WOOD DEFECT DETECTION AND CLASSIFICATION USING DEEP LEARNING

YAP YI REN

FACULTY OF ENGINEERING
UNIVERSITY OF MALAYA
KUALA LUMPUR

2019

WOOD DEFECT DETECTION AND CLASSIFICATION USING DEEP LEARNING

YAP YI REN

SUBMITTED TO THE FACULTY OF ENGINEERING UNIVERSITY OF MALAYA IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER MECHATRONICS OF ENGINEEING

FACULTY OF ENGINEERING
UNIVERSITY OF MALAYA
KUALA LUMPUR

2019

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: YAP YI REN

Matric No: KQF170006

Name of Degree: MASTER OF MECHATRONICS ENGINEERING

Title of Project Paper/Research Report: WOOD DEFECT DETECTION AND CLASSIFICATION USING DEEP LEARNING

Field of Study:

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date:

Subscribed and solemnly declared before,

Witness's Signature Date:

Name:

Designation:

ABSTRACT

In the timber and wood industry, natural defects on wood and timber are always one of the main issues. In many timber and wood industry, the quality assurance of the board is still controlled by a human. This is because the defects can vary in many ways likes amount, shape, area and colour. The quality checking process can be very tedious and worker may easily makes mistakes in judgement. To reduce the human mistakes, this study focuses on designing a wood defect detection and classification by using the artificial intelligence technique of Convolutional Neural Network (CNN) in MATLAB. Convolutional Neural Network (CNN) is one of the deep neural networks used in two-dimensional data. It mainly used to classify objects in images, cluster them by similarity and execute object recognition. This technology can identify faces, street sign, tumours, human, etc. The CNN model consists of input images, Convolution Layers, Activation Function (ReLU), Pooling, Fully Connected layers and Output layer. Three sets of input data such as Knots, Crack and Normal are prepared for training and testing the CNN model by using different parameters. The results of the different configurations are compared and analysed. The accuracy of overall classification is 97.2%.

ABSTRAK

Dalam industri kayu, kecacatan semula jadi pada kayu adalah salah satu isu utama. Dalam banyak industri kayu, kualiti lembaga masih dikawal oleh manusia. Ini kerana kecacatan kayu boleh berbeza-beza daripada jumlah, bentuk, ruang dan warna. Proses oenerujsaab kuali boleh menjadi sangat membosankan dan ia boleh dijangka bahawa pekerja melakukan kesilapan. Untuk mengurangkan kesilapan manusia, kajian ini memberi tumpuan kepada reka bentuk pengesanan dan klasifikasi kecacatan kayu dengan menggunakan Teknik Rangkaian Konvolusi Neural (CNN) dalam MATLAB. Rangkaian neural convolutional (CNN) adalah salah satu rangkaian saraf yang digunakan dalam data dua dimensi. Ia digunakan untuk mengelaskan objek dalam imej, cluster mereka dengan kesamaan dan melaksanakan pengiktirafan objek. Teknologi ini dapat mengenalpasti wajah, tanda jalan, tumor, manusia, dan sebagainya. Rangkaian ini terdiri daripada Convolution Layer, Activation Function (ReLU), Pooling Layer, Fully Connected Layer dan Output Layer. Tiga set data seperti Knots, Crack and Normal disediakan untuk melatih dan menguji model Teknik Rangkaian Konvolusi Neural (CNN) dengan menggunakan parameter yang berbeza. Keputusan konfigurasi yang berbeza telah disbandingkan and dianalisis. Ketepatan pengkelasn untuk semua variati adalah 97.2%.

ACKNOWLEDGEMENT

First of all, I would like to thank Universiti Malaya (UM) for providing me such a comfortable study environment and facilities throughout my Master program at Faculty of Engineering. I sincerely appreciate the opportunity given by University Malaya.

Furthermore, I would like to express my utmost gratitude to my supervisor Ir. Dr. Chuah Joon Huang of VIP Research Laboratory, Department of Electrical Engineering for continuous support of my Master study and research project. He has helped immensely me throughout the course of the research project and writing of this research report. A sincere acknowledgement is dedicated to Ms. Lim Kim Lee from Classic Scenic for her inspiration and encouragement for this research project. Immense appreciations also go to the Saw and Mill Department at Forest Research Institute Malaysia (FRIM) and wood image database from University of Oulu for providing images for this project.

Special thanks to my family members for their support. Lastly, I would like to thank all my coursemates and lecturers who provided technical support and meaningful lectures. I perceive this achievement as a successful milestone during my life

Table of Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

AI: Artificial Intelligence

AIP: Analogue Image Processing

ANN: Artificial Neural Network

BP: Back Propagation

CNN: Convolutional Neural Network

DTWCT: Dual-Tress Complex Wavelet Transform

DIP: Digital Image Processing

GA: Genetic Algorithm

GLCM: Gray Level Co-Occurance Matrix

LBP: Local Binary Pattern

MLP: Multi-Layer Perceptron

PSO-GA: Particular Swarm- Genetic Hybrid Algorithm

RGB: Red, Green Blue

ReLU: Rectified Linear Unit

RNN: Recurrent Neural Network

NN: Neural Network

SURF: Speeded-Up Robust Features

SVM: Support Vector Machine

LIST OF APPENDICES

Chapter 1: Introduction

1.1 Introduction

This chapter involves a project introduction such as research background, challenges of this research, problem statement, objectives and scope of this research.

1.2 Research Background

In the timber and wood industry, the automation in sawmills production-lines has been available in many decades. However, in some areas, it is difficult to have automation, especially by utilizing a robot with the ability to visually inspect the product and rate it with right pace of work. In timber and wood industry, the quality assurance of board is still controlled by a human because the defects can vary in many ways such as area, colour, shape and amount. It is important for timber and wood industry to deliver right quality of boards to their clients. The manual checking process can be very tedious and it is prone to mistakes by human judgement. The small amount of human's mistake can lead to losses the income for sawmills.

The purpose of artificial intelligence (AI) is to make an intelligent machine that can reach the human level of intelligence, so that the machine can understand the problem and solve it. The main aim of this research is to automatically detect and classify wood defects from images captured. With the method of deep learning, it can help to resolve the automatic wood defect detection and classification problem.

1.3 Challenges in Wood Defect Classification

In the wood industry, there are various type of defects arising from many different causes. For instance, there are mainly three groups of defect such as natural defect, defect after conversion and defect after seasoning. Conversion is the process of converting raw timber to form suitable for woodworking or construction projects (Ron Smith 2016). Seasoning is the process of drying lumber to an appropriate level of moisture for woodworking (Ron Smith 2016).

Natural defect:	Defect after conversion	Defect after Seasoning
- Bark pockets - Burls - Fungal damage (Blue Stain) - Knot - Shake (crack)	- Chip mark - Diagonal grain - Torn grain - Wane - Machine burn - Machine gouge	- Bowing - Check - Crook - Twisting - Cupping - Spring - Case hardening

Table 1. 1: Types of wood defects

The main challenges in wood defect is the image database for the natural defect of wood. The images are difficult to look for due to in timber and wood industry, those timber with natural defect will be excluded or rejected from the manufacturer. This is because the timber with defect will affect the strength and performance of the timber.

Other than that, the hidden layers in Convolution Neural Network (CNN) also one of the main challenges. In CNN, convolution layers and pooling layers are considered as hidden layers and mainly used to train the model, these layers could be more than one by arrange them repeatedly. By increasing the hidden layers in CNN could increase the accuracy of the final results yet the more hidden layers might reduce the accuracy depends on the situation.

1.4 Objectives

- To detect and classify wood defects using Convolutional Neural Network.

- To achieve high accuracy detection and classification of wood detect using Convolutional Neural Network.

- To evaluate the performance of detection and compare the outcomes of the classification using different parameters.

1.5 Scope of research

This research project aims to design and train a wood defect detection and classification model by using a CNN Model. This system is developed using MATLAB. A set of images are used in training and testing of this system.

1.6 Problem Statement

The main aim of this research project is to implement a deep learning method for a visual inspection of wood & timber. This method is used to measure the natural defect and visual inspection of each board and this is where this research project will be implemented.

Based on a traditional way, a worker distinguishes the wood defects by looking at the board, i.e. visual inspection. However, recognitions are based on experience that the worker gained from training and experience. The worker has to distinguish the board features such as, colour, size, weight, shape, pattern and more. The manual task can be very tedious and it is prone to human errors.

