Pertanika J. Sci. & Technol. 25 (S): 119 - 128 (2017)

SCIENCE & TECHNOLOGY

PERTANIKA JOURNALS

Journal homepage: http://www.pertanika.upm.edu.my/

Classification of Aromatic Herbs using Artificial Intelligent Technique

A. Che Soh* , U. K. Mohamad Yusof, N. F. M. Radzi, A. J. Ishak and M. K. Hassan

Department of Electrical and Electronic Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

ABSTRACT

Herbs have unique characteristics such as colour, texture and odour. In general, herb identification is through organoleptic methods and is heavily dependent on botanists. It is becoming more difficult to identify different herb species in the same family based only on their aroma . It is because of their similar physical appearance and smell. Artificial technology, unlike humans, is thought to have the capacity to identify different species with precision. An instrument used to identify aroma is the electronic nose. It is used in many sector including agriculture. The electronic nose in this project was to identify the odour of 12 species such as lauraceae, myrtaceae and zingiberaceae families. The output captured by the electronic nose gas sensors were classified using two types of artificial intelligent techniques: Artificial Neural Network (ANN) and Adaptive Neuro-Fuzzy Inference System (ANFIS). From the result, ANFIS has 94.8% accuracy compared with ANN at 91.7%.

Keywords: Artificial Neural Network, Adaptive Neuro-Fuzzy Inference System

INTRODUCTION

The leaves of the plant that do not develop persistent woody tissue are called herbs (Chen et al., 2012). What make herbs valuable is their taste, aroma, health and medicinal properties, commercial significance, pesticide properties and colour sources (Fischer, 2010; Haddi et al., 2013; Konduru et al., 2015). The presence of phytochemical in the form of volatile compound gives herbs their characteristic aroma. Beneficial properties found in herbs are terpenes, steroids, phenolic compounds, amino acids, lipids, and alkaloids (Ganora, 2008). Researchers recognise plant species by analysing the physical form or the texture of the herbs (Husin et al., 2012; Ishak et al., 2009; Zalikha, 2011).

ARTICLE INFO

Article history:
Received: 24 August 2016
Accepted: 02 December 2016

E-mail addresses:
azuracs@upm.edu.my (A. Che Soh),
umi.mpputm@gmail.com (U. K. Mohamad Yusof),
nfmr86@gmail.com (N. F. M. Radzi),
asnorji@upm.edu.my (A. J. Ishak),
khair@upm.edu.my (M. K. Hassan)
*Corresponding Author

ISSN: 0128-7680

Humans identify different herb species by using their sensory panels. However, critics have pointed out that our sense of smell is subjective and is usually inaccurate. Hence, this severely limits our identification capability. Among the factors that can influence the human sensory system are physical, mental health, tiredness and other conditions of the body (Tudu et al., 2009). On the other hand, the main problem in identifying different herbs in the same family is their physical appearance (they may look and smell alike). Even botanists face difficulty identifying different herbs based on their aroma.

In recent years, researchers have used chemical gas and liquid to differentiate the aroma of different herbs. It is a complex procedure and an expensive one involving an aroma-detecting equipment (Fischer 2010). Volatile gas from herbs are analysed using complicated and expensive experiment involving gas chromatography (GC) with a selective mass spectrometric (MS) detector. The result is accurate but involves various experiments, time consuming, and costly (Fischer, 2010). There is a demand for new technology that provides good results in real time with low cost, simple procedures and user friendly. Consequently, the electronic nose sensor was invented to detect aroma (Wilson, 2013). The device is popular in the herbs industry because it has several advantages such as the ability to provide chemical and physical information of the plant in real time (Dinrifo, 2011; Haddi et al., 2013; Wilson, 2013). It can also detect simple or complicated smell. An electronic nose consists of an array of electronic chemical sensors with partial specificity and an applicable pattern-recognition system (Wilson & Baietto, 2009). In biological olfactory system, the odour of sample will be obtained from the smell process before that information is processed by the brain. The neural system will recognise the sample and identify the odour. In an electronic nose system on the other hand, raw data from the odour signal is captured by the gas sensor array and processed using algorithm formulated such as neural network to identify the odour. The output or the result will be acquired from the database of the system.

