SEMARAK ILMU PUBLISHING

Journal of Advanced Research in Applied Mechanics

Journal homepage:
https://semarakilmu.com.my/journals/index.php/appl_mech/index
ISSN: 2289-7895

Sound Classification for Non-Destructive Diagnosis of Basal Stem Rot Disease based on Stem Density in Oil Palm Trunks

Chalatorn Augsornthip[1], Paramin Neranon[1,*], Pornchai Phukpattaranont[2], Arisara Romyen[3]

[1] Department of Mechanical and Mechatronics Engineering, Faculty of Engineering, Prince of Songkla University, Songkhla 90110, Thailand
[2] Department of Electrical and Biomedical Engineering, Faculty of Engineering, Prince of Songkla University, Songkhla 90110, Thailand
[3] Research Center for Fundamental Economic Development and Agricultural Economic Potential in the Southern Region, Faculty of Economics, Prince of Songkla University, Songkhla 90110, Thailand

ARTICLE INFO

Article history:
Received 18 June 2024
Received in revised form 31 July 2024
Accepted 8 August 2024
Available online 30 August 2024

Keywords:
BSR disease; Ganoderma; Sound recognition; Convolutional Neural Network (CNN); Support Vector Machine Classifier (SVC); Multi-Layer Perceptron (MLP)

ABSTRACT

Oil palm trees play a crucial role in the economies of Southeast Asian countries, providing palm oil that is essential for various products and consumption. However, the Basal Stem Rot (BSR) disease caused by the white-rot fungus Ganoderma spp. poses a significant challenge, leading to wilting and death of the trees. This study proposes a practical and accessible approach for the routine detection of Basal Stem Rot (BSR) disease in oil palm trees caused by Ganoderma spp. A prototype system for initial disease diagnosis is developed by utilizing the local wisdom technique of tapping and assessing knocking sounds based on stem density in oil palm trunks. Three algorithms, Convolutional Neural Network (CNN), Support Vector Machine Classifier (SVC), and Multi-Layer Perceptron (MLP) models are compared for accuracy in two sound recognition experiments. The first experiment involves a three-class classification model (healthy, boundary, and infected areas), while the second excludes the boundary area. Results demonstrate that the CNN model outperforms SVC and MLP, achieving the highest accuracy of 90.73% in the two-class scenario and 84.97% in the three-class scenario, along with superior precision, recall, and F1 score. The CNN model proves to be the most effective in accurately classifying the data in both scenarios compared to the other models. This prototype system offers promising implications for the early detection and proactive management of Ganoderma spp. disease, contributing to the preservation and sustainability of oil palm plantations in Southeast Asia.

1. Introduction

The palm oil industry holds significant importance as an agro-industry in Southeast Asia, particularly in countries such as Indonesia, Malaysia, and Thailand. It plays a vital role in boosting the economy and contributing to the GDP of these tropical-climate nations. In Thailand, a substantial portion, approximately 30 percent, of the agricultural land is dedicated to oil palm cultivation. The global market exhibits a growing demand for palm oil production. This increased demand can be

* *Corresponding author.*
E-mail address: paramin.n@psu.ac.th

https://doi.org/10.37934/aram.124.1.1938

attributed to the high-quality oil produced by oil palm trees, which finds extensive application in healthy cooking due to its rich content of vitamins and antioxidants. Additionally, from Asian Agri's research [1]. Palm oil serves as a raw material for the production of various products, including biofuel, butter, ice cream, detergents, soaps, cosmetics, etc.

Oil palm trees are generally defenceless to various fungal infections, and one of the most destructive diseases they face is known as basal stem rot (BSR) caused by Ganoderma spp. as shown in Figure 1(a) from [2]. BSR significantly impacts oil palm trees, leading to direct stand loss, reduced palm yields, and ultimately, tree death (as shown in Figure 1(b)). Technically, the disease's effect penetrates the palm tree from the root to the top. Symptoms of BSR include pale leaves, rotting leaflets, and decay of the stem. This extensive damage to the stem results in the formation of cavities within the palm trunk, further damaging the structural integrity of the oil palm tree.

Several methods have been employed for the identification of BSR disease. As extensively documented [3,4], infected oil palm trees can be diagnosed through field observation and visual assessment. Regular inspections are conducted to detect symptoms such as leaf yellowing and wilting. Moreover, careful examination of Ganoderma's basidiocarp on the tree trunk's base is essential. The initial appearance of Ganoderma spp. is characterized by the presence of brown or decaying tissue. However, it is crucial to emphasize that once oil palm trees reach the stage of internal decay, treatment becomes much more challenging. In such cases, farmers often resort to the removal of infected trees to prevent the further spread of the disease through wind-dispersed spores.

(a) (b)

Fig. 1. (a) Full-growth pathogen (b) Infected oil palm tree

In the field of agricultural research, innovative methods for diagnosing the health of oil palm trees have been explored with notable success. A significant contribution in this area was made by Khairunniza *et al.,* [4] who experimented with differentiating between healthy and diseased oil palm trees. Moreover, they advanced their study by categorizing the stages of the disease into four distinct levels. By utilizing Terrestrial Laser Scanning (TLS), they captured high-angle images of the trees, and combined this with measurements of the Diameter at Breast Height (DBH) - the diameter of the tree trunk measured at approximately 1.3 meters from the ground. This ground-breaking approach not only provided a novel method for diagnosing the health of oil palm trees but also paved the way for more detailed and accurate disease staging, offering valuable insights for the management and treatment of these vital agricultural assets.

An advanced diagnostic tool, Polymerase Chain Reaction (PCR). Amiri *et al.,* [5] has been developed for early detection. PCR is a molecular biology technique invented by Kary Mullis in 1983, used to amplify specific DNA sequences in a sample. This technique works by repeatedly separating the DNA strands and binding specific primers to the Ganoderma spp. DNA sequence if pathogens are present in the tested sample. After detection, the amplified DNA can be visualized through gel

electrophoresis, where the presence of Ganoderma spp. is indicated by a visible band corresponding to the DNA sequence. Nevertheless, PCR is a powerful technique that requires specialized knowledge, equipment and laboratory facilities to perform accurately. For farmers, PCR might not be easily accessible or practical for their routine detection of plant diseases such as BSR disease, since it involves a certain level of expertise in molecular biology. Therefore, while the PCR technique is a powerful diagnostic tool, its widespread implementation among farmers may be challenging.

Researchers have widely made significant advancements in the development of smart tools for diagnosing BSR disease. One notable advancement is the utilization of Electrical Resistance Tomography (ERT) for early detection. It was also conducted by Elliott *et al.,* [6] and Hamidon and Mukhlisin [7]. ERT is a geophysical technique employed to image sub-surface structures within affected palm trunks by measuring electrical resistivity. The objective of the study was to identify internal decay and structural abnormalities in infected oil palm trees at an early stage. The experiment and subsequent analysis demonstrated the potential of ERT as a viable device for primary detection of Ganoderma spp. disease.

However, In Thailand, the widespread implementation of TLS, PCR or ERT methods faces significant challenges due to limited availability and budgetary constraints. Technically, Thai farmers resort to a local wisdom technique called sound classification for detecting Ganoderma spp. disease. Following folk wisdom, experts proficiently knock on the tree trunk and classify the resulting knocking sound to distinguish between infected and healthy oil palm trees affected by BSR disease. Similarly, Mr. Prasert [8] employed local wisdom to categorize fruit ripeness through sound analysis. The knocking sounds produced by different trees vary, depending on their health status. In addition, Kharamat *et al.,* [9] introduced a method for classifying the ripeness of durian fruit based on the knocking sounds using a Convolutional Neural Network (CNN). The study demonstrates the successful application of CNN in accurately categorizing durian fruit ripeness based on the sound produced when tapped. This research contributes to the development of automated techniques for fruit ripeness classification, offering potential benefits for the agricultural industry.

In this method mentioned above, experts carefully tap the fruit and assess the resulting knocking sound. This method offers a practical solution by utilizing simple equipment, effectively solving limitations related to restricted availability and budget constraints. However, due to varying levels of experience and listening capability, not all farmers can use this method. Consequently, only a limited number of farmers can effectively detect BSR disease. Several farmers can start identifying BSR disease when it has reached a severe stage. This is characterized by observable symptoms such as yellowing and wilting of the leaves, undersized leaves, as well as the presence of brown or decaying tissue. Additionally, the detection of Ganoderma spp. root rot or butt rot further confirms the presence of the disease. The recovery of the tree at the severe stage becomes significantly challenging and necessitates substantial financial investment.

