Deep learning- and infrared thermography-based subsurface damage detection in a steel bridge

By

Rahmat Ali

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfillment of the requirements of the degree of

Master of Science

Department of Civil Engineering

Faculty of Engineering

University of Manitoba

Winnipeg

ABTRACT

The aging and deterioration of bridge infrastructure is becoming a serious issue around the world. In this research, a new deep-learning-based method is proposed to detect subsurface damage in the steel elements of a bridge using thermography, without physical contact. Thermal images of the structural steel elements on the Arlington Bridge in Winnipeg, Manitoba, were captured using an uncooled microbolometer and were then used to train and validate a deep inception neural network; a maximum testing accuracy of 96% was achieved. Next, ultrasonic pulse velocity tests were conducted for validating thermal-image-based subsurface damage detection, and a contour map of the sections was plotted and compared with a deep-learning-based results. This research demonstrates that the combination of infrared thermal technology with a deep neural network is a practical approach to autonomously detecting subsurface damage in the elements of a steel bridge, with minimum human intervention.

ACKNOWLEDGEMENTS

I am greatly thankful to my thesis advisor Prof. Dr. Young-Jin Cha for providing me constant encouragement and continuous support during my MSc study. His patience, motivation, and immense knowledge support me in many ways during my study period.

Beside my advisor, I would like to acknowledge the consistent support, insightful suggestions and excellent guidance of my committee member Prof. Dr. Dagmar Svecova, Head of the Department of Civil Engineering. I would like to take the opportunity to thanks my committee member Prof. Mohammad T. Araji for his guidance and encouragement throughout this work.

I am also grateful to the Structures Lab Manager Dr. Chad Klowak, Graudate Student Advisor Jeniffer Merrell and my colleagues for their technical assistance and guidance.

The financial support provided by the Government of Manitoba through the Manitoba Graduate Scholarship is gratefully acknowledged.

Finally, thanks to my parents, my brother Sher Shah, family members and my best friends Chloe and Jiangyu Zeng for their constant support and encouragement.

Rahmat Ali

Table of Contents

List of Tables

List of Figures

Copyright

This thesis includes a detail explanation of a journal paper submitted to “Construction and Building Material” journal under the title “Ali, R. Cha, Y. J. (2018). Subsurface damage detection of a steel bridge using deep learning and uncooled micro-bolometer”. The submitted journal paper are still under review. All the relevant sections are referenced in each chapter accordingly.

Chapter 1. Introduction

This chapter provides background information on the subject of structural health monitoring (SHM). A summary of the relevant literature is provided, as well as the problem statement and research objectives. Lastly, the thesis organization is outlined.

1.1 Overview

SHM has drawn the attention of wide variety of engineering and other research communities in the last two decades, and various approaches for dealing with complex problems in SHM have been developed. All structures, including buildings, bridges, tunnels, dams, oil and gas pipelines, pavements, rails, wind energy plants, trains, ships, planes and myriad others, are subjected to various external and internal factors which may lead to deterioration or malfunction. These unwanted deterioration malfunction are usually a result of inferior material, low quality inspection, improper construction process, extreme loading or environmental conditions, or an accident.

To rapidly detect variations in material properties and to respond to them in a more sophisticated way, it is imperative to implement damage detection systems. SHM technology can detect anomalies, thereby increasing the efficiency of maintenance and repair, and reducing operating and maintenance costs. To ensure the safety of the public, non-destructive testing (NDT) has been conducted. NDT is an extensive analysis technique used to assess the properties of a materials, structural members or entire structural system without causing damage [1]. NDT technique ensures that the structures are examined and monitored without damaging the structure. Numerous NDT techniques are studied for various structures, such as hammering and chain dragging for bridge deck [2], fibre-optic sensors for concrete beams and prestressed concrete structures [3, 4], acoustic emission sensor for concrete cube samples [5, 6], ground penetration

radar for pavements [7], air-coupled and impact echo techniques for reinforced concrete [8], and computer vision techniques for concrete cracks and loosened bolts [9-12]. The computer vison-based method uses image processing techniques for damage detection in structures and have been aimed to make the detection process easier and more economical. Four different edge detection methods, including fast Fourier Transform (FFT), Canny edge detector, fast Haar Transform (FHT) and Sobel edge detector, have been evaluated in a comparative study for detecting surface cracks in concrete [13]. Among these methods, FHT was found as the best alternative for surface crack detection in concrete. It was found in literature that in the last two decades, research have been conducted on surface and visible damage detection [9-11, 13, 14].

Moreover, image processing techniques had been incorporated with a sliding window technique for detecting concrete cracks to localize the detected damage within an input image, and the result was quite good [9]. In order to further increase the accuracy and overall performance of image processing techniques as a damage detection method, many researchers have been actively involved by employing machine learning techniques [15]. Despite these improvements, these methods require time-consuming pre-processing and post-processing steps, which is a major drawback. Additionally, these methods can only detect a single damage type at a time [16]. To resolve these issues, Cha et al [11] approached a new innovative idea by implementing a deep learning method for detecting surface cracks in concrete. This new method is autonomous in nature and maintained a high accuracy of approximately 98%, without any pre-processing and post-processing steps [11].

Recently, Cha et al [16] used faster region-based convolution neural network (Faster R-CNN) [17] for detecting damage. This method has been proven to easily detect five types of damage, namely concrete cracks, bolt corrosion, high steel corrosion, medium steel corrosion, and

steel delamination, and could potentially also detect multiple other types of damage. The method was found to be robust and satisfactory. The average precision for the above five damage types were 90.6%, 98.1%, 83.4%, 82.1% and 84.7%, respectively [16]. It was derived from the aforementioned research that a digital camera can be successfully incorporated with deep learning methods to detect multiple surface damage in a single image. Despite such great research, it is still imperative to concentrate on subsurface damage in structures. Subsurface damage is hidden and is more unpredictable. Therefore, it is extremely important to detect this damage before it leads to severe results. One way to spot this interior and subsurface damage is to use infrared thermography (IRT). IRT is a widely used non-contact and in-situ method.

IRT functions by detecting the infrared energy emitted from objects, and then converts this energy into temperature, and displays the temperature distribution image of the surface. One of the advantages of the IRT camera is that it can detect things which a normal camera or naked eye cannot easily detect [18, 19]. Active and passive thermography are the two common techniques of IRT. Passive thermography utilizes natural sources of heat, such as the sun. On the other hand, active thermography requires an external heating source during the test [20]. IRT has been applied in damage detection in the last several years. Several researchers have applied IRT for subsurface delamination detection in concrete bridge deck slabs [21-23]. Vaghefi et al [22] compared infrared thermography with chain dragging, and found that IRT can detect damage with 80% accuracy, whereas chain dragging can detect damage with 40% accuracy in concrete bridges. However, another study conducted in the same year by Oh et al [23] discovered that impact echo and IRT achieved a very good accuracy of almost 100% for subsurface delamination in concrete bridges, whereas chain dragging achieved 75% accuracy.

Gucunki et al [21] concluded that IRT has high potential for detecting subsurface damage. Omar and Nehdi used infrared thermography with ground penetration radar for detecting subsurface damage in bridge deck slabs [24]. Many experiments have focused on the detection of various types of subsurface damage in concrete bridge decks, such as delamination, cracks, and corrosion. However, since steel bridges are ubiquitous in Canada, and since approximately 40% of these bridges were built over 50 years ago [25], it is extremely important to further develop techniques for detecting subsurface damage in steel elements of bridges. Common types of subsurface damage which occur in steel elements include delamination, coating delamination, and corrosion below surface paints. Even with the application of various galvanizing techniques, the delamination and corrosion of the steel below the coating are difficult to completely prevent. Therefore, it is very important to detect subsurface delamination and corrosion below coatings or paint in the early stages to avoid unwanted damage.

To date, studies have been made on surface corrosion. However, subsurface corrosion and subsurface delamination is still unexplored [26]. Johannah [26] studied corrosion mechanisms under organic coatings. Marcus and Amirudin elaborated on mechanisms of corrosion under paints, and categorized subsurface coating delamination into two types: anodic delamination, which causes dissolution of the steel under coatings, and cathodic delamination, which causes coating delamination between subsurface and coating [27, 28]. Recently, an experimental study was carried out by introducing defects artificially in a piece of real pipeline system for oil transportation, and thermography was used for detecting damage. They concluded that thermography is an effective solutions for detecting subsurface damage.

However, to improve the results further, an active heat source such as a flash or a lamp should be used as an active excitation [29]. The active excitation helps to differentiate between the

damaged portion and intact portion more clearly, with high accuracy. The previous literature demonstrates that IRT technology can be utilized to detect subsurface damage in steel. In this research, the IRT method was chosen to investigate subsurface damage in steel elements of a bridge, because of the availability of good quality uncooled microbolometers that can effectively detect subsurface damage. However, in case of large structures, it is often difficult to apply IRT efficiently for detecting subsurface damage. Analysis of the vast amount of thermal image data requires an experienced interpreter, and requires more time to accurately identify subsurface damage. Therefore, for this research, a deep learning technique is incorporated with IRT to automatically detect this damage. Deep learning has become increasingly predominant in various fields. Since SHM is a field that largely demands NDT techniques, therefore recently deep learning has become an essential tool in SHM. In this study, an inception neural network of GoogleNet [30] was used to detect subsurface damage in thermal images collected from the steel bridge.

1.2 Problem definition

This research is an attempt to detect subsurface damage in steel elements of a bridge by obtaining thermal images using an uncooled microbolometer. Although surface damage such as cracks or surface corrosion can be investigated manually, internal damage such as subsurface delamination, debonding, and corrosion below coatings requires a more careful investigation using monitoring devices. All the manual inspections are often erroneous and expensive. Moreover, applying IRT to large structures requires extensive data analysis and is labor intensive. The development of an automated technique for data analysis will reduce human involvement, and will expedite the inspection process of a structure. The use of an inception neural network [30] will significantly reduce the time required for the classification of subsurface damage and intact surface. The key

advantage of the proposed method based on inception neural network and transfer learning is its fast, accurate and efficient detection of subsurface damage without pre or post-processing steps.

1.3 Objective statement

The purpose of this study is to ameliorate the damage inspection and assessment of large-scale steel bridges. This research is an attempt to develop a system for an efficient identification of subsurface delamination, debonding, or corrosion below the coatings of the steel elements of a bridge. The information obtained during the assessment process could be used for decision making regarding maintenance of structures in a timely manner. To accomplish these objectives, the steps followed are as under:

- Subsurface damage of steel elements of a bridge are observed using infrared thermography.

- Passive thermography is used to detect subsurface damage without using any external excitation, the ideal time window for data collection was investigated.

- An inception neural network was used to build a robust classifier for autonomously detecting interior damage such as subsurface delamination, debonding and corrosion below paint, with minimal human intervention.

- To validate the results obtained from the deep learning-based subsurface damage detection, an Ultrasonic Pulse Velocity test is conducted and surface maps are produced using a surfer contour software.

1.4 Thesis organization

This thesis contains six chapters. Chapter 1 presents the overall theme and summarizes the problem and objective of the research. Chapter 2 gives an overview of various methods of damage detection in the field of SHM, and the current relevant areas of SHM research. It also reviews the limitations of the methods currently available for structure health monitoring. Chapter 3 provides the overview of the proposed method and the detailed working procedure of the inception neural network. This chapter also covers all the relevant concepts and terminology used in inception neural networks. Chapter 4 explains the data collection procedure in the field using infrared thermography. A data preparation method is also discussed in detail in this chapter. Chapter 5 explains the training and validation results of the given network. Additionally, this section covers the network used for the prediction of the subsurface damage in the steel elements of the bridge. The results and discussion of test images are also included in this section. The results obtained from thermal images and deep learning are included in this section, and are compared and validated with the ultrasonic pulse velocity (UPV) results. Chapter 6 summarizes the thesis and includes future recommendations.

Chapter 2. Literature review

Most infrastructures are difficult to investigate by traditional visual inspection by trained engineers, due to the inaccessibility of the structures and potential risks to the inspectors [31]. In order to find proper method available for internal damage of steel members, extensive literature review is conducted in this Chapter.

2.1 Non-destructive testing

NDT is the process of inspecting, testing, or assessing materials or assemblies without damaging the serviceability of the part or system [32]. The key objective of NDT is to study the quality of materials, members of a structure, or entire assemblies, without disrupting their ability to perform their intended functions. All testing procedures that do not affect the short term and long-term integrity of a structure or system, and which maintain the anticipated life of the structure or system, are known as NDT. In general, destructive testing usually determines the failure mechanism of a material by exploring its yield strength, tensile strength, ductility, toughness and compressive strength. Conversely, NDT techniques indicate the properties of materials without reaching the failure level of the materials [33].

Extensive investigation has been carried out in the past decades to improve NDT methods which can identify physical, mechanical, chemical, acoustical and magnetic properties. Recently, advanced, reliable, and sensitive NDT methods have been emerging. For any NDT method to be successfully applied, the properties and application of the tested material must be properly understood. Moreover, the main issues associated with the applied NDT method is very essential. There are a wide number of NDT methods which are currently in practice that are used in civil,

structural and mechanical industries. In the section below, the most widely used NDT methods are explained in detail.

2.1.1 Advantages of NDT

All NDT techniques feature a unique set of advantages [34]. The key advantage of NDT is that the overall structural integrity of the structure or its elements is maintained. Additionally, most NDT based assessments can be carried out multiple times without any damage to the structure. Many NDT techniques can provide timely information that presents the current performance of a structure. Bridges, pipelines, buildings, pressure vessels, aircraft, and railway tracks are examples of infrastructure that are regularly monitored. Additionally, NDT can be used to enhance output and profitability by maintaining the useful life of structures [35].

NDT can be carried out throughout the entire life cycle of a structure. For example, an oil offshore platform is constructed on a seabed. The structural sections used for the construction process will be tested with NDT at several stages of the construction process, such as after forming, welding, and when the sections are placed in their specified positions. Similarly, NDT methods can be employed to monitor the structure during its life time. Essentially, NDT can be applied at each stage of the structure's life span, from the construction stage up to the end of its useful life.

2.1.2 Limitations of NDT

The main disadvantage or limitation of NDT is that no single method can be used to collect all the information necessary to conduct a complete, comprehensive analysis of the structure to be investigated. To do a complete analysis, multiple NDT techniques will be required. Therefore, it directly raises the overall cost and complexity of the analysis. However, the cost of implementing an NDT system can only be justified if it maintains the anticipated lifespan of the infrastructure. All NDT methods have some sort of limitation. For example, it is hard to interpret the complex

results obtained from ground penetration radar (GPR), and therefore a skilled interpreter is required for data assessment [36]. Similarly, for NDT tests, such as the rebound hammer test, results are not reliable, and various factors, such as the age and moisture content of concrete, can greatly affect the final test results. Likewise, most of the NDT techniques are time consuming and require expertise to accurately analyze the data. It is very difficult to analyze all the images obtained from the structure manually. Therefore, the demand for automated options has greatly increased in the last decade.

2.2 NDT methods

This section explains various NDT methods used in the recent years for SHM.

2.2.1 Ultrasonic pulse velocity

A UPV is an in-situ test to check the quality of concrete. The strength and quality of concrete structures such as beam, column, slab etc. can be examined by the measurement of velocity of an ultrasonic pulse passing through structural element [37]. The data can be easily interpreted by considering the velocity. Higher velocities shows good quality and uniformity of a material. On the other hand, lower velocities indicates the presence of defects such as cracks, delamination or voids. It can be utilized to identify abnormal areas in the material [37-41].

The UPV device consists of a transducer which transmits and sends a signal and the signal is received by a receiver and display the results in terms of travel time. When a wave encounter certain defects in a section, certain portion of the energy emitted is reversed to the initial point. In such condition defects are considered as deviations from the normal intact element to be tested. However, zones of high deterioration or interior damage such as cracks will have a comparatively very low velocity. The UPV device requirements are covered in ASTM C597-09 [40]. The main hindrance in using the UPV is the process of coupling the equipment transmitter and receiver to

the surface to be detected. However, the use of viscous material such as vaseline or any other lubricants can improve the coupling. Shah et al [42] studied parameters such as non- linear and linear to detect various level of defects in concrete. They investigated micro level damage as well as macro level damage. The author investigated that wave depletion was highly affected due to variation in power and damage level. In comparison with low voltage level, high voltages diminished significantly with the increasing damage level. Additionally, it was found that higher wave velocities are generally resulted from high power. The authors recommended that more detailed investigation should be carried out on different damage levels and its relationship with pulse velocity [42].

Another detailed research were conducted on approximately 84 different sample by Bogas et al. [39]. The compressive strength of all samples used for experimental work ranges from 30 MPa to 80 MPa. They determined the relationship between compressive strength of concrete and UPV. Moreover, the influence of various parameters such as water content, cement content, aggregate volume, type of cement, type of aggregate, and type of admixture were studied. Additionally, they conducted experiments by replacing normal weight aggregates such as fine and coarse aggregate by a new light weight aggregate. The experimental results showed that in light weight concrete, aggregate volume affects the relationship between compression strength and UPV lesser than the normal weight concrete [39].

2.2.2 Acoustic emission

The acoustic emission (AE) method is a commonly used technique for detecting defects [43-45, 46]. The AE technique was first applied to monitor a bridge by Hopwood [43]. The AE test is based on elastic waves, which are transient in nature. These waves are received and recorded by transducers. The AE technique requires two parts. Firstly, elastic waves are produced from

material deformation, and secondly, these transient elastic waves are received by a transducer. This technique is extensively applied for finding the location of origin of the damage, and for evaluating the severity of the damage [44]. This technique is highly distinguished in NDT due to its potential for automated source location [45].

The AE technique functions by sensing the high frequency sound waves that are emitted during the formation of cracks or any other interior defects. These waves travel through the material from the location of the defect and are received by sensors to be recorded [47]. The key advantage of this method is that it can be utilized for monitoring damage both globally and locally. Additionally, it can be utilized for detection over a range of desirable distances [48]. Moreover, wireless sensing technology based on radio frequency transmission was introduced [49].

Global and local monitoring are two procedures of AE monitoring [50]. In case of bridge structures, flaws and interior damage are stimulated by either a regular traffic load or a heavy truck load. In general, the global AE technique can monitor the whole structure. However, if more specific information is mandatory then a local monitoring AE technique is the best alternative [48]. The flow of traffic is uninterrupted during AE experimental work, which is another advantage of this method [47]. This technique requires more expertise and involves several trials to eliminate noise from the data. Moreover, the AE method has no perfect standard procedure available that can be used as a reference for the application of this method to bridges and all other structures [47]. In a composite steel concrete bridge, the location of damage was investigated using the AE technique [44]. The time difference between the occurrences of AE events was used to locate the damage. Portions of the structure where the damage was suspected were initially detected by global AE monitoring. On the other hand, for more detailed information about the emitted waveforms and to locate the defects, a local AE monitoring was used. AE sensors can only observe changes

in structure conditions and therefore it can be used for monitoring in time but cannot be used for maintenance checks after the deformation has already occurred.

2.2.3 Ground penetration radar

The ground penetration radar (GPR) is the most widely used NDT technique for the assessment of damage in reinforced concrete bridge deck slabs and pavements [37]. It is a fast NDT technique which provides an electromagnetic (EM) wave reflection survey. In addition to concrete bridge deck slabs, it can also be used for other concrete structures such as buildings and tunnels [52]. The GPR method is generally used for the assessment of interior characteristics of subsurface layers, such as delamination, voids, cracks, member thickness, rebar positioning in concrete, zones of high moisture content and deformation [53, 54]. Generally, conducting a GPR test requires a gridding on the surface, and then data are collected on all points of a mesh [37]. The GPR antenna transmits pulses of an EM energy and then a part of the transmitted energy is reversed back to an antenna when it encounters different materials while penetrating through layers.

In real field conditions, the accuracy and reliability of the GPR method were validated on a bridge structure which was later demolished [55]. From the experimental work it was found that the quality of the results was highly dependant on two factors, including the experience of operators and the object under testing. When the bridge was demolished, the results obtained from GPR experiments were compared to the actual state of the bridge. It was found that the difference between the actual and GPR results was 10mm for the concrete on the upper rebar layer. The author further mentioned that the difference was due to a resolution problem. Additionally, the GPR based results for pavement thickness were also compared with the actual results, and it was found that the difference was 9mm. It was mentioned that the main reason behind the difference

was the presence of small concrete cover which overlapped reflections from the bottom of the pavement and rebar [55].

A mobile GPR system was also used for concrete bridge deck slab inspection [56]. To determine the accuracy of the mobile and manual GPR methods, the results of the two methods were compared. The accuracy of mobile GPR for layer thickness was in the range of 5 – 15mm if the velocity of the signal is assumed constant in horizontal way. Additionally, this accuracy level can be easily ameliorated if the variation in the signal velocity is considered. Moreover, the flow of traffic is less interrupted when using mobile GPR. Therefore, the mobile GPR method is more cost effective and easy to handle in comparison to a manual method. This method is comparatively expensive and unable to provide sufficient information about the presence of corrosion [57]. Additionally, it is cumbersome to use GPR for vertical structural members.

2.2.4 Impact echo method

The impact echo (IE) method is one of the oldest techniques that is applied for defects detection. It is also used for measuring thickness of bridge member and other section. It was first used for shaft foundation and concrete pipes evaluation [51]. The use of IE technique was highly emerged in 1980 for concrete member assessment such as concrete crack depth determination and assessment of bond in overlays [58, 59]. For conducting IE experiments a sketch of gridding plan is carried out on the member surface and IE tests are conducted on the required location. A 2ft distance was used between the two points in gridding for IE experiments [60].

