MEKATRONIKA
e-ISSN: 2637-0883
VOL. 2, ISSUE 1, 28 – 37
DOI: https://doi.org/10.15282/mekatronika.v2i1.6722

ORIGINAL ARTICLE

Intelligent Classification of Palm Oil Tree Pollination Using E-Nose

Abdul Rahim Jalil[1], Muhammad Sharfi Najib[1,*], Suhaimi Mohd Daud[1], Mujahid Mohamad[1], Saiful Nizam Tajuddin[2], Che Mohd Aizal Che Mohd[2], Razam Abd Latip[3], Mohammad Saiful Nizam Ismail[3], Mohamad Nor Azizi Shabudin[3], Wan Rusyidah W. Rusik[3]

[1]Faculty of Manufacturing and Mechatronic Engineering Technology, Universiti Malaysia Pahang, 26600 Pahang, Malaysia.
[2]Pusat Bioaromatik, Universiti Malaysia Pahang, 26300 Pahang, Malaysia.
[3]Sime Darby Plantation Research Sdn Bhd, 42960 Selangor, Malaysia.

ABSTRACT – The pollination period is one of the crucial steps needed to ensure crop yield increases, especially in palm oil palm plantations. Most of the research has difficulty determining the pollination period of palm oil. Many problems contribute to this problem, such as difficut to reach and depedency of the polination insect as the insect activity is influenced by the surrounding enviroment.E-Nose can help determine the period by classifiy odour pattern of the male and female palm oil flower. The pattern of each of the flowers were classified using cased – based reasoning artificial intelligent technique. This paper shows the research of the palm oil pollination flower odour profile pattern using case-based reasoning (CBR) classifier.

ARTICLE HISTORY
Received: 22nd April 2020
Revised: 27th May 2020
Accepted: 6th June 2020

KEYWORDS
Pollination Palm Oil
E-Nose
Male and Female Flower
Odour Pattern
Cased Based Reasoning

INTRODUCTION

Agriculture in palm oil is still one of the vital sectors in Malaysia due to the increase in palm oil products such as cooking oil and other product. There are approximately 2450 species throughout the entire world, and most can be found in tropical regions such as Africa, with about 65 species [1]. In the past century, the pollination of the palm oil tree is mainly done by the biotic and abiotic pollination agents. Biotic pollination is done by wind and rain whereas, Abiotic pollination is insects such as bees and bats. To increase the pollination rate of palm oil, some of the palm oil plantations use the assist pollination. The effectiveness of pollination is highly dependent on the pollination period of the flower. If the introduction of the pollen is too early or too late, it will result in unfertilised flowers, hence resulting in a poor pollinated bunch [2].

Both the male and female flowers of the palm oil emit unique and strong smells [3]. Some expert researchers in palm oil state that the smell is almost similar to the anise smell. The strong and pungent smell is essential as it is finction as attractor to other pollination insects [4]. As for palm oil, the weevil or Elaeidobius kamerunicus Faust (Coleoptera: Curculionidae) is the most attractive to the unique smell of the palm oil flower [5]. The odour that releases by the flower produce by a multitude of different chemical compounds [6]. For palm oil tree flowers, the pollen smell release by a phenyl propane compound called methyl chavicol. This compound is also known as estragole [7], and it is more concentrated toward lignin formation. Phenylpropanes, a C6 – C3 carbon skeleton, also serve as pollinators and aid in pathogen defence since they derive from methyl eugenol and isoeugenol. When flowers are ready to be pollinated, they emit an increased amount of volatile compounds. After pollination, flowers reduce the synthesis of volatiles to prevent further visitors to non-pollinated flowers[8].

There are lots of classification methods that are widely being used. In this research, the classification using case-based reasoning (CBR) is introduced as the CBR technique is proven as one of the excellent classifications based on the research by [9],[10]. CBR is a technique that utilised a simple solving problem method that uses the past cases library. These cases will be comparing their similarity value to solve the current cases. It is a widespread method used by humans to determine the solution for a new problem. As the CBR only calculate the similarity value between the store case and the current case, this makes the algorithm very simple as the CBR do not need any testing and training data to determine the output of the cases.