After an extensive review [9-11], it has been observed that there is a lack of research studies on detecting BSR disease utilizing the local wisdom technique of tapping and assessing the resulting knocking sound. To address the notable research gap, this paper consequently proposes a novel approach to diagnose BSR disease in oil palm trees by emulating the sound-based classification derived from local wisdom. To replicate the expertise of local wisdom, machine learning techniques are employed, utilizing data and algorithms to mimic the human learning process. The integration of machine learning technology enables the establishment of a new, simplified technique for diagnosing BSR disease in oil palm trees.

The subsequent sections of the research present the finer details of the proposed study. These sections encompass various aspects, including the development of a prototype algorithm, the integration of local wisdom and machine learning, the recording and extraction of knocking sounds,

the transformation of sounds into frequency-domain spectrograms, and a comparative analysis of machine learning models. These detailed sections contribute to a comprehensive understanding of the research methodology and its potential implications.

2. Methodology

As extensively reviewed, it has been identified that the density within the trunks is a key distinguishing characteristic between healthy and unhealthy oil palm trees. Ganoderma spp. disease initially develops inside the trunk and progresses outward, causing decay and brown rot within the trunk. The study by Najmie *et al.,* [10] significantly contributes to the understanding of how Ganoderma spp. disease impacts the density and ultrasonic properties of oil palm trunks. The research findings have important implications for the detection and monitoring of Ganoderma spp. disease in oil palm plantations, enabling the implementation of timely disease management strategies. Abu-zanona *et al.,* [11] demonstrates a substantial 50% reduction in stem density in oil palm trunks infected by Ganoderma spp. disease compared to healthy trunks. However, the technique employed in the study may not be easily accessible or practical for routine detection of BSR disease among Thai farmers due to the reliance on a bulky ultrasonic measurement system that requires collaboration with an external microsecond timer device and an AC signal. Thus, its applicability in oil palm fields may be limited.

In addition, Tan *et al.,* [12] propose a non-destructive and efficient method for detecting Ganoderma spp. disease in oil palm trees using near-infrared spectroscopy (NIRS) classification. This review paper enhances our understanding of the potential of NIRS for disease detection in oil palm plantations. However, it is important to note that NIRS classification may not be suitable for Thai farmers due to their bulky systems and the need for multiple components that are difficult to transport.

Furthermore, Koh *et al.,* [13] have discovered that electrical impedance can be effective in measuring the density of fruits. However, when applied to palm trees, which exhibit high impedance, significant challenges emerge. These challenges include the necessity for amplification circuits to address low electrical currents and the requirement for oscilloscopes to measure the modified current. Consequently, this method becomes expensive and impractical, particularly when used by Thai farmers to identify Ganoderma spp. disease.

Therefore, to address the limitations above, this research aims to develop a sound classification system for detecting Ganoderma spp. disease in oil palm trees. By analyzing and classifying specific sounds, this system offers a potential solution for early detection and proactive management of Ganoderma spp. disease. Machine learning techniques are utilized to replicate the expertise of local wisdom, enabling the establishment of a simplified technique for diagnosing stem density in oil palm trunks. The following sub-topics cover data acquisition, system design, signal processing, machine learning algorithms, parameter selection, and evaluation indicators, respectively, for oil palm disease detection.

Machine learning was invented to make the computing system of the computer replicate human intelligence by putting the input data and answers to gain rules. From the literature review, some studies have explored the use of machine learning techniques to classify oil palm disease based on spectrum reflection from the leaves. Abu-zanona *et al,* [11] created a diagnosis model to classify four common diseases threatening palms, Bacterial leaf blight, Brown spots, Leaf smut, and white scale by importing labelled pictures to Convolutional Neural Network (CNN), which gives 99.10% accuracy. Not only those four common diseases but also the BSR has been used to create a diagnosis model.

Tan *et al.,* [12] investigate the use of machine learning techniques to classify Ganoderma spp. using near-infrared spectral data. K-Nearest Neighbor (KNN), Naïve Bayes (NB), Support Vector Machine Classifier (SVC), and Decision Tree (DT) approaches were used to diagnose BSR disease by measuring spectrum reflection from the leaves. The study demonstrates the potential of this approach for accurate identification of the fungus responsible for BSR disease in oil palm trees. The research contributes to improved disease management strategies in oil palm plantations. Lee *et al.,* [14] also proposed a method for early detection of BSR disease in oil palm trees using hyperspectral images, which were captured from a top-down perspective. The approach utilizes a Multi-Layer Perceptron (MLP) model trained on hyperspectral image data to classify healthy and diseased oil palm trees. The study highlights the potential of machine learning in achieving more accurate and efficient detection of BSR disease. To complete the aim of this research, three supervised machine learning algorithms were selected: CNN, SVC, and MLP approaches.

2.1 Data Acquisition

The knocking sound data collected from both healthy and unhealthy oil palm trees have been divided into in-sample and out-sample groups. To collect this data, the palm trunk was repeatedly tapped using a rubber mallet measuring 6.5 cm in diameter and weighing a total of 715g. For its durability, versatility, and capability to handle high sound pressure levels, the Apple AirPods Pro was selected as the recording device for capturing the knocking sounds. Prior to recording the sound, it is essential to remove any dead palm fronds and clean the tree trunk.

The experimental apparatus was set up, as depicted in Figure 2. In case, a variety of magnitudes of hitting sounds is allowed to design the machine learning system with a diverse range of knocking sound inputs. This approach ensures that the machine algorithms can effectively classify and identify Ganoderma spp. disease in oil palm trees, accommodating different intensities of knocking sounds commonly encountered in real-world scenarios.

In this research, the data were collected from three different southern provinces in Thailand, including Satun and Nakhon Si Thammarat, respectively. The dataset comprises 2075 knocking sounds obtained from approximately 60 oil palm trees, divided equally between 30 trees in Satun and 30 in Nakhon Si Thammarat. Each tree yielded approximately 30 to 40 distinct knocking sound files. These recordings are categorized into three classes: 748 files from the healthy class, 717 files from the infected class, and 610 files from the boundary class. From 2075 sound files is our dataset, the cross-validation process involves partitioning it into three separate segments: 60% for training, 15% for validation, and 25% for testing purposes.

(a) (b)

Fig. 2. The experimental setup when collecting the knocking sound data from (a) healthy area (b) the infected area

The expert categorized the BSR disease into three types:

i. Healthy: Represents the healthy area with a solid trunk density, producing a distinct knocking sound when hit;
ii. Infected: Represents the area that has been damaged by the disease, resulting in a softer trunk compared to other areas;
iii. Boundary: Represents the borderline between Healthy and Infected areas. The boundary area has a similar solidness to the healthy area, but the knocking sound is more similar to the infected area.

This diverse dataset allows the machine learning system to learn and differentiate the various knocking sounds associated with different stages of Ganoderma spp. disease, improving the accuracy and effectiveness of disease detection.

2.2 System Design

Designing the sound classification system for detecting Ganoderma spp. disease in oil palm trees involves several key steps that can be illustrated in Figure 3. This research paper focuses on constructing prototype machine-learning models that replicate human sound perception, similar to the local wisdom mentioned in the introduction. The knocking sound data captured from healthy and unhealthy oil palm trees can be separated into two groups, i.e., in-sample and out-sample groups. To achieve durability, versatility, and ability to handle high sound pressure levels, making them suitable for capturing knocking sounds, a dynamic microphone is generally a good choice for recording knocking sounds was used. Before analysis, the dataset undergoes noise cancellation to ensure accuracy and make the model more corrected. The transformed data, converted into MFCCs (Mel frequency cepstral coefficient), is then subjected to three different algorithms (CNN, SVC, and MLP) for efficiency comparison, elaborated in the next section. Additionally, continuous improvement and validation processes ensure that the system remains accurate and reliable in real-world scenarios.