In the 21st century, object classification is very common in various industries. Object recognition is used to identify object in images or videos by using vision technology. It is also an important key output of deep learning and machine leaning algorithm. The main purpose is to make an intelligent machine that can reach the human level of intelligence, so that machines can accurately differentiate and understand the objects in the images. There are three types of board used in this research which are boards with knot, boards with shake (crack) and boards with good quality.

Chapter 2 Literature Review

2.1 Introduction

This chapter will cover the introduction of artificial intelligence (AI), basic AI technologies used in industry, image pre-processing techniques, artificial neural network (ANN) and discuss the technologies used in the research field and the significance of it. Then following by the training process of ANN, algorithm of ANN, comparison of the algorithm of ANN and the brief explanation of natural wood defect.

2.2 Artificial Intelligence (AI)

Artificial Intelligence (AI) is a branch of computer science concerning smart machines that are capable of performing tasks like human being. AI has played an important key role in a Fourth Industrial Revolution. AI can be defined as a computer system that allows to perform tasks that such as learning, recognise, self-correction, etc. These AI systems can be powered by a few groups such as machine learning, deep learning, some of them are powered by some specific rules, etc. In short, AI can be defined as a system which thinks like a human, processes like a human, works like a human, performs like a human and analyses like a human.

Basically, the architecture of Artificial Intelligence (AI) can be separated into two groups such as AI Paradigms and AI Problem Domains (Francesco Corea 2018). The AI Paradigms are used to solve specific problems. For AI Problem Domains are defined as the types of problem that can be solve by using AI. The AI Paradigms and AI Problem Domains are shown in the Table 2.1 and 2.2.

AI Paradigms	Definition
Logic -based tools	Mainly used for problem solving
Knowledge - based tools	Mainly used in huge databases and ontologies
Probabilistic methods	Tools that used for analysis and prediction
Machine learning	Tools that used for data analysis
Embodied intelligence	Tools that use for designing an intelligence robotics system
Search and optimization	Tools that used for intelligent search and optimization

Table 2. 1: Types of AI Paradigms and Definition

AI Problem Domains	Definition
Causes	The capability to solve problem
Knowledge	The capability to understand the knowledges
Plan	The capability to perform decision making
Communication	The capability to understand and learn from the human communications and languages.
Perception	The capability to perform sensation through various type of sensors.

Table 2. 2: AI Problem Domains & Definition

2.3Artificial Intelligence Technologies

Various types of AI technologies are developed and implemented in industry and research. These technologies are developed by using the architecture of AI, some of the technologies are using standalone architecture method and some of the technologies are combining the architecture together to achieve a better results. Table 2.3 shows a list of technologies in industry.

AI Technologies	Definition
Robotic Process Automation (RPA)	A technology that allows the high volume and repetitive tasks that required humans replaced by robotic automation system through digital system control.
Expert System	A technology that allow the computer system to do decision making and solve complex problem.
Computer Vision (CV)	A technology that allow the system to do image analysis by using industry camera.
Natural Language Processing (NLP)	A technology that allow the system to do analysis and synthesis on human language and speech.
Artificial Neural Networks (ANN)	A technology are developed based on the biological neural network. Thus this allow the system to learns and make decision like human.
Autonomous System	A technology used to control a groups of the networks under a single administrative control.
Distributed Artificial Intelligence (DAI)	A technology used to solve complex learning method, large-scale planning and decision making on large datasets.

Affective Computing	A technology used to recognize, interpret, process and simulate emotional of human by using external devices like sensors, microphone and others as input data.
Evolutionary Algorithms (EA)	It is a subset of a wide ranging computer science domain called *evolutionary computation* that uses mechanisms inspired by biology to look for optimal solutions (Franceso Corea 2018)
Ambient Intelligence	A technology that allow the system to interact, respond and perform specific tasks that commanded by human.
Machine Learning	A technology that allow the system to perform learning and self-improvement to the outcome without being explicitly programmed.
Deep Learning	A technology that teaches computers what to do from the examples that trained. It is considered part of a machine learning.
Pattern Recognition	A technology that allow the system to study and learn the features based on the given circumstances or images.

Table 2. 3: AI Technologies & Definition

2.4 Image Pre-Processing

In previous years, techniques on image processing are developed rapidly in the area of computer vision and machine learning. They are used to perform some operations on the images, either to enhance the images or to extract features from the images. Image processing is applicable to the medical imaging, aeronautics, military, geographic system and more. Image processing consists of Analogue Image Processing (AIP) and Digital Image Processing (DIP).

Analogue Image Processing (AIP) is usually applied on analog signal and it only processes two-dimensional signal. For instance, the images operates by electric signal like television images, photographs, painting, etc. Digital Image Processing (DIP) is used on digital signal that processing in two – dimensional signal, these digital signal can be defined as pixels. Various computer algorithms or software are used to perform DIP. DIP has more advantages than AIP as it allows a wide range of algorithms to be applied to the input data. Image pre-processing is an important tasks which allows the input images being normalise and filter unwanted distortion in order to enhance the important features from the input images.

The main purpose of image pre-processing is to enhance and strengthen the images features by using image recognition, filtering unwanted distortions, enhancing important features, segmentation, normalised, etc. First of all, Image recognition involves the use of algorithm and machine learning concept to recognise certain objects, peoples, building, animal and other variables in the images,

Image recognition is generally used to perform an enormous number of visual tasks such as guiding independent robots, labelling and content search in the pictures, accident prevention and self-drive system.

Secondly, several elements affecting images quality, for instance, camera quality, resolution, the effect on environment, visual angle, etc. Thus, image enhancement is used to enhance or filtering unwanted noise or distortion on the images. The techniques of image enhancement are divided into two groups which are the spatial domain and frequency domain.

In the field of spatial domain, there are 4 kinds of processing techniques mainly including point processing, median filtering, image subtraction and histogram equalisation, to improve the pixel value of the images.

Point processing is the process of contrast enhancement, it produced a higher contrast image than the original image by scaling the grey level of each pixel. Median filtering is a useful smoothing technique used to filter noise. It considers each pixel in the image and calculated the median pixel value from the surrounding and replaced them with the median value that calculated. Image subtraction is used to perform pixel subtraction from two images and produce an output image by using the value from subtraction. The output images are used to differentiate the differences between the two input images. Histogram equalization is used to adjust the contrast of the images.

In the frequency domain, the images are transferred into frequency domain, which enhance the images by using Fourier transform and inverse Fourier transform techniques in order to obtain resultant images. There are two types of processing technology in this frequency domain, image smoothing and image sharpening. The image smooth is used to reduce the camera noise, missing pixel, spurious pixel value, and others by using low pass or high pass filter. Image sharpening is used to highlight the features or details of the images and enhance the details that have been blurred in the images.

Third, Image segmentation is used to processing and analysis by dividing the images into multiple parts or segments containing each pixel with similar attributes.

Image segmentation techniques consist of separating foreground from the background (thresholding) and clustering regions of pixels based on similarities in colour or shape (clustering).

In research fields and industry, there are several image pre-processing techniques are introduced and utilised. Table 2.4 shows the several pre-processing techniques that been used in the industries.

Image Pre-processing Techniques	**Definition**
Image Restoration	To restore the images information that lost during the blurred process.
Independent Component Analysis (ICA)	To separates a multivariate signal computationally into additive subcomponents
Anisotropic Diffusion	To reduce the noise without reducing the important features of the images
Principal Component Analysis (PCA)	To summarise the features of the images without losing the significant part of the original images

Table 2. 4: Image Pre-Processing & Definition

2.5 Review of Wood Defect Detection and Classification

In previous years, a lot of researchers have involved in the wood defect classification topic. A lot of image processing techniques, classification techniques and algorithm are being introduced. Neural Network (NN) is the most famous classifier among others. Hongbo Mu, Li Li, Lei Yu, Mingming Zhang and Dawei Qi (2006) used X-ray as an alternative way to detect the wood defect non-destructively and Back-Propagation (BP) neural network to perform classification.

Zhen-Nan Ke, Qi-Jie Zhao, Chun Hui Huang, Pu Ai and Jin Gang Yi (2016) introduced a particular swarm-genetic hybrid algorithm (PSO-GA) to perform wood defect detection. A comparison was made between traditional genetic algorithm (GA) and PSO-GA which show that the PSO-GA have higher accuracy and better detection results.

Si yuan Wu, Zhao Zhang and Liang Feng (2009) used few types of image processing method such as non-negative matrix factorization (NMF) to breakdown the defects images, Local binary pattern (LBP) to extract the original spatial local structure features and Dual-tree complex wavelet transform (DTWCT) to extract the energy based on statistical features. At the end of the research, the Support Vector Machine (SVM) classification method are combining with the LBP and DTWCT to generate average accuracy with 90% of final results.

