad [81,100].	Fu et al (2016).



(a)



(b)



(c)

Figure 35.Pie chart showing scores for (a) Brisque (b) Nique and (c) Piqe IQA results for IRT images.

4.3.3 Signals data quality for Impact Echo.

Some of the quality check for NDE time series data are:

1) Observation of signal output at the point of testing.
2) Value checks: The values of some of the raw data for any voidness. Checked, assorted out bad and good signals.
3) Null value: Zero value appearing instead of real integers. Checked, sorted out bad from good signals.
4) Duplicate values: Checked for figures repeating incorrectly.
5) Statistical: Descriptive statistical parameters of raw data.

6) Correlations: Mathematical relationship for likely expected trend. (Shrivastava et al 2019).

The output of the IE signals is time and acceleration (g) with 204800 rows, each testing point saved in a 'lvm' file format. This was checked for all files to confirm there were no missing or null values for the files test points. Overall, 2,275 IE test points (files) were taken for the entire bridge sections tested. Also, the time duration varied from minimum value zero (0) to a maximum value of 1.99990 having a time step of 9.765625E-6sec (Approximately 0.00001sec).

As part of our quality procedures to ensure that final signals depict the condition of the concrete deck, our crew observed every sound during the IE testing and separated bad signals from the observed good signals. These signals, good and bad have been shown in the Figure 36 and 37.

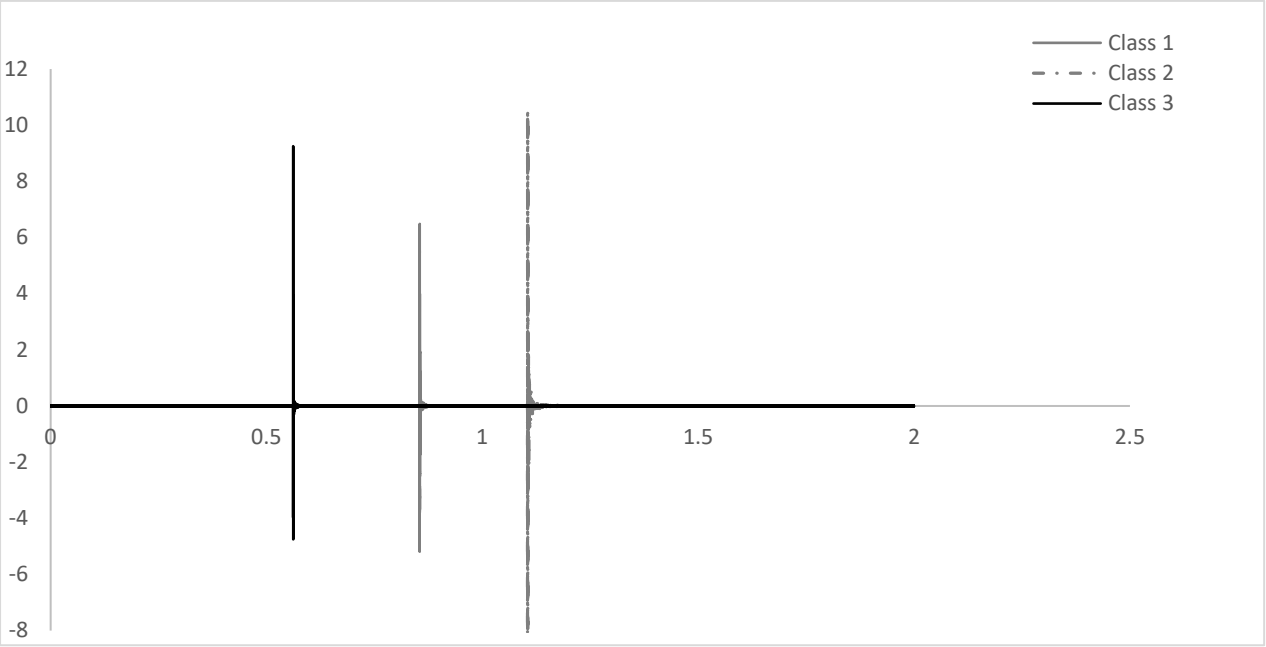


Figure 36.IE signals class 1 (sound), class 2 (shallow delamination) and class 3 (deep delamination) for Park River Median bridge.

4.3.4 Signals data quality for GPR.

Same quality check as IE was conducted for the GPR data for correctness of raw data. Each testing signal is saved in a 'csv', 'DZT' and 'DZX' formats having 209 files for each format. Each of the signals has the time duration varied from minimum value zero (0) to a maximum value of 11.97656sec with a time step of 0.023438sec. The longitudinal signals scans along the length of the bridges giving an output file having 16,383 columns [512 16,383]. The transverse signals which scan along the width of the bridges gives an output file having an average of 1,225 amplitude columns. [512 1225]. All files checked to confirm there were no missing or null values for the files test signals. A graphical display of cases 1, 2 and 3 of signal file_004 is shown in Figure 37.

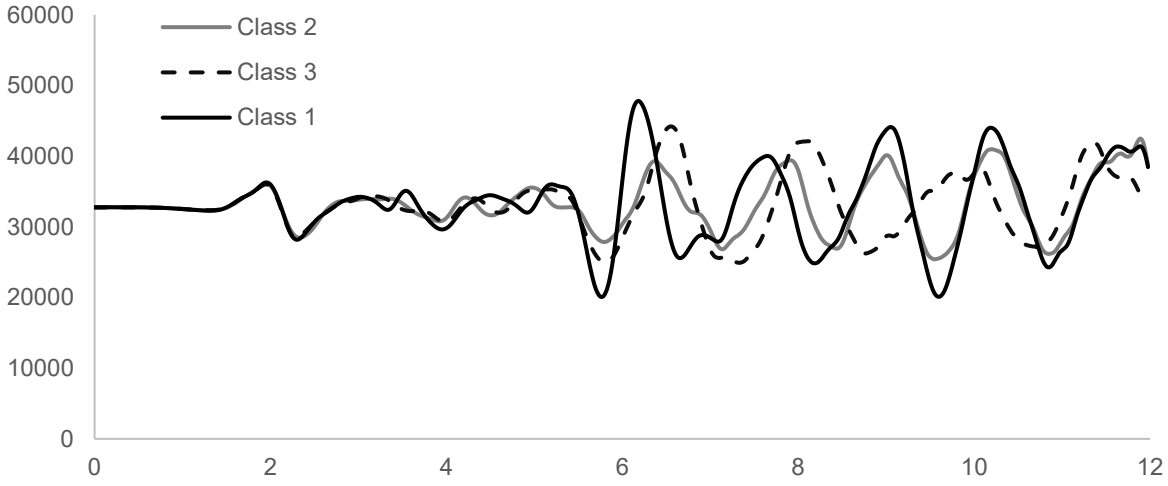


Figure 37.GPR signal for Park River Median for class 1, class 2 and class 3 signal file_004

5. DATA ANNOTATION

5.1 Background and overview of data Annotation

Building an AI model for the purpose of making intelligent predictions in place of human interpretation, requires large volumes of training data. For a model to make decisions and act, it must be trained to understand specific information. Training data for AI purposes must be properly categorized and annotated for a specific action. This implies that metadata must be assigned to the images in the form of identifiers, captions, or keywords. Data annotation is therefore the categorization and labeling of data for AI applications. Big, annotated image datasets now play an important role in Computer Vision research.

5.1.1 Image Annotation

Image annotation is simply the process of attaching labels to an image. Annotation can range from one label for the entire image or numerous labels for every group of pixels within the image. The annotated images can serve as sources of ground truth data, which could be fed to a computer vision algorithm for processing. Through training, the model would then be able to distinguish animals from unannotated images. Other image annotation methods used in computer vision are:

- **Bounding Boxes:** drawing a box around certain objects within the image, having only length and width.
- **3D cuboids:** drawing a box around objects in an image labeling length, width, and approximate depth.
- **Polygons:** used to label objects which does not fit well in a bounding box or 3D cuboid due to their shape, size, or orientation within the image. With polygons, annotators draw lines by placing dots around the outer edge of the object. The points are connected and the space within the area surrounded by the dots are annotated using a predetermined set of classes. When objects are assigned more than one class of annotation, it is called a multi-class annotation.
- **Lines and Splines:** mainly used to train machines to recognize lanes and boundaries. Annotators draw lines along the boundaries required to label.
- **Semantic segmentation/Pixel-wise:** Semantic segmentation is much more precise and specific assigning required index category (0 or 1), metadata, tag or label to every single pixel in an entire image.

Cheng et al. (2018) discussed different Automatic image Annotation methods in five categories, namely:

1) Generative model-based image annotation.
2) Nearest neighbor-based image annotation.
3) Discriminative model-based image annotation.
4) Tag completion-based image annotation.
5) Deep Learning-based image annotation.

5.2 Region-based pixel-wise semi-automatic IRT image annotation.

In image segmentation and visual tracking, well annotated image and video ground truth are essential for performance evaluation and comparison of methods (Aguilar et al 2012). Although many ground-truth datasets have been published, there are still few compared with the diversity of images and applications of interest in the real world. Fast and accurate image annotation remains an open problem in computer vision and related fields. Image annotation tools seek to maximize labeling accuracy while minimizing human labor and time (Wang et al. 2014 and Qin et al. 2018).

In our study, a region-based pixel-wise annotation method was devised in labelling the delaminated portions of the IRT images collected during the NDE survey, e.g. ground truth.

5.2.1 Innovative annotation

Unlike other semantically segmented image datasets, an innovative methodology has been developed to map the ground truth. In carrying out an annotated data set, our team generated a stitched image of the image dataset for the concerned bridges using a student version image processing software, Agisoft 2021 © professional Version. The essence of this is to automatically annotate the individual images in the dataset based on pixels based. In a bid to implement this, it was imperative to ensure that the distortion in the stitched images was corrected and undistorted. Therefore, the orthogonal stitched images, mapped with the ground truth layout provides an image data with geometric properties like that of the reference layout, the delaminated survey layout. This corrected and undistorted stitched image is therefore mapped to the individual images on pixel basis. A mapping function is therefore developed between the ground truth and collected image dataset, thereby transferring the classes of removal (sub-surface delamination) to stitched maps. This algorithm provided a more reliable method of annotating IRT dataset in place of the laborious human annotation usually adopted in image annotation.

Below are the major steps devised for annotation of the IRT images collected (Figure 38).

- Image Stitching using commercial software Agisoft metaphase professional.
- Alignment of stitched and ground truth delamination survey layout.

- Segmentation, morphological operations, connected components and region props algorithms.
- Pixel–wise Annotation.

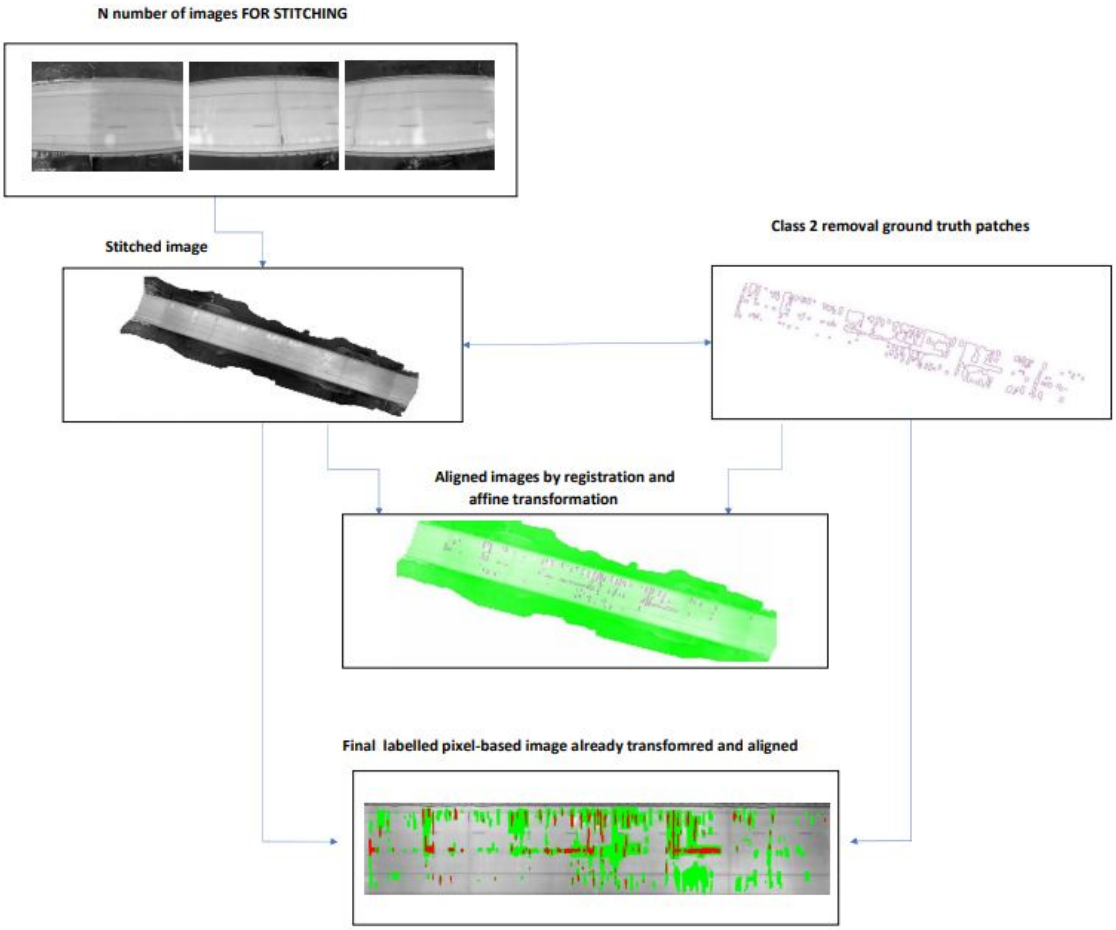


Figure 38. Steps for image pixel-based Annotation steps

5.2.2 Image Stitching.

The Agisoft software as mentioned earlier was used in stitching the images automatically. The

Processing of images with Metaphase includes the following main steps:

- Loading selected overlapping photos into Metaphase
- Inspecting loaded images, removing unnecessary images.
- Aligning photos.
- Building dense point cloud.
- Building mesh (3D polygonal model).