IE involved excitation or impactor and a receiver. The most widely used tools for IE testing are the impact hammers. The head of the impact hammer incorporate movable parts which is employed to impact the structural member and a force sensor for determining the amplitude of the impact force. The amplitude of the wave is dependant on the magnitude of force therefore access

to the impact force is very important. Based on the availability of the impact force, it is very easy to use the magnitude of the impact force to simply scaled the data and compare the post processed results at various points.

An impactor is used as an excitation source to generate waves and a receiver is installed on the same side to receive the waves. The obtained signals are further processed to measure the current state of the structure. A large number of interpretation procedure were used by various authors with IE technique for delamination detection [21]. The location in concrete bridge deck slab where there is a delamination will exhibit a shift to a comparatively higher values in the returning frequency. One reason of such a shift in returning frequency is because wave reflections usually occurred at lesser depth [61]. It was found in [61] that the IE technique can accurately detect cracks, large voids and thickness of concrete structural member. The authors observed that the reliability of the method is very high and it is anticipated that the current method will be more effective in damage detection in future [61]. It was observed that IE method is very sensitive to slab dimensions [38]. Additionally, IE methods sensitivity to other parameters such as slab defects, size of slab were investigated. It was postulated that the reflected waves will interrupt the outgoing waves and may largely affect the thickness of the member measurements as well as frequency response of impact echo. The authors recommended that for damage detection and for defects localization in concrete slabs, the IE method should be used in conjunction with ultrasonic method. The use of two different method will increase the overall testing accuracy [38].

2.2.5 Impulse response

The impulse response techniques was initially used for deep foundations [62]. In this method concrete are strikes by the hammer which produce compressive stress waves. The frequency range of theses waves are generated from low strain impact and varies from 0 - 3000 Hz. The frequency

range is dependant on the type of the hammer head such as metal or rubber hammer [37]. The force and velocity are measured by the load cell and receiver, respectively. The data acquisition system record the time based signals and are then transformed to frequency domain using FFT algorithm [37]. The velocity response to impact spectra represents mobility spectrum [21], whereas displacement to impact spectra ratio is termed as flexibility spectrum. Similarly, the ratio of the force at a particular point to the subsequent velocity is known as mechanical impedance. It is the inverse of mobility spectrum. The impulse response method can be used for checking the pile length, cross sectional area and length of the piers. This method was also considered effective for detecting hidden voids under concrete pavements. Moreover, it has the potential for damage detection in concrete tunnels and concrete bridge decks [37].

The idea of an IE method and an impact response method seems slightly similar in general, however, there are greater differences between the two techniques in case of bridge deck testing. The frequency range of IE method is 3 – 40 kHz and therefore this method is based on specific wave generation mode between top and bottom or with in deck. Conversely, frequency range of impact response method is 0-1 kHz which is significantly lower than IE method. Moreover, impulse response technique is based on response of structure in the locality of impact [21]. The impact response method has used for different concrete structure which includes bridge deck, piers, retaining structure, freeways, arch bridge, box beam pre-stress bridge and cooling towers.

It was found that this technique is effective in evaluating concrete bridge decks, piers, pavements, slabs, silos and retaining structures. This method can be used for quick assessment and identification of damage areas that can be used as a primary information for extensive investigations. Therefore, this method is more economical for rapid assessment of structures on site [62]. The piles were also tested using Impulse Response Mash's technique. This method was

used to rapidly determine continuity of concrete in precast piles and also the length of the piles. The data obtained from the impact response method were compared with the designed length and was found similar. This method is extremely helpful for rapid onsite inspection. However, this technique required adequate knowledge about stress wave's signals which is a major obstruction in implementing this method.

2.2.6 Electromagnetic conductivity

This technique is based on the principal of electromagnetism and is used for examination of crack detection, coating thickness, material identification and thickness measurement. It is used for inspection of aircraft skin and for heat exchanger tubes in nuclear industry. It provides information about the material such as geometrical properties of the material, degree of saturation, and electrical information [56]. The damage in concrete can be determined by observing alteration in conductivity [56, 63, 64]. The transmitting coils generate an electromagnetic fields and are supervise when pass through the material. The information about the geometrical and electrical properties of the material can be obtained from the differences between transmitting and receiving fields where both amplitude and phase are different [64].

Gridding technique is used on the structural member surface which helps to take measurements at specific point. To improve the accuracy of the results, the reading should be overlapped for every half meter [64]. The data obtained is recorded and a contour map is plotted. A study about the relationship between electrical characteristic of concrete with concrete mix were conducted [63]. The concrete quality and strength were determined by indirect procedure. The properties such as water cement ratio, cement paste hydration rate and degree of hydration rate were used as a base for conductivity of a mix. The author concluded that this method has the potential for real time structure inspection [63]. This method is low cost, quick and non-contact

and was used for masonry arch bridge investigation. The assessment of bridge includes moisture detection as well moisture movement with respect to time. The method can be applied for identifying inhomogeneity in material. It was found that moisture level, clay content and salinity are the controlling parameters of this method [64].

2.2.7 Electrical resistivity

The electrical resistivity (ER) is used to measure the durability and quality of concrete. The ER of concrete is associated with pore structure, pore size distribution, and porosity of the cement matrix. The hydration of cement paste can control all of the above mentioned characteristics of concrete. The ER of concrete increased with the passage of time. Several other factors that influence the ER include temperature, relative humidity, and ions concentration [65]. The ER of concrete is affected by the temperature, moisture, and quality of concrete. The range of ER varies from 10^1 Ωm to 10^6 Ωm. In general, ER in a cell can be defined as the voltage applied and the resulting current ratio [65]. A minimum of two electrodes are used in this technique in which voltage / current is applied, and then the resulting voltage / current is carefully recorded. For calculating resistivity, a factor which is termed as cell constant is multiplied to the resistance which is the ratio of voltage and current.

In reinforced concrete, various factors such as steel location and depth to surface value highly affect the overall distribution of resistivity of concrete [65]. Additionally, other parameters such as surface layer of concrete also affect the resistivity distributions. The resistivity will be lower if the surface of concrete is wet. However, if the surface is dried then it will be comparatively greater than the normal concrete [21]. It was found that the corrosion rate was higher in areas of low resistivity than areas of high resistivity [65]. Therefore, it was concluded that ER technique can be used for detecting steel corrosion. Moreover, presence of moisture in concrete and its

quantity can be analyzed by using the same technique. The fluctuation in temperature also plays an important role in determining resistivity and therefore temperature correction factors were applied to obtain an accurate result [65]. As such, surface resistivity method can be implemented for chloride ion penetration risk in concrete instead of rapid chloride method. Also, such method can be used to measure the susceptibility of concrete by determining its resistance to chloride penetration, and it was found that surface resistivity technique is more robust and reliable in comparison to rapid chloride ion penetration method [66]. The interpretation of the raw data obtained using this method is more challenging and therefore, it limits its use for large structures [57].

2.2.8 Computer vision

The demand for computer vision techniques is increasing rapidly for the assessment of anomalies in various fields, such as SHM, bio-medical engineering, and myriad others. In recent years, digital images have been used to assess damage in structures. Advancements in computer technology and progresses in image processing algorithms have an important role in the improvement of damage detection methods for structures. Computer vision techniques have a high potential to overcome the deficiencies of the damage detection methods described in the previous subsections. Computer vision-based methods require a camera, a computer, and an algorithm for processing images to detect damage. A high resolution camera and a suitable algorithm can detect thin cracks and other anomalies in the structures. Numerous computer vision techniques have been developed recently for detecting cracks, corrosion, bolt loosening, deflection, and delamination [9-12, 14]. Cha et al [11] used a convolutional neural network (CNN) [67] technique for crack detection and compared the results with Canny and Sobel edge detection methods. The performance of the CNN was tested for various conditions, including blurred images, low light, and bright images. A maximum

accuracy of 97% was recorded. Recently, Gustavo et al used a depth camera with faster- R-CNN for detecting spalling in concrete which obtained a high average precision [68]. This deep learning techniques have not been applied to subsurface damage detection.

2.2.9 Radiography

This method has the potential to detect concrete voids, location of reinforcement in concrete, honeycombing in post-tensioning ducts, and layers of different materials [37, 56, 69]. It is a nuclear non-destructive technique, therefore it requires a special license and safety training for implementation. High electromagnetic radiation energy passes through the object or structural member while on the other side results are recorded in the form of a photographic film [37].

The optimum source of radiation and material absorption of energy largely depends on the materials density and its thickness. Additionally, exposure time can also influence the selection of radiation source. The radiation based photograph represents high density material such as reinforcement bars as light areas. Conversely, the voids which is low density region appear as a dark area [37]. The damaged portion in concrete canal lining was tested using x-ray computed tomography (CT). A sample of concrete was obtained and crack patterns were investigated in the concrete cores obtained from the existing concrete canal. Helical CT scans were used for monitoring purposes. Based on data visualization, it was observed that the dense portion was represented by white segment and the air or void area was represented by a dark colour [70]. This method requires a special certificate or license for safe implementation and therefore it is not the optimum alternative for bridge health monitoring.

2.2.10 Infrared thermography

An infrared thermography (IRT) is an NDT method used for subsurface damage detection. This method monitors variation in temperature on a surface of material using a thermal camera. The

color difference of each pixel in thermal image enables the inspector to differentiate between defected surface and intact surface with appreciable accuracy. The defects include subsurface delamination, inner voids, interior cracks and spalled. The surface can be detected from a maximum distance with high accuracy. IRT is classified into two main categories, i.e. active thermography (AT) and passive thermography (PT). The difference between the two categories is based on the external heating source or excitation. In AT, various types of external heating sources are required. Conversely, no external stimulations are required in case of PT. The basic concept of IRT for NDT is that the unusual temperature which is comparatively different than the surrounding temperature will represent zones of subsurface delamination in concrete. In PT, temperature difference can be seen under natural condition such as solar radiation. This method is applicable because of its ease and fast procedure of collecting data. The PT method has been used for delamination detection in concrete bridge deck. On the other hand, AT used an external heating source to improve thermal image characteristics by increasing differences in temperature values between defected and intact regions [71].

Additionally this method is a global inspection method and can cover a wide area of the structure and reduce the overall inspection time. The results obtained from the IRT provide a reliable information which can be used to either conduct more detailed investigation or maintenance procedure. In case of bridge deck inspection, the thermal images show the region of defects such as delamination in preliminary survey. The survey is then further used for more detailed investigation. The field inspector can focused more effectively on those defected regions for further analysis. The IRT experiments based on two mechanism of heat transfer for concrete exploration. These includes conduction and radiation [37].

In conduction mechanism, the temperature of the surface is determined during heat flow condition. Subsurface defects affect the flow of heat, therefore, this effect shows a difference in thermal image which can be used for detection of interior anomalies. In radiation mechanism, Stefan-Botzmann law can be used to calculate the surface temperature. The radiation such as electro-magnetic are emitted from the surface and therefore the rate of the emitted energy from a unit area can be determined very easily using the aforementioned law. Different types of sensors are used to measure these radiation because these radiation are not visible by naked eye [37].

The IRT equipment consists of three main components. The scanner component is used for IR transmission. The data are recorded and analyzed with a data acquisition component. Similarly, a visual recorder is used to record the scanned data. Thermography has been used for heat loss detection in buildings. The updated version of thermal camera have the potential to detect differences in temperature to a smaller value of 0.1 $^{\circ}$C. IRT can provide warning of damage well an advance, before it leads severe condition. Moreover, it can also be used for providing historic maps of the overall trend of structure [72].

The IRT method is that it is computationally efficient, non-contact and a safe method, and can be used to inspect objects in several sizes. The distance between the target and the camera range can be modified to a desired level by installing appropriate lens. However, there are some limitation of the IRT technique can be used only to detect subsurface damage to a certain depth. The deeper subsurface damage cannot be accurately detected in case of thick concrete section. The good quality thermal cameras are costly, and experienced operators are required. Moreover, the thermal results are largely influenced by surrounding wind speed, emissivity, solar radiation and moisture content [37]. This method was selected for this research because it can detect subsurface damage with only thermal camera and it has the potential to gather large amount of data without

any physical contact. The thermal data obtained can be input to the DINN network to detect zones of subsurface damage which further increases its applicability.

Chapter 3. Methodology

This chapter introduces the overall method for detecting subsurface damage in steel members of a bridge. Deep learning technique and the relevant concepts are used for detection and localization purposes explained in the following subsections.

3.1 Overview of the proposed method

An automated subsurface damage detection process is outlined in the following section. A flow diagram of the infrared thermography and deep learning-based method are presented in Figure 1. To generate a data bank, thermal videos are collected from the famous century-old Arlington Bridge in Winnipeg, Canada. The bridge site is shown in the map in Figure 2. The network in this research was based on the deep inception neural network (DINN) [30, 73].

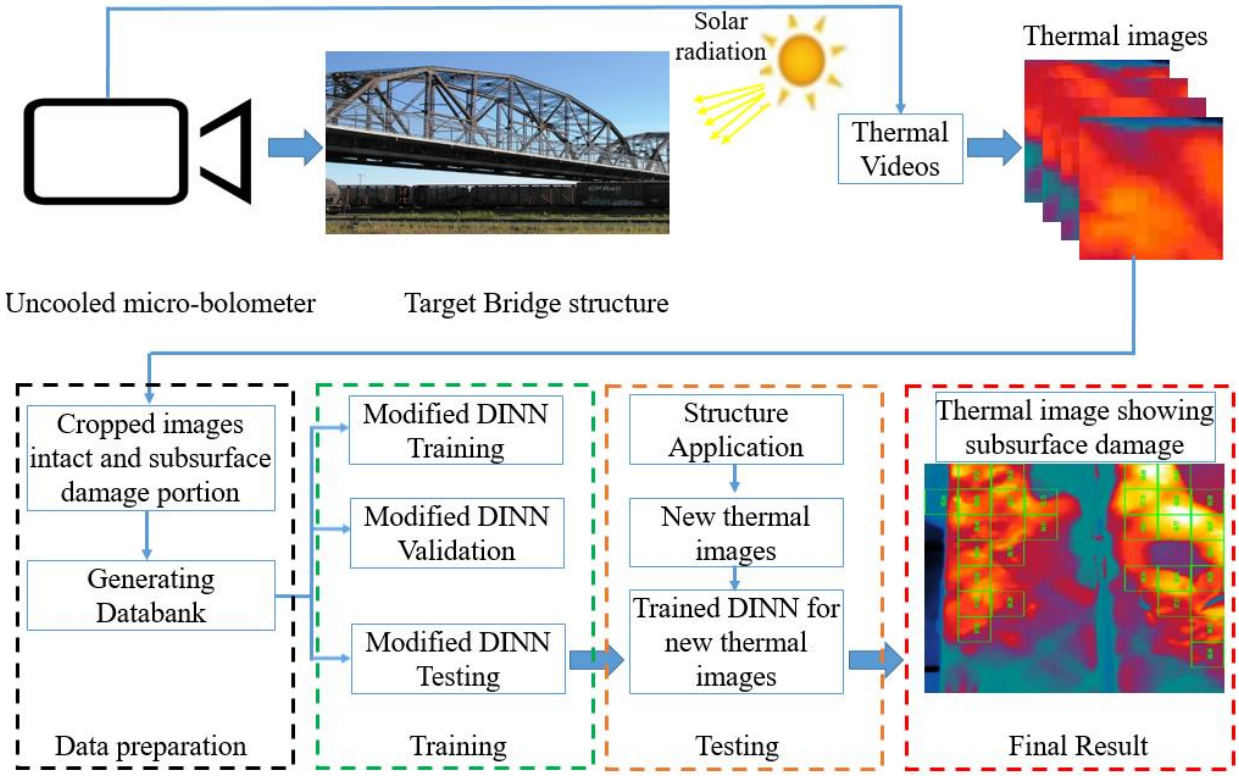


Figure 1: Flowchart of detection of subsurface damage in steel elements of a bridge [73]

A total of 34 thermal videos of steel elements of the bridge were collected. All the videos were converted into thermal images. The thermal images were collected using a Forward Looking Infrared (FLIR) thermal camera, which is equipped with an uncooled micro-bolometer detector. The thermal images obtained from the steel bridge were used for training the network. All thermal images were divided into two categories: thermal images showing single or multiple portions of subsurface damage, and thermal images without subsurface damage. The data were collected over the summer of 2018, from May to August. The thermal images were cropped into sub-images, showing zones of subsurface damage and intact regions. The prepared data were used for training and validating the modified DINN. The detection accuracy of the network was determined by testing new thermal images obtained from the steel bridge. These new thermal images were only applied for testing purposes, and were not used for any training or validation. The trained modified DINN network successfully detected and localized subsurface damage in steel elements of a bridge. Regions of subsurface damage were delineated using green bounding boxes. An Ultrasonic Pulse Velocity (UPV) tests were conducted to generate ground truth for validation purposes. The regions of subsurface damage indicated by the thermal images were validated by the UPV tests.

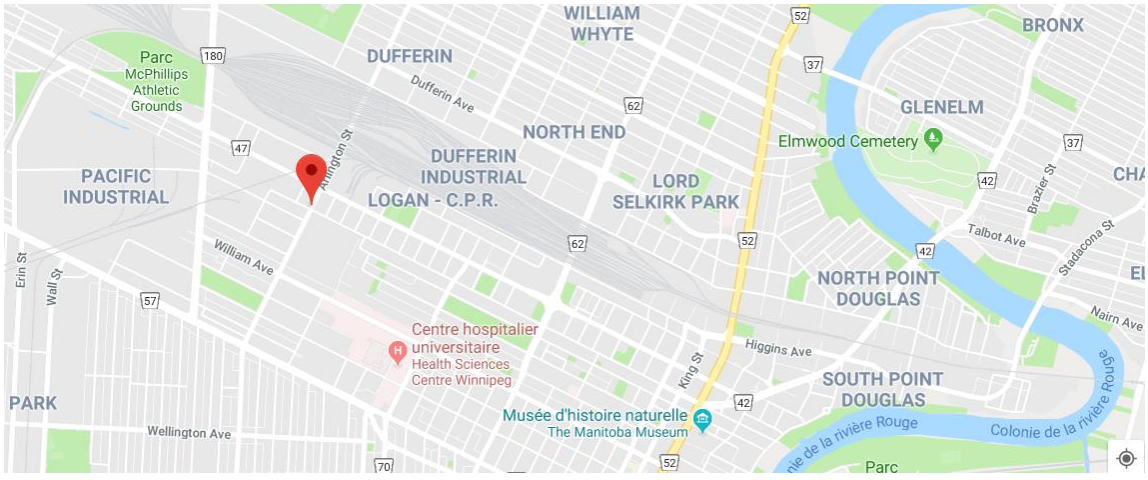


Figure 2: Location of Arlington Steel Through Truss Bridge, Winnipeg, Canada

3.2 Deep inception neural network

This section provides a detailed explanation of the overall architecture of the neural network, including each layer used in the network, and the conceptual background of each layer. A general inception neural network includes convolution layers, pooling layers, mixed layers, a fully connected layer and a softmax layer. The mixed layers contains both convolution layers and pooling layers, arranged in several different orders. The detail of each layer is explained in the following sections. In this research, the original deep inception network was modified and incorporated with transfer learning techniques. A new fully connected layer and a softmax layer were added to the transfer learning section to minimize the overall training time of the network.

Figure 3 shows a thermal image being fed to the DINN architecture [73] and passing through several layers, to obtain a compressed feature map. In general, the input to the current layer is the output obtained from the previous layer, and the input to the initial layer is the resized image, which has a dimension of 299 × 299 × 3. The detailed information about the input layer and output layer sizes are listed in the appendix, in Table 4 and Table 5, respectively. The DINN network generates a compressed feature map for each thermal image. This compressed feature map gathers the information that is necessary for the classification of the thermal image. The compressed feature map is transferred to the fully connected layer, and then to the adjacent softmax for further classification. In the training section, the ground truth for the subsurface damage and intact regions is prepared. The outcome of the network is compared with the ground truth, and the cross entropy is determined. The optimizer layer is used to optimize the differences between the values.