Yi Xiang Zhang, Ya Qin Zhao, Ying Liu, Lin Quan Jiang and Zhen Wei Chen (2016) used Local Binary Pattern (LBP) algorithm to differentiate the local texture feature of the images and Back Propagation (BP) neural network to perform classification. A comparison table was made between the image processing algorithm which is Gray Level Co-occurrence Matrix (GLCM) algorithm and Local Binary Pattern (LBP), these algorithms was combined with BP neural network for classification purpose. Based on the comparison table, it is proved that LBP is much effective than GLCM. By using the LBP algorithm with BP neural network, the accuracy of classification is 93.3% and the processing time is shorter.

Mohamad Mazen Hittawe, Satya M. Muddamsetty, Desire Sidibe and Fabrice Meriaudeau (2015) introduced a combination of multiple features extraction for image processing such as Speeded-Up Robust Features (SURF) and Local Binary Pattern (LBP). Support Vector Machine (SVM) is used to classify the knots and cracks, the final results show an average accuracy of 93%.

R.Qayyum, K. Kamal, T. Zafar and S. Mathavan (2016) used Gray Level Co-occurrence Matrix (GLCM) algorithm for image processing and Particle Swarm Optimization (PSO) trained feedforward neural network as a classifier. The accuracy of the classification is 78.26%.

2.6 Artificial Neural Network (ANN)

A neural network is a part of the categories in artificial intelligence (AI). The AI Expert Maureen Caudill defined neural network as "a computing system made up of several simple, highly interconnected processing element, which process information by their dynamic state response to external inputs" (Tharani 2018). A common application of ANN generally focuses on solving complicated problems like signal processing and pattern recognition. ANN can be defined as a computer that works like the human brain..

A basic neural network is defined as a computing system that usually consists of a large number of processors (**neurons**) and highly interconnected in parallel or series arrangement. Each of the processors (**neurons**) are connected to the layers on either side. The first layer (**Input Layer**) can be defined as input neuron, these units are used to receive the various form of information from the outside world. The last layer (**Output Layer**) can be defined as output neuron, these units are responded to the information learned and show the results of learning and training. Besides, there are one or more hidden units (**Hidden Layer**) in between the input and output layer. These hidden neurons can be defined as the brain of the neural network, it is where all the processing, learning and training happened through a neural network system. The connections between one neuron with another are represented by a number called **weights or biases**. For instance, the input layer received the input data, the neuron will calculate a weighted sum added with bias, after that based on the results and a pre-set activation function (**ReLU**), neural network system will decide whether it should be ignored or activated. In the end, the last hidden layer is linked to the output layer which used to show the final results.

2.7 Training Process

Many different kinds of algorithms were used in the training ANN system, with their advantages and disadvantage. There is mainly three types of training and learning method which are reinforcement learning supervised learning and unsupervised learning.

2.7.1 Supervised learning

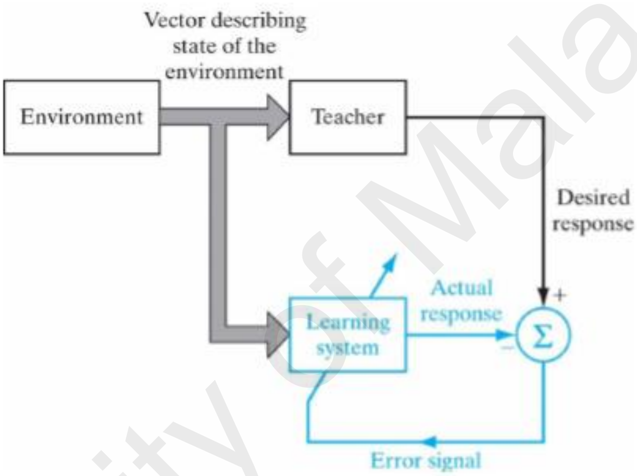


Figure 2. 1: Flow Diagram of supervised learning. (Haykin, Simon, 2009)

Based on supervised learning, both inputs and outputs are assigned in the system. The neural network will start processing the inputs and comparing the results output against the desired outputs. Besides, supervised learning can perform back propagation which allows the error signal feedback to the system and due to this, the system is allowed to perform weights adjustment. This process is repeated over and over again until the system has been trained correctly.

2.7.2 Unsupervised learning

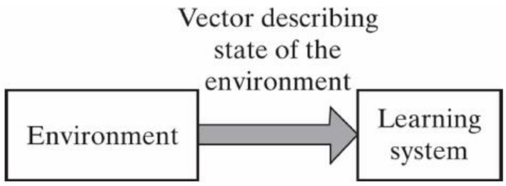


Figure 2. 2: Flow diagram of unsupervised learning. (Haykin, Simon, 2009)

In unsupervised learning, the inputs are provided in the system but not with the desired output. This system will decide what features and characteristics will be use to differentiate the input data. In general, it defined as an adaption or self-organization system.

2.7.3 Reinforcement learning

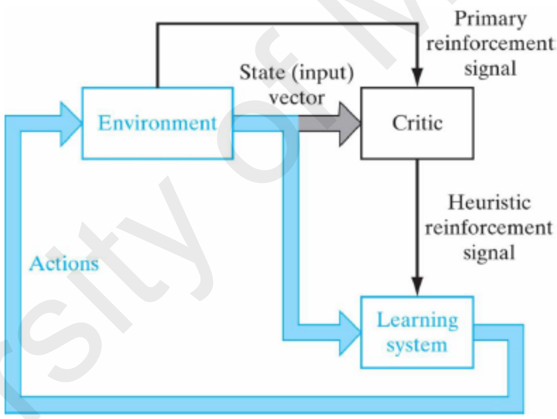


Figure 2. 3: Flow diagram of reinforcement learning (Haykin, Simon, 2009)

Reinforcement learning is similar to supervised learning. It is a type of learning algorithm in which the system will decide on what action to take. Reinforcement learning refers to goal-oriented algorithms which learn how to achieve a complex objective over many steps.

2.8 Types of ANN (Algorithm)

In the neural network, training is the process used to minimise the prediction error and increase the accuracy. There is mainly two types of ANN which is feedforward and backpropagation. It is possible to make the neural system more flexible and powerful by adding hidden layers.

2.8.1 Multilayer Perceptron (MLP)

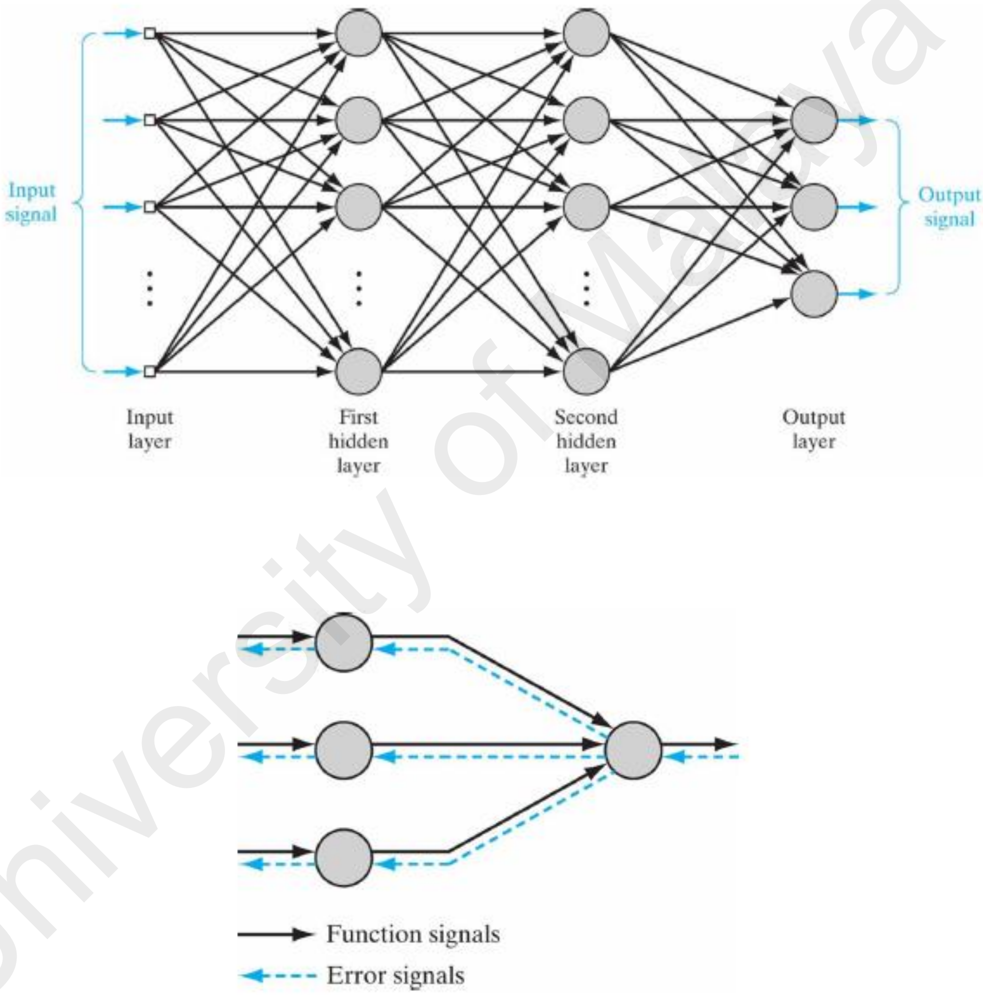


Figure 2. 4: Architecture of MLP & Direction of Flow (Haykin, Simon, 2009).