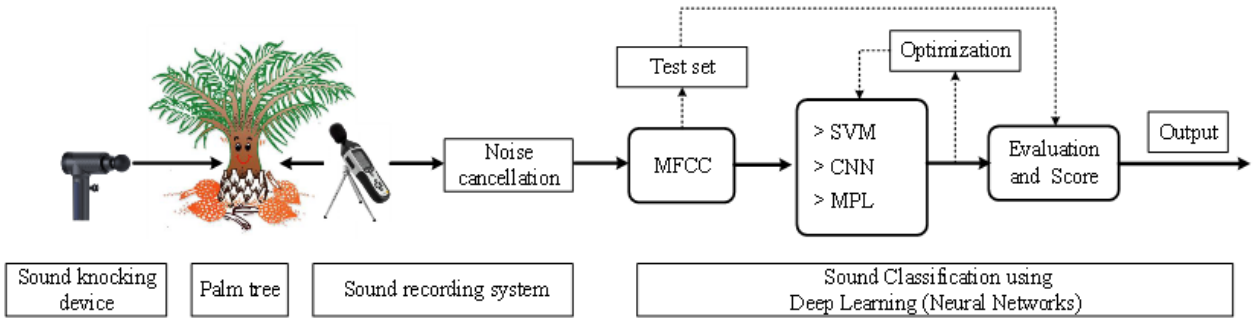

Fig. 3. System block diagram of the sound classification of basal stem rot disease based on stem density in oil palm trunks

2.3 Noise Cancellation & MFCCs Extraction

Integrating noise cancellation in machine learning-based sound classification enhances the Signal-to-Noise Ratio (SNR), crucial for extracting meaningful features and achieving accurate classification. This approach, valuable in preprocessing, also mitigates overfitting risks and improves model interpretability by removing irrelevant noise. In recent developments, Audacity, an open-

source audio editing software, effectively employs noise cancellation by combining the SNR method with the Noise Gate function. The SNR method enhances signal clarity by improving the ratio between the relevant signal and background noise. The Noise Gate function complements this by reducing sounds below a certain threshold, thus minimizing ambient noise. This combined approach creates cleaner audio inputs, crucial for machine learning models. Raja *et al.,* [15] utilized this technique in Audacity, optimizing the Noise Gate function to match the specific noise characteristics of their recorded area, aiming to obtain the cleanest sound files for their models. This demonstrates the importance of tailored noise reduction in preparing high-quality datasets for machine learning applications.

Mel-Frequency Cepstral Coefficients (MFCCs) are feature extraction techniques employed in speech and sound signal processing to extract relevant information from a sound signal. They are based on the Mel-frequency scale, a nonlinear scale approximating the human auditory system's frequency response. The process of extracting MFCCs involves dividing the sound signal into short frames and applying a filter bank to transform the magnitude spectrum into the Mel-frequency scale. Then it takes the inverse Discrete Fourier Transform (DFT) of the logarithm of the Mel-frequency spectrum to obtain the cepstral coefficients. These resulting MFCCs offer a concise representation of the sound signal's spectral characteristics, making them valuable input features for machine learning models and other signal processing algorithms.

The transformation of all DFT coefficients into MFCCs is accomplished using Eq. (1). Subsequently, the Mel-scale coefficients undergo logarithmic compression, followed by the application of the discrete cosine transform to obtain the cepstral coefficients. In this study, the selection of 13 cepstral coefficients was influenced by extensive research [16-18], where a range from 12 to 20 coefficients was considered. Given that the complexity of our signals is not as high as in speech recognition tasks, 13 coefficients were deemed appropriate for our research. The number of coefficients has a significant impact on a model's ability to capture the complexity of the signals. Higher numbers of coefficients can provide more detailed information and enable more precise feature extraction. However, excessively high numbers may lead to overfitting and result in performance issues for the model.

$$M(k) = 2595 * log_{10}\left(1 + \frac{f(k)}{700}\right) \quad (1)$$

2.4 Machine Learning Algorithms

This research employs supervised machine learning algorithms, which are computational statistical models enabling computers to learn from provided data and establish relationships for predictions or decision-making in various tasks. Through experience and feedback, these algorithms can enhance their performance over time. Specifically, the study focuses on classifying the BSR disease in oil palm trees using sound classification. Supervised machine learning involves training the algorithm on labelled data, where input features (knocking sounds) are paired with corresponding target outputs (healthy, infected, and boundary classes). The algorithm learns from these labelled examples to recognize patterns and classify new or unseen data. For this purpose, three supervised machine learning algorithms were selected: CNN, SVC, and MLP approaches.

2.4.1 Convolution Neural Network (CNN)

CNN is a deep learning algorithm comprising convolutional layers that apply filters to input data, capturing relevant features. To reduce dimensionality, extracted features are then processed through pooling layers. It was initially developed by Yann LeCun *et al.,* [19]. The CNN algorithm is a type of neural network consisting of multiple layers called deep learning, including convolutional layers, pooling layers, and fully connected layers. Convolutional layers utilize learnable filters or kernels to extract local features such as edges, corners, and textures from the input image, as represented by Eq. (2). This process enables CNNs to effectively recognize and interpret visual patterns, making them highly suitable for tasks like image classification and computer vision.

$$y_{m,n} = x_{m,n} * h_{m,n} = \sum_{j=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} x_{i,j} \cdot h_{m-i,n-j} \tag{2}$$

From Eq. (2), convolved matrix y position [m,n] is calculated by the sum of kernel x multiplied by the input matrix h, while "i" and "j" sequentially vary from -∞ to the size of row and column in kernel x. However, the pooling layers downsample the feature maps produced by the convolutional layers to reduce the spatial resolution and extract the most salient features.

2.4.2 Multilayer Perceptron (MLP)

The definition of MLP is a type of neural network that comprises at least three layers: the input layer, the hidden layer(s), and the output layer (Goodfellow *et al.,* [20]). Unlike CNN, MLPs focus on direct data processing without involving specialized data extraction techniques. This process enables the MLP to learn and represent complex relationships within the data, making it capable of solving various tasks such as classification, regression, and other machine-learning problems. The input data undergoes a series of operations within the hidden layers, where each node multiplies the input data by weight and adds a bias term, as represented by Eq. (3) Below:

$$h_i = a\left(\sum_{i=1}^{n} x_i w_i + b\right) \tag{3}$$

where h_i is the weighted values computed by the sum of input x_i multiply by the weight value from each node in the hidden layer plus bias. Then the values of this summation will pass to the activation function in the hidden layer. The softmax function is used in the output layer to obtain probabilities for class labels. The network adjusts weights and biases to minimize the loss and improve predictions through iterative iterations and epochs.

2.4.3 Support Vector Machine Classifier (SVC)

SVC is a supervised learning algorithm invented by Vladimir Vapnik and Alexey Chervonenkis in the 1960s and 1970s. It is a variant of the Support Vector Machine (SVM) algorithm and is widely used for classification tasks. SVC aims to find the best hyperplane that separates data points into different classes based on their features. It achieves this by transforming the input features into a higher-dimensional space using kernel functions like linear, polynomial, or radial basis function (RBF) kernels. SVC has advantages such as handling non-linearly separable data, performing well with high-

dimensional features, and having good generalization performance. However, SVC can be sensitive to the choice of kernel function and its parameters and can be computationally intensive for large datasets. Based on the information, SVC is regarded as one of the most popular machine learning algorithms. Therefore, we intend to include this algorithm in our study for comparison with CNNs and MLPs. The key SVC equation used to find the optimal hyperplane for classification can be expressed as:

$$f(x) = sign(wx + b) \tag{4}$$

where f(x) is the predicted class label for the input data point x; w is the weight vector, representing the orientation of the hyperplane; x is the feature vector of the input data point; b is the bias term, which shifts the hyperplane from the origin; sign() is the sign function, which returns +1 if the argument is positive or zero, and -1 if it is negative (Long *et al.,* [21]). During the training process, SVC aims to find the optimal values for the weight vector w and bias term b that define the hyperplane, separating the data points of different classes with the maximum margin.

2.5 Machine Learning Parameters Selection

Machine learning parameter selection is key to achieving optimal model performance and generalization. Different machine learning algorithms have parameters that need to be suitably set before training the model. The selection of these parameters directly influences how the model learns and represents the underlying patterns in the data. In this study, we conducted a pilot experiment to optimize the three machine-learning parameter selections. The pilot experiment aimed to find the best combination of parameter values for each machine-learning model. By conducting this preliminary experiment, we aimed to identify the optimal configurations that would yield improved performance and generalization in the subsequent phases of the study. In addition, optimizing or tuning hyperparameters in machine learning tasks was employed in controlling the learning process to achieve an optimized model.

