Research Article

Application of Electronic Nose to Predict the Optimum Fermentation Time for Low-Country Sri Lankan Tea

Tharaga Sharmilan ⊗,1 Iresha Premarathne,1 Indika Wanniarachchi ⊗,2 Sandya Kumari ⊗,3 and Dakshika Wanniarachchi ⊗,4

1Instrument Center, Faculty of Applied Sciences, University of Sri Jayewardenepura, Nugegoda 10250, Gangodawila, Sri Lanka
2Department of Physics, Faculty of Applied Sciences, University of Sri Jayewardenepura, Nugegoda 10250, Gangodawila, Sri Lanka
3Department of Science & Technology, Faculty of Applied Sciences, Uva Wellassa University of Sri Lanka, Badulla 90000, Sri Lanka
4Department of Chemistry, Faculty of Science, University of Kelaniya 11300, Colombo, Sri Lanka

Correspondence should be addressed to Dakshika Wanniarachchi; dakshikacw@sjp.ac.lk

Received 11 October 2021; Revised 28 November 2021; Accepted 25 January 2022; Published 23 February 2022

Academic Editor: Igor Tomasevic

Copyright © 2022 Tharaga Sharmilan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The fermentation stage is vital during the black tea manufacturing process to produce the best-quality tea. The oxidation of tea biochemical compounds results in the appearance of characteristic smell peaks during the fermentation stage. These subtle changes in tea aroma are hard to detect unless one is a trained personnel. Here for the first time, we applied e-nose to monitor the fermentation process of Sri Lankan low-country tea. In this study, detection of smell peaks during fermentation was conducted by a custom-made e-nose (Digi-Nose) with four gas sensors. Singular value decomposition (SVD) is applied to eliminate the noise and dimensionality reduction in the sensor responses observed. The prediction of the time of appearance of smell peaks was conducted with a support-vector machine (SVM). Finally, theaflavin content with time was compared to validate the optimum fermentation times observed with an e-nose.

1. Introduction

Tea is a popular beverage due to its stimulating effects and health benefits consumed by people all around the world. There are several different tea varieties such as black tea, green tea, and oolong tea based on the level of oxidation of polyphenolic compounds. Black tea is fully oxidized tea. The production of black tea has several steps such as withering, rolling, fermentation, drying, and finally sorting and packaging. Among them, fermentation is the critical stage and it plays an important role in determining the quality of final black tea.

In this stage, chemical constituents and enzymes react in the presence of oxygen to produce polyphenolic compounds due to the stress initiated by plant cell rupture [1]. In addition, physical parameters (humidity and temperature) and thicknesses of the fermentation bed have a significant impact on the quality of tea produced [2]. The fermentation stage results increase in theaflavin (TF) and thearubigin (TR) with time due to the oxidation of catechins and their gallates. [3, 4]. TR content of tea is always considerably higher than the TF, and the preferred ratio of TR: TF is 10:1 [4]. The levels of TF increase during fermentation and reach a maximum point and then decrease while thearubigin is increasing [3, 4] with time. The golden color of tea liquor corresponds to high TF content, while the dull color in tea liquor is a result of over fermentation [5]. If the tea particles are underfermented or overfermented, it leads to less tea quality. Therefore, finding optimum fermentation time is important [3, 4].

The aroma of the brewed tea is another important factor setting the price for tea. The aroma of tea particles during the
fermentation stage consists of a wide variety of compounds such as aldehydes, ketones, alcohols, alkanes, alkenes, and esters. These compounds originate from fatty acid derivatives, monoterpenes, carotenoids, phenylpropanoids, etc. [7–10] The most abundant aroma compounds in black tea are (E)-2-hexenal, hexanal, (E)-geraniol, linalool, linalool oxide II, benzeneacetaldehyde, linalool oxide I, benzaldehyde, methyl salicylate, and 3,7-dimethyl-1,5,7-octatriene-3-ol [9, 11]. The most prominent variation in the volatile compounds of black tea aroma is observed with the alcohol. In addition, the aldehydes such as (E)-2-hexenal and hexanal and the alcohol compounds of black tea aroma are observed with the aldehyde, hyde, methyl salicylate, and 3,7-dimethyl-1,5,7-octatriene-3-ol II, benzeneacetaldehyde, linalool oxide I, benzaldehyde, methyl salicylate, and 3,7-dimethyl-1,5,7-octatriene-3-ol [7–10].

Identification of aroma compounds that presents during tea fermentation was performed with gas chromatography-mass spectrometry (GC-MS). Their characteristics (floral, grassy, sweet, etc.) were identified using gas chromatography-olfactometry (GC-O) [6–9]. GCMS has high sensitivity, and over 70 different chemical compounds present in tea aroma were identified in different studies [4, 6, 12]. However, GCMS instruments are not found in many tea factories due to the high price of the instrument. Since oxidation will continue on the way from the factory to the laboratory (unless the sample is frozen), GC-MS has limitations to adapt into regular quality checking in the production line. Nowadays, these conventional techniques are used to compare the results of electronic devices when calibration is required. At present, optimum fermentation time is detected by humans observing the color changes in tea particles from green to copper brown and by smelling tea particles to detect the development of a fruity smell that appears after optimum fermentation in Sri Lankan tea factories. But detecting optimum fermentation time will vary from person to person; thus, consistency cannot be maintained. Therefore, there is a need to develop a system to monitor tea fermentation with minimum human intervention.

