Classification of Agarwood using ANN

M. S. Najib, N. A. Mohd Ali, M. N. Mat Arip, M. Abd Jalil and M. N. Taib, *Member, IEEE*

***Abstract*—An artifical neural network (ANN) has been modeled for the classification of Agarwood region. The target regions were from Melaka, Pagoh, Super Pagoh, Ulu Tembeling and Indonesia. The data analysis using Principal Component Analysis (PCA) was done to find significant input selection from 32 sensors of the E-nose and to recognize pattern variations from different number of Agarwood samples as inputs to ANN training. The network developed based on three layers feed forward network and the back propagation learning algorithm was used in executing the network training. Five input neurons, two hidden layer sizes and one output neurons were found to be the optimized combination for the network. The experimental results reveal that the proposed method is effective and significant to the classification of Agarwood region.**

Keywords-component; Agarwood, Classificastion, ANN

I. INTRODUCTION

ESSENTIAL oils produced previously and currently containing volatile aroma compounds from numerous plants. One of the local plants found in Malaysia is Agarwood plant. In [1], Agarwood is a fragrant wood that is usually derived from the diseased timber of the genus *Aquilaria* (Thymelaeceae). Hiroaki *et. al* [2] reported that a good quality Agarwood is widely used as incense while the low quality ones are used for essential oil extraction [2-3]. Agarwood is traded worldwide in significant volume and its quality very much depending on the wood resin content, aroma and origin mainly from Agarwood producing countries such as Malaysia, Indonesia and India [3-4]. Several medication benefits of Gahar u has been found recently. Two of the benefits are a pharmacological effect on the central nervous system achieved by oral administration or abdominal injection previously and human health [5-6].There are two ways generally employed for plants' detection or measurement of their associated properties [7-26]. There are either by examining the plants' chemical constituents or by direct evaluation of the plants' physical characteristics. Human beings can identify volatile component based on smell, touch, vision and test. It also gives many advantages to a robot application to operate in a human world. A humanoid robot could use its sense of smell to help identify food items and to assess their freshness [27].

One of the methods to sense and record odor based on chemical detection is through Gas chromatography (GC) [27]. The derivatives GC instrument is gas chromatography/mass spectrometry (GC/MS) is commonly to separate and analyze liquid samples [23] and gas mixture [23-24]. One of the applications of GC/MS is in detecting drug odor. However, GC/MS was limited as an analytical tool for this application for two reasons. First, since the thresholds for odors are often very low (parts per billion) (Dubois, 1976), there was not enough chemical mass in a given canister to identify the offending

Manuscript received November 28, 2011. This work was partially supported by the the Ministry of Science and Technology Malaysia under eSciencefund Grant, code: 05-03-10-SF0070).

Muhammad Sharfi is with Control and Instrumentation Group, Faculty of Electrical & Electronics Engineering, UMP, Pahang, Malaysia (e-mail: sharfi@ ump.edu.my).

Dr Nor Azah Md Ali is with Herbal Product Development Programme and senior scientist at Natural Products Division, Forest Research Institute Malaysia (FRIM), Kepong, Malaysia (e-mail: norazah@frim.gov.my).

Abd Majid Jalil is with Natural Producs Division, Forest Research Institute Malaysia (FRIM), Kepong, Malaysia (e-mail: majid@frim.gov.my).

Mohd Nasir Mat Arip is with Forest Products Division, Forest Research Institute Malaysia (FRIM), Kepong, Malaysia (e-mail: mnasir@frim.gov.my).

Mohd Nasir Taib is with Center for Engineering System Studies, Faculty of Electrical Engineering, UiTM, Shah Alam, Malaysia (e-mail: dr.nasir@ieee.org).

agents. Second, lot-to-lot variations in the product precluded precise identification of the off-odor producing chemical components [7]. Similar way as man extracts overall feature from an odor, instead of analyzing the chemical composition of an odor as GC/MS does [10]. This type of sensing requires two major processes for identification, extraction process and separation of essential oil components using GC/MS.

Alternatives to chemical-based detection, the physical of plant can be identified using a non-destructive experiment. This type of identification can sense the physical of the wood directly without physical damage. One of the techniques can be done using image recognition [28-29]. However, image recognition has drawbacks during complex weather and lighting condition whereby few techniques could do the work. Another technique that does not involve chemical composition identification is an electronic nose (e-nose). E-nose is able to detect odor in numerous application such as in medication [7, 13], quality of portable water [8], classification of reagents, liquor and perfumes [30], food quality monitoring [9], virgin olive oils [10], fruits [11], hazardous [12], plant malaise detection [14], organic compound [15], adulteration of Virgin Coconut Oil [17] and basal stem rot detection [18]. It consists of an array of sensors [7-8]. E-nose is used as an alternative to instruments, such as liquid chromatography or optical microscopy [8]. The range of e-nose applications include sewage [31], grading identification of tea [32-33], biodiesel [34], bacteria research [35], chemicals [19, 36-37], flavor discrimination [38], fragrance [39-40], water quality [8, 41-42] and also spoiled beef classification [43]. In many applications, several methods were employed in E-nose as inputs to classification, i.e., Principal Component Analysis (PCA) [18], Discriminant Factor Analysis (DFA) [11] and also ANN [22, 40]. Interestingly, it is reported that in [40], the results of discriminating different brands of cigarettes odors employing ANN has shown better results than direct discrimination process from the E-nose. There are various specifications of e-nose available in the market for wide range of applications. However, as the results of e-nose application employed with e-nose is better [40] than direct discrimination process, a vital step that must be taken into consideration is the data analysis and pre-processing or feature extraction techniques before employing the collected sample data into classifier as significant inputs. There are various feature extraction technique were identified that can be utilized such as Cluster Analysis (CA), Probabilistic Analysis (PA) [33], standard statistical tool such as mean and standard deviation [44-45], and PCA [31, 34]. The PCA value were extracted as the input [31, 33-34, 46] for BPNN classification analysis in classification of sewage odors, grading identification of green tea, biodiesel, tea storage times for leaf sample, beverage sample and residue using e-nose. There are many e-nose standard models available in market.

PCA was employed for feature selection in [18, 47-49] with slightly different results. In [47, 50], there were only two particle components considered while in [18, 47-49], three particle components are significant. No result can be drawn as mentioned in [47, 49]. However in [48], the least correlated sensors were selected among the highly correlated sensors as the best representation. Another technique was implemented by Markom *et. al* [18] by taking the high coefficient value as the significant sensors. Instead of analyzing with several particle components, one principle component is sufficient enough to be examined by looking the total variance explained [51]. The high coefficient values were selected and directly employed to ANN which produced 100% classification rate as proven by Markom *et. al* [18].

II. BACKGROUND THEORY

A. Principal Component Analysis

Principal component analysis (PCA) can be defined as a process of orthogonal transformation that involves a mathematical procedure transforming a vector of possibly correlated variables into a number of uncorrelated variables from the original dimensional space to a new space spanned by orthogonal principal axes. This transformation will result several number of uncorrelated principal components where the first principal components has as a high variance as possible that significantly contributing to the overall data variations.

Let X denote an N x P data matrix, where N is the number of data samples, which can be regarded as N realizations of a P-dimensional random vector X, which has been normalized to

zero-mean and unit variance. PCA is a linear transformation, from R^P to an M-dimensional vector space, where $M \leq P$. The optimal linear mapping in the least mean square sense is the one formed by the eigenvectors of the correlation matrix S_X of X, where $S_X = (1/(N-1))X^TX$. Let Z denote the N x M transformed data matrix. The PCA transforms X to Z by the following equation:

$$Z = XV_M \quad (1)$$

where, V_M, the P x M weight matrix, consists of M eigenvectors corresponding to the first M largest eigenvalues of the correlation matrix S_X, or V_M corresponds to the first M column vectors of matrix V, while V is obtained through singular value decomposition (SVD) of a scaled matrix $T=1/((N-1)^{1/2})X$, i.e., $T=UDV^T$. Here, both U and V are unitary matrices, and D is a P x P diagonal matrix with nonnegative diagonal elements d_i in decreasing order. Note that the correlation matrix of Z is a diagonal, which is given by $diag = \{d_1^2, d_2^2, \ldots\}$; i.e, columns of Z are mutually uncorrelated. Let x denote a row of X (one of the N samples, or one NAV), $z_j=xv_j$ is referred to as the j-th principal component coefficient or score (or the j-th PC score). The column vectors v_j of V is called the weight score of z_j or the j-th feature vector. The minimum mean-square error due to dimension reduction is $\sum_{i=M+1}^{P} d_i^2$. The best sensors selected from PCA were then applied into ANN [18] as input for classification process.

