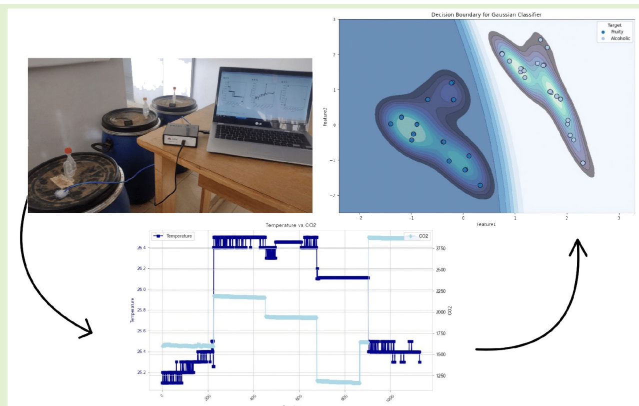
Monitoring and Characterization Technology for Coffee Fermentation Aromas: A Data-Driven Approach

Renato Faraco-Filho, João V. de Castro, Felipe Oliveira Barino, Deivid E. Campos, and Alexandre Bessa dos Santos

***Abstract*—The coffee fermentation process is a crucial step in the production of high-quality coffee. However, the traditional methods of monitoring the fermentation process are subjective, expensive, and time-consuming. In this study, we developed a novel approach using electronic CO_2 and temperature sensors, combined with a 3-D-printed encapsulation, to monitor and characterize the coffee fermentation process. Our methodology involved the following steps: Development of a sensor probe for aroma characterization during the coffee fermentation process, implementation of a robust data processing pipeline, including clustering techniques and principal component analysis (PCA), to effectively identify and classify aromas, collection of data from the sensor probe throughout the fermentation process, and analysis of the collected data to identify patterns and trends that correlate with different stages of the fermentation process. The results of our study demonstrate that our proposed approach is an effective and reliable tool for monitoring and characterizing the coffee fermentation process. Our findings suggest that the use of electronic sensors and data processing techniques has the potential to significantly improve the quality and efficiency of coffee production.**



***Index Terms*—Aroma classification, coffee fermentation, data processing, electronic sensors, principal component analysis (PCA), process monitoring.**

I. Introduction

THE incorporation of advanced technologies into industrial processes has propelled the operational efficiency of companies by enabling more effective data collection and analysis. This innovative industrial environment is driven by advancements in machine learning, the expansion of the Internet of Things (IoT), data mining practices, and the development of high-capacity instrumentation. These elements interact symbiotically, where the increasing demand for data drives the enhancement of high-performance sensors, which, in turn, enables new information processing paradigms.

According to [1], the agricultural sector is undergoing significant transformation due to the adoption of Industry 4.0 principles. In this context, research in agricultural sensors plays a crucial role in enhancing production efficiency and quality. This technological evolution extends not only to agriculture but also to highly influential sectors, such as the coffee industry. Coffee is one of the most globally traded commodities, with coffee consumption reaching 175.6 million bags in 2022 [2]. Furthermore, global coffee consumption has consistently increased, with an annual average of 1.9% over the past 50 years [3]. This scenario underscores the importance of technological innovation not only in agriculture in general but also in key sectors such as coffee, where the pursuit of production improvement is constant and essential for the global economy. Moreover, coffee plays an important role in the Brazilian economy, as the country is the biggest coffee producer and exporter [4].

Manuscript received 5 December 2023; revised 12 January 2024; accepted 14 January 2024. Date of publication 29 January 2024; date of current version 14 March 2024. The work of Renato Faraco-Filho, João V. de Castro, and Deivid E. Campos was supported by the Programa de Bolsas de Pós-Graduação Universidade Federal de Juiz de Fora (PBPG-UFJF) through the Programa de Pós-Graduação em Engenharia Elétrica (PPEE-UFJF), CAPES, and CNPq. The associate editor coordinating the review of this article and approving it for publication was Prof. Xiaofeng Yuan. *(Corresponding author: Renato Faraco-Filho.)*

Renato Faraco-Filho, João V. de Castro, Felipe Oliveira Barino, and Alexandre Bessa dos Santos are with the Department of Circuits, Federal University of Juiz de Fora (UFJF), Juiz de Fora 36036-900, Brazil (e-mail: renato.luiz@engenharia.ufjf.br).

Deivid E. Campos is with the Department of Computational Modeling, Federal University of Juiz de Fora (UFJF), Juiz de Fora 36036-900, Brazil.

Digital Object Identifier 10.1109/JSEN.2024.3357294

Moreover, the intricate relationship between the coffee bean fermentation process and the aroma profile of the beverage is a crucial element to consider. Fermentation plays a vital role in shaping the sensory attributes of coffee, contributing to the diversification and enhancement of its aroma and flavor [5], [6], [7], [8], [9]. This, in turn, directly impacts the value and appreciation of coffee in the global market. However, it is important to note that coffee fermentation is a challenging process due to the lack of precise control. Factors such as fermentation time, temperature, and pH have a significant influence on the final outcome and, therefore, represent important challenges to be addressed to ensure the quality and consistency of the coffee produced [10], [11].