Innovative technologies were developed by several research groups using the electronic eye method [5, 11, 13–18], the electronic tongue method [19–22], and the electronic nose method [23–35] that aim at minimizing human intervention. At present, the electronic nose methods are often utilized to classify different types of tea or different tea grades [24, 26, 31, 35]. The e-nose systems developed so far are based on metal oxide gas sensors, which are inherently difficult to adapt to the high-humidity environment of tea manufacturing. Thus, studies conducted under factory conditions are limited. [28, 29]. The pioneering work of Bhattacharya and coworkers to detect first- and second-nose smell peaks was validated with a colorimetric test of tea infusion [25]. The detection of changes in aroma profiles during the fermentation stage requires experienced staff, yet there are greater possibilities to make errors. Therefore, e-nose technology offers a way of maintaining uniformity and consistency in fermentation monitoring. However, studies conducted on finding smell peaks and their associations with TF values are limited and such studies were not conducted in Sri Lankan black tea. [29] Therefore, this study is focused on monitoring black tea fermentation in a low-country wet zone tea factory aiming at the identification of smell peaks during the fermentation process.

A custom-built e-nose system is used in this regard with an array of metal oxide gas sensors. The gas sensors selected were based on the aroma of tea compounds; hence, gas sensors specific for alcohols, such as alkane, were used here. The dimensionality reduction in the data collected was performed with the singular value decomposition method, and peaks were identified with the order distance filter method. The main objective of this study is to identify the presence of smell peaks during the fermentation stage, which replaces human detection. Correct identification of the optimum position where the TF level is maximum is critical to produce quality tea. The first and second smell peaks were identified, and correlation analysis was conducted for selected batches with the TF content.

2. Materials and Methods

2.1. Sensor Selection. The main compounds present in tea aroma and flavor are linalool, geraniol, phenylacetaldehyde, benzaldehyde, methyl salicylate, and hexanal [7–9]. These are mostly aldehydes, ketones, esters, hydrocarbons, and esters. Therefore, when developing the “e-nose” system, it is important to choose sensors that are responsive to the abovementioned chemical compounds with excellent sensitivities to detect subtle changes such as “first nose” and “second nose” as practiced in the factory. Our custom-made e-nose system contains an array of four MOS sensors to record the sensor profile. Sensors used in the sensor array are given in Table 1.

In order to validate the sensitivity of the sensor array, a series of organic solvents were selected and the effect of the functional group was investigated on the sensitivity of the sensor array in the previous study [36, 37]. In that study, it was concluded that sensors in the e-nose system can classify different chemicals with different functional groups.

2.2. Development of Digi-Nose. Custom-developed [38] electronic nose (Digi-Nose), as shown in Figure 1, has been used to monitor the emission of volatile compounds during the fermentation process. The electronic nose system consists of (a) data acquisition hardware, (b) gas sensor chamber, and (c) vacuum pumps. Arduino-related hardware is used to design the data acquisition system, and the outputs of the sensors are acquired in the SD card. The sensor chamber is an airtight waterproof case that houses the sensors. As mentioned in the previous section, four MOS gas sensors have been used to capture the odor from the fermentation process. Sensor chamber and vacuum pumps are connected using transparent pipes in order to supply the tea aroma to the gas sensors. Two inlets have been used to insert the tea aroma and environment air into the sensor chamber. Vacuum pumps (12 V) have been used to draw the tea aroma and environment air to be analyzed. The sensor chamber is cleaned by environment air between two consecutive sample collections.
2.3. Tea Fermentation Monitoring Using Digi-Nose. Tea
samples are collected from Sri Lankan low-country tea
factory in the Avissawella region, which manufactures or-
thodox tea. Only dhool 1 (first set of tea particles after the
rolling stage) of each batch has been selected for this study.
Leaves undergo 10–12 hr withering followed by rolling for
20 min and separated as dhool 1. Experimental sniffing cycle
was limited to 3 minutes and contains the following stages: 1
minute for sniffing process, 1 minute for odor lock, and 1
minute for sensor cleaning. One minute cleaning time is
given at the end of each sniffing cycle to clear any residual
particles in the sensor chamber. Environment air was used to
clean the sensor chamber, and the sample air inlet was closed
while the sensor chamber was cleaning. The device continues
to collect the data until the experimental process is ter-
ninated by the user.

A random tea sample of dhool 1 is immediately collected
after the rolling stage and placed in a beaker (200 ml) with
10 cm thickness to identify the smell peaks by electronic nose.
Altogether, 48 fermentation cycles have been collected during
this process. Data collection is continued until the batch is sent
for firing. Aroma detection was performed using developed
Digi-Nose outside of the fermentation area. The details of
collected samples carried out in this study are listed in Table 2.
A sample sniffing cycle of Digi-Nose is given in Figure 2.

2.4. Singular Value Decomposition. SVD is a statistical tool
for dimensionality reduction and noise elimination in signal
processing. The SVD represents an expansion of the original
data in a coordinate system where the covariance matrix is
diagonal. Calculating the SVD consists of finding the ei-
genvalues and eigenvectors of AAT and AT A. The eigen-
values of AT A make up the columns of V, and the
eigenvectors of AAT make up the columns of U. Also, the
singular values in S are square roots of eigenvalues from
AAT or AT A. The singular values are the diagonal entries of
the S matrix and are arranged in descending order. The
singular values are always real numbers. If matrix A is a real
matrix, then U and V are also real. [39, 40].

