Computers and Electronics in Agriculture 157 (2019) 305–321

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Noise filtering framework for electronic nose signals: An application for beef quality monitoring

Dedy Rahman Wijaya[a,b,1,2,*], Riyanarto Sarno[a,1], Enny Zulaika[c,1]

[a] Department of Informatics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
[b] School of Applied Science, Telkom University, Bandung, Indonesia
[c] Department of Biology, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

ARTICLE INFO

Keywords:
Electronic nose
Noise filtering
Beef quality
Classification
Regression

ABSTRACT

Beef is one of the most popular and widely consumed foodstuffs in the world. Nevertheless, it can easily decay if not properly treated during distribution and storage. The consumption of low quality beef causes a serious health hazard. The electronic nose (e-nose) is a rapid and low-cost instrument for beef quality classification. Hence, the development of a mobile e-nose for online meat quality monitoring is appealing. In the last few years, e-noses have been used to classify different grades of beef and to predict the number of the microbial population in beef samples. Several methods are used to deal with these classification and regression problems. Especially in multiclass beef classification and regression, signals contaminated with noise can significantly degrade the performance of the pattern recognition module. Therefore, the presence of internal and external noise in e-nose signals is a major challenge in beef quality monitoring. In this study, a noise filtering framework based on a fine-tuned discrete wavelet transform (DWT) was developed to handle noisy signals generated by an e-nose sensor array. To the best of our knowledge this is the first time the problem of e-nose signal noise in beef quality classification is tackled. The proposed framework was integrated and tested on several machine learning algorithms that were used in previous studies, i.e. k-nearest neighbor (k-NN), support vector machine (SVM), quadratic discriminant analysis (QDA), artificial neural network (ANN), and adaptive neuro fuzzy inference system (ANFIS). Furthermore, the effect of noise filtering was investigated in the classification with two, three, and four classes of beef. The effect of noise filtering was also observed in regression tasks to predict the size of microbial population in beef samples. The experimental results showed that the proposed framework provides a significant improvement in multiclass classification and regression tasks.

1. Introduction

Animal-based protein (ABP) consumption rose worldwide from 23.1 kg/person/year to 42.20 kg/person/year from 1961 to 2011 (Sans and Combris, 2015). The Food and Agriculture Organization of the United Nations estimates that beef will still be a popular ABP source globally in 2050. Their report demonstrates that the total demand will increase in developed countries as well as in developing countries (Alexandratos and Bruinsma, 2012). Meanwhile, beef is an ideal medium for the growth of pathogenic microorganisms. The consumption of infected meat can lead to serious health problems. Therefore, it is indispensable to ensure the quality of meat. Meat quality is defined as a set of properties that together identify what we appreciate about meat when we use it as a raw material to be processed into meat products. In other words, meat quality refers to attributes that determine the feasibility of meat consumption, fresh or stored without deteriorating (Elmasry et al., 2012; Purslow, 2017). If not handled properly, meat is a highly perishable food. For instance, the quality of beef stored at room temperature in open air will quickly degrade. In addition, meat quality degradation can also occur during distribution. Apart from intrinsic biotic factors, it can be affected by external factors, including temperature, meat chill chain, and transportation (Nychas et al., 2008).

Several microbiological and chemical methods have been developed for meat quality assessment (e.g. sensory panels, total count of bacteria, total volatile basic nitrogen (TVB-N), and gas chromatography) but most of them need special skills, laborious, and time-consuming (Wojnowski et al., 2017). The primary standard of beef freshness is the total count of bacteria, but its measurement is time-consuming, especially in relation to sample preparation and period of incubation. For decades, e-noses have been projected as a method for rapid detection of

* Corresponding author at: School of Applied Science, Telkom University, Jalan Telekomunikasi, Terusan Buah Batu, 40257 Bandung, West Java, Indonesia.
E-mail addresses: dedyrw@tass.telkomuniversity.ac.id (D.R. Wijaya), riyanarto@if.its.ac.id (R. Sarno), 1.enny@bio.its.ac.id (E. Zulaika).
[1] Jalan Raya ITS, Keputih, Sukolilo, Surabaya City, East Java 60111, Indonesia.
[2] Jalan Telekomunikasi, Terusan Buah Batu, Bandung City, West Java 40257, Indonesia.

https://doi.org/10.1016/j.compag.2019.01.001
Received 21 January 2018; Received in revised form 2 December 2018; Accepted 1 January 2019

microbial spoilage of muscle foods (Ellis and Goodacre, 2001). Until now, significant progress has been made by innovations in smart packaging (SP) systems. SP must meet user requirements, such as providing a rapid, accurate, cost-effective, and user-friendly system (Ahmed et al., 2018). In addition, the integration of smart sensors in consumer electronics such as a refrigerators is promising for food quality monitoring in the Internet of Things era (Wijaya and Sarno, 2015). In the last few years, several studies have been conducted to develop devices and methods to detect beef spoilage. Fourier transform infrared (FTIR) spectroscopy and e-noses are among the most widely used analytical instruments for beef quality detection (Argyri et al., 2010; Balasubramanian et al., 2009, 2004; El Barbri et al., 2008; Kodogiannis, 2017; Kodogiannis et al., 2015, 2014; Kodogiannis and Alshejari, 2016, 2014; Mohareb et al., 2016; Najam ul Hasan et al., 2012; Panigrahi et al., 2006; Papadopoulou et al., 2011, 2013; Zaragozá et al., 2014). The overall performance of e-noses is not too much different compared to FTIR spectroscopy but e-nose technology is cheaper, simpler, faster, and more suitable for real-time and online analysis. Therefore, the development of the e-nose as an apparatus for beef quality monitoring is relevant.

However, beside the aforementioned advantages, a major drawback of e-noses is the instability of the sensor response. This is caused by changes in ambient conditions (temperature and humidity), ambient air variations, sensor aging, and other physical/chemical processes. In the e-nose community, the term *sensor drift* is often used to refer to variation in the sensor response in identical measurement conditions. From the perspective of pattern recognition, the terms *sensor drift* and *noise* refer to the same phenomenon because they both lead to inconsistent data, causing performance degradation. Henceforth, we use the term *noise* to refer to instability of the sensor response. In e-nose signal processing, noise handling is necessary. Several studies have reported that e-nose signals can be considerably contaminated by noise in severe conditions (Tian et al., 2005; Wijaya et al., 2016a). The existence of noise in e-nose signals is inevitable, especially in continuous monitoring. It is a major challenge in computation for classification and regression tasks. Referring to the explanation above, there were several motivations to conduct this study:

1. The application of e-noses for beef quality monitoring faces several challenges. First, the beef spoilage process produces water vapor, which changes the humidity levels in the sample chamber. These changes in humidity level can make the sensor response unstable. Moreover, the sensors are susceptible to poisoning due to exposure to ethanol and sulphur compounds generated by protein decomposition (Schaller et al., 1998). Another problem is the possibility of sensor saturation due to exposure to various gases for a long period of time. In addition, a recent study found that ambient air variations largely influence MOS gas sensor stability (Kiselev et al., 2018). This condition continues during beef storage, so the existence of noise in e-nose signals is ineluctable.
2. Basically, the probability of success in classification tasks is *1/m* where *m* is the number of classes. In other words, having a large number of classes in multiclass classification and having a continuous output in regression are intricate problems. Performance degradation occurs when the number of classes increases so the classifier produces low precision and recall in multiclass classification tasks (Wijaya et al., 2017b). Furthermore, signals contaminated with noise make these tasks more complicated.
3. Hitherto, no study has addressed or focused on the effect of noise filtering in e-nose signals for beef quality monitoring. On the other hand, several studies have demonstrated the performance of machine learning algorithms with data that were contaminated by artificial noise (Nettleton et al., 2010; Sáez et al., 2014; Zhu and Wu, 2004). Instead of using artificial noise, this study used naturally generated noise caused by uncontrolled ambient conditions in beef quality monitoring.

In this study, a noise filtering framework was developed to deal with e-nose signals contaminated with noise in beef quality monitoring. As the baseline, methods from previous works were used, i.e. k-NN, SVM, ANN, ANFIS, and QDA, and their performances were compared when using the proposed framework. The effect of noise filtering was investigated in binary and multiclass (three and four classes) classification using the meat quality standard issued by the Agricultural and Resource Management Council of Australia and New Zealand (CSIRO Food and Nutritional Sciences, 2003). In addition, the effect of noise filtering was also observed in regression tasks. Neural network regression (NNR), support vector regression (SVR), and ANFIS regression were used to predict the number of microbial population in beef samples.

The remainder of this paper is structured as follows: Section 2 discusses related works on beef quality classification, wavelet transform, and the impact of noise on the performance of machine learning algorithms. Section 3 specifies the materials and methods used in the experiment, including the experimental set-up, the proposed noise filtering framework, and a brief explanation of the machine learning algorithms that were used in previous works as well as their parameter settings. Section 4 explains the results of the experiment, the condition of the sample chamber, the result of signal reconstruction, and the advantages of using noise filtering in classification and regression tasks. Finally, Section 5 draws the conclusion of this work.

2. Related works

The first study on the use of e-noses for beef quality assessment was reported in 2004 (Balasubramanian et al., 2004). A large number of sensors (28 gas sensors) from a commercial e-nose device were used to classify fresh and spoiled beef. This study showed that the QDA classifier could solve this binary classification problem with 98.48% accuracy. Further research was performed using a custom e-nose with only seven gas sensors accompanied by a temperature and humidity sensor (Panigrahi et al., 2006). This study showed that QDA still had the highest classification accuracy (96%). Moreover, SVM could successfully detect red meat spoilage (El Barbri et al., 2008); it achieved 98.81% and 96.43% classification accuracy for beef and sheep meat, respectively. Radial basis function neural network (RBFNN) has been utilized to detect the freshness of beef, with 92.2% accuracy (Balasubramanian et al., 2009). In another study, a wireless e-nose was used to recognize the freshness of beef and fish using k-NN classifiers (Najam ul Hasan et al., 2012). Multiclass classification has been performed since 2013 (Papadopoulou et al., 2013). In this study, three sensory classes were distinguished ('fresh', 'semi-fresh', and 'spoiled'). Performing multiclass classification is necessary because it is important to know when the beef has reached the semi-fresh stage before it spoils completely. SVM has been reported to be able to differentiate three classes of beef with more than 85% accuracy (Mohareb et al., 2016; Papadopoulou et al., 2013). Neuro-fuzzy approaches have been reported to be able to distinguish the quality of beef fillet with 94.28% accuracy (Kodogiannis and Alshejari, 2016) and 95.71% accuracy (Kodogiannis, 2017). Related to sensory class determination, Ellis and Goodacre have stated that the odor was described as 'dairy/buttery/fatty/cheesy' at 10^7 cfu cm^{-2} and as 'spoiled' when the microbial population had reached 10^8 cfu cm^{-2} (Ellis and Goodacre, 2001). Moreover, slime and off-odors are correlated with the population of pseudomonads to the arbitrary level of $10^{7\text{-}8}$ cfu/g (Nychas et al., 2008). In the e-nose community, various ranges are used to determine beef sensory classes (Balasubramanian et al., 2009, 2004; El Barbri et al., 2008; Kodogiannis, 2017; Panigrahi et al., 2006; Papadopoulou et al., 2013). A more rigorous standard is used for process monitoring in the meat industry, as shown in Table 3. Four sensory classes are distinguished and meat is considered safe for consumption if TVC $< 5 \log_{10}$ cfu/g.

In the field of agriculture, metal-oxide semiconductor (MOS) gas sensors are used for various purposes, such as detection of tea plant damage caused by pests (Sun et al., 2017), monitoring of soil gas

Table 1
Comparison of the present study with previous studies.

Comparison aspect	Present study	Previous studies
Noise handling	Considered	Not considered (Balasubramanian et al., 2009, 2004; El Barbri et al., 2008; Kodogiannis, 2017; Kodogiannis and Alshejari, 2016; Mohareb et al., 2016; Najam ul Hasan et al., 2012; Panigrahi et al., 2006; Papadopoulou et al., 2013)
Standard of beef quality	Meat industry guidelines published by Meat Standards Committee of ARMCANZ (Agricultural and Resource Management Council of Australia and New Zealand) (CSIRO Food and Nutritional Sciences, 2003; PrimeSafe, 2002)	Unclear standard (Balasubramanian et al., 2009, 2004; El Barbri et al., 2008; Kodogiannis, 2017; Kodogiannis and Alshejari, 2016; Mohareb et al., 2016; Najam ul Hasan et al., 2012; Panigrahi et al., 2006; Papadopoulou et al., 2013)
Number of classes	Four classes (regrouped into two and three classes for comparison)	Two classes (Balasubramanian et al., 2009, 2004; El Barbri et al., 2008; Najam ul Hasan et al., 2012; Panigrahi et al., 2006) and three classes (Kodogiannis, 2017; Kodogiannis and Alshejari, 2016; Mohareb et al., 2016; Papadopoulou et al., 2013)
Source of noise	Generated by uncontrolled environment	Artificial noise (Nettleton et al., 2010; Sáez et al., 2016, 2014; Zhu and Wu, 2004)
Type of noise handled based on time of occurrence	Short-term noise (unstable sensor response)	Long-term noise (gradual changes of sensor response) (Liu et al., 2014; Vergara et al., 2012; Yan and Zhang, 2016; Zhang et al., 2017; Zhang and Zhang, 2015)
Investigation of different number of classes	Considered	Not considered (Nettleton et al., 2010; Sáez et al., 2016, 2014; Zhu and Wu, 2004)

emission (Pineda and Pérez, 2017), classifying beef and pork (Wijaya et al., 2017a), differentiating fish species (Güney and Atasoy, 2015), and detecting wood decay (Baietto et al., 2013). In addition, MOS gas sensors are low-cost and less sensitive to moisture than cold gas sensors (Balasubramanian et al., 2016). Nevertheless, their signal is still susceptible to contamination with noise. The noise that exists in the gas sensor output originates from both internal and external sources. Moreover, external noise can arise in extreme conditions, such as fluctuating ambient humidity and exposure to ethanol and sulphur compounds in long-term beef quality monitoring. In a previous study, the noise magnitude reached 20% in severe conditions (Tian et al., 2005). Commonly, in the machine olfaction community and the sensory field, the term *sensor drift* is used to refer to unpredictable temporal variations of the sensor response when a sensor array is used in identical conditions. This can be caused by several factors, such as sensor aging, sensor poisoning, and environmental changes. Several studies have been conducted to address sensor drift. A dataset from three years of observation has been used to investigate the sensor drift phenomenon (Vergara et al., 2012). The authors used SVM as an ensemble classifier, which was trained at different points of time. The combination kernel function has been used in SVM (Liu et al., 2014). A combination of sparse autoencoder (SAE) and deep belief network (DBN) has been investigated (Liu et al., 2015). The development of extreme learning machine (ELM) was proposed for sensor drift compensation (Zhang and Zhang, 2015). A semi-supervised approach can achieve favorable performance in dealing with sensor drift (Liu et al., 2014). Another study not only concerned with sensor drift over time but also investigated different responses from a multi-device dataset (Yan and Zhang, 2016). Furthermore, a domain regularized component analysis (DRCA) method has been introduced to avoid the difference in probability distribution between source and target dataset causing machine learning failure (Zhang et al., 2017). Generally speaking, all of these studies used a long-term dataset with tightly controlled operating conditions and then developed drift compensation methods. However, an unstable sensor response does not only occur in long-term utilization (long-term noise) but also arises when a gas sensor is operated in an uncontrolled environment in many real measurements.

