Mehrab Mahdian

SIGNAL SEPARATION ALGORITHMS OF AN ELECTRONIC NOSE

SUPERVISOR

László Grad-Gyenge

BUDAPEST, 2023

Contents

Summary

Electronic Nose technology has been rapidly emerging and extended in recent years, thanks to the advances in materials, sensors, and machine learning technologies. The electronic nose (E-Nose) is a sensor fusion device consisting of an array of Metal-Oxide (MOx) sensors that detect Volatile Organic Compounds (VOCs) in its environment. The electric resistance of the sensors generates electric signals that change when they encounter and absorb different types of molecules, and afterward, the signals are processed by machine learning algorithms. The edge advantages of such devices are rapidity, portability, and compactness. As a result, these devices are becoming widely used in the food and beverage industry, agriculture and forestry, pharma, security, and environmental monitoring.

In a real environment, sensors detect a mixture of odors. Hence, it is beneficial to identify the individual components of such a mixture. This process is called Blind Source Separation (BSS), where there is no (or little) information about the sources or the mixing process. One approach to this problem is to have a good understanding of the possible individual sources. Then, by using this pre-obtained information, we can segregate the sources, hopefully separating at least one of the sources. In this context, machine learning models and more advanced ones, deep learning models, data analysis, and pattern recognition can be used.

This research aims to employ and evaluate different models regarding their effectiveness, suitability, and performance. Besides, several other samples relating to the food industry are to be measured using the Electronic Nose device.

Kivonat

Az elektronikus orr-technológia az utóbbi években gyorsan fejlődött és bővült az anyagok, érzékelők és gépi tanulási technológiák fejlődésének köszönhetően. Az elektronikus orr (Electronic Nose) egy érzékelőfúziós eszköz, amely fém-oxid (Metal-Oxide) érzékelők sorából áll, amelyek érzékelik az illékony szerves vegyületeket (Volatile Organic Compound) a környezetében. Az érzékelők elektromos ellenállása elektromos jeleket generál, amikor különböző típusú molekulákat abszorbeálnak. Ezt követően a jeleket gépi tanulási algoritmusok dolgozzák fel. Az ilyen eszközök előnye a gyorsaság, a hordozhatóság és a kompaktság. Ennek eredményeként ezek az eszközök széles körben használatosak az élelmiszer- és italiparban, a mezőgazdaságban és az erdészetben, a gyógyszeriparban, a biztonságban és a környezetfelügyeletben.

Valós környezetben az érzékelők szagok keverékét érzékelik. Ezért előnyös az ilyen keverék egyes összetevőinek azonosítása. Ezt a folyamatot BSS-nek (Blind Source Separation) nevezik, ahol nincs (vagy kevés) információ a forrásokról vagy a keverési folyamatról. Ennek a problémának az egyik megközelítése a lehetséges egyéni források megfelelő ismerete. Ezután az előre megszerzett információk felhasználásával a források, vagy legalábbis az egyik forrás, remélhetőleg elkülöníthetők lesznek. Ebben az összefüggésben gépi tanulási modelleket és fejlettebbeket, mély tanulási modelleket, valamint adatelemzési és mintafelismerő eszközöket fognak használni.

Ebben a kutatásban különböző modellek alkalmazását és értékelését célozzuk meg azok hatékonysága, alkalmassága és teljesítménye szempontjából. Emellett számos különböző élelmiszeripari mintát kell mérni az Electronic Nose készülékkel.

1. Introduction

Thanks to advances in sensor technologies and computer algorithms, Electronic noses (E-noses) are increasing the margin of attention. They have numerous applications in the food industry, pharma, forestry, agriculture, etc. E-noses are rapid and accurate tools to detect and identify odors and volatile organic compounds, and they are developed to replicate the human olfactory system.

Despite the extensive application of E-noses in the automatic detection of odors, one crucial underdeveloped area is separating individual signal components from a mixture signal. This process is called signal separation, source separation, or blind source separation. Nonetheless, in E-nose research, the term odor approximation is also used. Odor approximation is necessary in E-nose systems, enabling accurate identification of odor components and improving the system performance. However, the computer algorithms to process E-nose signal in that way is not trivial. An E-nose equipped with precise computer algorithms to decompose complex and overlapping odors and then accurately detect compositions may benefit food quality control, environmental monitoring, medical diagnosis, etc.

This research focuses on developing an odor approximation model, which can be utilized in the case of E-noses to extract the subcomponent of a mixture signal. In this study, we gather various samples and analyze our models on these samples. Our findings can contribute to broadening the E-nose application circle, improving its performance and reliability, and opening doors for odor recognition in diverse applications since we rarely have a single component to be detected in the real environment.

This study will first develop an AE model to capture the underlying pattern in the data by encoding it and then recreating it from the latent space. We will find the optimal structure of the AE through individual component analysis. Next, we will use the developed AE network for odor approximation. We will examine its performance by testing this network on combination samples.

To increase the accuracy and specificity of the model, we will introduce the domain adaptation concept. We will examine this new model by analyzing its performance on multiple combination samples taken over time.

The following chapter of this thesis is as follows: Chapter 2 comprehensively reviews E-nose technology, its application, and existing solutions in the literature. Chapter 3 outlines the methodologies in this study concerning the E-nose device sample collection. In Chapter 4, we present our findings by analyzing the model performance. Finally, in Chapter 5, we summarize our results and highlight ways for further research.

Keywords: electronic nose, signal separation, odor approximation.

2. Literature Review

2.1. Related Works

The first E-Nose for classification and identification purposes was created using three Figaro semiconducting gas sensors acting as transducers by (Persaud and Dodd, 1982) from the University of Warwick, England. This device demonstrated the capability of discriminating odors like the mammalian olfactory system. Continuous studies have been conducted in this field until now, and these devices are becoming more and more utilized in the food industry, pharma, security, and environmental monitoring.

The application of an electronic nose for industrial odors and gaseous emissions measurement and monitoring was reviewed by (Deshmukh *et al.*, 2015). They examined the electronic nose's applications in industrial emissions monitoring, emphasizing its advantages compared to traditional models. Despite its potential, its industrial use remains underexplored. The study discusses applications, calibration methods, and sampling and explores the electronic mucosa system as an improvement. The review concludes by discussing the pros and cons of artificial olfaction in industrial contexts.

In another study, (Tan and Xu, 2020) reviewed the use of electronic noses (E-noses) and electronic tongues (E-tongues) in assessing food quality. The study discusses their principles, applications, and limitations, comparing widely used pattern recognition algorithms such as artificial neural networks and support vector machines. The review concludes that e-nose and e-tongue, when combined with these algorithms, are powerful, cost-effective tools for rapid and accurate food quality analysis. They are applicable in both in-line and off-line measurements, aiding food processing monitoring and final product quality detection. They also carefully control sample preparation, sampling, and data processing.

The application of E-Noses in medical settings was reviewed in a study (Thaler and Hanson, 2005). Their reviews conclude promising results in medical applications such as detecting sinusitis, pneumonia, and cancer. Nonetheless, Challenges include improving data analysis and gaining medical community acceptance.

Regarding how E-Nose devices work, these devices work on the principle of mimicking a human nose's olfactory function. According to (Dunkel *et al.*, 2014), the

human olfactory system consists of three parts. First is the odor receiver, which comprises olfactory receptor glands and scent delivery systems. Second, the nervous systems are responsible for signal transmission to the brain. Third, a decision system to recognize and identify the smell. An E-Nose, likewise, has three main parts: a sample delivery system, a detection system, and a computing system, according to (Karami, Rasekh, and Mirzaee-Ghaleh, 2020). In this regard, the definition of an E-Nose must be restricted to intelligent chemical array sensor systems or devices designed to detect odorant molecules like the human nose.

According to (Gardner and Bartlett, 1994), "*An* electronic *nose is an instrument, which comprises an array of electronic chemical sensors with partial specificity and an appropriate pattern-recognition system, capable of recognizing simple or complex odors.*"
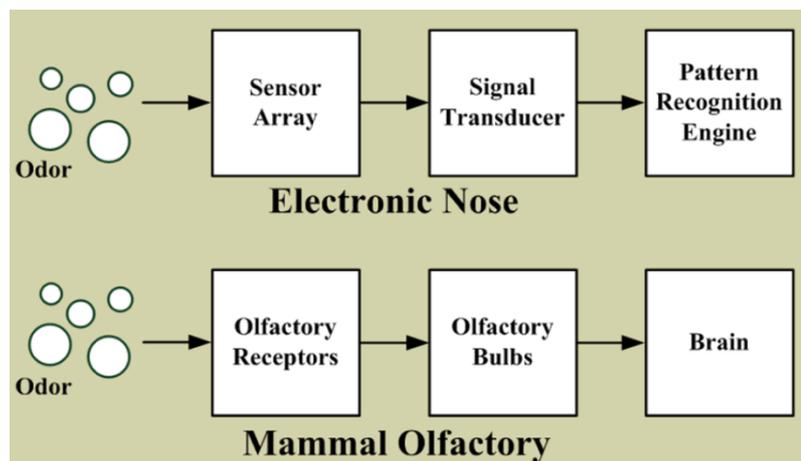
Figure 2.1 An analogy between the human olfactory system and E-Nose.
This image was taken from a study by Chiu and Tang published in 2013 (Chiu and Tang, 2013)

In the E-nose system, the olfactory receptor segment incorporates the odor delivery unit, which consists of pipes, pumps, and valves, facilitating aroma delivery into the sensor chamber. The sensor array, the heart of any E-nose, includes several different gas sensors acting as human olfactory receptors.

