

METHOD FOR A REFINED SCENT PROFILE DETECTION AND ANALYSIS OF A GAS MIXTURE

Technical field

The present invention relates to the field of signal processing of gas sensor array signals.

- 5 More specifically the invention relates to electronic gas sensing technology, machine learning algorithms for data analysis, and optimization of sensor operational parameters to enhance the accuracy and efficiency of scent profile identification.

Background art

- 10 A gas sensor array, often referred to as an electronic nose, is a system comprising multiple sensors that respond to different gases. Each sensor in the array produces a signal when exposed to a particular gas or a mixture of gases. The pattern of responses from these sensors can then be used to identify and quantify the gases present.

Signal processing of gas sensor array data involves several steps:

- 15 1. Signal Acquisition: This is the initial step where the electrical signals from the sensors are collected. These signals are usually in the form of current or voltage changes corresponding to gas concentration.
2. Preprocessing: The raw signals often contain noise and may be affected by factors like temperature and humidity. Preprocessing techniques like filtering, normalization, and baseline correction are used to clean and standardize the data.
- 20 3. Feature Extraction: This step involves extracting meaningful information from the preprocessed signals. The goal in this step is to reduce the dimensionality of the data (i.e., reduce the number of variables under consideration) to improve the performance of subsequent analysis, while still preserving the essential information. Features might be statistical measures like the mean or variance of the sensor signals, specific patterns or
- 25 shapes in the signal, or even frequency components from a Fourier transform.
4. Pattern Recognition: In this step, the extracted features are analyzed to classify and recognize the type and concentration of gases present. This can involve various statistical and machine learning techniques, including linear classification based on statistical theory or nonlinear classification based on neural networks.
- 30 5. Model Training and Testing: The pattern recognition algorithms are trained using known data samples, and then their performance is tested and evaluated with new samples to verify the accuracy and reliability of the system.

6. Decision Making: Based on the recognition and classification results, decisions or predictions are made regarding the types and concentrations of gases detected by the sensor array.

Applications

- 5 The applications of gas sensor arrays with robust signal processing are vast:
- Environmental Monitoring: Detecting pollutants and harmful gases in the air.
 - Medical Diagnostics: Analyzing breath samples for markers of diseases.
 - Food Quality Control: Identifying spoilage or contamination in food products.
 - Safety: Detecting gas leaks in industrial settings or hazardous environments.

Challenges and Future Directions

While gas sensor arrays coupled with advanced signal processing techniques offer immense potential, there are challenges:

- Selectivity and Sensitivity: Ensuring the sensors are highly selective to specific gases and sensitive enough to detect low concentrations.
- 15 • Environmental Variability: The sensors must be robust against changes in temperature, humidity, and other environmental factors.
- Data Analysis: As sensor technology advances, the data becomes more complex, requiring more sophisticated algorithms for accurate analysis.

20 Researchers are continually exploring new materials for sensors, advanced signal processing algorithms, and machine learning techniques to overcome these challenges.

Detailed description

Coarse and refined analysis

Coarse analysis is the first stage of evaluating the gas mixture, where the sensor array's initial readings are broadly analyzed to identify a wide range of potential scent profiles. The scent
25 profiles may include not just the target scents but also background scents or any other scent profiles that could influence the selection of refinement parameters. The coarse analysis may or may not employ machine learning models. In some embodiments the coarse analysis may employ signal normalization, signal smoothing, noise removal filtering and draft reduction.

Refined analysis is a deeper, more targeted examination of the gas mixture, conducted after the
30 sensor array has been fine-tuned based on the refinement parameters. This phase aims to isolate and accurately identify the subset scent profiles, particularly focusing on the target scents. The refined analysis may or may not employ machine learning models. Preferable technics for the refined analysis include Kalman filtering and Wiener filtering (adaptive filtering-based

methods), and/or machine learning-based feature extraction and dimensionality reduction with for example neural networks (e.g. auto-encoders, convolutional neural networks, etc).

Initial and expanded data set

- 5 The "initial data set" refers to the first collection of readings obtained from the gas sensor array when it is exposed to the gas mixture for analysis. This data set is gathered under initial, predefined operational settings of the sensor array before any refinements are made based on the analysis. The operational setting may be predefined in an adaptive manner by being based on previous readings made on previously measured gas mixtures, e.g., a gas mixture measured at the same location but at a different point in time.
- 10 The "expanded data set" is a more comprehensive collection of readings acquired from the gas sensor array after adjusting the operational parameters based on the refinement parameters derived from the initial analysis. This data set is collected under modified operational conditions that are specifically tuned to enhance the detection and analysis of the subset scent profiles identified as relevant.
- 15 It is important to note that both the initial and expanded data set comprise reading on the same mixture of gas.

Superset and subset of scent profiles

- The "superset of scent profiles" refers to a range of scent profiles detected by the gas sensor array in the gas mixture. It serves as the foundation upon which more focused analysis is built, enabling the method to cover all bases before narrowing down to specific scent profiles of interest.
- 20

- A "subset of scent profiles" refers to one or several scent profiles within the superset. These subsets are identified based on their significance to the end goal, such as target scents, background scents, or indicative scents that may influence the refinement of the operational parameters.
- 25

- A scent profile within the context of the application may refer to target scents or background scents. Target scents are the scents which the analysis aims to detect. Background scents are present in the environment but are not the main focus of the analysis. Their identification, however, can be crucial for context and accuracy as they might provide important information for refining the analysis (e.g., scents that signal the presence of specific target compounds or conditions).
- 30

Operational parameter

Sampling Rate: This refers to the frequency at which the sensor array collects data from the gas mixture. For more volatile compounds, a higher sampling rate might be necessary to capture

rapid changes in concentration. Conversely, for more stable compounds, a lower sampling rate could suffice.

Duration of Sensing: This is the total time period over which data is collected. longer duration may be required for detecting scent profiles that manifest over time or for compounds with lower volatility. Shorter durations could be used for more potent or easily detectable scents.

Selection of Sensors: This involves choosing specific sensors within the array that are most sensitive to the target scent profiles. If the target scent profile is known to contain specific compounds, sensors that are particularly responsive to these compounds would be selected.

