Decoding the Inverse Problem of Gas Dispersion with Physics-aware Deep Learning Approach

TIAN CHANGHAO

SCHOOL OF MATERIALS SCIENCE AND ENGINEERING

2025

Decoding the Inverse Problem of Gas Dispersion with Physics-aware Deep Learning Approach

TIAN CHANGHAO

SCHOOL OF MATERIALS SCIENCE AND ENGINEERING

A thesis submitted to the Nanyang Technological University
in partial fulfilment of the requirement for the degree of
Doctor of Philosophy

2025

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research, is free of plagiarised materials, and has not been submitted for a higher degree to any other University or Institution.

04/02/2026

. . . .

Date

Changhao Tian

. . . .

TIAN CHANGHAO

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it is free of plagiarism and of sufficient grammatical clarity to be examined. To the best of my knowledge, the research and writing are those of the candidate except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

04/02/2026

. . . .
Date

. . . .
Prof. CHEN XIAODONG

Authorship Attribution Statement

This thesis contains material from 1 paper(s) published in the following peer-reviewed journal(s) / from papers accepted at conferences in which I am listed as an author.

Chapter 4 is published as Changhao Tian, Annan Wang, Han Fan, Thomas Wiedemann, Yifei Luo, Le Yang, Weisi Lin, Achim J. Lilienthal, Xiaodong Chen. Deep Learning Based Topography Aware Gas Source Localization with Mobile Robot. International Conference on Robotics and Automation International Conference on Robotics and Automation (ICRA) 2025.

The contributions of the co-authors are as follows:

- Prof. Xiaodong Chen provided the initial project direction and edited the manuscript drafts.
- I prepared the manuscript drafts. The manuscript was revised by Dr Han Fan, Dr. Thomas Wiedemann and Dr. Yifei Luo.
- I co-designed the study with Annan Wang and performed all the laboratory work at School of Materials Science, Nanyang Technological University, Singapore and Chair: Perception for Intelligent Systems, Munich Institute of Robotics and Machine Intelligence (MIRMI) at Technical University of Munich, Germany. I also analyzed the data.
- Prof. Achim J. Lilienthal and Prof. Weisi Lin provided guidance and assistance during the project and also edited the manuscript.

04/02/2026

. . . .

Date

Changhao Tian

. . . .

TIAN CHANGHAO

Abstract

Gas sensing plays a critical role in ensuring workplace safety and accurately assessing environmental impacts. As gases are often colorless, exist at ambient temperatures, and do not produce readily detectable physical signals, remote sensing of gases at small to medium scales remains challenging. Current gas monitoring systems rely on fixed, passive sensor networks, which are inadequate for dynamic scenarios such as emergency response or environmental exploration. Robotic olfaction aims to enable deployable and adaptive gas monitoring solutions capable of operating in unknown and dynamic environments. Within this domain, robotic spatial olfaction seeks to reconstruct the spatial distribution of gas-related features—such as source location and concentration—using limited, sparse observations, representing a process of high-dimensional inference from low-dimensional measurements.

This thesis addresses two complementary tasks in robotic spatial olfaction that correspond to different stages of gas dispersion. In the early stage, gas plumes are typically turbulent and intermittent, and concentration gradients are weak or unstable. To handle this, a gas source localization (GSL) method is developed that integrates spatial context from SLAM-derived maps and wind vector information to infer likely source locations, even in obstacle-rich environments where wind-topography interactions introduce significant uncertainty. In the later, more stable stage of dispersion, where gas fields tend to smooth out under diffusion and convection, the focus shifts to gas distribution mapping (GDM). A Transformer-based model is proposed to reconstruct near-ground gas and wind fields from sparse and irregular sensor observations. Leveraging attention mechanisms and positional encoding, the model captures long-range spatial dependencies and adapts to unstructured input patterns, outperforming traditional interpolation and CNN-based approaches under sparse sampling conditions.