More than a mere technical stage in production, fermentation is a sensory alchemy that transforms raw coffee beans into a truly exceptional tasting experience. Currently, the primary techniques used in the coffee fermentation process rely on human sensory methods [12], [13], as evidenced by a study involving six certified Q-Grader coffee tasters [14]. This evaluation takes place after the completion of the fermentation and roasting processes, where coffee beverage samples corresponding to each treatment are assessed following the Fine Robusta Cupping protocol. Each taster evaluates the coffee beverage five times, assigning scores for characteristics such as fragrance/aroma, flavor, aftertaste, acidity, bitterness/sweetness, and mouthfeel, on a scale of 6–10 points.

Furthermore, Gonzalez-Rios et al. [15] proposed the use of headspace solid-phase microextraction/gas chromatography–mass spectroscopy (HS-SPME/GC–MS) and headspace solid-phase microextraction/gas chromatography–olfactometry (HS-SPME/GC–O) for the characterization of aromas during the coffee fermentation process. Beyond the requisite analytical expertise and the associated relatively high costs of sample preparation, the applicability of these methodologies for continuous monitoring within the food industry is constrained [16], [17].

In this challenging and critical coffee production scenario, the search for innovative and accurate monitoring solutions becomes imperative. It is in this context that electronic odor sensors [18], also known as e-noses, emerge as a promising tool for the coffee industry and other agricultural sectors [19]. Since various studies have shown correlations between chemical properties and odor descriptors [20], [21], [22]. It is important to highlight that e-noses are typically composed of a set of gas sensors that use conductive polymers, metal oxide semiconductors, quartz crystal microbalances, and surface acoustic wave sensors designed to detect a wide variety of volatile organic compounds (VOCs) [23].

Indeed, odor sensors have been successfully used both in the food industry [24], [25], [26] and in the characterization of the coffee roasting process, as evidenced in the study [27], as well as in the coffee fermentation process [28]. In addition, in this context, investment in technological research is imperative to stimulate the development of sustainable solutions for the agricultural sector. Advances in the IoT, sensors and sensor networks, robotics, artificial intelligence, big data, cloud computing, etc. foster the transition toward the Agriculture 4.0 era [29], [30]. In addition, methods such as PCA and clustering have been widely used for process monitoring [31], [32] and also in the food industry [33]. However, it is crucial to emphasize that in the context of coffee fermentation, the use of automated low-cost sensors coupled with data processing represents a novel approach. This is particularly noteworthy considering the various limitations associated with human sensory analysis, including issues related to human fallibility such as physical, physiological, psychological, and health-related factors [34]. In addition, traditional chromatography methods are expensive and typically require laboratory settings for execution.

Fig. 1. Probe used in the experiment, featuring a 3-D-printed PLA enclosure to protect the sensors while allowing the entry of gases generated during the fermentation process.

This article plays a fundamental role in introducing the use of electronic odor sensors in the coffee fermentation process, setting a new standard of accuracy and reliability in this area of study. Our research's main objective is to explore the capabilities of these odor sensors, based on commercial temperature and CO_2 sensors, to monitor and enhance the fermentation process, as these variables play a fundamental role in this process [35]. In addition, we seek to identify how these sensors can contribute to achieving more consistent and high-quality results in coffee production. Throughout this article, we will present details of the methodology, results, and conclusions that illustrate the positive impact of these sensors on coffee production.

II. Methods

A. Device Assembly

The device used in this study was created using electronic temperature and CO_2 sensors, enabling the construction of a sensor array for the e-nose assembly. In addition, a 3-D-printed probe was designed using polylactic acid (PLA) to protect the sensors from coffee beans while allowing the entry of gases generated during the fermentation process. Fig. 1 shows the probe used in the experiment.

For data acquisition, a hardware system was designed to interact with the probe, capturing signals from the temperature and humidity sensors and storing them in a database for future

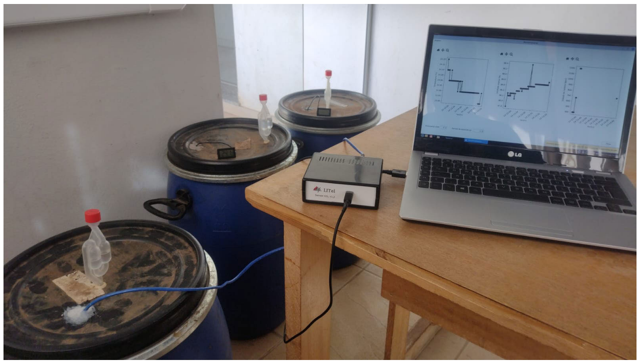


Fig. 2. Experimental setup showing the fermentation barrel and the devices used in this study.

analysis. In conjunction with this equipment, software was also developed to enable real-time data visualization for the operator.

B. Experimental Setup and Data Acquisition

The setup for monitoring the fermentation process included only the designed devices. The probe was positioned in the middle of the barrel, approximately 10 cm below the grain line. In addition, a 50-L coffee fermentation barrel was used, with coffee beans and must already added and ready for fermentation. Fig. 2 depicts the fermentation barrel and the devices used in this study.

After installing temperature and CO_2 concentration sensors, and inoculating the yeast following established best practices for coffee fermentation, we meticulously sealed the fermentation barrel. Throughout the initial 24 h, we maintained a fully closed environment, adhering to the traditional coffee fermentation procedures. Subsequently, we implemented a dual-monitoring strategy, collecting internal temperature and continuous CO_2 data using sensors both in the morning and afternoon. Each acquisition session lasted approximately 3 h, spanning a duration of 96 h. The monitoring parameters were carefully chosen to encompass crucial aspects of the fermentation process, and the fermentation time was correlated with the process, as shown in [28].