A particularly significant and interesting aspect of electronic nose system is the classification of herb species. Pattern recognition systems focus on recognise patterns and regularities in data. Basically, the system was trained from labelled training data in supervised learning. Unlabelled data is identified by the formulated algorithm to discover previously unknown patterns in unsupervised learning. Prediction problems in pattern recognition relate to classification, regression and clustering (Guterriez, 2002). The pattern recognition is a problem of assigning an object to a class. The most common classification algorithm used in artificial olfactory system is artificial neural network (Amari et al., 2006; Dinrifo, 2011; Husin, 2012; Ihsan et al., 2009; Li, 2007). Artificial neural networks (ANN) technique used in chemical vapour recognition have proven to be suitable in analysing and recognising patterns for complex data. The design of artificial neural network is inspired by the human brain. The structure of ANN consists of a pyramid of layers, where the neurons are organised and linked to the external environment by input and output layers. Every neuron is a basic information-processing unit that can calculate its activation level given the inputs and numerical weights. The weights are modified to bring the network input and output behaviour in line with the environment (Dinrifo, 2011; Guterriez, 2002; Li et al., 2007).

Comparison of artificial neural networks (ANN) with fuzzy inference systems (FIS), showed that the neural network was difficult to use due to prior knowledge rule or it has to be learnt from scratch. Among the disadvantages of the neural network system are complex learning algorithms and difficult to extract knowledge. Compared with fuzzy inference systems, it can incorporate prior rule-base, interpretable by if-then rules, simple interpretation and implementation. However, the fuzzy system is unable to acquire linguistic knowledge. Additionally, knowledge must be provided. Therefore, an integrated system that combines the FIS and ANN modelling concept is an advantage and complements each other (Gulbag & Temurtas, 2006).

There are not many studies that look at identification of herbs in the same family based on their aroma. Therefore, the aim of this study is to explore, analyse and show the difference between herbs based on their aroma. Mohamad Yusof et al. (2015) employed an electronic nose to classify 12 herb species from three aromatic herbs families, namely *Lauraceae, Myrtaceae*, and *Zingiberaceae* was studied by. This artificial intelligence was effective in acquiring signals and advantageous for sample preparation compared with other systems. Raw data from the odour signal is captured by the gas sensor array in electronic nose system. The signal = is processed using several standard normalisation techniques to give better interpretation of data. The objective of this paper is to compare the performance of two types of artificial intelligent techniques. The artificial neural network (ANN) and adaptive neuro-fuzzy inference system (ANFIS) is employed using normalised data for herbs classification. The accuracy of classification of both techniques is presented for 12 herb species in three families. The performance of both techniques will be evaluated based on of the accuracy of the system to classify the herbs species.

THEORY AND METHODS

Experiment Overview

Electronic nose was used to classify 12 aromatic herbs species from Lauraceae, Myrtaceae, and Zingiberaceae family by Mohamad Yusof et al. (2015). The list of herbs was chosen and collected with the consultation of botanist from Bioscience Institute based on the availability of samples from Agricultural Conservatory Park, Universiti Putra Malaysia. The scientific name of the sample is listed in Table 1. Due to fast response, affordable cost, low power consumption and large number of target gas detection, multiple metal oxide gas sensors from Figaro were selected as shown in Table 2 to detect a broad range of chemical compound according to the phytochemical of the herbs 15 g applied for each species as a sample.

Artificial Neural Network (ANN)

The ANN model is designed from two to five inputs to find the best result of herbs classification. Training was done by using scaled conjugate gradient backpropagation method. Data was divided into 70% training, 15% testing and of the rest for validation. Sigmoid activation function was used in neural network of the study. The architecture of neural network in this research is

