Doctor of Philosophy

Bridge Damage Identification and Its Severity Estimation Using Artificial Intelligence

August 2019

Department of Civil and Environmental Engineering
The Graduate School of
Seoul National University

Soram Lim

Bridge Damage Identification and Its Severity Estimation Using Artificial Intelligence

A dissertation submitted to the Graduate School of
Seoul National University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

by
Soram Lim

July 2019

Approval Signatures of Dissertation Committee

Hyun-Soo Lee

Seokho Chi

Moonseo Park

Junho Song

Ki Tae Park

Bridge Damage Identification and Its Severity Estimation Using Artificial Intelligence

지도교수 지 석 호

이 논문을 박사학위논문으로 제출함

2019년 3월

서울대학교 대학원

건설환경공학부

임 소 람

임소람의 박사학위논문을 인준함

2019년 7월

위 원 장 이현수 (인)

부 위 원 장 지석호 (인)

위 원 (인)

위 원 송준호 (인)

위 원 박기태 (인)

DEDICATION

To all of my friends and family

ACKNOWLEDGMENT

In closing my doctoral dissertation, I would like to express my gratitude to many of the people who have supported and encouraged me.

First of all, I sincerely appreciate my supervisor, Prof. Seokho Chi, who guided me to the end of the Ph.D. course with patience and support. If I hadn't met Prof. Chi, I wouldn't have even started or finished this course. It was an honor to watch and learn from my great professor and senior in life. I would also like to thank Prof. Hyun-Soo Lee, the chief of the dissertation committee, who has constantly reminded me of not only knowledge but also the mindset of a scholar since I entered the graduate course. Also, many thanks to Prof. Moonseo Park for his great support and sincere advice for research as well as his valuable life advice to have a cold head and a warm heart. I also feel grateful to Junho Song for his advice and guidance as a dissertation committee with a gentle smile. While busy with your work, I would also like to express gratitude to Ki Tae Park, the head of the sustainable infrastructure research center, for giving his generous advice to help keep my dissertation balanced between academy and practice. Many thanks are given to Prof. Duyon Kim for watching and cheering me on from the start of my master's degree.

Next, I sincerely thank Kyung-Hoon Park and Jong-Wan Sun of the Korea Institute of Construction Technology, and Byung-Chul Lee, the CEO of G3WAY, for allowing me to research on this topic and giving me generous advice.

I am sincerely grateful to the members of the Construction Innovation Lab. with whom I spent many moments of works and fun. I was very encouraged by you guys' supports and considerations. Thank you for decorating a page of my life. I would also like to express my deepest gratitude and love to the Choam family (i.e., Bo Mi, Hong Eui, Jae Ho, Ji Won, Seung Mok), Hye In, and Ji Hye

for helping me not lose my direction whenever I am exhausted and go to the trouble. I also thank all my friends in the graduate course of the same university entrance year in 2009 for sharing joys and sorrows with them in the building number 35. Many thanks to Min-Woo Lee, Junehyeong Park, and Soo Hyun Koo for their generous advice and support as seniors who walked the path of graduate life. Also, I would like to express my great gratitude to Eun-Kyung Lee, who was beside me in the most difficult time of my life.

Lastly, I would like to express my most profound gratitude and endless love to my family. I send my sincere gratitude to my uncle and aunt, Seung-Ryul Kim, and Jung-Hyun Lee, who send me warm support and love. My special thanks are given to my brother, Seokbin Lim, who always share my trivial stories and research concerns. I give all these honors of this dissertation to my beloved and esteemed parents, Hyesook Kim and Woojin Lim. Their endless love and support enabled me to finish my long journey of Ph.D. course and dissertation. I finally send love and gratitude my grandmother, Sungyoon Yoon, who is watching from heaven.

Once again, I would like to extend my heartfelt thanks to all of you above.

Gwanak
July 2019
Soram Lim

ABSTRACT

Bridge Damage Identification and Its Severity Estimation Using Artificial Intelligence

Soram Lim

Department of Civil and Environmental Engineering
The Graduate School of Seoul National University

Bridge inspection is a fundamental step in obtaining infrastructure condition information for the bridge management cycle. In recent years, the number of aged bridges has increased rapidly, so the quality of inspection can be reduced due to the limits on the budget, time, and the number of qualified inspectors. Therefore, the aim of this research was to provide information to inspectors in advance of their inspections, such as the type, location, grade, and severity of the damage that has occurred on the bridge to be inspected at a specific time. To accomplish this goal, the factors that influence bridge damage was identified, and a model was developed to estimate damage using artificial intelligence. The research target was the deck of a pre-stressed concrete I-type (PSCI) bridge focusing on seven types of damage, i.e., cracking, map cracking, scaling, breakage, leakage, efflorescence, and corrosion of exposed rebar.

First, the identification, structural, traffic, and inspection data were obtained from the Korean Bridge Management System (KOBMS), which is

managed by the Korea Institute of Civil Engineering and Building Technology. Then, weather data were collected from the Korea Meteorological Administration. After preprocessing, the input data consisted of 59 independent variables and two dependent variables. Next, a correlation analysis was performed to remove 11 variables, taking into account the higher importance among the pairs with similar impact on the grade of the damage. Using the artificial intelligence approach, three kinds of decision tree methods were applied, and Extreme Gradient Boosting (XGBoost), which had the best performance, was used to derive the influencing factors for each type of damage. In general, dead load, live load, age, girder strength, and relative humidity were determined to be the most frequent influencing variables. Then, identification, structural and environmental factors associated with each type of damage were extracted. Using Deep Neural Networks (DNN) and XGBoost, artificial intelligence models were developed to estimate the severity levels of the various types of damage using the identified influencing variables. As a result, the XGBoost model, which was composed of 22 detailed models and showed the highest accuracy, was selected as the final model, and it was used to generate portfolios of damage individually and regionally that indicated the location-specific damage that may occur. The same methodology was used to assess the girders of PSCI bridges to confirm the possibility of expanding the model.

In this dissertation, the scattered quantitative information about the factors that influence damage to bridges was assembled and clarified by using deck inspection data. This is a pioneering attempt to use artificial intelligence techniques to produce information in advance of inspections to support those inspections. The developed methodology can provide information about existing managed bridges as well as bridges that have no past inspection history. By providing this information about bridges before they are inspected, the time required for the inspections and the risk of the quality degradation of bridge

inspections can be reduced. In addition, it will be possible to distribute inspection times, budgets, and human resources efficiently by providing inspection priorities among the number of bridges to be inspected. This research can contribute to the reduction of the lifecycle cost of the bridges and the extension of the life of the bridges through the expansion of this research with various components and structural forms, which shows the potential for leading a safer society.

Keywords: Bridge Inspection, Inspection Support, Bridge Damage Influencing Variables Identification, Bridge Damage Location Estimation, Bridge Damage Severity Estimation, Artificial Intelligence, XGBoost, Deep Neural Networks

Student Number: 2015-31060

CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1. Introduction

1.1. Research Backgrounds and Motivation

Proper bridge management to maintain operational traffic quality and to ensure traffic safety is essential for transportation agencies. The management lifecycle starts with a field inspection, followed by condition diagnosis, maintenance decision making, and maintenance actions. As the first step of the cycle, bridge inspection is a fundamental and vitally important aspect of the bridge management process. Currently, there is a substantial number of bridges that are more than 30 years old that should be inspected in several countries, including the United States (U.S.) and South Korea, since many bridges around the world were built during times of rapid economic growth. In 2017, the average age of 614,387 bridges in the U.S. was 43 years old (The American Society of Civil Engineers (ASCE), 2017), and more than 11,000 bridges in South Korea will be over 30 years old in 2028, whereas the number of bridges more than 30 years was only 4,471 bridges in 2018 (Korea Infrastructure Safety Corporation (KISTEC), 2018). Given the need to inspect such an increasing number of bridges, the limited budget and the limited professional manpower are threatening the quality of the inspections (Kušar, 2017), as shown in Figure 1-1. For example, three novice inspectors were asked to conduct biannual

inspections of 204 bridges within a three-month period in the second half of 2017 (Incheon Metropolitan City, 2018), and they had to check 126 bridges in one month during the first half of 2018 (Jeju Special Self-Governing Province, 2018). Thus, the lack of sufficient inspection time per bridge and the increasing risk of missing some significant damages during the inspections can decrease the quality of the inspections. Therefore, providing information on the possible occurrences of damage before inspections could reduce the inspection time and errors in the inspection, thereby promoting more effective and efficient inspections that have enhanced quality (Campbell et al., 2016).

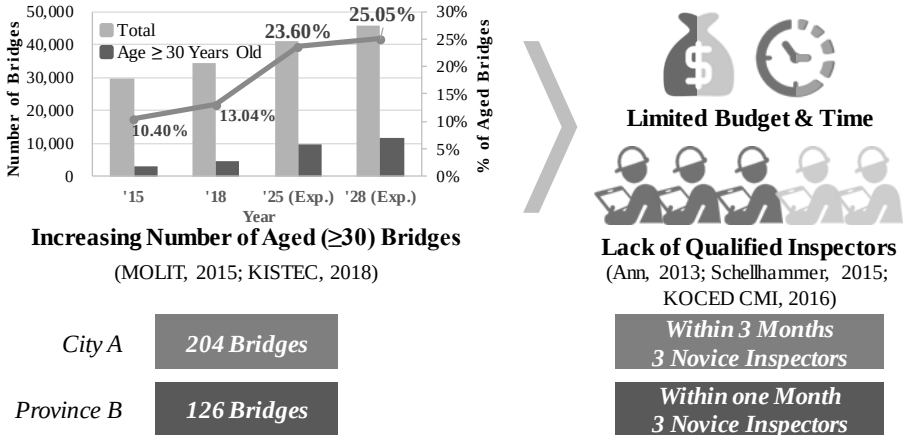


Figure 1-1. Lack of Inspection Resources with Increasing Number of Aged Bridges

In recent years, many decisions have been made based on estimates using the data that have been accumulated in different fields. In the field of crime prevention, PredPol (Predictive Policing), which has been used extensively in cities in California and Georgia since 2012, is a software that includes an estimation model which trains past crime cases every six months and estimates when and what type of crimes will occur in a particular location. The crime reduction rate after the use of this system began in these regions ranged from 8 to 20% (PredPol, 2014). Similarly, in South Korea, GeoPros (Geographical Profiling System) has been used since 2014 to identify crime risk areas. GeoPros calculates the crime risk index by using floating population, age distribution, income level, weather information, past crime occurrence status, and the police focus their patrol efforts on the areas that have high indices. The effect of introducing this system was a 13 - 45% reduction in crime rates in certain districts (Korea Crime Scene Investigation (KCSI), 2014). In the field of medicine, the artificial intelligence system called 'Watson' was developed by IBM. Watson estimates possible diseases by integrating the patient's genetic information, disease history, lifestyle, and present symptoms (IBM, 2019). In the bridge management filed, the Bridge Management System (BMS) has accumulated information related to bridges (i.e., their identification and structural information), traffic information, inspection records, and maintenance records. Therefore, it is possible to use BMS data to estimate when serious damage to a bridge in a specific location may occur and what type of

damage it may be, thereby providing more efficient inspection and maintenance.

Many studies have conducted to predict the conditions of bridges using the accumulated inspection data from BMSs (Scherer and Glagola, 1994; Cattan and Mohammadi, 1997; Tokdemir et al., 2000; Morcous et al., 2002; Zhao and Chen, 2002; Melhem et al., 2003; Bektas et al., 2013; Huang et al., 2015). However, these studies often are limited in their ability to provide detailed estimations of the levels of damage (e.g., cracking, scaling, and leakage). Previous studies have focused mainly on predicting the conditions of entire bridges or their major components, such as the deck, superstructure, and substructure (e.g., grade B of the entire bridge, rating 8 of the superstructure) (Su, 2003; Huang et al., 2010; Creary and Fang, 2013; Shan et al., 2016).

In practice, inspection guidance and manuals are provided around the world (Federal Highway Administration (FHWA), 2012; Ministry of Land, Infrastructure and Transport of Korea (MOLIT) and Korea Infrastructure Safety Corporation (KISTEC), 2012). But they mainly represent the average tendency of damage occurrences based on structural and material engineering information, so it is difficult to identify the time, location, and severity of the damages that have occurred for individual bridges.

1.2. Problem Description

Numerous researchers have conducted studies to quantify the conditions of bridges for use in making decisions related to the management of bridges. The results from previous studies are helpful in identifying the budget allocation aspects associated with prior bridges or components that had to be managed. Moreover, inspection guidelines and manuals contain only general theories concerning the damage mechanisms, and, as a result, limitations exist in the ability to provide specific information concerning the time, location, and severity of damage for individual bridges.

The first cause is that most studies have focused on network-level analyses. This approach considers a group of bridges or components that are subjected to the same conditions, e.g., traffic, humidity, amount of chloride, and span length as a network (Korea Institute of Civil Engineering and Building Technology (KICT) and Korea Infrastructure Safety Corporation (KISTEC), 2016), and predicts the average condition of particular types of bridges and particular components (Melhem and Cheng, 2003). Earlier studies related to the network analysis approach targeted the condition ratings of the entire bridge, such as a level C of a concrete slab bridge (Scherer and Glagola, 1994; Tokdemir et al., 2000) or the evaluation of the conditions of major components of the bridge, including the deck, superstructure, and substructure, such as a rating 9 bridge

concrete deck (Cattan and Mohammadi, 1997; Bektas et al., 2013; Creary and Fang, 2013). The network analysis has focused on calculating the maintenance cost to estimate the budget for bridge management. However, to obtain specific information on damage for an individual bridge, it is necessary to have a project-level approach in which individual bridges can be maintained and their damage-inducing characteristics can be analyzed without averaging them.

Inspection guidelines and manuals present the average trends of damage based on the aspects of structural and material engineering when damage is caused frequently by the structural form of the bridge and the environmental conditions. For example, in the case of the concrete deck of girder-type bridges, cracking may occur between the girders (Ministry of Land, Infrastructure and Transport of Korea (MOLIT) and Korea Infrastructure Safety Corporation (KISTEC), 2012), and extreme temperatures can induce non-structural cracks in concrete components (Federal Highway Administration (FHWA), 2012). The structure type of the girder and temperature are the key factors associated with damage to concrete components, but there is a lack of analysis and utilization based on actual inspection data to identify the internal and external conditions of individual bridges.

In summary, to date, analyzing systems that are focused on the network level and inspection guidelines and manuals have proven to be insufficient for generating specific information for individual bridges, such as “five cracks in grade B in the middle-left span area.” Therefore, using a data-driven approach,

a model is needed that can estimate the condition of the bridge at the damage level based on the factors that contribute to various types of damages, thereby providing support for inspections of individual bridges. For this purpose, the collection of damage level inspection data and various data that influence the variables was conducted, and the identification of the various causes of damage including the empirical factors was performed. Based on the identified variables, an estimation model that provides information at the damage level was developed for practical usage, and the model was used to generate the information that is needed for the inspections.

1.3. Research Objectives and Scope

The overarching goal of this dissertation is to provide the information that bridge inspectors need on the day they inspect bridges by developing a process using artificial intelligence to identify damage influencing variables and to estimate the type, location, grade, and severity of the damage that has occurred on the components of the bridge. To achieve this goal, it was determined that the following four specific objectives must be achieved:

1) To prepare the damage-level data for artificial intelligence models: Based on the information from a literature review of previous models for estimating the conditions of bridges, bridge data and traffic data were obtained from the Korean BMS, and weather data were obtained from the Meteorological Administration. The collected data were processed and used as input for the following artificial intelligence models of damage influencing factor identification and its severity estimation.

2) To identify the variables that influence the different types of damage to the bridges: First, the redundant variables were removed by correlation analysis to improve the performance of the estimation model and to prevent overfitting. Next, to consider the different mechanisms by which damage

occurred, different combinations of the variables that had more serious effects on the bridge conditions were identified by tree-based artificial intelligence methods.

3) To develop a model that estimate the damage in order to generate the inspection support information: Using the identified variables by different types of damage, a decision tree model and a neural networks model were developed to estimate the severity of the damage, and the model that provided the better performance was chosen. To validate the developed model, damage portfolios at both the project level and the network level were generated to support the inspections of the bridges.

4) To confirm the expandability of the model: Since it is necessary to understand the damage to various components of bridges to support the actual inspections, the same process was conducted for another component, and the model was confirmed to be extended and generalized.

Figure 1-2 shows the overall framework of this research.

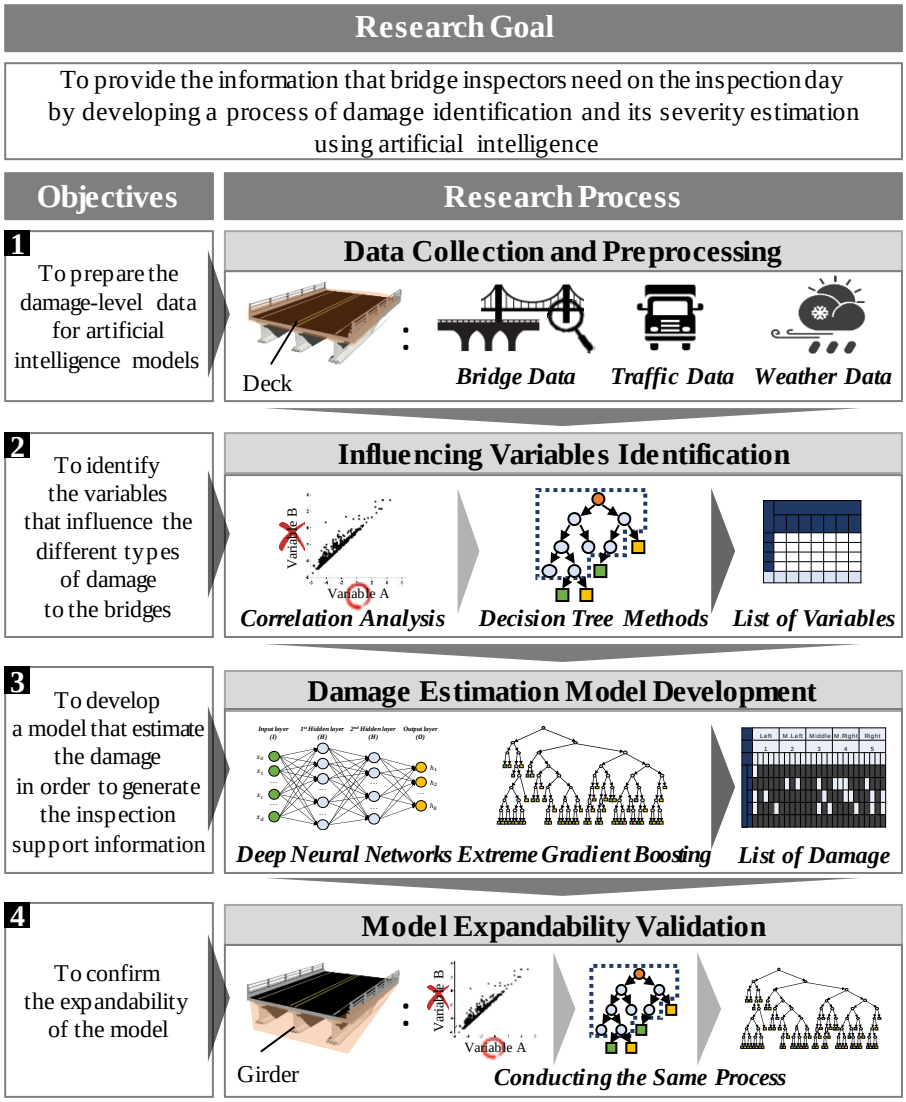


Figure 1-2. Research Framework

This research was focused mainly on pre-stressed concrete I-type (PSCI) decks; PSCI was the largest proportion (27%) in KOBMS compared to the other three major types of main structure in South Korea, i.e., reinforced concrete slab (15%), rahmen (24%), and steel box girder (22%). The deck directly supports the load on the bridge and has the highest rate of deterioration, which ultimately determines the serviceability (Freyermuth et al., 1970; Morcous et al., 2002; Scott et al., 2003; Huang, 2010). In the model expansion section, the focus was on the girder because it is the second most important component that divides and supports the load on the deck. In this study, credible inspection data were used that were obtained from detailed inspections and precise safety diagnoses conducted by professionals.

"Estimation" in this research implies the sense of calculating several condition grades at a specific time, whereas "prediction" is generally used to calculate the change of conditions at different times.

1.4. Dissertation Outline

This dissertation consists of seven chapters, including this Introduction. Descriptions of the content of each chapter are presented below.

Chapter 1, ***Introduction*****:** The backgrounds and motivations of this research are introduced, and the limitations of the previous studies and the current situations related to estimating the conditions of bridges based on the level of damage are specified. Consequently, the objectives and scope of this study are addressed.

Chapter 2, ***Preliminary Research*****:** This chapter describes the backgrounds of the current bridge inspection and bridge management systems and presents information concerning the estimation of the condition of bridges in previous studies.

Chapter 3, ***Research Methodology:*** This chapter explains the process of collecting and processing the dataset and exploring the overall characteristics of the data; it also introduces the methodology of this research that was developed after reviewing the methods used in previous studies.

Chapter 4, ***Damage Influencing Variables Identification:*** After conducting the correlation analysis to reduce dimensions, in this chapter, the influencing factors of seven types of damage to the PSCI bridge deck are derived using three kinds of tree-based artificial intelligence models; the method with the best performance is selected, and the variables that influence the damage are discussed considering the structural and material engineering aspects.

Chapter 5, ***Damage Estimation Model Development:*** In this chapter, a model is developed to estimate the severity of the damage using deep neural networks and extreme gradient boosting, and the model that produced the better results is chosen. The model is validated by generating bridge portfolios that describe the information related to the estimated damage at the project level and at the network level.

Chapter 6, ***Validation of the Model Expandability:*** The same process as described above is performed on the girders to confirm the expandability of the model. The variables that have distinct effects on the different types of damages to the girders are identified, and a model is developed that can provide highly-accurate estimates of the damage to the girders.

Chapter 7, ***Conclusion:*** This chapter describes the research results, the contributions of this study, and identifies future work in which the research results can be applied to improve bridge inspection practice.

Figure 1-3 shows the structure of this research.

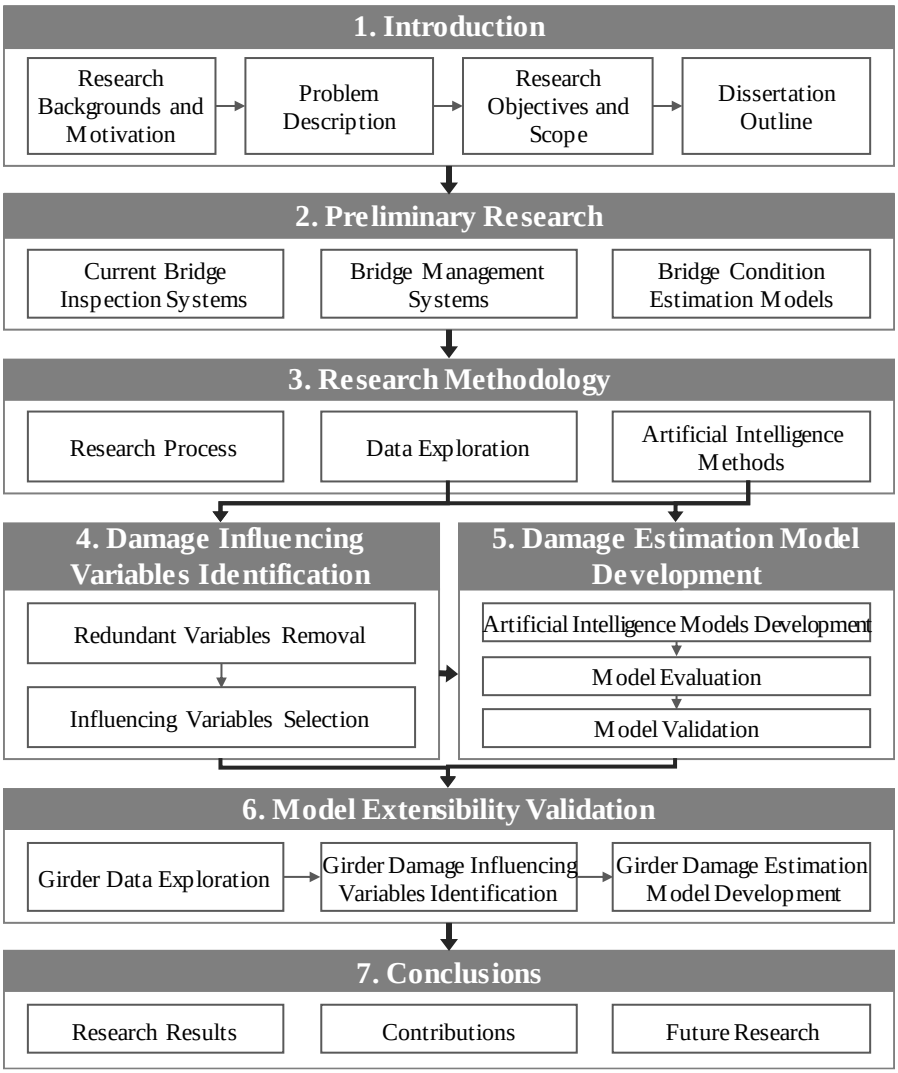


Figure 1-3. Dissertation Structure

Chapter 2. Preliminary Research

Numerous countries around the world have enacted laws and regulations that require the inspection of bridges. This chapter introduces current bridge inspection systems and bridge management systems in South Korea and foreign countries, especially the U.S., which has a Federal Department of Transportation and separate Departments of Transportation in each of the 50 States. This chapter also reviews the literature that addresses the targets of the estimation of the conditions of bridges and the methods that support the research backgrounds and problem description.

2.1. Current Bridge Inspection Systems

2.1.1. Types of Inspections

Since bridges are infrastructure facilities, they are generally managed by government agencies. Each country has its own laws and regulations for regular inspections. Bridge inspection programs in Denmark, Finland, France, Germany, South Africa, South Korea, Sweden, the United Kingdom, and the United States require 1) both short-interval visits and annual checks by maintenance contractors; 2) inspection of known defects at medium intervals

(e.g., 1 to 3 years), usually by agency-certified inspectors; and 3) thorough inspection of bridges at long-term intervals (i.e., 5 or 6 years), generally by agency-licensed professional engineers (Hearn, 2007).

In South Korea, the government passed *the Special Act on the Safety Control for Infrastructure* in 1995, after the collapse of the Seongsu Bridge in Seoul in 1994. In 2019, this law was revised as the Special Act on the Safety Control and Maintenance of Establishments. The aim of the Act is to enforce periodic inspection of major infrastructure facilities to provide timely repairs.

With the exception of urgent inspections, three types of inspections are performed according to the task, level, interval, and the inspector's qualifications, and the database that is saved includes 1) routine inspections at visible ranges, conducted every six months to one year by a novice inspector; 2) detailed inspections, accompanied by visual inspection and simple material testing, at two-year to four-year intervals by an Agency qualified inspector; and 3) precise inspections, conducted through structural analysis and safety assessment every 4 to 6 years (Table 2-1).

Table 2-1. Types of Inspections in South Korea

Types of Inspections	**Routine Inspections**	**Detailed Inspections**	**Precise Inspections**
Task	Visual inspection	Visual inspection Simple material testing	Visual inspection Non-destructive testing Structural analysis
Level	Damage on the Entire Bridge	Damage on Primary Components of the Bridge	Damage on All Components of the Bridge
Interval	6 months	Grade a - 3 years Grades b, c - 2 years Grades d, e - 1 year	Grade a - 6 years Grades b, c - 5 years Grades d, e - 4 years
Inspector's Qualifications	Novice technician	Agency-certified inspector	Agency-certified inspector
Saved as a Database	No (Saved as a text file)	Yes	Yes

Note: The above information is a summary of the contents from the Ministry of Land, Infrastructure and Transport of Korea (MOLIT) and the Korea Infrastructure Safety and Technology Corporation (KISTEC) (2017).

In the U.S., the *National Bridge Inspection Standards* were implemented by Federal Highway Administration (FHWA) after the collapse in 1967 of the Silver Bridge that connected Point Pleasant, West Virginia and Gallipolis, Ohio. Although the FHWA requires every Department of Transportation (DOT) in the U.S. to submit basic specification data and inspection information on bridges, each DOT in the 50 states has different systems of inspection cycles, methods, and inspectors' qualifications. Table 2-2 shows the general trends of inspections in several states in the U.S.

Table 2-2. Types of Inspections in the U.S.

Types of Inspections	**Routine Inspections**	**Fracture-critical Member Inspections**	**Underwater Inspections**
Task	Visual inspection	Visual inspection	Wading and Probing Or Dive inspection
Level	Specific components	Fracture-critical member	Underwater portion of a bridge substructure and the surrounding channel
Interval	24 months (predefined intervals by component or previous condition)	24 months (predefined intervals by previous condition)	60 months (within 72-month intervals)
Inspector's Qualifications	Agency-certified inspector	Agency-certified inspector	Agency-certified inspector
Save as Database	Yes	Yes	Yes

Note: The contents were summarized from Hearn (2007).

The U.S. is also concentrated on visual inspection. Most inspections are done by certified inspectors, and the results of all inspections are recorded in the database. In order to utilize the limited resources more efficiently to deal with the large number of bridges to be checked, the inspection intervals are defined differently by the types of components or by the previous conditions. For example, in the case of the box girder bridges in Oregon, a component that has active corrosion should be checked during every cycle, while the curve girders that are subject to out-of-plane distortion are checked at 48-month intervals (Hearn, 2007). Thus, the U.S. inspection system is complex and it is different in the 50 DOTs.