2.5. Peak Detection Algorithm. The “scipy.signal.argrelex-
trema” algorithm has been used to find the peak points in this
study. It is called as order (distance) filter algorithm. The
minimum distance is used as a filter in this algorithm. It is a new
peak detection algorithm from Scipy scikit-learn version 0.11.0.
It calculates the relative extrema of data. Data, comparator, axis,
order, mode, and extrema are the parameters of this algorithm
[41]. It includes an order parameter that can serve as a kind of
minimum distance filter. Parameters, returns, and explanation
of those parameters are given as follows:

“scipy.signal.argrelextrema (data, comparator, axis = 0,
order = 1, and mode = “clip”)

(i) Data: array in which to find the relative extrema.
(ii) Comparator: function to use to compare two data
points. Two arrays should be taken as arguments.
(iii) Axis: axis over which to select from data. Default is
0.
(iv) Order: how many points on each side to use for the
comparison to consider comparator(n, n + x) to be
true.
(v) Mode: how the edges of the vector are treated. "wrap"
(wrap around) or "clip" (treat overflow as the same as
the last (or first) element). Default is "clip."
(vi) Extrema: indices of the maxima in arrays of integers.
Extrema[k] is the array of indices of axis k of data.
The return value is a tuple even when data are 1-D.

![Table 1: Sensors used in the sensor array.](image)

| Sensor | Sensitive substances | Detection range |
|--------|----------------------|-----------------|
| MQ2    | LPG, isobutane, propane, methane, hydrogen, smoke, and alcohol | 200 ppm–5000 ppm (LPG and propane), 300 ppm–5000 ppm (butane), 5000 ppm–20000 ppm (methane), 300 ppm–5000 ppm (H2), and 100 ppm–2000 ppm (alcohol) |
| MQ3    | Alcohol and benzene  | 0.05 mg/L–10 mg/L (alcohol) |
| MQ4    | Methane              | 300–10000 ppm (CH4) |
| MQ5    | H2, LPG, and alcohol | CH4, CO 200–10000 ppm |

![Figure 1: (a) E-nose system developed in this study. (b) The sensor chamber [37].](image)
The filtering behavior is customizable through the comparator parameter, which can make it customizable for building our own filtering algorithm over it. Therefore, order 3 is selected for building the algorithm for the detection peak points in this study.

2.6. Theaflavin Content Analysis. Sodium carbonate (anhydrous) (assay 99.9%, Thermo Fisher Scientific), gallic acid (Sisco Research Laboratories, India, 98%), and methanol (assay 99.8%, Sisco Research Laboratories, India) were used to determine the total polyphenol content in tea samples. UV-visible spectrophotometer (Implen GmbH, Germany) was used in spectroscopic analysis. The moisture content of the tea sample was measured using grounded and sieved tea particles (size of 595 \( \mu \text{m} \) – 841 \( \mu \text{m} \)). Then, 2.000 ± 0.001 g of tea sample (mi) was oven-dried at 103°C until constant weight (mf) was obtained, which is used in the calculation of theaflavin.

The measurement of the theaflavin content was conducted by adopting the method established by Robert and Smith [42], where a known amount (2.25 g) of tea sample was added to 250 mL of volumetric flask and 100 mL of hot distilled water was poured into it. The sample was boiled for 10 min under an 85°C water bath. The extraction was filtered through cotton wool and allowed it to reach room temperature. The extracted sample (25 mL) was shaken with 25 mL of ethyl acetate using the separation funnel and allowed to separate. The separated ethyl acetate layer (12.5 mL) was vigorously shaken with 2.5% of aqueous sodium hydrogen carbonate (12.5 mL) for 30 seconds and allowed to separate the layer. About 4 mL portion of ethyl acetate was diluted with methanol up to 25 mL in the volumetric flask. The blank sample was prepared using the same procedure without a tea sample. The absorbance of the solution was obtained using the UV-Vis spectrophotometer at (380 nm) to calculate the value of TF.

2.7. Correlation Analysis. Correlation analysis was conducted to find the existence of the relationship between the peaks observed with Digi-Nose and biochemical analysis. Python programming language was used using "scikit-learn" free software machine learning library for the correlation analysis. From the results, the Pearson correlation coefficient is used to examine the strength and direction of the linear relationship between two variables. The correlation coefficient can range in value from \(-1\) to \(+1\). The larger the absolute value of the coefficient, the stronger the relationship between the variables. The Pearson correlation, with an absolute value of 1, indicates a perfect linear relationship. A correlation close to 0 indicates no linear relationship.