In one study to develop a wearable E-nose carried out by (Seesaard, Lorwongtragool, and Kerdcharoen, 2015), they proposed using conducting polymers/SWNT-COOH nanocomposites. They found that the sensors could detect numerous volatile compounds prevalent in human body wastes, and by incorporating simple pattern recognition, they could discriminate the human body odor from two

persons. In another study by (Seekaew *et al.*, 2014), they developed an E-nose to detect ammonia using the same conducting polymer nanocomposite sensors.

Another type of sensor made from carbon-based nanomaterial was proposed in studies conducted by (Kondee *et al.*, 2022),(Seekaew *et al.*, 2019), and (Kerdcharoen and Wongchoosuk, 2013). Carbon-based materials, including nitrogen-doped carbon oxide quantum dots (NCQDs), graphene, and carbon nanotubes (CNTs), were utilized in research due to their exceptional properties. NCQDs demonstrated high humidity sensitivity, stability, and linear response. Graphene and its derivatives offered versatility in applications like sensors, energy storage, and solar cells, benefiting from their unique electronic properties. CNTs, especially with metal oxides, enhanced gas sensor sensitivity and versatility through various preparation and characterization techniques.

Metal-oxide (MOx) sensors are another type of sensor, which is proposed by (Traiwatcharanon, Timsorn, and Wongchoosuk, 2017), (Arayawut, Kerdcharoen and Wongchoosuk, 2022), (Traiwatcharanon *et al.*, 2022), and (Chaloeipote *et al.*, 2021). Researchers utilized metal oxide sensors, including silver nanoparticles for humidity sensing, lead oxide nanoparticles for paraquat detection, and copper oxide via 3D printing for ammonia gas sensing. These sensors demonstrated high sensitivity, stability, and selectivity in their respective applications. The studies showcased innovative fabrication methods and optimized sensor performance, emphasizing the practical potential of metal oxide sensors in scientific and environmental monitoring.

In 2005, (James et al. 2005) reviewed chemical sensors for E-Nose. This study investigated conducting polymers, metal oxide semiconductors, piezoelectric, optical fluorescence, and amperometric gas sensors.

Despite the wide usage of these sensors, a few studies have focused on the specific type of sensor for odor separation tasks, also called odor approximation. The term odor approximation refers to identifying or reproducing a particular scent from a mixture of scents. In a study by (Muñoz-Aguirre *et al.*, 2007), they utilized quartz crystal microbalance (QCM) sensors to approximate complex fruit flavors. QCM sensors were coated with specific materials to enhance sensitivity and selectivity. The sensors were employed to measure responses to individual odor components and mixtures in orange and melon flavors. This approach allowed quantitative analysis, aiding in identifying and evaluating key odor components in a mixture.

These sensors absorb odorous molecules based on physisorption and chemisorption principles. Upon the presence of molecules on the material's surface, expediting expansions of volume, transfer of charges, ion exchange, and interaction with ion species results in a change in electrical conductivity or resistivity. Afterward, the generated signal by the sensor arrays is undergone through data analysis performed via supervised or unsupervised machine learning with statistical methods.

In one study on the discrimination of pathogenic bacterial volatile compounds, (Seesaard *et al.*, 2020) used Principal Component Analysis (PCA). This analysis showed that the first principal component (PC1) explained 92.4% of the variance, while the second principal component (PC2) explained 7.2%. This finding highlights the critical role of PC1 in effectively distinguishing between various bacterial species. In another study on discriminating red wine conducted by (García et al., 2006), they incorporated PCA. Researchers successfully differentiated red wines from the same cellar, grape variety, and geographic origin. Nonetheless, the PCA method had an 80% success rate due to interference from water and ethanol vapors.

Other machine learning models applicable in research on E-nose delivering favorable result is as follows: Hierarchical Cluster Analysis (CA) (da Silva Torres, Garbelotti and Moita Neto, 2006), (Le Thanh-Blicharz and Lewandowicz, 2020), Linear Discriminant Analysis (LDA) (Kiselev *et al.*, 2018), partial least squares discriminant analysis (PLS-DA) (Huang *et al.*, 2017) multivariate data analysis (Buvé et al., 2022), and Artificial Neural Network (ANN) (Thazin, Pobkrut and Kerdcharoen, 2018), (Zhang and Tian, 2014) etc.

In 2020, a survey on "Electronic Nose and Its Application" (Karakaya, Ulucan, and Turkan, 2020) reviewed the recent technologies applicable to E-Nose devices. At first, they opened an introduction to E-nose technology, how it works, and its applications. Then, they precisely elaborated the E-Nose components and the types of sensors, detailing eight different kinds of sensors. In the following, they featured the most used machine learning (ML) and Deep learning (DL) algorithms, elaborated eight algorithms, and concluded the frequently used data analysis methods based on each sensor type. They compared five algorithms, followed by a system performance evaluation. Next, they listed 12 commercial E-Nose- devices and their usage area. In the rest of the paper, they enumerated the application of E-Nose. Based on their gatherings, E-Noses are widely

utilized in the food and beverage industry (including meat quality assessment, chocolate and cocoa beans, alcoholic beverages, tea quality, coffee quality, oil and vinegar, dairy products, and water quality and monitoring), forestry and agriculture, medicine and healthcare (including the detection of asthma, cancer, tuberculosis, sinusitis, cyst fibrosis, etc.), indoor and outdoor monitoring, system security, and packaging. Ultimately, they pointed out the main challenges in the E-Nose system, an outlook, and concluded. According to them, the main challenges are sensor sensitivity, sensor selectivity, humidity, sensor shift, sensor stability, reproducibility, sensor fault tolerance, cross-sensitivity, sensor array size, material selection, algorithms, parameter selection, and lack of data.

Despite extensive research on E-Nose devices and their commonly used applications, a few studies focused on the signal separation of the given odor mixture by an E-Nose device. Signal separation, also called Blind Source Separation (BSS), is a technique in signal processing aimed at separating signals and decomposing them into their subcomponents given an unknown mixed signal.

In one study (Phaisangittisagul and Nagle, 2011), they strived to approximate the components in a mixture. Their data comprised regular coffee, Sumatra coffee, and green tea. They proposed wavelet analysis and support vector machine (SVM) to predict the waveform of the odor mixtures. Then, they evaluate the performance by how close the prediction result to the measured individual waveform is. Their findings were significant because they presented algorithm solutions instead of special hardware.

Another study on odor approximation (Prasetywan and Takamichi, 2020) used the essential oil mass spectrum database. They used non-negative matrix factorization (NMF) to extract the basis vectors, and then they used the non-negative least-squares technique to determine the recipe. Their studies resulted in approximating essential oils using 30 odor components.

2.2. Auto-encoder

An Auto-encoder (AE) is a kind of unsupervised learning artificial neural network. The AE consists of two main parts: an encoder and a decoder. The encoder transforms the input data and represents it in a lower dimension, and the decoder recreates the input data from the reduced dimensions. The output of the encoder is called latent space. (Kramer, 1991).

AE has numerous variations and applications. For instance, we have regularized AE for classification tasks, variational AEs as generative models, feature detection, anomaly detection, and facial recognition. Figure 2.2 shows the general structure of the AE.
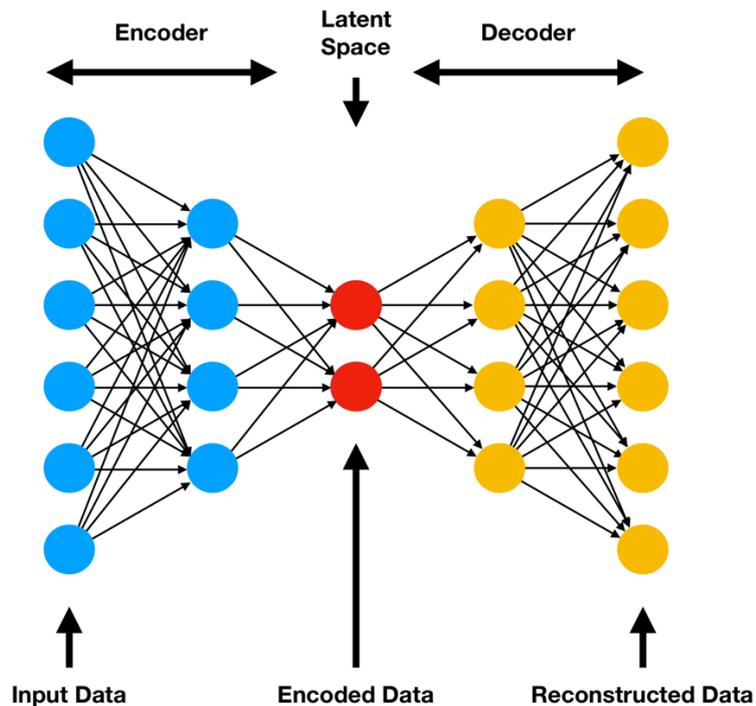
Figure 2.2 The General Structure of Auto-encoder[1]

---

[1] Author: Steven Flores, Year: 2019, " Variational Autoencoders are Beautiful", accessed Oct 2023, https://www.compthree.com/blog/autoencoder/.

