2019

Enhanced Concrete Bridge Assessment Using Artificial Intelligence and Mixed Reality

Enes Karaaslan
University of Central Florida

ENHANCED CONCRETE BRIDGE ASSESSMENT USING ARTIFICIAL INTELLIGENCE AND MIXED REALITY

by

ENES KARAASLAN
B.S. Bogazici University, 2012
M.S. Middle East Technical University, 2015

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Civil, Environmental, and Construction Engineering
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Fall Term
2019

Major Professor: F. Necati Catbas

ABSTRACT

Conventional methods for visual assessment of civil infrastructures have certain limitations, such as subjectivity of the collected data, long inspection time, and high cost of labor. Although some new technologies (i.e. robotic techniques) that are currently in practice can collect objective, quantified data, the inspector's own expertise is still critical in many instances since these technologies are not designed to work interactively with human inspector. This study aims to create a smart, human-centered method that offers significant contributions to infrastructure inspection, maintenance, management practice, and safety for the bridge owners. By developing a smart Mixed Reality (MR) framework, which can be integrated into a wearable holographic headset device, a bridge inspector, for example, can automatically analyze a certain defect such as a crack that he or she sees on an element, display its dimension information in real-time along with the condition state. Such systems can potentially decrease the time and cost of infrastructure inspections by accelerating essential tasks of the inspector such as defect measurement, condition assessment and data processing to management systems. The human centered artificial intelligence (AI) will help the inspector collect more quantified and objective data while incorporating inspector's professional judgment. This study explains in detail the described system and related methodologies of implementing attention guided semi-supervised deep learning into mixed reality technology, which interacts with the human inspector during assessment. Thereby, the inspector and the AI will collaborate/communicate for improved visual inspection.

Dedicated to;

My mother, Fatma Karaaslan
My father, Ramazan Karaaslan
My brother, Burak Karaaslan
And my beloved friends who supported, encouraged, and inspired me

ACKNOWLEDGMENTS

I would like to express my profound appreciation to my committee chair and advisor Professor F. Necati Catbas, who has given me the opportunity to be a part of innovative research projects, enlightened me with entrepreneurial vision and taught me the ways of being a proactive, conscientious, and hardworking person. He continually and convincingly conveyed a spirit of adventure in regard to novel research and proposal for funding, and technology transfer. Without his guidance and persistent help, this dissertation would not have been possible.

To my committee members, Dr. Ulas Bagci, Dr. Hae-bum Yun and Dr. Joseph LaViola, for their time and valuable feedback; Dr. Ozan Celik, Dr. Shuhei Hiasa, Mr. Chuan-Zhi Dong and Mr. Kevin Pfeil for their help and insight in exchanging ideas and conducting experiments; our former and current research team members of Civil Infrastructure Technologies for Resilience and Safety (CITRS) along with all the undergraduate and graduate researchers who took part in the creation of this work.

I also would like to acknowledge the financial support for this research that is provided by UCF College of Graduate Studies and NexcoWest USA.

And last but most certainly not the least, I would like to express my deepest love and gratitude to my family and my great friends for their timeless and priceless support and encouragement through many years of personal, professional, and academic challenges.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE: INTRODUCTION

Federal Highway Administration provides annual statistics on structurally deficient bridges. According to 2017 statistics, 54,560 bridges are structurally deficient among total number of 54,000 bridges [1]. Utilizing novel technologies for better management of such aged and deteriorated civil infrastructures is becoming more critical. To prevent the impending degradation of civil infrastructure, utilizing novel technologies for periodic inspection and assessment for long term monitoring is the most promising solution [2]. Although the inclination to use conventional inspection methods still persists, advanced sensing technologies have the ability to better understand the current condition with more resolution and accuracy [3]. Conventional methods for visual assessment of infrastructures have certain limitations, such as subjectivity of the collected data, long inspection time, and high cost of labor. On the other hand, imaging technologies allow collecting quantified data and performing objective condition assessment. These techniques are now receiving a breakthrough improvement with the employment of the state-of-the-art Artificial Intelligence (AI) models. Instead of post-processing of the collected inspection data, an AI system can detect the damages in real-time and analyze for condition assessment at a reasonable accuracy. The main objective of the AI integrated Mixed Reality (MR) system introduced in this dissertation study is to assist the inspector by accelerating certain routine tasks such as measuring all cracks in a defect region or calculating area of spalling. In this system, the human-centered AI interacts with the inspector instead of completely replacing the human involvement during the inspection. This collective work will lead to quantified assessment, reduced labor time while also ensuring human verified results.

For a complete methodology in civil infrastructure management, generating optimal decision strategy for rehabilitation, repair, and maintenance is essential. The inspection data collected using the MR system is processed in a novel decision support system that can effectively utilize the data to generate multi-objective decision alternatives. For effective use of the limited capital for maintenance and repair, the bridges are prioritized based on a number of factors from structural condition to importance of bridge to the connected transportation network. It is also important to take into account the entire network of bridges rather making decisions based on the individual ranking scores. The overall objective of the inspection system that integrates AI assistance, MR device and data integration to decision-support is illustrated in Figure 1.

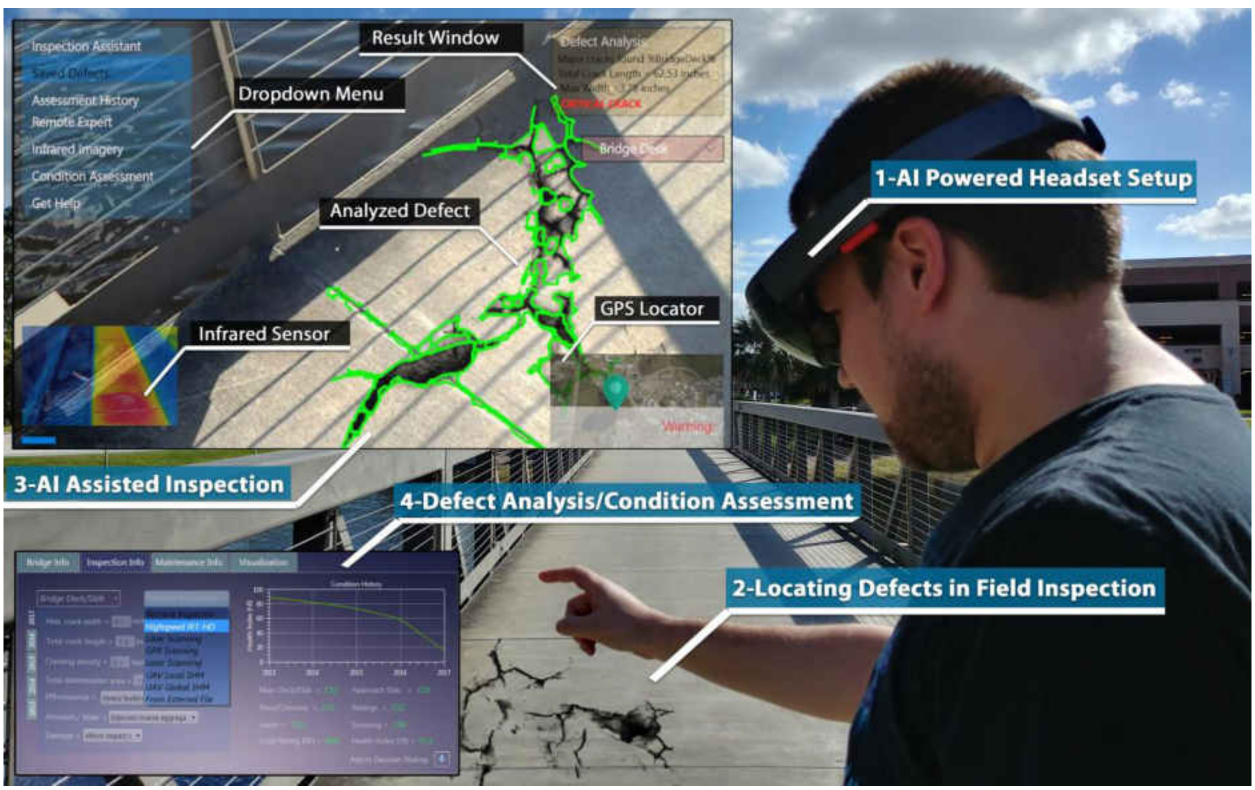


Figure 1: Overall objectives of the described system

The scope of this dissertation study involved a novel deep learning based method to quantify the concrete spalling and cracks on the bridge elements. The deep learning model was also integrated in mixed reality system to enable a new type of inspection methodology that benefits from human-AI collaboration instead of robotic data collection. A holographic wearable device with described methodology will assist bridge inspectors to collect and analyze data in real-time. As inspector make adjustments in the analyzed data, the system sends a semi-supervised training data to fine-tune the deep learning model. The model is periodically updated in the cloud and the updated model weights are sent back to the headset. Integration of the collected data that was uploaded to the server into the bridge decision support was also explored in this dissertation study. The decision support system analyzes all types of image-based non-destructive evaluation (NDE) data and performs time-history predictions based on the condition of the concrete defects. The system uses another deep-learning model to make the time-dependent predictions. Lastly, a novel decision ranking methodology was explored in this study; a fundamental work was completed; yet detailed results evaluating the system performance will be achieved in a future study. Similarly, the maintenance strategy generation was also investigated in the scope of the dissertation; yet more focused research will be conducted in the continuation of this work. The objective and the scope of the dissertation was summarized in Figure 2.

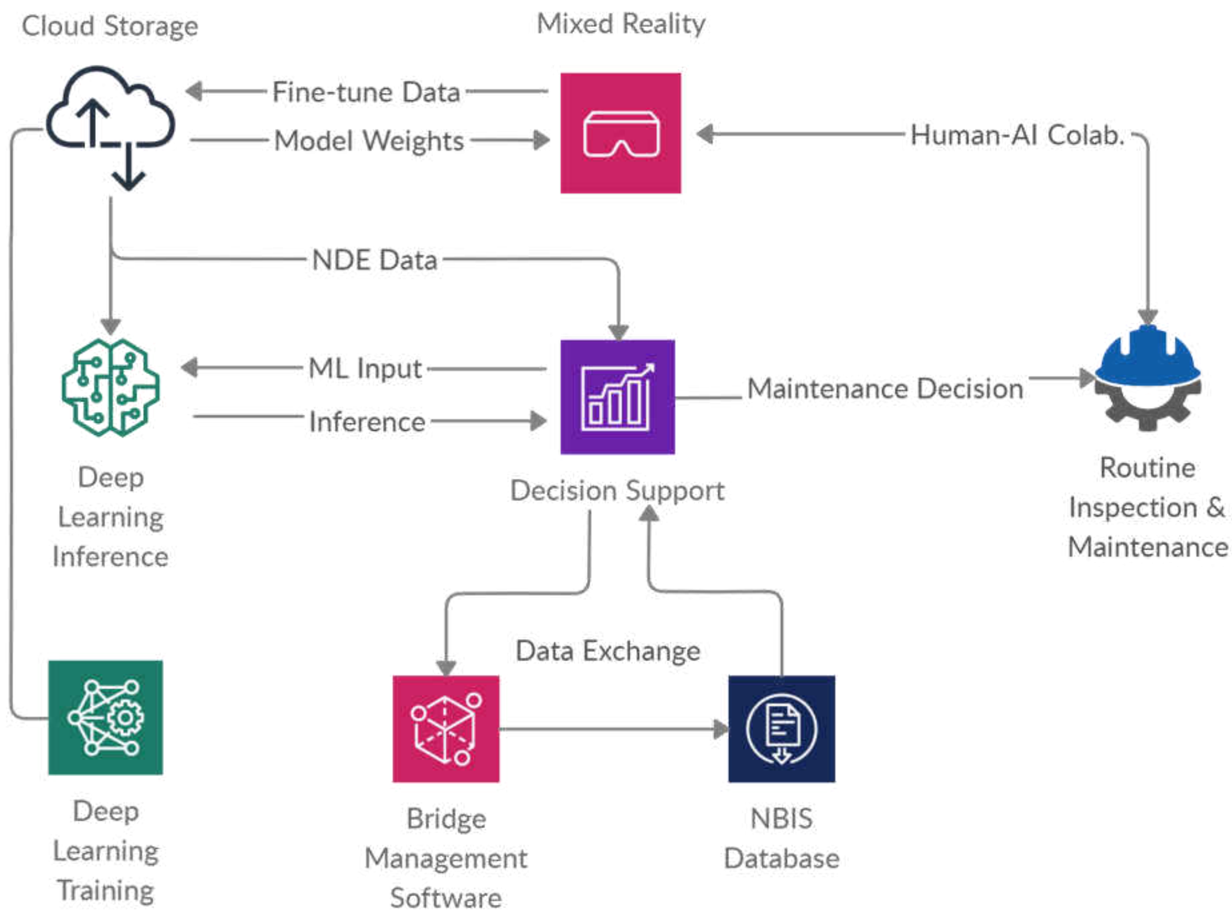


Figure 2: Objective and scope of the dissertation

To accomplish the objectives of the dissertation research, each research component was studied separately; yet the connections between the components were maintained at every step to create a complete methodology for bridge health monitoring. First step was to explore current bridge inspection practice that uses both traditional and novel technologies. A detailed literature review was carried out, and several field visits were made to observe periodic inspections on-site. In addition, meetings with decision makers were arranged to collect preliminary data on how different bridge parameter can affect maintenance/repair decisions. In the second step, structural health monitoring in general concepts was studied. Novel vision-based technologies including infrared and other camera-based methods were explored; effective utilization of collected data in the decision-making was investigated. In the third step, state-off-the-art methods were included in the analysis of data. Deep learning algorithms to analyze the collected vision-based data were investigated. However, the prior research on current inspection practice showed that the full automation in processing the bridge inspection data was not desirable, even with the most advanced algorithms, human involvement in the process was still preferred. Hence, the last step of the dissertation focused on a unique human-centered AI approach using mixed reality system to enable a collaboration structure between AI and the inspector.

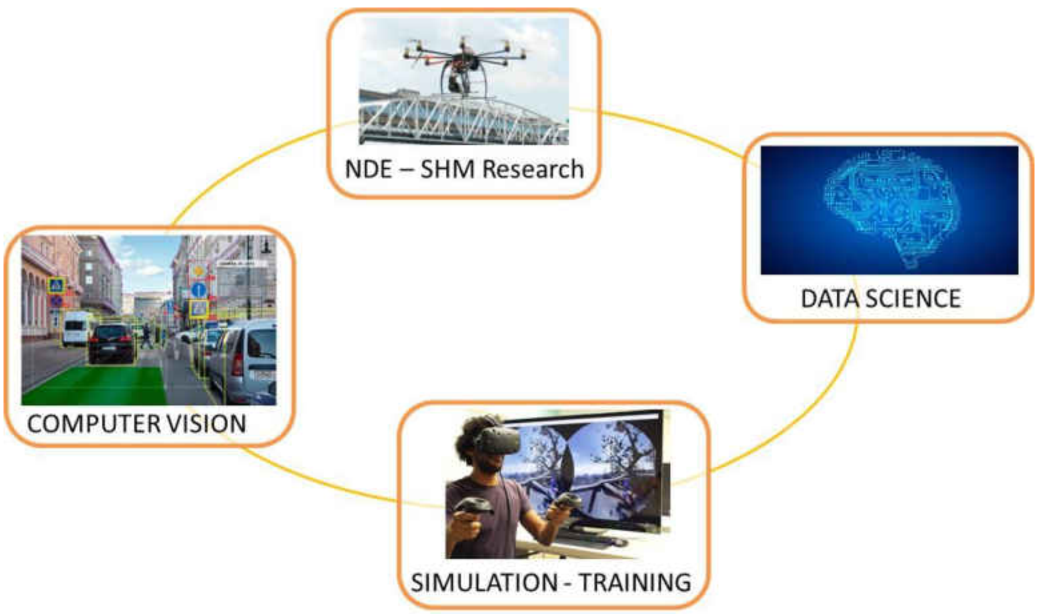


Figure 3: Interdisciplinary collaboration of the dissertation research

As shown in Figure 3, this dissertation study provides a comprehensive scope of multidisciplinary work including real-time deep learning, mixed reality system, and structural health monitoring. The interdisciplinary research yielded substantial scientific contribution with many novelties and innovations:

- This study is the first approach to integrate artificial intelligence with mixed reality system in civil engineering.

- First application of real-time artificial intelligence - human communication in civil engineering.
- A functional and trustable artificial intelligence system for a high-risk application such as bridge inspections.
- A true attention guide mechanism for deep learning since a human inspector provides the attention to the system.
- Practical integration collected inspection data and effective utilization in bridge decision-making.
- Adaptive ranking methodology for bridge prioritization using a novel deep learning approach.

Virtual, Augmented and Mixed Reality

The arrival of Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR) technologies is shaping a new environment where physical and virtual objects are integrated at different levels. Due to the development of mobile and embedded devices, together with interactive physical-virtual connections, the customer experience landscape is evolving into new types of hybrid experiences. However, the boundaries between these new realities, technologies and experiences have not yet been clearly established by researchers and practitioners. A brief discussion is given here to define these technologies illustrated in Figure 4.

Figure 4: Different of types of immersive technologies

Virtual Reality (VR) is a computer-simulated reality that replicates a physical environment or imaginary world through an immersive technology. VR replaces the user's physical world with a completely virtual environment and isolates the user's sensory receptors (eyes and ears) from the real world [4]. The VR is observed through a system that displays the objects and allows interaction, thus creating virtual presence [5]. Nowadays, VR headsets have gained vast popularity especially in gaming industry. The Augmented Reality (AR), on the other hand, is an integrated technique that often leverages image processing, real-time computing, motion tracking, pattern recognition, image projection and feature extraction. It overlays computer generated content onto the real world. An AR system combines real and virtual objects in a real environment by registering virtual objects to the real objects interactively in real time [6].

The beginning of AR dates back to Ivan Sutherland's see-through head-mounted display to view 3D virtual objects [7]. The initial prototype was only able to render few small line objects. Yet, the AR research has recently gained dramatic increase and now it is possible to visualize very complex virtual objects in the augmented environment. The recent developments of AR/VR technology helped companies produce holographic headsets that benefits Mixed Reality (MR) technology, in which one can experience hybrid reality where physical and digital objects co-exist and interact in real time. The term Mixed Reality was originally introduced in a 1994 paper "A Taxonomy of Mixed Reality Visual Displays" [8]. In the paper, a Virtuality Continuum (VC), in other words, mixed reality spectrum was explained in detail. A schematic representation is shown in Figure 5.

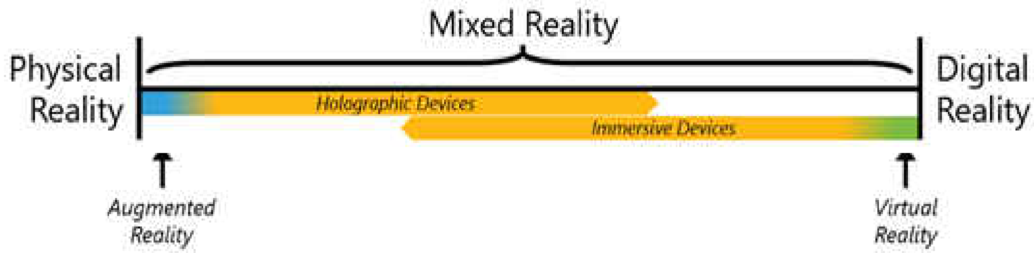


Figure 5 Mixed reality spectrum and device technologies [9]

The MR technology has breakthrough applications especially with successful deployment of 3D user interfaces such as in computer-aided design, radiation therapy, surgical simulation and data visualization [10]. The next generation of computer games, mobile devices, and desktop applications also will feature 3D interaction [11]. There are also some other efforts for using MR technology in construction industry and maintenance operations. Kamat and El-Tawil (2007) discusses the feasibility of using AR to evaluate earthquake-induced building damage. Behzadan and Kamat (2007) investigated the application of the global positioning system and 3 degree-of-freedom (3-DOF) angular tracking to address the registration problem during interactive visualization of construction graphics in outdoor AR environments [13]. The vision-based mobile AR systems are vastly used in 3D reconstruction of scenes for architectural, engineering, construction and facility management applications. Bae et al. (2013) developed a context-aware AR system that generates 3D reconstruction from 3D point cloud. Important effort for use of AR in infrastructure inspection is also shown by several researchers [14]. Researchers in University of Cambridge currently collaborate with Microsoft to develop an effective bridge inspection practice in which the data collected from the field is visualized in MR environment in the office [15]. Moreu et al. (2017) developed a conceptual design for novel structural inspection tools for structural inspection applications based on HoloLens [17] device [16]. The experiments conducted with the HoloLens for taking measurements and benchmarking the obtained measurements are shown in the study. The proposed methodology takes even a further step and combines AI implementation with MR technology. In this system, the embedded AI architecture not only predicts the location/region of cracks and spalling on the infrastructure in real-time along with condition information but also augments the information in the holographic headset for improved human inspector - AI interaction.

Overview of Deep Learning Approaches in Damage Detection and Analysis

For more than a decade, researchers have been investigating employing the techniques in the Computer Vision field to analyze cracks, spalls and other types of damages. The early approaches mostly used edge detection, segmentation and morphology operations. Yet, the recent advances in AI yielded very promising accuracy and possessed a wide range of applicability. A review paper on computer vision based defect detection and condition assessment of concrete infrastructures emphasizes the importance of sufficiently large, publically available and standardized datasets to leverage the power of existing supervised machine learning methods for damage detection [18]. According to the study, learning based methods can be reliably used for defect assessment. For the processing of defect images, many researchers in the literature implemented Convolutional Neural Network (CNN) to perform automatic crack detection on concrete surfaces. Combined with transfer learning and data augmentation, CNN can offer highly accurate input for structural assessment. Yokoyama and Matsumoto (2017) developed a CNN based crack detector with 2000 training images [19]. The main challenge of the detector was that the system often classifies stains as cracks. Yet, the detection is successful for even very minor cracks. Similarly, Jahanshahi and Masri (2012) developed a crack detection algorithm that however uses an adaptive method from 3D reconstructed scenes [20]. The algorithm extracts the whole crack from its background, where the regular edge detection based approaches just segment the crack edges; thereby offering a more feasible solution for crack thickness identification. Adhikari et al. (2014) used 3D visualization of crack density by projecting digital images and neural network models to predict crack depth, necessary information for condition assessment of concrete components [21].

For detection of spalls and cracks, German et al. (2012) used entropy-based thresholding algorithm in conjunction with image processing methods in template matching and morphological operations [22]. In addition to detection of local defects of structures, there are also studies on identifying global damages of the structures. Zaurin et al. (2015) performed motion tracking algorithms to measure the mid-span deflections of bridges under the live traffic load [23]. Computer Vision is also used to process ground penetration radar (GPR) and infrared thermography (IRT) images that are useful to identify delamination formed inside the concrete structures. Hiasa et al. (2016) processed the IRT images of bridge decks taken with high-speed vehicles [25]. In identifying damages, many different techniques are useful for specific purposes. However, a more generalized deep learning approach is introduced in this study so that the methods can be expanded toward identifying almost any type of damage if sufficient amount of training data is available.

The CNN models are mostly composed of convolutional and pooling layers. In the convolutional layers, the input images are multiplied by small distinct feature matrices that are attained from the input images (corners, edges etc.) and their summations are normalized by matrix size (i.e. kernel size). By convolving images, basically similarity scores between every region of the image and the distinct features are assigned. After convolution, the negative values of similarity in the image matrix are removed in the activation layer by using the rectified linear unit (ReLU) transformation operation. After the activation layer, the resultant image matrix is reduced to a very small size and added together to form a single vector in the pooling layer. This vector is then inserted in fully connected neural network where actual classification happens. The image vectors of the trained images are compared with the input image and a correspondence score is calculated for each classification label. The highest number will indicate the classified label. A summary of the described procedure is shown in Figure 6.

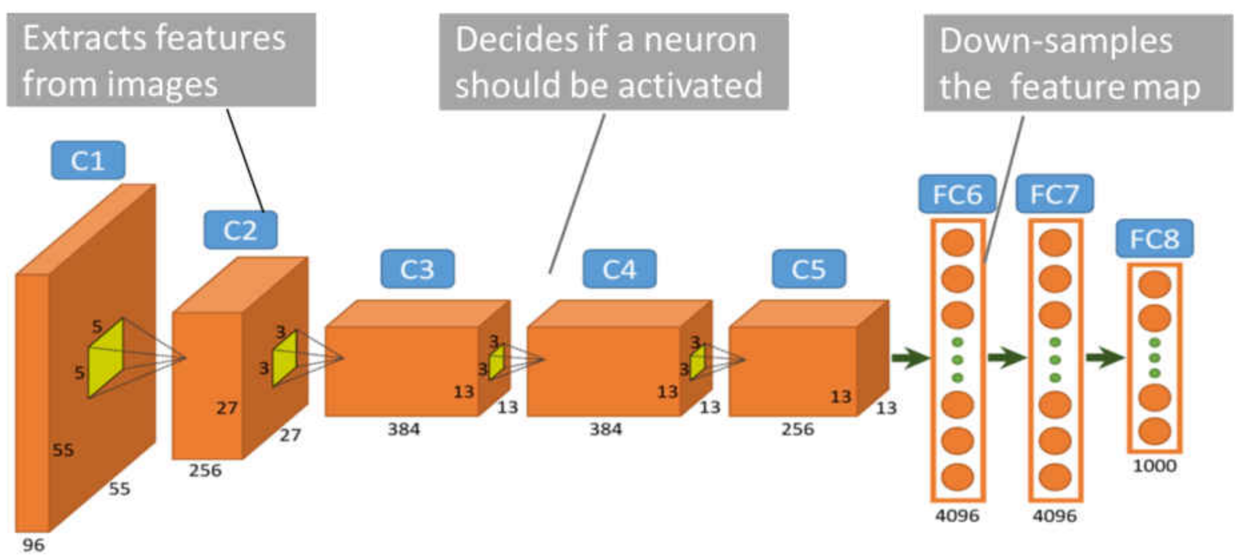


Figure 6: A commonly used Convolutional Neural Network – AlexNet [26]

Long-term Decision Making on Bridge Rehabilitation, Maintenance and Repair

Structural health monitoring (SHM) can be considered at local and global level monitoring. Local SHM (LSHM) evaluates serviceability of bridges by monitoring local level defects such as cracks, delamination, corrosion and roughness. Global SHM (GSHM), on the other hand, assesses soundness of bridges by measuring vibration, deflection and loading with respect to expected behavior or in comparison to its past performance. Generally, periodic LSHM data is employed for repair and maintenance work to recover the serviceability of bridges while GSHM is conducted to make decision for rehabilitation and replacement. Traditionally, SHM is conducted by means of visual, sound, and force-based methods or when they are in service. Both methods require direct access to bridges, causing extra field work and time as well as potential danger to inspectors. With the growing potential of camera-based methods, a complete non-contact SHM using NDE along with effective utilization in decision-making is possible.