### Table 2: Experiment details of fermentation cycles.

| Location and tea type          | Year | Month | No. of batches collected |
|-------------------------------|------|-------|-------------------------|
| Avissawella tea factory, Sri Lanka | 2019 | October | 12                      |
|                               |      | November | 05                      |
|                               |      | December | 18                      |
| Type-orthodox                 | 2020 | July    | 13                      |

![Figure 2: A sample single sniffing cycle of the Digi-Nose system [37].](image-url)
between the variables. The sign of the coefficient indicates the direction of the relationship. If both variables tend to increase or decrease together, the coefficient is positive, and the line that represents the correlation slopes upward. If one variable tends to increase as the other decreases, the coefficient is negative, and the line that represents the correlation slopes downward. To determine whether the correlation between variables is significant, the p-value is compared to the significance level. Usually, a significance level (denoted as or alpha) of 0.05 works well. An alpha of 0.05 indicates the risk of concluding that a correlation exists. If the p-value is less than or equal to the significance level, the correlation is statistically significant. If the p-value is greater than the significance level, then the correlation is not statistically significant [43].

2.8. Principal Component Analysis (PCA). It is one of the feature extraction techniques and is used for the dimensionality reduction process. Jupyter Notebook was used to analyze the dataset in this study. The dataset should be scalable when performing PCA. Therefore, data were standardized on to unit scale for the optimal performance of the machine learning algorithm. About 70% of the dataset has been used to make the model. About 30% were used to test the model. The total explained variance ratio was found to provide the amount of variance each principal component has after doing dimensionality reduction. Then, the visualization of data has been plotted to identify the distinguished peaks. Principal components can be visualized according to the distribution of the peaks. Then, scatter plots have been created from the principal components to see the separation of three peaks.

2.9. Support-Vector Machine (SVM). It is a supervised machine learning algorithm and a very good tool for the classification of problems. It is effective in high-dimensional spaces, and different kernel functions can be specified for the decision function.

The SVM is usually implemented using kernel as it transforms input data space into the required form. Therefore, this kernel trick helps to build a more accurate classifier. There are three common kernels used in the classification, such as linear kernel, polynomial kernel, and radial basis function kernel. The linear kernel can be used as a normal dot product of any two given observations, and the polynomial kernel is a more generalized form of the linear kernel. The radial basis function (RBF) kernel is a popular one and is commonly used in SVM classification problems as it can map an input space in infinite-dimensional space [44].

When training the model with radial basis function (RBF) kernel, two hyperparameters should be set before training the model, such as C and gamma. The parameter C is common to all SVM kernels, which ranges from 0.1 to 100. It is used to control the error as it maintains regularization. Therefore, misclassification can be avoided. The parameter gamma is used to give the curvature weight of the decision boundary. Gamma can be tuned. It depends upon the data, which ranges from 0.0001 to 10. A higher value of gamma will perfectly fit the training dataset, which causes over-fitting. Therefore, good values for the C and gamma need to be found out to specify the learning algorithm. The classification has been implemented in Python using the scikit-learn library to estimate how accurately the model can predict the smell peaks.

2.10. Model Evaluation Metrics. These are used to assess the algorithm’s performance in supervised learning. The measured performance is interpreted in terms of accuracy, recall, precision, and F1 score. A confusion matrix is one of the methods used to calculate the metrics (Figure 3) [36].

As shown in Table 3, true positive (TP), false positive (FP), true negative (TN), and false negative (FN) parameters are used in the confusion matrix.

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}
\]
\[
\text{Precision} = \frac{TP}{TP + FP}
\]
\[
\text{Recall} = \frac{TP}{TP + FN}
\]
\[
\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}
\]

The accuracy of an algorithm is represented as the ratio of correctly classified sensor data (TP + TN) to the total number of sensor data (TP + TN + FP + FN). Precision is the capability of a classifier not to tag a false positive instance of sensor data points that is really negative. Another term is recall, which is the ability of our algorithm to detect all positive sensor data. F1 score is a weighted harmonic mean of the abovedescribed precision and recall.

3. Results and Discussion

3.1. Digi-Nose Data. The data collection was initiated as soon as the dhool 1 tea particles are laid on containers for fermentation. Digi-Nose sensor dataset has four columns as the device has four sensors. Each sensor values contain the 3-min interval as the device has 3-min time period for the sniffing cycle. In the beginning, the first two sniffing cycles have been skipped to reduce the errors coming from the device. Sensor resistance values have been collected during the entire fermentation process, with a 5-second interval during sniffing, odor lock, and cleaning.

Static change in sensor resistance was used \((\Delta R = R_{\text{Environment Air}} - R_{\text{Sample Air}})\) to preprocess the collected data. Then, the complete sets of sensor data were normalized. The stable sensor data are obtained during the odor lock region. Thus, an average of every 5-s data during the odor lock was used to represent one sniffing cycle. Therefore, this process can eliminate the noise. Figure 3 illustrates averaged odor lock sensor data of a representative batch. All four sensor values indicate a similar pattern except MQ5, which shows minimum variation.
These normalized odor lock data values were used for the singular value decomposition (SVD) process for further noise elimination. Figure 4 shows SVD-processed data for all four sensors. SVD is applied here for the purpose of dimensionality reduction and noise reduction. All these signals' pre-processing was conducted using Python software. But, the intensity of smell peaks is important to identify the optimum fermentation time. The peaks present in SVD-processed data were extracted with the order (distance) filter algorithm. The first three peaks detected for the 48 fermentation cycles are given in Table S1 Supporting Information.

The clustering of peaks was conducted with principal component analysis that was then carried out for the three peaks isolated from each batch. Figure 5 indicates the smell peak classification.