- Generating texture.
- Building tiled model.
- Building digital elevation model.
- Building orthomosaic.
- Exporting results.

Agisoft Metaphase is an advanced image-based 3D modeling solution aimed at creating professional quality 3D content from still images. Based on the latest multi-view 3D reconstruction technology, it operates with arbitrary images. Photos can be taken from any position, providing that the object to be reconstructed is visible on at least two photos. Both image alignment and 3D model reconstruction are fully automated.

Generally, the final goal of photographs processing with Metaphase is to build 3D surface, Ortho mosaic and Digital Elevation Model (DEM). The processing procedure includes four main stages.

1) The first stage is camera alignment. At this stage Metaphase searches for common points on photographs and matches them, as well as it finds the position of the camera for each picture and refines camera calibration parameters. As a result, a sparse point cloud and a set of camera positions are formed. The sparse point cloud represents the results of photo alignment.
2) The next stage is generating dense point cloud, that is built by Metaphase based on the estimated camera positions and pictures themselves.
3) The third stage is generation of a surface: Mesh and/or DEM. 3D polygonal mesh model represents the object surface based on the dense or sparse point cloud, this type of surface representation is not always required, so the user may choose to skip mesh model generation step.
4) After the surface is reconstructed, it can be textured (relevant for mesh model only) or an Orth mosaic can be generated.

Figure 39 shows an example of ortho-mosaic stitched image using Agisoft Professional Metaphase.

Figure 39.Sample of stitched images generated from 20-sub images using Agisoft Professional Metaphase (Park River North Bound, ND).

5.2.3 Alignment of stitched and ground truth delamination survey layout.

Alignment of ground truth delamination survey generated from the study was aligned with the stitched image using image registration and affine transformation, such as scale, translation, and rotation. The affine transformation technique is used to correct lens distortion, panorama stitching, and image registration. Transforming and fusing the images to a large, flat coordinate system is desirable to eliminate distortion and align images.

Table 17 shows transformation matrixes for executing affine transformation and registration of the stitched and ground truth delamination survey layout. In aligning these images for Park River North Bound, the transformation matrix adopted for Translation (T), Rotation (R) and Scale (S) are shown below. The final satisfactory matrix values were gotten from repeated iterations until the images are nearly perfectly aligned. Below are the applied values of T, R and S after several repeated iterations.

$$T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -321 & -112 & 1 \end{bmatrix} \qquad R = \begin{bmatrix} cos87.4 & sin87.4 & 0 \\ -sin87.4 & cos87.4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

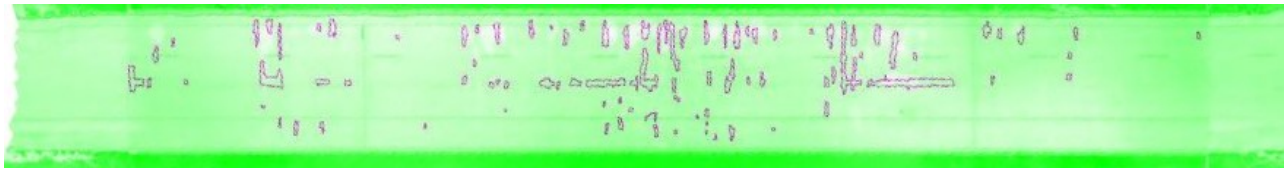
$$S = \begin{bmatrix} 2.6 & 0 & 0 \\ 0 & 2.6 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Table 17 . Affine 2D transformation.

Affine Transform	Example	Transformation Matrix	
Translation		$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ t_x & t_y & 1 \end{bmatrix}$	t_x specifies the displacement along the x axis t_y specifies the displacement along the y axis.
Scale		$\begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$	s_x specifies the scale factor along the x axis s_y specifies the scale factor along the y axis.
Shear		$\begin{bmatrix} 1 & sh_y & 0 \\ sh_x & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	sh_x specifies the shear factor along the x axis sh_y specifies the shear factor along the y axis.
Rotation		$\begin{bmatrix} \cos(q) & \sin(q) & 0 \\ -\sin(q) & \cos(q) & 0 \\ 0 & 0 & 1 \end{bmatrix}$	q specifies the angle of rotation.

The affine transformation was applied on the stitched image as the moving image with respect to the stitched image (Figure 40). The MATLAB code for this is shown in the Appendix D.

(a)



(b)

Figure 40.Image registration and alignment by affine transformation technique for (a) class 2 and (b) class 3 removal.

5.2.4 Binarization, Segmentation and morphological operations

The flowchart adopted for processing and annotating images in this study is shown in Figure 41. Once the ground truth and stitched images have been successfully aligned as discussed in an earlier section, the images are therefore ready as input for further processing.

In simple terms, binarization can be achieved by thresholding. This means assigning all pixels with gray-level lower than a given threshold to either background or the foreground, and the remaining pixels to the other set (Sezgin et al 2004). A thresholding procedure attempts to determine an intensity value, called the threshold, which separates the desired classes. The segmentation is then achieved by grouping all pixels with intensities greater than the threshold into one class and all other pixels into another class. Thresholding is a simple yet often effective means for obtaining a segmentation of images in which different structures have contrasting intensities or other quantifiable features (Sahoo et al. 1988). For example, one way to find regions in an image is to look for abrupt discontinuities in pixel values, which typically indicate edges. These edges can define regions. Morphology is a broad set of image processing operations that process images based on shapes. In a morphological operation, each pixel in the image is adjusted based on the value of other pixels in its neighborhood. Some of the operations used were erode image, dilate image, morphologically open image, morphologically close image, morphological operations on binary images, filling holes/regions in images etc. The segmented images for class 2 and 3 removal are shown (Figure 42).

Assume that the size of the image is M×M and the gray level is L. The gray value of the pixel with coordinates (x, y) is defined as f(x, y), and its average gray value of neighborhood is defined as g(x, y), whose values are between 0 and L-I. If the frequency of (i, j) is Cij, the corresponding joint-probability density Pij is:

$$Pij = Cij / (M \times M),\ 0 \leq I, j \leq (L\text{-}1),\ \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} Pij = 1 \qquad (4)$$

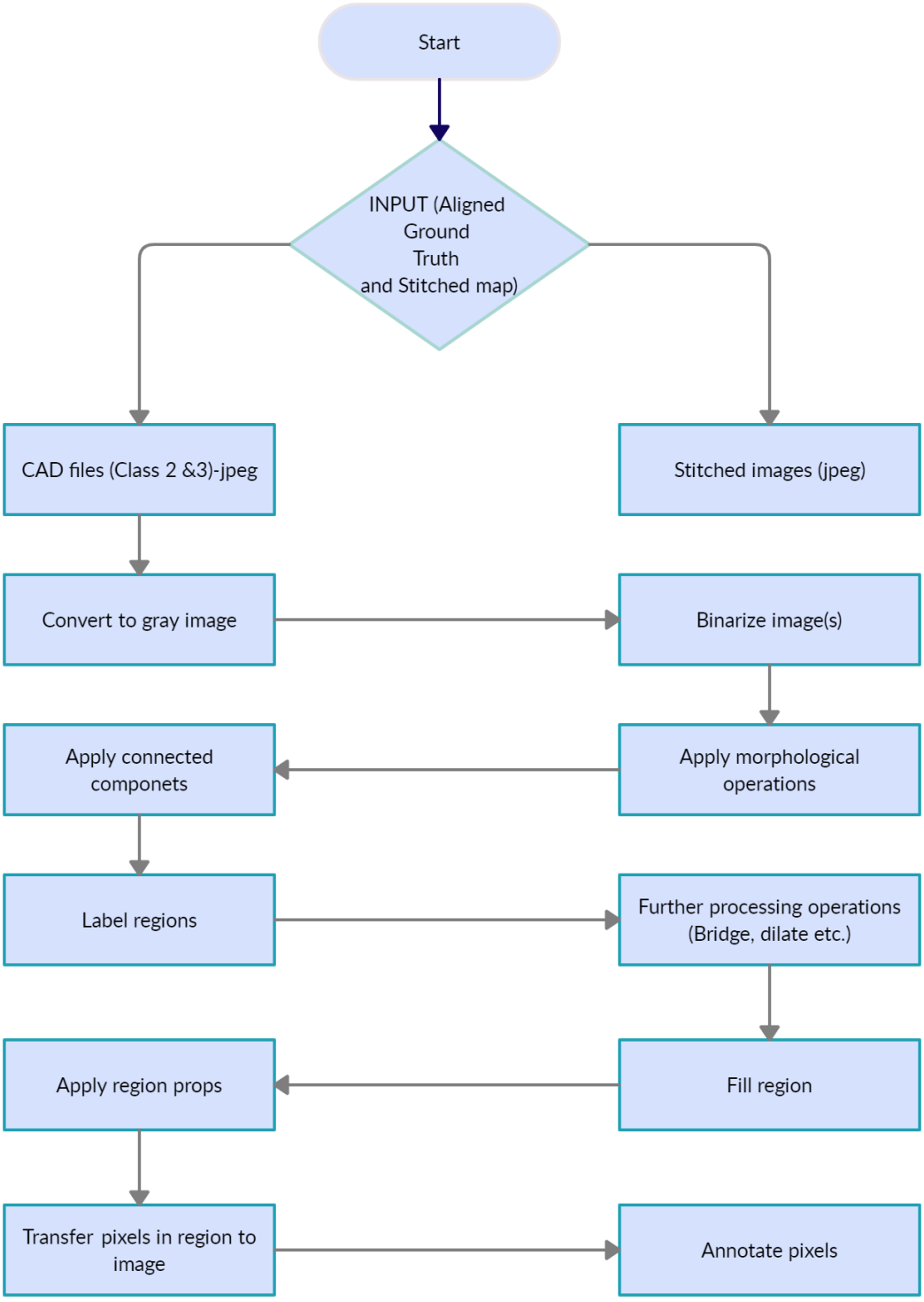
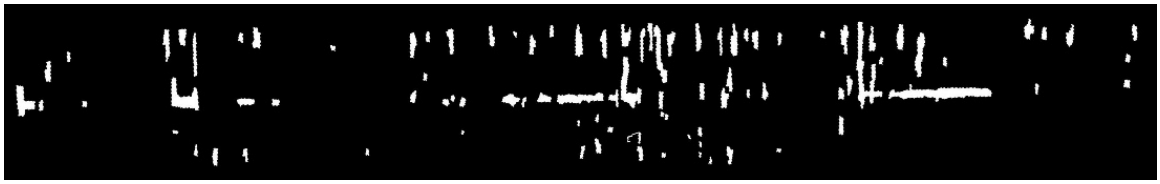


Figure 41.Flow chart for image processing and annotation of IRT

5.2.5 Connected components and region props algorithms.

Pixel connectivity specified as 4 or 8, connects pixels if their edges or corners touch. Four (4)-connected pixels are connected if their edges touch. Two adjoining pixels are part of the same object if they are both on and are connected along the horizontal or vertical direction. Similarly, eight (8)-connected pixels are connected if their edges or corners touch. Two adjoining pixels are part of the same object if they are both on and are connected along the horizontal, vertical, or diagonal direction. And so, the connected components are labelled, filled, and forming region props that will be an input for labelling of the pixels within each region.

(a)



(b)

Figure 42.Image segmentation, binarization of delaminated portions for (a) class 2 and (b) class 3 removal

5.2.6 Pixel –wise Annotation

Afterwards, these pixels within the labelled regions in Figure 43 are annotated as Green (255) for class 2 removal and Red (255) for class 3 removal within the RGB color space. These labelled pixels are thereafter transferred to the stitched images which had earlier been aligned and having similar geometric properties and coordinates.

