E-nose based rapid prediction of early mouldy grain using probabilistic neural networks

Xiaoguo Ying, Wei Liu, Guohua Hui*, and Jun Fu

School of Information Engineering; Key Laboratory of Forestry Intelligent Monitoring and Information Technology of Zhejiang Province; Zhejiang A & F University; Linan, China

Keywords: early mouldy grain, electronic nose, non-linear, probabilistic neural network, rapid prediction

In this paper, early mouldy grain rapid prediction method using probabilistic neural network (PNN) and electronic nose (e-nose) was studied. E-nose responses to rice, red bean, and oat samples with different qualities were measured and recorded. E-nose data was analyzed using principal component analysis (PCA), back propagation (BP) network, and PNN, respectively. Results indicated that PCA and BP network could not clearly discriminate grain samples with different mouldy status and showed poor predicting accuracy. PNN showed satisfying discriminating abilities to grain samples with an accuracy of 93.75%. E-nose combined with PNN is effective for early mouldy grain prediction.

Introduction

Grain is vital to human’s life. According to previous reports, fungal growth not only brings great losses to farmers, but also induces food safety problems for human being.¹⁻³ Grain mouldy spoilage is a complex procedure from microbes by consuming large amounts of nutritional substances. Early mouldy spoilage is a stage where microbes precede normal metabolisms and produce potent mycotoxin under proper circumstances.³⁻⁵ Microbes build a saprophytic relationship with grain and its products. If it comes, grain’s color and odor change slightly, along with producing heat and the increases of fatty acid value, pH and mucedine amount, etc. If there are some effective analysis methods, this stage can be effectively prevented and commercial loss will be well controlled.

In recent years, non-destructive measurement has been paid great attention in food research. In 2012, Kaya-Celiker et al. successfully realized the discrimination of mouldy peanuts using FTIR-ATR system.⁵ In 2008, Defilippi et al. predicted Castlebrite apricots quality during storage by aroma detection using electronic nose.⁶ In 2009, Laureati et al. well discriminated and characterize 3 cultivars of Perilla frutescens by means of sensory descriptors, electronic nose and tongue analysis.⁷ E-nose detection, as a new aroma analytical technique, has been widely applied in food areas, such as fruits,⁷⁻⁸ vegetables,⁹ meats,¹⁰⁻¹¹ drinks,¹² etc. E-nose system could real-time detect sample’s quality change by gas component analysis, which imitates human olfactory sense. If grain deteriorates during storage, the gas components volatilized by fungi will change accordingly in terms of species and concentration.¹³⁻¹⁴

PCA is one commonly used method in e-nose data processing.¹⁵⁻¹⁶ As a mathematical conversion method, it transforms some related variables into unrelated variables and the permutation of these new variables is conducted according to the decrease of variance. BP network, a pattern recognition algorithm, also has broad applications in data processing.¹⁷⁻¹⁸ One of the most mentioning merits is that it gives a non-linear reflection between the inputs and the outputs. PNN, embodying Bayesian decision criterion, is first proposed by D.F. Specht based on radial basis function (RBF) neural network. This method has been confirmed a useful tool for solving pattern classification problems and shows many significant advantages including simple structure, rapid training, etc.¹⁹⁻²¹

Here, e-nose responses to mouldy grain samples were measured and analyzed by PCA, BP network, and PNN. Results indicated that PCA or BP network could not totally discriminate these samples. PNN successfully predicted all samples with an accuracy of 93.75%.

Results and Discussion

E-nose sensor original responses and PCA results

E-nose original responses to fresh rice sample are displayed in Figure 1(a). The volatile gases existing in the headspace of sample are inhaled into e-nose gas chambers and sensed by the functional materials settled in gas sensors. The specific absorption of functional materials for specific gases induces changes in material’s electrical characteristics. And sensor responses increase with the increase of gas concentrations. So signals induced by electrical changes can be used to characterize gas concentrations. Additionally, 8 gas sensors present different responses due to their different sensing abilities for specific gas species. E-nose sensor array forms different responding pattern for various samples with
different qualities. E-nose response to sample is reported as conductivity value (V). All sensors’ initiative responses to sample are close to zero. Sensors’ responses gradually increase and finally reach their stable values. S4, S1 and S8 have big response values, and their final stable values are about 0.33, 0.19 and 0.14 V, respectively. S6 has a response value about 0.09 V, while S5, S7 and S3 show similar response about 0.07 V. S2 hardly presents responses.