A typical architecture graph of an MLP and the directions of two basic signal (Feed Forward function signal and back propagation (BP) error signal) are shown in Figure 2.4. An MLP consist of three or more layers. It utilized a non-linear activation function that allow it to classify the data that are not linearly separable.

2.8.1.1 Function Signal (Feed forward)

A function signal is considered as feed-forward input signal that received from the input data, move forward through the layer and neurons, and generate the output signal at the end of the layer.

2.8.1.2 Error Signal (Back propagation)

An error signal is considered as back propagation signal that will move backwards (layer by layer) from the output through the network. Error signal occurred due to its computation by every neuron of the network involves an error-dependent function in one form or another.

2.8.2 Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) is one of the deep neural networks that used in two-dimensional data. It mainly used to classify objects in images, cluster them by similarity and execute object recognition. This technology can identify faces, street sign, tumours, human, etc.

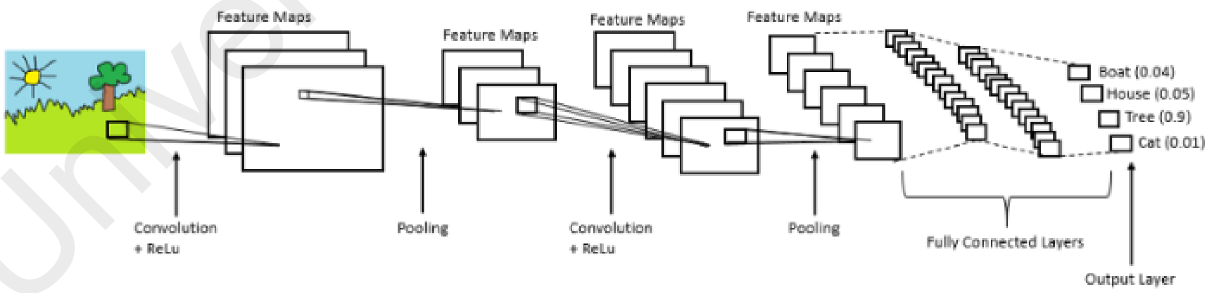


Figure 2. 5: Structure of CNN model (Prabhu, 2018)

Based on the Figure 2.5, a typical CNN architecture consists of input images, convolution layers, activation function (ReLU), pooling, fully connected layers and output layer.

Input layer – A raw image is used as input data.

Convolution layers + ReLU – To extract the specification of input images and retain the relationship between pixel by learning image features using certain parts of input data. The activation function (ReLU) is used to perform non-linearity to a system that just been computing linear operation during the convolution layers.

Pooling layer –To make sure that the subsequent layer of convolution layer can pick up the larger scale of details rather than just edges and curves

Fully connected layers – To classify the input images into assigned classes based on the input data.

2.8.3 Recurrent Neural Network

Recurrent Neural Network (RNN) is part of the artificial neural network that normally used in speech recognition and natural language processing. RNN is developed to recognise data's sequential characteristics and use patterns to predict the next likely scenario (Margaret Rouse 2018). It also a type of neural network where the output from the previous neurons are fed as input to the current neurons. For example, RNN remembers the past results and its final decision are influenced by what it has trained from the past.

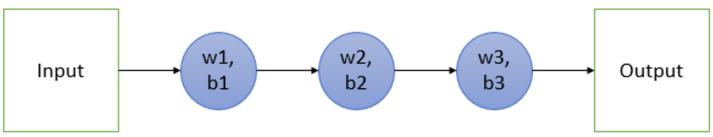


Figure 2. 6: Structure of typical ANN model

A block diagram of a typical neural network is shown in Figure 2.6. An input layer, three hidden layers and an output layer are involved. Each of the hidden layers will have its own set of weight, biases and independent to each other.

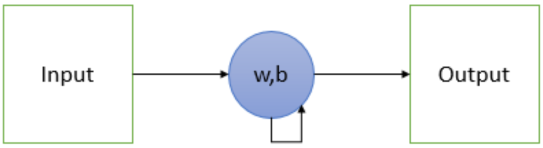


Figure 2. 7: Structure of RNN system

By using RNN system, it will convert the independent activation to dependent activation by providing same weights and biases to all the hidden layer, this will reduce the complexity of the parameters and memorizing the previous output as the input to next hidden layer. Thus, the three hidden layers can be joined together and sharing the same weights and biases.

2.8.4 Comparison of ANN algorithm

Based on the explanation of MLP, CNN and RNN above, a comparison can be made. In industries, it is possible to use both MLP and CNN for image classification but CNN is preferable. This is because MLP takes a vector as input and CNN take tensor as input thus CNN can understand spatial relation between pixels of images (Dhiraj Kumar 2018). RNN is used to do recognition of data characteristics and use patterns or algorithm to estimate the next scenario.

In short, MLP is mainly used for simple image classification, tabular dataset and regressive prediction problems. CNN is suitable for complicated classification of images and RNN is used for sequence processing.

2.9 Natural Defect

In the wood industry, it is important to have strong and good quality wood. Wood is sold in various form, size, type and cuts. There are some characteristics are used to check and describe the wood board such as density, texture, colour, defects and wood grain. If the wood contains a lot of knots and crack, then it can be unstable. Some of the parts will become weak and break when overweight, thus it is important to recognise the defect.

2.9.1 Blue stain

Blue stain is caused by microscopic fungi that caused the colour of board change into blue-black and grey. It is a common cause for the discolouration of wood and the blue stain does not affect the performance, the strength of the wood and do not cause decay. Due to this, it can be used for the exact same purpose as non-stained wood in construction or wood industry.

Figure 2. 8: Sample of Blue Stain

2.9.2 Wood decay (Rot)

The main factor of causing Rot in the wood board is fungi. The fungus will spread through the whole wood board and the part that contains fungus is weakened and it is eliminated to use in construction and wood industry. There is mainly three types of rot which are brown rot, white rot and soft rot.

2.9.3 Shake

Shake is one of the natural defects in the wood industry. It occurred around the growth ring of wood. In common, cracks or split in the wood are considered shakes as well. It can be used in construction or other wood industry depending upon depth and uses. The main problem of shake is aesthetic whereby the appearance and the surface are important.

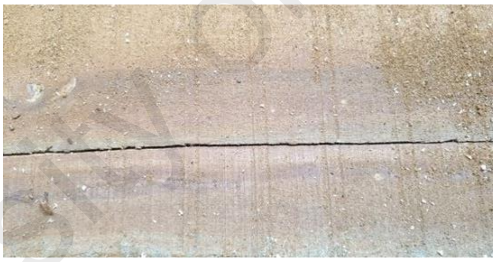


Figure 2. 9: Sample of Shake

2.9.4 Knot

Knots are the most common defect in the wood industry. The knot is caused by tree branches tend to die off and their bases become overgrown and enclosed by the subsequent layer of wood. Thus, knots in wood can be defined as the part where tree branches once were. It very difficult to remove knot on the wood, knot might compromise the technical properties like strength, the performance of the wood depending upon uses and depth. The other problem of knot is aesthetic whereby the appearance and the surface are important.

Figure 2. 10: Sample of Knot. (Adapted from website of University OULU)

2.10 Summary

There are many types of approaches that can be used in detection and classification of wood defects based on literature reviews. Through the reviews, most researchers preferred to use Neural Network as classifier because NN can learn the extraction features, support a large amount of dataset as input, duration of processing is shorter than the traditional machine learning algorithm. In this study, a CNN model for wood defect detection and classification is implemented by using a different amount of layers and parameter to achieve and increase the performance of this system.

