Estimation of concrete strength based on deep learning-based image segmentation coupled with IR thermal extraction

Graduate School of Chosun University

Department of Architectural Engineering

Major: Building and Material Engineering

Minwuye Mesfin, Woldeamanuel

Estimation of concrete strength based on deep learning-based image segmentation coupled with IR thermal extraction

심층학습 기반 화상 분할과 연동된 열화상 정보를 활용한 콘크리트의 강도예측

25, February 2022

Graduate School of Chosun University

Department of Architectural Engineering

Major: Building and Material Engineering

Minwuye Mesfin, Woldeamanuel

Estimation of concrete strength based on deep learning-based image segmentation coupled with IR thermal extraction

Advisor: Professor Hyeong-Ki Kim

This thesis is submitted to graduate School of Chosun University in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In

Architectural Engineering

October 2021

Graduate School of Chosun University

Department of Architectural Engineering

Major: Building and Material Engineering

Minwuye Mesfin, Woldeamanuel

The Master of Science thesis committee for Minwuye Mesfin Woldeamanuel certifies that this is the approved version

Committee:

Professor Chang-geun Cho, Chosun University   (sign)
(Chair Committee)

Professor Taehoon Kim, Chosun University   (sign)
(Committee Member)

Professor, Hyeong-Ki Kim, Chosun University   (sign)
(Committee Member)

December 2021

Graduate School of Chosun University

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

On estimation of concrete strength based on deep learning-based image segmentation coupled with IR thermal extraction

Minwuye Mesfin Woldeamanuel, B.Sc.

Advisor: Prof. Hyeong-Ki, Kim, PhD.

Department of Architectural Engineering

Graduate School of Chosun University

In this research, a method was proposed that is used to predict the strength of concrete at any time interval. This method is composed of two major schemes. The first one is deep learning-based concrete image segmentation technique which employs state-of-the-art neural network architectures to accurately classify the concrete area from any captured image. Segmentation models with a convolutional neural network (CNN) backbone were trained, tested and evaluated by varying different hyperparameters with in the training environment to come up with a model that best fit the target of the research. Then, this method together with thermal extraction was used to predict the strength of the concrete. The second scheme is the application of a thermal imaging (IR) camera to detect the concrete's surface temperature. The method's reliability was first investigated by varying some parameters. Then the IR camera is used on the construction site to collect concrete temperature data from the segmented image. Maturity method was then used to interpret the relationship between temperature, age, and strength of concrete.

The result shows that the method can estimate the strength of concrete at any time. In addition to this, it is also very helpful in increasing the efficiency of construction projects in regards to both time and economic standpoint.

한 글 요 약

심층학습 기반 화상 분할과 연동된 열화상 정보를 활용한

콘크리트의 강도예측

Minwuye Mesfin Woldeamanuel, B.Sc.

Advisor: Prof. Hyeong-Ki, Kim, PhD.

Department of Architectural Engineering

Graduate School of Chosun University

본 연구에서는 임의의 시간 간격으로 콘크리트의 강도를 예측하는 방법을 제안하였다. 이 방법은 크게 두 가지 방식으로 구성된다. 첫 번째는 촬용된 모든 이미지에 대해 신경망 아키텍처를 사용하여 콘크리트 영역을 정확하게 분류하는 심화학습 기반의 이미지 분할 기술이다. 연구 대상에 가장 적합한 CNN (Convolutional Neural Network) 백본 모델을 찾기 위해 여러 환경에서 다양한 가상의 계수를 사용하여 훈련, 테스트 및 평가되었습니다. 그런 다음 이 방법을 열적 정보 추출과 함께 사용하여 콘크리트의 강도를 예측했다. 열화상 카메라를 적용하여 콘크리트 표면 온도를 감지하였다. 방법의 신뢰성은 일부 환경변수를 변경하여 사전에 조사하였다. 그런 다음 건설 현장에서 열화상 카메라를 사용하여 분할된 이미지에서 구체적인 온도 데이터를 수집하였다. 이를 성숙도(Maturity)식을 사용하여 콘크리트의 온도, 재령, 강도 간의 관계를 해석했다. 실험결과 이 방법이 콘크리트의 강도를 범용적으로 추정할 수 있음을 확인하였다. 뿐만 아니라 시간적, 경제적 측면에서도 공사의 효율성을 높이는 데 큰 도움이 될 수 있음을 알 수 있었다.

1. INTRODUCTION

1.1.　Background

In this age, where concrete is the core material used for construction over steel and other related materials, environmentally sustainable, mechanically stable, and economically affordable concrete is a significant concern. Especially the close ties it created with society and its environment makes it a center of attention for perfection. With this situation comes the continuous development of techniques to improve every bit of the variables encountered in the industry. This paved the way to find different alternatives and methods to improve its quality and reliability. In the past century, various research has been made to develop easy, reliable, and effective ways to advance the construction industry. The studies involve the design and construction phase up to the maintenance and inception by incorporating different techniques to reach a better quality of concrete. Most of them focus on using different by-products and supplementary materials in the construction phase, mix proportioning optimization methods, and detailed microstructure investigations (Nandhini & Karthikeyan, 2021). These areas are significant and influence the output of any concrete construction industry. The construction industry is changing faster than anticipated, not just in quantity and quality but also in technological advancement. Smart and advanced technologies are being applied material-wise, optimizing scheduling, labor and machinery output, monitoring, maintenance, inspection, and overall cost minimization. Therefore, both the materials and the methodologies have to be compatible with the construction work because these are all aspects that come down to when selecting any construction project.

Deep learning is one of these recently applied techniques getting much recognition due to its reliable output in most of the used areas. It is a part of AI, more specifically machine learning, which uses mathematical models called neural networks to learn and then predict the outcome

of a given data. This method has several applications ranging from concrete defect and failure detection to construction cost prediction, workforce optimization models, and even more so regarding safety protocols. This explains how vast and deep the usage runs in the construction sector.

One of the subdivisions of deep learning applications is using neural networks, especially Convolutional Neural Networks (CNNs) in image classification and segmentation. CNNs are being used in a number of aspects in the construction industry, one of which is assisting in predicting the strength of concrete(Jang, Ahn, & Kim, 2019; Lee, Ahn, & Kim, 2020). However, there are still some drawbacks to these methods. The first one is that these methods use laboratory concrete samples, which entirely contradicts the point of automatic concrete strength and condition monitoring. They used microscopical and standard images of the concrete samples and used CNN model to predict the strength of concrete. They achieved a good result, but there was still a lack of applicability in the construction sites. The second drawback and reason for this investigation is the advantage of semantic segmentation. Semantic segmentation is one of the advanced image classification techniques with higher accuracy than other image classifications. Therefore, this method was implemented as a concrete area classification technique. The other reason is the advantage of adopting a temperature-based model, which gives another set of use in monitoring the concrete. These and others were the main reasons for this method's investigation and development.

The other technique often used as a testing method is Infrared Thermography (IRT), also called thermal imaging. It is a powerful and one of the non-destructive testing techniques used to investigate various experiments in many fields. It is a technique based on gathering thermal information from any object to assess any variations and defects. Traditionally thermal sensors were being used to conduct these tests, but the non-destructive and non-contact behavior makes it the easiest choice. In the construction industry, problems such as Moisture damage, corrosion

of reinforcement bars, cracks, and overall structural distress are areas that IRT has been applied as an identifying method.

Therefore, in this research, the two methods mentioned above are used in combination to develop a model that predicts the strength of concrete at a given time. This is an essential point because knowing the strength of concrete at any time, especially in the early ages, is crucial in terms of time management, work schedule, and cost reduction. Furthermore, this method can be applied in any concrete construction, making it versatile and reliable.

1.2. Objective of the research

The main objective of this study is to:

- Apply a deep learning-based image segmentation model that can effectively segment the concrete area from an image for further applications,

- Investigate the reliability of thermal imaging (IR) cameras in extracting concrete's surface temperature and how it can be applied in the construction site,

- Develop an integrated network (a method) that comprises deep learning and maturity models that can effectively predict the in-place strength of concrete,

- Compare the applicability and reliability of this proposed model with other sensors, and

- Identify how the proposed model can be used to monitor the condition of concrete and optimize the output of construction concerning formwork striking time.

1.3. Outline of the research

This research comprises seven chapters. The first chapter will briefly describe the general introduction of the study. It mainly focuses on the overview of the two methods used in this research and their importance in the construction industry. The objectives and the general outline of the research are also presented in this section.

Chapter two will cover deep learning and some of the known applications. It will describe deep learning, its components, and how it functions. Several neural networks are presented, and Convolution neural networks (CNN) and Semantic Segmentation are also emphasized in this section.

Chapter three will be the continuation of the previous chapter, where experiments were done to develop the model for concrete area prediction. It will briefly describe the methodology and the steps taken to develop our concrete area recognition model. The results and discussion were also presented along with visual and graphical presentations and examples.

Chapter four will investigate the concept of infrared thermography (IRT). It will cover the basic features of IRT and the main factors that influence the output of temperature data from infrared images. The reliability of thermal imaging cameras in capturing the surface temperature of concrete was also done in this chapter by doing some experiments. Different examples were also presented in this chapter to explain the effect of some of the investigated variables.

Chapter five will cover the experiments done on the integrated method of the two schemes: deep learning-based image segmentation and thermal imaging. First, it will introduce a concept called maturity of concrete by mentioning some previous works. Then, it will also discuss the evolution of the method and how it will function as an essential tool for estimating the mechanical properties of the concrete. Then, laboratory and field experiments are included in this section to check the applicability and reliability of the proposed method. Comparison with other methods was also the other point in this chapter.

Finally, chapter six will conclude the key takeaways and the significant points of the research. It will also recommend some ideas and suggest a plan for further research.

2. DEEP LEARNING: A LITERATURE REVIEW

2.1. Introduction

The way we process and make the best decision has always been a unique point that explains what a human being is and separates us from other living things. This is similarly noticed in artificial intelligence (AI) technologies, especially machine learning methods, which are gaining more attention due to their advanced perceptive and problem-solving capabilities. These days AI runs into every aspect of society, from using different search engines, pattern recognition, image processing, and recognition to advanced techniques such as autonomous driving by making life easy and accessible. One of the principal subdivisions of AI is Machine learning (ML), where data analysis and statistical methods are applied to automate and improve any system. It is a concept drawn where data can learn and adapt to make decisions without external human intervention. Traditionally, problems were dealt with more or less in a primeval way where only human beings are the only solutions.

Nevertheless, time is changing, and different approaches are being created to consider easy solutions, ML being the top of the list. In addition to this, the emergence of advanced computing technologies, high-rated GPUs, and the availability and access to data makes ML, especially Deep Learning (DL), very convenient to apply even beyond the scope of computer-based applications. Currently, every field of study in some way applies ML and DL to solve problems. During this research, the focus will be on deep learning and how deep learning-based methods are essential in solving problems related to construction. Deep learning is one of the few subdivisions of ML and AI, which mainly depend on networks to perform tasks. Figure 1 shows the hierarchy of AI, ML, and DL. Usually, there is confusion between supervised learning and

DL because they both have the same concept. They both use training data and previously known information to predict the outcome. Supervised learning primarily performs classification and regression models, but DL involves much deeper and more data-intensive methods such as localization and segmentation. In addition to this, DL requires an extensive amount of data and uses neural networks and models, making it the ideal choice for solving these complex issues. Therefore, this chapter will briefly introduce deep learning and its derivatives. It will also cover the specific models used in this research along with these advantages and limitations.
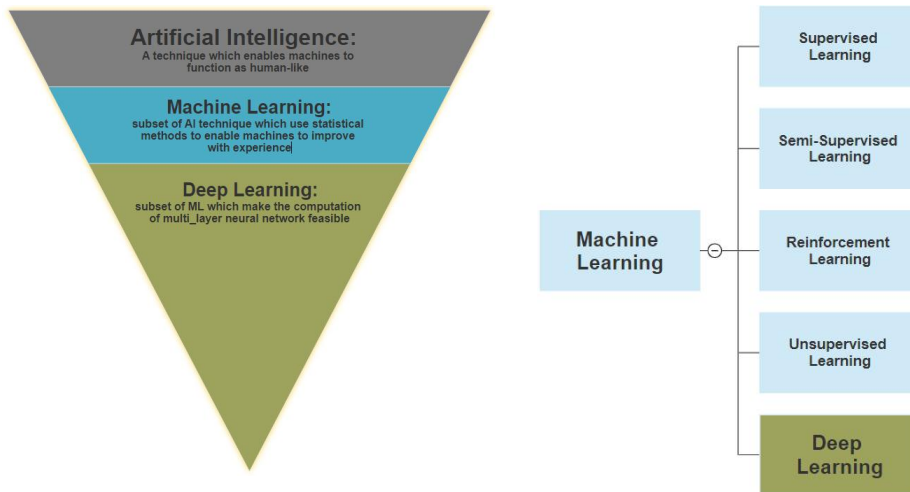
Figure 1: Hierarchy of AI, ML, and DL (redrawn from (Alom et al., 2019; Nassif, Shahin, Attili, Azzeh, & Shaalan, 2019))

2.2. Deep Learning (DL)

Deep Learning is a machine learning division focused on the notion of how a system will learn and interpret its surroundings based on previously existing data. It is an algorithm that learns tasks directly from data and then uses that knowledge to build and forecast models for a wide range of applications. The entire concept of deep learning is built on the human brain's perception and how it can identify anything based on various parameters. It is the process through which data passes through millions of neurons (nodes) to connect the dots and arrive at a better and

more precise result. It is a comprehensive machine learning technique that incorporates neural networks in successive layers to learn from data iteratively.

Neuron, also called Perceptron (Rosenblatt, 1958), is the building block of deep learning. It is a single network that includes the input layer, its weights, multiple functions, and the output layer. Figure 2 shows a typical perceptron that includes the essential components and their connection. X_i and Y represent the number of features in the input data and the output, respectively, W_i, stands for the weights of each input feature, W_0 is the bias used to shift the result of the functions, and F is the non-linear (activation) function. Even though these are the basic components, other layers are sometimes involved depending on the model's purpose. When multiple perceptrons are connected, it will result in a neural network which can be seen in Figure 3. The inside layers labeled as hidden will determine how deep the neural network is and even the type of neural network itself.

$$Y = F\left(W_o + \sum_{i=1}^{n} X_i W_i\right)$$
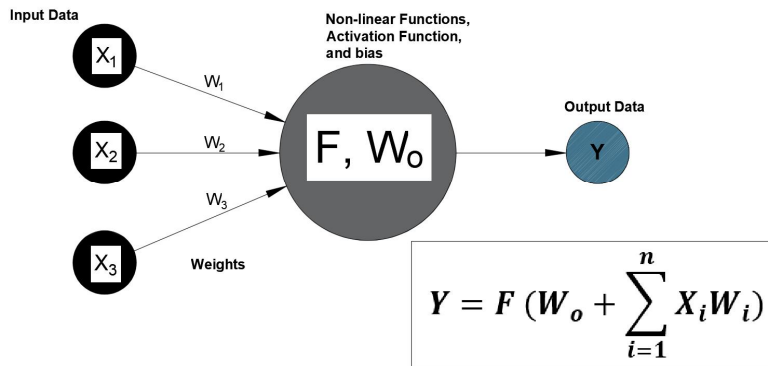
Figure 2: An illustration of a perceptron (redrawn from(Liu, Deng, & Yang, 2019))
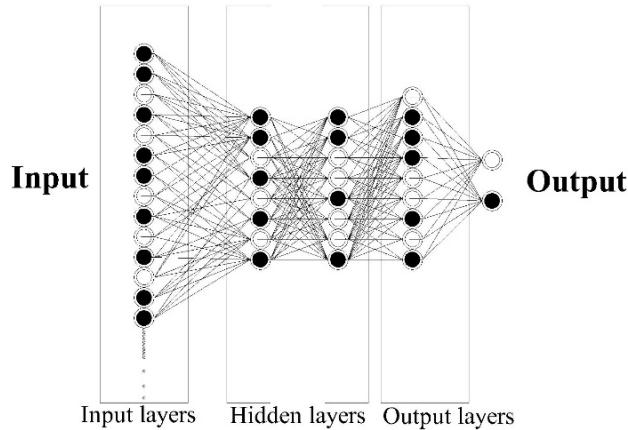
Figure 3: An illustration of a typical neural network

2.3. Deep Learning Neural Networks

From the emergence of the deep learning concept, many networks and architectures have been developed for various data types and applications. These networks are composed of multiple layers with different distinguishing features and parameters. Their purpose varies from simple binary or digits recognition to simulation and autonomous driving (Garcia-Garcia, Orts-Escolano, Oprea, Villena-Martinez, & Garcia-Rodriguez, 2017), which requires models that have a complicated structure. These models can be used by themselves or in combination with others to boost their performance and applicability. This section will explain some of the conventional and frequently used deep neural networks with their key features and differences.

2.3.1. Artificial Neural Networks (ANNs)

One of the most common and frequently used neural networks is called Artificial Neural Network (ANN), also called Feed-Forward Neural Network. This network is the stepping-stone for the other advanced networks. It consists of an input layer, a hidden layer, and an output layer. The inner and outer layers are passive layers that do not take part in the basic learning process of the network. The inner layers provide the data needed for the active layers, hidden layers, to exploit

and construct the output through the output layers. The data can be any type, such as an image, numbers, or video, as long as it can be representable in some format. The hidden layers are the place where real learning takes place. They comprise the different functions and parameters necessary to find patterns from the input data and transfer them to the output layer. Figure 3 in the previous section can be a good example of ANN. Although no particular range or quantity of layers govern a network's depth, increasing the number of hidden layers and linking functions allows the ANN to be referred to as a 'Deep' ANN, or Deep neural network.

2.3.2. *Recurrent Neural Networks (RNNs)*

Recurrent neural networks (RNNs) are models developed to process and interpret sequential data such as speech, videos, or time-dependent variables (Akinosho et al., 2020). In RNN models, the time variable is crucial since the output is determined not only by the operations of the functions directed to it but also by the occurrence of the previous version. This signifies that the sequence and result of previous data impact the current output, distinguishing it from other types of models. The hidden layers in RNN use the previously hidden layer as an input and the input data to predict the subsequent output. This function acts as a sort of loop where the previous result will determine the next output. Figure 4 shows the basic network arrangement of a typical RNN. However, when the sequential data is substantial, the model has difficulty generalizing its results, but it is still the best option for handling time and position-related problems.
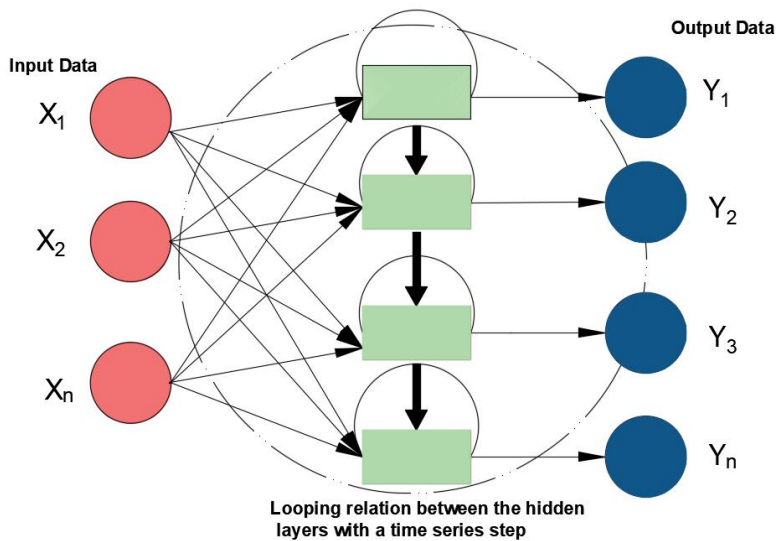
Figure 4: A typical RNN architecture (redrawn from(Alom et al., 2019))

2.3.3.   *Generative Adversarial Networks (GANs)*

Generative Adversarial Network (GAN) was first proposed by Goodfellow et al. (Goodfellow et al., 2014) in 2014. It is a new approach based on improving the accuracy and predictive capability by learning from comparing real and false data sets. As the name "adversarial" suggests, it is based on analyzing and comparing real and fake data to adapt and improve the model's accuracy. This model consists of two parts: generator and discriminator. Generator maps noised false data that resembles the real data input. The discriminator is a filtering layer that distinguishes between the generated false data with the actual data and represents it as a cost/loss function. Figure 5 shows the connection between these two parts. The model is set up to learn which one is fake or real and generate another set of data. GAN is usually applied to image-based deep learning techniques. It is highly preferred when the labeled image data is considerably low. Many studies found this network crucial in image generation and dataset creation (Radford, Metz, & Chintala, 2015) for deep learning architectures. It is also applied in combination with CNNs for image classification with a limited dataset.
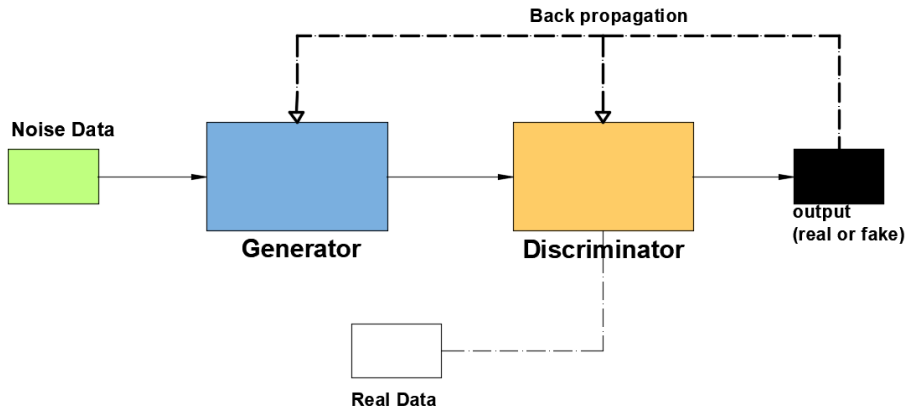
Figure 5: A typical GAN architecture with backpropagation (redrawn from(Feng et al., 2020))

2.3.4.   Convolutional Neural Networks (CNNs)

Convolutional neural networks (CNN, ConvNet) (Yu Songa, 2020) are a class of deep, one-directional artificial neural networks applied to analyzing visual imagery. They are an interconnected assembly of simple processing elements, units, or nodes, whose processing ability of the network is stored in the interunit connection strengths or weights, obtained by adapting to, or learning from, a set of training patterns. The detailed structure of CNN consists of convolutional layers, pooling layers, regularization layers, dropout layers, and other specific and targeted hidden layers. Figure 6 below shows a typical CNN network orientation which includes the major layers that define the network. Therefore, these layers mentioned above are responsible for the whole function of CNNs. CNNs are the most established and deep-rooted networks that are usually applied in image processing and classification problems.

11
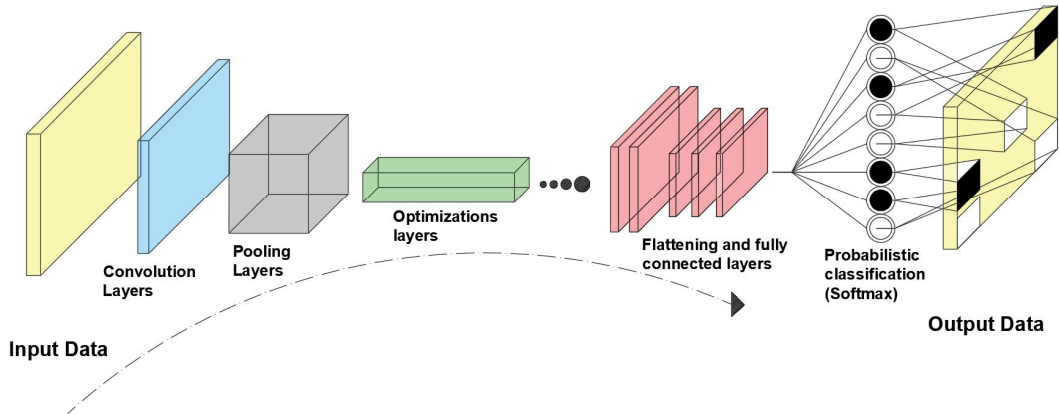
Figure 6: A typical CNN architecture (redrawn from(Tabian, Fu, & Sharif Khodaei, 2019))

Overall, the above-mentioned neural networks are the major ones frequently used in aspects involving deep learning methods. Of course, other models are also available, but due to the scope of this research, CNN models are dealt with in detail which is presented in the next section.