Heating Profile: This relates to the temperature conditions under which the sensors operate. Certain sensors might require higher temperatures to detect specific compounds effectively. Adjusting the heating profile can help in improving the sensor's responsiveness to particular scent profiles.

Energy Feed to Selected Sensors: This parameter dictates the amount of power supplied to the sensors. Adjusting the energy feed can influence the sensor's sensitivity and response time.

Higher energy might enhance the sensor's ability to detect low-concentration compounds.

Refinement parameters

Refinement parameters play a role in transitioning from a general, broad-spectrum analysis to a more targeted and precise investigation. They are essentially the insights gained from the initial data that guide the fine-tuning of the gas sensor array's operational parameters.

These parameters are typically based on the initial set of potential scent profiles identified during the coarse analysis. They can include information about the types of compounds present, their concentration ranges, their chemical properties, and how they interact or change over time. The primary purpose of these parameters is to provide a focused direction for the subsequent analysis. They inform decisions on how to adjust the operational parameters of the sensor array (like sampling rate, sensor selection, heating profiles, etc.) for a more efficient and accurate detection of specific scent profiles. For instance, if the coarse analysis suggests the presence of certain volatile organic compounds, the refinement parameters might include specific sensitivities and response characteristics required to detect these compounds accurately.

Examples of Refinement Parameters:

- Identifying the most relevant sensors within the array that are responsive to the predicted compounds.
- Determining the optimal temperature range for sensor operation based on the volatilization points of the compounds.

- Establishing the appropriate duration of sensing to capture the complete profile of slowly evolving or reacting compounds.
- Deciding on the sampling rate based on the volatility and reaction kinetics of the identified compounds.

5

Different types of sensors

Different types of sensors that can be commonly integrated into a gas sensor array include:

10 Metal Oxide Semiconductor (MOS) Sensors: These sensors detect gases by a change in resistance when a gas interacts with a heated metal oxide surface. They are widely used for detecting a variety of gases, including carbon monoxide, hydrogen, and volatile organic compounds.

Electrochemical Gas Sensors: These sensors measure the concentration of a target gas by oxidizing or reducing the gas at an electrode and measuring the resulting current. They are particularly effective for detecting gases like carbon monoxide, chlorine, and nitrogen oxides.

15 Infrared (IR) Sensors: These sensors work by measuring the absorption of infrared light by gas molecules. They are commonly used for detecting gases that absorb infrared light, such as carbon dioxide and hydrocarbons.

Photoionization Detectors (PID): PID sensors use ultraviolet light to ionize gas molecules and measure the resulting current. They are effective for detecting volatile organic compounds.

20 Catalytic Bead Sensors: These sensors detect flammable gases by measuring the heat generated when the gas oxidizes on an active bead surface. They are commonly used for detecting hydrocarbons and other combustible gases.

Solid-State Sensors: These sensors use a solid electrolyte to detect specific gases. They are known for their stability and are used for gases like ammonia and hydrogen sulfide.

25 Conductive Polymer Sensors: These sensors change their electrical resistance when exposed to certain gases and are used for detecting gases like ammonia and organic vapors.

Optimizing operational parameters for decreased drift

Sampling Rate:

30 **High Sampling Rate:** Continuously operating the sensors at a high sampling rate, especially for volatile compounds, might lead to faster degradation or changes in the sensor materials due to more frequent exposure to the target gases. This can potentially accelerate drift.

Low Sampling Rate: Operating at a lower sampling rate could reduce the wear and tear on the sensor elements, potentially slowing down the drift process. However, it might not capture rapid changes in gas concentration effectively.

Duration of Sensing:

5 **Longer Duration:** Prolonged exposure to the target gases or the environment can cause gradual changes in the sensor's characteristics, leading to drift. This is especially true for sensors that might be sensitive to the accumulation of contaminants or degradation over time.

10 **Shorter Duration:** Shorter sensing durations might mitigate some of the long-term effects that contribute to drift. However, they may not be as effective for detecting slow-emerging patterns in gas concentrations.

Selection of Sensors:

15 Different sensors have different propensities for drift. Choosing sensors that are specifically tailored to the target gases can minimize drift, especially if these sensors are inherently more stable or less prone to environmental influences.

Heating Profile:

20 **Higher Temperatures:** Operating sensors at higher temperatures can increase their sensitivity but may also accelerate sensor degradation or changes in the sensor material, leading to drift.

20 **Optimized Temperature Conditions:** Maintaining an optimal heating profile that balances sensitivity with longevity can help in minimizing drift. This involves avoiding excessively high temperatures that might hasten sensor degradation.

Energy Feed to Selected Sensors:

25 **Higher Energy Feed:** Providing more power to the sensors can improve their responsiveness and sensitivity but might also increase the rate of sensor degradation, contributing to drift.

Balanced Energy Feed: Adjusting the energy feed to an optimal level that ensures effective sensing without overstraining the sensor materials can help in reducing the rate of drift.

30 **Optimizing operational parameter based on gas properties**

Molecular Weight and Size:

- Sensing Duration: Heavier and larger molecules may diffuse more slowly, potentially requiring longer sensing durations to achieve accurate measurements.
- Number of Sensors: Larger molecules might require sensors with larger surface areas or multiple sensors to capture sufficient interactions.
- 5 – Types of Sensors: Certain sensor types may be more effective for larger or heavier molecules due to their interaction mechanisms.
- Sampling Rate: Slower diffusion rates might allow for lower sampling rates, as changes in concentration levels occur more slowly.
- Sensor Energy Feed: Larger or heavier molecules might require sensors with higher
- 10 energy inputs to facilitate sufficient interactions for detection.
- Heating Profile: Molecules with larger sizes may diffuse more slowly, possibly necessitating a specific heating profile to maintain an optimal reaction rate and sensor response.

Reactivity:

- 15 – Sensing Duration: Highly reactive gases may require shorter sensing durations to prevent sensor degradation or to capture rapid reactions.
- Number of Sensors: Multiple sensors might be needed to differentiate between the target gas and other reactive species present.
- Types of Sensors: Sensors resistant to chemical degradation or those specifically
- 20 designed for reactive gases might be necessary.
- Sampling Rate: High reactivity might necessitate a higher sampling rate to capture rapid concentration changes.
- Sensor Energy Feed: Highly reactive gases might require sensors with controlled energy feeds to prevent rapid degradation or unwanted reactions.
- 25 – Heating Profile: Temperature regulation is crucial for reactive gases to maintain the integrity of the sensor and to control the reaction kinetics.