2.1.2. Inspection Guidelines and Manuals

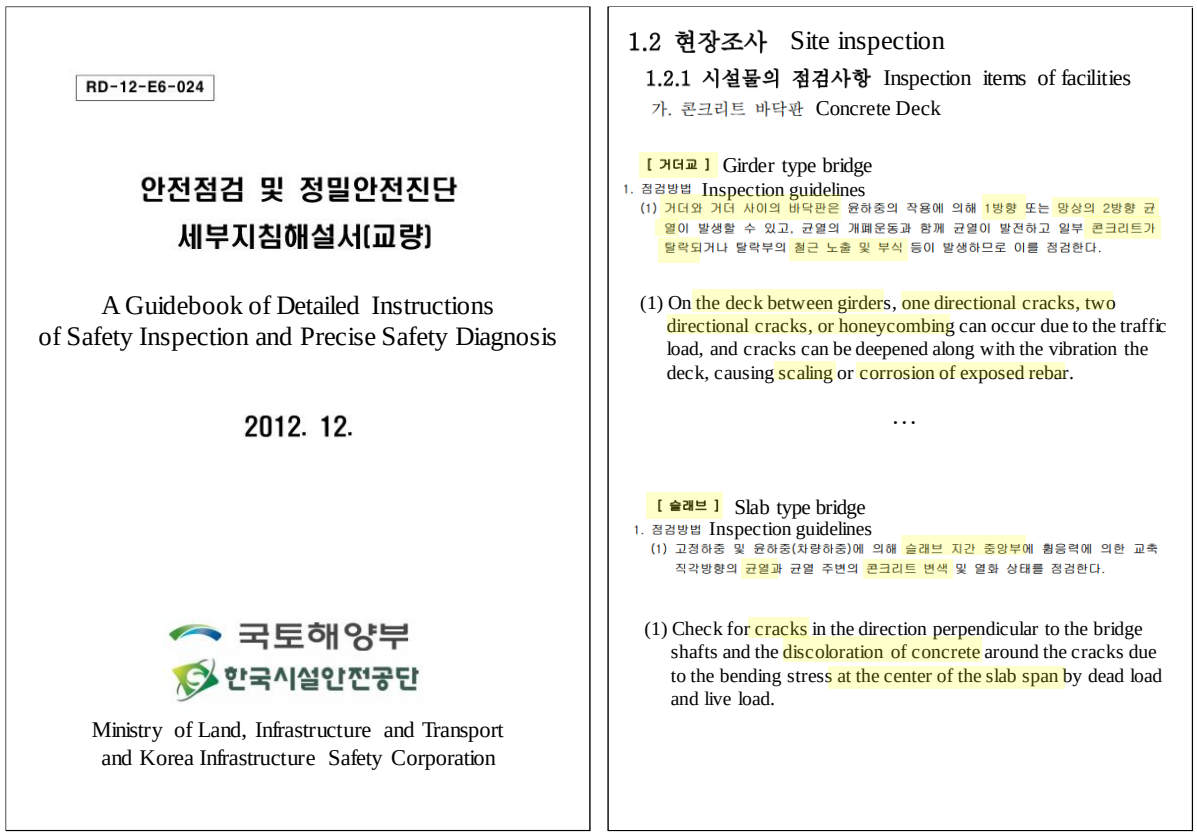
Since almost all types of inspections are fundamentally based on visual inspection, guidelines and manuals have been published in countries around the world to support inspectors, and some of them are described in Table 2-3. Most of the documents in Table 2-3 use both text and figures to describe frequent damage by the material used to construct the bridge, the type of structure, and types of components. However, those contents show the generalized tendency of the occurrences of damage based on structural and materials engineering knowledge and on the knowledge of the experienced inspectors. Therefore, environmental exposures, including the traffic and the weather associated with an individual bridge, can be difficult to consider, and it is hard to provide the location and severity of any damage that has been detected.

Based on a guidebook of detailed instructions for safety inspections and precise safety diagnoses in South Korea, Figure 2-1 shows the parts and components that should be the main focus of inspections by different structural types of bridges. This guidebook describes the areas on components in which damage is likely to occur within a span of each type of bridge. Even though specific areas are identified where damage frequently occurs based on the type of structure, it is difficult to apply the information to a particular bridge using the guidelines alone.

Table 2-3. Inspection Guidelines and Manuals from National Agencies

Nation	Inspection Guidelines and Manuals
Australia	Routine Visual Bridge Inspection Guidelines (Level 1 Inspections) for Bridges, 6706-02-2234 (2009), Main Roads Western Australia. Detailed Visual Bridge Inspection Guidelines for Concrete and Steel Bridges (Level 2 Inspections), 6706-02-2233 (2008), Main Roads Western Australia.
Canada	BIM Inspection Manual, Version 3.1 (2008), Alberta Infrastructure and Transportation. Bridge Inspection and Maintenance System: Level 2 Inspection Manual, Version 1.1 (2007), Alberta Transportation. Ontario Structure Inspection Manual (2008), Ontario Ministry of Transportation.
Germany	Highway Structures Testing and Inspection, DIN 1076 (1999), Deutsche Norm. Recording and Assessment of Damages, RI-EBW-PRÜF (2017), Bundesanstalt für Straßenwesen. ASB Structure Inventory, (Coding Manual for ASB-ING) (2013), Bundesanstalt für Straßenwesen.
Norway	Handbook for Bridge Inspections (2005), Norwegian Public Roads Administration.
South Korea	A Guide Book of Detailed Instructions of Safety Inspection and Precise Diagnosis (2012), Ministry of Land, Infrastructure and Transport of Korea (MOLIT) and Korea Infrastructure Safety Corporation (KISTEC). Detailed Instructions of Safety Inspection and Precise Diagnosis (2017), MOLIT and KISTEC.
United Kingdom	Requirements for Inspection and Management of Bridges BD 62/07 (2007) and BD 63/17 (2017), The Highways Agency, Transport Scotland, Welsh Assembly Government, and the Department for Regional Development Northern Ireland.
United States	Commonly Recognized (CoRe) Structural Elements (2001), American Association of State Highway and Transportation Officials (AASHTO). The Manual for Bridge Evaluation, 3rd ed. (2017), AASHTO. Bridge Inspector's Reference Manual, FHWA NHI 12-049 (2012), Federal Highway Administration (FHWA). Recording and Coding Guide for the Structural Inventory and Appraisal of the Nation's Bridges, FHWA-PD-96-001 (1995), FHWA.

Note: The contents were modified and updated from TABLE 2 INSPECTION MANUALS—FOREIGN SOURCES (Hearn, 2007).



RD-12-E6-024

안전점검 및 정밀안전진단
세부지침해설서(교량)

A Guidebook of Detailed Instructions
of Safety Inspection and Precise Safety Diagnosis

2012. 12.

국토해양부
한국시설안전공단

Ministry of Land, Infrastructure and Transport
and Korea Infrastructure Safety Corporation

1.2 현장조사 Site inspection

1.2.1 시설물의 점검사항 Inspection items of facilities

가. 콘크리트 바닥판 Concrete Deck

[거더교] Girder type bridge

1. 점검방법 Inspection guidelines

(1) 거더와 거더 사이의 바닥판은 윤하중의 작용에 의해 1방향 또는 망상의 2방향 균열이 발생할 수 있고, 균열의 개폐운동과 함께 균열이 발전하고 일부 콘크리트가 탈락되거나 탈락부의 철근 노출 및 부식 등이 발생하므로 이를 점검한다.

(1) On the deck between girders, one directional cracks, two directional cracks, or honeycombing can occur due to the traffic load, and cracks can be deepened along with the vibration the deck, causing scaling or corrosion of exposed rebar.

...

[슬래브] Slab type bridge

1. 점검방법 Inspection guidelines

(1) 고정하중 및 윤하중(차량하중)에 의해 슬래브 지간 중앙부에 휨응력에 의한 교축직각방향의 균열과 균열 주변의 콘크리트 변색 및 열화 상태를 점검한다.

(1) Check for cracks in the direction perpendicular to the bridge shafts and the discoloration of concrete around the cracks due to the bending stress at the center of the slab span by dead load and live load.

Figure 2-1. Inspection Guidelines in the Official Guidebook of South Korea

In the U.S., FHWA provides a manual to support inspectors that describes the inspection program, the inspection methods, the reporting method, and the structural backgrounds by components in different materials (Federal Highway Administration (FHWA), 2012). Figure 2-2 shows this manual guides the inspection of common deficiencies on a particular type of component by considering the locations of frequent occurrences. As in the case of South Korea, it is difficult to use the information in this guideline to extract information for a specific bridge, which makes it difficult to consider the different external environments in which individual bridges are located.

In Australia, the guideline instructs the user to check the points of frequent damage specified in the text and figure and to determine their cause and location according to the type of material, e.g., timber, concrete, and steel (Figure 2-3). But still, it is difficult to identify the specific locations of the occurrence of damage in the entire bridge.

In summary, general information about frequently-occurring damage on specific materials for various components or on a specific structural form and component was provided by guidelines and manuals of the transportation agencies throughout the world. However, it is difficult to provide inspectors with specific information on the location and severity of damage that occurs on individual bridges using these documents.

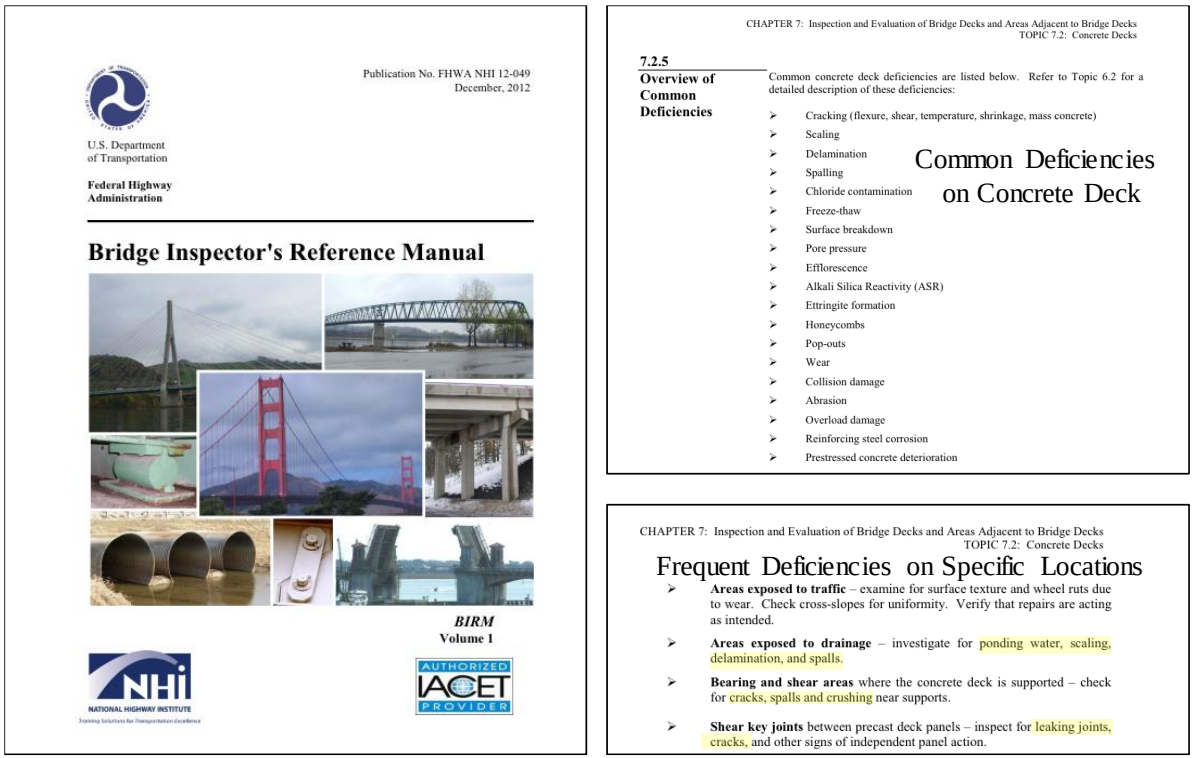


Figure 2-2. Inspection Guidelines in the Official Manual of the United States

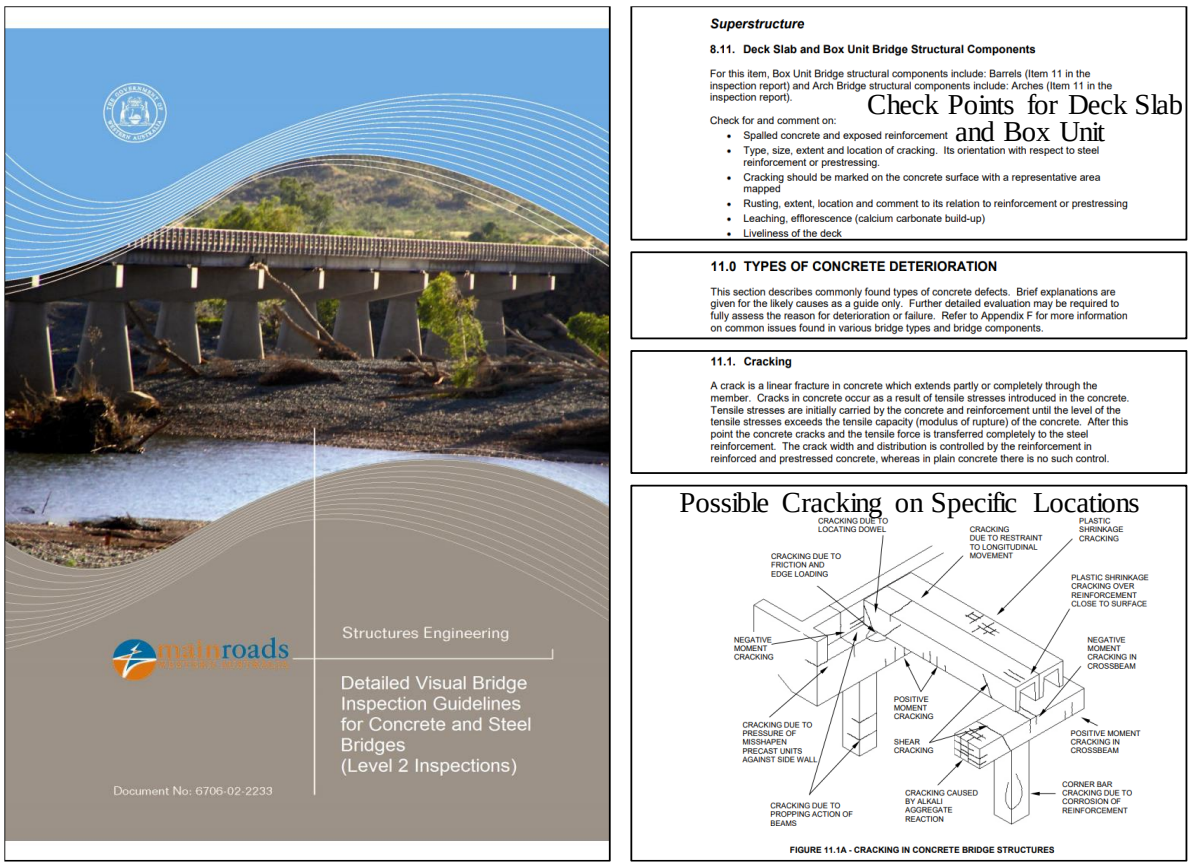


Figure 2-3. Inspection References in the Official Guidelines of Australia

2.1.3. Condition Rating Systems

There is a great variety of criteria that can be used to assess the condition of bridges, their components, and the damage they have incurred in numerous countries. Most authorities in Denmark, Finland, France, Norway, Sweden, the United Kingdom, and the United States collect the component level data; some authorities in Germany, South Africa, South Korea, Spain, and several DOTs in the U.S. including California, Florida, Iowa, Kansas, Kentucky, North Dakota, Texas, Wisconsin gathers the damage level data (Winn, 2011; Mirzaei et al., 2012; Kotze et al., 2015; Nsabimana, 2015). Almost all rating systems were three, four, or five levels except nine-level condition ratings of the U.S. National Bridge Inspection Program (NBIP) and 100 points of scoring system in Span.

In South Korea, for example, visual inspection is used to assign one of the five-level condition grades, i.e., “A” (excellent), “B” (good), “C” (normal), “D” (inadequate), and “E” (poor). Table 2-4 indicates that the states of “C”, “D”, and “E”, respectively, indicate that repairing, strengthening, and rehabilitating are required to maintain functionality.

Table 2-4. Condition Grades Guidelines in South Korea

Condition Grade	Description
A (Excellent)	The best condition without damage
B (Good)	The condition that minor damage, not affecting the functionality, occurred on secondary components and partial repairs are required to enhance the durability
C (Normal)	The condition that minor damage on primary components or extensive damage on secondary components occurred but the overall safety is not affected, and repairs for primary components or reinforcements for secondary components are required to maintain the durability and functionality ➔ Repair
D (Inadequate)	The condition that damage occurred on primary components are required, and usage restriction should be decided ➔ Strengthened
E (Poor)	The condition that severe faults threatening the safety of facility occurred on primary members, the service of facility must be restricted immediately, and reinforcements or a reconstruction is required ➔ Rehabilitation

Note: The contents were translated from Ministry of Land, Infrastructure and Transport of Korea (MOLIT) and Korea Infrastructure Safety and Technology Corporation (KISTEC) (2017).

A condition assessment of the entire bridge, which is referred to as the "diagnosis", begins at the level of the damage. The lowest values of the condition states for each type of damage, collected at the inspection stage, become the condition ratings of the components. The lowest grade of the components within one span becomes the representative ratings of the span,

and their average becomes the condition ratings of components in all of the bridges. The weighted sum of the different components represents the condition grade of the entire bridge (Figure 2-4). Again, it is important to collect and analyze data on the level of the damage, which is the smallest and the basic unit of this rating.

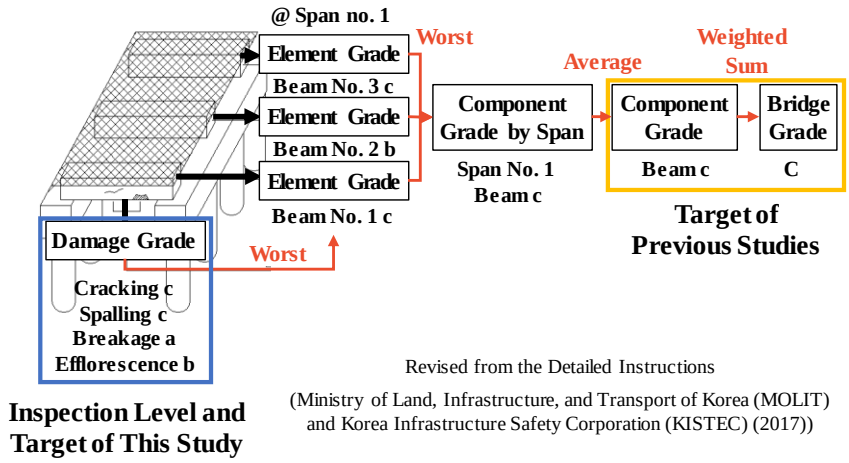


Figure 2-4. Principle of Condition Diagnosis in South Korea

In the U.S., the rating system in the NBIP is built at the component level for three major subsystems of the bridge (i.e., the deck, the superstructure, and the substructure) and culverts. The condition rating in the U.S. is based on the types and severity of the damage rather than on the rating of the entire member at a glance, as shown in Table 2-5. There is no index of damage level, but the condition of damage is reflected implicitly. For example, the fifth state includes the check points regarding section loss, cracking, spalling, and scour, and the third state reflects fatigue cracks, which are more serious damage.

Table 2-5. Condition Rating Guidelines in the U.S.

State	Description
9	Excellent condition
8	Very good condition: no problems noted.
7	Good condition: some minor problems.
6	Satisfactory condition: structural components show some minor deterioration.
5	Fair condition: all primary structural components are sound but may have minor section loss, cracking, spalling, or scour.
4	Poor condition: advanced section loss, deterioration, spalling, scour.
3	Serious condition: loss of section, deterioration, spalling, or scour have seriously affected primary structural components. Local failures are possible. Fatigue cracks in steel or shear cracks in concrete may be present.
2	Critical condition: advanced deterioration of primary structural components. Fatigue cracks in steel or shear cracks in concrete may be present or scour may have removed substructure support. Unless closely monitored, it may be necessary to close the bridge until corrective action is taken.
1	'Imminent' failure condition: major deterioration or section loss present in critical structural components or obvious vertical or horizontal movement affecting structural stability. Bridge is closed to traffic but corrective action may put back in light service.

Note: The contents were obtained from Kotze et al. (2015).

2.2. Bridge Management Systems

2.2.1. Information in BMS

A BMS is defined as a computerized management system to handle the amount of information which is required to organizing and conducting all of the activities related to managing a network of bridges (Scherer and Glagola, 1994; Mirzaei et al., 2012). According to the survey on 21 BMSs in 16 countries, BMS generally includes basic inventory information, inspection information, intervention information, and prediction information, as described in Table 2-6 (Mirzaei et al., 2012). Basic inventory information includes identification information (e.g., location of the bridge and number of lanes) for the bridge itself and detailed structural information (e.g., deck thickness and number of girders). Inspection information is collected according to type of inspections and the target level of inspection which can be damage, component, structure type, and bridge level, in each country. Inspection data can be collected directly through the recently developed mobile application, typed into the database by the officer after the inspection, or stored in the form of a report in text format. Intervention information includes various maintenance methodologies and the corresponding cost information, and prediction information contains condition degradation and performance improvements after maintenances.

Table 2-6. Types of Information in Bridge Management System

Deck Damage Type	Description
Basic Inventory Information	Information on the infrastructure objects owned or managed by the user of the BMS, including structure types, numbers of structures per structure type, and archives, as well as how the location information, loading information and use information is entered
Inspection Information	Information about inspections where the information obtained is either entered into or used by the BMS, such as the information collected and how it is collected
Intervention information	Information about maintenance and preservation activities such as repair, rehabilitation and reconstruction activities, that is either entered into or used by the management system
Prediction Information	Information on the aspects being predicted by the BMS, e.g. change in physical condition and performance indicators due to deterioration and the execution of interventions

Note: The contents were extracted from Mirzaei et al. (2012).

In South Korea, detailed inspections and precise inspections are not performed as frequently as the regular inspections, and they are conducted by agent-certified inspectors, which produces highly-reliable information. These reliable inspections produce detailed information because all of the components of each span are checked, and this information is entered in tabular format in the Facility Management System (FMS) and compared to the results of the routine inspection, which remain in the text report (Table 2-1). Some of the FMS data in particular regions are managed thoroughly in KOBMS by the

Korea Institute of Civil Engineering and Building Technology (KICT). This reliable and detailed information was used in this study to estimate the damages on bridges.

In the U.S., the entire lifecycle of bridge management has been accumulated in the National Bridge Inventory (NBI) since the 1970s. Every state government is required to submit information regarding the conditions of its bridges to NBI regularly. In the 1990s, the American Association of State Highway and Transportation Officials (AASHTO) developed the commonly-recognized (CoRe) elements for bridge inspection, which were used to standardize the collection of element-level condition data in the United States. Two other major sources of data concerning the conditions of bridges are the NBI and CoRe element condition data (Bektaş, 2017).

2.2.2. Functions of BMS

The primary objective of BMS is to support bridge managers to make consistent and cost-effective decisions for maintenance, repair, and rehabilitation (MR&R) of bridges and under limits on resources (Morcous, 2000). Numerous functions of BMSs include (1) assessing and predicting the conditions of bridges; (2) evaluating and analyzing the improvement in their conditions after treatments; and (3) analyzing life-cycle cost of maintenance strategies and recommending the optimal strategy given the budget constraints (Morcous, 2000).

As shown in Figure 2-5, each function matches condition diagnosis, maintenance decision making, and maintenance actions of the bridge management cycle and is supported by the model, i.e., the deterioration model, the improvement model, and the optimization model established based on the data accumulated in the BMS (Morcous, 2000; Winn, 2011). The deterioration model is defined as the relationship between measures of bridge condition and a vector of explanatory variables (Ben-Akiva and Gopinath, 1995). This model predicts the future condition of different bridge components, which results in assisting the determination of the optimal maintenance strategy and the estimation of future funding requirements. Bridge deterioration models will be discussed minutely in the next section. The improvement model includes the

assessment of the change of condition of bridge components after conducting different maintenance methods, and the cost estimation of each method for determining labor, materials and equipment. The optimization model calculates the total cost of MR&R strategies for different bridge components, and finally suggests the best strategy. This model may use life cycle cost analysis to explain the maintenance conducted on a bridge throughout its entire service life (Morcous, 2000; Winn, 2011).

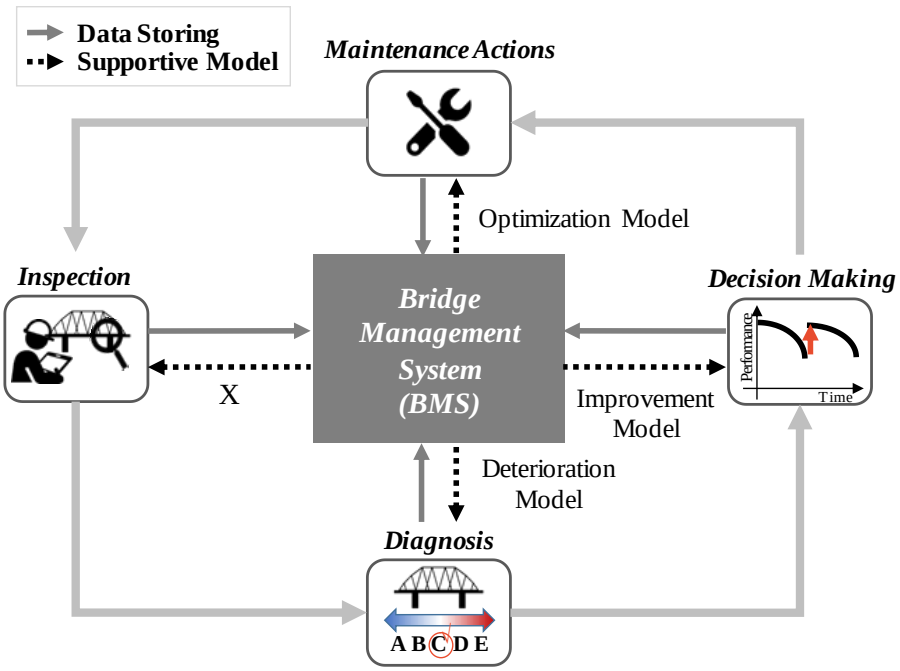


Figure 2-5. Relationships between the Bridge Management Tasks and the Functions of the Bridge Management System

The inspection step, however, rarely has been supported by any kinds of model. This situation has occurred because BMS has concentrated on optimizing maintenance actions with minimizing costs. Since bridge inspection is the first step of the bridge management cycle, the model to estimate damage and to provide information before inspections can be expected to improve the quality of inspection as well as the quality of bridge management.

2.3. Bridge Condition Estimation Models

2.3.1. Targets of Previous Models

A significant number of studies have sought to quantify the profile of bridge deterioration at the whole level, e.g., grade of the whole bridge after a specific time interval, or at the component level, e.g., a deterioration curve of the deck; however, this has resulted in a gap when considering inspections at the damage level, e.g., cracking of the deck at a specific time. In the earlier stages of predicting the conditions of bridges, several studies attempted to estimate the condition of the entire bridge. Scherer and Glagola (1994) developed a Markov chain model that ranged from 3 (poor) to 9 (new) to predict the condition of an entire bridge in Virginia, and Tokdemir et al. (2000) estimated sufficiency ratings of bridges in California using an artificial neural network (ANN) and genetic algorithms. Su (2003) analyzed the relevance of bridge deterioration of the entire bridge and its environment in Taiwan using logistic regression.

In succession, Cattan and Mohammadi (1997), Bektas et al. (2013), and Creary and Fang (2013) conducted research studies on the major components of bridges, including the deck, superstructure, and substructure. Among the major components, many studies have focused on analyzing the condition of

the deck because it has the highest degeneration rate due to its exposure to environmental conditions, including traffic and weather (Freyermuth et al., 1970; Morcous et al., 2002). Morcous et al. (2002) conducted case-based reasoning for modeling deck deterioration using inventory data, inspection data and maintenance data from Quebec, Canada. The condition of the deck of the entire bridge normally is defined as a weighted sum of the number of components of the deck that are in each condition. Thus, Melhem et al. (2003) developed a decision tree model using the wrapper method to predict deck condition ratings from 1 (protected) to 5 (failed), using the Commonly Recognized (CoRe) elements data in the U.S. Morcous (2005) also utilized the decision tree method to model deck deterioration and a performance of the tree model was better than Markov chain. Similarly, Huang (2010) developed an ANN model to predict the condition of the deck probabilistically at a certain point in time, and Huang and Chen (2012) analyzed bridge inspection data in Florida and formed five clusters by different bridge characteristics derived in the form of fuzzy rules; then, they used these rules to determine the average deterioration rate of decks in each bridge cluster. Recently, in addition to the deck, Shan et al. (2016) identified superstructure deterioration of the U.S. bridges using logistic regression, and Huang et al. (2015) also conducted research on secondary components such as abutments, railings, approaches, and waterways, focusing on identifying rainfall effects.

Few studies have been conducted at the damage level. Zhao and Chen (2002) predicted the occurrence of cracking and spalling damage on superstructures and substructures, but not on decks, and they used bridge data from Singapore for their probabilistic neural networks. To ensure that inspections are conducted in a timely manner and at a reasonable cost, investigations are still needed to provide detailed estimations of the number of defects to be inspected, such as "five cracks and two corrosion areas," and their severity. Table 2-7 summarizes different target levels in previous condition estimation model.

Table 2-7. Different Target Levels in Previous Condition Estimation Model

Scope	Literature
Entire bridge	Scherer and Glagola (1994); Tokdemir (2000); Su (2003)
Component	• **Major Structure (Deck, Superstructure, and Substructure)**: Cattan and Mohammadi (1997); Bektas et al. (2013); Creary and Fang (2013) • **Deck**: Morcous et al. (2002); Melhem et al. (2003); Morcous (2005); Huang (2010); Huang and Chen (2012) • **Superstructure:** Shan et al. (2016) • **Abutment, Railing, Barrier, Approach, Waterway:** Huang et al. (2015)
Damage	• **Cracking, Spalling on the Superstructure and Substructure**: Zhao and Chen (2002)

2.3.2. Methods of Previous Models

Previous estimation models can be categorized into three groups that are not mutually exclusive (Morcous et al., 2002): deterministic, stochastic, and artificial intelligence models. A representative method of deterministic models is regression. Su (2003), for example, analyzed the relevance of the entire bridge deterioration and its environment in Taiwan using logistic regression. Similarly, Shan et al. (2016) used logistic regression to calculate the probability of a bridge having a deficient superstructure of steel bridges in Oklahoma. Regression models are simple, intuitive, and efficient for the analysis of networks with a large population, however, has limitations in that: (1) the method needs statistical assumptions including linear relationship between variables; (2) the method neglects random errors from unobserved explanatory variables in the estimation; and (3) the method predicts average condition regardless of the current condition and the condition history of individual facilities.