This experimental approach faced limitations due to the commercial nature of the barrels used. As a result, we were unable to extend the fermentation process to the point where acetic and putrid odors typically develop, Fig. 3 shows the original obtained dataset.

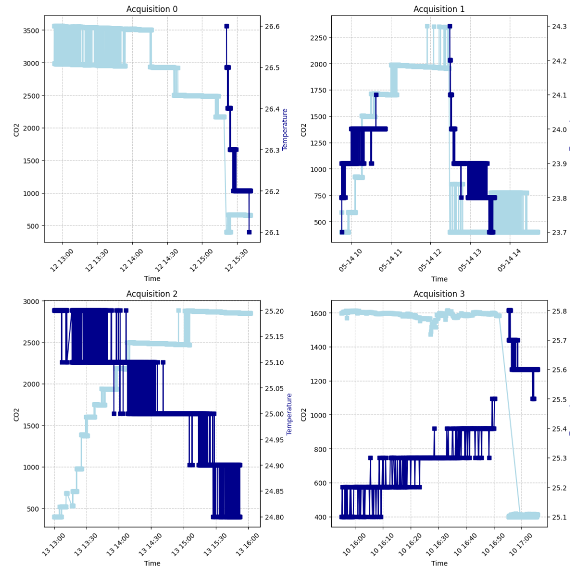


Fig. 3. Original obtained dataset.

C. Data Processing

Data processing plays a fundamental role in the analysis and interpretation of information in many scientific studies, including agriculture [36]. Raw sensor data were read from a CSV file and organized into a data frame with columns for time, CO_2 concentration, and temperature. Subsequently, to improve the quality of the data collected during coffee fermentation process monitoring, the density-based spatial clustering of applications with noise (DBSCAN) clustering technique proposed by [37] was used to perform effective filtering. Noisy and corrupted data, which can occur in specific situations, pose a challenge in obtaining accurate information. For this purpose, parameters such as a minimum sample distance of 300 and a minimum of two samples per neighborhood were used. DBSCAN acted as a valuable tool to identify and isolate data points that did not fit clear patterns or could be considered noise. It is important to note that continuous data collection was not feasible, and CO_2 sensor has a response time. The data were grouped and filtered to obtain measurements after the establishment of the sensor's steady state. Different clusters, using the same parameters, were created for various segments of the temperature and CO_2 data. Subsequently, the filtered data were unified into a single time series. Fig. 4 shows the data extraction process using the clustering method, while Fig. 5 displays the CO_2 and temperature data after processing.

Furthermore, principal component analysis (PCA) [38] was used to investigate possible relationships between the collected data (CO_2 and temperature) and the different phases of the fermentation process. These analyses played a crucial role in understanding the patterns and trends present in the data, contributing to the identification of significant correlations between the monitored variables and specific stages of coffee fermentation.

However, after completing the PCA, an imbalance in the data pertaining to the two fermentation stages was identified. To ensure a uniform distribution between the two data classes, undersampling was applied. This approach involved reducing the majority class, resulting in a set of observations with balanced target variables.

Next, the k-means algorithm was applied to cluster the scattered datasets derived from the variations in the daily means of the monitored variables. This allowed the calculation of the mean value for each cluster. It is worth noting that the number of clusters was determined using the elbow method, as exemplified in Fig. 6. Based on these clusters and their

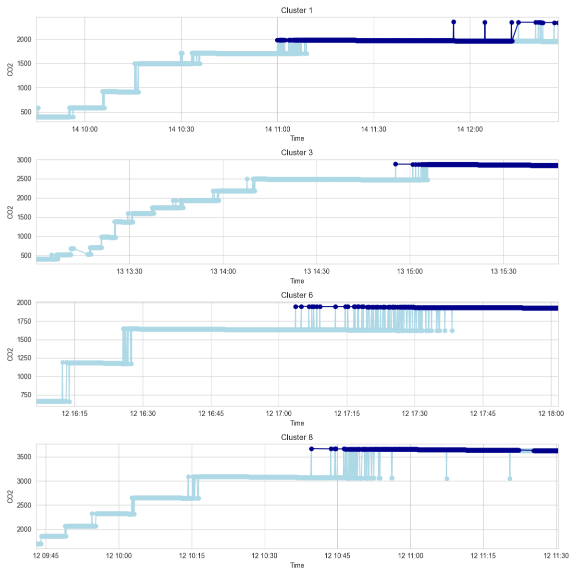


Fig. 4. Clusters used for noise reduction in the data.

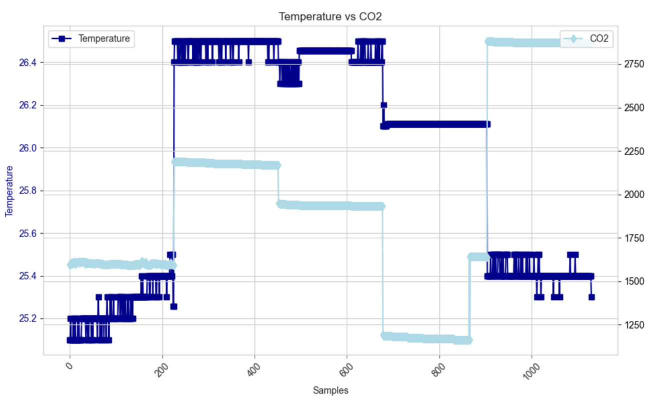


Fig. 5. Processed CO_2 and temperature data after clustering.

mean values, a new dataset was constructed. Fig. 7 presents the overlap of the mean values for each cluster, along with their respective error bars (average standard deviation of 0.062 in PC1 and 0.064 in PC2).