In the experiments, we utilized BayesianSearchCV for Architecture and Hyperparameter selection. This powerful approach leverages probabilistic predictions to identify the set of hyperparameter values that yield the highest accuracy among the numerous architectures and hyperparameter options specified. Technically, BayesianSearchCV is utilized for architecture and hyperparameter selection in machine learning models. The Bayesian optimization technique efficiently searches the hyperparameter space, helping to identify the optimal architecture and hyperparameters for enhanced model performance. The selection of the architecture and hyperparameters for BayesianSearchCV was guided by insights from various reviews [22-29]. These reviews provided valuable information on best practices and trends in model design, enabling the researchers to define a comprehensive and promising scope for the optimization process. BayesianSearchCV for architecture and hyperparameter selection involves defining the search space, initializing the optimization, performing cross-validation, updating the Bayesian model, suggesting new hyperparameters, and repeating the process to find the best configuration for improved model performance.

2.6 Performance Evaluation

Data splitting involves dividing a knocking sound dataset into subsets for training, validation, and testing sessions. The training set is used to teach the model, the validation set to fine-tune it, and the

testing set to assess its performance on unseen data. BaysiensearchCV, incorporating cross-validation in its process, was used for model development. Machine Learning Evaluation Indicator refers to the metrics used to assess the performance of machine learning models. These indicators provide valuable insights into the model's accuracy, precision, recall, F1 score, and other measures, helping to evaluate and compare the model's effectiveness in solving specific tasks. Various machine learning techniques were designed for specific purposes, and each one has a unique technique to generate human decisions. Consequently, the machine-learning model scoring indication must serve as the benchmark for comparing each model's score. For more detail, accuracy represents the proportion of correctly classified instances out of the total instances in the dataset.

Precision is the ratio of true positive predictions to total positive predictions. In other words, it measures how many predicted positive instances were true positives. Higher precision indicates fewer false positives. Recall, also known as sensitivity or true positive rate, is the ratio of true positive predictions to the total number of actual positive instances. It measures the model's ability to correctly identify positive instances. Higher recall indicates fewer false negatives. The F1 score is the harmonic mean of precision and recall and provides a balance between the two metrics. It is a useful metric when there is an imbalanced class distribution. It can evaluate the performance of a model more comprehensively and consider the trade-off between false positives and false negatives. The evaluation metrics made up of Accuracy, Precision, Recall, and F1 score can be calculated by Eq. (5)–Eq. (8) (Hastie *et al.,* [30]), whereas true (T) and false (F) represent labelled data, and positive (P) and negative (N) represent the validity of label data, respectively.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{5}$$

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

$$F1\ score = \frac{2 \times Precision\ \times Recall}{Precision + Recall} \tag{8}$$

3. Experimental Results and Discussion

This section presents the study's outcomes, evaluating the performance of machine learning models in sound classification based on stem density in oil palm trunks. It discusses the experimental results, and analysis, which draws discussion on the models' effectiveness and potential applications. According to the experiment, each model building, training, and testing step was performed using Google Colaboratory as the computational resource. The system utilized was equipped with a 2vCPU with a clock speed of 2.2GHz and 13GB of RAM. For the machine learning process, no pre-trained models were employed. Instead, all models were developed from scratch using PyTorch 2.0.1. The key purpose of this approach was to gain insights into the dataset's nature and characteristics, paving the way for future enhancements and improvements.

The experimental results can be categorized into two main groups: sound characteristics (section 3.1) and performance comparison (section 3.2). In the Sound Characteristic section, we analyze the sound features extracted from the oil palm trunks, providing insights into the sound characteristics relevant to disease detection. In the Performance Comparison section, we compare the accuracy,

precision, recall, and F1 score of the machine learning models used, evaluating their effectiveness in classifying healthy and infected areas based on sound data.

3.1. Sound Characteristics

The experimental result of sound characteristics provides an analysis of the sound features extracted from the oil palm trunks, revealing key sound characteristics relevant to disease detection. This section presents three key results: (1) noise cancellation results: producing the outcomes of noise cancellation techniques and demonstrating the effectiveness of reducing unwanted noise and enhancing the signal quality; (2) classification power spectrum results: providing insights into the frequency distribution and spectral characteristics of the sound data; and finally (3) transformed knocking sound data to MFCCs: transforming the knocking sound data to MFCCs, which allows for further analysis.

The dataset used in the experiment was recorded at a sample rate of 24 kHz, capturing various sounds, including the knocking sound along with background noise from unrelated sources such as people chatting, footsteps, and wind blowing. In Figure 4(a), the power spectrum of the signals clearly shows the presence of this noise, which can interfere with the accurate detection of the knocking sound. To enhance the quality of the knocking sound and remove the unwanted noise, noise cancellation techniques were applied, as shown in Figure 4(b). The cleaned knocking signal is now ready for further processing, ensuring a more accurate representation of the original knocking sound.

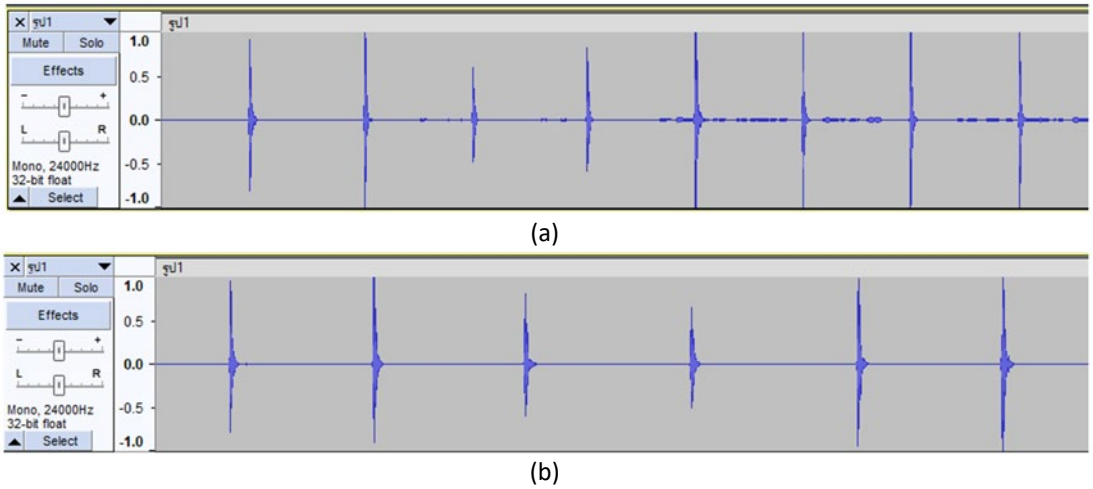

Fig. 4. result of sound characteristic (a) a knocking sound signal before using noise cancellation (b) a knocking sound signal after using noise cancellation

After obtaining the cleaned sound signals, each sound was split into 0.5-second intervals to prepare them as inputs for the machine-learning models. The 0.5-second interval time was chosen to precisely locate the knocking sound within each signal. Overall, these steps in data preprocessing and cleaning are crucial for providing the machine learning models with accurate and relevant input data, enabling them to perform sound classification more effectively and accurately.

The power spectrum is a Fourier transform tool used to analyze the frequency content of a signal. It shows how the power or energy of the signal is distributed across different frequencies, revealing the strength of various frequency components. Figure 5 displays a recorded sound file capturing the sound of hitting an oil palm tree, which contains 100 instances of hitting. The dominant frequency components and their strengths are identified in a set of a hundred infected knocking sounds.

These signals were subjected to the Fourier transform, and the power spectrum of the results in which the analysis confirms that a diverse range of knocking sounds with varying magnitudes produce similar spectrum profiles. The pattern in the knocking sound spectrum in Figure 5 appears consistent, with minor differences in the amplitudes of the curve. If the spectral patterns of the impact sounds are alike, it follows that controlling the force of impact is unnecessary for this project. Additionally, based on the findings of Kaiming *et al.,* [31] 's studies, the CNN machine learning algorithm demonstrates accurate sound classification even when the sounds have varying amplitudes since the CNN algorithm's performance remains unaffected by the amplitude of the signals.