The implementation of a proper infrastructure management has become crucial due to the fact that the US infrastructure has deteriorated significantly in the last decade. Advanced remediation strategies for deteriorated infrastructures are being developed using certain decision making models in order to maintain the optimal funding use and remediation time [27]. To prevent the impending degradation of bridges, utilizing novel technologies for periodic inspection, assessment and better management for proper maintenance has become more critical. Although the inclination to use conventional inspection methods still persists, advanced sensing technologies have the ability to better understand the current condition with more resolution and accuracy [28]. For this reason, better utilization of NDE as routine inspection practice becomes necessary. Optimized decision making based on the NDE input can be carried out by integrating utilization solutions to bridge management frameworks.

The infrastructure support framework in this study is designed to retrieve information from novel NDE techniques including vision-based technologies (e.g. infrared thermography, other imagery data) and perform network level decision analysis using both NBI's inputs and automatically retrieved inspection data form NDE. The system also performs condition prediction based on the historical data and retrieved NDE input. A schematic representation of NDE integration with bridge management is shown in Figure 7.

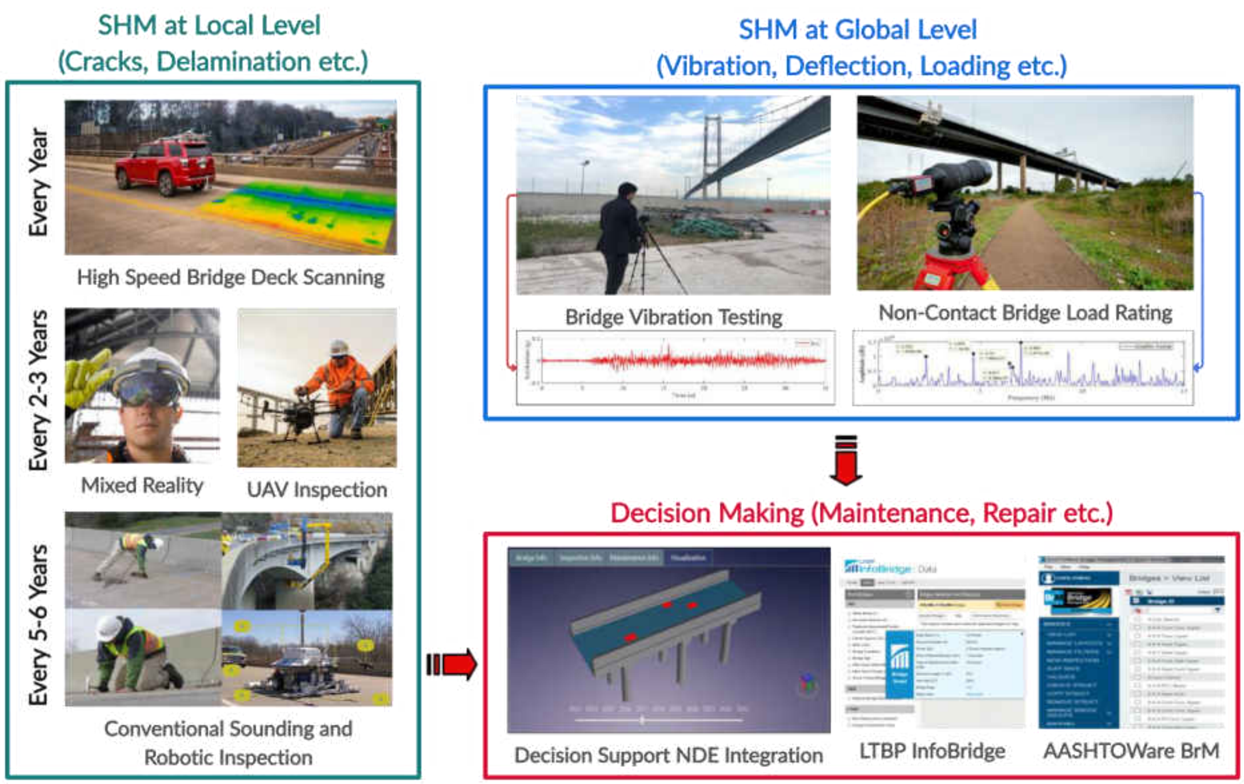


Figure 7: Complete SHM methodology using non-contact methods and advanced decision making

CHAPTER TWO: CURRENT PRACTICES IN BRIDGE INSPECTIONS

Deterioration of road infrastructure arises from aging and various other factors. The aging of bridges is one of the most critical factors for large number of underperforming bridges in the US. According to Federal Highway Administration (FHWA) and Federal Transit Administration (FTA) [29], the total number of bridges listed in the National Bridge Inventory (NBI) was 588,844 in 2000; approximately 67% of them were more than 25 years old, and 26% of them were over 50 years old. By 2015, the number increased to 611,845 bridges, and 72% of them were older than 25 years, and 38% were over 50 years old [30]. Thus, structural systems have aged to an extent that critical decisions such as repair or replacement should be made effectively. To prevent the impending degradation of these bridges, utilizing novel technologies for periodic inspection, assessment and better management for proper maintenance has become more critical. Therefore, innovative technologies and procedures are needed to allow infrastructure's owners to monitor their bridges more effectively and create optimal maintenance strategies. However, the progressive improvement on the aforementioned needs is slow, even though the existing status of the US civil infrastructure is well documented (e.g. ASCE report card [31]). One of the challenges in better managing of bridges is not only the use novel technologies but also to integrate these technologies into the current bridge inspection and management systems to utilize the data for optimal decision making. Sometimes the use of additional data, if not managed properly, may become burden to the State Departments of Transportation (DOTs), leading additional management and labor costs. When inspection data is utilized effectively, long-term decision can be made by monitoring the trend of bridge health. An example utilization of NDE data, described in Figure 8, was proposed by Catbas et. al [32].

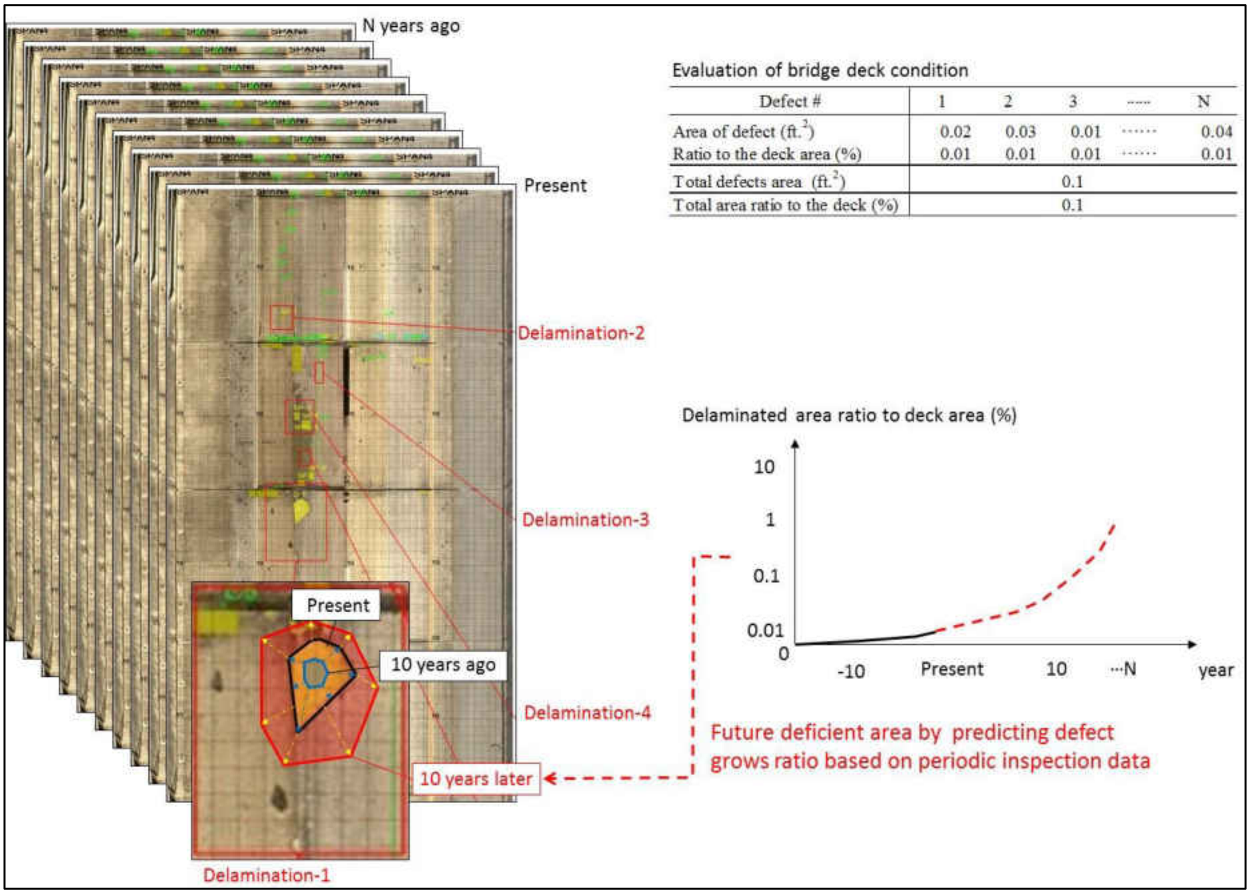


Figure 8: Example utilization of NDE for bridge management [33].

When Infrared Thermography (IRT) data from the deck surface is collected over a period of 10 years, it will provide very critical information on how a local delamination can potentially the impact the integrity of the overall structure. It will be possible to conduct time history prediction on the data to determine the optimal timeline of the necessary maintenance/repair actions. Hence, periodic NDE data can be utilized very effectively in infrastructure decision-making.

Conventional Inspection Techniques

By the U.S. federal regulations, all structures, which carry traffic and have a span greater than 20 feet are subject to comprehensive inspection at least every two years. The regulations define eight types of bridge inspection and three of them are periodic: routine inspection, fracture-critical member inspection, and underwater inspection. State DOTs, on the other hand, establish more detailed guidelines for periodic use of hands-on inspection, close-up access, and collection of quantitative data. State DOTs define guidelines for short-interval, interim inspections in response to bridge defects, conditions, or load posting. State DOTs also establish guidelines for long-interval, in-depth inspections for selected bridge types and bridge elements [34].

In a routine biennial inspection, trained bridge inspectors check for obvious damage, which may take such varied forms as spalled concrete, corroded steel, and even insect and fungus attack on timber elements; they also examine bearings, deck drains, and expansion joints for proper operation, evaluate the serviceability of bridge substructures, decks, approaches, and appurtenances, and, for waterway crossings, inspect the channel for scour and obstructions to flow. All of these elements are inspected visually, using basic hand tools where appropriate (Figure 9-a). If the side-deck or under-deck needs to be inspected, as show in Figure 9-b, a snooper truck is commonly used.

a-) Crack Measurement | b-) Side Inspection with Snooper Truck

Figure 9: Visual Inspection using basic hand tools (a) and snooper truck (b)

The goal of these federally mandated inspections is to assess and document the condition of essential bridge elements to ensure safety and serviceability and to facilitate the timely programming of maintenance and repairs. Some bridges are also subject to special inspections, in-depth evaluations of the safety and serviceability of particular elements known to have specific problems or present particular risks. For example, special inspections are conducted on fracture-critical bridges: those with non-redundant steel tension components, the fracture of which would likely cause catastrophic failure. Traditional slow and subjective methods used in assessing deck condition, which include impact sounding (Figure 10-b), chain dragging (Figure 10-a), half-cell potential, and core analysis, are being replaced by more modern techniques.

a-) Chain Drag	b-) Hammer Sounding

Figure 10: Traditional methods of assessing deck condition

Novel Non-Destructive Evaluation Methods

Conventional SHM methods have offered indispensable measures for condition assessment of civil infrastructure systems. The information derived from NDE methods served to the purposes such as damage detection and localization, assessment of overloading, effects of aging and eventually decision making. Most of these methods comprise the utilization of sensors such as accelerometers, strain gauges, displacement gauges, tiltmeters, etc. to measure the response over time and to provide high quality data for a sound evaluation [32]. Most of the time, either short term or long term, the instrumentation of the sensors requires long hours of labor, excessive amount of time and money. At times, structures might have locations that are hard to reach and instrument or they might need to be closed to operation, which may create chaos especially in highly populated areas. The aforementioned factors necessitate the use of novel technologies that could remedy the difficulties [35]. Integration of image processing and computer vision techniques in place of traditional sensing methodologies offers an efficient tool to monitor the structures under operational conditions as a complementary method to the load tests. Several studies prove the possible use and efficiency on the issues such as load quantification, dynamic displacement measurement and damage detection [36]. Figure 11 shows a computer-vision based method for dynamic structural identification [37].

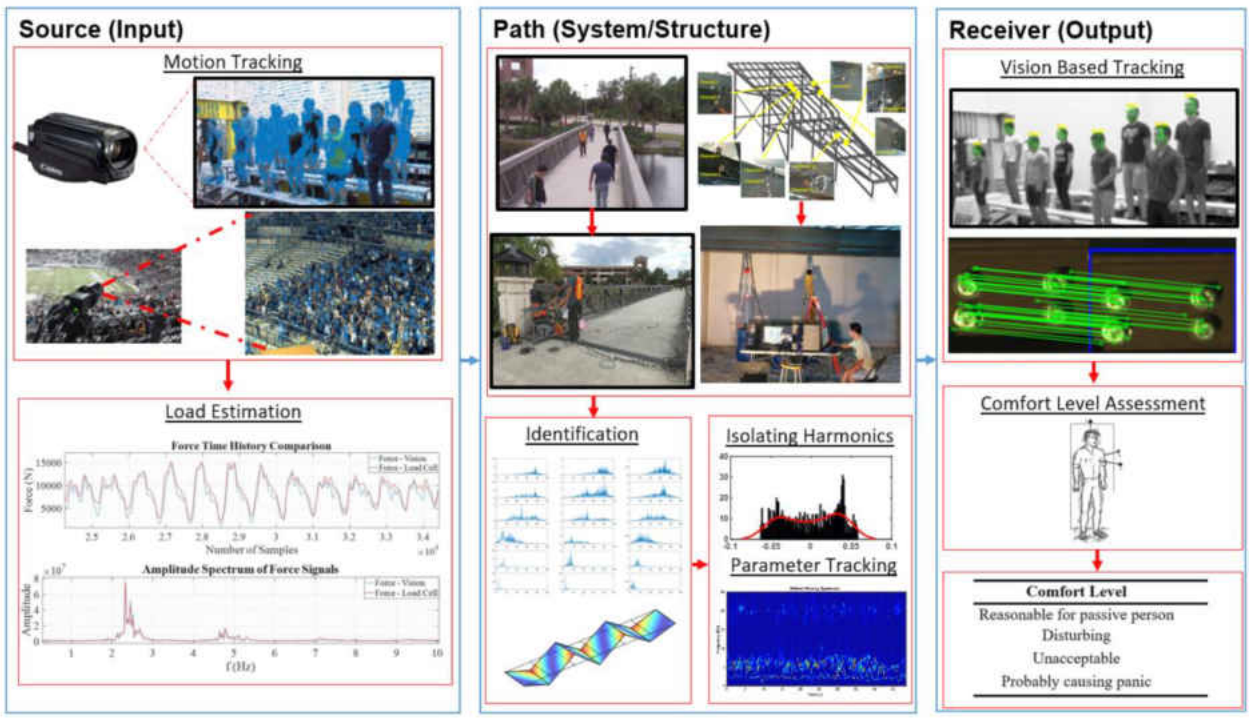


Figure 11: Dynamic load estimation, structural identification and human comfort assessment [37].

Concrete bridge decks deteriorate faster than other bridge components due to direct exposure to traffic along with environmental effects and others such as salting for deicing. Thus, the FHWA's LTBP Program regards them as the highest priority issues for bridge performance [38]. Since most DOTs spend 50-80 % of their expenditures on their bridges for repair, rehabilitation and replacement of concrete bridge decks, better methods are needed to detect defects and quantify the extent and severity of bridge deck conditions early, accurately, and rapidly with minimal traffic impact, ideally, without lane closures for bridge deck inspections

[38]. Under these circumstances, NDE techniques such as HD image-based crack detection, impact echo (IE), ultrasonic surface waves (USW), electrical resistivity (ER), ground-penetrating radar (GPR) and IRT have been developed to inspect and monitor aging and deteriorating structures rapidly and effectively in place of visual and sounding inspection methods. Some of these NDE techniques are shown in Figure 12.

Figure 12: Different NDE technologies used in bridge deck inspection

Since vision-based technologies are non-contact methods and both IRT and HD images can instantly portray a wide range of concrete structures at one time, the combination of IRT and HD technologies can be the fastest and easiest NDE methodology [39]. Specifically, the great potential of the combined method lies in removing direct access and contact as well as its capability for bridge deck scanning at normal driving speeds without lane closures as shown in Figure 13.

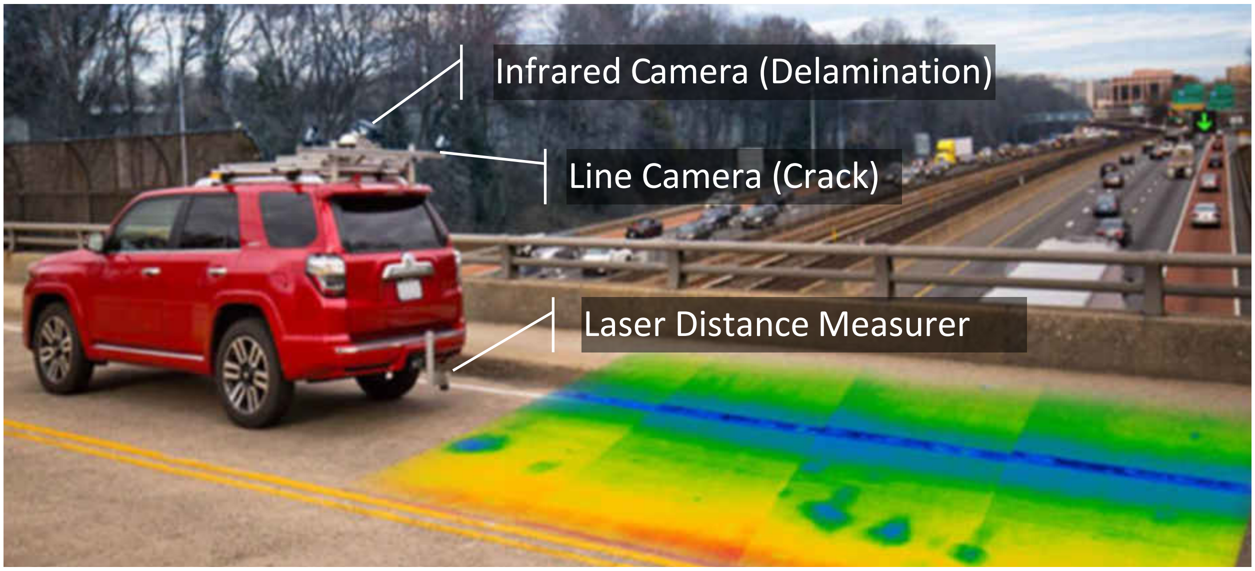


Figure 13: High-speed bridge deck scanning using IRT [40]

The high-speed deck scanning technologies are utilized with a vehicle driving at 50 mph (80 km/h), the data collection speed is 800 times faster than an integrated robotic system, which has been developed to conduct more efficient and effective bridge deck inspections than traditional methods with some NDE technologies. An average data collection time of 350 m2 of bridge deck area is one hour by the robotic system [24], while the combination of IRT and HD system can scan the same area of deck in 4.5 seconds. If preparation and traffic control time are considered, the productivity becomes much greater. Furthermore, as the number of bridges to be inspected increases, the productivity becomes at least 1,000 times greater than the other NDE methods. Therefore, less cost, less labor, better utilization of inspectors and eventually an increase of proper bridge management and maintenance of minor repairs are possible.

Long Term Bridge Management

The bridge management practice in the United States has improved significantly over the last 40 years both at the federal and state levels. At the federal level, the National Bridge

Inspection Standards (NBIS) unifies the method of collecting data and condition assessment on the public highway bridges [41]. The collected inspection data by the state departments of transportation (DOTs) is submitted to Federal Highway Administration (FHWA) annually in a nationwide reporting/coding format that is later entered to National Bridge Inventory (NBI) database [42]. Based on NBI, bridge owners are able to monitor condition and performance of their bridges to make accurate management decisions. FHWA imposes an appraisal rating to all government owned bridges through routine inspections that are recorded to NBI. Bridge appraisals are carried out by calculating scores in three categories: Structural adequacy and safety, serviceability and functional obsolescence, and essentiality for public use. After the scores in these categories are summed, special reductions are made. The resultant score will give the sufficiency rating that could be used for ranking bridges for infrastructure management. The sufficiency rating in NBI's bridge appraisals basically receives input from local and global assessments as well as some additional parameters. At the state level, state DOTs may have different procedures regarding bridge asset management, funding, maintenance considerations and resource allocation. A comprehensive National Cooperative Highway Research Program (NCHRP) synthesis report published by the Transportation Research Board (TRB) puts out the differences in state practices and explains the reasons of the variety in the bridge management practices mainly on the following issues: The differences in the policy, financial, technical and institutional operations as well as the different approaches to planning, programming and budgeting [43].

Today, all state DOTs have a bridge management process. Most employ some type of automated BMS with an associated database of bridge-related information, including NBIS data and ratings, but often incorporating detailed element-level inspection data. Agencies use economic methods to varying degrees in bridge management, but overall, the practices do not represent wide use. Common examples of applications to individual structures include the use of benefit-cost analysis for major bridge projects, and life cycle cost comparisons of rehabilitation versus replacement options for specific structures. Agencies that have full-featured BMSs are more likely to employ economic analyses in network level bridge management, but the practice is not yet widespread.

Development of new BMSs with more advanced decision support capabilities began in the United States in the 1980s. BMS designs and implementations were pursued independently by several DOTs, including North Carolina, Pennsylvania, Kansas, New York, Indiana, and Texas and Florida. The FHWA sponsored a demonstration project that led to the development of Pontis. Pontis is now further improved and renamed as Bridge Management (BrM) as an AASHTOWare product maintained as part of AASHTO's BRIDGEWare suite, and is used by more than 40 state DOTs plus other transportation agencies. AASHTOWare product breakdown is shown in Figure 14.

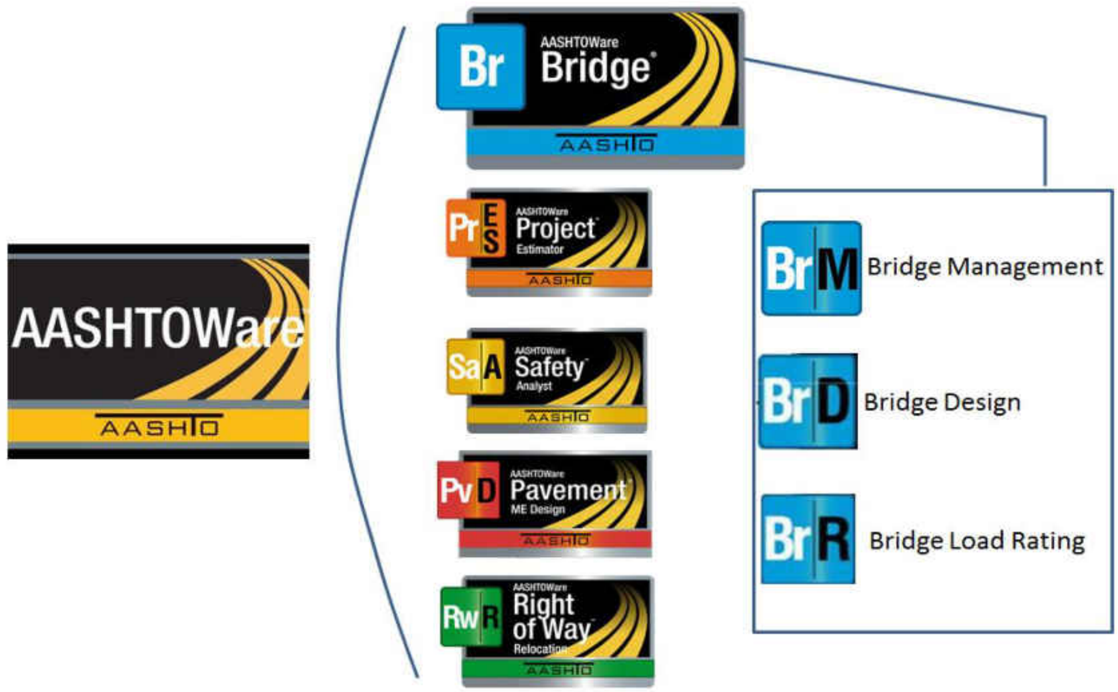


Figure 14: AASHTOWare product breakdown

Modern bridge management systems, including AASHTOWare BrM, have multi-objective performance frameworks for project evaluation, priority setting, and resource allocation. The objectives to be maximized, such as those presented in legislation and agency strategic plans, include safety, mobility, condition, and environmental sustainability. At the same time, agencies are continually called upon to minimize life-cycle costs (LCC) and manage risk. The graphical user interface of BrM is shown in Figure 15.