In addition, to avoid excessive duplication of identical data and minimize the potential problem of overfitting, the synthetic minority over-sampling technique (SMOTE) technique was applied. This approach synthesizes new records based on the existing ones, creating "synthetic" data that are similar but not identical to real data [4]. This strategy is advantageous as it maintains data diversity and helps avoid information loss. Fig. 8 illustrates the dataset after k-means clustering, and Fig. 9 shows the final dataset after over-sampling using the SMOTE technique. In addition, Fig. 10 illustrates the flowchart depicting the data processing method used.

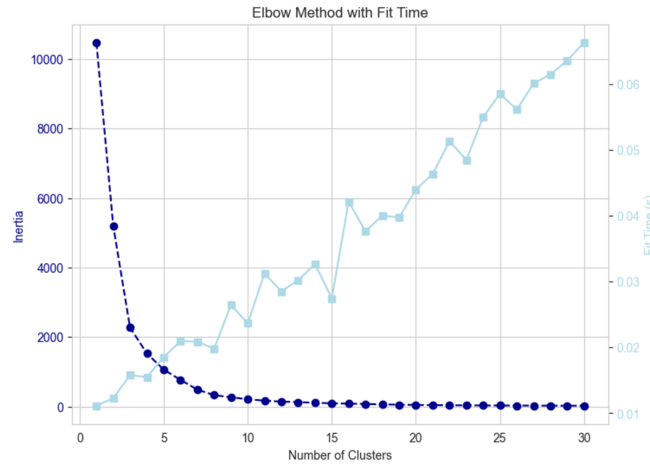


Fig. 6. Elbow method used to determine the number of clusters.

Fig. 7. Construction of the new dataset based on cluster values.

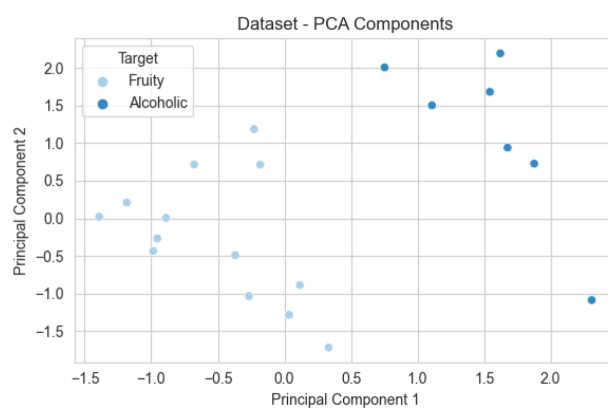


Fig. 8. Dataset after clustering.

III. Results and Discussion

To obtain a comprehensive preliminary visualization, the mean concentrations of each parameter per day were calculated, establishing associations with the fermentation stages of coffee, which can be seen in Fig. 11, where the dark blue background represents the stage where coffee has a fruity aroma, and the light blue background represents the stage of alcoholic aroma. It was observed that during the transition between fruity and alcoholic aromas, there was a significant increase in CO_2 and temperature levels, which then decreased until the optimal alcoholic aroma was reached. It is important to note that CO_2 levels reduced more abruptly throughout the process.

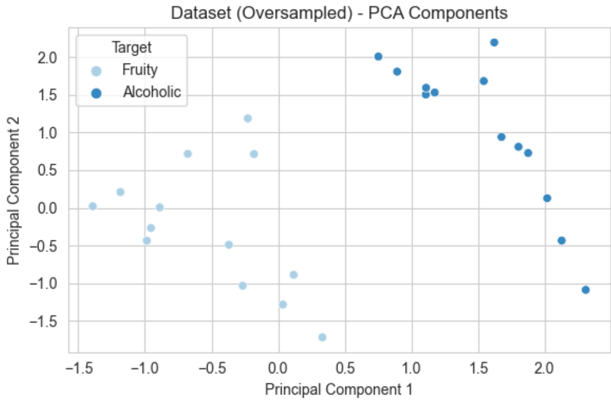


Fig. 9. Dataset after oversampling using the SMOTE technique.

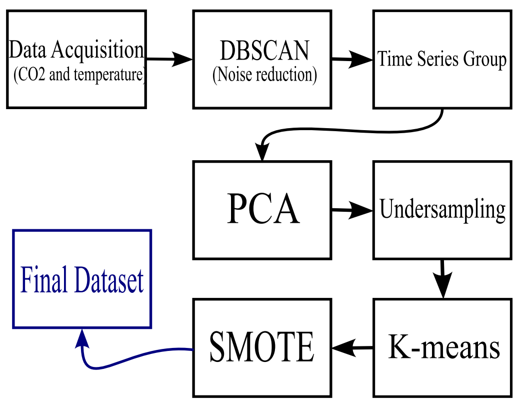


Fig. 10. Data processing method.