This paper shows the palm oil flower odour pattern using e-nose and performs the classification of both of the male and female flower pollen smell by using a cased based reasoning classifier. The CBR classifier will undergo performance measurement for each of the samples to determine the classifier performance.

METHODOLOGY

Experiment Flow Chart

Figure 1 shows the whole process of analysing and interpreting the data from the sample collection until the performance measurement. The palm oil of male and female flowers was used in the experiment with eight samples. This

*CORRESPONDING AUTHOR |Muhammad Sharfi Najib | ✉ sharfi@ump.edu.my

sample needs to be kept in an airtight container, and the experiment needs to be conducted as fast as possible as the flower odour will decrease over time. The e-nose setup was conducted in a veltilated chamber to increase detection of the sample odour, as in Figure 2.

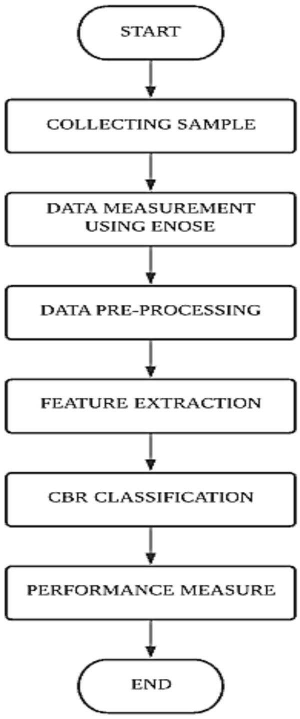

Figure 1. Methodology Flowchart

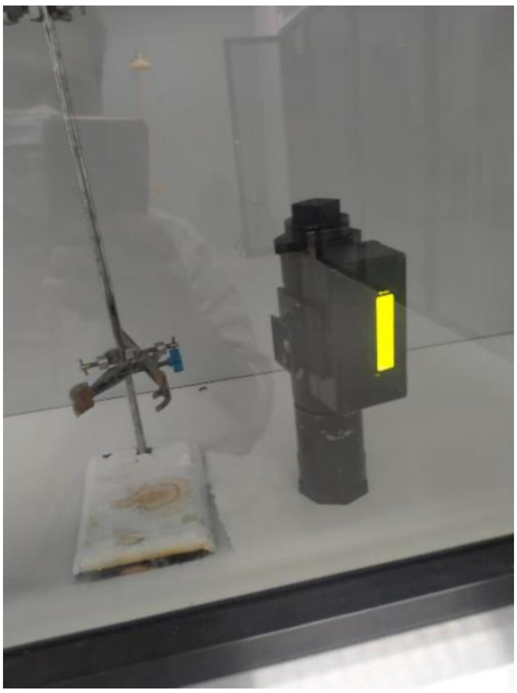

Figure 2. Experimental setup

Each of the data samples needs to undergo all this processing to achieve the CBR intelligent classification. The data measurement includes neutralising the sensor of the e-nose as this is a crucial step to get more accurate data as the sensor's neutralisation will help minimise the contamination of the odour. In data pre-processing, the raw data need to normalise first before proceeding to another step. The normalised data will have a value between 0 and 1 through the mean calculation technique. Data pre-processing stage will enhance the reliability of the data measured before proceed to feature extraction and CBR classification, henceforth would provide the performance measurement in terms of accuracy and consistency.

Data Measurement

Table 1 shows the raw experiment data set from each sensor (i.e. S1 is the first sensor). Each sample will contain 200 rows of data, and each of the samples needs to be repeated five times. At the end of the experiment, every sample will have five sets of experiment raw data set.

Table 1. Raw datasets by the sensors.