Several studies have investigated the impact of noise on the performance of machine learning algorithms. From a machine learning point of view, the type of noise can be differentiated into class noise and attribute noise (Zhu and Wu, 2004). Class noise is caused by misclassification of instances into a wrong class label or the same instance classified into different classes. Attribute noise is caused by erroneous attribute values and incomplete attribute values. Unfortunately, dealing with attribute noise is more complicated because we cannot intuitively change an attribute value rather than re-labeling an instance to tackle class noise. The authors state that the highest classification accuracy can be obtained by using both clean training data and clean testing data. A study on the impact of noise on the precision of four supervised learning algorithms (Naïve Bayes, C4.5, k-Nearest-Neighbor, and Support Vector Machine) in binary classification tasks has been reported (Nettleton et al., 2010). The experimental results showed that noise in the training dataset caused the worst performance for all methods. Another study proposed the use of a decomposition strategy for multiclass classification to deal with noisy data (Sáez et al., 2014). The results showed that one-vs-one decomposition yields higher accuracy and is more robust for multiclass classification of noisy datasets. In another study, noise filtering methods to deal with class noise in medical data have been discussed (Sáez et al., 2016). This study confirmed that SVM is a robust classifier even without noise filtering. Nevertheless, noise filtering is still needed when the noise level is high. It is also stated that the utilization of noise filtering does not always provide better performance. Hence, proper noise filtering methods are necessary. For noise filtering, the wavelet transform is a commonly used method in various areas, for example electroencephalography (Wang et al., 2011), audio processing (Ridoean et al., 2018; Sarno et al., 2018), and e-nose signal processing (Hariyanto et al., 2017; Wijaya et al., 2017c, 2016b; Zanchettin and Ludermir, 2007).

Our study focused on how to deal with noise in e-nose signals and investigated the influence of noise filtering in beef quality classification and regression tasks, which has not been addressed in previous studies. In the experiment, the noise was naturally generated as a result of changes in ambient conditions to demonstrate a real case instead of using artificial noise like other studies related to noise impact investigation (Nettleton et al., 2010; Sáez et al., 2014; Zhu and Wu, 2004). Our hypothesis was: noise filtering can present a significant improvement on beef quality classification and regression tasks because the success rate of the classifier depends not only on the learning algorithm but also on the quality of the input data. In this study, a noise filtering framework was developed based on proper adjustment of DWT. Table 1 shows a comparison of our study against previous works.

3. Materials and methods

3.1. Materials

MOS gas sensors and a temperature-humidity sensor were used in the experiment. MOS gas sensors are a type of hot sensor and are less

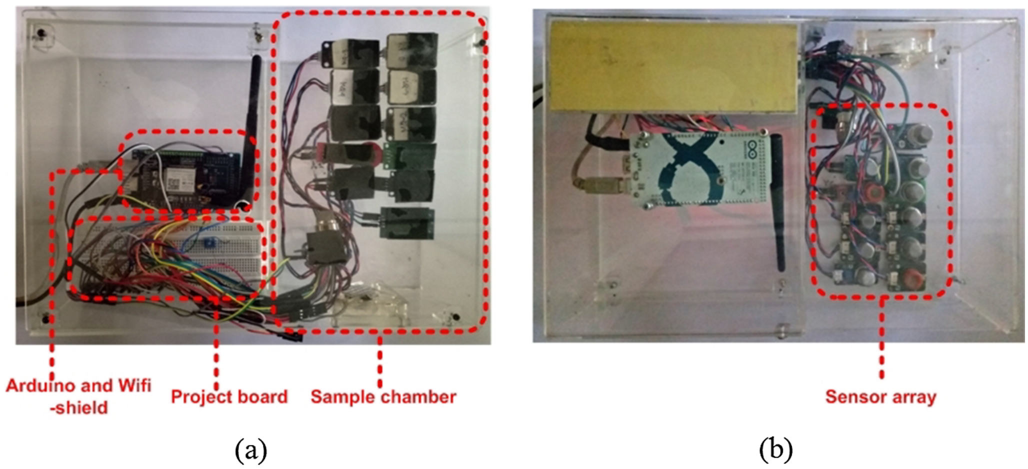

Fig. 1. Sensor box: (a) top view; (b) bottom view.

sensitive to moisture than cold sensors. Hence, they are more suitable for beef quality monitoring because of the high humidity in the sample chamber. All sensors and the WiFi-shield for the communication module were assembled on an Arduino Mega ADK microcontroller. The casing of the sensor box was made of transparent acrylic material so that the condition of the device and sample could easily be observed. The sample chamber was located on the bottom of the sensor array for faster sensor response. Fig. 1 shows the sensor box used in this experiment.

In the initial experiment, eleven sensors were used to assemble a custom e-nose sensor box. The sensor array was optimized, reducing the number of sensors to seven, in our previous work (Wijaya et al., 2016a). Hence, data produced by only seven sensors were used for the further processes. Table 2 shows the list of sensors in the optimized sensor array.

This custom e-nose was designed to detect gasses produced by mesophilic bacteria in the process of beef spoilage. Mesophilic bacteria grow optimally at 20–45 °C. Hence, the experiment was conducted at room temperature, to accelerate the process of beef decay. The data from the sensor array were recorded and sent to a server every minute for continuous beef quality monitoring. Hence, the e-nose could record the beef spoilage process precisely in every stage. It was used to monitor 500 g of fresh beef, which was placed in the sample chamber. In this experiment, we used an extra-lean beef as the sample. Fig. 2 shows a scheme of the experimental set-up for data acquisition.

The output of the gas sensors was in the form of analog sensor resistance values. When sent to the server, these analog values were converted into digital values. The size of the data from each sensor was 1 to 5 bytes. The smallest size occurred when the sensor produced a single-digit integer. The largest data size occurred for tens with two numbers behind the comma. The data from the sensor array were transferred to the WiFi-shield module via a wireless access point to the computer server during three days of observation. The raw data were received by a TCP/IP socket server on a virtual port of the computer server. Then, the data were parsed and stored by self-developed middleware in a MySQL database for further analysis. Based on this mechanism, the process of beef spoilage could be monitored in real-time and data for training and testing could be collected. The total number of data from this experiment was 4720 records. The beef sample began to spoil at 680 min, so the data were dominated by the 'spoiled' beef class (74%). The utilization of a dataset with class imbalance causes high accuracy but low recall and precision in the minority classes. We strived to make sure that no particular class had distribution greater than 50%. Therefore, we performed undersampling by taking 1400 data, from when the meat was still fresh until it was rotting, so that no class dominated the distribution over 50%. In addition, this way we could also make sure that all data were within class boundaries because they are susceptible to misclassification. The dataset consisted of 1400 records/instances from one beef sample, divided into training data and testing data. In this study, the meat standard issued by the Agricultural and Resource Management Council of Australia and New Zealand (CSIRO Food and Nutritional Sciences, 2003) was used. In this standard, the grade of meat is distinguished into four categories according to the total viable count (TVC), as shown in Table 3.

Table 2
Gas sensors used in the experiment.

Alias	Sensor	Selectivity
S1	MQ135	Carbon dioxide, alcohol, ammonia, NO_x, smoke, benzene
S2	MQ136	Hydrogen sulfide
S3	MQ2	Alcohol, hydrogen, smoke, Liquefied petroleum gas (LPG), methane, i-butane, propane
S4	MQ4	Methane
S5	MQ6	Iso-butane, propane, LPG
S6	MQ9	Carbon monoxide, methane, and propane
S7	DHT22	Temperature and humidity

The signals that were generated by the sensor array during the beef spoilage process were labeled using the procedure described below. Suppose we have an instance of feature subset X with m attributes from the sensor array given by:

$$X = \langle x_1 \quad \cdots \quad \cdots \quad x_m \rangle \in class_j. \tag{1}$$

where $class_j$ is a particular class label from the class domain. For classification, an instance can have c possible class labels ($Class$), which is expressed as follows:

$$Class = \{class_1 \;\;, \cdots \;\; class_c\},\; c \in Z^+ \tag{2}$$

where $c = 2$ means binary classification and $c > 2$ means multiclass classification. In this experiment, the effect of noise filtering on electronic nose signals was observed for various numbers of classes in a beef classification task. E-nose signals are time series data that represent the quality of the beef according to its shelf-life. In this study, the holdout method was used to divide the dataset in training data and testing data. In the i-th class, the training data for every feature subset (X_i) was about half ($\frac{1}{2}X_i$) and the remaining half were used as testing data ($X_i - \frac{1}{2}X_i$), where ($X_i \in class_i$). The data were randomly selected from each class for training and testing. The training data and the testing data were rotated in two different experiments. This was to ensure that the results of the experiment would be fair and objective. Fig. 3 illustrates the partition of the dataset in this experiment.

In accordance with previous studies, the dataset was also regrouped for classification into two and three classes. Thus, we investigated the

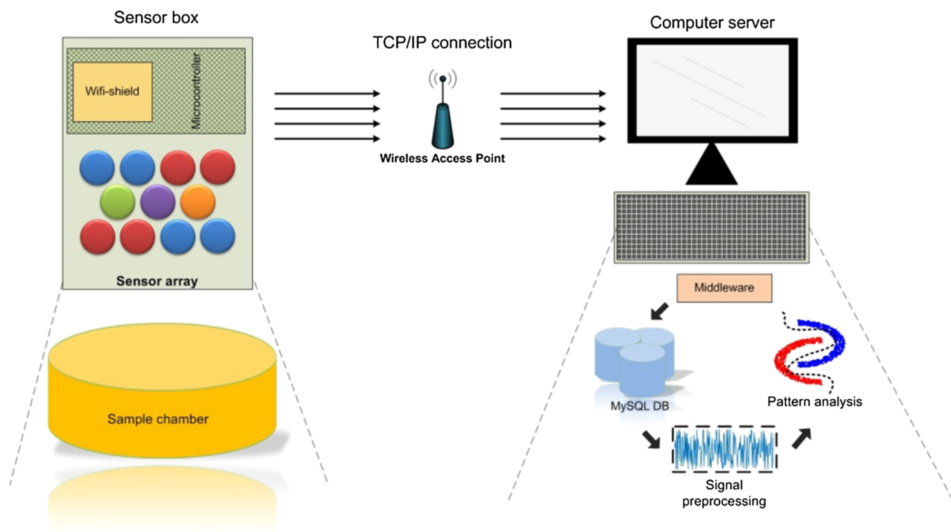

Fig. 2. The scheme of experiment.

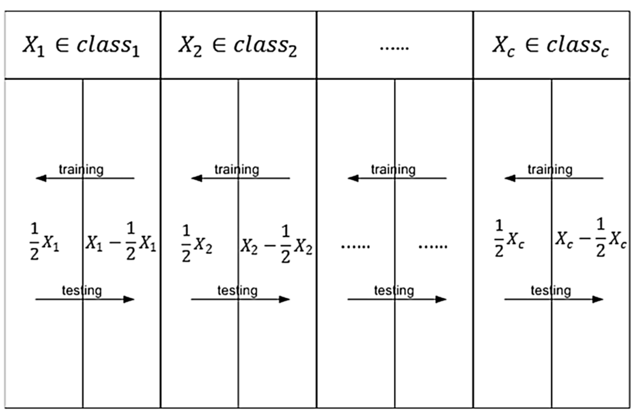

Fig. 3. Partition of the dataset.

ability of our proposed framework to improve the performances from past studies for binary and three-class classification. The ground truth data refer to the size of the microbial population in the gold standard of beef quality. In this experiment, optical density was measured by a spectrophotometer with 1000x dilution. After this, a hemocytometer was used to quantify the microbial population. The combination of the classical and the two-hour method was used as a rule of thumb (Harley and Prescott, 2002). The ground truth data were utilized to determine the class labels. The scenario of class labeling can be seen in Fig. 4.

According to Fig. 4, beef quality can be divided into 'fresh' and 'spoiled' for binary classification. The members of the 'fresh' class were obtained by grouping together the classes 'excellent' (E), 'good' (G), and 'acceptable' (A) ($\{E, G, A\} \in fresh$), while the members of the 'spoiled' (S) class were kept into the same class ($S \in spoiled$). Three classes ('fresh', 'semi-fresh', and 'spoiled') were formulated by grouping together the members of 'good' and 'acceptable' classes into 'semi-fresh' ($\{G, A\} \in semifresh$), while the remaining classes were mapped into 'fresh' and 'spoiled' ($E \in fresh$ and $S \in spoiled$).