Chapter 3 Methodology

3.1 Introduction

The objective of this project is to design a system which able to classify wood defect in images. In this project, the convolutional neural network (CNN) model is used to do the identification and classification of the wood defect. As stated in chapter two, Convolutional Neural Network (CNN) is one of the deep neural networks used in two-dimensional data. It mainly used to classify objects in images, cluster them by similarity and execute object recognition.

3.2 Architecture of CNN

A CNN algorithm is allowed to take in an input raw images, by assigning learnable weights and biases to the object in the images, CNN can differentiate or classify one from the other. In other words, CNN can be trained to understand the sophistication of the image. By comparing the duration of the pre-processing requirement in CNN and other neural network models, CNN is much lower.

3.2.1 Input image

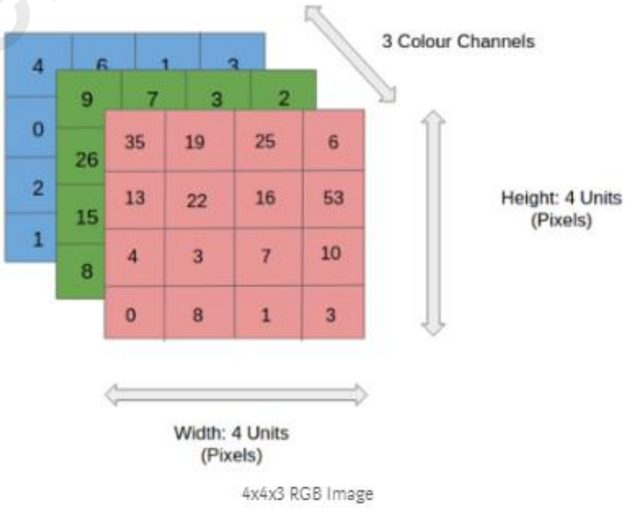


Figure 3. 1: Sample of Input Images in Matrix form (Saha, Sumit 2018)

Based on the Figure 3.1, the RGB images are shown and separated into three colour panels such as Red, Green, and Blue. Other than RGB, grayscale, HSV, CMYK and others are included in the image as well. These images are represented in the form of pixels.

3.2.2 Convolution Layer

Convolution layer is used to reduce the size of the images into a proper form which is easier to do processing without losing any important feature which is critical for getting a good prediction. The element that carries out the convolution operation in the first part of the convolutional layer is called **Kernel** or **Filter**. The Kernel is shown in Figure 3.2 in a matrix form of 3x3x1.

-1	-2	-1
0	0	0
1	2	1

Figure 3. 2: Sample of Kernel (Robinson, Rob 2017)

A kernel is placed in the image and started at the top-left corner. The pixel of the input image is multiplied with the kernel and the product are summated. After that, the result of the summation is placed in the new image at the point whereby corresponding to the centre of the kernel. This kernel process is moved over by one pixel and this process is repeated until the entire image is traversed. The Kernel is moved over by one pixel due to the stride length is one, when performing matrix multiplication.

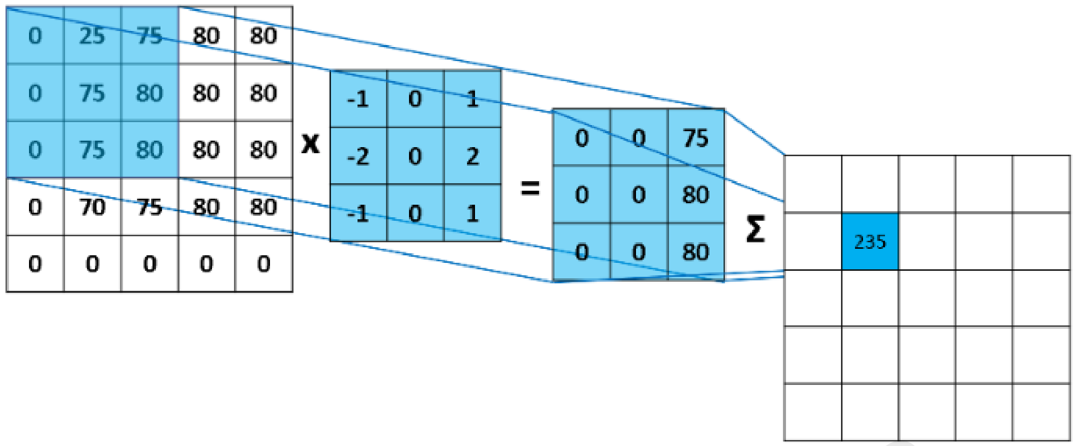


Figure 3. 3: Process of convolution layer and kernel(Robinson, Rob 2017)

The purpose of this process is to extract the high-level features of the input images. There could be many convolution layers in a system. For example, the first convolution layer is used to capture low-level features such as colour, edges orientation, etc. Following added layer by layer, the model will recognise the high-level features and provide a network that understanding the images in the dataset. Furthermore, activation functions are added on each convolution layer. The activation function is used to perform non-linearity to a system that just been computing linear operation during the convolution layers.

There are few types of activation function used in industry such as **linear, step, sigmoid, tanh** and **rectified linear unit (ReLU)**. In machine learning industry, **ReLu** is more preferable due to it is simple and consistent technique and it used to avoid and rectifies vanishing gradient problem.

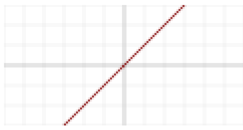
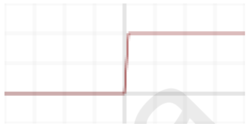
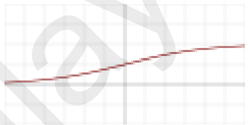
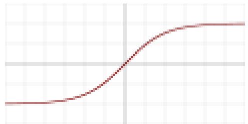
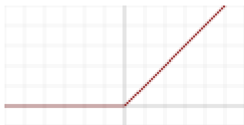
Activation Function	Formula	Graph
Linear	$f(x) = ax$	
Step	$f(x) = \begin{cases} 0 & for\ x < 0 \\ 1 & for\ x \geq 0 \end{cases}$	
Sigmoid	$f(x) = \dfrac{1}{1 + e^{-x}}$	
Tanh	$f(x) = \tanh(x) = \dfrac{2}{1 + e^{-2x}} - 1$	
ReLU	$f(x) = \begin{cases} 0 & for\ x < 0 \\ x & for\ x \geq 0 \end{cases}$	

Table 3. 1: Types of Activation functions (Adapted from Davydova, Olga 2017)

3.2.3 Padding

During the convolution process, there are some empty values (0) which surround the convolved images. This is due to the final result of the convolution is placed at the centre of the kernel. To solve this problem, a padding process is applied in the system. After adding a process of padding, the results of convolution will produce an output image that has the same size to the input image.

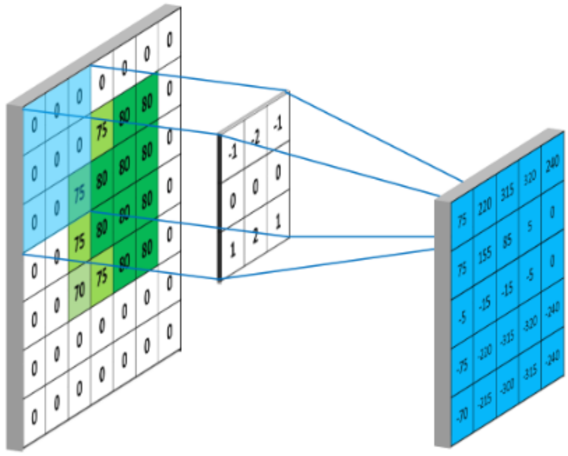


Figure 3. 4: Sample of Padding process (Robinson, Rob 2017).

3.2.4 Pooling layer

A pooling layer is added after the convolution layers and the padding layers. Pooling layer is used to make sure that the subsequent layer of convolution layer can pick up the larger scale of details rather than just edges and curves. Besides it is also used to reduce the spatial size of convolved features. This is to minimise the computational power that required to process the data by reducing the dimensionality. Max pooling and average pooling are involved in this system. Max pooling is used to return the maximum value from the input image that covered by Kernel. Average pooling is used to return the average value from the input image covered by the Kernel.

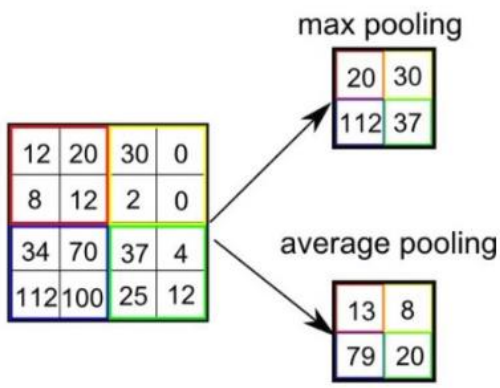


Figure 3. 5: Sample of Max Pooling and Average Pooling (Saha, Sumit 2018)

A Convolutional Neural Network (CNN) is formed through the process above and able to understand and learn the features of the input images. The number of layers will increase by depending on the complexities of the input images.