B. Artificicial Neural Network (ANN)

ANN plays an increasing role in various applications such as in plantation [18, 32], fragrance industry [52], and medication [13]. The ANN training requires very long task. It requires several repeated training sessions followed by an exhaustive process evaluation in order to select the training with the best result [53]. The ANN training can be categorized into two categories supervised and unsupervised learning. Supervised learning uses pattern classifier to relate specific class to specific sensor output while unsupervised learning does not require predetermined classes for training [36]. Artificial Neural Network (ANN) or Simulated Neural Network (SNN) is also called as Neural Network (NN) in general term. It is comprises of group of processing unit, which is modeled based on human neural network system. ANN is an adaptive system that can change its structure to classify a problem based on external or internal information flowing via neural network. ANN can also be categorized as a non-linear modeling instrument. ANN can be used to model complex input and output network so that there exist several connections between nodes of the data. Nerve is a type of tissue. The nerve tissues composed of two cells, neurons and glials. The function of neurons is to transmit messages to cell body. Whereas, glials is a cell having a direct contact surround the neurons. The idea of neural network development was established based on biological neurons. The biological neurons shows in Fig. 1 grant a model for computation. There are several major elements in biological neurons that are connected to each other which are dendrites, axons and synaptic terminals. Mathematical term of dendrites is equivalent to inputs, axon is an output or signals transmitter and a terminal that connects to dendrites of other neurons is called synapse. Signals to the dendrites (inputs) or in another term called receiver generates an electrical pulse conducted along the axon (outputs). The propagation of electrical pulse from input via outputs to synapses is a complete process of neural network. A feed forward NN is made up of one or more neurons layer interconnection between input layer, hidden layer and output layer as shown in Fig. 1. Depending on the network structure design, there are a number of nodes in each layer connected to other nodes in the other layer. The arrangement of the lower nodes is then connected to the other nodes in the other layer. The propagation of information flow in single direction from input to the neurons output layer in the network training. Inputs are received by the input layer, which performs as a data distribution centre and fans out the inputs to the first hidden layer. Each hidden layer will first activate and transform the data before propagating them to the next layer. This process is repeated through each hidden layer until finally all of the outputs from the last hidden layer will be reunited in the output layer to produce the network outputs. For the case of one hidden layer as shown in Fig. 1, the data,

which is activated and transformed by the hidden layer, will be forwarded to the output layer immediately. A hidden layer is so-named because the network can be regarded as a black box with inputs and outputs that can be seen but the hidden layers cannot be seen. This network is very powerful and has been shown in many cases the ability to learn any arbitrarily complex and non-linear input-output mapping. It also has a capacity much greater than the dimensionality of its input and output layers.

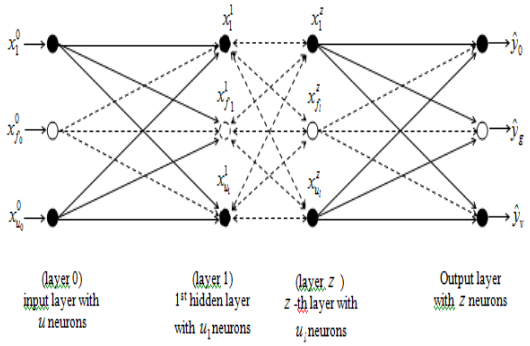

Fig. 1. Multilayer feed-forward network structure for general hidden layer.

Consider a single hidden layer feed-forward network in Fig. 2.

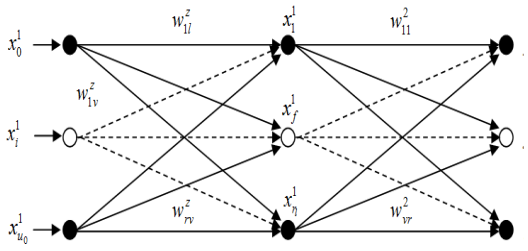

Fig. 2. Multilayer Multilayer feed-forward network structure for general hidden layer.

In this network there are u input neurons, r hidden neurons and v output neurons. The propagation of the inputs to the hidden layer can be described as follows:

$$\begin{bmatrix} x_1^1 \\ x_2^1 \\ \vdots \\ x_r^1 \end{bmatrix} = f\left\{ \begin{bmatrix} w_{11}^1 & w_{12}^1 & w_{13}^1 & \cdots & w_{1u}^1 \\ w_{21}^1 & w_{22}^1 & w_{23}^1 & \cdots & w_{2u}^1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ w_{r1}^1 & w_{r2}^1 & \cdots & \cdots & \cdots \end{bmatrix} \begin{bmatrix} x_1^0 \\ x_2^0 \\ \vdots \\ x_u^0 \end{bmatrix} \right\} \quad (2)$$

Where x_j^0, $j = 1,2,...,u$ is the number of input neurons, x_i^0, $i = 1,2,...,r$ is the number of hidden neurons, w_{ru}^1 is the weights between the input and the hidden layers, f is the activation function of the hidden neurons, and r and u are the size of the hidden and the input layers respectively. f is a non-linear transfer function makes an ANN highly flexible for adjusting weights and biases to different learning situations in order to capture the non-linear features underlying input-output data. The sigmoidal function, f , is represented as:

$$f(x) = \frac{1}{1+e^{-x}} \quad (3)$$

In many cases, biases are also included in the network in order to increase the efficiency during network training. By incorporating the bias units in the hidden layer, Equation 2 can be rewritten as:

$$\begin{bmatrix} x_1^1 \\ x_2^1 \\ \vdots \\ x_r^1 \end{bmatrix} = f\left\{ \begin{bmatrix} w_{11}^1 & w_{12}^1 & w_{13}^1 & \cdots & w_{1u}^1 \\ w_{21}^1 & w_{22}^1 & w_{23}^1 & \cdots & w_{2u}^1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ w_{r1}^1 & w_{r2}^1 & \cdots & \cdots & \cdots \end{bmatrix} \begin{bmatrix} x_1^0 \\ x_2^0 \\ \vdots \\ x_u^0 \end{bmatrix} \right\} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_r \end{bmatrix} \quad (4)$$

Where b_r are the biases.

This shows that the outputs of the network can be expressed as deterministic functions of the inputs, and the whole network represents a multivariate non-linear functional mapping. Simplifying Equation 3 gives:

$$x_i^1 = f\left\{ \sum_{j=1}^{u} w_{ij} x_j + b_i \right\} \quad (5)$$

where $i = 1,2,\cdots,r$ and $j = 1,2,\cdots,u$.

The output of the ith hidden neuron is obtained by first froming a weighted linear

combination of the n input values, followed by the addition of a bias and transforming the total with the activation function, f .