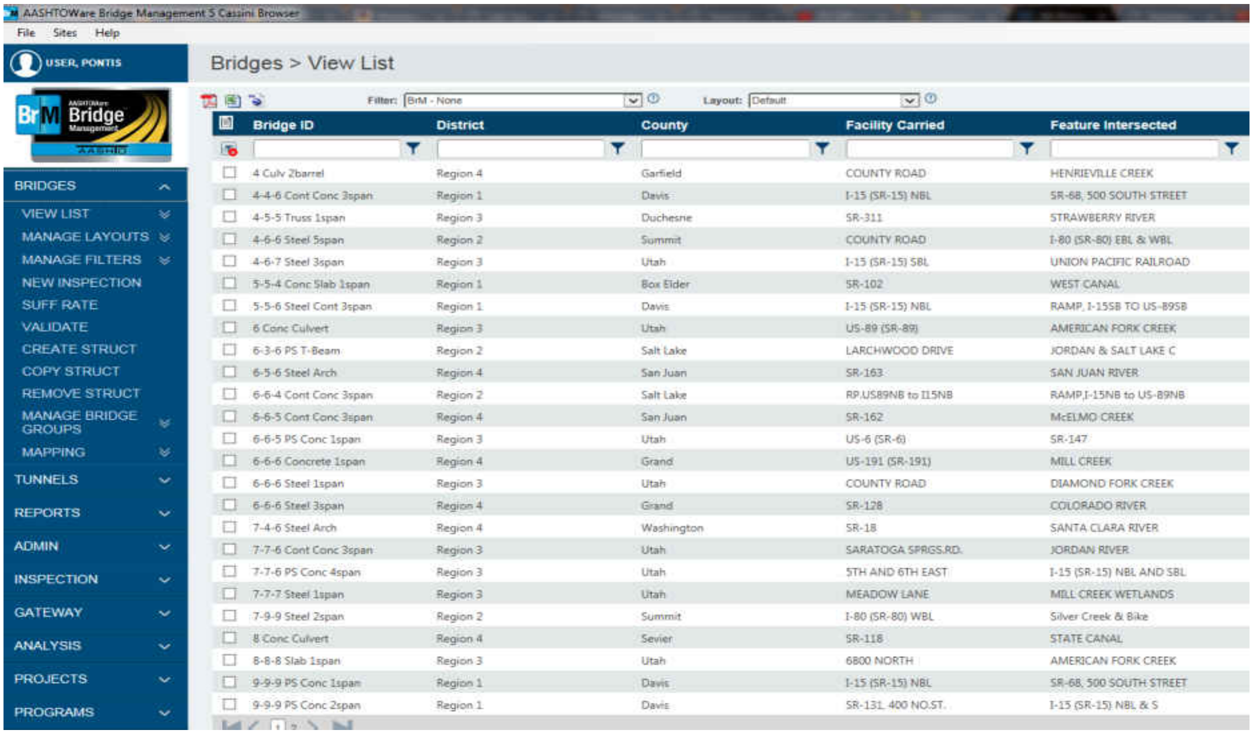


Figure 15: AASHTOWare BrM bridge search module

After Pontis version 4 series, the bridge management system was renamed as BrM with the version 5.1. The new version allows relatively flexible web-based approach with the inclusion of new AASHTO Elements including protective systems and defects. BrM mainly performs a multi objective tradeoff analysis using the following constraints: mobility, life cycle cost, condition and risk. An example of the analysis is given in Figure 16. In this example, bride condition is the most heavily weighted component meaning that it is more important than risk, mobility and life cycle. The agency is able to see how every component exactly impacts the overall utility of the asset.

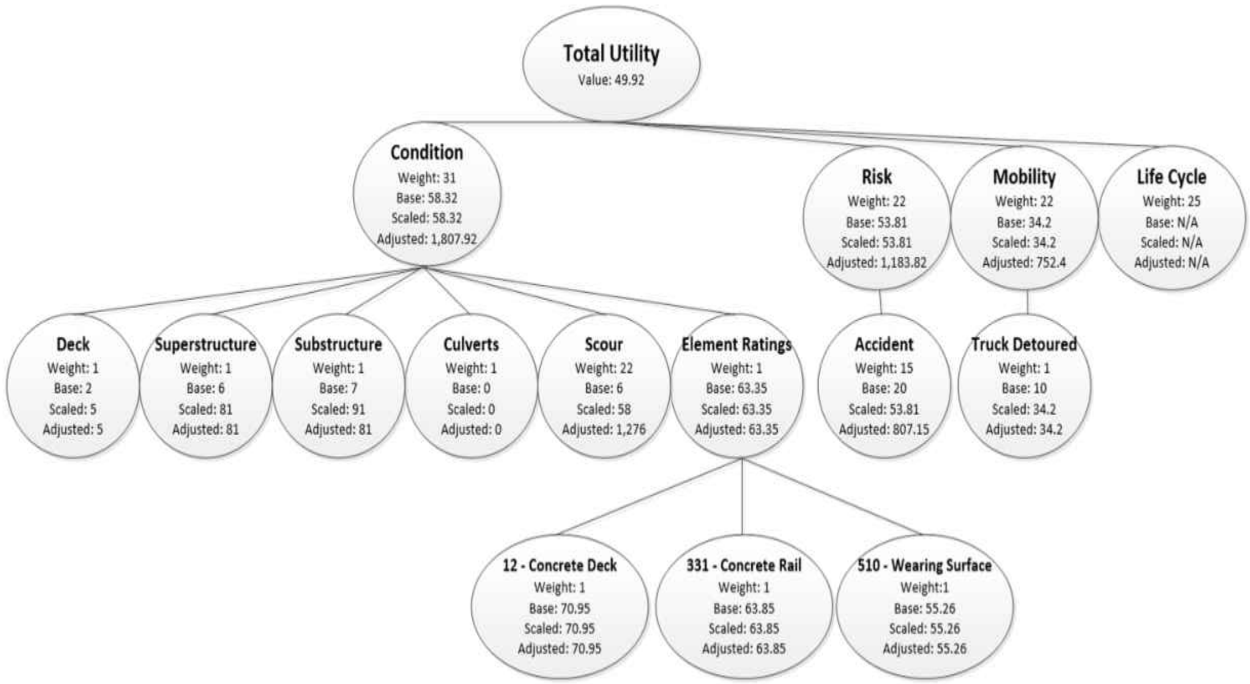


Figure 16: A multi objective analysis example in BrM

BrM however lacks support for NDE methods, consideration of local DOT practice, accurate calibration of bridge ranking (sometimes yields unrealistic ranking of bridges in the network analysis) and automation in inputting inspection data.

A basic limitation of both the NBI and the element level approach is that the data collected relies upon visual inspection techniques. The subjective, variable, and generalized nature of this data makes it less desirable for comprehensive long-term decision making [46]. For this reason, The Federal Highway Administration (FHWA) has initiated the LTBP Program, which was authorized in the Safe, Accountable, Flexible, and Efficient Transportation Equity Act: A Legacy for Users, signed into law in August 2005. The LTBP Program is a minimum 20-year, multifaceted research effort that is strategic in nature and has both specific short-term and long-range goals. The overall objective of the LTBP Program is to inspect, evaluate, and periodically monitor representative samples of bridges nationwide in order to collect, document, maintain, and manage high-quality, quantitative performance data over an extended period of time. The program will employ sensing technologies and non-destructive evaluation and testing tools in addition to conventional bridge inspection approaches [47]. An important tool, LTBP InfoBridge was also recently launched to support the bridge data management (Figure 17).

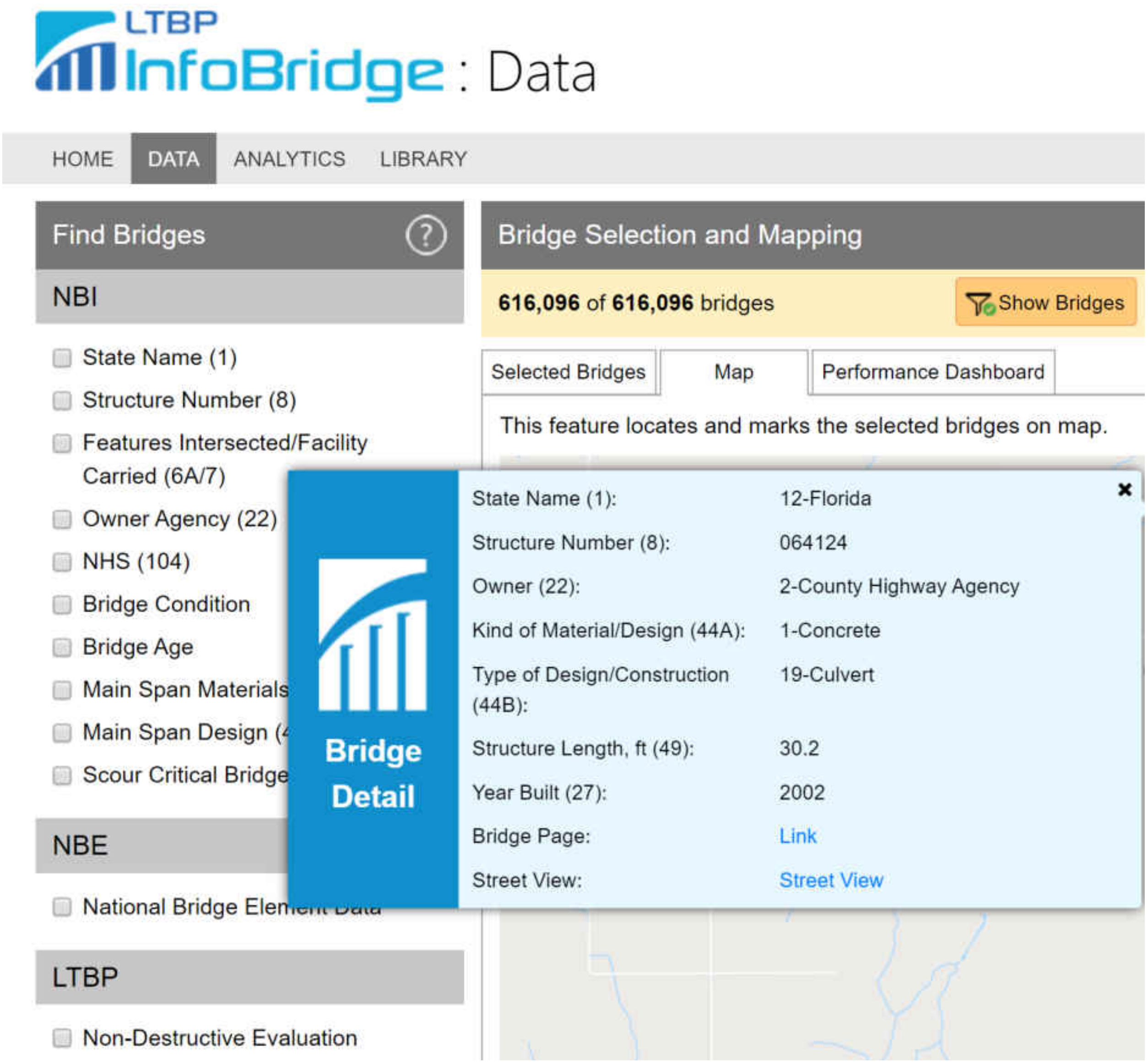
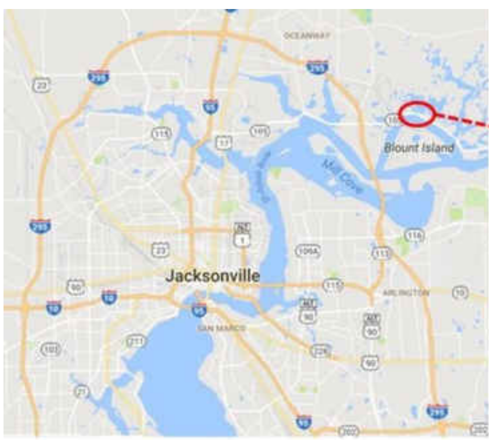


Figure 17: LTBP InfoBridge, an online data management tool for bridges

A Case Study on Non-Contact Bridge Inspection

A case study from a bridge in Jacksonville, FL carried out by CITRS researchers is provided in this section to demonstrate the performance of novel NDE methods especially Infrared Thermography. This case study involves basic equipment such as a handheld infrared camera, hammer and measurer to investigate the challenges of infrared and associated image-based operations. The study focuses on both the successful detection results and the uncertainties in detections for different bridge elements including the deck, underside, piers, and pier caps. The locations of the bridge site and the inspections are shown in Figure 18.



(a) Location of the test bridge

(b) Satellite image of the test bridge

(c) Bridge deck (Sidewalk)

(d) Underside of deck, piers & pier cap

Figure 18: Case study location and bridge testing site

On the bridge deck and the sidewalk, both sound and delaminated areas were validated by hammer sounding method. Some indications of thermal contrast were caused by discoloration of the surface even though those areas were found intact. However, they can still be distinguished easily by visually examining them at the site or from the visual image. Therefore, comparing the visual and IR images is imperative when using IRT to correctly identify the cause of the thermal contrast (i.e. delamination, discoloration and debris) especially for daytime IRT application. If the surface is a different color than the surrounding, that area should be evaluated carefully, since that part might not be delaminated. On the other hand, if the surface color is uniform on the visual image, but the IR image shows thermal contrast, that area has a high probability of being delaminated. In this study, delaminated areas of the bridge deck, underside of the deck and the pier were detected successfully by IRT and validated by hammer sounding as shown in Figure 19.

(a) Bridge deck (b) Bridge pier (c) Underside of deck

Figure 19: Verification of concrete defects detected from infrared camera

The uncertainties associated with IRT application, such as the effect of reflections from surrounding objects and solar radiation were also investigated during the case study. These areas were also cross-checked by visual examination of the physical site or the visual image of the surface. The delaminated areas of the bridge deck, underside of the deck and the pier were successfully detected by IRT and validated by hammer sounding.

Even though IRT showed very successful results in locating a sub-concrete defects in an ideal environment, certain conditions created obvious false positive results. As shown in Figure 20, IRT can be affected by reflections from surrounding objects. It can be seen in the infrared images, the person’s body temperature was reflected on the concrete surface, causing false positive temperature spots for delamination.

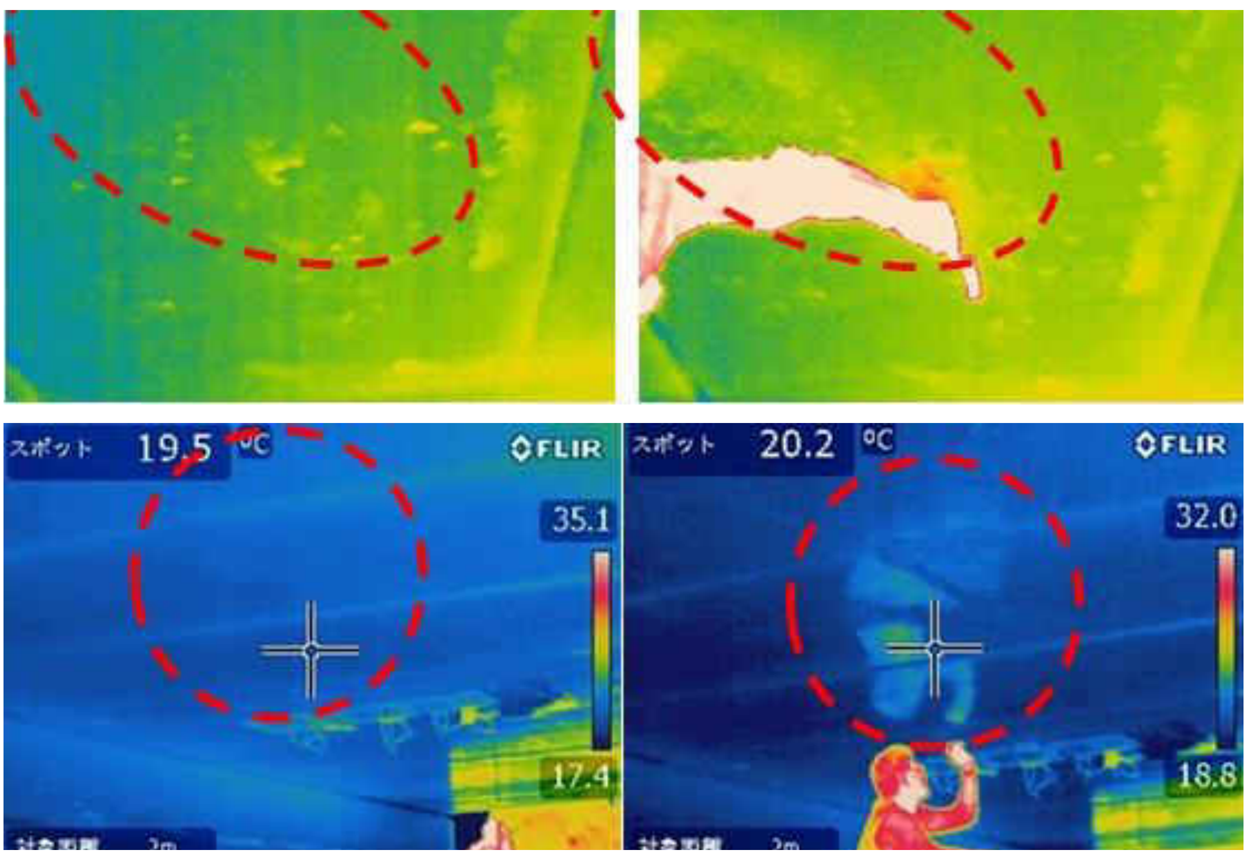


Figure 20: Effect of reflection from human body temperature

The effect of solar radiation affects the boundary between exposed and shaded area as shown in Figure 21. When both areas are captured in the same IR image, it becomes difficult to detect possible delamination within the exposed or shaded areas in the IR image, making condition assessment challenging. Therefore, nighttime is preferable for IRT application under passive condition, especially for bridge decks since they experience intense solar radiation during the day.

Figure 21: Effect of solar radiation on IRT detection result

This case study generated very important results about the strength and challenges of a novel NDE technology, infrared thermography. The NDE method showed significant potential in identifying and locating sub-concrete defects; yet the results could be easily impacted by external conditions such as solar radiation, reflection from human body temperature, shiny or colored objects. However, cross-analysis with visual imagery could tackle most of these problems, helps to determine the false positive results in IRT.

CHAPTER THREE: ATTENTION BASED REAL-TIME DEEP LEARNING FOR CONCRETE INSPECTION

Deep learning approaches have been shown robust in identifying damages; yet these methods require precisely labeled, large amount of training data for high accuracy complementary to visual assessment of inspectors. Especially in image segmentation operations, in which damages are subtracted from the image background for further analysis, there is a strong need to localize the damaged region prior to segmentation operation. However, available segmentation methods mostly focus on the latter step (i.e., delineation), and mislocalization of damaged regions causes accuracy drops. Inspired by the superiority of human cognitive system, where recognizing objects is simpler and more efficient than machine learning algorithms, which are superior to human in local tasks, this dissertation study describes a novel method to dramatically improve the accuracy of the damage quantification (detection + segmentation) using an attention-guided technique. In the proposed method, a fast object detection model, Single Shot Detector (SSD) trained on VGG-16 base classifier architecture, performs a real-time crack and spall detection while working interactively with the human inspector to ensure recognition of the region of interest is well-performed. Upon the inspector's verification, happening in real-time, the detected damage region is used for damage segmentation for further analysis. This initial region of interest selection drastically lowers the computational cost, required amount of training data and reduces number of outliers. For optimal performance, a modified version of SegNet architecture was used for damage segmentation. Based on various performance criteria, the proposed attention-guided infrastructure damage analysis technique provides 30% more precision with a very minor sacrifice in computational speed compared to analysis without using attention guide.

The proposed attention based real-time deep learning methodology focused on cracks and spalls in this dissertation and will expand the system scope in the future with more defect types. The explored AI system will first perform real-time detection of concrete defects with a minimal assistance by a human inspector; then the system will use the refined region as an attention mechanism to restrict semantic segmentation, which will be held by a CNN based segmentation operation. Due to the attention-guidance, the segmentation will be performed much more accurately with much less false positives and leakages. For the supervised training of the CNN-based detection and segmentation models; the training data is first prepared by collecting images from different sources, annotated in a unified format and augmented for further robustness. After CNN models were trained using clusters of Graphical Processing Units (GPUs), an extensive evaluation was carried out to test the performance and the accuracy of the AI system.

Data Preparation, Augmentation and Annotation

The available defect images were gathered from various sources including industry partners, transportation agencies and other academic institutions. Some of the data were only categorized but not annotated; considerable portion of the data were annotated with bounding boxes and a smaller dataset was annotated for segmentation. An extensive data augmentation was however applied to the datasets to further increase AI prediction accuracy. The data augmentation included rotation, scaling, translation and Gaussian noise as shown in Figure 22.

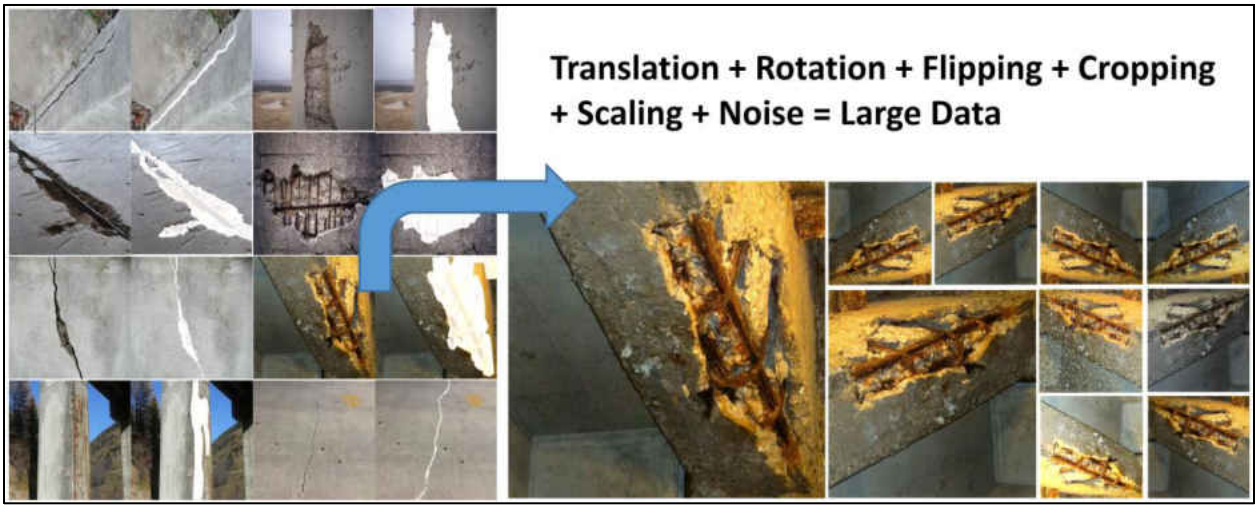


Figure 22: Annotation of dataset and augmentation with translation, scaling, rotation and noise.

The annotation styles of all of the training datasets were unified and converted to Pascal VOC 2012 annotation format [48]. The information of the training datasets is summarized in Table 1.

Table 1 Summary of the training datasets

Dataset Annotation	Class Types	Dataset Size	Source
Sub-cropped, labeled but not annotated	Cracking and intact concrete	40,000 images (with large data augmentation)	Concrete crack dataset [49]
Labeled and annotated for boundary boxes	Line crack, alligator crack, joint failure, spalling	9000 images, 15500 labels (no data augmentation)	Road damage dataset [50]
Labeled and annotated for segmentation	Cracking and spalling	2000 images (with little data augmentation)	Bridge inspection dataset [51]
Labeled and annotated for segmentation	Cracking and spalling	300 images (with no data augmentation)	Image scrapping and some field data

Since, the datasets obtained from other studies were sieved into a clean, relevant and compatible dataset for the deep learning methodology described here; the total size of the dataset was reduced to 34102 labels (27186 labels with bounding box annotation, 6919 labels with segmentation annotation). The total available data was split into 70% training, 15% validation and 15% test data. The test data was never used during training to validate the performance metrics; therefore, the models would have no familiarity with the test dataset prior to final evaluation.

Model Training and Hyperparameters

The supervised trainings of the AI models were performed in the Newton Visualization Cluster operated by UCF Advanced Research Computing Center (2014). The Newton cluster includes 10 compute nodes with 32 cores and 192GB memory in each node; two Nvidia V100 GPUs are available in each compute node totaling 320 cores and 20 GPUs. A single training takes about 22 hours on the GPU cluster computer for total of 1,000,000 steps for the detection model and 31 hours for the segmentation model. Keras v2.2 Python wrapper for Python with TensorFlow v2.11 backend framework is used for model creation and performing the model trainings [53]. The model fit function in Keras performs cross validation between training batches and gives a performance score at every new batch. At the end of each epoch (i.e. after all images are input once), the performance is validated on the validation set and a validation score is calculated. After training finishes, the model is evaluated again on the test dataset.