The first two principal components explain the majority of the variance in this analysis (93.54%) as given in Table 4. Therefore, this is an indication of the total information represented compared to the original data. Then, SVM was used to build the model to classify the smell peaks. The RBF was chosen as the kernel function of the SVM model. A grid search method was used to search the two important parameters with the best performance using 10-fold cross-validation on the training dataset. The parameters included the penalty factor [0.1, 0.5, 1, 5, 10, 50, 100] and the gamma [0.0001, 0.0005, 0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1, 5, 10]. Accuracy was computed here by comparing actual test set values and predicted values by tuning the hyperparameters values. Thus, the penalty factor and gamma were set to 1 and 0.001, respectively. Four evaluation metrics, including accuracy, recall, precision, and F1 score, were computed for the classification model. The SVM achieved an accuracy of 83%, a recall of 83%, a precision of 85%, and an F1 score of 83%, respectively.

The obtained confusion matrix of the classification is given in Figure 6. Looking at the first, second, and third peak columns, first peak smell, second peak smell, and third peak smells are predicted by the model 100%, 75%, and 73%, respectively. Thus, the SVM model provided satisfactory performance for the smell peak classification using the Digi-Nose system.

It is clear that the Digi-Nose system is capable of distinguishing smell peaks during the fermentation process. In previous studies, PCA was utilized for sensors responses at specific times to conduct the PCA. [25, 27, 45]. However, the aroma of tea particles changes due to the variation of aldehydes, alcohols, ketones, and ester compounds during the fermentation stage that results in two different smell peaks. [9, 11] In previous studies, PCA was utilized for fixed time without a peak detection algorithm. [25, 27, 45] Since the fermentation process is sensitive to weather, condition peaks appear at different times. In this study, peaks were detected in comparatively lower time in previous studies, which could be due to differences in tea clones used, climate factors, and processing methods [38, 46–48]. Three separate clusters can be identified based on the three peaks, and the classification rate was 83%. However, the optimum fermentation time strongly depends on the TF level.

Cross-validation of the Digi-Nose results with theaflavin content.

Among 48 batches of e-nose data, only selected fermentation cycles have been used for theaflavin (TF) analysis. The major polyphenolic compounds present in tea leaves are TF and thearubigins (TRs), which contribute to the characteristic color, taste, and aroma of tea [2, 3]. According to the previous studies, TF content and the tea price have a significant relationship. The variation of TF content with fermentation time illustrated a compatible pattern with previous studies [2, 3]. Initially, a rapid increase toward maxima was followed by a decline in fermentation time due to enzyme reaction activity. [3] The oxidative enzymatic reaction is more favorable to form TF at the initial stages of fermentation [5]. However, due to
high enzymatic reaction TF undergoes oxidative polymerization to form TR [2, 3]. Figure S1 in the supporting information gives average TF variation with time for batches selected in this study. It shows a significant maximum peak of around 45 min. However, this time strongly depends on the withering time [49], fermentation condition temperature, and humidity [50]. The optimum fermentation time is considered when the TF: TR ratio reaches 1:10. In previous studies, this optimum fermentation time is observed after the maximum TF peak [2, 3, 28]. Figure 7 indicates a comparison of theaflavin content with each sensor response for the representative batch. In most of the batches, the smell peak 2 appears after the TF maximum, which ensures the reach of optimum fermentation time. In this study, the tea samples were collected from a low-country tea factory that accepts tea from small tea garden owners. There is a considerable variation in the leaf quality. Such that in addition to the two leaves and bud, it is common to find few mature leaves as well. According to a previous study that analyzed the optimum fermentation time, most of the low-country tea gardens have varieties TRI2023, TRI2025, and TRI2043 that range between 45 and 75 mins. [46].

The time where the peak maximum TF appears was noted and compared with smell peaks observed with DigiNose values that are given in Table 5.

3.2. Correlation Analysis. In order to further analyze the significance of each sensor response toward the theaflavin content, correlation analysis was conducted for the batches listed in Table 5. The summary of the correlation analysis of

![Figure 4: Sample of SVD-processed sensor data (comparing level): (a) MQ2; (b) MQ3; (c) MQ4; and (d) MQ5.](image)
the smell peak one with TF content is given in Figure 8. Accordingly, MQ3 has positive correlation with TF \((r = -0.51, p = 0.38)\) while MQ2 \((r = -0.5, p = 0.28)\) and MQ5 \((r = -0.77, p = 0.13)\) have negative correlation with TF. MQ2 has negative intercorrelation with MQ5 \((r = 0.69, p = 0.19)\), while MQ4 has positive intercorrelation with

---

**Table 4: Interpretation of total variance based on principal components.**

| Principal components | Eigenvalue | Variance contribution rate | Contribution rate of cumulative variance |
|----------------------|------------|----------------------------|------------------------------------------|
| Principal component 1 | 3.42       | 84.21                      | 84.21                                    |
| Principal component 2 | 0.38       | 9.33                       | 93.54                                    |
| Principal component 3 | 0.09       | 2.26                       | 95.80                                    |

---

**Figure 5: Classification of smell peaks.**

**Figure 6: Confusion matrix of the SVM classification.**
MQ5 \( (r = 0.54, p = 0.35) \). According to correlation analysis, the MQ2 sensor and MQ5 sensor indicate a strong correlation with TF.