Concentration:

- Sensing Duration: Higher concentrations might be detected more quickly, potentially reducing the necessary sensing duration.
- 30 – Number of Sensors: In environments with low concentrations, more sensors might be needed to ensure reliable detection.
- Types of Sensors: Sensors with higher sensitivity might be selected for low-concentration environments.

- **Sampling Rate:** Variable concentrations might require adaptive sampling rates to ensure accurate detection at all times.
- **Sensor Energy Feed:** Sensors detecting low-concentration gases might need higher energy inputs to increase sensitivity and signal-to-noise ratios.
- 5 – **Heating Profile:** The optimal temperature might vary with concentration to either enhance the sensor's response or prevent saturation.

Boiling and Melting Points:

- **Sensing Duration:** Gases with temperatures close to their boiling or melting points might exhibit different behaviors, affecting the optimal sensing duration.
- 10 – **Number of Sensors:** Temperature-controlled sensors or multiple sensors might be needed to account for phase changes.
- **Types of Sensors:** Sensors capable of operating effectively in the temperature range of the gas's boiling or melting points might be necessary.
- **Sampling Rate:** Rapid phase changes might necessitate higher sampling rates to capture transient behaviors.
- 15 – **Sensor Energy Feed:** Sensors for gases near their boiling or melting points might require energy adjustments to accommodate phase changes and ensure consistent measurements.
- **Heating Profile:** Precise temperature control is necessary to avoid condensation or vaporization of the gas, which could significantly affect sensor readings.
- 20

Solubility and Humidity Sensitivity:

- **Sensing Duration:** Gases that are highly soluble in water may require longer durations or special considerations in humid conditions.
- **Number of Sensors:** Additional sensors might be required to measure and compensate for humidity levels.
- 25 – **Types of Sensors:** Sensors that are either unaffected by humidity or can measure it alongside the gas concentration might be chosen.
- **Sampling Rate:** Changes in humidity might affect the gas concentration, requiring an adjusted sampling rate to maintain accuracy.
- **Sensor Energy Feed:** Adjusting the energy feed might be necessary to account for the effects of humidity on gas solubility and sensor response.
- 30 – **Heating Profile:** Temperature regulation can help mitigate the effects of humidity by controlling condensation and adsorption processes.

Specificity and Cross-Sensitivity:

- Sensing Duration: Gases with high specificity requirements might need longer sensing durations to accurately distinguish them from other gases.
- Number of Sensors: More sensors might be needed to differentiate the target gas from other similar gases.
- 5 – Types of Sensors: Selective sensors that are less prone to interference from other gases might be preferred.
- Sampling Rate: A higher sampling rate might be needed to distinguish between closely related gases quickly.
- Sensor Energy Feed: Energy inputs might be fine-tuned to enhance the specificity of
- 10 sensors and reduce cross-sensitivity to other gases.
- Heating Profile: Careful temperature regulation can help improve specificity by optimizing the sensor's response to the target gas while minimizing responses to others.

Stability and Decay Rate:

- Sensing Duration: Gases that decay or react over time might require quicker
- 15 measurements for accurate readings.
- Number of Sensors: In cases of unstable gases, redundant sensors might be used to ensure consistent detection.
- Types of Sensors: Sensors designed for rapid detection and those less affected by the decay products might be optimal.
- 20 – Sampling Rate: Faster decay rates might necessitate higher sampling rates to capture accurate data before significant changes occur.
- Sensor Energy Feed: For unstable gases, the energy feed might be adjusted to ensure rapid detection before significant decay occurs.
- Heating Profile: A tailored heating profile might be necessary to stabilize the gas or to
- 25 accelerate the sensing process for gases with high decay rates.

Iterative improvement of coarse analysis

To improve the classification of supersets in the coarse analysis using machine learning (ML), informed by the results from a refined analysis tailored to precisely determine concentrations and/or presence of subset scent profiles, it is possible to implement a feedback loop that

30 enhances the ML model. The following is a detailed explanation of how this could be implemented:

1. Data Integration from Refined Analysis:

- **Approach:** Gather detailed data from the refined analysis, which focuses on accurately determining the concentrations and presence of specific subset scent profiles.
- **Application:** Analyze this data to identify patterns, correlations, or characteristics that were not evident or misclassified in the initial coarse analysis.

2. Feature Enhancement for the ML Model:

- **Approach:** Use insights from the refined analysis to enhance the feature set used by the ML model. This could include adding new features or modifying existing ones to better capture the nuances of the scent profiles.
- **Application:** If the refined analysis reveals that certain scent profiles have unique signatures at specific concentrations or under certain conditions, these characteristics can be incorporated as new features in the ML model.

3. Re-training the ML Model:

- **Approach:** Employ the augmented data and enhanced features to re-train the ML model. This process involves using machine learning algorithms to update the model based on the new insights.
- **Application:** The re-trained model should now be better at identifying the superset of scent profiles during the coarse analysis, as it has learned from the more precise data obtained in the refined analysis.

4. Validation and Continuous Learning:

- **Approach:** Continuously validate the ML model's performance using new datasets and refine it as needed. This ensures the model remains accurate over time and under varying conditions.
- **Application:** Regularly compare the ML model's superset classifications with results from refined analyses to identify any discrepancies or areas for improvement.

5. Cross-Referencing and Pattern Recognition:

- **Approach:** Implement advanced pattern recognition techniques that can cross-reference various scent profiles and their concentrations to identify complex patterns.

- **Application:** Use techniques like deep learning, which can handle high-dimensional data and uncover intricate patterns that might be missed by simpler ML models.

6. Feedback Loop Optimization:

- **Approach:** Optimize the feedback loop to ensure that the ML model receives relevant and high-quality data from the refined analysis.
- **Application:** Develop a systematic approach for selecting which data from the refined analysis should be used to update the ML model, focusing on data that provides the most value in improving superset classification.

By integrating these steps, the machine learning model used in the coarse analysis becomes more sophisticated and accurate in classifying the superset of scent profiles. This approach leverages the detailed and precise information gathered during the refined analysis to continually enhance the ML model's ability to classify complex scent profiles more effectively.