The output of the hidden neurons will be propagated to the neurons in the output layer. Traditionally, the characteristics of the neurons in the output layer are taken to be the same as that for the hidden layer, but this will limit the dynamic range of the output neurons to be less than 1.0. Since in practical applications the dynamic range of the output may be greater than 1.0, the activation function of the output neurons can be set to linear. Thus the outputs of the forward propagation of the hidden layer neurons to the output layer can be calculated as:

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_v \end{bmatrix} = f\left\{ \begin{bmatrix} w_{11}^2 & w_{12}^2 & w_{13}^2 & \cdots & w_{1u}^2 \\ w_{21}^2 & w_{22}^2 & w_{23}^2 & \cdots & w_{2u}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ w_{r1}^2 & w_{r2}^2 & \cdots & \cdots & w_{vr}^2 \end{bmatrix} \begin{bmatrix} x_1^1 \\ x_2^1 \\ \vdots \\ x_r^1 \end{bmatrix} \right\} \quad (1.6)$$

Where $\hat{y}_v$, $i = 1,2,...,v$ is the output neurons, w_{vr}^2 is the weight that connects the hidden neuron and output neurons and v is the output layer size. Given that the activation function is linear, e.g. set to 1.0, then the outputs of the network is similar to the output of the neurons in the output layer. Thus, using equation 10, the prediction of the i th neuron in the output layer is obtained by:

$$\hat{y}_i = \sum_{j=1}^{r} w_{ij} x_j \quad (7)$$

Where $i = 1,2,\cdots,v$ and $j = 1,2,\cdots,r$

Combination of equations 4 and 5, provide an explicit expression for the complete function represented by the network with a single hidden layer in Fig. 2 can be realized in the form:

$$\hat{y}_i = \sum_{j=1}^{r} w_{ij}^1 f\left\{ \sum_{k=1}^{u} w_{jk}^1 x_k + b_j \right\}, 1 \le i \le v \quad (8)$$

w_{ij}^1 , w_{jk}^1 and b_j the adaptive variables to be optimized and their values are changed many times during the network training process.

If we let $\Theta = [\theta_1 \ldots \theta_{n_\theta}]^T$ represent all of the unknown weights and biases of the network where n_θ denotes the dimension of the parameter vector Θ and is defined as $n_\theta = (n+1)r + r \times v$, then the network in Equation 6 can be represented as a function of Θ :

$$\hat{y}_i(\Theta) = \sum_{j=1}^{r} w_{ij}^1 f\left\{ \sum_{k=1}^{u} w_{jk}^1 x_k + b_j \right\}, \quad (9)$$

Where $1 \le i \le v$

ANN modeling can be used in many way of training method. Several training method in ANN were introduced including Radial Basis Function (RBF) [54], Counter Propagation Network (CPN) [55-58], Recursive Prediction Error (RPE) [59-61], Kohenen Network [62-63], Porbablistic Neural Network (PNN) [64-66]. Another training algorithm method that is famously used is Backpropagation Neural Network (BPNN). There several interesting research which employed BPNN in ANN training in several classification such as in mechanical failure [67], study on customer [68], sugarcane spot [69] and [70]. In this work, BPNN training algorithm is employed in the multilayered perceptron network (MLP) architecture. This method can be described by well-defined and explicit set of equations for weight corrections. These equations are applied throughout the layers, beginning with the correction of the weights between hidden and output layers and then moving backwards to correct the weights between the input and hidden layers.

The Backpropagation method was well described based on the assumption that all weights in a given network contribute to some portion of the output error, and that weight corrections should be proportional to the output error contributed by each weight. The weight updating or network learning is based on the definition of a suitable error function to be minimized with respect to the weights and biases of the network. Consider the sum-squared error function below:

$$J = \frac{1}{2}\sum_i \left(y_i - \hat{y}_i\right)^2 \tag{10}$$

Where y_i is the desired output and $\hat{y}_i$ is the calculated output produced by the ANN. The goal here is to find a procedure for evaluating the derivatives of the error function with respect to all adaptable parameters. In this case (the steepest descent approach), the parameter vector Θ is adjusted using the increment vector $\lfloor \Delta\theta_1 \quad \Delta\theta_1 \quad ... \quad \Delta\theta_{n_\theta} \rfloor$ defined along the negative gradient direction of *J*, *i.e.*

$$\Delta\theta_i = \eta \frac{\partial J}{\partial \theta_i} = \eta \nabla_i J \tag{11}$$

Although the one-hidden-layer-model is used in the present application, it is useful to derive the gradient of J for the general case and the result for the one-hidden-layer model can be readily be obtained as a special case. The application of the chain rule can be written as follows:

$$\frac{\partial J}{\partial w_{ij}^l} = \frac{\partial J}{\partial \hat{y}_i} \times \frac{\partial \hat{y}_i}{\partial w_{ij}^l}, \tag{12}$$

where $l = 0,1 ...$ is the layer's number

Differentiating Equation 7

$$\frac{\partial J}{\partial \hat{y}_i} = -\left(y_i - \hat{y}_i\right) = -\delta_i^l \tag{13}$$

where δ_i^l is called the error signal of the *i*-th neuron in the *l*-th layer. From Equation 7,

$$\frac{\partial \hat{y}_i}{\partial w_{ij}^l} = x_j^{l-1} \tag{14}$$

Hence,

$$\frac{\partial \hat{y}_i}{\partial w_{ij}^l} = -\delta_i^l x_j^{l-1} \tag{15}$$

Next consider the $(l-1)$th layer. Using the chain rule yields

$$\frac{\partial J}{\partial w_{ij}^{l-1}} = \sum_k \frac{\partial J}{\partial \hat{y}_k} \times \frac{\partial \hat{y}_k}{\partial x_{ij}^{l-1}} \times \frac{\delta_i^l x_j^{l-1}}{\partial z_i^{l-1}} \times \frac{\partial z_i^{l-1}}{\partial w_{ij}^l} \tag{16}$$

Where

$$\frac{\partial \hat{y}_i}{\partial w_{ij}^{l-1}} = -w_{ki}^l \tag{17}$$

$$\frac{\partial x_i^{l-1}}{\partial z_i^{l-1}} = \frac{\partial f(z)}{\partial z} = f'\left(z_i^{l-1}\right) \tag{18}$$

$$\frac{\partial z_i^{l-1}}{\partial w_{ij}^{l-1}} = -x_j^{l-2} \tag{19}$$

By defining the error signal for the i th neuron of the $(l-1)$th layer as

$$-\delta_i^{l-1} = f'\left(z_i^{l-1}\right)\sum_k \delta_k^l w_{ki}^l \tag{20}$$

Equation 18 can be rewritten as:

$$\frac{\partial J}{\partial w_{ij}^l} = -\delta_i^{l-1} x_j^{l-2} \tag{21}$$

Similarly it can be shown that

$$\frac{\partial J}{\partial b_i^{l-1}} = -\delta_i^{l-1} \tag{22}$$

By carrying this procedure, the following general results are obtained for $q = l, l-1, ..., 2$

$$-\delta_i^{q-1} = f'\left(z_i^{q-1}\right)\sum_k \delta_k^q w_{ki}^q \tag{23}$$

$$\frac{\partial J}{\partial w_{ij}^{q-1}} = -\delta_i^{q-1} x_j^{q-2} \tag{24}$$

$$\frac{\partial J}{\partial b_i^{q-1}} = -\delta_i^{q-1} \tag{25}$$

Equation 22 shows the error signal propagation backwards. The steepest descent minimization of the energy function defined in Equation 9 produces the following increments for upgrading Θ

$$\Delta w_{ij}^k = \eta_w \delta_i^k x_j^{k-1} \quad (26)$$

$$\Delta b_i^k = \eta_b \delta_i^k \quad (27)$$

Where

$$\delta_i^l = y_i - \hat{y}_i \quad (28)$$

$$-\delta_i^k = f'\left(z_i^k\right)\sum_k \delta_q^{k+1} w_{qi}^{k+1} \quad (29)$$

Where $k = l-1,..., 2,1$

The constant η_w and η_b represent the learning rates (typically a number between 0 to 1) for weights and biases respectively. In practice, a large value of the learning rate would be preferable because this would result in rapid learning. Unfortunately a large value of learning rate can also lead to oscillation or even divergence. To overcome this problem, a momentum term is usually included so that Equations 26 and 27 become:

$$\Delta w_{ij}^k(t) = \eta_w \delta_i^k x_j^{k-1} + \alpha_w \Delta w_{ij}^k(t-1) \quad (30)$$

$$\Delta b_i^k(t) = \eta_b \delta_i^k + \alpha_b \Delta b_i^k(t-1) \quad (31)$$

Where α_w and α_b are momentum constants which correspondingly resolve the effect of past changes of $\Delta w_{ij}^k(t-1)$ and $\Delta b_i^k(t-1)$ on the current updating direction in the weight $\Delta w_{ij}^k(t)$ and biases $\Delta b_i^k(t)$. In essence, during each training cycle, the weight changes in the network are subject to the weight changes of the previous pass. These changes are determined by the momentum constant (typically it ranges between 0 to 1) . Time variable, t, is explicitly shown in Equations 30 and 31 to differentiate between the previous values and current values.