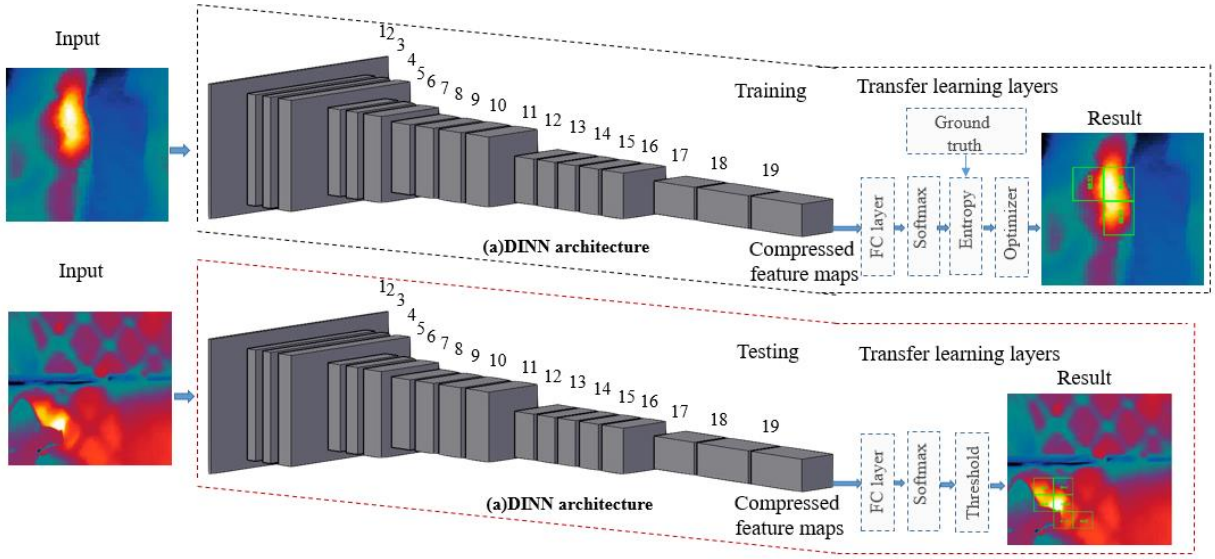


Figure 3: Overall architecture for training and testing [73]

The main reason for using a deep inception neural network in this research is that it reduces computational cost. It was found that the overall computational cost of the inception network is comparatively lesser than that of other networks, such as AlexNet [78]. For instance, the inception architecture of GoogleNet used 5 million parameters, while AlexNet employed 60 million parameters, or twelve times more than the inception architecture [30]. Additionally different sizes of convolution are used to collect the detail information at different scales. These include 1×1 convolution, 3×3 convolution and 5×5 convolution. Moreover, the idea of using 1×1 convolution layer significantly reduced the computation cost. The transfer learning layers were used to minimize the computation cost and to obtain a greater accuracy with limited available training data. The details are explained in the following sections.

3.2.1 Convolution layers

Convolution layers are the important layers of the network and it considerably increase the computation cost [17]. The convolution layers parameters are based on filter which has learning

abilities and are known as kernels. Each filter covered a portion of the image and the specific area is known as receptive field. The depth of the filter is equal to the depth of the input. The main function of the convolution layer is to develop the feature maps i.e. to generate the output of the input by conducting three consecutive operations. The three operations of convolution layer are expressed in Figure 4. A dot product is first executed between a 3×3 subarray and a 3×3 convolution filter or kernel. The size of the convolution kernel determines the receptive field which must match to the size of the subarray. Secondly the dot product of all the values are summed and finally a bias of 1 is added to the summed value to obtain the feature map of the original subarray.

In initial stage the value of weights and biases are generated randomly and these random values are further tuned in the training process. Once one subarray is completed, then the subarray moves towards right and this motion of the subarray is termed as stride. When the subarray move one pixel at a time then the stride is equal to one. And when the stride move two pixels at a time then the stride is equal to two. When the subarray reached to the far right of the input then the subarray moves one stride down and begin again from the left. In Figure 4 the depth of the input is equal to 1, but the depth of the input volume can be more than 1. To ensure the dot product between the subarrays and kernels, the depth of convolution kernel must be consistent with the depth of the input volume. In addition, the number of kernels varies based on needs and determines the depth of the output volume i.e. the feature maps depth is match to the kernels number. The relationship between input and output sizes are expressed in Equation (2).

$$W_2 = (W_1 - F)/S + 1, \tag{1}$$

where W_2 is width of output (feature map), W_1 is width of input, F is kernel width and S is stride. The output volume size is dependant on the stride, depth and zero padding.

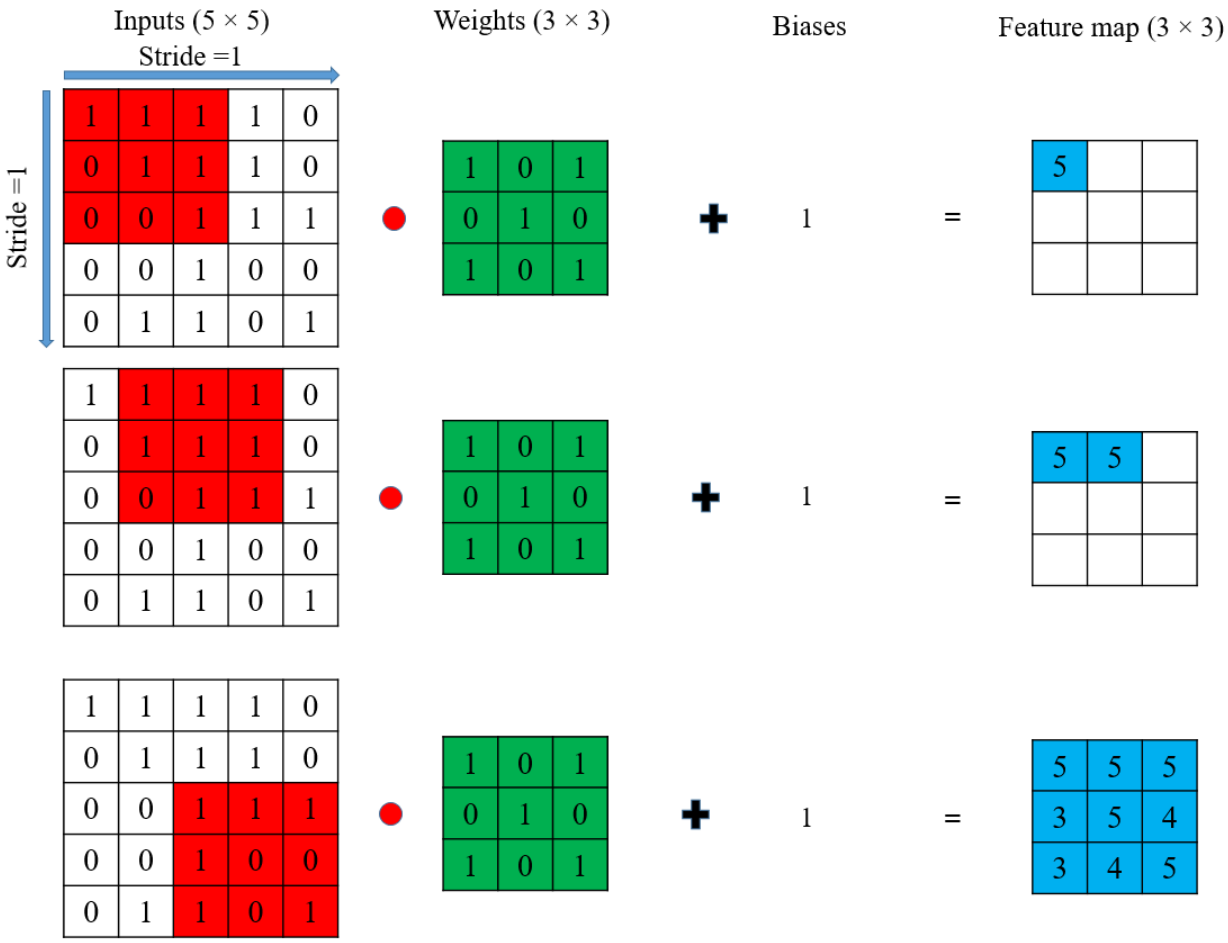


Figure 4: Convolution layer

In most practical situations, zero-paddings are provided around the border of the original input volume to enhance the edges of feature maps. Zero-padding size is a hyper parameter which manipulate the spatial size of the output. The application of convolution layers decrease the output size. Therefore, it is essential to maintain more information from the input volume for tiny level feature extraction from the edges.

$$W_2 = (W_1 - F + 2P)/S + 1, \quad (2)$$

where W_2 is width of output (feature map), W_1 is width of input, F is width of kernel, S is stride and P is padding. An example of zero-padding is shown in Figure 5, where a zero-padding is provided around a 5×5 input array with a stride equal to 2. The width of output volume of a feature can be determined using the Equation (2).

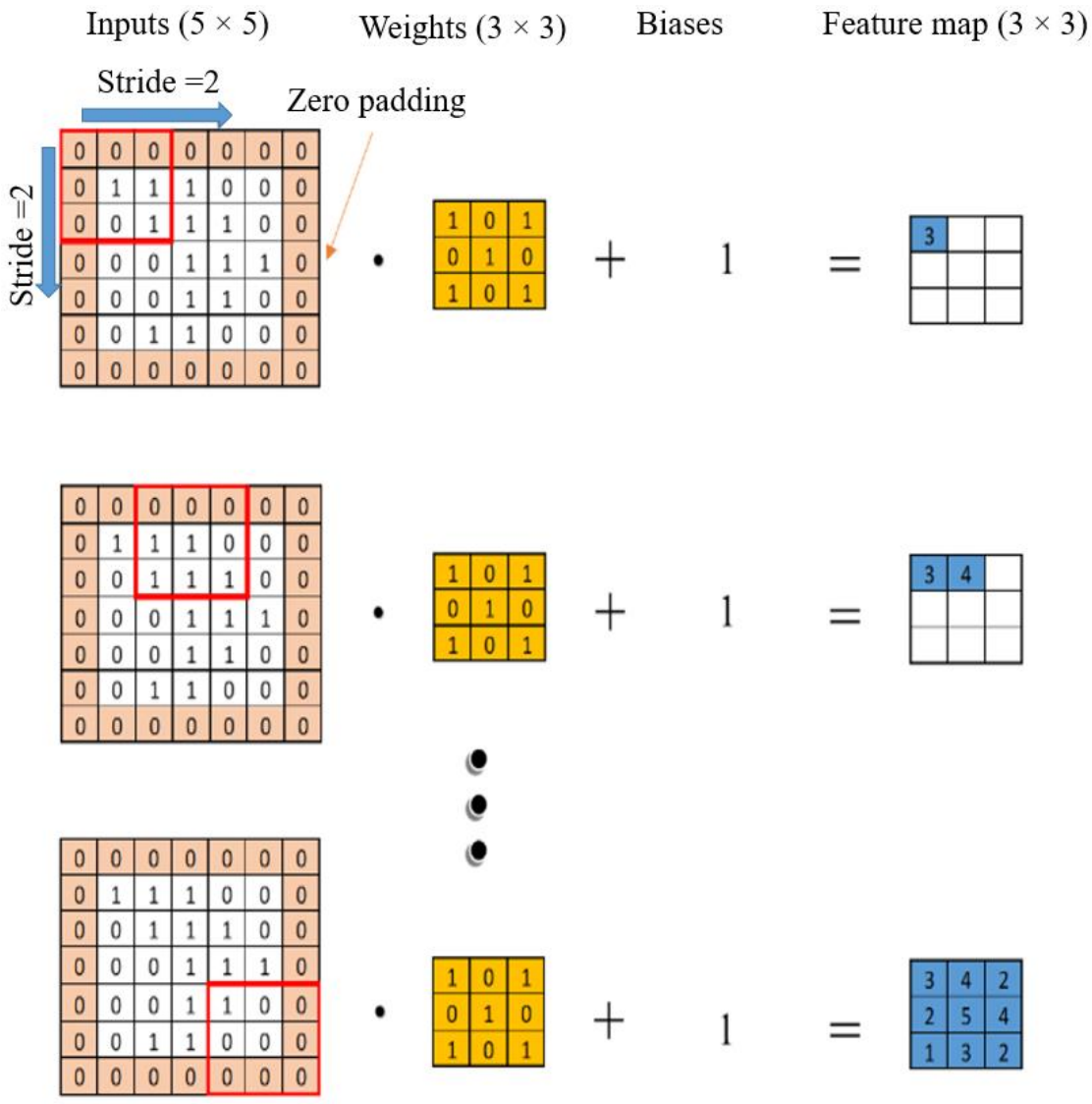


Figure 5: Convolution layer with zero padding

3.2.2 Pooling layers

Pooling layers are provided after convolution layer and are used to compress a feature maps along the direction of width and height by down sampling [79]. The major features are retained and the number of parameters as well as the computation costs are reduced significantly. The process of pooling can be conducted by two different ways i.e. max pooling and mean pooling. Similar to convolution, there are parameters of size and stride, however, there is no parameter of depth in pooling layer. The pooling itself is an operation without any weights or biases, it can be operated on each layer of input volumes independently. The max pooling process is explained in Figure 6. Moreover, zero-padding can be applied when pooling and the relationships between output sizes and input sizes are same.

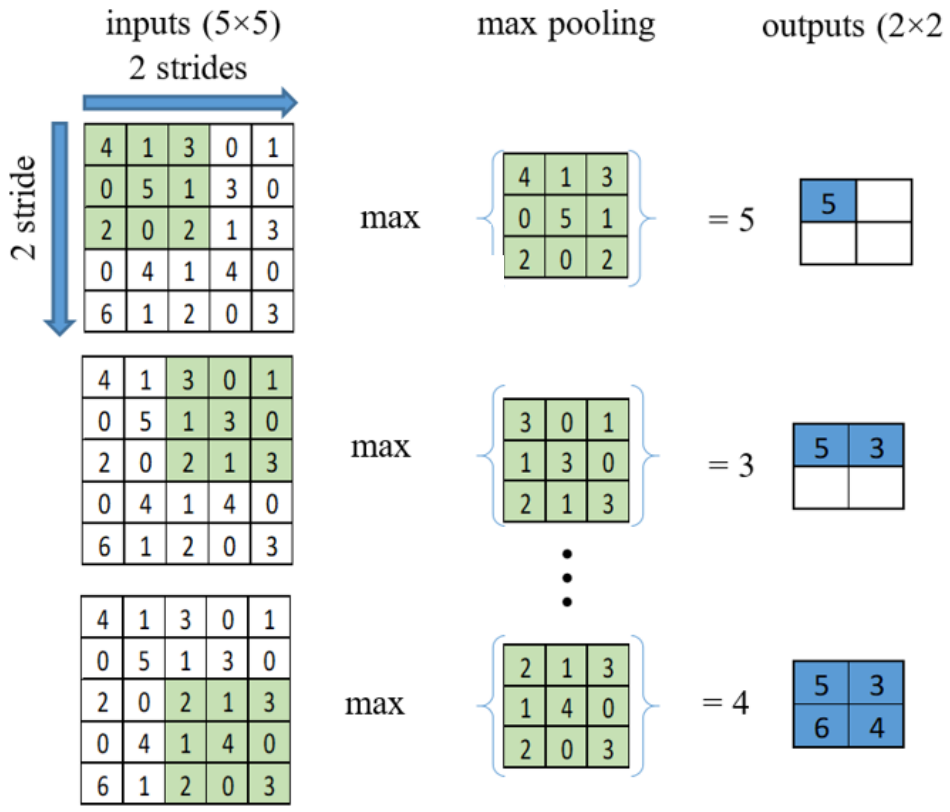
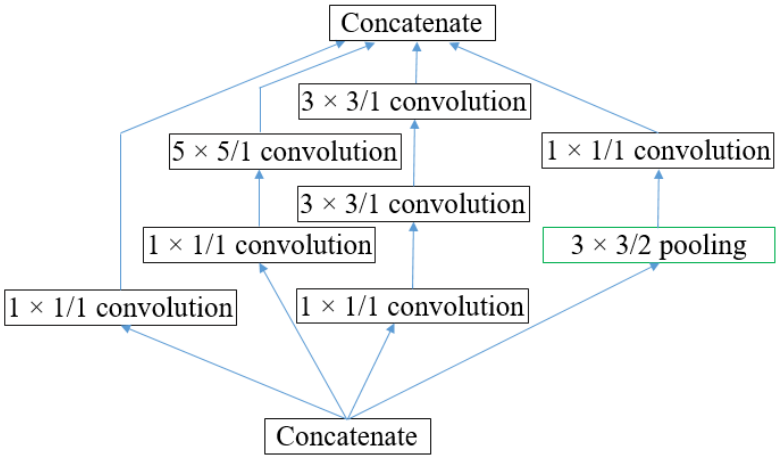


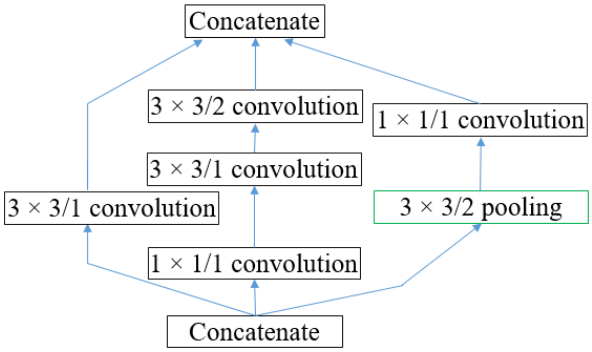
Figure 6: Pooling layer

3.2.3 Mixed layers

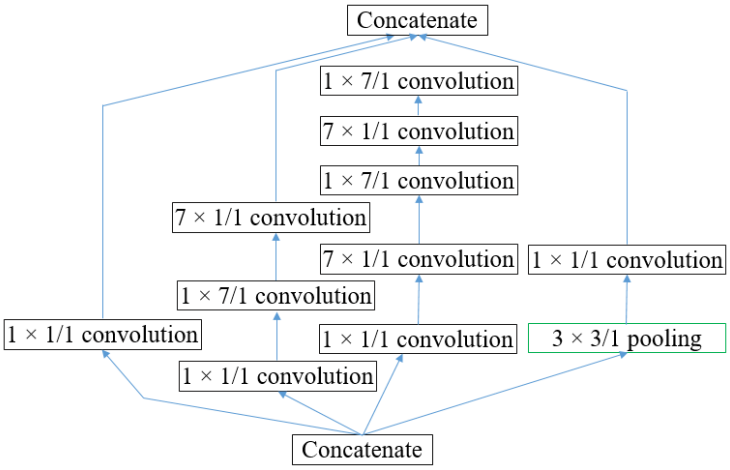
Mixed layers include both convolution and pooling layers together [30, 80]. The convolution layers and pooling layers are provided in different arrangement both in series or parallel. The larger convolutions are converted into smaller one and the symmetric convolutions are converted into asymmetric ones. There are five different types of mixed layers based on their arrangement in the network. And the details of these five different mixed layers are presented in Figure 7 [73]. These layers include 1 × 1, 3 × 3, 1 × 3, 3 × 1, 5 × 5, 1 × 7, 7 × 1 convolution, and pooling layers with a stride equal 1 or 2.



Mixed layer-a



Mixed layer-b



Mixed layer-c

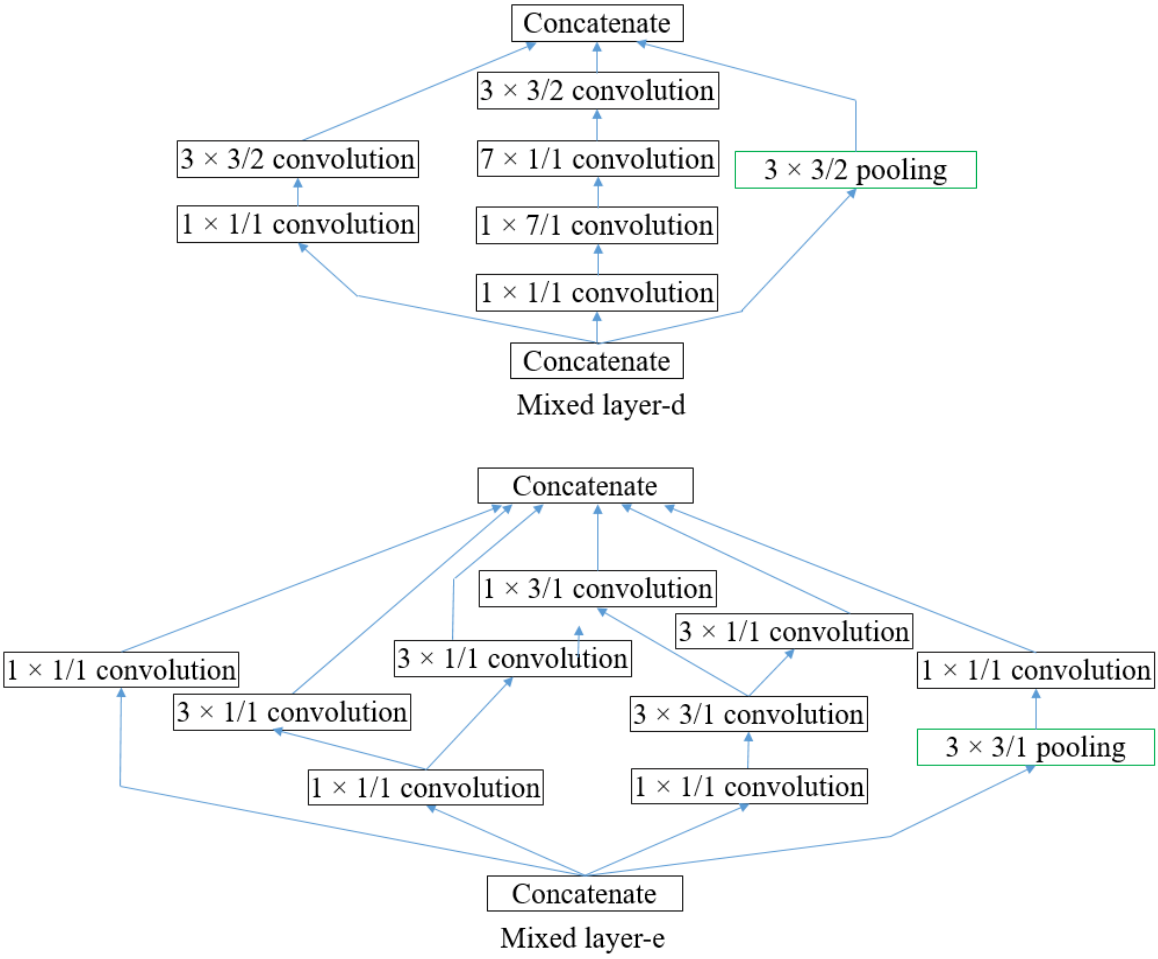
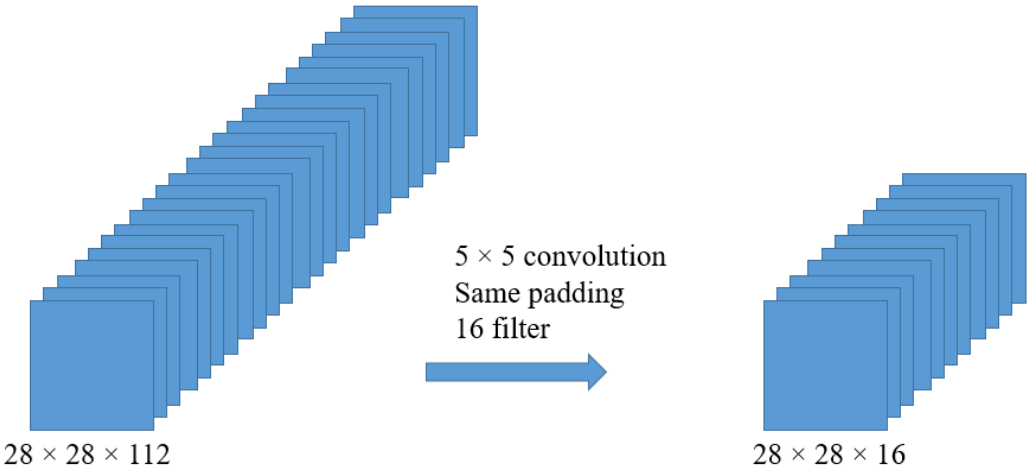