There are critical challenges in effectively training a deep learning model. One major challenge is overfitting. Overfitting usually occurs when the data is either too small compared to the size of the neural network architecture or when the data is large but not diverse enough. If dataset is too small, more data can be added to the training using data augmentation techniques or the complexity of the network architecture may be reduced (e.g. decreasing the number of convolutional layers). If the dataset is large but not diverse enough, then regularization may be added to the network (e.g. adding dropout layers or L1/L2 regularization) [54]. Early stopping is another way to tackle the overfitting problem.

Another cause of poor training performance is the wrong selection of training hyperparameters. Learning rate is often chosen to be too small in order to obtain reduced training loss. However, validation loss then becomes much larger. This indicates that the model is fitting over the branch of the loss function and not converge around the local minima [55]. Increasing the learning rate, on the other hand, yields to nonconverging loss function. To overcome these challenges, learning rate scheduling and early stopping method were used during the training. The optimal hyperparameters found for detection and segmentation models are shown in Table 2.

Table 2 Summary of the training datasets

Hyper Parameters	**Detection Model**			**Segmentation Model**		
Learning rate scheduling	*Initial* 0.001	*Epochs<100* 0.0001	*Epochs<200* 0.00001	*Initial* 0.01	*Epochs<25* 0.001	*Epochs<50* 0.0001
Model Optimizer	Adam Optimizer ($\beta=0.9$, $\varepsilon=10^{-8}$)			RMSprop Optimizer		
Batch Size	Training = 32, Evaluation = 4			Training = 16, Evaluation = 4		
Validation Loss	Categorical Cross-entropy			Binary Cross-entropy		
Regularization	Dropout, L2 Regularization			Batch Normalization		
Monitoring	Early Stopping			Early Stopping		

The effect of batch size selection on the training performance was also investigated during trainings. Even though some studies indicated that large batch sizes (more than 64) in a Stochastic Gradient Descent based model might impact the training performance by causing less gradient noise in the training, hence leading to poor generalization behavior [56]; a contrary result was observed when comparing the batch sizes smaller than 32. During training of a limited dataset, when very small batches were fed through the network, the network failed to provide a stable enough estimate of what the gradient of the full dataset would be when averaging the gradients of small batches [57]. The comparison of different batch sizes in terms of total loss values (the summation of classification loss and localization loss) for the damage detection model is shown in Figure 23.

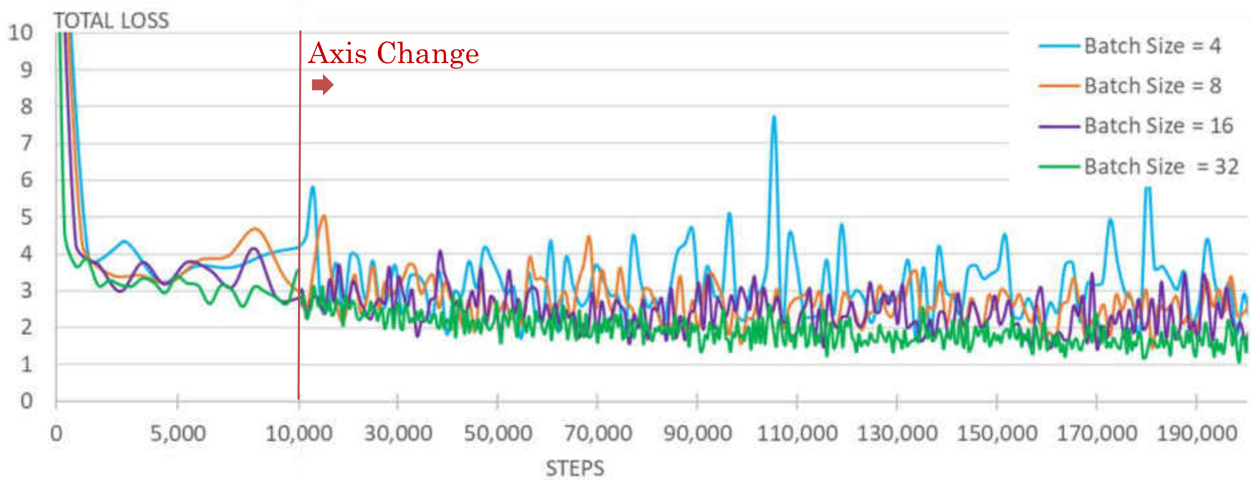


Figure 23: Effect of batch size selection on the training performance (the detection model is used for reference and the steps are shown 4 times more frequently after the 10,000th step).

Selection of the model optimizer is also very important. Model optimizers search for local minima and the maxima points of the training model's cost function. Commonly used optimizers are Adam Optimizer, Stochastic Gradient Descent (SGD) and RMSprop Optimizer [58]. All optimizers were experimented on both models and the best performance was observed when Adam Optimizer was used in the detection model and RMSprop Optimizer was used in the segmentation model. The training performance comparison of these model optimizers on this particular dataset is shown in Table 3. For damage detection, minimum obtained classification and localization loss values; for damage segmentation, minimum obtained Dice loss value were used as comparison metrics.

Table 3: Comparison of the model optimizers

	Adam Optimizer	**Stochastic Gradient Descent (SGD)**	**RMSprop Optimizer**
Damage Detection[1]	classification loss = 0.69 localization loss = 0.48	classification loss = 0.92 localization loss = 0.64	classification loss = 0.78 localization loss = 0.57
Damage Segmentation[2]	Dice loss = 0.55	Training Failed	Dice loss = 0.31

[1]Evaluation loss values obtained at the end of 193rd epoch (114,835 training steps)

[2]Evaluation loss calculated by Dice loss function obtained at the end of 112th epoch (33,936 training steps)

As shown in Table 3, Adam Optimizer yielded the lowest evaluation loss values when the concrete defect dataset was fully trained on the object detection model. Adam Optimizer uses exponentially weighted averages just like RMSprop but also integrates the idea of momentum optimization; so, it converges faster but at the same time maintains its stability [59]. However, for more complex networks, Adam Optimizer won't converge to an optimal solution as opposed to RMSprop as shown in the segmentation model's results despite the long training duration. Stochastic Gradient Descent (SGD) performed poorly in both models and even failed to train the segmentation model. SGD uses random search to escape the local minima/maxima points but sometimes causes major pikes when converging the cost function [60].

Real Time Damage Detection

For real time detection of damages, a lightweight architecture that can run on mobile GPUs was selected. SSD: Single Shot MultiBox Detector (SSD) is a relatively new, fast pipeline developed by Liu et al. (2016). SSD uses multi boxes in multiple layers of convolutional network and therefore has an accurate region proposal without requiring many extra feature layers. SSD predicts very fast while sacrificing very little accuracy, as opposed to other models in which significantly increased speed comes only at the cost of significantly decreased detection accuracy [62]. The network architecture of the original SSD model is shown in Figure 24.

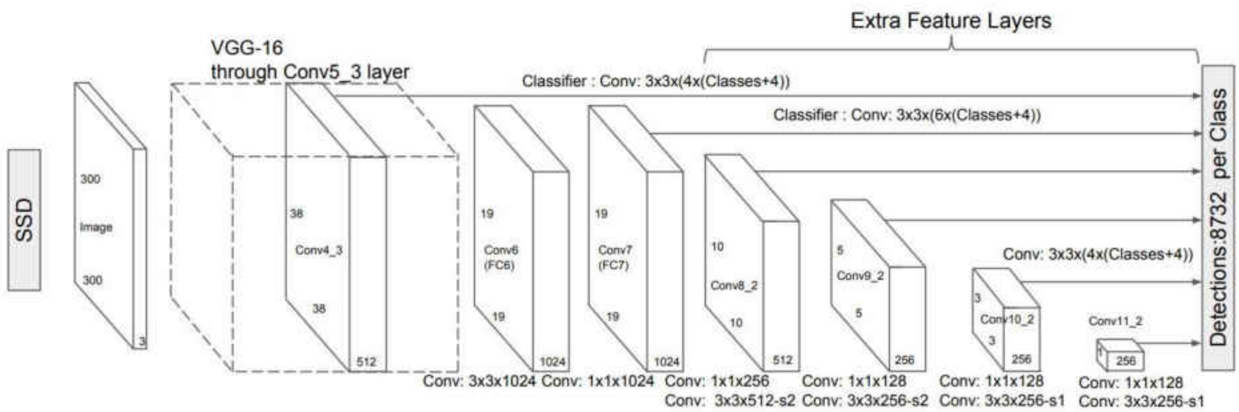


Figure 24: Original SSD network architecture [61].

Original SSD paper uses *VGG-16* as a base architecture. VGG has become widely adopted classifier after it won the 2015 ImageNet competition [63]. Although newer classifiers such as *MobileNetV2* offers faster prediction speeds at comparable accuracy [62], VGG is a better choice to benefit transfer learning in this study due to bottleneck connections of *MobileNetV2* making the weight transfer difficult. Transfer learning allows employing the weights of already trained networks by fine-tuning only the certain classifier layers based on the size of the available dataset. Figure 25 shows challenging cases where damage detection algorithm from real-world images show promising results.

Figure 25: Damage detections on real-world images (Left: Spalling on a beam at far location, Right: Longitudinal crack located at a building wall).

Attention-Guided Damage Segmentation

For concrete defect assessment, it is not solely enough to detect the damage in a bounding box; but the damage also needs to be segmented from intact regions in order to perform quantification including necessary defect measurements for understanding the extent of the defects. Therefore, another complementary deep learning model was implemented in parallel to the SSD to perform segmentation of the damage regions. Popular segmentation models such as *FCN, UNet, SegNet* and *SegCaps* [64] were investigated; however their architectures were found to be too large for the small annotated dataset used in this study. To overcome this challenge, the VGG weights that were fine-tuned in SSD architecture were used in a relatively small, customized segmentation architecture that is inspired by the *SegNet* model [65]. The architecture of the modified *SegNet* model is shown in Figure 26 and

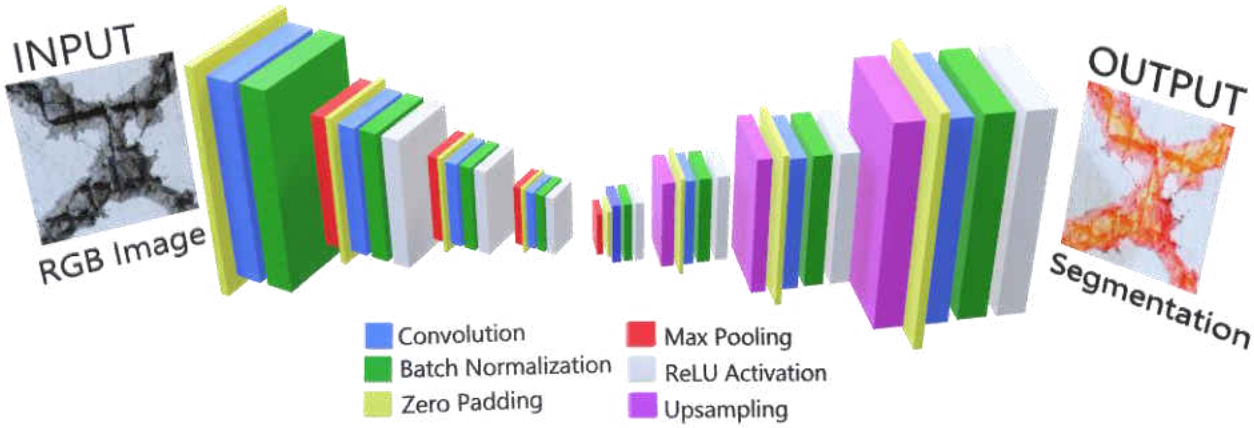


Figure 26: An illustration of the modified *SegNet* architecture.

Table 4: Serialized architecture of the modified SegNet model

CNN Layer	Output Shape	# of Parameters	Layer Type
Input Layer	(224, 224, 3)	0	
Zero Padding	(226, 226, 3)	0	
Convolution	(224, 224, 3)	1792	
Batch Normalization	(224, 224, 64)	256	
Max Pooling	(112, 112, 64)	0	
Zero Padding	(114, 114, 64)	0	
Convolution	(112, 112, 128)	73856	
Batch Normalization	(112, 112, 128)	512	
ReLU Activation	(112, 112, 128)	0	
Max Pooling	(56, 56, 128)	0	
Zero Padding	(58, 58, 128)	0	ENCODERS
Convolution	(56, 56, 256)	295168	
Batch Normalization	(56, 56, 256)	1024	
ReLU Activation	(56, 56, 256)	0	
Max Pooling	(28, 28, 256)	0	
Zero Padding	(30, 30, 256)	0	
Convolution	(28, 28, 512)	1180160	
Batch Normalization	(28, 28, 512)	2048	
ReLU Activation	(28, 28, 512)	0	
Max Pooling	(14, 14, 512)	0	
Zero Padding	(16, 16, 512)	0	
Convolution	(28, 28, 512)	2359808	
Batch Normalization	(28, 28, 512)	2048	
ReLU Activation	(28, 28, 512)	0	
Up Sampling	(56, 56, 512)	0	
Zero Padding	(58, 58, 512)	0	
Convolution	(56, 56, 256)	1179904	
Batch Normalization	(56, 56, 256)	1024	
ReLU Activation	(56, 56, 256)	0	
Up Sampling	(112, 112, 256)	0	
Zero Padding	(114, 114, 256)	0	DECODERS
Convolution	(112, 112, 128)	295040	
Batch Normalization	(112, 112, 128)	512	
ReLU Activation	(112, 112, 128)	0	
Up Sampling	(224, 224, 128)	0	
Zero Padding	(226, 226, 128)	0	
Convolution	(224, 224, 64)	73792	
Batch Normalization	(224, 224, 64)	256	
Convolution	(224, 224, 1)	65	
ReLU Activation	(224, 224, 1)	0	
Total params: 5,467,265 **Trainable params: 5,463,425** **Non-trainable params: 3,840**			

As a unique approach for damage segmentation, an attention guidance approach was proposed in this study. A sequential connection was created between detection and segmentation models. First, images were fed into damage detection pipeline and when the human-inspector verified the bounding box, damage segmentation was executed only for the region inside the detected bounding box. This approach significantly improved the accuracy of segmentation and successfully prevents outliers. Figure 27 shows qualitatively how attention guided segmentation was superior to the segmentation without attention guidance. On the left image in the figure, the model performs only pixel wise segmentation operation to find the damage regions and subtract them from the background. On the right image, the model first performs detection and then immediately inputs the detection results into the segmentation pipeline.

Figure 27: Effectiveness of attention guided segmentation (Left: Defect analysis resulted in some false positive results when only the segmentation model was used; Right: Misclassified pixels were readily removed when the detection pipeline created attention guidance for segmentation).

The shown example in Figure 27 substantiates the performance improvement of the segmentation model on challenging scene when the model inputs only the detected regions into segmentation architecture instead of putting the entire image. The segmentation architecture this time classifies the damage pixels at much higher accuracy when there is an initial region of interest selection.

Human-Centered AI and Semi-Supervised Learning

The proposed deep learning methodology for concrete defect analysis is designed for human-computer interaction environments such as wearable holographic headset and handheld mixed reality (MR) devices. Using these technologies that integrated the proposed methods, an inspector can continuously communicate with the AI system. The human-computer interaction in MR will entail practical human-AI collaboration to create a collective intelligence. The AI models for damage detection and segmentation in the proposed methodology will allow the inspector to adjust in real-time the prediction threshold values, model inference parameters and even the attention regions. This type of human-centered system can easily outperform a fully automated robotic technique [66]. Similar systems are commonly seen in automated vehicle technologies, visualization systems in the health industry and video game engines. The semi-supervised approach referred in this methodology consists of automatically generating the fine-tuning data by using inspector's adjustments in the detection boxes and segmentation regions during the routine inspection performance. Inspector only provides minimal input to the system during inspection and the annotated training data is automatically generated. The system periodically schedules a fine-tuning in the cloud and optimizes the weights of the last six convolutional layers. The fine-tuned model weights are automatically updated in the deployed device. The AI framework thereby improves its prediction accuracy as the inspector uses the device without doing any data preparation. Figure 28 describes the overall system.

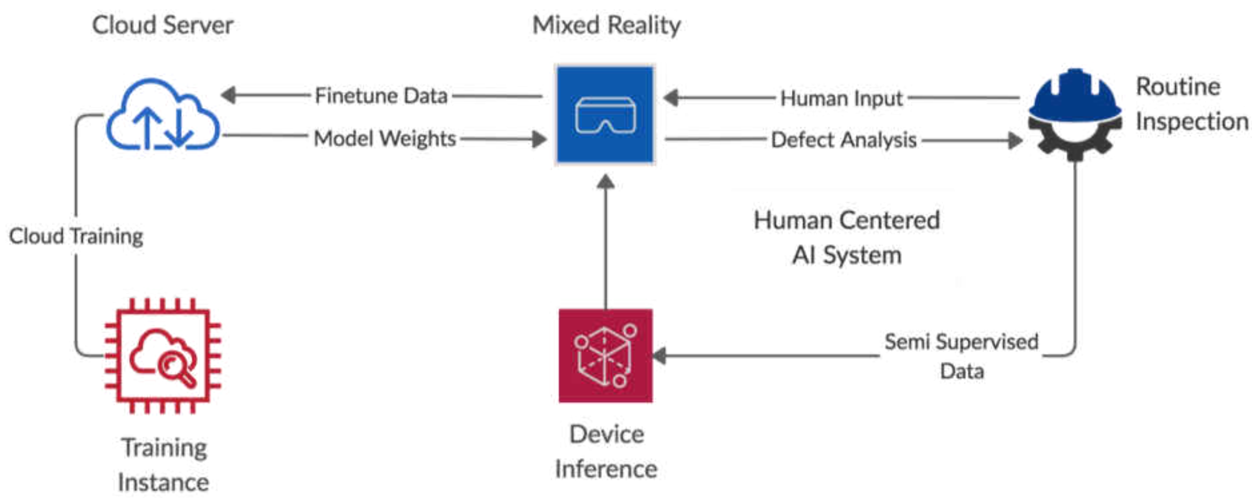


Figure 28: The interaction diagram of the system components in the proposed methodology.

During a bridge inspection, by asking the human inspector to modify prediction threshold will help improving the accuracy of the detection and determining the boundary region of the segmentation. In Figure 29, real-time damage detection is not showing one of the spall regions to the inspector when the prediction threshold is set to 0.8; when the inspector adjusts the value to 0.7; the missing spall region is also detected. (The value represents probability of accurate prediction.)

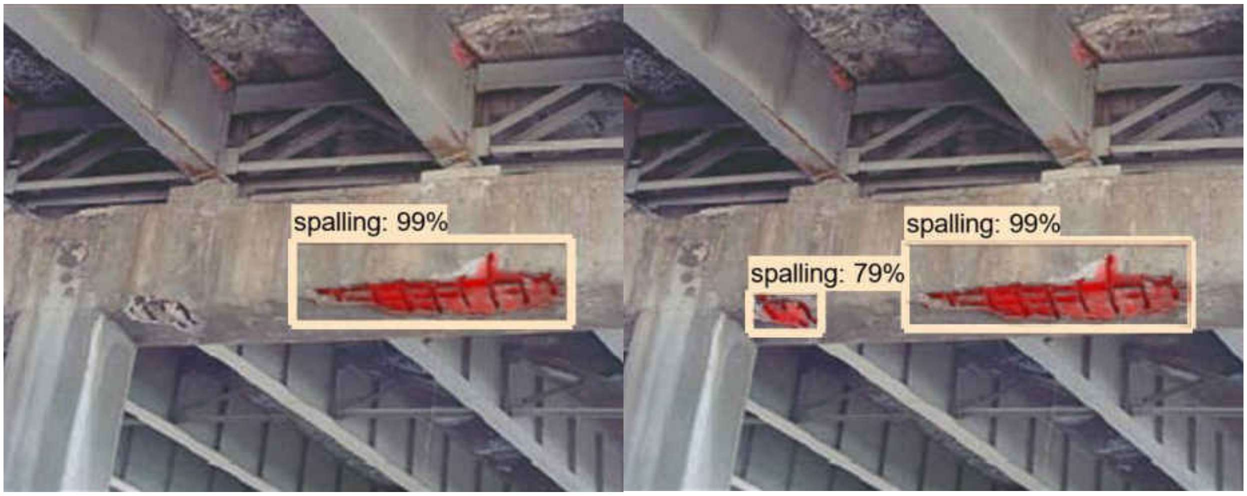


Figure 29: Example of human-AI collaboration in the proposed methodology (detection AI alone on the left misses a spall, while human-assisted AI detects all spalls on the right with threshold adjustment by the inspector).

Similarly, the human inspector can also fine-tune the segmentation boundary by adjusting the prediction threshold. Thus, the damage area can be calculated at higher accuracy. The fine-tuned segmentations along with the corresponding bounding box coordinates are recorded for future re-training while benefitting from semi-supervised learning. The threshold adjustment is only one of the several ways to achieve a human centered system. The user may also provide a manual region of interest by quickly drawing a target area, which will be used in the proposed attention guide procedure. Any wearable device with MR capability would be ideal to facilitate this kind of human-computer interaction. Some example results from the proposed human-AI collaborative damage detection and segmentation are shown Figure 29.

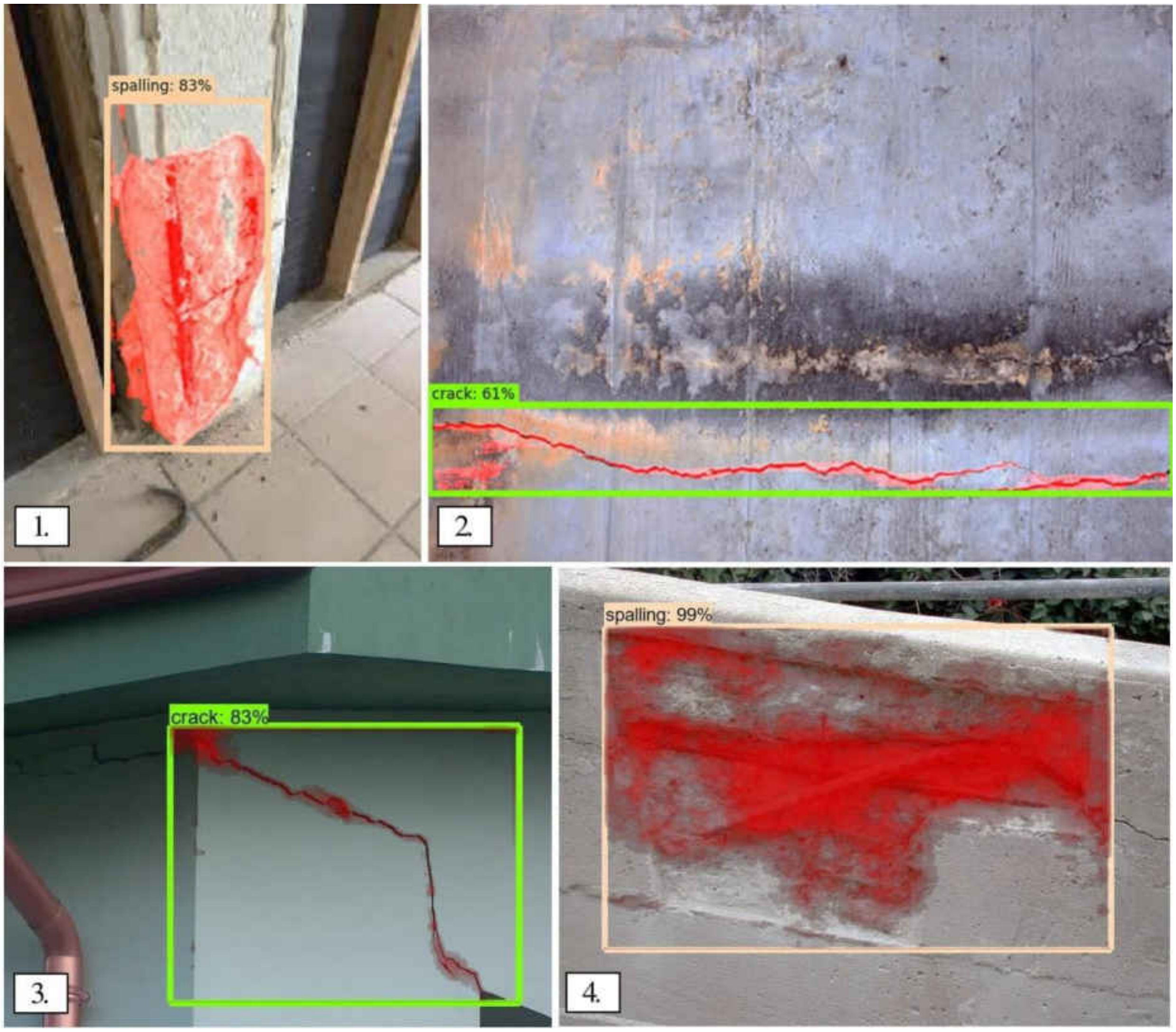


Figure 29: Example results from the proposed human-AI collaborative damage detection and segmentation (images 1-4 show results from the test dataset using only threshold adjustments in the segmentation pipeline; images 5-6 show implementation results from the deployed MR headset in which the inspector sees the analysis output projected onto real world).

Deep Learning Model Evaluation

The performances of the damage detection and segmentation AI models were evaluated using the accepted evaluation procedures used in the literature. The evaluations were carried on only the AI models without benefiting the human centric framework in which the human inspector assists the AI with minimal input to improve the prediction accuracy. Therefore, real-life performance results from the human-AI collaboration are expected to be superior than the evaluation results of the individual AI models in this study.