In summary, the Backpropagation method corrects for weights and biases in the network according to

$$\Delta w_{ij}^k(t) = \Delta w_{ij}^k(t-1)\Delta w_{ij}^k(t) \quad (32)$$

$$\Delta b_i^k(t) = \Delta b_i^k(t-1) + \Delta b_i^k(t) \quad (33)$$

There are numerous application can employed with BPNN. One of the applications of BPNN is classifications in various fields [67, 71-72]. It can be initiated in two types, multiple classification and binary classification [73]. Binary classifications are defined in various ways such as one-aginst-all (OAA), one-against-one (OAO), direct acyclic graph SVM (DAGSVM), the hierarchical tree-based methods and error correcting output codes (ECOCS) methods [74]. Binary classification in neural network is more efficient as compared to multiple classification because weight output is compared by one node [73]. Implementations of binary classification found in neural network were introduced in several approaches. One of the approach introduced in [75] is a mapping of the input element vector to the network output consisting of two classes. Another approach found in [76] is a pairwise binary classification, where n(n-1)/2 classifiers are created and n being the number of several classes. This technique can be applied to the classification of multiple input and single output. In order to measure the performance, three major algorithm have been implemented in [77] using metric precision, sensibility and specificity. This three analysis tools will be very effective tools to compare classification results.

Therefore, this paper proposes a new technique, employing E-nose and artificial neural network (ANN), for detecting the origin of the Agarwood region.

III. METHODOLOGY

Fig. 1 explains the research design for ANN model design. The research comprises of two main experiments. The first experiment was an experimental setup and sample preparation process whereby the second one is the software part using ANN. Samples were collected from five identified regions (Super Pagoh, Ulu Tembeling, Pagoh, and Melaka) in the form of unprocessed resin. The resins from each region were first validated by authorized Forest

Research Institutes Malaysia (FRIM) members. Samples of one kilogram per group were grinded into fine chips which is then divided into ten sample vials. A hundred gram for each sample set. Hence, all the sample sets from all regions contributed to 50 sample sets. Each sample set were filled in the several closed vials. The closed vials were ensured so that there were no leakages found for preventive procedure. The closed vials that were filled with grinded sample set were preheated in special heater in the constant temperature value of $80\ ^{0}C$. All these steps were procedure were carried out at the controlled environment conditions at FRIM research lab laboratory. E-nose and sample vials were connected with special tubes and pipes for sample inlet, purge inlet and exhaust. Before measuring the Agarwood, the E-Nose nozzle is inserted into flushing tube coming from nitrogen gas source (Fig. 3) to flush the measurement chamber of any contamination, this is repeated 10 times.

For measurement, the tip of E-Nose nozzle is inserted through the septum of the closed vial (Fig. 3) and reading of the Agarwood odor will be monitored over 13 minutes. This procedure is repeated ten times and all readings are recorded by a PC in Microsoft Excel 2007 spreadsheet format. Note that the readings are in terms of the resistant values of the E-nose sensors. Once the odor measurements have been completed, the readings are taken for an off-line signal processing using two platform, Microsoft Excel, MATLAB and SPSS software. The Microsoft Excel had been utilized to acquire data sensed by the Cyronose and convert it into *.CSV file while MATLAB package was applied as a platform to normalize the data and imported to SPSS software. This procedure consists of the following steps. The first step is data preprocessing where each readings are normalized. By defining the regional normalized data

$$ND_R = \frac{\sum_{j=1}^{m}\sum_{i=1}^{n} UND}{\max(\sum_{j=1}^{m}\sum_{i=1}^{n} UND)}, \quad R = 1, 2, 3, \ldots S \quad (34)$$

Where, n is the number of row, m is the number of column, and R is the region number and S is the integer number of final region and UND is the un-normalized data.

After that, they are analyzed statistically using SPSS so that the most significant sensor readings could be picked up for ANN applications. These sensors were selected based on number of sensors reduction process to find the most significant sensors of the total sensor arrays. In this work, Principle Component Analysis (PCA) was applied as the number of sensors reduction technique which is commonly employed with e-nose data signal processing [18]. PCA is an effective linear method for discriminating between the e-nose measurement different ranges of odors complexity [49]. After sensor selection was done using PCA, the pattern variations based on the selected sensors were represented in graphical representation in two and three dimension view. These representations are to show the variations influences of the significant selected sensors. Then, selected sensors were used as inputs to ANN model. Then, the significant sensor readings were arranged such that the training and validation data will be interlaced with the ratio of 25:20 from 45 data sets. In MATLAB coding, the data sets were loaded from the excel CSV file and they were converted into MAT file.

The training data was arranged such that the datasets of Melaka, Pagoh, Ulu Tembeling, Super Pagoh and Indonesia are in ascending order with each region is a set of (5 rows x 10 column) matrix dimension. For testing data, the output target (1, 2, 3, 4 and 5) respectively was arranged in a set of (4 rows x 1 column) matrix dimension with repeated number of output in each region. Half of these training and testing data were divided into fifty percent each using interlacing technique. Then, neural network simulation using BPNN was done using feed-forward method. At this initial stage, the ANN design parameters (hidden layer, learning rate and momentum rate) were randomly selected. Then the parameters were optimized beginning with varying hidden layer by fixing the other two parameters (learning rate and momentum rate) with the mean of squared error (MSE) had been the ANN model performance. After the optimization of the first parameter (hidden layer), this procedure was followed with the optimized hidden layer and fixed momentum rate and fixed learning rate. The last one is the final ANN design parameter with optimized hidden layer and learning rate with varying momentum rate. After decided the final

ANN design parameters, the model were used to classify the testing inputs.

IV. RESULTS AND DISCUSSION

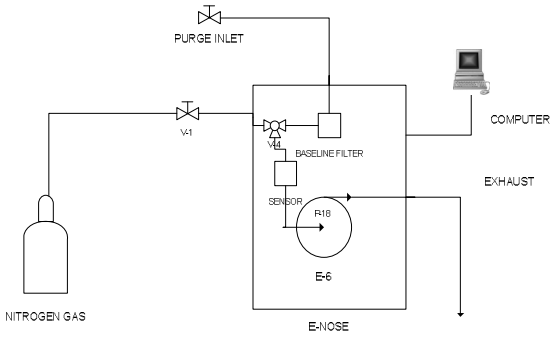

Fig. 3. Experimental Setup.

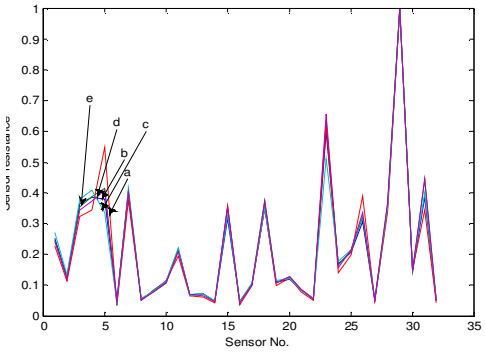

Fig. 4. Mean of normalised readings of the Agarwood odor for region (a) Melaka, (b) Pagoh, (c) Ulu Tembeling, (d) Super Pagoh, (e) Indonesia.

Fig. 4 shows a sample measurement of the E-nose 32 sensors array. It can be observed that each of the Agarwood exhibits almost similar pattern across the regions via direct observation. However, from the direct observation (Fig. 4), the readings of the Agarwood odor for regions (a, b, c, d and e respectively) are very difficult to be discriminated since it contributes greatly non-unique patterns from sensor 5 (S5) to sensor 32 (S32). The significant variance can only be observed at readings of sensor 3 (S3), sensor 4 (S4) and sensor 5 (S5) respectively. The results obtained in that manner because the normalization was done without taking into consideration of the significant sensor array reading. However, the direct observation without statistical technique is not a conclusive discussion. By applying the PCA analysis, the total variance explained shows only one principal component that can be examined and nine out of thirty two sensors were found significant based on their high PCA coefficient as tabulated Table 1. The selected sensors are S6, S8, S12, S13, S14, S16, S22, S27 and S32.