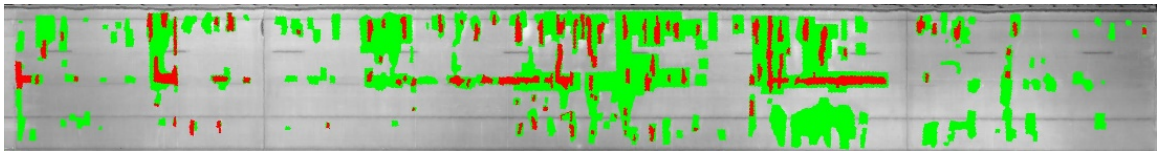


Figure 43. Pixel-based annotated image for Park River North Bound

5.2.7 Saving annotated images.

The already annotated images are saved for the five (5) bridges being assessed and investigated by our team. Images for different bridges are saved with unique file names having two (2) sets of the raw/original mosaic and annotated images for each folders of the bridge decks.

5.3 Impact Echo (IE) Annotation

The collected IE data was annotated and validated by cross referencing the location of IE signals on the ground truth delamination survey layout. Figure 44 shows the bridge deck layout for Forest River NB and the embedded delamination classes 2 and 3 in pink and blue color respectively. For this deck, the overall deck area was measured to be 7,745 sq. ft. while the class 2 and 3 repairs areas were 1,956 sq.ft. and 481 sq.ft respectively. The IE tests taken at each of the 10ft x 10ft grid amounting to at least 100 IE test points per region. Magnified sections of each of the origins 2A, 2B and 2C are shown in figure 45 (a-c).

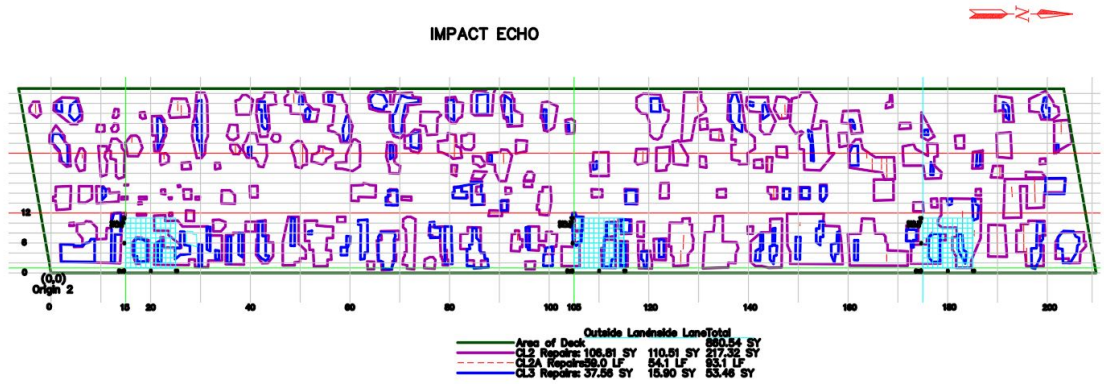


Figure 44.Impact Echo test points 2A, 2B and 2C for Forest River NB

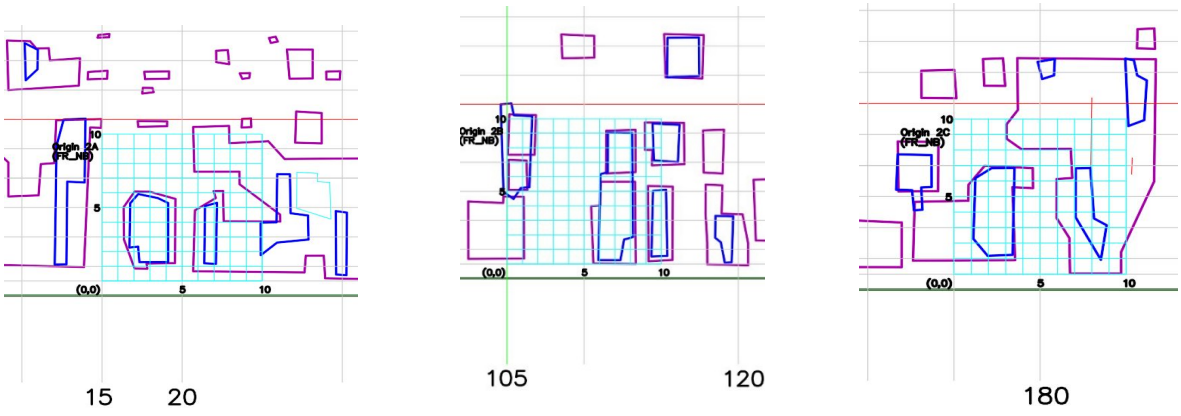


Figure 45.Magnified up sections of IE test points (a)Origin 2A (b) Origin 2B (c) Origin 2C showing a 10ft-by-10ft region at 1ft by 1ft grid lines.

The output IE data are saved in a 'lvm' as earlier discussed while a summary presentation of the annotated test points is shown in Table 17. The data of test, file number, file name, global X and Y offset, local X and Y offset, data quality and the class removal are shown below. Class 1 indicates the IE test points were taken on a sound concrete surface. The annotated IE data can be found in Appendix A.

Table 18. showing summary of selected IE annotation of each test result for Forest River NB bridge at Origin 2A

File No.	FileName_LVM	Global X (ft)	Global Y (ft)	Origin	Local_X_ft	Local_Y_ft	Data Quality	Removal Class
0	Metal.lvm	15	1	2A	0	0	BAD	1
1	Metal_1.lvm	15	1	2A	0	0	GOOD	1
2	Metal_2.lvm	15	1	2A	1	0	GOOD	1
3	Metal_3.lvm	15	1	2A	2	0	GOOD	1
4	Metal_4.lvm	15	1	2A	3	0	GOOD	1
5	Metal_5.lvm	15	1	2A	4	0	GOOD	1
6	Metal_6.lvm	15	1	2A	5	0	GOOD	1
7	Metal_7.lvm	15	1	2A	6	0	GOOD	1
8	Metal_8.lvm	15	1	2A	7	0	GOOD	1
9	Metal_9.lvm	15	1	2A	8	0	GOOD	1

5.4 Ground penetrating Radar (GPR) Annotation

From the field recoding of the GPR test data collected for the bridges under study, we plotted the scan lines on the ground truth delamination survey layout in CAD format. The grids in both directions (X & Y) were also plotted in the CAD drawing. A typical GPR scan test is shown in Figure 46, a magnified section of the bridge deck. A longitudinal scan is shown along grid-line y_i. The coordinates of the scan along the constant y_i

Assuming Length of scan, = L

Number of signals = n

Increment, dx = $\frac{L}{n}$,

Therefore the signals (1, 2, 3, 4, …….n) are discretized and assigned coordinates points according to the representation;

$x_1 = dx$, $x_2 = x_1 + 1.dx$, $x_3 = x_1 + 2.dx$, $x_4 = x_1 + 3.dx$, ……$x_n = x_1 + (n-1)dx$,

and therefore, the coordinates of the signals are represented as (x_1, y_i), (x_2, y_i), (x_3, y_i), (x_4, y_i), (x_5, y_i)…….(x_n, y_i), In annotating the signal, all the signals whose coordinates falls between the ground truth coordinates (X_0, y_I) and (X_1, y_I) are labelled as class 1, those that fall between (X_1, y_I) and (X_2, y_I) are labelled class 3, those that fall between between (X_2, y_I) and (X_3, y_I) are labelled class 2, those that fall between (X_3, y_I) and (X_4, y_I) are labelled class 3, those that fall between (X_4, y_I) and (X_5, y_I) are labelled class 2 and so on. This is also repeated for all longitudinal and the transverse signals. Table 18 shows a typical GPR annotated data for the investigated bridges.

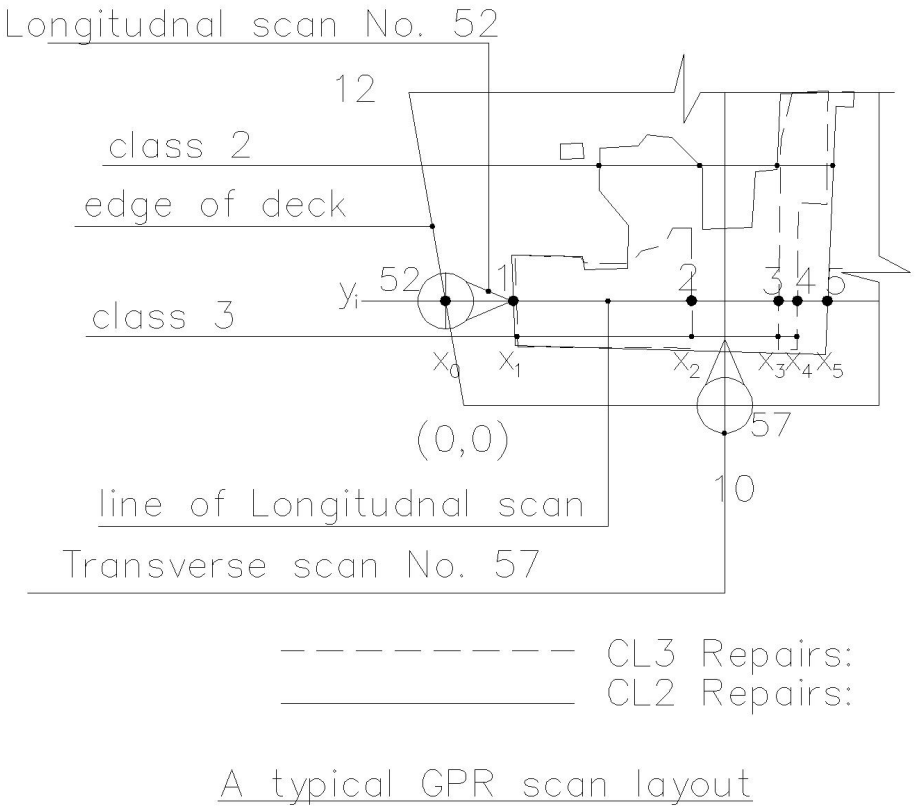


Figure 46.Typical GPR scan test layout

In annotating signal 52 for Forest River NOTH Bound, the X-axis has an increment of 10ft from the origin while the Y-axis has an increment of 2 ft from the origin to 12 ft. The entire width of this bridge is about 36ft. while the length is about 210 ft. The part of the Forest River NB investigated for GPR is the Right lane as shown in Figure 47 and 48. The delaminated portions are regions shown in pink and blue for classes 2 and 3 while the remaining portion are sound concrete portions depicted with class 1.

Table 19. Excerpt of a GPR signal annotation for Park river median, File_003 (The signal continues to [512 16282])

n	**1**	**2**	**3**	**4**	**5**	**6**
X(ft)	0.022278	0.044556	0.066834	0.089112	0.11139	0.133668
Y(ft)	4	4	4	4	4	4
Class	1	1	1	1	1	1
Time_ns	Amp_1	Amp_2	Amp_3	Amp_4	Amp_5	Amp_6
0	32762	32757	32767	32766	32765	32769
0.0234375	32762	32757	32767	32766	32765	32769
0.046875	32762	32757	32767	32766	32765	32769
0.0703125	32759	32758	32769	32766	32768	32768

0.09375	32755	32761	32770	32765	32772	32767
0.1171875	32753	32765	32772	32765	32776	32766
0.140625	32751	32769	32772	32765	32779	32764
0.1640625	32751	32773	32772	32765	32780	32762
0.1875	32751	32775	32772	32765	32781	32762

The scans are carried out for the longitudinal test starting from the origin (0,0) with a GPR scan number 50, in the positive X direction (X+). The next scan starts from the next Y-coordinate at 2ft towards the negative X direction (X-). These measurements are taken with respect to the origin (0,0). The scans along the longitudinal direction continues to signal number 55. Then field note showing the summary of the scan direction is in the GPR folder in the One drive link provided in Appendix A.

Similarly, scans were carried out in the transverse direction starting from the origin (0,0) with a scan number 56 along the width to the extent of 12ft in the positive +Y direction. The next scan starts from the next X-coordinate at 10ft also towards the positive +Y direction. This transverse scan continues to scan 77.

To annotate the GPR signals, the coordinates of the points at which the signal lines intercepts with the classes of removal were found. These points were generated in CIVIL-3D and their coordinates (Eastings and Northings) were automatically generated for these points and exported and saved in MS Excel for each bridge decks. The points of interception were saved in the PENZ format.