In the regression task, several machine learning algorithms were used to predict continuous values that correspond to the size of the microbial population instead of the discrete output from the classification task. Moreover, the effect of noise filtering on the regression is more evident when a large amount of data is used. Hence, a large amount of detailed ground truth data were needed to predict the size of the microbial population in the beef sample. In the experiment, the observed number of the microbial population was based on the data generated by DMFit (www.combase.cc), a tool that has adopted the primary model of Baranyi (Baranyi and Roberts, 1994). Thus, almost the same dataset was used except for the continuous label. This dataset was used to investigate the effect of noise filtering on the regression task. The holdout method was used to split the dataset into data for training and data for testing in a balanced way. Our complete raw dataset is also available online (Wijaya et al., 2018).

3.2. Methods

3.2.1. Proposed noise filtering framework

In this subsection, we discuss the steps of e-nose signal processing. Here, noise filtering is key to improving the quality of input data. Fig. 5 shows the proposed noise-filtering framework for e-nose signals.

DWT was used to reconstruct the raw signals contaminated with noise. The DWT of signal $x(t)$ can be mathematically expressed by the following equation (Gao and Yan, 2011):

$$DWT(v, w) = \langle x(t), \psi_{v,w}(t) \rangle = \frac{1}{\sqrt{2^v}} \int_{-\infty}^{\infty} x(t)\psi * \left(\frac{t - 2^v w}{2^v}\right) dt \tag{3}$$

where v and w are the scaling parameter and the shifting parameter, respectively. The scaling parameter establishes the time and frequency resolution of the scaled mother wavelet (MWT). The scaled MWT is represented by $\psi((t - 2^v w)/2^v)$. The value of the scaling parameter is inversely proportional to the frequency. A higher value of v means a lower frequency and vice versa. Shifting parameter w moves the scaled MWT along the time axis. When performing DWT, the level of decomposition and the type of wavelet function are the two most influential parameters on the changes of the signal's structure (Wijaya et al., 2016c). Fig. 6 depicts the detailed steps of parameters adjustment for noise filtering to determine the wavelet decomposition levels and the best-suited MWTs.

The scaling parameter corresponds to the decomposition level in signal reconstruction. A lower frequency needs more decomposition. Hence, the first thing to do is to set the correct decomposition level, because an improper decomposition level leads to signal defects caused by a mismatch between the reference frequency range and the frequency characteristic of the signal. The proper decomposition level can be determined by the following rule (Gao and Yan, 2011):

$$\frac{F_{sample}}{2^{level+1}} \leq F_{char} \leq \frac{F_{sample}}{2^{level}} \tag{4}$$

where $F_{char}, F_{sample}, level$ indicate frequency characteristic, frequency sampling, and decomposition level, respectively. The advantage of the wavelet transform is the abundance of MWTs that can be used.

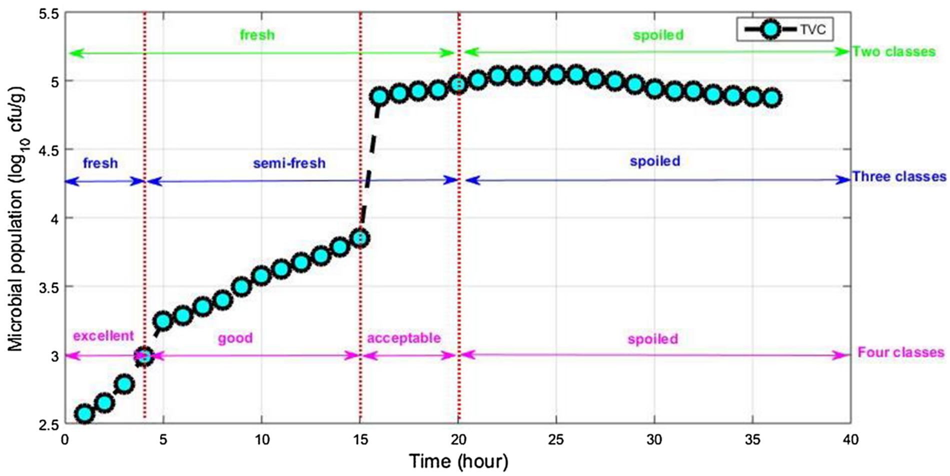

Fig. 4. Scenario of beef class labeling (two, three, and four classes).

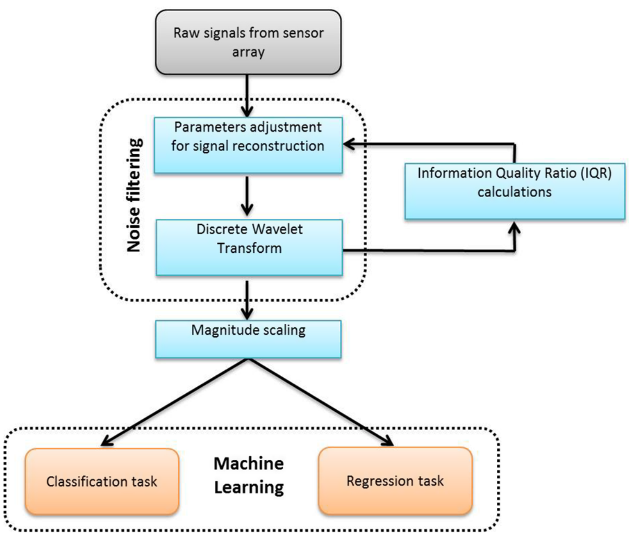

Fig. 5. The proposed noise filtering framework. The rounded rectangles represent input/output data. The blue rectangles are algorithms. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

However, this will raise the question which MWT to choose for signal reconstruction. Thus, the second problem is: which MWT is the best-suited. In the present study, the main principle of noise filtering is signal reconstruction without losing essential information. The information quality ratio (IQR) was used to find the best-suited MWT for a particular signal (Wijaya et al., 2016c). The IQR value between the original/raw signal ($x(t)$) and the reconstructed signal ($y(t)$) is given by:

$$IQR(x(t), y(t)) = \frac{\sum_{x,y} p(x, y) log_2(p(x)p(y))}{\sum_{x,y} p(x, y) log_2(p(x, y))} - 1 \quad (5)$$

x, y, $p(x)$, $p(y)$, $p(x, y)$ are the element of the original signal, the element of the reconstructed signal, the marginal probability of x, the marginal probability of y, and the joint probability of x and y, respectively. Several MWTs are compared, where the highest IQR value indicates the MWT that is best-suited for the signal. In the sensor array signal processing, the IQR values of m signals reconstructed by n different MWTs can be illustrated by the following matrix.

$$IQR_{m\times n} = \begin{bmatrix} IQR_{11} & IQR_{12} & \dots & \dots & IQR_{1n} \\ IQR_{21} & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & IQR_{m-1n} \\ IQR_{m1} & \dots & \dots & IQR_{mn-1} & IQR_{mn} \end{bmatrix}$$

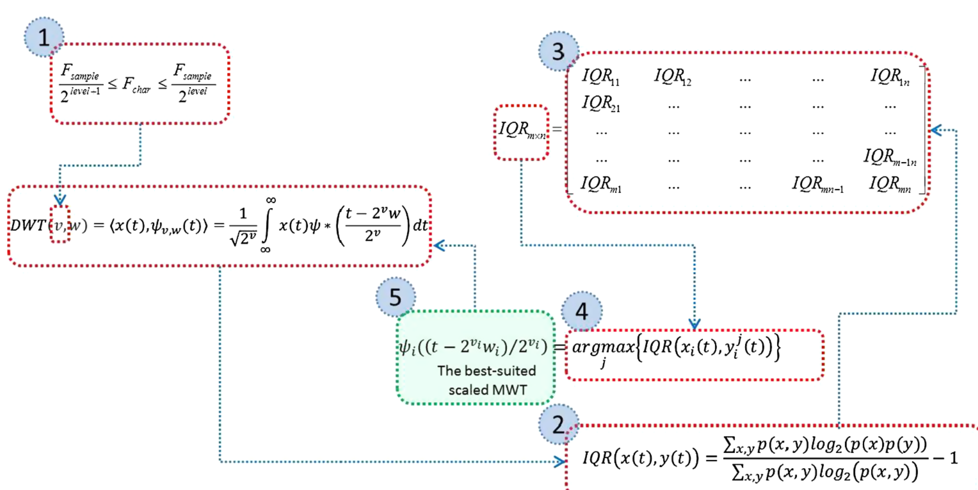

Fig. 6. Detailed steps of noise filtering: (1) set the proper wavelet decomposition levels (determine scaling parameter (v)); (2) IQR calculation between original ($x(t)$) signal and reconstructed ($y(t)$) signal; (3) build the IQR matrix ($IQR_{m\times n}$); (4) determine the best-suited scaled MWTs using the argument of maxima for each signal based on the IQR matrix; (5) signal reconstruction using the best-suited scaled MWTs.

Based on above matrix, the chosen scaled MWT for a particular signal ($x_i(t)$) is determined by the largest IQR value against reconstructed signal ($y_i^j(t)$). Where, i and j are the index of the signals and the index refers to the MWT used to reconstruct the signal, respectively. The scaled MWTs of each signal ($\psi_i((t-2^{v_i}w_i)/2^{v_i})$) are influenced by translation parameter (w_i) and scaling parameter (v_i), corresponding to the level of wavelet decomposition. Hence, the best-suited scaled MWTs for each signal can be determined according to argument of maxima of the IQR function as follows:

$$\psi_i((t-2^{v_i}w_i)/2^{v_i}) = \underset{j}{\operatorname{argmax}}\,\{IQR(x_i(t), y_i^j(t))\},\ i \in m, j \in n \tag{6}$$

Based on these adjustments, the proper noise filtering was performed to ensure the quality of the input data. The wavelet decomposition level, the best-suited MWT, and the result of signal reconstruction are explained in the next section. In this study, the magnitude of a particular value of the reconstructed signal ($y(t)$) was computed according to the following equation:

$$y' = \frac{y - \min(y(t))}{\max(y(t)) - \min(y(t))} \times (c-1) \tag{7}$$

where y' and c correspond to the new scaled value of y and the number of possible class labels, respectively. The complete pseudocode of the proposed noise filtering framework is as follows:

Algorithm 1 (*Noise filtering framework*).

```
Algorithm 1: Noise filtering framework
Input: m-Dimensional signal contaminated with noise (original signal)
Output: m-Dimensional reconstructed signal
/* Load original signal and get the length of signal */
X ← load(original_signal)
L ← length(X)
Fs ← 1024 // Set frequency sampling

/* Get the proper decomposition level based on Eq.(4) */
for i ← 1 to m do // Loop until m-Dimensional signal
/* Get frequency characteristic using Fast Fourier Transform (FFT) */
    fft_abs ← abs(fft(X[i]-mean(X[i])))
    indexmax ← argmax_{l∈L} F(fft_abs,l) // Find the index of maximum fft_abs
    f_char ← indexmax*Fs/L
    dlevel[i] ← decLevel(f_char) // Get decomposition level
end for

/* Get the best-suited MWT */
mwt ← ['haar','daubechies','symlet','coiflet','biorthogonal','meyer']
N ← length(mwt)
/* Build IQR matrix */
for i ← 1 to m do // Loop until m-Dimensional signal
    for j ← 1 to n do // Loop until n MWTs
        Y[i][j] ← waveletDenoising(X[i],mwt[j],dLevel[i])
        // Calculate IQR value between i-th original signal (X) and i-th denoised
        signal (Y) that reconstructed by j-th MWT
        IQR[i][j] ← calculateIQR(X[i], Y[i][j])
    end for
    /* Get the best i-th reconstructed signal (Y) based on maximum value of IQR*/
    k ← argmax_{j∈n} F(IQR[i],j) // Find the index of maximum IQR value
    best_Y[i] ← Y[i][k] // Get the best reconstructed signal for feature matrix
end for
NormalizerMinMaxScaler(best_Y, 0, c-1) // Magnitude scaling (Normalization)

/* Using feature matrix to fed machine learning algorithm for training and testing */
MachineLearning.Training(best_Y)
MachineLearning.Testing(best_Y)
```

3.2.2. k-Nearest Neighbor

k-NN is a type of instance-based learning that compares testing data with training data to determine their similarity, where similarity means the 'closeness' between both data sets, which is measured using a distance metric, e.g. Euclidean, city block, Chebyshev, Mahalanobis, etc. In this experiment, we set $k = 5$ and used Euclidean distance referring to a previous work (Najam ul Hasan et al., 2012).

3.2.3. Support vector Machine

SVM uses an optimal hyperplane to separate data with different classes. Several kernel functions can be used, such as a second-order polynomial or a radial basis function (RBF), to deal with the beef classification problem (El Barbri et al., 2008; Mohareb et al., 2016; Najam ul Hasan et al., 2012; Papadopoulou et al., 2013). In this experiment, multiclass SVM with RBF kernel was used to investigate the effect of noise filtering, where the value of regularization parameter (C) and gamma (γ) were set to 1 and 2, respectively. In MATLAB, the default values of C and γ are 1. We set $C = 1$ and $\gamma = 2$ because this yields more stable and better results than the default value ($\gamma = 1$) based on cross validation.

3.2.4. Quadratic discriminant analysis

In previous works, the use of QDA to differentiate fresh and spoiled beef has been reported (Balasubramanian et al., 2004; Panigrahi et al., 2006). QDA is a type of discriminant classifier akin to linear discriminant analysis (LDA). However, unlike LDA, QDA assumes that different classes have different covariance matrices (James et al., 2013). In this method, two or more classes ($class_i$, $i = 1, 2, \cdots, c$) are separated by a quadratic surface. The quadratic discriminant function is computed by:

$$\delta(x)_{class_i} = -\frac{1}{2}log\ |\Sigma_{class_i}| - \frac{1}{2}(x-\mu_{class_i})^T\Sigma_{class_i}^{-1}(x-\mu_{class_i}) + log(\pi_{class_i}) \tag{8}$$

where Σ_{class_i}, μ_{class_i}, π_{class_i} are the covariance matrix, the mean vector, and the prior probability of membership of the i-th class, respectively. The QDA classifier involves plugging estimates for Σ_{class_i}, μ_{class_i}, π_{class_i} and assigning x to the class with the highest value of $\delta(x)_{class_i}$. Hence, the class label ($class$) of x can be determined by:

$$class = \underset{class_i}{\operatorname{argmax}}\delta(x)_{class_i},\ i = 1, 2, \cdots, c \tag{9}$$

where c is the total amount of possible class labels.