PCA results about 3 grains are displayed in Figure 1(b–c). From Figure 1(b), the first principal components (PC1) and the second principal components (PC2) capture data variance of 94.37%. However, sample 2, 3 and 5 overlap with each other. So, the 2-dimension PCA can not totally discriminate these samples. From Figure 1(c), the first 3 principal components (PC1, PC2 and PC3) capture data variance of 95.63%. So, it can be considered that it incorporates sample’s most original data. But there are still many overlapping samples. As a result, PCA method cannot discriminate all grain samples.

BP network analysis result
In BP network analysis, training to identify was performed toward 64 groups of eigen parameters at first. Then it was utilized to classify new testing samples of 48 groups. BP network recognition results toward testing samples are displayed in Table 1. It is clear that BP network exhibits a general satisfying classification result toward training samples, while it shows a poor result toward testing samples by neither changing hidden units nor improving weight arithmetic. Results indicate that BP network can not recognize all grain samples.

PNN analysis result
In PNN analysis, distribution density (spread) plays an important role. When spread value is close to zero, PNN becomes a neighbor classifier. When spread value increases, nearby design vector should be considered. Before using PNN to classify, spread value is supposed to be determined. PNN recognition results to training samples at different spread values are displayed in Table 2. It is clear that the recognition rate is the highest at the spread between 0.003 and 0.005. So, spread value of 0.005 was adopted for discriminating testing samples. PNN recognition results to testing samples at the spread of 0.005 are displayed in Table 3. PNN method shows an predicting accuracy of 100% toward the first 3 samples, while presents an accuracy of 87.5% toward early mouldy oat, fresh red bean and early mouldy red bean. Red bean presents hard texture and appearance. It is not easy for red bean to become deteriorated. To summary, PNN exhibits a satisfying discrimination toward these samples with different species and quality.

In this paper, e-nose responses to samples with different mouldy status were measured. PCA, BP network and PNN were used to analyze e-nose data. Results indicated e-nose presented sensitive responses to rice, oat, and red bean fresh and early mouldy samples. The developed method based on e-nose and PNN technique showed satisfying predicting accuracy of 93.75%.

Table 2. BP network recognition results

| Sample No. | Training recognition rate (%) | Sample recognition rate (%) |
|------------|-------------------------------|----------------------------|
| 1          | 87.50                         | 88.50                      |
| 2          | 88.50                         | 83.50                      |
| 3          | 87.50                         | 88.50                      |
| 4          | 88.50                         | 82.50                      |
| 5          | 100.00                        | 65.00                      |
| 6          | 100.00                        | 72.5                       |

Table 3. PNN recognition rate at different spread values

| Sample No. | Spread 1 | Spread 2 | Spread 3 | Spread 4 | Spread 5 | Spread 6 |
|------------|----------|----------|----------|----------|----------|----------|
| 1          | 100.0    | 100.0    | 87.5     | 75.0     | 100.0    | 87.5     |
| 2          | 75.0     | 100.0    | 100.0    | 87.5     | 100.0    | 87.5     |
| 3          | 100.0    | 100.0    | 100.0    | 100.0    | 100.0    | 100.0    |
| 4          | 100.0    | 100.0    | 100.0    | 100.0    | 100.0    | 100.0    |
| 5          | 100.0    | 100.0    | 100.0    | 100.0    | 100.0    | 100.0    |
| 6          | 100.0    | 100.0    | 100.0    | 100.0    | 100.0    | 100.0    |
Materials and Methods

Samples and preparations
Rice, oat, and red bean samples were purchased in an agriculture market in Hangzhou. The samples had 2 types, that was, fresh and early mouldy. Each sample of 50 g was stored at a biochemical incubator under different storage conditions. Storage condition of fresh samples was set at a temperature of 10°C and a relative humidity of 15%. While the storage condition of early mouldy samples was set at a temperature of 25°C and a relative humidity of 55%. All samples were numbered from 1 to 6, corresponding to fresh rice, early mouldy rice, fresh oat, early mouldy oat, fresh red bean and early mouldy red bean, respectively. After 5 days' storage, these samples were taken for e-nose detection under room temperature.

E-nose apparatus and measurement
A portable e-nose was used for the experiment. The diagram structure of e-nose is displayed in Figure 2. It consists of 3 main parts: data acquisition, modulating and transmitting unit (U1); sensor array and chamber unit (U2); power and gas supply unit (U3). It consists of 8 M.O.S gas sensors with different sensitive species, which is shown in Table 4.