Table 1
Scientific name of twelve herb species

Family Name	Abbreviation	Scientific Name
Lauraceae	LCI	1. *Cinnamomum Iners*
	LCV	2. *Cinnamomum Verum*
	LCP	3. *Cinnamomum Porrectum*
	LLE	4. *Litsea Elliptica*
Myrtaceae	MSA	5. *Syzygium Aromaticum*
	MSP	6. *Syzygium Polyanthum*
	MMA	7. *Melaleuca Alternifolia*
	MRT	8. *Rhodomyrtus Tomentosa*
Zingiberaceae	ZSK	9. *Scaphoclamys Kunstleri*
	ZET	10. *Etlingera Terengganuensis*
	ZZZ	11. *Zingiber Zerumbet*
	ZEC	12. *Elettariopsis Curtisii*

Table 2
The selected FIGARO MOS gas sensor for electronic nose

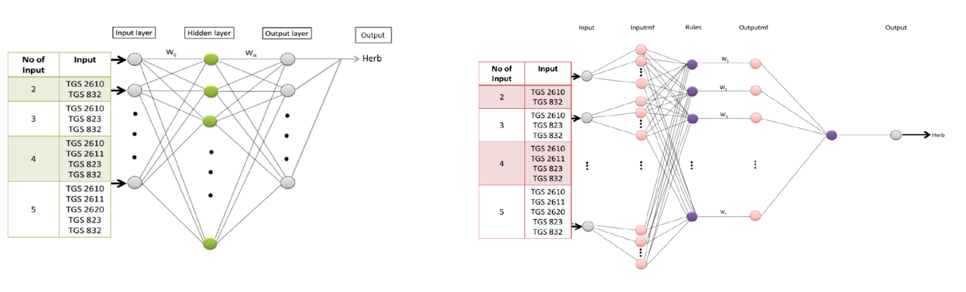
Sensor Type	Abbreviation	Type of gas detection
TGS 2610	Sensor 1	Butane, propane, liquefied petroleum gas
TGS 2611	Sensor 2	Methane, natural gas
TGS 2620	Sensor 3	Alcohol, toluene, xylene, volatile organic compound
TGS 823	Sensor 4	Organic solvent vapours
TGS 832	Sensor 5	Halocarbon, Chlorofluorocarbon

made up of the input from sensor 1, sensor 2, sensor 3, sensor 4 and sensor 5, 20 hidden layers and 12 species of herbs as the output. Every node from the input layer is connected to a node from the hidden layer. On the other hand, every node from the hidden layer is connected to a node in the output layer. Each link is associated with the weight *wij*.

The input layer represents the raw information that is fed into the network. Every single input to the network is duplicated and sent to the nodes in hidden layer. Data is accepted from the input layer by the hidden layer. It uses the input values and is modified based on weight, *wij* value. The new value will be sent to the output layer. After that, there will be another modification based on the weight, *wjk* from the connection between hidden and output layer. Finally, the output layer processes the information received from the hidden layer and produces the output for 23 herbs species. The two layers feed-forward back propagation structure is illustrated in Figure 1(a).

Adaptive Neuro-Fuzzy Inference System (ANFIS)

The ANFIS is another classification technique used in this study. To build the fuzzy inference system into the structure, the subtractive clustering was selected. For better performance in



(a) (b)

Figure 1. (a) Structure of ANN; b) Structure of ANFIS

the training phase, hybrid optimisation method was adopted. This is because it is faster and the results are closest compared with the back propagation gradient descent optimisation method. Sensor 1 to sensor n input layer is the first layer of the ANFIS structure. Premise or antecedent parameters of the ANFIS are contained in the second layer. It is dedicated to the fuzzy sub-space. The consequent parameters of the fifth layer were used to optimise the network. In the hybrid learning algorithm, the node outputs go forward until layer five and the consequent parameters are identified by least-square method during the forward pass. In the backward pass, error signals propagate backwards and the premise parameters are updated by gradient descent method. The structure of ANFIS in this project is shown in Figure 1(b).

Throughout the learning process, the parameters associated with the membership functions changed. A gradient vector facilitates the computation of these parameters. It provides a measurement of the fuzzy inference system modelling the input or output data for a given set of parameters. When the gradient vector is obtained, any of several optimisation routines can be applied in order to adjust the parameters to reduce measurement errors. This error measure is usually defined by the sum of the squared difference between actual and desired outputs. The ANFIS uses either back propagation or a combination of least squares estimation and back propagation to estimate membership function parameter.