Based on the characteristics of the knocking sound signals, the choice of 13 coefficients strikes a balance between capturing relevant information and mitigating potential overfitting. In Figure 6, we visualize examples of the knocking sound signals transformed into MFCCs for each class: healthy, infected, and boundary. The MFCC representations provide a concise and informative representation of the knocking sound, allowing for easy visual comparison between the different classes. This helps in understanding the distinctive features and patterns present in each class of knocking sounds.

Additionally, to ensure the MFCC representations are more robust and comparable, cepstral mean normalization is applied. This normalization process centres the MFCC coefficients on zero, reducing the impact of differences in amplitude or volume among the sound signals. It standardizes the coefficients, making them more suitable for feeding into the machine learning models. Once the dataset is transformed into MFCCs with 13 coefficients and normalized, it is ready for the machine learning model-building process in the subsequent section.

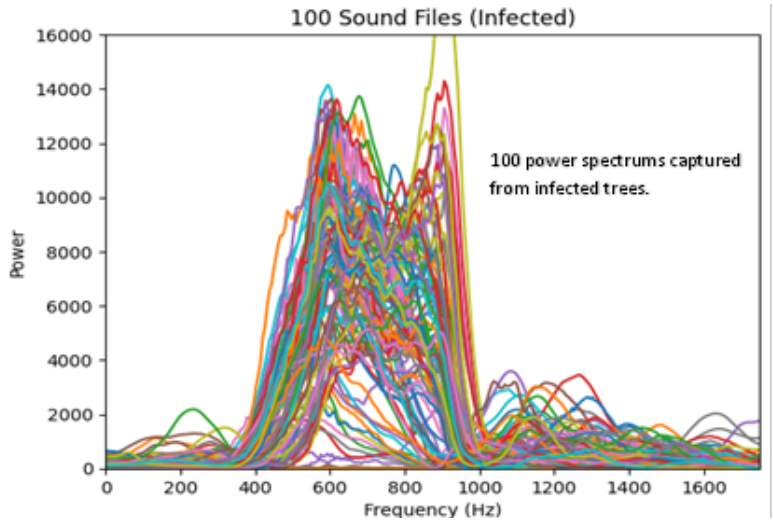

Fig. 5. The plot displays the power spectrum characteristic of the infected knocking sound signals which are not control the amplitudes, obtained from a dataset of 100 files

3.2. Performance Comparison

The primary objective of this study is to compare the performance of three classifiers: CNN, MLP, and SVC, in classifying knocking sounds based on the density within oil palm trunks. This density serves as a crucial distinguishing characteristic between healthy and unhealthy oil palm trees. Additionally, the study aims to compare the classifiers' performance using different input features: one scenario involving all classes and another scenario involving two classes. To achieve this, the experiments were conducted and divided into two main sessions: 1) the experiment capable of

classifying all classes (Healthy, Boundary, Infected scenarios) and 2) the experiment capable of classifying two classes (Healthy, Infected scenarios).

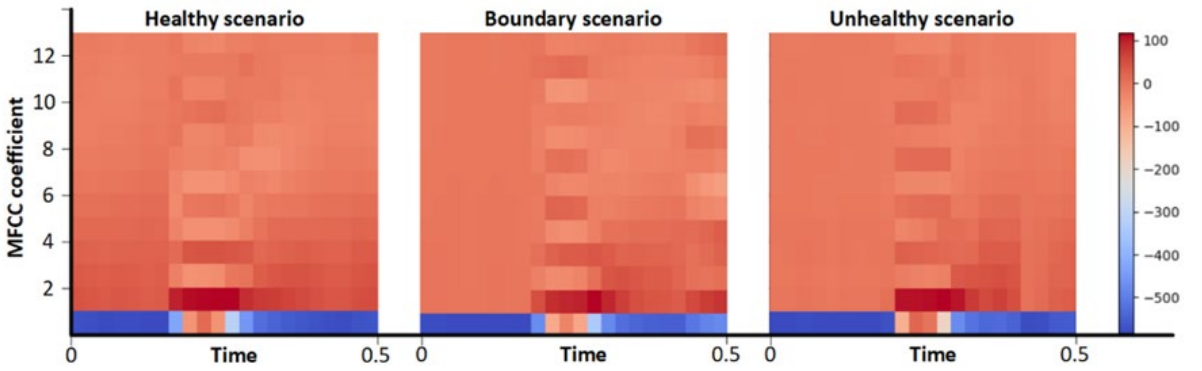

Fig. 6. Three Examples of transformed data to MFCCs from each class (Healthy, Infected, and Boundary)

In the context of machine-learning parameter selection, preliminary pilot experiments were initially conducted to identify the optimal combination of parameter values for each machine-learning model. The findings from these initial experiments revealed that the CNN model exhibits superior performance and is composed of multiple layers, including convolutional and max pooling layers. As shown in Figure 7, it starts with two convolutional layers, each consisting of 32 filters with a kernel size of 3x3. Subsequently, a max pooling layer with a pool size of 2x2 follows. The model then continues with two more convolutional layers, each containing 64 filters with a 3x3 kernel size. Afterwards, the model has a dense layer with 128 neurons using the ReLU activation function, and another dense layer with 3 neurons using the softmax activation function. The model utilizes the Adam optimizer and employs categorical_crossentropy as the loss function.

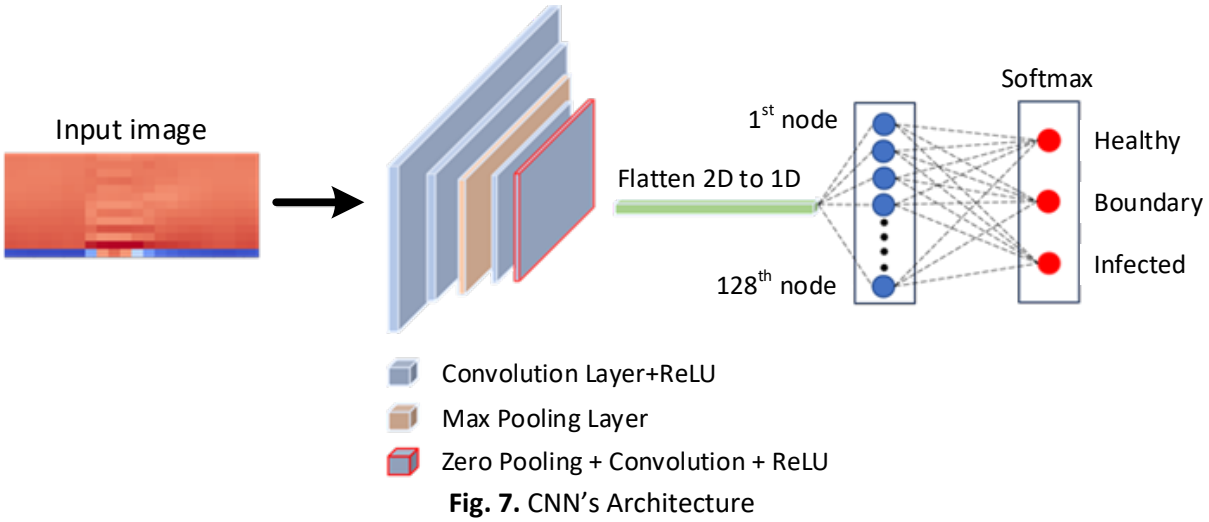

Fig. 7. CNN's Architecture

Figure 8 displaying the MLP's Architecture that consists of two hidden layers. Each layer comprises 256 neurons with the ReLU activation function. Softmax is the activation function for the output layer. The model's alpha value is set to 0.01, and the initial learning rate is 0.001. The model incorporates a warm start. The BaysiensearchCV utilized Adam as the solver, which has categorical_crossentropy to be the loss function. The SVC model employs the Radial basis function (RBF) as the kernel with a regularization parameter (C) set to 10. Scaling gamma was chosen in the

model. The stopping criterion for the model is determined by a tolerance value of 1e-3. The decision function follows the one-vs-rest shape for multiclass classification.