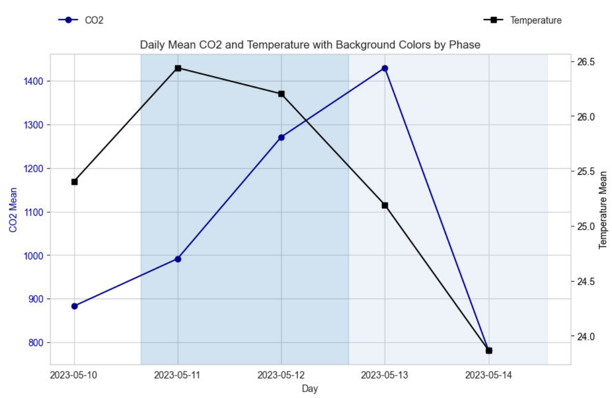


Fig. 11. Daily average temperature and CO_2 concentrations throughout the fermentation days.

It is important to emphasize that the eigenvalues for the principal components highlighted that Principal Component 1 (PC1) had an eigenvalue of 1.0294, while Principal Component 2 (PC2) had an eigenvalue of 0.9707. These eigenvalues reflect that PC1 slightly contributes more to the total variance compared with PC2, emphasizing the importance of PC1 in representing the data.

Digging deeper into the individual contributions, the variance explained by PC1 is notably high at 65.67%, underscoring its pivotal role in capturing and representing the dominant patterns within the data. Simultaneously, PC2 contributes to the variance by 34.33%, providing additional

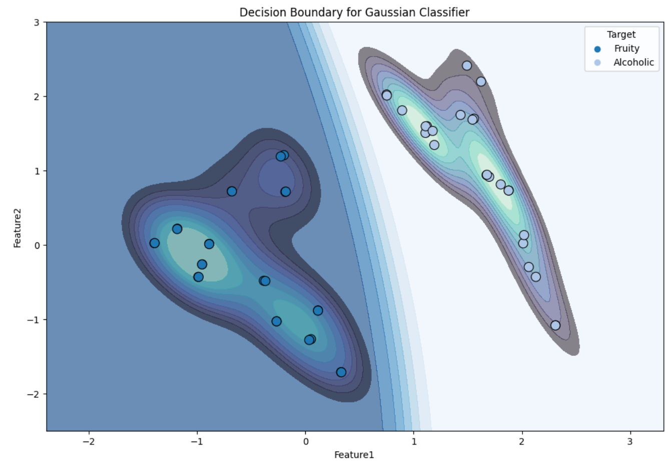


Fig. 12. Decision boundary obtained using a Gaussian Naive Bayes classifier, representing the odor descriptor.

insights into the underlying structures of the dataset. The distinct variances explained by PC1 and PC2 collectively highlight the comprehensive coverage and efficacy of the PCA approach in retaining essential information while simplifying the overall dataset.

In the k-means clustering analysis, several metrics were calculated to assess the quality of the obtained clusters. The Silhouette score, a measure of how well-defined the clusters are, was 0.7614, suggesting effective and well-defined clustering. The Davies–Bouldin index, which evaluates the distinction between clusters, had a value of 0.4106, indicating well-defined and distinct clusters. To evaluate cluster consistency, the average standard deviation of clusters in our dataset was calculated, resulting in a value of 0.0626. This low average standard deviation suggests that points within each cluster are similar and consistent.

Finally, a Gaussian Naive Bayes classifier was used on the obtained dataset. The model was built considering two features, PC1 and PC2, representing essential information for categorization. Based on this information, the classifier was trained to learn underlying patterns in the data and make accurate predictions about the target classes. The decision boundary obtained can be seen in Fig. 12, which presents the odor descriptor.

This integration of a Gaussian Naive Bayes classifier adds a layer of sophistication to our analysis, as it goes beyond clustering to predict specific class labels. The model's ability to learn and discern patterns from the data not only contributes to our understanding of the dataset but also provides a valuable tool for future predictions and categorizations based on odor descriptors.

While acknowledging the potential of commercial sensors in odor perception during coffee fermentation, it is essential to recognize that our analyses were conducted offline post-fermentation, and the current dataset's volume is not extensive. Despite these constraints, our results underscore the promise of commercial sensors as valuable tools in the context of coffee fermentation odor detection.

Moreover, it is important to highlight that despite the advantages presented by PCA, such as the lack of redundancy of data given the orthogonal components, smaller database representation, and reduction of noise [38], this method also comes with key disadvantages. The covariance matrix is challenging to be evaluated accurately [38], and even the simplest invariance may not be captured by PCA unless the training data explicitly provide this information [39].

Similarly, *k*-means clustering, despite being relatively scalable and simple, and suitable for datasets with compact spherical clusters that are well-separated [40], has its drawbacks. It exhibits severe effectiveness degradation in high-dimensional spaces, as almost all the pairs of points are about as far away as average, and the concept of distance between points in high-dimensional spaces is ill-defined [40]. As for DBSCAN, despite its resistance to noise and outliers, it is highly sensitive to the setting of input parameters.

These limitations are crucial considerations in the application of these techniques to real-time monitoring systems in our study. Furthermore, deploying techniques such as PCA and clustering for real-time monitoring systems poses notable challenges, as evidenced in the existing literature [41], [42], [43]. Low-end microcontrollers, due to their restrictions, may face difficulties in executing computationally intensive algorithms such as PCA efficiently, requiring higher processing capabilities. Memory limitations can impact the application of these techniques, especially with large datasets. The complexity of clustering algorithms and their resource requirements may also be restrictive [44], [45].