3.2.5. Artificial neural network

ANN is a layered model classifier that processes data layer-by-layer to determine the predicted class label. Each layer consists of several neurons, which transform the data based on their connection weights and activation function. In this study, a multilayer perceptron with a single hidden layer was used. The size of the input, the hidden, and the output layer were 7, 24, and 4, respectively. The sigmoid activation function was used in the hidden layer, while the softmax activation function was used in the output layer. This neural network architecture

Table 3
Beef quality standard.

Class	Total viable count ($\log_{10}$ cfu/g)
Excellent (E)	< 3
Good (G)	3–4
Acceptable (A)	4–5
Spoiled (S)	< 5

*cfu/g: colony forming unit of bacteria in one gram of meat

refers to the architecture that was used in a previous work to investigate the effect of noise filtering (Najam ul Hasan et al., 2012).

3.2.6. Adaptive Neuro-Fuzzy inference system

In previous studies, the combination of an e-nose and ANFIS has been reported for satisfactory performance to classify three grades of beef quality (Kodogiannis, 2017; Kodogiannis and Alshejari, 2016). In this work, a first-order Sugeno model and a Gaussian membership function were used to construct the ANFIS.

3.2.7. Neural network regression

The architecture of NNR is the same as a neural network for classification that consists of an input layer, one hidden layer, and an output layer. The difference is only one neuron in the hidden layer with a linear activation function to accommodate the continuous output for the regression task.

3.2.8. Support vector regression

The regression problem is more intricate for signals that are contaminated by noise. In this experiment, a large value of parameter C was used to penalize any misclassified data points. In other words, the regressor is pushed harder to deal with noisy data in the regression task. Meanwhile, γ is kept low to avoid overfitting by ensuring a broader decision region. For the ε parameter, the default value was used. Epsilon support vector regression (ε-SVR) was used to predict the number of the microbial population. The RBF kernel was utilized to deal with non-linear data. The values of C, ε, and γ were set to 1000, 0.1, and 0.1, respectively.

3.2.9. ANFIS regression

The use of ANFIS for regression is similar to its application in the classification problem. The ANFIS final layer is a summation of all inputs from the previous layer, which naturally generates real numbers. Hence, we do not provide a further explanation about the use of ANFIS in the regression tasks.

In this experiment, DWT, k-NN, SVM, QDA, and ANFIS were implemented in MATLAB 2015a. DWT uses the Wavelet Toolbox. k-NN, SVM, QDA use the Statistics and Machine Learning Toolbox. ANFIS uses the Fuzzy Logic Toolbox. The neural networks were built based on open source framework Deeplearning4j (DL4J) for more complete and flexible implementation (Deeplearning4j Development Team, 2017). SVR was implemented using LIBSVM as a library for support vector machine in MATLAB 2015a (Chang and Lin, 2011).

3.2.10. Performance measures

Several metrics were used to measure the performance of the machine learning algorithms in the classification and regression tasks. For the classification problem, accuracy, precision, and recall were used to measure the performance of the classifiers. Table 4 shows a detailed explanation of these metrics. In this table, tp, tn, fp, fn denote ***true positive***, ***true negative***, ***false positive***, and ***false negative***, respectively. Based on these metrics, the performance can be measured not only to know the overall accuracy but also how good the system can detect a particular class of beef.

To measure the performance of the machine learning algorithms in the regression task, the following metrics were used: root mean squared error (RMSE), R-squared (R^2), adjusted R-squared (adjusted R^2), bias factor (B_f), accuracy factor (A_f), mean relative percentage residual (MRPR), and mean absolute percentage residual (MAPR). A detailed explanation of these metrics is given in Table 5. The symbols y, $\hat{y}$, p, L denote actual value vector, prediction result vector, number of predictors, and total number of samples, respectively. In this study, actual value vector refers to the observed number of the microbial population and prediction result vector refers to the outputs of the regression analysis.

4. Results and discussion

In this section, the conditions in the sample chamber are described in relation to changes in temperature and humidity. Subsequently, the signal reconstruction results are discussed. Furthermore, the performances of the machine learning algorithms using the noise filtering framework and without using the noise filtering framework (non-framework) are compared based on the dataset discussed above. In the case of non-framework, the signals contaminated with noise were only normalized to equalize the scale before being consumed by the machine learning algorithms.

4.1. Changes in temperature and humidity in the sample chamber

In the experiment, we also monitored the temperature and humidity during the beef spoilage process. Fig. 7 shows the changes of temperature and humidity in the sample chamber. The starting temperature was 33.5 °C. The temperature tended to increase during the spoilage process, reaching a maximum of 39.8 °C. The temperature rise was also affected by the gas sensors generating heat. Apart from that, temperature changes during day and night also affected the temperature, making the changes in temperature unstable. The initial relative humidity was 55.1%. The changes in humidity were affected by water vapor produced in the beef spoilage process, causing the humidity to rise to 82.9% when the beef sample was rotting. Hence, this uncontrolled environment was a source of noise because it affected the sensor response.

4.2. Signal reconstruction

First, the most appropriate decomposition level has to be determined. Then, the IQR values are calculated and compared to each

Table 4
Performance metrics for classification.

Metric	Equation	Description
Accuracy (%)	$accuracy = \frac{tp+tn}{tp+tn+fp+fn} \times 100$	Accuracy is used to measure the ratio of identification success for all classes against overall testing results.
Precision (%)	$precision = \frac{tp}{tp+fp} \times 100$	Precision is used to measure the capability of the classifier to detect a valid prediction against the number of all predictions of this class. For instance, the numbers of correct predictions of fresh beef are divided by all samples that are identified as fresh beef even though some of them are actually spoiled.
Recall (%)	$recall = \frac{tp}{tp+fn} \times 100$	Recall measures the ratio between the number of correct predictions for a particular class and the actual number of this class. For example, the number of valid predictions as fresh beef is divided by the total number of samples of real fresh beef.

Table 5
Performance metrics for regression.

Metric	Equation	Description
RMSE	$RMSE(y,\hat{y}) = \sqrt{\frac{\sum_{i=1}^{L}(y_i-\hat{y}_i)^2}{L}}$	RMSE is used to measure the difference/error between actual and prediction vectors. A lower RMSE value means less difference between the actual and the prediction value.
R^2	$R^2(y,\hat{y}) = 1 - \frac{\sum_{i=1}^{L}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{L}(y_i-\bar{y})^2}$	R-squared represents the parts of the variance of the actual vector that can be predicted by the regression model. Typically, it ranges from 0 to 1. R^2 equals 1 means that the regression model can correctly predict the actual value and vice versa. If R^2 is negative, it means that the regression model does not follow the trend of the actual data.
Adjusted R^2	$\bar{R}^2(y,\hat{y}) = R^2(y,\hat{y}) - \frac{(1-R^2(y,\hat{y}))p}{L-p-1}$	Adjusted R^2 is an extension of R^2 to avoid bias caused by an increase of the number of predictors. The range is the same as for R^2 but the value is usually lower than R^2.
Bf	$B_f(y,\hat{y}) = exp\left[\frac{\sum_{i=1}^{L}(\ln(y_i)-\ln(\hat{y}_i))}{L}\right]$	Bias factor denotes whether the predictions are under- or overestimated against the actual values. An unbiased prediction is indicated by B_f equals 1. $B_f < 1$ means the prediction result is lower than the actual value (underestimated) and vice versa (Baranyi et al., 1999).
Af	$A_f(y,\hat{y}) = exp\left[\sqrt{\frac{\sum_{i=1}^{L}(\ln(y_i)-\ln(\hat{y}_i))^2}{L}}\right]$	The accuracy factor measures the average accuracy of the prediction model. The value of A_f is equal to or greater than 1. A value larger than 1 indicates less accurate prediction results (Baranyi et al., 1999).
MRPR (%)	$MRPR(y,\hat{y}) = \frac{1}{L}\sum_{i=1}^{L}\frac{(y_i-\hat{y}_i)\times 100}{y_i}$	MRPR verifies the percentage of over- or under-prediction from the residual point of view. If the value is 0 it means there is no residual between the actual and the predicted value. MRPR < 0 indicates over-prediction and vice versa.
MAPR (%)	$MAPR(y,\hat{y}) = \frac{1}{L}\sum_{i=1}^{L}\frac{\lvert y_i-\hat{y}_i\rvert\times 100}{y_i}$	The absolute value of MAPR provides information about the percentage of average deviation between the actual and the predicted values. A lower MAPR indicates a lower absolute residual between the actual and the predicted values.

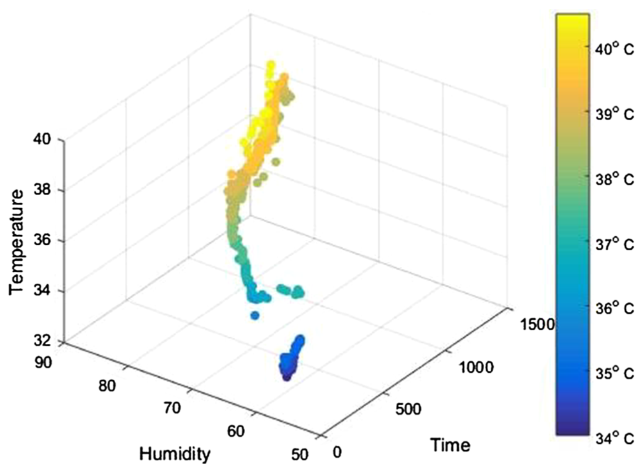

Fig. 7. Changes in temperature and humidity in the sample chamber.

Table 6
MWTs and decomposition levels of DWT for e-nose signals noise filtering.

Sensor	MWT	Decomposition level
S1	bior2.4	11
S2	bior3.3	11
S3	db1	11
S4	db1	10
S5	bior2.2	11
S6	db1	10
S7	db6	10

reconstructed signal using a particular MWT. In this experiment, 38 MWTs were compared, i.e. biorthogonal (bior1.1, bior1.3, bior1.5, bior2.2, bior2.4, bior2.6, bior2.8, bior3.1, bior3.3, bior3.5, bior3.7, bior3.9, bior4.4, bior5.5, bior6.8), symlet (sym2-sym8), coiflet (coif1-coif5), daubechies (db1-db10), and dmey. With Eqs. (4) and (5), the wavelet decomposition level and the best-suited MWT can be determined. The results of signal reconstruction based on DWT were used in the further processes. Table 6 shows the best-suited MWTs and decomposition levels for each signal corresponding to each sensor.

Fig. 8 shows the original and the reconstructed signals corresponding to the gas sensors. All of the sensor responses changed progressively during the beef spoilage process. The existence of biomarkers such as CH_4, H_2S, NH_3, and CO_2 caused decreased sensor resistance. However, the signals were contaminated with a great deal of noise. Hence, signal reconstruction was performed by considering the quality of the information that corresponds to a proper wavelet decomposition level and the best-suited MWT for each signal. The meat spoilage biomarkers had increasing concentrations during the spoiling process (800–1000 min) as shown in Fig. 8a–d, and f. The ambient humidity also increased during the beef spoilage process, as shown in Fig. 8g. Conversely, the resistance of S5 increased, which means the concentrations of LPG, *iso*-butane, and propane became smaller when the beef was decaying.

4.3. Binary classification

Binary classification divides the beef samples into two classes: 'fresh' and 'spoiled'. The indicators to differentiate fresh from rotten beef are high concentrations of CH_4, H_2S, NH_3, and CO_2 produced by spoiled beef. Hence, binary classification is the simplest problem in beef quality classification because there is only one class boundary with significant differences in concentration of the biomarkers (Wijaya et al., 2016b).

Table 7 shows a performance comparison between several machine learning algorithms with and without using the proposed noise filtering framework. The noise filtering framework gave a small improvement of classification accuracy for SVM but had no impact on k-NN and produced a slight performance decrease in ANN and ANFIS. QDA had the worst performance compared to the other algorithms. Noise filtering could not be performed on QDA because it causes a singular covariance matrix in multivariate data (determinant equals zero). Another obstacle is the near-zero variance predictors for a particular class caused by noise reduction. Nevertheless, the accuracies were satisfactory, either using or not using the noise filtering (greater than 90%), except for QDA (77.79%). The utilization of the noise filtering framework in k-NN had no impact on its accuracy, which remained at 93.64%. In SVM it resulted in a small accuracy enhancement of only 1.07%. Performing noise filtering in ANN and ANFIS even slightly decreased the performance (1.29% and 3.85% respectively). These results indicate that noise filtering does not provide satisfactory performance improvements in binary classification tasks.

4.4. Three-class classification

In the three-class classification task, the beef quality was divided into 'fresh', 'semi-fresh', and 'spoiled'. Having more classes and more boundaries implies that this task is more complicated than binary classification. The performance of the noise filtering framework is detailed in Table 7. Without noise filtering, the recall values of 'fresh' were low (ranging between 48.86% and 64.69%) and so were the precision values of 'semi-fresh' (ranging between 60.83% and 71.73%).