When evaluating machine-learning models, classifying the predictions into four categories: true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) is a common practice. However, evaluating object detection segmentation models will require additional metrics to measure the accuracy of the detected or segmented areas. Mean Average Precision (mAP) is a performance indicator that finds the average of maximum precisions at different recall values based on a confidence threshold. Average precision (F1 Score) is calculated by as follows:

$$Presicion = \frac{TP}{TP + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$Average\ Precision\ (F1) = 2 \times \frac{presicion \times recall}{precision + recall} \tag{3}$$

To calculate the precision and recall; TP, FP and FN need to be determined from an evaluation metrics. One of the common metrics is the Intersection over Union (IoU). IoU is also known as Jaccard Coefficient, first introduced by Jaccard (1912). IoU measures region overlap without concentrating on true boundaries, hence, IoU is preferable method for detection measures, not for segmentation. This metric is simply the ratio between the intersection and the union of the predicted boxes and the ground truth boxes.

$$IoU = \frac{area\ of\ overlap}{area\ of\ union} = \frac{TP}{TP + FP + FN} \tag{4}$$

A similar approach is the Dice Similarity Coefficient (DSC) first proposed by Dice (1945). DSC is a reliable and the most commonly used metric for delineation accuracy. It measures regional overlap in the boundary of segmented and ground truth contours. DSC was simply calculated by the Equation (5).

$$DSC = \frac{2TP}{2TP + FP + FN} \tag{5}$$

The evaluation of the concrete damage detection and segmentation was carried out using IoU and DSC metrics in mAP calculations. In addition to the mean average precision calculation, the average prediction speed (in milliseconds) was also monitored during the evaluations. The results are shown in Table 5.

Table 5: Performance Evaluation of Segmentation Model

	mAP using IoU	mAP using DSC	Speed (ms)
Defect detection using SSD with VGG16 backbone	0.74	-	0.17
Segmentation with no attention guide (*SegNet* only)	-	0.64	0.68
Segmentation with attention guide (SSD + *SegNet*)	-	0.88	0.72

According to the evaluation results; the damage detection model could predict the damage boundaries in 74% mean average precision using the IoU metric and 85% precision using the DSC metric. The segmentation model, on the other hand, showed significant improvement when the detected boundaries were used as initial region of interest. Segmentation without attention guide predicts the damage regions in 52% mAP using IoU metric and 68% mAP using DSC metric. When segmentation model was coupled with the detection model (attention guided segmentation) the precision increased up to 79% with IoU metric and 88% with DSC metric. As for the inference speeds; the models were tested on the Pascal GPU, a mobile chipset commonly used by mixed reality devices. The damage detections performed very fast at 0.17 milliseconds per frame. The segmentation was not close to real time but still could predict under a second. The segmentation model operated at 0.68ms average speed when executed alone on the input image. However, it operated 13ms faster when the attention guide provided an initial region of interest (attention). Therefore, the sacrifice on the computational speed was minimal when damage detection and segmentation were coupled (only 0.04ms).

The evaluation results show a clear comparison of *Segmentation Only* operation and the *Sequential Operation* (Detection + Segmentation). The sequential pipeline resulted in approximately 30% higher precision in subtracting the spall and crack regions than the segmentation pipeline alone with only slight sacrifice in the computation time.

CHAPTER FOUR: INTEGRATION OF THE DEEP LEARNING METHODOLOGY INTO MIXED REALITY

The proposed AI assisted infrastructure assessment using mixed-reality (MR) technology employs the state-of-the-art methods and algorithms from interdisciplinary practices. Machine learning is vastly used for robust detection of cracks and spalls on infrastructures whereas human-computer interaction concepts are employed for improving the assessment performance by including the professional judgment of the human inspector. MR is an excellent platform to maintain this interaction since it augments virtual information into the real environment and allows the user to alter the information in real-time. In this proposed methodology, bridge inspector uses MR headset during routine inspection of infrastructure. While the inspector performs routine inspection tasks, the AI system integrated into the headset continuously guides the inspector and shows possible defect locations. If a defect location is confirmed by the human inspector, the AI system starts analyzing it by first executing defect segmentation, then characterization to determine the specific type of the defect. If the defect boundaries need any correction or segmentation needs to be fine-tuned, the human inspector can intervene and calibrate the analysis. The alterations made by the human inspector (e.g. change of defect boundary, minimum predicted defect probability etc.) will be used later for retraining of the AI model by following a semi-supervised learning approach. Thereby, the accuracy of AI is improved over time as the inspector corrects the system.

Another advantage of the system is that the inspector can analyze defects in a remote location while reducing need for access equipment. Even though in some cases, hands-on access is evitable (i.e. determining sub-concrete defects); the system can be still effective for quick assessments in the remote location. If the defect location is far or in a hard to reach location, the headset can zoom in and still perform assessment without needing any access equipment such as snooper truck or ladder. The proposed framework is illustrated in Figure 30.

Figure 30: Visual representation of the AI powered mixed reality system. (The headset user interface and analysis environment are shown for illustration purposes.)

The MR technology has breakthrough applications especially with successful deployment of 3D user interfaces such as in computer-aided design, radiation therapy, surgical simulation and data visualization [10]. The next generation of computer games, mobile devices, and desktop applications also will feature 3D interaction [11]. There are also some other efforts for using MR technology in construction industry and maintenance operations. Kamat and El-Tawil (2007) discusses the feasibility of using AR to evaluate earthquake-induced building damage. Behzadan and Kamat (2007) investigated the application of the global positioning system and 3 degree-of-freedom (3-DOF) angular tracking to address the registration problem during interactive visualization of construction graphics in outdoor AR environments [13]. The vision-based mobile AR systems are vastly used in 3D reconstruction of scenes for architectural, engineering, construction and facility management applications. Bae et al. (2013) developed a context-aware AR system that generates 3D reconstruction from 3D point cloud. Important effort for use of AR in infrastructure inspection is also shown by several researchers [14]. Researchers in University of Cambridge currently collaborate with Microsoft to develop an effective bridge inspection practice in which the data collected from the field is visualized in MR environment in the office [15]. Moreu et al. (2017) developed a conceptual design for novel structural inspection tools for structural inspection applications based on HoloLens [17] device [16].

The proposed methodology of AI assisted infrastructure assessment using MR systems differs from the state-of–practice of current machine learning-based approaches and mixed reality implementations in several aspects. Table 6 shows comparison of the proposed method with major literature work. The major difference of the proposed method from the current mixed reality approaches is that the system performs automatic detection and segmentation of the defect regions using real-time deep learning operations instead of manually marking the defect regions in the MR platform. In this way, the system can save significant amount of time in defect assessment as opposed to marking all these defects in the current MR implementations.

Table 6: Comparison of the proposed research with the major literature work

Ioannis (2017)	**Moreu et al. (2017)**	**Bae et al. (2013)**	**Xie et al. (2017)**	**Proposed Method**
Remote bridge inspections with HoloLens	Structural inspection and measurement using HoloLens	Mixed reality for structure 3D scene reconstruction	CNN based crack detection	Mixed reality assisted bridge condition assessment
Data collections is monitored from an remote location	On-site measurement of defects	Image data is reconstructed after the data collection	Aims post processing of images to identify defects	On site system to augment bridge inspector's performance
Focuses on visualization and post-processing of data	Relies on human mostly while obtaining measurements	No detection of defects is implemented, 3D model is used for inspection	Detection performance relies on AI system only	Aims creating a collective intelligence with human - AI collaboration
Views high -resolution defect images on real size bridge model	Uses 3D projective geometry for measurement estimation	Uses widely 3D projective geometry to register images	Uses basic data augmentation techniques to increase training dataset	Uses an extensive data augmentation that generates many variations of defect images

Camera Calibration and Pose Estimation

The condition assessment methodology based on the AI system's damage analysis will require answers to following: "how wide is this crack?" "Which one of the bridge piers is closer?" "What is the camera height, rotation or focal length?" This information is required for identifying actual measures of defects for accurate assessment of infrastructures and also for augmenting a certain object onto 3D view or highlighting defects in an MR headset. Using projective geometry and camera calibration models, it is possible to perform correct projections of objects onto 3D, achieve scene reconstruction and accurately predict actual dimension of the objects. However performing transformations in 3D spaces requires use of 4D projective geometry instead of conventional 3D Euclidian geometry [70]. The projection matrix allowing camera rotation is defined as in Equation (6).

$$x = K[R \quad t]X \tag{6}$$

Where, x: Image coordinates, K: Intrinsic matrix, R: Rotation matrix, t: Translation, X: World coordinates. The projected coordinate vector x is calculated by multiplying the world coordinates by the rotation and translation free projection matrix. The coordinate parameters are then put into system of equation as in Equation (7).

$$w\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K\begin{bmatrix} \alpha & s & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \tag{7}$$

The local coordinates on image plane are represented by u and v; w defines the scale of the projected object. α and β stand for rotation angles with respect to coordinate axes and s is short for sinus function. Unity allows camera control that help developers perform correct projection onto image plane form 3D view. The projection is described in Figure 31.

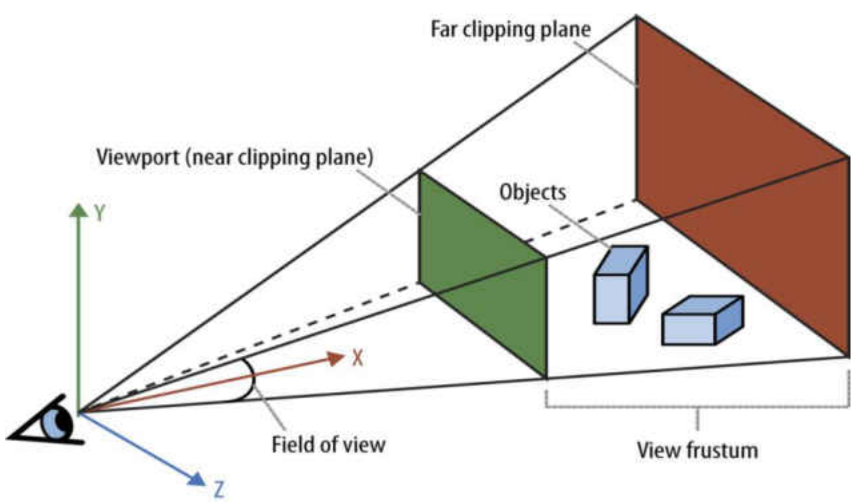


Figure 31: Camera, viewport, and projection of real world objects onto 2D image plane [71]

Real-time Image Target Tracking

Conventionally, the camera localization for augmented reality (AR) relies on detecting a known pattern within the captured images, namely a marker. The first AR tracking applications used markers placed on the object to robustly register image at different camera angles [72]. However, placing markers in the scene was not always possible depending on the applications. Simon et al. (2000) used planar structures in the camera scene to perform markerless AR tracking [73]. In a different study, Ferrari et al. (2001) introduced markerless AR with a real-time affine region tracker [74]. Recent works in the AR tracking uses Visual SLAM algorithm (simultaneous localization and mapping) to perform robust markerless tracking [75]–[77]. In this study, an open source AR tracking library called EasyAR was used to perform markerless tracking [78]. EasyAR has third party plugin for Unity 3D, a widely used platform development environment for cross-platform applications [79]. To implement the deep learning models in Unity, ML-Agents plugin was also configured [80].

After a crack or spall region is detected and accurately segmented from the scene, an image target is automatically created in the platform environment. The image targets work with feature-based 3D pose estimation using the calculated projection matrix [81]. The projection matrix can be calculated by following the stereo camera calibration procedure provided by the headset manufacturers. After successful calibration, camera intrinsic and extrinsic parameters such as camera focal length, location and orientation of the camera are retrieved in Unity using the headset sensors gyroscope and head-position-tracker. EasyAR in Unity is capable of creating on-the-fly image targets from the damage-detection output and perform fast, robust markerless tracking using Visual SLAM. 3D pose is estimated accurately at different angles and distances; the inspector still sees the overlay information on correct location as shown in Figure 32.

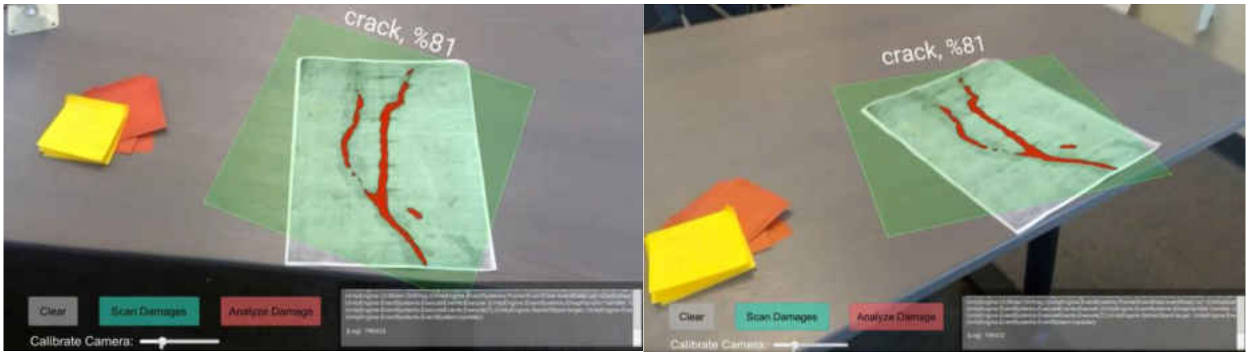


Figure 32: Markerless tracking of on-the-fly image target created from AI analysis results

One of the important limitations of the AR tracking is that segmentation results are only projected onto planar surfaces since the created image targets are two-dimensional. Therefore, volumetric calculations from curved surfaces (e.g. circular columns) such have intrinsically large error. In the future improvement of this work, 3D image targets will be created to perform AR projection onto curved concrete surfaces.

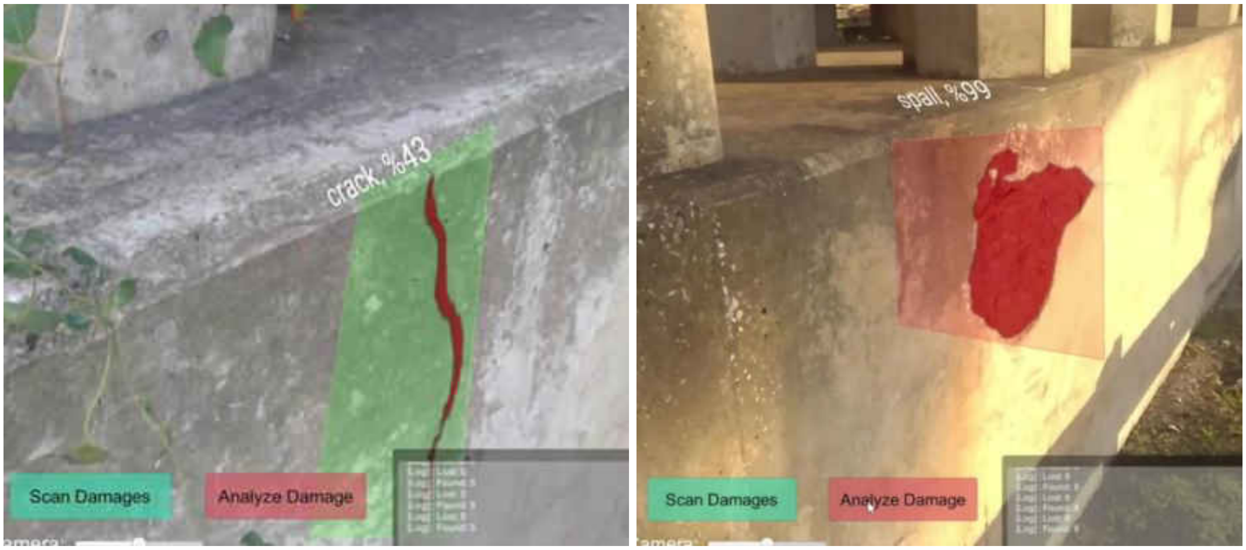


Figure 33: Real-world examples from the headset showing AI analysis results projected on the concrete defects (left: concrete crack; right: concrete spalling).

Retrieval of Dimensional Properties

Accurate retrieval of real-world dimensional properties from the AR projection is the most important step in estimating the condition of the concrete damage in the proposed system. Depending on the stereoscopic video see-through technology used, several techniques can be used to estimate the real geometric distance. The proposed methodology in this dissertation study uses spatial mapping techniques based on the binocular disparity.

After a successful calibration, basic proportioning of image pixel size to a known real-world dimension (camera offset from eye focus is known) is used to calculate the area of a spall or length of a crack. Figure 34 show calibrated image targets in the Unity platform.

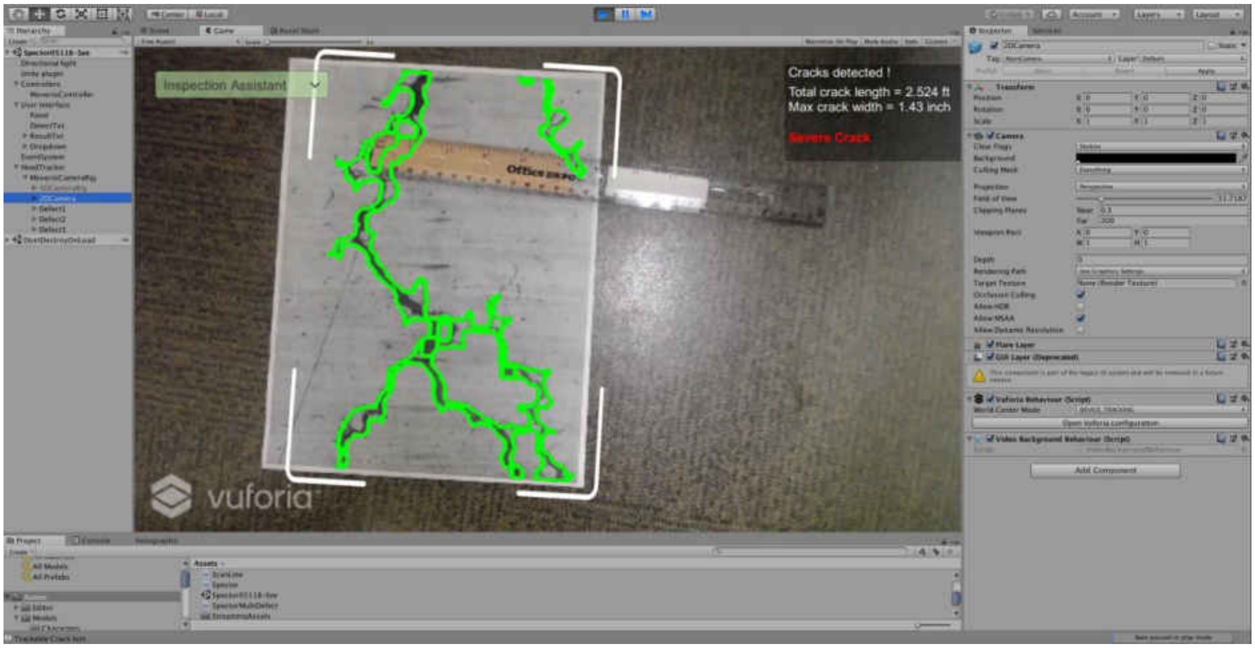


Figure 34: Calibrated image target that estimates maximum crack width in Unity

In order to improve the estimation accuracy of the geometric properties, the AR target object projected onto the defect surface is continuously calibrated using non-linear least square fitting. The necessary data points were obtained in real-time from the headset's different camera positions as the inspector gets closer to the object or looks at the defect from different angles. Figure 35 shows the details of the non-linear least square fit calculation of an example calibration of image target to estimated area of spalling.

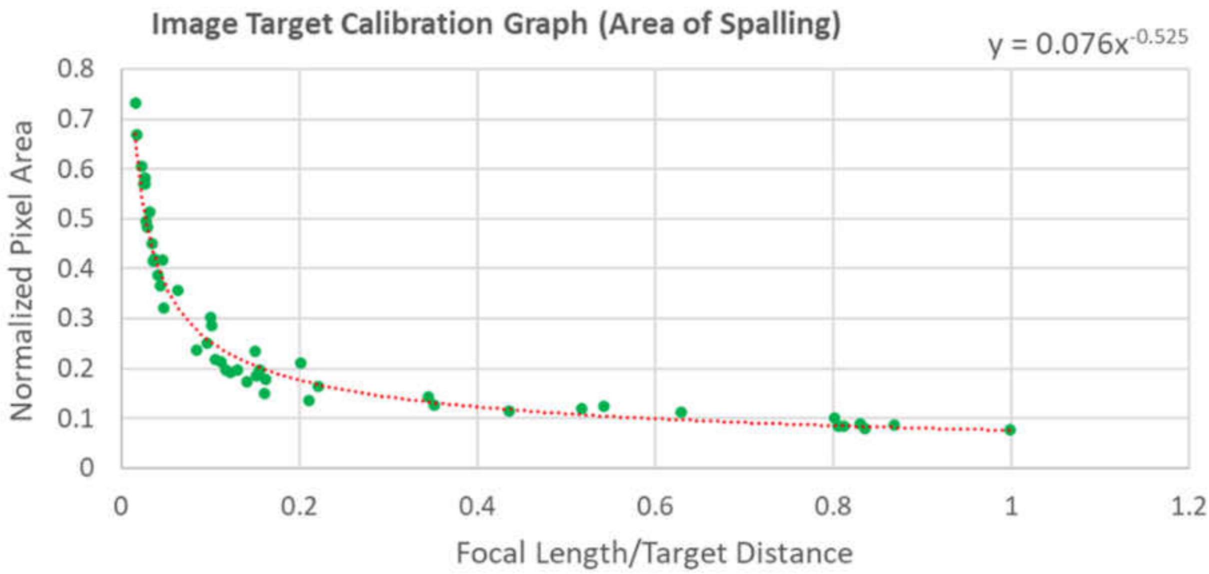


Figure 35: Calibration of image target for more accurate prediction of geometrical property

In the horizontal axis of the calibration, the estimated target distance normalized by the focal length, and in the vertical axis, the pixel area of the target normalized by the camera resolution were used. The fit equation corrects the known distance parameter in the dimension proportion to predict the area at higher accuracy.

Evaluation of Factors Affecting the Geometric Estimation

An experimental work was conducted in CITRS Lab in order to determine the factors affecting the accuracy of the retrieval of the dimensional properties. The experiment's objective was to investigate effect of both the environmental factors and the camera specifications on the geometric estimation. In the conducted experiment, the described calibration method was repeated multiple times for different ambient illumination, crack width, target distance and camera resolution using Moverio BT-300 smart glasses in a laboratory environment. A set of synthetically generated crack images with different thicknesses, brightness and patterns were printed on letter size papers and placed on a white platform. The experiment setup is shown in Figure 36 and the results of the laboratory experiment were tabulated in Table 7. In the experiment, it was assumed that the cracks were perfectly segmented using the methodology described in Chapter 3.

Figure 36: Experiment setup to evaluate factors affecting the performance of the geometric estimation in the MR headsets.

Table 7: Calculation of average error in geometric estimation under different conditions

Illumination	**Camera Res.**	**Target Dist. (ft)**	**Crack Width (in)**	**Average Error (%)**
1 light on	720p	3ft	1/6”	8.62%
1 light on	720p	3ft	1/2”	5.04%
1 light on	720p	3ft	1”	3.10%
1 light on	720p	5ft	1/6”	18.17%
1 light on	720p	5ft	1/2”	8.60%
1 light on	720p	5ft	1”	6.09%
1 light on	1080p	3ft	1/6”	5.35%
1 light on	1080p	3ft	1/2”	4.04%
1 light on	1080p	3ft	1”	2.76%
1 light on	1080p	5ft	1/6”	13.10%
1 light on	1080p	5ft	1/2”	6.27%
1 light on	1080p	5ft	1”	3.55%
2 lights on	720p	3ft	1/6”	8.02%
2 lights on	720p	3ft	1/2”	4.85%
2 lights on	720p	3ft	1”	2.87%
2 lights on	720p	5ft	1/6”	16.99%
2 lights on	720p	5ft	1/2”	7.07%
2 lights on	720p	5ft	1”	4.71%
2 lights on	1080p	3ft	1/6”	4.14%
2 lights on	1080p	3ft	1/2”	2.45%
2 lights on	1080p	3ft	1”	1.23%
2 lights on	1080p	5ft	1/6”	9.92%
2 lights on	1080p	5ft	1/2”	5.43%
2 lights on	1080p	5ft	1”	3.03%

The experiment showed that the accuracy of the geometric estimation is largely dependent on the distance of the crack from the headset camera. Due to limited capabilities of the headset used in the experiment, the procedure was repeated only at short distances of 3ft and 5ft. In larger distances, the headset could not create the AR tracker objects of the targets. The camera resolution was also found to be an important factor affecting the accuracy of the geometric estimation. Setting the camera resolution to 1080p (1920x1080) yielded acceptable results at both 3ft and 5ft. According to results, minor cracks (crack width less than 1/6") can be reliably measured only at 3ft distance and 1080p resolution. The illumination level was also found to be signification factor affecting the performance of the geometric estimation in the MR system. In the simulated environment, one of the two florescent lamps were turned off to create a darker ambient. However, the experiments were not repeated when both lights were turned off; since the targets become completely invisible in the camera and therefore the headset fails to create the AR trackers. When the both lights were turned on and the resolution was set to 1080p, 1-inch crack can be measured at approximately 98% accuracy from 3ft distance.

CHAPTER FIVE: EFFECTIVE UTILIZATION OF NON-DESTRUCTIVE EVALUATION DATA AND DECISION MAKING

Developing a bridge management strategy in the network level with efficient use of capital is very important for optimal infrastructure remediation. This study introduces a novel decision support framework that considers many aspects of bridge management and successfully implements the investigated methodology in a web-based platform. The proposed decision support system uses advanced prediction models, decision trees and incremental machine learning to generate the optimal decision strategy. The system aims to achieve an adaptive and flexible decision making while entailing powerful utilization of nondestructive evaluation (NDE) methods. The NDE data integration and visualization allow automatic retrieval of inspection results and overlaying the defects on a 3D bridge model. Furthermore, a deep learning-based damage growth prediction model will estimate the future condition of the bridge elements and utilize this information in the decision-making process. The decision ranking takes into account a wide-range factors including the structural safety, serviceability, rehabilitation cost, life-cycle cost, societal and political factors to generate optimal maintenance strategies with multiple decision alternatives. This study aims to bring a complementary solution to currently in-use systems with utilization of advanced machine-learning models and NDE data integration, but it is still equipped with main functions of those systems and capable of transferring data to them.