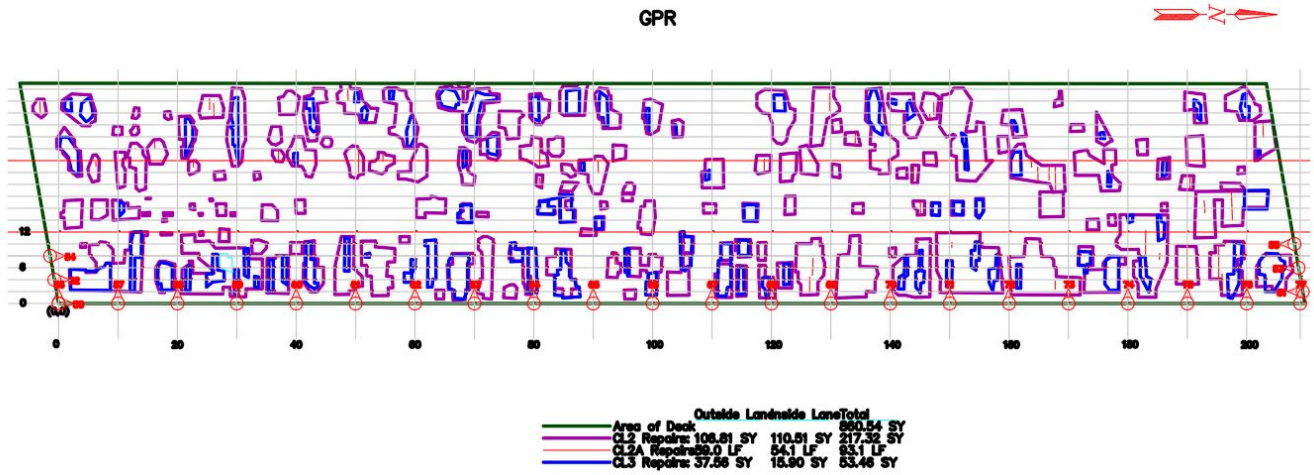


Figure 47.GPR test points 2A, 2B and 2C for Forest River NB

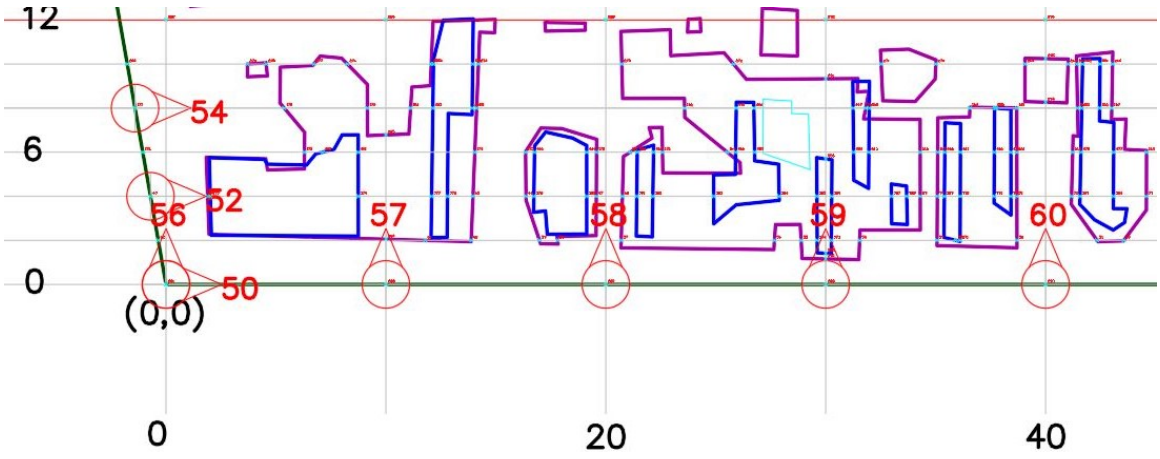


Figure 48.Magnified section of some GPR test points for Forest River NB

The coordinate regions generated is required for annotating the GPR test data result. From earlier sections, the structure of the GPR test files has the structures: [rows columns] = [512 16,383] for the longitudinal scans and [rows columns] = [512 1225] for the transverse scans along the width of the bridge. The scan output is therefore discretized by dividing the overall length of each scan by the total number of the columns. The overall length of scan 52 is 209.79 ft. Therefore, the scan distance increment considered for the bridge under study is represented as:

Increment (du) = $\frac{209.79}{16383}$, du = 0.012805ft.

The cumulative scan distance is represented for each discretized increment of each signal starting from 0.01280 ft for point 1 to 209.79 ft for point 16,383 and is entered in the concerned file. This represents the X-coordinate for the signal considered while the Y-coordinate is along 4ft on the Y-axis. The signal columns are therefore labelled with each removal class (1, 2 or 3) using the exported Excel having the coordinate regions of the classes of removal. For instance, from Figure 48, signal 52 starts with class 1 (starts)-class 2&3-class 2-class 2&3-class 2-class 1…. class 1(ends). At points where class 2 and 3 overlap, the deeper delamination is shown, therefore class 3 delamination is shown. (Figure 47 and 48). This process is repeated for all longitudinal and transverse signals, a total of 209 scan files and saved in their respective folders. The full annotated GPR data can be found in Appendix A.

5.5 Significance of dataset and future works

Bridge inspectors face three challenges when they use NDE for bridge deck evaluation.

Proof of concepts are performed under controlled environment: Most studies proving NDE applications for deck evaluation have been performed on specimens made in laboratories, with artificial defects, and under controlled conditions. Therefore, these case studies do not mimic real

inspections scenarios and challenges associated with them. Sometimes NDE methods resulted in less conclusive evaluations compared to the proof-of-concept studies when applied to in-service bridge decks. Promoting NDE for concrete deck evaluation in North Dakota requires field investigations to identify and address challenges associated with field inspection.

Occasional field investigations using NDE are not validated with the ground truth: Validation of NDE methods require coring or destructive testing of bridge decks. Cores are normally taken to measure concrete properties of bridge decks not to validate NDE. In addition, bridge owners may not see a reason to inspect a deck using NDE that is being destructed or replaced. Therefore, it is very rare to see NDE validation through ground truth in State and Federal repositories. Some efforts have been made in the past to compare the results of one NDE method to another, or to conventional sounding techniques. However, either very few or no comparisons between NDE and ground truth of bridge decks have been published publicly in the past. The uncertainty associated with NDE methods can be eliminated or reduced significantly if their results are compared to the actual condition of bridge decks.

Data interpretation: Unlike traditional sounding techniques, deck evaluation using NDE methods often require trained operators for processing and decision making that could potentially increase the cost of the deck inspections. Without trusting NDE data, some bridge owners cannot justify covering the cost associated with NDE inspections. Collecting high resolution and ground truth validated NDE data would be the first step in using advanced artificial intelligence for bridge evaluation. The role of trained operators could be significantly minimized if proper and robust artificial intelligence solutions are developed for NDE data analysis. However, without NDE datasets that are accurately annotated, artificial intelligence solutions will not be robust.

To its capacity this dataset will address the lack of field validated dataset. The collected NDE data have been annotated by cross-referencing the geospatial locations associated with defects (classes of removal) with NDE data collected from same location. The dataset will be a useful resource for benchmarking, developing, and training advanced AI models such as CNN’s to minimize the role of expert’s opinion in NDE data quality control and analysis.

6. CONCLUSION AND RECOMMENDATIONS

6.1 Conclusions

Bridge repairs and maintenance is expected to be on the rise in years to come based on the age, structural condition, and deterioration of bridges in North Dakota and the US at large. Traditional methods of bridge evaluation for assessment of deteriorations such as cracks, delamination, corrosion, excessive deformation etc. are time consuming, prone to error, disrupts traffic, have safety concerns etc. Non-destructive evaluation (NDE) methods can provide essential information for bridge stakeholders however, they often require trained operators to interpret the data. It is possible to eliminate the role of trained operators though development and application of Artificial Intelligence (AI) models to NDE data for bridge evaluation. Effective AI models cannot be developed without annotated and realistic field datasets which are extremely rare in bridge studies, if not nonexistent.

In this study, the research team collected and annotated Infrared Thermography (IRT), Impact Echo (IE), and Ground Penetrating Radar (GPR) from five in-service bridges decks located in Walsh county in North Dakota before they underwent extensive repair by North Dakota Department of Transportation. The procedure used to capture the NDE data has been selected through the review of past collection. A total of 1,064 thermal images, 2275 impact echo, and 655,818 GPR signals were collected in this study.

To ensure the collected data had the sufficient quality, several data quality metrics were considered and applied on the data. For IRT images, Perception Based image quality evaluator (Piqe), Naturalness image quality evaluator (Niqe), Blindness/Reference less Image Spatial Quality Evaluator (Brisque), were adopted in classifying the image quality. The result of image quality control showed that the data collected in this study had the required quality. Other quality assessment metrics were considered for the IE and GPR signals including value checks, null value, duplicate values, and descriptive statistical checks. The results of signal quality control were also satisfactory for the GPR and IE data.

The bridge decks were prepared for ground truth data collection by marking and removing delaminated portions of the deck identified by chain-dragging and other visual investigation methods. Depending on the extent of sub-surface delamination, they were classified as class 1 (sound concrete removal with no delamination), class 2a (concrete removal around corroded reinforcing top bars up to 0.75 inches clearance around bars) class 2 (concrete removal to the top of the reinforcing bars) and class 3 (delaminated removal up to the top of the bottom bar of the top mat). Coordinates and layout of the delamination in form of removed delaminated patches, photographs of exposed decks, classes of deck exposures, presence of rebar corrosion etc. were collected and documented prior to repair of

deck. A delamination survey layout was prepared from the information gathered to serve as ground truth data. The IRT, IE and GPR field test data were therefore annotated based on these classes of delaminated sub-surface removal. For the purpose of AI, the data was annotated as class 1, 2 and 3.

Prior to the actual IRT image annotation, a set of IRT images were stitched for each of the bridge decks. An image-processing technique was developed using affine transformation and image registration, segmentation, morphological operations, connected components, region props etc. to align the stitched maps with the ground truth and to execute a semi-automatic pixel–wise annotation of the IRT images. The coordinate of the classes of delamination (region props) for each of the delamination survey were transferred onto the stitched images and therefore annotated pixel-wise with pixel values and color for each class of delamination: Green for class 2 (G=255) delamination removal and Red for class 3 (R=255) delamination removal in Red, Green, Blue (RGB) color space.

Similarly, the collected IE and GPR signal data were also annotated and validated with the ground truth delamination survey layout. The output IE data are saved in a 'lvm.' The annotation of all the test points (2275 IE test points) was done by cross referencing the ground truth maps with the global coordinates of IE. To annotate the GPR signals, the coordinates of the points at which the signal lines intercepts with the classes of removal were noted. These points were generated in CIVIL-3D and their coordinates (Eastings and Northings) were automatically generated for these points and exported and saved in MS Excel for each bridge decks. The points of interception were saved in the PENZ format. The signal columns of the entire 209 files were therefore labelled with each removal class (1, 2 or 3) using the exported Excel having the coordinate regions of the classes of removal.

Reviewing the contact and non-contact NDE methods adopted in the assessment, one can conclude the following:

- A set of high-quality IRT, GPR, and IE data for five bridges decks were successfully generated.
- The annotations of the IE and GPR were completed for all signals by assigning a class of removal and local coordinate with respect to a bottom-left corner of each bridge deck.
- The annotation for the IRT images were completed through matching ground truth maps with stitched thermal images and assigning a class to each pixel.

These set of data will eventually be benchmarked and correlated in future studies with the validated ground truth data generated from on-site chain dragging by experienced inspector. The ground truth depicts the exact condition of the bridge deck. An on-going preliminary study results on the IRT data revealed an average of 69% detection accuracy in comparison with the ground truth. (Ichi et al, 2021).

6.2 Recommendations and Future works

According to the results of this study, the following recommendations regarding use of advanced NDE methods such as UAS's (IRT) IE and GPR for NDE are highlighted as the following:

6.2.1 Subcontractors

From review of past studies and further studies our team is embarking on, it is seen that weather and other known parameters have significant effect and impact on the quality of data being collected. These have been discussed and viable recommendations suggested for future bridge investigation.

- **Condition of deck prior to evaluation**

Past studies suggest that the presence of water-filled voids makes subsurface defect detection easier than air-filed voids using GPR and IRT. This could be because of water having a higher thermal emissivity than air. Therefore, acquiring NDE data after a rainy day (or wetting the deck) and allowing the deck's surface to dry up would most likely yield a higher contrast between sound concrete and delaminated concrete. However, initial investigation on GPR and IRT data collected in this study showed that these methods are still effective in evaluation of dry decks. Even though presence of water in concrete can change the properties of IE signals; however, there is no evidence of altered accuracy in delamination detection as a result of water content variation.

- **Weather Condition**

Environmental conditions such as changes in ambient temperature (or solar loading), humidity and wind speed will change surface emissivity and therefore, will affect the thermal radiation of concrete bridge decks. Favorable environmental conditions include sufficient solar loading or changes in ambient temperature to produce a thermal gradient in the concrete. Commonly, four hours of sun exposure is recommended to generate a sufficient thermal gradient between defected and sound concrete. Preliminary findings based on the ambient conditions (temperature, humidity, and wind speed) collected reveals that increase in temperature is favorable to IRT data collection if the objective is to detect both surface and subsurface defects via thermography. Weather information from nearby weather stations should be collected before and during each inspection for proper weather monitoring. An ideal situation for UAS-assisted IRT data collection in North Dakota would be in a sunny day in summer with relatively low wind speed. The peak temperature in a sunny day and not rainy day should be foremostly considered and thereafter the humidity since thermal cameras function based on thermal properties. Ideally, IRT data can be collected any time of the year, it should be noted that the magnitude of temperature change (Δt) will be smaller during the winter months. It is therefore recommended that testing should not be carried out when ambient air temperatures are less than 32°F as ice in delamination will cause false readings. Standards and past studies recommend that as a guide, an ambient temperature rise of 20°F or higher after sunrise and winds speed less than 15

mph, will allow favorable data collection on Portland-cement concrete or asphalt surfaces. It is also strongly recommended to avoid testing when the wind velocity exceeds 30 mph. In addition, in line with the Federal Aviation Administration (FAA) flight restriction at nights as a result of safety concerns of personnel, equipment and operations, the day heating, passive IRT is recommended. It is therefore suggested from our investigation that the day and time with the highest temperature difference from the time of sunrise, least prevailing wind speed and favorable humidity is appropriate for IRT data collection. For IE and GPR, there was no tangible change in the environmental effect since all the data was collected in three consecutive days in summer.