TABLE I
SENSOR SELECTIONS

PCA Results	
Sensors	Coefficient
S8	0.000443
S32	0.000495
S16	0.000534
S14	0.000625
S27	0.000643
S22	0.000870
S6	0.000920
S13	0.001307
S12	0.001646
S9	0.002274
S21	0.002275
S19	0.002509
S17	0.002807
S2	0.003111
S10	0.003611
S30	0.005087
S20	0.005631
S24	0.006153
S1	0.012875
S25	0.012898
S11	0.013777
S3	0.029807
S26	0.036920
S4	0.037714
S7	0.041967
S28	0.044028
S15	0.045744
S18	0.049807
S31	0.073679
S5	0.090309
S23	0.184309
S29	0.298826

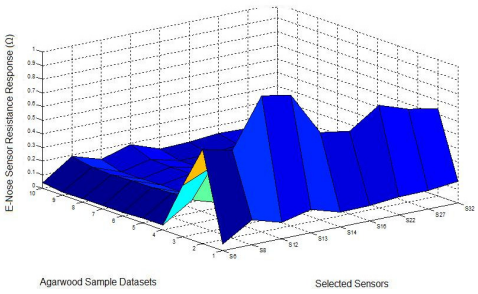

Fig. 5. Mean of normalized data sample from S6, S8, S12, S13, S14, S16, S22, S27 and S32 from Melaka region.

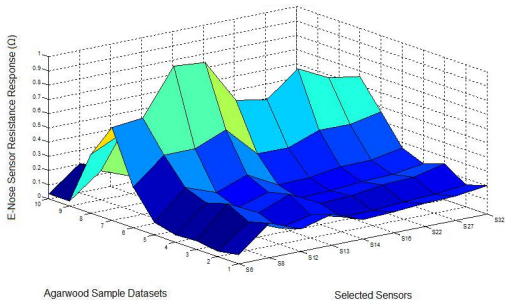

Fig. 6. Normalized data sample S6, S8, S12, S13, S14, S16, S22, S27 and S32 from Super Pagoh region.

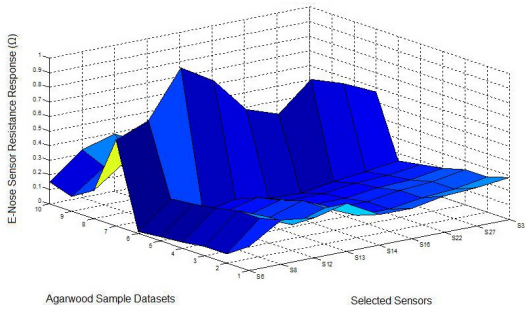

Fig. 7. Normalized data sample from S6, S8, S12, S13, S14, S16, S22, S27 and S32 from Ulu Tembeling region.

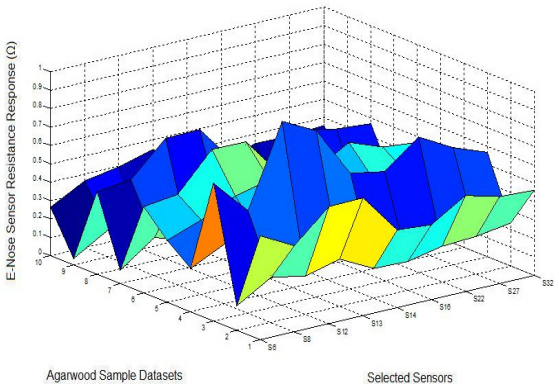

Fig. 8. Mean of normalized data sample from S6, S8, S12, S13, S14, S16, S22, S27 and S32 from Pagoh region.

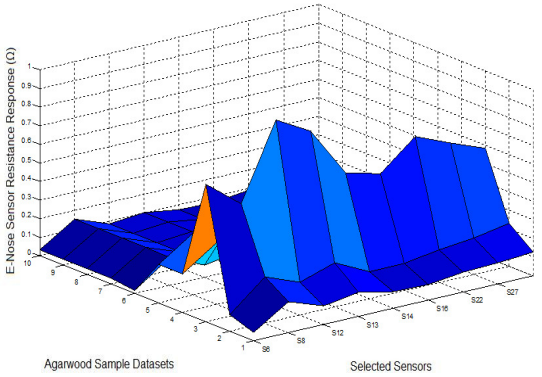

Fig. 9. Mean of normalized data sample from S6, S8, S12, S13, S14, S16, S22, S27 and S32 from Indonesia region.

Fig. 5 until 9 show the pattern variations and the standard deviation of mean respectively from S6, S8, S12, S13, S14, S16, S22, S27 and S32 from five Agarwood regions. It was found that combination of S6, S8, S12, S13, S14, S16, S22, S27 and S32were having high variance as compared to other combinations. This is due to the mean of normalized data sample having high variance across the region. This implies that the data is significant, can be trained and validated into ANN model.

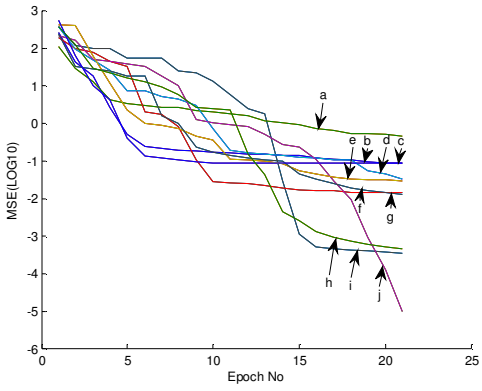

Fig. 10.Training mean of square error (MSE) plot over 4 epocl networks with (a) 8; (b) 9; (c) 5; (d) 10; (e) 7; (f) 6; (g) 3; (h) 4; (i) 1; (j) 2 hidden layer size.

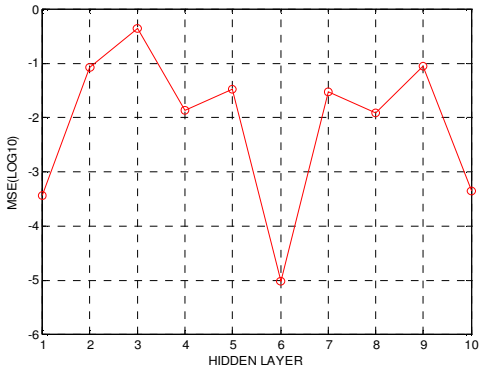

Fig. 11. ANN final trainig errors with varying hidden layer

In Fig. 10, the variance of speed of convergence among 10 hidden layer sizes can be seen and compared from the slowest and fastest with log of mean of squared error (MSE) versus epoch number. The MSE value can be compared clearly in Fig. 11. The slowest convergence is hidden layer size 6. The different of the speed of convergence greatly different with the speed belong to hidden layer size 2 which is the fastest convergence. This result shows that, the decision to find the optimal hidden layer size is promising through varying the hidden layer size 1 to 10. Hence, it indicates that the trained network for classification of Agarwood with 3 hidden layer sizes provides best architecture. This network architecture was then validated with the rest of the testing data.

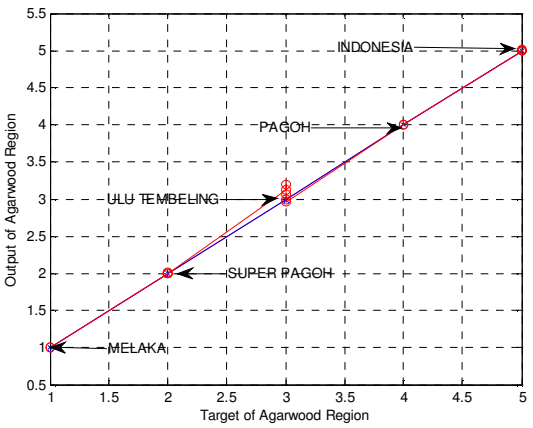

Fig. 12. ANN Model of Agarwood Classification.