However, despite the mentioned limitations, it is crucial to underscore the significance of the method, particularly when compared with older techniques. This distinction becomes particularly noteworthy given the various constraints associated with human sensory analysis and the traditional chromatography methods.

So we believe that the use of electronic sensors and automated analytical techniques represents a significant advancement, overcoming the limitations of the traditional approaches. The implementation of techniques such as PCA and clustering, even with challenges on hardware, stands out as a valuable innovation in process monitoring, offering a more efficient and accurate alternative. These approaches not only provide gains in terms of efficiency and cost but also reduce reliance on methods that may be susceptible to human error and prohibitive costs.

This transition to automated and technologically advanced methods not only opens doors for more precise and efficient analyses but also lays the groundwork for future research and advancements in integrating innovative technologies into industrial process monitoring.

IV. Conclusion

In this work, we developed a probe using commercial CO_2 and temperature sensors, combined with a 3-D-printed encapsulation, with the purpose of monitoring and characterizing the coffee fermentation process. We demonstrated an effective data processing method that was successfully applied in aroma classification throughout the coffee fermentation process. Our results highlight that the use of commercial sensors is a promising tool for aroma characterization and coffee fermentation process monitoring.

Building upon the obtained results, we find it compelling to delve deeper into this approach in future studies. An envisioned extension of this work involves integrating supplementary sensors to measure additional pertinent volatile compounds in coffee. This augmentation would facilitate a more thorough analysis of aromas during fermentation, offering a comprehensive understanding of the process.

Moreover, we propose the exploration of more advanced machine learning algorithms to augment the aroma classification capability. This advancement holds the potential to refine and broaden the scope of our findings, providing a more nuanced and accurate interpretation of the complex aroma profiles in coffee fermentation. The integration of such enhancements in future research endeavors could significantly contribute to the advancement of our understanding in this domain.

Furthermore, we consider that this work opens up exciting prospects for the application of sensors and monitoring techniques in industrial processes, especially in the food production sector. The development of real-time monitoring systems and quality control systems based on sensory data can significantly contribute to improving the quality and efficiency of industrial processes. Therefore, we believe that this research can serve as a starting point for future investigations in this area.

In summary, this study successfully demonstrated the feasibility of using commercial sensors and data processing techniques to characterize aromas during the coffee fermentation process. This approach has the potential to positively impact the food industry and open new horizons in process monitoring research.

Acknowledgment

The authors would like to acknowledge the invaluable non-financial support from Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and Instituto Nacional de Energia Elétrica (Inerge-UFJF).

References

[1] Y. Liu, X. Ma, L. Shu, G. P. Hancke, and A. M. Abu-Mahfouz, “From Industry 4.0 to agriculture 4.0: Current status, enabling technologies, and research challenges,” *IEEE Trans. Ind. Informat.*, vol. 17, no. 6, pp. 4322–4334, Jun. 2021.

[2] (Sep. 2023). *Organization, I. Monthly Coffee Market Report (International Coffee Organization)*. [Online]. Available: https://www.icocoffee.org/documents/cy2022-23/cmr-0923-e.pdf

[3] L. W. Lee, M. W. Cheong, P. Curran, B. Yu, and S. Q. Liu, “Coffee fermentation and flavor—An intricate and delicate relationship,” *Food Chem.*, vol. 185, pp. 182–191, Oct. 2015.

[4] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, “SMOTE: Synthetic minority over-sampling technique,” *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, Jun. 2002.

[5] M. Haile and W. H. Kang, “The role of microbes in coffee fermentation and their impact on coffee quality,” *J. Food Qual.*, vol. 2019, pp. 1–6, Mar. 2019.

[6] N. Cordoba, M. Fernandez-Alduenda, F. L. Moreno, and Y. Ruiz, “Coffee extraction: A review of parameters and their influence on the physicochemical characteristics and flavour of coffee brews,” *Trends Food Sci. Technol.*, vol. 96, pp. 45–60, Feb. 2020.

[7] A. E. Peñuela-Martínez, S. Moreno-Riascos, and R. Medina-Rivera, “Influence of temperature-controlled fermentation on the quality of mild coffee (*Coffea arabica* L.) cultivated at different elevations,” *Agriculture*, vol. 13, no. 6, p. 1132, May 2023.

[8] A. F. S. Pino, Z. Y. D. Espinosa, and E. V. R. Cabrera, "Characterization of the rhizosphere bacterial microbiome and coffee bean fermentation in the Castillo-Tambo and bourbon varieties in the Popayán-Colombia Plateau," *BMC Plant Biol.*, vol. 23, no. 1, p. 217, Apr. 2023.

[9] I. Prakash et al., "Metabolomics and volatile fingerprint of yeast fermented robusta coffee: A value added coffee," *LWT*, vol. 154, Jan. 2022, Art. no. 112717.

[10] K. N. Aswathi, A. Shirke, A. Praveen, S. R. Chaudhari, and P. S. Murthy, "Pulped natural/honey robusta coffee fermentation metabolites, physico-chemical and sensory profiles," *Food Chem.*, vol. 429, Dec. 2023, Art. no. 136897.

[11] R. Mahingsapun et al., "Enhancement of arabica coffee quality with selected potential microbial starter culture under controlled fermentation in wet process," *Food Bioscience*, vol. 48, Aug. 2022, Art. no. 101819.