Data Measurement No.	S_1	S_2	S_3	S_4
DM_1	$RD_{1,1}$	$RD_{1,2}$	$RD_{1,3}$	$RD_{1,4}$
DM_2	$RD_{2,1}$	$RD_{2,2}$	$RD_{2,3}$	$RD_{2,4}$
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
DM_{200}	$RD_{200,1}$	$RD_{200,2}$	$RD_{200,3}$	$RD_{200,4}$

Data Pre-Processing

The experiment raw data set was normalised using Equation (1) to normalise all the value between 0 and 1. The normalisation technique reduces the data fluctuation and restores the measurement data based on the concentration levels to make the features invariant. It is essential to normalise the data from the raw data before further analysing because unnormalised data can affect the classifier performance. Each raw data will be normalised by dividing with the highest value in every row of data where R is the raw data and R_{max} is the maximum raw data.

$$R' = \frac{R}{Rmax} \tag{1}$$

Table 2 show the normalising data from the raw data obtain based on the experiment conducted. Every experiment raw data set will undergo normalising data denoted as *ND*.

Table 2. Normalised data.

Normalized Data Measurement No.	S_1	S_2	S_3	S_4
nDM_1	$ND_{1,1}$	$ND_{1,2}$	$ND_{1,3}$	$ND_{1,4}$
nDM_2	$ND_{2,1}$	$ND_{2,2}$	$ND_{2,3}$	$ND_{2,4}$
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
nDM_{200}	$ND_{200,1}$	$ND_{200,2}$	$ND_{200,3}$	$ND_{200,4}$

Feature Extraction

The feature or the odour pattern for each sample for palm oil pollination flower can be extracted using the normalised data. Before the odour pattern can be illustrated using a graphical method, each data set from each sample (1000 row of data) need to be reduced to 200 rows of data. These can be performed using the mean calculation technique. The data will be further process by clustering the data into 10 clusters. In the end, every sample will have 10 clusters (data case) of data. Hence, it will produce 80 data (B1, B2, B3, B4, J1, J2, J3 and J4). From these cases, the odour pattern for each of the samples can be illustrated using graphical method.

Cased Based Reasoning

CBR classification used is used in this research to classify the odour pattern for each sample that has been used. CBR mainly uses past cases or store cases that function as libraries that are further used to classify the expected sample in the future. The CBR classification uses one of the cases as current cases, and the other cases will remain as store cases. In this research, one out of the 80 will be set as the current case, and the other 79 store cases will calculate its similarity percentage compare to the current cases. The similarity formula that is used in this research is given in Equation 2 where fiI is the feature i input cases, fiR as the feature i retrieved cases, sim is the similarity function and wi as the important weight of feature i.

$$sim = \frac{\sum_{i=1}^{n} w_i \, x \, sim(f_i^I, f_i^R)}{\sum_{i=1}^{n} w_i} \tag{2}$$

Peformance Measure

The CBR classifier of the palm oil pollination will be further analysed or evaluated using a confusion matrix as the confusion matrix able to measure the classifier system's sensitivity, specificity, and accuracy based on the palm oil pollination flower sample. To determine the prediction output and actual output, each sample's three highest values will be extracted based on the similarity comparison between the actual and the predicted cases, which will help increase the CBR classifier's performance measurement.

EXPERIMENTAL RESULTS

Figure 3 shows the normalised data of odour feature pattern for each sample B1, B2, B3, B4. The x-axis is the sensor number utilised in the electronic nose, and Y-axis representing the normalised value for each sample. The normalised data for each sample will be further used as store cases for cased based reasoning classification. Figure 4 shows the normalised data of odour feature pattern for each sample J1, J2, J3 and J4.