Fig. 12 shows the training and validation data respectively. It indicates that the classification of Agarwood sample from five (5) regions produces very minima error whereby almost all regions are slightly misalignment at region from Melaka. However it clearly shows that the ANN model for Agarwood effectively classify the samples from Melaka, Super Pagoh, Ulu Tembeling, Pagoh, and Indonesia as region 1, 2, 3, 4 and 5, respectively.

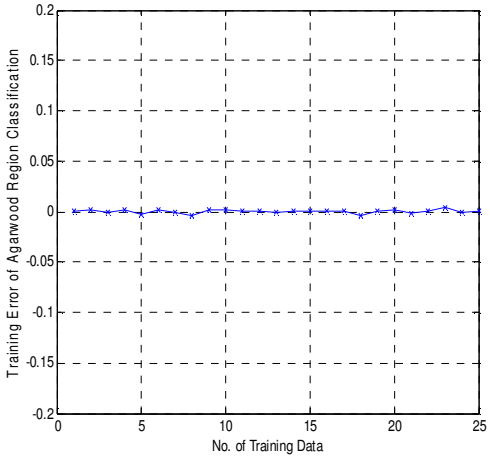

Fig. 13. Training Error.

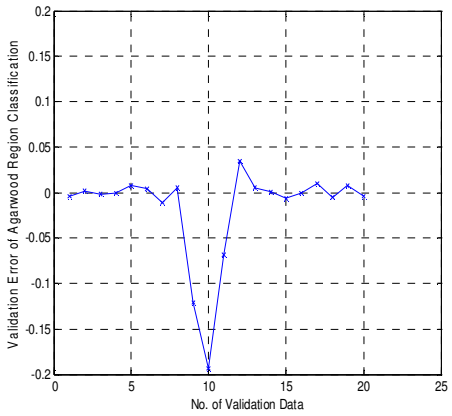

Fig. 14. .Validation Error.

Finally, Fig. 13 and 14 show the training and validation data error respectively. At all regions, the maximum training data error (≈ 0%) which is less than that of the maximum validation data error (≤ 20 %) because it was effectively fitted modeled network. It also shows that one cannot depend on the fitting performance of the training network single-handedly but must be validated with testing method as exercised in this work.

A consequence of the results obtained in the testing of the ANN model is the results of classification matrix of desired true region which is arranged in Table 1. This Table shows how the numbers of outputs match the number of targets as explained in Table 1, hence, all outputs are consistent with its target by undertaking the threshold value where it was preset for the accuracy. In this case, the threshold value was set at 0.2. Thus, the fit value for each region of this level of accuracy must lie between ± 0.2 from the target value while the unfit one is the value which cannot achieve the target between the predetermined range. Hence, the outcomes of the classification is tabulated in Table 1 demonstrate the high accuracy of classification rate.

TABLE I
CLASSIFICATION MATRIX WHICH SUMMARIZES THE CLASSIFICATION OF 100% ACCURACY

Predicted Region	True Region					
	TR 1	TR 2	TR 3	TR 4	TR 5	Total (True)
TR1	4	0	0	0	0	4
TR2	0	4	0	0	0	4
TR3	0	0	4	0	0	4
TR4	0	0	0	4	0	4
TR5	0	0	0	0	4	4
Total (Predicted)	4	4	4	4	4	20

V. CONCLUSION

A neural network has been modeled to classify Agarwood into five regions; Melaka, Super Pagoh, Ulu Tembeling, Pagoh and Indonesia using E-Nose employed with ANN. The raw data has been normalized and pre-processed in order to have unique data features from each Agarwood region as inputs to the ANN model. The architecture of the network designed with 5 input neurons, 2 hidden layer sizes and 1 output neuron is compatible with the objective of this work. The network model training was performed using Backpropagation Neural Network (BPNN) training algorithm. The algorithm has shown high accuracy in classifying the Agarwood from 5 different regions. The training and validation errors were analyzed based on MSE value. For future improvement, the model may be trained and optimized by varying other parameters such as learning rate and momentum rate to have less MSE error. Another improvement that can be made is by examining the sample data to be preprocessed and introducing randomized noise into the network for robustness guarantee of the network model. Based on the experimental result, the eventual consequence of realizing the ANN of the classification work, it can be seen that the Agarwood sample is linearly separable among various Agarwood sample regions. It is feasible to use the system to accurately classify the Agarwood region.

ACKNOWLEDGMENT

This work was using the data gathered at the Forest Research Institute Malaysia (FRIM) with collaboration of Advance Signal Processing (ASP) research group Faculty of Electrical Engineering UiTM. The authors would like to thank all ASP research group UiTM and FRIM for supporting this research.

REFERENCES

[1] J. Ueda, L. Imamura,Y. Tezuka,Q. La Tran,M. Tsuda, and S. Kadota,, "New sesquiterpene from Vietnamese agarwood and its induction effect on brain-derived neutrophic factor mRNA expression in vitro," *Bioorganic & Medicinal Chemistry,* vol. 14, pp. 3571-3574, 2006.

[2] T. Hiroaki, M. Ito, T. Shiraki, T. Yagura and G. Honda, *Sedative effects of vapor inhalation of agarwood oil and spikenard extract and identification of their active components* vol. 62: Springer, 2008.

[3] M. A. Nor Azah, J.Mailina, A. Abu Said, J. Abd. Majid, S. Saidatul Husni, H. Nor Hasnida, and Y. Nik Yasmin, "Comparison of chemical profiles of selected Gaharu oils from peninsular Malaysia," *Malaysian Journal of Analytical Sciences,* vol. 12, pp. 338-340, 2008.

[4] Q. Shu-yuan, *III Aquilaria species: in vitro culture and the production of eaglewood (agarwood)* vol. 33: Bristish Library, 1995.

[5] Z. Jinhua, Z. Chunshan, J. Xinyu and X. Lianwu, "Extraction of essential oil from shaddock peel and analysis of its components by gas chromatography mass spectometry," *J. Cent. South Uni. Technol,* vol. 13, 2006.

[6] S. A. Rezzoug*, et al.*, "Optimization of operating conditions of rosemary essential oil extraction by a fast controlled pressure drop process using response surface methodology," *Journal of Food Engineering,* vol. 71, pp. 9-17, 2005.

[7] S. S. Susan, G.K. Bahram and H. Troy Nagle *Analysis of Medication Off-odors Using an electronic nose*: Oxford University Press, 1997.

[8] J. W. Gardner*, et al.*, "An electronic nose system for monitoring the quality of potable water," *Sensors and Actuators B: Chemical,* vol. 69, pp. 336-341, 2000.

[9] N. Barié*, et al.*, "A novel electronic nose based on miniaturized SAW sensor arrays coupled with SPME enhanced headspace-analysis and its use for rapid determination of volatile organic compounds in food quality monitoring," *Sensors and Actuators B: Chemical,* vol. 114, pp. 482-488, 2006.

[10] A. Cimato*, et al.*, "Analysis of single-cultivar extra virgin olive oils by means of an Electronic Nose and HS-SPME/GC/MS methods," *Sensors and Actuators B: Chemical,* vol. 114, pp. 674-680, 2006.

[11] M. Lebrun*, et al.*, "Discrimination of mango fruit maturity by volatiles using the electronic nose and gas chromatography," *Postharvest Biology and Technology,* vol. 48, pp. 122-131, 2008.

[12] P. U. Kurup, "An Electronic Nose for Detecting Hazardous Chemicals and Explosives," in *Technologies for Homeland Security, 2008 IEEE Conference on*, 2008, pp. 144-149.

[13] W. Ping*, et al.*, "Development of electronic nose for diagnosis of lung cancer at early atage," in *Information Technology and Applications in Biomedicine, 2008. ITAB 2008. International Conference on*, 2008, pp. 588-591.

[14] M. A. Markom, A. Y. Md Shakaff, A. H. Adom, M.N. Ahmad and A.H. Abdullah, "The Feasibility Study of Utilising Electronic Nose and ANN for Plant Malaise Detection," in *Proceedings of MUCET2008 Malaysian Universities Conferences on Engineering and Technology*, Perlis, Malaysia, 2008.