Camera sensor and specifications

Some parameters of IRT cameras should be considered for delamination detection such as the spectral range, spatial resolution, temperature range and frame rate. The spatial resolution of an IRT camera primarily depends on the object-to-camera distance, lens system and detector size. The spatial resolution decreases with increasing object-to-camera distance, while lens systems with a small field of view (FOV) have higher spatial resolution. At least an IRT camera having a low focal length and resolution of 640 by 480 or 640 by 512 that can detect the IR waves within the IR spectral band is recommended. Cameras with higher spatial resolutions are preferable as they become more commercially available.

IRT acquisition requirement

The image acquisition plan depends on the:

- Type of terrain to be stitched and processed.
- Ground Sampling Distance (GSD): The GSD required by the project specifications will define the distance (flight height) at which the images have to be taken. For example, a GSD of 5 cm means that one pixel in the image represents linearly 5 cm on the ground (25 square centimeters for a square pixel).
- Overlap: The overlap depends on the type of terrain that is mapped and will determine the rate at which the images have to be taken. It is usually recommended that images collected should have sufficient overlap.

It is recommended to acquire the images with a regular grid pattern. The recommended overlap is at least 75% frontal overlap (with respect to the flight direction) and at least 60% side overlap (between flying tracks). The camera should be maintained as much possible at a constant height over the terrain or object to ensure a desired GSD. Lower GSD can show delaminated regions in better details. Currently, there are no required or recommended GSD for IRT. The research team will investigate the effect of GSD on size of delamination in the future. For projects with multiple flights there should be

overlap between the different flights and the conditions (sun direction, weather conditions, no new buildings, etc.) should be similar to each other.

FAA flight regulations

We highly recommend that contractors ensure that they follow the Certificated Remote Pilots including Commercial Operators regulations and certification process stated by the Federal Aviation Authority (FAA). To fly under Part 107 rules, there are 3 main steps:

Step 1: Learn the Rules

Step 2: Become an FAA-Certified UAS Pilot by Passing the Knowledge Test

Step 3: Register the UAS with the FAA.

IE and GPR data collection.

IE and GPR methods have been seen to be complementary in depth of detection. The acoustic scanning method provides rapid mapping of shallow delamination (up to 5–6 cm). GPR antennas used for concrete structure or bridge scans have penetration depths of 30-50 cm. The maximum depth of delamination in this study was limited to class 3 which is immediately below the top rebar of the bridges inspected. In addition, data should be collected at close grid spacings of minimum of 30 by 30 cm to maximum 100 by 100 cm in both directions for impact echo for better representation of the deck area. Although a smaller grid space will consume more time during data collection. Depending on the width and length of the bridge deck under study, GPR grid spacings should be as close as possible in both longitudinal and transverse directions. In this study we have adopted a grid spacing 60 cm increment along the transverse section and a 300cm-600cm along the longitudinal section of the bridge deck. However, closer grid spacing is recommended depending on the site and project constraints. In future studies, it is also recommended that the precise coordinate location of the IE and GPR field data points are collected with a GPS equipment. These coordinates should be referenced to the coordinates of the delamination and the bridge deck.

6.2.2 NDDOT

- Past studies and ours show that UAS-assisted evaluation of bridge decks were 4 to 8 times faster than contact methods. A priority for NDDOT could be to invest in improving UAS capabilities for full evaluation of concrete bridge decks via AI. Investing in autonomous platforms for data collection without possible traffic disruption and safety concerns will tremendously help improve bridge inspection in ND.

- NDDOT is recommended to collect visual and IRT data using UAS on every bridge before and after repair. The images, along with the ground truth maps, can be used to create a sustainable model for concrete bridge deck delamination detection and asset management in ND.
- More research initiatives are required to develop AI models for interpretation and classification of bridge deck data. AI virtual libraries can be added for bridge robust bridge evaluation.
- North Dakota is the home of extreme weather (low/high temperature, wind speed, and precipitation) which affect the thermal emissivity. It is recommended to investigate the effects of extreme weather on the collected data via UAS for bridge deck evaluation.
- Investigating the relationship between concrete deck surface and subsurface defects using AI models could be another intriguing aspect that can be further explored by NDDOT.
- Establishment of NDE data quality control protocols for successful defect detection based on the ground truth maps generated in this study.
- NDDOT in collaboration with research institutions should set up a sustainable and running laboratory for further NDE studies and research endeavors aimed at improving the capabilities of advanced NDE methods for improved results.

In conclusion, the possibilities of IRT and other advanced techniques in bridge evaluation leaves no doubt that if well explored and implemented in future bridge studies and projects, will help to accrue greater benefits in terms of overall cost reduction, time, safety, effective structural health monitoring (SHM) and longevity of bridge infrastructures.

REFERENCES

Abouhamad, M., Dawood, T., Jabri, A., Alsharqawi, M., & Zayed, T. (2017). Corrosiveness mapping of bridge decks using image-based analysis of GPR data. Automation in Construction, 80, 104-117.

Ahmadvand, M., Dorafshan, S., Azari, H., & Shams, S. (2021, March). 1D-CNNs for autonomous defect detection in bridge decks using ground penetrating radar. In Health Monitoring of Structural and Biological Systems XV (Vol. 11593, p. 115930M). International Society for Optics and Photonics.

Ahmed, H., La, H. M., & Tran, K. (2020). Rebar detection and localization for bridge deck inspection and evaluation using deep residual networks. Automation in Construction, 120, 103393.

Aguilar, F. D., & Hirata, N. S. (2012, March). ExpressMatch: a system for creating ground-truthed datasets of online mathematical expressions. In 2012 10th IAPR International Workshop on Document Analysis Systems (pp. 155-159). IEEE.

Agaian, S. S., Silver, B., & Panetta, K. A. (2007). Transform coefficient histogram-based image enhancement algorithms using contrast entropy. IEEE transactions on image processing, 16(3), 741-758.

Agaian, S. S., Lentz, K. P., & Grigoryan, A. M. (2000, September). A new measure of image enhancement. In IASTED International Conference on Signal Processing & Communication (pp. 19-22).

American Institute of Physics. Kylili, A., Fokaides, P. A., Christou, P., & Kalogirou, S. A. (2014). Infrared thermography (IRT) applications for building diagnostics: A review. Applied Energy, 134, 531-549.

ASCE (American Society of Civil Engineers), 2019, ASCE report card for America's infrastructure. http:// www.infrastructurereportcard.org/ (accessed on March 29, 2021).

ASTM, Standard Test Method for Evaluating Asphalt-Covered Concrete Bridge Decks Using Ground Penetrating Radar, in ASTM D6087-08, West Conshohocken, PA., 2010.

ASTM, Standard Test Method for Detecting Delamination in Bridge Decks Using Infrared Thermography, in ASTM D4788, West Conshohocken, PA., 2014.

ASTM Standard D4580-03(2007), 2007, "Standard Practice for Measuring Delaminations in Concrete Bridge Decks by Sounding," ASTM International, West Conshohocken, PA, 2007, 4 pp.

ASTM, C. (2000). Test method for measuring the P-wave speed and the thickness of concrete plates using the impact-echo method. Annual Book of ASTM Standards, C1383-15.

ASTM AD (2013) 4788: standard test method for detecting delaminations in bridge decks using infrared thermography. ASTM International, West Conshohocken 31.

ASTM, “Standard Practice for Measuring Delaminations in Concrete Bridge Decks by Sounding,” D4580-03, ASTM International, West Conshohocken, PA, 2007.

Apel, D., Behme, W., Eberlein, R., & Merighi, C. (2015). Successfully control data quality: Practice solutions for business intelligence projects. Heidelberg: dpunkt. verlag.

Arbesser, C., Spechtenhauser, F., Mühlbacher, T., & Piringer, H. (2016). Visplause: Visual data quality assessment of many time series using plausibility checks. IEEE transactions on visualization and computer graphics, 23(1), 641-650.

Arbesser, C., Spechtenhauser, F., Mühlbacher, T., & Piringer, H. (2016). Visplause: Visual data quality assessment of many time series using plausibility checks. IEEE transactions on visualization and computer graphics, 23(1), 641-650.

Azeroual, O., Saake, G., & Wastl, J. (2018). Data measurement in research information systems: metrics for the evaluation of data quality. Scientometrics, 115(3), 1271-1290.

Azari, H., & Lin, S. (2019). Evaluation of the impact echo method for concrete bridge decks with asphalt overlays. Transportation Research Record, 2673(2), 436-444.

Bringier, B., Richard, N., Larabi, M. C., & Fernandez-Maloigne, C. (2006, September). No-reference perceptual quality assessment of colour image. In 2006 14th European Signal Processing Conference (pp. 1-5). IEEE.

Caleb Hing, C. L. Ph.D.; P.E., M.ASCE1 ; and U. B. Halabe, Ph.D., P.E., F.ASCE, Non-destructive Testing of GFRP Bridge Decks Using Ground Penetrating Radar and Infrared Thermography, 2010.

Catbas, F. N., Hiasa, S., Khuc, T., Matsumoto, M., & Mitani, K. (2015). Development, implementation and evaluation of image-based technologies for civil infrastructure systems. Report Submitted to West Nippon Expressway Company limited (NEXCO-West), Osaka, Japan.

Chen, S. D., & Ramli, A. R. (2003). Minimum mean brightness error bi-histogram equalization in contrast enhancement. IEEE transactions on Consumer Electronics, 49(4), 1310-1319.

Chen, S.; Laefer, D. F.; Byrne, J.; and Natanzi, A. S., The Effect of Angles and Distance on Image-Based, Three-Dimensional Re-constructions, in Proc., 2nd Int. Conf. on Engineering Sciences and Technologies, 2757–2761. Boca Raton, FL: CRC Press., 2017.

Cheng, Q., Zhang, Q., Fu, P., Tu, C., & Li, S. (2018). A survey and analysis on automatic image annotation. Pattern Recognition, 79, 242-259.

Caviedes, J., & Oberti, F. (2004). A new sharpness metric based on local kurtosis, edge and energy information. Signal Processing: Image Communication, 19(2), 147-161

Colla, C., & Lausch, R. (2003). Influence of source frequency on impact-echo data quality for testing concrete structures. Ndt & E International, 36(4), 203-213.

Colla, C., & Lausch, R. (2003). Influence of source frequency on impact-echo data quality for testing concrete structures. Ndt & E International, 36(4), 203-213.

Dorafshan, S., & Azari, H. (2020a). Evaluation of bridge decks with overlays using impact echo, a deep learning approach. Automation in Construction, 113, 103133.

Dorafshan, S., & Azari, H. (2020b). Deep learning models for bridge deck evaluation using impact echo. Construction and Building Materials, 263, 120109.

Dorafshan, S., & Maguire, M. (2018). Bridge inspection: human performance, unmanned aerial systems and automation. Journal of Civil Structural Health Monitoring, 8(3), 443-476.

Dorafshan, S., Campbell, L. E., Maguire, M., & Connor, R. J. (2021). Benchmarking Unmanned Aerial Systems-Assisted Inspection of Steel Bridges for Fatigue Cracks. Transportation Research Record, 03611981211001073.

Dorafshan, S., Thomas, R. J., & Maguire, M. (2018). Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete. Construction and Building Materials, 186, 1031-1045.

Federal Aviation administration, Certificated Remote Pilots including Commercial Operators - [online] Available at: http:// www.faa.gov/uas/commercial_operators/ (accessed on June 21, 2021).

Farrag, S., Yehia, S., & Qaddoumi, N. (2016). Investigation of mix-variation effect on defect-detection ability using infrared thermography as a nondestructive evaluation technique. Journal of Bridge Engineering, 21(3), 04015055.

Federal Highway Administration Research and Technology Nondestructive Evaluation (NDE) Web Manual [online] Available at: https://fhwaapps.fhwa.dot.gov/ndep/TechnologyMenu.aspx

Ferzli, R., & Karam, L. J. (2009). A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB). IEEE transactions on image processing, 18(4), 717-728.

Fhwaapps.fhwa.dot.gov. 2020. NDE Technology. [online] Available at: <https://fhwaapps.fhwa.dot.gov/ndep/DisplayTechnology.aspx?tech_id=9> [Accessed 1 October 2020].