2.3. Domain Adaptation

Domain adaptation is a field between machine learning and transfer learning. This method is beneficial when we want to model and apply source data distribution to different but related data.

The domain adaptation models are trained on one or more source domains to different target domains. This transformation is called domain shift, also known as distributional shift. Figure 2.3 shows the domain adaptation.
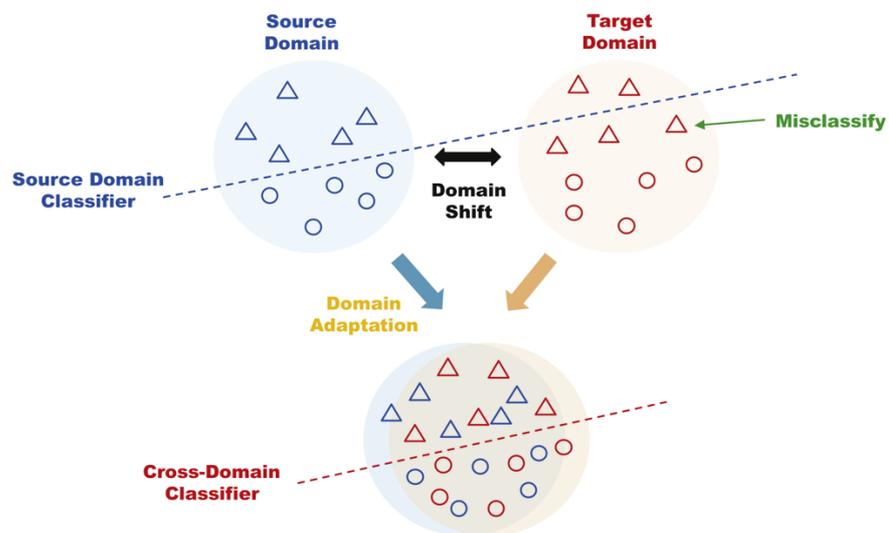
Figure 2.3 Illustration for domain adaptation.

This picture was taken from a study by (Li *et al.*, 2019)

3. Methodology

This chapter will introduce the E-nose device, its components, sample chamber, sample collections, and data processing.

3.1. The Electronic Nose Device Design

The body of the electronic nose device is constructed using a 3D printer with a pipe diameter of 2 cols. There is a fan to inhale the o placed on the top of the body. The device consists of sensor arrays to detect volatile organic compounds (VOC), an MCU board, a microcontroller to read sensor values and transmit the data to the backend, and an analog multiplexer. An analog multiplexer instead of an adorning analog-to-digital converter (ADC) is necessary due to the surplus of sensors compared to available ADC ports. These sensors are linked to the MCU board through a multiplexer. Figure 3.1 shows the electronic nose device.
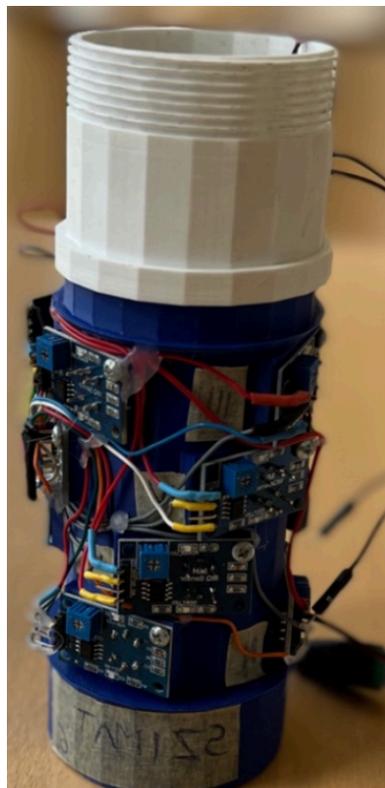
Figure 3.1 The Electronic Nose Device

3.2. Sensor Array

The sensors embedded in the E-Nose device are MOx sensors; see Figure 3.2. MOx sensors consist of a metal oxide film, typically made of materials like tin dioxide ($SnO2$), tungsten oxide ($WO3$), or zinc oxide (ZnO), which acts as the sensing element. The metal oxide film is heated to a high temperature, usually around 200-500 degrees Celsius, increasing its gas sensitivity.

When the metal oxide film is exposed to specific volatile organic compounds (VOCs), the VOC molecules interact with the film's surface, leading to changes in the electrical conductivity of the sensor. These changes are then measured and converted into a gas concentration reading.

MOx sensors are known for their low cost, compact size, and fast response time. They can detect various gases and are particularly sensitive to certain pollutants and toxic gases. However, they may also exhibit cross-sensitivity to other gases, meaning they can respond to multiple gases simultaneously, complicating their readings in certain situations. Figure 3.1 illustrates the sensors used in the E-nose device, and Table 3.1 provides a comprehensive overview and applications for each sensor type.

Figure 3.2 MOx Sensors Embedded in E-Nose Device[2]

---

[2] Unknown Author, Unspecified Year, "Gas Detection Module," Aliexpress, accessed Oct 2023, https://www.aliexpress.com/item/1005004315461088.html.

Model	Detection Type	Target Gases	Applications
MQ-2	Combustible gases	Methane (CH4), Propane (C3H8), Butane (C4H10), and other flammable gases	Gas leak detection, safety systems
MQ-3	Alcohol vapors	Alcohol vapors in the air	Breathalyzers, alcohol detection systems, safety devices
MQ-4	Alcohol vapors	Alcohol vapors in the air	Breathalyzers, alcohol detection systems, safety devices
MQ-5	Flammable gases	Liquefied Petroleum Gas (LPG), Methane (CH4)	Gas leak detection, safety systems
MQ-6	Liquefied Petroleum Gas (LPG)	Liquefied Petroleum Gas (LPG)	Gas detection, safety equipment, industrial systems
MQ-7	Carbon Monoxide (CO)	Carbon Monoxide (CO) gas	Gas detection, safety equipment, environmental monitoring
MQ-8	Coal, Gas, Hydrogen	Coal, Gas, Hydrogen	Gas detection, industrial applications
MQ-9	Carbon Monoxide (CO), Flammable Gases	Carbon Monoxide (CO), Methane (CH4), Propane (C3H8)	Gas detection, safety equipment, industrial monitoring
MQ-135	Various air pollutants and gases	Ammonia (NH3), Nitrogen oxides (NOx), Alcohol vapors, Benzene (C6H6), Carbon monoxide (CO), Volatile organic compounds (VOCs)	Air quality monitoring, environmental assessment, safety applications

Table 3.1 Gas Detection Sensors: A Comprehensive Overview for Diverse Applications

3.3. MCU

The MCU is utilized to gather sensor signals and transmit the data to the backend through the Internet. The project utilizes an ESP32-WROOM-32S board. It is used to digitize sensor data, collect sensor data, connect to Wi-Fi, connect to the backend, and send data to the backend.

3.4. Sample Collection

The sample collection for this study comprised five everyday food items: Vinegar, Bread, Beef, and Chicken. All samples were procured from a local market and stored at room temperature throughout the measurement process to simulate real-world conditions.

Measurements were taken at regular intervals of 12 hours and conducted four times to ensure accuracy and reliability. During each measurement session, individual items, along with the combinations of two items, were taken. Each measurement session lasted for 3 minutes. Table 3.2 presents an overview of the sample collection, indicating the items studied and their combinations during the measurement process.

In subsequent sections of this paper, each measurement will be denoted as 'Mx.' For instance, the combination of Vinegar and Beef for the first measurement will be represented as 'Vinegar+Beef M1', and for the second measurement, it will be denoted as 'Vinegar+Beef M2', where 'M2' indicates the second measurement session.

Samples	replicates
Bread	4
Vinegar	4
Beef	4
Chicken	4
Bread+Beef	4
Bread+Chicken	4
Vinegar+Beef	4
Vinegar+Chicken	4

Table 3.2 Sample Collection

3.5. Experiment Procedure

The E-Nose device was positioned adjacent to a chamber containing the items to prepare the device for measurements. A fan was placed in the center of the chamber, and there were holes on all four sides to maintain the airflow to the E-Nose device.