As another approach, Markov chain is a typical method of stochastic models, and was utilized in PONTIS BMS in the U.S. While Scherer and Glagola (1994) developed a deterioration model of the entire bridge using Virginia data from the PONTIS, Huang et al. (2015) applied Markov Chain Monte Carlo (MCMC) simulation, which added the procedure to find best

transition probability by Monte Carlo simulation, to develop deterioration model of bridges in Victoria, Australia. By the Markov chain model, a next condition rating after a certain time interval is calculated by the previous condition rating times transition probability. Owing to this concept, this method overcome the three limitations of regression by the transition probability to treat uncertainty and randomness using non-linear programming, and consideration of the current state. Nevertheless, the Markov method still has limitations in that (1) transition probabilities are independent by time; (2) the method neglects the condition history and assumes discrete time interval; (3) the method assumes independency of condition ratings in the 1-st order model in BMS (e.g., PONTIS and BRIDGIT).

Recently, many studies have applied artificial intelligence techniques using the accumulated BMS data. For example, Melhem et al. (2003) and Morcous (2005) used a decision tree to discover damage-influencing identification factors and predict deck condition simultaneously. Several studies focusing on the estimation itself, rather than identifying factors, have applied a neural network, which is known to yield high accuracy (Li et al., 1996; Cattan and Mohammadi, 1997; Zhao and Chen, 2002; Elhag and Wang, 2007; Huang, 2010; Creary and Fang, 2013; Lee et al., 2014). The artificial intelligence method also has advantages of not assuming statistical distributions among variables and not requiring the predetermined relationships between inputs and outputs. This method, however, is computationally intensive and

needs a certain amount of data to train a model.

Therefore, the second objective of this study, i.e., to derive the factors that influence the damage to bridges, was to be achieved by using methods based on decision trees, and the third objective, i.e., building the model for estimating damage, was achieved by using the tree-based classification method and neural networks. In order to overcome the limitations of the artificial intelligence method mentioned above, a suitable target, i.e., the deck of PSCI, was selected in order to obtain a sufficient amount of data for analysis, and a high-performance computer with a large graphics processing unit was used to reduce the computation time.

2.4. Summary

In this chapter, first, the current bridge inspection systems around the world were reviewed, especially in South Korea and the U.S. After the bridge collapse accidents that occurred in both countries, laws and regulations were established to force regular inspections, i.e., mainly visual inspections. The guidelines and the manual used at the time of the visual inspections describe the average tendencies of damage to occur by the main types of structures and material, and, therefore, it is difficult to produce information for each bridge by considering the characteristics and external environments of individual bridges. The rating systems have 5 grades in South Korea and 9 grades in the U.S

Then, types of information in BMSs, i.e., basic inventory information, inspection information, intervention information, and prediction information, were presented and the particular BMSs in in South Korea and the U.S., i.e., FMS and KOBMS in South Korea, and NBI and the CoRe element condition in the U.S., were introduced. Inspection, which is the first step in the bridge management cycle, did not receive sufficient support with functions that used data stored in BMS compared with the other three steps in the management cycle. Therefore, providing information on possible damage before inspections would expand the functionality of BMS and consequently enhance the quality of inspection as well as that of the entire bridge management cycle.

The previous research efforts that had been conducted on estimating the conditions of bridges were investigated. The gap between the inspection level, damage, and the target level of the previous studies with respect to the components or the entire bridge was confirmed, and this was the reason the artificial intelligence approach was required to accomplish the goal and the objectives of this research.

Chapter 3. Research Methodology

This chapter introduces the process, material, and methods of this research. This section explains four steps to accomplish the four objectives of this research. In the first step, the procedures for collecting and processing data are described, and the artificial intelligence methods used in this study are introduced.

3.1. Research Process

This research was performed in four steps to achieve four objectives as shown in Figure 3-1. First, the kinds of data to be collect were decided through a literature review, and the corresponding data were obtained. The data that were collected from the Korean Bridge Management System included identification, structural, traffic, and inspection data, and weather data were collected from the Korea Meteorological Administration. Since this study focused on the deck of PSCI bridges, the data that were collected were filtered to complete the dataset. The preprocessing stage included generating new variables (i.e., data transformation), which included severe weather conditions, previous conditions of the decks, locations of the damage, and the severity of the damage.

In the second stage of identifying the influencing variables, first, a correlation analysis was used to remove the less important variables to affect the condition grades of the damage among each pair of highly correlated variables. Then, decision tree models, i.e., one of the artificial intelligence techniques, were used to classify the degree of damage. Using the decision trees that were generated, the values of relative importance of the input variables were calculated, and the variables that influenced the various types of damage were selected using the rank of relative importance.

Next, using the identified variables, models to estimate the severity of the damage were developed based on two methods from artificial intelligence, i.e., deep neural networks and extreme gradient boosting. By comparing the performances of the two models, the final model to be used was selected, and lists of damage (i.e., a damage portfolio) for each bridge and region were generated to validate the model.

Then, the entire process was applied to the PSCI girder data to derive the variables that influence the damage to girders and to develop a model to estimate the severity of the damage to girders. Based on the results, the extensibility of the model was validated.

The process of removing redundant variables was implemented using R software, version 3.3.4, and the selection of variables and the development of estimation models were trained on Python 3.6.8 with TensorFlow 1.11.0 by Google and with computing libraries, including Keras 2.2.4, Numpy 1.15.4, Pandas 0.23.4, Scikit-learn 0.20.2, and Xgboost 0.81.

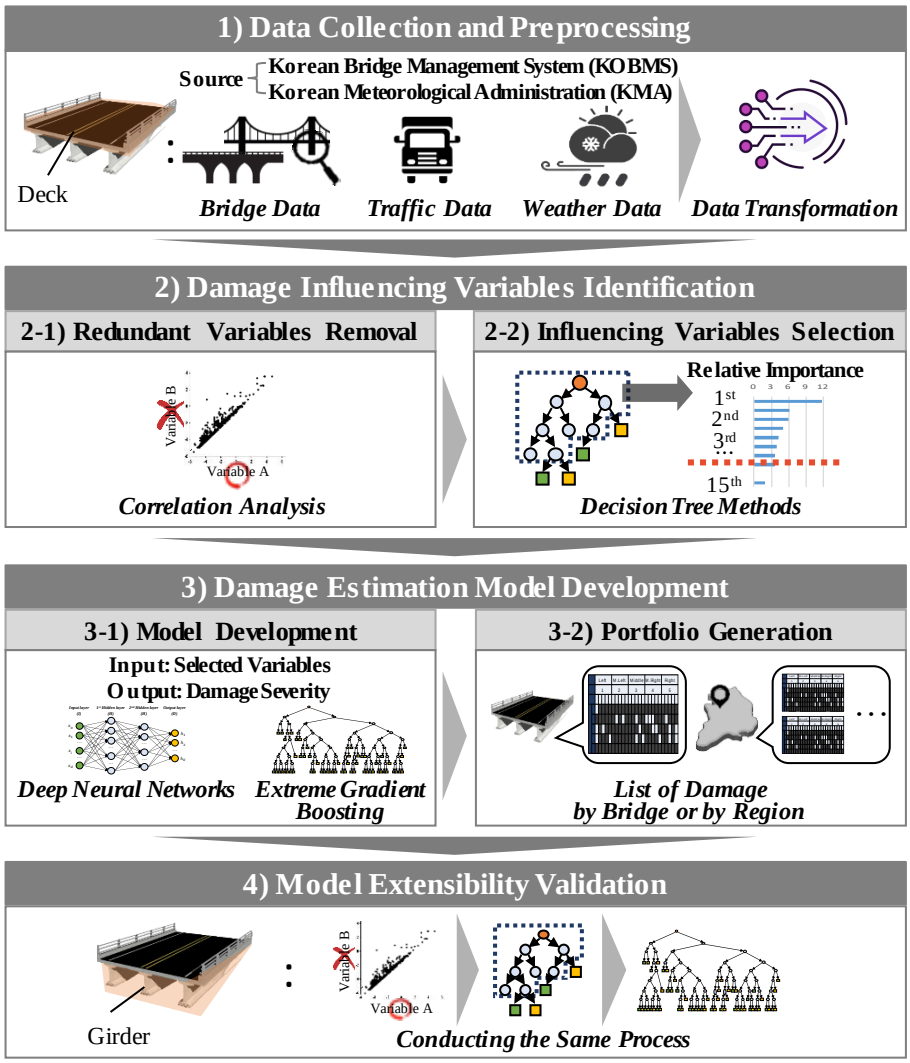


Figure 3-1. Research Process

3.2. Data Exploration

3.2.1. Data Collection and Preprocessing

(1) Data Used in Previous Research

Before collecting data, factors that influenced on the condition degradation of the entire bridge, components, and damage was investigated from literature. Earlier studies have specified the identification factors (e.g., number of spans and main structure types) that affect bridge damage. Su (2003) identified structural factors, such as a beam-type girder and the distance to the nearest coast, had impacts on the entire bridge deterioration. Shan et al. (2016) found that the region was significant because bridges located in particular states that were managed by specific authorities had higher probabilities of damage. Scherer and Glagola (1994) determined a few specific influencing variables for condition of the entire bridge, such as road system type, number of spans, deck structure type, and traffic from the PONTIS BMS. Huang et al. (2015) were more concerned about identification factors, such as region, crossing type, and the structural characteristics of the surrounding components, such as abutment, railing, barrier, approach, and waterway.

In addition to identification factors, recent studies have added structural and environmental factors to reflect the complex nature of such damage. Numerous studies have agreed that the degree of bridge deterioration is dependent on the exposure environment in addition to identification and structural factors (Wang et al., 2007; Yianni et al., 2016). The factors generally considered in previous studies include: traffic, weather (i.e., temperature, precipitation, humidity, and wind), and chemicals (e.g., chloride and deicing material) (Dadson et al., 2002; Morcous et al., 2002; Wang et al., 2007; Agrawal et al., 2010; Zhang et al., 2016).

Since traffic volume is one of the major loads on a bridge deck, many studies have determined the total traffic volume and truck traffic volume to be critical deterioration factors (Zhao and Chen, 2002; Huang, 2010; Huang et al., 2010; Creary and Fang, 2013; Lee et al., 2014; Shan et al., 2016). For example, Scherer and Glagola (Scherer and Glagola, 1994) used average daily traffic (ADT) and average daily truck traffic (ADTT) to develop the Markov chain prediction model for analyzing the conditions of bridges in Virginia. Similarly, Melhem et al. (2003) and Morcous (2005) considered ADT and the percentage of trucks in the traffic, i.e., the ratio ADTT to ADT, as inputs to a decision tree for rating the condition of bridges.

Weather conditions also have been considered to be a major factor in the deterioration of bridges. For instance, Huang et al. (2015) developed a Markov chain model by adding rainfall data to more accurately predict the conditions

of bridge components. In doing so, they used 189,247 inspection records from Victoria, Australia, and they determined that rainfall that exceeded 600 mm had a significant adverse impact on the conditions of railings and barriers. Similarly, Lee et al. (2014) studied 25 bridges located in Queensland, Australia to investigate the effects of humidity, temperature, and rainfall on the conditions of concrete slabs, and they found that peak climate conditions, such as the highest temperature or the maximum amount of rainfall deteriorated the conditions of slabs to a greater extent than annual average values. Dadson et al. (2002) also conducted analysis of variance (ANOVA), t-test, and cluster analyses to estimate the service life of paint on steel girder bridges in Virginia and, subsequently, to determine the effects of environmental conditions, e.g., temperature, rainfall, snowfall, and humidity, on service life. They found that paint in the region adjacent to the Atlantic Ocean and paint in regions with high humidity and high temperatures had the lowest service lives.

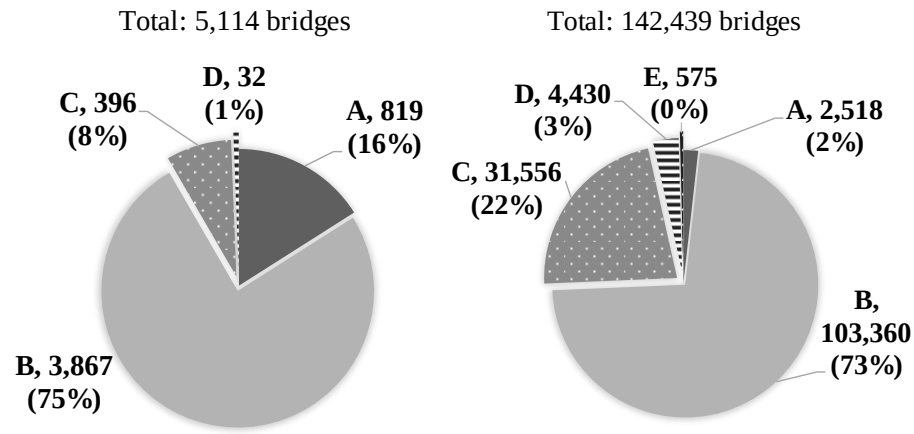
Based on the previous studies, diverse identification factors, detailed structural factors related to adjacent components, environmental factors including traffic and weather considering severe condition, and bridge age were need to be collected and considered to estimate bridge damage.

(2) Data Collection

The data about the bridge were collected from KOBMS, which is administrated by KICT. KOBMS includes 443,553 records collected from 10,187 detailed inspections and the precise safety diagnosis of 2,388 bridges built from 1966 to 2017 in South Korea. The inspections were conducted from 2003 to 2017. Matching the literature review findings, the research considered factors in four categories at different levels from KOBMS: i.e., identification factors (e.g., total length and the number of lanes), traffic factors (e.g., average daily traffic) at the entire bridge level, structural factors (e.g., the sizes and strengths of the girders) at the span level, and inspection factors at the damage level (e.g., damage type and condition level). The identification and structural factors are constant by time, since they are innate after construction, whereas the traffic and inspection factors change and accumulate with time. Among major bridge components such as decks, superstructures, and substructures, this research focused on the condition status of damage on bridge decks since they directly support dead and live loads on the bridge and consequently determine safety and serviceability of the bridge (Freyermuth et al., 1970; Morcous et al., 2002; Scott et al., 2003; Huang, 2010). Thus, the corresponding 142,439 (32%) deck inspection records of 2,388 bridges were extracted from the whole data set. Regarding their inspection factors, on the inspection date, the age of each bridge was calculated as the date of its last inspection minus the date its construction was completed. Even if the inspection data were discontinuous for each bridge, inspection data were attained for various ages based on the

sufficient number of inspections for numerous bridges.

The condition grades of an entire bridge were distributed as “A” (16.0%), “B” (75.6%), “C” (7.7%), “D” (0.6%), and “E” (0%) (Figure 3-2(a)). At the bridge damage level, condition states were distributed as “A” (1.8%), “B” (72.6%), “C” (22.2%), “D” (3.1%), and “E” (0.4%) (Figure 3-2(b)). The “C”, “D”, and “E” condition grades for an entire bridge occurred in quite small proportions (i.e., 8.3%) compared to the “A” and “B”; however, the proportion of the “C”, “D”, and “E” condition grades in the damage level was 25.7%, more than three times larger. This difference was due to the condition ratings being calculated as the weighted average of the condition state of damage.



(a) Condition Distribution at the Entire Bridge Level

(b) Condition Distribution at the Damage Level

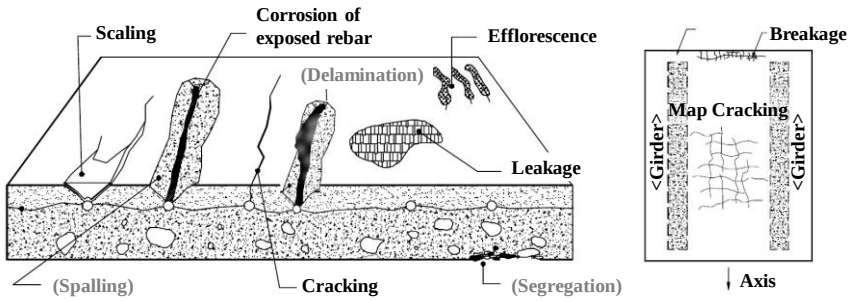
Figure 3-2. Distribution of Condition Grades

Seven types of deck damage were chosen: cracking, map cracking, scaling, breakage, leakage, efflorescence, and corrosion of exposed rebar. Detailed descriptions of each damage type was explained in Table 3-1. The shapes of selected damage are shown in Figure 3-3. Other damage types including contamination and material segregation were excluded due to the limited number of data for analysis. As described in Figure 3-4, the predominant damage was deck cracking in every structure type (over 49%).

Table 3-1. Descriptions of Types of Deck Damage

Deck Damage Type	Description
Cracking	A uni-directional linear fracture in concrete which extends partly or completely through the member
Map Cracking	Bi-directional and inter-connected cracks that form networks of varying size
Scaling	Also known as surface breakdown, the gradual and continuing loss of surface mortar and aggregate over an area
Breakage	A developed spalling in losing the original shape
Leakage	A water seeping through cracks or voids in the hardened concrete
Efflorescence	A combination of calcium carbonate leached out of the cement paste and other recrystalized carbonate and chloride compounds
Corrosion of Exposed Rebar	A rust seeks to occupy a much greater volume than the original steel and this expansion process leads to rebar exposure because of concrete damage and disintegration

Note: The contents were extracted from Main Roads Western Australia (2008) and Federal Highway Administration (FHWA) (2012).



Revised from the Guidebook (Ministry of Land, Infrastructure, and Transport of Korea (MOLIT) and Korea Infrastructure Safety Corporation (KISTEC), 2012)

Figure 3-3. Seven Types of Deck Damage

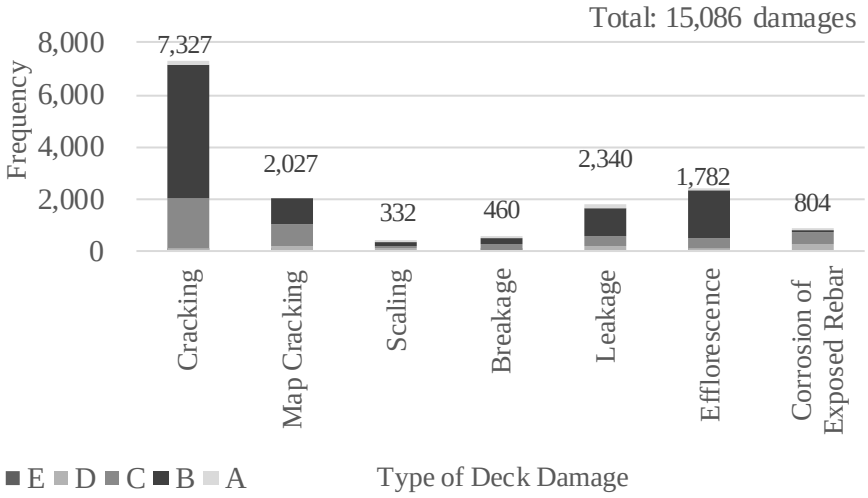


Figure 3-4. Condition Grade Distribution by Type of Deck Damage of Pre-stressed Concrete I Type Bridges

Weather data were obtained from the Korea Meteorological Administration (KMA). From 1973 to 2017, the monthly average of temperature, rainfall and snowfall, relative humidity, and daily wind speed were collected. The weather data from the nearest weather station were matched for each bridge. Moreover, the amount of chloride was collected based on the distance from the coastline (Jung et al., 2008).

Since the damage of bridges becomes more severe in extreme environments, such as heat waves, than in the average environment, the summer average and the winter average of temperature and relative humidity were calculated, and the number of days when heat waves, tropical nights, strong wind occurred were counted. Moreover, the differences between summer and winter average were also computed with respect to temperature and relative humidity. Various studies have directly or indirectly reported traffic volume and weather factors and considered severe weather conditions, such as the maximum temperature in a year and excessive rainfall in a year (Lee et al., 2014), which affects bridge components' deficiencies. There remains a need to study the effects of truck traffic and various kinds of severe weather conditions on different types of damage.

(3) Data Preprocessing

Collected data was preprocessed by data reduction, cleaning, and integration to improve data quality, and thus obtain more accurate analysis results (Han et al., 2012). First, data reduction was conducted to eliminate variables with a high missing rate (e.g., median barrier type with 38% missing values), variables whose values are same for entire data (e.g., road system type of 'general national highway'), and duplicated variables (e.g., main structure type duplicated by a code '23'). Variables related to rib flange (e.g., width of top rib flange) which are not present in PSCI were also removed. As a result, the original 112 variables were reduced into 61 variables consisted of 52 numerical variables (e.g., height, girder strength, average daily traffic) and 9 categorical variables (e.g., region, pavement type, damage condition grade). In the data cleaning process, noise was eliminated based on the distribution, and the 5 inspection data after the 2016 Gyeongju earthquake (2016.9.12.) was removed in the adjacent 8 cities. On the other hand, missing values were remained when using tree-based artificial intelligence algorithms due to the method has its own processing algorithm for missing values; they were replaced as the code '999' or value 0 to maintain their property. The last preprocessing step was data integration to the final generate dataset for analyses. The general and structural variables were combined according to the representative bridge number and the inspection variables were added based on the span or numbers of the bridge having the same bridge number.

Next, using variables that already existed, data transformation was conducted to generate new variables including severe weather conditions, location, damage severity extents. First, to represent severe weather conditions, the original weather data was transformed into the summer average (i.e., the average of data from June, July, and August), and the winter average (i.e., the average of data from December, January, and February) of temperature, precipitation, and relative humidity. The number of days when heat waves and tropical nights occurred (i.e., temperatures above 33 ℃ temperatures above 25 ℃, respectively), and when the strong wind over 25 m/s of an instant maximum wind speed happened, which value indicates a weak typhoon, were counted. Moreover, the differences between the summer and winter averages were additionally considered. In structural engineering, one of the causes of damage is repeated shrinkage and expansion, and therefore, the temperature difference should be considered as one variable in addition to the lowest and highest temperature.

In order to consider the maintenance history indirectly, the latest condition grades of the deck were taken into account. Since the maintenance history had been collected at the component level in the form of the texts of reports, it was difficult to match this record to the collected dataset at the damage level in tabular format. Therefore, three variables were generated to reflect the latest condition grades of the deck considering the time, i.e., 1) the representative condition grade of the entire deck at the latest inspection, 2) the condition grade

of a span of the deck at the latest inspection, and 3) the time difference between the end of the latest inspection and the current inspection.

A location variable was defined using the span number data from the KOBMS. Single span bridges were treated separately, and for multi-span bridges, the entire bridge was divided into five sections, i.e., the left, middle-left, center, middle-right, and right sections. For example, bridges with 16 spans were divided into five sections, i.e., four, three, two, three, and four spans, respectively, based on the equal-frequency binning of damage, i.e., placing damages that occurred with similar frequencies in separate bins.

The severity of damage was defined as the number of defects in a given range. The widths and lengths of the cracking and map cracking (e.g., less than a width of 0.1 mm for level A cracking and a width in the range of 0.1 to 0.3 mm for level B cracking). The sizes of the damaged areas were used to identify the ratings of the other types of damage (e.g., smaller than 2% area for level A efflorescence, and the area from 2% to 10% for level B efflorescence). Based on these criteria, defects on the deck can be counted. When multiple damages occurred in a given span, an evaluation record of each damage was stored in the KOBMS. That is, when three cracks were observed, three rows of data were stored. In the KOBMS data, the number of defects that was observed in one span varied from one to a maximum of 77, and, as the number of defects increased, the frequency of the occurrences decreased. Therefore, the severity of the damage is defined using equal-frequency binning (Table 3-2).

Table 3-2. Definition of the Damage Severity

Number of Defects	Damage Severity
0	Level 0
1	Level 1
2	Level 2
3-5	Level 3
≥ 6	Level 4

3.2.2. The Final Dataset for Analysis

Table 3-3 lists the 61 variables utilized in this study, along with their types and utilization. Compared to previous studies, this research considered more detailed identification, structural, and environmental factors (Table 3-3). Since damage type and damage condition grade were used to make partitions and act as target data (i.e., dependent variable), the final dataset of predictors (i.e., independent variables) comprised a tabular format of 59 variables (51 numerical variables and 8 categorical variables) and 142,439 records.

- 22 identification factors (e.g., region, competent authority, owner, crossing type, minimum radius of curve, design live load, vehicle weight limit, vehicle height limit, height, depth of water, total length, total width, effective with, total pavement area, skew angle at the beginning and at the end, number of upward and downward lanes, number of spans, span length between the center of supports, maximum span length between supports, and substructure type),
- 17 structural factors (e.g., thickness and strength of deck; type and thickness of pavement; waterproofing type; strength, diameter, and spacing of main reinforcing rebar; strength, height, quantity, and spacing of girders; thickness and width of top flange; thickness and width of bottom flange; web thickness)

- 18 environmental factors: traffic factors (e.g., ADT and ADTT), weather factors (e.g., averages of annual, summer, and winter temperatures and relative humidity; difference in average temperature and humidity between summer and winter; the number of days with heat waves and tropical nights; average of annual precipitation; summer rainfall; snowfall; the annual average wind speed; the number of days with strong wind; the amount of chloride), and
- 2 inspection factors (e.g., span number, age).

Table 3-3. List of Input Variables

Domain	Variables	Actual Values	Litera-ture	Type
Identi-fica-tion Factor	Region	6 metropolitan cities and 8 provinces	*	C
	Competent authority	Central/Local government, Professional institute	*	C
	Owner	5 Agencies and 19 offices	*	C
	Crossing type	None, Over water, Over land	*	C
	Minimum radius of curve	0–6500 (m)	+	N
	Design live load	DB–13.5 (24.3 ton), DB–18 (32.4 ton), DB–24 (43.2 ton), Foreign standards	*	C
	Vehicle weight limit	23–40 (ton)	+	N
	Vehicle height limit	2–4.61 (m)	+	N
	Height	2.3–54 (m)	+	N
	Depth of water	0–23 (m)	+	N
	Total length	5.9–2,380 (m)	*	N
	Total width	3.6–56.0 (m)	*	N
	Effective width	3.1–45.0 (m)	*	N
	Total pavement area	16–29,798 (m^2)	*	N
	Skew angle at the beginning	0–90 (deg.)	*	N
	Skew angle at the end	0–90 (deg.)	*	N
	Number of upward lanes	1–5	*	N
	Number of downward lanes	1–5	*	N
	Number of spans	1–48	*	N
	Span length between the centers of supports	2.9–210 (m)	*	N
	Maximum span length between supports	1–180 (m)	*	N
	Substructure type	8 types of pier and 5 types of abutment	*	C
Struc-tural Factor	Deck thickness	18–35 (cm)	*	N
	Deck strength	200–5200 (kg/mm^2)	*	N
	Pavement type	None, Concrete, Asphalt, Latex modified concrete (LMC), Etc.	*	C
	Pavement thickness	0–88 (mm)	+	N
	Waterproof type	None, Mortar, Asphalt, Sheet, Resin, Etc.	+	C
	Strength of main reinforcing rebar	1,600–5,700 (kgf/mm^2)	+	N
	Diameter of main reinforcing rebar	12.5–82 (mm)	+	N
	Spacing of main reinforcing rebar	100–480 (mm)	+	N
	Girder strength	210–5,700 (kg/mm^2)	*	N
	Girder height	23–600 (mm)	+	N
	Girder quantity	1–153	*	N
	Girder spacing	106–7,700 (mm)	*	N

Note: The asterisk denotes the variables used in the literature; The plus sign (+) indicates the variables additionally considered in this study.

Table 3-3. List of Input Variables (Continue)

Domain	Variables	Actual Values	Literature	Type
Structural Factor	Thickness of top flange	10–300 (mm)	+	N
	Width of top flange	20–700 (mm)	+	N
	Thickness of bottom flange	10–300 (mm)	+	N
	Width of bottom flange	25–700 (mm)	+	N
	Web thickness	10–520 (mm)	+	N
Environmental Factor	Average daily traffic (ADT)	277–68,072	*	N
	Average daily truck traffic (ADTT)	11–8741	*	N
	Annual average temperature	8.0-15.4 (℃)	*	N
	Summer average temperature	18.3-26.0 (℃)	+	N
	Winter average temperature	-5.8-5.6 (℃)	+	N
	Difference in average temperature between summer and winter	16.9-29.2 (℃)	+	N
	The number of heat waves	1-24 (days)	+	N
	The number of tropical nights	0-16 (days)	+	N
	Annual average precipitation	589-1789 (mm)	*	N
	Summer average rainfall	69-237 (mm)	*	N
	Snowfall	34.9-270.8 (mm)	*	N
	Annual average relative humidity	56.7-81.8 (%)	*	N
	Summer average relative humidity	61.6-91.3 (%)	+	N
	Winter average relative humidity	47.6-79.7 (%)	+	N
	Difference in average humidity between summer and winter	0.3-29.3 (%)	+	N
	Annual average wind speed	1.1-4.7 (m/s)	*	N
	The number of strong wind days	0-3 (days)	+	N
	Amount of chloride	1.8-5.3 (km/m^3)	+	N
Inspection Factor	Span number	1–48	+	N
	Age	1–58 (year)	*	N

Note: The asterisk denotes the variables used in the literature; the plus sign (+) indicates the variables additionally considered in this study.

3.3. Artificial Intelligence Methods

In this study, the artificial intelligence (AI) approach was used to connect the inspection data of similar conditions without averaging the conditions of individual bridges, which was what limited the previous research. As a white box model, the decision tree method classifies labels and simultaneously identifies the variables that affect the classification in a range of specific numerical values. Therefore, this method has been used frequently to determine the influencing factors in both classification and estimation problems, and it already had been used with bridge data (Melhem et al., 2003; Morcous, 2005; Lim and Chi, 2019a), so the possibility of using it in this study was confirmed. Deep Neural Networks (DNNs) have gained in popularity due to the unique training method involved, the structure for detecting nonlinear relationships between features, and their excellent performance in estimating results (Haykin, 1999). DNNs already have been used successfully to estimate the levels of various conditions by considering the complex nature of bridge-related variables with high accuracy (Tokdemir et al., 2000; Huang, 2010; Lim and Chi, 2019b).

3.3.1. Decision Tree Methods

A decision tree is a flowchart-like tree structure in which each internal node (non-leaf node) denotes a test of an attribute, each branch represents an outcome of the test, and each leaf node (or terminal node) holds a class label. Each internal node in the decision tree is labeled with an input feature. The leaves are labeled with different possible output values (Han et al., 2012).

The trees grow deeper with splitting at a node, which is determined by a certain generation algorithm. A single decision tree model, such as the Classification and Regression Tree (CART), must grow extensively to learn highly-irregular patterns, which can lead to overfitting and produce low bias, but it also produces high variance classification results. To avoid overfitting, Random Forest (RF) and Extreme Gradient Boosting (XGBoost) include ensemble strategies, i.e., bagging and gradient boosting, respectively (Zhang et al., 2018). The training concepts of each method are described in Figure 3-5. Since the appropriate methodology differs depending on the distribution of the data, various decision tree methods were compared in this study to identify the optimal method.