BACKGROUND – SECOND ASPECT

Infectious animal diseases caused by pathogenic microorganisms such as bacteria, fungi, and viruses threaten the health and well-being of farm animals, limit their productivity, and significantly increase economic losses to many sectors. Detection of pathogen and other causes of diseases is an important step for the diagnosis and successful treatment of the animal diseases. The conventional techniques employed to diagnose the pathogens in farm animals are generally time-consuming. Also, these techniques often provide inconclusive or inconsistent results owing to their reliance on human expertise and/or chances of human errors. Therefore, it is advantageous to have an autonomous mechanism for automated diagnosis of diseases, which reduces human involvement and makes accurate diagnosis of animal diseases.

One of the existing automated diagnosis techniques is disclosed in the patent document WO2017212437A1. The said document discloses use of sensor devices to detect enteric diseases in farm animals. Particularly, this document discloses a detection system comprising a sensor device, which in turn comprises a suction means for suctioning air from an environment with said animals, and a sensor means configured to determine information about the type and concentration of a plurality of smelling molecules in the air and then to emit a signal representative of the information of smelling molecules. Further, the detection system includes a transmission device, which is configured to convey the signal representative of smelling molecules. A processing device of self-learning type is configured to receive and process the signal representative of the smelling molecules, to detect a risk of enteric diseases in the animals associated with the information about the type and concentration of the smelling molecules.

Further, a signalling device is configured to report to a user about the risk of enteric diseases in the animals before the enteric diseases arise.

However, the said document focuses only on a single type of diseases and does not hint the user about a probable cause of the disease, and thus, fails to completely address the limitations in the existing animal disease diagnostic techniques. Moreover, in the said document, the classification and analysis of the enteric diseases is done using a Principal Component Analysis (PCA) model and a Linear Discriminant Analysis (LDA) model, which are two most used methods for classification. As a result, the classification disclosed by the said document is extremely weak and inaccurate, particularly in scenarios where the data is obtained from different farms, which are in different geographical locations and are associated different climates.

In addition, the said document suggests use of Metal Oxide Semiconductor (MOS) gas sensors for the diagnosis. However, the MOS gas sensors fail to create repeatable results, since the temperature, humidity and other environmental conditions presented in the environment affect the results of these sensors. Thus, in order to eliminate any variations in the results of these sensors, the results of these sensors need to be studied in different environmental conditions and the deviations/drift in the results need to be compensated. Thus, the idea of relying on gas sensors for VOCs is both unreliable and complex. Therefore, the factors such as humidity, temperature and other randomly occurring environmental changes in the farm, make the diagnosis method suggested in the said document completely unreliable, since the device needs to be calibrated first based on data collected from countless farms and in many different environmental conditions. Finally, the said document also does not suggest considering the health profile/disease history of the animals as a factor for diagnosing the animals. The disease history of the animals plays an extremely vital role in the accuracy of diagnosis, since the animals emit different odours for the same disease based on their age, gender, and health history.

In view of the limitations in the existing technologies, it is evident that a stronger and more robust classification and analysis technique is needed to be able to accurately detect different diseases, as soon as they emit from the body of the animal, and before affecting other animals in the herd. Additionally, an accurate and extensive data collection mechanism is needed to create a real-time database for analysing different diseases occurring in different geographical locations / climatic conditions.

The information disclosed in this background of the disclosure section is only for enhancement of understanding of the general background of the invention and should not be taken as an acknowledgement or any form of suggestion that this information forms the prior art already known to a person skilled in the art.

SUMMARY – SECOND ASPECT

Disclosed herein is an AI based VOC analyser for diagnosing diseases in the farm animals by capturing and analysing odour of the animals, captured from body, breath, urine, saliva, and other discharges of the animal, thereby offering a reliable diagnostic tool for the farmers and veterinarians.

In an embodiment, the proposed VOC analyser (alternatively referred to as a diagnostic device) may be configured to continuously collect and analyse the odour data and inform the farmers as soon as there is a change in the odour of the animal. Also, the said diagnostic device may be configured to predict the probable cause of the odour changes and determine if it is due to a disease or disorder, with the help of an AI engine that is extensively trained on data including, without limitation, age, gender, breed and behaviour of the animal, environment conditions associated with the animal and animal sickness history.

In an embodiment, as stated above, in addition to the odour data, the diagnostic device may consider various other factors such as, without limitation, breath analysis and headspace sampling for accurate diagnosis of the diseases. Moreover, to further enhance the accuracy of the diagnosis, the device does not rely only on the odour of the animals, but also evaluate health statuses and symptoms observed in the animals. For example, information such as, without limiting to, animal health and disease history, farm disease history, age, gender, and breed information of the animal and even disease statistics and survey information related to the geographical location of the animal may be considered for the analysis. Since a wide range of data is utilized for the analysis, the proposed diagnostic device may be used for diagnosing a numerous diseases including, without limitation, Ketosis, Metritis, Salmonella, Mastitis, Campylobacter, cancer, diabetes, and the like.

In some implementations, the proposed diagnostic device may be also used for diagnosing diseases in human beings, by analysing human body odour, saliva, urine, and breath samples of the humans. For instance, the device may be implemented for detecting diseases/disorders such as Covid-19, flue, diabetes, Tuberculosis (TB), urinary tract infections, cancer etc. in the humans.

The foregoing summary is illustrative only and is not intended to be in any way limiting. In addition to the illustrative aspects, embodiments, and features described above, further aspects, embodiments, and features will become apparent by reference to the drawings and the following detailed description. For a better understanding of exemplary embodiments of the present invention, together with other and further features and advantages thereof, reference is made to the following description, taken in conjunction with the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS – SECOND ASPECT

The embodiments of the disclosure itself, as well as a preferred mode of use, further objectives, and advantages thereof, will best be understood by reference to the following detailed
5 description of an illustrative embodiment when read in conjunction with the accompanying drawings. One or more embodiments are now described, by way of example only, with reference to the accompanying drawings in which:

FIG. 1A illustrates an exemplary arrangement of the proposed VOC analyser or diagnostic device in accordance with various embodiments of the present disclosure.

10 **FIG. 1B** illustrates interaction between the AI engine and various sources of information in accordance with various embodiments of the present disclosure.