Fhwaapps.fhwa.dot.gov. Long-Term Bridge Performance (LTBP) Program Protocols, Version 1 [online] Available at: https://www.fhwa.dot.gov/publications/research/infrastructure/structures/ltbp/16007/036.cfm

FLIR Systems, Thermal Imaging Guidebook for Building and Renewable Energy Applications, http://www.flirmedia.com/MMC/THG/Brochures/T820325/ T820325_EN.pdf, (2013) [Accessed 04-06-2017].

Fu, Y. Y. (2006). Color image quality measures and retrieval.

Fu, Y., & Wang, S. (2016). A no reference image quality assessment metric based on visual perception. Algorithms, 9(4), 87.

Gassman, J. J., Owen, W. W., Kuntz, T. E., Martin, J. P., & Amoroso, W. P. (1995). Data quality assurance, monitoring, and reporting. Controlled clinical trials, 16(2), 104-136.

Gao, C., Panetta, K., & Agaian, S. (2012, April). A new color contrast enhancement algorithm for robotic applications. In 2012 IEEE International Conference on Technologies for Practical Robot Applications (TePRA) (pp. 42-47). IEEE.

Gebauer, M., & Windheuser, U. (2015). Structured data analysis, profiling and business rules. Wiesbaden: Springer Fachmedien Wiesbaden.

Gucunski, N., Slabaugh, G., Wang, Z., Fang, T., & Maher, A. (2008). Visualization and interpretation of impact echo data from bridge deck testing. Transportation Research Record, 2050, 11-21.

Gucunski, N., Kee, S., La, H., Basily, B., & Maher, A. (2015). Delamination and concrete quality assessment of concrete bridge decks using a fully autonomous RABIT platform. Structural Monitoring and Maintenance, 2(1), 19-34.\

Gucunski, N., & National Research Council. (2013). Nondestructive testing to identify concrete bridge deck deterioration. Strategic Highway Research Program (SHRP) 2 Report S2-R06A-RR-1-Transportation Research Board.

Gucunski, N., Romero, F., Kruschwitz, S., Feldmann, R., Abu-Hawash, A., & Dunn, M. (2010). Multiple complementary nondestructive evaluation technologies for condition assessment of concrete bridge decks. Transportation Research Record, 2201(1), 34-44.

Hasler, D., & Suesstrunk, S. E. (2003, June). Measuring colorfulness in natural images. In Human vision and electronic imaging VIII (Vol. 5007, pp. 87-95). International Society for Optics and Photonics.

Halabe, U. B., Hing, C. L., Klinkhachorn, P., & GangaRao, H. V. (2007, March). Subsurface defect detection in FRP bridge decks using ground penetrating radar. In AIP Conference Proceedings (Vol. 894, No. 1, pp. 1443-1452).

Halonen, R., Westman, S., & Oittinen, P. (2011, January). Naturalness and interestingness of test images for visual quality evaluation. In Image Quality and System Performance VIII (Vol. 7867, p. 78670Z). International Society for Optics and Photonics.

Heinrich, B., Hristova, D., Klier, M., Schiller, A., & Szubartowicz, M. (2018). Requirements for data quality metrics. Journal of Data and Information Quality (JDIQ), 9(2), 1-32.Hiasa, S., Birgul, R., & Catbas, F. N. (2016). Infrared thermography for civil structural assessment: demonstrations with laboratory and field studies. Journal of Civil Structural Health Monitoring, 6(3), 619-636.

Hing, C. L. (2006). Nondestructive evaluation of fiber reinforced polymer bridge decks using ground penetrating radar and infrared thermography.

Hing, C. C., & Halabe, U. B. (2010). Nondestructive testing of GFRP bridge decks using ground penetrating radar and infrared thermography. Journal of Bridge Engineering, 15(4), 391-398.

Hinrichs, H. (2002). Datenqualitätsmanagement in data warehouse-systemen (Doctoral dissertation, Universität Oldenburg).

Kee, S. H., Oh, T., Popovics, J. S., Arndt, R. W., & Zhu, J. (2012). Nondestructive bridge deck testing with air-coupled impact-echo and infrared thermography. Journal of Bridge Engineering, 17(6), 928-939.

Khatereh Vaghefi, Ph.D., S.M.ASCE ; Theresa (Tess) M. Ahlborn, Ph.D., P.E., M.ASCE ; Devin K. Harris, Ph.D., A.M.ASCE ; and Colin N. Brooks, Combined Imaging Technologies for Concrete, 2013.

Kim, M., & Chung, M. G. (2008). Recursively separated and weighted histogram equalization for brightness preservation and contrast enhancement. IEEE Transactions on Consumer Electronics, 54(3), 1389-1397.

Krcmar, H. (2015). Informationsmanagement. In Informationsmanagement (pp. 85-111). Springer Gabler, Berlin, Heidelberg.

La, H. M., Gucunski, N., Kee, S. H., Yi, J., Senlet, T., & Nguyen, L. (2014, September). Autonomous robotic system for bridge deck data collection and analysis. In 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems (pp. 1950-1955). IEEE.

Lee S, Kalos N (2014) Non-destructive testing methods in the US for bridge inspection and maintenance. KSCE J Civ Eng 18(5):1322–1331. https://doi.org/10.1007/s12205-014-0633-9

Marziliano, P., Dufaux, F., Winkler, S., & Ebrahimi, T. (2002, September). A no-reference perceptual blur metric. In Proceedings. International conference on image processing (Vol. 3, pp. III-III). IEEE.

Maalouf, A., & Larabi, M. C. (2010, August). A no reference objective color image sharpness metric. In 2010 18th European Signal Processing Conference (pp. 1019-1022). IEEE.

Mittal, A., Moorthy, A. K., & Bovik, A. C. (2011, November). Blind/referenceless image spatial quality evaluator. In 2011 conference record of the forty fifth asilomar conference on signals, systems and computers (ASILOMAR) (pp. 723-727). IEEE.

Mittal, A., Soundararajan, R., & Bovik, A. C. (2012). Making a “completely blind” image quality analyzer. IEEE Signal processing letters, 20(3), 209-212.

Moorthy, A. K., & Bovik, A. C. (2009, February). Perceptually significant spatial pooling techniques for image quality assessment. In Human Vision and Electronic Imaging XIV (Vol. 7240, p. 724012). International Society for Optics and Photonics.

Moorthy, A. K., & Bovik, A. C. (2011). Blind image quality assessment: From natural scene statistics to perceptual quality. IEEE transactions on Image Processing, 20(12), 3350-3364.

North Dakota Section of the American Society of Civil Engineers Infrastructurereportcard.org/north-dakota 2019.Oh, T., Kee, S. H., Arndt, R. W., Popovics, J. S., & Zhu, J. (2013). Comparison of NDT methods for assessment of a concrete bridge deck. Journal of Engineering Mechanics, 139(3), 305-314.

Ooi, C. H., Kong, N. S. P., & Ibrahim, H. (2009). Bi-histogram equalization with a plateau limit for digital image enhancement. IEEE transactions on consumer electronics, 55(4), 2072-2080.

Ong, E., Lin, W., Lu, Z., Yao, S., Yang, X., & Jiang, L. (2003, July). No-reference JPEG-2000 image quality metric. In 2003 International Conference on Multimedia and Expo. ICME'03. Proceedings (Cat. No. 03TH8698) (Vol. 1, pp. I-545). IEEE.

Panetta, K., Gao, C., & Agaian, S. (2013). No reference color image contrast and quality measures. IEEE transactions on Consumer Electronics, 59(3), 643-651.

Panetta, K. A., Wharton, E. J., & Agaian, S. S. (2008). Human visual system-based image enhancement and logarithmic contrast measure. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 38(1), 174-188.

Pipino, L. L., Lee, Y. W., & Wang, R. Y. (2002). Data quality assessment. Communications of the ACM, 45(4), 211-218.

Pozzer, S., Rezazadeh Azar, E., Dalla Rosa, F., & Chamberlain Pravia, Z. M. (2021). Semantic Segmentation of Defects in Infrared Thermographic Images of Highly Damaged Concrete Structures. Journal of Performance of Constructed Facilities, 35(1), 04020131.

Qin, X., He, S., Zhang, Z., Dehghan, M., & Jagersand, M. (2018, March). Bylabel: A boundary based semi-automatic image annotation tool. In 2018 IEEE Winter Conference on Applications of Computer Vision (WACV) (pp. 1804-1813). IEEE.

Rakha, T., & Gorodetsky, A. (2018). Review of Unmanned Aerial System (UAS) applications in the built environment: Towards automated building inspection procedures using drones. Automation in Construction, 93, 252-264.

Rens KL, Wipf TJ, Klaiber FW (1997) Review of nondestructive evaluation techniques of civil infrastructure. J Perform Construct Facil 11(4):152–160. https://doi.org/10.1061/ (ASCE)0887-3828(1997)11:4(152).

Rhee, J. Y., Kee, S. H., Kim, H. S., & Choi, J. J. (2018). Seasonal Variation and Age-related Changes in the Relative Permittivity of Concrete Bridge Decks on Korea Expressways. International Journal of Concrete Structures and Materials, 12(1), 2.

Rhee, J. Y., Choi, J. J., & Kee, S. H. (2019). Evaluation of the depth of deteriorations in concrete bridge decks with asphalt overlays using air-coupled GPR: A case study from a pilot bridge on Korean expressway. International Journal of Concrete Structures and Materials, 13(1), 23.

Ribeiro, F., Florencio, D., & Nascimento, V. (2011, September). Crowdsourcing subjective image quality evaluation. In 2011 18th IEEE International Conference on Image Processing (pp. 3097-3100). IEEE.

Robert, M. (1982). Science behind thermography-thermal infrared sensing for diagnostics and control. J. of Thermosense, 371, 2-9.

Rolander D, Phares B, Graybeal B, Moore M, Washer G (2001) Highway bridge inspection: state-of-the-practice survey. J Transport Res Board 1749:73–81. https://doi.org/10.3141/ 1749-12.

Ryan TW, Hartle RA, Mann JE, Danovich LJ (2012) Bridge inspector's reference manual. FHWA NHI 03-001, FHWA, U.S. Department of Transportation. https://www.fhwa.dot.gov/ bridge/nbis/pubs/nhi12049.pdf.

Sahoo PK, Soltani S, Wong AKC. 1988. A survey of thresholding techniques. Comput. Vis. Graph. Image Proc. 41:233–60

Scott, M., Rezaizadeh, A., Delahaza, A., Santos, C. G., Moore, M., Graybeal, B., & Washer, G. (2003). A comparison of nondestructive evaluation methods for bridge deck assessment. ndt & E International, 36(4), 245-255.

Sezgin, M., & Sankur, B. (2004). Survey over image thresholding techniques and quantitative performance evaluation. Journal of Electronic imaging, 13(1), 146-165.

Shrivastava, S., Patel, D., Bhamidipaty, A., Gifford, W. M., Siegel, S. A., Ganapavarapu, V. S., & Kalagnanam, J. R. (2019, December). DQA: Scalable, Automated and Interactive Data Quality Advisor. In 2019 IEEE International Conference on Big Data (Big Data) (pp. 2913-2922). IEEE.

Strong, D. M., Lee, Y. W., & Wang, R. Y. (1997). Data quality in context. Communications of the ACM, 40(5), 103-110.

Tran, Q. H., Han, D., Kang, C., Haldar, A., & Huh, J. (2017). Effects of ambient temperature and relative humidity on subsurface defect detection in concrete structures by active thermal imaging. Sensors, 17(8), 1718.

Union, I. T. (2008). ITU-T Recommendation P. 910: Subjective video quality assessment methods for multimedia applications.

Vaghefi K, Oats RC, Harris DK, Ahlborn TTM, Brooks CN, Endsley KA, Dobson R (2011) Evaluation of commercially available remote sensors for highway bridge condition assessment. J Bridge Eng 17(6):886–895. https://doi.org/10.1061/ (ASCE)BE.1943-5592.0000303

Vaghefi K, Ahlborn TTM, Harris DK, Brooks CN (2015) Combined imaging technologies for concrete bridge deck condition assessment. J Perform Construct Facil 29(4):04014102. https://doi.org/10.1061/(ASCE)CF.1943-5509.0000465

Washer, G., Fenwick, R., Nelson, S., & Rumbayan, R. (2013). Guidelines for thermographic inspection of concrete bridge components in shaded conditions. Transportation research record, 2360(1), 13-20.

Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. Journal of Management Information Systems, 12(4), 5–33.

Wang, Y., Chen, Q., & Zhang, B. (1999). Image enhancement based on equal area dualistic sub-image histogram equalization method. IEEE Transactions on Consumer Electronics, 45(1), 68-75.

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing, 13(4), 600-612.

Wang, C., & Ye, Z. (2005). Brightness preserving histogram equalization with maximum entropy: a variational perspective. IEEE Transactions on Consumer Electronics, 51(4), 1326-1334.

Wang, T., Han, B., & Collomosse, J. (2014). Touchcut: Fast image and video segmentation using single-touch interaction. Computer Vision and Image Understanding, 120, 14-30.

Würthele, V. G. (2003). Datenqualitätsmetrik für Informationsprozesse: Datenqualitätsmanagement mittels ganzheitlicher Messung der Datenqualität (Doctoral dissertation, ETH Zurich).