2.4.   Architecture of a CNN

Convolutional neural networks have grown popular in image classification, image segmentation, and object localization tasks. The availability of computational resources (mainly GPUs) and open access data pushed convolutional networks to improve whole-image classification and data segmentation techniques to a higher level. As a result, they are extensively used in several image and video processing studies. A great deal of investigation is done on CNN models due to the focus of this research, which utilizes images in performing segmentation for further application. As mentioned above, a typical CNN comprises layers with their function, optimizing the next layer that comes before it.

The primary layer and one of the building blocks of CNNs is a convolutional layer. This layer is a crucial section of CNN that separates it from other neural networks. It contains a set of filters

called kernels which extract a particular feature from the input data. They are composed of matrices of a small volume than the input. These filters do operations on the input data to create a feature map (Tabian et al., 2019). The height and weight of these filters are smaller than those of the input volume. Each filter is convolved with the input volume to compute an activation map made of neurons. Figure 7 shows the operation of a convolutional layer. They use a vector dot product operation to calculate an activation function using weights of the inputs and a bias value. The activation function determines the value of each pixel according to the distinct features that the convolution layer has. These filters can be fixed values such as Sobel and Scharr filters (Kaur & Kant, 2014) for edge detection where their values are known, or the network will adjust these filters by learning from every iteration. The second option is preferable, especially for deep networks, to increase the adaptability of the filter to the input that it uses and the outcome expected. Based on this, the new pixel set will have a more converged distinct and predictable pixel value. The more the convolutional layers are available, the deeper the network gets, the more features it will recognize and classify (Xincong Yang, 2018).
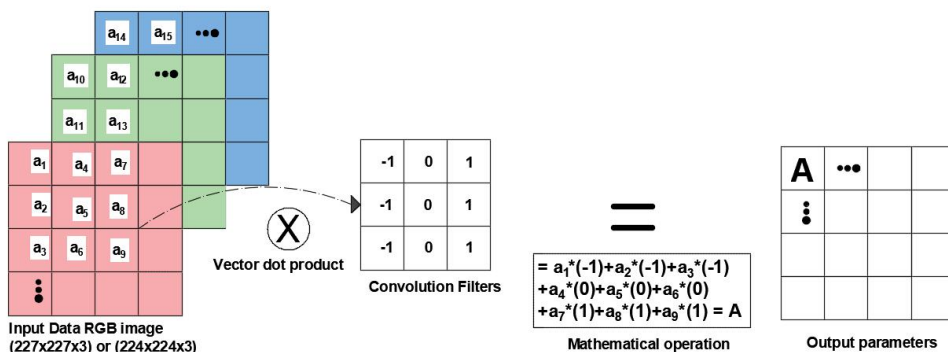
Figure 7: Convolution layer mathematical operation

After the computation is done under the convolution layer, a function is needed to normalize the calculated values. These functions are called activation functions. These are functions that introduce non-linearity to the input values and optimize the network's decision. Several functions

can perform this operation. The choice of these functions depends on the type of networks they are implemented on. Previously sigmoid, and hyperbolic functions (Szandała, 2021) have been used. But due to their complicated result and slow learning process, the rectified linear unit (ReLU) is currently the best option. ReLU function works as a thresholding function which puts the negative value to zero and the positive ones as their own value (Agarap, 2018; Chigozie Nwankpa, Nov, 2018). It is usually inserted next to the convolutional layer to optimize the result and speed up the training process. It is also applicable in other types of neural networks besides CNNs. Figure 8 shows some of the non-linear functions and their graphical representations, along with their equations. Modified versions of ReLU such as Leaky ReLU, PReLU activation functions are also available to use depending on the purpose and the quality of result expected.

$$\sigma(x) = \frac{1}{1+e^{-x}}$$

$$y = \begin{cases} 0, x \leq 0 \\ x, x > 0 \end{cases}$$

$$f(x) = tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

— Sigmoid Function

— · — Hyperbolic (Tanh) Function

— ReLU Function
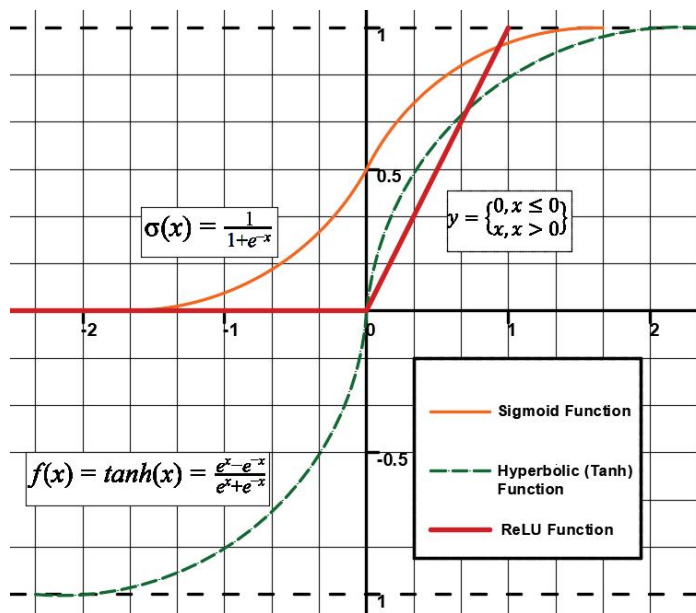
Figure 8: some of the often-used activation functions (redrawn from (Nwankpa, Ijomah, Gachagan, & Marshall, 2018))

ince the calculated input parameters are extensive, downsampling is necessary to control the complexity of the network. These sampling layers are called pooling layers. These layers are

responsible for the resizing of the data that was derived from the previous layers. They serve as a filtering mechanism for selecting the crucial data and discarding the rest. They have two essential purposes: reducing some layers to decrease the complexity and the computational time of the network, and the second is to avoid overfitting that is usually caused by irrelevant data. There are many types of pooling layers out of which the most popular and frequently used are max pooling (Ranzato, Boureau, & LeCun, 2007), average pooling (LeCun, Bottou, Bengio, & Haffner, 1998), mixed pooling (Yu, Wang, Chen, & Wei, 2014), and region of interest (ROI) pooling. As their name suggests, the max-pooling uses the maximum activation value from the data set, and the average pooling implements the mean value. Mixed pooling incorporates both average and max pooling simultaneously to downsample the activation values. ROI pooling is usually used in object detection and image segmentation, where it focuses on a specific ROI based on the expected output. Figure 9 shows the most common pooling layers and how there operate with the convolutional and non-linear layers.
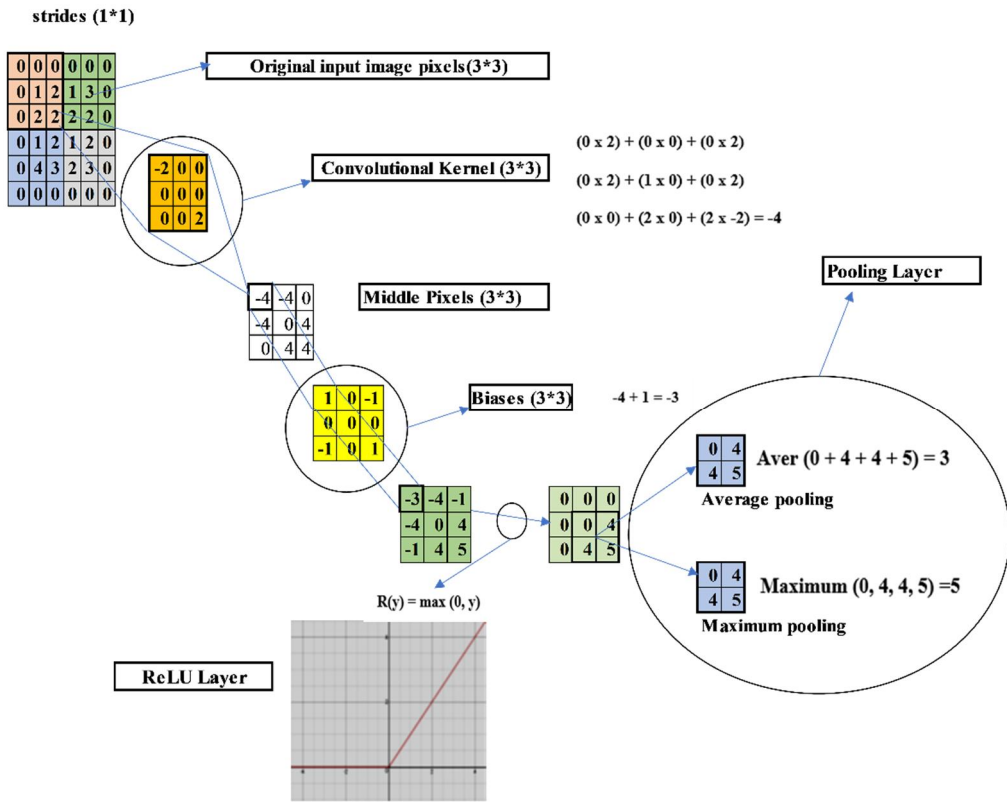
Figure 9: Convolutional and pooling layer operation example (redrawn from (Agarap, 2018; X. Yang et al., 2018) (Xincong Yang, 2018)).

Other layers such as dropout, normalization, and crop layers are also available depending on the CNN development. These layers are usually used to increase the CNN's optimization capacity and efficiency. For example, the dropout layer (Srivastava, Hinton, Krizhevsky, Sutskever, & Salakhutdinov, 2014) randomly cut layers coming to it with a given set of parameters to minimize overfitting and increase the generalization capacity of the network. It is one way of regularizing a network that helps to increase the model's accuracy. Then comes a fully connected layer used to connect the dots from the outputs of optimized activation functions and predict the outcome. It is a classification layer that analyses the results from all the previous parameters. It usually comes with a probabilistic activation function called the Softmax layer (Qiuyu Zhu, April 24,

16

2020). This layer is based on probabilistic distribution, which weights the result to classify data. It compares the result from the network with the groundTruth (gTruth) data used during the training stage and give a probabilistic value for input. Finally, the output/classification layer classifies based on the classes provided during training.

In general, the layers mentioned earlier are important in achieving an efficient, accurate, and precise CNN model. However, some of the layers might not be present in some models. The choice of selecting some of the layers also depends on how accurate the intended network is expected and the financial requirement and costs behind it.

2.5.   State-of-the-art CNN architectures

Since the introduction of the LeNet (Badrinarayanan, Kendall, & Cipolla) CNN architecture more than 20 years ago, CNN models have progressed into different structures and configurations. A groundbreaking change happened when Alexnet (Krizhevsky, Sutskever, & Hinton, 2012) was introduced a decade ago. Alexnet is a deep convolutional neural network architecture that was developed by Krizhevsky and his team which is used for image classification and localization purposes. This neural network has 60 million parameters and 650,000 neurons. It has five convolutional layers following max-pooling layers, batch normalization, and rectified linear units (ReLU). Three fully connected layers following a dropout layer are also connected at the end of the convolutional layers. Finally, a one thousand softmax classifier is put at the end to classify the input values. Figure 10 shows the general framework of the architecture. The development of this model advanced the scope of CNN and revolutionized the research for better options.

As a result, many CNN models emerged, each with its concept and improvements. Visual Geometry Group (VGG) (Simonyan & Zisserman, 2014) is one of these models that was developed and presented by Simonyan and Zisserman in 2015. It is a 19-layer version of Alexnet

that simulates the effect of raising a network's depth in order to maximize output. VGG is available in two versions: VGG 16, which has 13 convolution layers and three completely connected layers, and VGG 19, which has three additional convolution layers. It has also substituted the 11x11 and 5x5 convolution layer filter sizes in Alexnet to 3x3 filter sizes. This reduced the model's computing complexity by lowering the parameters. However, the total number of parameters is nearly 138 million, which is a disadvantage in processing time and cost. Figure 11 shows the architecture of VGG.

ResNet (He, Zhang, Ren, & Sun, 2016) is the other model that was developed by He et al. in 2015. This model changed the whole understanding of deep neural networks by introducing a concept called residual learning. The idea of this concept is that the next layer not only learns from the computation of the previous layer but also from the previous layer itself. This creates a skip connection of the input layer to the next layer, which widens the scope and depth of learning. Figure 12 shows the framework of a typical ResNet Architecture. This model has 152 layers where it can be modified in certain situations and depending on the user's expectation.

Even though it is a deeper network than the other above-mentioned models, it has a lesser computational complexity, making it time-efficient to train the model. ResNet also has different versions depending on the number of layers it has. ResNet 18/50/101/152 are frequently employed Architectures.
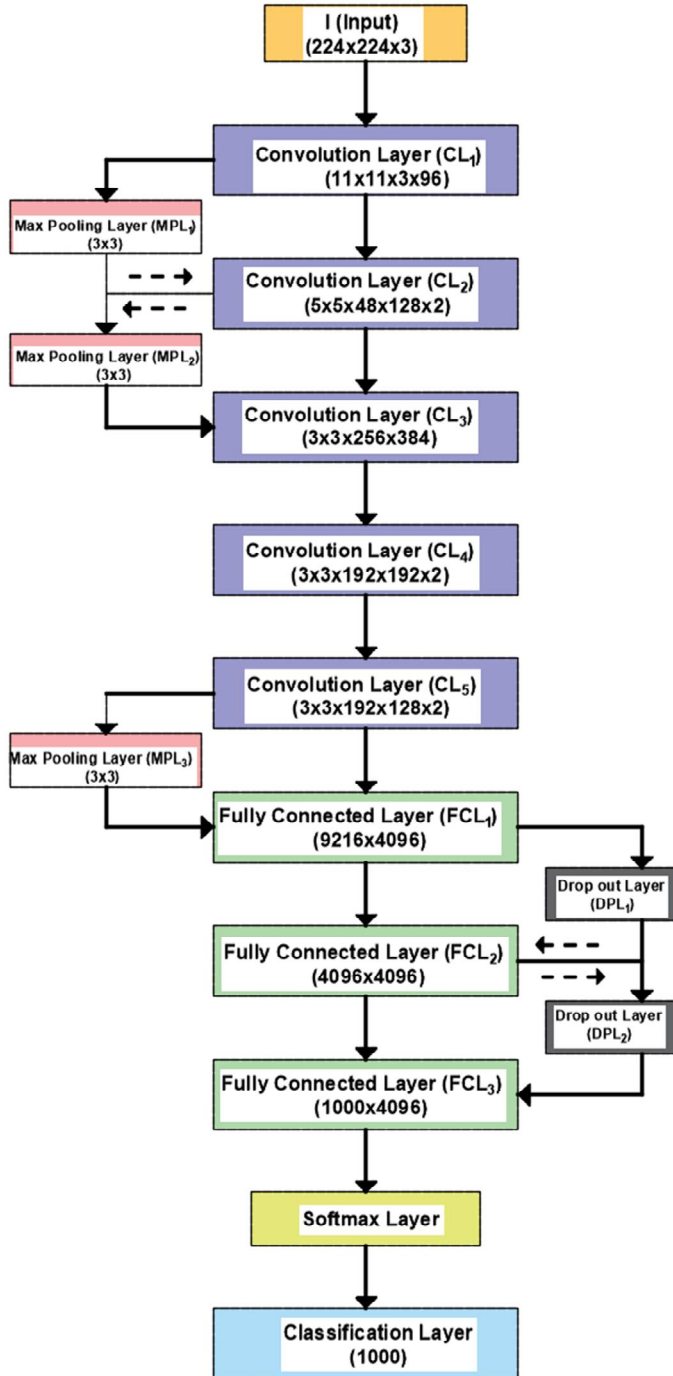
Figure 10: Frame work of Alexnet CNN architecture ((Krizhevsky et al., 2012))
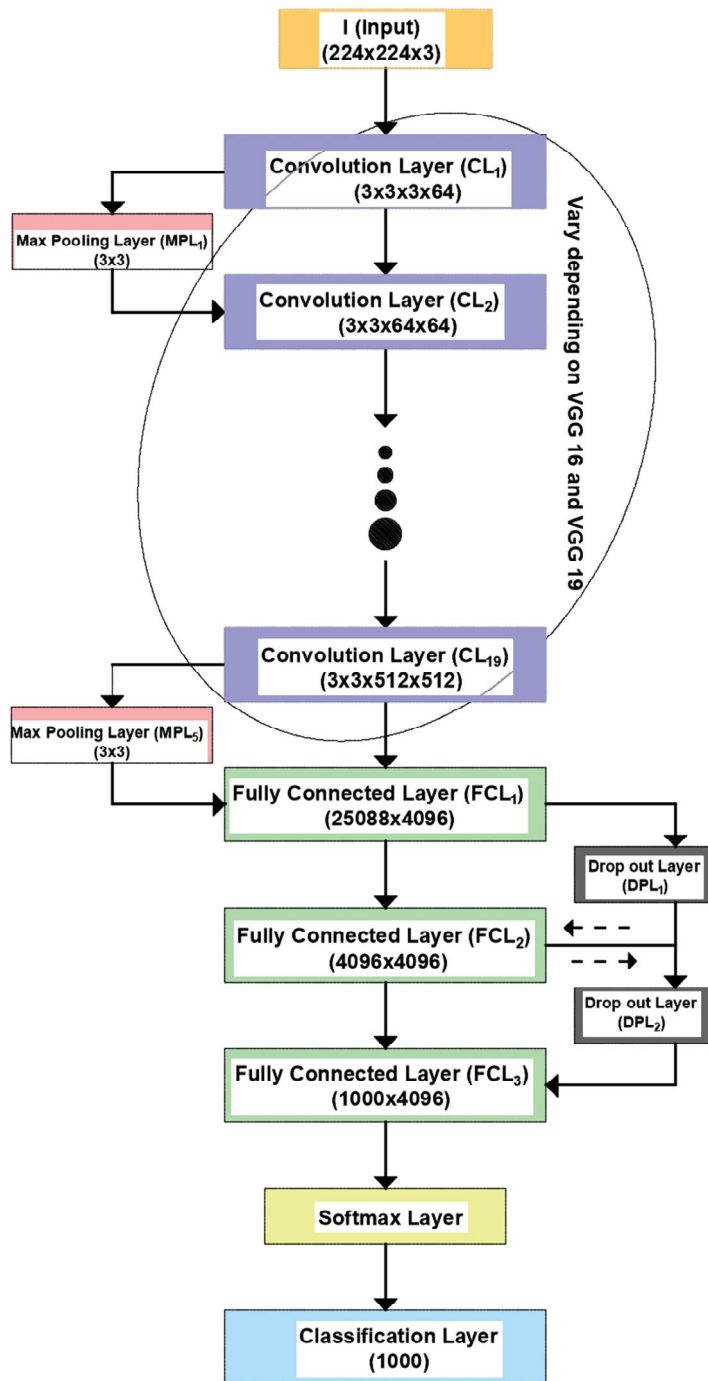
Figure 11: Frame work of VGG CNN architecture ((Simonyan & Zisserman, 2014))
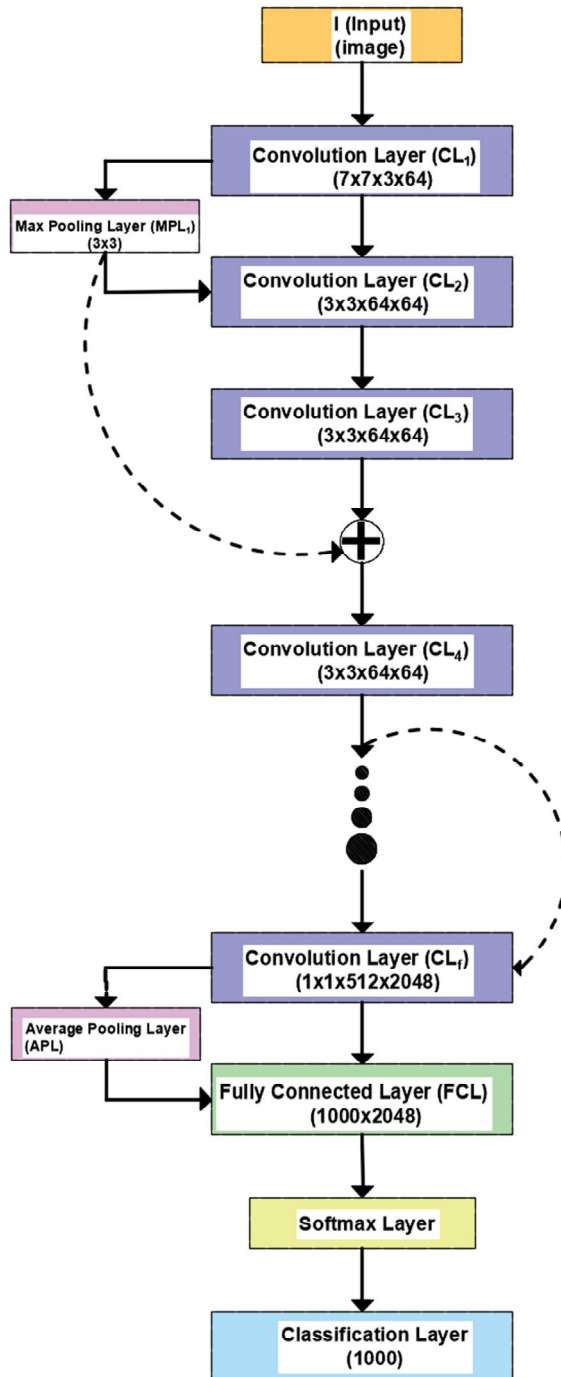
Figure 12: Frame work of ResNet CNN architecture ((He et al., 2016))

In general, the models stated above are just a few of the many CNN models available. These are the most widely used and have been validated on a variety of open-access datasets. As a result, they are perfect for a wide range of deep learning-based applications in various industries. Similarly, by modifying some of the parameters and layers in these CNN models, deep learning-based image processing was attempted in this study, which has proven to be a promising approach in the future.

2.6. Transfer Learning

Regardless of its form or volume, any data has a significant impact on the learning process. Data mining techniques have significantly evolved due to the internet's free access, which results in a massive volume of data collection. However, this does not imply that training a network with all the readily available data is trivial. Collecting, filtering, and labeling training data are time-consuming and labor-intensive. When networks are trained from the ground up, it can take years to achieve the desired result. Due to this, another approach is needed to elevate this obstacle.

Transfer learning has been the primary beneficiary of deep learning and segmentation works. It is extracting knowledge from previously classified tasks and applying it to the targeted classes (Huang, Pan, & Lei, 2017). It provides an effective way of training an extensive network using scarce training data without overfitting. This is very important in boosting the efficiency of the training time and the unnecessary architecture development for different purposes. In a clearer sense, it is easier to train an already trained or learnable model than to develop a new one.