3.2.5 Classification – Fully Connected Layer

In classification, a fully connected layer is usually been used. It used to learn the non-linear output from CNN. The neurons in the fully connected layers will detect a certain feature, after that it will preserve its value and communicate the value with the classes that assign by the programmer. All the classes will check the features and decided where it belongs to. Basically, the fully connected layers will look at the high-level features that most correlate to a particular class and particular weights, thus the correct probabilities of the final result of classification is more accurate. A softmax function is added at the end of the fully connected layer due to this system has multiple classification logistic regression model. Softmax function help to calculate the probabilities distribution, the calculated probabilities will be in the range of 0 to 1.

3.3 Modelling and Programming

A large number of defect samples were collected in this system for training and testing. In this study, MATLAB is used to perform the process of training, testing and classification. It is a fourth generation programming language, which includes many engineering function such as image and signal processing, neural network, communication, control system, design, robotics, etc. C++ language is the main programming language that is used in MATLAB.

Three types of board are used in this study which are boards with the knot, boards with the shake (crack) and boards in good quality. All pictures are captured from the saw and mill department at FRIM and website of the University of OULU. Different wood defect is stored separately in 3 folders, the system can detect the defect based on the folders. Besides, training samples and test samples are stored in different folders.

A total of 2500 are used as training samples and 700 for test samples. Each of the categories consists of 830 to train and 230 to test the accuracy. The size of the samples are in different pixels, all the samples need to pre-process before training.

3.4 Image Pre- Processing

The image dataset has a different pixel, thus images are resizing to a 128x128 pixel for training purpose. The size of images was resized to 128x128 pixels, thus the data matrix is reduced which also reduces the processing time. Besides, data augmentation also included in this stage for better classification. Data augmentation will transform the image dataset by flipping, rotating, zooming and cropping. The purpose of data augmentation is to decrease the overfitting problem between training dataset and test dataset. The data augmentation only used in training dataset, training dataset will undergo a data augmentation process to create more data by using augmentation techniques.

3.5 CNN Block Diagram

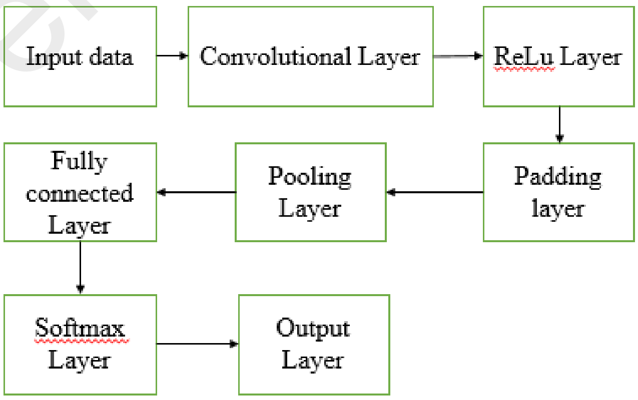


Figure 3. 6: Block diagram of CNN of this system

Block diagram of this system is shown in Figure 3.6. In this study, an input layer, convolution layer, padding layer, ReLU layers, pooling layer, softmax layer and fully-connected layer are used for the classification purpose.

This study uses 3 layers of convolution, 4 layers of ReLU, 1 layer of max pooling, 2 layers of average pooling and 2 fully-connected layers. The MATLAB function of CNN is shown in Table3.2.

Layers	Function
Convolution Layer	convolution2dLayer (filerSize, numFilters)
Max Pooling Layer	maxPooling2dLayer (poolSize, Name, Value)
Average Pooling Layer	averagePooling2dLayer (poolSize, Name, Value)
ReLu	reluLayer()
Fully connected Layer	fullyConnectedLayer (outputSize, Name, Value)
Softmax layer	softmaxLayer()

Table 3. 2: MATLAB Functions of different layers

The convolution layers are formed to get the features, in this system a Kernel size of [5 5] is used. There are multiple feature maps for each convolution layer and each feature map is connected to the previous layer. The convolved results will go to the fully-connected layer after the convolution and pooling layers. The system can learn the features of each defect by repeatedly extracting the high-level features of the entire training samples.

This system also uses supervised learning. As discussed in Chapter 2, both inputs (training samples) and output (test sample) are provided in the system based on supervised learning. CNN will start processing the input samples and comparing the output results to the desired output. Besides, CNN involves back propagation in the training process due to back propagation allow the error signal feedback to the system thus the system can adjust the weights. This process is repeated over and over again by adjusting the correct parameter, until the system is properly trained.

3.6 Variables

A CNN model consists of two types of variable which are constant and calculate variables. The constant variables will affect the training and testing results. Constant variables consist of the number of input images, the pixel value of input images, number of batch size, learning rate, number of epoch and kernel size. The output layer is considered as the calculate variables due to it shows the final results of this system.

Input Images – 2500 of images are used as input data, 700 of images are used as testing samples.

Pixel value – 128 x 128 pixels

Batch size – 100 training samples are used utilise in one iteration.

Epoch – 5, 10 & 100 epochs are used in this model

Learning rate – 0.001 learning rate is used to controls control how much to change the model in response to the estimated error each time the model weights are updated.

Kernel Size – 5x5 matrix is used as a filter size in this system

A confusion matrix table is used to describe the performance of a classification model for the calculated variable. In the confusion matrix, the percentage and the number of prediction are shown.

3.7 Summary

A CNN model is suggested as the classifier in this study. It used 2300 images as input dataset and 830 images for testing purpose. The results are generated and presented in Chapter 4 by passing through the processes (convolution layer, pooling layer, padding, activation layer, fully connected layer and softmax layer) and parameter tuning based on the model performance.

Chapter 4 Results and Discussions

4.1 Introduction

This chapter reveals the results of the identification and classification of the wood defect. The purpose of this study is to get the accuracy of the final results as high as possible. There was a total of 2500 images of defects are used as training purpose and a total of 700 images of different defects are being used in this study for testing purpose. These images are all in different angles, shapes, light exposure and without overlapping. This system has learned and trained 3 different types of natural defect of wood and obtained the feature of each defect.

4.2 Results

To build a CNN model, different parameters need to be set including learning rate, the number of epoch, number of batch size, kernel size and more. The accuracy of the final results is depending on the parameter setting. The different number of epoch are tested to verify the accuracy of the results. The Table 4.1 shows the accuracy rate of the different number of epochs and the time taken for training. The Table 4.1 is clearly shown that the highest accuracy is 97.2% which 25 and 50 epochs are used.

Number of Epoch	Training time	Accuracy
10	*96s*	85%
15	*141s*	95.9%
25	*237s*	97.2%
50	*474s*	97.2%
100	*932s*	95.8%

Table 4. 1: Results of CNN based on different epoch

The training graph of epoch 50 and 100 are shown in Figure 4.1 and 4.2. Based on the graph, it shows that the gradient becames smooth when the epoch more than 50. From that, the epoch 50 is suitable to use as the parameter.