[12] L. L. Pereira et al., "Improvement of the quality of Brazilian Conilon through wet processing: A sensorial perspective," *Agricult. Sci.*, vol. 10, no. 3, pp. 395–411, 2019.

[13] T. R. Lingle and S. N. Menon, "Cupping and grading—Discovering character and quality," *Craft Sci. Coffee*, vol. 1, p. 181, Jan. 2017.

[14] D. M. D. J. Cassimiro et al., "Wet fermentation of Coffea canephora by lactic acid bacteria and yeasts using the self-induced anaerobic fermentation (SIAF) method enhances the coffee quality," *Food Microbiol.*, vol. 110, Apr. 2023, Art. no. 104161.

[15] O. Gonzalez-Rios et al., "Impact of 'ecological' post-harvest processing on the volatile fraction of coffee beans: I. Green coffee," *J. Food Composition Anal.*, vol. 20, nos. 3–4, pp. 289–296, May 2007.

[16] M. Ghasemi-Varnamkhasti et al., "Classification of non-alcoholic beer based on aftertaste sensory evaluation by chemometric tools," *Expert Syst. Appl.*, vol. 39, no. 4, pp. 4315–4327, Mar. 2012.

[17] M. Ghasemi-Varnamkhasti et al., "Monitoring the aging of beers using a bioelectronic tongue," *Food Control*, vol. 25, no. 1, pp. 216–224, May 2012.

[18] G. Iatropoulos, P. Herman, A. Lansner, J. Karlgren, M. Larsson, and J. K. Olofsson, "The language of smell: Connecting linguistic and psychophysical properties of odor descriptors," *Cognition*, vol. 178, pp. 37–49, Sep. 2018.

[19] M. Mohd Ali, N. Hashim, S. Abd Aziz, and O. Lasekan, "Principles and recent advances in electronic nose for quality inspection of agricultural and food products," *Trends Food Sci. Technol.*, vol. 99, pp. 1–10, May 2020.

[20] S. Boesveldt, M. J. Olsson, and J. N. Lundström, "Carbon chain length and the stimulus problem in olfaction," *Behavioural Brain Res.*, vol. 215, no. 1, pp. 110–113, Dec. 2010.

[21] A. Keller et al., "Predicting human olfactory perception from chemical features of odor molecules," *Science*, vol. 355, no. 6327, pp. 820–826, Feb. 2017.

[22] T. Debnath and T. Nakamoto, "Predicting human odor perception represented by continuous values from mass spectra of essential oils resembling chemical mixtures," *PLoS ONE*, vol. 15, no. 6, Jun. 2020, Art. no. e0234688.

[23] A. D. Wilson, "Review of electronic-nose technologies and algorithms to detect hazardous chemicals in the environment," *Proc. Technol.*, vol. 1, pp. 453–463, Dec. 2012.

[24] X. Yang et al., "Discriminant research for identifying aromas of non-fermented Pu-erh tea from different storage years using an electronic nose," *J. Food Process. Preservation*, vol. 42, no. 10, Oct. 2018, Art. no. e13721.

[25] J. Tan and W. L. Kerr, "Characterizing cocoa refining by electronic nose using a Kernel distribution model," *LWT*, vol. 104, pp. 1–7, May 2019.

[26] J. Chen, J. Gu, R. Zhang, Y. Mao, and S. Tian, "Freshness evaluation of three kinds of meats based on the electronic nose," *Sensors*, vol. 19, no. 3, p. 605, Jan. 2019.

[27] W. Dong et al., "Comparative evaluation of the volatile profiles and taste properties of roasted coffee beans as affected by drying method and detected by electronic nose, electronic tongue, and HS-SPME-GC-MS," *Food Chem.*, vol. 272, pp. 723–731, Jan. 2019.

[28] R. L. F. Filho, F. Oliveira Barino, J. Calderano, Ì. F. V. Alvarenga, D. Campos, and A. B. dos Santos, "In-fiber Mach–Zehnder interferometer as a promising tool for optical nose and odor prediction during the fermentation process," *Opt. Lett.*, vol. 48, no. 15, p. 3905, Aug. 2023.

[29] S. O. Araújo, R. S. Peres, J. Barata, F. Lidon, and J. C. Ramalho, "Characterising the agriculture 4.0 landscape—Emerging trends, challenges and opportunities," *Agronomy*, vol. 11, no. 4, p. 667, Apr. 2021.

[30] M. Raj et al., "A survey on the role of Internet of Things for adopting and promoting agriculture 4.0," *J. Netw. Comput. Appl.*, vol. 187, Aug. 2021, Art. no. 103107.

[31] K. R. Kini, F. Harrou, M. Madakyaru, F. Kadri, and Y. Sun, "Automatic detection of unbalanced sitting postures in wheelchairs using unlabeled sensor data," *IEEE Sensors Lett.*, vol. 6, no. 8, pp. 1–4, Aug. 2022.

[32] K. R. Kini, M. Madakyaru, F. Harrou, and Y. Sun, "Detecting pediatric foot deformities using plantar pressure measurements: A semisupervised approach," *IEEE Design Test*, vol. 40, no. 4, pp. 61–68, Aug. 2023.