The graphical illustration of how transfer learning works from a pre-trained model is shown in Figure 13. The word "pre-trained" refers to a network built from the ground up for a particular purpose. Alexnet, for example, is a pre-trained CNN model that has been trained using 1,000,000+ training datasets to identify 1000 classifications. Therefore, retraining this model with a smaller data set and using the knowledge from the pre-training stage is the idea behind

transfer learning. Sometimes this can be done by adjusting the last layer (Figure 13 Blue line) or multiple layers (Figure 13 red line) to fine-tune the network to boost performance. Transfer learning has made significant development over the last decade and a half on a wide range of applications in several fields. In fact, using this strategy is more effective than training models from scratch for a single target dataset (Farahani, Pourshojae, Rasheed, & Arabnia, 2021).
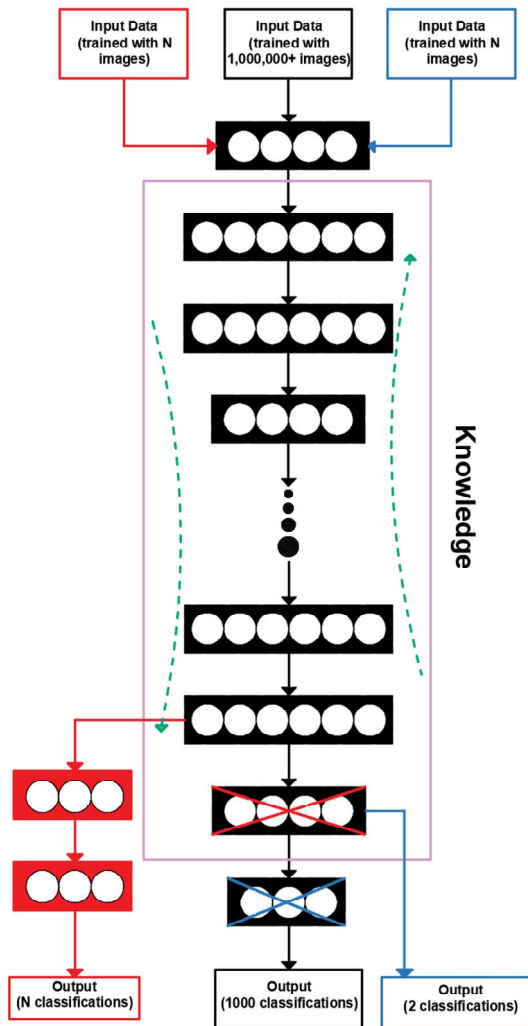
Figure 13: Graphical representation of Transfer Learning

2.7. Semantic Segmentation

Image processing has advanced these recent years from image recognition to deeper classification and localization techniques. Semantic segmentation was one of the most challenging issues to define during this stage. However, due to the exponential growth in computing capability and the development of deeper neural networks, the concept is at its highest investigation and implementation stage. Semantic image segmentation, also known as pixel-level image classification, is to group together sections of an image that correspond to the same object class (Liu et al., 2019). It is a way of categorizing every pixel on the image to a certain class that is previously determined during training. Semantic segmentation has various applications ranging from detecting regions of interest (ROI) in medical studies to autonomous driving. Figure 14 shows the difference between image classification, object localization, and semantic segmentation. The key point that separates semantic segmentation from the others is that it has a higher advantage of extracting the boundaries of objects or scenes better and paints a clear picture of the classes in an image. In addition to this, it is more detailed where every pixel of the images belongs to a certain class, making it more accurate and reliable.

In the past, image segmentation was a difficult task because the models available were based on the classification that considers the whole image, not the actual content. But, nowadays, different approaches are being developed, which are now beginning to investigate and tackle semantic segmentation. One of the milestones which changed the semantic segmentation techniques was Fully Convolution Network (FCN) (Long, Shelhamer, & Darrell, 2015). This network is built on the backbone of CNN architecture, with convolution layers adapting to upsample the prediction output instead of fully connected layers like in a traditional CNN. It is a pixel-to-pixel prediction in which deconvolution layers are used to upsample the result, extrapolate, and generate a dense per-pixel labeled output (Garcia-Garcia et al., 2017).

(a)
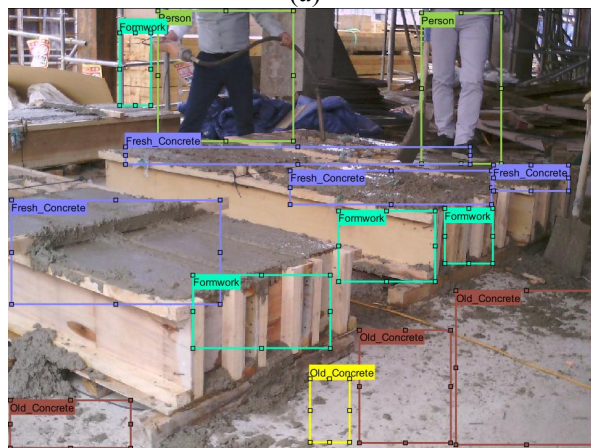
(b)
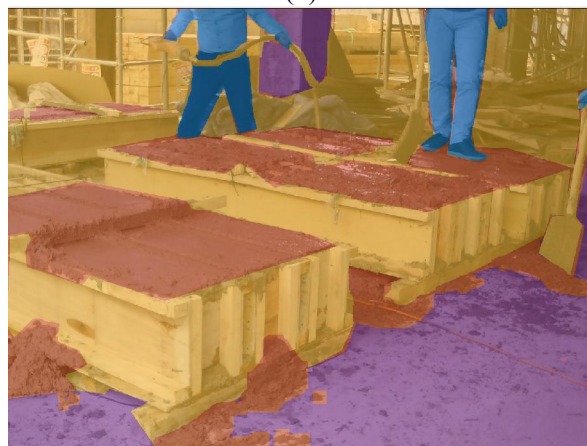
(c)

Figure 14 Comparison of image classification (a), object localization (b), and semantic segmentation (c).

This model involves fully convolutional models using well-known state-of-the-art CNN architectures such as Alexnet, VGG, ResNet, and others. They are then upsampled to segment the input data based on the classes given. Figure 15 shows a typical framework of FCN, which illustrates how the convolution and deconvolution layers are connected and by un-pooling how the segmented output is reconstructed.
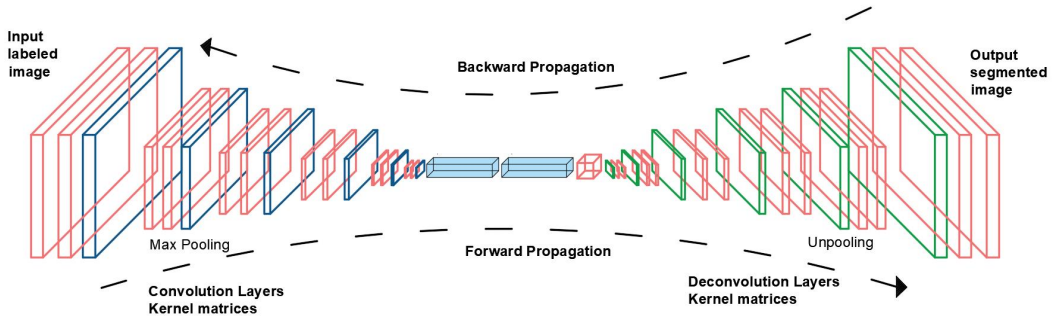
Figure 15: Framework of FCN (redrawn from(Long et al., 2015))

Even though FCN is the most often used model, other models also work well for semantic segmentation. Overall, the principle of removing fully connected layers and replacing them with convolution layers applies to both models. Segnet (Badrinarayanan et al., 2017) is one popular and widely used image segmentation designs. This model has an encoder-decoder type of architecture where the encoder is a convolutional network, usually VGG 16 and 19, without the fully connected layers. After every convolutional layer, it has two-by-two max-pooling layers on both encoder and decoder architectures. This helps pool and upsampling the features, making it distinctive compared to FCN. In the end, a softmax classifier is also used to classify the final output from the convolutional layers. Because Segnet comprises the decoder architecture, it is slower than FCN. It also requires less memory for both training and testing. In addition, the model is substantially smaller than FCN.
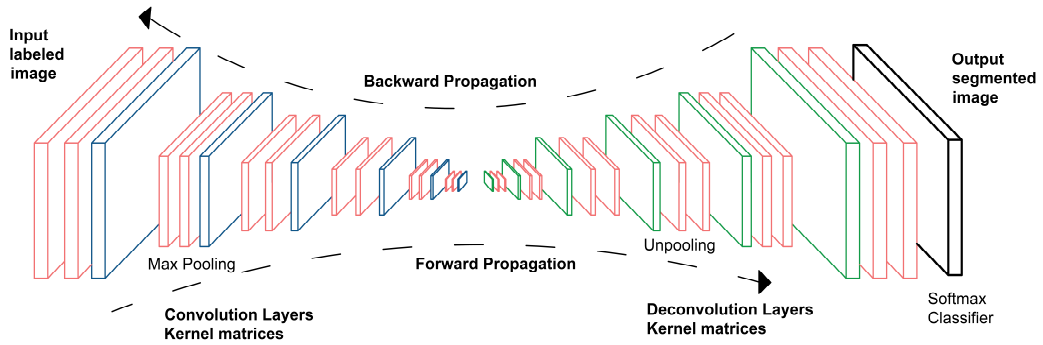
Figure 16: Framework of SegNet architecture (redrawn from(Badrinarayanan et al., 2017))

DeepLab (Chen, Papandreou, Kokkinos, Murphy, & Yuille, 2017) is also another impactful architecture in the image segmentation that has progressed from DeepLab1 to DeepLabv3+ (Chen, Zhu, Papandreou, Schroff, & Adam, 2018), which showed how robust the model is in maximizing the pixel label output. One of the issues under consideration was how the resolution would be affected when the up-sampling techniques are performed. Especially when the network is getting deeper and deeper, the resolution usually will be affected when trying to recover it. In FCN, a bi-linear interpolation was performed to up-sample the features. However, this resulted in the reduction of the resolution. For this purpose, a convolution and up-sampling method called atrous or dilated convolution was introduced on this model. This type of convolution inserts zeros between the elements in the filters to increase the field of view of the predicted image. This approach is a way to up-sample the feature maps without compromising the image's resolution. This increases the size of the convolutional kernels, which helps capture a large amount of context of that specific feature in the image without affecting the total number of parameters in the network. Figure 17 shows how the feature map is rated and up-sampled in atrous convolution. The equation in the figure is how the rate, r changes the size of the feature map. As a result of this and other changes in upsampling the image's features, its performance and accuracy have improved, making it the number one choice for image segmentation.

27

$$Y[i] = \sum_k (X(i + r.k)w[k]$$
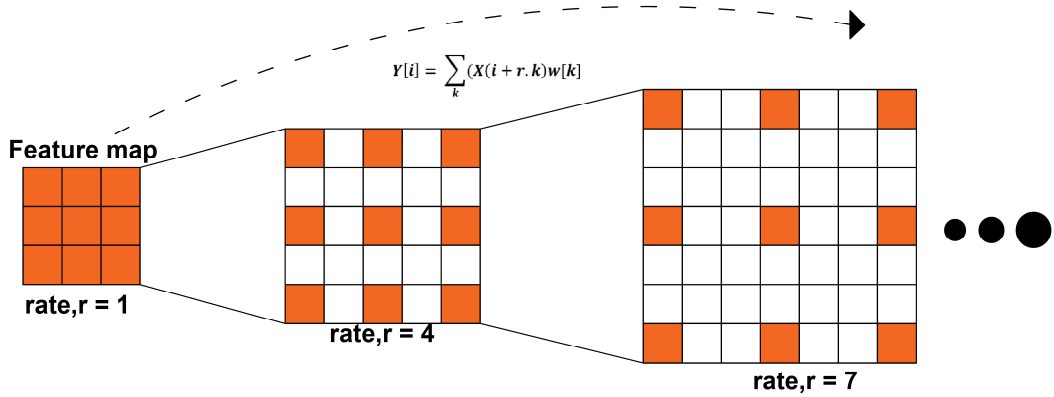
Figure 17: Atrous (dilated) convolution in DeepLab architecture (redrawn from (Chen et al., 2018)

The architectures outlined above and others are crucial in image segmentation and have made substantial progress in recent years. In addition, recent research, primarily based on deep learning approaches, has yielded groundbreaking results that have the potential to make significant contributions throughout the broad field of visual comprehension. However, there is still a lag, particularly in the construction business, in incorporating different challenges in the sector for solving and optimizing the output. Therefore, the primary goal of this study is to demonstrate how vital deep learning-based methodologies are in the construction industry.

3. APPLICATION OF DEEP LEARNING MODELS IN CONCRETE AREA SEGMENTATION[1]

3.1. Introduction

In this section, a model was developed that can detect and segment the area of a concrete section from an image. This was accomplished through the use of CNN-based segmentation models. Four segmentation models with CNN backbones were selected with different structures to investigate and compare how they identify and segment the area of a concrete section from an image. These models were trained with different images of construction sites that are labeled as "concrete" and "not concrete." These models passed into training hyperparameters and data augmentation techniques that helped achieve accurate and unbiased models. They are then tested and validated with statistical matrices and visual image representations. The main purpose of this research is to investigate the tendency of these architectures and their implementation to a database composed of a series of images that will objectively know what is concrete or not. On the other hand, it aims to determine the architecture that detects the concrete area well. It also allows CNN models to enter the construction sector, where they will be applied as an automated system to address problems both independently and in conjunction with other approaches.

3.2. Experimental Methodology

3.2.1. Data Acquisition and Labeling

Data collection is one of the crucial variables that must be carefully considered. This is because the network's output will highly depend on the type and number of data that it uses to learn. In

---

[1] Mesfin, W. M., Cho, S., Lee, J., Kim, H. K., & Kim, T. (2021). Deep-Learning-Based Segmentation of Fresh or Young Concrete Sections from Images of Construction Sites. *Materials*, *14*(21), 6311. *(Editor's choice)*

addition to this, the selected training data has to include all the possible scenarios that will probably occur when applying the model. Therefore, it is imperative to select the correct data type related to the study and cover every possible aspect. For this study, 1000 RGB images were gathered from different construction sites in Gwangju, South Korea. These images are at a different concrete casting stage, making it appropriate to investigate every possible way to develop a viable model. Figure 18 shows some of the selected images for training the model.
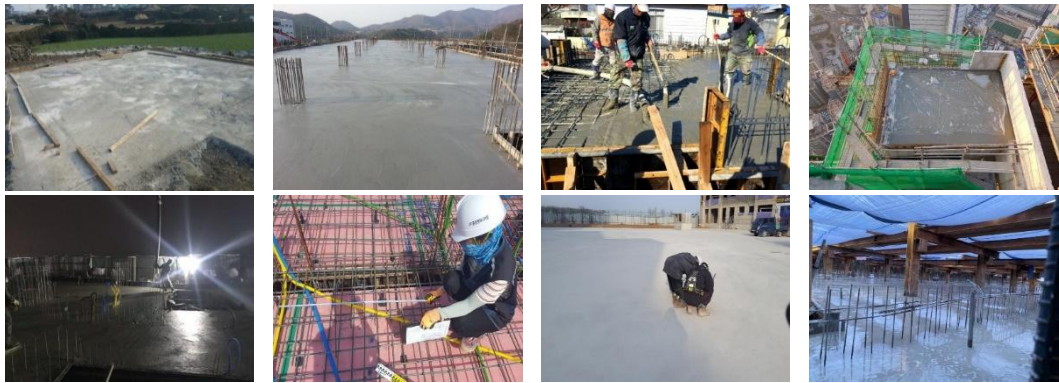
Figure 18: Some of the RGB images used for training

The images were of different sizes, which were then resized based on the requirement of the CNN model. According to the provisions of the CNN models, any image resolution size greater than 224 x 224-pixel size is viable as input data. In addition to this, the size of the thermal image is 240 x 320-pixel size. Therefore, the thermal image size was selected based on the two conditions, and all the images were resized. This procedure sometimes can be done while training the architecture simultaneously, which is a matter of preference. However, all the images were resized for this study before training them. Once the pixel adjustment was made, the images were labeled.

In image processing, data (image) labeling is a procedure of categorizing objects or scenes in an image. This is an essential step in acquiring a higher quality of training data set and influencing the outcome of the trained model. Different types of image labeling and annotations can be

applied for different applications. Bounding box labeling, line labeling, polygonal-based labeling, and pixel-based labeling are the ones that are frequently used. Bound box and polygonal-based labeling are usually used for object detection and image classification. They are used when the category being identified has a definitive shape and pattern that can be covered. Line labeling is ideally used for lane detection and recognition in autonomous driving. Pixel-based labeling is usually used for image segmentation and scene labeling. Because semantic segmentation is being attempted, pixel-based labeling was used for this investigation. This labeling approach assigns a category or class to each image pixel. As a result, it is more exact and accurate than other labeling methods. Furthermore, because concrete lacks a precise pattern and form to specify, it is more difficult to employ the other bounded labeling approaches.

Figure 19 shows some of the labeled images with different conditions. As it can be seen, the images are of different conditions and scenarios. This variation is necessary to train and develop a model that can adapt to different environments and situations. In addition, this will increase the model's reliability when it is introduced into another surrounding. Only two classifications were applied to the labeling procedure. The first one is labeled 'Concrete,' the main region of interest (ROI). All possible regions that show fresh and young concrete were labeled under this category. The second one is labeled as 'Not concrete,' which is all except the regions labeled as 'concrete.' This includes the old concrete, surrounding environment, people, equipment, and everything that is not concrete.

After all the images were labeled based on the conditions, a training dataset was generated. A dataset is a categorized directory in which each data (image) is annotated to its respective class. Classes are the categories on the images that the trained model will classify, in this case being 'concrete' and 'Not concrete.' At this point, the dataset is ready to be trained.

31

Every concrete area was labeled, even the shadow and the dark spots. 	The whole image was labeled as 'not concrete' since there were no traces of concrete.
	The reinforcement bars were avoided as much as possible.
	Person's interruption was tried to be avoided.
Only the fresh and young concrete was labeled as concrete. Old concrete and concrete covered by paint were avoided. 	When the concrete was covered with a transparent sheet, the area that the sheet was contacted with concrete were labeled as 'concrete'.
Wet regions or concrete covered with water is also labeled as 'concrete'. 	When the area covered with concrete is in extreme darkness, it was labeled as 'nonconcrete.'

Figure 19: Different conditions used in data labeling (the orange color represents 'not concrete', the blue color is 'concrete')

32

3.2.2.   *Data training*

Data training is teaching the model by providing it with an input dataset based on a specified parameter. However, training a deep neural architecture from the ground up is a massive undertaking that is not feasible due to the requirement of a vast number of datasets. Therefore, as mentioned in the above section, transfer learning will be the best option in training models. Taking this advantage, three state-of-the-art CNN architectures coupled with three backbone encoders were trained for this research. A fully convolutional network (FcN) with Alexnet backbone encoder, Segnet with VGG 16 backbone, and DeepLabV3+ with Resnet 18 and 50 as a backbone were used to investigate the capacity of automatic concrete area detection from the labeled dataset. Table 1 shows the structures and the content of these CNN architectures.

Table 1: Summary of the pretrained CNN models used

Pretrained Backbone Encoder	Alexnet (Krizhevsky et al., 2012)	VGG 16 (Simonyan & Zisserman, 2014)	Resnet 18 (He et al., 2016)	Resnet 50 (He et al., 2016)
CNN Model	FcN	Segnet	DeepLabV3+	DeepLabV3+
Parameters	~60 million	~138 million	~11 million	~23 Million
Neurons	650,000	5504	4096	-
Input image size	227 x 227	224 x 224	224 x 224	224 x 224
Convolutional layer	5	3	5	5
Fully connected layer	3	3	1	1
Max pooling layer	5	5	2 (1 average)	2 (1 average)
Total Connections	24	40	78	192
Total layers	25	41	71	177
Output type	classification	classification	classification	classification

After these models were selected, a training environment was set, consisting of hyperparameters that are crucial in shaping and fine-tuning the models in the right way. Table 2 shows these hyperparameters with their set values. One of these was the Stochastic Gradient Decent (SGD) optimization algorithm. This algorithm is one of the most dominant optimizers that is used to update the model's parameter to minimize the loss fiction, false detection. It trains every element or input of the network repeatedly until it converges into a specific minimum loss value. The

momentum also works with SGD by adding a fraction of the previous training parameters to adjust the current results. It also shapes the SGD in an easier and faster direction, making the algorithm converge quicker.

Other parameters such as learning rate, regularization, shuffling method, and execution environment were also used. These are important in preventing overfitting and lowering the generalization gap. But, the main parameters that were the variables for this investigation were mini-batch size and epoch. Mini-batch size is the size of input values that goes through the training step at one time. It is the number of data the is used to train the model in a single training process. An epoch is a collection of samples in the training data that can update the models' parameters. It is the number of passes through the training process. Therefore, finding the correct values for these hyperparameters is significant to determine error progress and avoid overfitting. In addition to this, these values vary for every network and application purpose, making them volatile. As a result, determining the accurate values is vital. This study trains the models by varying three mini-batch sizes, 10, 30, 60, and three epochs' values, 1, 4, 8. Unfortunately, values above these were not possible due to GPU limitations.

Table 2: Training hyperparameters used in the training process of the dataset

Training parameters	Options used
Optimization algorithm	SGDM
Momentum	0.9
Execution environment	Single GPU
L2 regularization	0.0005
Shuffle	Every epoch
Initial learn rate	0.003
Learn rate schedule	Piece-wise
Mini batch size	Varied
Max epochs	Varied
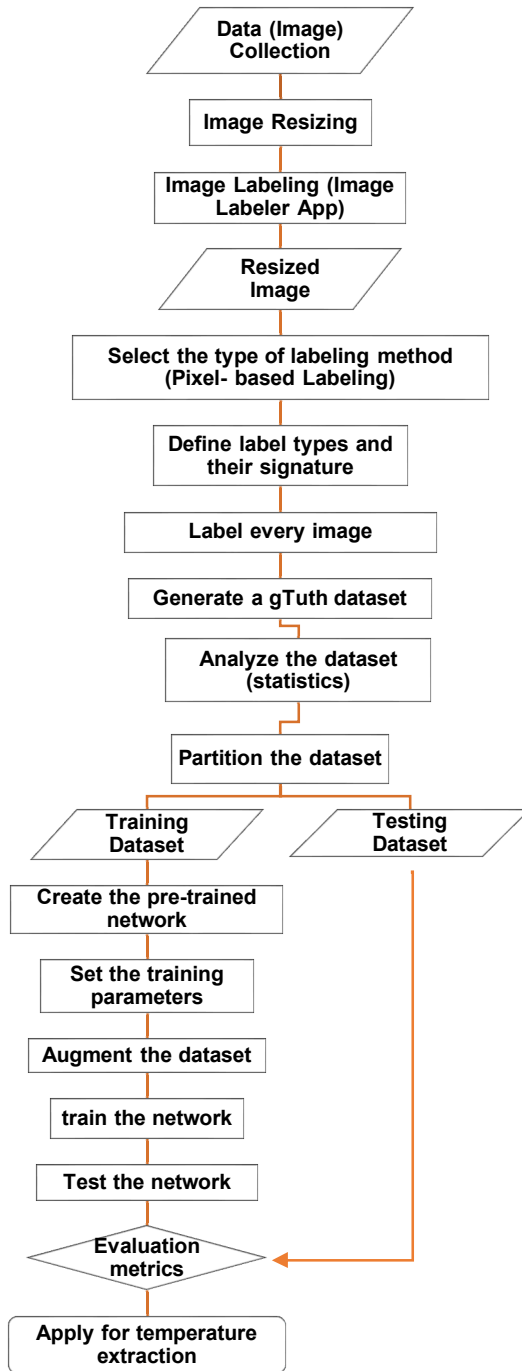
Figure 20: Flow chart of the process of training the CNN-based segmentation model

The other issue that was of the focus was the possibility of the model to be biased due to repetitive training using the same data. This is a very vital issue, especially since fine-tuning a model is pervasive and requires many training sessions. To avoid this problem, a method called data augmentation was employed. Data augmentation is the process of manipulating and varying the training data to increase the model versatility and minimize overfitting without changing the labeling of the training data. It creates a diverse data set without additional data, increasing the model's accuracy and reliability. There are different ways to augment data, from manual transformations to deep learning approaches. In this research, geometric transformations were used to alter the training images. The training images were reflected and translated in both x and y-directions every time the model was trained without changing the labels. This increased the scope of the data set where the model will encounter a different dataset with the same labeled classes.

After these values were set and the data was augmented, training was conducted in a workstation with a MATLAB, 2021a, and A NVIDIA GeForce RTX 2080 Ti 11GB GPU execution environment. Figure 20 shows the whole process of training the model.

3.2.3. *Evaluation and Validation*

The performance of the proposed method is evaluated by comparing the pre-trained models with different evaluation metrics. The quality of the learning algorithms is generally evaluated by analyzing how well they perform on test data (Valeria Maeda-Gutiérrez, 2020). In our case, we used five evaluation metrics to determine the capability and tendency of detection of the architectures. Accuracy, mean accuracy, was the first one which for our case represents the percentage of correctly identified pixels for each class. It is the ratio of correctly classified pixels to the total number of pixels in that class, according to the ground truth. It is the ratio of true positives to the sum of true positives and false negatives in a simpler term. The second one it the global accuracy, which is the ratio of correctly classified pixels, regardless of class, to the total

number of pixels. The main difference between mean and global accuracy is that global accuracy does not consider the class, making it a generalized evaluation method than mean accuracy.

