اونيۆرسيتي تيكنيكل مليسيا ملاك

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

WOOD DEFECT IDENTIFICATION USING CONVOLUTIONAL NEURAL NETWORK FEATURES WITH SUPPORT VECTOR MACHINE CLASSIFIER.

MARTINA BINTI ALI

MASTER OF SCIENCE IN INFORMATION TECHNOLOGY

2025

Faculty of Information and Communications Technology

WOOD DEFECT IDENTIFICATION USING CONVOLUTIONAL NEURAL NETWORK FEATURES WITH SUPPORT VECTOR MACHINE CLASSIFIER.

Martina Binti Ali

Master of Science in Information Technology

2025

WOOD DEFECT IDENTIFICATION USING CONVOLUTIONAL NEURAL NETWORK FEATURES WITH SUPPORT VECTOR MACHINE CLASSIFIER.

MARTINA BINTI ALI

**A thesis submitted
in fulfillment of the requirements for the degree of
Master of Science in Information Technology**

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

Faculty of Information and Communications Technology

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2025

DECLARATION

I declare that this thesis entitled " Wood Defect Identification Using Convolutional Neural Network Features With Support Vector Machine Classifier. " is the result of my own research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature :...

Name : MARTINA BINTI ALI

Date : 30/6/2025

APPROVAL

I hereby declare that I have read this thesis and in my opinion this thesis is sufficient in terms of scope and quality for the award of Master of Science in Information Technology

Signature : ....

Supervisor Name : TS DR UMMI RABA'AH HASHIM

Date : 7/7/2025

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

DEDICATION

This thesis is dedicated to my beloved family, whose unwavering love and guidance have been my greatest strength. To my precious daughter, Nur Annisa Insyirah, whose presence brings endless joy and purpose to my life. To my respected supervisor, Dr. Ummi Raba'ah Hashim, and dear friends, for their invaluable support and encouragement throughout this journey. And to someone who has always stood by me, offering motivation, a listening ear, and steadfast companionship, I am truly grateful. Thank you all for being my constant inspiration and source of strength.

UTeM
اونيۏرسيتي تيكنيكل مليسيا ملاك
UNIVERSITI TEKNIKAL MALAYSIA MELAKA

ACKNOWLEDGEMENT

In the Name of Allah, the Most Gracious, the Most Merciful. First and foremost, I would like to express my sincere gratitude to my supervisor, Dr. Ummi Raba'ah Hashim, for her invaluable guidance, encouragement, and support throughout this research journey. My heartfelt thanks also go to my beloved family for their unwavering love, patience, and encouragement, which have been my greatest source of strength. I am deeply grateful to the Ministry of Higher Education (MOHE), Malaysia, for supporting this research through the Fundamental Research Grant Scheme (FRGS/1/2022/ICT02/UTEM/02/2). My appreciation extends to Fakulti Teknologi Maklumat dan Komunikasi, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia, for providing the resources and support necessary to complete this work. Alhamdulillah, this achievement would not have been possible without the countless blessings and guidance from Allah (SWT).

ABSTRACT

Accurate classification of wood surface defects is essential for maintaining product quality and minimizing material waste in the timber industry. However, achieving high classification accuracy is challenging due to the limited availability of labeled datasets, particularly across diverse wood species. This study proposes a Convolutional Neural Network–Support Vector Machine (CNN-SVM) approach that leverages transfer learning for feature extraction and multi-class wood defect classification. Pre-trained CNN models were employed to extract discriminative features from wood surface images, which were then classified using a Support Vector Machine to enhance accuracy across nine defect classes. The method was evaluated based on classification accuracy and statistical validation. Among the tested models, the ResNet50-SVM combination demonstrated the most consistent and accurate performance. These findings suggest that the CNN-SVM approach offers a viable solution for improving automated wood defect classification.

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

PENGENALPASTIAN KECACATAN KAYU MENGGUNAKAN CIRI RANGKAIAN NEURAL KONVOLUSIONAL DENGAN PENGKLASIFIKASI MESIN SOKONGAN VEKTOR