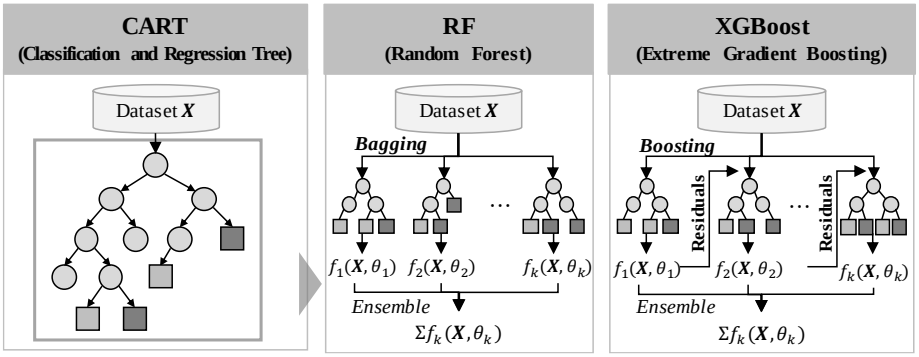


Figure 3-5. Training Concepts of CART, RF, and XGBoost

(1) Classification and Regression Tree

CART, introduced by Breiman et al. (1984), is constructed by splitting the dataset into subsets, and using all predictors to create two child nodes, beginning with the entire dataset. Using a Gini index, the best predictor is chosen to make the data subsets as homogeneous as possible with respect to the target variable (Ture et al., 2009). This study used CART algorithm for three reasons: 1) CART deals with both continuous and categorical attributes simultaneously to build a decision tree; 2) CART can handle both numerous variables and missing values without treatment; and 3) dissimilar to ID3 and C4.5 algorithms, CART does not use probabilistic assumptions owing to the use of a Gini index (Patidar and Tiwari, 2013).

CART algorithm first generates trees based on splitting criteria. Starting from the root node, trees are grown by repeatedly splitting the data with only one explanatory variable at each split. A Gini index acts as the stopping criterion

(Bektas et al., 2013). The index measures the impurity of dataset D, as in Equation (3.1):

$$Gini(D) = 1 - \sum_{i=1}^{m} {p_i}^2 \qquad (3.1)$$

where p_i is the probability that a tuple (i.e. row in a table) in D belongs to class C_i and is estimated by $n(C_{i,D})/n(D)$. $C_{i,D}$ means the set of tuples of class C_i in D, and $n(C_{i,D})$ indicates the number of tuples. The sum is computed over m classes. When considering a binary split, the Gini index is computed as a weighted sum of the impurity of each resulting partition. For instance, if a binary split on A divides D into D_1 and D_2, the Gini index of D_2 given that partitioning is in Equation (3.2) (Han et al., 2012). The value of the index increases as the impurity grows, and the Gini index is equal to zero for completely homogeneous nodes (Bektas et al., 2013) as in Equation (3.2):

$$Gini_A(D) = \frac{n(D_1)}{n(D)} Gini(D_1) + \frac{n(D_2)}{n(D)} Gini(D_2) \qquad (3.2)$$

(2) Random Forest

RF, one of the embedded methods, was used since it is known to generate stable performance with less risk of overfitting. As well as numerical variables, categorical variables can be dealt with the random forest after one-hot encoding, i.e., the number of categories creates the same number of columns, which are filled with one for the given category and zeros for the others. The random forest indicated a forest of random decision trees by bootstrapping samples using the Classification and Regression Trees classifier (Tuv et al., 2009). Given a training set, D, of d tuples, a training set D_i of d' ($< d$) tuples are sampled, while replacing D for each iteration i (i = 1, 2, …, k). The depth of tree is defined as the number of edges from the tree's root node. The process of generating the random forest includes: (1) growing a given number (i.e., *ntree*) of trees to the maximum depth using bootstrap samples from the training set; (2) finding the best split of variables by adjusting the number of randomly selected variables at each node with the predefined maximum number (i.e., *mtry*); and (3) utilizing the best split from the optimum number of possible variables.

(3) Extreme Gradient Boosting

XGBoost, proposed by Chen and Guestrin (2016), was also applied as one of the embedded methods due to its high accuracy and low risk of overfitting with the application of a boosting algorithm. XGBoost is based on gradient boosting, which generates a strong classifier by iteratively updating parameters of the former classifier to decrease the gradient of loss function (Friedman, 2001) as shown in Figure 3-6. A simple gradient boosting method has the advantage of high estimation accuracy, but is prone to overfitting.

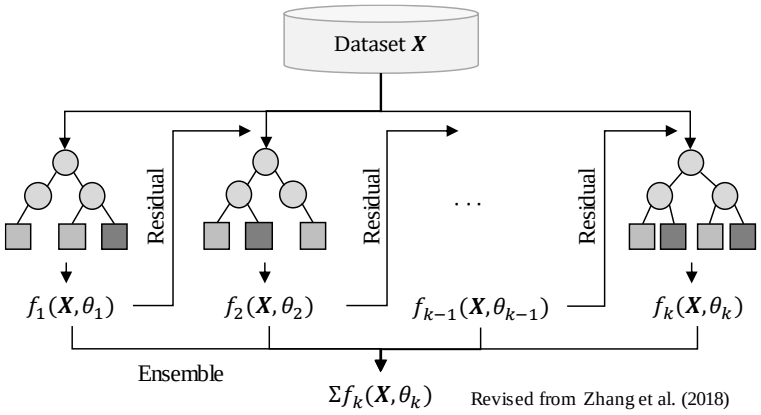


Figure 3-6. Concept of Extreme Gradient Boosting

To overcome this limitation, regularization is added to the loss function in the objective function of XGBoost, as in Equation (3.3):

$$J(\theta) = L(\theta) + \Omega(\theta) \tag{3.3}$$

where θ indicates the parameters trained from a set of given data. L is the training loss function, e.g., square loss or logistic loss, which measures the model's predictive ability. Ω is the regularization term, e.g., $L1$ norm or $L2$ norm, which measures the complexity of the model. The output of a model, $\hat{y}_i$, is averaged or voted by a collection F of K trees, as in Equation (3.4):

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), \qquad f_k \in F \tag{3.4}$$

where f is a scoring function to estimate output in the CART algorithm, and F is the set of all possible CARTs (Zhang et al., 2018). The figures in the leaves mean f values. During the boosting step, output at the t time iteration is defined as the sum of the scores, and the output is updated by adding a score of one new tree at a time, as in Equation (3.5):

$$\hat{y}_i^{(t)} = \sum_{k=1}^{t} f_k(x_i) = \hat{y}_i^{(t-1)} + f_k(x_i) \tag{3.5}$$

Therefore, the objective function at the t time iteration can be given as in Equation (3.6):

$$J^{(t)} = \sum_{i=1}^{n} L(y_i, \hat{y}_i) + \sum_{k=1}^{t} \Omega(f_k)$$
$$= \sum_{i=1}^{n} L(y_i, \hat{y}_i{}^{(t-1)} + f_k(x_i)) + \Omega(f_k) + constant \tag{3.6}$$

where n is the number of predictions. In XGBoost, the regularization term $\Omega(f_k)$ is given as in Equation (3.7):

$$\Omega(f_k) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} {w_j}^2 \tag{3.7}$$

where γ is the complexity parameter of each leaf. T is the number of leaves. λ is a parameter to scale the penalty. w is the vector of scores on leaves. Unlike the first-order in general gradient boosting, the second-order Taylor expansion is taken to the loss function in XGBoost. When using mean square error (MSE) as a loss function, the objective function of the t-th tree can be derived with removing constants as in Equation (3.8):

$$J^{(t)} \approx \sum_{i=1}^{n} \left[g_i w_{q(x_i)} + \frac{1}{2}\left(h_i {w^2}_{q(x_i)}\right) \right] + \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} {w_j}^2$$
$$= \sum_{j=1}^{T} \left[\left(\sum_{i \in I_j} g_i \right) w_j + \frac{1}{2} \left(\sum_{i \in I_j} h_i + \lambda \right) {w_j}^2 \right] + \gamma T \tag{3.8}$$

where $I_j = \{i|q(x_i) = j\}$ is the set of indices of data points assigned to the j-th leaf. Since all the data points on the same leaf get the same score, the index of the summation in the second line is changed. g_i and h_i are the first and second derivative of MSE loss function. Let $G_j = \sum_{i \in I_j} g_i$ and $H_j = \sum_{i \in I_j} h_i$, and the final objective function is transformed into a quadratic function as in Equation (3.9):

$$J^{(t)} = \sum_{j=1}^{T} \left[G_j w_j + \frac{1}{2}(H_j + \lambda) {w_j}^2 \right] + \gamma T \tag{3.9}$$

Therefore, the best w_j for a given $q(x_i)$ and the best objective reduction is derived as in Equation (3.10) and Equation (3.11), respectively.

$$w_j{}^* = -\frac{G_j}{H_j + \lambda} \tag{3.10}$$

$$J^* = -\frac{1}{2} \sum_{j=1}^{T} \frac{{G_j}^2}{H_j + \lambda} + \gamma T \tag{3.11}$$

Equation (3.11) measures the performance of the tree structure $q(x_i)$ (Zhang et al., 2018). The CARTs continue to branch off until the tree grows to the maximum depth. When to stop branching is determined by the following equation. When a tree splits a leaf into two leaves, the score of gain is derived as in Equation (3.12):

$$Gain = \frac{1}{2}\left[\frac{{G_L}^2}{H_L + \lambda} + \frac{{G_R}^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda}\right] - \gamma \qquad (3.12)$$

This formula is composed of four terms: 1) the score on the new left leaf, 2) the score on the new right leaf, 3) the score on the original leaf, and 4) regularization on the additional leaf. If the gain is smaller than γ, the tree stop growing (Xgboost developers, 2016).

3.3.2. Deep Neural Networks

A neural network allows the processing of information that is initiated by a densely interconnected and parallel distributed structure, as occurs in the brain (Aleksander and Morton, 1990; Huang, 2010). Figure 3-7 shows that a basic neural network is comprised of an input layer (I), a hidden layer (H), and an output layer (O) that are connected by neurons. A "deep" neural network indicates that the hidden layer consists of two or more layers (Winn, 2011).

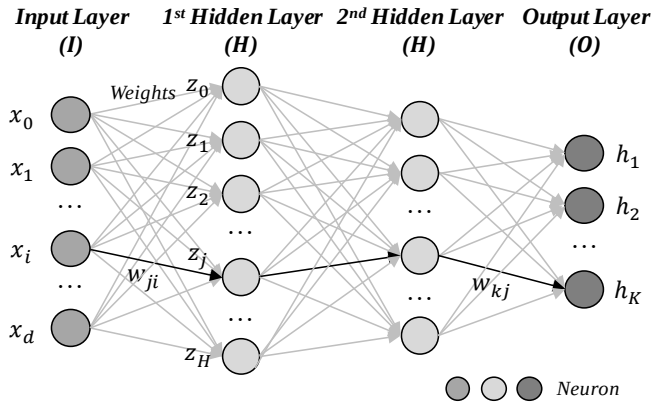


Figure 3-7. Concept of Deep Neural Networks

The nodes in the input layer collect the input signals. Each node in the hidden layers receives an activation signal, which is the weighted sum of all the inputs from the connected neuron, as presented by Equation (3.13):

$$z_j = \sigma\left(\sum_{i=0}^{d} w_{ij} x_i\right) \tag{3.13}$$

where z_j is the output of neuron j in the hidden layer (where j = 1, 2, ⋯, H); σ is the activation function; w_{ij} is the weight of the connection between j-th neuron in the hidden layer and i-th neuron in the input layer; and x_i is the i-th input in the input layer (i = 1, 2, ⋯, d). The neurons in the output layer produce output signals that are transformed again using a weighted sum of activation signals from the hidden neurons, as shown in Equation (3.14):

$$h_k = \sigma\left(\sum_{j=0}^{H} w_{jk} z_j\right) \tag{3.14}$$

where h_k is the predicted values of the k-th output (k = 1, 2, ⋯, K); and w_{jk} is the weight of the connection between the j-th and k-th neurons in the hidden and output layers. The error function at the output neurons is defined as the difference between the predicted value and target value, as shown in Equation (3.15):

$$e_k = h_k - y_k \tag{3.15}$$

where e_k is the error of the k-th output, and w_{jk} is the weight of the connection between the j-th and k-th neurons in the hidden and output layers.

Several algorithms have been developed to minimize errors, and the most popular algorithm, i.e., the back-propagation algorithm, was used in this study. The back-propagation process is divided into forward and backward phases. In the forward phase, the input values pass through the networks, and the activation values of all hidden and output neurons are computed. In the backward phase, all of the weights in the connections, initialized to small random numbers (e.g., between - 0.1 and 0.1), are updated by iterations until the error reaches the minimum value. The error calculated by a loss function , which is defined as the difference between the predicted values and the actual values (Bishop, 2006; Elhag and Wang, 2007). Loss function was fixed as a categorical cross entropy in this study to estimate the multiclass of the target.

To reduce the risk of overfitting, regularization that includes any technique to modify the complexity of the model to reduce errors in the testing phase is generally conducted. Several regularization techniques, including L2 regularization, batch normalization, dropout, and early stopping were applied in this study. L2 regularization, also called weight decay, simplifies the complexity of the model by reducing the weight of less important features. Before the input signals enter the activation function, batch normalization reduces the weight range through normalization using the weighted sum of the input signals. Dropout reduces the complexity of the model by omitting parts of the network. Early stopping is a technique that terminates the training of a model when performance is stabilized or has degrades (Goodfellow et al., 2016).

L2 regularization was applied to the input layer, and batch normalization and dropout were applied to hidden layers. Early stopping based on a training error was conducted.

3.4. Summary

In this chapter, the research methodology used in the research was introduced, the collection of data was described, and related methodologies were explored. First, the step-by-step methodology and the results required to achieve the four research objectives were presented, i.e., 1) data were collected from two sources, and multiple new variables were created; 2) the less influential variables were removed through correlation analysis, and the decision tree method was used to determine the influencing variables by damage-type factors; 3) a way of estimating the severity of damages was developed using DNN and XGBoost, and it was used to produce a damage portfolio with the specified final model; and 4) the same process was applied to the PSCI girder data to validate the possibility of expanding the model.

In order to collect data, previous research results were investigated to determine the variables that influence the deterioration of the condition of the components of the bridge as well as the entire bridge, and related variables, i.e., identification, structural, environmental, and inspection factors, were collected from KOBMS and KMA, and preprocessing was conducted to tune the input data. As a result, an input dataset was generated with 59 independent variables (52 numerical variables and 7 categorical variables) and 2 dependent variables (i.e., damage type and damage condition grade) for seven types of damage to

the decks of bridges.

Artificial intelligence methods were used to accomplish the second and third objectives. Three kinds of decision tree methods, i.e., CART, RF, and XGBoost, were reviewed to extract the variables that influenced the types of damage, and DNN was introduced as one of the approaches to develop a model for estimating the severity of the damage to bridges.

Chapter 4. Damage Influencing Variables Identification

In this chapter, a correlation analysis was conducted to reduce dimensions, and the influencing variables were identified for seven types of damage to PSCI bridges using either the CART, RF, or XGBoost decision tree methods, depending on which method had the best performance.

4.1. Redundant Variables Removal

4.1.1. Correlation Analysis Design

Feature selection, or dimensionality reduction, is a reduction of the size of the dataset in order to conduct time-efficient analysis, to improve performance, and to prevent overfitting in the model estimation (Han et al., 2012). There are three types of methods, i.e., the filter, wrapper, and embedded methods. The filter method calculates a defined numerical value (e.g., correlation coefficient, Gini coefficient), and it deletes redundant variables that have low values. The wrapper method repeatedly selects a different combination of variables based on accuracy when the combination is input to model estimation (i.e., genetic algorithm). The embedded method (e.g., random forest) includes the process of calculating the importance of variables in model estimation (Genuer et al., 2010; Jović et al., 2015).

To give decision makers intuitive insights related to bridge management, the correlation-based method was used to focus on removing redundant predictors without changing the original characteristics of the variables, which is a different approach from that used in Principal Component Analysis (Huang et al., 2012). Consequently, multicollinearity, which indicates a state of high inter-correlations among the independent variables, was eliminated, and the effects of individual variables were explained clearly (Reddy et al., 2013). The feature selection method was applied to the numerical predictors due to the possibility of redundancy.

Using correlation analysis, one of the highly correlated pair of variables that had less impact on the condition grades of the deck damage was removed. Because of the large number of data, Pearson's correlation coefficient can be used under the assumption of normality by the central limit theorem. According to Huang et al. (2012), the procedure for removing redundant predictors includes ranking, calculating, comparing, and removing. First, a total of 51 numerical variables were ranked according to their values of correlation with the damage condition grades, which were converted from categorical values, "A" to "E", into numerical values, 1 to 5. Then, the correlation coefficients between two numerical variables, r_{ij} (i, j = 1, ⋯, number of predictors), were calculated. The threshold was defined as 0.6 based on previous studies to prevent multicollinearity (Reddy et al., 2013), and, when $|r_{ij}| > 0.6$, the

lower-ranked variable was removed.

For categorical variables, Cramer's coefficient (V), which is a measure of the association between categorical variables (Cramér, 1946), was calculated for 7 categorical variables. Regional and operational variables (i.e., region, competent authority, owner) were excluded from analysis and remained to serve as labels for classifying the estimation results in future portfolios. Thus, five categorical variables, i.e., crossing type, design live load, substructure type, pavement type, waterproof type was analyzed. Based on Pearson's chi-square test, V measures the frequency difference between levels of a categorical variable (e.g., concrete and asphalt of pavement type) when a level of another categorical variable was fixed. For example, when a waterproof type is mortar, the frequencies difference between concrete and asphalt of payment type are compared and the hypothesis test is conducted using chi-square value. It can be concluded that the waterproof type is not related to the pavement type. In the case of the $k \times r$ contingency table, V is defined as shown in Equation (4.1):

$$V = \sqrt{\frac{\chi^2}{n * \min(k-1, r-1)}} \tag{4.1}$$

where χ^2 is derived from Pearson's chi-squared test, n is the number of samples, and k and r are the number of levels in each categorical variable (Colignatus, 2007).

4.1.2. Removal Results and Discussions

Table 4-1 shows the combinations of factors among 51 numerical variables where the correlation coefficients exceeded the threshold value of 0.6. The variables in bold font with asterisks in Table 4-1 were chosen because they ranked higher than the other variables. 11 variables were eliminated, i.e., total width, total pavement area, skew angle at the end, span length between the centers of supports, girder height, ADT, annual average temperature, summer average temperature, annual average precipitation, summer average relative humidity, and winter average relative humidity. The total area of pavement was ranked slightly higher than the total length, but, finally, the total length was selected. The reason was that the total area of the pavement was represented as width times length, but this was inappropriate because the entire width cannot be part of the total pavement area since the correlation coefficient between the effective width and the total pavement area was 0.255 where the value between the effective width and the total length was 0.071. Note that the variables with $r_{ij} < 0.6$ also were the inputs for the damage estimation model. Therefore, 40 numerical variables with 8 categorical variables were remained after this step.

Table 4-1. Redundant Variables Removal for the Deck by Correlation Analysis

Domain	Variables	Rank	Correlation Coefficient
Identification Factors	*Total width*	27	0.973
	Effective width*	13	
	Total length*	45	0.693
	Total pavement area	39	
	Skew angle at the beginning*	36	0.973
	Skew angle at the end	38	
	Maximum span length between supports*	21	0.897
	Span length between the centers of supports	30	
Structural Factors	**Maximum span length between supports***	21	0.703
	Girder height	44	
Environmental Factors	*Average daily traffic (ADT)*	15	0.818
	Average daily truck traffic (ADTT)*	5	
	Annual average temperature	45	0.691
	The number of tropical nights*	18	
	Annual average temperature	45	0.641
	Winter average temperature*	24	
	The number of heat waves*	10	0.638
	Summer average temperature	39	
	Annual average precipitation	32	0.744
	Summer average rainfall*	23	
	Annual average relative humidity*	31	0.840
	Summer average relative humidity	32	
	Annual average relative humidity*	32	0.769
	Winter average relative humidity	39	

Note: The italic font denotes the removed variables; the bold font with asterisk indicates the remained variables.

Conceptually similar variables had high correlation and variables with greater variation and severe conditions remained. Among the identification factors, the pairs with high correlations reflected the fact that the effective width is defined as excluding the width of the curb and rail from the total length, and the maximum length of a span is determined as excluding the width of support

from the span length between the centers of the supports. Among the variables related to the skew angle, the skew angle at the beginning, which had a larger coefficient of variation (i.e., the standard deviation divided by the mean) than the skew angle at the end, had more influence on the condition of the damage to the deck. In the same way, total width and span length between the centers of supports were removed. Despite the higher rank, total pavement area was removed instead of total length, because the pavement area, which is calculated by effective width times total length, cannot represent the effective width due to the small value of the correlation coefficient between them (0.225). For structural factors, the design basis for the ratio of span length to girder height were assigned to the highly-correlated pair. The variables associated with span length was removed instead of the girder related variables, which implied the importance of the girder which directly supports the load on the deck. With respect to the environmental factors, ADT, which has a smaller unit load compared to ADTT with larger unit load by trucks, was deleted and several variables representing the average values of climate conditions were removed. The number of tropical nights and heatwaves, which were more extreme temperature conditions, were retained when the annual and summer average temperature were eliminated. The average rainfall during the summer remained after the annual average precipitation was removed. However, for relative humidity, the annual average relative humidity was found to be more important than the summer and winter averages. Age in inspection factor remained as it

obviously affects deck damage according to previous structural and material studies.

As described in Table 4-2, the V values ranged from 0.033 to 0.371, indicating weak relationships, so all categorical variables remained to be analyzed.

Table 4-2. Cramer's Coefficient between Categorical Variables

Categorical Variables	Crossing Type	Design Live Load	Substruc-ture Type	Pavement Type	Waterproof Type
Crossing Type	1.000	-	-	-	-
Design Live Load	0.141	1.000	-	-	-
Substructure Type	0.195	0.115	1.000	-	-
Pavement Type	0.074	0.033	0.138	1.000	-
Waterproof Type	0.190	0.269	0.238	0.371	1.000

4.2. Influencing Variables Selection

4.2.1. Decision Tree Models Development

After using correlation analysis to remove the redundant variables, three kinds of decision tree methods, i.e., CART, RF, and XGBoost, were developed to select the influencing factors by the type of damage. The decision tree method was selected because of its transparency of interpretation and because it is known to produce stable performance with less risk of overfitting (Patidar and Tiwari, 2013).

(1) Model Design

The trees developed in this paper estimated damage condition grades (i.e., target) based on identification, structure, traffic, and inspection factors (i.e., predictors). Damage condition grades were divided into two classes, i.e., 1) less-damaged states including “A” and “B”, and 2) maintenance-required states from “C” to “E”, to transform the estimation problem into a binary classification problem to derive the damage influencing factors that cause more serious damage (Figure 4-1).

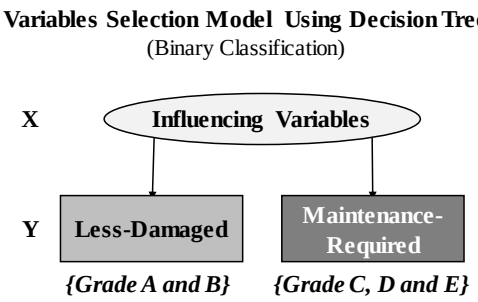


Figure 4-1. Model Design for Variable Selection

The dataset was split into training data (70%) and testing data (30%) while maintaining the ratio of maintenance-required states to less-damaged states, i.e., stratified splitting, and three-fold cross validation was conducted. Three-fold cross validation, one of the model validation technique, divides data samples into three parts in equal size; a third is reserved as the testing set and the remain partitions are collectively utilized to train the model. These steps are repeated three times (Han et al. 2012). Categorical variables were used as they were for CART, and converted by one-hot encoding for RF and XGBoost. Scalability and low performance issues of one-hot encoding were overcome by the nature of a decision tree. In the decision tree, each node is independent and its value is computed separately, so there is no collision between numerical variables and one-hot encoded variables.

The evaluation of the model was conducted by accuracy and F1 score. The accuracy of the model is the percentage of tuples on the given testing set that is classified correctly by the model. That is, accuracy focuses on the correctness

of the model irrespective of whether it is positive or negative. The F1 score simultaneously measured that the misclassification rate was low and that the correct classification rate was high. The threshold used in this study for evaluating the three indicators was 0.7 (good) (Kohavi and John, 1997).

(2) Model Verification

CART is vulnerable to overfitting due to noise. Breiman et al. (1984) and De'ath and Fabricius (De'ath and Fabricius, 2000) introduced three approaches for producing the best tree. The first approach is simple and involves growing an overlarge tree and then cutting it back (pruning). This method can be infeasible since the number of splits is often very large (Bektas et al., 2013). The second alternative considers the cost complexity of a tree, which is a function that consists of the quantity of leaves in the tree, $|T|$, and the tree's misclassification rate, $R(T)$. For any number α (≥ 0), which is called a complexity parameter, there is a unique smallest tree that minimizes $R(T) + \alpha|T|$. A high α value represents a small tree, whereas a low value leads to a complex tree (Therneau and Atkinson, 1997; De'ath and Fabricius, 2000). The third method selects the tree size by cross-validation to find the tree with the minimum error by comparing different training sets (Bektas et al., 2013). Using this method, the entire dataset is divided into a training set to develop the model and a testing set to evaluate the model. Commonly, the ratio of the training set to testing set is seven to three for a large dataset over 300 tuples (Urbanowicz and Browne, 2017). Practically, the second and third approaches are combined to find the best tree. Since it is computationally hard to calculate all tree cases, α was calculated for the limited number of cases based on the number of folds of cross-validation. The 1-SE (i.e., the standard error of the

minimal error) rule helps to identify the smallest tree, where α is less or equal to the sum of the minimal error and 1-SE (Therneau and Atkinson, 1997).

In the case of RF, for every tree grown in a random forest, about a third of the cases are out-of-bag (OOB), i.e., a set of samples that is not used to train the current tree. The OOB samples can be used to evaluate the prediction performance. The key hyperparameters of RF were *ntrees* (i.e., the number of trees in RF) and *mtry* (i.e., the number of randomly selected for split), which were determined by grid search (i.e., finding the best performance parameter by changing each parameter within an acceptable testing range) to minimize the prediction error of the OOB samples (Wang et al., 2016). First, to determine the baseline of the search, the grid search options were set in a wide range of *ntree* = [10, 50, 100, 200] and *mtry* = [5, 10, 20, 50, 88]. In the first trial, random forest generation was performed 20 times with four options for *ntree* and five options for *mtry*. Then, the generation of the random forest was repeated within a range around the value that showed the best performance of the previous trial until the range converged to a point. For instance, *ntrees* and *mtry* were determined to be 500 and 10, respectively, for cracking.

Seven hyperparameters of XGBoost, i.e., maximum depth, minimum child weight, gamma, subsample, colsample by tree, alpha for L1 regularization, and learning rate, were tuned for each damage type. The optimum values were calculated by grid search (Jain, 2016), and Table 4-3 provides descriptions of

all of the parameters and options for the grid search. The maximum number of iterations was 1,000. To reduce the risk of overfitting, early stopping was applied; when the error was the same for 50 times, the model stopped learning. For example, in the case of cracking, 69 times of trials in grid search were conducted and the best combination, which had the highest performance measures, was from the 56-th trial, and it was calculated as 10 for maximum depth, 1 for minimum child weight, 0.2 for gamma, 0.90 for subsample, 0.85 for colsample by tree, 0 for alpha, 0.1 for learning rate, and 179 iterations.

Table 4-3. Grid Search Inputs for Tuning Hyperparameters of XGBoost

Hyper-parameter	**Description**	**From**	**To**	**Interval**	
				1st	**2nd**
Maximum Depth	Maximum number of edges from the node to the tree's root node	4	10	2	1
Minimum Child Weight	Minimum sum of weights of all observations required in a child	1	5	2	1
Gamma	Minimum loss reduction required to make a split	0	0.5	0.1	-
Subsample	Ratio of observations for sampling to construct each tree	0.5	1	0.1	0.05
Colsample by Tree	Ratio of columns for sampling to construct each tree	0.5	1	0.1	0.05
Regularization Alpha	L1 regularization term on weights	0	0.3	0.01	-
Learning Rate	Aamount of updating weights	0.01	0.3	0.01	-

(3) Variable Importance Measures

To understand the drivers of damage estimation and compare the importance of individual variables, this research computed Shapley additive explanation (SHAP) values proposed by Lundberg et al. (2018). Although XGBoost supports traditional feature importance reporting, such as gain and split counts, these values have consistency issues; a feature with a high value in one tree tends to show a low value in another tree (Papadopoulos and Kontokosta, 2019). Based on game theory and conditional expectation, SHAP values are advantageous to show consistency for different individual trees (Calle and Urrea, 2010).

SHAP is an additive feature attribution method, in which a model's output is defined as a sum of real values attributed to each input variable. That is, an explanation model g is defined as a linear function of binary features as in Equation (4.2):

$$g(z') = \phi_0 + \sum_{i=1}^{M} \phi_i z_i' \tag{4.2}$$

where $z_i' \in \{0,1\}^M$, which represents a feature being observed ($z_i' = 0$) or unknown ($z_i' = 1$). M is the number of input variables. ϕ_i is the feature attribution values in the real range $\mathbb{R}$. To compute SHAP values, the output of a tree conditioned on a feature subset S is defined as $f_x(S) = [E(f(x)|x_s)]$. SHAP values average all possible conditional expectations based on game

theory, that is, prediction accuracy before and after the application of one variable, as defined in Equation (4.3):

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!\,(M - |S| - 1)!}{M!} [f_x(S \cup \{i\}) - f_x(S)] \qquad (4.3)$$

where S is the set of non-zero z' indexes. N is the set of all input features. According to the second term of Equation (4.3), a SHAP value can be negative, and a SHAP value is calculated for each tree model. Then, the average of the absolute SHAP values was used as a representative value for the entire ensemble tree (Papadopoulos and Kontokosta, 2019). A variable with the bigger SHAP value is more important than others. In order to better understand individual influence among the derived variables, this research used the relative value divided by the total sum (Lysenko et al., 2018). The research finally selected variables that had more than two percent of relative SHAP values by a set model trials.

