Sound Classification using Convolutional Neural Network: A Prototyping Tool for Ganoderma Oil Palm Disease Detection

CHALATORN AUGSORNTHIP[†1] PARAMIN NERANON[†1]
PORNCHAI PHUKPATTARANONT[†2]

Abstract: Oil palm trees are one of the most important crops in Southeast Asia. Palm oil is capable of creating products that are important for the economy and citizens' consumption. One of the most major problems of oil palm diseases is Ganoderma disease, which is a white-rot fungus causing oil palm trees to wilt and die. Conventionally, local wisdom is initially executed to classify the oil palm tree disease using the sound technique by knocking the trunk. This is because the Ganoderma disease can massively damage a stem of the oil palm tree. Consequently, this research is aiming to develop a prototype device to be able to initially diagnose the oil palm disease. The results show the optimized experimental setup for the sound recording process i.e. the knocking force acting to the tree trunk, the position of the microphone and the trunk position of the impact point. Finally, the Convolutional Neural Network (CNN) was successfully implemented for classifying the wood-knocking sounds based on the generated spectrograms of these sounds with an achieved accuracy of approximately 70-80%. It can be concluded that the proposed method based on the CNN approach in the sound classification and recognition systems can be used to initially diagnose the Ganoderma disease.

Keywords: Sound Classification; Convolutional Neural Network (CNN); Ganoderma Oil Palm Disease

1. Introduction

The palm industry is one of the most important agro-industries in Southeast Asia as it can boost the economy and raise the GDP of the tropical-climate countries, especially for Indonesia, Malaysia and southern Thailand. In Thailand, there is approximately 70 per cent of land is managed by farmers for oil palm cultivation. The global market demands palm oil production increasingly. This is because the oil palm trees can produce high-quality oil which is wildly used for healthy cooking to enhance the human's immune system (as it offers high vitamins and antioxidants). Alternatively, it can also be processed to produce biofuel, butter, ice cream, detergents, soaps, cosmetics, etc.

However, oil palm trees are normally defenseless to various fungal infections. One of the most crucial devastating diseases in the oil palm trees is called Ganoderma boninense [1], which is a major affecting the oil palm trees with their direct loss of the stand, reducing palm yields and being killed eventually. Technically, the effect of the disease is penetrating to palm tree from the root to the top. It makes the leaf to be pale, a leaflet is going to be rotten, and the stem is going to be rotten inside. This will slightly make a hollow in the palm trunk due to massive damage of a stem of the oil palm tree. Conventionally, to initially detect the Ganoderma oil palm disease, local wisdom is used to classify it using the sound technique, whereas an expert will carefully listen to the knocking sound of the tree trunk.

Thu, this research focuses on the development of a new prototype that can initially detect the Ganoderma disease of the oil palm trees by using Convolutional Neural Network (CNN) for sound classification. As the symptom Ganoderma significantly causes the stem tissue to be destroyed and leads to a hollow stem trunk, this work-in-progress paper presents how to design the optimized test setup for the sound recording and classification processes. These are made up of the preliminary experiments on i.e. how to apply a constant knocking force on the tree trunk; how to set up a microphone position on the test rig; how to find a suitable hitting point on the trunk; and finally how to roughly estimate an optimized CNN algorithm.

2. Process of Knocking Sound Classification

The sound knocking process (in Figure 1) is simply affected surrounding noises such as strong wind, birds, bugs, and so on. Consequently, it necessitates the elimination of the inferences to enhance the classification performance, and a digital filter was utilized. We also applied segmentation to divide a first-touch knocking sound of each test (it is to be noted that a set of knocking sounds is generated by using a pendulum hitting the palm tree trunk). In fact, the redundancy sound is not allowed pass-through. Then, we have to extract the features used to train the CNN model.

The CNN is capable to create effective classifiers, particularly for image classification due to their feature extraction and classification parts [2]. A visual representation of each knocking sound sample is then needed to identify the captures for classification, and this technique is the same method employed to classify images with high accuracy. Applying Mel-frequency cepstral coefficients (MFCCs) provides spectrograms delivered for visualizing the short-term power spectrum of the hitting audio. The MFCC uses a quasi-logarithmic spaced frequency scale, which is closed to how the human auditory system always processes audio sounds [3].

†1 Department of Mechanical and Mechatronics Engineering,
†2 Department Electrical Engineering , Faculty of Engineering
Prince of Songkla University, Hat Yai, Songkhla, 90112, Thailand