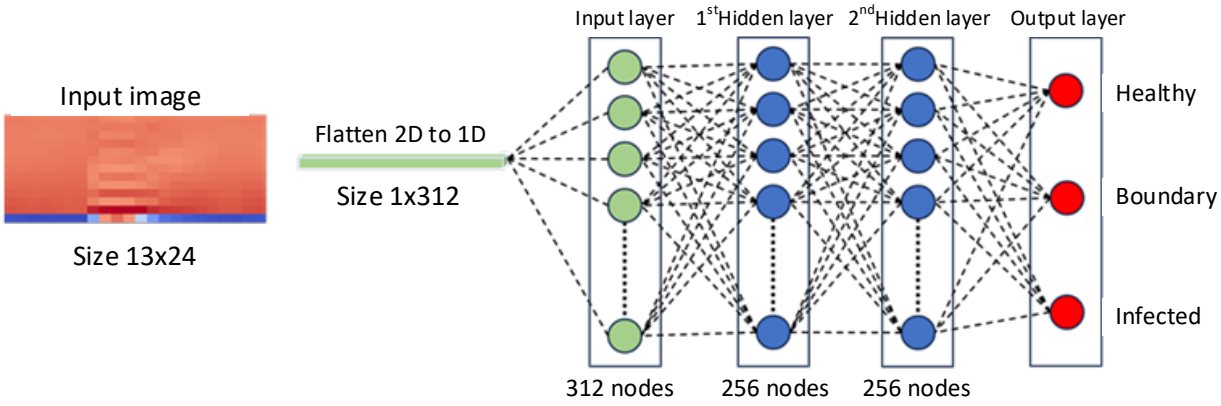

Fig. 8. MLP's Architecture

After completing the parameter selection for CNN, MLP, and SVC models, we proceeded to utilize the optimized machine learning models to evaluate their effectiveness in the task of knocking sound classification in oil palm trees. The objective was to identify the most suitable approach for accurate classification. The evaluation was conducted based on metrics such as accuracy, precision, recall, and F1 score in all the given scenarios. These results provided insights into the performance of each classifier, enabling us to make informed conclusions about their efficacy for the specific classification task as presented in Table 1.

Table 1
Accuracy, Precision, Recall and F1-score for every model

Model evaluation	2 classes (Boundary excluded)			3 classes		
	CNN	MLP	SVC	CNN	MLP	SVC
Accuracy	90.73%	83.38%	84.21%	84.97%	72.25%	69.64%
Precision	0.908	0.847	0.843	0.851	0.760	0.699
Recall	0.907	0.837	0.842	0.848	0.750	0.692
F1 score	0.908	0.842	0.843	0.849	0.755	0.696

According to the test results, in the two-class (boundary excluded) scenario, the CNN model achieved the highest accuracy of 90.73% (its confusion matrix is shown in Figure 9(a)), followed by the MLP model with 83.38% and the SVC model with 84.21%. The CNN model also obtained the highest precision, recall, and F1 score, indicating its overall better performance in distinguishing between the two classes. In the three-class scenario, the CNN model again obtained the highest accuracy of 84.97% (its confusion matrix is shown in Figure 9(b)), while the MLP and SVC models achieved lower accuracies of 72.25% and 69.64%, respectively. However, the precision, recall, and F1 score for the three-class scenario were highest for the CNN model, suggesting its superiority in classifying all three classes compared to the other models.

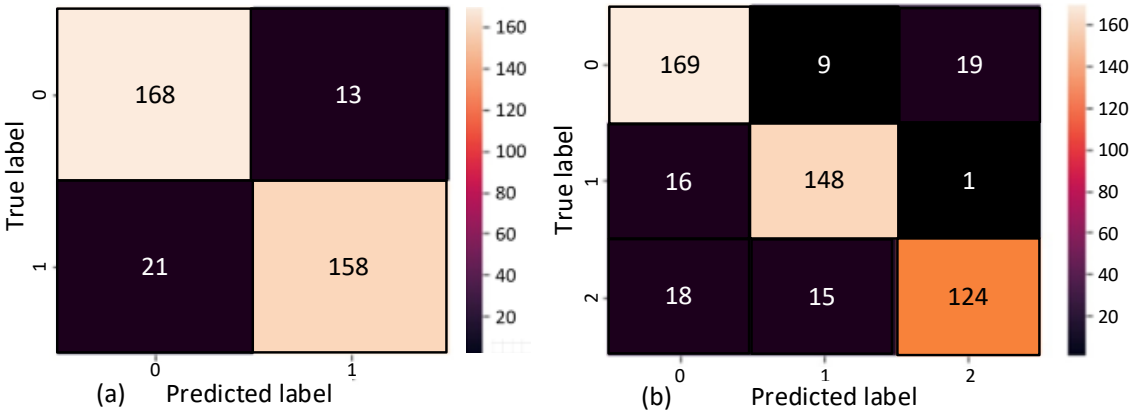

Fig. 9. (a) CNN confusion matrix from 2-classes scenario (b) CNN confusion matrix from 3-classes scenario

In summary, the table provides a comprehensive overview of the model's performance in both the two-class and three-class scenarios, utilizing various evaluation metrics. These metrics offer valuable insights into the models' accuracy, precision, recall, and F1 score for the respective classification tasks. Notably, the CNN model consistently exhibited superior performance in both scenarios, surpassing the MLP and SVC models in terms of accuracy, precision, recall, and F1 score. However, it is essential to note that the accuracy was higher in the two-class scenario compared to the three-class scenario. The subsequent section will discuss the reasons behind the observed decrease in accuracy for the three-class scenario.

3.3 Discussion

Based on the performance comparison, it is evident that the primary reason for lower accuracy in the three-class scenario for all algorithms is the presence of the boundary class. The sound characteristics of the boundary class are similar to both the healthy and infected classes, causing confusion during the classification process. The MFCC spectrogram of the boundary class becomes entangled with the healthy and infected classes, leading to misclassification and adversely affecting the algorithms' performance. Despite the visual reparability, the boundary class presents challenges for accurate classification, impacting the performance of the algorithms significantly.

In Figure 10, the mean power spectrum of each class is displayed, providing insights into the frequency characteristics of the knocking sound signals. The power spectrum density is represented by the magnitude of the peaks at different frequencies. Three classes are considered: infected, healthy, and boundary. The infected dataset exhibits the highest peak in power spectrum density, with a magnitude of approximately 7,000 at a frequency of around 750 Hz. On the other hand, the healthy dataset displays another prominent peak with a magnitude level of around 6,000 at a frequency of approximately 1250 Hz. The boundary dataset, representing the area between healthy and infected classes, shows an interesting characteristic. Its power spectrum overlaps between the infected and healthy datasets, with energy concentrations at frequencies close to both 1000 Hz and 1200 Hz. This overlapping feature makes the boundary dataset more challenging to classify accurately, as its sound characteristics share similarities with both healthy and infected classes.

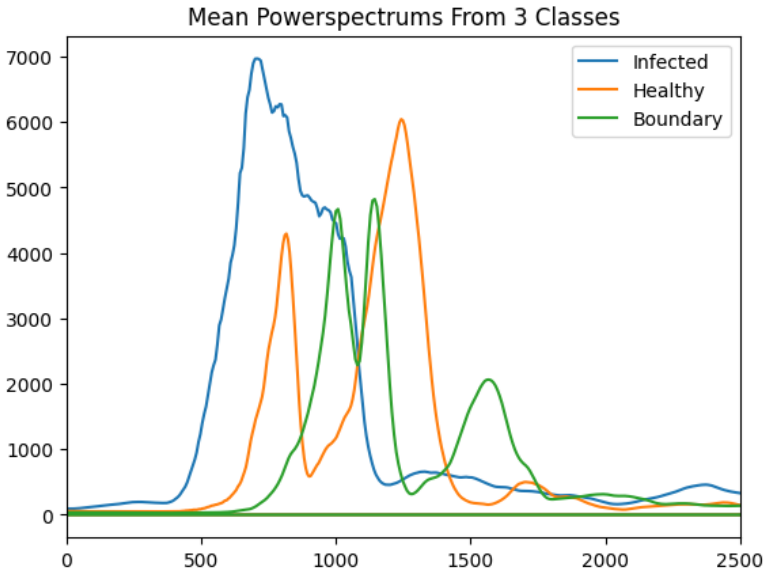

Fig. 10 The mean power-spectrum plot

Figure 11 further corroborates our findings, highlighting the influential role of the boundary class in the observed inaccuracies. The box plot of the mean frequency power spectrum vividly illustrates the distinct characteristics of each class. The healthy class demonstrates a prominent cluster lying between 1200 – 1450 Hz, while the infected class exhibits another distinct cluster spanning approximately 700-1100 Hz, positioned below the healthy class. However, the presence of the boundary class introduces a perplexing scenario, with its distribution chaotically overlapping both the healthy and infected classes, spanning approximately 900 – 1300 Hz. This intricate overlap significantly contributes to the classification errors, as the boundary class exhibits sound and signal properties akin to both healthy and infected data. The visual evidence presented in Figure 11 reinforces our assertion that the accurate differentiation of the boundary class from the other two classes poses considerable challenges for classification algorithms, thereby impacting overall accuracy and performance.