4.2.2. Model Evaluation and Variables Selection

Table 4-4 shows the performance of the tree generation results using the CART, RF, and XGBoost methods. Most of the performance indicators were greater than the threshold of 0.7; the exception was the cracking F1 scores in both training and testing when using CART.

Table 4-4. Performance Results of CART, RF, and XGBoost for the Deck

Decision Tree Methods	Dataset	Measure	Damage Type							Average
			Cracking	Map Cracking	Scaling	Breakage	Leakage	Efflorescence	Corrosion of Exposed Rebar	
CART	Training	Acc.	0.88	0.85	0.82	0.76	0.87	0.89	0.88	0.85
		F1	0.60*	0.81	0.76	0.80	0.74	0.82	0.92	0.78
	Testing	Acc.	0.87	0.77	0.81	0.72	0.87	0.88	0.83	0.82
		F1	0.59*	0.75	0.73	0.77	0.72	0.81	0.89	0.75
RF	Training	Acc.	0.94	0.96	0.90	0.87	0.94	0.94	0.93	0.93
		F1	0.85	0.94	0.87	0.89	0.87	0.92	0.96	0.90
	Testing	Acc.	0.94	0.93	0.85	0.78	0.91	0.93	0.88	0.89
		F1	0.84	0.91	0.80	0.80	0.81	0.91	0.92	0.86
XGBoost	Training	Acc.	0.99	1.00	0.99	1.00	0.99	0.99	0.99	0.99
		F1	0.97	0.99	0.95	0.90	0.98	0.98	0.95	0.96
	Testing	Acc.	0.94	0.97	0.89	0.83	0.94	0.94	0.92	0.92
		F1	0.85	0.96	0.86	0.84	0.88	0.92	0.95	0.89

Note: The asterisk denotes a value less than 0.7.

The F1 score is a higher criterion that is harder to satisfy than the accuracy, and the performance of testing process, which is closer to the actual predicted case, is more important than the performance of training. Therefore, the F1 scores of three methods in the testing process were compared, and the performance of XGBoost was better than the other two tree-based methods, i.e., CART and RF (Figure 4-2). Hence, the XGBoost was the chosen method to select the influencing factors.

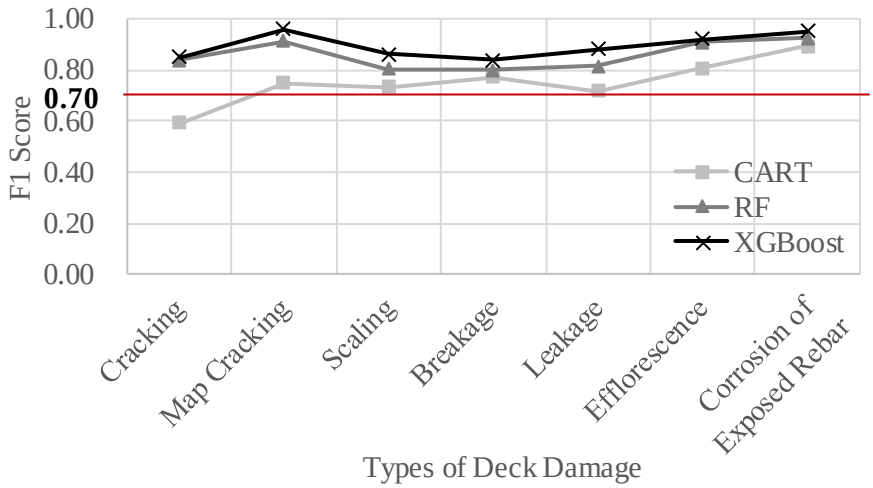


Figure 4-2. F1 Scores of CART, RF, and XGBoost in Testing of the Deck Data

Figure 4-3 shows an example of the trees that were generated for deck cracking. In the case of the 5-th iteration, the variables in a particular range were determined, such as the thickness of the deck being less than 21 centimeters. This reflected the significance of the change of the design criterion

for the thickness of the deck from 18 cm in 1977 to 22 cm in 2005 for the same design live-load criterion as DB-24 (i.e., 96 kN supported by the rear wheels). The predicted values in the squares indicate the probabilities of the classes. The positive values approach the maintenance-required class, and the negative values approach the less-damaged class. If a case entered the generated XGBoost trees, the output value was calculated by putting a sum of the values of the corresponding leaves into the logistic function. The class is determined as the maintenance-required class when the output value is greater than 0.5 and as the less-damaged class when the output value is less than 0.5. Several iterations were performed until the error was minimized, which was 250 iterations for cracking. When a case is entered for the model, the predicted values in 250 trees are combined with the sigmoid function to derive a value of 0 or 1, i.e., the final class is the less-damaged class (0) or the maintenance-required class (1).

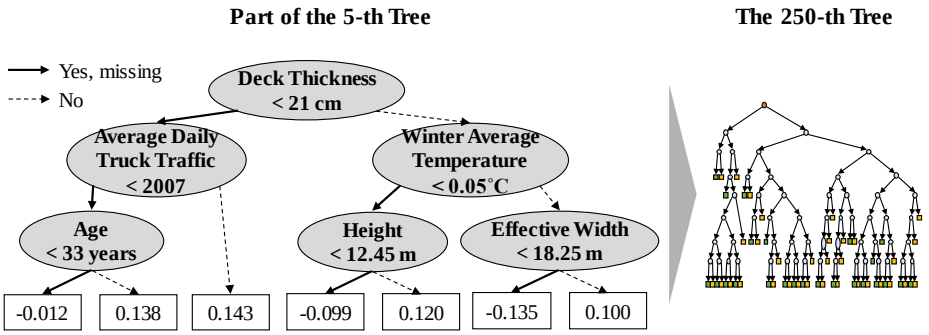


Figure 4-3. Tree Example of Deck Cracking Using XGBoost

After the generation of the tree, the values of the variables importance were calculated using the relative SHAP value. Table 4-5 shows the selected variables with the relative SHAP value in bold font with asterisks of 8 variables for deck cracking: vehicle weight limit, height, total length, deck thickness, spacing of main reinforcing rebar, ADTT, average annual relative humidity, and age.

Table 4-5. Relative SHAP Values (%) and Selected Variables of Deck Cracking

Domain	Variables	Relative SHAP (%)
Identification Factor	*Region*	0.43
	Competent authority type	0.38
	Owner	0.15
	Crossing type	0.76
	Minimum radius of curve	1.96
	Design live load	0.05
	Vehicle weight limit*	**7.92**
	Vehicle height limit	1.12
	Height*	**3.54**
	Depth of water	1.21
	Total length*	**4.04**
	Effective width	1.68
	Skew angle at the beginning	1.13
	Number of upward lanes	1.23
	Number of downward lanes	0.78
	Number of spans	0.97
	Maximum span length between supports	0.7
	Substructure type	0.96

Note: The bold font with asterisk denotes the selected variables; the italic font indicates the unchosen variables.

Table 4-5. Relative SHAP Values (%) and Selected Variables of Deck Cracking (Continue)

Domain	Variables	Relative SHAP (%)
Structural Factor	**Deck thickness***	**3.96**
	Deck strength	0.41
	Pavement type	0.05
	Pavement thickness	0.51
	Waterproof type	0.44
	Strength of main reinforcing rebar	0.65
	Diameter of main reinforcing rebar	1.74
	Spacing of main reinforcing rebar*	**3.55**
	Girder strength	1.15
	Girder quantity	1.10
	Girder spacing	1.80
	Thickness of top flange	0.63
	Width of top flange	1.05
	Thickness of bottom flange	0.44
	Width of bottom flange	0.26
	Web thickness	0.26
Environ-mental Factor	**Average daily truck traffic (ADTT)***	**9.37**
	Winter average temperature	1.79
	Difference in average temperature between summer and winter	1.28
	The number of heat waves	1.19
	The number of tropical nights	1.52
	Summer average rainfall	1.82
	Snowfall	1.66
	Annual average relative humidity*	**2.09**
	Difference in average humidity between summer and winter	1.34
	Annual average wind speed	1.65
	The number of strong wind days	0.01
	Amount of chloride	0.56
Inspection Factor	**Age***	**21.98**

Note: The bold font with asterisk denotes the selected variables; the italic font indicates the unchosen variables.

4.2.3. Selected Damage Influencing Variables

As a result of the XGBoost method for different types of damage, various number of different influencing variables was derived for each domain (Table 4-6). Cracking had the smallest number of influencing variables, i.e., 8, whereas corrosion of exposed rebar had the largest number of influencing variables, i.e., 17. Overall, a relatively larger number of influencing variables was derived from the identification factor and the environmental factor. The age from the inspection factor was derived as the top influencers of all types of damage except leakage.

As described in Table 4-7, the variables that were derived most frequently were vehicle weight limit, total length, annual average relative humidity, and age. The variables derived from 5 or 6 kinds of damage were height, girder strength, and ADTT. Therefore, it was confirmed that time (i.e., age), dead load (i.e., total length), live load (i.e., vehicle weight limit and average daily traffic), and relative humidity were important to induce damage to the concrete components of the bridge, including the deck (Main Roads Western Australia, 2008; Federal Highway Administration (FHWA), 2012). Even though the bridge is made of concrete, its weight increases due to the increase in the amount of material with a long length, and the weights associated with high traffic loads and the large number of trucks are added to the weight of the deck

of the bridge, ultimately causing damage. In addition, three variables among the five variables used in the deterioration model of the existing KOBMS were reconfirmed, i.e., age, ADTT, and average humidity. However, other variables, i.e., the maximum length of the span and the amount of chloride on the surface were identified as influencing variables only for the breakage and corrosion of exposed rebar. Since the girder is a distinctive component of PSCI bridges, the strength of a girder made of concrete was derived by a factor that is more important than the quantity and the spacing of the girders, which stand and divide the load on the deck.

Table 4-6. Number of Selected Influencing Variables by the Type of Deck Damage and the Domain

Domain	**Damage Type**						
	Cracking	**Map Cracking**	**Scaling**	**Breakage**	**Leakage**	**Efflorescence**	**Corrosion of Exposed Rebar**
Identification Factor	3	4	8	5	4	6	9
Structural Factor	2	4	3	4	2	1	3
Environmental Factor	2	5	4	3	6	5	4
Inspection Factor	1	1	1	1	1	1	1
Number of Influencing Variables	8	14	16	13	13	13	17

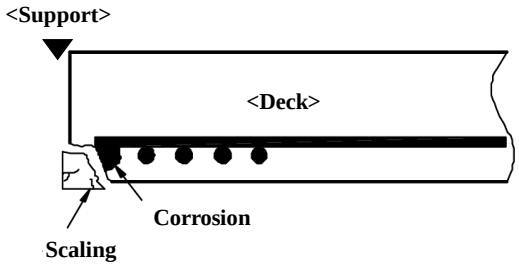
In terms of damage type, the results of the influencing variables that were extracted are discussed by groups of types of damage with similar properties, i.e., cracking and map cracking, scaling and breakage, and efflorescence and corrosion or exposed rebar. Leakage will be explained separately.

In addition to the common factors, the thickness of the deck and the spacing of the main reinforcing rebars were the major influencing factors for cracking. Thus, the structural strength of the deck was an important factor for the occurrence of one-directional cracks. In the case of map cracking, i.e., bi-directional cracks, several girder-related factors were derived, i.e., strength, quantity, spacing, and the width of the top flange. Among the environmental factors, the average temperature in the winter and snowfall were identified as the influencing variables for map cracking, and this corresponded to the results of previous research, which indicated that freezing and thawing, moisture, and the calcium chloride in the deicing chemicals that were related to the effects of snowfall on the condition of low temperatures (Federal Highway Administration (FHWA), 2002; Deschenes, Jr. et al., 2018). In summary, for one-directional cracks, the strength of the deck itself was important, but for more complex and larger cracks, i.e., map cracking, the structural conditions of the girder that distributes the load of the deck were significant. In addition, the winter conditions can result in the generation of map cracking and the enlargement of cracks that already exist.

Table 4-7. List of Influencing Variables by the Type of Deck Damage

Domain	Variables	Cracking	Map cracking	Scaling	Breakage	Leakage	Efflorescence	Corrosion of exposed rebar	Count
Identification Factor	Region							3	1
	Owner							16	1
	Competent authority						2		1
	Minimum radius of curve			9			13	2	3
	Vehicle weight limit	3	4	5	6	9	4	13	7
	Height	7	9	8	3	8	9		6
	Depth of water			16			10	11	3
	Total length	4	8	6	11	11	5	6	7
	Effective width			14		13		9	3
	Skew angle at the beginning		2						1
	Number of upward lanes			10	12				2
	Number of spans			11				10	2
	Maximum span length between supports				13			8	2
Structural Factor	Deck thickness	5				3			2
	Pavement thickness			15	2				2
	Girder strength		3	2	10	1	3		5
	Girder quantity		6						1
	Girder spacing		7		4			15	3
	Width of top flange		13						1
	Thickness of bottom flange				8				1
	Strength of main reinforcing rebar							7	1
	Diameter of main reinforcing rebar			4				4	2
	Spacing of main reinforcing rebar	6							1
Environmental Factor	Average daily truck traffic (ADTT)	2	14	7	7	6	6		6
	Winter average temperature		10					12	2
	Difference in average temperature between summer and winter			13		12			2
	The number of heat waves					10			1
	The number of tropical nights						11	14	2
	Summer rainfall		11				12		2
	Snowfall		5			5	8		3
	Annual average relative humidity	8	12	3	9	7	7	5	7
	Difference in average humidity between summer and winter			12	5	4			3
	Amount of chloride							17	1
Inspection Factor	Age	1	1	1	1	2	1	1	7
Number of influencing variables		8	14	16	13	13	13	17	34

Scaling and breakage are phenomena in which concrete is physically separated. For scaling, the radius of the curve and the number of spans were derived, which was supported by the scaling occurrence mechanism; when a bridge is curved, the strength of the surface concrete tends to weaken, and scaling can occur at the supports that meet the decks, as shown in Figure 4-4 (Ministry of Land, Infrastructure, and Transport of Korea (MOLIT) and Korea Infrastructure Safety Corporation (KISTEC), 2012). A large number of spans indicates a large number of supports, all of which have the possibility of the occurrence of scaling. The influencing factors in the case of breakage were derived, and they included girder-related factors, such as the spacing of the girders and thickness of the bottom flange of the girders. Since the girder-related variables were significant in map cracking as the developed damage of cracking, the girder-related variables were derived in breakage as the damage of scaling developed. In addition, the difference in the average humidity between summer and winter was derived for both scaling and breakage as compared to the other types of damage. This result corresponded to the fact that the strength of concrete particles can be weakened when they are repeatedly saturated with the moisture and then dried (Portland Cement Assosciation, 2002; Federal Highway Administration (FHWA), 2012).



Revised from the Guidebook (Ministry of Land, Infrastructure, and Transport of Korea (MOLIT) and Korea Infrastructure Safety Corporation (KISTEC), 2012)

Figure 4-4. Position of Scaling Occurrence

Leakage is the escape of water through a crack in the deck, so cracking should be prevented beforehand (Song et al., 2009; Federal Highway Administration (FHWA), 2012). Therefore, the thickness of the deck and the strength of the girder, which are influencing factors of cracking and map cracking, were determined to be the important variables that are associated with leakage. In addition, it was found that severe weather conditions can cause leakage, and these conditions include the temperature and humidity differences in summer and winter, the number of heat waves, and the amount of snowfall.

The selected influencing variables for the efflorescence and corrosion of exposed rebar, e.g., region, competent authority, and owner were different from those of other types of damage. A particular city with high temperature and high relative humidity was identified for efflorescence, and a metropolitan city on the coast was derived for corrosion of exposed rebar, which indicated the effect of chloride from sea salt and humidity. In addition, the depth of water, which

indicated the effects of moisture on concrete components, was derived for both damage. In the case of efflorescence, the number of tropical nights, summer average rainfall, and snowfall were selected by the model, which represented the movements of water molecules caused by high temperature in the summer and chloride ions from deicing chemicals in the winter. Variables related to the main reinforcing rebar were found to be associated with the corrosion of the exposed rebar, and the maximum length of the span between supports was derived, which is in proportion to the length of rebar and the radius of the curve that can influence the bending of the rebar. Regarding environmental conditions, the average temperature in the winter which leads to the shrinkage of rebar and the amount of chloride which accelerate corrosion were found to be influence factors for the corrosion of exposed rebar (Main Roads Western Australia, 2008; Federal Highway Administration (FHWA), 2000; Federal Highway Administration (FHWA), 2012).

4.3. Summary

In this chapter, the variables that influence the seven types of damage to the deck of PSCI bridges were identified by two steps. In the first step, a correlation analysis was performed to remove 11 redundant variables. Among the highly correlated variables, the variables with greater variation and severe conditions were remained.

Next, influencing factors by the different types of damages were selected using three different decision tree methods, i.e., CART, RF, and XGBoost, and XGBoost, which had the best performance, was chosen to derive the these factors. The most frequently identified variables that influenced most of target damage were time, dead load, live load, and relative humidity, and the strength of the girders was found to be a significant variable because it reflected the structural characteristics of the PSCI bridges. For the individual types of damage, e.g., cracking, the variables related to the strength of the deck itself were implicated, and, in the case of map cracking, girder-related variables, the temperature during the winter, and snowfall were implicated. For scaling, curves that can reduce the strength of the surfaces of the bridges and the number of spans were proportional to the number of potential locations for the occurrence of damage. The case of breakage was similar to map cracking in that girder-related variables were implicated in the development of damage

from scaling. For both scaling and breakage, the difference in the humidity levels in the summer and in the winter was identified as a major influencing variable. According to the fact that leakage arises after the occurrence of cracking, the variables related to the strength of the deck, which were derived as the influencing factors of cracking and map cracking, were also extracted as the important factors. Regional variables, the depth of the water, and high temperatures were identified as the causes of efflorescence and the corrosion of exposed rebar, and water and various ions also could be secondary causes. In the case of the corrosion of exposed rebar, the variables related to the length, diameter, and strength of the rebar, as well as the amount of chloride, were identified as variables that influenced the damage. In summary, various variables were derived that can represent the structural and material engineering information for each type of damage. The identified influencing variables were used to develop the damage estimation model in the following procedure.

Chapter 5. Damage Estimation Model Development

This chapter describes the process of developing a model to estimate the severity of damage, and the model was validated by generating lists of damages. To develop a model that can be used to estimate the severity of damages, two kinds of AI-based classification methods were used, i.e., DNN and XGBoost, and the evaluation of the performance of the model is discussed here. To validate the model, damage portfolios that were generated at the project level (i.e., an individual bridge) and at the network level (i.e., a region) are presented.

5.1. Artificial Intelligence Models Development

5.1.1. Model Design

The model was designed to estimate damage severity extents in two stages, which estimates the damage occurrence firstly and then estimated the severity of damage in the event of damage occurrence as shown in Figure 5-1. The damage severity extents were defined as the number of damage in particular ranges; a defect indicated the level 1, and more than six defects was the level 4. The range was determined by equal-frequency binning.

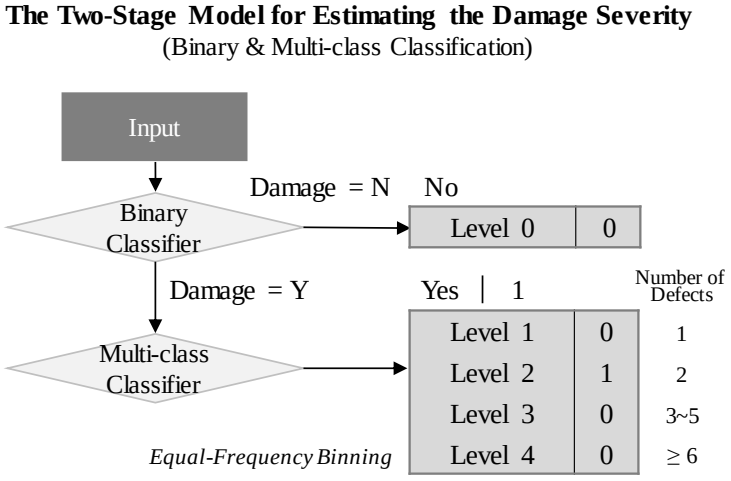


Figure 5-1. Model Design for Damage Severity Estimation

To prepare the input to the model, the format of the inspection data was reshaped from long type to wide type, as described in Figure 5-2. The tabular format of the inspection data from the KOBMS was recorded as one row that represented each damage (i.e., long-type format), and the KOBMS data did not include cases in which there was no damage. In order to calculate the number of defects and to represent the no-damage cases simultaneously, the data were transformed into the wide-type format, as shown in Figure 5-2. The wide format was designed to represent multiple rows of data in the long format in a single row for the particular inspection data associated with a particular span of a bridge. Figure 5-2 shows that the number of tuples in the long format of the same bridge, span, inspection period, type of damage, and condition grades was counted, and all of them were filled into the wide format. For example, two

rows of cracking in grade B were changed into the number 2 in the cell of cracking in grade B. The 142,439 data in the long format were converted into 8,397 data in the wide format. After tuning the dataset, the number of defects was converted into severity levels, ranging from level 0 to level 4.

Long Format

Bridge No.	Span No.	Inspection Period	Deck Damage Type	Condition Grade
25512	3	2016-05-24	**Cracking**	**A**
25512	3	2016-05-24	**Cracking**	**B**
25512	3	2016-05-24	**Cracking**	**B**
25512	3	2016-05-24	**Cracking**	**C**
25512	3	2016-05-24	**Mab Cracking**	**B**
…	…	…	…	…
25512	3	2016-05-24	**Corrosion of Exposed Rebar**	**B**
25512	3	2016-05-24	**Corrosion of Exposed Rebar**	**B**
…	…	…	…	…

Wide Format

Bridge No.	Span No.	Inspection Period	Cracking					Map Cracking					…	Corrosion of Exposed Rebar				
			A	B	C	D	E	A	B	C	D	E	…	A	B	C	D	E
25512	3	2016-05-24	**1**	**2**	**1**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	…	**0**	**2**	**0**	**0**	**0**
…	…	…	…															

Figure 5-2. Reshaping the Dataset from the Long Format to the Wide Format

Submodels were developed to estimate the severity of the damage in five classes, i.e., level 0 to level 4, by seven types of damage and five grades of condition. Therefore, the number of possible submodels was 35, and the entire model was referred as “the model” (Figure 5-3).

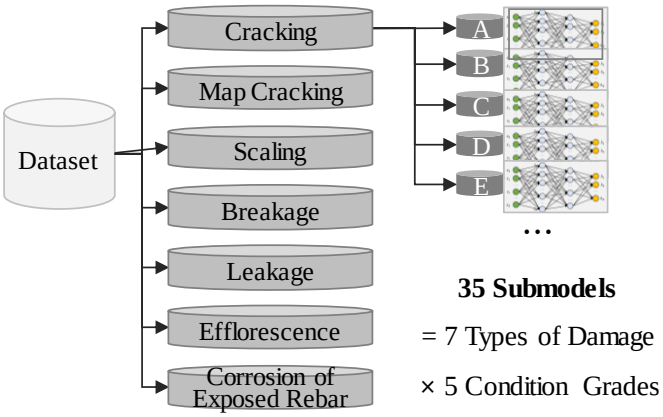


Figure 5-3. Structure of the Damage Estimation Model

Since the datasets in this study were highly biased, performance evaluation metrics that were independent of the class distribution were selected (Hammerla et al., 2016). For multi-class classification, in addition to accuracy, the weighted values of precision, recall, and F1 score were used, as defined in Equations (5.1), (5.2), and (5.3), where N_c is the number of data in class c (Hammerla et al., 2016; Scikit-Learn Developers, 2019a).

$$P_w = \frac{1}{N_{total}} \sum_c N_c \, Precision_c \tag{5.1}$$

$$R_w = \frac{1}{N_{total}} \sum_c N_c \, Recall_c \tag{5.2}$$

$$F_w = 2 \sum_c \frac{N_c}{N_{total}} \frac{Precision_c \times Recall_c}{Precision_c + Recall_c} \tag{5.3}$$

5.1.2. Model Verification

After each dataset was divided into training and testing sets in the ratio of 7:3 (Lee et al., 2008), the optimum values of the seven DNN hyperparameters (i.e., depth, node, epoch, activation function, optimizer, learning rate, and decay) were determined by a grid search using three-fold cross validation. Table 5-1 describes the detailed parameters. Depth and node determine the complexity of the network structure. Depth, i.e., the number of hidden layers, was considered to be two or more, which made the neural networks model "deep". Batch size and epoch decided the size and number of iterations for the training units that are parts of the training data set. For example, batch size and epoch were selected as 10 and 50 respectively with 200 samples of a dataset. That is, the dataset is divided into 20 batches, and each batch contains 10 samples. The model weights are updated 20 times in each batch, and the total training is 1,000 times. Activation function transfers the input values to an output value, and the function of the last hidden layer was determined as the softmax function for classification. To find the optimal performance model, different activation functions with different output ranges (e.g., (0, 1) for sigmoid, (-1, 1) for softsign and tanh, (0, ∞) for softplus, and [0, ∞) for ReLU) were applied in the grid search. Optimizer indicates an algorithm that updates weights using parameters, learning rate and decay (Goodfellow et al., 2016). The training

process was stopped when one of the following conditions was met, i.e., (1) the maximum number of epochs was reached; (2) the loss function was minimized; or (3) the performance gradient fell below the minimum specified value (Winn, 2011).

XGBoost is used both for feature selection by taking advantage of the white box model and for classification or estimation (Zhang et al., 2018). Therefore, XGBoost also was used to develop a model that could be used to estimate damage. To optimize hyperparameters of XGBoost, the same process of verification, that was explained in Section 4.2.1. Decision Tree Models Development, was performed.

Table 5-1. Grid Search Inputs for Tuning Hyperparameters of Deep Neural Networks

Hyper-parameter	**Description**	**Trials**
Depth	Number of hidden layers	[2, 3, 4, 5]
Node	Number of nodes in each hidden layer	[32, 50, 70, 100, 150, 200]
Batch Size	Number of samples propagated through the network	[32, 50, 70, 100, 150, 200]
Epoch	Number of iterations	[30, 50, 70, 100, 150, 200]
Activation Function	A function defines the output of that node given an input or set of inputs	Sigmoid, Softsign, Tanh, Softplus, ReLU
Optimizer	An algorithm for updating the weight parameters to minimize the loss function	Stochastic Gradient Descent (SGD), Adaptive Gradient (Adagrad), RMSprop (Root-Mean-Square prop), Adadelta (Adaptive Delta), Adam (Adaptive Moment Estimation)
Learning Rate	Amount of updating weights	[0.001, 0.01, 0.1]
Decay	Regularization term of updating weights	[0.001, 0.01, 0.1]

5.2. Model Evaluation

5.2.1. Model Performance Comparison

(1) Estimation Results

Table 5-2 shows examples of the test results from the model that was developed. In the first row, cracking was estimated to have severity level 4 in grade B and severity level 3 in grade C, and both estimates were correct. In the last row, efflorescence was estimated to have severity level 4 in grades B and C, but the actual values were level 3 in grade B and level 2 in grade C, which showed two cases of incorrect estimates. As in the third case of Efflorescence, there was a case in which the occurrence of damage was estimated, but, in reality, there was no damage (i.e., a false alarm). Conversely, cases that estimated no damage when damage actually had occurred (i.e., miss) occurred less frequently than the false alarm. That is, the model was conservative in its estimation of damage, and this will be addressed in Section 5.2.2. Model Performance Evaluation.

Table 5-2. Example of the Estimation Results

Damage Type	Bridge No.	Inspection Date	Age	Location	Estimation					Actual					Total Difference
					A	B	C	D	E	A	B	C	D	E	
Cracking	030039	2015-06-24	19	Middle Left		L4	L3				L4	L3			0
	032058	2014-08-30	15	Right					L1					L1	0
	028151	2012-07-15	9	Middle Right				L4					L3		1
	033150	2004-06-15	2	Middle Right	L2	L1				L1	L2				2
Efflorescence	001342	2015-06-09	12	Left				L4					L4		0
	028191	2016-06-21	12	Middle Right			L3					L2			1
	032005	2006-06-26	10	Right		L1									1
	002545	2007-10-23	21	Middle Left		L4	L4				L3	L2			2

Note: The gray cells indicate Level 0 (no damage).

(2) Performance Comparison of DNN and XGBoost

As a result of developing models to estimate the severity of damage, using DNN and XGBoost, 22 submodels were generated out of a total of 35 possible combinations of seven types of damage and five condition grades. Comparing the average performance when both models were tested, it was apparent that XGBoost had better overall performance than DNN (Table 5-3).

Table 5-3. Performance Comparison (%) of DNN and XGBoost Models in Estimating Deck Damage

Method / Measure	DNN	XGBoost
Accuracy	62.90	95.13
Precision	91.19	94.16
Recall	62.89	95.17
F1 Score	73.56	94.48

Note: Precision, Recall, and F1 score were weighted values.

Among the four average performance measures, the precision values (i.e., true positive ratio to actual results) were similar, while the recall of the DNN model (i.e., true positive ratio to predicted results) was much lower than that of the XGBoost model, which indicated that the DNN model tended to overestimate. As a result, the average accuracy and the average recall of DNN were around 20 - 30% lower than these values for XGBoost, and this phenomenon also was observed in the submodels (Table 5-4, Table 5-5, Table 5-6, and Table 5-7).