ABSTRAK

Pengelasan kecacatan permukaan kayu yang tepat adalah penting untuk mengekalkan kualiti produk dan meminimumkan pembaziran bahan dalam industri perkayuan. Namun begitu, pencapaian ketepatan pengelasan yang tinggi sering terjejas oleh kekurangan set data berlabel, terutamanya merentas pelbagai spesies kayu. Kajian ini mencadangkan kaedah gabungan Rangkaian Neural Konvolusi dan Mesin Vektor Sokongan (CNN-SVM) yang memanfaatkan pembelajaran pemindahan (transfer learning) untuk pengekstrakan ciri dan pengelasan pelbagai kelas kecacatan kayu. Model CNN pralatih digunakan untuk mengekstrak ciri diskriminatif daripada imej permukaan kayu, yang kemudiannya diklasifikasikan menggunakan Mesin Vektor Sokongan bagi meningkatkan ketepatan pengelasan bagi sembilan jenis kecacatan. Kaedah ini dinilai berdasarkan ketepatan pengelasan dan pengesahan statistik. Dalam kalangan model yang diuji, gabungan ResNet50-SVM menunjukkan prestasi yang paling konsisten dan tepat. Dapatan kajian ini mencadangkan bahawa pendekatan CNN-SVM berpotensi sebagai satu penyelesaian yang berkesan bagi meningkatkan ketepatan sistem pengelasan kecacatan kayu secara automatik.

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

TABLE OF CONTENTS

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

LIST OF TABLES

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

LIST OF FIGURES

LIST OF ABBREVIATIONS

$UTeM$	-	Universiti Teknikal Malaysia Melaka
MTIB	-	Malaysian Timber Industry Board
AVI	-	Automated Vision Inspection
CNN	-	Convolutional Neural Network
SVM	-	Support Vector Machine
KSK	-	Kembang Semangkuk
ReLU	-	Recitified Linear Unit
GDP	-	Gross Domestic Product

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

LIST OF PUBLICATIONS

The following is the list of publications related to the work of this thesis:

1. Martina Ali, Ummi Raba'ah Hashim, Kasturi Kanchymalay, Aji Prasetya Wibawa, Lizawati Salahuddin, Rahillda Nadhirah Norizzaty Rahiddin. "A Review of Recent Deep Learning Applications in Wood Surface Defect Identification." IAES International Journal of Artificial Intelligence (IJ-AI). 2024.

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

CHAPTER 1

INTRODUCTION

1.1 Background

According to the Malaysian Wood Council (2023), Malaysia's wood industry significantly contributes to the national economy, accounting for over 2% of GDP and providing employment to more than 200,000 individuals. In 2022, the industry generated over RM25 billion in export revenue, underscoring its importance to Malaysia's economic framework. Since the adoption of the Paris Agreement by the United Nations in 2016, global efforts have emphasized resource conservation, energy efficiency, and carbon emission reduction. These principles have become pivotal in the wood industry, as sustainable forest management and optimized wood utilization align closely with the goals of mitigating climate change and promoting environmental sustainability. Historically, wood has been indispensable for various human needs, including construction, tools, and fuel. While modern alternatives have supplanted wood in some applications, recent innovations in wood-based products have increased demand. This increase is driven by global economic growth and a rising aspiration for improved quality of life. However, Malaysia faces challenges in meeting the growing demand for wood and wood products. The current limitations in wood supply, low utilization rates, and insufficient storage and processing capacities are significant barriers to the industry's development. Addressing these issues requires a comprehensive assessment of log and board processing quality to enhance efficiency and improve the quality of wood products. In this context, Malaysia's commitment to the Paris Agreement is a guiding framework for balancing economic growth with environmental stewardship. Malaysia's wood industry can contribute to global climate goals by promoting sustainable practices and advancing wood-processing technologies while sustaining its vital economic role.