Figure 9 indicates the summary of the smell peak 2, where MQ3 \( (r = -0.46, p = 0.43) \) and MQ2 \( (r = -0.76, p = 0.13) \) have a negative correlation between TF content. MQ3 has multi-intercorrelation with MQ2 \( (r = 0.86, p = 0.06) \) and MQ4 \( (r = 0.56, p = 0.33) \). MQ4 has intercorrelation with MQ5 \( (r = 0.68, p = 0.2) \) and MQ2 \( (r = 0.59, p = 0.29) \). When compared the correlation coefficients obtained with peak 1, the overall correlation of TF with sensor response is weak except for the MQ2 for peak 2.

Furthermore, previous studies conducted indicate an initial grassy smell that appears due to the byproducts of lipid degradation such as (Z)-3-hexenol, hexanal, and (E)-2-hexenal. [8–10, 12, 34] The smell peaks are considered more sweet smell due to compounds originated from glycosides such as linalool, geraniol, and related species. [8–10, 12, 34] According to the intensity of the peaks observed, MQ3 and MQ2 indicated relatively higher intensity in SVD-processed data compared to other sensors since these are for alcohol
detection. However, a detailed study is required correlating the content of esters, aldehydes, ketones, alcohols, and various hydrocarbons present with each of the respective sensors in addition to the comparison with TF value.

In this study, we were able to identify the appearance of smell peaks during the fermentation stage of black tea manufacturing. This is the pioneering study in implementing the e-nose technology to monitor tea production. Furthermore, this technology can be expanded to integrate with the tea tasting and evaluate the quality of the produced tea exported.

This aroma sensing technology can be extended to other industries such as confectioneries, cosmetics, essential oil, and industries where a rapid detection of food quality by aroma is needed.

4. Conclusions

In this study, the fermentation stage of black tea was monitored using a custom e-nose system (Digi-Nose) for a low-country Sri Lankan tea. Sri Lanka is a major supplier to the world tea market, but studies conducted with e-nose devices to monitor the quality of tea produced are not available. Therefore, this serves as a pioneering study to introduce an e-nose system to monitor the fermentation stage. The study recorded the e-nose sensor profile of tea aroma of 48 batches with an average peak of 1 appearing at 21 minutes, peak 2 at 43 minutes, and peak 3 at 65 minutes considering all four sensor values. The system was able to classify the smell peaks detected with 83% accuracy as peak 1, peak 2, and peak 3 with a support-vector machine algorithm. A correlation study of peaks 1 and 2 of each sensor with maximum TF content observed in each batch found a higher correlation with MQ2 and MQ5 sensors. Furthermore, when considering the time of smell peaks, peak 2 appeared past the time of TF maximum. Thus, it can be suggested that the optimum time for fermentation time is past the second smell peak detected by the Digi-Nose. However, it is worth noting the fact that the optimum time depends on the unique climatic and processing conditions in the factory and the demand for a particular tea quality in the market. Hence, the Digi-Nose can be successfully utilized for other tea factories after substantiated with tea biochemical parameters.

Data Availability

Data related to this research (Digi-Nose data and chemical analysis) will be made available upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors would like to thank the Department of Physics, University of Sri Jayewardenepura, Instrument Center, University of Sri Jayewardenepura, and Avissawella tea factory staff for their support to effectively carry out the research. This research was funded by the National Research Council of Sri Lanka, under grant number 17-038 (for research assistant S. Tharaga) and the University of Sri Jayewardenepura, Sri Lanka, under grant number ASP/01/RE/SCI/2019/28 (for research assistant Iresha Premaratne).
Supplementary Materials

Table S1: smell peaks recorded on each sensor. Figure S1: average variation of theaflavin with time. (Supplementary Materials)