FIG. 2 shows a flowchart illustrating a method of diagnosing diseases in animals using a VOC analyser in accordance with some embodiments of the present disclosure.

The figures depict embodiments of the disclosure for purposes of illustration only. One skilled in
15 the art will readily recognize from the following description that alternative embodiments of the structures and methods illustrated herein may be employed without departing from the principles of the disclosure described herein.

DETAILED DESCRIPTION – SECOND ASPECT

In the present document, the word "exemplary" is used herein to mean "serving as an example,
20 instance, or illustration." Any embodiment or implementation of the present subject matter described herein as "exemplary" is not necessarily to be construed as preferred or advantageous over other embodiments.

While the disclosure is susceptible to various modifications and alternative forms, specific embodiment thereof has been shown by way of example in the drawings and will be described in
25 detail below. It should be understood, however that it is not intended to limit the disclosure to the particular forms disclosed, but on the contrary, the disclosure is to cover all modifications, equivalents, and alternative falling within the spirit and the scope of the disclosure.

The terms "comprises", "comprising", or any other variations thereof, are intended to cover a non-exclusive inclusion, such that a setup, device or method that comprises a list of components
30 or steps does not include only those components or steps but may include other components or steps not expressly listed or inherent to such setup or device or method. In other words, one or more elements in a device or system or apparatus preceded by "comprises... a" does not,

without more constraints, preclude the existence of other elements or additional elements in the device or system or apparatus.

In the following detailed description of the embodiments of the disclosure, reference is made to the accompanying drawings that form a part hereof, and in which are shown by way of

5 illustration specific embodiments in which the disclosure may be practiced. These embodiments are described in sufficient detail to enable those skilled in the art to practice the disclosure, and it is to be understood that other embodiments may be utilized and that changes may be made without departing from the scope of the present disclosure. The following description is, therefore, not to be taken in a limiting sense.

10 In an embodiment, the proposed VOC analyser studies each animal based on odours emitted from the body of the animal and/or from samples of milk, urine, manure, and saliva of the animal. The device studies the disease related odours much closer to each animal individually and combines the collected information with other information such as the animal health history, feeding status, vaccination, age, and disease history of the animals, in order to offer
15 comprehensive diagnostics to the farmers and/or veterinarians.

In an embodiment, the VOC analyser is configured to continuously collect (i.e., in real-time) the odour data to detect and analyse even the smallest changes in the animal body odour. In an embodiment, the real-time odour data collected is compared with the odour emitted from the breath, milk, urine, manure, and other secretions of the animal to ensure the accuracy. Finally,
20 the collected data is compared with the animal health journal (that includes disease history, age, gender, health profile etc. information related to the animal) in order to offer accurate diagnosis to the farmers in real time.

As stated in the above paragraph, the animal health journal plays a major role in accurate diagnosis of the diseases. In an embodiment, the animal health journal is created, for example, by
25 studying and creating disease databases from over 100 different farms across 5 different countries, using different devices at random times of day. Such a detailed analysis helps in creating a disease odour fingerprint for each possible disease. Thereafter, each animal is individually studied every day, and as soon as the odours of the animal get closer to the disease fingerprint, the concerned farmers are notified about a possibility of disease occurring in the
30 animals.

In an embodiment, the self-learning AI engine analyses the disease history of each animal, disease history of the farm and even the disease history / health statistics of the surrounding geographical region to enhance the accuracy with which occurrence of a disease in the farm is predicted and notified to the farmers. Generally, it may be observed that pattern and

characteristics of each disease correspond only to a specifically designed classification algorithm. In the proposed method, the AI engine uses a combination of many different classification algorithms to predict and/or detect a particular disease most accurately. Thus, the AI engine increases the detection accuracy by incorporating the animal health history, age,
5 gender, feeding status, and other information related to common issues in the farm and/or the entire country.

Thus, in summary, the proposed VOC analyser offers a fully automated and AI-driven diagnostic technique for continuous, accurate and reliable diagnosis and control of diseases in the farm animals. The uniqueness of the proposed VOC analyser lies in the aspect of combining the real-
10 time VOC information obtained from the farm with the animal health journal to create a pattern and using a pre-training AI engine to analyse the pattern (i.e., pattern recognition) to predict and diagnose the potential diseases in the animals.

The other uniqueness of the proposed VOC analyser is in the specific design / hardware construction of the device. For instance, the proposed VOC analyser is configured to take 30
15 seconds of sample odour and 30 seconds of clean air to study variations in the sensor resistance, which in turn helps in detecting variations in the VOC concentrations. None of the existing technologies seem to suggest a device with same or similar design factors.

FIG. 1 illustrates an exemplary arrangement of the proposed VOC analyser or diagnostic device in accordance with various embodiments of the present disclosure.

20 In an embodiment, the present disclosure makes a first of its kind attempt to combine a VOC analyser and Artificial Intelligence (AI) technology for use in veterinary to diagnose animal diseases in the farms. In an embodiment, the VOC analyser is a smart device comprised primarily of Gas Sensor Arrays (GSAs), which mimic the human's olfactory system. In an embodiment, the odour data is gathered and measured using the VOC analyser and it is provided to an analytics
25 application, installed on an external computing system such as a smartphone, for creating a database of odours that could be used during the development of an AI based pattern recognition algorithm for detection of diseases in the animals.

In an embodiment, the VOC analyser may be a handheld device. Further, as shown in FIG. 1, the VOC analyser may comprise a power inlet, a first inlet, a second inlet and a first outlet. In an
30 implementation, the power inlet may be used to connect the VOC analyser to a continuous power supply and/or alternatively for recharging a battery configured in the VOC analyser. As an example, the battery of the VOC analyser may be charged using a regular USB charger or a type-C USB charger. In an embodiment, it is recommended to connect the VOC analyser to a power source while using it for the analysis.

In an embodiment, the first inlet may be used supplying fresh air to the gas sensors (not shown in FIG. 1) configured in the VOC analyser from a surrounding environment. In an embodiment, the fresh air being supplied may be filtered using an active carbon filter for preventing any contaminants from entering the sensory module.