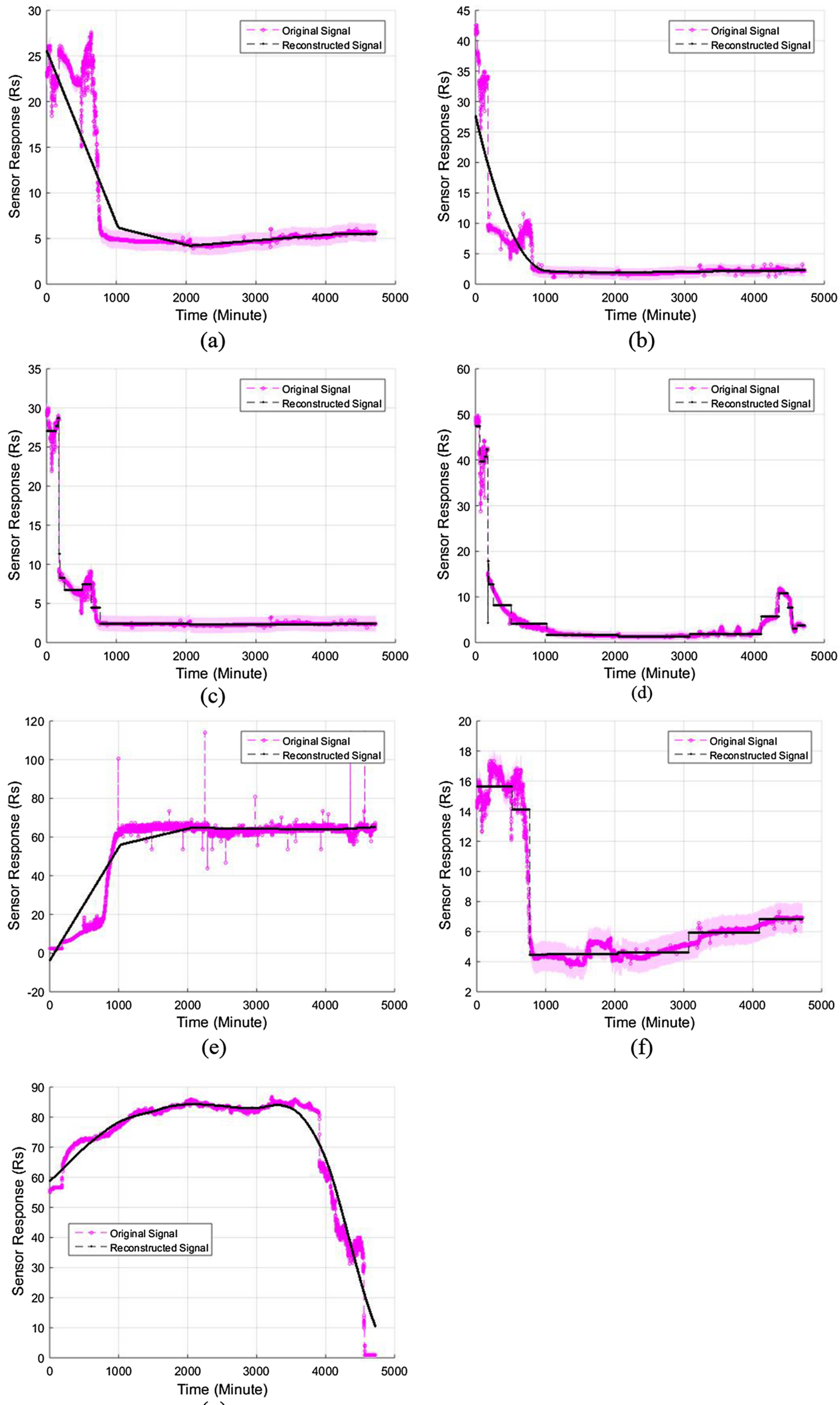

Fig. 8. Comparison of original and reconstructed signals. The shaded magenta line and the black line are the original and the reconstructed signal, respectively. (a) S1 (b) S2 (c) S3 (d) S4 (e) S5 (f) S6 (g) S7. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

These low precision and recall values were caused by many 'fresh' beef samples being misclassified as 'semi-fresh'. On the other hand, the recall and precision values of spoiled beef were relatively higher without noise filtering. In three-class classification, the effect of noise filtering was more significant than for binary classification. Almost all the machine learning algorithms got a performance improvement; most of the

Table 7
Performance comparison of machine learning algorithms without and with proposed framework (average accuracy, precision, and recall).

The number of classes	Class labels	Performance metrics (%)	k-NN without framework	k-NN with framework	SVM without framework	SVM with framework	QDA without framework	QDA with framework	ANN without framework	ANN with framework	ANFIS without framework	ANFIS with framework
Two classes	fresh	precision	89.95	↓ 88.38	87.31	↑ 88.38	79.28	N/A	90.50	↓ 88.38	89.95	↓ 88.38
		recall	98.05	↑ 100	99.48	↑ 100	86.62	N/A	99.74	↑ 100	100	↓ 94.94
	spoiled	precision	97.73	↑ 100	99.37	↑ 100	87.68	N/A	99.68	↑ 100	100	↓ 94.49
		recall	90.89	↓ 89.04	87.68	↑ 89.04	74.38	N/A	91.50	↓ 89.04	90.89	↓ 89.04
	accuracy		93.64	= 93.64	92.57	↑ 93.64	77.79	N/A	94.93	↓ 93.64	94.71	↓ 90.86
Three classes	fresh	precision	100	= 100	100	= 100	92.51	N/A	100	= 100	70.86	↑ 100
		recall	62.59	= 62.59	62.59	= 62.59	62.24	N/A	64.69	↑ 89.86	48.86	↑ 66.78
	semi-fresh	precision	67.61	↓ 66.21	60.83	↑ 66.21	67.12	N/A	71.73	↑ 76.18	68.50	↑ 84.06
		recall	96.90	↑ 100	99.79	↑ 100	67.36	N/A	99.59	↑ 100	75.14	↑ 91.94
	spoiled	precision	97.73	↑ 100	99.84	↑ 100	85.16	N/A	99.68	↑ 100	96.24	↓ 94.49
		recall	90.89	↓ 89.04	83.00	↑ 89.04	93.49	N/A	93.10	↓ 89.04	100	= 100
	accuracy		86.00	= 86.00	82.43	↑ 86	78.07	N/A	88.64	↑ 91.57	83.64	↑ 90.43
Four classes	excellent	precision	100	= 100	100	= 100	89.44	N/A	100	= 100	85.71	↑ 100
		recall	62.59	= 62.59	62.59	= 62.59	62.59	N/A	65.03	↑ 89.86	54.84	↑ 97.55
	good	precision	47.42	↑ 73.97	27.97	↑ 71.88	0.00	N/A	15.29	↑ 76.38	35.67	↑ 88.7
		recall	55.65	↑ 100	83.49	↑ 100	32.08	N/A	49.52	↑ 99.53	56.36	↑ 100
	acceptable	precision	35.54	↑ 67.04	67.31	↓ 64.11	0	N/A	12.15	↑ 70.91	8.21	↑ 62.61
		recall	47.47	↑ 97.45	10.88	↑ 89.78	0	N/A	46.74	↑ 74.09	4.01	↑ 66.92
	spoiled	precision	97.73	↑ 100	96.31	↑ 100	77.73	N/A	99.29	↓ 99.07	86.37	↑ 94.49
		recall	90.89	↓ 89.04	78.94	↑ 85.59	83.74	N/A	96.55	↓ 92.24	94.33	↓ 89.04
	accuracy		70.50	↑ 85.5	65.43	↑ 82	50.14	N/A	73.21	↑ 88.29	68.64	↑ 88.93

Notes: ↓ is lower result. ↑ is higher result. = is equal result.

accuracy, precision, and recall values increased. According to the experiment, the recall values of the ‘fresh’ class were significantly improved for ANN (from 64.69% to 89.86%) and ANFIS (from 48.86% to 66.78%). Moreover, the precision values of the ‘semi-fresh’ class also increased for SVM (from 60.83% to 66.21%), ANN (from 71.73% to 76.18%), and ANFIS (from 68.5% to 84.06%). With the exception of k-NN, the accuracy of all machine-learning algorithms increased: SVM from 82.43% to 86%, ANN from 88.64% to 91.57%, and ANFIS from 83.64% to 90.43%. These results indicate that the accuracy improvements of SVM, ANN, and ANFIS were 3.28%, 2.93%, and 6.79%, respectively. Thus, the proposed noise filtering framework gives more improvement in three-class classification tasks than in binary classification tasks. In the case of QDA, the same happens as in binary classification, namely noise filtering leading to a singular covariance matrix.

4.5. Four-class classification

In four-class classification, the quality of the beef was categorized into ‘excellent’, ‘good’, ‘acceptable’, and ‘spoiled’. This is more complicated than binary and three-class classification because there are more boundaries between classes, which increases the possibility of wrong predictions. Table 7 details the performance of noise filtering in four-class classification. Without noise filtering, the recall values for ‘excellent’ ranged between 54.84% and 65.03%, the recall values for ‘good’ ranged between 32.08% and 83.49%, and the recall values for ‘acceptable’ ranged between 0% and 47.47%. Overall, the recall and precision values for ‘spoiled’ were relatively high. The recall values ranged between 78.94% and 96.55% and the precision values ranged between 77.73% and 99.29%. These results indicate that all machine learning algorithms perform badly when distinguishing between ‘excellent’, ‘good’, and ‘acceptable’ beef if the signal is contaminated with noise. The noise filtering framework can be used successfully to deal with this issue. According to the experimental results, performance improvements were obtained by all machine learning algorithms, except QDA because, as in the previous two cases, the noise filtering leads to a singular covariance matrix, or near-zero variance predictors for a particular class. The increased accuracy for k-NN, SVM, ANN, and ANFIS was 15%, 16.57%, 15.08%, 20.29%, respectively. These results imply that the influence of the noise filtering framework in four-class classification is more significant than for binary and three-class classification. It could significantly improve the recall values of ‘excellent’ for ANN (from 65.03% to 89.86%) and ANFIS (from 54.84% to 97.55%). Improvements also occurred for the recall of ‘acceptable’. Large recall improvements occurred for k-NN (from 47.47% to 97.45%), SVM (from 10.88% to 89.78%), ANN (from 46.74% to 74.09%), and ANFIS (from 4.01% to 66.92%). The precision improvements of ‘good’ were also significant for k-NN (from 47.42% to 73.97%), SVM (from 27.97% to 71.88%), ANN (from 15.29% to 76.38%), and ANFIS (from 35.67% to 88.7%). The performances of the layered model machine learning algorithms (ANN and ANFIS) had the highest classification accuracy (ca. 88%) and thus are promising for multiclass classification. Moreover, there is still some space for improvement regarding the tuning of the hyperparameter in the layered models.

4.6. Trend of machine learning algorithm performance in classification tasks

Fig. 9 shows the performance trend of the machine learning algorithms. It also compares the performance when using the framework and when not using the framework for two, three, and four classes of beef. The performance difference between framework and non-framework was small in binary classification. There was a slight increase for one algorithm but a small decrease or no effect for the others. Typically, if the number of classes increases, the performance will deteriorate. This is shown by the non-framework trend of accuracy, precision, and recall: the performance indicates a downward trend. It decreased with three classes and got worse with four classes. This happens to all machine learning algorithms when consuming e-nose signals that are badly contaminated with noise. The performance gaps between framework and non-framework were higher in multiclass classification. The performance improvements were more significant and more stable when the number of classes increased. The noise filtering framework improved accuracy, precision, and recall. These results indicate that the noise filtering framework can mitigate severe performance degradation when the number of classes increases. For multiclass classification, ANFIS and ANN provide the best performance to classify several grades of beef. In contrast, QDA produces the worst performance and applying noise filtering in this case is unfavorable because it can lead to a singular covariance matrix, which means that the determinant equals zero.

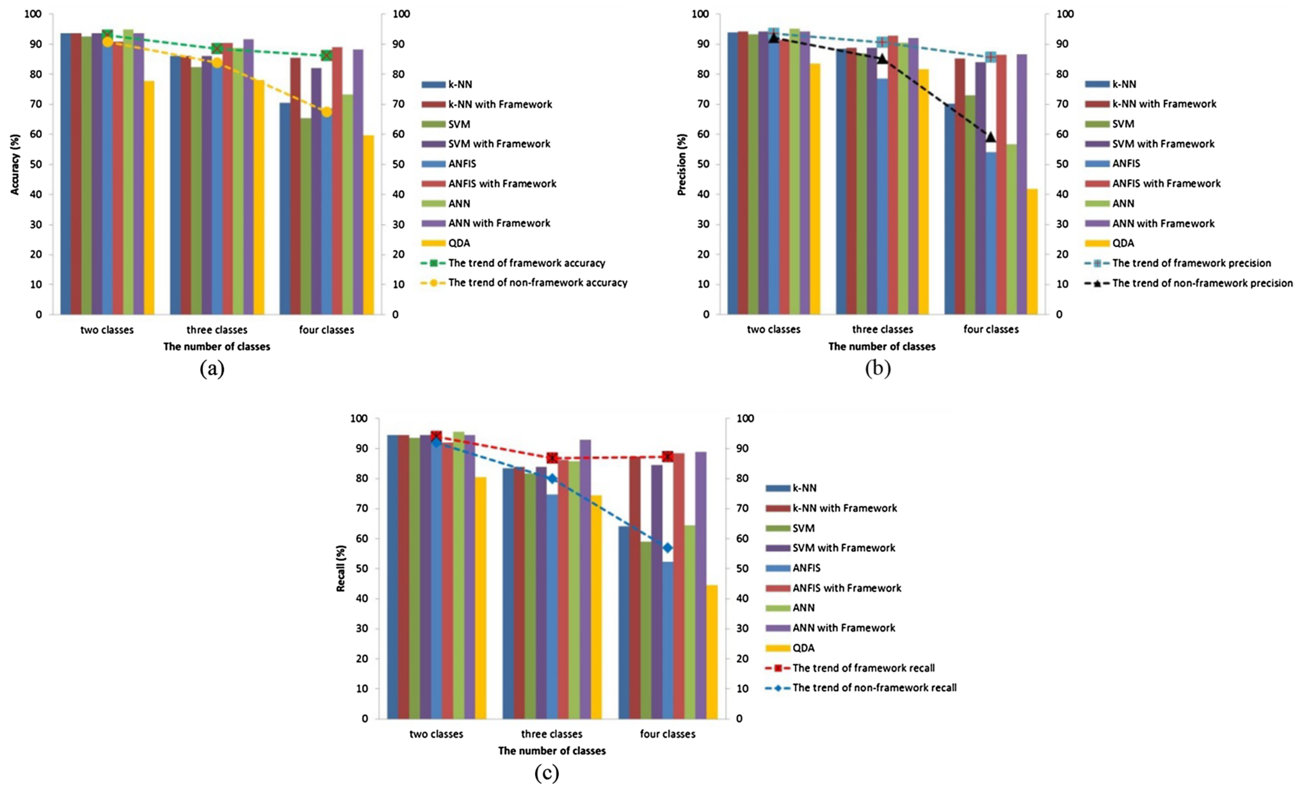

Fig. 9. Performance comparison and trends of various machine learning algorithms in beef classification tasks: (a) accuracy (b) precision (c) recall.