*Corresponding author, Email address: paramin.n@psu.ac.th

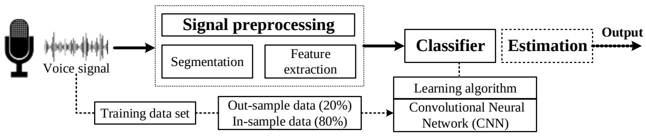

Figure 1 The diagrom of wood-knocking sound classification

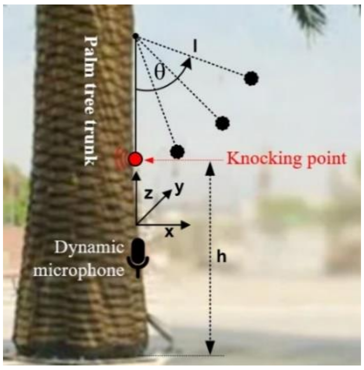

Figure 2 The experimental setup and notifications

Initially, a sequential model was used because of its ease in order to easily construct the CNN model layer by layer. Convolution 2 dimensions were applied to deal with various images of the short-term frequency spectrum in 2-dimensional matrices. The CNN model was roughly proposed as 1 hidden layer and 128 nodes. Kernel size or the size of the filter matrix for our convolution is assumed as a 3x3 filter matrix. An activation function for the layer is the Rectified Linear Activation (ReLU). A Flatten layer is served as a connection between the convolution and dense (our output layer) layers.

3. Experiments and Results

3.1 Experimental setup

A set of preliminary experiments involving the knocking sound classification task has been developed. It is to be noted that the CNN as proposed in the previous section is implemented throughout the tests. The schematic of the test apparatus designed for this study is illustrated in Figure 2. Constant knocking sounds can be repeatably generated using a pendulum technique to hit the palm tree trunk at the knocking point. The key objectives of the experiments are to strategically design the optimized test setup for the audio classification: (1) to find a suitable length (l) of the pendulum rope and an angle (θ) of the release point of the pendulum, which provide different potential energy of the knocking force, (2) to estimate an appropriate sound recording position (microphone location), and (3) to determine a suitable hitting point on the oil palm trunk (h) referenced to ground as depicted in the figure.

In the test procedure, the experimental demonstrator was required to perform constantly by releasing the pendulum under the following condition as mentioned, whereas the rope is strictly taut while the pendulum swings all the time. The set of hitting sounds was simultaneously recorded and evaluated, in which the results have been addressed in the following section.

3.2 Experimental results

This section explains the preliminary test outcomes which were used to evaluate the performance of the sound classification based on the CNN algorithm. Table 1 illustrates the results of Test 1: rope lengths, Test 2: the angle of the release point, Test 3: microphone position, and Test 4: the hitting point on the oil palm trunk, respectively.

Table 1 Experimental results

Test 1		Test 2		Test 3				Test 4	
l (cm)	Acc (%)	θ (°)	Acc (%)	y (cm)	Acc (%)	z (cm)	Acc (%)	h (cm)	Acc (%)
15	68.0	20	64.0	5	72.0	5	76.0	20	72.0
30	64.0	45	68.0	10	64.0	10	68.0	40	76.0
60	72.0	85	76.0	15	64.0	15	64.0	60	84.0

According to Test 1, it can be seen that the highest accuracy of the tests is about 72% at the rope length of 60 cm. This means the potential energy affects the accuracy of the knocking sound classification process; however, the longer rope could bring a harder setting up process. Similarly, to the first test, Test 2 shows that the optimum angle of dropping position of the pendulum is around 85° which give the highest accuracy of 76%. As shown in Test 3, the optimized position of the microphone setup at y = 5 cm and z = 5 cm with an accuracy of approximately 74%. Moreover, the results of Test 4 demonstrate the suitable hitting point on the trunk of h = 60 cm, which yields the highest accuracy of 84%.

This set of preliminary results has become useful in the future study of the optimization of the CNN model. The generalized CNN model should be designed carefully to achieve high accuracy with small architecture and also avoid limitations of overfitting and underfitting in machine learning.

4. Concluding Remarks

Convolutional Neural Network (CNN) is capable to create effective classifiers, particularly precise in sound classification, which can be transformed to the short-term power spectrum of the hitting audio for visualizing. This work in progress paper expresses how to optimize the experimental setup of the knocking sound on the oil palm tree trunk. Based on several preliminary tests (Tests 1-4), the performance of the wood-knocking sound classification using CNN contributes an achieved accuracy of approximately 70-80%. Consequently, this knowledge will be accepted and undertaken to develop a new prototype device to be able to initially diagnose the Ganoderma disease. In the future, we will optimize the CNN model which will offer the high performance of sound classification.

Reference

[1] R. Malinee, D. Stratoulias, N. Nuthammachot, Detection of Oil Palm Disease in Plantations in Krabi Province, Thailand with High Spatial Resolution Satellite Imagery. *Agriculture*. 2021; 11(3):251.

[2] D. Utebayeva, M. Alduraibi, L. Ilipbayeva and Y. Temirgaliyev, Stacked BiLSTM - CNN for Multiple label UAV sound classification, *2020 Fourth IEEE International Conference on Robotic Computing (IRC)*, 2020, pp. 470-474,

[3] S. Jayalakshmy, B. L. Priya and N. Kavya, "CNN based Categorization of respiratory sounds using spectral descriptors," *2020 International Conference on Communication, Computing and Industry 4.0 (C2I4)*, 2020, pp. 1-5, doi: 10.1109/C2I451079.2020.9368933.