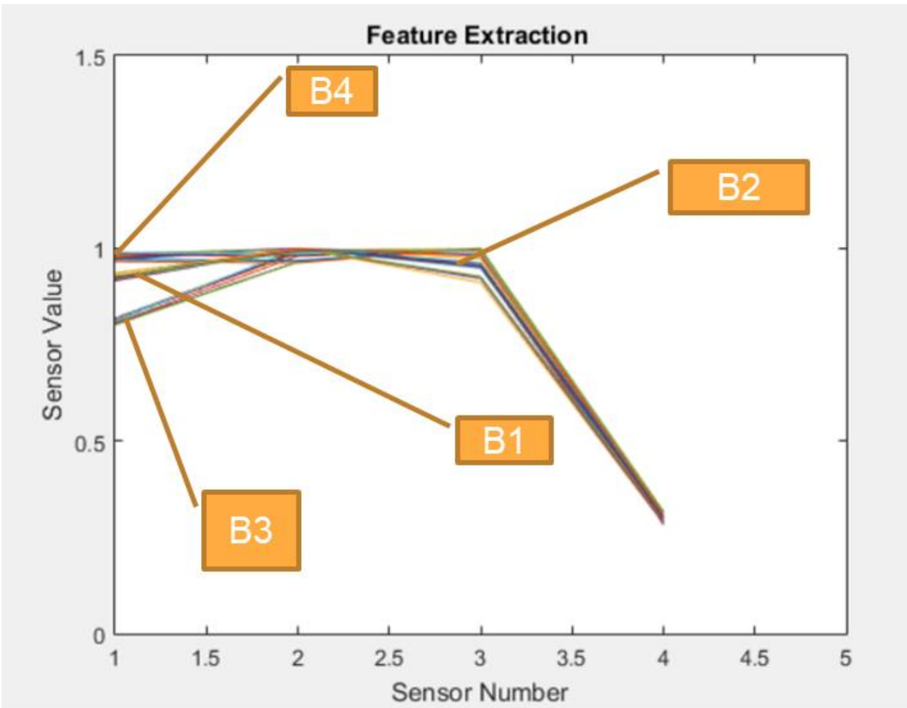

Figure 3. Odour Pattern Sample B

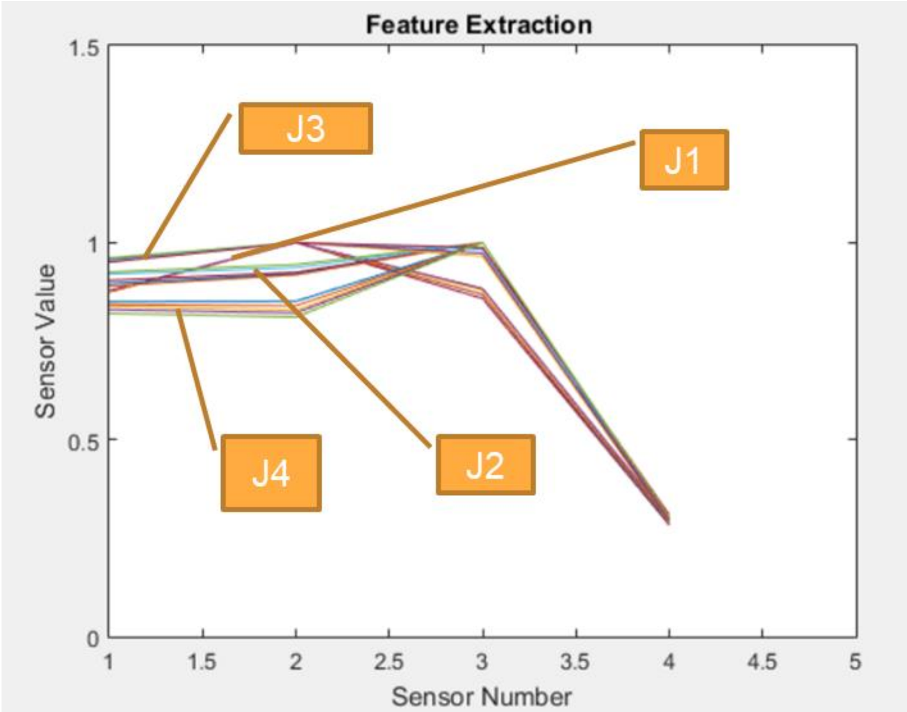

Figure 4. Odour Pattern Sample J

Table 3 shows the CBR case library for flower samples: B1, B2, B3, B4, J1, J2, J3 and J4. Each sample has 10 cases, in which Table 4 contain 80 cases for all the sample then the case library is stored into CBR memory as stored cases.

Table 3. CBR Cases.

Sample	Cases
B1	K1 until K10
B2	K11 until K20
B3	K21 until K30
B4	K31 until K40
J1	K41 until K50
J2	K51 until K60
J3	K61 until K70
J4	K71 until K80

Table 4. CBR Cases.