In Eq. (8), the log of a zero determinant means that the function cannot be executed. An increasing number of classes will be a problem because QDA has a separate covariance matrix for every class. This assumption is more likely to generate singular covariance matrices than LDA. Hence, QDA is not recommended for multiclass classification.

4.7. Comparison of different parameters

This study did not focus on parameter optimization but a comparison of different parameters used in k-NN, SVM, and ANN is provided as proof of the noise filtering framework's effectiveness. For k-NN, the value of k was tested on $[1, 2, \cdots, 15]$. We found that the changes of k did not have a significant effect when the noise filtering framework was applied. Thus, we set $k = 5$ to investigate this effect with different distance metrics. It is enough to represent the results of the experiment in general. Table 8 shows a comparison of the accuracy for four distance metrics (Euclidean, city block, Chebyshev, and Mahalanobis). The results indicate that the noise filtering framework can provide a significant improvement in four-class classification. For SVM, we tested $C = [10^{-2}, 10^{-1}, \cdots, 10^{7}]$. The changes of C did not have a significant impact when noise filtering was used. Hence, we set $C = 1$ to compare several kernel functions, i.e. linear, quadratic, third-degree polynomial and RBF, as shown in Table 9. The results show that the noise filtering framework could improve the accuracy in binary and multiclass classification for SVM classifiers. Satisfactory improvement was obtained in multiclass classification. Changes of k in k-NN and C in SVM could not provide significant improvements because the data were well-separated as a result of the noise filtering. Table 10 shows the performance of ANN using several activation functions, i.e. sigmoid, hyperbolic tangent (tanh), rectified linear unit (ReLU), and softplus. The results show that the noise filtering framework could not provide improvement in binary classification, whereas it could produce better performance in multiclass classification. These results confirm that the proposed noise filtering framework can provide performance improvement of different parameters in the machine learning algorithms.

4.8. Effect of noise filtering on regression task

In this study, regression analysis was utilized to predict the number of bacteria in beef samples (TVC) in accordance with beef quality. Fig. 10 shows a plot of actual and predicted TVC values using NNR. Fig. 10a shows the performance of NNR without noise filtering (non-framework). It shows that there is no proper distribution around the line of equity ($y = x$) within the $\pm$ 1 log unit area. The residual plot in Fig. 10b also indicates a linear pattern of the residual values. This indicates that the model cannot capture linear information during prediction of the size of the microbial population. The regression result shows that NNR cannot correctly model the signals contaminated with noise. The utilization of the noise filtering framework provides better performance, as shown in Fig. 10c. The trend follows the line of equity, even though not perfectly. The residual plot in Fig. 10d also shows a more random pattern compared with Fig. 10b. This means that there is not much leakage of linear

Table 8
Effect of noise filtering framework on k-NN with different distance metrics.

Accuracy (%)	Two classes				Three classes				Four classes			
	Euclidean	City block	Chebyshev	Mahalanobis	Euclidean	City block	Chebyshev	Mahalanobis	Euclidean	City block	Chebyshev	Mahalanobis
Non-framework	93.64	93.36	94.14	63.21	86.00	85.71	86.93	44.43	70.50	70.14	70.86	44.14
Framework	≈ 93.64	↑ 93.64	↓ 93.64	↑ 75.29	≈ 86.00	↑ 86.86	↓ 85.86	↑ 58.00	↑ 85.50	↑ 86.50	↑ 85.86	↑ 54.71

Table 9
Effect of noise filtering framework on SVM with different kernel functions.

Accuracy (%)	Two classes				Three classes				Four classes			
	Linear	Quadratic	Polynomial-3	RBF	Linear	Quadratic	Polynomial-3	RBF	Linear	Quadratic	Polynomial-3	RBF
Non-framework	92.71	88.50	85.64	92.57	76.07	71.29	78.43	82.43	50.29	58.71	70.93	65.43
Framework	↑93.64	↑93.64	↑86.29	↑93.64	↑86.00	↑73.93	↓77.71	↑86.00	↑82.21	↑81.79	↑80.36	↑82.00

Table 10
Effect of noise filtering framework on ANN with different activation functions.

Accuracy (%)	Two classes				Three classes				Four classes			
	Sigmoid	Tanh	ReLU	Softplus	Sigmoid	Tanh	ReLU	Softplus	Sigmoid	Tanh	ReLU	Softplus
Non-framework	94.93	94.93	95.36	94.93	88.64	88.00	88.08	87.29	73.21	78.07	78.50	79.22
Framework	↓93.64	↓93.65	↓93.65	↓93.65	↑91.57	↓87.50	↑89.36	↑88.79	↑88.29	↑85.36	↑85.50	↑85.65

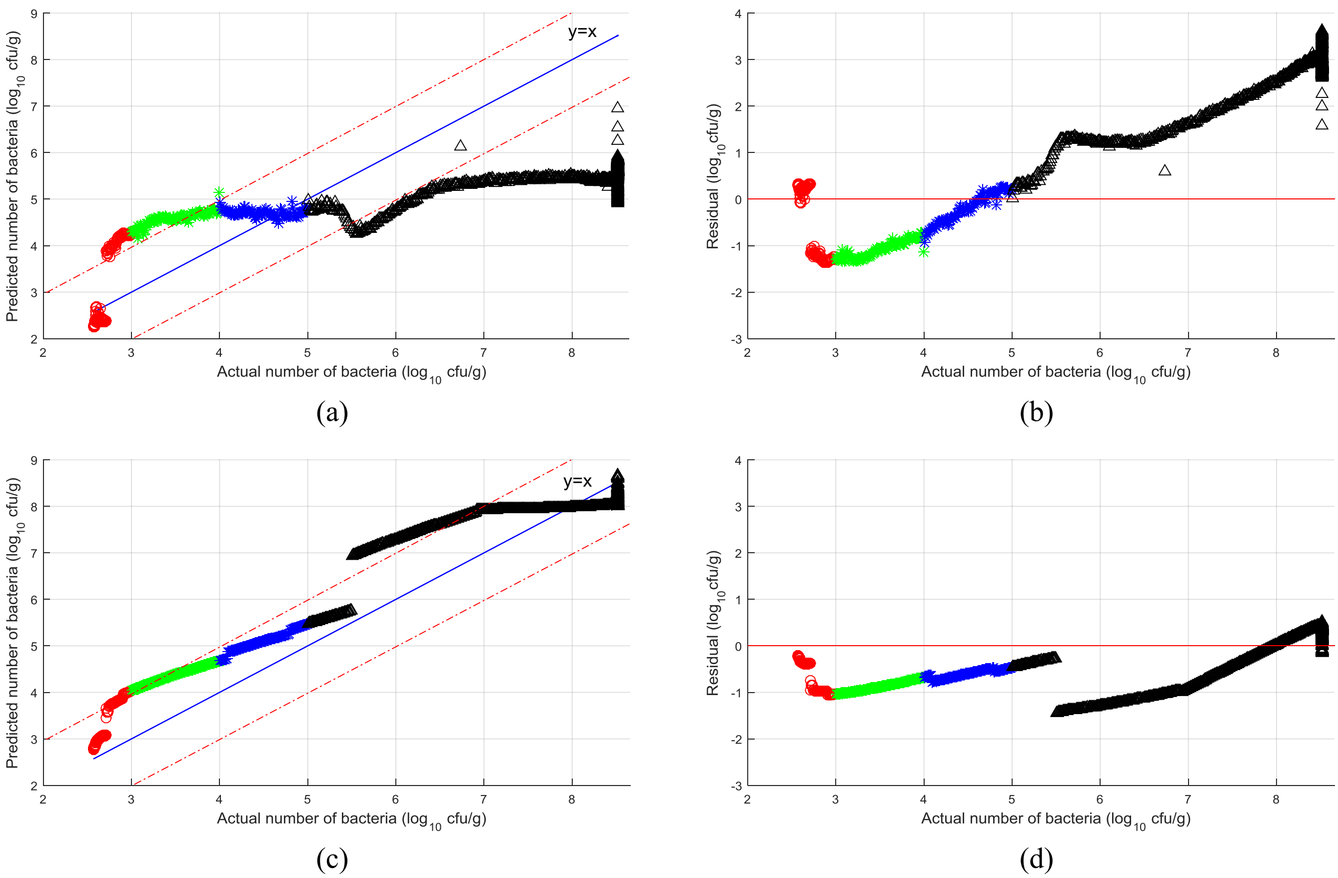

Fig. 10. Prediction of the number of the microbial population using NNR. Red, green, blue, and black marks indicate 'excellent', 'good', 'acceptable', and 'spoiled' beef, respectively. (a) non-framework NNR, (b) residual of non-framework NNR, (c) NNR with framework, (d) residual of NNR with framework. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

information to the residual compared to non-framework. Table 11 shows a quantitative comparison between non-framework NNR and NNR with framework. RMSE denotes the prediction error, which shows the difference between actual and predicted values. RMSE was reduced from 2.556 to 0.519 for overall prediction. Furthermore, R^2 and adjusted R^2 also showed superior results for NNR with framework (0.953) compared to non-framework NNR (0.771). These results indicate that noise filtering can reduce errors and improve the agreement between predicted and actual data. The B_f of non-framework NNR indicates that the regression model produced overestimated results for 'excellent', 'good', and 'acceptable' but yielded underestimated predictions after the beef had started to spoil. Meanwhile, NNR with framework also had overestimated results for 'excellent', 'good', and 'acceptable' but nearly perfect results in the spoiled stage (0.982) as well as overall B_f (1.011). According to the experimental results, NNR with framework had better A_f than non-framework NNR, except for the 'acceptable' class. Overall, A_f improved from 1.458 to 1.1. Moreover, MRPR exhibited vast underprediction for non-framework NNR (25.249%) but slight overprediction when using the noise filtering framework (-1.586%). The overall MAPR value for non-framework NNR was also higher (30.574%) than for NNR framework (7.388%). This indicates that NNR with framework has a lower deviation.

Table 11
Effect of noise filtering framework on NNR.

	Non-framework					Framework				
	excellent	good	acceptable	spoiled	overall	excellent	good	acceptable	spoiled	overall
RMSE	0.784	1.111	0.373	2.741	2.556	0.652	0.894	0.611	0.476	0.519
R^2	0.836	0.762	0.047	0.599	0.771	0.926	0.997	0.971	0.731	0.953
Adjusted R^2	0.835	0.759	0.037	0.599	0.771	0.925	0.997	0.971	0.731	0.953
B_f	1.078	1.32	1.047	0.677	0.729	1.201	1.259	1.136	0.982	1.011
A_f	1.266	1.328	1.087	1.488	1.458	1.226	1.264	1.139	1.066	1.1
MRPR (%)	−10.555	−32.247	−4.94	32.139	25.249	−20.565	−26.041	−13.621	1.572	−1.586
MAPR (%)	22.099	32.247	7.093	32.139	30.574	20.565	26.041	13.621	5.203	7.388

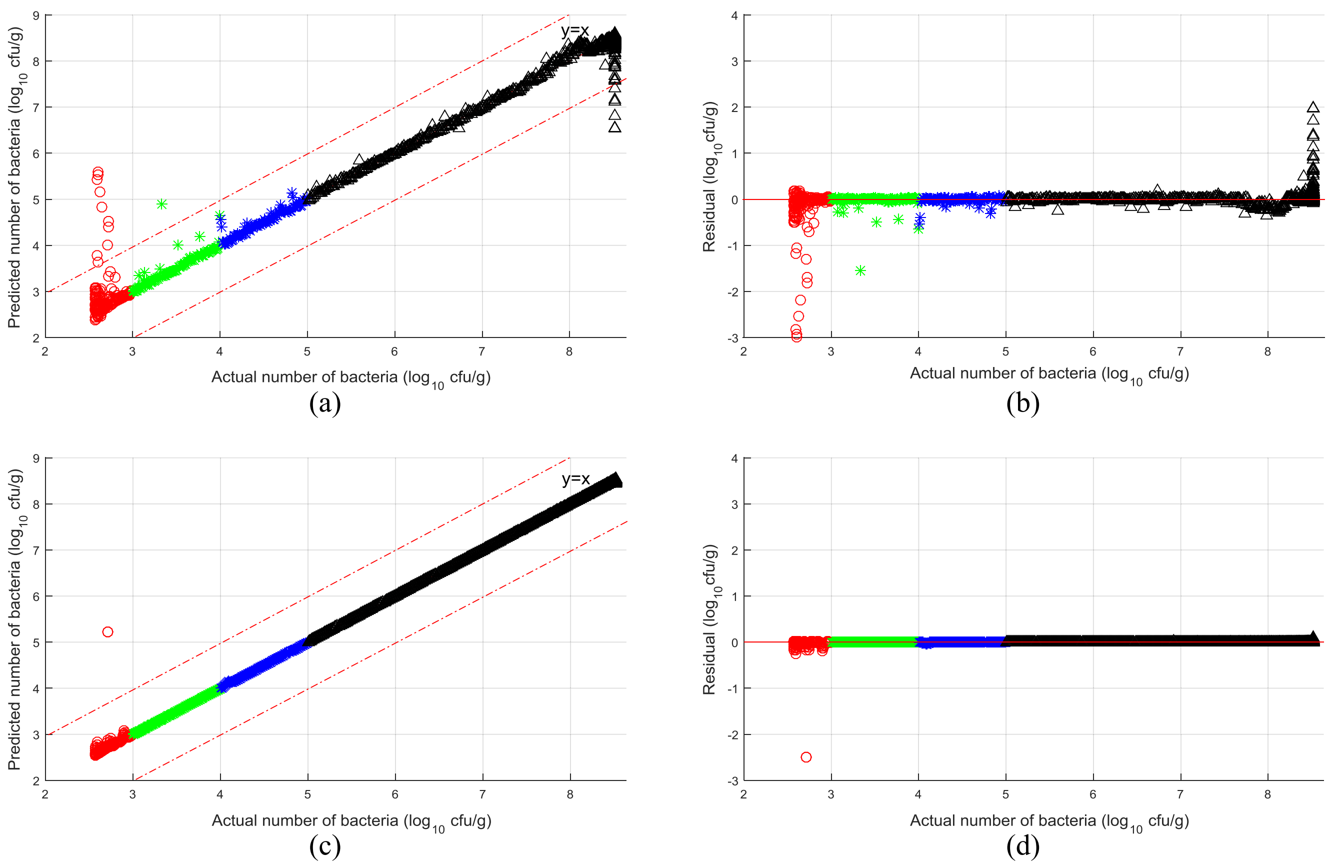

Fig. 11. Prediction of the number of the microbial population using SVR. Red, green, blue, and black marks indicate 'excellent', 'good', 'acceptable', and 'spoiled' beef, respectively. (a) non-framework SVR, (b) residual of non-framework SVR, (c) SVR with framework, (d) residual of SVR with framework. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 12
Effect of noise filtering framework on SVR.