Intersection over Union (IoU), sometimes called Jaccard Index, is another parameter used to evaluate neural networks. It is a metric for measuring the overlap between the ground truth class and the predicted class. The lower the value, the more the prediction capability decreases. In other terms, it is the ratio of true positives to the sum of true positives, false positives, and false negatives. The boundary factor (BF) score is another metric that describes how much each segmentation's boundary aligns with the boundary of the ground truth. Moreover, running time is also one parameter for evaluating the efficiency of the networks.

3.3.   Results and Discussion

The evaluation matrices mentioned in the above sections were computed and plotted as shown in Figure 21 below. The values of these matrices were shown on Y-axis, and the independent variables (X-axis) are epoch and mini-batch size. The labels on the x-axis denote mini-batch size and epoch; for example, "10Ep-1" represents a mini-batch size of 1 and an epoch size of 10. Based on this arrangement, these matrices were used to check how the models performed in detecting the concrete sections.

In terms of these evaluation matrices, except for the training time, the DeepLabV3+ models with Resnet backbone architecture, both Resnet 18 and 50, have a greater detection performance than the others. The accuracy of these models was greater than 80%, with specific models, such as the model trained with 30 epochs and a batch size of 8, having a 94.13 percent accuracy. The average IoU values and the boundary factor scores were likewise greater than the other two models. One of the key reasons for this was that Resnet CNN models feature a skip connection that generates additional parameters from prior layers in addition to the original path from the

input data. This will aid in the creation of more preceptive input in the depth of learning, ultimately increasing the model's accuracy.
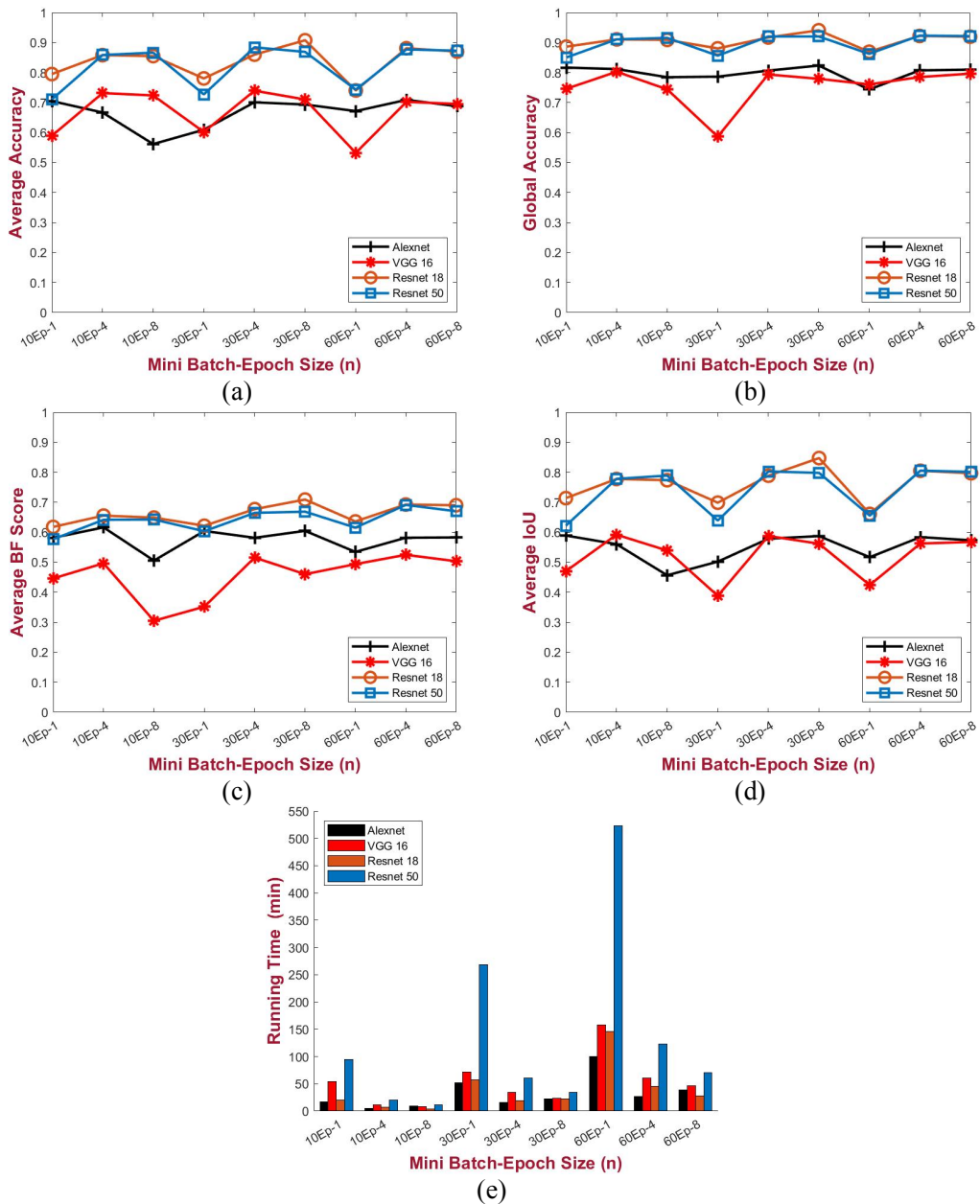
Figure 21 Evaluation of the segmentation models: global accuracy (a), average accuracy (b), average IoU (c), average BF score (d), and running time (e)
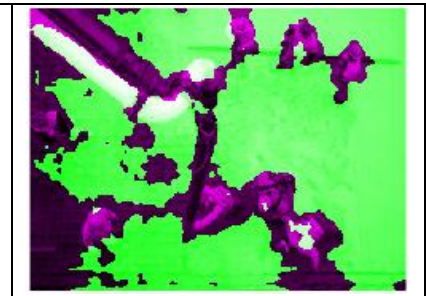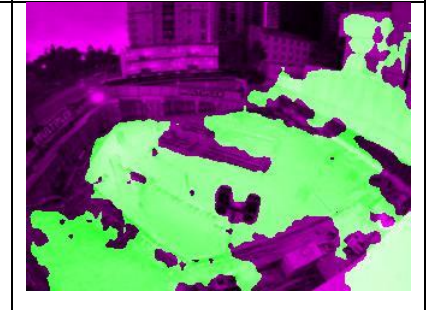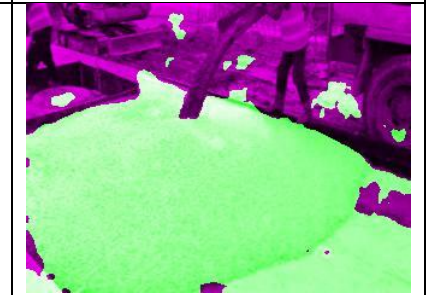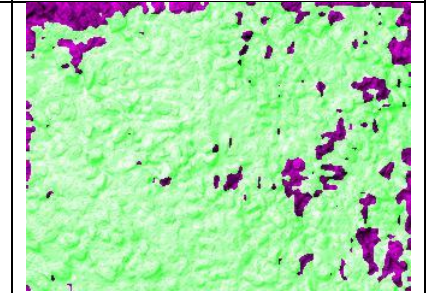
The other explanation is the vanishing gradient problem. This is a critical issue that most models face throughout training. When training neural networks, the model's accuracy begins to accrue as the learning process progresses from layer to layer and eventually reaches a point when any increase will result in the loss of the accumulated values. This is due to the vast number of layers included in most neural networks and the activation functions adopted by the models. To reduce this issue, all the models in this study employed a ReLU activation function, a non-bounded function that updates learning parameters from layers. However, because the models used were pre-trained networks, the number of layers was not modified. As a result, compared to Resnet-based models, the massive number of layers Alexnet and VGG have resulted in a loss in final accuracy. This caused a gap between these segmentation models and the DeepLabV3+ Resnet-based models.

Regarding the training time, contrary to the high performance in terms of accuracy and recall, the training time for the DeepLabV3+ Resnet models was very much higher than the other models. Especially when the number of epochs increases, the training time will increase by keeping a minimum batch size because the model will train the input data individually with small batch size. This will take longer to train even without using a larger epoch size. As shown in Figure 21(e), the running time increases as the epoch size increases with fixed batch size, in the opposite fixing the epoch size, when the mini-batch size increases, the running time increase because the number of input data that are processed at the same time increase when large batch size is taken. However, overall, running time is only a problem when training and fine-tuning the model. After the appropriate accuracy and segmentation level has reached, training was stopped. Therefore, training time will not be an issue during the application of the model.

In addition to the mathematical evaluations, test images with different cases were also presented to check the actual applicability of the models. For the test images, the highest performed model from the evaluation matrices, which is the DeepLabV3+ with Resnet backbone trained with an

epoch of 30 and a batch size of 8, was used. Figure 22 depicts eight cases with test images and their respective segmented image. The bright green represents the area covered by concrete, whereas the deep purple shows the not concrete sections. Each case is different with its own condition and set of surrounding environments. Case 1 in the figure below shows an image of a concrete slab in a casting stage. Some of the reinforcements were not covered by the concrete, but the model accurately segmented only the area that is covered by concrete. But the area that is covered by the shadow of the person was not segmented as concrete. In case 2, the model segmented the area of the concrete even though the image was crowded with construction workers. But exact edge detection between the concrete and the person was not clearly distinguished. Case 3 provides an image of fresh and old concrete. But, the model segmented the old concrete as concrete, which was not the target of the model. This is because there is a little difference between old and fresh concrete other than the depth of their color contrast. Therefore, additional investigation is needed to solve this misdetection. Case 4 indicates an image with interruption by formwork. The model segmented the formwork panel as concrete. This is because one of the parameters that the model checks is the texture of the concrete, whether it is smooth or not. The smooth surface of the panel misled the model to specify it as concrete.

Case 1: Reinforcement interruption, shadow		

Case 2: Person Interruption,		
Case 3: Old and new concrete		
Case 4: Material and color difference,		
Case 5: Smoothness, and texture		

Case 6: Brightness,		
Case 7: Color		
Case 8: No concrete,		
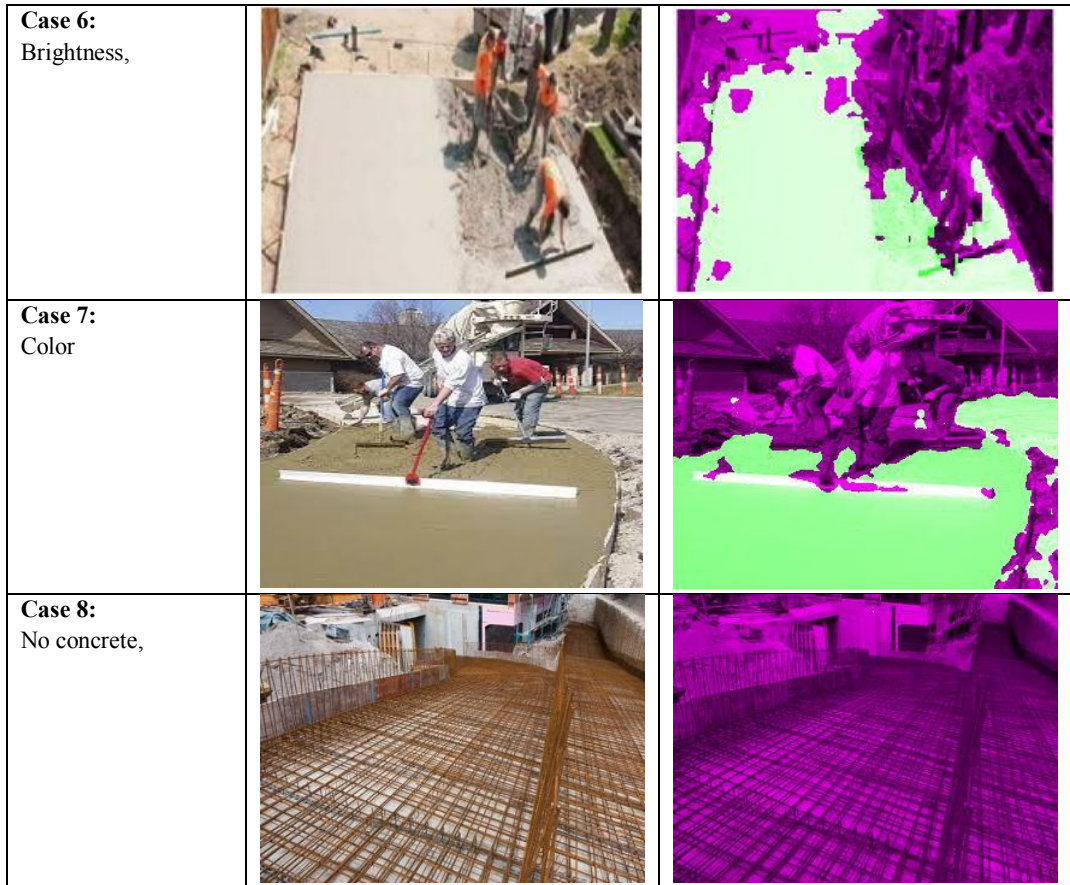
Figure 22 Visual representations of some of the case studies (bright green: concrete; deep purple: not concrete)

In case 5 of Figure 22, the image is all concrete, but the texture is different. The concrete was not smoothed, which some of the coarse aggregates were showing. The model segmented most of the area as concrete with a minor noise as not concrete. This is due to most of the training images were finished concrete sections with a smooth surface. This problem can be avoided by adding additional training images with a variety of textures. Case 6 and 7 show the variables of brightness and color. Due to these variables, the regular gray color of concrete was changed to a different color. But the model segment both images well with a negligible false segmentation. Case 8 shows an image with no traces of concrete sections which the model effectively segments

the whole image as not concrete. This case is as crucial as the others since accurately detecting the area as not concrete is half of the task.

The above-mentioned cases generally show some of the variables that will influence the model's accuracy and reliability in real-time applications. From the results of the segmented images, it can be said that the model has a very high potential for being used as a source of input in generating real-time concrete sections from images. In addition to this, for the purpose of this study, its output is acceptable regarding the extraction of temperature values from the\ segmented concrete sections. However, some adjustments are still needed, especially in some of the variables. Integration with other image processing methods is also another option to refine the quality of the segmentation and increase its accuracy.

4. INFRARED THERMOGRAPHY (IRT) AND APPLICATION ON THE CONCRETE INDUSTRY

4.1. Introduction

Several automatic vision systems have advanced significantly in supporting the visibility of objects in a broader spectrum of wavelength. In the previous days, visible light was the only factor governing how images are generated and their corresponding qualities. For example, a standard digital camera captures the reflection of visible light of the region of interest (ROI) to give a three-channel (RGB) colored image. Without a visible light, this image cannot be produced. However, the development of infrared cameras paved the way to exploit images and data into a broader range of wavelength spectrums. In addition to this, factors other than visible light, such as radiation and temperature, become important in representing objects as images. Any object emits infrared radiation if its temperature is greater than absolute zero (0K, -273 OC). Therefore, these infrared cameras capture the radiations emitted and relate them which an infrared wavelength spectrum, which is between 0.7 to 1000 μm, to define objects and their temperature on an image. This method is called Infrared Thermography (IRT), also referred to as thermal imaging. It is an approach that utilizes the infrared radiation emitted from an object to develop an image of an object and its thermal information (Sarawade & Charniya, 2018). It is a non-contact measurement of the temperature of objects from infrared radiations by using a thermal camera. Since infrared rays cannot be seen by a human, a device is necessary to capture these rays and convert them to temperature values. One of the significant advantages of a thermal imaging camera is that it can give a clear picture when there is no visible light (in the dark). Figure 23 shows that a thermal imaging camera gives a perfect image of ROI at night compared to a visible-light image. This increases the range of application of the thermal camera at any condition of 24 hours a day.
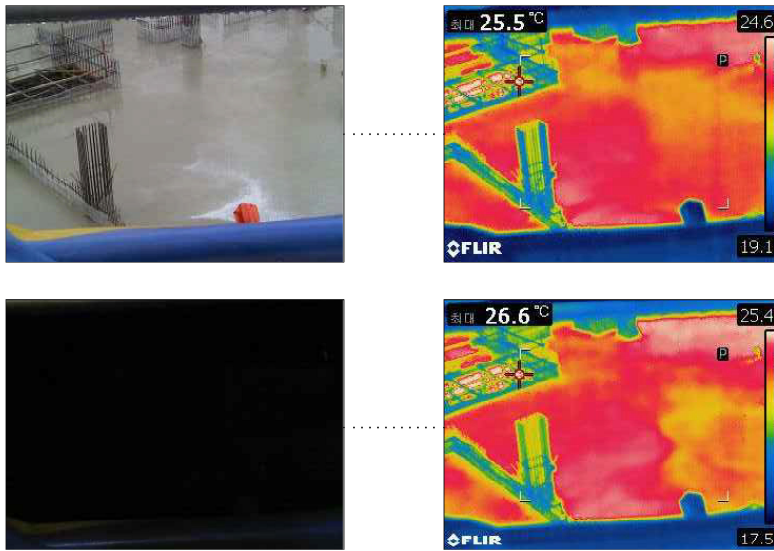
Figure 23: Difference between visible light RGB image and IR image with and without visible light

4.2. Basic Principles and Features of Thermal imaging (IRT) camera

IRT functions based on electromagnetic radiation called infrared or thermal radiation, which has a longer wavelength than visible light (Gade & Moeslund, 2014). This radiation is a heat transfer mechanism generated and transferred to and from any object with a temperature value above absolute zero. Figure 24 shows the general scenario of how detectors capture infrared radiation. This radiation can come in two different ways. The first one is that an external heat source might heat the object, which in turn the body emits this energy for the thermal cameras to detect. Heat sources such as the sun, lamps are some of the external heat sources to consider. The second one is that the body might itself generate its heat. Bodies with an exothermic reaction might emit their heat sources due to the transfer of energy from chemical reactions of their composition parts. Thus, these infrared cameras capture the emitted radiation and relate it to a spectrum of infrared wavelengths to identify objects and their temperatures in an image.
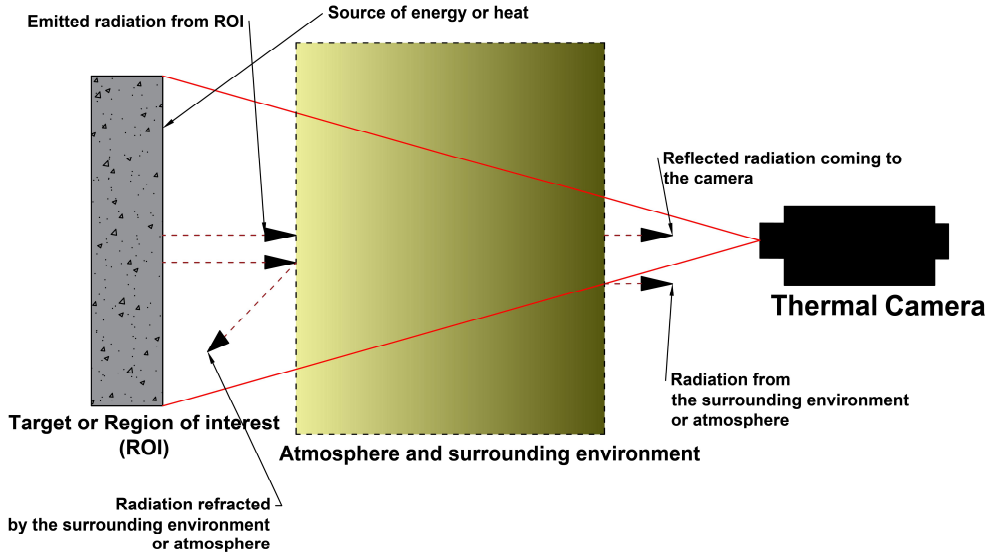
Figure 24 process of transmittance of infrared radiation

Infrared light has longer wavelengths than ultraviolet or visible light and encounters less reflection or absorption due to particles. As a result, it can penetrate the atmosphere significantly and acquire observations across a vast range of distances. This spectrum range is classified into five regions, as shown in Figure 25 However, all the infrared radiation is often not detected by thermal cameras. Some factors impact the detectability of thermal data from an object. The temperature of the object itself, the surrounding environment, and the type of instrument (thermal camera) (Rao, 2008) are some of them, which will be discussed below.
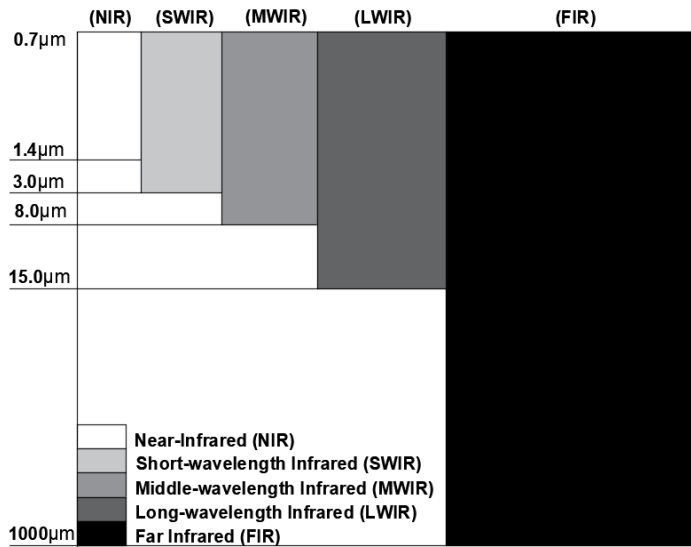
Figure 25: Range of Infrared spectrums (redrawn from (Sarawade & Charniya, 2018))

The temperature of the object and the surrounding environment are the most important factors that affect the output of thermal cameras since they are directly related to the intensity of thermal radiation. As the temperature of the body increases, the intensity of the thermal radiation increases, which will be easy to capture and detect. However, this is not always the case. There is a certain limit where the temperature and the intensity of thermal radiation are directly related. Both will increase to a point where increasing the temperature will inversely affect the intensity after this point. This depends on the wavelength of the emitted radiation. There is a peak limit where further increase in temperature will increase the vibration of molecules, decreasing the wavelength. This will shift the radiation to a visible light spectrum. Therefore, the selection of the optimum portion of the spectrum is important in achieving good reading by infrared cameras.

The surrounding environment is also one of the major factors influencing the output. Humidity, wind, precipitation, and snow affect the transmittance of infrared radiation through the atmosphere. A number of researches have been done to show that the atmospheric condition has a greater impact on the IR radiation, especially in sensitive situations where a small deviation of

47

temperature causes problems (Hernández Alonso, 2019). This is also somehow directly related to the greenhouse effect. Water and carbon dioxide molecules are the most compounds that trap infrared radiation coming from the sun. Therefore, when the concentration of these molecules increases, the IR transmittance will be hindered. This is because water and carbon dioxide molecules take a considerable amount of time to dissipate the IR radiation absorbed by them (Pierrehumbert, 2011; Smirnov, 2019). Thus, humid and rainy atmospheric conditions will affect the accuracy of the output of the IR camera.

Selecting the proper thermal measuring instrument is also another point to consider. Different types of infrared thermometers and cameras are available with different specifications and formats. Each of them differs based on the application they are intended for. Most of the commercially available IR cameras are configured at a middle wavelength IR (MWIR) ($\sim3 - 5.5$ μm) and long-wavelength IR (LWIR) ($\sim8 - 14$ μm) range of spectrums (Sarawade & Charniya, 2018). Cameras with LWIR spectral range can only cover low-temperature radiation, which is up to 500OC. These cameras are used predominantly in various application areas, especially in diagnosis and maintenance purposes. This is because they have higher efficiency in terms of interference by the atmosphere condition and accessing a lower temperature range. On the other hand, MWIR cameras are used for specific purposes where the range of temperature is higher for the LWIR cameras to access.