The following are the rules for Sugeno-type fuzzy-rule-based model for five-input ANFIS:

- If (TGS2610 is LCI) and (TGS2611 is LCI) and (TGS2620 is LCI) and (TGS823 is LCI) and (TGS832 is LCI) then (HERBSPECIES is CinnamomumIners) (1)
- If (TGS2610 is LCP) and (TGS2611 is LCP) and (TGS2620 is LCP) and (TGS823 is LCP) and (TGS832 is LCP) then (HERBSPECIES is CinnamomumVerum) (1)
- If (TGS2610 is LCV) and (TGS2611 is LCV) and (TGS2620 is LCV) and (TGS823 is LCV) and (TGS832 is LCV) then (HERBSPECIES is CinnamomumPorrectum) (1)
- If (TGS2610 is LLE) and (TGS2611 is LLE) and (TGS2620 is LLE) and (TGS823 is LLE) and (TGS832 is LLE) then (HERBSPECIES is LitseaElliptica) (1)
- If (TGS2610 is MMA) and (TGS2611 is MMA) and (TGS2620 is MMA) and (TGS823 is MMA) and (TGS832 is MMA) then (HERBSPECIES is MelaleucaAlternifolia) (1)

- If (TGS2610 is MRT) and (TGS2611 is MRT) and (TGS2620 is MRT) and (TGS823 is MRT) and (TGS832 is MRT) then (HERBSPECIES is RhodomyrtusTomentosa) (1)
- If (TGS2610 is MSA) and (TGS2611 is MSA) and (TGS2620 is MSA) and (TGS823 is MSA) and (TGS832 is MSA) then (HERBSPECIES is SyzygiumAromaticum) (1)
- If (TGS2610 is MSP) and (TGS2611 is MSP) and (TGS2620 is MSP) and (TGS823 is MSP) and (TGS832 is MSP) then (HERBSPECIES is SyzygiumPolyanthum) (1)
- If (TGS2610 is ZEC) and (TGS2611 is ZEC) and (TGS2620 is ZEC) and (TGS823 is ZEC) and (TGS832 is ZEC) then (HERBSPECIES is ElettariopsisCurtisii) (1)
- If (TGS2610 is ZET) and (TGS2611 is ZET) and (TGS2620 is ZET) and (TGS823 is ZET) and (TGS832 is ZET) then (HERBSPECIES is EtlingeraTerengganuensis) (1)
- If (TGS2610 is ZSK) and (TGS2611 is ZSK) and (TGS2620 is ZSK) and (TGS823 is ZSK) and (TGS832 is ZSK) then (HERBSPECIES is ScaphoclamysKunstleri) (1)
- If (TGS2610 is ZZZ) and (TGS2611 is ZZZ) and (TGS2620 is ZZZ) and (TGS823 is ZZZ) and (TGS832 is ZZZ) then (HERBSPECIES is ZingiberZerumbet) (1)

The ANFIS training process first determines the fuzzy sets and the number of sets of each input variable and shape of their membership function. Training data passes through the neural network and with adjusts the input parameters to identify the relationship between input and output, and to minimise the errors. The expected output of the ANFIS will be the 12 herb species. The best structure was determined by the lowest value of the error given by the ANFIS model.

RESULTS AND DISCUSSIONS

The classification techniques in this research were implemented using ANN and ANFIS. The result of classification using ANN is shown in Table 3. Two and Five inputs of classification were done and the percentage of accuracy is given to indicate the performance of the system. The lowest percentage for 83.4% of accuracy obtained from two inputs structure of ANN. Three input systems, TGS 25610, TGS 823 and TGS 832, yielded 85.8% accuracy and less error value for 9.769E-3 compared with two inputs. With four inputs, the accuracy increases to 90.2% and 8.868E-3 value of training error. Meanwhile, five inputs show provides the highest accuracy at 91.7% to classify the sample herbs and the lowest value of training error among the other inputs. Increasing the number of inputs increases value of accuracy and with less training errors.