	Non-framework					Framework				
	excellent	good	acceptable	spoiled	overall	excellent	good	acceptable	spoiled	overall
RMSE	0.602	0.182	0.092	0.131	0.196	0.302	0.001	0.01	0.005	0.074
R^2	0	0.732	0.906	0.971	0.989	0.154	1	0.999	1	0.998
Adjusted R^2	−0.007	0.729	0.904	0.971	0.989	0.148	1	0.999	1	0.998
B_f	1.068	1.01	1.005	0.998	1.003	1.018	1	1.001	1	1.001
A_f	1.182	1.047	1.021	1.017	1.046	1.083	1	1.002	1.001	1.02
MRPR (%)	−8.287	−1.116	−0.481	0.142	−0.446	−2.163	−0.004	−0.07	0.007	−0.127
MAPR (%)	9.571	1.582	1.059	0.413	1.042	2.445	0.013	0.094	0.014	0.163

Fig. 11a and b show the baseline performance of non-framework SVR, which is visually better than that of NNR. However, the influence of noise filtering SVR's performance is more interesting. Fig. 11c and d indicate that the noise filtering framework had a significant influence on SVR's performance. It yielded almost perfect prediction results with a distribution pattern that perfectly follows the line of equity. Fig. 11c

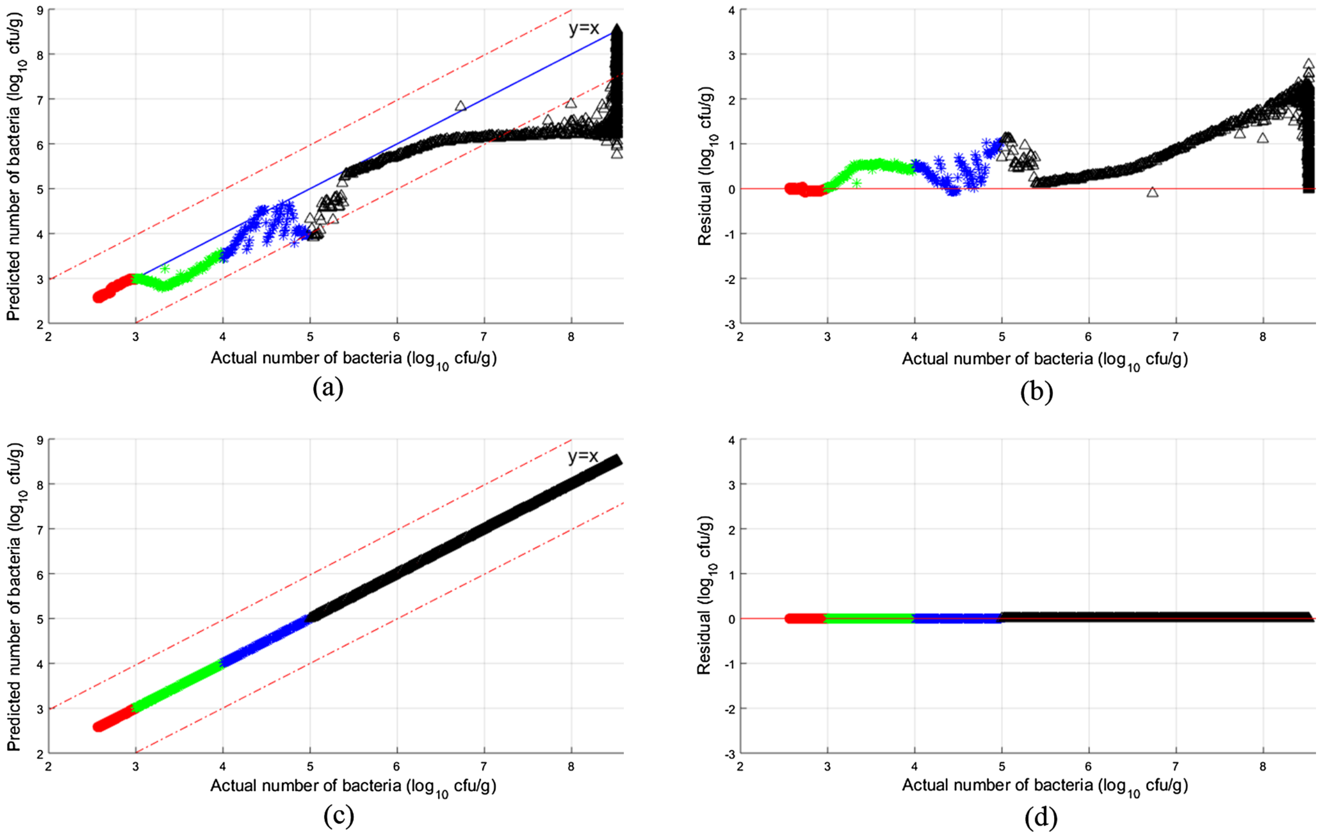

Fig. 12. Prediction of the number of the microbial population using ANFIS regression. Red, green, blue, and black marks indicate 'excellent', 'good', 'acceptable', and 'spoiled' beef, respectively. (a) non-framework ANFIS regression, (b) residual of non-framework ANFIS regression, (c) ANFIS regression with framework, (d) residual of ANFIS regression with framework. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 13
Effect of noise filtering framework on ANFIS regression.

	Non-framework					Framework				
	excellent	good	acceptable	spoiled	overall	excellent	good	acceptable	spoiled	overall
RMSE	0.032	0.432	0.513	0.818	0.769	0.006	0.003	0.001	0	0.002
R^2	0.978	0.678	0.228	0.6	0.886	1	1	1	1	1
Adjusted R^2	0.977	0.675	0.219	0.6	0.886	1	1	1	1	1
B_f	1.006	0.886	0.907	0.939	0.94	1.002	1.001	1	1	1
A_f	1.011	1.14	1.127	1.119	1.117	1.002	1.001	1	1	1.001
MRPR (%)	−0.612	11.238	9.042	5.656	5.66	−0.197	−0.093	0.014	0.001	−0.015
MAPR (%)	0.828	11.238	9.247	5.677	5.773	0.197	0.093	0.02	0.001	0.018

indicates a small difference between the actual and the predicted values; there is only one point that is significantly different, which is in the 'excellent' class. Table 12 shows more details of the performance. The overall RMSE value is 0.074, which explains the very low error. Furthermore, R^2 and adjusted R^2 also performed better. The values for all classes are high (0.999–1) except for 'excellent' with an R^2 of only 0.154 of and an adjusted R^2 of 0.148. This is because the baseline values of 'excellent' (non-framework SVR) are extremely low ($R^2 = 0$ and adjusted $R^2 = -0.007$). However, overall R^2 and adjusted R^2 still showed high agreement between the actual and the predicted values (0.998). Moreover, the overall B_f value is 1.001, which shows very low bias, with only 0.001 overestimates in the prediction results. In terms of accuracy, $A_f \approx 1$ implies better accuracy than non-framework. Fig. 11d shows that the residual is closer to zero compared to Fig. 11b. This can be explained by the lower MRPR and MAPR values.

Fig. 12 shows the prediction of the number of the microbial population using ANFIS with framework and non-framework ANFIS. Fig. 12a and b show that ANFIS produced unsatisfactory performance when it had to deal with noisy signals. The majority of the prediction results are below the actual values, which means that the results were underestimated. The advantage of using the noise filtering framework is shown in Fig. 12c and d. They indicate a significant improvement when using the noise filtering framework. The plots of prediction of the number of the microbial population almost precisely follow the line of equity as well as the low residual. Moreover, Table 13 also shows a very low RMSE (ca. 0.002), which means small differences between the actual values and the predicted values from the regression model. Moreover, the R^2 and adjusted R^2 values indicate perfect agreement between the actual and the predicted values ($R^2 = 1$). The noise filtering also successfully reduced the bias ($B_f = 1$). When not using framework, it yielded an underestimated value ($B_f = 0.94$). The value of A_f decreased from 1.117 to 1.001, which means an improvement of model accuracy. MRPR was −0.015% and MAPR was 0.018%, indicating low residuals generated by the regression model when using the noise filtering framework.

The above results demonstrate the advantage of using the proposed noise filtering framework in regression tasks. Basically, SVR is more robust than NNR and ANFIS against noisy signals. In SVR, the soft margin, which is represented by a slack variable, can give better performance than in NNR and ANFIS. Nevertheless, noise filtering is still needed in SVR. NNR and ANFIS yield unsatisfactory performance with noisy signals. This is because they do not have a particular mechanism to deal with noise. They get a significant improvement when noise filtering is applied. Finally, according to the experimental results, utilization of the proposed noise filtering framework presents a significant improvement in regression tasks. This was proved by reduced number of errors (RMSE), bias (B_f), residual (MRPR and MAPR) as well as increased accuracy (A_f) and agreement between actual and prediction results, as shown by better appropriate values (R^2 and adjusted R^2).

5. Conclusions

The use of analytical instruments is necessary for food safety and in particular for highly perishable foods such as beef, lamb, pork, fish, etc. The utilization of an e-nose device for online food quality monitoring is a prospective concept to monitor shelf-life during meat storage. However, the presence of noise in the e-nose signal generated by the gas sensors is a major challenge. In this study, a framework for noise filtering was developed to deal with e-nose signals in beef quality monitoring that are contaminated with noise. The effects of the framework have been investigated for classification of two, three, and four classes of beef grade based on several well-known machine learning algorithms that have been used in previous studies on beef quality monitoring. The results showed that k-NN, SVM, ANN, and ANFIS got significant improvement in multiclass classification, while the impact on binary classification was not very distinct. In multiclass classification tasks, the proposed noise filtering framework could also provide performance improvement on various parameters of the machine learning algorithms. Meanwhile, the utilization of the noise filtering framework in QDA caused a singular covariance matrix problem. The layered models, e.g., ANN and ANFIS produced the best performance when using the framework, especially in multiclass classification. Meanwhile, QDA had the worst performance in all cases.

In regression tasks, NNR, SVR, and ANFIS were used to predict the number of the microbial population in a beef sample. The experimental results demonstrated that the use of the noise filtering framework can reduce the number of errors, bias and residual in regression tasks. Moreover, it can increase the parts of the variance of the actual data that can be predicted by the regression model, as indicated by the enhancement of R^2 and adjusted R^2. The prediction accuracy also increased, as shown by $A_f \approx 1$. According to the experimental results, SVR and ANFIS performed better than NNR. Overall, the results show the advantage of using the proposed noise filtering framework. This is proved by the performance improvement of the machine learning algorithms that were used in previous studies. Hence, the proposed noise filtering framework provides a solution to deal with e-nose signals in beef quality monitoring that are contaminated with noise caused by uncontrolled ambient conditions. Finally, performing noise filtering on e-nose signal is highly recommended, especially for multiclass classification and regression tasks.

Acknowledgements

This work was supported by Institut Teknologi Sepuluh Nopember, Telkom University, and Ministry of Research, Technology, and Higher Education Republic of Indonesia under contract 014/PNLT3/PPM/2018.

References

Ahmed, I., Lin, H., Zou, L., Li, Z., Brody, A.L., Mabood Qazi, I., Lv, L., Ramesh Pavase, T., Ullah Khan, M., Khan, S., Sun, L., Zhenxing Li, C., 2018. An overview of smart packaging technologies for monitoring safety and quality of meat and meat products. Packaging Technol. Sci. 31, 449–471. https://doi.org/10.1002/pts.2380.

Alexandratos, N., Bruinsma, J., 2012. WORLD AGRICULTURE TOWARDS 2030/2050 The 2012 Revision. Rome.

Argyri, A.A., Panagou, E.Z., Tarantilis, P.A., Polysiou, M., Nychas, G.J.E., 2010. Rapid qualitative and quantitative detection of beef fillets spoilage based on Fourier transform infrared spectroscopy data and artificial neural networks. Sens. Actuat. B 145, 146–154. https://doi.org/10.1016/j.snb.2009.11.052.

Baietto, M., Pozzi, L., Wilson, A.D., Bassi, D., 2013. Evaluation of a portable MOS electronic nose to detect root rots in shade tree species. Comput. Electron. Agric. 96, 117–125. https://doi.org/10.1016/j.compag.2013.05.002.

Balasubramanian, S., Amamcharla, J., Shin, J.E., 2016. Possible application of electronic nose systems for meat safety: an overview. In: Méndez, M.L.R. (Ed.), Electronic Noses and Tongues in Food Science. Elsevier Inc., London, pp. 59–71. https://doi.org/10.1016/B978-0-12-800243-8.00007-X.

Balasubramanian, S., Logue, C.M., Marchello, M., 2004. Spoilage identification of beef using an electronic nose system. Transactions of the ASAE 47, 1625–1633. https://doi.org/10.13031/2013.17593.