There are different types of thermal cameras with different specifications and built-in applications. These days most of them are readily available and compact in size. Thermal cameras use different pallets to represent objects and their temperature difference. Figure 26 shows some of the pallets used in thermal cameras. The borderline colormap of the pallets represents the range of temperature detected on the ROI, with boundaries being maximum and minimum temperatures. The colormap can be in a range of black and white as in the pallet called

GRAYSCALE and sometimes can be multiple colors as for the format called RAINBOW. The choice of using the formats depends on the user's preference and the purpose of the study.
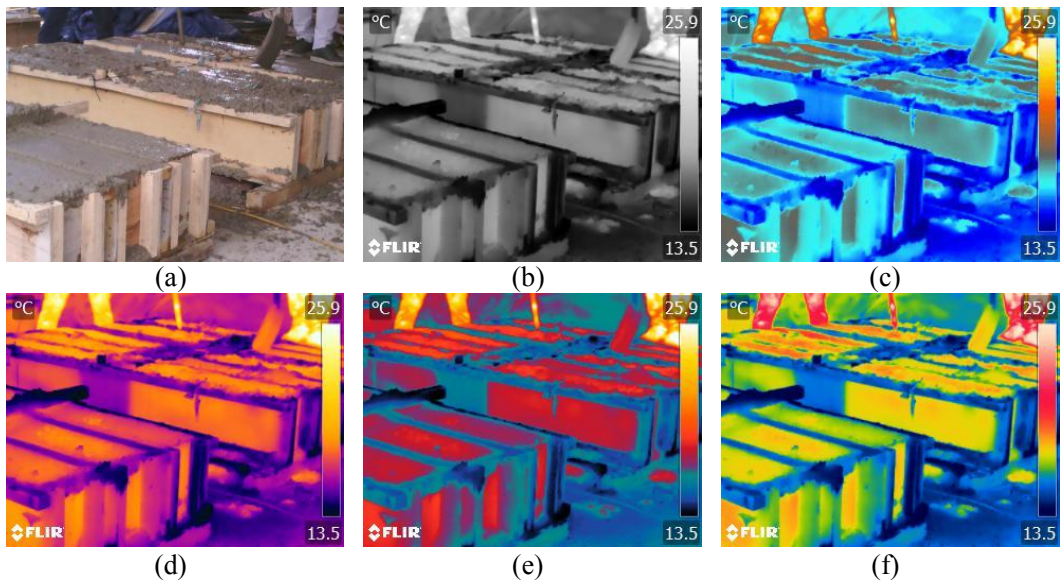
Figure 26: Some of the formats used on IR camera (a. original image, b. GRAYSCALE, c. ARCTIC, d. IRON, e. LAVA, and f. RAINBOW)

Thermal cameras also have different features and parameters that govern the type and the purpose. Camera resolution, thermal sensitivity, accuracy, and detection range are some of the major features that influence the quality of the thermal cameras as well as the output from them. Higher camera resolution will give a clear and precise representation of the object on the image. Higher thermal sensitivity also allows detecting fractional temperature changes, which increases data accuracy. Wide detection range and accuracy also give the thermal camera an extensive scope of temperature coverage which will be applicable in worse situations. The features mentioned above and others define which type of instruments and cameras to use and how accurate and precise the expected result is. From an economic point of view, maximizing the efficiency and features for no extra purpose will not be viable.

4.3. Investigation on the reliability of IRT in extracting temperature from concrete

Since the introduction of infrared rays in the 1800s, the application of IRT has extended significantly from the production of electronic appliances to military surveillance and nuclear industries. The application also transcends in the construction industry, especially in detecting and classifying failures and defects in concrete and related structures (Sirca Jr & Adeli, 2018). A number of researches have been done to support the application of IRT in the construction industry. Corvaglia and Largo (Corvaglia & Largo, 2008) used IRT for quality control of fiber-reinforced polymers (FRPs) installed on degraded columns and beams. They used it to self-detect FRPs failures during installation, operation, and bonding with structural members. They set up different parameters to investigate the accuracy and efficiency of IRT techniques and come to a conclusion that it is one of the fastest and most effective NDT methods to apply under these kinds of conditions. Omar and Nehdi (Omar & Nehdi, 2016) investigated the applicability of IRT, primarily passive IRT, to detect delamination on concrete bridge deck elements. They also compared this method with other NDT tests to observe the technique's viability and gave promising results. Solla et al. (Solla, Lagüela, Fernández, & Garrido, 2019) used IRT as an option for assessing corrosion of reinforcement combined with a geophysical method called ground-penetrating radar (GPR). The results showed that IRT is an important tool to find and classify the reasons for reinforcement corrosion. Al-Hadhrami et al. (Al-Hadhrami, Maslehuddin, Shameem, & Ali, 2012) studied the application of IRT in assessing concrete density by testing on different compositions of concrete mixtures. Eva et al. (Barreira, Almeida, & Ferreira, 2017) also assessed the effect of moisture in lightweight concrete using IRT.

Because of the findings of those above and other studies, IRT can be regarded as one of the nondestructive testing (NDT) procedures used in the concrete sector (Usamentiaga et al., 2014)

(Rao, 2008). It has come a long way from flaw detection to monitoring and managing concrete structures. It is now being used in the construction industry both alone and in conjunction with other approaches to improve its quality and efficiency. However, most of the research done on IRT technology was structure monitoring and fault detections such as rebar corrosion, cracks, and surface deterioration on concrete surfaces (Pla-Rucki & Eberhard, 1995). However, a different approach is used in this research by applying thermal imaging on concrete surfaces coupled with the deep learning technique to determine the mechanical strength of the concrete. In order to apply the proposed technique, the validity of the IRT has to be adequately verified. Two aspects that influence the output of an IR camera were explored in this study: instrument consistency, distance range from ROI.

4.3.1.   *Type of equipment used*

The consistency of the equipment is one of the aspects that influences the results from thermal imaging cameras. As a result, two thermal imaging cameras, the FLIR T530 and the FLIR C2 were used, as shown in Figure 27. This is to check how much the difference in equipment affects the temperature data. Both have a spectral range of 7.5 to 14μm, object temperature of -10oc to 150oc, operating temperature range of -10oc to 50oc, an accuracy range of ±20c/2% variation. But the IR sensors and image frequency are different. The T350 has an IR sensor of 320 x 240 and an image frequency of 30Hz. C2 has an IR sensor and image frequency of 80 x 60 and 9Hz, respectively. The difference in the image frequency is very crucial since the refresh rate of the cameras is very important. Especially if it is applied in motion-based detection, the image frequency has to be higher to find a higher accuracy and precise value.

<center>(a)   (b)</center>

Figure 27: Type of equipment used, (a) FLIR T530, (b) FLIR C2

When considering the IR sensors, the higher the IR sensors, means the higher number of measured pixels. The T530 IR camera has 76800 pixels that measure the temperature of the ROI, which is higher than 4800 pixels in the case of C2. This means that since the image output is equal to the measurement pixels, every bit of the image data has the actual measured temperature values, which is accurate and realistic.

A standard concrete mix proportion was used to cast six conventional concrete cubes. The experiment's setup is shown in Figure 28. The humidity and temperature of the curing chambers were kept at 55±1% and 21±0.5°C, respectively. The thermal camera was kept at a distance of 0.70 meters from the concrete's surface. Both cameras monitored the temperatures every hour for ten hours to see how they differed. After taking both the RGB and IR images of the samples, they were analyzed using the software FLIR Tools+ to extract temperature data of the concrete cubes' sections from the images. As shown in Figure 28 (c) the sections of the concrete cubes we selected first and their temperature value was extracted in tabular format. Then their temperature distribution histogram was plotted. Finally, their temperature distribution at 1, 4, and 10 hours after casting was compared by both cameras.

<center>52</center>

(a)

(b)   (c)
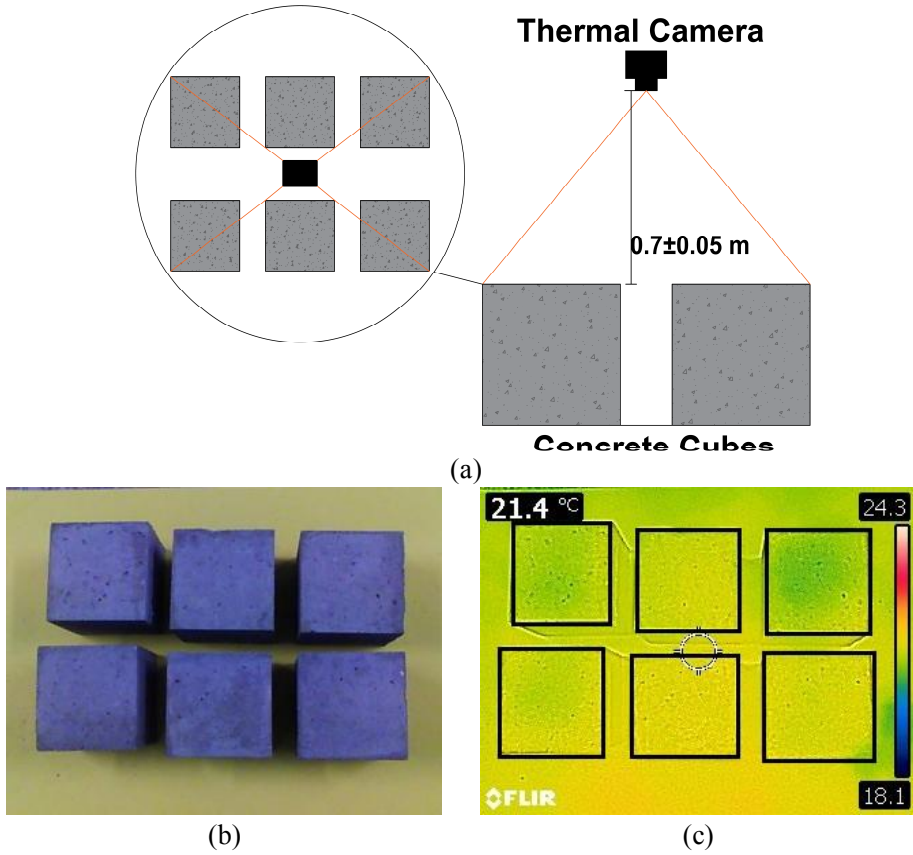
Figure 28: (a) Set up of the experiment (b) the RGB image of the concrete cubes and (b) their corresponding thermal images

(a)   (b)
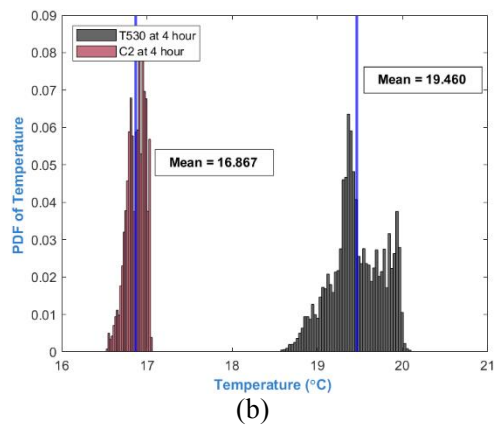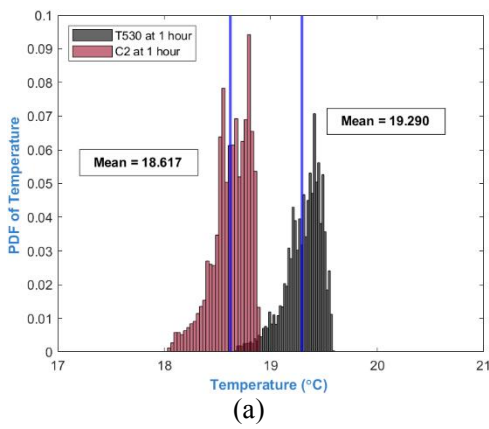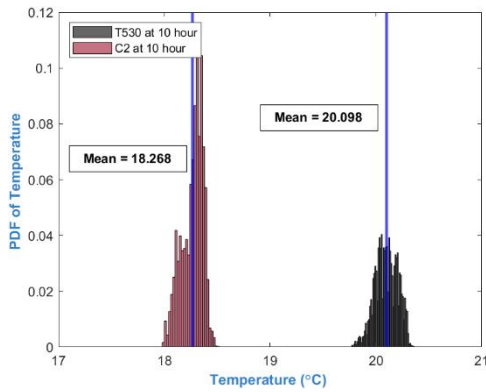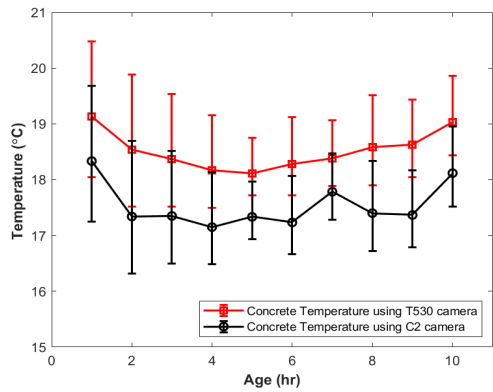
(c)   (d)

Figure 29 Effect of equipment on temperature fluctuation at 1 hour (a), 4 hour (b), and 10 hour (c) and temperature progress (d)

A comparison of the two cameras reveals a difference, despite the fact that every other variable was maintained. Figure 29 shows the plot of the difference in equipment. There is a 1-3°C of average temperature difference between the two cameras. But this variation is within the parameters' tolerance which can be considered acceptable. The experiment was also done on old concretes of retaining walls outside of the laboratory. Figure 30 shows the probability density functions of the temperature values of some examples of old concrete structures. A common area was chosen from the images, and the temperature distribution was plotted. The findings are consistent with those obtained in a laboratory experiment with a temperature variation of less than 2°C. As a result, the choice of each of the thermal cameras is viable because the slight temperature difference will not affect the study's further investigation.

54

(a)

(b)
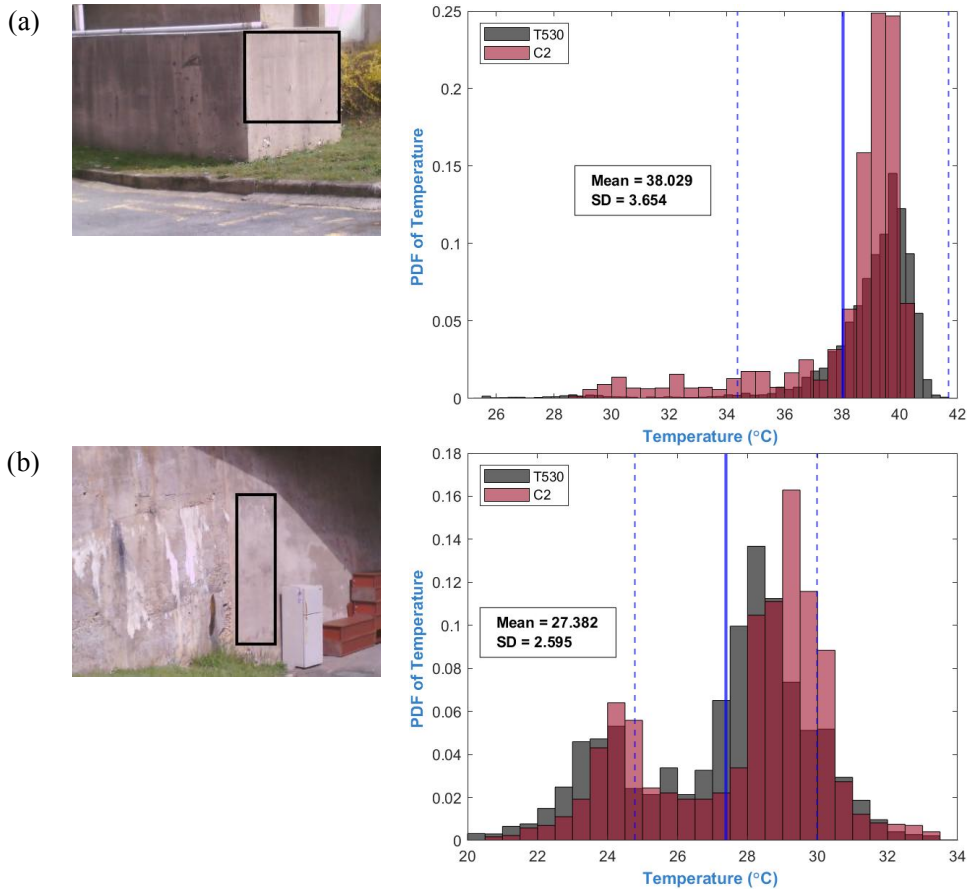
Figure 30: Difference between the two cameras old concrete structures

4.3.2.   *Effect of distance range from ROI*

The second factor that was investigated was the effect of the distance range between the object and the camera. The previously casted concrete cubes were placed at various distances from the thermal camera, as Figure 31 below. This was first done in a laboratory setting with temperature and humidity controlled up to a 20-meter range. The thermal images of the samples at various distances from the thermal camera are shown in Figure 32. As can be seen, as the distance increases, the resolution decreases significantly. This is even more pronounced on the C2 camera, where the samples are difficult to discern after a 10-meter interval.

55
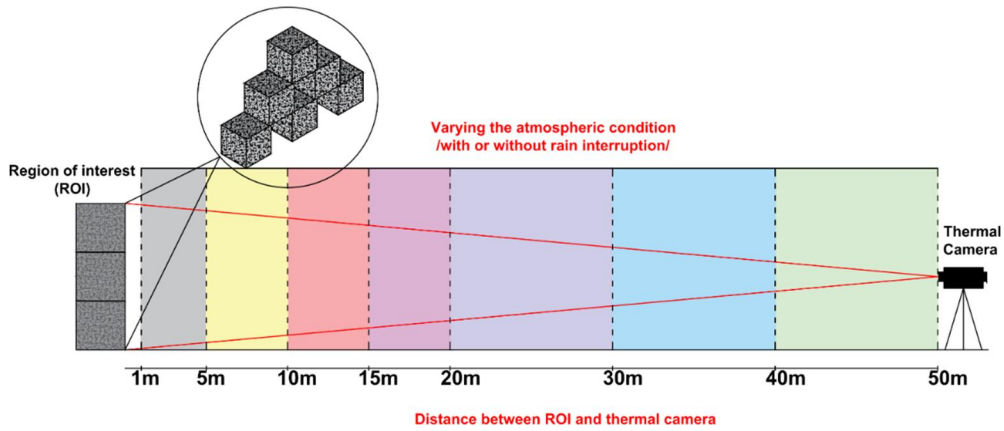
Figure 31: Test setup for investigating the effect of distance between ROI and thermal camera
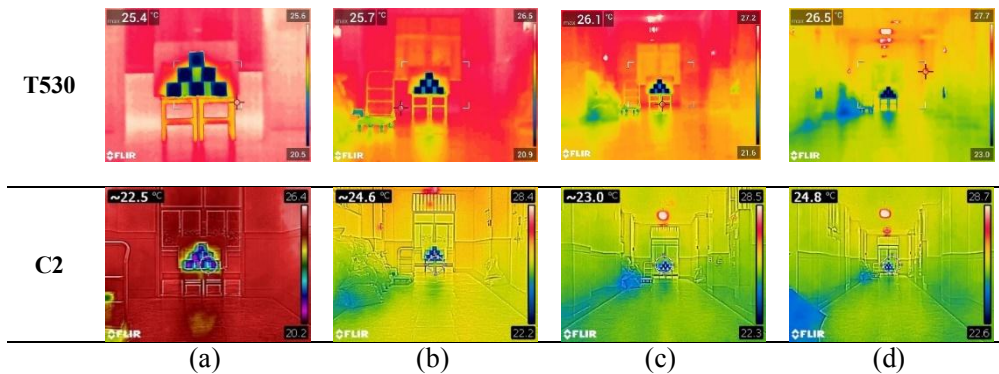
Figure 32: Examples of IR images taken at a (a) 5m, (b) 10m, (c) 15m, and (d) 20m distance from the thermal camera in the laboratory.

Figure 33 shows how the distance between the cameras affects the temperature data from the cameras. The data collected using T530 is fairly consistent, with only a ±1°C difference on average. However, C2 data shows a greater difference than T530. It was impossible to collect data after the 10 meters, implying that the distance range has a significant impact on C2. However, the temperature data from T530 shows higher reliability than C2.

The concrete samples were then taken into an outside environment, and their thermal images were captured. During this experiment, the environmental conditions such as humidity, wind speed, air temperature, and rain were not controlled. Figure 34 shows the difference in loss of

56

resolution from both cameras due to distance increment. The results were consistent with the laboratory findings with a less than 2°C variation when the T530 camera was used. The average temperature values of both cameras at various distances can be seen in Figure 35 below. Since the C2 camera cannot generate thermal images from more than 10 meters, the T530 camera was used to check the possibility and extent of distance that the camera can identify and generate thermal images. For this, thermal images of the same concrete samples were taken from a distance up to 50 meters. Their temperature was plotted in Figure 36, which substantiates the above conclusions.

The same procedure was followed on the old concretes as seen in Figure 37 confirming that the distance has a lower effect on the temperature data. This is a crucial aspect to consider, mainly because it affects the output of the thermal cameras, especially when sensitive and small variation data are needed. However, in this case, a small variation will not affect the results, making it acceptable to use.
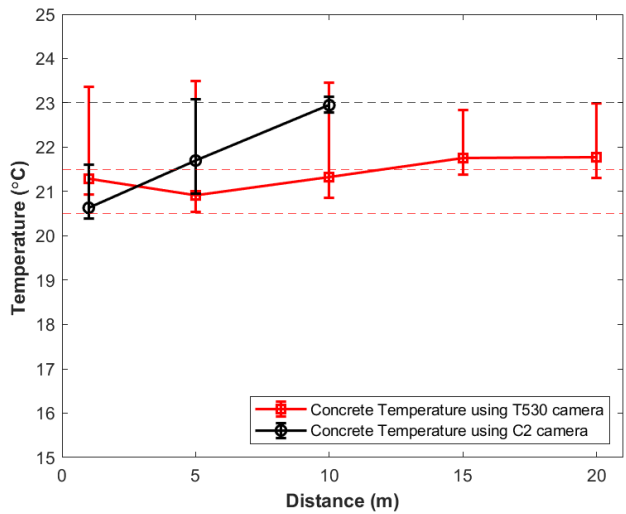
Figure 33: the effect of distance on the temperature data in the laboratory