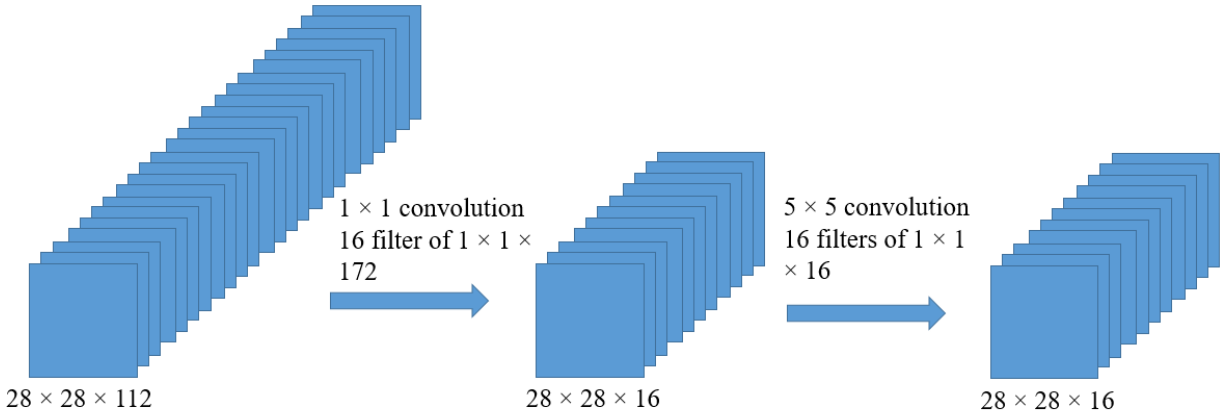
Figure 7: Different arrangements of mixed layers [73]

The main purpose of arranging different layers are to collect extensive information and to reduce the computation cost. For instance in mixed layer-a and mixed layer-b in Figure 7, 1 × 1 convolution is provided before 3 × 3 convolution to minimize the computation cost. Similarly in mixed layer-c 1 × 1 convolution layer is provided before 1 × 7 convolution and at the same time 1 × 7 and its asymmetric 7 × 1 are provided. In case of mixed layer-d and mixed layer-e 1 × 1 layer is also introduced as shown in Figure 7. In general, a pooling layer is used to reduce the height and width of the layer. On the other hand to minimize the depth of the channel a 1 × 1 convolution is provided.



$= 28 \times 28 \times 16 \times 5 \times 5 \times 112 = 35{,}123{,}200$

a) In the absence of 1 × 1 convolution



$= 28 \times 28 \times 16 \times 1 \times 1 \times 112 + 28 \times 28 \times 16 \times 5 \times 16 = 6{,}422{,}528$

b) In the presence of 1 × 1 convolution

Figure 8: Computational cost comparison with and without 1 × 1 convolution

To understand the effect of 1 × 1 convolution on computational cost. The computational cost of the two conditions having same input and output values with the absence and presence of

1 × 1 convolution are presented in Figure 8 [73]. The input and output are 28 × 28 × 112 and 28 × 28 × 16 in both cases. It was found that the computational cost without 1 × 1 convolution is 35 million as shown in Figure 8 (a) which is approximately 6 times more than that of Figure 8 (b) where a 1 × 1 convolution is applied. Moreover, this difference in computational cost due to the use of 1 × 1 convolution will increased dramatically when the range between the input and output increased.

3.3 Transfer learning section

In section 3.2 it is discussed that each thermal image fed to the network will generate a compressed feature map (CFM) and all the CFM are saved and transferred to the next layer for further classification. These CFM contain the summary of the training feature maps and are used for further recognition of the image [81]. The transfer learning section is presented in Figure 3. The major function is to optimize the weights and biases and to develop better prediction results. The accuracy of each iterations are indicated in this section. To perform the major function, an input of ground truth and compressed feature map is necessary. The ground truth is the true label of the image, whereas the compressed feature is obtained from DINN network. A CFM is transferred to a fully connected layer and then followed by the softmax layer. A softmax layer receives the operated values and convert it into the probabilities of which the image is belong to the class. Then the loss between the predicted probabilities and the ground truth is determined by cross entropy. Similarly, an optimizer collects all the values such as ground truth, bottleneck, weights, biases, and cross entropy to minimize the loss by updating and tuning the weights and variables.

3.3.1 Compressed feature maps

A compressed feature map is an array from DINN network which contains the feature of the input image. All the CFM are saved in a file and then transferred to transfer learning section. A

compressed feature maps is an array of 2048 values, and it is not human decodable. A sample of one array is reshaped to a new size of 64 × 32 and visualize in Figure 9.

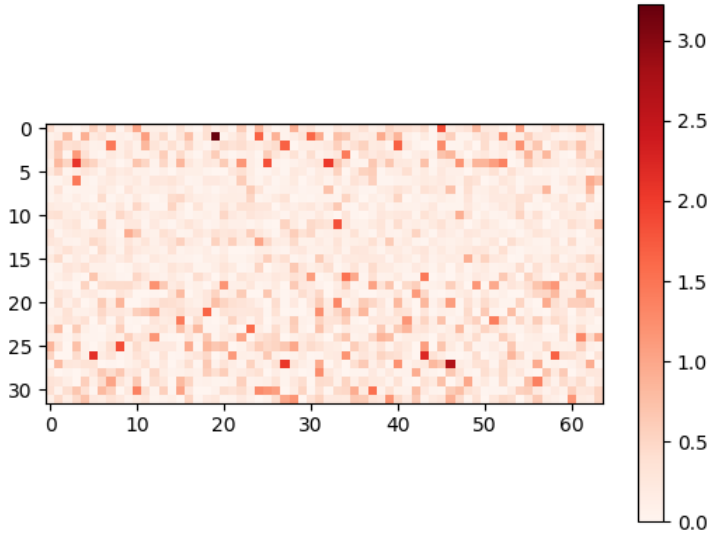


Figure 9: Compressed feature maps

3.3.2 Fully connected layer

The fully connected layer represents a series of operations operated on the CFM [16, 80]. The function of fully connected layer is to extract the features of CFM by connecting each node in fully connected layer with all the nodes in CFM. Such as, in Figure 10, both y_1 and y_2 in the outputs array have relationship with the whole input CFM array. The entire shapes and operations are shown below in Figure 10. Weighs and biases in final operations are the most important variables which are shared by all the CFM and are tuned to develop a better prediction. The shape of weights, biases and outputs are determined by the number of classes, the network is requested to recognize. In this research, the network is required to distinguish two classes, subsurface damage and intact

portion. Therefore, the shape of weights is 2048 × 2, the shape of biases is 1 × 2 and the shape of outputs is 1 × 2.

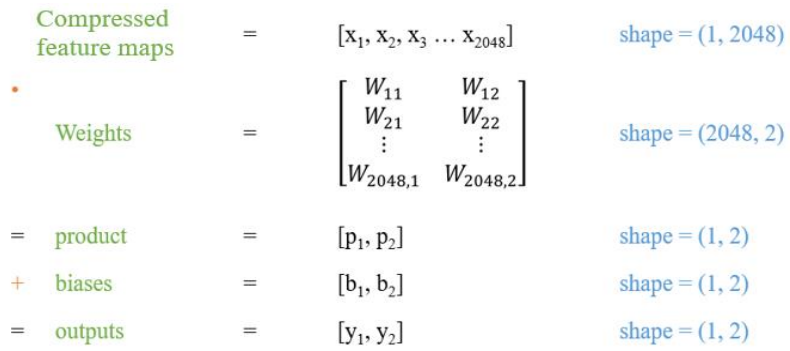


Figure 10: Fully connected layer final operations

3.3.3 Softmax layer

The Softmax layer maps the outputs from the fully connected layer into a series of probabilities which are between 0 and 1 [16, 80]. The probabilities are the prediction of the network that distinguish among all the classifications. The softmax layer is expressed in Equation (3) and an example is shown in Figure 11.

$$P_i = \frac{e^{y_i}}{\sum_i^K e^{y_i}}, \tag{3}$$

where P_i is probability of belonging to i classification, K is number of classifications, e is Euler's number, and y_i is value from fully connected layer

$$\text{input} = [y_1, y_2] = [2, -1]$$

$$P_1 = \frac{e^{y_1}}{e^{y_1}+e^{y_2}} = \frac{e^{2}}{e^{2}+e^{-1}} = 0.95$$

$$P_2 = \frac{e^{y_2}}{e^{y_1}+e^{y_2}} = \frac{e^{-1}}{e^{2}+e^{-1}} = 0.05$$

$$\text{prediction} = [P_1, P_2] = [0.95, 0.05]$$

Figure 11: Softmax Layer

3.3.4 Ground truth

Ground truth is an array which is the true label of the image. The shape of ground truth is similar to the prediction of softmax, therefore ground truth and prediction are compared to calculate the loss which is known as cross entropy. Based on cross entropy results the network tune the weights and biases to minimize the cross entropy. Ground truth contains values of either 0 or 1. The value 1 represents the classification of the image. For example, a ground truth of [1, 0] means the image is belonging to class 1 which means the thermal image shows subsurface damage and [0, 1] means class 2 which means the thermal image shows intact portion.

3.3.5 Cross entropy

Cross entropy [80] is the most widely used method for calculating the loss between the ground truth and the prediction obtained from the neural network. One image has one cross entropy. The cross entropy is expressed in Equation (4) [73].

$$C = -\sum_i^K y_i \log(P_i)\,, \tag{4}$$

where C is cross entropy, K is number of classifications, y_i is ground truth value and P_i is prediction value. There are two important features of cross entropy. First, the output is always positive. The probability P_i is a value between 0 and 1 which makes log (P_i) negative, and the negative sign at the beginning converts the sum to positive. Second, the cross entropy shrinks if the difference between the ground truth and prediction decreases. Item $y_i$log (P_i) is 0 when y_i is 0, which does not influence the result; but if y_i is 1, then a larger P_i makes the sum smaller which means the loss is reduced. An example is shown in Figure 12 for more explanation.

$$\text{ground truth} = [y_1, y_2] = [1, 0]$$
$$\text{prediction} = [P_1, P_2] = [0.95, 0.05]$$

$$C = -\sum_{i}^{K} y_i \log(P_i)$$
$$= -[y_1 \log(P_1) + y_2 \log(P_2)]$$
$$= -[1 \times \log(0.95) + 0 \times \log(0.05)]$$
$$= 0.0223$$

Figure 12: Cross entropy explanation

3.3.6 Optimizer

A function is required to show the optimum relationship between the CFM and the ground truth [80]. An operation called optimizer is used to update the weights and biases to get the best function. The model used for this function is presented below in Equation (5) [73].

$$f(P) = Wx + b, \quad (5)$$

where P is prediction/probability, x is an array of compressed feature maps, W is a weight and b is a bias. In other words the best function is the minimum cross entropy. However, when the amount of training images changes, the mean value of all cross entropies is used instead of sum of cross entropies. In mathematic form it is expressed in Equation (6) [73].

$$C_{mean} = Mean(\sum_{j}^{N} C_j), \tag{6}$$

where C_{mean} is mean cross entropy, j is image number, N is number of images, C_j is cross entropy of image j. Based on the above mentioned explanation the mean cross entropy is a function that varies with the weights and biases and those weights and biases are shared by all the images. In addition, the initial inputs of CFM and ground truths are constants, thus the function varies with weights and biases only and are expressed in Equations (7) and (8) [73]:

$$C_{mean} = Mean(\sum_{j}^{N} C_j) = C(W, b), \tag{7}$$

$$C_{mean} = C\left(W_{11},\ W_{12},\ \ldots W_{1,2048},\ b_1,\ W_{21},\ W_{22},\ \ldots\ W_{2,2048},\ b_2\right), \tag{8}$$

where W and b are arrays but W_i and b_i are numbers. Therefore, two formats are used to represent mean cross entropy. In our network, an optimizer called gradient descent optimizer is used to minimize mean cross entropy. The processes to find the optimum weights and biases is represented in mathematic form in Equations (9) and (10):

$$W^*,\ \ b^* = \arg \min_{W,b} C_{mean} = \arg \min_{W,b} C(W,b), \tag{9}$$

$$W^*,\ b^* = \arg \min_{W,b} C\left(W_{11}, W_{12}, \ldots W_{1,2048}, b_1, W_{21}, W_{22}, \ldots W_{2,2048}, b_2\right), \tag{10}$$

The gradient descent optimizer process is presented in Figure 13 [73]. First, a random value W^0 and b^0 (or $W_{11}^0, W_{12}^0, \ldots W_{1,2048}^0, b_1^0, W_{21}^0, W_{22}^0, \ldots W_{2,2048}^0, b_2^0$) are picked for initial weights and biases. Then it compute the partial differential for each variable and tune the variable along the gradient. And the direction is downward because it is anticipated that the minimum mean cross entropy. The step it moves is called learning rate, which can be denoted as ŋ. Sudden variation in the weights can obstruct to minimize the loss. It is possible that such sudden large steps may miss the optimal value. To avoid missing optimal value a variable called learning rate is used for this purpose. On the other hand, very small learning rate may leads to local minimum. Therefore, an appropriate learning rate should be selected to obtain the optimal weights and biases. Thus, a group of new variables W^1 and b^1 is achieved. And W^1, b^1 are indicated in Equation (11) and Equation (13), and it is expressed in number format as shown in Equation (12).

$$W^1 = W^0 - ŋ \frac{\partial C_{mean}}{\partial W} | W^0, b^0 \text{ and } b^1 = b^0 - ŋ \frac{\partial C_{mean}}{\partial b} | W^0, b^0, \tag{11}$$

$$W_{11}^1 = W_{11}^0 - ŋ \frac{\partial C_{mean}}{\partial W_{11}} | W_{11}^0, W_{12}^0, \ldots W_{1,2048}^0, b_1^0, W_{21}^0, W_{22}^0, \ldots W_{2,2048}^0, b_2^0,$$

$$W_{12}^1 = W_{12}^0 - ŋ \frac{\partial C_{mean}}{\partial W_{12}} | W_{11}^0, W_{12}^0, \ldots W_{1,2048}^0, b_1^0, W_{21}^0, W_{22}^0, \ldots W_{2,2048}^0, b_2^0, \tag{12}$$

$$\dots W_{1,2048}^{1} = W_{1,2048}^{0} - \eta \frac{\partial C_{\text{mean}}}{\partial W_{1,2048}} | W_{11}^{0}, W_{12}^{0}, \dots W_{1,2048}^{0}, b_{1}^{0}, W_{21}^{0}, W_{22}^{0}, \dots W_{2,2048}^{0}, b_{2}^{0},$$

$$b_{1}^{1} = b_{1}^{0} - \eta \frac{\partial C_{\text{mean}}}{\partial b_{1}} | W_{11}^{0}, W_{12}^{0}, \dots W_{1,2048}^{0}, b_{1}^{0}, W_{21}^{0}, W_{22}^{0}, \dots W_{2,2048}^{0}, b_{2}^{0}, \quad (13)$$

All the weights and biases are updated during each iteration or training step. In ideal case, the minimum mean cross entropy is obtained when all the partial differentials for all weights and biases reach zero. But in reality, it is almost impossible to happen, thus a certain number of training steps is needed to stop the training loops and the final partial differentials are very close to zero. Theoretically, there is a concern about if the mean cross entropy is local minimum or global minimum. However, the probability increases dramatically that the local minimums are close enough to the global minimum with the increase of network size and insisting on global minimum may result in overfitting [82]. The training network in this research is based on DINN and transfer learning layer, thus it is large enough to neglect the error between local and global minimum.

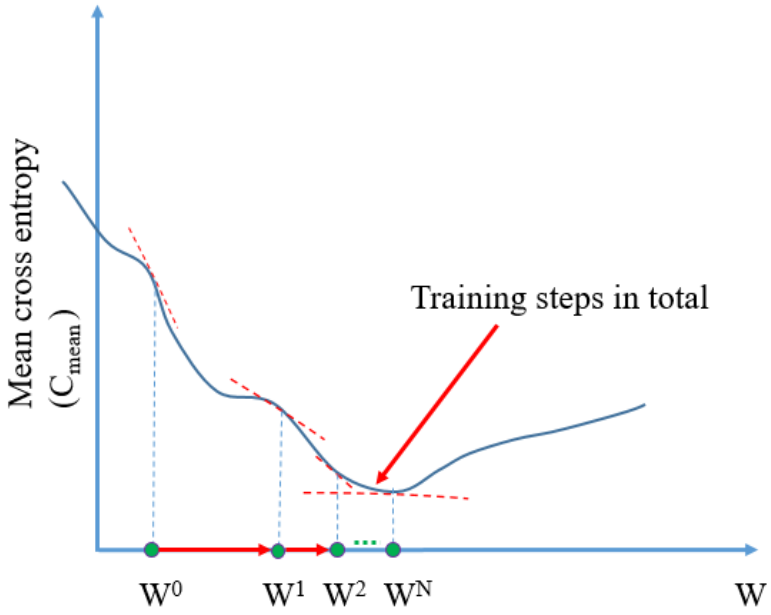


Figure 13: Gradient descent [73]

Chapter 4. Infrared thermography and data collection

This chapter provides detailed information about infrared thermography and other relevant concepts. The procedure followed for data collection is also discussed.

4.1 Infrared thermography background

IRT is a technique which observes infrared thermal energy emitted from different materials and measures its temperature value [71]. It displays a thermal image showing the temperature distribution of the materials. All those materials emit infrared radiant energy when their temperature is higher than absolute zero. The radiant energy emitted by the surface is a function of its emissivity and its temperature. An emissivity can be defined as the material's capability to emit radiant energy in comparison to a black body. The values of emissivity vary between 0 and 1. The typical emissivity values of materials are available in thermal cameras and can be adjusted accordingly. In recent years, considerable improvements and developments have been made in infrared thermography. The IR detectors in the thermal camera sense the infrared energy emitted by the materials, and have the potential to measure minor variations in temperature. There are two different IRT approaches, classified based on the type of heat source: active IRT and passive IRT [83]. These types of IRT are explained in detail below.

Active thermography uses a heating source to generate a temporary heat flow during inspection [84, 85]. The material is heated by an external or internal source. Internal excitation includes electrical, ultrasound, induction and optical techniques, whereas external excitation includes thermal or optical methods. Generally, a flash lamp is used as an external excitation, as shown in Figure 14, which is comparatively economical and easy to use. The use of external excitation helps to produce heat on the surface of the material, which can facilitate the localization

of internal defects. The surface of the internal defect heats up sooner than the surrounding region because the void or vacuum obstructs the conductive heat transfer.

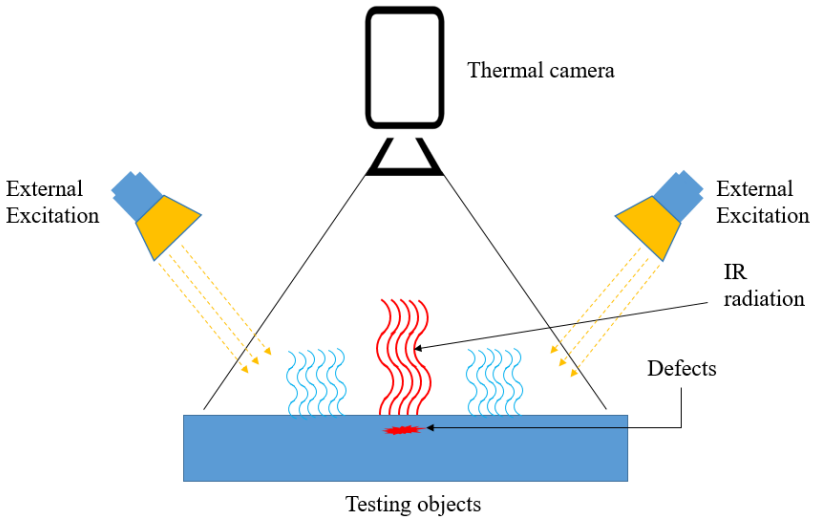


Figure 14: Active infrared thermography

Conversely, passive thermography is dependant on solar heating for excitation [84]. This technique is more economical but is highly susceptible to weather conditions. Generally, the temperature increases due to solar heating during the day. The subsurface damage in a section obstructs the flow of heat during the day, which causes the temperature of the surface above the damaged region to increase more rapidly than the surface above the intact regions. Using a thermal camera, the surface above the interior damage shows a greater temperature than the sound surface, as shown in Figure 15. On the other hand, after sunset, the surface above the subsurface defects is cooler than the surfaces above the intact region. During the night-time, the regions above the interior defects are indicated as cold spots.