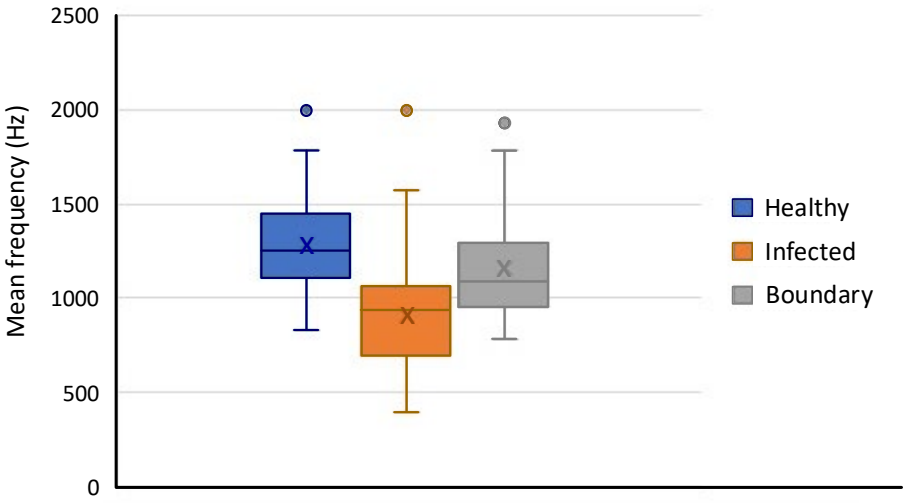

Fig. 11. The box plot of the mean frequency power spectrum for all classes (Healthy, Boundary, Infected scenarios)

In order to improve the performance of classification or decrease the misclassification made by the boundary data. Expanding the dataset by acquiring more samples for the boundary class can

significantly improve classification performance. A larger and more diverse dataset would enable the algorithm to better capture the distinctive features of the boundary class and reduce the entanglement with healthy and infected classes. Therefore, collecting additional boundary class data from various environments and conditions will help the model generalize better and improve its ability to differentiate between classes.

To confirm this, the inclusion of 311 additional sound files collected from Trang province significantly enhanced the model's performance, as indicated by the results showcased in Figure 12, illustrating the post-improvement confusion matrix of the model. These new files are distributed among the three classes, with 101 in the healthy class, 110 in the boundary class, and 100 in the infected class. The presence of the confusion matrix in Figure 12 underscores the advantages of this expanded dataset. It clearly demonstrates the improved capability of the model to accurately classify sounds into their respective categories. With a more balanced distribution of data across healthy, boundary, and infected classes, the model displays reduced bias and heightened precision in distinguishing among these categories.

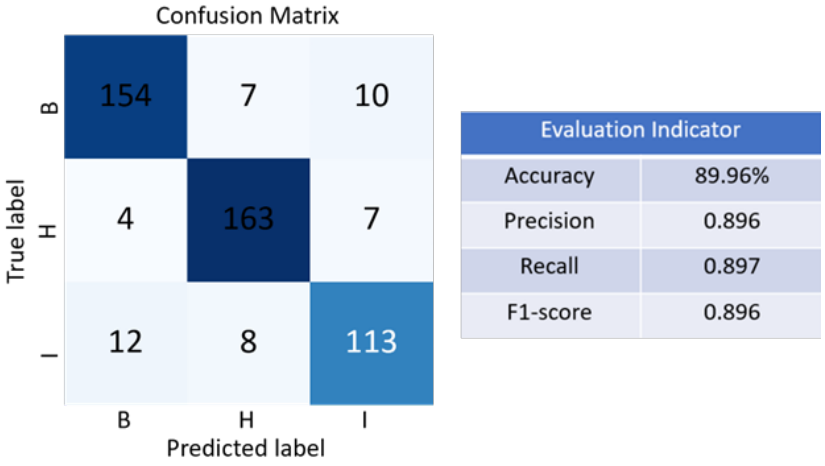

Fig. 12. The performance of the model with the additional dataset from Trang province

This diverse exposure to various sound samples, including the unique features from the Trang province, enhances the model's capacity to handle real-world variations. Consequently, the model showcases increased accuracy and resilience, establishing itself as a more dependable tool for practical applications like detecting diseases in oil palm trees. The inclusion of additional data not only enriches the training environment but also substantially elevates the model's overall predictive performance.

In the context of the algorithm selection literature [32-36] and in consideration of the preceding discussion, it is notable that a majority of the works examined make use of pre-trained models. These models necessitate substantial computational resources for the construction of an effective classification system. To address the imperative of cost minimization in the domain of BSR diagnosis, we outline the next phase of our research efforts. In this upcoming stage, we aim to seamlessly integrate our developed system into an affordable hardware platform, such as the Raspberry Pi. In alignment with this objective, our commitment lies in developing models that do not rely on pre-existing counterparts, thereby aligning our approach with the parameters of resource frugality and self-sufficiency.

4. Conclusion

In this research, we developed a sound classification system aimed at leveraging local wisdom to diagnose BSR disease in oil palm trees through a non-destructive method: knocking on the tree trunk with a rubber mallet. The system leverages local wisdom and considers two output class setups: one with three classes (healthy, boundary, and infected), and the other with two classes (healthy and infected), excluding the boundary class to investigate the potential improvement in model accuracy by removing an ambiguous class. Among the tested models, the CNN model proved to be the most effective, achieving an accuracy of 90.73% for the two-class model and 84.98% for the three-class model. The precision, recall, and F1-score for both models also demonstrated excellent performance.

In conclusion, our research demonstrates the potential of using local wisdom for BSR disease diagnosis in oil palm trees. While the CNN model showed promising results, continued efforts in data collection and model refinement will further improve the system's accuracy and prevent overfitting situations to be more applicable in real-world settings. However, it is important to acknowledge the limitations of this study. The model's generalizability may be constrained due to the relative dataset of 2,075 sound files from 60 oil palm trees in three different southern provinces of Thailand. To enhance the system's performance and applicability, future research will focus on expanding the dataset with a more diverse and extensive collection of sound files, encompassing a broader range of oil palm trees and environments. This will enable the model to better generalize and adapt to various scenarios in real-world settings.

5. Acknowledgement

This research was supported by Natural Science, Research and Innovation Fund (NSRF) and Prince of Songkla University (Grant No ENG6701223S).

References

[1] Asian Agri, 2019. The Benefits of Palm Oil. (accessed 02.20.23)

[2] Malinee, Rachane, Dimitris Stratoulias, and Narissara Nuthammachot. "Detection of oil palm disease in plantations in krabi province, thailand with high spatial resolution satellite imagery." *Agriculture* 11, no. 3 (2021): 251. https://doi.org/10.3390/agriculture11030251

[3] Fee, Chung Gait. "Management of Ganoderma diseases in oil palm plantations." *The Planter* 87, no. 1022 (2011): 325-339.

[4] Khairunniza-Bejo, Siti, and Chin Nee Vong. "Detection of basal stem rot (BSR) infected oil palm tree using laser scanning data." *Agriculture and Agricultural Science Procedia* 2 (2014): 156-164. https://doi.org/10.1016/j.aaspro.2014.11.023

[5] Amiri, S. S., Ahmad, Z. A. A., Ismail, S. I., 2021. "A review on the molecular detection of Ganoderma boninense from oil palm." *Journal of Oil Palm Research* 33, no. 1 (2021): 1-14.