T530

C2

| | | | | |
| (a) | (b) | (c) | (d) | (e) |
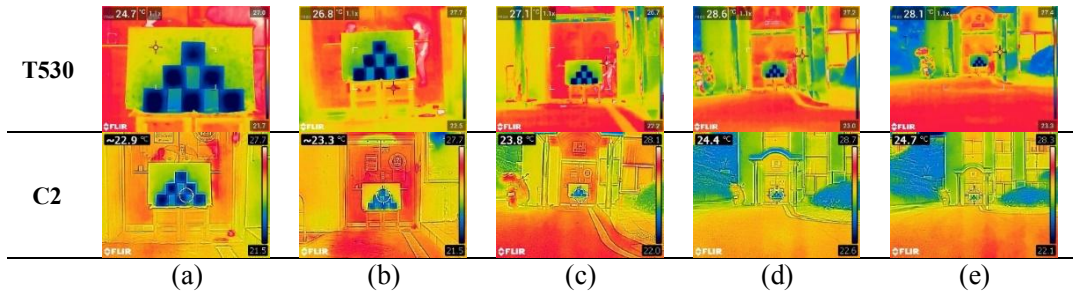
Figure 34: Examples of IR images taken at a (a) 3m, (b) 5m, (c) 10m, (d) 15m, and (e) 20m distance from the thermal camera in the field.
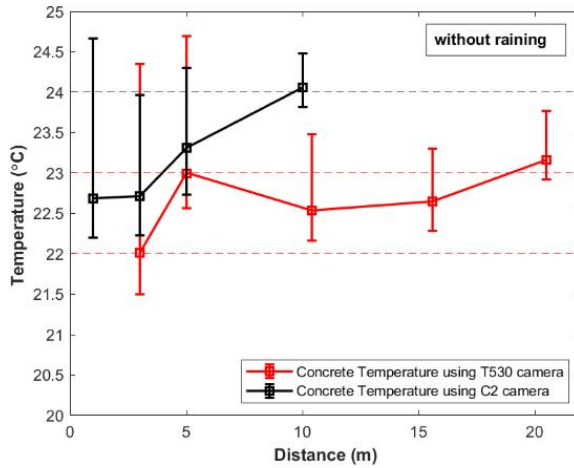
Figure 35: Effect of distance from the thermal camera in the field
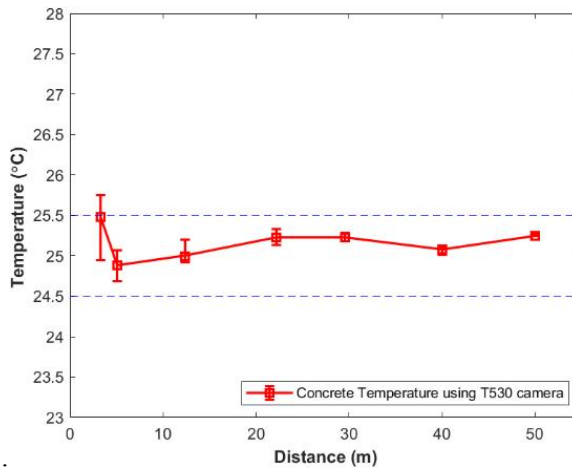
Figure 36 average temperature of concrete samples with a range of distance by T530 camera
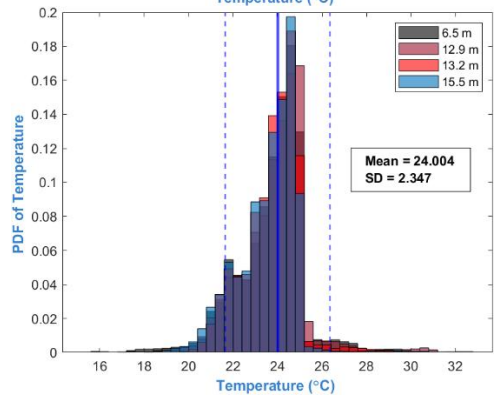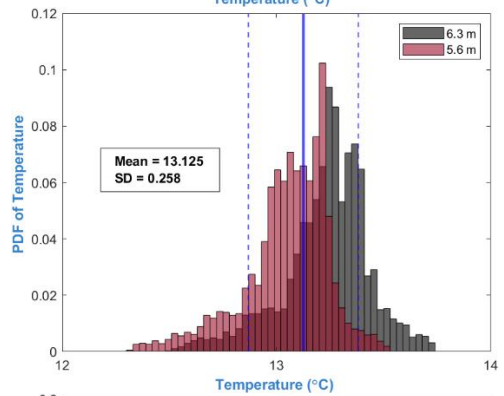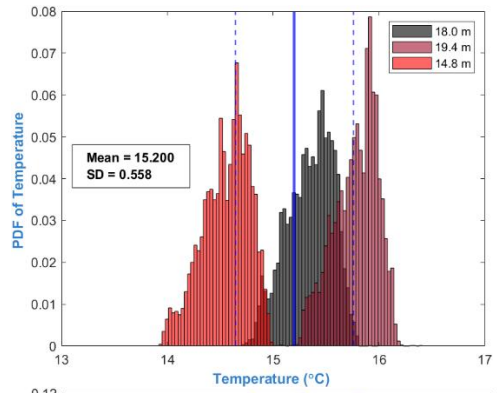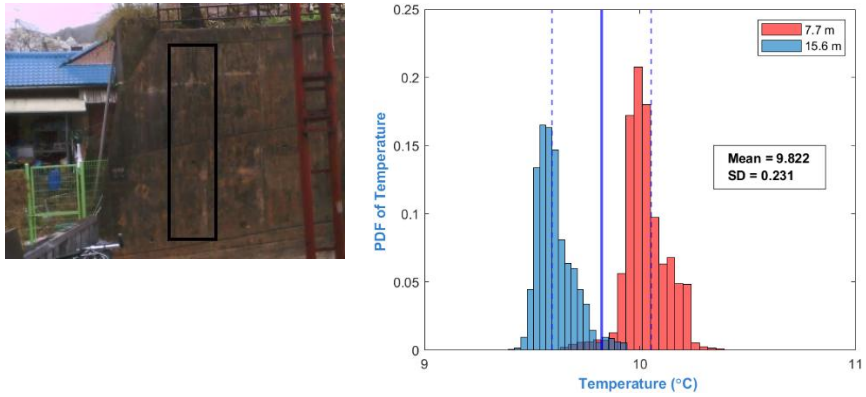
Figure 37: Examples of thermal images of old concrete structures from a certain distance

Based on the experiment done, it can be understood that IR resolution plays a pivotal role in the image quality and accuracy of the output of thermal imaging cameras. Higher resolutions provide precise and reliable measurements of smaller targets from further distances, creating sharper thermal images. The higher the detector resolution, the more accurate the camera. The difference in the image frequency is very crucial since the refresh rate of the cameras is significant. Overall, IRT is an essential and reliable testing technique applied on concrete structures.

5. CORRELATION OF DEEP LEARNING AND IRT IN PREDICTING STRENGTH DEVELOPMENT OF CONCRETE: PROPOSED MODEL

5.1. Introduction

Two separate concepts were raised in the previous chapters: deep learning-based image segmentation and IRT. Both were investigated to check if they are reliable and can optimize their respective application areas. From the results, it can be said that they can perform their tasks better than their other constituents. Therefore, in this chapter, the combination of these two approaches was applied in both field and laboratory experiments. Furthermore, their reliability in predicting the strength of concrete at any time was investigated by comparing it with practical experimental results. But first, the relationship between temperature and mechanical properties will be discussed since the goal is to relate the temperature value of concrete in estimating its strength.

The mechanical strength of concrete can be affected by many internal and external conditions. Among them, temperature is one of the governing variables responsible for developing the mechanical properties of concrete. External (surrounding) temperature and internal temperature influence concrete up to the microstructure level. From the standpoint of hydration of cement, at the early stages of contact with water, a large amount of heat is evolved where the major hydration products are formed. These products are responsible for attaining the expected mechanical properties of the concrete. Therefore, heat evolution (temperature history) and time (age) are major variables in determining the concrete's status. In order to relate the temperature history of concrete with its strength development, a concept called concrete maturity was drawn, which is discussed in the next section.

61

5.2. Concrete Maturity

Concrete maturity is a term that relates the combined effect of temperature and concrete age to describe the development of the mechanical properties of concrete. It is one of the non-destructive ways of evaluating the properties of concrete based on temperature and time. The temperature history of the concrete during the curing phase is used to find the in-place strength and other target parameters. This concept can be traced back to the industrial revolution in the 1940s during the introduction of stem curing (N. J. Carino, 1984; Nurse, 1949; Saul, 1951). Since then, there have been different researches done to validate the prediction capacity of this method. In 1951 Saul (Saul, 1951) proposed what is known to be the first 'maturity rule' by expressing time and temperature-dependent function. He proposed that above a certain datum temperature, -10.5oc for his case, the strength development is a factor of the temperature and the age of concrete. However, he emphasized that this will be valid if the concrete temperature did not reach above 50oc within 2 hours or 100oc within 6 hours after mixing has started. Saul presented his working theory with the following equation.

$$M = \sum_{0}^{n} (T^c - T^c{}_0)\Delta t \tag{1}$$

Where,

- M is maturity at age t,
- T^c is the average temperature of concrete at a time interval of Δt, and,
- $T^c{}_0$ is the datum temperature.

Since then, different studies have been drawn proposing different methodologies and equations to develop concrete maturity with their drawbacks and limitations. However, a major concept was proposed by Hansen and Pedersen (Hansen & Pedersen, 1977) in 1977,

which based the maturity of concrete on the Arrhenius equation. The general equation is expressed in the following equation.

$$M = \sum_{0}^{n} A e^{\frac{-E}{RT}} \Delta t \tag{2}$$

Where,

- T, represents the temperature of the concrete at a time interval of Δt,
- E is the apparent activation energy,
- R is the gas constant, and
- A is the constant of proportionality.

Investigations were done to validate whether the Nurse-Saul method or the Arrhenius method of calculating the maturity of concrete is viable. It was found that the Arrhenius method is more accurate over a wide range of curing temperatures and gives a better relation of temperature and age than the Nurse-Saul method (Naik, 1985). But there are still some limitations in selecting the appropriate activation energy, E, the constant of proportionality, A, and datum temperature. Therefore, different recommendations were made on these variables (N. Carino, 1981; N. J. Carino, 1984; Chengju, 1989; Tank & Carino, 1991) to suggest that these variables are dependent on the type of concrete, the curing conditions, the surrounding environment, and the type of constituents in the concrete.   In addition to the above-mentioned studies, several other research was done to develop an accurate concrete maturity model. Some of them are listed in Table 3 below. According to the building standards ASTM C1074 (S. ASTM, 2004), both the Nurse-Saul and Arrhenius methods are options to develop the maturity model of concrete.

Table 3: List of some the maturity equations developed along with their reference

Proposed by	Maturity function
Rastrup, 1954 (Rastrup, 1954)	$$t_e = \sum_0^n 2^{\frac{T^c}{10}} \Delta t$$
Weaver and Sadgrove, 1971 (Weaver & Sadgrove, 1971)	$$t_e = \sum_0^n \frac{(T^c + 16)^2}{(T_r - 16)^2} \Delta t$$
Hansen and Pedersen, 1977 (Hansen & Pedersen, 1977)	$$t_e = \sum_0^n A e^{\frac{-E}{RT}} \Delta t$$ E = 33.5KJ/mole, for $T^c \geq 20°C$ E = (33.5 +1470(20-T^c)) KJ/mole, for $T^c < 20°C$
Carino, 1984 (N. J. Carino, 1984)	$$t_e = \sum_0^n e^{B(T-T_r)} \Delta t$$ Where, B is a temperature sensitivity factor (1/°C)
Jonassen, 1985 (Jonasson, 1985)	$$t_e = \sum_0^n A e^{\frac{-E}{RT}} \Delta t$$, where $E = 4600R(\frac{30}{T^c+10})^{0.39}$
Helland, 1987 (Helland, 1987)	$$t_e = \sum_0^n A e^{\frac{-E}{RT}} \Delta t$$ E = 31.0KJ/mole, for $T^c \geq 20°C$ E = (31 +1200(20-T^c)) KJ/mole, for $T^c < 20°C$
Fib CEB-FIP model code, 2010 (Walraven, 2012)	$$t_e = \sum_0^n e^{[13.65 - \frac{4000}{273+T(\Delta t)}]} \Delta t$$

5.3. Strength-Maturity Relationships

The main reason for developing the appropriate maturity model is to develop a relationship with the mechanical properties of concrete. Plowman (Plowman, 1956), first proposed the relationship of strength and maturity based on the Nurse-Saul function. After being subjected to conventional curing for 24 hours, concrete cube specimens were exposed to temperatures ranging from -11.5 to +18 °C to develop the interdependence. He developed the assumption that a straight line will

64

be fitted while describing strength using a logarithmic maturity function. The equation below describes the strength maturity relationship where S stands for the compressive strength, a and b are constants, and M is the maturity based on the Nurse-Saul method.

$$S = a + blog(M) \qquad (3)$$

He also suggested that the equation above is valid irrespective of the influencing factors such as water to cement ratio, cement type, and aggregate used. He proposed the reference temperature where the maturity of the concrete starts be -11.7°C. This temperature range was considered to be the point where the strength of the concrete remains constant irrespective of time. Since this development, researchers have done different investigations on advancing the relationships. In 1968, what is called 'the crossover effect' was proposed by Verbeck and George (Verbeck, 1968), which explains how the curing temperature affects the strength development despite having the same maturity. This means that concrete cured at a higher temperature will develop a higher strength at the early stages than lower temperature cured concrete with the same maturity. However, at a later stage, the high temperature cured concrete will have a lower strength due to the early development of reaction products which halted further strength development. This scenario was one of the grounds for Tank and Carino (Tank & Carino, 1991) to formulate a rate constant model that stipulates the relationship of compressive strength of the concrete and its temperature. They stated that the compressive strength at a given time is a function of the current strength and the concrete temperature at that specific time. Knudsen (Knudsen, 1984) also supported this model by adding some assumptions. He considered the reactivity (rate constant) and particle size distribution of the cement particles as the governing factors in the strength development of concrete. Table 4 shows some of the studies on the strength–maturity relationships and their equations.

65

$$s = s_u(\frac{k(t - t_o)}{1 + k(t - t_o)})$$ (4)

$$s = s_u(\frac{k_r(t_e - t_{or})}{1 + k_r(t_e - t_{or})})$$ (5)

where,

- S is the strength at age t,

- S_u is the ultimate strength,

- k and k_r are the rate constants and,

- t_o and t_{oe} are the temperature at the start of strength development.

Table 4: List of some the maturity-based strength development equations developed along with their reference

Proposed by	Maturity-based strength development function
Plowman, 1956 (Plowman, 1956)	$s = a + b \log(M)$
Chin, 1971 (Kee, 1971)	$s = \dfrac{M}{\dfrac{1}{A} + \dfrac{M}{S_u}}$
Lewand and Reichard, 1978 (Lew & Reichard, 1978)	$s = \dfrac{K}{1 + Ka[\log(M - 30)]^b}$
Carino, 1983 (N. J. Carino, Lew, & Volz, 1983)	$s = \dfrac{(M - M_o)}{\dfrac{1}{B} + \dfrac{(M - M_o)}{S_\infty}}$
Knudsen, 1984 (Knudsen, 1984), Geiker, 1983 (Geiker, 1983)	$s = s_\infty(\dfrac{\sqrt{k(t - t_o)}}{1 + \sqrt{k(t - t_o)}})$
Carino, 1984 (N. J. Carino, 1984)	$s = s_\infty(\dfrac{k(t - t_o)}{1 + k(t - t_o)})$
Hansen and Pedersen, 1984 (Hansen & Pedersen, 1984)	$s = s_u e^{-(\frac{\tau}{t})^\alpha}$
Weisstein, 2017 (Weisstein, 2019)	$s = \dfrac{a}{1 + be^{-cM}}$

5.4. Experimental Methodology

In the laboratory experiment, extracting temperature from IR camera was investigated by comparing its results with other thermal sensors, thermocouples. Based on the temperature data from both equipment, a strength development curve was drawn and compared with the experimental data to define how well both the IR camera and thermocouple sensors can define the strength of concrete based on the maturity method. The second and major experiment that was done was the investigation of the application of this proposed method in the construction site. This was done on the field where an IR camera was used to monitor and gather the temperature of a concrete section and combined with the deep learning-based segmentation technique to develop a strength development curve. This was an important experiment, especially since this method's application is crucial in optimizing the output of construction sites in terms of time and cost.

5.4.3. Laboratory-Based investigation

In this experiment, an IR camera is used to extract the temperature of concrete samples, which will help predict the strength of concrete based on maturity function. Comparison between other thermal sensors such as thermocouples was also made to check the variation and deviation in the temperature value. Constituent materials were selected for this, and their respective properties were examined. Figure 38 and Table 5 how the properties of the materials used in this experiment. After materials were selected, mixing was done based on the specification provided in ASTM C192 (A. ASTM, 2013). Water to binder ratio of 0.35 and sand to an aggregate ratio of 0.4 was used as a mix proportion. Thirty cylindrical samples were prepared to measure the compressive strength of the concrete starting from 12 hours to 28 days. The samples containers were modified to insert the thermocouple sensors in three positions: top, middle, and bottom. Figure 39 shows the arrangement of the thermocouple sensors in the sample containers. After the modification was done, the prepared concrete was poured and cured at a standard atmospheric temperature.

| (a) | (b) | (c) |
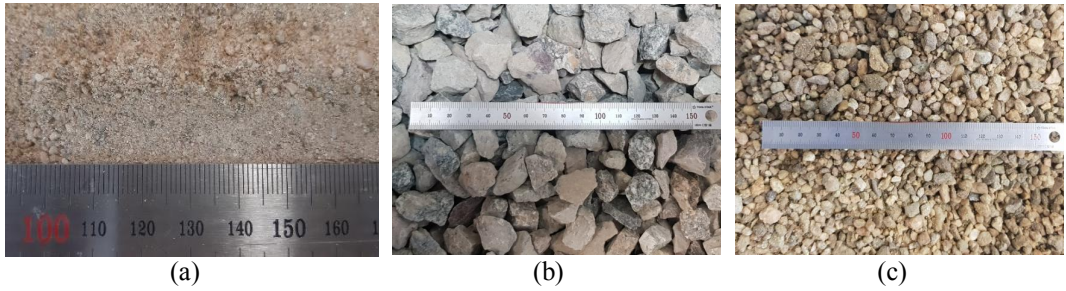
Figure 38: Materials used for preparing the concrete (a) sand, (b) coarse agg. 1, and (c) coarse agg. 2

Table 5: Properties of the materials used

	Fine Agg.	Coarse Agg. 1	Coarse Agg. 2
Specific Gravity	2.45	2.46	2.62
Loose Bulk Density (kg/L)	-	1401.7	1291.8
Compacted Bulk Density (kg/L)	-	1507.7	1424.4
Water Absorption (%)	2.19	11.32	
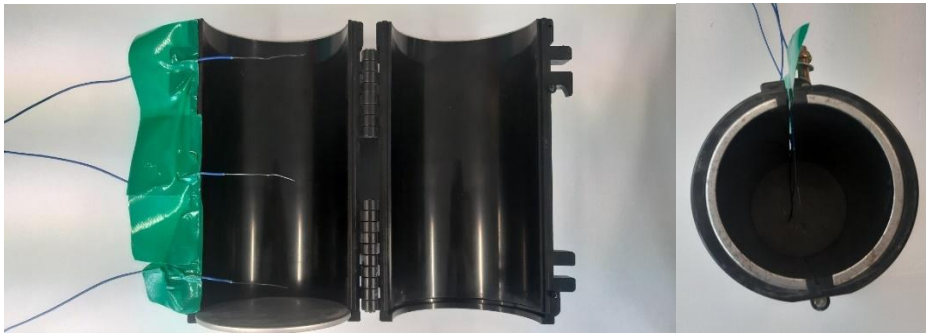
Figure 39: Sample container modifications

After they are cured, the compressive strength of the samples was measured along with the core temperature of the samples at their respective ages. Figure 40 shows the compressive strength of the samples with their respective atmospheric condition.
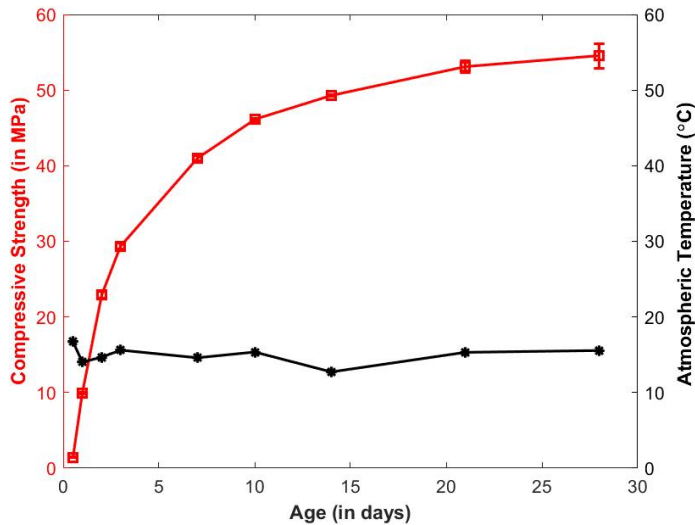
Figure 40: Compressive strength of laboratory prepared samples with their surrounding temperature

The thermocouples and the thermo-imaging camera were used to measure the temperature as prearranged. When measuring the temperature with the thermocouples, the wires inserted in the concrete were directly connected to a 34972A LXI data acquisition and data logger switch unit shown in Figure 41 below. This unit has a type-K thermocouple measuring module that reads the temperature of any connected material in degree Celsius. The reading was done every 20 minutes for one day and every 2 hours for the next two days after the concrete was poured into the modified cylindrical container. A smaller reading interval was necessary at the early stages because the initial temperature variation was very high. This change determines how the strength will develop in the concrete.

69

Figure 41: Temperature measuring equipment (data logger)

For the IR images, FLIR C2 thermo-imaging camera was used to capture the thermal images of the samples. Figure 42 shows the process of extracting the temperature values from the thermal images of the cylindrical samples. The IR camera used in this experiment can capture both thermal and RGB images. Therefore, both images were generated with the same time interval as the thermocouples. After this, the RGB images were taken, and the concrete samples' surface was selected using a FLIR Tools+ software. This was done to avoid the interruption of the background and the surrounding environment, which is responsible for shifting the temperature value. After the concrete samples' area was selected from the RGB images, it overlapped with its corresponding thermal image to extract the temperature of only those sectioned areas. The temperature distribution was then plotted, and statistical analysis was performed. The temperature extraction and statistical analysis were done in a MATLAB environment. Finally, the results were compared and used as an input for strength prediction based on the maturity function.

IR camera (FLIR C2)

Thermal image of the concrete cylindrical samples

RGB image of the concrete cylindrical samples

Selected concrete surface of the cylinders

Overlapped image

Temperature distribution
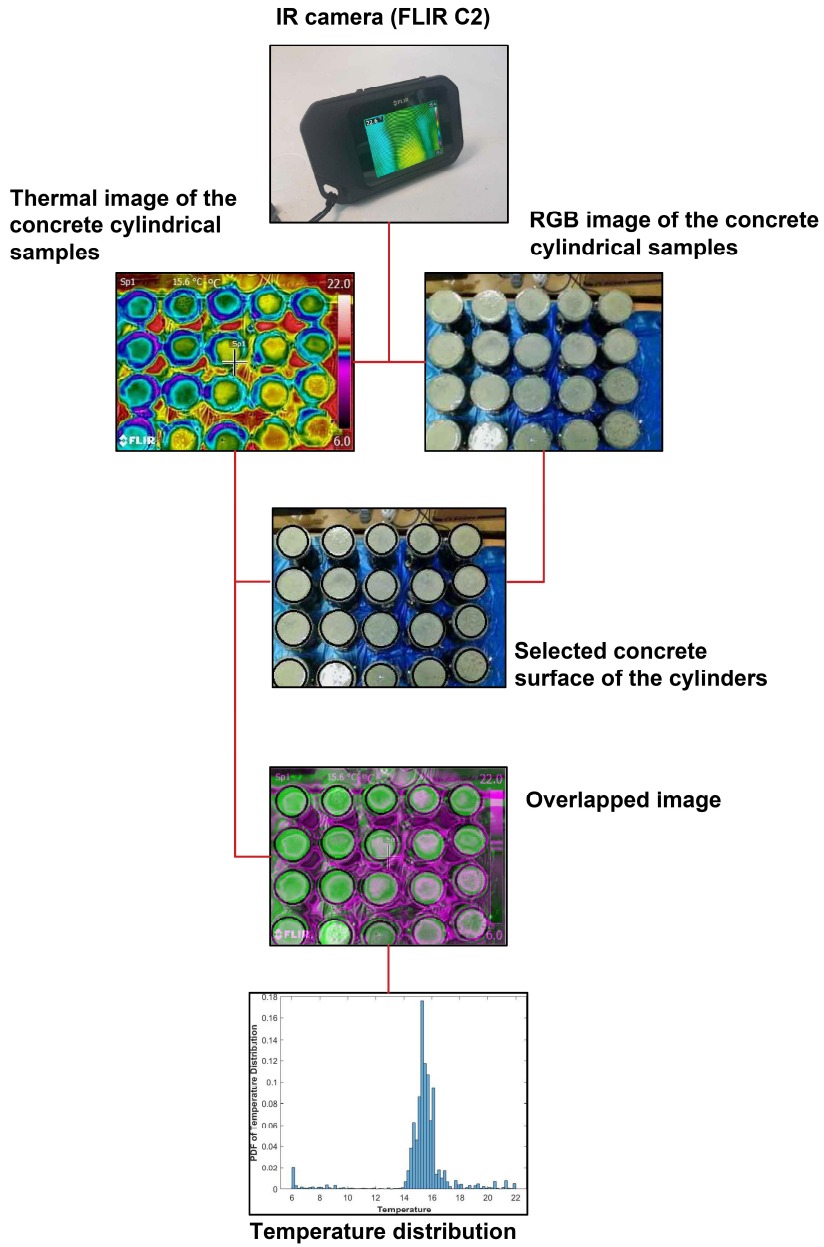
Figure 42: Process of extracting temperature from the thermal images of the cylindrical samples

5.4.4. *Field-Based Investigation*

The field experiment was necessary to check the applicability of the proposed model as well as confirm the results from the laboratory investigations. For this purpose, a construction site for an apartment building in Gwangju, South Korea, was selected. The site consists of several high-rise residential buildings at different construction stages. Most of them were under the concrete casting stage, making the site suitable for the testing environment. Furthermore, since the temperature was the primary variable, the construction site was visited in two extreme conditions: summer and winter.