References

[1] Z. Naheed, R. Barech, M. Sajid, N. A. Khan, and R. Hussain, "Effect of rolling, fermentation and drying on the quality of black tea," Sarhad Journal of Agriculture, vol. 23, no. 3, 2007.
[2] O. Martin, P. O. Owuor, R. Mang‘oka, and M. M. Kavoi, "Changes in thearubigin fractions and theaflavin levels due to variations in processing conditions and their influence on black tea liquor brightness and total colour," Food Chemistry, vol. 85, no. 2, pp. 163–173, 2004.
[3] K. R. Jolvis Pou, "Fermentation: the key step in the processing of black tea," Journal of Biosystems Engineering, vol. 41, no. 2, pp. 85–92, 2016.
[4] T. Muthumani and R. S. S. Kumar, "Influence of fermentation time on the development of compounds responsible for quality in black tea," Food Chemistry, vol. 101, no. 1, pp. 98–102, 2006.
[5] A. Akuli, A. Pal, B. Gopinath et al., "A machine vision system for estimation of theaflavins and thearubigins in orthodox black tea," International Journal on Smart Sensing and Intelligent Systems, vol. 9, no. 2, 2016.
[6] H. Wu, W. Huang, Z. Chen et al., "GC-MS-based metabolomic study reveals dynamic changes of chemical compositions during black tea processing," Food Research International, vol. 120, pp. 330–338, 2019.
[7] Z. Yang, S. Baldermann, and N. Watanabe, "Recent studies of the volatile compounds in tea," Food Research International, vol. 53, no. 2, pp. 585–599, 2013.
[8] C.-T. Ho, X. Zheng, and S. Li, "Tea aroma formation," Food Science and Human Wellness, vol. 4, no. 1, pp. 9–27, 2015.
[9] X. Q. Zheng, Q. S. Li, L. P. Xiang, and Y. R. Liang, "Recent advances in volatiles of teas," Molecules, vol. 21, no. 3, pp. 1–12, 2016.
[10] X. Chen, "Aroma characterization of Hanzhong black tea (Camellia sinensis) using solid phase extraction coupled with gas chromatography-mass spectrometry and olfactometry and sensory analysis," Food Chemistry, vol. 274, pp. 130–136, 2018.
[11] G. S. Gill, A. Kumar, and R. Agarwal, "Monitoring and grading of tea by computer vision - a review," Journal of Food Engineering, vol. 106, no. 1, pp. 13–19, 2011.
[12] E. Sheibani, S. E. Duncan, D. D. Kuhn, A. M. Dietrich, J. J. Newkirk, and S. F. O’Keefe, "Changes in flavor volatile composition of oolong tea after panning during tea processing," Food Sciences and Nutrition, vol. 4, no. 3, pp. 456–468, 2016.
[13] S. Borah and M. Bhuyan, "Quality indexing by machine vision during fermentation in black tea manufacture- ing," in Proceedings of the Quality Control by Artificial Vision, Gatlinburg, TN, US, 2003.
[14] S. Borah, Machine Vision for Tea Quality Monitoring with Special Emphasis on Fermentation and Grading Emphasis on Fermentation and Grading, Ph.D. Dissertation, 2005.
[15] S. Saranka, T. Kartheeswaran, D. D. C. Wanniarchchi, and W. Wanniarchchi, "Monitoring fermentation of black tea with image processing techniques," Proc. Tech. Sess, vol. 32, pp. 31–37, 2016.
[16] G. Singh and N. Kamal, "Machine vision system for tea quality determination - tea quality index (TQI)," IOSR Journal of Engineering, vol. 3, no. 7, pp. 2250–3021, 2013.
[17] C. Dong, G. Liang, B. Hu et al., "Prediction of congou black tea fermentation quality indices from color features using non-linear regression methods," Scientific Reports, vol. 8, no. 1, pp. 10535–10611, 2018.
[18] Y. Wang, Y. Liu, Q. Cai, L. Li, J. Ning, and Z. Zhang, "Monitoring the withering condition of leaves during black tea processing via the fusion of electronic eye (E-eye), colorimetric sensing array (CSA), and micro-near-infrared spectroscopy (NIRS)," Journal of Food Engineering, vol. 300, Article ID 110534, 2021.
[19] N. Liu, Y. Liang, J. Bin et al., "Classification of green and black teas by PCA and SVM analysis of cyclic voltammetric signals from metallic oxide-modified electrode," Food Analytical Methods, vol. 7, no. 2, pp. 472–480, 2014.
[20] T. Liu, Y. Chen, D. Li, and M. Wu, "An active feature selection strategy for DWT in artificial taste," J. Sensors, vol. 2018, Article ID 9709505, 2018.
[21] P. Saha, S. Ghorai, B. Tudu, R. Bandypadhyay, and N. Bhattacharyya, "A novel technique of black tea quality prediction using electronic tongue signals," IEEE Transactions on Instrumentation and Measurement, vol. 63, no. 10, pp. 2472–2479, 2014.
[22] P. Saha, S. Ghorai, B. Tudu, R. Bandypadhyay, and N. Bhattacharyya, "Tea quality prediction by autoregressive modeling of electronic tongue signals," IEEE Sensors Journal, vol. 16, no. 11, pp. 4470–4477, 2016.
[23] B. Tozlu, H. I. Okumus, and C. Simsek, "Online quality classifying with electronic nose for black tea production," International Journal of Academic Research in Psychology A, vol. 6, no. 4, 2014.
[24] M. Xu, J. Wang, and S. Gu, "Rapid identification of tea quality by E-nose and computer vision combining with a synergistic data fusion strategy," Journal of Food Engineering, vol. 241, pp. 10–17, 2018.
[25] N. Bhattacharyya, S. Seth, B. Tudu et al., "Monitoring of black tea fermentation process using electronic nose," Journal of Food Engineering, vol. 80, no. 4, pp. 1146–1156, 2007.
[26] J. B. Lazaro, A. Ballado, F. P. F Bautista, J. K. B. So, and J. M. J. Villegas, "Chromometric data analysis for black tea fermentation using principal component analysis," AIP Conference Proceedings, vol. 2045, pp. 1–6, 2018.
[27] M. Sharma, D. Ghosh, and N. Bhattacharya, "Electronic Nose-a new way for predicting the optimum point of fermentation of Black Tea," International Journal of Science and Engineering Invention, vol. 2, no. 3, pp. 56–60, 2013.
[28] B. H. Tozlu and H. I. Okumus, "A new approach to automation of black tea fermentation process with electronic nose," Automatika, vol. 59, 2018.
[29] B. Tozlu, C. H I Okumus, and O Aydemir, "On-line monitoring of theaflavins and thearubigins ratio in Turkish black tea using electronic nose," International Journal of Engineering and Applied Sciences, vol. 7, no. 05, pp. 2305–8269, 2015.
[30] S. Buratti, A. Casiraghi, P. Minghetti, and G. Giovanelile, "The joint use of electronic nose and electronic tongue for the evaluation of the sensorial properties of green and black tea infusions as related to their chemical composition," Food and Nutrition Sciences, vol. 04, no. 06, pp. 605–615, 2013.
[31] Z. Qin, X. Pang, D. Chen, H. Cheng, X. Hu, and J. Wu, "Evaluation of Chinese tea by the electronic nose and gas chromatography-mass spectrometry: correlation with sensory
properties and classification according to grade level,” *Food Research International*, vol. 53, no. 2, pp. 864–874, 2013.