[6] Elliott, Monica L., Timothy K. Broschat, and Lothar Göcke. "Preliminary evaluation of electrical resistance tomography for imaging palm trunks." *Arboriculture & Urban Forestry (AUF)* 42, no. 2 (2016): 111-119. https://doi.org/10.48044/jauf.2016.010

[7] Hamidon, Nur Aqilah, and Muhammad Mukhlisin. "A review of application of computed tomography on early detection of basal stem rot disease." *Jurnal Teknologi* 70, no. 3 (2014). https://doi.org/10.11113/jt.v70.3461

[8] Janpen, P., 2022. Co-operation with Palm Thai Pattana Garden [Personal interview, 30 November, 2] Satun 2022 (Unpublished).

[9] Kharamat, Weangchai, Manop Wongsaisuwan, and Norrarat Wattanamongkhol. "Durian ripeness classification from the knocking sounds using convolutional neural network." In *2020 8th International Electrical Engineering Congress (iEECON)*, pp. 1-4. IEEE, 2020. https://doi.org/10.1109/iEECON48109.2020.229571

[10] Najmie, M., K. Khalid, A. Sidek, and M. Jusoh. "Density and ultrasonic characterization of oil palm trunk infected by Ganoderma boninense disease." *Measurement Science Review* 11, no. 5 (2011): 160-164. https://doi.org/10.2478/v10048-011-0026-x

[11] Abu-zanona, Marwan, Said Elaiwat, S. Younis, N. Innab, and M. M. Kamruzzaman. "Classification of palm trees diseases using convolution neural network." *International Journal of Advanced Computer Science and Applications* 13, no. 6 (2022): 10-14569. https://doi.org/10.14569/IJACSA.2022.01306111
[12] Tan, Mas Ira Syafila Mohd Hilmi, Mohd Faizal Jamlos, Ahmad Fairuz Omar, Kamarulzaman Kamarudin, and Mohd Aminudin Jamlos. "Ganoderma boninense classification based on near-infrared spectral data using machine learning techniques." *Chemometrics and Intelligent Laboratory Systems* 232 (2023): 104718. https://doi.org/10.1016/j.chemolab.2022.104718
[13] Koh, T. H., Melton, L. D., 1994. Current techniques in the analysis of fruit tissues. Food Science Department, University of Otago, Dunedin, New Zealand.
[14] Lee, Chee Cheong, Voon Chet Koo, Tien Sze Lim, Yang Ping Lee, and Haryati Abidin. "A multi-layer perceptron-based approach for early detection of BSR disease in oil palm trees using hyperspectral images." *Heliyon* 8, no. 4 (2022). https://doi.org/10.1016/j.heliyon.2022.e09252
[15] Raja, B. M., Ravindran, V., Nair, V. J., Pillai, A. R., 2021. Evaluation of Noise Reduction Techniques in Audacity. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 10(6), 1756-1763.
[16] Deng, Li, and Dong Yu. "Deep learning: methods and applications." *Foundations and trends® in signal processing* 7, no. 3–4 (2014): 197-387. https://doi.org/10.1561/2000000039
[17] Lin, Yi-Lin, and Gang Wei. "Speech emotion recognition based on HMM and SVM." In *2005 international conference on machine learning and cybernetics*, vol. 8, pp. 4898-4901. IEEE, 2005. https://doi.org/10.1109/ICMLC.2005.1527805
[18] Weninger, Felix, Hakan Erdogan, Shinji Watanabe, Emmanuel Vincent, Jonathan Le Roux, John R. Hershey, and Björn Schuller. "Speech enhancement with LSTM recurrent neural networks and its application to noise-robust ASR." In *Latent Variable Analysis and Signal Separation: 12th International Conference, LVA/ICA 2015, Liberec, Czech Republic, August 25-28, 2015, Proceedings 12*, pp. 91-99. Springer International Publishing, 2015. https://doi.org/10.1007/978-3-319-22482-4_11
[19] LeCun, Yann, Bernhard Boser, John S. Denker, Donnie Henderson, Richard E. Howard, Wayne Hubbard, and Lawrence D. Jackel. "Backpropagation applied to handwritten zip code recognition." *Neural computation* 1, no. 4 (1989): 541-551. https://doi.org/10.1162/neco.1989.1.4.541
[20] Goodfellow, Ian. *Deep Learning*. MIT Press, 2016.
[21] Long, Jonathan, Evan Shelhamer, and Trevor Darrell. "Fully convolutional networks for semantic segmentation." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3431-3440. 2015. https://doi.org/10.1109/CVPR.2015.7298965
[22] Piczak, Karol J. "Environmental sound classification with convolutional neural networks." In *2015 IEEE 25th international workshop on machine learning for signal processing (MLSP)*, pp. 1-6. IEEE, 2015. https://doi.org/10.1109/MLSP.2015.7324337
[23] Huzaifah, Muhammad. "Comparison of time-frequency representations for environmental sound classification using convolutional neural networks." *arXiv preprint arXiv:1706.07156* (2017).
[24] Kahl, Stefan, Thomas Wilhelm-Stein, Hussein Hussein, Holger Klinck, Danny Kowerko, Marc Ritter, and Maximilian Eibl. "Large-Scale Bird Sound Classification using Convolutional Neural Networks." *CLEF (working notes)* 1866 (2017).
[25] Malik, Hassaan, Umair Bashir, and Adnan Ahmad. "Multi-classification neural network model for detection of abnormal heartbeat audio signals." *Biomedical Engineering Advances* 4 (2022): 100048. https://doi.org/10.1016/j.bea.2022.100048
[26] Bugatti, Alessandro, Alessandra Flammini, and Pierangelo Migliorati. "Audio classification in speech and music: a comparison between a statistical and a neural approach." *EURASIP Journal on Advances in Signal Processing* 2002 (2002): 1-7. https://doi.org/10.1155/S1110865702000720
[27] Valles, Damian, and Rezwan Matin. "An audio processing approach using ensemble learning for speech-emotion recognition for children with ASD." In *2021 IEEE World AI IoT Congress (AIIoT)*, pp. 0055-0061. IEEE, 2021. https://doi.org/10.1109/AIIoT52608.2021.9454174
[28] Grama, Lacrimioara, Lorena Muscar, and Corneliu Rusu. "Sound Classification Algorithms for Indoor Human Activities." In *2021 16th International Conference on Engineering of Modern Electric Systems (EMES)*, pp. 1-4. IEEE, 2021. https://doi.org/10.1109/EMES52337.2021.9484121
[29] Zhao, Dong, Huadong Ma, and Liang Liu. "Event classification for living environment surveillance using audio sensor networks." In *2010 IEEE International Conference on Multimedia and Expo*, pp. 528-533. IEEE, 2010. https://doi.org/10.1109/ICME.2010.5583889
[30] Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. "The elements of statistical learning: data mining, inference, and prediction." (2009).

[31] He, Kaiming, and Jian Sun. "Convolutional neural networks at constrained time cost." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 5353-5360. 2015. https://doi.org/10.1109/CVPR.2015.7299173

[32] Paymode, Ananda S., and Vandana B. Malode. "Transfer learning for multi-crop leaf disease image classification using convolutional neural network VGG." *Artificial Intelligence in Agriculture* 6 (2022): 23-33. https://doi.org/10.1016/j.aiia.2021.12.002

[33] Lachgar, Mohamed, Hamid Hrimech, and Ali Kartit. "Optimization techniques in deep convolutional neuronal networks applied to olive diseases classification." *Artificial Intelligence in Agriculture* 6 (2022): 77-89. https://doi.org/10.1016/j.aiia.2022.06.001

[34] Ahad, Md Taimur, Yan Li, Bo Song, and Touhid Bhuiyan. "Comparison of CNN-based deep learning architectures for rice diseases classification." *Artificial Intelligence in Agriculture* 9 (2023): 22-35. https://doi.org/10.1016/j.aiia.2023.07.001

[35] Jiang, Bo, Jinrong He, Shuqin Yang, Hongfei Fu, Tong Li, Huaibo Song, and Dongjian He. "Fusion of machine vision technology and AlexNet-CNNs deep learning network for the detection of postharvest apple pesticide residues." *Artificial Intelligence in Agriculture* 1 (2019): 1-8. https://doi.org/10.1016/j.aiia.2019.02.001

[36] Rizzo, Matteo, Matteo Marcuzzo, Alessandro Zangari, Andrea Gasparetto, and Andrea Albarelli. "Fruit ripeness classification: A survey." *Artificial Intelligence in Agriculture* 7 (2023): 44-57. https://doi.org/10.1016/j.aiia.2023.02.004