The E-Nose device is connected to Wi-Fi, enabling access to the backend website for recording measurements. Finally, the recorded data is stored in a CSV file, as

illustrated in Table 3.3, showcasing the initial 12 measurements from 9 MOx sensors. The signal generated by the E-nose for Vinegar M1 is shown in Figure 3.3.

Ticks	mq2	mq3	mq4	mq5	mq6	mq7	mq8	mq9
5525514807	142	1130	990	1773	142	558	1021	1248
5525519684	142	1130	990	1773	142	558	1021	1248
5525529684	142	1130	990	1773	142	558	1021	1248
5525539688	142	1130	990	1773	142	558	1021	1248
5525549684	142	1130	990	1773	142	558	1021	1248
5525559684	142	1130	990	1773	142	558	1021	1248
5525569684	142	1130	990	1773	142	558	1021	1248
5525579684	142	1130	990	1773	142	558	1021	1248
5525589715	142	1130	990	1773	142	558	1021	1248
5525599684	142	1130	990	1773	142	558	1021	1248

Table 3.3 A Snippet of CSV file for Vinegar
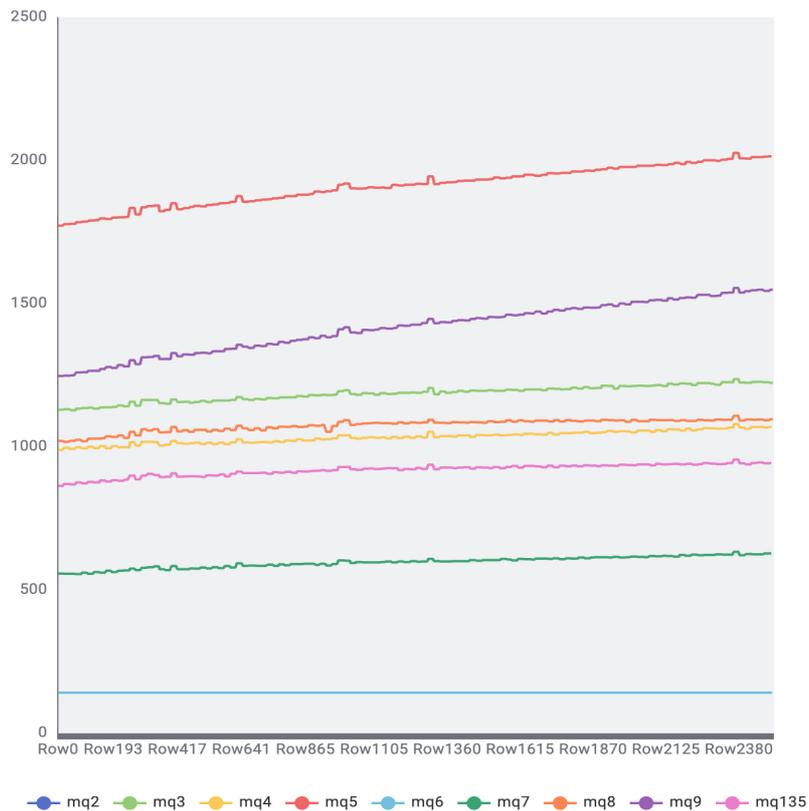
Figure 3.3 Signal Generated by E-nose for Vinegar M1

Before each measurement, the E-nose was warmed up for 30 minutes. The measurements were conducted at room temperature, with special care to minimize other

odors. Samples were also separated from each other to prevent interference. Table 3.4 provides information regarding the environmental condition during the measurements.

Measurement	Mean Temperature	Mean Humidity
1st	25° C	56%
2nd	24° C	59%
3rd	24° C	58%
4th	25° C	64%

Table 3.4 Environment Condition during the Measurement

4. Results

4.1. Auto-Encoder

Autoencoders (AE) are unsupervised machine learning algorithms designed to reconstruct input data. This model is constructed with an encoder network, which reduces the input data's dimensionality, followed by a decoder network that reconstructs the original data from the reduced representation. Essentially, an AE comprises an encoder and a decoder interconnected. The primary objective in designing an AE is to identify a bottleneck in the architecture where dimensions can be reduced without substantial information loss. The output of this reduced dimensionality, produced by the encoder, is referred to as the latent space.

In this study, we utilized an AE encoder trained on a single sample, which was subsequently validated with other samples. The validation sample demonstrating the minimum reconstruction loss is considered the closest approximation to our ground truth. This method aligns with the application of AE in anomaly detection. In the subsequent analysis, we aim to identify the ideal latent space dimension, optimizing it to minimize loss without compromising data integrity.

In the AE's structure, ReLU activation functions are employed in encoder layers to introduce non-linearity and counter the vanishing gradient problem. Leaky Rectified Linear Unit (LeakyReLU) activation functions are used in decoder layers (except the last layer) to prevent neuron inactivity. The final layer of the AE utilizes the sigmoid activation function for reconstructing the original data. The model is optimized using the mean square error (MSE) loss function and the Adaptive Moment Estimation (Adam) optimizer.

We have developed a deep autoencoder (AE) architecture, achieving optimal results by reducing the dimensionality to 4. The architecture comprises layers arranged as follows: (9-8-7-6-5-4-5-6-7-8-9), which satisfies the step-by-step reduction and subsequent expansion of dimensions within the AE network.

Table 4.1, Table 4.2, Table 4.3 and Table 4.4 display the reconstruction losses for the tested data. The first column corresponds to the various training sets in each table, while the first row represents the respective test sets. Notably, the minimum loss occurs

when the autoencoder (AE) is tested with its corresponding true data, indicating an impressive accuracy in reconstructing the original samples.

This analysis highlights the AE's ability to discern between different samples and individual measurements within the samples. For instance, in Table 4.1, the minimum loss for Bread M1 is 0.000079, while the loss increases incrementally from Bread M1 to Bread M4, reaching 0.121675. This observation underscores the AE's discriminative power in capturing subtle variations within the data.

Test Data / Train Data	Bread M1	Bread M2	Bread M3	Bread M4
Bread M1	**0,00008**	0,01183	0,01771	0,12167
Bread M2	0,00756	**0,00011**	0,01436	0,05782
Bread M3	0,00561	0,00832	**0,00005**	0,02699
Bread M4	0,04692	0,04452	0,03020	**0,00028**
Vinegar M1	0,04408	0,01400	0,02305	0,01083
Vinegar M2	0,00517	0,08635	0,00729	0,03400
Vinegar M3	0,00442	0,00184	0,00677	0,01209
Vinegar M4	0,01132	0,01441	0,02244	0,03811
Beef M1	0,00586	0,02700	0,02030	0,15280
Beef M2	0,00374	0,01298	0,03448	0,30164
Beef M3	0,01133	0,01066	0,04598	0,22751
Beef M4	0,01653	0,03406	0,06296	0,27046
Chicken M1	0,00246	0,02980	0,03440	0,15738
Chicken M2	0,00923	0,01973	0,10838	0,33760
Chicken M3	0,00925	0,02520	0,02138	0,11644
Chicken M4	0,02442	0,03756	0,06231	0,19387

Minimum Loss
3 Decimal Loss
2 Decimal Loss
1 Decimal Loss

Table 4.1 Reconstruction Losses of Autoencoder for Bread Samples

Train Data \ Test Data	Vinegar M1	Vinegar M2	Vinegar M3	Vinegar M4
Bread M1	0,06458	0,03491	0,03024	0,02876
Bread M2	0,02438	0,00608	0,01052	0,01350
Bread M3	0,03804	0,01277	0,01276	0,02367
Bread M4	0,00924	0,03159	0,03831	0,03385
Vinegar M1	**0,00002**	0,01042	0,01520	0,01693
Vinegar M2	0,00476	**0,00300**	0,00928	0,00508
Vinegar M3	0,00846	0,01142	**0,00031**	0,00779
Vinegar M4	0,01037	0,01872	0,00729	**0,000031**
Beef M1	0,09499	0,04252	0,04574	0,03109
Beef M2	0,07727	0,07408	0,04849	0,04883
Beef M3	0,07411	0,02285	0,02061	0,05238
Beef M4	0,15807	0,05083	0,06954	0,03459
Chicken M1	0,12856	0,03275	0,05873	0,03272
Chicken M2	0,13473	0,11366	0,06384	0,09866
Chicken M3	0,12373	0,04316	0,04243	0,04780
Chicken M4	0,11279	0,11639	0,05628	0,03222