- 5 In an embodiment, the second inlet of the VOC analyser may be used for supplying odour from a test sample container to the sensory module of the VOC analyser. As an example, the test sample container may be a plastic and/or metal container, which may be used for holding a test sample such as, without limiting to, saliva, urine, manure etc. of the animal, whose odour data need to be analysed. In an embodiment, the test sample container may be replaced after each analysis to
10 prevent any contamination in the test sample.

In an implementation, the test sample container may be connected to the fresh air supply from the active carbon filter through a T-junction hose, as shown in FIG. 1 (specifically, the directional arrows shown in FIG. 1 indicate a direction of movement of air between the active carbon filter, first inlet, test sample container and the second inlet of the VOC analyser). In other words, the
15 fresh air sucked-in through the active carbon filter is supplied both to the first inlet and the test sample container. This enables the VOC analyser to accurately measure a deviation in the odour data of the test sample with respect to the fresh air in the environment. Thus, any deviation detected in the odour of the test sample may be considered as a deviation and/or disorder in the health of the animal, whose test sample is being analysed.

- 20 In an embodiment, as soon as the VOC analyser is powered ON (using a power switch on the VOC analyser, which is not shown in FIG. 1), the VOC analyser starts receiving/sucking the fresh air through the first inlet connected to the active carbon filter. At the same time, the fresh air is also supplied to the test sample container through the T-junction. Thereafter, the fresh air supplied to the test sample container gets mixed with the test sample, causing the test sample to produce an
25 odour corresponding to the test sample. Further, the test sample odour is supplied to the VOC analyser through the second inlet. After receiving both the fresh air and the test sample odour, the sensory circuit (i.e., the GSAs) of the VOC analyser analyses both the fresh air and the sample odour. During the analysis, the volatile compounds comprised in the test sample odour affect/change the resistance of the sensors. The deviations/changes in the sensor resistance are
30 then translated to corresponding VOC concentration values. Subsequently, the VOC information (i.e., concentration of each VOC component) calculated by the sensory circuit is transmitted to an external computing unit for further analysis and diagnosis. In an embodiment, the exhaust air (i.e., the analysed air) is released from the VOC analyser through the first outlet.

In an embodiment, the external computing unit may be configured with pre-trained AI based
35 pattern recognition algorithms, which analyse the real-time VOC information received from the

VOC analyser to predict a possibility of disease in the animal. In an embodiment, the possibility of disease is predicted by combining the real-time VOC information with the animal health journal and creating a pattern corresponding to the deviations in the VOC information.

Subsequently, the generated patterns are recognized using the preconfigured AI pattern

- 5 recognition algorithms to confirm occurrence of the diseases in the animal, whose sample is being analysed.

In an embodiment, once the possibility of disease is detected, the external computing unit may notify the same to the farmer and/or concerned veterinarian, so that the farmer/veterinarian may take suitable measures to handle the diseases. For example, handling the disease may
10 include treating the disease and preventing the disease from spreading to other animals in the herd/farm.

In an embodiment, the external computing unit and the VOC analyser may be connected using at least one of a wired or a wireless communication network. As an example, the external computing unit may be a smartphone of the farmer and/or the veterinarian.

- 15 In an embodiment, the operation of the AI based pattern recognition algorithm starts by pre-processing of the VOC information generated by the VOC analyser, to remove any errors or inconsistencies in the VOC information, that might lead to other issues while constructing data frames for each of the identified odours. Once the data has been pre-processed, predefined data normalization methods will be applied on each data frame of the odours. The methods such as
20 Principal Component Analysis (PCA) and Autoencoder may be used for handling the noise issues in the collected data.

In an embodiment, the data classification may be seen as a predicting modelling in Machine Learning (ML), with which a class label can be predicted based on a given labelled data as input. As an example, the classifiers that have been examined in the present disclosure include, without
25 limiting to, Support Vector Machine (SVM), Linear Discriminant Analysis (LDA), Logistic Regression (LR), Random Forest (RFC) and Multilayer perceptron (MLP). The prediction accuracy of the odours and diagnosis have been validated by cross-validation, accuracy averaging, learning curve, train and test accuracy, train and test precision, train and test recall and train and test.

- 30 It may be noted that, several researchers and companies work on better and faster diagnostic methods for detecting different diseases in domestic and animal farms and for providing better assistance to the veterinarians. With the arrival of artificial intelligence, machine learning, and Internet of things, the necessity of designing intelligent diagnostic devices is more and more important. This allows farmers to easily observe their farms and diagnose sicknesses in the

animals in a short time. Additionally, the device also assists the veterinarians to stay connected with the farmers, whenever it is necessary.

Also, it shall be noted that the accuracy of the AI engine depends on the strength of the available dataset. The proposed AI engine is associated with a massive database that includes a list of
5 various diseases and their symptoms for various breeds of animals including, without limitation, cows, goats, horses, pigs, and chickens. This allows the AI engine to prepare a well distributed and high volume database for better training of the pattern recognition algorithms.

In an embodiment, the AI engine may be coupled with an AI chatbot for systematically assisting the veterinarians whenever a diagnosis is made. One of the objectives of this chatbot is to extract
10 an organized symptoms report by communicating with the farmers. This is an example of communication between the AI engine and the farmer.

FIG. 1B shows an overview of interaction between the AI engine and various sources of information. In an embodiment, the AI engine (i.e., the diagnostic application installed in the computing unit) may receive and study information from various sources related to the animals
15 and/or the farm. For example, the AI engine may receive the real-time VOC information from the VOC analyser. Further, the AI engine may combine the received VOC information with various other information including, without limitation, information from a health journal of the animal, a database storing information related to the animal and various other IoT based devices deployed in the environment of the animal.

20 In an embodiment, the health journal of the animal may include information such as, without limiting to, disease history of the animal, earlier medications/treatments, disease symptoms, feed status, age, gender, breed and other information related to the animal. The database may comprise a plurality of pictures of the animal, including pictures of the diseased region of the animal, as the pictures help in better diagnosis of symptoms and/or diseases of the animals.
25 Further, the IoT devices may include any device that is deployed in the farm and/or an environment around the animal. As an example, the IoT devices from which the AI engine may receive data may include, without limitation, a milking device/robot, a temperature sensor, a humidity sensor and the like. The readings from these IoT devices help in assessing the overall condition of the environment around the animal, and thereby make it possible for the AI engine
30 to produce most accurate predictions on the nature and type of diseases.