A thermal imaging camera was installed at the construction site to monitor and record the thermal information of the selected concrete sections. FLIR T530 thermo-imaging camera was used to capture the thermal images of the concrete sections. The IR camera was set to take images every 30 minutes for the first three days just after the start of concrete casting. In addition, RGB images of the same thermal images were also taken for the main purpose of integrating with the deep learning-based image segmentation technique. Figure 43 shows some examples of the images taken from the construction site in a day with a six-hour interval. As seen from this figure, at night, the RGB images show nothing but darkness, but their corresponding thermal images show an explicit detail. This was one of the advantages of the IR cameras, which was mentioned in the previous sections.

After both types of images were taken, they were processed in the proposed deep learning-based image segmentation method to find the temperature of the concrete sections. First, the RGB image is passed through the proposed trained network to segment the correct concrete area automatically. It is then overlapped on the thermal image to get the intersection thermal map of the concrete area. Then the overlapped image is analyzed to give a thermal histogram of the concrete area. This was done for every image taken from the IR camera. The one problem during overlapping the two images was that the RGB image was not accessible at night due to the

72

absence of visible light. This makes the segmentation of concrete area difficult. However, a looping system was designed to use the previously captured image for segmentation and then overlap it with the thermal image of the darker RGB image. This way, the temperature at the 30-minute interval was collected in both weather conditions. This way, the temperature of the concrete sections is gathered for further application in developing the maturity-based strength development curve. Figure 44 shows the proposed method's process that combines both the image segmentation and IR techniques to analyze the data and develop a strength development model.
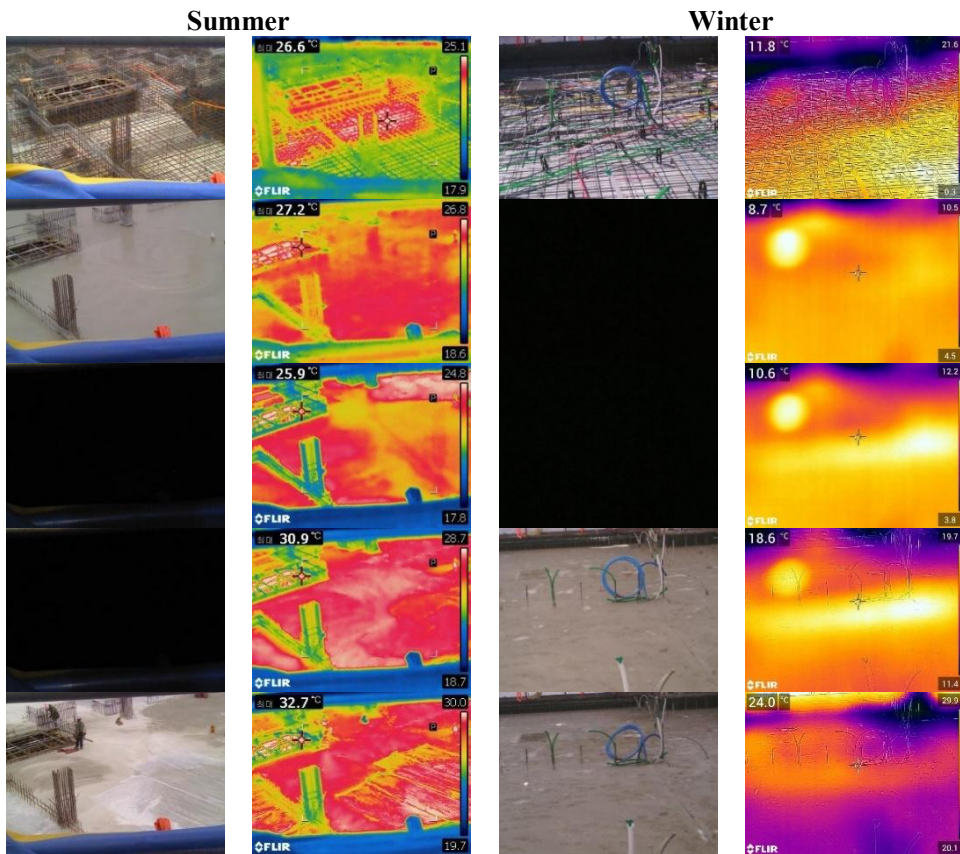
Figure 43: Examples of RGB and their corresponding thermal images of concrete section taken from the construction sites in 24 hours of a day with 6 hours of interval.
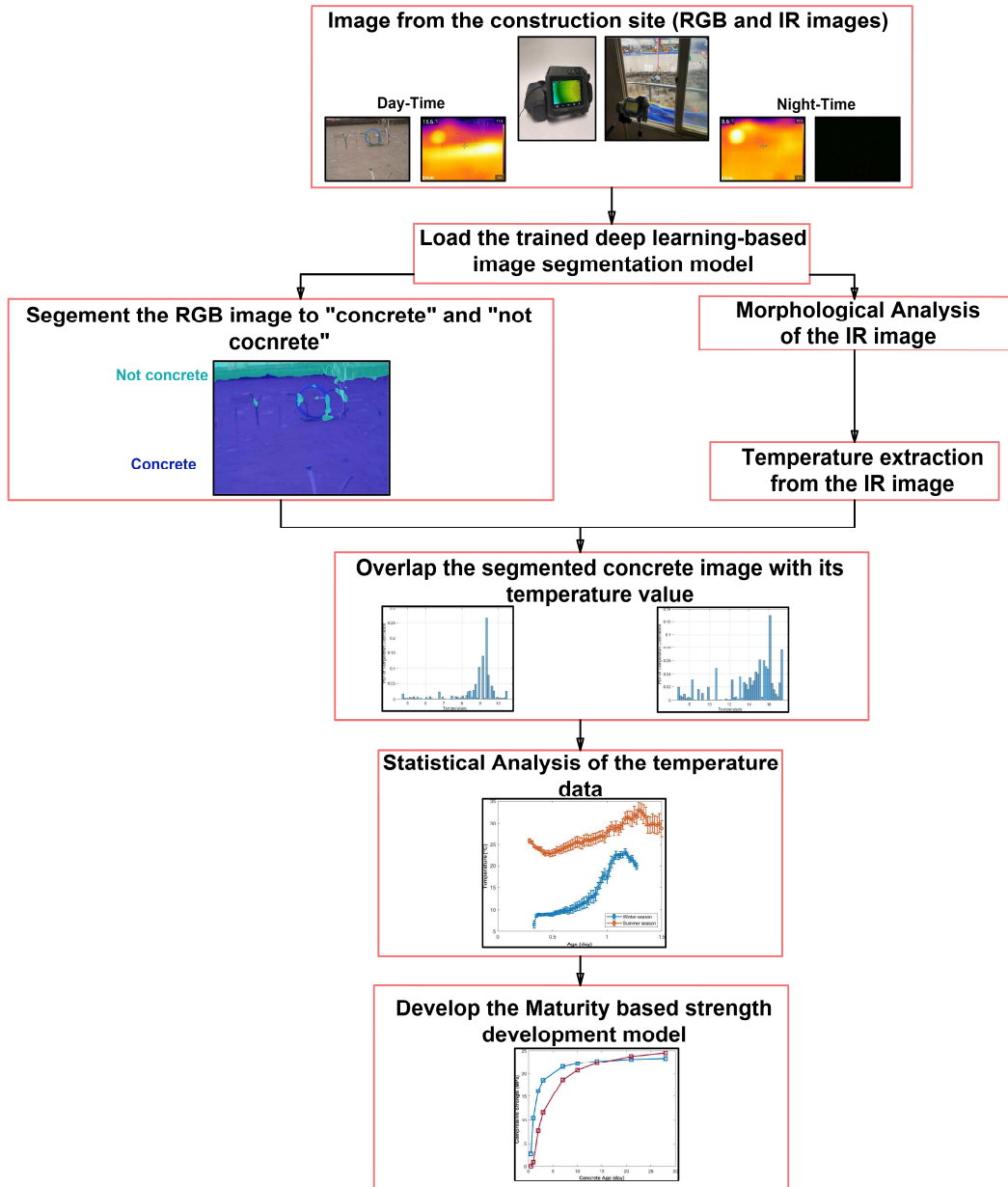
Figure 44: general process of the proposed method

One of the other objectives of this work was to compare the output of the proposed method with the actual results. As a result, cylindrical samples for the specified concrete sections were gathered concurrently with applying the proposed method. Table 6 displays the concrete

specifications for both summer and winter situations. These samples were cured under the same circumstances as the selected concrete sections to maintain uniformity. The compressive strength of the materials was then evaluated for a period ranging from 12 hours to 28 days. Finally, the findings were examined and compared to the proposed technique, presented in the next section.

Table 6: Specifications of the concrete casted on site

Description	Winter	Summer
Concrete Temperature (oc)	18	19.2
Air temp (oc)	2	28.2
Slump (mm)	135	160
Air content (%)	4.80	4.40
chloride content (kg/mm3)	0.025	0.06
Expected strength, 28 Days (MPa)	27	26
Max. Aggregate Size (mm)	25	25

5.5. Result and Discussion

5.5.1. Comparison of thermocouples and thermal imaging camera

During the laboratory experiment, one of the aspects that was looked into was how the data from the thermocouple and the thermal imaging camera differed. In Figure 45, the temperature data from both testing methods were displayed alongside the ambient temperature of the surrounding area. The data appears to show some convergence during the first 9-hour interval, with an average temperature differential of ±1.25°C. However, this gap tends to widen at the next 12-hour period, reaching +3.06°C to be exact. One of the reasons is that most of the cement hydration occurs at the first 9-12 hours, which during this time, the heat evolved is relatively higher. As a result, the internal and surface temperatures are nearly identical, with minor variations.

The other reason is the surrounding environment (the curing temperature). The first 9-hour time span occurs during the day, where the ambient temperature is moderate, and the inside and surface temperatures are similar. However, the second period occurred at night, causing the air

75

temperature to drop dramatically. As a result, a significant amount of heat is lost through convection between the concrete's surface temperature and the atmosphere, which lowers its value compared to the internal temperature. But yet another thing that was noted was that the temperature drop was proportional to the temperature difference between the internal and surface, as well as the external (atmosperic) and surface temperature. This proportionality is helpful in scaling up the surface temperature as an input to the maturity function. After 24 hours after the casting, a temperature differential of 1.07°C between internal and surface and +0.93°C between internal and surrounding was recorded, indicating that the atmospheric condition of the concrete is a critical factor in determining the surface temperature.

Overall, between a three-day time interval, the difference between the thermocouple and the IR images is ±1.86°C. This variation can be considered acceptable regarding their applicability in the maturity equations. Thus, the thermal imaging (IR) camera is an appropriate alternative for the thermocouple as non-destructive testing (NDT) method for in-place concrete strength prediction.
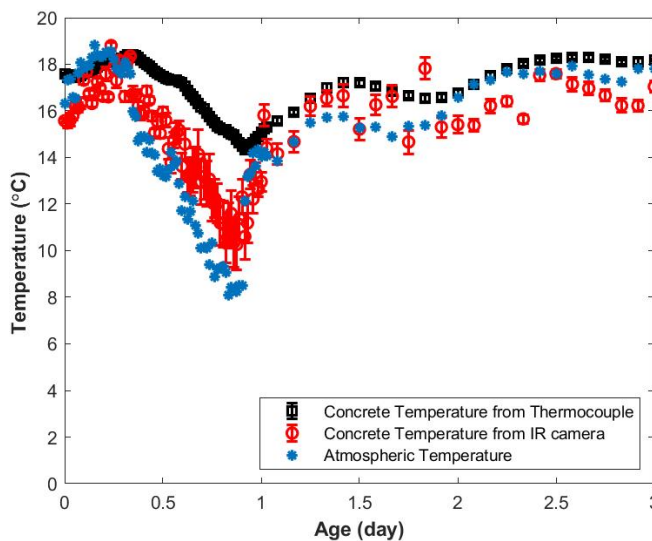
Figure 45: Temperature variation between thermocouples, thermal imaging camera, and atmospheric temperature

In addition to temperature data analysis, maturity functions were also developed based on the temperature data collected from both testing instruments. As noted in previous sections, building codes (ASTM C1074 and ACI 228) offer two methods for calculating maturity functions: the maturity index (temperature-time factor) method and the equivalent age method. In this experiment, The equivalent age (Arrhenius) approach was used to determine the correlation between time and temperature. The building standards recommend a specific temperature of 20°C, but other choices are also viable as long as they are reported. The other important variable that was an input in the equivalent age calculation was the activation energy,E. This energy is the minimum energy required to start any chemical reaction, for this case, cement hydration. This energy depends on several factors, such as the type of cement, the curing conditions of the concrete, and the type of supplementary materials used. Based on this, the ASTM C1074 recommends values ranging from 38,000 to 45,000 J/mol for a Type I cement. Therefore an average value of 41,500 J/mol Equation 6 and 7 show the basic equations for developing equivalent age and strength development based on temperature.

$$t_e = \sum_0^n A e^{\frac{-E}{RT}} \Delta t \tag{6}$$

$$S = s_u \left(\frac{k(t - t_o)}{1 + k(t - t_o)} \right) \tag{7}$$

Figure 46 (a) shows the relation between the compressive strength with the equivalent age at a specific temperature of 20°C. It can be seen that a smaller temperature gap has a little effect at the early ages and tends to increase in the long run in the strength prediction of concrete. This is because the maturity equation is a cumulative calculation process in which the deviation of the variable at early stages will not have a notable difference. Nevertheless, as time passes, it tends to vary greatly. The upside of this is that the thermal imaging-based prediction is a very

77

conservative method of production of the compressive strength of concrete with a wide range of a factor of safety.
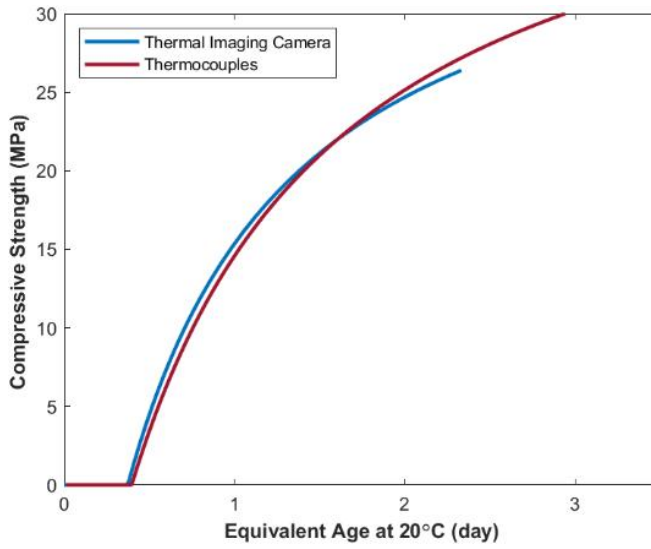
(a)
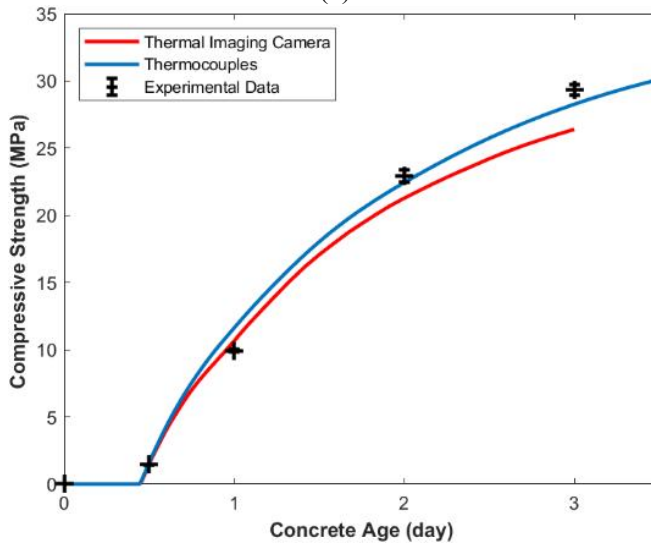
(b)

Figure 46: Relationship between actual concrete age and compressive strength with equivalent age

Figure 46 (b) depicts the predicted and actual experimental values obtained from the test cylinders. According to the figure, the data from both thermocouples and thermal pictures were

fairly similar to the experimental data, implying that thermal images are equally effective in generating the precise concrete temperature. This, in turn, suggests that maturity-based strength estimation using temperature data from thermal images is one way of predicting the actual strength with acceptable accuracy.

5.5.2. *Combined effect of concrete area prediction and Strength Prediction based on maturity method*

The second thing that was investigated was the integration of deep learning-based image detection and IR-based temperature extraction for determining in-place concrete strength. From the initial findings, it is enough to say that the thermal imaging camera can effectively replace other forms of temperature sensors and detectors. As previously stated in the preceding part, a thermal camera was installed on the construction site, and the experiment was carried out. The temperature pattern in both winter and summer is depicted in Figure 47. As expected, a considerable temperature differential was observed due to the weather conditions. The flow of the temperature history is similarly consistent with the expectations indicated in the building standards. Because the rate of cement hydration increases with age, the amount of heat emitted increases, resulting in a temperature rise, as seen in the figure.
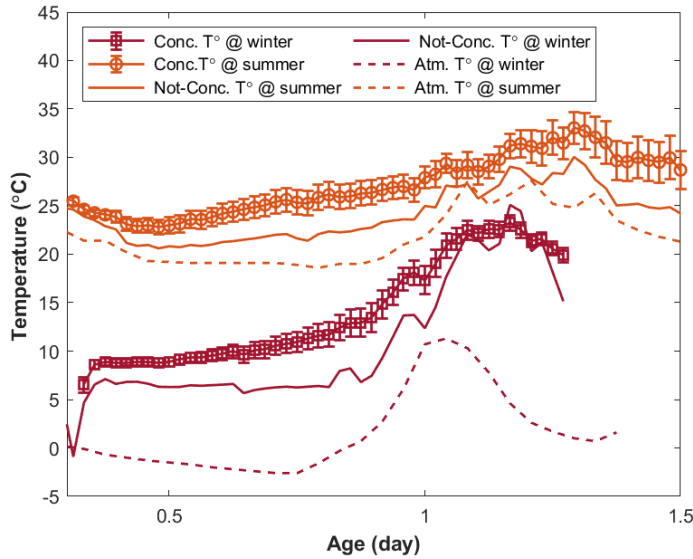
Figure 47: Temperature history in summer and winter conditions

After the temperature data is analyzed, the maturity-based strength development was drawn for both weather conditions. Since the material composition is different from the laboratory experiments, different values are assigned for the variables for calculating the equivalent age and the strength development. The same reference temperature of 20°C was used as a datum temperature to derive the equivalent age and strength development equations. Table 7 shows some of the parameters that were used for both weather conditions as they are crucial in developing the prediction equations.

Table 7: Strength prediction parameters

Parameters	Summer	Winter	Remark
Specific temperature, Tr	20°C	20°C	As per ASTM C1074
Activation energy, E	42,300 J/mol	42,300 J/mol	(Soutsos, Kanavaris, & Hatzitheodorou, 2018)
Rate constant, R	8.314 J/mol-K	8.314 J/mol-K	As per ASTM C1074
Rate constant, k	0.24(1/day)	0.114(1/day)	Regression Analysis, (H.-M. Yang et al., 2020), (Soutsos et al., 2018)
Start of strength development, to	0 (day)	0.53 (day)	
Ultimate strength, Su	24 MPa	27 MPa	
Constants for strength development	a = 4.2936 b = 2.9139	a = 5.8269 b = -0.6553	

Equations 1 and 2 show the maturity-based strength development equation used for both weather conditions. Some of the parameters used for both weather conditions are crucial in developing the prediction equations. As shown in Figure 48 the data from the proposed method is closely aligned with the experimental data, especially at the early ages of the concrete. Therefore, this method is applicable in predicting the strength of concrete based on the surface temperature data it acquires from the thermal imaging camera.

(a)

(b)
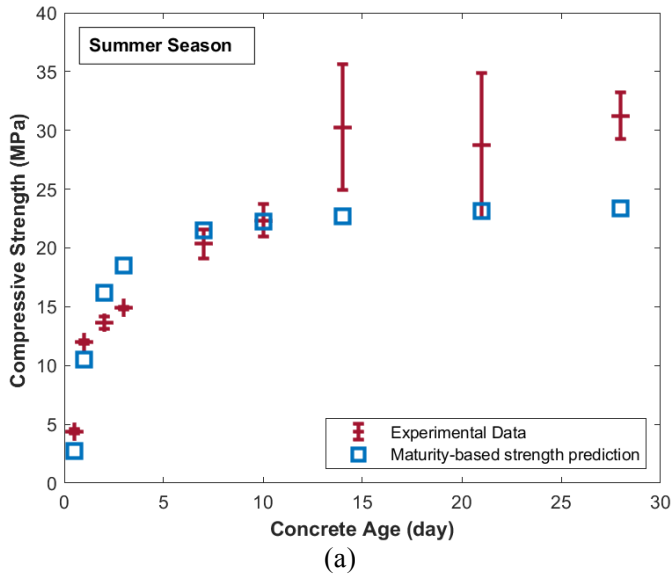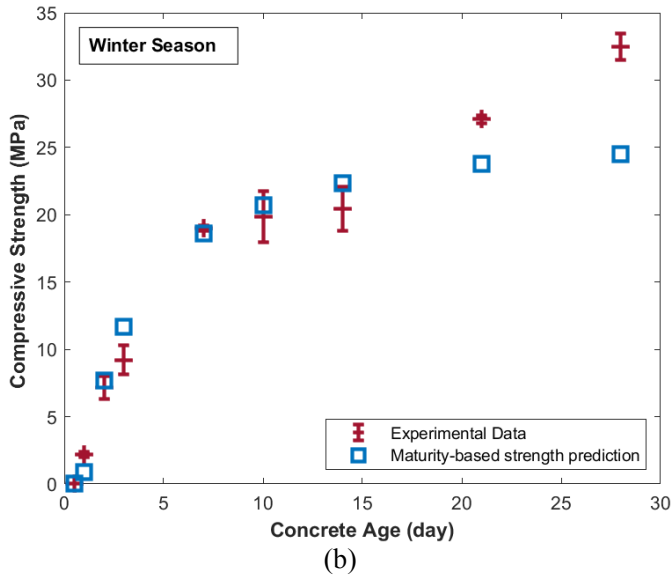
Figure 48: Relationship between experimental data based to maturity function-based strength development in summer condition

5.5.3.　　　Comparison in formwork removal time

Construction scheduling and proper timing are crucial in the construction industry, for which time and money are the prominent factors to consider. Therefore, to maintain and fast track these conditions, different variables have to be considered. From this formwork striking or removing tome is a very important point. During this investigation, one of many issues investigated was how the information gathered from the thermal imaging cameras helps optimize the duration for formwork removal.

Currently, different building standards recommend different time frames and minimum strength requirements for the safe removal of formworks. Table 8 shows some of the most used building standards and their recommendations. Most of the recommendations from the building standards depend on a number of factors. Some of the major ones are the type of cement used, the expected concrete grade, the pozzolanic or supplementary materials used, and the main area of this study, the temperature of the concrete. In addition, the quality and consistency of construction and

sectional dimensions of the structural members are also other issues considered in formulating these standards.

These standards are very beneficial in maintaining the uniformity and reliability of the construction work. They are sometimes helpful in covering the construction site's mismanagement and workmanship errors. However, usually, there is a misunderstanding in knowing whether the specific minimum time or minimum strength governing factor. In the construction site, to avoid confusion, it is usually the minimum time requirement that is applied, which in most cases, the strength will be achieved way earlier than the minimum time. This method is usually beneficial, especially in small to medium scale constructions where the formwork removal time slightly affects the whole construction schedule. However, when massive projects are performed, time will become a valuable variable more than other issues. Therefore, the concrete in the construction site might achieve its required attainable strength capacity pretty early than expected before the standard formwork removal time was reached. At this stage, knowing the exact time where the required strength is achieved will be crucial because it will help avoid scheduling problems and increase the output of the work based on time and money. Therefore, one of the advantages of this method is to check if the strength expected is achieved by monitoring its surface temperature.