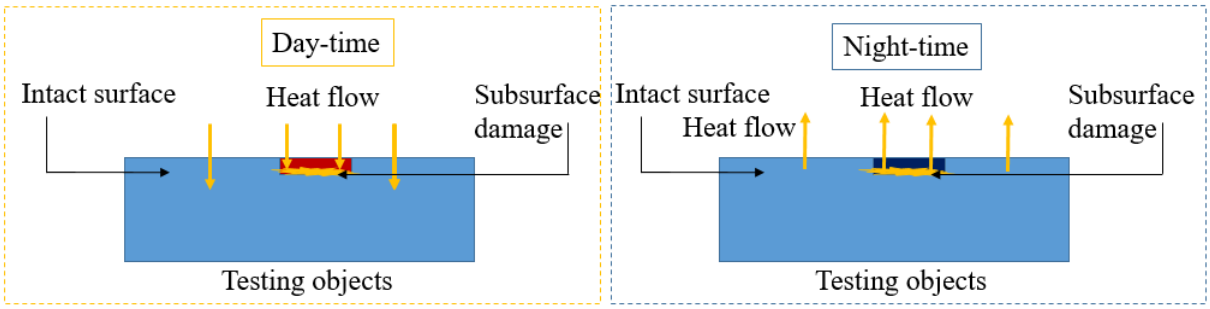


Figure 15: Passive infrared thermography

In passive thermography, it is important to choose the best time window for collecting data to localize subsurface defects using a thermal camera. Therefore, it is extremely important to investigate the best time for data collection. To date, numerous experiments have been conducted for concrete delamination detection. It was found that a well defined and clear thermal image of concrete can be obtained about 40 minutes after sunrise. The intact zones and interior damaged zones can be easily distinguished [86]. Another report provides a more detailed explanation about concrete subsurface damage. It was found that the best time for collecting thermal data to detect subsurface damage depends on the depth of delamination. They conducted experiments on concrete samples with delamination at different depths below the surface. The delamination was artificially induced in the testing sample. The delamination in the tested samples was located at depths of one inch, two inches, three inches and five inches below the surface. They found that the best data collection time is five hours and forty minutes after sunrise for shallow delamination, six hours and twenty minutes after sunrise for delamination at two inches below the surface, seven hours and forty-one minutes after sunrise for delamination at three inches below the surface, and nine hours after sunrise for delamination at five inches below the surface [87].

More detailed information about concrete delamination is provided in ASTM D4788 [88]. It was found that the temperature difference should be 0.5 °C to carefully differentiate between the

delaminated and intact zones. They recommend to collect thermal data both during the day and the night. However, there are no appropriate guidelines about the best time for data collection. Moreover, the standard also specifies that the direct radiation from the sun on the surface of concrete is helpful to generate a temperature difference of 0.5 °C. Based on the temperature difference, it is easy to differentiate between the sound concrete and the adjacent delaminated concrete areas [88]. Despite all these experiments on concrete, there are numerous disagreements among various researchers. In this research we focus on steel subsurface damage in a bridge structure. To find the optimum time for data collection, a total of 34 thermal videos and images were collected in the summer of 2018 on the Arlington Bridge. Based on the experimental work, it was found that the detection of subsurface damage in steel elements of a bridge highly depends on the orientation of a bridge with respect to the sun radiation. In the Arlington Bridge study, one side of the bridge was investigated in the morning approximately one hour after sunrise, while data was collected for the opposite side of the bridge about one to two hours before sunset. The best time window for data collection was during sunlight hours. Moreover, the quality of the thermal images reduced significantly on cloudy and windy conditions.

4.2 Electromagnetic waves

Electromagnetic waves are the synchronized vibrations of a magnetic field and an electric field, which travel at the speed of light [89]. These waves carry radiant energy and can travel through space. The frequency and wavelength characterizing different types of electromagnetic waves in the electromagnetic spectrum are presented in Figure 16 [90].

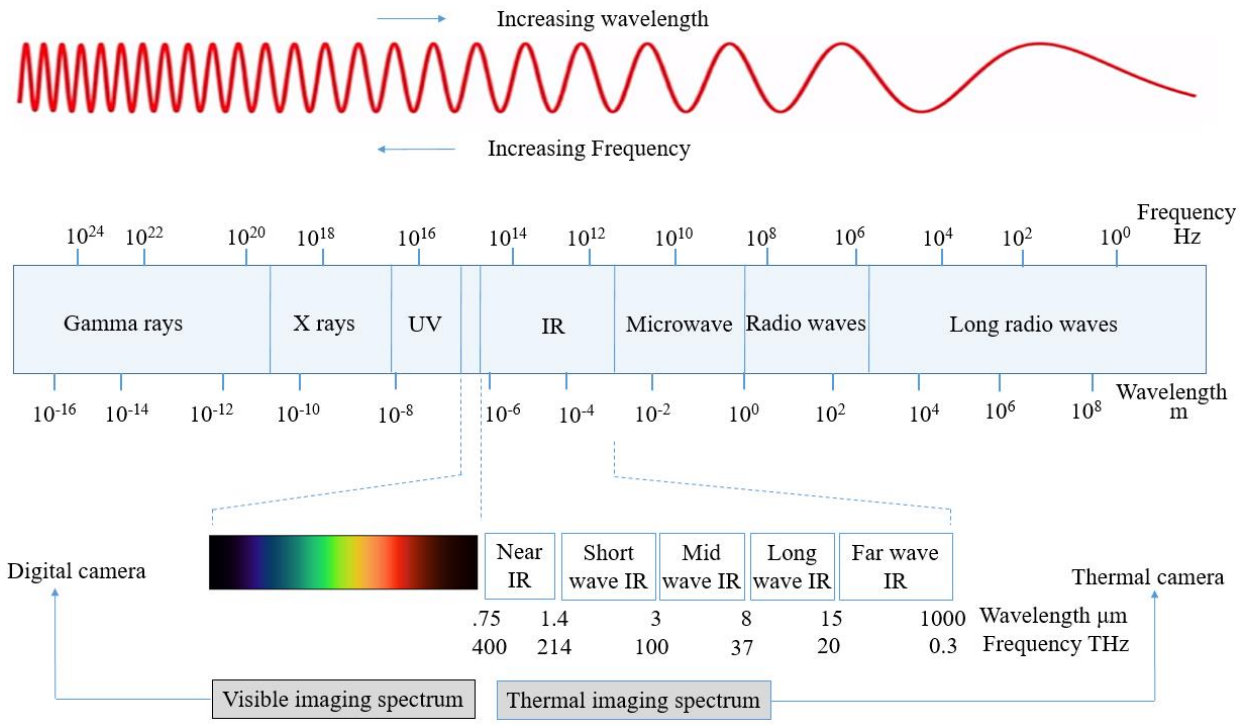


Figure 16: Electromagnetic spectrum [90]

4.3 Data collection

The thermal image data used in this study were gathered from the old Arlington Bridge using a Forward Looking Infrared Thermal (FLIR T650SC) camera [91]. The camera has excellent features for collecting both visible and thermal images. The camera has an uncooled micro-bolometer (uncooled vanadium oxide) and can capture a thermal image of 640 × 480 pixels. The detail characteristics of the thermal camera are presented in Table 1 [73]. A total of 34 thermal videos of steel elements of a bridge were collected. All the videos were converted into thermal images to develop a data bank. All thermal images were captured at a distance of 3ft to 13ft from the steel members of the bridge. The camera has the potential to adjust the distance between the target position and the camera. Additionally different lenses can be used to cover small and large areas.

In this research, two different lenses, T197909 and T197524, with focal lengths of 24.6mm and 41.3mm, respectively, were used. The camera can also produce the multi-spectral dynamic images (MSX), which combine both digital and thermal images at the same time, using a fusion algorithm that incorporates visible spectrum details from the digital image on top of a thermal image. The obtained MSX image shows both the thermal information of the material surface and the visible image, which provides the sharper edges of the target element. Therefore, surface corrosions, stains and discoloration can be distinguished. Moreover, various environmental conditions, such as temperature, humidity, wind speed, material emissivity and myriad others can also be input into the camera. The training and validation processes are explained in the next chapter.

Table 1: FLIR camera specifications [73]

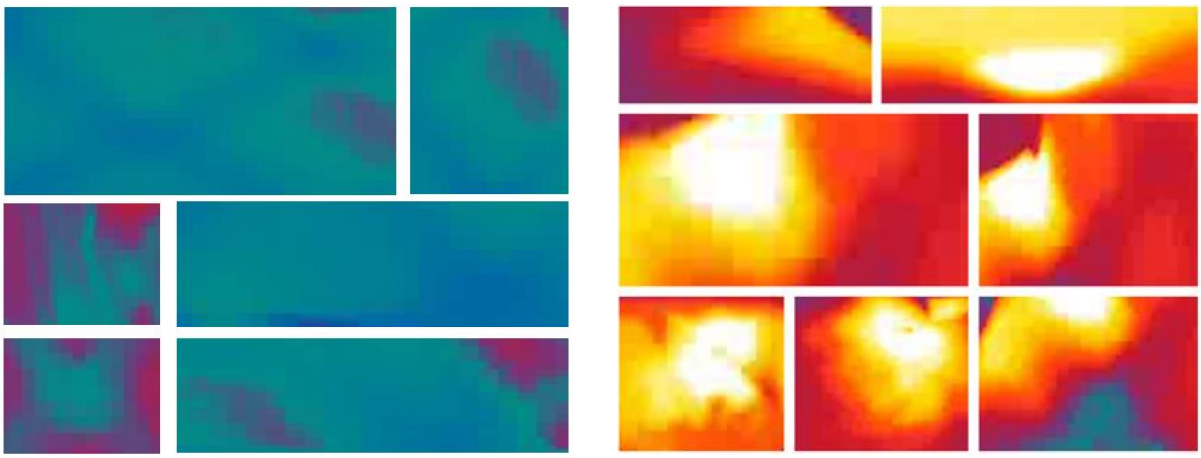
Camera	FLIR T650SC
Detector type	Uncooled micro-bolometer
Image frequency	30 Hz
Field of view	25°×19°
Resolution	640×480
Spectral range	7.5–14 μm
Emissivity correction	0.01 to 1.0 or from materials list
Laser	Manually operated
Operating temperature	−15 °C to 50 °C
Weight	2.87 lb

Chapter 5. Training and Validation

The proposed method requires thermal image data for detecting sub-surface damage in steel elements of a bridge. The training, validation and testing of the thermal images are explained in the following subsections. In section 5.1, table 2, and in section 5.2, Figure 18(a) are reproduced from a journal paper submitted to construction and building material [73].

5.1. Data bank generation

The samples of thermal images used for the purpose of training and validation are depicted in Figure 17 [73]. Different batches of thermal images, both for subsurface damage and intact sections, were cropped and labelled. The data were divided into two different categories: thermal images containing subsurface damage and thermal images without subsurface damage. A thermal image was classified as containing subsurface damage if any portion of the image featured a zone of subsurface damage. Such thermal images were cropped and labelled as subsurface damage, as shown in Figure 17(b). Conversely, for an intact image, the whole image should be without any subsurface defects, as shown in Figure 17(a). This principle was used to generate the entire data bank that was used in the training process.



a) Intact thermal images b) Sub-surface damage thermal images

Figure 17: Sample of thermal images for training [73]

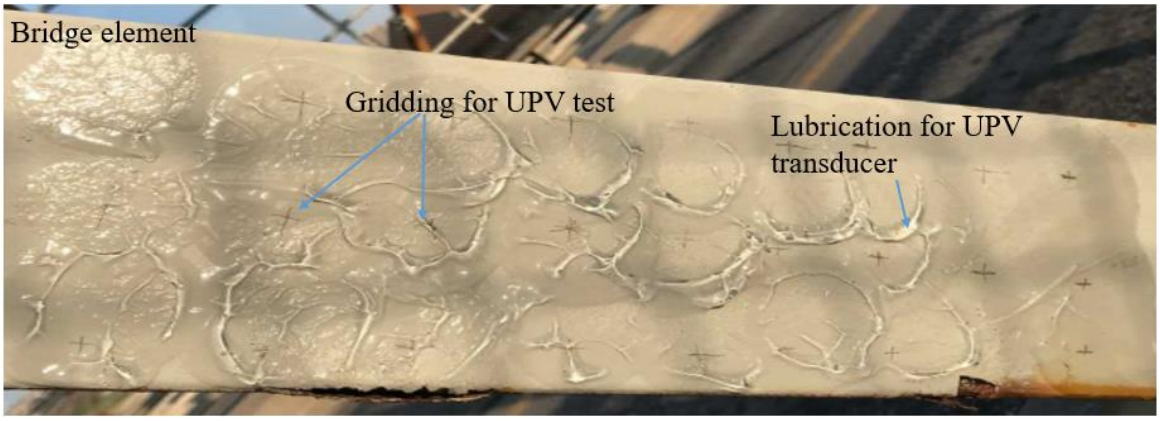
This process of sampling enables the neural network to easily distinguish between intact and subsurface damage zones. The samples of thermal images were validated using a UPV test method, which is described in the following subsections. The total number of images used for training and validation are listed in Table 2 [73]. Additionally, an augmentation technique was used in the network to increase the amount of training data by automatically rotating, scaling, and cropping the images, and by changing the brightness of the images.

Table 2: Thermal image data for training and validation [73]

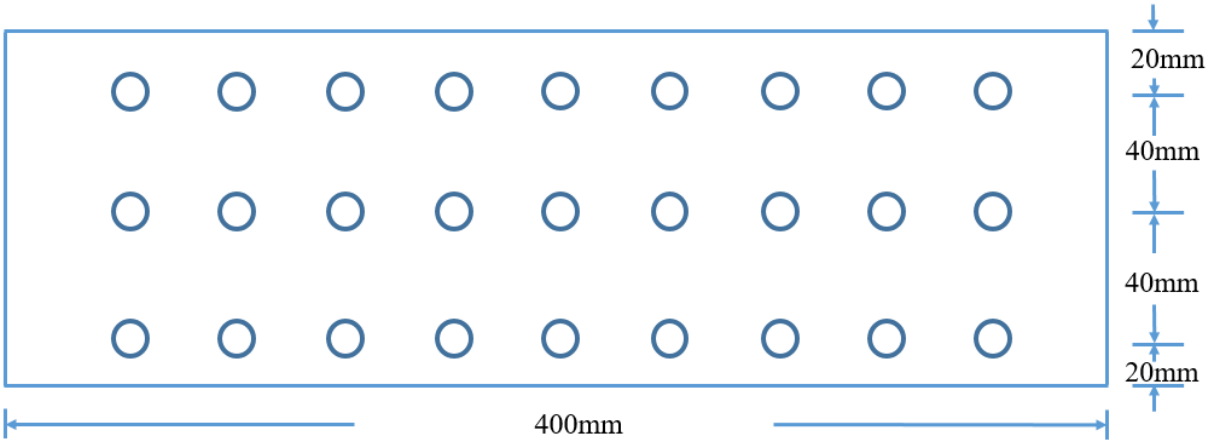
Number of Thermal images	Training (pixels)	Validation (pixels)
Intact	800 (299×299)	200 (299×299)
subsurface damage	800 (299×299)	200 (299×299)

5.2. Validation of data with a UPV tests

The thermal images provided sufficient data about the subsurface defects in the steel elements of the bridge. However, to further validate the thermal image results, a UPV test method was used [73]. This UPV uses ultrasonic pulses that pass through the member. By determining the time required for the pulses to travel through the section, the interior state of the section is studied. This information helps to characterize the interior defects within the material by distinguishing between the intact and subsurface damage portions [74]. The thermal images or videos of the steel members of the bridge were captured by a thermal camera and the same exact locations were reinvestigated by a UPV test [75]. To ensure that the UPV results and thermal images corresponded to the exact same locations, the evaluated sections were divided into several grids for UPV tests, as shown in Figure 18.



a) Gridding on steel elements at site



b) Gridding plan detail dimensions

Figure 18: UPV test gridding for training thermal data validation [73]

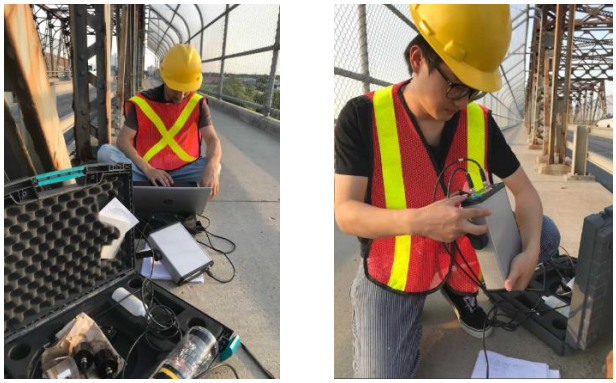
For instance, one section of a steel element of the bridge with dimensions of 400mm × 120mm × 10mm was used, as shown in Figure 18(a), with a gridding detail that is depicted in Figure 18(b). On each node, a non-destructive digital indicator UPV test was conducted and recorded. In order to transfer the recorded UPV data into a surface map, a surface mapping software was used. This software-generated map helps to easily compare the UPV map with the thermal results. A Surfer software (Surfer 15.5.382; Golden, CO, USA) was used to prepare UPV based contour maps [76]. For creating a grid file, a Surfer software randomly takes the grid data and utilizes it to produce a uniformly spaced data composed of nodes. It is difficult to choose the best gridding method for the Surfer software, because algorithms are computed in the software internally.

Surfer software has different gridding options, such as the Kriging method. Other gridding techniques include 1) Local polynomial method, 2) Nearest Neighbor methods, 3) Radial Basis function method, 4) Natural Neighbor method, and 5) Triangulation with linear interpolation method. However, for preparing UPV data contour surfaces, the Kriging grid method was used

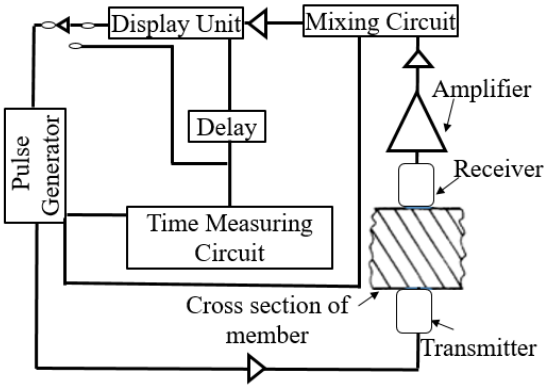
because it is more effective and can correct a dense data [77]. Additionally, it can also extrapolate values by preparing more flexible and accurate surfaces. Each node value of a grid is dependant on the adjacent node. The adjacent node value should be known and as used as a reference to calculate the given node value. The weights are provided to the given point based on its distance from the reference node. Data points are weighted based on the distance from the node. The distance between the two points is the factor that is used for augmenting the particular weights. The nearby points are given more weights than the points which are farther from the given node and therefore, the contribution of the farther values are considered less during estimation. For instance, to calculate the UPV value, Equation (14) is used [73].

$$G_A = \sum_{i=0}^{l} R_i G_i, \tag{14}$$

where G_A is the grid value of node A, G_i is the value at location i with weightR_i, and l is the total number of the adjacent values used in the calculation. Figure 19(a) shows the field experimental test and Figure 19(b) [73] shows the typical UPV circuitry requirement.