Table 5-4. Accuracy (%) of the DNN Model for the Deck in Testing

Damage Types	A	B	C	D	E	Average
Cracking	65.5	39.1	61.0	59.7	68.3	58.7
Map Cracking	N	64.3	59.6	69.6	N	64.5
Scaling	N	62.3	67.6	68.8	N	66.2
Breakage	N	61.6	58.3	N	N	60.0
Leakage	N	56.7	66.2	64.0	N	55.7
Efflorescence	N	43.5	58.2	65.3	N	62.3
Corrosion of Exposed Rebar	N	60.2	55.5	58.9	N	58.2
Average	65.5	55.4	60.9	64.4	68.3	62.9

Note: N means not generated. [0,85) [85,100] (%)

Table 5-5. Accuracy (%) of the XGBoost Model for the Deck in Testing

Damage Types	A	B	C	D	E	Average
Cracking	98.1	63.4	89.7	99.3	99.6	90.0
Map Cracking	N	90.3	93.9	99.3	N	94.5
Scaling	N	93.2	97.3	98.4	N	96.3
Breakage	N	93.8	93.5	N	N	93.6
Leakage	N	81.8	95.9	99.4	N	92.4
Efflorescence	N	79.0	94.3	99.0	N	90.8
Corrosion of Exposed Rebar	N	96.5	93.0	96.1	N	95.2
Average	98.1	85.4	93.9	98.6	99.6	95.1

Note: N means not generated. [0,85) [85,100] (%)

Table 5-6. Recall Rate (%) of the DNN Model for the Deck in Testing

Damage Types	A	B	C	D	E	Average
Cracking	65.5	39.1	61.0	59.7	68.3	58.7
Map Cracking	N	64.3	59.6	69.6	N	64.5
Scaling	N	62.3	67.5	68.8	N	66.2
Breakage	N	61.6	58.3	N	N	60.0
Leakage	N	56.7	66.2	64.0	N	62.3
Efflorescence	N	43.5	58.1	65.3	N	55.6
Corrosion of Exposed Rebar	N	60.2	55.5	58.9	N	58.2
Average	65.5	55.4	60.9	64.4	68.3	62.9

Note: N means not generated. [0,85) [85,100] (%)

Table 5-7. Recall Rate (%) of the XGBoost Model for the Deck in Testing

Damage Types	A	B	C	D	E	Average
Cracking	98.3	63.5	89.6	99.3	99.7	90.1
Map Cracking	N	90.2	93.7	99.4	N	94.4
Scaling	N	93.5	97.3	98.4	N	96.4
Breakage	N	93.1	93.5	N	N	93.3
Leakage	N	82.0	95.8	99.4	N	92.4
Efflorescence	N	79.2	93.9	99.1	N	90.7
Corrosion of Exposed Rebar	N	96.3	93.0	96.1	N	95.1
Average	98.3	85.4	93.8	98.6	99.7	95.2

Note: N means not generated. [0,85) [85,100] (%)

The performance of the XGBoost model was superior to that of DNN, because the learning algorithms were different. XGBoost had better performance than DNN on the class-imbalanced data of this study, where the damage occurrence data (i.e., the damage severity from level 1 to level 4) were less than 11.9% as shown in Table 5-8. These results originated from the boosting step of XGBoost in the learning process. Boosting is iteratively generating a tree by assigning more weight to the misclassified samples to be selected for a next tree, as shown in Figure 3-6, so the XGBoost model trained repeatedly to classify the minor classes. However, the DNN model requires a certain amount of data to train complex networks with layers and nodes, after which it showed low performance.

Table 5-8. Percentage (%) of Damage Data for the Deck

Damage Types	A	B	C	D	E
Cracking	3.6	43.2	11.9	0.7	0.3
Map Cracking	0.1	10.8	6.4	0.8	0.3
Scaling	0.1	6.5	2.7	1.5	0.0
Breakage	0.1	6.1	6.5	0.3	0.0
Leakage	3.0	19.9	5.1	0.6	0.0
Efflorescence	0.0	21.0	7.7	1.2	0.3
Corrosion of Exposed Rebar	0.1	3.5	7.1	4.2	0.0

Note: The total number of data in the wide format was 8,397.

[0,15) [15,100] (%)

5.2.2. Model Performance Evaluation

After selecting the XGBoost model, the performances of the submodels were evaluated, and the causes of the differences in performance were analyzed. Since the F1 score is defined by a harmonic average of precision and recall, in this research, this measure was mainly considered to simultaneously consider precision and recall. The F1 scores of the submodels are presented in Table 5-9. Submodels were not generated for most of grade A and E cases. Almost all of 22 submodels showed high f1 score over the 85% except three submodels including cracking, leakage, and efflorescence in grade B.

Table 5-9. F1 Score (%) of the XGBoost Model for the Deck in Testing

Damage Types	A	B	C	D	E	Average
Cracking	98.0	62.7	88.5	99.0	99.6	89.56
Map Cracking	N	88.8	92.9	99.1	N	93.60
Scaling	N	91.5	96.6	97.6	N	95.23
Breakage	N	91.9	91.0	N	N	91.45
Leakage	N	80.4	95.0	99.2	N	89.83
Efflorescence	N	77.3	93.4	98.8	N	91.53
Corrosion of Exposed Rebar	N	95.1	91.7	95.1	N	93.97
Average	98.00	83.96	92.73	98.13	99.60	94.48

Note: N means not generated. [0,85) [85,100] (%)

There are two reasons why submodels were not created as follows. The first reason was that the minimum number of data in severity classes was not satisfied. During the training process, three-fold cross validation was conducted, which requires a minimum number of six tuples, which was to come from triplicate values of two phases (i.e., training and testing) because of three-fold cross-validation. Therefore, if there were classes that had less than six tuples of the severity of damage, a model was not generated. As described in Table 5-10, data of map cracking in grade A did not have level 4 of severity data.

The second reason that there is a certain range of damage grades, which was present in the actual inspection, depending on the type of damage. That is, map cracking, scaling, breakage, efflorescence, and corrosion of exposed rebar are unlikely to occur on a small scale. In addition, because the maintenance was done before developing to grade E, the aforementioned damage and leakage are difficult to find. It also can explain the case of breakage in grade D. For these reasons, it is difficult to generate submodels with a small amount of data.

Table 5-10. Distribution of the Deck Damage Data for Cases without Submodels

Occurrence	Grade	N	Y				Grade	N	Y			
Severity Extents		L0	L1	L2	L3	L4		L0	L1	L2	L3	L4
Map cracking	**A**	8388	9				**E**	8374	23			
			2	1	6	0			12	1	5	5
Scaling	**A**	8385	12				**E**	8396	1			
			4	0	2	6			0	0	0	1
Breakage	**A**	8392	5				**E**	8397	0			
			4	0	1	0						
Leakage	**A**	8141	256				**E**	8397	0			
			235	17	3	1						
Efflorescence	**A**	8393	4				**E**	8370	27			
			1	2	1	0			20	0	0	7
Corrosion of Exposed Rebar	**A**	8390	7				**E**	8396	1			
			6	0	1	0			1	0	0	0
Breakage	**D**	8369	28									
			20	3	5	0						

Note: The total number of data in the wide format was 8,397.

The differences in the performances of the submodels were explained using silhouette coefficients. In the clustering analysis for a dataset, D, each object, $i \in D$, is assigned into k clusters, i.e., $C_1, \cdots, C_k$. The silhouette coefficient is a measure of how close each point in one cluster is to the points in adjacent clusters, as defined in Equation (5.4).

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \tag{5.4}$$

where $a(i)$ is the average distance between i and all other objects in the cluster to which i belongs, and $b(i)$ is the minimum average distance from i to all clusters to which i does not belong. The range of the coefficient is [-1, 1]. Values near the +1 value indicate that the object is far away from the neighboring clusters (i.e., well clustered), and a value of 0 means that the object lies on or very close to the decision boundary between two neighboring clusters. Negative values indicate that those objects might have been assigned to the wrong cluster (Han et al., 2012; Scikit-Learn Developers, 2019b).

In this study, the silhouette coefficient was used to confirm that clusters of the damage severity were well separated based on the distribution of their predictors. As described in Table 5-11, the absolute silhouette values of cracking, leakage, and efflorescence in grade B were 0.03, 0.04, and 0.02, respectively, which were relatively lower than the average values of 0.19 for the other cases, which corresponded to the cases with the relatively low F1 scores in Table 5-9.

Table 5-11. Silhouette Coefficient of Severity Classes of the Deck Data

Damage Types	A	B	C	D	E
Cracking	-0.30	-0.03	-0.13	-0.43	-0.15
Map Cracking	-0.33	-0.10	-0.10	-0.23	-0.47
Scaling	-0.23	-0.26	-0.09	-0.17	-0.36
Breakage	-0.30	-0.13	-0.12	-0.18	-
Leakage	-0.20	-0.04	-0.10	-0.13	-
Efflorescence	-0.42	-0.02	-0.06	-0.19	-0.39
Corrosion of Exposed Rebar	-0.29	-0.13	0.06	0.06	0.18

$|s| < 0.05$ $|s| \geq 0.05$

This was because small absolute values of the silhouette coefficient indicate that the objects were located near the boundary of the severity clusters, which makes it difficult to determine the distinctive characteristics of each severity class. Most of the silhouette values were negative, and some of the absolute values were relatively large, i.e., as large as 0.47, but these cases were well learned by XGBoost due to the boosting algorithm.

5.3. Model Validation

5.3.1. Damage Portfolio Generation

Using the damage estimation model that was developed, a system was designed to generate inspection support information at the project level (i.e., individual bridges) and at the network level (i.e., a region). The user interface was designed so that users who require damage information in advance easily can use the model. The estimation results for the entire bridge were evaluated to validate the model.

At the project level, a damage portfolio for a particular bridge was computed, as shown in Table 5-12. Basic information and the expected inspection date, bridge number, region, main structure type of the bridge, and age are displayed, and the estimated levels of severity are described in a tabular format by the type of damage and condition grade for each location. Users can check the distribution of damage by location and the importance of each damage by the associated level of severity.

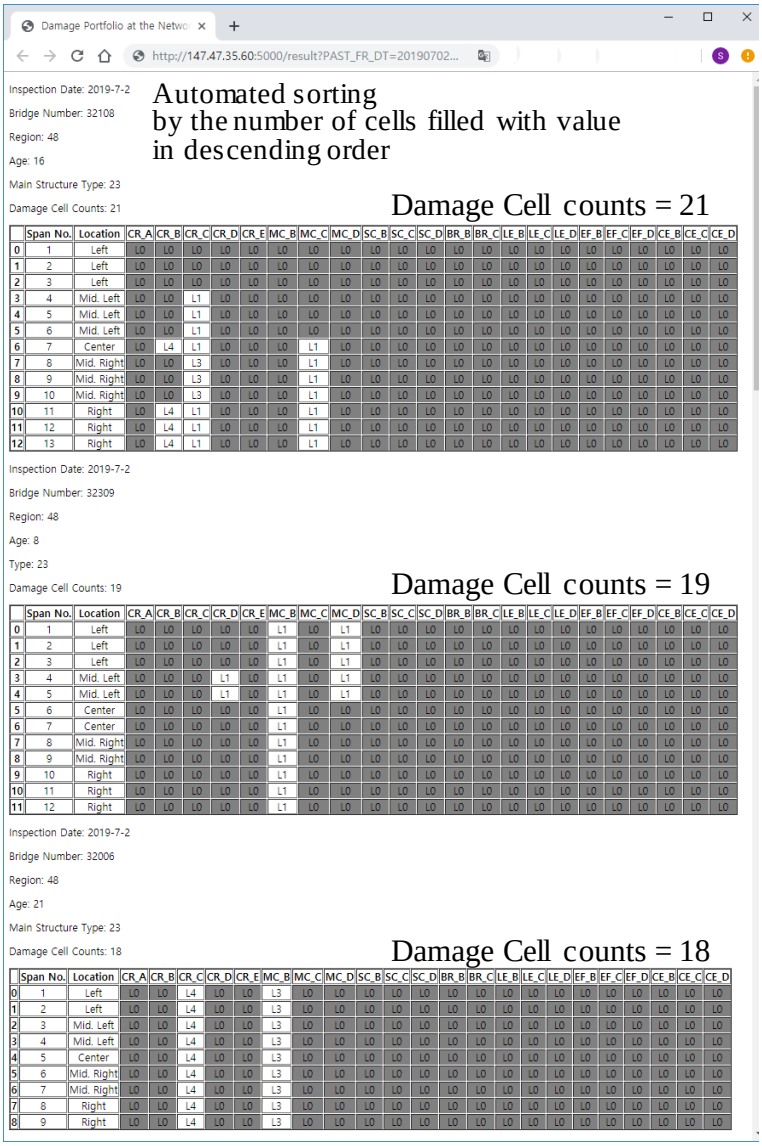
Table 5-12. Example of Deck Damage Portfolio for a Bridge (Project Level)

- Expected Inspection: 2019.7.2.
- Bridge Number: 000495
- Region: Gangwon-Do
- Main Structure Type: Pre-Stressed Concrete I
- Age: 28 years

Location		Left							Middle Left							Center							Middle Right							Right						
Span Number		1							2							3							4							5						
Damage Type		CR	MC	SC	BR	LE	EF	CE	CR	MC	SC	BR	LE	EF	CE	CR	MC	SC	BR	LE	EF	CE	CR	MC	SC	BR	LE	EF	CE	CR	MC	SC	BR	LE	EF	CE
Condition Grade	**A**																																			
	B									L3				L1							L1			L1		L3		L1					L3			
	C		L1				L1													L1							L1	L1							L1	L1
	D		L2					L1													L4	L4							L1						L4	L1
	E																																			

Note: Cracking (CR), Map Cracking (MC), Scaling (SC), Breakage (BR), Leakage (LE), Efflorescence (EF), Corrosion of Exposed Rebar (CE).
The cells with diagonal lines indicate the submodels were not generated. The gray cells indicate the L(evel) 0 (no damage). The damage severity (number of defects) is represented by: L1 (1), L2 (2), L3 (3-5), L4 (≥6).

When performing actual inspections, an inspector will be in charge of multiple bridges, so in order to enhance the quality of the inspections within a limited time, information about the most-damaged bridge among the bridges to be inspected will be helpful to the inspector. Therefore, the system was designed to generate multiple portfolios in a specific region where inspections are scheduled, which indicates a network level analysis. Figure 5-4 shows that the portfolios were sorted automatically by the number of cells filed with the values in descending order. The bridges with higher marks are more likely to be damaged, which requires additional investment in inspection resources (i.e., time, cost, and inspectors).



Note: Cracking (CR), Map Cracking (MC), Scaling (SC), Breakage (BR), Efflorescence (EF), Leakage (LE), Corrosion of Exposed Rebar (CE).
The gray cells indicate the L(evel) 0 (no damage). The damage severity (number of defects) is represented by: L1 (1), L2 (2), L3 (3-5), L4 ($\geq$6).

Figure 5-4. Example of Damage Portfolio for Bridges in a Specific Region (Network Level)

5.3.2. Estimation Performance Evaluation

To confirm that the model reflects actual situations, the estimated results were compared with the actual results using the entire model, which is composed of 22 submodels. A total of 3,001 tuples organized by span number of 705 bridges was tested, which was about 30% of the total of 2,388 bridges. The format of the tuples was the same as the wide format in Figure 5-2. Following the comparison method described in Table 5-2, the estimated value of level 2 and the actual value of level 3 were judged to be different, so this comparison was quite conservative. The result of calculating the difference between actual levels and estimated levels among 22 submodels was ranged from zero to eight, and the frequencies of each difference bins are shown in Table 5-13.

As a result of the comparison, it reached 67% in the case in which zero to two errors occurred, which was only about 5% when estimated perfectly. The model showed that the estimation performance of the model is secured to some extent by showing fluctuations of about zero to three errors (87.54%) and zero to four errors (95.57%) for all 22 submodels. Therefore, it can be concluded that the model to estimate damage was successfully developed with acceptable accuracy.

Table 5-13. Difference between the Estimated Value and the Actual Value in Estimating Deck Damage

Difference	Frequency	Percentage (%)	Cumulative Percentage (%)
0	142	4.73	4.73
1	795	26.49	31.22
2	1,074	35.79	67.01
3	616	20.53	87.54
4	241	8.03	95.57
5	90	3.00	98.57
6	29	0.97	99.53
7	12	0.40	99.93
8	2	0.07	**100.00**
Total	**3,001**	**100.00**	**-**

5.4. Summary

In this chapter, models were developed to estimate the severity of the damage using DNN and XGBoost. A two-stage model was designed to classify the occurrence of damage in the first stage and to categorize four severity extents in the next stage. As a result of comparing the performances of the derived models, the XGBoost model was selected because it provided about 95% of the performance measures compared with the 60 - 70% of the performance measures except recall provided by the DNN model, which indicated then DNN model provided conservative estimates of the inputs in this study. XGBoost performed better than DNN because the boosting algorithm provided the advantage of learning unbalanced data, whereas DNN required a sufficient amount of data in order to learn the complex structure that had numerous nodes and layers. After selecting the XGBoost model, its performance was evaluated. Some of submodels for almost all damage types, except cracking, in grades A and E were not generated due to a lack of data, and cracking, leakage, and efflorescence in grade B had a low F1 score because the boundary of the severity of the damage was vague.

Then, the XGBoost model was validated by generating a damage portfolio at the project level (i.e., an individual bridge) and at the network level (i.e., a region). It was possible to check the distribution of the severity by the type of

damage type and by the location in each portfolio. The portfolios of bridges that were generated for particular areas were sorted in descending order of the number of damages expected, which helped in determining the priorities when assigning time, cost, and inspectors. Using the generated portfolios, the developed model was validated by comparing the estimated results with the actual values at the location level of individual bridges. The results of the difference showed acceptable levels of around 90% accuracy of zero to three errors for all 22 submodels. Therefore, the developed model was confirmed its validity.

Chapter 6. Model Expandability Validation

This chapter describes the validation of the expandability of the model by applying the damage estimation process that was developed in this study to another component, i.e., the girder. First, the characteristics of girder data were explored, and the variables that influenced damage were identified by correlation analysis and XGBoost. Using the identified variables, DNN and XGBoost were used to develop models for use in estimating the damage to girders, and the extensibility was confirmed based on the results of the performance evaluation.

6.1. Girder Data Exploration

The girder was selected because of the status of the next most important component following the deck. The input data composed of 59 predictors and 2 targets (51 numerical variables and 8 categorical variables) was the same as that of the deck. The data of 12,867 tuples of 492 bridges in the long format was used to identify influencing variables and the transformed data of 3,607 rows in the wide format was utilized to develop damage estimation model. Six types of damage (i.e., cracking, map cracking, scaling, leakage and efflorescence, breakage of anchoring zone, and corrosion of exposed rebar)

were considered as estimation targets. Other damage types including exposed aggregate and insufficient thickness of covering were excluded since they are originated form construction error. As described in Figure 6-1, the predominant damage was girder cracking (40%) similar the deck of PSCI bridges.

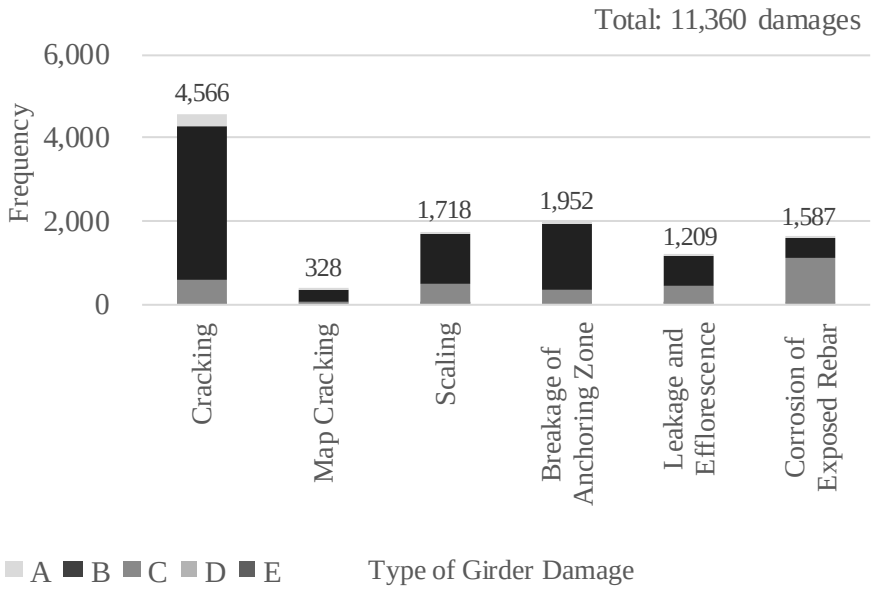


Figure 6-1. Condition Grade Distribution by Type of Girder Damage of Pre-stressed Concrete I Type Bridges

6.2. Girder Damage Influencing Variables Identification

6.2.1. Redundant Variables Removal

The first step of variable identification, redundant variables removal using correlation analysis, was performed and 15 variables were eliminated: vehicle height limit, total pavement area, total width, girder quantity, number of upward lanes, skew angle at the beginning, deck thickness, width of bottom flange, ADT, average annual temperature, summer average temperature, difference in average temperature between summer and winter, summer rainfall, annual average relative humidity, and difference in average humidity between summer and winter (Table 6-1). Similar to the deck case, the remained variables with greater variation and severe environmental conditions were mainly remained, such as effective width, ADTT, and summer and winter average temperature. For five categorical variables, Cremer's coefficient values ranged from 0.033 to 0.325, indicating weak relationships similar to the deck results, so all categorical variables left to be analyzed.

Table 6-1. Redundant Variables Removal for the Girder by Correlation Analysis

Domain	Variables	Rank	Correlation Coefficient
Identification Factors	*Vehicle height limit*	25	0.636
	Average daily truck traffic (ADTT)*	8	
	Total length*	49	0.712
	Total pavement area	55	
	Total width	12	0.970
	Effective width*	7	
	Effective width*	7	0.911
	Girder quantity	11	
	Effective width*	7	0.644
	Number of upward lanes	14	
	Skew angle at the beginning	22	0.969
	Skew angle at the end*	14	
Structural Factors	*Deck thickness*	31	0.624
	Pavement thickness*	6	
	Width of top flange*	44	0.822
	Width of bottom flange	60	
Environ-mental Factors	*Average daily traffic (ADT)*	9	0.834
	Average daily truck traffic (ADTT)*	8	
	Summer average temperature	34	0.618
	The number of heat waves*	28	
	Winter average temperature*	40	0.932
	Difference in average temperature between summer and winter	45	
	Annual average precipitation*	39	0.739
	Summer rainfall	54	
	Annual average relative humidity	24	0.838
	Summer average relative humidity*	18	
	Winter average relative humidity*	31	0.662
	Difference in average humidity between summer and winter	43	

Note: The italic font denotes the removed variables; the bold font with asterisk indicates the remained variables.

6.2.2. Influencing Variables Selection

The second step of the identification of the variables was the use of the decision tree methods, i.e., CART, RF, and XGBoost, to select the variables. The accuracy and F1 score of the binary classification model were shown in Table 6-2 and Figure 6-2.

Table 6-2. Performance Results of CART, RF, and XGBoost for the Grider

Decision Tree Methods	Dataset	Meas-ure	Damage Type						Aver-age
			Cracking	Map Cracking	Scaling	Breakage of Anchoring Zone	Leakage and Efflorescence	Corrosion of Exposed Rebar	
CART	Training	Acc.	0.91	0.90	0.89	0.92	0.90	0.89	0.90
		F1	0.51*	0.47*	0.77	0.89	0.65*	0.92	0.70
	Testing	Acc.	0.90	0.83	0.80	0.85	0.83	0.80	0.83
		F1	0.52*	0.29*	0.66*	0.81	-	0.87	0.63*
RF	Training	Acc.	0.94	0.96	0.90	0.90	0.87	0.91	0.91
		F1	0.74	0.84	0.84	0.88	0.49*	0.93	0.79
	Testing	Acc.	0.95	0.91	0.90	0.87	0.86	0.94	0.91
		F1	0.78	0.55*	0.85	0.84	0.32*	0.95	0.72
XGBoost	Training	Acc.	0.96	0.99	0.97	0.96	0.95	0.98	0.97
		F1	0.84	0.95	0.94	0.95	0.86	0.98	0.92
	Testing	Acc.	0.94	0.92	0.91	0.93	0.92	0.92	0.92
		F1	0.75	0.67*	0.84	0.91	0.76	0.94	0.81

Note: The asterisk denotes a value less than 0.7.

In the case of CART, the average F1 score in testing process was found to be lower than the threshold of 0.7. By type of damage, CART did not meet the criteria for cracking, map cracking and scaling. For leakage and efflorescence, TP and FP were all estimated to 0 and precision was not defined, so F1 score was not calculated consequently. As for RF, the F1 scores of map cracking and leakage and efflorescence were lower than 0.7, whereas XGBoost showed higher performance except map cracking (0.67). Even in the case of map cracking, this situation can be improved when a little more data are collected. Therefore, XGBoost was again selected for variable selection, and this method will be used for future expansion of these processes without comparing three different types of decision tree methods.

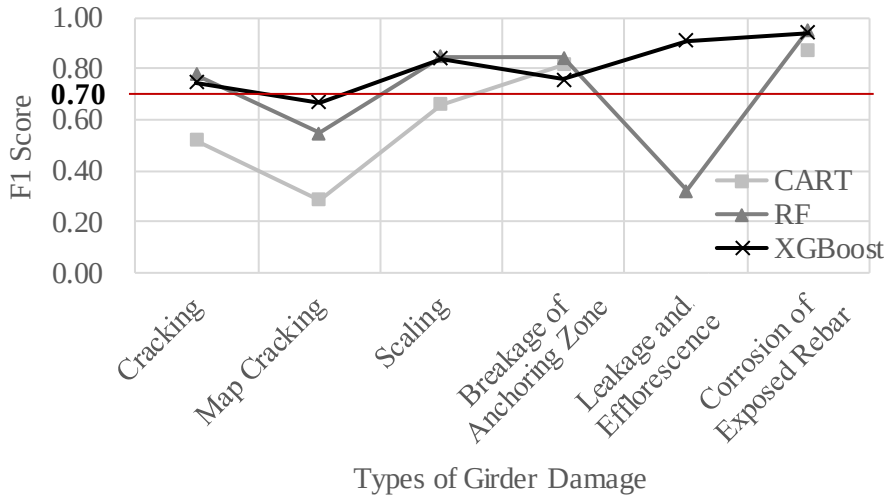


Figure 6-2. F1 Scores of CART, RF, and XGBoost in Testing of the Girder Data

Table 6-3 indicates that the number of influencing variables that were selected for the girder varied from 9 to 16; numerous identification factors were derived for cracking and for the correlation of the exposed rebar, and environmental factors were more derived than structural factors.

Table 6-3. Number of Selected Influencing Variables by the Type of Girder Damage and the Domain

Domain	**Damage Type**					
	Cracking	**Map Cracking**	**Scaling**	**Breakage of Anchoring Zone**	**Leakage and Efflorescence**	**Corrosion of Exposed Rebar**
Identification factor	7	3	4	4	4	8
Structural factor	3	4	1	0	3	2
Environmental factor	3	5	5	4	6	5
Inspection factor	1	1	1	1	1	1
Number of influencing variables	14	13	11	9	14	16

Table 6-4 provides the list of variables selected for the damage to girders. The most frequent variables were ADTT and age, and the variables derived from 4 to 5 different kinds of damage were the weight limit of the bridge and its height, total length, the spacing of its girders, annual average precipitation,

snowfall, and the average relative humidity during the summer. Likewise, the deck, that time (i.e., age), and dead load (i.e., total length), live load (i.e., vehicle weight limit and ADTT) were important factors with respect to damage to the girders. For the deck, the strengths of the girders and annual average relative humidity were found frequently, whereas for the girder, the spacing of the girders and average relative humidity during the summer were extracted frequently. Unlike the types of damages to the deck, the annual average precipitation and snowfall, which keep the bridge wet for a significant amount of time, had impacts on almost all types of girder damage except cracking.

By the type of damage, the results of the selected influencing variables were analyzed by three groups of types of damage with similar properties, i.e., cracking and map cracking, scaling and breakage of the anchoring zone, and leakage, efflorescence, and corrosion of the exposed rebar.

Table 6-4. List of Influencing Variables by the Type of Girder Damage

Domain	Variables	Cracking	Map Cracking	Scaling	Breakage of Anchoring Zone	Leakage and Efflorescence	Corrosion of Exposed rebar	Count
Identification factor	Owner						8	1
	Competent authority		12					1
	Minimum radius of curve						6	1
	Design live load			8				1
	Vehicle weight limit	11		3	2	12	4	5
	Height	12	4		6		3	4
	Depth of water		9				5	2
	Total length	5		4	7	9	15	5
	Effective width	7		9		2		3
	Skew angle at the end				4			1
	Number of spans	4					9	2
	Span length between the centers of supports	14				14	13	3
	Substructure type	10						1
Structural factor	Pavement thickness	2	8					2
	Girder strength		6					1
	Girder height					4		1
	Girder spacing	3	10			13	14	4
	Width of top flange			10				1
	Strength of main reinforcing rebar					10		1
	Spacing of main reinforcing rebar	8	7				7	3
Environmental factor	Average daily truck traffic (ADTT)	6	1	5	8	8	16	6
	Winter average temperature		2	2		6		3
	The number of tropical nights				9			1
	Annual average precipitation		3	6		5	11	4
	Snowfall		5	11	1	11	10	5
	Summer average relative humidity	9	13		5	3	12	5
	Average wind speed	13				7		2
	Amount of Chloride			7			2	2
Inspection factor	Age	1	11	1	3	1	1	6
Number of influencing variables		14	13	11	9	14	16	29

In the case of cracking, the influencing factors, including effective width, number of spans, and the lengths of the spans between the centers of the supports implied the size and the weight (i.e., the dead load) of the deck in a span, which is divided and supported by girders. In the same context, the type of substructure that supports the dead load of the girder and the dead and live loads from the deck was identified as an influencing factor. This explanation was similar to the case in which the strength of the girder, which supports the load of the deck, was an important factor in the cracking of the deck. For map cracking, it can be explained that it is difficult for bi-directional cracks to occur given the significant strength of the girder. As previous studies have indicated, freezing and thawing, moisture, and the calcium chloride in the deicing chemicals are the causes of map cracking (Federal Highway Administration (FHWA), 2002; Deschenes Jr. et al., 2018), and map cracking also can be the result of the low average temperature in the winter and the moisture-related variables, i.e., annual average precipitation, snowfall, and average humidity during the summer.

The commonly-derived information for the scaling and breakage of the anchoring zone include the effects of live loads and dead loads, which are represented by weight limit of the vehicles and the total length of the bridge, respectively, and the effects of moisture and ions as the result of snowfall and the deicing chemicals. Scaling, which is the damage that develops due to cracks, showed that design live load and effective width, which indicated that the live

and dead loads of the deck were important factors, with considering how the width of the top flange of the girder can divide this load well. The factors associated with scaling that were derived in this study indicated that scaling can be promoted by the physicochemical effects of moisture (i.e., annual average precipitation), shrinkage in the winter (i.e., the low average temperature in the winter), snowfall and the subsequent use of deicing chemicals, and chloride ions (i.e., the amount of chloride). Also, breakage of the anchoring zone can occur and be accelerated when internal corrosion occurs due to the effects of the high temperature in the summer (i.e., the number of tropical nights) and humidity (i.e., the average relative humidity in the summer), snowfall, and deicing chemicals.