Table 3
Result of classification using ANN

Input	Network	MSE	Accuracy
TGS 2610, TGS 832	[2 20 12]	1.948E-2	83.4 %
TGS 2610, TGS 823, TGS 832	[3 20 12]	9.769E-3	85.8 %
TGS 2610, TGS 2611, TGS 823, TGS 832	[4 20 12]	8.868E-3	90.2 %
TGS 2610, TGS 2611, TGS 2620, TGS 823, TGS 832	[5 20 12]	7.554E-3	91.7 %

Next, the ANN network was evaluated with testing data for five inputs to show performance of the system. From the results in Table 4, ANN yielded 84.2% accuracy for classification process. The testing data is independent from the training data whereby the former contains more noise compared with training data resulting in the accuracy to be slightly lower and the MSE error to be slightly higher.

Table 4
Result of classification using ANN

Dataset	MSE	Accuracy
Training	7.554E-3	91.7 %
Testing	2.742E-2	84.2 %

In the ANFIS classification method, the highest percentage of classification was given by five inputs of ANFIS structured as 94.8% of accuracy with RMSE value 2.472E-4. The classification accuracy for four inputs was achieved at 94.7% with a difference of only 0.1% from five inputs. Furthermore, the three inputs produced 92.7% of accuracy and 4.301E-4 of RMSE. The lowest percentage is made up from two inputs which showed 85.4% accuracy. Table 5 shows classification using ANFIS technique. The ANFIS structure was also evaluated using testing data. The result as in Table 6 shows that percentage of accuracy is getting lower. Human error that may have occurred during the experimental procedure may influence data collection and lower the quality of testing data.

Table 5
Result of classification using ANFIS

Input	FIS	MSE	Accuracy
TGS 2610, TGS 832		8.6912E-4	85.4 %
TGS 2610, TGS 823, TGS 832	Fuzzy	4.301E-4	92.7 %
TGS 2610, TGS 2611, TGS 823, TGS 832	Subtractive	2.713E-4	94.7 %
TGS 2610, TGS 2611, TGS 2620, TGS 823, TGS 832	Clustering	2.472E-4	94.8 %

Table 6
Comparison of error and accuracy for training set and testing set

Dataset	RMSE	Accuracy
Training	2.472E-4	94.8 %
Testing	3.965E-4	92.7 %

The k-fold cross-validation method was used to validate the performance of the classifier in the electronic nose system as shown in Table 7. The K-fold cross validation is run several times, each with a different random arrangement in order to obtain an accurate estimate to the accuracy of a classifier. The ANFIS showed higher accuracy compared with ANN with lower value of true error where the former reported 94.8 % of true error while the latter (ANN) reported 91.2 % of true error. Hence, ANFIS was validated as better classifier compared with ANN in this research to classify 12 herb species of three families.

Table 7
K-fold cross-validation results for ANN and ANFIS

ACCURACY	NO. OF EXPERIMENT			
	EXP. 1	EXP. 2	EXP. 3	TRUE ERROR
ANN	90.9 %	91.6 %	91.1 %	91.2 %
ANFIS	92.1 %	94.8 %	97.6 %	94.8 %

CONCLUSION

In this study, we have considered two types of artificial intelligent techniques for classification purposes. The classification of the 12 herbs was successfully done using artificial neural network and adaptive neuro-fuzzy inference system. The ANFIS technique gives better performance with higher percentage of accuracy (94.8%) to classify the herb species compared with the ANN technique (91.7%). Cross-validation showed the best classifier was ANFIS by comparing the true error for both ANN and ANFIS. The results showed the proposed structure of electronic nose system was viable and hence, the objective of this study was achieved. The study had also contributed to improving the artificial olfactory system.

ACKNOWLEDGEMENT

The authors acknowledge with gratitude the financial support supported by Fundamental Research Grant Scheme, Ministry of Higher Education, Malaysia, FRGS/2/2013/TK02/UPM/02/5, and Project Title: Formulation of Algorithm to Classify Distinctive Odors Pattern of Aromatic Plant Species using Hybrid Artificial Intelligence Techniques and Universiti Putra Malaysia Grant Scheme (Geran Putra IPS) (Project Title: Development of E-Tongue Device for Herb Recognition System). Special thanks also goes to Institute of Bioscience, Universiti Putra Malaysia for providing samples of herbs.