State of Practice in Bridge Management

The bridge management practice in the United States has improved significantly over the last 40 years both at the federal and state levels. At the federal level, the National Bridge

Inspection Standards (NBIS) unifies the method of collecting data and condition assessment on the public highway bridges [3]. The collected inspection data by the state departments of transportation (DOTs) is submitted to Federal Highway Administration (FHWA) annually in a nationwide reporting/coding format that is later entered to National Bridge Inventory (NBI) database [4]. Based on NBI, bridge owners are able to monitor condition and performance of their bridges to make accurate management decisions. FHWA imposes an appraisal rating to all government owned bridges through routine inspections that are recorded to NBI. The sufficiency rating in NBI's bridge appraisals receives input from local and global assessments as well as some additional parameters. At the state level, state DOTs may have different procedures regarding bridge asset management, funding, maintenance considerations and resource allocation. A comprehensive National Cooperative Highway Research Program (NCHRP) synthesis report published by the Transportation Research Board (TRB) puts out the differences in state practices and explains the reasons of the variety in the bridge management practices mainly on the following issues: The differences in the policy, financial, technical and institutional operations as well as the different approaches to planning, programming and budgeting [5]. According to the interviews conducted within the synthesis study, a mixed centralized and decentralized management strategies are followed in many agencies. Whereas the bridge replacement and rehabilitation projects, which can be funded by Federal Highway Bridge Program, are centralized; the maintenance and repair projects are decentralized by being funded internally. In order to maintain communication between centralized and decentralized decisions, many states employed a Bridge Management System (BMS), which incorporates detailed state procedures at the element level and NBIS requirements at federal level [6]. Although BMS has limited use toward decision-making, state agencies find it helpful in terms of compilation of data and display of short-term and long-term information [7]. According to NCHRP study, the characteristic use of BMS for state DOT decision-making is analyzed as follows [5]:

- Technical aspects in the decision making such as condition assessment and performance assessment are mainly held in BMS rather than economic and social analyses that involve life cycle cost analysis, social impact analysis etc.
- The decision-making based on BMS output is generally for short-term rather than long-term and the recommended actions are not proactive of future predicted conditions by lacking predictive models and scenario analysis.
- The decision making usually do not recommend multiple action strategies with a comparative analysis.

Enhanced Decision Support System

The decision support systems have undergone dramatic improvements in the last decade with the employment of intelligent systems and sensor monitoring. Machine learning (ML) and artificial intelligence (AI) started to play an important role in decision-making. Many researchers have recently proposed AI based decision support systems for infrastructure management. For instance, Yin (2010) developed an intelligent decision support system which does not only quantify the inspection data and evaluate the deterioration of the existing bridges but also provide optimum bridge monitoring plan for advanced management according to the project budget and timeline [7]. Quintela (2007), for another example, presents a real-time decision support system for civil engineering structures that makes use of prediction models using artificial neural network and data mining techniques. The system occupies real-time sensors to verify the accuracy of the employed prediction models [8]. In a different angle, Jiao (2013) proposed an unsupervised performance evaluation strategy for bridges using fuzzy clustering on health monitoring data. With the proposed strategy, bridge condition state can be assessed by calculating the fuzzy nearness. Lee (2008) addresses the problems of slow adoption of bridge management systems and impractical future prediction of bridge conditions. The study proposes artificial neural network-based prediction algorithm called backward prediction model to treats the inconsistency BMS inputs and bridge agencies' existing data. Bocchini (2013) develops a simple Markov chain model for life-cycle analysis of bridge networks. The proposed model includes the effect of deterioration, maintenance actions, bridge failures, and rehabilitations. These studies aim to solve bridge management problem using pre-deep learning techniques (i.e. classical machine learning and clustering techniques).

The decision support system proposed in this study has a multi-component structure in which different bridge management operations will communicate with each other. As described in Figure 37, the proposed system uses the condition assessment data obtained with NDE and predicts the future condition of the bridge by utilizing the historical information. The prediction model is trained on a cloud server in the attached deep learning instance. The decision support component utilizes an adaptive decision ranking methodology, which prioritizes the bridges based on a variety of factors. The presented ranking system initializes with the default parameters; then will gradually adapt the practice of the infrastructure owner by fine-tuning the decision weights. Synchronization with NBIS database and bridge management software entails data generation for fine-tuning the deep learning models. Finally, the decision support component generates multiple maintenance decision strategies while optimizing the cost and performance.

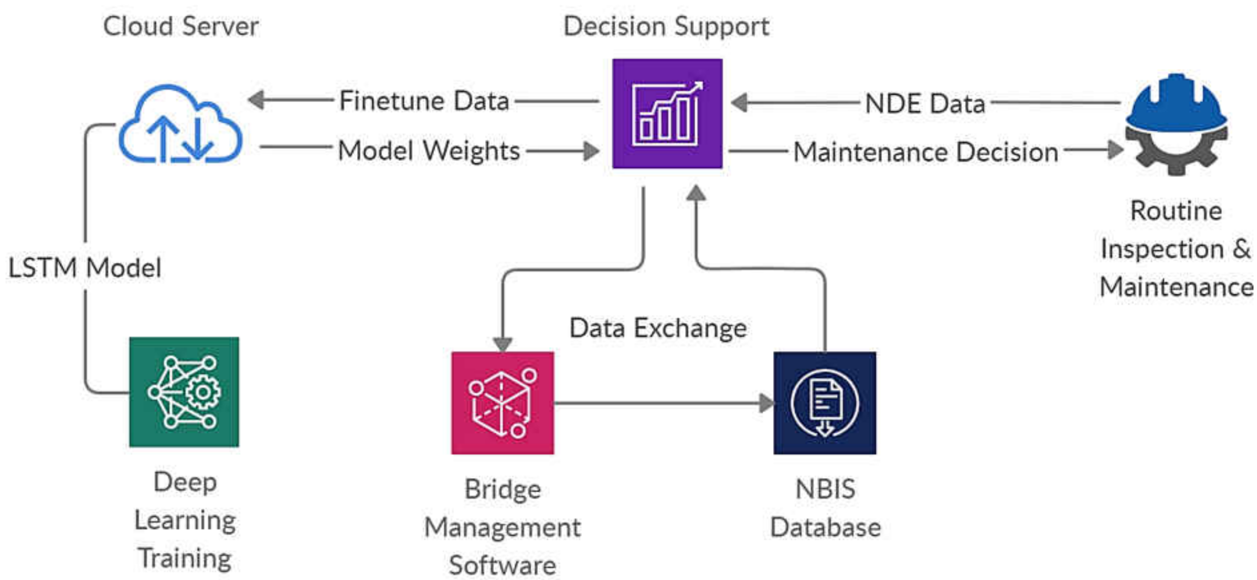


Figure 37: System component diagram of the proposed decision support framework

The proposed infrastructure support methodology was implemented in web-based framework that uses a powerful backend platform named *SageMaker*, a cloud compute service operated by Amazon, to perform the deep learning based predictions [9]. A fully functional graphical user interface was also developed to test the system components as shown in Figure 38.

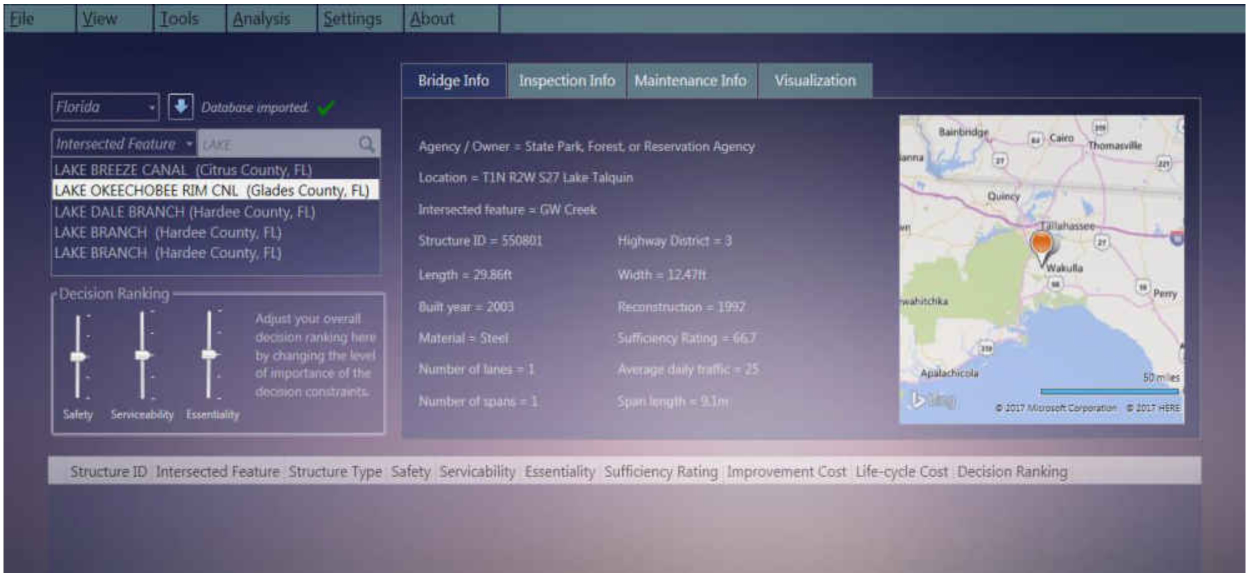


Figure 38: Web-based software implementation of the decision support framework (bridge information retrieval)

The proposed system is aimed to serve as more than a decision-making tool, but also an integration system that can make NDE much more beneficent and effective by retrieving data automatically and transferring it to the widely used bridge management software. The enhanced the decision support system aims to accomplish the following:

- Processing the bridge inventory data of both public and private agencies to retrieve necessary bridge information used in decision support components (e.g. bridge condition, historical data, geolocation)
- Retrieving local element inspection data directly from NDEs such as Infrared Thermography (IRT), Ground Penetration Radar (GPR), laser scanning, remote sensing and drone inspections.
- Element condition assessment based on the quantified damage information and Health index (HI) calculation of the structure. Analysis of historical element

condition states to predict the future condition using a time series forecasting model that estimates the damage growth.

- A novel, adaptive decision ranking implementation for bridge maintenance decisions using bridge appraisals and deep learning-based ranking algorithm.
- Adapting the infrastructure owner's maintenance practice through periodic model updates to fine-tune the decision ranking weights using automatically generated data from users' decision actions.
- Decision tree implementation to produce maintenance/repair strategies with alternative actions and associated cost calculation.
- Damage visualization on realistic 3D bridge model with a timeline feature demonstration both the past and future condition.
- Data exchange and synchronization with infrastructure owner's bridge management software and the NBI database.

Integration of Non-destructive Evaluation Data

Effective utilization of NDE data is important for the decision support system to generate objective and reliably decision strategies. The quantified information from NDE can be utilized by machine learning models to make very accurate predictions. To integrate the NDE data in the system, the data frames are encoded in XML format which is also compatible with FHWA's RABIT™ (Robotics Assisted Bridge Inspection Tool) [10]. As shown in Figure 39, the data frame is composed of header information about the structure and the readings with x and y locations.


```xml
<BridgeInformation>
    <ltbpcmn:StructureID>790174</ltbpcmn:StructureID>
    <ltbpcmn:LTBPBridgeName>SR-430 WB and IWW HALIFAX RIVER</ltbpcmn:LTBPBridgeName>
    <ltbpcmn:State>12</ltbpcmn:State>
    <ltbpcmn:ProtocolName>NDE003</ltbpcmn:ProtocolName>
</BridgeInformation>
<NDE003-TestInfo>
    <TestDate>2011-02-22</TestDate>
    <EquipmentUsed>
        <ltbpcmn:Name>Corrosion Analyzing Instrument - CANIN+</ltbpcmn:Name>
        <ltbpcmn:Model>Wheel Electrodes</ltbpcmn:Model>
        <ltbpcmn:Manufacturer>Proceq</ltbpcmn:Manufacturer>
    </EquipmentUsed>
    <AmbientAirTemperature>75</AmbientAirTemperature>
    <DeckSurfaceTemperature>75</DeckSurfaceTemperature>
            <YLocationUnit>feet</YLocationUnit>
            <XLocationUnit>feet</XLocationUnit>
    <TestSite>Complete bridge deck</TestSite>
    <NDE003-Readings>
        <YLocation>1</YLocation>
        <XLocation>0</XLocation>
        <HCPReading>-38</HCPReading>
    </NDE003-Readings>
```


Figure 39: Example NDE data collected from a bridge deck (Electrical Resistivity).

A standardized format for NDE data is important for successful communication between different technologies. Furthermore, visualizing this data is also important for inspectors to make better decisions. Therefore, a decision support system should also be capable of reading the point data to overlay damage information on a 3D bridge model. In the implemented web-based platform, NDE inspection data can be retrieved by either uploading exported NDE files from IRT, GPR, Ultrasound, LIDAR, UAV etc. or via manual entry of visual inspection reports. The visualization module in the software implementation displays the imported NDE data in a inspection timeline allowing the infrastructure owner investigate the damage condition at different time steps (see Figure 40).

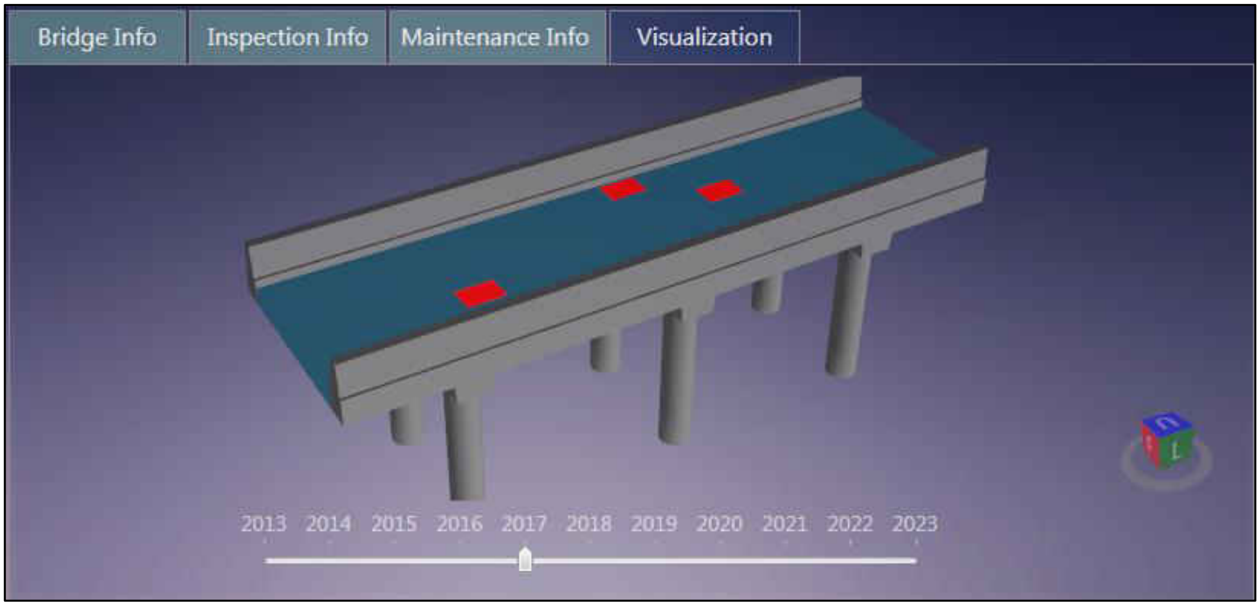


Figure 40: Visualization of NDE data and overlay of damage information

For a rational and quantifiable condition assessment, it is very important to use a bridge management that can generate remediation strategy for different bridge inspection practices [11]. The proposed decision support system can integrate NDE data and calculate element condition states according to major bridge inspection guides (e.g. AASHTO, FHWA, State DOT inspection guides). However, for most inspection guides, the condition state limits need to be quantified for effective utilization of NDE. After NDE data is successfully imported and condition states are determined, the proposed system calculates an important decision ranking parameter, bridge Health Index to determine overall structural health of the bridge. Health Index is calculated for each year using the Equation (8). This equation was first introduced in TRNews article [12].

$$HI = (\sum CEV) \ / \ (\sum TEV) \times 100 \tag{8}$$

The implemented web-based tool uses element condition state limits defined by AASHTO Guide Manual for Bridge Element Inspection [13] and calculates Health Index for each inspection year as shown in Figure 41. In the implemented system, the element inspection data can easily be exported to formats compatible with federally used bridge management software such as AASHTOWare BrM [14] and LTBP InfoBridge [15].

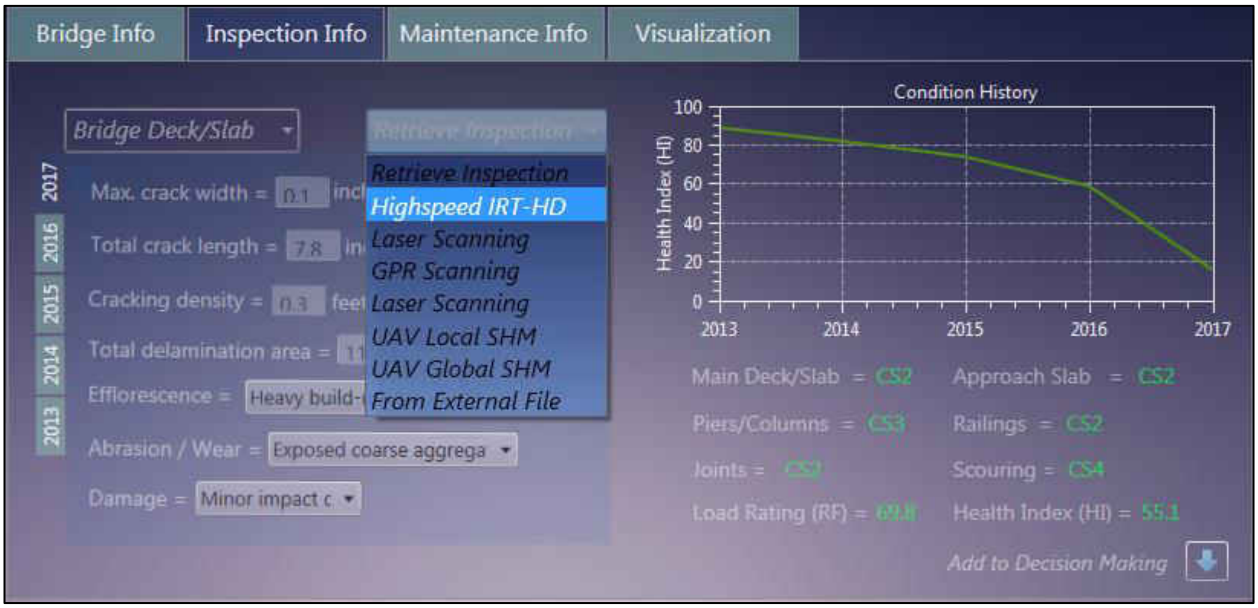


Figure 41: Web-based implementation analyzing historical NDE data.

Deep Learning Based Prediction of Deterioration Growth

Past bridge inspection data along with maintenance/repair information constitutes the basis of predicting future conditions of bridge elements or components. The condition of bridge elements innately possesses significant amount of uncertainty partly due to inaccurately entered or missing inspection records[16]. In the proposed methodology, a deep learning model tackles the uncertainty related future condition prediction problem by using time history prediction on the NDE data. In the proposed methodology, hybrid architecture of Convolutional Neural Networks (CNN) and Long-Term Short Memory (LSTM) models is periodically trained on the historical NDE data. The proposed network architecture is called CNN-LSTM, a deep learning model that fuses CNN and LSTM to predict image-based information in a future time step [17]. LSTMs are a very promising solution to sequence and time series related problems [18]. They can effectively handle time lags between data points as opposed to Recurrent Neural Networks (RNN) [19]. Hence, the data doesn't have to be collected at a fixed time step. As shown in Figure 42 a common LSTM architecture is composed of a cell (the memory part of the LSTM unit) and three regulators, usually called *gates*, maintaining the flow of information through an input, output gate and a forget gate [20]. These gates can learn which data in a sequence is important to keep or throw away. Hereby, it can pass relevant information down the long chain of sequences to make predictions. LSTMs have been successfully used in many real-life applications such as speech recognition [21], forecasting stock prices [22], and estimating cancer growth [23].

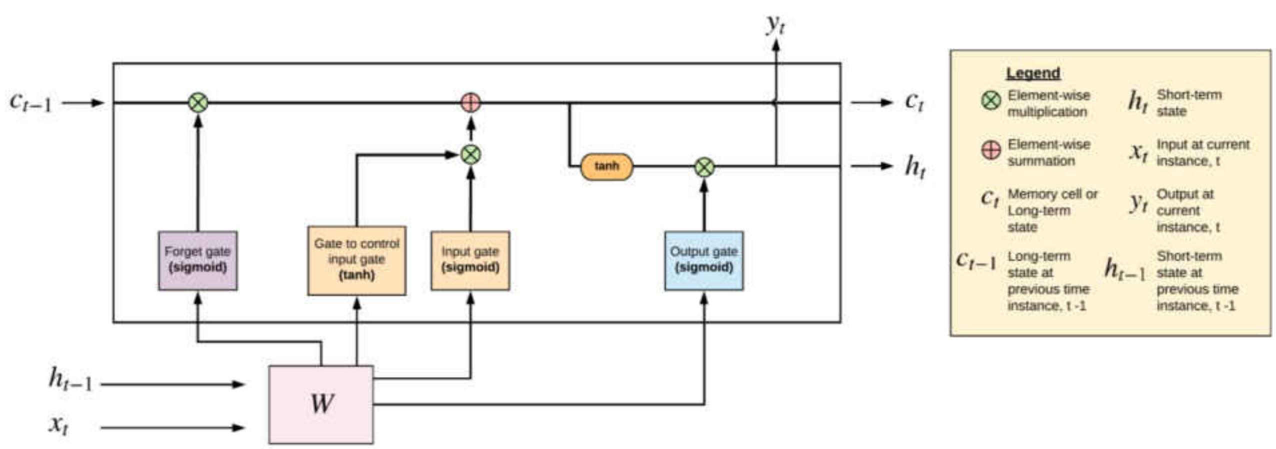


Figure 42: Diagram of a typical LSTM cell [24]

CNNs are on the other hand works very well with image data [25]. They created breakthrough success in image classification tasks [26]. CNN based image analysis of infrastructure damage has been vastly studied in the past [27, 28]. CNN-LSTM, the hybrid architecture of CNN and LSTM networks integrated in the decision support system, inputs the

NDE data as an image format, and passes it through the convolutional layers of CNN, in which the underlying spatial features are extracted and stacked in one-dimensional vector. Then, the LSTM cells receive these feature vectors and make a prediction at a point of time. While CNN branch of the model extracts the spatial relationships, LSTM creates a temporal context. Thus, these hybrid architectures are often called spatial-temporal neural networks. An important advantage of CNN-LSTM is that a time series prediction of spatial data can be performed in a single, end-to-end system instead of using two separate suboptimal systems. In the example shown in Figure 43, crack images belonging to inspections at different years are processed first in the CNN layers. The spatial features of crack deterioration extracted from CNN are then stored in LSTM cells. The output gate of the cell gives the predicted condition of the crack at the requested point of time.

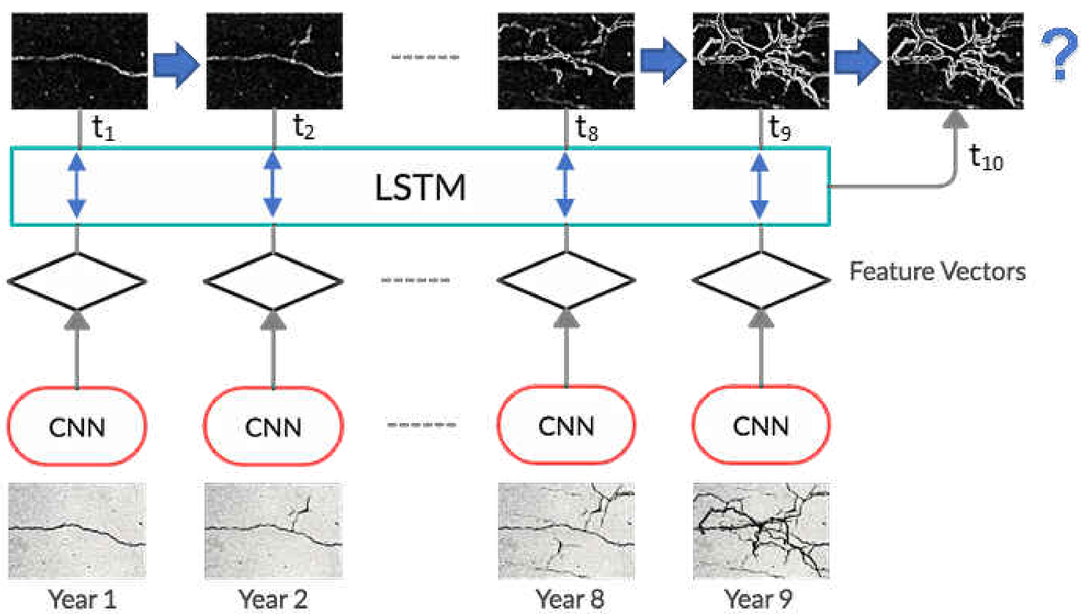


Figure 43: CNN-LSTM model for time history prediction of future damage condition.