Minimum Loss
3 Decimal Loss
2 Decimal Loss
1 Decimal Loss

Table 4.2 Reconstruction Losses of Autoencoder for Vinegar Samples

Train Data \ Test Data	Beef M1	Beef M2	Beef M3	Beef M4
Bread M1	0,00504	0,00308	0,00552	0,01024
Bread M2	0,01489	0,00438	0,00903	0,00737
Bread M3	0,00656	0,00525	0,00323	0,03342
Bread M4	0,04965	0,06754	0,08542	0,07457
Vinegar M1	0,03865	0,03315	0,03566	0,02883
Vinegar M2	0,01193	0,00436	0,02754	0,00397
Vinegar M3	0,00922	0,00616	0,00444	0,00240
Vinegar M4	0,01397	0,00895	0,02910	0,00293
Beef M1	**0,00025**	0,02884	0,01751	0,01606
Beef M2	0,01310	**0,00010**	0,00553	0,01431
Beef M3	0,00782	0,00358	**0,000048**	0,00659
Beef M4	0,02489	0,01969	0,01264	**0,00059**
Chicken M1	0,00476	0,00389	0,00348	0,01257
Chicken M2	0,02047	0,00108	0,00570	0,01285
Chicken M3	0,00793	0,01302	0,00650	0,02369
Chicken M4	0,03916	0,03223	0,02879	0,00612

Minimum Loss
3 Decimal Loss
2 Decimal Loss
1 Decimal Loss

Table 4.3 Reconstruction Losses of Autoencoder for Beef Samples

Test Data / Train Data	Chicken M1	Chicken M2	Chicken M3	Chicken M4
Bread M1	0,00856	0,00371	0,01784	0,01706
Bread M2	0,01393	0,00485	0,02448	0,01066
Bread M3	0,00556	0,05451	0,00522	0,02247
Bread M4	0,06448	0,09345	0,07425	0,07806
Vinegar M1	0,04106	0,08587	0,07964	0,03554
Vinegar M2	0,00796	0,00443	0,01661	0,02995
Vinegar M3	0,01714	0,03549	0,01445	0,00571
Vinegar M4	0,01264	0,00605	0,02930	0,02065
Beef M1	0,00394	0,03005	0,01961	0,03253
Beef M2	0,00751	0,00446	0,01574	0,02209
Beef M3	0,00791	0,00798	0,01033	0,01636
Beef M4	0,01661	0,01315	0,04170	0,01184
Chicken M1	**0,00022**	0,00354	0,01009	0,02981
Chicken M2	0,00468	**0,00257**	0,02140	0,02184
Chicken M3	0,00653	0,01489	**0,00015**	0,03380
Chicken M4	0,03382	0,02517	0,06969	**0,00318**

Minimum Loss
3 Decimal Loss
2 Decimal Loss
1 Decimal Loss

Table 4.4 Reconstruction Losses of Autoencoder for Chicken Samples

Now, we extend our studies to combinations after finding the optimal structure and hyperparameter for the AE. Here, the AE is first trained on individual components and then tested on mixtures. We aim to assess the AE's ability to capture the underlying patterns observed in single components when given a combination, especially when one or both components of the mixture are present.

Table 4.5, Table 4.6, Table 4.7, and Table 4.8 show the reconstruction losses of the autoencoder (AE) trained on individual samples and tested on combinations. Overall, to some extent, the model can reconstruct samples within their relevant classes, but not precisely to the exact measurements.

In Table 4.5, in the case that the AE is trained on Bread M1, it shows the minimum loss when tested on "Bread+Beef M1". However, when tested on "Bread+Beef M2", the minimum reconstruction loss still corresponds to the model trained on "Bread M1". Notably, an incorrect minimum reconstruction loss occurs outside its relevant class for "Bread+Beef M3" when the AE is trained on "Chicken M3".

Noticeably, the Vinegar samples show substantially better results, as shown in Tables 4.7 and 4.8. The AE effectively identifies the underlying patterns of Vinegar samples, which enables the decomposition and discrimination of the signals in the mixture and even for each measurement. In this case, the vinegar showed to be the predominant odor. This finding suggests the significance of considering odor intensity and its masking effect in future studies concerning signal decomposition or odor approximation.

Test Data / Train Data	Bread+Beef M1	Bread+Beef M2	Bread+Beef M3	Bread+Beef M4
Bread M1	0,00065	0,00367	0,03579	0,03164
Bread M2	0,02061	0,01366	0,03978	0,03117
Bread M3	0,02169	0,02035	0,01116	0,01480
Bread M4	0,05299	0,04848	0,05519	0,03901
Vinegar M1	0,04793	0,04169	0,06154	0,03437
Vinegar M2	0,00971	0,00929	0,03204	0,01871
Vinegar M3	0,00962	0,00814	0,02365	0,01729
Vinegar M4	0,01796	0,01879	0,04084	0,02654
Beef M1	0,01136	0,00873	0,01546	0,01672
Beef M2	0,00580	0,00851	0,03806	0,03803
Beef M3	0,00507	0,00359	0,03229	0,02986
Beef M4	0,02396	0,01962	0,07091	0,00708
Chicken M1	0,00643	0,00513	0,01616	0,07460
Chicken M2	0,01530	0,00971	0,04625	0,04485
Chicken M3	0,00699	0,00378	0,00456	0,03981
Chicken M4	0,05710	0,04713	0,09150	0,11572

Minimum Loss
3 Decimal Loss
2 Decimal Loss
1 Decimal Loss

Table 4.5 Reconstruction Losses of Autoencoder Trained on Individual Samples and Tested on Bread+Beef Combinations

Test Data / Train Data	Bread+Chicken M1	Bread+Chicken M2	Bread+Chicken M3	Bread+Chicken M4
Bread M1	0,04684	0,05437	0,02983	0,03094
Bread M2	0,00526	0,00936	0,07429	0,10931
Bread M3	0,00717	0,00636	0,02086	0,01294
Bread M4	0,01516	0,04008	0,60130	0,25693
Vinegar M1	0,03997	0,04147	0,05374	0,04883
Vinegar M2	0,00707	0,01212	0,02726	0,02940
Vinegar M3	0,00728	0,02603	0,02328	0,02584
Vinegar M4	0,01393	0,02354	0,05865	0,04149
Beef M1	0,01776	0,01409	0,05421	0,06736
Beef M2	0,00881	0,01199	0,18893	0,09931
Beef M3	0,00544	0,00561	0,08905	0,09679
Beef M4	0,01324	0,03066	0,14600	0,38314
Chicken M1	0,00335	0,00663	0,02047	0,04939
Chicken M2	0,00537	0,02820	0,09847	0,27093
Chicken M3	0,00771	0,00220	0,07147	0,07306

Minimum Loss
3 Decimal Loss
2 Decimal Loss
1 Decimal Loss

Table 4.6 Reconstruction Losses of Autoencoder Trained on Individual Samples and Tested on Bread+Chicken Combination

Test Data / Train Data	Vinegar+Beef M1	Vinegar+Beef M2	Vinegar+Beef M3	Vinegar+Beef M4
Bread M1	0,01165	0,01367	0,03579	0,03164
Bread M2	0,02061	0,01366	0,03978	0,03117
Bread M3	0,02169	0,02035	0,01116	0,01480
Bread M4	0,05299	0,04848	0,05519	0,03901
Vinegar M1	0,00079	0,04169	0,06154	0,03437
Vinegar M2	0,00971	0,00129	0,03204	0,01871
Vinegar M3	0,00962	0,00814	0,00237	0,01729
Vinegar M4	0,01796	0,01879	0,04084	0,00265
Beef M1	0,01136	0,00873	0,01546	0,01672
Beef M2	0,00580	0,00851	0,03806	0,03803
Beef M3	0,00507	0,00359	0,03229	0,02986
Beef M4	0,02396	0,01962	0,07091	0,00708
Chicken M1	0,00643	0,00513	0,01616	0,07460
Chicken M2	0,01530	0,00971	0,04625	0,04485
Chicken M3	0,00699	0,00378	0,00456	0,03981
Chicken M4	0,05710	0,04713	0,09150	0,11572

Minimum Loss
3 Decimal Loss
2 Decimal Loss
1 Decimal Loss

Table 4.7 Reconstruction Losses of Autoencoder Trained on Individual Samples and Tested on Vinegar+Beef Combinations

Test Data / Train Data	Vinegar+Chicken M1	Vinegar+Chicken M2	Vinegar+Chicken M3	Vinegar+Chicken M4
Bread M1	0,08636	0,15700	0,10232	0,11659
Bread M2	0,04186	0,06882	0,05123	0,06690
Bread M3	0,05017	0,02244	0,03878	0,04091
Bread M4	0,01188	0,09200	0,01352	0,02664
Vinegar M1	0,01130	0,01079	0,02174	0,03971
Vinegar M2	0,00813	0,00937	0,05258	0,05682
Vinegar M3	0,00347	0,00569	0,06828	0,03471
Vinegar M4	0,05621	0,02122	0,00194	0,02746
Beef M1	0,08950	0,10170	0,10999	0,09923
Beef M2	0,16221	0,11001	0,16127	0,19360
Beef M3	0,18484	0,19921	0,08079	0,21422
Beef M4	0,32128	0,22959	0,24024	0,38386
Chicken M1	0,15171	0,00733	0,01278	0,18160
Chicken M2	0,17484	0,19092	0,17984	0,20755
Chicken M3	0,11891	0,14178	0,15194	0,02278
Chicken M4	0,22692	0,19344	0,19888	0,23648

Minimum Loss
3 Decimal Loss
2 Decimal Loss
1 Decimal Loss

Table 4.8 Reconstruction Losses of Autoencoder Trained on Individual Samples and Tested on Vinegar+Chicken Combinations

4.2. Domain Adaptation Auto-encoder

As our next step, to bring the capability to discriminate between each measurement, even in the case of mixtures, we introduced domain adaptation to the AE. Domain adaptation is a method to improve the model performance on a target domain when there is a lack of sufficient annotated data.