REFERENCES

Amari, A., Barbri, N. E., Llobet, E., Bari, N. E., Correig, X., & Bouchikhi, B. (2006). Monitoring the Freshness of Moroccan Sardines with a Neural-Network Based Electronic Nose. *Sensors (Basel, Switzerland), 6*(10), 1209–1223.

Chen, C., Muhamad, A., & Ooi, F. (2012). Herbs in exercise and sports. *Journal of Physiological Anthropology, 31*(1), 4.

Dinrifo, R.R. (2011) Neural network-based electronic nose for cocoa beans quality assessment. *Agricultural Engineering International: CIGR Journal, 13*(4), 1–17.

Fischer, D.C.H. (2010). Handbook of herbs and spices. *Brazilian Journal of Pharmaceutical Sciences, 46*(4), 821–822.

Ganora, L. (2008). Herbal Constituents Foundations of Phytochemistry. *Louisville, Colorado: Herbalchem Press*, 1–15.

Gulbag, A., & Temurtas, F. (2006). A study on quantitative classification of binary gas mixture using neural networks and adaptive neuro-fuzzy inference systems. *Sensors Actuators, B Chem., 115*(1), 252–262.

Guterriez. (2002). LECTURE 1: Pattern Recognition Course Introduction. *Texas A&M University*, 1–20.

Haddi, Z., Alami, H., El Bari, N., Tounsi, M., Barhoumi, H., Maaref, A., Jaffrezic-Renault, N., & Bouchikhi, B. (2013). Electronic nose and tongue combination for improved classification of Moroccan virgin olive oil profiles. *Food Research International, 54*(2), 1488-1498.

Husin, Z., Shakaff, A.Y.M., Aziz, A.H.A., Farook, R.S.M., Jaafar, M.N., Hashim, U., & Harun, A. (2012). Embedded portable device for herb leaves recognition using image processing techniques and neural network algorithm. *Computers and Electronics in Agriculture, 89*, 18-29.

İhsan Ömür Bucak, & Bekir Karlık. (2009). Hazardous Odor Recognition by CMAC Based Neural Networks. *Sensors*, 7308–7319.

Ishak, A.J., Hussain, A., & Mustafa, M.M. (2009). Weed image classification using Gabor wavelet and gradient field distribution. *Comput. Electron. Agric., 66*(1), 53–61.

Konduru, T., Rains, G.C., & Li, C. (2015). A customized metal oxide semiconductor-based gas sensor array for onion quality evaluation: system development and characterization. *Sensors (Basel), 15*(1), 1252–1273.

Li, C., Heinemann, P., & Sherry, R. (2007). Neural network and Bayesian network fusion models to fuse electronic nose and surface acoustic wave sensor data for apple defect detection. *Sensors Actuators B Chem., 125*(1), 301–310.

Mohamad Yusof, U.K., Che Soh, A., Radzi, N.F.M., Ishak, A.J., Hassan, M.K., Ahmad, S.A., & Khamis, S. (2015). Selection of Feature Analysis Electronic Nose Signals Based on the Correlation Between Gas Sensor and Herbal Phytochemical. *Australian Journal of Basic and Applied Sciences, 9*(5), 360-367.

Tudu, B., Kow, B., Bhattacharyya, N. & Bandyopadhyay, R. (2009). Normalization techniques for gas sensor array as applied to classification of black tea. *International Journal on Smart Sensing and Intelligent Systems, 2*, 1-14.

Wilson, A.D. (2013). Diverse applications of electronic-nose technologies in agriculture and forestry. *Sensors (Basel), 13*(2), 2295–2348.

Wilson, A.D., & Baietto, M. (2009). Applications and advances in electronic-nose technologies. *Sensors (Basel), 9*(7), 5099–5148.

Zalikha, Z. (2011). Plant Leaf Identification Using Moment Invariants & General Regression Neural Network. In *11th IEEE International Conference on Hybrid Intelligent Systems,* (pp. 430-435).