a) UPV test at site

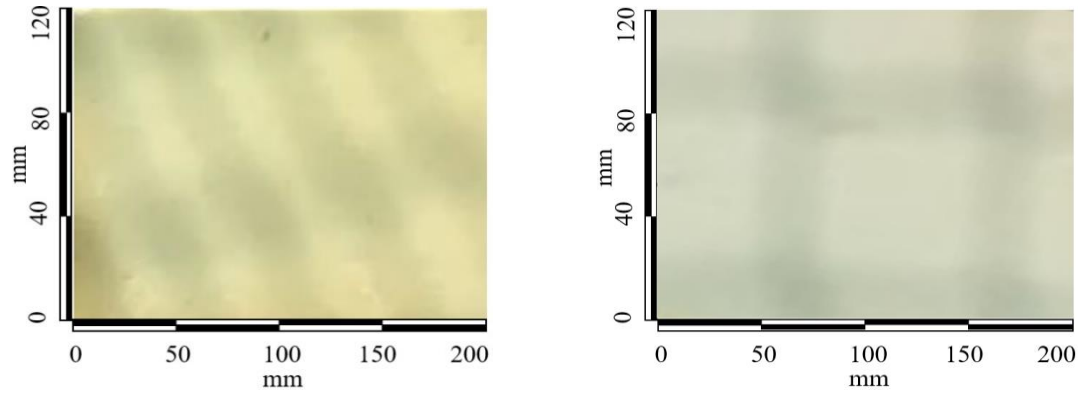


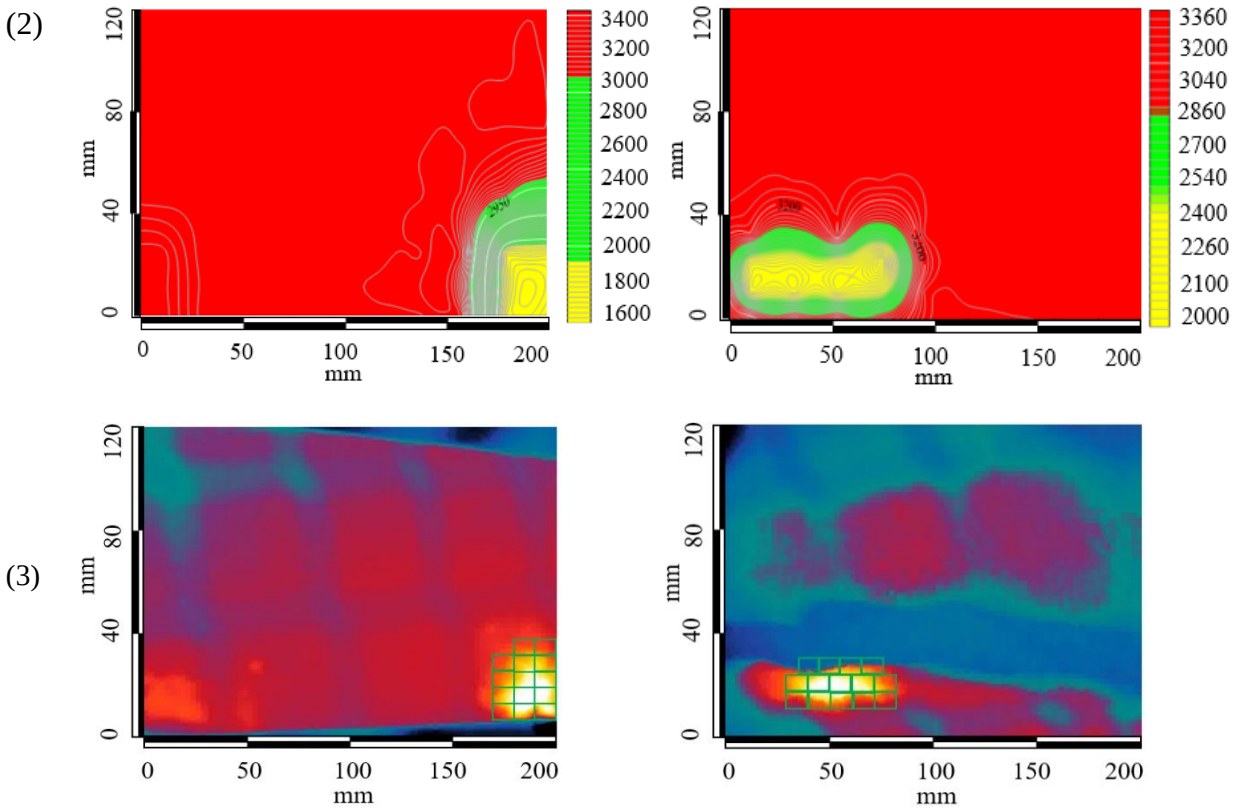
b) UPV circuitry requirement

Figure 19: Field experiment and UPV circuitry requirement [73]

The visible image, UPV based results, and thermal images of the subsurface damage detection are shown in Figure 20 [73]. Different types of structural sections, such as angles and channels sections, were used for validation purposes. Figure 20(a)(3) shows subsurface damage in the bottom left corner of a thermal image with green boxes. The subsurface damage is validated by a UPV test and the results are shown in Figure 20(a)(2). The yellow surface in Figure 20(a)(2) shows zones of high subsurface damage. The red surface shows intact surface while the green surface is the transition between intact and subsurface damage zones. Based on UPV tests, all thermal images were cropped and labelled as depicted in Figure 20(a)(3). A thermal image of one side of a channel section is shown in Figure 20(b). A UPV based contour map was drawn to validate the thermal image results. The UPV based results were used as a reference to label the thermal image using green bounding boxes as shown in Figure 20(b)(3). The labelled thermal image was used in the training process.

(1)





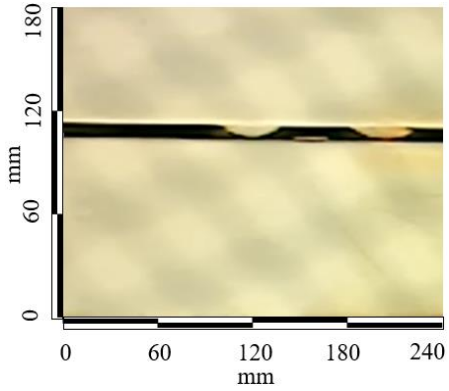
(a) Angle section (b) Channel section

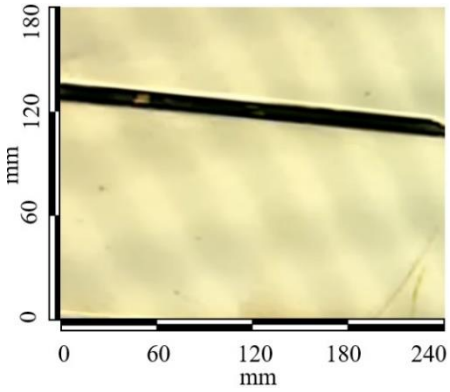
Figure 20: [73]

Figure 21 shows digital images, UPV based results, and thermal images for detecting subsurface damage [73]. Figure 21(a)(2) shows a UPV contour map of a back-to-back double channel section, and its thermal image is depicted in Figure 21(a)(3). The yellow colour in the UPV contour map is used to label thermal images for training the DINN. Another example of a double channel section is shown in Figure 21(b), which depicts subsurface damage in the lower right portion. A thermal image also shows impressive heat at the same location. Based on the previous literature, it is evident that subsurface damage zones of a member usually obstruct the flow of heat during the day, causing noticeable hot spots. Figure 21(b)(3) shows a thermal image that distinguishes the subsurface defects from the intact surface. For validating thermal image

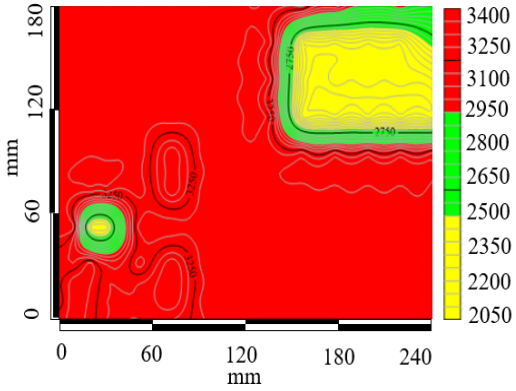
results, UPV based contour maps were used as shown in Figure 21(b)(2). Overall, the UPV and the thermal images identified the same regions of subsurface damage. However, slight differences were observed. The main causes of these differences were the variation in data collection speed and data collection time, and other environmental conditions such as temperature, humidity and wind speed, which significantly affect thermal images. On the other hand, UPV is based on interpolation and extrapolation which also slightly affects the results.

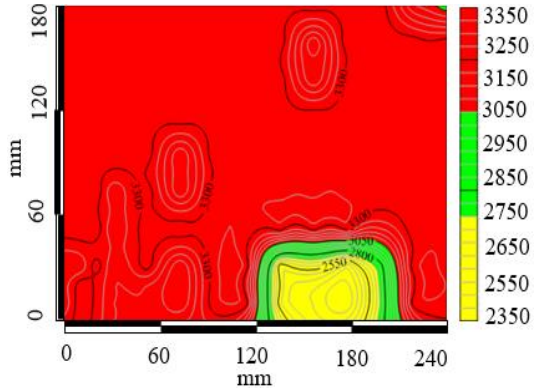
(1)



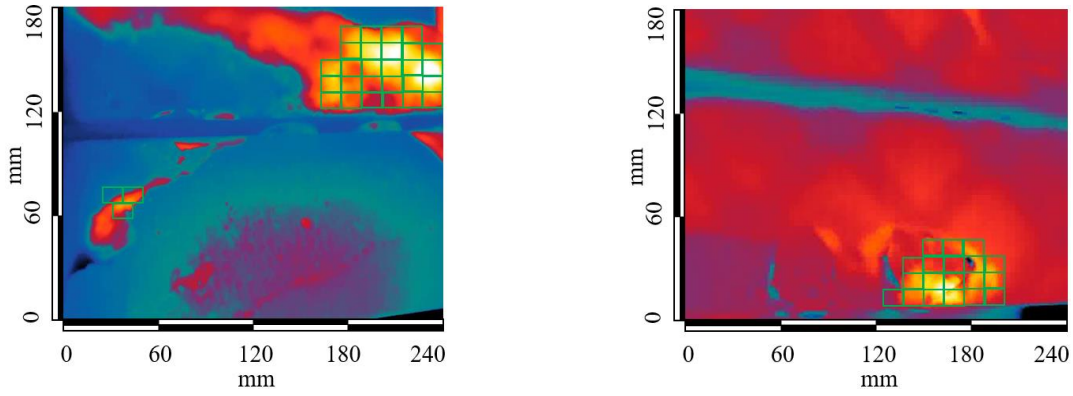


(2)





(3)



(a) double channel section (b) double channel section

Figure 21: Validation of subsurface damage using UPV [73]

5.3 Training and validation results

The modified DINN network was trained using UPV validation for all the data tabulated in Table 2. A group of 1600 thermal images with both intact and subsurface damage zones were used for training purposes. During the training process, a mini-batch of 100 thermal images was randomly selected and fed to the network. The training and validation accuracies were plotted as shown in Figure 22 [73]. Overall, the difference between training and validation accuracy is relatively small. A training accuracy of 96.5% was achieved at 100 iterations and then reduced gradually to 95.5%. A maximum training accuracy of 97.5% was achieved when the iterations reached to 300. The graph shows a continuous fluctuation in the training accuracy between 300 and 800 iterations. A training accuracy of 97.5% was maintained between 800 and 1200 iterations, and then again between 1500 and 2000 iterations. On the other hand, the validation accuracy was measured by testing 400 thermal images. The validation accuracy is similar to the training accuracy at several training steps. Based on the experimental results, 1200 iterations were used to test new thermal images. A maximum and minimum validation accuracy of 97.5 % and 95.5 % were achieved during the network validation. The training and validation processes were completed using a

workstation equipped with an Intel Core i7-6700K (CPU) at 4.00GHz, a 32 GB of RAM, and NVidia GeForce GTX 1080 version 385.54 GPU.

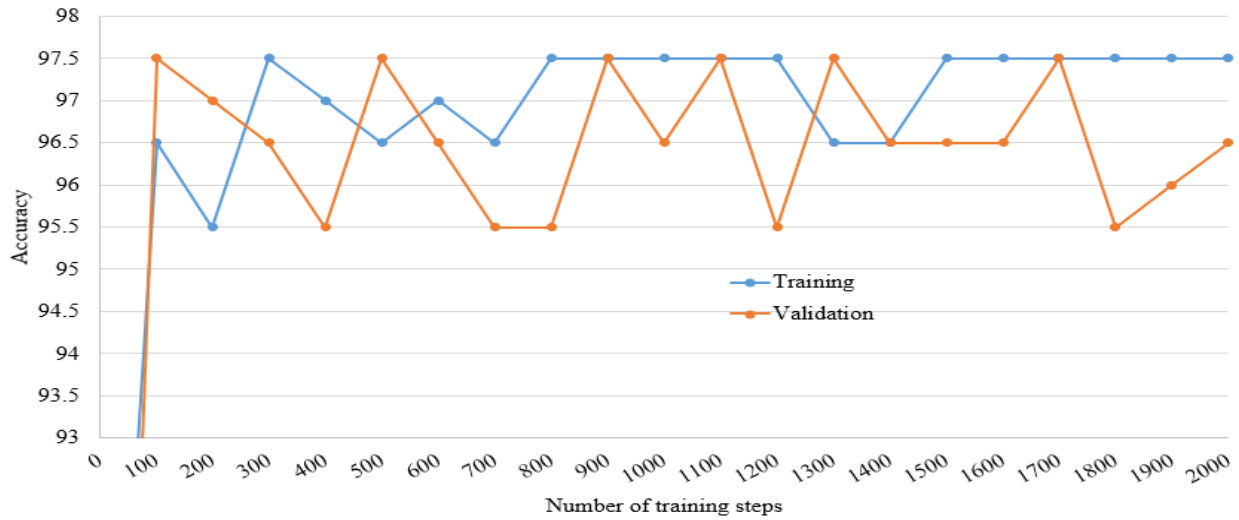


Figure 22: Accuracy vs training steps [73]

5.4 Prediction of new testing images

The trained network was used for testing new thermal images. A total of 200 new thermal images of steel elements of a bridge were used. The main purpose of the experiment was to define a threshold, i.e. to determine the optimum probability that could be used for testing new thermal images using the DINN network. The thermal images used for testing were not used for training and validation operations. To determine the optimum probability value, initially a probability value of 0.5 was selected. All thermal images used in the testing process were tested with a 50% probability value. The testing accuracy was calculated based on the results. The true positive (TP) and true negative (TN) were used for determining the accuracy. The false positive (FP) as well as the false negative (FN) were used for determining other statistical measures. The probability threshold of 0.5 for detecting subsurface damage means that if the similarity between the ground

truth and the subsurface damage area is 50% or more, the area would be categorized as containing subsurface damage and it would be localized by a green-bounding box. Similarly, more experiments were conducted using a range of probabilities, including 55%, 60%, 65%, 70%, 75%, and 80%. Testing accuracies of 95.65% and 95.80% were achieved by setting the probability to 50% and 55%, respectively.

A maximum testing accuracy of 96% was attained using a probability threshold value of 0.6. For probabilities above 0.6, the testing accuracy dropped steadily, reaching a minimum accuracy of 94% for probability of 0.8. As indicated in Figure 23, the testing accuracy is clearly declining as the probability exceeds 80%, and therefore no more experiments were conducted for probabilities above this value. The accuracy was calculated based on the Equation (15) below.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \tag{15}$$

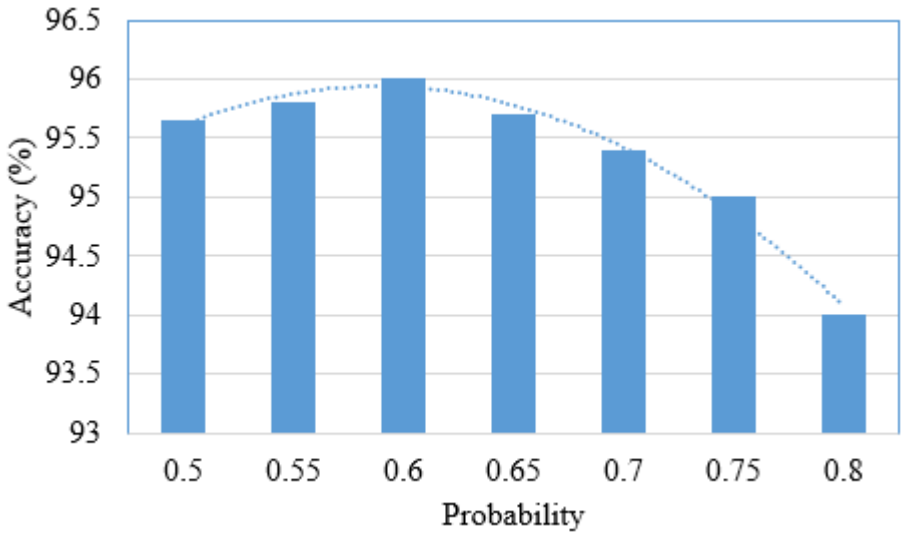


Figure 23: Probability threshold for subsurface damage detection vs testing accuracy

The overall outcomes of the experimental trials are tabulated in Table 3. The data were used to determine other statistical information such as sensitivity, specificity, precision and F1-score. The sensitivity was calculated based on Equation (16) and the results are expressed in Figure 24. The maximum sensitivity was achieved with a probability value of 0.6.

$$\text{Sensitivity} = \frac{TP}{TP + FN}, \tag{16}$$

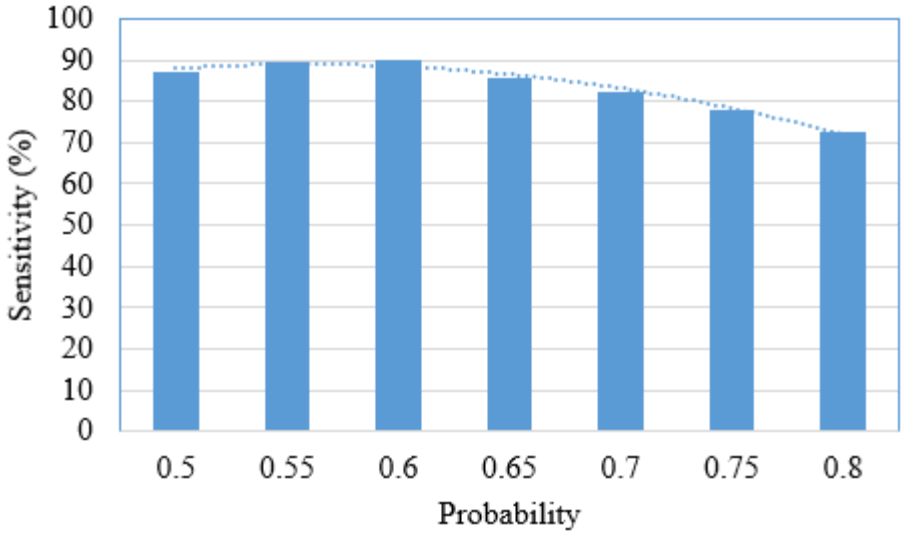


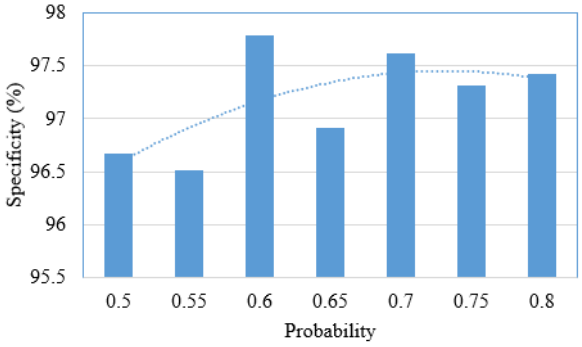
Figure 24: Probability threshold for subsurface damage detection vs sensitivity

The specificity, precision and F1-score were calculated based on Equation (17), (18) and (19), respectively. The specificity results are presented in Figure 25. Based on the statistical measures, it was found that the probability of 0.6 yielded high specificity, sensitivity, and accuracy, as well as a high F1-score. Therefore, this probability was deemed the optimal probability for testing new thermal images. However, a probability value of 0.6 resulted in a relatively low precision. The relationship between the F1-score and probability is depicted in Figure 26.