AE network is primarily intended to reconstruct its input data. Nonetheless, the AE is trained to construct a different output when incorporating domain adaptation. In other words, it transforms the input and shifts it to another domain. We trained the AE on combination samples to construct the individual components in our experiments. That is the minimum prior knowledge about the source. In this case, we have a dedicated AE network for each data. For instance, we train AE on "Bread+Beef M1" to construct the "Bread M1" and another AE to construct 'Beef M1'. Afterwards, the Euclidean distance between the transformed "Bread M1" and the ground truth "Bread M1" is calculated. The minimum Euclidean distance shows the closest transformation to the true values in a hyperplane.

The structure and hyperparameters of the AE network are similar to that of in section 4.1. Nonetheless, this network requires extensive training iterations, and consequently, it increases the amount of time for training. In this experiment, we set the epochs equal to 10000. However, even more iteration delivers better results.

The Euclidean distance between the transformed combinations and the ground truth for each data is illustrated in Table 4.9, Table 4.10, Table 4.11, and Table 4.12. The first rows show the transformations, that is, the data that the AE model was trained on, in which each one is a dedicated model. The first column represents the test data, the combination, which was tested to assess the model performance for transformation in each case. Lastly, the last rows represent the ground truth, the individual components. The objective is to test each combination on the dedicated models, finding the transformation and then calculating the Euclidean distance between the transformations and the ground truth.

Table 4.9 displays the results for the "Bread+Beef" samples. We can observe that the minimum Euclidean distances correspond to each class and each measurement. There are a few erroneous results in Table 4.11 and Table 4.12. The red cell represents the erroneous result, and the yellow cells represent the true negatives. This error can be due to insufficient training iterations.

Overall, the dedicated domain adaptation AE network can be a powerful method in decomposing one signal component of the mixture when minimum prior information regarding the source is available. Nonetheless, the model requires extensive training data and numerous iterations, which can be computationally expensive. Hence, we propose this method when the working data in actual use cases is limited, and there can be minimum prior information about the data.

Table 4.9 header abbreviations: columns are Bread+Beef M1 -> Bread M1, BreadBeef M2 -> Bread M2, Bread+Beef M3 -> Bread M3, Bread+Beef M4 -> Bread M4, Bread+Beef M1 -> Beef M1, Bread+Beef M2 -> Beef M2, Bread+Beef M3 -> Beef M3, Bread+Beef M4 -> Beef M4.

Test Data	Bread+Beef M1 -> Bread M1	BreadBeef M2 -> Bread M2	Bread+Beef M3 -> Bread M3	Bread+Beef M4 -> Bread M4	Bread+Beef M1 -> Beef M1	Bread+Beef M2 -> Beef M2	Bread+Beef M3 -> Beef M3	Bread+Beef M4 -> Beef M4
Bread+Beef M1	0,0216	0,3484	0,2312	0,4941	0,0166	1,3377	0,3982	0,3214
Bread+Beef M2	0,2245	0,0153	0,2799	0,4665	0,1363	0,1282	0,3899	0,3190
Bread+Beef M2	0,2279	0,6785	0,0929	0,4910	0,1240	1,5387	0,0694	0,3238
Bread+Beef M2	0,2836	0,2545	0,4654	0,0213	0,1416	1,5127	0,4789	0,0286
Bread+Chicken M1	0,4125	0,2972	0,3664	0,4686	0,1539	0,8619	0,4074	0,3180
Bread+Chicken M2	0,8196	0,9079	0,3682	0,4620	0,1498	1,1626	0,4067	0,3209
Bread+Chicken M3	0,3246	0,2324	0,5522	1,0092	0,1338	1,5299	0,4226	0,3272
Bread+Chicken M4	0,2699	0,2300	0,3987	0,9223	0,1397	1,5212	0,6816	0,3286
Vinegar+Beef M1	0,5905	0,2186	0,8162	7,4563	0,1723	1,5138	0,3596	0,3286
Vinegar+Beef M2	0,3949	0,2511	0,4137	6,7361	0,1619	1,5188	0,5284	0,3286
Vinegar+Beef M3	0,4114	0,2375	0,7070	4,9963	0,1528	1,5225	0,4214	0,3285
Vinegar+Beef M4	0,2993	0,2443	0,4507	2,3318	0,1509	1,5172	0,6663	0,3286
Vinegar+Chicken M1	0,5410	0,2286	0,8139	6,1857	0,1673	1,5168	0,3642	0,3285
Vinegar+Chicken M2	0,5934	0,3098	0,5387	8,4704	0,1798	1,5076	0,4178	0,3286
Vinegar+Chicken M3	0,5413	0,2508	0,7671	7,2111	0,1636	1,5132	0,4139	0,3286
Vinegar+Chicken M4	0,3104	0,2440	0,4551	3,1958	0,1553	1,5191	0,5595	0,3286
Ground Truth	Bread M1	Bread M2	Bread M3	Bread M4	Beef M1	Beef M2	Beef M3	Beef M4

Table 4.9 Euclidean Distances Between Transformed Combinations and Ground Truth in Domain Adaptation Autoencoder to Decompose Bread+Beef Signals

Table 4.10 header abbreviations: columns are Bread+Chicken M1 -> Bread M1, Bread+Chicken M2 -> Bread M2, Bread+Chicken M3 -> Bread M3, Bread+Chicken M4 -> Bread M4, BreadChicken M1 -> Chicken M1, BreadChicken M2 -> Chicken M2, BreadChicken M3 -> Chicken M3, BreadChicken M4 -> Chicken M4.

Test Data	Bread+Chicken M1 -> Bread M1	Bread+Chicken M2 -> Bread M2	Bread+Chicken M3 -> Bread M3	Bread+Chicken M4 -> Bread M4	BreadChicken M1 -> Chicken M1	BreadChicken M2 -> Chicken M2	BreadChicken M3 -> Chicken M3	BreadChicken M4 -> Chicken M4
Bread+Beef M1	0,0995	0,1050	0,2327	0,3955	0,06985	0,31512	0,27307	1,52580
Bread+Beef M2	0,2175	0,2330	0,3333	0,3955	0,12446	0,42455	0,36918	1,57847
Bread+Beef M2	0,0673	0,2463	0,2741	0,3956	0,07013	0,34009	0,29931	1,00245
Bread+Beef M2	0,5379	1,2334	0,2294	0,3956	0,35756	0,29203	0,25326	0,80650
Bread+Chicken M1	0,0369	0,0915	0,2254	0,3954	0,05559	0,33844	0,28515	1,52852
Bread+Chicken M2	0,0605	0,0642	0,3061	0,3953	0,05960	0,23201	0,28993	1,50642
Bread+Chicken M3	0,1165	0,8220	0,0348	0,3956	0,23733	0,28315	0,25068	1,51321
Bread+Chicken M4	0,2161	0,9168	0,2931	0,3856	0,18601	0,31336	0,26632	0,75059
Vinegar+Beef M1	0,3445	1,2334	0,4487	0,3949	0,30216	0,31999	0,46460	1,54328
Vinegar+Beef M2	0,2237	1,2329	0,6215	0,3955	0,34665	0,31924	0,39591	1,55458
Vinegar+Beef M3	0,0738	0,9841	0,4066	0,3952	0,39725	0,27098	0,28488	1,55092
Vinegar+Beef M4	0,1909	1,0107	0,3576	0,3956	0,27327	0,30524	0,27343	1,54170
Vinegar+Chicken M1	0,2557	1,2197	0,3796	0,3950	0,31547	0,30157	0,39924	1,53722
Vinegar+Chicken M2	0,4494	1,2334	0,6630	0,3953	0,25788	0,41454	0,56485	1,54368
Vinegar+Chicken M3	0,2894	1,0140	0,5001	0,3949	0,34418	0,30063	0,41087	1,54621
Vinegar+Chicken M4	0,1321	1,0129	0,4076	0,3956	0,32736	0,30214	0,28555	1,01397
Ground Truth	Bread M1	Bread M2	Bread M3	Bread M4	Chicken M1	Chicken M2	Chicken M3	Chicken M4