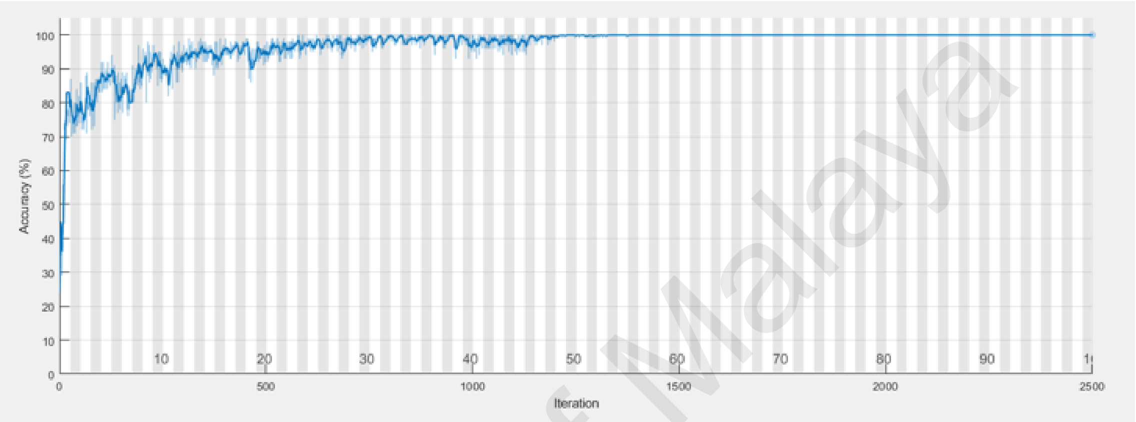


Figure 4. 1: Training graph of epoch 100

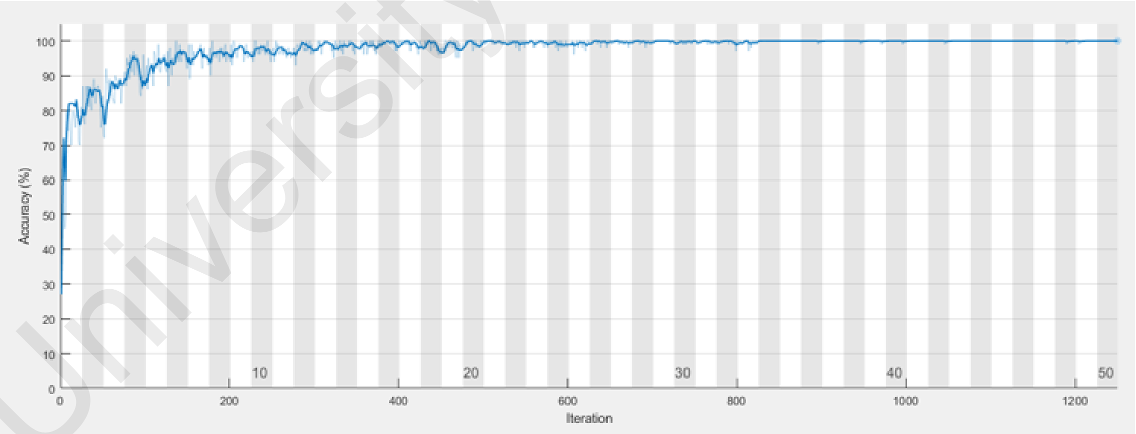


Figure 4. 2: Training graph of epoch 50

From the Figure 4.3 and 4.4, the loss graph of 25 and 50 epoch are shown. Based on the graph, at the beginning of the training process, the loss of both epochs are very high and after a certain time, the loss becomes low and close to zero. This is due to the CNN model started to understand and learned the features of the input dataset.

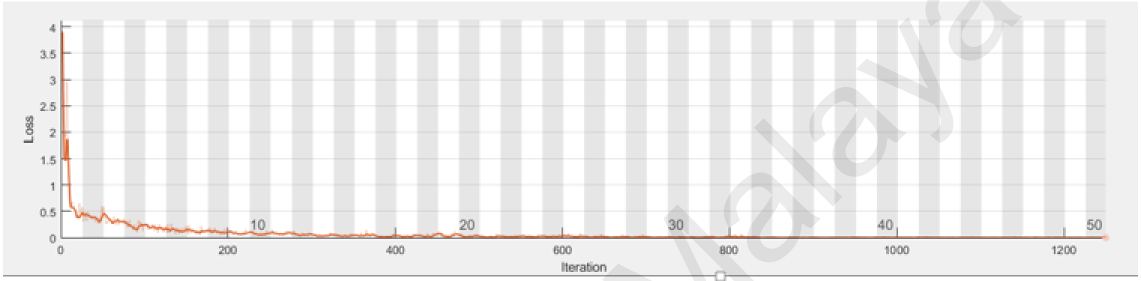


Figure 4. 3: Loss graph of epoch 50

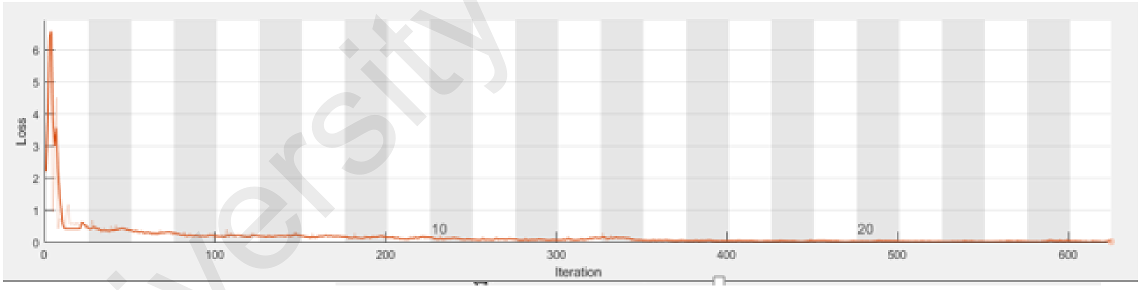


Figure 4. 4: Loss graph of epoch 25

The confusion matrix of 25 and 50 epochs are shown in Figure 4.5 and 4.6. The purpose of the confusion matrix is to describe the performance or results of the classification model on a set of test data.

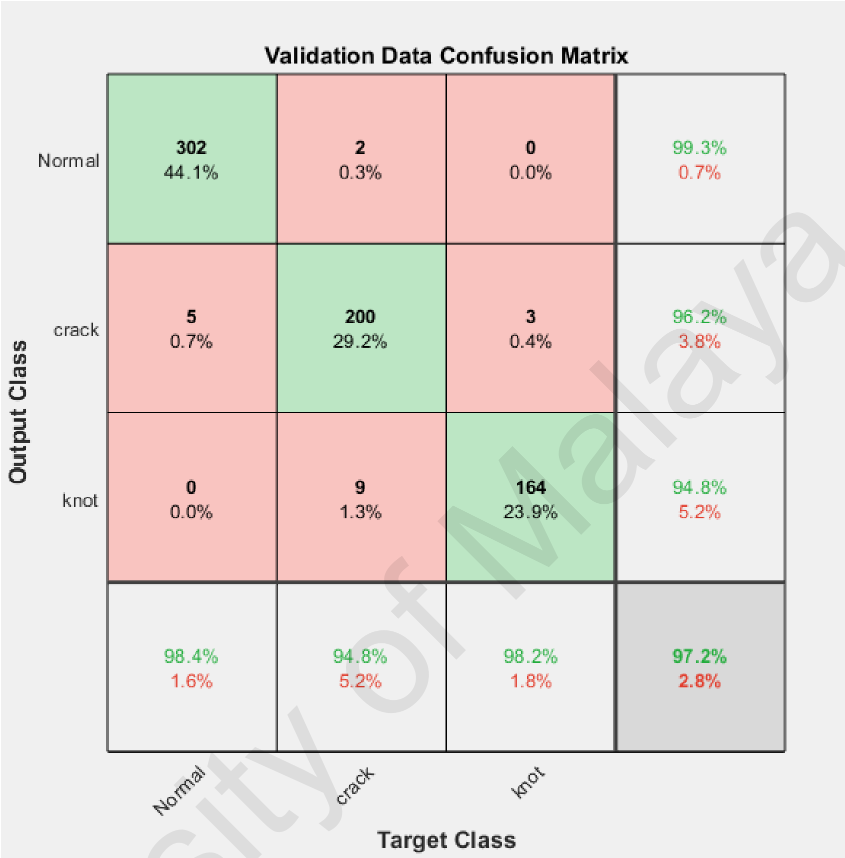


Figure 4. 5: Confusion matrix of testing output for epoch 50

From the Figure 4.5, there are 3 defects mainly evaluated in this CNN model which is normal, crack and knot. With an overall accuracy of 99.3% of the correct prediction and 0.7% of the wrong prediction, the output class can classify normal wood. For crack, 96.2% of the correct prediction and 3.8% is a false prediction, there are a total of eight images that predict into the normal and crack group wrongly. For knots, 94.8% of correct prediction and 5.2% of false prediction, a total of 9 images are incorrectly classified into the crack group. The overall accuracy is 97.2%.