Balasubramanian, S., Panigrahi, S., Logue, C.M., Gu, H., Marchello, M., 2009. Neural networks-integrated metal oxide-based artificial olfactory system for meat spoilage identification. J. Food Eng. 91, 91–98. https://doi.org/10.1016/j.jfoodeng.2008.08.008.

Baranyi, J., Pin, C., Ross, T., 1999. Validating and comparing predictive models. Int. J. Food Microbiol. 48, 159–166. https://doi.org/10.1016/S0168-1605(99)00035-5.

Baranyi, J., Roberts, T.A., 1994. A dynamic approach to predicting bacterial growth in food. Int. J. Food Microbiol. 23, 277–294. https://doi.org/10.1016/0168-1605(94)90157-0.

Chang, C.-C., Lin, C.-J., 2011. LIBSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technol. 2, 1–27. https://doi.org/10.1145/1961189.1961199.

CSIRO Food and Nutritional Sciences, 2003. Vacuum-packed meat : storage life and spoilage.

Deeplearning4j Development Team, 2017. Deeplearning4j: Open-source distributed deep learning for the JVM, Apache Software Foundation License 2.0.

El Barbri, N., Llobet, E., El Bari, N., Correig, X., Bouchikhi, B., 2008. Electronic nose based on metal oxide semiconductor sensors as an alternative technique for the spoilage classification of red meat. Sensors 8, 142–156. https://doi.org/10.3390/s8010142.

Ellis, D.I., Goodacre, R., 2001. Rapid and quantitative detection of the microbial spoilage of muscle foods: current status and future trends. Trends Food Sci. Technol. 12, 414–424.

Elmasry, G., Barbin, D.F., Sun, D.W., Allen, P., 2012. Meat quality evaluation by hyperspectral imaging technique: an overview. Crit. Rev. Food Sci. Nutr. 52, 689–711. https://doi.org/10.1080/10408398.2010.507908.

Gao, R.X., Yan, R., 2011. Wavelets: theory and applications for manufacturing. Springer, New York, Dordrecht Heidelberg London, New York 10.1007/978.

Güney, S., Atasoy, A., 2015. Study of fish species discrimination via electronic nose. Comput. Electron. Agric. 119, 83–91. https://doi.org/10.1016/j.compag.2015.10.005.

Hariyanto, Sarno R., Wijaya, D.R., 2017. Detection of diabetes from gas analysis of human breath using e-Nose. In: 2017 11th International Conference on Information & Communication Technology and System (ICTS). IEEE, Surabaya, pp. 241–246.

Harley, J.P., Prescott, L.M., 2002. Microbiology. In: Laboratory Exercise in Microbiology, Fifth Ed. McGraw-Hill, New York. https://doi.org/10.1007/s13398-014-0173-7.2.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An Introduction to Statistical Learning with Applications in R. In: Springer Texts in Statistics, 2017th ed.,. Springer. https://doi.org/10.1007/978-1-4614-7138-7.

Kiselev, I., Sysoev, V., Kaikov, I., Koronczi, I., Tegin, R.A.A., Smanalieva, J., Sommer, M., Ilicali, C., Hauptmannl, M., 2018. On the temporal stability of analyte recognition with an e-nose based on a metal oxide sensor array in practical applications. Sensors (Switzerland) 18. https://doi.org/10.3390/s18020550.

Kodogiannis, V.S., 2017. Application of an electronic nose coupled with fuzzy-wavelet network for the detection of meat spoilage. Food Bioprocess Technol. https://doi.org/10.1007/s11947-016-1851-6.

Kodogiannis, V.S., Alshejari, A., 2016. Neuro-fuzzy based identification of meat spoilage using an electronic nose. In: 2016 IEEE 8th International Conference on Intelligent Systems. IEEE, pp. 710–717. https://doi.org/10.1109/IS.2016.7737406.

Kodogiannis, V.S., Alshejari, A., 2014. An adaptive neuro-fuzzy identification model for the detection of meat spoilage. Appl. Soft Comput. J. 23, 483–497. https://doi.org/10.1016/j.asoc.2014.06.009.

Kodogiannis, V.S., Petrounias, I., Kontogianni, E., 2014. Identification of meat spoilage by FTIR spectroscopy and neural networks. In: Proceedings of the International Joint Conference on Neural Networks, pp. 1644–1651. https://doi.org/10.1109/IJCNN.2014.6889395.

Kodogiannis, V.S., Petrounias, I., Lygouras, J., 2015. An intelligent based decision support system for the detection of meat spoilage. Eng. Appl. Artif. Intell. 323, 303–310. https://doi.org/10.1007/978-3-319-11310-4_26.

Liu, Q., Hu, X., Ye, M., Cheng, X., Li, F., 2015. Gas recognition under sensor drift by using deep learning. Int. J. Intell. Syst. 30, 907–922. https://doi.org/10.1002/int.21731.

Liu, Q., Li, X., Ye, M., Ge, S.S., Du, X., 2014. Drift compensation for electronic nose by semi-supervised domain adaption. IEEE Sens. J. 14, 657–665. https://doi.org/10.1109/JSEN.2013.2285919.

Mohareb, F., Papadopoulou, O., Panagou, E., Nychas, G.-J., Bessant, C., 2016. Ensemble-based support vector machine classifiers as an efficient tool for quality assessment of beef fillets from electronic nose data. Anal. Methods 8, 3711–3721. https://doi.org/10.1039/c6ay00147e.

Najam ul Hasan, Ejaz N., Ejaz, W., Kim, H.S., 2012. Meat and fish freshness inspection system based on odor sensing. Sensors (Switzerland) 12, 15542–15557. https://doi.org/10.3390/s121115542.

Nettleton, D.F., Orriols-Puig, A., Fornells, A., 2010. A study of the effect of different types of noise on the precision of supervised learning techniques. Artif. Intell. Rev. 33, 275–306. https://doi.org/10.1007/s10462-010-9156-z.

Nychas, G.J.E., Skandamis, P.N., Tassou, C.C., Koutsoumanis, K.P., 2008. Meat spoilage during distribution. Meat Sci. 78, 77–89. https://doi.org/10.1016/j.meatsci.2007.06.020.

Panigrahi, S., Balasubramanian, S., Gu, H., Logue, C.M., Marchello, M., 2006. Design and development of a metal oxide based electronic nose for spoilage classification of beef. Sens. Actuat. B 119, 2–14. https://doi.org/10.1016/j.snb.2005.03.120.

Papadopoulou, O., Panagou, E.Z., Tassou, C.C., Nychas, G.J.E., 2011. Contribution of Fourier transform infrared (FTIR) spectroscopy data on the quantitative determination of minced pork meat spoilage. Food Res. Int. 44, 3264–3271. https://doi.org/10.1016/j.foodres.2011.09.012.

Papadopoulou, O.S., Panagou, E.Z., Mohareb, F.R., Nychas, G.J.E., 2013. Sensory and microbiological quality assessment of beef fillets using a portable electronic nose in tandem with support vector machine analysis. Food Res. Int. 50, 241–249. https://doi.org/10.1016/j.foodres.2012.10.020.

Pineda, D.M., Pérez, J.C., 2017. SENose: An under U$50 electronic nose for the monitoring of soil gas emissions. Comput. Electron. Agric. 133, 15–21. https://doi.org/10.1016/j.compag.2016.12.004.

PrimeSafe, 2002. Microbiological testing for process monitoring in the meat industry.

Purslow, P.P., 2017. Chapter 1 - Introduction. In: Purslow, P.P. (Ed.), New Aspects of Meat Quality: From Genes to Ethics. Elsevier Ltd, pp. 1–9. https://doi.org/10.1016/B978-0-08-100593-4.00001-1.

Ridoean, J.A., Sarno, R., Sunaryo, D., Wijaya, D.R., 2018. Music mood classification using audio power and audio harmonicity based on MPEG-7 audio features and Support Vector Machine, in: Proceeding - 2017 3rd International Conference on Science in Information Technology: Theory and Application of IT for Education, Industry and Society in Big Data Era, ICSITech 2017. IEEE, Bandung. doi:10.1109/ICSITech.2017.8257088.

Sáez, J.A., Galar, M., Luengo, J., Herrera, F., 2014. Analyzing the presence of noise in multi-class problems: alleviating its influence with the one-vs-one decomposition. Knowl. Inf. Syst. 38, 179–206. https://doi.org/10.1007/s10115-012-0570-1.

Sáez, J.A., Krawczyk, B., Woźniak, M., 2016. On the influence of class noise in medical data classification : treatment using noise filtering methods on the influence of class noise in medical data classification : treatment using noise filtering methods. Appl. Artificial Intelligence 9514. https://doi.org/10.1080/08839514.2016.1193719.

Sans, P., Combris, P., 2015. World meat consumption patterns : An overview of the last fifty years (1961 – 2011).

Sarno, R., Ridoean, J.A., Sunaryono, D., Wijaya, D.R., 2018. Classification of music mood using MPEG-7 audio features and SVM with confidence interval. Int. J. Artif. Intell. Tools 27. https://doi.org/10.1142/S0218213018500161.

Schaller, E., Bosset, J.O., Escher, F., 1998. "Electronic Noses" and their application to food. Lebensmittel-Wissenschaft und-Technologie 31, 305–316. https://doi.org/10.1006/fstl.1998.0376.

Sun, Y., Wang, J., Cheng, S., 2017. Discrimination among tea plants either with different invasive severities or different invasive times using MOS electronic nose combined with a new feature extraction method. Comput. Electron. Agric. 143, 293–301. https://doi.org/10.1016/j.compag.2017.11.007.

Tian, F., Yang, S.X., Dong, K., 2005. Circuit and noise analysis of odorant gas sensors in an E-nose. Sensors 5, 85–96. https://doi.org/10.3390/s5010085.

Vergara, A., Vembu, S., Ayhan, T., Ryan, M.A., Homer, M.L., Huerta, R., 2012. Chemical gas sensor drift compensation using classifier ensembles. Sens. Actuat. B 166–167, 320–329. https://doi.org/10.1016/j.snb.2012.01.074.

Wang, D., Miao, D., Xie, C., 2011. Best basis-based wavelet packet entropy feature extraction and hierarchical EEG classification for epileptic detection. Expert Syst. Appl. 38, 14314–14320. https://doi.org/10.1016/j.eswa.2011.05.096.

Wijaya, D.R., Sarno, R., 2015. Mobile electronic nose architecture for beef quality detection based on internet of things technology. GlobalIlluminators, Bandung, pp. 655–663.

Wijaya, D.R., Sarno, R., Daiva, A.F., 2017a. Electronic Nose for Classifying Beef and Pork using Naïve Bayes. International Seminar on Sensor, Instrumentation, Measurement and Metrology (ISSIMM) Surabaya.

Wijaya, D.R., Sarno, R., Zulaika, E., 2018. Electronic nose dataset for beef quality monitoring in uncontrolled ambient conditions. Data in Brief. https://doi.org/10.1016/j.dib.2018.11.091.

Wijaya, D.R., Sarno, R., Zulaika, E., 2016a. Sensor array optimization for mobile electronic nose: wavelet transform and filter based feature selection approach. International Review Comput. Software 11, 659–671. https://doi.org/10.15866/irecos.v11i8.9425.

Wijaya, D.R., Sarno, R., Zulaika, E., 2016b. Gas concentration analysis of resistive gas sensor array. In: 2016 IEEE International Symposium on Electronics and Smart Devices. IEEE, Bandung, pp. 337–342. https://doi.org/10.1109/ISESD.2016.7886744.

Wijaya, D.R., Sarno, R., Zulaika, E., 2016c. Information quality ratio as a novel metric for mother wavelet selection. Chemometrics Intelligent Laboratory Syst. 160, 59–71. https://doi.org/10.1016/j.chemolab.2016.11.012.

Wijaya, D.R., Sarno, R., Zulaika, E., Sabila, S.I., 2017b. Development of mobile electronic nose for beef quality monitoring. In: 4th Information Systems International Conference 2017, ISICO 2017, Procedia Computer Science. Elsevier B.V, Bali, pp. 728–735. https://doi.org/10.1016/j.procs.2017.12.211.

Wijaya, D.R., Sarno, R., Zulaika, E., Sabila, S.I., 2017c. Development of mobile electronic nose for beef quality monitoring. Procedia Comput. Sci. 124211, 728–735. https://doi.org/10.1016/j.procs.2017.12.211.

Wojnowski, W., Majchrzak, T., Dymerski, T., Gębicki, J., Namieśnik, J., 2017. Electronic noses: powerful tools in meat quality assessment. Meat Sci. 131, 119–131. https://doi.org/10.1016/j.meatsci.2017.04.240.

Yan, K., Zhang, D., 2016. Calibration transfer and drift compensation of e-noses via coupled task learning. Sens. Actuat. B 225, 288–297. https://doi.org/10.1016/j.snb.2015.11.058.

Zanchettin, C., Ludermir, T.B., 2007. Wavelet filter for noise reduction and signal compression in an artificial nose. Appl. Soft Comput. J. 7, 246–256. https://doi.org/10.1016/j.asoc.2005.06.004.

Zaragozá, P., Fernández-Segovia, I., Fuentes, A., Vivancos, J.L., Ros-Lis, J.V., Barat, J.M., Martínez-Máñez, R., 2014. Monitorization of Atlantic salmon (Salmo salar) spoilage using an optoelectronic nose. Sens. Actuat. B 195, 478–485. https://doi.org/10.1016/j.snb.2014.01.017.

Zhang, L., Liu, Y., He, Z., Liu, J., Deng, P., Zhou, X., 2017. Anti-drift in E-nose: a subspace projection approach with drift reduction. Sens. Actuat. B 253, 407–417. https://doi.org/10.1016/j.snb.2017.06.156.

Zhang, L., Zhang, D., 2015. Domain adaptation extreme learning machines for drift compensation in E-nose systems. IEEE Trans. Instrum. Meas. 64, 1790–1801. https://doi.org/10.1109/TIM.2014.2367775.

Zhu, X., Wu, X., 2004. Class noise vs. attribute noise: a quantitative study. Artif. Intell. Rev. 22, 177–210. https://doi.org/10.1007/s10462-004-0751-8.