The proposed methodology for damage condition prediction was implemented using Keras, a high level deep learning library written in Python [29]. However, the implementation lacked a training data due to unavailability of periodic NDE dataset over a certain period of time. Therefore, the implemented model used pre-trained weights from known data sets from public data sets followed by an updated schema with accurate weights as NDE data is entered to the system and the model is fine-tuned (i.e. transfer learning).

Adaptive Bridge Decision Ranking

For optimal bridge management, it is very important to make network level decisions that take into account all bridges in a transportation network. There is always a limited capital to be spent in maintenance and repair. Therefore, the bridges have to be prioritized based on their importance and the capital needs for improvement. In the proposed decision support methodology, all bridges in a defined transportation network are ranked according to a wide range of criteria such as past inspection history, average daily traffic, number of alternative routes, condition growth rate and life cycle cost. The ranking methodology also accounts for the available maintenance budget and allowed timeline as these external factors will also have significant impact on the decision-making,

Due to complexity of optimizing the decisions within a large number of ranking criteria, a deep learning-based model was used in this study to tackle the multivariate problem in the groupwise scoring. To develop and train the model, TF-Ranking, a scalable deep-learning library recently published by Google TensorFlow team was used [97], [98]. The same ranking model was already deployed in Google's major software platforms such as Google Drive and Gmail

[99]. The proposed TF-Ranking model for bridge prioritization first inputs the multivariate data in *LBSVM* format (Library for Support Vector Machines [100]); then extract features from each decision ranking criteria (i.e. factors affecting the bridge prioritization such as structural condition, repair cost, importance of location, life-cycle cost etc.). From these extracted features, a scoring function is created in the hidden layer of the neural network. During training, the weights of the scoring function are optimized using *Softmax Cross Entropy*, a commonly used listwise loss metric. Finally, the model is served in the *SageMaker* to make predictions from the raw data entries. Figure 44 describes the TF-Ranking model used in the proposed decision support system.

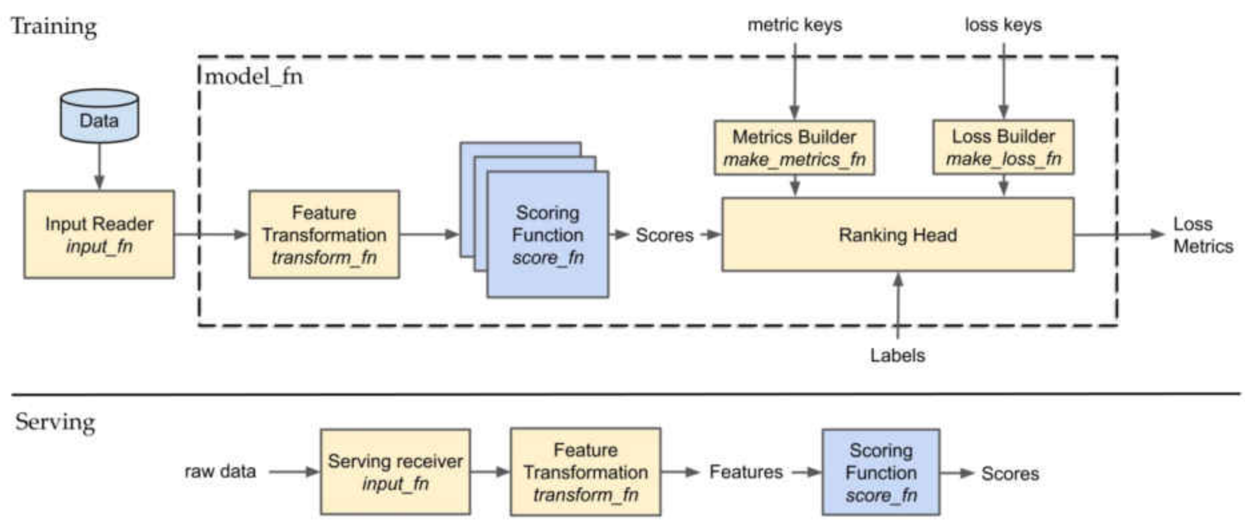


Figure 44: TF-Ranking model to prioritize bridges based on multi-criteria decision factors [98].

The major challenge in deep learning based ranking models for infrastructure decision-making is the unavailability of supervised training data. Such data would require example rankings prepared by a large number of inspectors from many different bridges. Therefore, a different strategy was followed in this study to generate the necessary training data. The TF-Ranking model was initialized by training with an artificial dataset. First, the bridges are scored based on the factors affecting the maintenance decision. The scoring method uses NBI's sufficiency rating as basis; but extends the ranking to include also bridge life cycle cost, current health index calculated from NDE, future predicted health index, and available maintenance/repair budget. The scored bridges are used as initial training data for TF-Ranking model. However, the model will not learn the covariate features in the bridge prioritization since the scoring was made independently for bridges. Therefore, the model needs to be fine-tuned via incremental learning. As inspectors use the decision support system, the inspector's alterations in the prioritizations are used as fine-tuning data. The system will then gradually improve and adapt the infrastructure owner's practice.

The sufficiency rating in NBI's bridge appraisals receives input from local and global assessments as well as some additional parameters [43]. Similarly, the appraisals of bridges in this study are carried out by scoring them in three categories: Safety, serviceability and essentiality. After the scores in these categories are summed, special reductions are made. The resultant score will give the sufficiency rating that could be used for ranking bridges for infrastructure management (priority ranking if resultant score is subtracted from 100). First structural adequacy score is calculated by subtracting the score reductions from overall condition rating and load capacity. The attained RF value from load rating test is input in Equation (9) to find the safety score.

$$S_1 = 55 - (32.4 - RF)^{1.5} \times 0.3254 - CR \tag{9}$$

S_1 cannot be less than zero and larger than 55. CR indicates for lowest condition rating of the bridge components. The equivalent condition rating in the scale of 0-100 is found as discussed by Sobanjo (2008). The study proposes a translation from element condition ratings to NBI's component ratings [44]. Thus, the value of CR is determined for varied conditions as below:

- Critical condition and worse (<2) → CR = 55,
- Serious condition (3) → CR = 40,
- Poor condition (4) → CR = 25,
- Fair condition (5) → CR = 10.

Serviceability score in Equation (10) is calculated based the geometry of the structure, deck condition, structural evaluation, average daily traffic and structure type.

$$S_2 = 30 - [SR + RS + VC] \tag{10}$$

S_2 cannot be less than zero and larger than 30. SR indicates structural rating in which the rating scores from deck condition, structural evaluation, deck geometry, under-clearances, waterway adequacy and approach road alignment are summed. RS is determined based on roadway sufficiency, which is calculated using average daily traffic (ADT) and road width. Finally, vertical clearance (VC) is another rating parameter for serviceability.

Bridge importance score S_3 is calculated based on the average daily traffic value, detour length, also S_1 and S_2 as in Equation (11).

$$S_3 = 15 - \left\{ 15 \left[\frac{ADT \times DL}{320{,}000 \times \frac{S_1 + S_2}{85}} \right] + 2 \right\} \tag{11}$$

Bridge importance score S_3 cannot be less than 0 and more than 15. ADT stands for average daily traffic and DL is the detour length. In addition, there is also special reduction score

S_4 which is calculated based detour length, structure type and traffic safety features of the bridge. Decision Ranking (DR) score extends the NBI's sufficiency rating in the way that it takes life-cycle cost (LCC) into account and also offers flexibility to infrastructure owners' decision practice. The resultant score is calculated by summing individual scores that are adjusted by the weight factors w_1, w_2, w_3 and w_4 and multiplying it by the bridge value index (VI). The default values of the weight factors are equal to 1.0. w_1 is calibrated based on the current and predicted future Health Index. However, the infrastructure owners can adjust them according to their own decision consideration (political pressure, higher serviceability concern etc.) and the decision support system will create a ranking score function and will optimize these parameters accordingly. The decision ranking is calculated as in Equation (12).

$$DR = (w_1 S_1 + w_2 S_2 + w_3 S_3 - w_4 S_4) \times VI \tag{12}$$

Using a life cycle cost analysis, it is possible to comprehensively evaluate the total generated environmental impact for a product and understand the trade-offs in impacts between different periods in the product's life cycle [45]. Catbas et al. (2008) investigated structural health monitoring approaches for life cycle management of bridges [101]. The life cycle cost analysis of bridges is explained in detail in the National Cooperative Highway Research Program (NHRP) report by Transportation Research Board [102]. Mohammadi et al. simplified the bridge life cycle cost (BLCC) and used a single parameter to quantify the bridge decision-making process in an optimal scheduling scheme [103]. Three major elements constitute the life cycle cost: (1) bridge condition rating, (2) costs associated with various bridge works, and (3) bridge service life expectancy. Equation (13) shows the calculation of life cycle cost of bridges.

$$VI = r \times t/c \tag{13}$$

In the Equation (13), VI is the bridge value index, r is the condition rating, t indicates bridge service life expectancy and c stands for maintenance cost. The calculated life cycle cost is an important decision parameter because infrastructure owners often prefer building a new bridge instead of repairing the old one in case the bridge repair cost is very high. Therefore, the life cycle cost was also included in the decision ranking.

The artificial data for initial training was created from the ranking methodology discussed throughout this section. An example data was created from Florida's NBI bridges. The data was split into different congressional districts to represent the separate transportation networks. The training data has the following features columns: structural adequacy, serviceability, bridge importance, bridge value index and available fund. The values for district available funds were obtained from the FDOT 2019 Work Program Instructions report, under District Bridge Repair & Rehabilitation Funds [104]. A small sample from the generated data is show in Table 8.

Table 8: A small sample from the generated data showing bridges from different districts

District	Intersected Feature	Type	Structural Adequacy	Serviceability	Bridge Importance	Bridge Value Index	Available Fund	Decision Ranking
FL-2	Brown Creek	Prestressed	37%	20%	10%	14.8	$18.3M	40
FL-4	Palm Avenue	Prestressed	45%	25%	12%	15.4	$15.9M	42
FL-5	Lake Jesup	Prestressed	54%	29%	18%	20.1	$9.2M	38
FL-1	Gum Creek	Concrete	55%	30%	15%	17.8	$10.3M	45
⋮								

The TF-Ranking model was trained on the generated dataset only for 15,000 steps to initialize the model without causing overfitting. The model will gradually improve in the deployed decision support system and incrementally learn the valuable deep features of the ranking methodology as the infrastructure owner make changes in the ranking orders. The training performance of the model is shown in Figure 45. Training Loss indicates how successfully the model converged with the training batches of the data at each step. The evaluation metric that was used in the model is Normalized Discounted Cumulative Gain (NDCG), a commonly used metric to measure ranking quality [105].

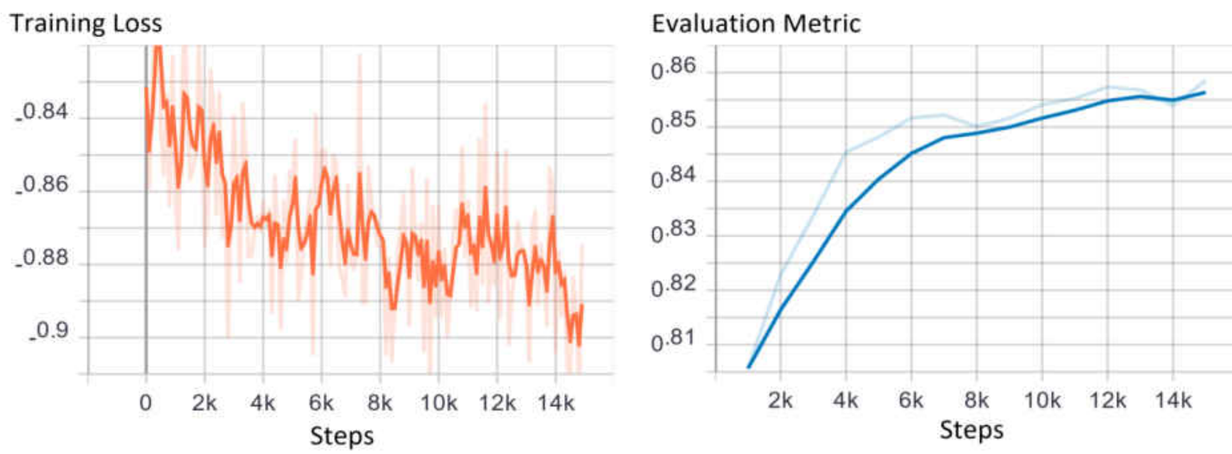


Figure 45: Training results of the TF-Ranking model using the generated dataset

Since the ranking data used in the training was generated from an analytically defined mythology, the TF-Ranking model easily fits the data during training; yet will not perform well on real-life ranking without a fine-tuning. As real data arrives, the model will improve incrementally.

Decision Strategy Generation

Maintenance, repair and rehabilitation of deteriorating bridge structures may require very costly remediation actions. Advanced decision support system aim to reduce remediation costs by preventing the costs that are associated with subjectivity of the decision making [92]. In the proposed decision support methodology, a multi-criteria maintenance strategy was used to generate the optimal maintenance actions that are specific to each infrasture owner's maintenance practice. Decesion trees are created for each single deterioration mode with multicriteria optimization. The multi-criteria selection strategy is the simplified implementation of the methodology introduced by S. A. Dabous and S. Alkas (2008). In the decision algorithm, multiple criteria are connected to four main action categories: Replacement, major rehabilitation, minor rehabilitation, routine maintenance [106]. The maintenance actions are determined after checking the associated criteria as described in Figure 46. The authors also introduced a ranking methodology that provides score values for each criterion based on based on the rehabilitation strategy options.

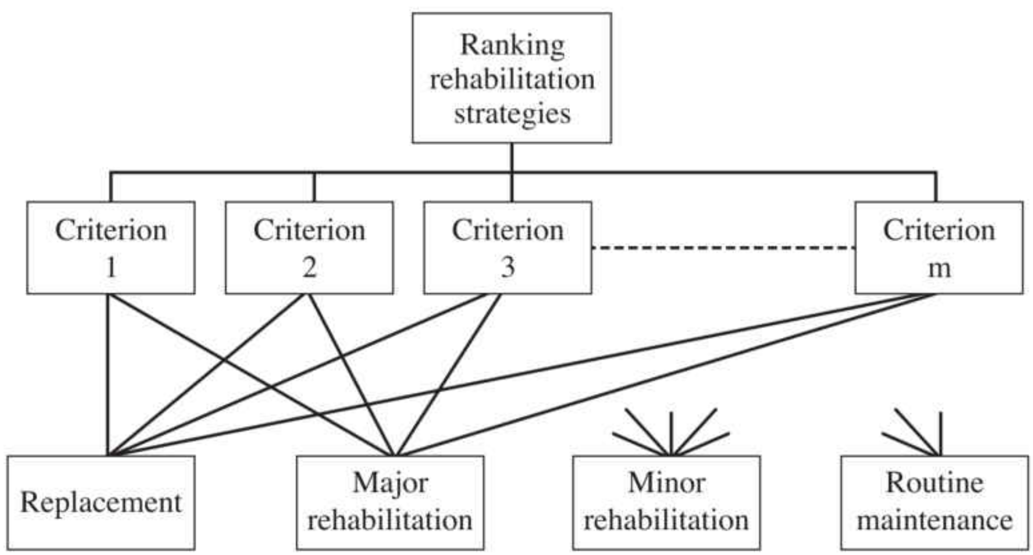


Figure 46: Multi-criteria selections of maintenance actions for bridge decks [107].

This study integrates the described multi-critera decision selection in more simplified way by taking into account the decesion rankings directly predicted from the deep learning-based model. For each rehabilitation strategy, the decision ranking score is updated and top strategies are suggested to the infrastructure owner. The decision selection was implemented inside the decision trees; allowing each deterioration mode to be analyzed inside decision loops. Once the strategy is selected and the funding availabilty is approved, detailed maintenance actions are automatically generated. An example decision action tree for bridge deck concrete cracking according to maintenance practice of Florida Department of Transportation (FDOT) is given in Figure 47**Error! Reference source not found.**. The maintenance practice is based on the FDOT's Bridge Maintenance and Repair Handbook [108]. In the decision tree, first the condition of the of the cracking is determined, then the availability of funding is checked through the decision ranking. Until bridge reaches the target priority, the associated maintenance actions are awaited. Once the funding becomes available (i.e. bridge falls inside the target ranking), conditional maintenance actions are suggested to the infrastructure owner. For instance, a minor cracking damage on the deck surface can be repaired using liquid sealer if there are many cracks. On the other hand, a moderate crack should be repaired with presure injection if the sealing was not previously made.

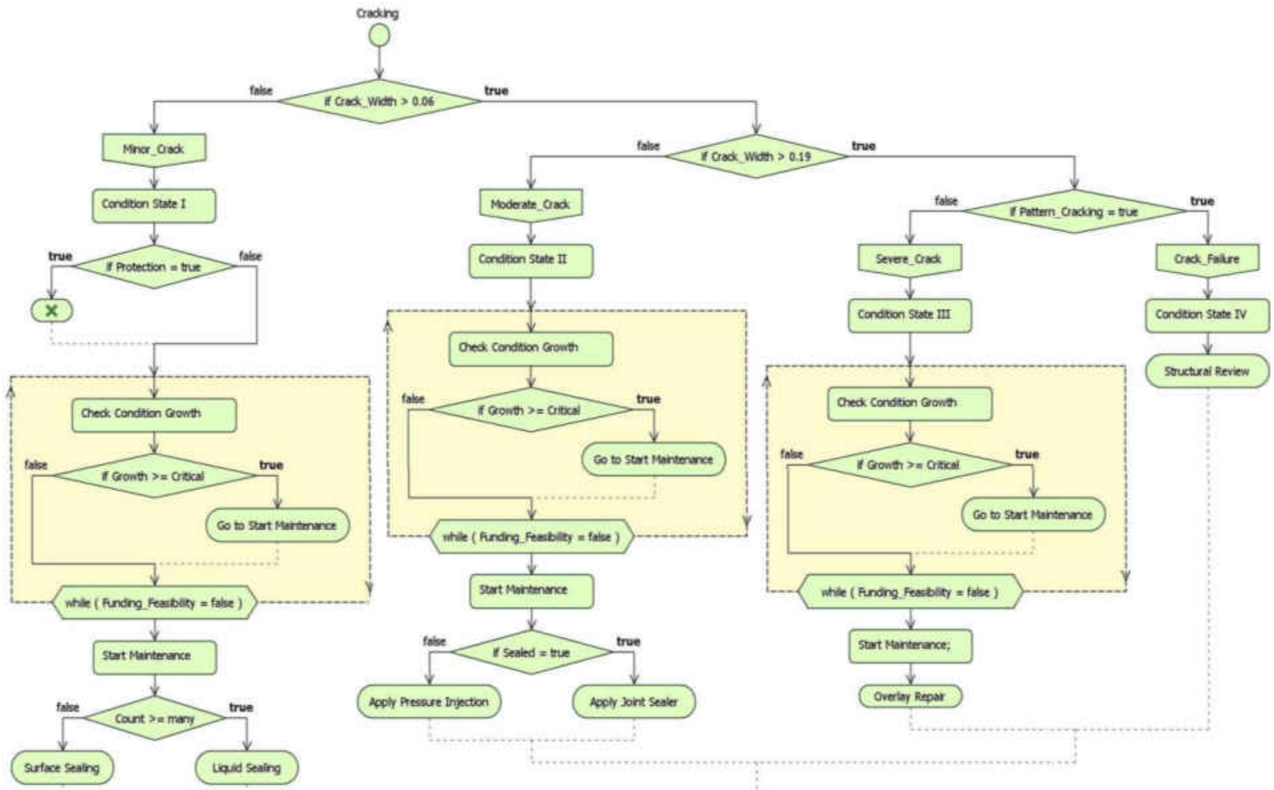
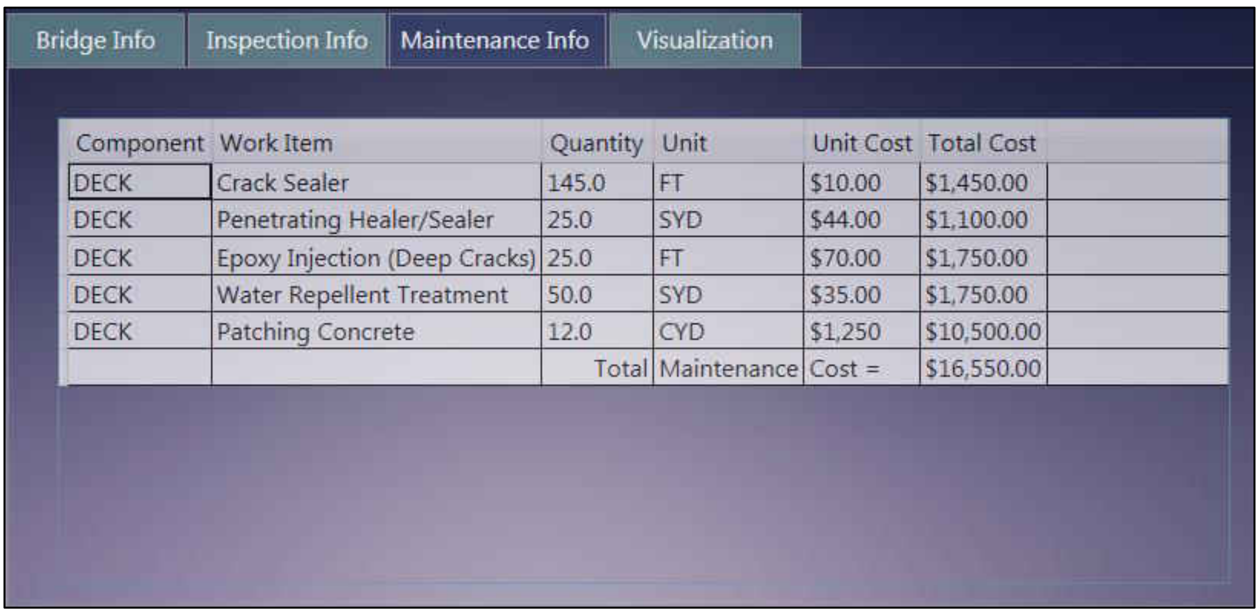


Figure 47: Example maintenance action tree for concrete bridge deck cracking based on FDOT's practice.

The maintenance strategy generation was also integrated in the web-based system implementation as show in Figure 48. The system automatically generates suggested maintenance/repair strategies based on the predefined repair pricing input and the decision tree algorithm used for each deterioration mode. However, it is possible to generate optional strategies based on different decision criteria selection. The system will modify the suggested maintenance actions. The user can also manually change a particular action in the maintenance suggestions and update the cost calculation according to the user-defined unit price list. The maintenance suggestions and the corresponding unit costs were obtained from NCHRP Final Report 668, Framework for a National Database System for Maintenance Actions on Highway

Bridges [109]. In the shown example, the NDE data belonging to high-speed IRT deck scanning was imported to the system and the suggested maintenance actions for cracks and spalling were automatically generated using the decision trees (it was assumed that this bridge was assumed high priority in the decision ranking). Then, a cost summary is shown to the user indicating the individual cost items for each maintenance action. In this particular NDE inspection, repair for spalling was shown as the major cost item.



Component	Work Item	Quantity	Unit	Unit Cost	Total Cost
DECK	Crack Sealer	145.0	FT	$10.00	$1,450.00
DECK	Penetrating Healer/Sealer	25.0	SYD	$44.00	$1,100.00
DECK	Epoxy Injection (Deep Cracks)	25.0	FT	$70.00	$1,750.00
DECK	Water Repellent Treatment	50.0	SYD	$35.00	$1,750.00
DECK	Patching Concrete	12.0	CYD	$1,250	$10,500.00
		Total	Maintenance	Cost =	$16,550.00

Figure 48: Maintenance-cost analysis module in the software implementation of the decision support system.

CHAPTER SIX: FINAL DISCUSSION AND CONCLUDING REMARKS

This dissertation study essentially searched for an answer to the following question: How can we establish a complete system of methodologies to enable a visual civil infrastructure assessment that is quantified, relatively accurate, faster, lower-cost, and automated, but still benefits the inspection engineer's judgement? The proposed approach extensively discussed in this study aimed to integrate and demonstrate novel deep learning detection and segmentation algorithms into mixed reality system by which a bridge inspector, for example, can benefit from this system during his/her routine inspection/assessment tasks. Using this system, the inspector can analyze a concrete crack or spalling in-real time and calculate its condition state without needing to perform any manual measurement. Furthermore, this dissertation also examined novel methods for effective use of collected inspection data in optimal decision making.

An Overview of the Dissertation Report

A very comprehensive scope of research was conducted in this dissertation study. The research involved interdisciplinary work in Structural Health Monitoring, Data Science, Computer Vision, and Human-Computer Interaction.

In Chapter 1, necessary introductory information and literature review about the deep learning techniques, different immersive technologies such as augmented, virtual and mixed realities, and bridge management practices were given.

In Chapter 2, current practices in bridge inspections in the U.S was discussed in detail. The structural health monitoring at local and global level was outlined; both conventional methods and non-destructive evaluation techniques were described. Lastly, a case study conducted by the CITRS researchers takes place in this chapter.

In Chapter 3, a comprehensive deep learning methodology for analysis of concrete bridge defects was explained thoroughly. The methods described in this chapter explained how a framework for collective human-AI intelligence could be created and how it could outperform the conventional or fully automated concrete inspections. The described human-centered AI asks only minimal input from the human inspector (e.g. modifying the prediction threshold, correcting the detection/segmentation boundaries) and gets its predictions verified before finalizing a damage assessment task. This kind of a collaboration layer between human expert and AI is unique approach of this study. The deep learning models employed in the proposed technique could detect a concrete defect in-real time on a mobile chipset and quantify it by performing pixel wise segmentation. To train these model, a sufficiently large database of concrete defect image database was gathered from various sources and annotated for model training. For damage detection, a single shot detector built on MobileNetV2 architecture (*SSDLite*) and for damage segmentation, an adapted version of SegNet architecture were trained using UCF's Newton cluster computers. Finally, the trained models were evaluated on a test database and inference speeds were calculated on a mobile device.