In an embodiment, the AI engine may be configured to seamlessly receive data related to the animal from any other source, which may not be explicitly indicated in FIG. 1B.

In an embodiment, the proposed AI engine can detect ten different diseases by predicting their odours with an accuracy of 90%. Moreover, the proposed AI engine has a short turnaround time,

which requires only 60 seconds of smelling for each test sample. This is enabled using a proprietary designed encoder-decoders architecture for extracting more informative features from the GSAs. Secondly, distribution-based (for example LDA) and neural network-based (for example MLP) classification approaches are used to predict a sickness label from the given GSA signals.

In other words, use of the encoder-decoders for learning a compact representation from gas-sensors is another uniqueness of the proposed disclosure. The encoder-decoder contains a list of parameters for the feature extraction from the gas sensors and a set of hyper-parameters for finding the optimal segmentation of the sensors and the optimal subset of the sensors to be used for the training purposes. This helps in finding the optimal time intervals for smelling a sample and predict its odour label.

It shall be also noted that, most of the existed machine learning approaches for similar devices are only applicable on data that is gathered in test laboratory environment, which is not the same as a real environment. Also, in the existing approaches, the gas-sensor data gathered from the stable or farm environment are mixed with other odours, and they are therefore too noisy. Therefore, the proposed method suggests pre-processing and cleaning of the data before further processing. For example, any unsupervised learning approach, such as clustering methods, may be used to remove the noisy training examples as outliers or anomalies.

In an embodiment, an adaptive instance normalization may be designed and implemented to build a denoiser that can regularize the extracted features. Also, transfer learning methods may be used to transfer the learned knowledge from the synthetic-noises to real noise denoiser. Based on the transfer learned model, as the synthetic noise denoiser can learn the general features from various synthetic-noises, the real-noise denoiser can learn the real-noise characteristics from real data as well.

FIG. 2 shows a flowchart illustrating a method of diagnosing diseases in animals using a VOC analyser in accordance with some embodiments of the present disclosure.

In an embodiment, at step 201, the method comprises collecting the odour data related to the test sample using the VOC analyser. As explained in the above sections, the VOC analyser is a smart device comprised of gas sensor arrays, which mimic the human's olfactory system. In an embodiment, the odour data is collected by directing the volatile organic compounds, comprised in the test sample, over the pre-heated gas sensors in the VOC analyser. The absorbed molecules change the conductivity properties of the gas sensors. These changes are then analysed and translated to corresponding VOC concentration levels in the air passing over the gas sensors.

In an embodiment, at step 203, the method comprises pre-processing the collected odour data for removing any errors or inconsistencies in the odour data. Also, the data frames corresponding to the odour data are constructed. Further, the collected odour data may be labelled with the names corresponding to the odour, such that, each odour data frame is labelled.

5 Furthermore, the method involves handling the noises in the odours data, reducing features, and performing signal segmentation for segmenting and presenting the odour data in a quantitative vector form to the AI pattern recognition algorithms. In an embodiment, the segmented signals are 3D matrices, and hence, the autoencoder architecture may be a best-suited model, as indicated in step 205.

10 In an embodiment, in step 207, the predetermined algorithms (i.e., Autoencoder) are applied on the pre-processed and segmented data for feature extraction of the odour data, which converts the existing dataset into a new dataset with more representative features. In an embodiment, the feature extraction may involve dimensionality reduction and/or removing redundant features. Additionally, in step 207, the method comprises classifying the features by grouping and
15 categorizing the data based on their common characteristics and features. As an example, the best prediction may be obtained using classifiers including, without limitation, Multilayer Perceptron (MLP) and Linear Discriminant Analysis (LDA).

Finally, at step 209, the method comprises validating the predictions made in the previous step. In an embodiment, the validation may include, without limiting to, cross validation, accuracy
20 score, balanced accuracy score, precision score, recall score, and learning curve. After successful validation, the possibility of a disease may be notified to the farmers and/or veterinarian for suitably handling the diseases in the animals.

Methods herein for detecting and analyzing volatile organic compounds (VOCs) may use an advanced electronic nose system. The methods may comprise acquiring high-definition data
25 from a custom-configured semiconductor gas sensor array, where each sensor is calibrated to detect specific VOCs based on unique electronic signatures. When analyzing data, a hybrid analytical framework may be used combed with convolutional neural networks (CNNs) and/or recurrent neural networks (RNNs) to enhance pattern recognition and temporal data processing. Refining detection algorithms in real-time may be performed via adaptive learning mechanisms
30 that update detection thresholds and parameters based on environmental feedback and detected scent profile complexities.

Unlike known devices, which may utilize more traditional chemical detection methods, methods herein introduce a combination of CNNs and RNNs for data analysis, emphasizing the use of advanced machine learning techniques for enhanced temporal and spatial resolution in VOC
35 detection. This approach addresses the problem of accurately identifying complex VOC profiles

in dynamic environmental conditions. It significantly enhances the system's ability to adapt to sudden changes in air quality, providing more reliable and precise detections. The integration of CNNs allows the system to effectively parse through spatial data to identify patterns linked to specific chemical signatures, while RNNs manage sequence data to track changes over time, thereby solving the problem of temporal data analysis in fluctuating environments.

The sensor array configuration may be performed using a unique configuration of semiconductor-based gas sensors, emphasizing their customization for specific VOC detection and the role of MEMS technology in enhancing sensor performance. Machine learning framework elaborates on the specifics of the hybrid analytical framework, where CNNs and RNNs may be applied to analyze sensor data. Advanced machine learning techniques may be utilized for real-time data processing. Real-time adaptive learning may be used where the adaptive learning mechanisms dynamically adjust detection algorithms based on real-time environmental feedback, thereby ensuring optimal performance. Data processing and cloud integration in a two-tier processing architecture, have the benefits of local preprocessing for reduced latency and cloud-based analysis for deeper learning capabilities. Environmental compensation algorithm may be used for the algorithm to adjust detection thresholds and calibrate the sensor based on environmental changes, supporting consistent performance across varied conditions. The hybrid analytical framework may utilize CNNs to spatially analyze sensor responses and RNNs to temporally refine the analysis based on previous scent detection events, thereby optimizing the system's predictive accuracy and response time for complex environments. Known devices and methods does not apply a hybrid analytical framework utilizing both CNNs and RNNs, which allows for both spatial and temporal data analysis, offering a significant enhancement in detecting complex VOC patterns over time.