In e-nose measurement, washing pump and valve 2 were turned on first. All sensors were purged by zero gas. Zero gas was obtained by filtering air with active carbon material. When sensors' responses returned to the baseline, washing pump and valve 2 were shut off. Testing sample was placed into a clean vial and sealed with sealing membrane. Then sampling pump and valve 1 were turned on. The gases in sample's headspace were inhaled into gas sensor chambers by sampling pump at a flux speed of 400 mL/min for 40 s. E-nose measurement interval was 0.05 s. E-nose real-time responses to samples were recorded. When measurement was over, gas sensors were recovered by zero gas at a flux speed of 1000 mL/min for 600 s, waiting for the next measurement.

Data analyses
Data pretreatments
To exclude some errors from sensor itself and external conditions, removing basis and smoothing filtration were performed to e-nose initiative data. And normalization pretreatment was also performed using following formula:

$$X'_{i} = (X_{i} - X_{i}^{min}) / (X_{i}^{max} - X_{i}^{min})$$

Where, $X_{i}$ represent $j$ sensor's instantaneous responding value. $X_{i}^{max}$, $X_{i}^{min}$ represent sensor's maximal instantaneous responding value and minimal instantaneous responding value in a complete period, respectively.

After normalization pretreatment, all sensors' instantaneous responding value ($X'_{i}$) ranges between 0 and 1, and maintains in the same level. Thus, it provides with useful data for the training of neural network and improves data's effectiveness.

Feature extractions
In e-nose data processing, feature extractions were performed to extract useful message for representing sample's whole message. Taking Figure 1(a) for example, each sensor shows their individual response with the increase of sampling time. And each curve's initiative value, peak value, mean value, stable value, the value at 10 s, 30 s, 35 s and 40 s were chosen as eigen values and marked as Ini, Pea, Mea, Sta, 10s, 30s, 35s and 40s, respectively. There eigen values detected by 8 sensors made up an original eigen vector quantity (R) and showed as follows:

$$\begin{array}{cccccccc}
R = \begin{bmatrix}
\text{Ini}_1 & \text{Pea}_1 & \text{Mea}_1 & \text{Sta}_1 & 10s_1 & 30s_1 & 35s_1 & 45s_1 \\
\text{Ini}_2 & \text{Pea}_2 & \text{Mea}_2 & \text{Sta}_2 & 10s_2 & 30s_2 & 35s_2 & 45s_2 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\text{Ini}_8 & \text{Pea}_8 & \text{Mea}_8 & \text{Sta}_8 & 10s_8 & 30s_8 & 35s_8 & 45s_8 \\
\end{bmatrix}
\end{array}$$

PCA
During the mathematical transformation, we should make sure the total variance of all variables remain unchanged and the variable that has maximum variance becomes the first variate (or the first PC). Similarly, the variable whose variance just behind the maximum should be the second variate (or the second PC), and make sure it has no relativity with the first variate and other variates. And by this analogy, the number of the variates is same.
as PCs, and there are no relativity between all variates. So PCA method is a useful tool for reduced data. \cite{15,16,22,23} We can show these variates on the diagram and find out the relationship. Therefore, PCA is also a method of equivalent substitute. \cite{24}

**BP network**

In the classification recognition of BP network, 8-dimension input and 8-dimension output were built. And the desired output form was \( (0, 1, 0, 0, 0, 0, 0, 0) \), where 1 represent the data belonged to this class, while 0 represent not. Hidden layer of S type simulating function was used in BP network, and neuron’s number was 16. Optimized Levenberg-Marquardt arithmetic was applied in network training function, and network’s maximal training step length was set at 5000 with an error of 0.001.

**PNN**

The structure of PNN is displayed in Figure 3. It consists of 3 main parts including input layer, radial layer and competitive layer. R is the dimension number of inputted sample vector, Q is the dimension number of inputted target vector, K is a classification number of inputted sample. The input layer delivers the target vector to the first radial base neuron, whose hidden neuron number is Q. Radial base was used to calculate the distance \(||\text{dist}||\) between the inputted sample vector and the inputted target vector. Non-linear reflection was realized via Gaussian kernel function as follows: \cite{25}

\[
p(x) = \exp\left(-\frac{||x - \omega||^2}{2\sigma^2}\right)
\]

\(k\) is a classification number from 1 to K. \(j\) is a number of training samples from 1 to R. \(\omega\) is a weight vector.