	S1	S2	S3	S4
K1	0.8177	0.9962	0.9976	0.2868
K2	0.8187	0.9901	0.9990	0.2819
K3	0.8171	0.9880	0.9985	0.2878
K4	0.8141	0.9917	0.9965	0.2915
K5	0.8122	0.9913	0.9978	0.2915
K6	0.8054	0.9804	1.0000	0.2991
K7	0.8059	0.9805	1.0000	0.2957
K8	0.8026	0.9648	1.0000	0.2969
K9	0.8042	0.9656	1.0000	0.3038
K10	0.8011	0.9644	1.0000	0.3049

K11	0.8105	0.9997	0.9864	0.3089
K12	0.8081	1.0000	0.9838	0.3106
K13	0.8059	0.9980	0.9884	0.3166
K14	0.8123	0.9971	0.9947	0.3191
K15	0.7991	0.9871	0.9994	0.3194
K16	0.7854	0.9820	1.0000	0.3235
K17	0.7931	0.9767	1.0000	0.3194
K18	0.7887	0.9600	1.0000	0.3194
K19	0.7848	0.9512	1.0000	0.3183
K20	0.7910	0.9424	1.0000	0.3170
K21	0.9756	1.0000	0.9499	0.2846
K22	0.9710	1.0000	0.9546	0.2891
K23	0.9786	1.0000	0.9601	0.2944
K24	0.9803	0.9999	0.9725	0.2971
K25	0.9846	0.9994	0.9820	0.3035
K26	0.9873	0.9957	0.9939	0.3042
K27	0.9889	0.9846	0.9981	0.3082
K28	0.9841	0.9799	0.9988	0.3035
K29	0.9713	0.9682	1.0000	0.3061
K30	0.9658	0.9624	1.0000	0.3043
K31	0.9193	1.0000	0.9112	0.2853
K32	0.9159	1.0000	0.9218	0.2883
K33	0.9208	1.0000	0.9270	0.2956
K34	0.9264	1.0000	0.9478	0.2987
K35	0.9239	1.0000	0.9529	0.2989
K36	0.9268	1.0000	0.9611	0.3052
K37	0.9347	1.0000	0.9800	0.3086
K38	0.9362	0.9961	0.9928	0.3147
K39	0.9274	0.9945	0.9963	0.3156
K40	0.9282	0.9897	0.9989	0.3172
K41	0.8775	1.0000	0.8581	0.2834
K42	0.8777	1.0000	0.8677	0.2863
K43	0.8739	1.0000	0.8687	0.2908
K44	0.8766	1.0000	0.8789	0.2969
K45	0.8783	1.0000	0.8832	0.2971
K46	0.8825	1.0000	0.9014	0.3053
K47	0.8852	1.0000	0.9148	0.3134
K48	0.8985	1.0000	0.9285	0.3180
K49	0.8941	1.0000	0.9352	0.3233
K50	0.9044	1.0000	0.9456	0.3291
K51	0.9253	0.9441	1.0000	0.2831
K52	0.9200	0.9362	1.0000	0.2847
K53	0.9042	0.9243	1.0000	0.2834
K54	0.8959	0.9204	1.0000	0.2878
K55	0.8895	0.9180	1.0000	0.2915
K56	0.8813	0.9051	1.0000	0.2931
K57	0.8727	0.8987	1.0000	0.2970
K58	0.8710	0.8924	1.0000	0.2972

K59	0.8640	0.8786	1.0000	0.3013
K60	0.8559	0.8606	1.0000	0.3013
K61	0.9588	1.0000	0.9654	0.2900
K62	0.9591	1.0000	0.9728	0.2932
K63	0.9608	1.0000	0.9854	0.3005
K64	0.9538	0.9998	0.9823	0.3030
K65	0.9504	0.9994	0.9872	0.3088
K66	0.9548	0.9988	0.9960	0.3157
K67	0.9520	0.9944	0.9988	0.3206
K68	0.9418	0.9900	0.9998	0.3204
K69	0.9342	0.9820	0.9999	0.3279
K70	0.9224	0.9801	0.9996	0.3295
K71	0.8500	0.8510	1.0000	0.3025
K72	0.8428	0.8395	1.0000	0.3023
K73	0.8387	0.8280	1.0000	0.3028
K74	0.8300	0.8213	1.0000	0.2972
K75	0.8202	0.8109	1.0000	0.3015
K76	0.8153	0.7957	1.0000	0.3020
K77	0.8132	0.7921	1.0000	0.3025
K78	0.8080	0.7793	1.0000	0.3005
K79	0.8037	0.7715	1.0000	0.2976
K80	0.8034	0.7680	1.0000	0.3031