As mentioned earlier, leakage is damage that occurs after cracking, and, therefore, it is important to prevent cracking first (Song et al., 2009; Federal Highway Administration (FHWA), 2012). The results of this study indicated that the variables related to the dead load of the deck (i.e., effective width and the length of the span between the centers of the supports), which also were found to be influencing factors for cracking, can influence leakage and efflorescence. Whereas this study explained that the leakage and efflorescence of the deck can be increased by the extreme conditions in the summer (e.g., the number of heat waves and tropical nights) and the differences in the temperature and humidity during the summer and the winter, the average temperature in the winter was found to be an influencing factor for the leakage

and efflorescence of girders. In addition, the average annual precipitation, snowfall, deicing chemicals, and average wind speed explained such physicochemical effects. Leakage and efflorescence can be promoted by moisture due to humidity, rainfall, and snowfall and the movement of moisture and chloride ions by the wind. The factors involved in the corrosion of exposed rebar included a specific owner, the minimum radius of the curve, the length of the span between the centers of the supports, and the spacing of the main reinforcing rebar. Similar to the damage to the deck, the characteristics of the main reinforcing rebar, including its bending, length, and the load per rebar, explained the occurrence and the acceleration of the corrosion. In addition, the variables related to the movement of moisture and chloride ions were derived, i.e., the depth of water, the average annual precipitation, snowfall, the average relative humidity in the summer, and the amount of chloride ions (Main Roads Western Australia, 2008; Federal Highway Administration (FHWA), 2000; Federal Highway Administration (FHWA), 2012).

6.3. Girder Damage Estimation Model Development

6.3.1. Model Evaluation

(1) Model Performance Comparison

Using the identified variables, DNN and XGBoost were used to develop a model for the prediction of damage to girders. As before, XGBoost had a better performance than DNN, so the XGBoost model will be used for future expansion of these assessments (Table 6-5). The overall performance of both models to estimate damage to the girder was about 6 - 9% lower than that of the model to estimate the deck damage. In particular, the DNN model had a larger decrease in overall performance than the XGBoost model. This was because the small number of the girder dataset with more serous distribution of class imbalance made it more difficult for DNN to learn its complex network structure.

Table 6-5. Performance Comparison (%) of DNN and XGBoost Models in Estimating Girder Damage

Measure \ Method	DNN	XGBoost
Accuracy	56.85 (-6.05)	89.31 (-5.82)
Precision	81.92 (-9.27)	86.78 (-7.38)
Recall	56.85 (-6.04)	89.31 (-5.86)
F1 Score	66.33 (-7.23)	87.69 (-6.79)

Note: Precision, Recall, and F1 score were weighted values; the values in parenthesis indicate the differences of performance between the girder model and the deck model.

The average accuracy and the average recall of DNN were around 32% lower than these values for XGBoost, and this situation also was shown in the submodels (Table 6-6, Table 6-7, Table 6-8, and Table 6-9).

Table 6-6. Accuracy (%) of the DNN Model for the Grider in Testing

Damage Types	A	B	C	D	E	Average
Cracking	56.0	49.1	60.8	N	N	55.3
Map Cracking	N	72.3	N	N	N	72.3
Scaling	N	55.0	61.1	N	N	58.1
Breakage of Anchoring Zone	N	66.9	N	N	N	66.9
Leakage and Efflorescence	N	47.9	57.1	N	N	52.5
Corrosion of Exposed Rebar	N	51.9	47.1	N	N	49.5
Average	56.0	57.2	56.5	-	-	56.8

Note: N means not generated. [0,85) [85,100] (%)

Table 6-7. Accuracy (%) of the XGBoost Model for the Grider in Testing

Damage Types	A	B	C	D	E	Average
Cracking	97.0	70.1	92.9	N	N	86.6
Map Cracking	N	95.8	N	N	N	95.8
Scaling	N	81.2	93.5	N	N	87.3
Breakage of Anchoring Zone	N	96.0	N	N	N	96.0
Leakage and Efflorescence	N	82.0	96.2	N	N	89.1
Corrosion of Exposed Rebar	N	93.8	83.9	N	N	88.9
Average	97.0	86.5	91.6	-	-	89.3

Note: N means not generated. [0,85) [85,100] (%)

Table 6-8. Recall Rate (%) of the DNN Model for the Grider in Testing

Damage Types	A	B	C	D	E	Average
Cracking	56.0	49.1	60.8	N	N	55.3
Map Cracking	N	72.3	N	N	N	72.3
Scaling	N	55.0	61.1	N	N	58.1
Breakage of Anchoring Zone	N	66.9	N	N	N	66.9
Leakage and Efflorescence	N	47.9	57.1	N	N	52.5
Corrosion of Exposed Rebar	N	51.9	47.1	N	N	49.5
Average	56.0	57.2	56.5	-	-	56.8

Note: N means not generated. [0,85) [85,100] (%)

Table 6-9. Recall Rate (%) of the XGBoost Model for the Grider in Testing

Damage Types	A	B	C	D	E	Average
Cracking	97.0	70.1	92.9	N	N	86.6
Map Cracking	N	95.8	N	N	N	95.8
Scaling	N	81.2	93.5	N	N	87.3
Breakage of Anchoring Zone	N	96.0	N	N	N	96.0
Leakage and Efflorescence	N	82.0	96.2	N	N	89.1
Corrosion of Exposed Rebar	N	93.8	83.9	N	N	88.9
Average	97.0	86.5	91.6	-	-	89.3

Note: N means not generated.

[0,85) [85,100] (%)

(2) Model Performance Evaluation

For submodels with accuracy and recall greater than 85%, the damage occurrence data (i.e., the damage severity from level 1 to level 4) were less than 7.4% as presented in Table 6-10. Even with the class-imbalanced data in this study, XGBoost showed superior performance compared to DNN because of the boosting algorithm for training the model, which was proved again. Therefore, XGBoost will be used for future expansion of the developed model without comparing DNN and XGBoost.

Table 6-10. Percentage (%) of Damage Data for the Girder

Damage Types	A	B	C	D	E
Cracking	2.8	34.8	7.4	0.5	0.0
Map Cracking	0.0	5.0	0.9	0.0	0.0
Scaling	0.3	19.1	7.3	0.1	0.2
Breakage of Anchoring Zone	0.3	4.0	0.7	0.2	0.0
Leakage and Efflorescence	0.4	18.1	3.9	0.3	0.0
Corrosion of Exposed Rebar	0.1	5.8	15.4	0.7	0.0

Note: The total number of data in the wide format was 3,607

[0,15) [15,100] (%)

The final model of XGBoost consisted of 11 submodels, and their F1 scores are presented in Table 6-11Table 6-7. For most of the grade A, D, and E cases, submodels were not created, and seven of the submodels showed high performance greater than 85%.

Table 6-11. F1 Score (%) of the XGBoost Model for the Girder in Testing

Damage Types	A	B	C	D	E	Average
Cracking	96.1	68.3	91.5	N	N	85.3
Map Cracking	N	94.6	N	N	N	94.6
Scaling	N	79.0	92.1	N	N	85.5
Breakage of Anchoring Zone	N	95.0	N	N	N	95.0
Leakage and Efflorescence	N	78.6	95.5	N	N	87.0
Corrosion of Exposed Rebar	N	91.8	82.2	N	N	87.0
Average	96.1	84.5	90.3	-	-	87.7

Note: N means Not generated. [0,85) [85,100] (%)

The reason why submodels were not generated in the case of the grade A, D, and E was the shortage of data, as previously evidenced in Section 5.2.2. Model Performance Evaluation. When one class of the damage severity had less than six tuples, a model was not generated. For example, data of map cracking in grade C did not have level 4 of severity data as described in Table 6-12, and then the submodel was out of existence. Form a practical point of view, the condition grades of A, D, and E for damage to the girder were hard to

be found. The girder is a key component that carries the dead load by the deck and live load by traffic, and is fixed before the damage reaches grade D or E. In addition, almost damage types except cracking and leakage and efflorescence were hard to occur in a small size, so then the grade A data is very small.

Table 6-12. Distribution of the Girder Damage Data for Cases without Submodels

Occurrence	**Grade**	**N**	**Y**				**Grade**	**N**	**Y**			
Severity Extents		**L0**	**L1**	**L2**	**L3**	**L4**		**L0**	**L1**	**L2**	**L3**	**L4**
Map cracking	**A**	3606	1				**C**	3573	34			
			1	0	0	0			29	3	2	0
Scaling	**A**	3597	10				**-**	-	-			
			7	3	0	0						
Breakage of Anchoring Zone	**A**	3597	10				**C**	3581	26			
			2	2	2	4			19	3	4	0
Leakage and Efflorescence	**A**	3591	16				**-**	-	-			
			12	3	1	0						
Corrosion of Exposed Rebar	**A**	3605	2				**-**	-	-			
			2	0	0	0						

Note: The total number of data in the wide format was 3,607.

Table 6-12. Distribution of the Girder Damage Data for Cases without Submodels (Continue)

Occurrence / Severity Extents	Grade	N	Y				Grade	N	Y			
		L0	L1	L2	L3	L4		L0	L1	L2	L3	L4
Cracking	**D**	3606	1				**E**	3606	1			
			1	0	0	0			1	0	0	0
Map cracking	**D**	3607	0				**E**	3607	0			
Scaling	**D**	3598	9				**E**	3598	9			
			1	0	8	0			1	0	8	0
Breakage of Anchoring Zone	**D**	3598	9				**E**	3607	0			
			3	6	0	0						
Leakage and Efflorescence	**D**	3607	0				**E**	3607	0			
Corrosion of Exposed Rebar	**D**	3583	24				**E**	3607	0			
			20	4	0	0						

Note: The total number of data in the wide format was 3,607.

As mentioned before, the silhouette coefficient was used to present data distribution. As shown in Table 6-13, the absolute silhouette values of cracking and scaling in grade B, and corrosion of exposed rebar in grade C were relatively lower than the values of the other cases, which was matched to the cases with the relatively low F1 scores in Table 6-11 except leakage and efflorescence in grade B. Small absolute values of silhouette coefficient meant that the data points were distributed at boundaries of the severity, so then the XGBoost was difficult to develop submodels by learning distinctive features of

each severity class. In the case of leakage and efflorescence, the absolute silhouette value was relatively large, so the data points were separated, which makes the learning step of XGBoost easier. However, this damage type includes two types of damages, i.e., leakage and efflorescence, which are related to physicochemical effects of water and icons, and this situation made it difficult to learn distinctive features of this mixed damage type with the similar occurrence mechanisms of leakage and efflorescence.

Table 6-13. Silhouette Coefficient of Severity Classes of the Girder Data

Damage Types	A	B	C	D	E
Cracking	-0.10	-0.03	-0.10	-0.11	0.07
Map Cracking	-0.27	-0.27	-0.28	-0.25	-
Scaling	-0.31	-0.04	-0.14	-0.33	-0.38
Breakage of Anchoring Zone	-0.27	-0.11	-0.28	-0.40	-
Leakage and Efflorescence	-0.33	-0.10	-0.05	-0.30	-
Corrosion of Exposed Rebar	-0.28	-0.12	-0.04	-0.10	-

$|s| < 0.05$ $|s| \geq 0.05$

6.3.2. Model Validation

Using the expanded model to estimate damage to the girder, a damage portfolio at the project level was successfully generated as presented in Table 6-14. To validate that the expanded model, a total of 931 tuples organized by span number of 145 bridges was tested, which was about 30% of the total of 492 bridges. The result of calculating the difference between actual levels and estimated levels among 11 submodels was ranged from 0 to 5, and the frequencies of each difference bins are shown in Table 6-15. As a result of the comparison, it reached 95% in the case in which zero to two errors occurred. To summarize, a model that can be used successfully to estimate the damage to girders was developed by applying the same methodology that was used for the deck, confirming the extensibility of the model to other components and other main types of structures.

Table 6-14. Example of Girder Damage Portfolio for a Bridge (Project Level)

- Expected Inspection: 2019.7.2.
- Bridge Number: 000495
- Region: Gangwon-Do
- Main Structure Type: Pre-Stressed Concrete I
- Age: 28 years

Location		Left						Middle Left						Center						Middle Right						Right					
Span Number		1						2						3						4						5					
Damage Type		CR	MC	SC	BA	LE	CE	CR	MC	SC	BA	LE	CE	CR	MC	SC	BA	LE	CE	CR	MC	SC	BA	LE	CE	CR	MC	SC	BA	LE	CE
Condition Grade	**A**																														
	B					L2				L3		L2				L1		L1						L2				L2		L2	
	C					L1						L1																L1		L2	
	D																														
	E																														

Note: Cracking (CR), Map Cracking (MC), Scaling (SC), Breakage of Anchoring Zone (BA), Leakage and Efflorescence (LE), Corrosion of Exposed Rebar (CE).
The cells with diagonal lines indicate the submodels were not generated. The gray cells indicate the L(evel) 0 (no damage).
The damage severity (number of defects) is represented by: L1 (1), L2 (2), L3 (3-5), L4 (≥6).

Table 6-15. Difference between the Estimated Value and the Actual Value in Estimating Girder Damage

Difference	Frequency	Percentage (%)	Cumulative Percentage (%)
0	348	37.34	37.34
1	383	41.14	78.49
2	157	16.91	95.40
3	33	3.49	98.89
4	9	0.91	99.81
5	2	0.19	**100.00**
Total	**931**	**100.00**	**-**

6.4. Summary

In this chapter, the expansion of the model to another component, i.e., girders, was performed, and its validity was confirmed. Initially, girder data were generated using the same dataset that was used for the deck. Using the dataset that was generated, two steps were performed to identify the causes of damage. After a correlation analysis was conducted, 15 redundant variables were removed, and the influencing variables were selected. As for most of the target damage of the girder, similar influencing variables related to the damage to the deck were derived, and they included time, dead load, live load, and relative humidity. Girder spacing also was found to be a key variable because it reflected the structural characteristics of PSCI bridges compared to the strength of the girders for the deck damage. In addition, the annual averages of precipitation and snowfall, which induce the physicochemical actions of water and chloride ions, were found to be important variables for almost all types of girder damage, including map cracking, scaling, breakage of the anchoring zone, leakage and efflorescence, and corrosion of the exposed rebar. For the individual types of damage, variables related to supporting the live and dead loads of the deck and transferring those loads of the girder were derived for cracking. For scaling, leakage, and efflorescence, the influencing variables related to the load applied to each girder were derived. Similar to the identified

influencing variables for the corrosion of exposed rebar to the deck, regional variable, the characteristics of the main reinforcing rebar, and variables related to physicochemical actions of water and chloride ion were also derived for the corrosion of exposed rebar to the girder. In summary, diverse variables were found that can provide structural and material engineering information for each type of girder damage.

As a result of developing the model to estimate damage to the girder, the XGBoost performance was superior to DNN as well as the deck, and, in the future, the XGBoost model can be used during the expansion of the model for use with another component or main type of structure. Almost all of submodels that were generated showed good performance except when there was a shortage of data. Then, the developed XGBoost model was validated by using the generated portfolios. The differences between the estimated results and the actual values presented acceptable levels of around 95% accuracy of zero to two errors for all 11 submodels. Thus, the model was confirmed its expandability.

Chapter 7. Conclusions

This chapter summarizes the research results and the contributions of this study. Also, future work is identified that is needed to support bridge inspection practices.

7.1. Research Results

Given the limited resources, including time, budget, and manpower, the quality of the inspection has risks that should be reduced. Therefore, there is a need to generate information in advance of an inspection, thus the goal of this dissertation was to develop a process using artificial intelligence to estimates the types, locations, grades, and severities of the damages that can occur on the components of bridges for use on the day of inspection. Four specific objectives were established to attain the research goal, i.e., 1) data collection and preprocessing, 2) identification of the influencing variables, 3) development of a model for estimating damage, and 4) validation of the expandability of the model.

First, bridge and inspection data from the Korean Bridge Management System and weather data from the Korea Meteorological Administration were collected and preprocessed. After preparing the dataset, a correlation analysis

was performed to remove redundant variables to resolve multicollinearity. Among the three decision tree methods, i.e., CART, RF, and XGBoost, XGBoost was selected to extract the variables that influenced the occurrence of damages to the components of bridges. In general, vehicle weight limit, total length, height, ADTT, annual average relative humidity, and the strength of the girder were derived frequently for overall types of damage. For each type of damage, diverse variables were derived that matched the corresponding mechanisms by which damage occurred that affected the structural and material properties. Then, damage estimation models were developed using the Deep Neural Networks and XGBoost, and the latter model was finally selected. Using the developed model, portfolios, i.e., lists of possible damages, were generated by bridge and by region. Finally, the same methodology was applied to the girders of PSCI-type bridges and, by obtaining a high accuracy and F1 score, thereby validating the expandability of the developed model.

7.2. Contributions

The results of this study are expected to contribute to academia, industry, and the social economy. From an academic perspective, this research quantitatively verified the scattered knowledge about the factors that influence the damage to bridges using actual inspection data. In addition, this was a pioneering attempt to produce information in advance to support inspections by analyzing prior inspection data using artificial intelligence techniques.

For the bridge management field, the developed methodology can be used to address existing bridges that have been managed as well as bridges that have no past inspection history. By providing estimated information before inspections, the inspection time and the risk of the quality degradation of the inspections can be reduced. In addition, it will be possible to distribute inspection time, budgets, and human resources efficiently by providing inspection priorities among the number of bridges to be inspected. The model also extends the functions of the existing KOBMS by adding inspection support modules, which enables the KOBMS to cover all four procedures of bridge management, i.e., inspection, condition diagnosis, maintenance decision-making, and maintenance actions.

Finally, this research has potential to reduce of the lifecycle costs of the bridge and the extension of the life of the bridge through the expansion of this

research with various components and structural types, and thereby it is expected to contribute to make a safer society by preventing accidents.

7.3. Future Research

Although it is noted that this research was performed based on the specific research scopes, i.e., the decks and girders of PSCI type bridges, additional research and examinations have to be performed to expand the scopes of such estimations to different types of main structures and different types of components. The methodology can be applied to different components, such as expansion joints, and different types of main structure types, such as reinforced slab (RCS) and steel box girder (STB), which are the other main types of structures in South Korea.

As the performance of the estimation model was evaluated and described, it was apparent that a sufficient amount of data is vitally important in generating a submodel and improving the performance of the model. The damage level data were obtained in the table format and they were typed directly by office workers or by the text data submitted in PDF file format by inspectors after detailed inspection and precise diagnosis. Thus, if the recently developed mobile application is used extensively, large amounts of inspection data will be accumulated more quickly and accurately, and, consequently, the performance of the model will be improved further.

For practical use, the generated model will be processed so that it can be applied to the current mobile inspection application or the web-based dashboard

with a function of automatic and periodic updating as the quantity of inspection data increases. As described in Figure 7-1, the visualization of the portfolios of a 3D model of a bridge would be help inspectors understand the information more easily. In addition, when producing network level portfolios, it would be more practical to provide inspection priorities that takes into account the distance between the bridges in the inspection area.

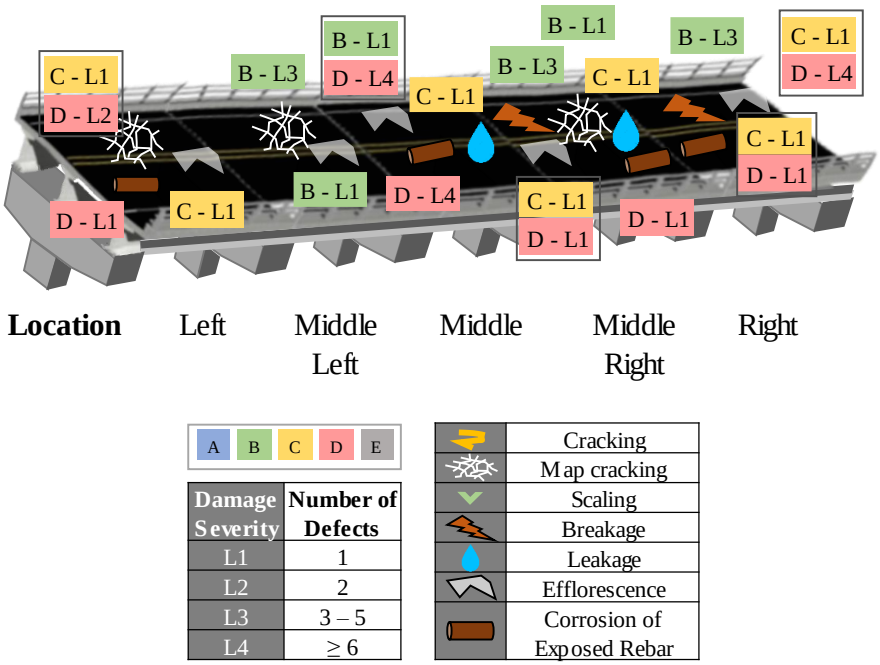


Figure 7-1. Visualization of the Estimation Results of a 3D model

Bibliography

Agrawal, A. K., Kawaguchi, A., and Chen, Z. (2010). Deterioration rates of typical bridge elements in new york. *Journal of Bridge Engineering*, *15*(4), 419–429. doi: http://doi.org/10.1061/(ASCE)BE.1943-5592.0000123

Ahn, Y. K. (2013). Forgotten facilities safety and things to look after. *Facility safety*, 42.

Ministry of Land, Infrastructure, and Transport of Korea (MOLIT). (2015). Yearbook of road bridges and tunnels, Sejong, Korea: MOLIT.

Alberta Infrastructure and Transportation. (2008). BIM inspection manual (Ver. 3.1). Retrieved January 4, 2019, from https://open.alberta.ca/publications/bim-inspection-manual-version-3-1

Alberta Transportation. (2007). Bridge inspection and maintenance system: level 2 inspection manual (Ver. 1.1). Retrieved January 4, 2019, from https://open.alberta.ca/publications/bridge-inspection-and-maintenance-system-level-2-inspection-manual

Aleksander, I. and Morton, H. (1990). *An introduction to neural computing*. Chapman and Hall: London.

American Association of State Highway and Transportation Officials (AASHTO). (2001). Commonly Recognized (CoRe) structural elements. Retrieved June 23, 2018, from https://trid.trb.org/view/473928

American Association of State Highway and Transportation Officials (AASHTO). (2017). The manual for bridge evaluation (3rd Ed.). Retrieved July 3, 2018, from https://store.transportation.org/Common/DownloadContentFiles?id=1712

Bektaş, B. A. (2017). Use of recursive partitioning to predict national bridge inventory condition ratings from national bridge elements condition data. *Transportation Research Record: Journal of the Transportation Research Board,* 2612(1), 29–38. doi: http://doi.org/10.3141/2612-04

Bektas, B. A., Carriquiry, A., and Smadi, O. (2013). Using classification trees for predicting national bridge inventory condition ratings. *Journal of Infrastructure Systems*, 19(4), 425–433. doi: http://doi.org/10.1061/(ASCE)IS.1943-555X.0000143

Ben-Akiva, M. and Gopinath, D. (1995). Modeling infrastructure performance and user costs. *Journal of Infrastructure Systems*, 1(1), 33–43.

Bishop, C. M. (2006). *Pattern recognition and machine learning* . New York: Springer.

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J. (1984). *Classification and regression trees*. Monterey, CA: Wadsworth & Brooks.

Bundesanstalt für Straßenwesen. (2013). ASB structure inventory, ASB-ING. Retrieved January 5, 2019, from https://www.bast.de/BASt_2017/DE/Ingenieurbau/Publikationen/Regelwerke/Erhaltung/ASB-ING.html

Bundesanstalt für Straßenwesen. (2017). Recording and assessment of damages, RI-EBW-PRÜ F. Retrieved January 5, 2019, from https://www.bast.de/BASt_2017/DE/Ingenieurbau/Fachthemen/b4-Bauwerkspruefung-RI-EBW-PRUEF/b4-Bauwerkspruefung-RI-EBW-PRUEF.html

Calle, M. L. and Urrea, V. (2010). Letter to the editor: stability of random forest importance measures. *Briefings in Bioinformatics*, 12(1), 86–89. doi: http://doi.org/10.1093/bib/bbq011

Campbell, L. E., Perry, C. N., Connor, R. J., and Lloyd, J. B. (2016). Element level bridge inspection: benefits and use of data for bridge management, FHWA/IN/JTRP-2016/13. Joint Transportation Research Program. doi: http://doi.org/10.5703/1288284316336

Cattan, J. and Mohammadi, J. (1997). Analysis of bridge condition rating data using neural networks. *Computer-Aided Civil and Infrastructure Engineering*, 12(6), 419–429.

Colignatus, T. (2007). A measure of association (correlation) in nominal data (contingency tables), using determinants. *Munich Personal RePEc Archive (MPRA)*, 2662.

Cramér, H. (1946). *Mathematical methods of statistics*. Princeton, NJ: Princeton University Press.

Creary, P. A. and Fang, F. C. (2013). The data mining approach for analyzing infrastructure operating conditions. *Procedia - Social and Behavioral Sciences*, 96, 2835–2845. doi: http://doi.org/10.1016/j.sbspro.2013.08.316

Dadson, D. K., Garza, J. M. de la, and Weyers, R. E. (2002). Service life and impact of virginia environmental exposure condition on paint on steel girder bridges. *Journal of Infrastructure Systems*, 8(4), 149–159. doi: http://doi.org/10.1061/(ASCE)1076-0342(2002)8:4(149)

De'ath, G. and Fabricius, K. E. (2000). Classification and regression trees: A powerful yet simple technique for ecological data analysis. *Ecology*, 81(11), 3178–3192. doi: http://doi.org/10.1890/0012-9658(2000)081[3178:CARTAP]2.0.CO;2

Deschenes Jr., R. A., Giannini, E. R., Drimalas, T., Fournier, B., and Hale, W. M. (2018). Effects of moisture, temperature, and freezing and thawing on alkali-silica reaction. *ACI Materials Journal*, 115(4), 575–584. American Concrete Institute.

Elhag, T. and Wang, Y. (2007). Risk assessment for bridge maintenance projects: neural networks versus regression techniques. *Journal of Computing in Civil Engineering*, 21(6), 402–409. doi: http://doi.org/10.1061/(ASCE)0887-3801(2007)21:6(402)

Federal Highway Administration (FHWA). (1995). Recording and coding guide for the structural inventory and appraisal of the nation's bridges, FHWA-PD-96-001. Retrieved April 2, 2018, from https://www.fhwa.dot.gov/bridge/nbi.cfm

Federal Highway Administration (FHWA). (2000). Materials and methods for corrosion control of reinforced and prestressed concrete structures in new construction, FHWA-RD-00-081.

Federal Highway Administration (FHWA). (2002). Guidelines for detection, analysis, and treatment of materials-related distress in concrete pavements, FHWA-RD-01-163. Retrieved April 3, 2019, from https://www.fhwa.dot.gov/publications/research/infrastructure/pavements/pccp/01163/

Federal Highway Administration (FHWA). (2012). Bridge inspector's reference manual, FHWA NHI 12-049.

Freyermuth, C. L., Klieger, P., Stark, D. C., and Wenke, H. N. (1970). Durability of concrete bridge decks-a review of cooperative studies. *Highway Research Record*, (328), 50–60.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 29(5), 1189–1232.

Genuer, R., Poggi, J.-M., and Tuleau-Malot, C. (2010). Variable selection using random forests. *Pattern Recognition Letters*, 31(14), 2225–2236.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep learning*. Cambridge, MA: The MIT Press.

Hammerla, N. Y., Halloran, S., and Plötz, T. (2016). Deep, convolutional, and recurrent models for human activity recognition using wearables. *arXiv preprint arXiv:1604.08880*.

Han, J., Pei, J., and Kamber, M. (2012). *Data mining: concepts and techniques* (3rd Ed.). Waltham, MA: Elsevier.

Haykin, S. (1999). *Neural networks: A comprehensive foundation* (2nd Ed.). Upper Saddle River, NJ: Prentice-Hall.

Hearn, G. (2007). Bridge inspection practices (Vol. 375). *Transportation Research Board*. Retrieved from https://www.nap.edu/download/14127

Huang, J., Huang, N., Zhang, L., and Xu, H. (2012). A method for feature selection based on the correlation analysis. *Proceedings of 2012 International Conference on Measurement, Information and Control (MIC)*, 529–532. Harbin, China: IEEE.

Huang, Q., Ong, K.-L., and Alahakoon, D. (2015). Improving bridge deterioration modelling using rainfall data from the bureau of meteorology. *Proceedings of the 13-th Australasian Data Mining Conference (AusDM 2015)*, 161–167. Sydney, Australia.

Huang, R.-Y., Mao, I. S., and Lee, H.-K. (2010). Exploring the deterioration factors of rc bridge decks: a rough set approach. *Computer-Aided Civil and Infrastructure Engineering*, 25(7), 517–529. doi: http://doi.org/10.1111/j.1467-8667.2010.00665.x

Huang, R. Y. and Chen, P. F. (2012). Analysis of influential factors and association rules for bridge deck deterioration with utilization of national bridge inventory. *Journal of Marine Science and Technology*, 20(3), 336–344.

Huang, Y. (2010). Artificial neural network model of bridge deterioration. *Journal of Performance of Constructed Facilities*, 24(6), 597–602. doi: http://doi.org/10.1061/(ASCE)CF.1943-5509.0000124

IBM. (2019). Medtronic: Predictive model for personalized care. Retrieved February 10, 2019, from https://www.ibm.com/case-studies/medtronic

Incheon Metropolitan City. (2018). Current status of budgeting and enforcement for bridges and tunnels in the second half of 2017. Retrieved February 10, 2019, from http://www.incheon.go.kr/program/fileDownload.do?fileNo=586550

Jeju Special Self-Governing Province. (2018). A result of periodical inspection of the bridge in the first half of 2018. Retrieved September 10, 2019, from https://www.jeju.go.kr/open/open/iopenboard.htm?category=1014&qType=title&q=

Jović, A., Brkić, K., and Bogunović, N. (2015). A review of feature selection methods with applications. *Proceedings of 2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia,* 1200–1205.

Jung, H.-J., Zi, G., Kong, J.-S., and Kang, J.-G. (2008). Durability prediction for concrete structures exposed to chloride attack using a bayesian approach. *Journal of the Korea Concrete Institute*, 20(1), 77–88.

Kohavi, R., and John, G. H. (1997). Wrappers for feature subset selection. *Artificial intelligence*, 97(1–2), 273–324.

Korea Construction Engineering Development Collaboratory Management Institute (KOCED CMI). (2016). The final report of development of basement for resesarch on bridge life extension by evaluating the performance of the deteriorated(demolished) bridge.

Korea Crime Scene Investigation (KCSI). (2014). Geographical profiling system (GeoPros). Retrieved September 10, 2019, from http://www.korea.kr/common/download.do?fileId=183606277&tblKey=GMN

Korea Infrastructure Safety Corporation (KISTEC). (2018). Annual report of facility management system.