The hybrid approach tackles the challenge of maintaining high accuracy and speed in environments where VOC concentrations and compositions change rapidly. CNNs effectively parse spatial variations across sensor arrays to identify VOC signatures, while RNNs track these signatures over time, adjusting the detection algorithms dynamically to ensure reliability even as environmental conditions change. Methods herein may further comprise local preprocessing of data on an embedded system within the sensor array to reduce latency and improve real-time responsiveness, followed by comprehensive data analysis on a cloud-based platform for deep learning enhancement and storage. The integration of local preprocessing and subsequent cloudbased analysis for deeper learning is not known. This bifurcated approach allows for rapid initial analysis and sophisticated secondary analysis, which is not achievable with traditional systems. This addresses the dual needs of immediate, on-site analysis and more complex, in-depth data processing that can adapt over time with new data inputs.

Local preprocessing allows for quick responses crucial in many industrial and safety-critical applications, while cloud-based analysis provides the computational power needed for ongoing improvement and learning from accumulated data. In methods herein operational parameters of the gas sensor array, including sampling rates and sensor activation sequences, may be

5 dynamically adjusted using a real-time decision engine that processes environmental sensors' input to maintain optimal sensitivity and specificity under varying conditions. Unlike known devices, a real-time decision engine is used that dynamically adjusts operational parameters based on environmental inputs, which is a novel approach for enhancing sensor array performance continuously. This solves the issue of sensor calibration and performance

10 degradation in fluctuating environmental conditions. The real-time decision engine may analyze environmental data to adjust the sensors' sampling rates and activation sequences, ensuring that the system remains highly accurate and effective regardless of external changes. Methods herein may incorporate an environmental compensation algorithm that adjusts VOC detection thresholds and sensor calibration based on detected changes in temperature, humidity, and air

15 quality, ensuring consistent system performance across different climates. This uniquely focuses on an environmental compensation algorithm that automatically recalibrates the system based on live environmental data, a feature not known. This ensures that the electronic nose system can operate with high accuracy across a wide range of environmental conditions without manual recalibration. By automatically adjusting detection thresholds and calibration, the system

20 maintains its sensitivity and specificity even when environmental factors such as temperature and humidity vary, which is critical for applications in diverse geographical locations.

Throughout this specification, plural instances may implement components, operations, or structures described as a single instance. Although individual operations of one or more methods are illustrated and described as separate operations, one or more of the individual

25 operations may be performed concurrently, and nothing requires that the operations be performed in the order illustrated. Structures and functionality presented as separate components in example configurations may be implemented as a combined structure or component.

The terms "an embodiment", "embodiment", "embodiments", "the embodiment", "the

30 embodiments", "one or more embodiments", "some embodiments", and "one embodiment" mean "one or more (but not all) embodiments of the invention(s)" unless expressly specified otherwise.

Similarly, structures and functionality presented as a single component may be implemented as separate components. These and other variations, modifications, additions, and improvements

35 fall within the scope of the subject matter herein. In addition, the term "each" used in the

specification does not imply that every or all elements in a group need to fit the description associated with the term “each”. For example, “each member is associated with element A” does not imply that all members are associated with an element A. Instead, the term “each” only implies that a member (of some of the members), in a singular form, is associated with an
5 element A.

Finally, the language used in the specification has been principally selected for readability and instructional purposes, and it may not have been selected to delineate or circumscribe the patent rights. It is therefore intended that the scope of the patent rights be limited not by this detailed description, but rather by any claims that are issued on an application based hereon.
10 Accordingly, the disclosure of the embodiments is intended to be illustrative, but not limiting, of the scope of the patent rights.

Claims

1. A method for analyzing a gas or a gas mixture, the method comprising:
 - acquiring an initial data set from a gas sensor array exposed to said gas or gas mixture;
 - performing a coarse analysis of the acquired initial data set;
 - 5 determining, based on the coarse analysis, a superset of associated scent profiles
 - determining a set of refinement parameters based on the superset of associated scent profiles;
 - adjusting a set of operational parameters for said sensor array based on the set of refinement parameters; and thereafter
 - 10 acquiring an expanded data set from said gas sensor array;
 - performing a refined analysis on the expanded data set; and
 - determining, based on the refined analysis, the presence of at least one target scent profile in the gas mixture.
- 15 2. The method according to claim 1, where the coarse analysis employ signal normalization.
3. The method according to claim 1 or 2, where the coarse analysis employ signal smoothing.
4. The method according to any of the preceding claims, where the coarse analysis employ noise
- 20 removal filtering.
5. The method according to any of the preceding claims, where the coarse analysis employ draft reduction.
- 25 6. The method according to any of the preceding claims, further comprising:
 - detecting volatile compounds, VOCs, in said gas or gas mixture.
7. The method according to any of the preceding claims, wherein the coarse analysis is performed using a machine learning model.

8. The method according to the preceding claim, wherein at least parts of the result from the refined analysis is used to improve the machine learning model used for coarse analysis.
- 5 9. The method according to any of the preceding claims, wherein the coarse analysis is performed locally within a scent detection system.
10. The method according to any of the preceding claims wherein the refined analysis is performed externally on cloud-based servers.
- 10 11. The method according to any of the previous claims, wherein the operational parameters comprise sampling rate.
12. The method according to any of the previous claims, wherein the operational parameters comprise duration of sensing.
- 15 13. The method according to any of the previous claims, wherein the operational parameters comprise heating profile.
- 20 14. The method according to any of the previous claims, wherein the operational parameters comprise energy feed to selected sensors.
15. The method according to any of the previous claims, the method further comprising:
wherein said operational parameters comprise at least one of: selection of sensors,
- 25 16. The method according of any the previous claims, wherein said operational parameters comprise heating profile.
17. The method according of any the previous claims, wherein the sensor array is fine-tuned based on the refinement parameters.
- 30

18. The method according of any the previous claims, wherein the refined analysis includes Kalman.filtering.

19. The method according of any the previous claims, wherein the refined analysis includes
5 Wiener-filtering.

10

15