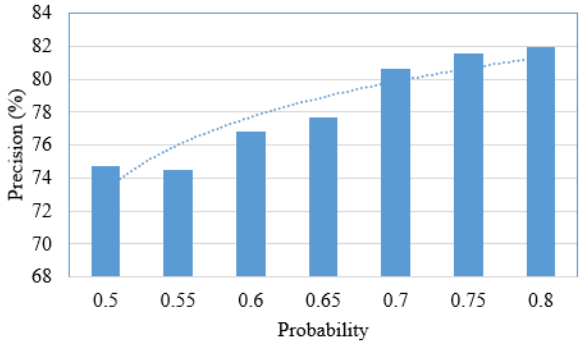
$$\text{Specificity} = \frac{TN}{TN + FP}, \tag{17}$$

$$\text{Precision} = \frac{TP}{TP + FP}, \tag{18}$$

$$\text{F1-score} = \frac{2TP}{2TP + FP + FN}, \tag{19}$$



a) Specificity vs probability threshold for subsurface damage detection



b) Precision vs probability threshold for subsurface damage detection

Figure 25: Statistical measures

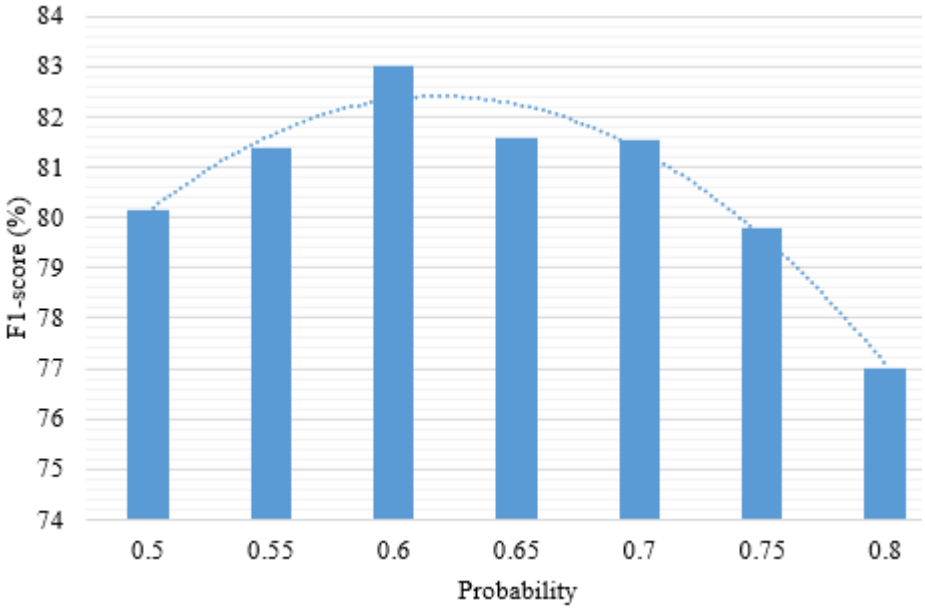
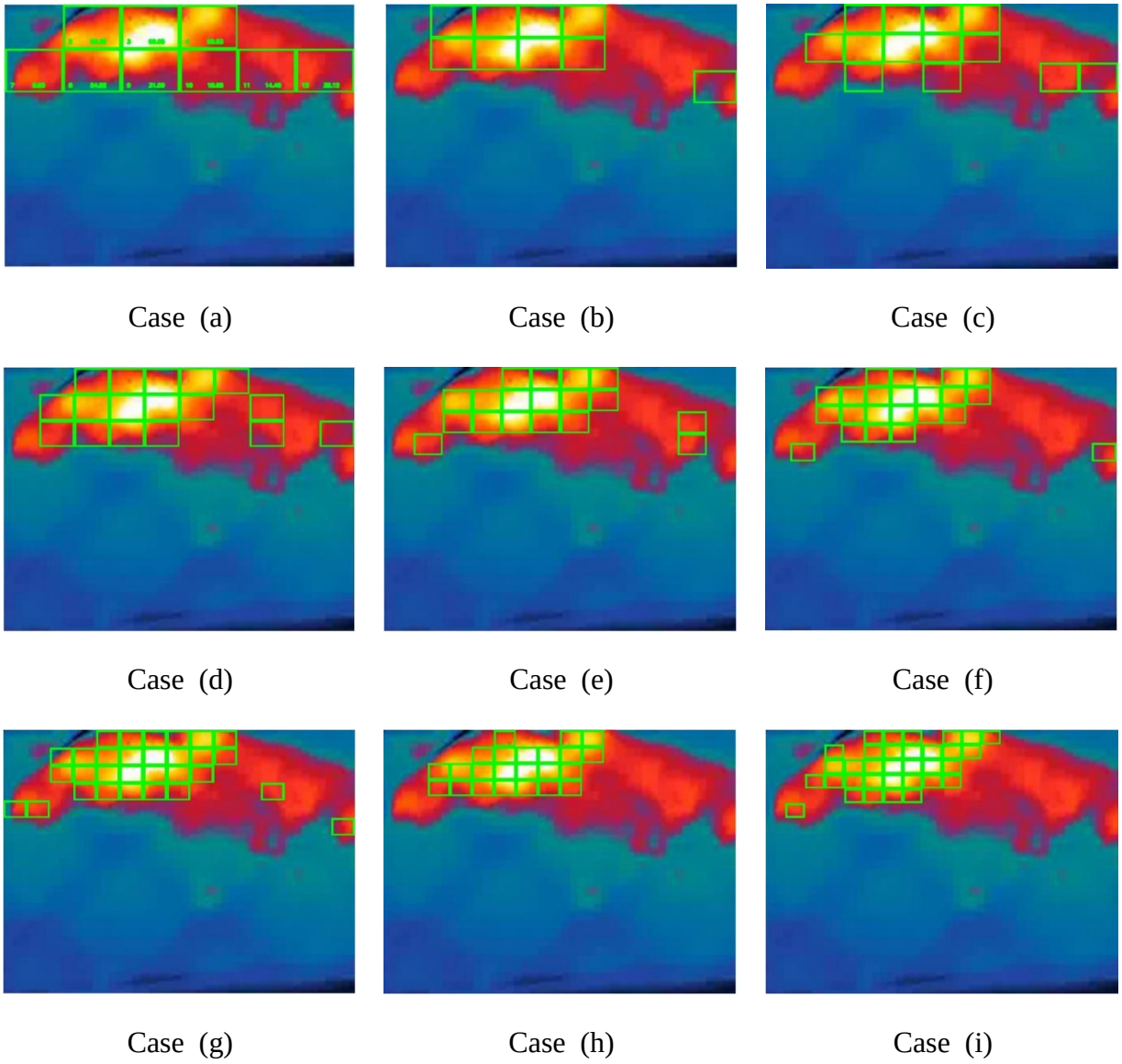


Figure 26: F1-score vs probability threshold for subsurface damage detection

The above statistical measures show that a probability of 0.6 should be used as a threshold for testing new thermal images. In order to test new thermal images the size of bounding box is

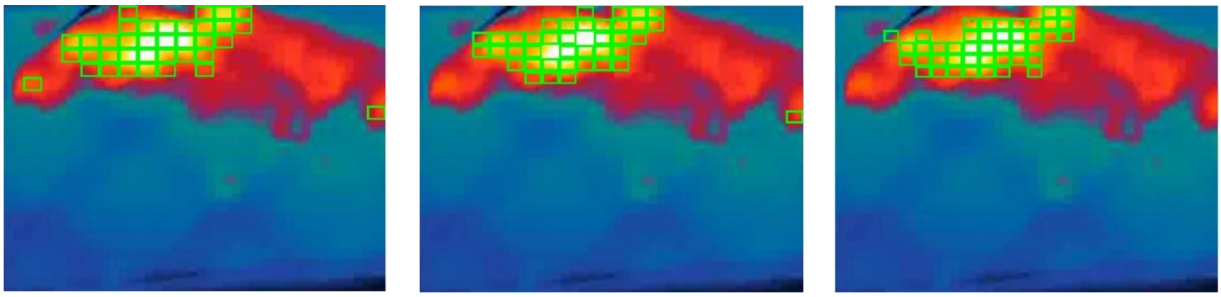
also important. Therefore, several experiments were conducted to find the best number of image divisions to provide high accuracy with low computational cost. A total of 12 different cases were studied to determine the optimal number of divisions, as shown in Figure 27. The details of the experiments for the 12 different cases are presented in Table 3.



Case (a) Case (b) Case (c)

Case (d) Case (e) Case (f)

Case (g) Case (h) Case (i)



Case (j) Case (k) Case (l)

Figure 27: Effect of number of division on computational cost and accuracy

Case (a) has the smallest number of divisions, as shown in Figure 27, and has the minimum computation cost. However, the false positive percentage for case (a) is higher in comparison to other cases. The minimum false positive percentage is achieved in case (l), but with a greater computation cost. In order to select the best number of divisions from these 12 cases, both the computation cost and false positive percentage were considered in the evaluation process.

In case (h) the large thermal image of 640 × 480 pixels was divided by 16 in both the vertical and horizontal direction, to obtain 256 images of 40 × 30 pixels. The total time required to analyze a new testing image in case (h) was only 55 seconds for a large image of 640 × 480 pixels or 0.21 seconds for a sub-image of 40 × 30 pixels.

Table 3: Different size of bounding boxes results [73]

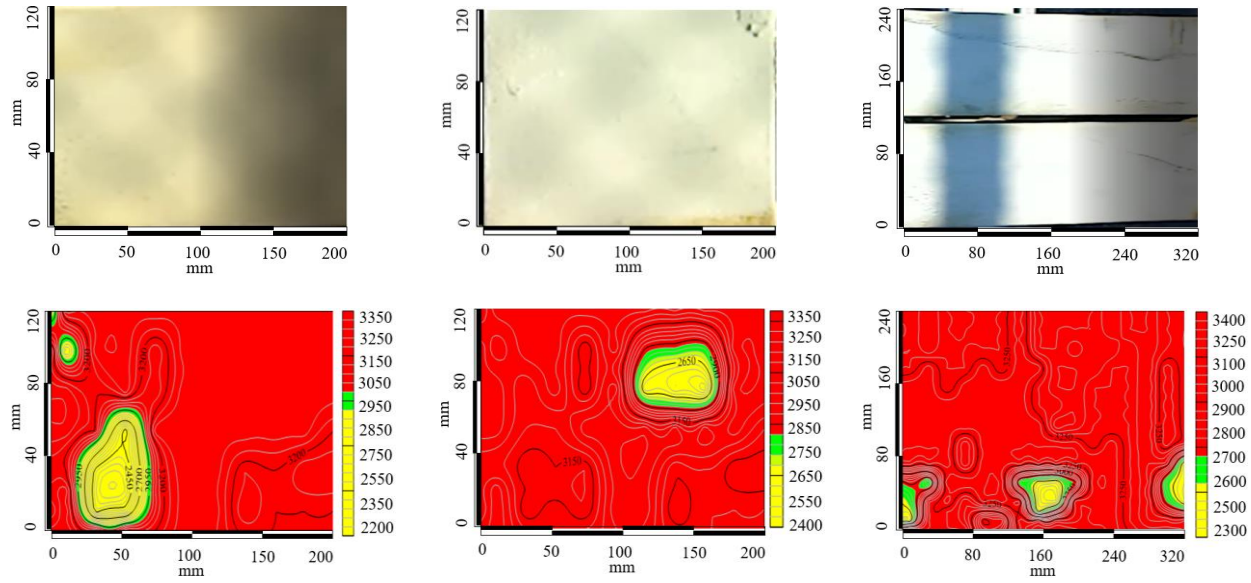
Case	Image size pixels	Divisions	Processing in seconds	Number of Sub-images	Size of sub-images	False positive (%)
(a)	640 × 480	6 × 6	9.29	36	107 × 80	11.14
(b)	640 × 480	8 × 8	15.75	64	80 × 40	6.25
(c)	640 × 480	9×9	18.84	81	71 × 53	4.89

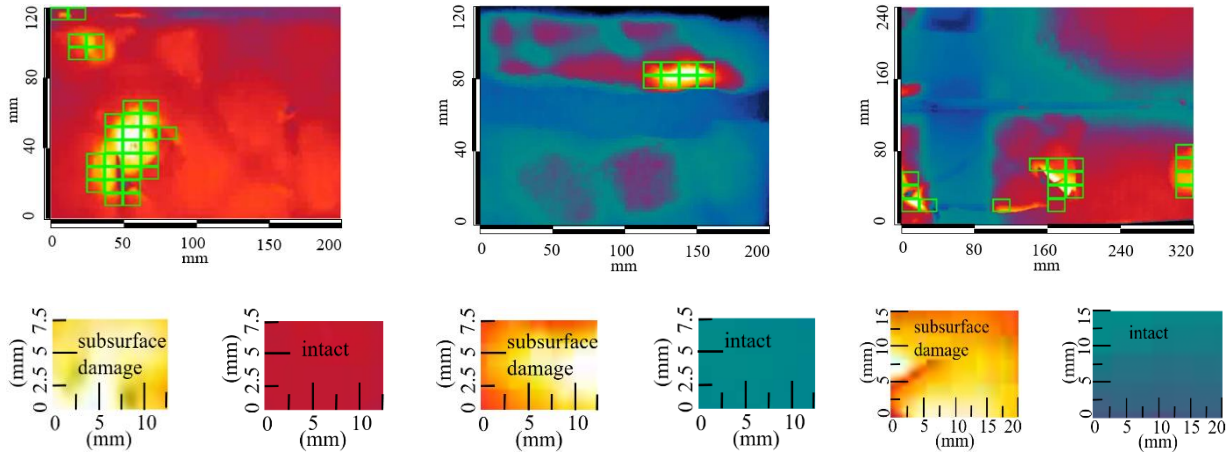
(d)	640 × 480	10×10	26.71	100	64 × 48	4.70
(e)	640 × 480	12×12	31.39	144	53 × 40	3.45
(f)	640 × 480	14×14	42.32	196	46 × 34	3.05
(g)	640 × 480	15×15	48.49	225	42 × 32	2.69
(h)	640 × 480	16×16	55.16	256	40 × 30	1.17
(i)	640 × 480	18×18	69.73	324	36 × 27	1.16
(j)	640 × 480	20×20	83.84	400	32 × 24	0.75
(k)	640 × 480	22×22	104.79	484	29 × 22	0.62
(l)	640 × 480	24×24	120.50	576	26 × 20	0.34

In order to determine the robustness of the network, several new thermal images were fed to the DINN and tested. The thermal image-based results were compared with the UPV-based surface maps as shown in Figure 28 and Figure 29 [73]. Figure 28(a) shows the results that were obtained for a single side of an angle section, including the digital image, the UPV-based results, and a thermal image in which the damage is delineated by green bounding boxes. The yellow colour in the thermal image represents zones of subsurface damage in the section. It was postulated that the subsurface defects may have been either debonding of coatings, corrosion below the coating, or steel subsurface delamination. The UPV tests were conducted to check the robustness of the DINN. Figure 28(b) presents a single side of another angle section which contains a subsurface damage in the upper right side of the section. The areas of subsurface damage are depicted by a yellow colour. The red colour shows the intact areas of the section.

The transition zone between intact regions and regions of subsurface damage is depicted by a green colour. The areas of subsurface damage in the UPV-based results are slightly bigger

than those depicted in the thermal images. One reason for this discrepancy is that the UPV-based results were prepared using a surface mapping software, which approximates the adjacent values of nodes using extrapolation and interpolation techniques. On the other hand, the thermal image-based results are highly affected by solar heating and weather conditions, and therefore vary considerably throughout the day. Figure 28(c) shows a double angle section which contains subsurface defects in four different locations. The section was selected to test the effect of the space between the two angle sections. The space between the two structural members shows different temperature values than the surrounding material. A thermal image accurately differentiates between the two different surfaces, and the DINN results successfully detect the zones of subsurface damage. The same section shown in Figure 28(c) was tested using UPV, at the exact same location. Both the results shows a good agreement for the double angle section. However, the UPV results did not display a small subsurface damage near the bottom edge, while the DINN detected this spot with a single bounding box.





a) Subsurface damage in angle section

b) Subsurface damage in angle section

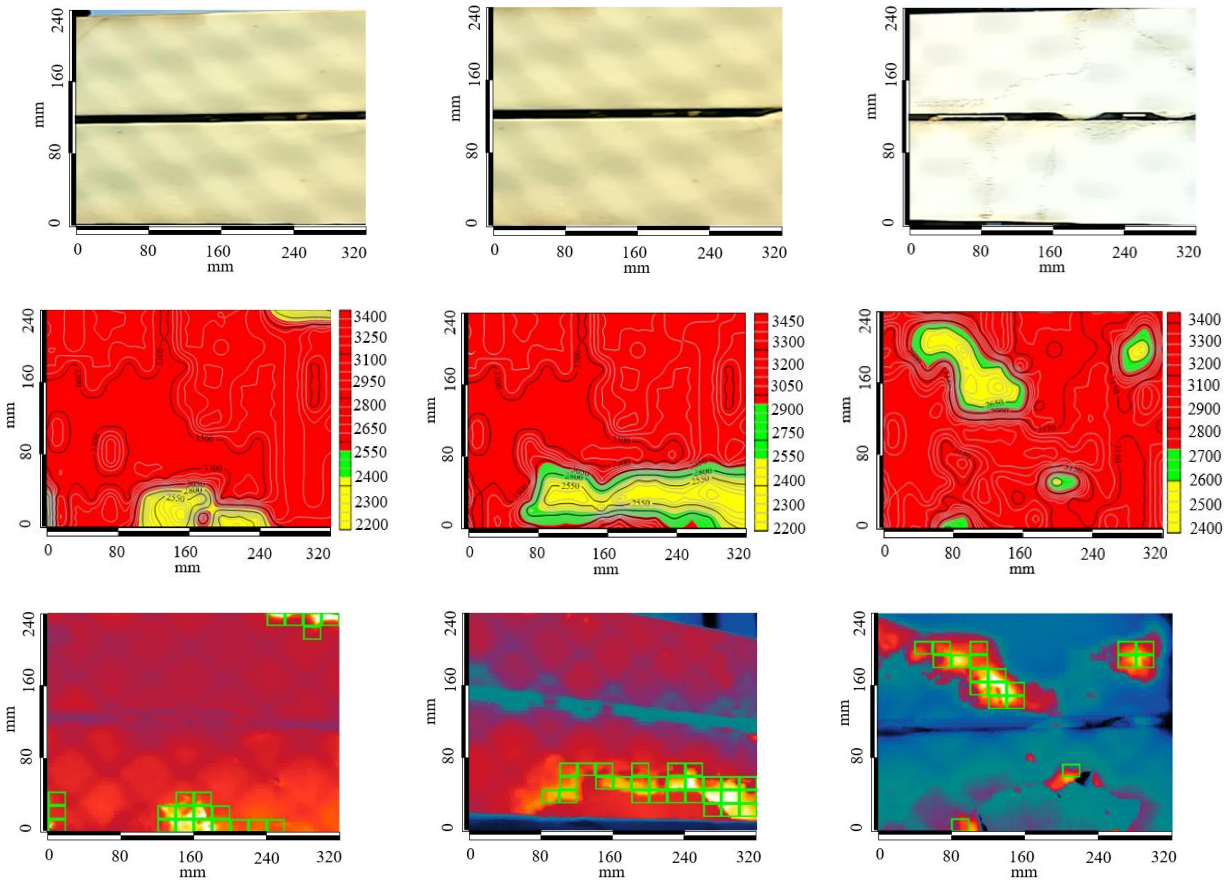
c) Subsurface damage in double-angle section

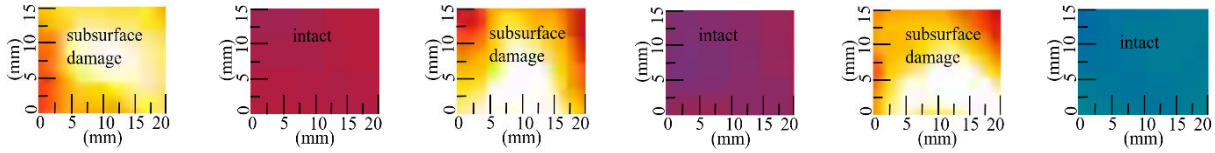
Figure 28: Testing different structural members of a bridge [73]

Figure 29 shows more examples of testing images. Figure 29(a) shows a digital red, green and blue (RGB) image, a UPV-based surface map, and a thermal image of a double channel section in which the damage is delineated by green bounding boxes. The thermal image shows a subsurface damage in both channel sections. The digital and thermal images both show the shadow of the nearby steel fence. The section was selected to determine the effects of a temporary shadow. The thermal image-based results were not effected by such temporary shadows. Figure 29(b) shows another example of a double channel section. One side of the section is completely intact. However, the other side shows longitudinal subsurface defects. The thermal image was fed to the network and damaged zones were detected and demarcated by green bounding boxes. The UPV-based results are congruent with the thermal image-based results.

Another double channel section was selected to study the effects of small structural fasteners such as rivets, as shown in Figure 29(c). The sides of the section where the two channel sections were connected by a rivet were captured in the thermal image. The results were not

affected by such minor structural fasteners. The DINN successfully detects large zones of subsurface damage. However, smaller zones of subsurface damage, which can be seen in the centre and bottom portion of the Figure 29(c), were indicated by only a single bounding box, although the bounding box not fully enclose the region of subsurface damage. It was found that more accurate results could be obtained by reducing the size of the bounding boxes. However, reducing the size of the bounding boxes also increases computational cost. The differences between the UPV-based results and thermal image-based results were insignificant. These differences can be attributed to two main factors: firstly, the UPV test results were visualized using a surface mapping software which is based on extrapolation. Secondly, the thermal image results are highly dependant on the environmental conditions.





a) Double channel section with shadow

b) Longitudinal subsurface defects in one section

c) Section with minor strutcural elements

Figure 29: Testing different structural members of a bridge [73]

Chapter 6. Conclusion and Future work

This chapter contains a summary of the proposed method used for subsurface damage detection. The limitations of the method are discussed and recommendations for future work are provided.

6.1 Conclusion

An approach for detecting subsurface damage in steel bridge elements is proposed. The infrared thermography was incorporated with a neural network to detect interior defects. In total, 2000 cropped thermal images of random sizes were fed to a deep inception neural network and 2000 corresponding compressed feature maps were generated. These include 1600 thermal images for the purpose of training and 400 thermal images for the purpose of validation. The compressed feature maps were classified in the fully connected layer and softmax layer in the transfer learning section. The network detects the subsurface damage portion by using green bounding boxes. The subsurface damage identified by the deep learning thermal image based method was validated by UPV tests. A surface mapping software was used to compare the UPV results with the deep learning thermal image based results. The robustness of the network was determined by testing 200 new thermal images. All the images were fed to the network and regions of subsurface damage were localized by green bounding boxes. The thermal images used for testing were also validated by a UPV test.

Thermal images of different subsurface damage conditions, including sections with a single zone of subsurface damage, sections with multiple zones of subsurface damage, and sections with a different background, were tested. Images of subsurface damage in different types of structural members, including double angle sections and double channel sections, both with and

without noise, were also tested. The noise was either in the form of minor structural elements, such as bolts and rivets, or the shadow of a protective fence on the bridge. All these images were fed to the network, and portions of subsurface damage were highlighted autonomously with green bounding boxes. It was found that the network can accurately detect subsurface damage in thermal images. Based on the experiments, the following conclusions were made:

- The data were collected in four different months, at different intervals of the day from dawn to dusk. The optimum time of day for detecting subsurface damage in thermal images was investigated. The experiments show that the quality of the collected data depends on the orientation of the bridge structure with respect to the sun. In the morning, good quality thermal images can be collected from the side of the bridge which is exposed to the solar radiation. On the other hand, the opposite side of the bridge can be best investigated just before sunset.
- The computational cost of the modified DINN network was reduced by introducing a new fully connected layer in the transfer learning section.
- A maximum training accuracy of 97% and a maximum validation accuracy of 95.5% were achieved at 1200 iterations. The validation accuracy was approximately equal to the training accuracy at various iterations.
- A total of 12 different cases were used to determine the optimum number of image divisions, considering both computational cost and testing accuracy. It was found that a high accuracy and an appropriate computational cost were obtained when a large thermal image of 640 × 480 pixels was divided into 256 sub-images of 40 × 30 pixels.