Table 5 shows the CBR calculation for similarity percentages for each target case. The similarity is the comparison between two target cases. As shown in the table, two cases are picked, and one of the cases is the store case, and the other is the current case. By using similarity calculation from Equation 2, the weighted similarity for each of the sensors is obtained. Finally, all the weighted similarity is added for every row to determine the similarity value between the two cases.

Table 5. CBR similarity calculation.

ATTRIBUTION	LOCAL WEIGHT	NORMALIZED WEIGHT	CURRENT CASE	STORED CASE	SIMILARITY FUNCTION	NORMALIZED SIMLIRAITY FUNCTION	WEIGHTED SIMILARITY
S1	1	0.25	0.817709	0.818696	0.999014	1	0.25
S2	1	0.25	0.996206	0.990072	0.993865	0.994847	0.248712
S3	1	0.25	0.997638	0.998972	0.998666	0.999652	0.249913
S4	1	0.25	0.286837	0.281883	0.995046	0.996029	0.249007
	4				MAX		SUM
					0.999014		0.997632
					SIMILARITY		99.76318

Table 6 show the confusion matrix table. The true positive, true negative, false positive and false negative can be extracted from the table. This extracted value from the table will be used to measure the classifier performance. This table is extracted from the voting table between the actual output and the predicted output from the CBR similarity. Three highest values from the similarity prediction result were selected from each of the samples. Therefore, it will further validate the performance of the classifier as more data is being analysed.

Table 6. Confusion matrix results.

		PREDICTED							
		B1	B2	B3	B4	J1	J2	J3	J4
ACTUAL RESULT	B1	30	0	0	0	0	0	0	0
	B2	0	30	0	0	0	0	0	0
	B3	0	0	28	0	0	0	2	0
	B4	0	0	0	27	1	0	5	0
	J1	0	0	0	0	29	0	0	0
	J2	0	0	0	0	0	27	0	2
	J3	0	0	2	3	0	1	23	0
	J4	0	0	0	0	0	2	0	28
TOTAL		30	30	30	30	30	30	30	30

Table 7 show the example of the CBR voting table similarity for every case and extract the data into the confusion matrix table.

Table 7. CBR confusion matrix for Sample B1 and B2.

		Predict		
Actual		*1ST*	*2ND*	*3RD*
SAMPLE B1	*K1*	*B1*	*B1*	*B1*
	K2	*B1*	*B1*	*B1*
	...	...	...	...
SAMPLE B2	K11	B2	B2	B2
	K12	B2	B2	B2

Table 8 show the classier performance for all of the sample. The sensitivity, specificity and accuracy of each of the sample is shown in the table.

Table 8. CBR Cases.

B1	SENSITIVITY	100.00
	SPECIFICITY	100.00
	ACCURACY	100.00
B2	SENSITIVITY	100.00
	SPECIFICITY	100.00
	ACCURACY	100.00
B3	SENSITIVITY	93.33
	SPECIFICITY	99.05
	ACCURACY	98.33
B4	SENSITIVITY	81.82
	SPECIFICITY	98.55
	ACCURACY	96.25
J1	SENSITIVITY	100.00

	SPECIFICITY	99.53
	ACCURACY	99.58
J2	SENSITIVITY	93.10
	SPECIFICITY	98.58
	ACCURACY	97.92
J3	SENSITIVITY	79.31
	SPECIFICITY	96.68
	ACCURACY	94.58
J4	SENSITIVITY	93.33
	SPECIFICITY	99.05
	ACCURACY	98.33

For most of the sample, the sensitivity is more than 90%, except for sample B4 with 81% and J3 with 79% sensitivity. As for specificity and accuracy, all samples have more than 90% accuracy and sensitivity.