In Chapter 4, the methods discussed in the previous chapter was in integrated in mixed reality system. In the described system, the inspector uses a holographic headset that satisfies certain hardware requirements (e.g. AI optimized chipset, high resolution camera and depth sensor) during routine inspection of infrastructure. With the help of mixed reality integration, the human inspector and AI can cooperate efficiently. While the inspector performs routine inspection tasks, the AI system performs the following: Continuously guides the inspector by showing possible defect locations (real-time detection); analyzes the verified defect locations and predicts damage regions (segmentation); estimates dimensional properties of the damage regions (e.g. crack length and spall area) and performs condition assessment (characterization); and creates visual mapping of defects in 3-dimensional space (positioning).

Finally, Chapter 6 introduced an enhanced decision support system that benefits from novel approaches at multiple levels. When the NDE data is integrated to the system, a powerful deep neural network, CNN-LSTM model predicted the future state of each concrete defect based on the historical NDE input. Another novel approach used in the system is that TF-Ranking, a deep learning based ranking model, prioritizes the bridges for maintenance/repair based on a large variety of factors including bridge assessment history, bridge importance for state traffic, structural reliability, serviceability, repair cost, life cycle cost, repair time and funding availability. These factors were grouped under structural adequacy, serviceability, importance and funding. The ranking system was developed in parallel with the NBI's sufficiency rating, yet it allowed infrastructure owners to calibrate certain weight factors also entails consideration of additional decision matrices such bridges' life cycle cost and predicted repair/maintenance cost etc., thereby gives more control over the decision. The TF-Ranking model quickly adapted the bridge maintenance practice of the infrastructure owner as new training data for fine-tuning was generated automatically from decision makers' priority adjustments.

Important Findings and Conclusions

In addition to important scientific contributions, this interdisciplinary dissertation study offers significant contributions to infrastructure inspection, maintenance, management practice, and safety for the transportation agencies in the US as well as other countries. The dissertation research generated the following important conclusions.

- Current scientific approaches have employed various learning-based methods for automatic detection of concrete defects while replacing human involvement in the process. However, the developed method aimed to merge engineer/inspector's expertise with AI assistance using a human-centric machine vision approach, thus yielding more reliable civil infrastructure visual assessment practice.
- In deep learning-based models, the availability of training data is the most critical aspect of developing a reliable system with good accuracy in recognition. Yet, in infrastructure assessment, creating a large image dataset is particularly a challenging task. The proposed method therefore used an advanced data augmentation technique to generate synthetically sufficient amount of crack and spall images from the available image data.
- The AI system follows a semi-supervised learning approach and consistently improves itself with use of verified detection and segmentation data in re-training. The use of semi-supervised learning addresses successfully the problems of small data in AI training particularly encountered in damage detection applications where a comprehensive, publicly available image dataset is unavailable.

- The attention guide approach (sequential detection and segmentation) yielded significant reduction in the computational cost of the segmentation operation since only a region of interest is used while other comparable models (e.g. MaskRCNN) performs segmentation on the entire image and performs localization in parallel. The sequential model also significantly improved the segmentation performance of concrete defects.
- Mixed reality system is an ideal environment to facilitate human – computer interaction. It enables the human-centered AI to interact with the inspector instead of completely replacing the human involvement during the inspection. This collective work will lead to quantified assessment, reduced labor time while also ensuring human verified results.
- Many advanced studies in the bridge management although provide good insight about optimal management of bridges, have unfortunately no real-life implementation. This dissertation study uses effectively some of these insights yet updates the overall methodology with recent advancements in Machine Learning. Furthermore, the proposed system was successfully implemented in a web-based platform that uses Amazon SageMaker to perform deep learning predictions in the backend. The implementation has a robust, functional user interface and a powerful visualization module that helps bridge inspectors gain more insight about the condition of the bridge. The visualization overlays NDE data on a 3D bridge model and even demonstrates the predicted future condition of the damage visually for a selected year.

- Lastly, this study aims to increase the adoption rate of NDE technologies by making the NDE output functional and useful for decision making. Within an acceptable uncertainty range, the proposed system shows that use of NDEs has the potential to become routine inspection practice when integrated with bridge management along with the benefits such as reduced cost and time of inspection. Although the methodology was investigated for bridges, it is applicable to all civil infrastructures.

Plans for Future Work

The enhanced infrastructure assessment methodology that uses artificial intelligence and mixed reality can be expanded in many ways in a future work of this dissertation study:

- The current scope of the study aimed a generic approach for infrastructure inspections even though the target damage types were only spalls and cracks. In the future work, the number of damage classes will be expanded to include rusting, efflorescence and sub-concrete defect (multi-channel input with infrared data). It is going to be possible to use the methods for steel and composite structures.
- A multichannel analysis method will be investigated in order fuse multiple sources of data (i.e. imagery data and infrared thermography). This new method will bring more capabilities such as robustly locating and quantifying subconcrete delamination and steel corrosion. Secondly, more defect types will be trained for the AI system.

- As these AI embedded MR devices are used in the field and more data is gathered from outside sources; the prediction accuracies can be improved considerably. Furthermore, to enable the system to evolve even when the devices are used offline. A unique incremental learning technique will be investigated and on-device training will be implemented.
- The methods used in the decision support system will be improved as real-life inspection data is collected and entered to the web-based implementation of the system. With availability of periodic NDE data and real-life decision-making information, the proposed deep learning models will be retrained, and performance evaluations of these models will be studied extensively.

LIST OF REFERENCES

[1] U.S. Department of Transportation Federal Highway Administation, “Deficient Bridges by Highway System,” 2017.

[2] F. Catbas, S. Ciloglu, A. Aktan, F. N. Catbas, S. K. Ciloglu{, and A. E. Aktan{, “Strategies for load rating of infrastructure populations: a case study on T-beam bridges,” 2017.

[3] M. C. Brown, J. P. Gomez, M. L. Hammer, and J. M. Hooks, “Long-Term Bridge Performance High Priority Bridge Performance Issues,” McLean, VA, 2014.

[4] A. H. Behzadan, S. Dong, and V. R. Kamat, “Augmented reality visualization: A review of civil infrastructure system applications,” *Adv. Eng. Informatics*, vol. 29, no. 2, pp. 252–267, 2015.

[5] M. Mihelj, D. Novak, and S. Beguš, *Virtual Reality Technology and Applications*, vol. 68, no. 11. 2014.

[6] R. Azuma, R. Behringer, S. Feiner, S. Julier, and B. Macintyre, “Recent Advances in Augmented Reality,” *IEEE Comput. Graph. Appl.*, vol. 2011, no. December, pp. 1–27, 2001.

[7] I. E. Sutherland, “The ultimate display,” in *Proceedings of the Congress of the Internation Federation of Information Processing (IFIP)*, 1965, pp. 506–508.

[8] P. Milgram and F. Kishino, “Taxonomy of mixed reality visual displays,” *IEICE Trans. Inf. Syst.*, vol. E77-D, no. 12, pp. 1321–1329, 1994.

[9] Windows Mixed Reality, “What is Mixed Reality ?,” 2018. .

[10] C. Coutrix and L. Nigay, “Mixed reality: A model of mixed interaction,” *Proc. Work. Conf. Adv. Vis. interfaces - AVI ’06*, pp. 43–50, 2006.

[11] J. LaViola, E. Kruijff, D. Bowman, and I. Poupyrev, *3D User Interfaces: Theory and Practice, Second Edition*. 2017.

[12] V. R. Kamat and S. El-Tawil, "Evaluation of Augmented Reality for Rapid Assessment of Earthquake-Induced Building Damage.," *J. Comput. Civ. Eng.*, vol. 21, no. 5, pp. 303–310, 2007.

[13] A. H. Behzadan and V. R. Kamat, "Georeferenced Registration of Construction Graphics in Mobile Outdoor Augmented Reality," *J. Comput. Civ. Eng.*, vol. 21, no. 4, pp. 247–258, 2007.

[14] H. Bae, M. Golparvar-Fard, and J. White, "High-precision vision-based mobile augmented reality system for context-aware architectural, engineering, construction and facility management (AEC/FM) applications," *Vis. Eng.*, vol. 1, no. 1, p. 3, Jun. 2013.

[15] B. Ioannis, "Mixed reality constructs a new frontier for maintaining the built environment," *Proc. Inst. Civ. Eng. - Civ. Eng.*, vol. 170, no. 2, p. 53, 2017.

[16] F. Moreu, B. Bleck, S. Vemuganti, D. Rogers, and D. Mascarenas, "Augmented reality tools for enhanced structural inspection," in *Structural Health Monitoring 2017: Real-Time Material State Awareness and Data-Driven Safety Assurance - Proceedings of the 11th International Workshop on Structural Health Monitoring, IWSHM 2017*, 2017, vol. 2, pp. 3124–3130.

[17] Microsoft, "The leader in Mixed Reality Technology | HoloLens," *Microsoft*, 2017. .

[18] C. Koch, K. Georgieva, V. Kasireddy, B. Akinci, and P. Fieguth, "A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure," *Adv. Eng. Informatics*, vol. 29, no. 2, pp. 196–210, 2015.

[19] S. Yokoyama and T. Matsumoto, "Development of an Automatic Detector of Cracks in

Concrete Using Machine Learning," *Procedia Eng.*, vol. 171, pp. 1250–1255, 2017.

[20] M. R. Jahanshahi and S. F. Masri, "Adaptive vision-based crack detection using 3D scene reconstruction for condition assessment of structures," *Autom. Constr.*, vol. 22, pp. 567–576, 2012.

[21] R. S. Adhikari, O. Moselhi, and A. Bagchi, "Image-based retrieval of concrete crack properties for bridge inspection," *Autom. Constr.*, vol. 39, pp. 180–194, 2014.

[22] S. German, I. Brilakis, and R. Desroches, "Rapid entropy-based detection and properties measurement of concrete spalling with machine vision for post-earthquake safety assessments," *Adv. Eng. Informatics*, vol. 26, no. 4, pp. 846–858, 2012.

[23] R. Zaurin, T. Khuc, F. N. Catbas, and F. Asce, "Hybrid Sensor-Camera Monitoring for Damage Detection: Case Study of a Real Bridge," *J. Bridg. Eng.*, vol. 21, no. 6, pp. 1–27, 2015.

[24] S. Hiasa, F. N. Catbas, M. Matsumoto, and K. Mitani, "Monitoring Concrete Bridge Decks using Infrared Thermography with High Speed Vehicles," *Struct. Monit. Maintenance, An Int. J.*, vol. 3, no. 3, pp. 277–296, 2016.

[25] S. Hiasa, "Investigation of Infrared Thermography for Subsurface Damage Detection of Concrete Structures," Electronic Theses and Dissertations. Paper 5063. <http://stars.library.ucf.edu/etd/5063>, 2016.

[26] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, 2012.

[27] M. Rashidi, B. Samali, and P. Sharafi, "A new model for bridge management: Part A: condition assessment and priority ranking of bridges," *Aust. J. Civ. Eng.*, vol. 14, no. 1, pp.

35–45, 2016.

[28] S. Hiasa, F. N. Catbas, M. Matsumoto, and K. Mitani, "Considerations and Issues in the Utilization of Infrared Thermography for Concrete Bridge Inspection at Normal Driving Speeds," *J. Bridg. Eng.*, vol. in press, 2017.

[29] U.S. Department of Transportation Federal Highway Administration and Federal Transit Administration, FHWA, and FTA, *2013 Status of the Nation's Highways, Bridges, and Transit: Conditions & Performance,*. 2014, pp. 1–482.

[30] FHWA, "Tables of Frequently Requested NBI Information," 2017. .

[31] ASCE, "2017 ASCE Infrastructure Report Card," 2017.

[32] F. N. Catbas, S. Hiasa, C. Dong, Y. Pan, O. Celik, and E. Karaaslan, "Comprehensive Structural Health Monitoring at Local and Global Level with Vision-based Technologies," *26th ASNT Research Symposium*. 2017.

[33] S. Hiasa, R. Birgul, and F. N. Catbas, "Investigation of Effective Utilization of Infrared Thermography (IRT) through Advanced Finite Element Modeling," *Constr. Build. Mater.*, vol. in press, 2017.

[34] G. Hearn, *Bridge Inspection Practices*, no. January 2007. 2016.

[35] F. N. Catbas *et al.*, "Long term bridge maintenance monitoring demonstration on a movable bridge," *Final Rep. Res. Proj.*, no. June, p. 186, 2010.

[36] F. N. Catbas, T. Kijewski-Correa, T. Kijewski-Correa, and A. M. Asce, "Structural Identification of Constructed Systems: Collective Effort toward an Integrated Approach That Reduces Barriers to Adoption."

[37] O. Celik, "Load Estimation, Structural Identification and Human Comfort Assessment of Flexible Structures," University of Central Florida, 2017.

[38] M. C. Brown, J. P. Gomez, M. L. Hammer, and J. M. Hooks, "Long-Term Bridge Performance High Priority Bridge Performance Issues," *Tech Brief*, no. October, 2014.

[39] S. Hiasa, F. N. Catbas, M. Matsumoto, and K. Mitani, "Feasibility of Infrared Thermography for Bridge Deck Inspection without Lane Closure, and The Uncertainties - Preliminary Report -," in *TRB 95th Annual Meeting*, 2016.

[40] F. N. Catbas, S. Hiasa, T. Khuc, M. Matsumoto, and K. Mitani, "Development , Implementation and Evaluation of Image-based Technologies for Civil Infrastructure Systems," 2015, p. 262.

[41] Z. Yin and Y. Li, "Intelligent decision support system for bridge monitoring," *2010 Int. Conf. Mach. Vis. Human-Machine Interface, MVHI 2010*, pp. 491–494, 2010.

[42] H. Quintela, M. F. Santos, and P. Cortez, "Real-Time Intelligent Decision Support System for Bridges Structures Behavior Prediction," in *Progress in Artificial Intelligence: 13th Portuguese Conference on Aritficial Intelligence, EPIA 2007, Workshops: GAIW, AIASTS, ALEA, AMITA, BAOSW, BI, CMBSB, IROBOT, MASTA, STCS, and TEMA, Guimar{ã}es, Portugal, December 3-7, 2007. Proceedings*, J. Neves, M. F. Santos, and J. M. Machado, Eds. Berlin, Heidelberg, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 124–132.

[43] B. A. Bektas, A. Carriquiry, and O. Smadi, "Using Classification Trees for Predicting National Bridge Inventory Condition Ratings," *J. Infrastruct. Syst.*, vol. 19, no. 4, pp. 425–433, 2014.

[44] J. O. Sobanjo, P. D. Thompson, R. Kerr, and T. R. Board, "Element-to-Component Translation of Bridge Condition Ratings," 2008, p. 25p.

[45] E. Karaaslan, Y. Zhao, and O. Tatari, "Comparative life cycle assessment of sport utility vehicles with different fuel options," *International Journal of Life Cycle Assessment*, pp.

1–15, 2017.

[46] S. Chase and H. Ghasemi, "Implications of the long term bridge performance program for life cycle costing in the United States," *Struct. Infrastruct. Eng.*, vol. 5, no. 1, pp. 3–10, Feb. 2009.

[47] J. Hooks and D. M. Frangopol, "LTBP bridge performance primer," *U.S. Bridg. Cond. Long-Term Bridg. Perform. A Prim.*, pp. 21–89, 2014.

[48] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The Pascal Visual Object Classes (VOC) Challenge," *Int. J. Comput. Vis.*, vol. 88, no. 2, pp. 303–338, Jun. 2010.

[49] P. Prasanna *et al.*, "Automated Crack Detection on Concrete Bridges," *IEEE Trans. Autom. Sci. Eng.*, vol. 13, no. 2, pp. 591–599, 2016.

[50] H. Maeda, Y. Sekimoto, T. Seto, T. Kashiyama, and H. Omata, "Road Damage Detection Using Deep Neural Networks with Images Captured Through a Smartphone," pp. 4–6, 2015.

[51] L. Yang, B. Li, W. Li, Z. Liu, G. Yang, and J. Xiao, "A robotic system towards concrete structure spalling and crack database," *2017 IEEE International Conference on Robotics and Biomimetics (ROBIO), Robotics and Biomimetics (ROBIO), 2017 IEEE International Conference on*. IEEE, p. 1276, 2017.

[52] UCF Advanced Research Computing Center, "Newton Visualization Cluster." Orlando, 2014.

[53] M. Abadi *et al.*, "TensorFlow: A System for Large-Scale Machine Learning," *12th USENIX Symp. Oper. Syst. Des. Implement. (OSDI '16)*, pp. 265–284, 2016.

[54] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A

simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, 2014.

[55] A. Geron, *Hands-On Machine Learing With Scikit-Learn & Tensor Flow*. 2017.

[56] S. De, A. Yadav, D. Jacobs, and T. Goldstein, "Automated Inference with Adaptive Batches," *Proc. 20th Int. Conf. Artif. Intell. Stat.*, 2017.

[57] N. Torosdagli, D. K. Liberton, P. Verma, M. Sincan, J. S. Lee, and U. Bagci, "Deep Geodesic Learning for Segmentation and Anatomical Landmarking," *IEEE Trans. Med. Imaging*, 2019.

[58] M. Andrychowicz *et al.*, "Learning to learn by gradient descent by gradient descent," in *Advances in Neural Information Processing Systems 29*, D. D. Lee, M. Sugiyama, U. V Luxburg, I. Guyon, and R. Garnett, Eds. Curran Associates, Inc., 2016, pp. 3981–3989.

[59] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," Dec. 2014.

[60] J. Schmidhuber, "Deep Learning in neural networks: An overview," *Neural Networks*. 2015.

[61] W. Liu *et al.*, "SSD: Single shot multibox detector," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9905 LNCS, pp. 21–37, 2016.

[62] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," 2018.

[63] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *CoRR*, vol. abs/1409.1, 2014.

[64] R. LaLonde and U. Bagci, "Capsules for Object Segmentation," no. Midl, pp. 1–9, 2018.

[65] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," *IEEE Trans. Pattern Anal. Mach.*

Intell., vol. 39, no. 12, pp. 2481–2495, 2017.

[66] M. Hill, "Overview of human-computer collaboration," vol. 8, no. June, pp. 67–81, 1995.

[67] P. Jaccard, "The Distribution of the Flora in the Alpine Zone," *New Phytol.*, 1912.

[68] L. R. Dice, "Measures of the Amount of Ecologic Association Between Species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.

[69] D. Xie, L. Zhang, and L. Bai, "Deep learning in visual computing and signal processing," pp. 1–18, 2017.

[70] S. Chen, L. Liang, W. Liang, and H. Foroosh, "3D Pose Tracking with Multitemplate Warping and SIFT Correspondences," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 11, pp. 2043–2055, 2016.

[71] T. Parisi, *Programming 3D Applications with HTML5 and WebGL*, no. 1. 2014.

[72] H. Kato, M. Billinghurst, I. Poupyrev, K. Imamoto, and K. Tachibana, "Virtual object manipulation on a table-top AR environment," in *Proceedings - IEEE and ACM International Symposium on Augmented Reality, ISAR 2000*, 2000.

[73] G. Simon, A. W. Fitzgibbon, and A. Zisserman, "Markerless tracking using planar structures in the scene," in *Proceedings - IEEE and ACM International Symposium on Augmented Reality, ISAR 2000*, 2000.

[74] V. Ferrari, T. Tuytelaars, and L. Van Gool, "Markerless augmented reality with a real-time affine region tracker," in *Proceedings - IEEE and ACM International Symposium on Augmented Reality, ISAR 2001*, 2001.

[75] H. Strasdat, J. M. M. Montiel, and A. J. Davison, "Visual SLAM: Why filter?," *Image Vis. Comput.*, 2012.

[76] Q. H. Gao, T. R. Wan, W. Tang, L. Chen, and K. B. Zhang, "An Improved Augmented

Reality Registration Method Based on Visual SLAM," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2017.

[77] C. H. Hsieh and J. Der Lee, "Markerless augmented reality via stereo video see-through head-mounted display device," *Math. Probl. Eng.*, 2015.

[78] VisionStar Information Technology, "EasyAR." Shanghai, 2017.

[79] Unity Technology, "Unity 3D," *Unity Technology*, 2018. .

[80] Arthur Juliani, "Introducing: Unity Machine Learning Agents – Unity Blog," *Unity*, 2017. [Online]. Available: https://blogs.unity3d.com/2017/09/19/introducing-unity-machine-learning-agents/.

[81] E. Marchand, H. Uchiyama, F. Spindler, E. Marchand, H. Uchiyama, and F. Spindler, "Pose estimation for augmented reality : a hands-on survey Pose estimation for augmented reality : a hands-on survey," vol. 22, no. 12, pp. 2633–2651, 2016.

[82] A. Mishra, " Using Google TensorFlow with Amazon SageMaker ," in *Machine Learning in the AWS Cloud*, 2019.

[83] N. Gucunski *et al.*, "RABIT: implementation, performance validation and integration with other robotic platforms for improved management of bridge decks," *Int. J. Intell. Robot. Appl.*, 2017.

[84] M. Bolukbasi, J. Mohammadi, and D. Arditi, "Estimating the Future Condition of Highway Bridge Components Using National Bridge Inventory Data," *Pract. Period. Struct. Des. Constr.*, vol. 9, no. 1, pp. 16–25, 2004.

[85] T. N. Sainath, O. Vinyals, A. Senior, and H. Sak, "Convolutional, Long Short-Term Memory, fully connected Deep Neural Networks," in *ICASSP, IEEE International*

Conference on Acoustics, Speech and Signal Processing - Proceedings, 2015.

[86] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, 1997.

[87] F. Gers, "Long short-term memory in recurrent neural networks," *Neural Comput.*, 2001.

[88] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, 2012, pp. 1097–1105.

[89] E. Karaaslan, U. Bagci, and C. F. Necati, "Attention-Guided Analysis of Infrastructure Damage with Semi-Supervised Deep Learning," *Autom. Constr.*, 2019.

[90] E. Karaaslan, U. Bagci, and F. N. Catbas, "Artificial Intelligence Assisted Infrastructure Assessment using Mixed Reality Systems," *Transp. Res. Rec. J. Transp. Res. Board*, p. 036119811983998, Jul. 2019.

[91] F. Chollet, "Keras: The Python Deep Learning library." 2015.

[92] M. Rashidi, B. Samali, and P. Sharafi, "A new model for bridge management: Part B: decision support system for remediation planning," *Aust. J. Civ. Eng.*, vol. 14, no. 1, pp. 46–53, 2016.

[93] R. W. Shepard, M. B. Johnson, and T. R. Board, "CALIFORNIA BRIDGE HEALTH INDEX: A DIAGNOSTIC TOOL TO MAXIMIZE BRIDGE LONGEVITY, INVESTMENT," *TR News*, no. 215, pp. 6–11, Jul. 2001.

[94] AASHTO, "Guide manual for bridge element inspection," *Bridg. Elem. Insp. Man.*, p. 172, 2011.

[95] American Association of State Highway and Transportation Officials, "AASHTOWare Bridge Management." Washington, DC, 1990.

[96] Federal Highway Administration, "LTBP InfoBridge: an intuitive and user-friendly

interface to access, visualize, and analyze bridge performance data," *Long Term Bridge Performance Program*, p. 2, 2019.

[97] Q. Ai, X. Wang, S. Bruch, N. Golbandi, M. Bendersky, and M. Najork, "Learning Groupwise Multivariate Scoring Functions Using Deep Neural Networks," Nov. 2018.

[98] R. K. Pasumarthi *et al.*, "TF-Ranking: Scalable TensorFlow Library for Learning-to-Rank," 2018.

[99] R. K. Pasumarthi, S. Bruch, M. Bendersky, and X. Wang, "Neural Learning to Rank Using TensorFlow Ranking: A Hands-on Tutorial," in *Proceedings of the 2019 ACM SIGIR International Conference on Theory of Information Retrieval*, 2019, pp. 253–254.

[100] C. C. Chang and C. J. Lin, "LIBSVM: A Library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, 2011.

[101] M. Gul, T. Terrell, N. Catbas, H. B. Gokce, D. Maier, and R. Zaurin, "Structural health monitoring for life cycle management of bridges," in *Life-Cycle Civil Engineering*, CRC Press, 2008, pp. 613–618.

[102] H. Hawk, "Bridge life-cycle cost Analysis, NCHRP Report 483," 2002.

[103] J. Mohammadi, S. A. Guralnick, and L. Yan, "Incorporating Life-Cycle Costs in Highway-Bridge Planning and Design," *J. Transp. Eng.*, vol. 121, no. 5, pp. 417–424, 1995.

[104] F. D. of Transportation, "FDOT Work Program Instructions," 2019.

[105] K. Jarvelin and J. Kekalainen, "IR evaluation methods for retrieving highly relevant documents," *SIGIR Forum (ACM Spec. Interes. Gr. Inf. Retrieval)*, 2000.

[106] S. Abu Dabous and S. Alkass, "Decision support method for multi-criteria selection of bridge rehabilitation strategy," *Constr. Manag. Econ.*, vol. 26, no. 786929861, pp. 883–

893, 2008.

[107] S. Abu Dabous and S. Alkass, “Decision support method for multi-criteria selection of bridge rehabilitation strategy,” *Constr. Manag. Econ.*, vol. 26, no. 8, pp. 881–891, 2008.

[108] FDOT Office of Maintenance, “Bridge Maintenance and Repair Handbook,” Florida.

[109] G. Hearn, P. D. Thompson, W. Mystkowski, and W. Hyman, *Framework for a National Database System for Maintenance Actions on Highway Bridges*. 2011.