[15] S. H. Saeed*, et al.*, "Experimental use of electronic nose for analysis of volatile organic compound (VOC)," in *Multimedia, Signal Processing and Communication Technologies, 2009. IMPACT '09. International*, 2009, pp. 113-115.

[16] F. K. Che Harun*, et al.*, "An electronic nose employing dual-channel odour separation columns with large chemosensor arrays for advanced odour discrimination," *Sensors and Actuators B: Chemical,* vol. 141, pp. 134-140, 2009.

[17] A. M. Marina, Y. B. Che Man and I. Amin, " Use of the SAW sensor electronic nose for detecting the adulteration of virgin coconut oil with RBD palm kernel olein," *Journal of the American Oil Chemists' Society,* 2009.

[18] M. A. Markom*, et al.*, "Intelligent electronic nose system for basal stem rot disease detection," *Computers and Electronics in Agriculture,* vol. 66, pp. 140-146, 2009.

[19] M. N. Taib*, et al.*, "Extending the response range of an optical fibre pH sensor using an artificial neural network," *Analytica Chimica Acta,* vol. 330, pp. 31-40, 1996.

[20] T. E. Brook*, et al.*, "Extending the range of a fibre-optic relative-humidity sensor," *Sensors and Actuators B: Chemical,* vol. 39, pp. 272-276, 1997/4//.

[21] M. N. Taib and R. Narayanaswamy, "Multichannel calibration technique for optical-fibre chemical sensor using artificial neural network," *Sensors and Actuators B: Chemical,* vol. 39, pp. 365-370, 1997/4//.

[22] F. B. M. Suah*, et al.*, "Optimisation of the range of an optical fibre pH sensor using feed-forward artificial neural network," *Sensors and Actuators B: Chemical,* vol. 90, pp. 175-181, 2003.

[23] N. Hong-seok*, et al.*, "Parylene gas chromatographic column for rapid thermal cycling," *Microelectromechanical Systems, Journal of,* vol. 11, pp. 718-725, 2002.

[24] E. S. Kolesar, Jr. and R. R. Reston, "Silicon micromachined gas chromatography system," in *Innovative Systems in Silicon, 1997. Proceedings., Second Annual IEEE International Conference on*, 1997, pp. 117-125.

[25] J. Sun*, et al.*, "A high resolution MEMS based gas chromatography column for the analysis of benzene and toluene gaseous mixtures," *Sensors and Actuators B: Chemical,* vol. 141, pp. 431-435, 2009.

[26] S. Fuerstenau*, et al.*, "Active pixel sensors for mass spectrometry," *International Journal of Mass Spectrometry,* vol. 215, pp. 101-111, 2002.

[27] G. M. Downes*, et al.*, "Measuring and modelling stem growth and wood formation: An overview," *Dendrochronologia,* vol. 27, pp. 147-157, 2009.

[28] P. Jiazhi and H. Yong, "Recognition of Plants by Leaves Digital Image and Neural Network," in *Computer Science and Software Engineering, 2008 International Conference on*, 2008, pp. 906-910.

[29] M. Hongbo and Q. Dawei, "recognition of wood defects based on artificial neural network," in *Advanced Intelligent Mechatronics, 2008. AIM 2008. IEEE/ASME International Conference on*, 2008, pp. 1232-1237.

[30] Y.-M. Yang, *et al.*, "Electronic nose based on SAWS array and its odor identification capability," *Sensors and Actuators B: Chemical,* vol. 66, pp. 167-170, 2000.

[31] G. Önkal-Engin, *et al.*, "e-NOSE Response Classification of Sewage Odors by Neural Networks and Fuzzy Clustering," in *Advances in Natural Computation*, ed, 2005, pp. 648-651.

[32] S. Borah, *et al.*, "Neural network based electronic nose for classification of tea aroma," *Sensing and Instrumentation for Food Quality and Safety,* vol. 2, pp. 7-14, 2008.

[33] H. Yu, *et al.*, "Quality grade identification of green tea using E-nose by CA and ANN," *LWT - Food Science and Technology,* vol. 41, pp. 1268-1273, 2008.

[34] D. S. Giordani, *et al.*, "Identification of the Biodiesel Source Using an Electronic Nose," *Energy & Fuels,* vol. 22, pp. 2743-2747, 2008.

[35] R. Dutta, *et al.*, "Stochastic resonance-based electronic nose: A novel way to classify bacteria," *Sensors and Actuators B: Chemical,* vol. 115, pp. 17-27, 2006.

[36] R. S. Hobson, *et al.*, "Temperature correction to chemoresistive sensors in an e-NOSE-ANN system," *Sensors Journal, IEEE,* vol. 3, pp. 484-489, 2003.

[37] G. Pioggia, *et al.*, "Towards a Real-Time Transduction and Classification of Chemoresistive Sensor Array Signals," *Sensors Journal, IEEE,* vol. 7, pp. 237-244, 2007.

[38] K. R. Kashwan and M. Bhuyan, "Robust electronic-nose system with temperature and humidity drift compensation for tea and spice flavour discrimination," in *Sensors and the International Conference on new Techniques in Pharmaceutical and Biomedical Research, 2005 Asian Conference on*, 2005, pp. 154-158.

[39] W. Hidayat, *et al.*, "Agarwood Grading Using an Electronic Nose and Artificial Neural Networks Based on Its Fragrance," in *Proceeding of International Conference on Mathematics and Natural Science (ICMNS 2006)*, Indonesia, 2006.

[40] D. Luo, *et al.*, "Application of ANN with extracted parameters from an electronic nose in cigarette brand identification," *Sensors and Actuators B: Chemical,* vol. 99, pp. 253-257, 2004.

[41] A. Catarina Bastos and N. Magan, "Potential of an electronic nose for the early detection and differentiation between Streptomyces in potable water," *Sensors and Actuators B: Chemical,* vol. 116, pp. 151-155, 2006.

[42] T. Dewettinck, *et al.*, "The electronic nose as a rapid sensor for volatile compounds in treated domestic wastewater," *Water Research,* vol. 35, pp. 2475-2483, 2001.

[43] S. Panigrahi, *et al.*, "Neural-network-integrated electronic nose system for identification of spoiled beef," *LWT - Food Science and Technology,* vol. 39, pp. 135-145, 2006.

[44] F. Maussang, *et al.*, "Automated segmentation of SAS images using the mean - standard deviation plane for the detection of underwater mines," in *OCEANS 2003. Proceedings*, 2003, pp. 2155-2160 Vol.4.

[45] D. Bellan, *et al.*, "Characterisation of crosstalk in terms of mean value and standard deviation [multiconductor transmission lines]," *Science, Measurement and Technology, IEE Proceedings -*, vol. 150, pp. 289-295, 2003.

[46] H. Yu, *et al.*, "Identification of Tea Storage Times by Linear Discrimination Analysis and Back-Propagation Neural Network Techniques Based on the Eigenvalues of Principal Components Analysis of E-Nose Sensor Signalsals," *Sensors,* vol. 9, pp. 8073-8082, 2009.

[47] P. Boilot, *et al.*, "Classification of bacteria responsible for ENT and eye infections using the Cyranose system," *Sensors Journal, IEEE,* vol. 2, pp. 247-253, 2002.

[48] J.Gardner, "http://www.ncbi.nlm.nih.gov/pmc/articles/PMC149373

[49] R. Dutta, *et al.*, "Identification of Staphylococcus aureus infections in hospital environment: Electronic nose based approach," *Sensors and Actuators, B: Chemical,* vol. 109, pp. 355-362, 2005.

[50] M. Falasconi, *et al.*, "Detection of toxigenic strains of Fusarium verticillioides in corn by electronic olfactory system," *Sensors and Actuators, B: Chemical,* vol. 108, pp. 250-257, 2005.

[51] M. Z. Z. Abiden and S. Z. Z. Abidin, "Introducing residual errors in accuracy assessment for remotely sensed change detection," in *Signal Processing & Its Applications, 2009. CSPA 2009. 5th International Colloquium on*, 2009, pp. 89-92.