Different classifier can be obtained by choosing different variance. If \(\sigma = \infty\), it gets close to a linear classifier. If \(\sigma = 0\), it gets close to a neighbor classifier. The output of radial base is a distance vector, as an input of the second competitive layer. Neuron number of the competitive layer is equal to the classification number \(k\) of inputted sample. Each neuron corresponds to a data category of training sample. After accepting the distance vector as an inputted vector by the competitive layer, each pattern’s frequency is calculated and recorded. Then the element with the maximal frequency is outputted as 1, which means it belongs to this class; if not, it is considered as other pattern.

**Conclusions**

In this paper, a rapid non-destructive method to identify early mouldy grain using e-nose was studied. E-nose responses to grains including rice, oat and red bean with 2 types (fresh and early mouldy) were detected and analyzed by PCA, BP network and PNN. Results indicated that e-nose sensitively responded to grain samples. Either two-dimension or 3-dimension PCA, and BP methods could not totally discriminate all samples. PNN successfully discriminated grain samples with an accuracy of 93.75%. The proposed method is effective for predicting early mouldy grain with some advantages including rapid response, high accuracy, low cost, etc.

**Disclosure of Potential Conflicts of Interest**

No potential conflicts of interest were disclosed.

**Funding**

This work is financially supported by National Natural Science Foundation (No. 81000645), and project supported by Key Laboratory of Forestry Intelligent Monitoring and Information Technology of Zhejiang Province.