Wood is widely used in manufacturing due to its strength, durability, and versatility, making it suitable for various applications. However, finding logs with flawless surfaces is rare in wood manufacturing. As a natural biological material, wood is susceptible to microorganisms that can damage its structure, resulting in defects. These defects, such as irregular tissue formations and structural damage, directly impact the quality and lifespan of wood products (Ding et al., 2020; Hashim et al., 2015). Wood defects are generally categorized into three types: growth defects (caused by physiological factors), pest damage defects (due to pathological factors), and processing defects (resulting from human error) (Y. Chen et al., 2023). Surface defects, in particular, significantly affect the quality of wood finishes, influencing structural strength and aesthetic appeal (Budakci and Cinar, 2004). Depending on their severity, defects are classified as permissible or non-permissible. Permissible defects are minor and acceptable for specific applications, while non-permissible defects render wood unsuitable for products requiring high structural integrity or visual precision (Hashim et al., 2015). The quality of wood has a direct impact on its suitability for specific applications. In the sawmill industry, visual inspection of wood is essential to ensure that it is sorted into grades based on quality. Wood grades are crucial in determining appropriate applications for general market specifications, strips, scantlings, or aesthetic purposes. Grading safeguards the interests of wood buyers and ensures they receive the desired quality. The quality of wood boards also directly determines the value of the resulting products. In Malaysia, the Malaysian Wood Industry Board (MTIB) publishes grading rules and trains wood graders under the Malaysian Grading Rules for Sawn Hardwood Wood (MGR). These rules include provisions for estimating strength and cutting yield based on wood dimensions and defects, such as type, characteristic, frequency, and location. Sorting wood by defect type, size, and frequency helps determine whether it can be repaired, recycled, or discarded. Effective defect identification is essential for maintaining high production quality and safety standards (Tan et al., 2022; Todoroki et al., 2010; P. Xu et al., 2021; Zhou et al., 2014). Studies indicate that defects can reduce wood yield by an average of 10% in manufacturing processes (Clément et al., 2006). Thus, early identification of defective items on the production line is critical to minimizing waste and ensuring product quality (Mohsin et al., 2022). Implementing robust quality control processes ensures that wood products meet performance, safety, and aesthetic standards. Early detection of defects enables manufacturers to take corrective action promptly, maintaining consistency and quality across production (Fang et al., 2017; Todoroki et al., 2010). The wood industry can optimize yield through these measures while upholding the standards demanded by its diverse applications.

Ensuring optimal wood utilization and product quality requires a thorough inspection of log and board processing. Traditionally, the wood industry relied on manual inspection methods, where human operators physically examined wood surfaces to identify defects. While widely used in primary and secondary wood industries, manual inspection had significant limitations, including low accuracy (around 70%), high costs for training, and susceptibility to human error due to fatigue and skill variability (Chun et al., 2021; Gu et al., 2009; Kryl et al., 2020; Urbonas et al., 2019). Moreover, manual inspections were slow, inconsistent, and insufficient to meet high production volume demands (Abdullah et al., 2020; Huber et al., 1985). To address these limitations, the wood industry has increasingly adopted technology powered by intelligent algorithms (Gergel et al., 2020). These advancements have revolutionized defect identification, with deep learning emerging as a powerful tool to improve accuracy and efficiency. Automated Visual Inspection (AVI)

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

systems, equipped with artificial intelligence (AI) and machine vision technologies, now provide faster and more reliable defect detection (Hashim et al., 2015; Kryl et al., 2020; Rahiddin et al., 2020). AVI systems eliminate human limitations, ensuring consistent quality standards while reducing costs and improving production volumes (Y. Huang et al., 2021; Shi et al., 2020; Urbonas et al., 2019). Studies highlight that AVI offers 25% higher identification accuracy than manual methods, resulting in a 5.3% increase in yield and substantial cost savings for rough mills (Buehlmann and Thomas, 2007). Automated grading surpasses traditional inspections by delivering greater precision and optimizing wood resource utilization (Kline et al., 2003). While challenges remain in further enhancing efficiency and yield, AVI is pivotal in improving quality control, benefiting the secondary wood industry by supporting sustainable resource management and consistent product reliability.

In this study, the CNN-SVM method is used, which combines the power of CNNs for feature extraction with the discriminative strength of Support Vector Machines (SVMs) for classification. CNN is employed to automatically extract features from wood surface images, and these features are then classified using an SVM, which is known for its effectiveness in high-dimensional spaces. The proposed method is aimed at improving classification accuracy across nine defect classes spanning four wood species.

This study will utilize the wood defect image dataset introduced by Hashim et al. (2015) to validate the CNN-SVM method. The performance of the method will be evaluated primarily based on classification accuracy, ensuring that it is capable of distinguishing between the various wood defects effectively.

Additionally, advancements in data augmentation and transfer learning (Chun et al., 2022) are incorporated to enhance the robustness of the model, especially in scenarios where labeled data is limited. These techniques help improve the generalization of the CNN-SVM model, thus ensuring better performance in real-world applications. This approach aims to provide an automated solution for wood defect classification, improving the accuracy and enhancing quality control in wood production processes.

1.2 Problem Statement

The wood industry faces significant challenges in efficiently classifying surface defects during production. Traditional manual inspection methods rely on human operators and are often inaccurate, time-consuming, and prone to errors due to visual fatigue, human error, and limited processing capacity (Rahiddin et al., 2020). These inefficiencies lead to reduced production quality, increased waste, and lower yields, ultimately affecting the overall efficiency of wood manufacturing processes. Surface defects, which can compromise wood products' strength, durability, and aesthetic value, are particularly challenging to classify due to the diverse and complex characteristics of different wood species and defect types (He et al., 2020).