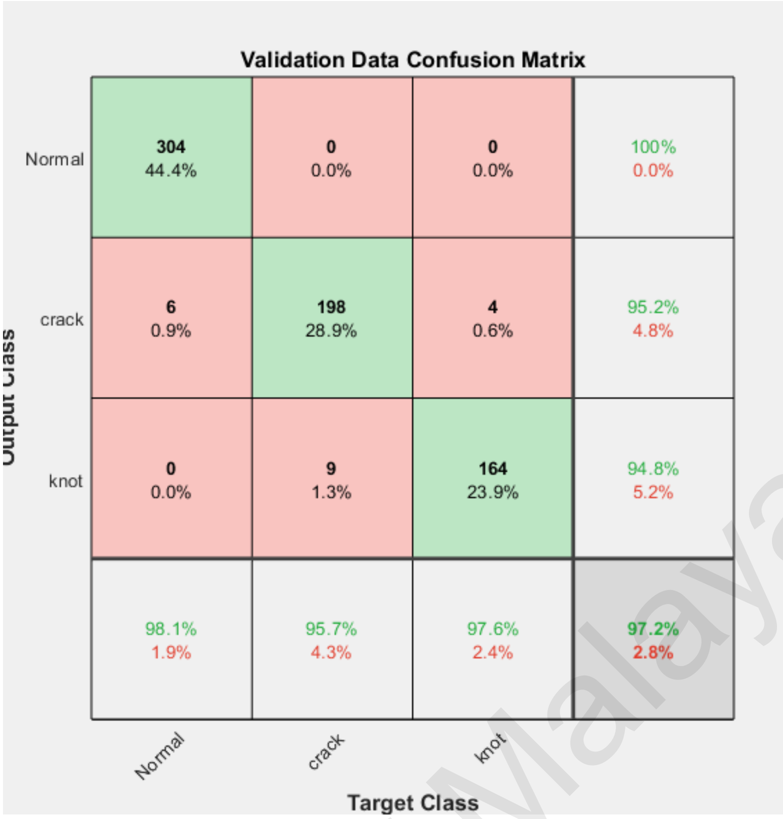


Figure 4. 6: Confusion matrix for testing output of epoch 25

From the Figure 4.6, the output class can classify the normal wood with 100% accuracy of correct prediction. For crack, 95.2% of correct prediction and 4.8% is a false prediction, there is a total of 10 images are predicted into the normal and crack group wrongly. For knots, 94.8% of correct prediction and 5.2% of false prediction, a total of 9 images are incorrectly classified into the group of crack.

4.3 Summary

According to results, CNN can very well learn and classify the features in wood defect by using appropriate parameters. However, only three defects are being assessed, thus the performance of this system is considered moderate due to there are many types of defect in wood and timber. Due to the processing time is lower, the model using 25 epochs is preferable

Chapter 5: Conclusion

5.1 Conclusion

In order to train a good and effective neural network system, the image dataset is playing an important role. The Small size of image dataset will cause a lot of training issues and the accuracy of prediction will be low. Besides image dataset, the algorithm or technique that is used for training and classifying is critical too.

In this project, a wood defect detection and classification system is designed using a CNN model. Three types of defect are considered for classification purpose. The results of the CNN model are used to do a comparison with the same parameter setting instead of the number of epochs. In this study, the 25 epochs with learning rate 0.001 given accuracy of 97.2% and the processing time is 237s, compare to the 50 epochs with learning rate 0.001 although the accuracy is same but the processing time is much longer. Thus, the result of CNN model with 25 epochs is chosen.

5.2 Recommendation and Future Work

Only three types of defects are used as input datasets based on this study. There are many different types of natural defects in the timber and wood industry, such as blue stain, rots and others. Future work focuses primarily on adding new defects, making this system more effective. Due to different settings, the parameters settings must also be considered instead of the images dataset.

Chapter 6: References

Auoriasoft (2017). *Image Recognition and Image Processing Techniques*. Retrieved from https://medium.com/@Adoriasoft/image-recognition-and-image-processing-techniques-fe3d35d58919

Browniee, Jason (2019). *Understanding the Impact of Learning Rate on Model Performance with Deep Learning Neural Networks*. Retrieved from https://machinelearningmastery.com/understand-the-dynamics-of-learning-rate-on-deep-learning-neural-networks/

Canuma, Prince (2018). *Image Pre-processing*. Retrieved from https://towardsdatascience.com/image-pre-processing-c1aec0be3edf

Corea, Francesco (2018). *AI Knowledge Map: How to Classify AI Technologies*. Retrieved from https://www.forbes.com/sites/cognitiveworld/2018/08/22/ai-knowledge-map-how-to-classify-ai-technologies/#232c60707773

Davydova, Olga (2017)*. 7 Types of Artifical Neural Networks for Natural Language Processing.* Retrieved from https://medium.com/@datamonsters/artificial-neural-networks-for-natural-language-processing-part-1-64ca9ebfa3b2

Goldstone, Kate (2015). *What is Wood Rot?* Retrieved from https://www.wood-finishes-direct.com/blog/what-is-wood-rot/

Gnanasegaram, Tharani (2018). *Aritifical Neural Network – A Brief Introduction.* Retrieved from https://medium.com/@tharanignanasegaram/artificial-neural-network-a-brief-introduction-572d462666f1

Hongbo Mu, Li Li, Lei Yu, Mingming Zhang and Dawei Qi (2006, 25-28 June 2006). *Detection and Classification of Wood Defects by ANN**. Paper presented at the 2006 IEEE International Conference on Mechatronics and Automation

Haykin, Simon (2009). *Neural Network and Learning Machine, Third Edition.* Retrieved from https://slideplayer.com/slide/7484497/

Inzaugarat, Euge (2018). *Understanding Neural Networks: What, How and Why?* Retrieved from https://towardsdatascience.com/understanding-neural-networks-what-how-and-why-18ec703ebd31

Kumar, Dhiraj (2018). *MLP vs CNN vs RNN Deep Learning, Machine Learning Model*. Retrieved from https://medium.com/@dhiraj8889_40289/mlp-vs-cnn-vs-rnn-deep-learning-machine-learning-model-52b945f76d9b

Mohamad Mazen Hittawe, Satya M. Muddamsetty, Desire Sidibe and Fabrice Meriaudeau (2015, 27-30 Sept 2015). *Multiple Features Extraction for Timber Defects Detection and Classification Using SVM*. Paper presented at 2015 IEEE International Conference on Image Processing (ICIP)

Perone, Christian.S (2017). *The Effective Receptive Field on CNNs.* Retrieved from http://blog.christianperone.com/2017/11/the-effective-receptive-field-on-cnns/

Prabhu (2018). *Understanding of Convolutional Neural Network (CNN)- Deep Learning.* Retrieved from https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148

Robinson, Rob (2017). *Convolutional Neural Network – Basics.* Retrieved from https://mlnotebook.github.io/post/CNN1/

Rouse, Margaret (2018). *Recurrent Neural Network*. Retrieved from https://searchenterpriseai.techtarget.com/definition/recurrent-neural-networks

R.Qayyum, K. Kamal, T. Zafar and S. Mathavan (2016, 7-8 Sept 2016). *Wood Defects Classification Using GLCM Based Features and PSO Trained Neural Network*. Paper presented at 2016 22nd International Conference on Automation and Computing (ICAC)

Saha, Sumit (2018). *A Comprehensive Guide to Convolutional Neural Network – The EL15 Way*. Retrieved from https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

Si yuan Wu, Zhao Zhang and Liang Feng (2009, 28-29 Nov 2009). *Statistical Feature Representations for Automatic Wood Defects Recognition Research and Applications*. Paper presented at 2009 Asia-Pacific Conference on Computational Intelligence and Industrial Applications (PACIIA)

Smith, Ron (2016). *Understanding & Working with wood defects*. Retrieved from https://www.woodworkingnetwork.com/best-practices-guide/solid-wood-machining/understanding-working-wood-defects

University of OULU. *Visual Inspection of Lumber*. Retrieved from http://www.ee.oulu.fi/~olli/Projects/Lumber.Grading.html

SuperDataScience (2018). *Convolutional Neural Network (CNN): Softmax & Cross-Entropy*. Retrieved from https://www.superdatascience.com/blogs/convolutional-neural-networks-cnn-softmax-crossentropy

Walia, Anish Singh (2017). *Activation Functions and its types-Which is better?* Retrieved from https://towardsdatascience.com/activation-functions-and-its-types-which-is-better-a9a5310cc8f

Woodford, Chris (2019). *Neural Network*. Retrieved from https://www.explainthatstuff.com/introduction-to-neural-networks.html

Yi Xiang Zhang, Ya Qin Zhao, Ying Liu, Lin Quan Jiang and Zhen Wei Chen (2016, 27-29 July 2016). *Identification of Wood Defects Based on LBP Feature.* Paper presented at 2016 35th Chinese Control Conference (CCC)

Zhen-Nan Ke, Qi-Jie Zhao, Chun Hui Huang, Pu Ai and Jin Gang Yi (2016, 11-12 July 2016). *Detection of Wood Surface Defects Based on Particle Swarm- Genetic Hybrid Algorithm.* Paper presented at 2016 International Conferences on Audio, Language and Image Processing (ICALIP)