Table 9. Overall performances.

OVERALL	SENSITIVITY	92.61
	SPECIFITY	98.93
	ACCURACY	98.13

The overall sensitivity, specificity and accuracy are excellent as this classifier can achieve more than 90%. The CBR performance measurement shows that the classification of the palm oil male and female flower pollination is highly performed. The CBR classification can be used to classify the flower's pollination with more than 90% success.

CONCLUSION

The different odour patterns of each of the samples B1, B2, B3, B4, J1, J2, J3 and J4 are shown as each palm oil pollen flower has different odours between males and females. The CBR classification technique used to classify the odour pattern is also proven as the overall sensitivity, specficty and accuracy for each sample is more than 90%.

ACKNOWLEDGEMENT

The authors would like to thank the Faculty of Manufacturing and Mechatronic Engineering Technology of Universiti Malaysia Pahang for the chances and support of this research.

REFERENCES

[1] M. B. Wahid, S. N. A. Abdullah, and I. E. Henson, "Oil palm - Achievements and potential," *Plant Prod. Sci.*, vol. 8, no. 3, pp. 288–297, 2005, doi: 10.1626/pps.8.288.

[2] M. R. M. Kassim and A. N. Harun, "Using wireless sensor networks to determine pollination readiness in palm oil plantation," *Smart Sensors, Meas. Instrum.*, vol. 23, pp. 137–155, 2017, doi: 10.1007/978-3-319-47322-2_7.

[3] Siswanto and D. Soetopo, “Population of oil palm pollinator insect (Elaeidobius kamerunicus faust.) at PTP Nusantara VIII Cisalak Baru, Rangkasbitung-Banten,” *IOP Conf. Ser. Earth Environ. Sci.*, vol. 418, no. 1, 2020, doi: 10.1088/1755-1315/418/1/012045.

[4] A. Vainstein, L. Lewinsohn, E. Pichersky, and D. Weiss, “Floral fragrance. New inroads into an old commodity," *Plant Physiol.*, vol. 127, no. 4, pp. 1383–1389, 2001, doi: 10.1104/pp.010706.

[5] P. Rekha, P. Prabhavathi, V. S. Murthy, and P. Jain, "A Survey of Floral Aroma Sensors," no. July, pp. 1155–1162, 2018.

[6] E. Bloem, S. Haneklaus, R. Daniels, and E. Schnug, “Influence of sulfur fertilization on floral scent patterns of crops in full bloom,” *Landbauforsch. Volkenrode*, vol. 60, no. 1, pp. 45–50, 2010.

[7] R. Li *et al.*, "Comparative transcriptome analysis of oil palm flowers reveals an EAR-motif-containing R2R3-MYB that modulates phenylpropene biosynthesis," *BMC Plant Biol.*, vol. 17, no. 1, pp. 1–17, 2017, doi: 10.1186/s12870-017-1174-4.

[8] M. Ramya, H. R. An, Y. S. Baek, K. E. Reddy, and P. H. Park, "Orchid floral volatiles: Biosynthesis genes and

transcriptional regulations," *Sci. Hortic. (Amsterdam).*, vol. 235, no. May, pp. 62–69, 2018, doi: 10.1016/j.scienta.2017.12.049.

[9] M. S. Najib, M. U. Ahmad, P. Funk, M. N. Taib, and N. A. M. Ali, "Agarwood classification: A case-based reasoning approach based on E-nose," *Proc. - 2012 IEEE 8th Int. Colloq. Signal Process. Its Appl. CSPA 2012*, pp. 120–126, 2012, doi: 10.1109/CSPA.2012.6194703.

[10] M. S. Najib, M. N. Taib, N. A. M. Ali, M. N. M. Arip, and A. M. Jalil, "Classification of agarwood grades using ANN," *InECCE 2011 - Int. Conf. Electr. Control Comput. Eng.*, pp. 367–372, 2011, doi: 10.1109/INECCE.2011.5953908.