Table 4.10 Euclidean Distances Between Transformed Combinations and Ground Truth in Domain Adaptation Autoencoder to Decompose Bread+Chicken Signals

Test Data	Vinegar+Beef M1 -> Vinegar M1	Vinegar+Beef M2 -> Vinegar M2	Vinegar+Beef M3 -> Vinegar M3	Vinegar+Beef M4 -> Vinegar M4	Vinegar+Beef M1 -> Beef M1	Vinegar+Beef M2 -> Beef M2	Vinegar+Beef M3 -> Beef M3	Vinegar+Beef M4 -> Beef M4
Bread+Beef M1	0,19513	1,11316	0,54652	1,05294	0,44022	1,97496	0,56482	1,08484
Bread+Beef M2	0,21145	0,65247	0,58710	1,04936	0,37725	1,94955	0,55319	1,08571
Bread+Beef M2	0,20306	0,37974	0,57033	1,05294	0,43842	1,87987	0,58921	1,08520
Bread+Beef M2	1,28471	2,77205	0,31688	0,31991	0,44740	0,99507	0,71260	0,61096
Bread+Chicken M1	0,19625	1,09765	0,59652	1,05294	0,44016	1,99233	0,57045	1,08439
Bread+Chicken M2	0,19700	0,73876	0,55854	1,05294	0,43374	1,99338	0,57575	1,08493
Bread+Chicken M3	0,60363	0,99186	0,31208	0,91119	0,43617	1,78644	0,52963	1,03829
Bread+Chicken M4	0,61227	0,87226	0,46954	0,96925	0,44402	1,01426	0,74198	0,96774
Vinegar+Beef M1	0,10664	1,56711	0,75542	0,41289	0,37503	1,71831	0,59283	1,17161
Vinegar+Beef M2	0,20083	0,10799	0,49643	0,43596	0,33973	1,06161	0,72204	1,07932
Vinegar+Beef M3	0,45961	1,54539	0,24599	0,35068	0,37367	1,66363	0,50777	0,69073
Vinegar+Beef M4	0,55893	0,88928	0,37659	0,11503	0,44027	0,99796	0,75631	0,59395
Vinegar+Chicken M1	0,34481	1,70530	0,64304	0,37862	0,38243	1,78326	0,56278	0,89574
Vinegar+Chicken M2	0,20550	0,77988	0,78694	0,50114	0,33903	1,18185	0,77440	1,33688
Vinegar+Chicken M3	0,19996	1,65530	0,61758	0,38936	0,35834	0,66158	0,60947	1,04312
Vinegar+Chicken M4	0,41919	0,53586	0,33176	0,26964	0,40152	0,99544	0,70988	0,61647
Ground Truth	Vinegar M1	Vinegar M2	Vinegar M3	Vinegar M4	Beef M1	Beef M2	Beef M3	Beef M4

Table 4.11 Euclidean Distances Between Transformed Combinations and Ground Truth in Domain Adaptation Autoencoder to Decompose Vinegar+Beef Signals

Test Data	Vinegar+Chicken M1 -> Vinegar M1	Vinegar+Chicken M2 -> Vinegar M2	Vinegar+Chicken M3 -> Vinegar M3	Vinegar+Chicken M4 -> Vinegar M4	Vinegar+Chicken M1 -> Chicken M1	Vinegar+Chicken M2 -> Chicken M2	Vinegar+Chicken M3 -> Chicken M3	Vinegar+Chicken M4 -> Chicken M4
Bread+Beef M1	0,21876	0,35729	0,44918	0,44134	0,79351	1,3896879	1,0116134	1,73358
Bread+Beef M2	0,22793	0,44628	0,55250	0,44134	0,82286	1,4346398	1,0621668	1,72549
Bread+Beef M2	0,22313	0,33438	0,53082	0,44134	0,78384	1,4698303	1,0569532	1,73127
Bread+Beef M2	0,28508	0,32295	0,82009	0,44123	0,63984	1,406739	1,0179622	1,85079
Bread+Chicken M1	0,22031	0,37096	0,46219	0,44134	0,83217	1,4233857	1,0335962	1,73517
Bread+Chicken M2	0,21881	0,36062	0,46228	0,44134	0,81954	1,425655	1,0286	1,73540
Bread+Chicken M3	0,20388	0,35446	0,38397	0,44126	0,63943	1,2492756	0,90524876	1,80239
Bread+Chicken M4	0,22589	0,34048	0,61042	0,44134	0,65650	1,4004908	1,0483528	1,73855
Vinegar+Beef M1	0,20350	0,34732	0,72272	0,44135	0,59912	1,093862	0,47397718	0,65079
Vinegar+Beef M2	0,21969	0,34194	0,35145	0,44122	0,58644	1,0388199	0,7191755	0,56619
Vinegar+Beef M3	0,21824	0,33407	0,33517	0,44134	0,63469	1,0198741	0,62725836	0,57408
Vinegar+Beef M4	0,22240	0,33378	0,55189	0,44137	0,60429	1,3346658	0,9969678	1,49560
Vinegar+Chicken M1	0,19587	0,35458	0,34762	0,44134	0,52195	1,051916	0,48437968	0,49263
Vinegar+Chicken M2	0,22077	0,25274	0,46857	0,44115	0,63305	1,00011	0,6107281	0,75324
Vinegar+Chicken M3	0,20668	0,35080	0,26788	0,44134	0,61955	1,0086931	0,31055294	0,44880
Vinegar+Chicken M4	0,22129	0,33377	0,49939	0,34134	0,58484	1,2686245	0,9311339	1,11428
Ground Truth	Vinegar M1	Vinegar M2	Vinegar M3	Vinegar M4	Chicken M1	Chicken M2	Chicken M3	Chicken M4

Table 4.12 Euclidean Distances Between Transformed Combinations and Ground Truth in Domain Adaptation Autoencoder to Decompose Vinegar+Chicken Signals

5. Conclusion

The Electronic nose (E-nose) device, blended with computer algorithms, is a powerful tool for accurately detecting odors or Volatile Organic Compounds (VOCs). It has a wide range of applications, and it is becoming more and more utilized. Even though it can detect and discriminate odors, a few studies have been conducted to approximate the odors in a mixture. In other words, approximating an odor in the mixture means separating/decomposing the signal generated by the E-nose device from a mixture of odors. That is called source/signal separation.

This study employed an Auto-encoder (AE) network to reduce dimensionality and reconstruct the data from the latent space. The latent space had a dimension of 4. Then, we calculated the reconstruction losses, i.e., Mean Squared Error (MSE). In these cases, these models are trained on individual components and tested on the individual components. The AE showed an outstanding result in detecting and discriminating different items and even discriminating between each measurement of the items. These results suggest a practical application in food quality control. However, it is essential to note that the detailed exploration of this practical application falls outside the scope of this research.

The optimal structure of the AE network from the individual component analysis is then incorporated to approximate the odors in a mixture. We trained the AE models on individual components, tested them on combination samples, and found the reconstruction loss. Ideally, the minimum reconstruction loss should represent the closeness of the data. The model's performance is not as good as AE's for detecting and discriminating the items. To some extent, this model could reconstruct most of the data within the same class. However, it was not able to discriminate between each measurement. The result of this experiment also brought the effect of odor intensity/strength. Some items have a more pungent odor, which may mask the other items in the mixture. Our model showed promising capability of approximating the predominant odors in a mixture (i.e., Vinegar samples in our study).

We conclude that AE network models can potentially approximate the odor classes in the mixture without prior knowledge of the source to some extent but not its corresponding measurement date, showing that this model cannot be utilized for odor

approximation and quality control in the case of food samples. However, its capability to approximate the predominant odor is effectual.

As a more advanced approach, we equipped the AE network with a domain adaptation technique to approximate the class odor in a mixture and its corresponding measurement occasion. We trained the AE on the combinations to construct its subcomponents. We have dedicated AE models for each sample to transform the data. This method calculates the Euclidean distance between the transformed data by AE and the ground truth (i.e., the individual components). The minimum Euclidean distance shows how close the transformed data to the ground truth is. This model showed promising results in approximating and decomposing the odor classes in the mixture and the measurement occasion.

Nevertheless, dedicated domain adaptation AE models require some prior information regarding the source, extensive training data, and intensive computational resources. However, it can be practically applied when working data in actual use cases is limited.

This study can be continued to make the domain adaptation AE independent from having prior knowledge. Eliminating the decoder part of the AE and incorporating the one-shot learning concept may potentially bring independence, allowing blind source separation. Having this level of autonomy in separating odors in a mixture has significant effects, and the ability to perform blind source separation could significantly impact fields such as food safety, which allows the precise detection of contaminants, resulting in an improvement of the overall quality of the food.