Yang, M., & Sowmya, A. (2015). An underwater color image quality evaluation metric. IEEE Transactions on Image Processing, 24(12), 6062-6071.

Zhang, J. K., Yan, W., & Cui, D. M. (2016). Concrete condition assessment using impact-echo method and extreme learning machines. Sensors, 16(4), 447.

Appendix A

The folders contain field data and annotated dataset from five in-service bridges. NDE field work was carried out between July -to-August 2020.

Name of parent folder – NDDOT summer 2020 bridge evaluation dataset and report.

Name of Sub-folders- There are four (4) sub-folders with its contents described below:

- **1.Ground Truth**

This folder contains eleven (11): five (5) DGN cad files, five (5) DXF cad files, and one (1) pdf file format

- **2.Impact Echo**

This folder contains two (2) sub-folder: Impact Echo Field Data and Impact Echo Annotated Data.

a. **01-Impact Echo Field Data:** This contains five (5) folders with the compressed folders in Rar forma and a Ms Excel file showing a tabulated format of the field data.

The five (5) sub-folders contain the raw data of the IE points collected on site for each of the bridges. Each of the 5 sub-folders contains sub-folders which contains the regions of test points A, B, C, D as appropriate. These regions are plotted on the deck layout and shown in the report. Each of the test regions contains two (2) folders of either bad or good signals. The folder named 'bad' contain signals deemed to be of a low/bad quality based on the inspector's judgement while the 'good' folders are deemed to be of good quality. The zipped/rar folder is a back-up for the five raw IE data

b. **02-Impact Echo Annotated Data:** This folder contains a Ms. Excel workbook format with name *"NDDOT Bridge_IE_ANNOTATION".* The workbook contains Five (5) sheets for each of the bridges. The columns of these sheets are Bridge, Date, File Number, FileName_LVM, Origin_X_Offset_ft, Origin_Y_Offset_ft, Origin, Local_Y_ft, Data Quality (Inspector) and sub-surface delamination (The removal class) which could be 1(sound surface), 2 (shallow delamination), or 3 (deep delamination).

Park River Median Impact Echo Test										
Bridge	Date	FileNumber	FileName_LVM	Origin_X_ Offset_ft	Origin_Y_ Offset_ft	Origin	Local_X_ft	Local_Y_ft	Data Quality (Inspector)	Sub-surface Delamination/Removal Class
UND-PR_Median	7/6/2020	0	Metal.lvm	38	14	1A	0	0	BAD	1
UND-PR_Median	7/6/2020	1	Metal_1.lvm	38	14	1A	0	0	GOOD	1
UND-PR_Median	7/6/2020	2	Metal_2.lvm	38	14	1A	1	0	GOOD	1
UND-PR_Median	7/6/2020	3	Metal_3.lvm	38	14	1A	2	0	BAD	1
UND-PR_Median	7/6/2020	4	Metal_4.lvm	38	14	1A	2	0	BAD	1
UND-PR_Median	7/6/2020	5	Metal_5.lvm	38	14	1A	2	0	GOOD	1
UND-PR_Median	7/6/2020	6	Metal_6.lvm	38	14	1A	3	0	GOOD	1
UND-PR_Median	7/6/2020	7	Metal_7.lvm	38	14	1A	4	0	BAD	1
UND-PR_Median	7/6/2020	8	Metal_8.lvm	38	14	1A	4	0	GOOD	1
UND-PR_Median	7/6/2020	9	Metal_9.lvm	38	14	1A	5	0	BAD	1
UND-PR_Median	7/6/2020	10	Metal_10.lvm	38	14	1A	5	0	GOOD	1
UND-PR_Median	7/6/2020	11	Metal_11.lvm	38	14	1A	6	0	GOOD	1
UND-PR_Median	7/6/2020	12	Metal_12.lvm	38	14	1A	7	0	GOOD	1
UND-PR_Median	7/6/2020	13	Metal_13.lvm	38	14	1A	8	0	GOOD	1
UND-PR_Median	7/6/2020	14	Metal_14.lvm	38	14	1A	9	0	GOOD	1
UND-PR_Median	7/6/2020	15	Metal_15.lvm	38	14	1A	10	0	GOOD	1
UND-PR_Median	7/6/2020	16	Metal_16.lvm	38	14	1A	0	1	GOOD	1
UND-PR_Median	7/6/2020	17	Metal_17.lvm	38	14	1A	1	1	GOOD	1
UND-PR_Median	7/6/2020	18	Metal_18.lvm	38	14	1A	2	1	GOOD	1
UND-PR_Median	7/6/2020	19	Metal_19.lvm	38	14	1A	3	1	GOOD	1
UND-PR_Median	7/6/2020	20	Metal_20.lvm	38	14	1A	4	1	GOOD	1
UND-PR_Median	7/6/2020	21	Metal_21.lvm	38	14	1A	5	1	GOOD	1
UND-PR_Median	7/6/2020	22	Metal_22.lvm	38	14	1A	6	1	GOOD	1
UND-PR_Median	7/6/2020	23	Metal_23.lvm	38	14	1A	7	1	GOOD	1
UND-PR_Median	7/6/2020	24	Metal_24.lvm	38	14	1A	8	1	GOOD	1
UND-PR_Median	7/6/2020	25	Metal_25.lvm	38	14	1A	9	1	GOOD	1
UND-PR_Median	7/6/2020	26	Metal_26.lvm	38	14	1A	10	1	GOOD	1
UND-PR_Median	7/6/2020	27	Metal_27.lvm	38	14	1A	0	2	GOOD	1
UND-PR_Median	7/6/2020	28	Metal_28.lvm	38	14	1A	1	2	GOOD	1
UND-PR_Median	7/6/2020	29	Metal_29.lvm	38	14	1A	2	2	BAD	1
UND-PR_Median	7/6/2020	30	Metal_30.lvm	38	14	1A	2	2	GOOD	1
UND-PR_Median	7/6/2020	31	Metal_31.lvm	38	14	1A	3	2	GOOD	1

Park River Median_IE_Annotation | Forest River_NB_IE_Annotation | Park River_SB_IE_Annotation | Park River_NB_IE_Annotation | Forest River_SB_IE_Annotation

Figure A-1. IE Annotation workbook and annotation format

- **3.GPR data**

Contains three (3) sub-folders: GPR Field data, GPR coordinates, GPR Annotated Data.

a. **GPR Field data:** The folder contains two sub-folders: 1.GPR SIGNAL DATA and 2.GPR FIELD PHOTOS. The *1.GPR SIGNAL DATA* folder contains five sub-folder for each of the bridges, a compressed rar of all the raw data and an Ms Excel GPR field note. Each bridge sub-folder contains the scan field test data collected for the bridge each in Ms Excel CSV, DZT and DZX for each scan. The Ms Excel CSV can be assessed with Ms Excel package while the other are not. Each file contains columns several hundreds to thousands of amplitude columns and a single time (ns) column. The GPR field note contains information of the field scan for the bridges. The *2.GPR FIELD PHOTOS* contains field photos collected during the GPR test.

b. **GPR coordinates:** This folder contains five Ms. Excel files for each of the bridges, each file having the coordinates of the classes of removal generated for the purpose of annotation. It contains columns for point number, eastings and northings.

c. **GPR Annotated Data.**

This contains the five (5) folders for the annotated bridges and their corresponding CAD layout.

- **4.Infrared Thermography data.**

This contains five (5) sub-folders for the bridges. Each folder contains two (2) sub-folders namely: Original images and Annotated images. The annotated images are named with the bridge names as the suffi with unique identity numbers

Dataset Repository

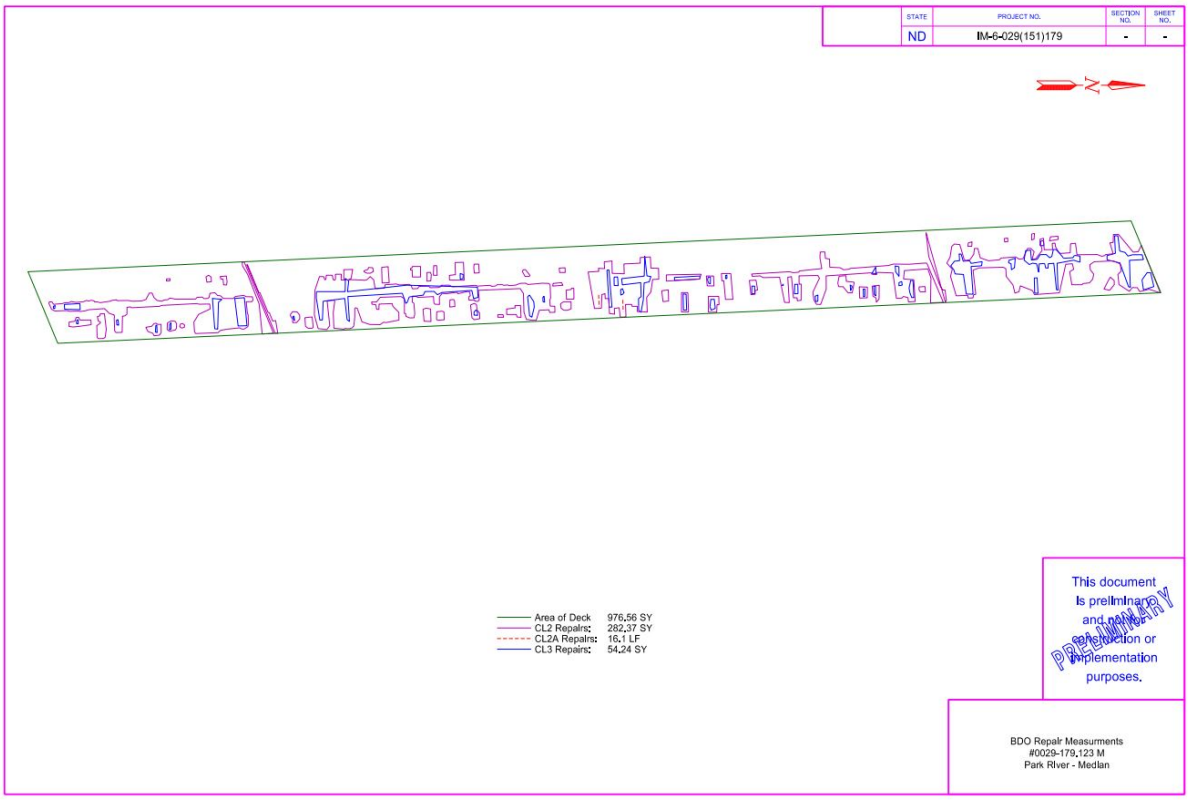
Bridge Deck NDE Annotated Dataset

Alternatively, readers are encouraged to use SDNET2021, an annotated NDE dataset for Structural Defects which can be found on UND's permanent repository:

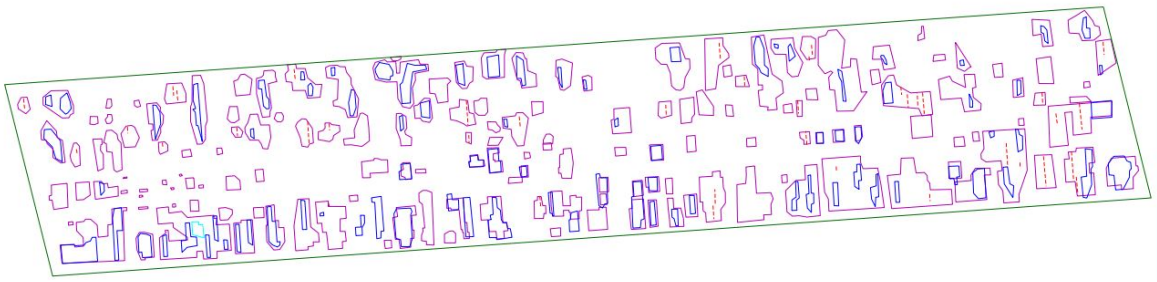
Ichi, Eberichi and Dorafshan, Sattar, "SDNET2021: Annotated NDE dataset for Structural Defects" (2021). *Datasets*. 19. https://commons.und.edu/data/19. https://doi.org/10.31356/data019.

Appendix B

Delamination Survey CAD layout



STATE	PROJECT NO.	SECTION NO.	SHEET NO.
ND	IM-6-029(148)168	-	-



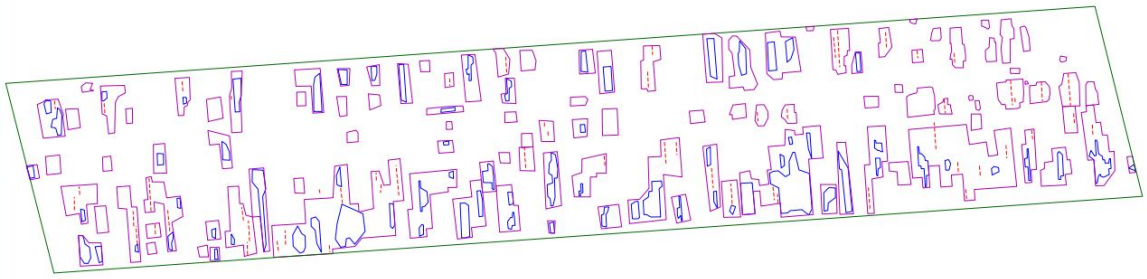
	Outside Lane	Inside Lane	Total
Area of Deck			860,54 SY
CL2 Repairs:	106,81 SY	110,51 SY	217,32 SY
CL2A Repairs:	39.0 LF	54.1 LF	93.1 LF
CL3 Repairs:	37,56 SY	15,90 SY	53,46 SY

This document is preliminary and not for construction or implementation purposes.