[32] H. Jiang and Q. Chen, “Development of electronic nose and near infrared spectroscopy analysis techniques to monitor the critical time in SSF process of feed protein,” *Sensors*, vol. 14, no. 10, pp. 19441–19456, 2014.

[33] S. Ghosh, B. Tudu, N. Bhattacharyya, and R. Bandyopadhyay, “A recurrent Elman network in conjunction with an electronic nose for fast prediction of optimum fermentation time of black tea,” *Neural Computing & Applications*, vol. 31, pp. 1–7, 2017.

[34] T. Sharmilan, I. Premaratne, I. Wanniarachchi, S. Kumari, and D. Wanniarachchi, “Review article electronic nose technologies in monitoring black tea manufacturing process,” *Sensors Hindawi*, vol. 2020, no. 8, 2020.

[35] M. Xu, J. Wang, and L. Zhu, “Tea quality evaluation by applying E-nose combined with chemometrics methods,” *Journal of Food Science and Technology*, vol. 58, no. 4, pp. 1549–1561, 2021.

[36] T. Sharmilan, D. Manohari, I. Wanniarachchi, S. Kumari, and D. Wanniarachchi, “Classification and identification of volatile organic solvents based on functional groups using electronic nose,” *International Journal of Sciences: Basic and Applied Research*, vol. 54, no. 3, pp. 158–173, 2020.

[37] E. A. H. Roberts and R. F. Smith, “Spectrophotometric measurements of theaflavins and thearubigins in black tea liquors in assessments of quality in teas,” *Analyst*, vol. 86, no. 1019, pp. 94–98, 1691.

[38] N. Bhattacharyya, S. Seth, B. Tudu et al., “Genotypic variation in biochemical compounds of the Sri Lankan tea (Camellia sinensis L.) accessions and their relationships to quality and biotic stresses,” *The Journal of Horticultural Science and Biotechnology*, vol. 92, no. 5, pp. 502–512, 2017.

[39] “Singular Value Decomposition (SVD) tutorial. (n.d.-b),” Web.Mit.Edu. Retrieved May, vol. 3, 2021.

[40] R. Bagheri: 2021, https://towardsdatascience.com/understand-singular-value-%20decomposition-and-its-application-in-data-science-388a54be95d.

[41] scipy.signal.findpeaks-SciPy v1.6.3 Reference Guide, The SciPy community, 2021, https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.find_peaks.html.

[42] E. A. H. Roberts and R. F. Smith, “Spectrophotometric measurements of theaflavins and thearubigins in black tea liquors in assessments of quality in teas,” *Analyst*, vol. 86, no. 1019, pp. 94–98, 1691.

[43] 2018, https://support.minitab.com/en-us/minitab-express/1/help-and-how-to/modeling-statistics/regression/how-to/correlation/interpret-the-results/.

[44] D. Yu and Y. Gu, “A machine learning method for the fine-grained classification of green tea with geographical indication using a mos-based electronic nose,” *Foods*, vol. 10, no. 4, 2021.

[45] N. Bhattacharyya, S. Seth, B. Tudu et al., “Detection of optimum fermentation time for black tea manufacturing using electronic nose,” *Sensors and Actuators B: Chemical*, vol. 122, no. 2, pp. 627–634, 2007.

[46] J. D. Kottawa-Arachchi, M. T. K. Gunasekare, M. A. B. Ranatunga, P. A. N. Punyasiri, and L. Jayasinghe, “Use of biochemical compounds in tea germplasm characterization and its implications in tea breeding in Sri Lanka,” *Journal of the National Science Foundation of Sri Lanka*, vol. 41, no. 4, 2013.

[47] T. Samanta, V. Cheeni, S. Das, A. B. Roy, B. C. Ghosh, and A. Mitra, “Assessing biochemical changes during standardization of fermentation time and temperature for manufacturing quality black tea,” *Journal of Food Science and Technology*, vol. 52, no. 4, pp. 2387–2393, 2015.

[48] S. Jayasekera, L. Kaur, A.-L. Molan, M. L. Garg, and P. J. Moughan, “Effects of season and plantation on phenolic content of unfermented and fermented Sri Lankan tea,” *Food Chemistry*, vol. 152, pp. 546–551, 2014.

[49] S. Deb and K. R. Jolivis Pou, “A review of withering in the processing of black tea,” *Journal of Biosystems Engineering*, vol. 41, no. 4, pp. 365–372, 2016.

[50] P. Tang, “Effect of fermentation conditions and plucking standards of tea leaves on the chemical compounds and sensory quality of fermented juice,” *Journal of Chemistry*, vol. 2018, Article ID 4312875, 2018.