- Several tests were conducted to define the optimum threshold of probability to be used for detecting delamination in thermal images. After careful investigation, it was found that a probability of 60% achieved a maximum testing accuracy of 96%.
- Specificity and F1-score were determined for a probability range of 50% to 80%. A maximum specificity of 97% and an F1-score of 83% were achieved at a probability of 60%.
- The UPV tests were used to validate thermal image-based results. It was found that surfer-based UPV maps show good agreement with the thermal image-based results.

6.2 Limitations

The following limitations were observed during the application of the proposed method:

- Since passive thermography was used, this method is applicable only in the summer season. In addition, environmental conditions, such as rain and cloud cover, also limit the process of data collection.
- The results are highly affected when the thermal image contain noise caused by vehicles in the background of the bridge member. Therefore, extra care should be taken to avoid such noise during data collection.
- The method is not applicable to structural members which are not exposed to the direct radiation of the sun, such as the interior girder below the bridge.

6.3 Future works

The aforementioned limitations provide an opportunity for future work in the field of structure health monitoring, such as the following:

- Active thermography should be studied in future to detect subsurface damage during the winter season as well. The active thermography would also be helpful to detect

subsurface damage in structural members which are not exposed to direct solar radiation, which is a problem even during the summer season.

- Multispectral dynamic images, which contain both visible and thermal information, should be further studied to detect multiple types of damage in structural members.

References

[1] Cartz, L. (1995). Non-destructive testing. ASM International. ISBN 978-0-87170-517-4.

[2] Clemena, G. G., & McKeel, W. T. (1978). Detection of delamination in bridge decks with infrared thermography. *Gerardo G. Clemena and Wallace T. McKeel, JR., Transportation Research Record*, (664), 180-182.

[3] Bao, Y., Hoehler, M. S., Smith, C. M., Bundy, M., & Chen, G. (2017). Temperature measurement and damage detection in concrete beams exposed to fire using PPP-BOTDA based fiber optic sensors. *Smart materials and structures*, *26*(10), 105034.

[4] Wolff, R. (1992). Monitoring of prestressed concrete structures with optical fiber sensors. In *First European Conference on Smart Structures and Materials* (Vol. 1777, p. 177706). International Society for Optics and Photonics.

[5] Ghiassi, B., Verstrynge, E., Lourenço, P. B., & Oliveira, D. V. (2014). Characterization of debonding in FRP-strengthened masonry using the acoustic emission technique. *Engineering Structures*, *66*, 24-34.

[6] Shah, A. A., Ali, R., Naseer, A., & Zhang, C. (2014). Assessment of Progressive Damage in Concrete with Acoustic Emission Technique. *Advances in Applied Acoustics (AIAAS)*, *3*.

[7] Huston, D. R., Pelczarski, N. V., Esser, B., & Maser, K. R. (2000). Damage detection in roadways with ground penetrating radar. In *Eighth International Conference on Ground Penetrating Radar* (Vol. 4084, pp. 91-95). International Society for Optics and Photonics.

[8] Epp, T., & Cha, Y. J. (2016). Air-coupled impact-echo damage detection in reinforced concrete using wavelet transforms. *Smart Materials and Structures*, *26*(2), 025018.

[9] Kim, B., & Cho, S. (2018). Automated Vision-Based Detection of Cracks on Concrete Surfaces Using a Deep Learning Technique. *Sensors*, *18*(10), 3452.

[10] Cha, Y. J., You, K., & Choi, W. (2016). Vision-based detection of loosened bolts using the Hough transform and support vector machines. *Automation in Construction, 71*, 181-188.

[11] Cha, Y. J., Choi, W., & Büyüköztürk, O. (2017). Deep learning-based crack damage detection using convolutional neural networks. *Computer-Aided Civil and Infrastructure Engineering*, *32*(5), 361-378.

[12] Cha, Y.J.; Chen, J.; Büyüköztürk, O. (2015). Motion magnification based damage detection using high speed video. In Proceedings of the *10th International Workshop on Structural Health Monitoring (IWSHM)*, Stanford, CA, USA, 1–3.

[13] Abdel-Qader, I., Abudayyeh, O. and Kelly, M. E., (2003). Analysis of Edge-Detection Techniques for Crack Identification in Bridges. Journal of Computing in Civil Engineering, 17(4): 255-263.

[14] Ali, R., Gopal, D. L., & Cha, Y. J. (2018). Vision-based concrete crack detection technique using cascade features. In *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2018* (Vol. 10598, p. 105980L). International Society for Optics and Photonics.

[15] Wu, L., Mokhtari, S., Nazef, A., Nam, B., & Yun, H. B. (2014). Improvement of crack-detection accuracy using a novel crack defragmentation technique in image-based road assessment. *Journal of Computing in Civil Engineering*, *30*(1), 04014118.

[16] Cha, Y. J., Choi, W., Suh, G., Mahmoudkhani, S., & Büyüköztürk, O. (2018). Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types. *Computer-Aided Civil and Infrastructure Engineering*, *33*(9), 731-747.

[17] Zeiler, M. D., & Fergus, R. (2014). Visualizing and understanding convolutional networks. In *European conference on computer vision* (pp. 818-833). Springer, Cham.

[18] Maldague, X., (2001). Theory and Practice of Infrared Technology for Nondestructive Testing.

[19] Bagavathiappan, S., Lahiri, B. B., Saravanan, T., Philip, J., & Jayakumar, T. (2013). Infrared thermography for condition monitoring–A review. *Infrared Physics & Technology*, *60*, 35-55.

[20] Robert, M. (1982). Science behind Thermography-Thermal Infrared Sensing for Diagnostics and Control. *J. of Thermosense,* Vol. 371, 2-9.

[21] Gucunski, N., & National Research Council. (2013). *Nondestructive testing to identify concrete bridge deck deterioration*. Transportation Research Board.

[22] Vaghefi, K., Ahlborn, T. T. M., Harris, D. K., & Brooks, C. N. (2013). Combined imaging technologies for concrete bridge deck condition assessment. *Journal of Performance of Constructed Facilities*, *29*(4), 04014102.

[23] Oh, T., Kee, S. H., Arndt, R. W., Popovics, J. S., & Zhu, J. (2012). Comparison of NDT methods for assessment of a concrete bridge deck. *Journal of Engineering Mechanics*, *139*(3), 305-314.

[24] Omar, T., & Nehdi, M. L. (2016). Non-Destructive Testing of Bridge Deck Using Passive Infrared Thermography and Ground Penetrating Radar. In *TAC 2016: Efficient Transportation-Managing the Demand-2016 Conference and Exhibition of the Transportation Association of Canada*.

[25] Canadian Institute of Steel Construction (CISC) (2013).

Available at: https://www.constructioncanada.net/specifying-structural-steel-for-bridges/.

[26] Karlsson, J. (2011). Corrosion Mechanisms under Organic Coatings-A study in relation to Next Generation's Pretretments.

[27] Amirudin, A., & Thierry, D. (1996). Corrosion mechanisms of phosphated zinc layers on steel as substrates for automotive coatings. *Progress in Organic Coatings*, *28*(1), 59-75.

[28] Marcus, P. (Ed.). (2011). *Corrosion mechanisms in theory and practice*. CRC press.

[29] Cadelano, G., Bortolin, A., Ferrarini, G., Molinas, B., Giantin, D., Zonta, P., & Bison, P. (2016). Corrosion detection in pipelines using infrared thermography: experiments and data processing methods. *Journal of Nondestructive Evaluation*, *35*(3), 49.

[30] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2818-2826).

[31] Hallermann, N., & Morgenthal, G. (2014). Visual inspection strategies for large bridges using Unmanned Aerial Vehicles (UAV). In *Proc. of 7th IABMAS, International Conference on Bridge Maintenance, Safety and Management* (pp. 661-667).

[32] Moore, W. M. (1972). *Detection of bridge deck deterioration* (No. Intrm Rpt). Texas Transportation Institute, Texas A & M University.

[33] Cawley, P. (2001). Non-destructive testing—current capabilities and future directions. *Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials: Design and Applications*, *215*(4), 213-223.

[34] Dwivedi, S. K., Vishwakarma, M., & Soni, A. (2018). Advances and Researches on Non Destructive Testing: A Review. *Materials Today: Proceedings*, *5*(2), 3690-3698.

[35] Gholizadeh, S. (2016). A review of non-destructive testing methods of composite materials. Procedia Structural Integrity, 1, 50-57.

[36] Morey, R. M. (1998). *Ground penetrating radar for evaluating subsurface conditions for transportation facilities*. Transportation Research Board.

[37] ACI 228.2R-98, (1998). Nondestructive Test Methods for Evaluation of Concrete in Structures. ACI, American Concrete Institute Report, Farmington Hills, Michigan, p. 62

[38] Azari, H., Nazarian, S., & Yuan, D. (2014). Assessing sensitivity of impact echo and ultrasonic surface wave's methods for nondestructive evaluation of concrete structures. *Construction and Building Materials*, *71*, 384-391.

[39] Bogas, J. A., Gomes, M. G., & Gomes, A. (2013). Compressive strength evaluation of structural lightweight concrete by non-destructive ultrasonic pulse velocity method. *Ultrasonics*, *53*(5), 962-972.

[40] ASTM C597-09, (2009). Standard Test Method for Pulse Velocity through concrete. American Standards for Testing Materials.

[41] Schickert, M. (2005). Progress in ultrasonic imaging of concrete. *Materials and Structures*, *38*(9), 807-815.

[42] Shah, A. A., Ribakov, Y., & Zhang, C. (2013). Efficiency and sensitivity of linear and non-linear ultrasonics to identifying micro and macro-scale defects in concrete. *Materials & Design*, *50*, 905-916.

[43] Hopwood, T. (1988). Acoustic emission inspection of steel bridges. *Public Works*, *119*(6).

[44] Holford, K. M., Davies, A. W., Pullin, R., & Carter, D. C. (2001). Damage location in steel bridges by acoustic emission. *Journal of intelligent material systems and structures*, *12*(8), 567-576.

[45] Holford, K. M., & Carter, D. C. (1999). Acoustic emission source location. In *Key Engineering Materials* (Vol. 167, pp. 162-171). Trans Tech Publications.

[46] Shateri, M., Ghaib, M., Svecova, D., & Thomson, D. (2017). On acoustic emission for damage detection and failure prediction in fiber reinforced polymer rods using pattern recognition analysis. *Smart Materials and Structures*, *26*(6), 065023.

[47] Nair, A., & Cai, C. S. (2010). Acoustic emission monitoring of bridges: Review and case studies. *Engineering structures*, *32*(6), 1704-1714.

[48] Carlos, M. F., Cole, P. T., Vahaviolos, S. J., Halkyard, T., & Alampalli, S. (2000). Acoustic emission bridge inspection/monitoring strategies. In *Structural Materials Technology IV An NDT Conference* (pp. 179-183).

[49] Grosse, C., Finck, R., Kurz, J., & Reinhardt, H. (2004). Monitoring techniques based on wireless AE sensors for large structures in civil engineering. In *Proceedings of the EWGAE 2004 Symposium in Berlin* (pp. 843-856).

[50] Carter, D. C., & Holford, K. M. (1998). Strategic considerations for the AE monitoring of bridges: a discussion and case study. *Insight-Non-Destructive Testing and Condition Monitoring*, *40*(2), 112-116.

[51] Steinbach, J., & Vey, E. (1975). Caisson evaluation by stress wave propagation method. *Journal of Geotechnical and Geoenvironmental Engineering*, *101*(ASCE# 11234 Proceeding).

[52] Barrile, V., & Pucinotti, R. (2005). Application of radar technology to reinforced concrete structures: a case study. *NDT & e International*, *38*(7), 596-604.

[53] Cantor, T. R. (1984). Review of penetrating radar as applied to nondestructive evaluation of concrete. *Special Publication*, *82*, 581-602.

[54] Carter, C. R., Chung, T., Holt, F. B., & Manning, D. G. (1986). An automated signal processing system for the signature analysis of radar waveforms from bridge decks. *Canadian Electrical Engineering Journal*, *11*(3), 128-137.

[55] Hugenschmidt, J., & Mastrangelo, R. (2006). GPR inspection of concrete bridges. *Cement and Concrete Composites*, *28*(4), 384-392.

[56] McCann, D. M., & Forde, M. C. (2001). Review of NDT methods in the assessment of concrete and masonry structures. *Ndt & E International*, *34*(2), 71-84.

[57] Rehman, S. K. U., Ibrahim, Z., Memon, S. A., & Jameel, M. (2016). Nondestructive test methods for concrete bridges: A review. *Construction and Building Materials*, *107*, 58-86.

[58] M. Sansalone, N.J. Carino, Impact-echo: A Method for Flaw Detection in Concrete using Transient Stress Waves, US Department of Commerce, National Bureau of Standards, Center for Building Technology, Structures Division, 1986.

[59] Lin, Y., & Su, W. C. (1996). Use of stress waves for determining the depth of surface-opening cracks in concrete structures. *Materials Journal*, *93*(5), 494-505.

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

[61] Grosse, C. U., Reinhardt, H. W., Krüger, M., & Beutel, R. (2005). Application of Impact-echo techniques for crack detection and crack parameter estimation in concrete. In *extended abstract, International Conference on Fracture ICF-11, Turin* (pp. 20-25).

[62] Davis, A. G. (2003). The nondestructive impulse response test in North America: 1985–2001. NDT & E International, 36(4), 185-193.

[63] Whittington, H. W., McCarter, J., & Forde, M. C. (1981). The conduction of electricity through concrete. *Magazine of concrete research*, *33*(114), 48-60.

[64] Colla, C., Das, P. C., McCann, D., & Forde, M. C. (1997). Sonic, electromagnetic and impulse radar investigation of stone masonry bridges. *Ndt & E International*, *30*(4), 249-254.

[65] Polder, R., Andrade, C., Elsener, B., Vennesland, Gulikers, J., Weidert, R., & Raupach, M. (2000). RILEM TC 154-EMC: electrochemical techniques for measuring metallic corrosion. *Materials and structures*, *33*, 603-611.

[66] Ryan, E., Burdette, E., Ankabrandt, R., Nidiffer, R., & Buchanan, B. (2013). Comparison of two methods to assess the resistance of concrete to chloride ion penetration. *Journal of Materials in Civil Engineering*, *26*(4), 698-704.

[67] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278-2324.

[68] Beckman, G. H., Polyzois, D., & Cha, Y. J. (2019). Deep learning-based automatic volumetric damage quantification using depth camera. *Automation in Construction*, *99*, 114-124.

[69] Malhotra, V. M., & Carino, N. J. (Eds.). (1991). *CRC handbook on nondestructive testing of concrete*. CRC press.

[70] Suzuki, T., Ogata, H., Takada, R., Aoki, M., & Ohtsu, M. (2010). Use of acoustic emission and X-ray computed tomography for damage evaluation of freeze-thawed concrete. *Construction and Building Materials*, *24*(12), 2347-2352.

[71] Maldague, X. (2001). Theory and practice of infrared technology for nondestructive testing.

[72] Stanley, C. C., & Balendran, R. V. (1994). Developments in assessing the structural integrity of applied surfaces to concrete buildings. *Structural Survey*, *12*(4), 4-9.

[73] Ali, R. Cha, Y. J. (2018). Subsurface damage detection of a steel bridge using deep learning and uncooled micro-bolometer. *Construction and Building Materials. (Submitted)*

[74] Z. Lin, F. Azarmi, Q. Alkaseasbeh, M. Azimi, F. Yan, (2014). Advanced Ultrasonic Testing Technologies with Applications to Evaluation of Steel Bridge Welding–an Overview, *Applied Mechanics & Materials*, https://doi.org/10.4028/www.scientific.net/amm.727–728. 785.

[75] PUNDIT (Portable Ultrasonic Non-destructive Digital Indicating Tester) UPV, Available at: http://www.utest.com.tr/en/23499/Pundit-Lab-Ultrasonic-Pulse-Velocity-Tester, (2010).

[76] Golden Software Surfer 15.5.382x64 (2018). Available at: http://getintopc.com/softwares/graphic-design/golden-software-surfer-15-5-382-x64-free-download/.

[77] Yang, C. S., Kao, S. P., Lee, F. B., & Hung, P. S. (2004). Twelve different interpolation methods: A case study of Surfer 8.0. In *Proceedings of the XXth ISPRS Congress* (Vol. 35, pp. 778-785).

[78] A. Krizhevsky, I. Sutskever, G. E. Hinton, (2012). Imagenet classification with deep convolutional neural networks, In *Advances in neural information processing systems*, 1097–1105, https://doi.org/10.1145/3065386.

[79] LeCun, Y., Kavukcuoglu, K., & Farabet, C. (2010). Convolutional networks and applications in vision. In *Proceedings of 2010 IEEE International Symposium on Circuits and Systems* (pp. 253-256). IEEE.

[80] Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., ... & Chen, T. (2018). Recent advances in convolutional neural networks. *Pattern Recognition*, *77*, 354-377.

[81] Tensor flow/hub Available at: https://github.com/tensorflow/hub/blob/master/docs/tutorials/image_retraining.md.

[82] Choromanska, A., Henaff, M., Mathieu, M., Arous, G. B., & LeCun, Y. (2015). The loss surfaces of multilayer networks. In *Artificial Intelligence and Statistics* (pp. 192-204).

[83] Wiecek, B. (2006). Review on thermal image processing for passive and active thermography. In *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference* (pp. 686-689). IEEE.

[84] Ibarra-Castanedo, C., Galmiche, F., Darabi, A., Pilla, M., Klein, M., Ziadi, A., ... & Maldague, X. P. (2003). Thermographic nondestructive evaluation: overview of recent progress. In *Thermosense XXV* (Vol. 5073, pp. 450-460). International Society for Optics and Photonics.

[85] Hung, Y. Y., Chen, Y. S., Ng, S. P., Liu, L., Huang, Y. H., Luk, B. L., ... & Chung, P. S. (2009). Review and comparison of shearography and active thermography for nondestructive evaluation. *Materials Science and Engineering: R: Reports*, *64*(5-6), 73-112.

[86] Gucunski, N., Nazarian, S., Yuan, D. and Kutrubes, D. (2013). Nondestructive Testing to Identify Concrete Bridge Deck Deterioration. Transportation Research Board, SHRP 2 Report S2-R06A-RR-1, Washington, D.C., USA.

[87] Washer, G., Fenwick, R., and Harper, J. (2009). Effects of Environmental Variables on Infrared Imaging of Subsurface Features in Concrete Bridges. *J. of the Transportation Research Board*. Vol. 2108/2109, pp: 107-114.

[88] ASTM D4788–03, (2007). Standard test method for detecting delaminations in bridge decks using infrared thermography. American Society of Testing Materials.

[89] Purcell, E. M., & Morin, D. J. (2013). *Electricity and magnetism*. Cambridge University Press.

[90] Mehta, A. (2011). Introduction to the electromagnetic spectrum and spectroscopy. Pharmaxchange. Available at: *http://pharmaxchange. info/press/2011/08/introduction-to-the-electromagnetic-spectrum-and-spectroscopy/. cit*, *25*.

[91] FLIR thermal camera (2018). Available at: https://www.flir.com/products/t650sc/.

APPENDIX A: Details input and output layers of DINN

Table 4. Size of input layers [73]

Layers	Type	Patch size/stride	Input size
1	conv	3 × 3/2	299 × 299 × 3
2	conv1	3 × 3/1	149 × 149 × 32
3	conv2	3 × 3/1	147 × 147 × 32
4	Pool	3 × 3/2	147 × 147 × 64
5	conv3	1 × 1/1	73 × 73 × 64
6	conv4	3 × 3/1	73 × 73 × 80
7	pool1	3 × 3/2	71 × 71 × 192
8	mixed_a	–	35 × 35 × 192
9	mixed_a1	–	35 × 35 × 256
10	mixed_a2	–	35 × 35 × 288
11	mixed_b	–	35 × 35 × 288
12	mixed_c	–	17 × 17 × 768
13	mixed_c1	–	17 × 17 × 768
14	mixed_c2	–	17 × 17 × 768
15	mixed_c3	–	17 × 17 × 768
16	mixed_d	–	17 × 17 × 768
17	mixed_e	–	8 × 8 × 1280
18	mixed_e1	–	8 × 8 × 2048
19	pool2	8 × 8/1	8 × 8 × 2048

Table 5. Size of output layers [73]

Layers	Type	Patch size/stride	Output size
1	conv	3 × 3/2	149 × 149 × 32
2	conv1	3 × 3/1	147 × 147 × 32
3	conv2	3 × 3/1	147 × 147 × 64
4	Pool	3 × 3/2	73 × 73 × 64
5	conv3	1 × 1/1	73 × 73 × 80
6	conv4	3 × 3/1	71 × 71 × 192
7	pool1	3 × 3/2	35 × 35 × 192
8	mixed_a	–	35 × 35 × 256
9	mixed_a1	–	35 × 35 × 288
10	mixed_a2	–	35 × 35 × 288
11	mixed_b	–	17 × 17 × 768
12	mixed_c	–	17 × 17 × 768
13	mixed_c1	–	17 × 17 × 768
14	mixed_c2	–	17 × 17 × 768
15	mixed_c3	–	17 × 17 × 768
16	mixed_d	–	8 × 8 × 1280
17	mixed_e	–	8 × 8 × 2048
18	mixed_e1	–	8 × 8 × 2048
19	pool2	8 × 8/1	1 × 1 × 2048