PRELIMINARY

BDO Repair Measurments
#0029-168,632 R
Forest River - NB

STATE	PROJECT NO.	SECTION NO.	SHEET NO.
ND	IM-6-029(148)168	-	-



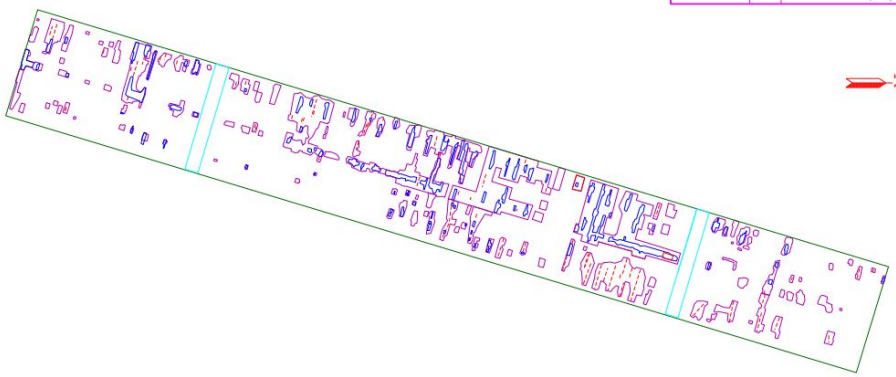
	Outside Lane	Inside Lane	Total
Area of Deck			861,56 SY
CL2 Repairs:	139,58 SY	164,19 SY	303,77 SY
CL2A Repairs:	57.5 LF	94.8 LF	152.3 LF
CL3 Repairs:	11,59 SY	32,59 SY	44,18 SY

This document is preliminary and not for construction or implementation purposes.

PRELIMINARY

BDO Repair Measurments
#0029-168,629L
Forest River - SB

STATE	PROJECT NO.	SECTION NO.	SHEET NO.
ND	IM-6-029(151)179	-	-



	Outside Lane	Inside Lane	Total
Area of Deck			2149.25 SY
CL2 Repairs:	105.96 SY	378.98 SY	484.94 SY
CL2A Repairs:	173.1 LF	112.4 LF	285.5 LF
CL3 Repairs:	5.33 SY	70.61 SY	75.94 SY

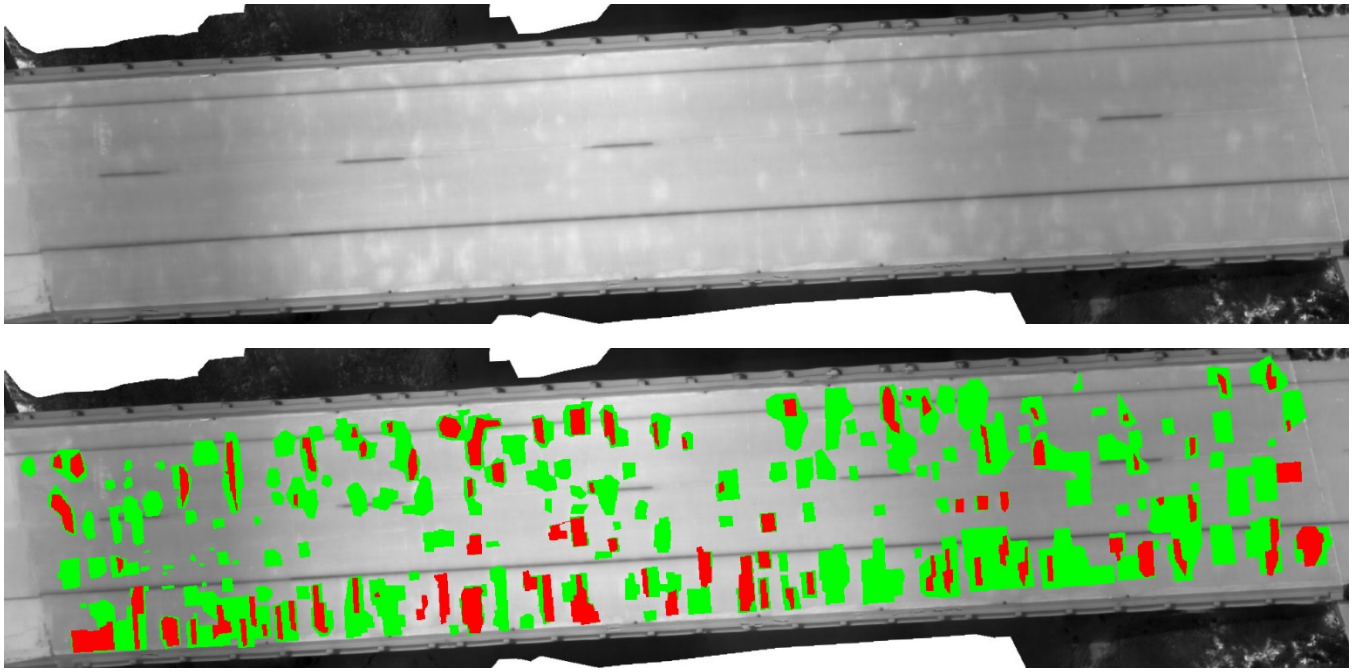
This document is preliminary and not for construction or implementation purposes.

PRELIMINARY

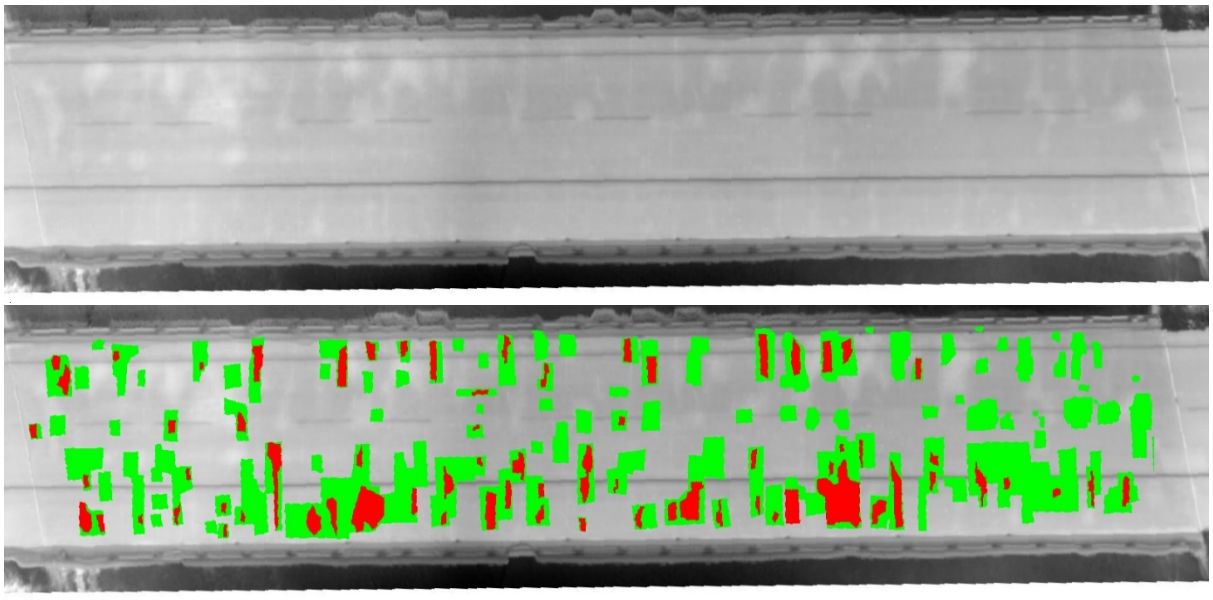
BDO Measurements
#0029-179,147 R
Park River - NB

Appendix C

Mosaic and annotated images



River North Bound



Forest River South Bound

Appendix D

Matlab code for Annotation

```
close all

clear all

clc


%1. Image Alignment..


cb = imread('FR_NB_Stitched.jpg');

cb_ref = imref2d(size(cb));


background = zeros(1);


T = [1 0 0;0 1 0;0 0 1];

tform_t = affine2d(T);


R = [cosd(-1.9) sind(-1.9) 0;-sind(-1.9) cosd(-1.9) 0;0 0 1];

tform_r = affine2d(R);


S = [1 0 0;0 1 0;0 0 1];

tform_s = affine2d(T);


% %Translation Followed by Rotation

% TR = T*R;

% tform_tr = affine2d(TR);

% [out_tr,out_tr_ref] = imwarp(cb,cb_ref,tform_tr);
```

```
% figure
% imshowpair(out_tr,out_tr_ref,background,imref2d(size(background)))

%Rotation Followed by Translation
RT = R*T;
tform_rt = affine2d(RT);
[out_rt,out_rt_ref] = imwarp(cb,cb_ref,tform_rt);
out_rt=imtranslate(out_rt,[0, -105]);
figure
imshowpair(out_rt,out_rt_ref,background,imref2d(size(background)))

%Filled region props for class 2 removal image beofre scaling up for effective filling
I_gray2=rgb2gray(imread('class 2A.jpg'));
bw2=im2bw(I_gray2);
bw22=imcomplement(bw2);
bw22=bwmorph(bw22,'bridge');
cc2=bwconncomp(bw22);

GT_C2=imfill(bw22,'holes');
imwrite(GT_C2, 'GT_C2.jpg');

%Scale Followed by Translation (Class 2 removal)
GT_C2A=imresize(GT_C2, 1.33);
GT_C2=imtranslate(GT_C2A, [-200, 105]);

figure
```

```
imshowpair(GT_C2,out_rt)

figure, imshow(GT_C2);

impixelinfo;

rgn2=regionprops(GT_C2,'PixelList');




%Scale Followed by Translation (Class 3 removal)

GT=imread('class 3A.jpg');

GT_C3A=imresize(GT, 1.33);

GT_C3=imtranslate(GT_C3A, [-200, 105]);


figure

imshowpair(GT_C3,out_rt)



StichedM_PRSB_2='2.Stitched_final1_PR_SB.jpg'

SM= out_rt;

GroundT_C2='class 2A.jpg'

GroundT_C3='class 3A.jpg'


figure, Aligned_2=imshowpair(GT_C2,SM);

figure, Aligned_3=imshowpair(GT_C3,SM);



%2. Class 3 Image Connected components..
```

```
I3=GT_C3;

I_gray3=rgb2gray(I3);


bw3=im2bw(I_gray3);


bwj3=imcomplement(bw3);


cc3=bwconncomp(bwj3);


rgn3=regionprops(bwj3,'PixelList');


lbl3=bwlabel(bwj3);


bw3=imfill(bwj3,'holes');


% 2 regions were not filled therefore use morph%%


bw33=bwmorph(bwj3,'bridge');


bw_3=imfill(bw33,'holes');


rgn3=regionprops(bw_3,'PixelList');
```

```
R=(SM(:,:,1));

G=(SM(:,:,2));

B=(SM(:,:,3));


for k = 1:numel(rgn2)
    data2 = rgn2(k).PixelList;
    m=size(data2,1);
    for j=1:m
        inx=data2(j,:);
        B(inx(2),inx(1))=0;
        G(inx(2),inx(1))=255;
        R(inx(2),inx(1))=0;
    end


end


newSM(:,:,1)=R;

newSM(:,:,2)=G;

newSM(:,:,3)=B;


R2=newSM(:,:,1);

G2=newSM(:,:,2);

B2=newSM(:,:,3);


for n = 1:numel(rgn3)
```

```
    data3 = rgn3(n).PixelList;
    k=size(data3,1);
    for i=1:k
        iny=data3(i,:);
        R2(iny(2),iny(1))=255;
        G2(iny(2),iny(1))=0;
        B2(iny(2),iny(1))=0;
    end
end

new_SM(:,:,1)=R2;
new_SM(:,:,2)=G2;
new_SM(:,:,3)=B2;

figure
imshow(new_SM)
impixelinfo;



theta = 0;
distorted = imrotate(new_SM,theta); % Try varying the angle, theta.
distorted_1 = imrotate(SM,theta);
figure, imshow(distorted)
figure, imshow(distorted_1)
impixelinfo;
```

```
final_SM=distorted(180:2140, 235:680,:);

SM_1=distorted_1(180:2140, 235:680,:);


theta1 = 85;

final_SM= imrotate(final_SM,theta1);

SM_1= imrotate(SM_1,theta1);

figure, imshow(final_SM)

figure, imshow(SM_1)

imwrite(final_SM, 'final_SM.jpg');

imwrite(SM_1, 'SM_1.jpg');
```