[52] T. Nakamoto, *et al.*, "Perfume and flavor identification by odor sensing system using quartz-resonator sensor array and neural-network pattern recognition," in *Solid-State Sensors and Actuators, 1991. Digest of Technical Papers, TRANSDUCERS '91., 1991 International Conference on*, 1991, pp. 355-358.

[53] M. A. Sovierzoski and F. I. M. Argoud, "Evaluation of ANN Classifiers During Supervised Training with ROC Analysis and Cross ValidationROC Analysis and Cross Validation," in *2008 International Conference on BioMedical Engineering and Informatics*, 2008.

[54] H. Chih-Cheng, *et al.*, "A comparative study of radial basis function neural networks and wavelet neural networks in classification of remotely sensed data," in *Automation Congress, 2002 Proceedings of the 5th Biannual World*, 2002, pp. 455-461.

[55] W. Guohua and Z. Xiaodong, "A Fast Audio Digital Watermark Method Based on Counter-Propagation Neural Networks," in *Computer Science and Software Engineering, 2008 International Conference on*, 2008.

[56] L. Kovacs and G. Terstyanszky, "Boundary region sensitive classification for the counter-propagation neural network," in *Neural Networks, 2000. IJCNN 2000, Proceedings of the IEEE-INNS-ENNS International Joint Conference on*, vol.1, 2000.

[57] P. Raviram, *et al.*Concurrency Control in CAD With KBMS Using Counter Propagation Neural Network," in *Advance Computing Conference, 2009. IACC 2009. IEEE International*, 2009, pp. 1521-1525.

[58] J. Wenbiao, *et al.*, "Audio Watermarking Algorithm Robust to TSM Based on Counter Propagation Neural Network," in *Image and Signal Processing, 2009. CISP '09. 2nd International Congress on*, 2009, pp. 1-4.

[59] H. Li, *et al.*, "An improved recursive prediction error algorithm for training recurrent neural networks," in *Intelligent Control and Automation, 2000. Proceedings of the 3rd World Congress on*, 2000, pp. 1043-1046 vol

[60] Y. Zhang and X. R. Li, "A fast and robust recursive prediction error learning algorithm for feedforward neural networks," in *Decision and Control, 1996., Proceedings of the 35th IEEE*, 1996, pp. 2036-2041 vol.2.

[61] W. Yongji, *et al.*, "A second order recursive prediction error algorithm for diagonal recurrent neural networks," in *Control Applications, 1998. Proceedings of the 1998 IEEE International Conference on*, 1998, pp. 172-176 vol.1.

[62] T. Srinivasan, *et al.*, "Knowledge discovery in clinical databases with neural network evidence combination," in *Intelligent Sensing and Information Processing, 2005. Pros of 2005 Int. Conference on*, 2005, pp. 512-517.

[63] M. R. Aghamohammadi, *et al.*, "Power system dynamic security classification using Kohenen neural networks," in *Power Systems Conference and Exposition, 2009. PSCE '09. IEEE/PES*, 2009, pp. 1-7.

[64] B. Comes and A. Kelemen, "Probabilistic neural network classification for microarraydata," in *Neural Networks, 2003. Proceedings of the International Joint Conference on*, 2003, pp. 1714-1717 vol.3.

[65] K. Z. Mao, *et al.*, "Probabilistic neural-network structure determination for pattern classification," *Neural Networks, IEEE Transactions on,* vol. 11, pp. 1009-1016, 2000.

[66] H. Wei-Bing, *et al.*, "A novel probabilistic neural network system for power quality classification based on different wavelet transform," in *Wavelet Analysis and Pattern Recognition, 2007. ICWAPR '07. International Conference on*, 2007, pp. 746-750.

[67] Z. Fang, *et al.*, "The Application of BP Neural Network on Mechanical Failure Classification," in *Intelligent Systems and Applications, 2009. ISA 2009. International Workshop on*, 2009, pp. 1-3.

[68] W. Hu and X. Yong, "Study on Customer Classification Based on BP Neural Networks," in *Wireless Communications, Networking and Mobile Computing, 2008. WiCOM '08. 4th International Conference on*, 2008, pp. 1-4.

[69] Z. Jinhui, *et al.*, "Application of BP Neural Network to Sugarcane Diseased Spots Classification," in *Intelligent Information Technology Application, 2008. IITA '08. Second International Symposium on*, 2008,.

[70] L. Suyi, *et al.*, "Classification of Fabric Defect Based on PSO-BP Neural Network," in *Genetic and Evolutionary Computing, 2008. WGEC '08. Second International Conference on*, 2008, pp. 137-140.

[71] M. Hong, *et al.*, "An Electronic Tongue Based on Genetic Algorithm Improving BP Neural Network," in *Computer Science and Software Engineering, 2008 International Conference on*, 2008, pp. 823-826.

[72] T. Kang, *et al.*, "A Study of Broccoli Grading System Based on Machine Vision and Neural Networks," in *Mechatronics and Automation, 2007. ICMA 2007. International Conference on*, 2007, pp. 2332-2336.

[73] Z. Wei, *et al.*, "Factoring tertiary classification into binary classification improves neural network for protein secondary structure prediction," in *Computational Intelligence in Bioinformatics and Computational Biology, 2004. CIBCB '04. Proceedings of the 2004 IEEE Symposium on*, 2004, pp. 175-181.

[74] J. Chen, *et al.*, "Adaptive binary tree for fast SVM multiclass classification," *Neurocomputing,* vol. 72, pp. 3370-3375, 2009.

[75] P. Kraipeerapun and C. C. Fung, "Binary classification using ensemble neural networks and interval neutrosophic sets," *Neurocomputing,* vol. 72, pp. 2845-2856, 2009.

[76] J. C. W. Chan, *et al.*, "Binary Classification Strategies for Mapping Urban Land Cover with Ensemble Classifiers," in *Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE Int.*, 2008, pp.1004 - 1007.

[77] A. L. D. Rossi and M. A. d. O. Camargo-Brunetto, "Protein Classification Using Artificial Neural Networks with Different Protein Encoding Methods," in *Seventh Int. Conference onIntelligent Systems Design an Intelligent Systems Design and Applications*, 2007.

Muhammad Sharfi Najib received B.Eng.(Hons) in Mechatronics Engineering from International Islamic University Malaysia (IIUM), Gombak in 2001 and M.Sc in Automation and Control from University of Newcastle Upon Tyne, UK in 2005. Currently he is a lecturer in Faculty of Electrical & Electronics Engineering, University Malaysia Pahang (UMP), Kuantan. His research interest includes instrumentation, control system, intelligent system.

Dr. Nor Azah Mohamad Ali received received her tertiary education from University Waterloo Canada (BSc Chemistry 1985), Universiti Kebangsaan Malaysia (MSc in chemistry 1996) and Natural Product Chemistry from University Putra Malaysia (Phd 2004). Currently she is the head of Herbal Product Development Programme and senior scientist at Natural Products Division, Forest Research Institute Malaysia (FRIM), Kepong, Malaysia. Her research interest includes natural product chemistry, essential oils, product development, instrumentation and sensor technology

Mohamad Nasir Mat Arip received B.Sc.(Hons) Upper Division (2003) and M.Sc in Chemistry (2005) from Universiti Kebangsaan Malaysia (UKM), Bangi. Currently he is a Research Officer in Forest Producs Division, Forest Research Institute Malaysia (FRIM), Kepong. His research interest includes Chemical Sensors and Biosensors

Abd Majid Bin Jalil received Dip.in Chemical Eng.(1999) from Universiti Teknologi Malaysia (UTM), Skudai Johor Bahru. Currently he is a Assistant Research Officer in Natural Producs Division, Forest Research Institute Malaysia (FRIM), Kepong

Mohd Nasir Taib received B.eng. (Electrical) from University of Tasmania,Australia , M.Sc (Control System) from University of Sheffield, UK and the PhD (Control & Instrumentation) from UMIST, UK. He is currently a Professor and Dean of Faculty of Electrical Engineering, UiTM, Shah Alam. He has been the chair of IEEE Control System Society, Malaysia Chapter, since 2009. His area of interest is advanced signal processing and control system