[33] D. Galvan, A. A. C. Tanamati, F. Casanova, E. Danieli, E. Bona, and M. H. M. Killner, "Compact low-field NMR spectroscopy and chemometrics applied to the analysis of edible oils," *Food Chem.*, vol. 365, Dec. 2021, Art. no. 130476.

[34] L. Lu, Z. Hu, X. Hu, D. Li, and S. Tian, "Electronic tongue and electronic nose for food quality and safety," *Food Res. Int.*, vol. 162, Dec. 2022, Art. no. 112214.

[35] D. G. dos Santos, C. C. D. S. Coelho, A. B. R. Ferreira, and O. Freitas-Silva, "Brazilian coffee production and the future microbiome and mycotoxin profile considering the climate change scenario," *Microorganisms*, vol. 9, no. 4, p. 858, Apr. 2021.

[36] A. Kamilaris, A. Kartakoullis, and F. X. Prenafeta-Boldú, "A review on the practice of big data analysis in agriculture," *Comput. Electron. Agricult.*, vol. 143, pp. 23–37, Dec. 2017.

[37] A. Ram, S. Jalal, A. S. Jalal, and M. Kumar, "A density based algorithm for discovering density varied clusters in large spatial databases," *Int. J. Comput. Appl.*, vol. 3, no. 6, pp. 1–4, Jun. 2010.

[38] S. Karamizadeh, S. Abdullah, A. Manaf, M. Zamani, and A. Hooman, "An overview of principal component analysis," *J. Signal Inf. Process.*, vol. 4, pp. 173–175, Aug. 2013.

[39] C. Li, Y. Diao, H. Ma, and Y. Li, "A statistical PCA method for face recognition," *Intell. Inf. Technol. Appl.*, vol. 3, pp. 376–380, 2008.

[40] D. Sisodia, L. Singh, S. Sisodia, and K. Saxena, "Clustering techniques: A brief survey of different clustering algorithms," *Int. J. Latest Trends Eng. Technol.*, vol. 1, pp. 82–87, 2012.

[41] C. Ündey, S. Ertunç, T. Mistretta, and B. Looze, "Applied advanced process analytics in biopharmaceutical manufacturing: Challenges and prospects in real-time monitoring and control," *J. Process Control*, vol. 20, no. 9, pp. 1009–1018, Oct. 2010.

[42] S. Gajjar, M. Kulahci, and A. Palazoglu, "Real-time fault detection and diagnosis using sparse principal component analysis," *J. Process Control*, vol. 67, pp. 112–128, Jul. 2018.

[43] N. H. Abdulghafoor and H. N. Abdullah, "A novel real-time multiple objects detection and tracking framework for different challenges," *Alexandria Eng. J.*, vol. 61, no. 12, pp. 9637–9647, Dec. 2022.

[44] R. Xu and D. WunschII, "Survey of clustering algorithms," *IEEE Trans. Neural Netw.*, vol. 16, p. 645, Mar. 2005.

[45] D. Xu and Y. Tian, "A comprehensive survey of clustering algorithms," *Ann. Data Sci.*, vol. 2, no. 2, pp. 165–193, Jun. 2015.

Renato Faraco-Filho was born in Conselheiro Lafaiete, Brazil, in 1997. He received the B.Sc. degree in electronics from the Federal University of Juiz de Fora (UFJF), Juiz de Fora, Minas Gerais, Brazil, in 2022, where he is currently pursuing the M.Sc. degree in electrical engineering with the Programa de Pós Graduação em Engenharia Elétrica (PPEE).

His current research interests include instrumentation, optical fiber sensors, and machine learning.

João V. de Castro was born in Santos Dumont Lafaiete, Brazil, in 2003. He is currently pursuing the B.Sc. degree in electrical engineering and robotics and industrial automation with the Federal University of Juiz de Fora, Juiz de Fora, Brazil.

His current research interests include instrumentation, optical fiber sensors, and computational methods.

Felipe Oliveira Barino was born in Juiz de Fora, Brazil, in 1996. He received the B.S. and M.Sc. degrees in electronics from the Federal University of Juiz de Fora (UFJF), Juiz de Fora, in 2019 and 2021, respectively, where he is currently pursuing the Ph.D. degree in electrical engineering with the Programa de Pós-Graduação em Engenharia Elétrica (PPEE).

His current research interests are related to instrumentation, metrology, machine learning, transfer learning, optical fiber sensors, and sensor packaging.

Deivid E. Campos was born in São João Nepomuceno, Brazil, in 1991. He received the bachelor's degree in exact sciences and the bachelor's degree in mechanical engineering from the Federal University of Juiz de Fora (UFJF), Juiz de Fora, Minas Gerais, Brazil, in 2014 and 2022, respectively.

Currently, he is immersed in a Master's Program in computational modeling with UFJF. His research interests encompass areas of significant relevance in the current scientific landscape, involving advanced instrumentation, high-tech fiber optic sensors, the intricacies of 3-D modeling with 3-D printing, and the complex nuances of machine learning.

Alexandre Bessa dos Santos was born in Juiz de Fora, Brazil, in 1975. He received the M.Sc. and Ph.D. degrees in electrical engineering from Pontifical Catholic University (PUC), Rio de Janeiro, Brazil, in 2001 and 2005, respectively.

Since 2010, he has been an Associate Professor with the Department of Circuits, Federal University of Juiz de Fora (UFJF), Juiz de Fora. His research interests include applied electromagnetism, optical sensors, computational methods, instrumentation, and metrology.