Korea Institute of Civil Engineering and Building Technology (KICT), and Korea Infrastructure Safety Corporation (KISTEC). (2016). The final report of operation of Bridge Management System in 2015.

Kotze, R., Ngo, H., and Seskis, J. (2015). Improved bridge deterioration models, predictive tools and costs. Retrieved from https://www.onlinepublications.austroads.com.au/items/AP-R487-15

Kušar, M. (2017). Bridge inspection quality improvement using standard inspection methods. *Proceedings of JOINT COST TU1402 - COST TU1406 - IABSE WC1 workshop*. Zagreb, Croatia.

Lee, J., Guan, H., Loo, Y., and Blumenstein, M. (2014). Development of a long-term bridge element performance model using elman neural networks. *Journal of Infrastructure Systems*, 20(3), 4014013. doi: http://doi.org/10.1061/(ASCE)IS.1943-555X.0000197

Lee, J., Sanmugarasa, K., Blumenstein, M., and Loo, Y.-C. (2008). Improving the reliability of a Bridge Management System (BMS) using an ANN-based Backward Prediction Model (BPM). *Automation in Construction*, 17(6), 758–772. doi: http://doi.org/10.1016/j.autcon.2008.02.008

Li, Z., Shi, Z. and Ososanya, E. T. (1996). Evaluation of bridge conditions using artificial neural networks. *Proceedings of SOUTHEASTCON'96*, 366–369. Tampa, FL.

Lim, S. and Chi, S. (2019a). Xgboost application on bridge management systems for proactive damage estimation. *Advanced Engineering Informatics*, 41, 100922.

Lim, S. and Chi, S. (2019b). Bridge Damage Prediction Using Deep Neural Network. *Proceedings of ASCE International Conference on Computing in Civil Engineering 2019*, 219-225. Atlanta, GA.

Lundberg, S. M., Erion, G. G., and Lee, S.-I. (2018). Consistent individualized feature attribution for tree ensembles. *arXiv preprint arXiv:1802.03888*.

Lysenko, A., Sharma, A., Boroevich, K. A., and Tsunoda, T. (2018). An integrative machine learning approach for prediction of toxicity-related drug safety. *Life Science Alliance*, 1(6), e201800098. doi: http://doi.org/10.26508/lsa.201800098 %J Life Science Alliance

Main Roads Western Australia. (2008). Detailed visual bridge inspection guidelines for concrete and steel bridges (level 2 inspections), 6706-02-2233. Retrieved January 4, 2019, from https://www.mainroads.wa.gov.au/Documents/Concrete %5E26amp%5E3B Steel Bridges Detailed Visual Inspection (Level 2) Guidelines Rev4A.RCN-D17%5E23337746.PDF

Main Roads Western Australia. (2009). Routine visual bridge inspection guidelines (level 1 inspections) for bridges, , 6706-02-2234. Retrieved January 4, 2019, from https://www.mainroads.wa.gov.au/Documents/Routine Visual Inspection Guidelines - 6706-02-2234 - Rev2(2).RCN-D13%5E23349420.PDF

Melhem, H. G. and Cheng, Y. (2003). Prediction of remaining service life of bridge decks using machine learning. *Journal of Computing in Civil Engineering*, 17(1), 1–9. doi: http://doi.org/10.1061/(ASCE)0887-3801(2003)17:1(1)

Melhem, H. G., Cheng, Y., Kossler, D., and Scherschligt, D. (2003). Wrapper methods for inductive learning: example application to bridge decks. *Journal of Computing in Civil Engineering*, 17(1), 46–57. doi: http://doi.org/10.1061/(ASCE)0887-3801(2003)17:1(46)

Ministry of Land, Infrastructure, and Transport of Korea (MOLIT). (2015). *Yearbook of road bridges and tunnels*, Sejong, Korea: MOLIT.

Ministry of Land, Infrastructure and Transport of Korea (MOLIT) and Korea Infrastructure Safety Corporation (KISTEC). (2012). A guide book of detailed instructions of safety inspection and precise diagnosis.

Ministry of Land, Infrastructure and Transport of Korea (MOLIT) and Korea Infrastructure Safety Corporation (KISTEC). (2017). Detailed instructions of safety inspection and precise diagnosis.

Mirzaei, Z., Adey, B. T., Klatter, L., and Kong, J. S. (2012). Overview of existing Bridge Management Systems-Report by the IABMAS Bridge Management Committee. *Proceedings of 6th International Conference on Bridge Maintenance, Safety and Management (IABMAS 2012)*. International Association for Bridge Maintenance And Safety (IABMAS).

Morcous, G. (2000). Case-Based Reasoning for Modeling Bridge Deterioration, Doctoral Dissertation. Quebec, Canada: Concordia University Montreal.

Morcous, G. (2005). Modeling bridge deck deterioration by using decision tree algorithms. *Transportation Research Record: Journal of the Transportation Research Board*, 11(S), 509–516. doi: http://doi.org/10.3141/trr.11s.e383j231l68k41h2

Morcous, G., Rivard, H., and Hanna, A. (2002). Modeling bridge deterioration using case-based reasoning. *Journal of Infrastructure Systems*, 8(3), 86–95. doi: http://doi.org/10.1061/(ASCE)1076-0342(2002)8:3(86)

Norwegian Public Roads Administration. (2005). Handbook for bridge inspections. Retrieved January 6, 2019, from https://www.tsp2.org/library-tsp2/uploads/48/Handbook_of_Bridge_Inspections_Part_1.pdf

Nsabimana, P. (2015). A method for prioritisation of concrete bridge inspections in South Africa, Master's thesis. Stellenbosch, South Africa: Stellenbosch University.

Ontario Ministry of Transportation. (2008). Ontario structure inspection manual. Retrieved January 4, 2019, from https://www.ogra.org/files/OSIM April 2008.pdf

Papadopoulos, S. and Kontokosta, C. E. (2019). Grading buildings on energy performance using city benchmarking data. *Applied Energy*, 233, 244–253. doi: http://doi.org/10.1016/j.apenergy.2018.10.053

Patidar, P. and Tiwari, A. (2013). Handling missing value in decision tree algorithm. *International Journal of Computer Applications*, 70(13), 31–36.

Portland Cement Assosciation. (2002). Types and causes of concrete deterioration. *Concrete Information*, IS536.

PredPol. (2014). Recent examples of crime reduction. Retrieved February 10, 2019, from https://www.predpol.com/results/

Reddy, M. C., Balasubramanyam, P., and Subbarayudu, M. (2013). An effective approach to resolve multicollinearity in agriculture data. *International Journal of Research in Electronics and Computer Engineering*, 1(1), 27–30.

Schellhammer, M. (2015.7.31.). Casey calling for more railroad bridge inspectors. Retrieved April 10, 2017, http://www.bradfordera.com/news/casey-calling-for-more-railroad-bridge-inspectors/article_d2cfc39c-372a-11e5-8d6d-3b47d0efa007.html

Scherer, W. T. and Glagola, D. M. (1994). Markovian models for bridge maintenance management. *Journal of Transportation Engineering*, 120(1), 37–51. doi: http://doi.org/10.1061/(ASCE)0733-947X(1994)120:1(37)

Scikit-Learn Developers. (2019a). 3.3.2.8.2. Multiclass and multilabel classification. Retrieved December 20, 2019, from https://scikit-learn.org/stable/modules/model_evaluation.html#multiclass-and-multilabel-classification

Scikit-Learn Developers. (2019b). Selecting the number of clusters with silhouette analysis on KMeans clustering. Retrieved April 12, 2019, from https://scikit-learn.org/stable/auto_examples/cluster/plot_kmeans_silhouette_analysis.html

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

Shan, Y., Contreras-Nieto, C., and Lewis, P. (2016). Using data analytics to characterize steel bridge deterioration. *Proceedings of Construction Research Congress 2016*, 1691–1699. San Juan, Puerto Rico. doi: http://doi.org/10.1061/9780784479827.169

Song, X. F., Wei, J. F., and He, T. S. H. (2009). A method to repair concrete leakage through cracks by synthesizing super-absorbent resin in situ. *Construction and Building Materials*, 23(1), 386–391. doi: http://doi.org/10.1016/j.conbuildmat.2007.11.009

Su, H. J. (2003). A correlation study of the existing bridges for failure analysis-case study of Taichung County, Master's thesis. Taichung City, Taiwan: Feng Chia University.

The American Society of Civil Engineers (ASCE). (2017). ASCE's 2017 infrastructure report card. Retrieved from https://www.infrastructure reportcard.org/cat-item/bridges/

The Highways Agency, Transport Scotland, Welsh Assembly Government, and The Department for Regional Development Northern Ireland. (2007). As built, operational and maintenance records for highway structures, BD 62/07. Retrieved January 4, 2019, from http://www. standardsforhighways.co.uk/ha/standards/dmrb/vol3/section2/bd6207.pdf

The Highways Agency, Transport Scotland, Welsh Assembly Government, and The Department for Regional Development Northern Ireland. (2017). Inspection of highway structures, BD 63/17. Retrieved January 4, 2019, from http://www.standardsforhighways.co.uk/ha/standards/ dmrb/vol3/section1/bd6317.pdf

Therneau, T. M. and Atkinson, E. J. (1997). *An introduction to recursive partitioning using the RPART routines*. Rochester, NY: Mayo Foundation.

Tokdemir, O. B., Ayvalik, C., and Mohammadi, J. (2000). Prediction of highway bridge performance by artificial neural networks and genetic algorithms. *Proceedings of the 17th International Association for Automation and Robotics in Construction (ISARC)*, 1091–1098. Taipei, Taiwan.

Ture, M., Tokatli, F., and Kurt, I. (2009). Using Kaplan–Meier analysis together with decision tree methods (C&RT, CHAID, QUEST, C4. 5 and ID3) in determining recurrence-free survival of breast cancer patients. *Expert Systems with Applications*, 36(2), 2017–2026.

Tuv, E., Borisov, A., Runger, G., and Torkkola, K. (2009). Feature selection with ensembles, artificial variables, and redundancy elimination. *Journal of Machine Learning Research*, 10(Jul), 1341–1366. Retrieved from http://www.jmlr.org/papers/volume10/tuv09a/tuv09a.pdf

Urbanowicz, R. J. and Browne, W. N. (2017). *Introduction to learning classifier systems*. Berlin, Germany: Springer.

Wang, H., Yang, F., and Luo, Z. (2016). An experimental study of the intrinsic stability of random forest variable importance measures. *BMC Bioinformatics*, 17(1), 60–77. BioMed Central. doi: http://doi.org/10.1186/s12859-016-0900-5

Wang, X., Nguyen, M., Foliente, G., and Ye, L. (2007). An approach to modelling concrete bridge condition deterioration using a statistical causal relationship based on inspection data. *Structure and Infrastructure Engineering*, 3(1), 3–15. doi: http://doi.org/10.1080/15732470500103682

Winn, E. K. (2011). Artificial neural network models for the prediction of bridge deck condition ratings, Master's Thesis. East Lansing, United States: Michigan State University.

Xgboost developers. (2016). XGBoost documentation. Retrieved from https://xgboost.readthedocs.io/en/latest/index.html

Yianni, P. C., Neves, L. C., Rama, D., Andrews, J. D., and Dean, R. (2016). Incorporating local environmental factors into railway bridge asset management. *Engineering Structures*, 128, 362–373. doi: http://doi.org/10.1016/j.engstruct.2016.09.038

Zhang, D., Qian, L., Mao, B., Huang, C., Huang, B., and Si, Y. (2018). A data-driven design for fault detection of wind turbines using random forests and xgboost. *IEEE Access*, 6, 21020–21031. doi: http://doi.org/10.1109/ACCESS.2018.2818678

Zhang, Y., O'Connor, S. M., Linden, G. W. van der, Prakash, A., and Lynch, J. P. (2016). SenStore: A scalable cyberinfrastructure platform for implementation of data-to-decision frameworks for infrastructure health management. *Journal of Computing in Civil Engineering*, 30(5), 4016011–4016012. doi: http://doi.org/10.1061/(ASCE)CP.1943-5487.0000560

Zhao, Z. and Chen, C. (2002). A fuzzy system for concrete bridge damage diagnosis. *Computers & Structures*, 80(7–8), 629–641. doi: http://doi.org/10.1016/S0045-7949(02)00031-7

Appendix

Appendix A: Bridge Damage Portfolio Generation System

A-I. User Interface Design

A-II. HTML Code of the User Interface

A-III. Python Code of the Portfolio Generation System

A-I. User Interface Design

At the project level, such inspectors and maintenance decision makers enters date of inspection and target bridge number (Figure A-1), and then they can obtain the estimated severity levels in a tabular format. In the case of multiple bridges at the network level, one types a code of a specific region with bridges to check, from district to cities or counties in a 5 digit or 2 digit area code respectively with inspection date (Figure A-2), and consequently he or she can get the multiple portfolios.

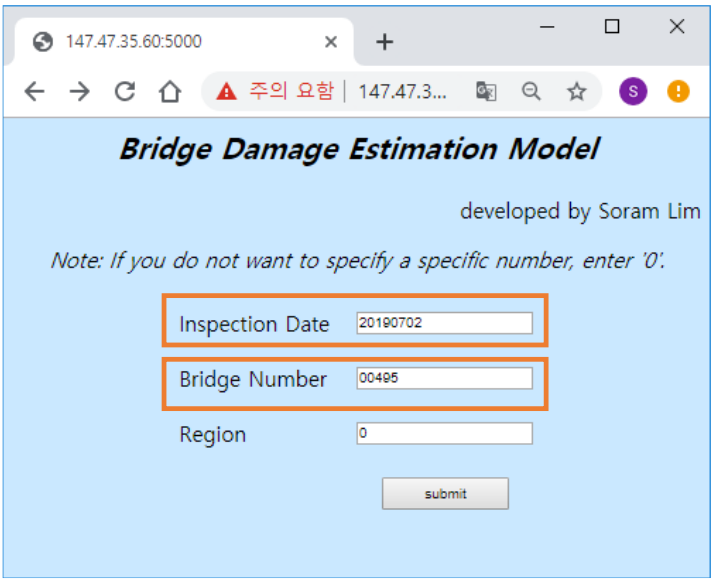


Figure A-1. User Interface Index Page Design (Project Level)

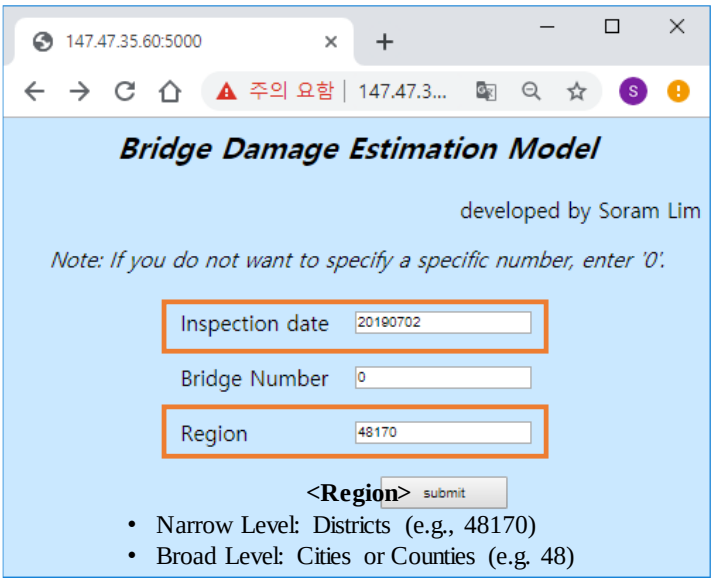


Figure A-2. User Interface Index Page Design (Network Level)

A-II. HTML Code of the User Interface

Code of Index.html

```
<html>
<style>
p {
   font-size: 20px;
}
.title {
   font-size: 28px;
   font-style: italic;
   font-weight: bold;
   }
body {
   background-image:url('/templates/bg2.png');
   background-color: #cae8ff;
   background-repeat:no-repeat;
   background-position: center;
   background-size:cover;}
</style>

    <body>
        <p class='title' style="text-align:center"> Bridge Damage Estimation Model </p>
        <p style="text-align:right"> developed by Soram Lim </p>
        <p style="text-align:center; font-style:italic;"> Note: If you do not want to specify a specific number, enter '0'. </p>
        <p>
        <p>
        <form action = "http://147.47.35.60:5000/result" method = "get">
```

```
        <center><div>
        <table>
                <tr>
                    <td><p>Inspection
                    date&amp;nbsp;&amp;nbsp;&amp;nbsp;</p></td>
                    <td height="50"><p><input type = "text" name =
                    "PAST_FR_DT" /></p></td>
                </tr>
                <tr>
                    <td><p>Bridge Number     </p></td>
                    <td height="50"><p><input type = "text" name =
                    "BRDG_NO" /></p></td>
                </tr>
                <tr>
                    <td><p>Region     </p></td>
                    <td height="50"><p><input type = "text" name =
                    "SIDO" /></p></td>
                </tr>
                <tr>
                    <td> </td>
                    <td height="80"><p><center><input type =
                    "submit" value = "submit" style="width: 120px;
                    height: 30px;"/></center></p></td>
                </tr>
        </table>
        </div></center>
    </form>

</body>
</html>
```

Code of Results.html

```
<!DOCTYPE html>
<html lang="en">

<style>
p {
   font-size: 20px;
}
body {
   background-color: #cae8ff;
}
</style>

<head>
      <meta charset="UTF-8">
      <title>Title</title>
</head>
<body>
{% for data in datalist %}
<p>Inspection Date: {{ datalist[loop.index0][0] }}</p>
<p>Bridge Number: {{ datalist[loop.index0][1] }}</p>
<p>Region: {{ datalist[loop.index0][2] }} </p>
<p>Age: {{ datalist[loop.index0][3] }}</p>
<p>Mainstructure Type: {{ datalist[loop.index0][4] }}</p>
<p>Cell Counts: {{ countcells[loop.index0] }}</p>
<p>{{ tables[loop.index0]|safe }}<p>
{% endfor %}
</body>
</html>
```

A-III. Python Code of the Portfolio Generation System

```
# Basic settings
import flask
import numpy as np
import tensorflow as tf
from keras.models import load_model

from flask import Flask, render_template, request

app = flask.Flask(__name__)

def sendResponse(responseObj):
    response = flask.jsonify(responseObj)
    response.headers.add('Access-Control-Allow-Origin', '*')
    response.headers.add('Access-Control-Allow-Methods', 'GET')
    response.headers.add('Access-Control-Allow-Headers',  'accept,content-
type,Origin,X-Requested-With,Content-
Type,access_token,Accept,Authorization,source')
    response.headers.add('Access-Control-Allow-Credentials', True)
    return response

# Loading developed model
## Intermediate omission

# List of birdge loading
unique_23 = []
```

```
with open('190612_unique_23.csv', newline='') as f:
        rawdata = csv.reader(f, dialect='excel')
        for row in rawdata:
                unique_23.append(row[0:5])

import pandas as pd

unique_23 = pd.DataFrame(unique_23)

header = unique_23.iloc[0]
unique_23 = unique_23[1:]
unique_23.rename(columns=header, inplace=True)

print(unique_23.columns, '\n')
print(unique_23.shape, '\n')
print(unique_23.head(7))

# API for estimation
@app.route('/')
def got():
    return render_template('Index.html')

from datetime import datetime

from flask import Flask, request, render_template, session, redirect
from collections import OrderedDict

@app.route('/result', methods=["POST", "GET"])
def html_table():
        if request.method == 'GET':
```

```
# for date = 0
PAST_FR_DT = flask.request.args.get('PAST_FR_DT')

if PAST_FR_DT == '0':
    PAST_FR_DT = datetime.now().strftime('%Y%m%d')

BRDG_NO = flask.request.args.get('BRDG_NO')
BRDG_NO = int(BRDG_NO)

Reg = flask.request.args.get('SIDO')

# for a bridge
if BRDG_NO > 0:
    INPUT2 = INPUT[(INPUT.BRDG_SEQ == BRDG_NO)]

    tbls = []
    datalist = []
    countcell = []

    # date input exist
    INPUT2['PAST_FR_DT'] = PAST_FR_DT
    INPUT2['PAST_TO_DT'] = PAST_FR_DT

    x0 = int(INPUT2['CMPLT_DT.1'].iloc[0])
    x0 = str(x0)

    x = datetime.strptime(x0, '%Y%m%d')
    y = datetime.strptime(PAST_FR_DT, '%Y%m%d')
```

```
z1 = y-x
z1 = z1.days
z2 = (y-x).days / 365.25
BRDG_AGE_YEAR = int(z2)

INPUT2['BRDG_AGE_DAY'] = z1
INPUT2['BRDG_AGE_YEAR'] = z2

Y_predict_sum, Y_actual_sum = atesting(INPUT2)

print(Y_predict_sum)

DATE = str(y.year)+"-"+str(y.month)+"-"+str(y.day)

print(INPUT2.shape)
print(INPUT2.head(3))
print(INPUT2['SIDO'])
SIDO = INPUT2['SIDO'].iloc[0]
tbls.append(Y_predict_sum.to_html(classes='data'))
data = [DATE, BRDG_NO, SIDO, BRDG_AGE_YEAR, '23']
datalist.append(data)

Y_predict_sum.replace({0:np.nan}, inplace=True)
ccell = Y_predict_sum.iloc[:,2:].count().sum()
countcell.append(ccell)

values = list(zip(datalist,tbls))
dictionary = dict(zip(countcell, values))
ordered_dict = dict(OrderedDict(sorted(dictionary.items(),
key=lambda t: t[0], reverse=True)))
```

```
        cells = list(ordered_dict.keys())
        values2 = list(ordered_dict.values())
        df2 = pd.DataFrame(values2)
        datalist2 = list(df2[df2.columns[0]])
        tbls2 = list(df2[df2.columns[1]])

# for a region
elif (BRDG_NO == 0) & (len(Reg) >= 2) :

        if len(Reg) == 2:
                SIDO = Reg
                unique_23_sel = unique_23[(unique_23.SIDO == SIDO)]

        elif len(Reg) == 5:
                CBJCODE = Reg
                unique_23_sel = unique_23[(unique_23.CBJCODE ==
                CBJCODE)]

        tbls = []

        datalist = []
        countcell = []

        for i in range(0, len(unique_23_sel)):
                BRDG_NO = unique_23_sel.iloc[i, 0]
                BRDG_NO = int(BRDG_NO)

                INPUT2 = INPUT[(INPUT.BRDG_SEQ == BRDG_NO)]
                INPUT2['PAST_FR_DT'] = PAST_FR_DT
                INPUT2['PAST_TO_DT'] = PAST_FR_DT
```

```
        x0 = int(INPUT2['CMPLT_DT.1'].iloc[0])
        x0 = str(x0)

        x = datetime.strptime(x0, '%Y%m%d')
        y = datetime.strptime(PAST_FR_DT, '%Y%m%d')

        z1 = y-x
        z1 = z1.days
        z2 = (y-x).days / 365.25
        BRDG_AGE_YEAR = int(z2)

        INPUT2['BRDG_AGE_DAY'] = z1
        INPUT2['BRDG_AGE_YEAR'] = z2

        Y_predict_sum, Y_actual_sum = atesting(INPUT2)

        print(Y_predict_sum)

        DATE = str(y.year)+"-"+str(y.month)+"-"+str(y.day)
        tbls.append(Y_predict_sum.to_html(classes='data'))
        data = [DATE, BRDG_NO, Reg, BRDG_AGE_YEAR,
        '23']
        datalist.append(data)

        Y_predict_sum.replace({0:np.nan}, inplace=True)
        ccell = Y_predict_sum.iloc[:,2:].count().sum()
        countcell.append(ccell)

values = list(zip(datalist,tbls))
dictionary = dict(zip(countcell, values))
```

```
        ordered_dict=dict(OrderedDict(sorted(dictionary.items(),
        key=lambda t: t[0], reverse=True)))

        cells = list(ordered_dict.keys())
        values2 = list(ordered_dict.values())
        df2 = pd.DataFrame(values2)
        datalist2 = list(df2[df2.columns[0]])
        tbls2 = list(df2[df2.columns[1]])

# for a specific date, all bridges, all regions
elif (BRDG_NO == 0) & (len(Reg) == 1) :

        tbls = []
        datalist = []
        countcell = []

        for i in range(0, len(unique_23)):
            BRDG_NO = unique_23.iloc[i, 0]
            BRDG_NO = int(BRDG_NO)

            INPUT2 = INPUT[(INPUT.BRDG_SEQ == BRDG_NO)]
            INPUT2['PAST_FR_DT'] = PAST_FR_DT
            INPUT2['PAST_TO_DT'] = PAST_FR_DT

            x0 = int(INPUT2['CMPLT_DT.1'].iloc[0])
            x0 = str(x0)

            x = datetime.strptime(x0, '%Y%m%d')
            y = datetime.strptime(PAST_FR_DT, '%Y%m%d')
```

```
        z1 = y-x
        z1 = z1.days
        z2 = (y-x).days / 365.25
        BRDG_AGE_YEAR = int(z2)

        INPUT2['BRDG_AGE_DAY'] = z1
        INPUT2['BRDG_AGE_YEAR'] = z2

        Y_predict_sum, Y_actual_sum = atesting(INPUT2)

        print(Y_predict_sum)

        DATE = str(y.year)+"-"+str(y.month)+"-"+str(y.day)
        tbls.append(Y_predict_sum.to_html(classes='data'))
        SIDO = INPUT2['SIDO'][0]
        data = [DATE, BRDG_NO, SIDO, BRDG_AGE_YEAR,
        '23']
        datalist.append(data)

        Y_predict_sum.replace({0:np.nan}, inplace=True)
        ccell = Y_predict_sum.iloc[:,2:].count().sum()
        countcell.append(ccell)

values = list(zip(datalist,tbls))
dictionary = dict(zip(countcell, values))
ordered_dict = dict(OrderedDict(sorted(dictionary.items(),
key=lambda t: t[0], reverse=True)))

cells = list(ordered_dict.keys())
values2 = list(ordered_dict.values())
```

```
            df2 = pd.DataFrame(values2)
            datalist2 = list(df2[df2.columns[0]])
            tbls2 = list(df2[df2.columns[1]])

        return render_template('Results.html', countcells=cells, tables=tbls2,
        datalist=datalist2)

if __name__ == "__main__":
    print(("* Loading Keras model and Flask starting server..."
    "please wait until server has fully started"))

    app.run(host='147.47.35.60', debug='True')
```

Abstract (Korean)

인공지능을 활용한 교량 손상 규명 및 심각도 추정

교량 점검은 교량 유지관리를 위한 기반 정보를 얻는 중요한 단계이다. 근래에 들어 노후화된 교량의 수가 급증하고 있으나 점검 예산 및 시간과 전문 인력의 수가 부족하여 점검의 품질이 저하될 위험이 있다. 본 연구에서는 인공지능을 활용하여 교량 손상의 영향 인자를 규명하고 특정 시점에 점검 대상 교량에서 발생하는 손상의 유형, 위치, 등급, 심각도 등 점검자에게 유용한 정보를 추정하는 기법을 개발하였다. 연구 대상은 프리스트레스트 콘크리트 I형 (Pre-Stressed Concrete I type, PSCI) 교량의 바닥판이며, 발생 가능한 구조적 균열, 망상균열, 박리, 파손, 백태, 철근 노출 및 부식 등 총 7가지 손상유형을 고려하였다.

먼저 한국건설기술연구원에서 관리하고 있는 교량유지관리 시스템 (Korean Bridge Management System, KOBMS)의 제원, 구조, 교통량, 점검데이터와 기상청으로부터 수집된 기상데이터를 가공하여 59개의 독립변수와 2개의 종속변수로 구성된 투입 데이터를 구성하였다. 다음으로 상관분석기법을 활용하여 유사한 영향력을 가진 각 변수 쌍에서 의미적 중요도를 고려하여 11개 변수를 제거하였고, 인공지능기법 중 하나인 의사결정나무방법 3가지를 비교하였으며, 그 결과 가장 높은 성능을 보인 Extreme Gradient Boosting (XGBoost)를 이용하여 손상유형별 영향인자를 도출하였다. 바닥판에 발생하는 손상유형 전반에 대해 고정하중과

활하중, 공용년수, 상대습도, 거더강도가 주요 영향인자로 규명되었다. 손상유형별로는 각 특성과 관련 있는 제원, 구조, 환경적 인자들이 도출되었다. 이러한 과정을 통해 도출한 7가지 손상유형별 영향인자를 투입하여 손상물량으로 정의된 심각도를 추정하는 인공지능 모델을 심층신경망 (Deep Neural Networks, DNN)기법과 XGBoost를 활용하여 개발하였다. 연구 결과, 더 높은 정확도를 보이는 22개의 세부모델로 구성된 XGBoost 모델을 최종 모델로 선정하였다. 개발한 모델을 활용하여 발생 가능한 손상을 위치별로 나타낸 손상 포트폴리오를 개별 및 지역별로 생성하는 시스템을 개발하였으며, 모델 확장 가능성 확인을 위해 PSCI 형식 교량의 거더 부재에 대해 동일한 방법론을 적용하여 손상 규명 및 심각도 추정 모델이 확장 가능함을 확인하였다.

본 연구는 산재되어 있었던 교량의 여러 손상유형별 영향인자에 대한 지식을 바닥판 점검 데이터를 활용하여 정량적으로 규명하였으며, 인공지능 기법을 활용하여 점검 지원 정보를 체계적으로 생산한 선구적인 시도이다. 본 연구에서 개발한 인공지능 기법을 활용하여 기존에 관리되고 있던 교량뿐만 아니라, 과거 점검 이력이 없는 교량에 대해서도 점검에 필요한 추정 정보를 사전에 제공함으로써 점검 시간과 점검 누락 위험을 감소시킬 수 있다. 또한 점검 대상 교량이 다수인 경우, 이에 대한 점검 우선 순위를 제공함으로써 점검 시간, 예산, 인력을 효율적으로 분배할 수 있도록 한다. 본 연구결과의 다양한 부재와 구조형식으로의 확장을 통해 교량의 생애주기 비용 감소와 교량 수명 연장에 기여할 수 있으며, 나아가 안전한 사회 구현에 기여할 수 있을 것으로 기대된다.

주요어: 교량 점검, 사전 정보, 교량 손상 영향 요인 규명, 교량 손상 위치 추정, 교량 손상 심각도 추정, 인공지능모델, XGBoost, 심층신경망

학 번: 2015-31060 **이름:** 임 소 람