Table 8: Minimum formwork striking strength and time according to different building standards

Building Standards	Structural Members	Minimum Strength	Minimum Time	Requirements
Eurocode 2 (BS 8110-1-1997) (BSI)	Members in compression, e.g.,columns,to protect against frost damage	5 MPa	-	This table can be applied to PC and SRPC of higher cement strength classes
	Flexural members	"10MPa or twice the stress to which it will be subjected to whichever is greater,	-	

83

				provided striking at that time will not result in unacceptable deflection".	
ACI 347-14 (ACI, 2014)	Members in compression, such as walls, columns, sides of beams and girders	-	12hrs	These periods are applicable provided that the surface temperature is ≥10°C. if not, they are subjected to be amended by the person in-charge	
	Flexural members	-	3 – 14 days Depending on the live and dead load, and the span of the members		
KASS 5 ,JASS 5 (Specification, 2018)	Members in compression, such as foundations, walls, columns, sides of beams and girders	5 MPa	2 – 5 days, if the mean air temperature is ≥20°C 3 – 8 days, if the mean air temperature is 10°C ≤ t < 20°C	These periods are applicable provided that the surface temperature is ≥10°C. In addition, the time frame also depends on the concrete quality and type of cement.	
	Flexural members	50% of the specified design strength	4 – 8 days, if the mean air temperature is ≥20°C 5 – 10 days, if the mean air temperature is 10°C ≤ t < 20°C		
CIRIA Report. 136 (Harrison, 1995)	Slabs and beam soffits	Varies Depends on the grade of the concrete	Depends on the concrete grade, type of cement, and the section dimensions.	These periods are applicable provided that the placing temperature is at least the mean air temperature or 5°C. whichever is greater.	
	Columns and sides of beams	≥2MPa To avoid frost action	-		

In order to check how and in what capacity it monitors, the data gathered from the summer and winter construction were used. After their strength is predicted by the maturity function based on the surface temperature of the concrete, the formwork striking time was calculated from them.

Figure 49 shows the standards of formwork striking time in relation to the building standards mentioned in the table above.
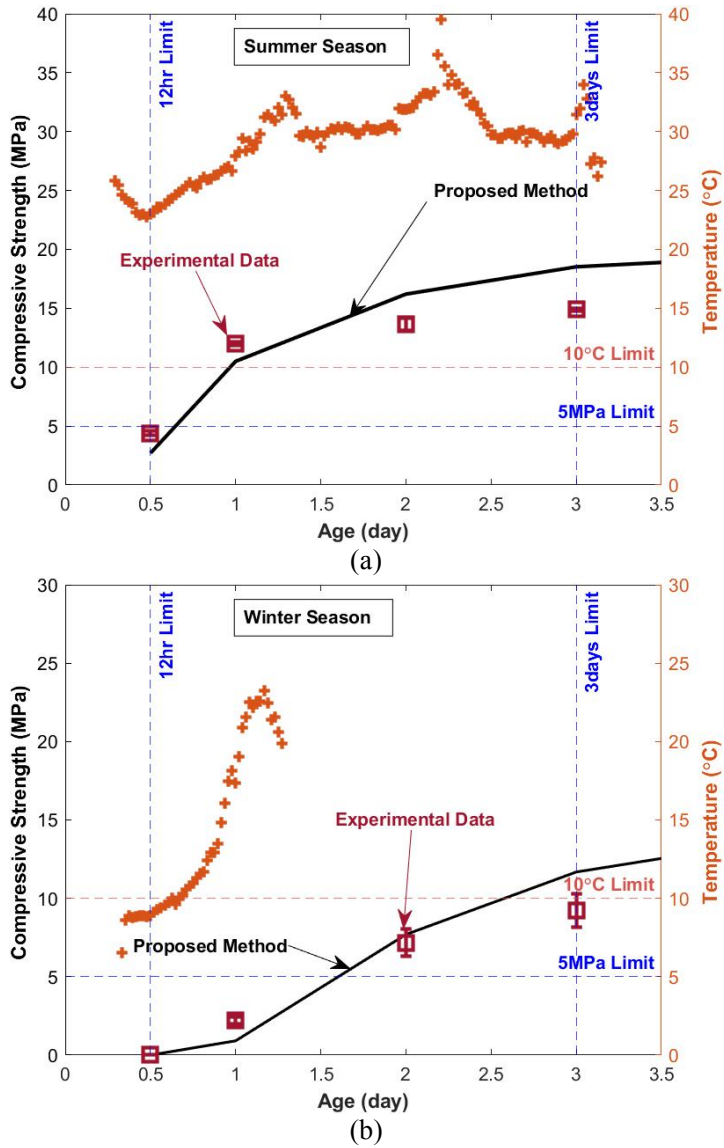
Figure 49: Variation in formwork removal time and strength for both summer and winter construction

In the summer construction, the 5MPa strength limit proposed by BS 8110 (BSI), JASS 5 (Specification, 2018), and KASS 5 was achieved before the concrete age reached three days both

in the experimental and the temperature based prediction. The same is true also for the winter condition. In the summer experiment, 5MPa was reached at 0.503 and 0.587 days based on experimental data and maturity-based prediction. Whereas in the winter, 1.584 and 1.632 days were when the concrete reached the required strength. This shows that the predicted strength development based on the surface temperature is a good way of predicting the time for formwork removal without compromising both the strength and time requirement.

In addition to this, since the thermal imaging camera gives the surface temperature of the concrete, it can also help maintain the general requirement that has to be maintained by the building standards. For example, the ACI 347 (ACI, 2014) requires that the mean air or concrete temperature be greater or equal to 10°C in order to apply the 12hr formwork removal time. Therefore, the proposed method also helps to monitor the surface temperature of the concrete to optimize this process. As it can be seen, during the summer experiment, the temperature is far more than 10°C throughout the experiment which the 12hr removal can be applied. But in the winter season the first 18hr was under the limit but after that the temperature increased exponentially which the 12hr is calculated from this point.

In general, the proposed method has shown to be more effective in monitoring and predicting the strength development of concrete. In addition to this, determining the formwork removal time was also shoe how optimized by using this technique. These additional advantages are crucial in increasing the productivity and efficiency of a construction project.

6. CONCLUSION

6.1. Conclusion

The implementation of computer-based detection methods progressed exponentially from filtering and thresholding techniques such as histogram analysis and brightness adjustment into deep learning-based systems. These days, deep learning, specifically convolutional neural networks (CNNs)-based image processing procedures, is being adapted to every aspect of concrete technology. In addition to deep learning, Infrared thermography has advanced dramatically in a concise period of time. Significant advancements have been made in various sectors, including in the concrete industry. It has evolved as excellent monitoring and defects detection technique or non-contact, non-invasive real-time temperature monitoring of targets throughout the years.

Based on this, a novel method was proposed that incorporates deep learning and a temperature-based maturity method to predict the mechanical properties of concrete. To accurately estimate the strength of concrete, this system integrates two interdependent schemes: deep learning-based concrete area detection and temperature monitoring. The first scheme used state-of-the-art architectures with a convolutional neural network (CNN) backbone to classify the area covered by concrete in an image. By adjusting a number of factors, these models were compared and verified to select the model with the best performance and accuracy. As a result, a DeepLab V3+ architecture with a Resnet backbone CNN has proven to perform better. This model was then used to segment the concrete area, and the segmented area was then combined with temperature data to create the maturity function. The second scheme used an imaging (IR) camera to detect the concrete's temperature. The method's reliability was first tested by adjusting some parameters such as the distance from the region of interest (ROI) and the environment's weather conditions. The camera is then used to collect concrete temperature data on the construction site.

Finally, the two techniques were combined to create a strength development function that can be used to estimate concrete strength at any time. Based on these investigations, an accuracy of more than 80% in strength prediction was achieved based on this proposed method compared to the actual experimental results. This suggests that proposed method can predict the strength of concrete which in turn help in determining the formwork removal time. It is also crucial in other aspects, the most essential of which increases the efficiency of construction projects regarding both time and economic standpoint. Furthermore, this method demonstrates the proclivity of deep learning architectures and their use to objectively judge whether something is concrete or not, as well as how an infrared camera can be utilized as non-destructive testing (NDT) tool to solve temperature-related issues.

6.2. Recommendations

This study combined two different methods that are crucial in developing a novel method of estimating the strength of concrete. This method was proven to be effective in predicting the strength of concrete at any time with an acceptable range of confidence. However, some key points have to be improved further to increase the model's applicability. An accurate and sufficient training data set for the deep learning image segmentation method is essential in developing a model. Increasing the number and the depth of the training dataset will create a broad scope for the trained model to recognize different aspects of the same ROI. In addition, introducing classes other than "concrete" and "not concrete" has to be implemented. This will help increase the model's depth in segmenting areas that resemble concrete sections but are not concrete. Training on a variety of hyperparameters with a higher GPU and adapting data augmentation methods other than geometric translations can also help discover a model with better accuracy.

Further investigation on the reliability of thermal imaging cameras is also required, particularly in some of the aspects that influence it. The effect of focusing, the quality of the thermal sensors on the camera, and interruption by obstacle are some of the variables that have to be studied. Finally, an easy and understandable interface system that incorporates both methods is necessary for applicability in any construction project. This eliminates the hassle of intricate computations, allowing the user to focus on the task at hand.

6.3. Future Studies

As discussed before, there is a lack of sufficient investigations and data in this method which incorporated deep learning and concrete strength prediction. Therefore, this research will be a significant asset in future studies. Based on this, these listed below are some of the plans for future investigations.

- Investigate other deep learning neural networks other than CNNs. This will be beneficial in increasing the model's scope and finding any model that can perform better than the one used in this study.

- Develop a structured graphical user interface (GUI) that incorporates the whole framework of the proposed method. This helps the user apply it in any construction project without any problems. Furthermore, any person without the knowledge can use it, broadening its versatility.

- Transform the image-based segmentation to real-time or motion-based segmentation. This will help in increasing the monitoring capability of the model.

- Apply this proposed model to other aspects of constructions, such as workforce monitoring and defect detection.

- Perform economic and feasibility studies to see if the method clearly is crucial in cost-saving as it is in increase the efficiency of construction and monitors the condition of the concrete.

REFERENCES

ACI. (2014). *ACI 347R-14. Guide to Formwork for Concrete*.

Agarap, A. F. (2018). Deep learning using rectified linear units (relu). (1803.08375).

Akinosho, T. D., Oyedele, L. O., Bilal, M., Ajayi, A. O., Delgado, M. D., Akinade, O. O., & Ahmed, A. A. (2020). Deep learning in the construction industry: A review of present status and future innovations. *Journal of Building Engineering*, 101827.

Al-Hadhrami, L. M., Maslehuddin, M., Shameem, M., & Ali, M. R. (2012). Assessing concrete density using infrared thermographic (IRT) images. *Infrared Physics & Technology, 55*(5), 442-448.

Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., . . . Asari, V. K. (2019). A state-of-the-art survey on deep learning theory and architectures. *Electronics, 8*(3), 292.

ASTM, A. (2013). *C192/C192M-13a-Standard practice for making and curing concrete test specimens in the laboratory*. Retrieved from

ASTM, S. (2004). Standard practice for estimating concrete strength by the maturity method. *ASTM C, 1074*, 1074-1093.

Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence, 39*(12), 2481-2495.

Barreira, E., Almeida, R. M., & Ferreira, J. P. (2017). Assessing the humidification process of lightweight concrete specimens through infrared thermography. *Energy Procedia, 132*, 213-218.

BSI, B. S. 8110-1:(1997). *Structural use of concrete–Part 1: Code of practice for design and construction*.

Carino, N. (1981). Temperature Effects on the Strength–Maturity Relation of Mortar, Report No. NBSIR 81-2244.

Carino, N. J. (1984). The maturity method: theory and application. *Cement, concrete and aggregates, 6*(2), 61-73.

Carino, N. J., Lew, H., & Volz, C. K. (1983). *Early age temperature effects on concrete strength prediction by the maturity method.* Paper presented at the Journal Proceedings.

Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2017). Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence, 40*(4), 834-848.

Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). *Encoder-decoder with atrous separable convolution for semantic image segmentation.* Paper presented at the Proceedings of the European conference on computer vision (ECCV).

Chengju, G. (1989). Maturity of concrete: method for predicting early-stage strength. *Materials Journal, 86*(4), 341-353.

Chigozie Nwankpa, W. I., Anthony Gachagan, Stephen Marshall. (Nov, 2018). *Activation Functions: Comparison of trends in Practice and Research for Deep Learning.* Retrieved from

Corvaglia, P., & Largo, A. (2008). *IRT survey for the quality control of FRP reinforced rc structures.* Paper presented at the Proceedings of the QIRT.

Farahani, A., Pourshojae, B., Rasheed, K., & Arabnia, H. R. (2021). A Concise Review of Transfer Learning. *arXiv preprint arXiv:2104.02144.*

Feng, J., Feng, X., Chen, J., Cao, X., Zhang, X., Jiao, L., & Yu, T. (2020). Generative adversarial networks based on collaborative learning and attention mechanism for hyperspectral image classification. *Remote Sensing, 12*(7), 1149.

Gade, R., & Moeslund, T. B. (2014). Thermal cameras and applications: a survey. *Machine vision and applications, 25*(1), 245-262.

Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., & Garcia-Rodriguez, J. (2017). A review on deep learning techniques applied to semantic segmentation. *arXiv preprint arXiv:1704.06857.*

Geiker, M. (1983). *Studies of Portland cement hydration by measurements of chemical shrinkage and a systematic evaluation of hydration curves by means of the dispersion model.* Technical University of Denmark,

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., . . . Bengio, Y. (2014). Generative adversarial networks. *arXiv preprint arXiv:1406.2661*.

Hansen, P. F., & Pedersen, E. (1984). *Curing of concrete structures*: BKI.

Hansen, P. F., & Pedersen, E. J. (1977). *Maturity computer for controlled curing and hardening of concrete*. Retrieved from

Harrison, T. (1995). *Formwork Striking Times: Criteria, Prediction and Methods of Assessment*: CIRIA London.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). *Deep residual learning for image recognition.* Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.

Helland, S. (1987). *Temperature and Strength Development in Concrete with W/C less than 0.40.* Paper presented at the Proceedings of Utilization of High Strength Concrete. Stavanger, Norway.

Hernández Alonso, A. (2019). Impact of climate conditions on the use of infrared thermography for the inspection of coated and uncoated unreinforced concrete samples.

Huang, Z., Pan, Z., & Lei, B. (2017). Transfer learning with deep convolutional neural network for SAR target classification with limited labeled data. *Remote Sensing, 9*(9), 907.

Jang, Y., Ahn, Y., & Kim, H. Y. (2019). Estimating compressive strength of concrete using deep convolutional neural networks with digital microscope images. *Journal of Computing in Civil Engineering, 33*(3), 04019018.

Jonasson, J.-E. (1985). *Early strength growth in concrete-preliminary test results concerning hardening at elevated temperatures.* Paper presented at the The 3^< rd> RILEM Symposium on Winter Concreting, Espoo Finland, 1985.

Kaur, P., & Kant, R. (2014). A review on: Comparison and analysis of edge detection techniques. *Int. Journal of Engineering Research and General Science, 2*(3).

Kee, C. F. (1971). *Relation between strength and maturity of concrete.* Paper presented at the Journal Proceedings.

Knudsen, T. (1984). The dispersion model for hydration of Portland cement I. General concepts. *Cement and Concrete Research, 14*(5), 622-630.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). *Imagenet classification with deep convolutional neural networks.* Paper presented at the Advances in neural information processing systems.

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

Lee, S., Ahn, Y., & Kim, H. Y. (2020). Predicting concrete compressive strength using deep convolutional neural network based on image characteristics. *CMC-COMPUTERS MATERIALS & CONTINUA, 65*(1), 1-17.

Lew, H., & Reichard, T. W. (1978). Prediction of strength of concrete from maturity. *Special Publication, 56*, 229-248.

Liu, X., Deng, Z., & Yang, Y. (2019). Recent progress in semantic image segmentation. *Artificial Intelligence Review, 52*(2), 1089-1106.

Long, J., Shelhamer, E., & Darrell, T. (2015). *Fully convolutional networks for semantic segmentation.* Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.

Naik, T. R. (1985). Maturity functions for concrete cured during winter conditions. In *Temperature Effects on Concrete*: ASTM International.

Nandhini, K., & Karthikeyan, J. (2021). The early-age prediction of concrete strength using maturity models: a review. *Journal of Building Pathology and Rehabilitation, 6*(1), 1-10.

Nassif, A. B., Shahin, I., Attili, I., Azzeh, M., & Shaalan, K. (2019). Speech recognition using deep neural networks: A systematic review. *IEEE Access, 7*, 19143-19165.

Nurse, R. (1949). Steam curing of concrete. *Magazine of Concrete Research, 1*(2), 79-88.

Nwankpa, C., Ijomah, W., Gachagan, A., & Marshall, S. (2018). Activation functions: Comparison of trends in practice and research for deep learning. *arXiv preprint arXiv:1811.03378.*

Omar, T., & Nehdi, M. L. (2016). *Application of passive infrared thermography for the detection of defects in concrete bridge elements.* Paper presented at the Proceedings of the International Conference of the Transporation Association of Canada, Toronto, ON, Canada.

Pierrehumbert, R. T. (2011). *Infrared radiation and planetary temperature.* Paper presented at the AIP Conference Proceedings.

Pla-Rucki, G. F., & Eberhard, M. O. (1995). Imaging of reinforced concrete: State-of-the-art review. *Journal of infrastructure systems, 1*(2), 134-141.

Plowman, J. (1956). Discussion: Maturity and the strength of concrete. *Magazine of Concrete Research, 8*(24), 169-183.

Qiuyu Zhu, Z. H., Tao Zhang, and Wennan Cui. (April 24, 2020). Improving Classification Performance of Softmax Loss Function Based on Scalable Batch-Normalization. *Applied Sciences, 10, 2950*, 1-8.

Radford, A., Metz, L., & Chintala, S. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*.

Ranzato, M. A., Boureau, Y.-L., & LeCun, Y. (2007). Sparse feature learning for deep belief networks. *Advances in neural information processing systems, 20*, 1185-1192.

Rao, D. P. (2008). Infrared thermography and its applications in civil engineering. *The Indian concrete journal, 82*(5), 41-50.

Rastrup, E. (1954). Heat of hydration in concrete. *Magazine of Concrete Research, 6*(17), 79-92.

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review, 65*(6), 386.

Sarawade, A. A., & Charniya, N. N. (2018). *Infrared thermography and its applications: a review.* Paper presented at the 2018 3rd International Conference on Communication and Electronics Systems (ICCES).

Saul, A. (1951). Principles underlying the steam curing of concrete at atmospheric pressure. *Magazine of Concrete Research, 2*(6), 127-140.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Sirca Jr, G. F., & Adeli, H. (2018). Infrared thermography for detecting defects in concrete structures. *Journal of Civil Engineering and Management, 24*(7), 508-515.

Smirnov, B. (2019). Infrared Radiation in the Energetics of the Atmosphere. *High temperature, 57*(4), 573-595.

Solla, M., Lagüela, S., Fernández, N., & Garrido, I. (2019). Assessing rebar corrosion through the combination of nondestructive GPR and IRT methodologies. *Remote Sensing, 11*(14), 1705.

Soutsos, M., Kanavaris, F., & Hatzitheodorou, A. (2018). Critical analysis of strength estimates from maturity functions. *Case Studies in Construction Materials, 9*, e00183.

Specification, J. A. S. (2018). JASS 5 Reinforced Concrete Work. *Architectural Institute of Japan (AIJ): Tokyo, Japan*.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research, 15*(1), 1929-1958.

Szandała, T. (2021). Review and Comparison of Commonly Used Activation Functions for Deep Neural Networks. In *Bio-inspired Neurocomputing* (pp. 203-224): Springer.

Tabian, I., Fu, H., & Sharif Khodaei, Z. (2019). A convolutional neural network for impact detection and characterization of complex composite structures. *Sensors, 19*(22), 4933.

Tank, R. C., & Carino, N. J. (1991). Rate constant functions for strength development of concrete. *Materials Journal, 88*(1), 74-83.

Usamentiaga, R., Venegas, P., Guerediaga, J., Vega, L., Molleda, J., & Bulnes, F. G. (2014). Infrared thermography for temperature measurement and non-destructive testing. *Sensors, 14*(7), 12305-12348.

Valeria Maeda-Gutiérrez, C. E. G.-T., Laura A. Zanella-Calzada, José M. Celaya-Padilla, Jorge I. Galván-Tejada, Hamurabi Gamboa-Rosales, Huizilopoztli Luna-García, Rafael Magallanes-Quintanar, Carlos A. Guerrero, Méndez. (2020). Comparison of Convolutional Neural Network Architectures for Classification of Tomato Plant Diseases. *Applied Sciences*.

Verbeck, G. J. (1968). *Structures and physical properties of cement paste*. Paper presented at the 5th International Congress on the Chemistry of Cement, 1968.

Walraven, J. C. (2012). *Model Code 2010-Final draft: Volume 1* (Vol. 65): fib Fédération internationale du béton.

Weaver, J., & Sadgrove, B. (1971). *Striking times of formwork-tables of curing periods to achieve given strength*. Retrieved from

Weisstein, E. (2019). Logistic Equation 2017. *URL http://mathworld. wolfram. com/LogisticEquation. html. Accessed*, 15-10.

Xincong Yang, H. L., Yantao Yu, Xiaochun Luo & Ting Huang. (2018). Automatic Pixel-Level Crack Detection and Measurement Using Fully Convolutional Network. *Computer-Aided Civil and Infrastructure Engineering*.

Yang, H.-M., Kwon, S.-J., Myung, N. V., Singh, J. K., Lee, H.-S., & Mandal, S. (2020). Evaluation of strength development in concrete with ground granulated blast furnace slag using apparent activation energy. *Materials, 13*(2), 442.

Yang, X., Li, H., Yu, Y., Luo, X., Huang, T., & Yang, X. (2018). Automatic pixel☐level crack detection and measurement using fully convolutional network. *Computer☐ Aided Civil and Infrastructure Engineering, 33*(12), 1090-1109.

Yu, D., Wang, H., Chen, P., & Wei, Z. (2014). *Mixed pooling for convolutional neural networks*. Paper presented at the International conference on rough sets and knowledge technology.

Yu Songa, Z. H., Chuanyue Shena, Humphrey Shic, David A. Langed. (2020). Deep learning-based automated image segmentation for concrete petrographic analysis. *Cement and Concrete Research, 135,106118*, 1-13.

PUBLICATIONS

1. Mesfin, W. M., Cho, S., Lee, J., Kim, H. K., & Kim, T. (2021). Deep-Learning-Based Segmentation of Fresh or Young Concrete Sections from Images of Construction Sites. *Materials*, *14*(21), 6311. (*Editor's Choice*)

ACKNOWLEDGEMENTS

First, I would like to thank the almighty God for letting me through all the difficulties. I thank him for giving me the strength and ability to learn, understand and complete this research.

Second, I want to express my deepest gratitude to my esteemed advisor, Professor Hyeong-Ki Kim (Ph.D.), for his sincerity and encouragement throughout the course of my master's degree. His immense knowledge and plentiful support have guided me in my academic research and daily life. I am thankful for the extraordinary experiences he arranged for me and for providing opportunities for me to grow professionally. This thesis would not have been possible without his guidance from the initial step, and it is an honor to learn from him.

Besides my advisor, I would like to thank my dissertation evaluation committee members, Prof. Chang-Geun Cho and Prof. Taehoon Kim, for their valuable comments and suggestions on my research. I would also like to give special gratitude to Dr. Hyo-Kyoung Lee for her treasured support, which was influential in my research as well as shaping my stay here in South Korea. I would like to extend my sincere thanks to my friends and Lab mates -Abel Shiferaw, Gebremichael Liyew, Elsa Tadege, Million Tafesse, and Jaesung Ahn, for their valuable support, inspiration, and assistance all through my studies.

Finally, my profound appreciation goes out to my beloved family and friends who directly or indirectly helped me with their prayers, encouragement, and emotional support throughout my years of study. This accomplishment would not have been possible without them. Thank you.