Automated Visual Inspection (AVI) systems have been introduced to address these limitations, offering improved scalability and consistency in defect classification. However, existing systems frequently with multi-class defect classification, especially when faced with imbalanced datasets and the scarcity of defect images required for training deep learning models (Hu et al., 2019). Data imbalance, a prevalent issue in wood defect datasets, often leads to biased models that favor majority classes, resulting in poor classification performance for minority classes (Dong et al., 2018; Johnson and Khoshgoftaar, 2019). While data augmentation and transfer learning techniques have shown promise in addressing these issues, a gap exists in leveraging advanced deep learning architectures tailored for the wood industry (Han et al., 2018; Ibrahim et al., 2018).

Recent advancements in machine learning have demonstrated the effectiveness of combining Convolutional Neural Networks (CNNs) for feature extraction with Support Vector Machines (SVMs) for classification, particularly in domains such as medical image analysis. CNNs extract high-level, discriminative features, while SVMs provide strong generalization capabilities for efficient classification (He et al., 2019). However, this robust framework has yet to be fully explored in the context of wood defect classification, presenting an opportunity to enhance classification accuracy and efficiency in this domain.

Despite the promise of CNN-SVM frameworks, several research gaps remain in the field of wood defect classification:

i). Limited Application of CNN-SVM in Wood Defect Classification: While CNN-SVM methods have succeeded in other areas (e.g., textile defect detection (Qiu et al., 2021), fruit fly classification (Peng et al., 2020), and cognitive disease diagnosis (Huang et al., 2021)), their potential to enhance accuracy in wood defect classification is largely unexplored.

ii). Underutilization of Transfer Learning: Although transfer learning helps improve model performance with small datasets, its integration with CNN-SVM approaches tailored for wood defect types and species is limited.

iii). Addressing Class Imbalance through Data Augmentation: Wood defect datasets are often imbalanced, which can reduce classification accuracy for less-represented defect types. This research addressed the issue by applying data augmentation techniques to balance the class distributions, enabling a more accurate evaluation of CNN-SVM performance across all defect categories

iv). Lack of Statistical Validation: Most studies report accuracy without statistical testing, which limits confidence in the significance of their results.

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

v). Absence of a Tailored Framework for Multi-Class Defect Classification in Wood: Many existing AVI systems struggle with multi-class defect classification in the wood industry, mainly when dealing with diverse wood species and defect types. There is a gap in developing frameworks specifically designed for this context.

This research addresses the above gaps by applying a CNN-SVM method that leverages transfer learning to classify nine wood defect classes across four wood species. It focuses solely on maximizing classification accuracy. The approach is validated through accuracy metrics and statistical analysis. Ultimately, this study contributes to advancing defect classification accuracy in automated wood inspection systems.

1.3 Research Question

This study seeks to answer the following research questions:

i). How can pre-trained CNN models be effectively utilized for feature extraction in combination with SVM to accurately classify multi-class natural defects on wood surfaces?

ii). How can the performance of the proposed CNN-SVM method be evaluated across different wood species based on classification accuracy and statistical analysis?

1.4 Research Objective

This study aims to address the challenge of accurately classifying multiple types of natural wood surface defects. The specific objectives are:

i). To implement a CNN-SVM method utilizing transfer learning for the accurate classification of multi-class natural defects on wood surfaces.

ii). To assess the classification performance of the proposed CNN-SVM method across different wood species based on accuracy and statistical evaluation.

1.5 Scope of Research

The scope of this research is as follows:

i. Data Sources

- The dataset used in this study is sourced from the UTeM Wood Defect Database (Hashim et al., 2015), which provides labeled images of wood defects for classification tasks.
- The dataset consists of nine defect categories: blue stain, brown stain, borer holes, knot, bark pocket, rot, split, wane, and clear wood (non-defective). These categories represent common surface defects observed in the wood industry.
- The dataset includes 2700 samples per species, with 300 images per defect class, ensuring a balanced dataset for training and evaluation.
- The wood defects in the dataset are represented across four Malaysian hardwood species: Merbau, Kembang Semangkuk (KSK), Rubberwood, and Meranti. These species were selected to reflect commonly used materials in the Malaysian wood industry, each exhibiting a range of surface defects.
- The wood samples were collected from secondary wood product factories in the Bukit Rambai Industrial Area, Melaka, Malaysia. This sourcing provides a realistic representation of wood defects typically encountered in industrial settings.

UNIVERSITI TEKNIKAL MALAYSIA MELAKA