References

Arayawut, O., Kerdcharoen, T. and Wongchoosuk, C. (2022) 'Structures, Electronic Properties, and Gas Permeability of 3D Pillared Silicon Carbide Nanostructures', *Nanomaterials*, 12(11), p. 1869. Available at: https://doi.org/10.3390/nano12111869.

Buvé, C. *et al.* (2022) 'Application of multivariate data analysis for food quality investigations: An example-based review,' *Food Research International*, 151, p. 110878. Available at: https://doi.org/10.1016/j.foodres.2021.110878.

Chaloeipote, G. *et al.* (2021) '3D printed CuO semiconducting gas sensor for ammonia detection at room temperature', *Materials Science in Semiconductor Processing*, 123, p. 105546. Available at: https://doi.org/10.1016/j.mssp.2020.105546.

Deshmukh, S. *et al.* (2015) 'Application of electronic nose for industrial odors and gaseous emissions measurement and monitoring – An overview,' *Talanta*, 144, pp. 329–340. Available at: https://doi.org/10.1016/j.talanta.2015.06.050.

Dunkel, A. *et al.* (2014) 'Nature's chemical signatures in human olfaction: a foodborne perspective for future biotechnology.', *Angewandte Chemie (International ed. in English)*, 53(28), pp. 7124–43. Available at: https://doi.org/10.1002/anie.201309508.

García, M. *et al.* (2006) 'Electronic nose for wine discrimination', *Sensors and Actuators B: Chemical*, 113(2), pp. 911–916. Available at: https://doi.org/10.1016/j.snb.2005.03.078.

Gardner, J.W. and Bartlett, P.N. (1994) 'A brief history of electronic noses,' *Sensors and Actuators B: Chemical*, 18(1–3), pp. 210–211. Available at: https://doi.org/10.1016/0925-4005(94)87085-3.

Huang, M. *et al.* (2017) 'Correlation of Volatile Compounds and Sensory Attributes of Chinese Traditional Sweet Fermented Flour Pastes Using Hierarchical Cluster Analysis and Partial Least Squares-Discriminant Analysis,' *Journal of Chemistry*, 2017, pp. 1–8. Available at: https://doi.org/10.1155/2017/3213492.

James, D. *et al.* (2005) 'Chemical Sensors for Electronic Nose Systems', *Microchimica Acta*, 149(1–2), pp. 1–17. Available at: https://doi.org/10.1007/s00604-004-0291-6.

Karakaya, D., Ulucan, O. and Turkan, M. (2020) 'Electronic Nose and Its Applications: A Survey,' *International Journal of Automation and Computing*, 17(2), pp. 179–209. Available at: https://doi.org/10.1007/s11633-019-1212-9.

Karami, H., Rasekh, M. and Mirzaee-Ghaleh, E. (2020) 'Qualitative analysis of edible oil oxidation using an olfactory machine,' *Journal of Food Measurement and Characterization*, 14(5), pp. 2600–2610. Available at: https://doi.org/10.1007/s11694-020-00506-0.

Kerdcharoen, T. and Wongchoosuk, C. (2013) 'Carbon nanotube and metal oxide hybrid materials for gas sensing,' in *Semiconductor Gas Sensors*. Elsevier, pp. 386–407. Available at: https://doi.org/10.1533/9780857098665.3.386.

Kiselev, I. *et al.* (2018) 'On the Temporal Stability of Analyte Recognition with an E-Nose Based on a Metal Oxide Sensor Array in Practical Applications,' *Sensors*, 18(2), p. 550. Available at: https://doi.org/10.3390/s18020550.

Kondee, S. *et al.* (2022) 'Nitrogen-doped carbon oxide quantum dots for flexible humidity sensor: Experimental and SCC-DFTB study,' *Vacuum*, 195, p. 110648. Available at: https://doi.org/10.1016/j.vacuum.2021.110648.

Kramer, M.A. (1991) 'Nonlinear principal component analysis using auto-associative neural networks', *AIChE Journal*, 37(2), pp. 233–243. Available at: https://doi.org/10.1002/aic.690370209.

Li, X. *et al.* (2019) Multilayer domain adaptation method for rolling bearing fault diagnosis,'*S Signal Processing*, 157, pp. 180–197. Available at: https://doi.org/10.1016/j.sigpro.2018.12.005.

Muñoz-Aguirre, S. *et al.* (2007) 'Odor approximation of fruit flavors using a QCM odor sensing system,' *Sensors and Actuators B: Chemical*, 123(2), pp. 1101–1106. Available at: https://doi.org/10.1016/j.snb.2006.11.025.

Persaud, K. and Dodd, G. (1982) 'Analysis of discrimination mechanisms in the mammalian olfactory system using a model nose,' *Nature*, 299(5881), pp. 352–355. Available at: https://doi.org/10.1038/299352a0.

Phaisangittisagul, E. and Nagle, H.T. (2011) 'Predicting odor mixture's responses on machine olfaction sensors,' *Sensors and Actuators B: Chemical*, 155(2), pp. 473–482. Available at: https://doi.org/10.1016/j.snb.2010.12.049.

Prasetyawan, D. and Takamichi, N. (2020) 'Sensory Evaluation of Odor Approximation Using NMF with Kullback-Leibler Divergence and Itakura-Saito Divergence in Mass Spectrum Space', *Journal of The Electrochemical Society*, 167(16), p. 167520. Available at: https://doi.org/10.1149/1945-7111/abd110.

Seekaew, Y. *et al.* (2014) 'Low-cost and flexible printed graphene–PEDOT:P SS gas sensor for ammonia detection, *Organic Electronics*, 15(11), pp. 2971–2981. Available at: https://doi.org/10.1016/j.orgel.2014.08.044.

Seekaew, Y. *et al.* (2019) 'Synthesis, Characterization, and Applications of Graphene and Derivatives', in *Carbon-Based Nanofillers and Their Rubber Nanocomposites*. Elsevier, pp. 259–283. Available at: https://doi.org/10.1016/B978-0-12-813248-7.00009-2.

Seesaard, T. *et al.* (2020) 'A hybrid electronic nose system for discrimination of pathogenic bacterial volatile compounds', *Analytical Methods*, 12(47), pp. 5671–5683. Available at: https://doi.org/10.1039/D0AY01255F.

Seesaard, T., Lorwongtragool, P. and Kerdcharoen, T. (2015) 'Development of fabric-based chemical gas sensors for use as wearable electronic noses.', *Sensors (Basel, Switzerland)*, 15(1), pp. 1885–902. Available at: https://doi.org/10.3390/s150101885.Daa Silva Torres, E.A.F., Garbelotti, M.L. and Moita Neto, J.M. (2006) 'The application of hierarchical clusters analysis to the study of the composition of foods', *Food Chemistry*, 99(3), pp. 622–629. Available at: https://doi.org/10.1016/j.foodchem.2005.08.032.

Tan, J. and Xu, J. (2020) 'Applications of electronic nose (e-nose) and electronic tongue (e-tongue) in food quality-related properties determination: A review,'' *Artificial Intelligence in Agriculture*, 4, pp. 104–115. Available at: https://doi.org/10.1016/j.aiia.2020.06.003.

Thaler, E.R. and Hanson, C.W. (2005) 'Medical applications of electronicnose technology' *Expert Review of Medical Devices*, 2(5), pp. 559–566. Available at: https://doi.org/10.1586/17434440.2.5.559.

Le Thanh-Blicharz, J. and Lewandowicz, J. (2020) 'Functionality of Native Starches in Food Systems: Cluster Analysis Grouping of Rheological Properties in Different Product Matric,'', *Foods*, 9(8), p. 1073. Available at: https://doi.org/10.3390/foods9081073.

Thazin, Y., Pobkrut, T. and Kerdcharoen, T. (2018) 'Prediction of Acidity Levels of Fresh Roasted Coffees Using E-nose and Artificial Neural NNetwork,''n *2018 10th International Conference on Knowledge and Smart Technology (KST)*. IEEE, pp. 210–215. Available at: https://doi.org/10.1109/KST.2018.8426206.

Traiwatcharanon, P. *et al.* (2022) 'Electrochemical paraquat sensor based on lead oxide nanoparticles,' *RSC Advances*, 12(25), pp. 16079–16092. Available at: https://doi.org/10.1039/D2RA02034C.

Traiwatcharanon, P., Timsorn, K. and Wongchoosuk, C. (2017) 'Flexible room-temperature resistive humidity sensor based on silver nanoparticles,'',' *Materials Research Express*, 4(8), p. 085038. Available at: https://doi.org/10.1088/2053-1591/aa85b6.

Zhang, L. and Tian, F. (2014) 'Performance Study Multilayer Perceptrons in a Low-Cost Electronic Nose,' *IEEE Transactions on Instrumentation and Measurement*, 63(7), pp. 1670–1679. Available at: https://doi.org/10.1109/TIM.2014.2298691.