---

### References

1. Roberto P, Adrian A, Martinelli E, Corrado DN, Arrullo DA, Maria GDE, et al. Detection of fungal contamination of cereal grain samples by an electronic nose. Sensors and Actuators B 2006; 119:425-30; http://dx.doi.org/10.1016/j.snb.2005.12.047
2. Magan N, Evans P. Volatiles as an indicator of fungal activity and differentiation between species, and the potential use of electronic nose technology for early detection of grain spoilage. J Stored Products Res 2000; 36:319-40; PMID:10880811; http://dx.doi.org/10.1016/S0260-8294(00)00057-0
3. Tsaia W-T, Mason JI, Woloshuk CP. Effect of three stored-grain fungi on the development of typhacea stercorae. J Stored Products Res 2007; 43:129-33; http://dx.doi.org/10.1016/j.japr.2006.01.001
4. Perkowski J, Buko M, Chmielowski J, Góral T, Tyra- kowska B. Content of trichodiene and analysis of fungal volatiles (electronic nose) in wheat and triticale grain naturally infected and inoculated with fusarium culmorum. Intl J Food Microbiol 2008; 126:127-34; PMID:18589811; http://dx.doi.org/10.1016/j.ijfoodmicro.2008.05.028
5. Kaya-Celiker H, Mallikarjunan PK, Schmale III D, Christie ME. Discrimination of mouldy peanuts with reference to aflatoxin using FTIR-ATR system. Food Control 2014; 44:64-71; http://dx.doi.org/10.1016/j.foodcont.2014.03.045
6. DeFilippi BG, Juan WS, Valdés H, Moya-León MA, Infante R, Campos-Vargas R. The aroma development during storage of castlebrite apricots as evaluated by gas chromatography, electronic nose, and sensory analysis. Postharvest Biol Technol 2009; 51:212-19; http://dx.doi.org/10.1016/j.postharvbio.2008.08.008
7. Laureati M, Buratti S, Bassoli A, Borgonovo G, Pagliarini E. Discrimination and characterisation of three cultivars of perilla frutescens by means of sensory descriptors and electronic nose and tongue analysis. Food Res Intl 2010; 43:959-64; http://dx.doi.org/10.1016/j.foodres.2010.01.024
8. Gutierrez-Martinez P, Schorr-Galindo S, Ragazzo-Sanchez JA. Discrimination of eight varieties of apricot...
(prunus armeniaca) by electronic nose, LLE and SPME using GC-MS and multivariate analysis. Sensors Actuators B 2007; 125:415-21; http://dx.doi.org/10.1016/j.snb.2007.02.035
9. Russo M, di Sarro R, Cefaly V, Carabeta S, Serra D, Fuda S. Non-destructive flavour evaluation of red onion (Allium cepa L.) ecotypes: an electronic-nose-based approach. Food Chem 2013; 141:896-99; PMID:23796064; http://dx.doi.org/10.1016/j.foodchem.2013.03.052
10. Han FK, Huang XY, Teye E, Gu FF, Gu HY. Nondestructive detection of fish freshness during its preservation by combining electronic nose and electronic tongue techniques in conjunction with chemometric analysis. Analytical Methods 2014; 6:529-36; http://dx.doi.org/10.1039/C3AY41579A
11. Huang L, Zhao JW, Chen QS, Zhang YH. Nondestructive measurement of total volatile basic nitrogen (TVB-N) in pork meat by integrating near infrared spectroscopy, computer vision and electronic nose techniques. Food Chem 2014; 145:228-36; PMID:24128472; http://dx.doi.org/10.1016/j.foodchem.2013.06.073
12. Hong XZ, Wang J. Detection of adulteration in cherry tomato juices based on electronic nose and tongue: comparison of different data fusion approaches. J Food Eng 2014; 126:89-97; http://dx.doi.org/10.1016/j.jfoodeng.2013.11.008
13. Zhang YN, Askim JR, Zhong WX, Orléanc P, Szulcik KS. Identification of pathogenic fungi with an optronic/electronic nose. Analyst 2014; 139:1922-8; PMID:24570999; http://dx.doi.org/10.1039/c3ana02112b
14. Balasubramanian S, Panigrahi S, Kortapalli B, Wolf-Hall CE. Evaluation of an artificial olfactory system for grain quality discrimination. JWTT-Food Science Technol 2007; 40:1815-25; http://dx.doi.org/10.1016/j.jwtt.2006.12.016
15. Yang CY, Wu TY. Diagnostics of gear deterioration using EEMD approach and PCA process. Measurement 2015; 61:75-87; http://dx.doi.org/10.1016/j.measurement.2014.10.026
16. Xu XZ, Xie L, Wang SQ. Multimode process monitoring with PCA mixture model. Computers and Electrical Engineering 2014; 40:2101-12; http://dx.doi.org/10.1016/j.compeleceng.2014.08.002
17. Jing GL, Du WT, Guo YY. Studies on prediction of separation percent in electrodialysis process via BP neural networks and improved BP algorithms. Desalination 2012; 291:78-93; http://dx.doi.org/10.1016/j.desal.2012.02.002
18. Zhang YX, Gao XD, Seiji K. Weld appearance prediction with BP neural network improved by genetic algorithm during disk laser welding. J Manufacturing Systems 2015; 34:53-9; http://dx.doi.org/10.1016/j.jmsy.2014.10.005
19. Delivopoulos E, Theocharis JB. A modified PNN algorithm with optimal PD modeling using the orthogonal least squares method. Information Sci 2004; 168:133-70; http://dx.doi.org/10.1016/j.ins.2004.02.001
20. Fatrah MA, Ren FL, GA, MR, FFNN, PNN and GMM based models for automatic text summarization. Computer Speech and Language 2009; 23:126-44; http://dx.doi.org/10.1016/j.csl.2008.04.002
21. Li YL, Ford W, Xiang K, Land W, Congdon R, Sadik O. Development of a complex adaptive PNN system for the rapid detection of Esoli. Proc Comp Sci 2013; 20:342-47; http://dx.doi.org/10.1016/j.procs.2013.09.283
22. Lan CH, Huang YL, Ho SH, Peng CY. Volatile organic compound identification and characterization by PCA and mapping at a high-technology science park. Environmental Pollution 2014; 193:156-64; PMID:25025736; http://dx.doi.org/10.1016/j.envpol.2014.06.014
23. Singh H, Raj VB, Kumar J, Mittal U, Mishra M, Nimal AT, et al. Metal oxide SAW E-nose employing PCA and ANN for the identification of binary mixture of DMMP and methanol. Sensors Actuators B 2014; 200:147-56; http://dx.doi.org/10.1016/j.snb.2014.04.065
24. Godoy JL, Vega JR, Marchetti JL. Relationships between PCA and PLS-regression. Chemometrics and Intelligent Lab System 2014; 130:182-91; http://dx.doi.org/10.1016/j.chemolab.2013.11.008
25. Yang B, Li XS, Luo J. A novel multi-human location method for distributed binary pyroelectric infrared sensor tracking system: region partition using PNN and bearing-crossing location. Infrared Physic Technol 2015; 68:35-43; http://dx.doi.org/10.1016/j.infrared.2014.10.006