Together, these two tasks form a unified framework for robotic spatial olfaction, covering both dynamic and steady-state dispersion regimes. This thesis contributes to the field by integrating environmental context, wind information, and sensor observations into learning-based models that are validated in both simulation and real-world experiments. By addressing the distinct challenges of each dispersion stage, the proposed approaches enhance the feasibility and robustness of mobile gas perception systems in real-world, dynamic environments.

Lay Summary

Gases, while often hazardous, are typically invisible, odorless, and difficult to detect, particularly in open, dynamic, or unstructured environments such as industrial facilities, disaster sites, or polluted areas. Conventional gas sensing systems rely on stationary sensor networks, which are limited in their ability to respond to changing conditions or to cover large, complex areas. This thesis investigates how mobile robots can be equipped to serve as active gas sensing platforms, capable of detecting and mapping harmful gases in real time, even with limited and sparse observations.

This thesis explores how mobile robots can estimate gases using sensors and smart algorithms, enabling them to detect, locate, and map gas distributions in real time. Specifically, the thesis focuses on two related challenges that arise at different stages of gas dispersion. When a gas first leaks or is released, it forms a chaotic and unpredictable plume that moves with the wind. In this early stage, robots need to estimate where the gas is coming from, a task called gas source localization. Later on, as the gas spreads out and stabilizes, it becomes important to estimate how the gas is distributed across a larger area, which is known as gas distribution mapping.

To address the first challenge, this work develops a method that allows robots to combine information from gas sensors, wind measurements, and a map of the environment to estimate the most likely location of the gas source, even when the wind is irregular or the space is cluttered with obstacles. For the second challenge, a more advanced model is designed using a type of neural network known as a Transformer, which can learn how gas spreads by analyzing sparse and unevenly collected sensor data. This helps the robot build a complete picture of the gas field, even when it only has a few scattered measurements.

Together, these tools allow mobile robots to make sense of complex gas environments with limited data, making them more effective in real-world applications. The results have been validated through both computer simulations and physical experiments, bringing us closer to practical, deployable gas-sensing robots that can operate in factories, disaster sites, or urban environments where fast, reliable gas monitoring is crucial.

Acknowledgements

I would like to express my deepest gratitude to my esteemed supervisors, Prof. Xiaodong Chen and Dr. Le Yang, as well as my cooperation supervisor Achim J Lilienthal in Technical University of Munich. for their invaluable guidance, unwavering support, and scholarly expertise throughout the course of my research. Their commitment to excellence, patience, and willingness to invest their time and knowledge in shaping my academic journey has been instrumental in my growth as a researcher.

Furthermore, I extend my heartfelt appreciation to the Agency for Science, Technology and Research (A*STAR) for their generous financial support throughout my doctoral studies, which has not only enabled me to pursue my research but has also opened doors to numerous academic and professional opportunities. I am grateful for their belief in my potential and their commitment to fostering research excellence.

I am also immensely grateful to the mentors who have provided me with guidance and assistance in various stages of my research. Their knowledge and expertise have been instrumental in shaping my research methodology, refining my analytical skills, and expanding my understanding of the field. I would like to specifically acknowledge Dr. Jianwu Wang, Dr. Han Fan, and Mr. Annan Wang for their contributions.

I would also like to thank my coworkers and colleagues for their valuable discussions, collaboration, and support. Their diverse perspectives and intellectual contributions have enriched my research journey, and I am truly fortunate to have had the opportunity to work alongside such talented individuals.

Lastly, I want to express my deep appreciation to my family and friends for their unwavering support, encouragement, and understanding throughout this demanding endeavor.

Table of Contents

Table Captions

Figure Captions

Figure 2.1 The MOX sensors' mechanism of gas sensing. (A). Gas reacts with the surface oxygen layer, inducing change of carrier density, reproduced with permission from 29. (B). Doping of factualization of MOX (e.g. with doping of carbon nanotube) can enhance sensors' sensitivity, reproduced with permission from 26. MOX sensing is fundamentally governed by surface chemistry, and sensor performance can be substantially tuned by material engineering. Both of which are critical considerations for robust sensor response and related applications.

Figure 2.2 Electrochemical-based gas sensor. (A). Liquid electrolyte gas sensor, reproduced with permission from 30. (B). Solid-state electrolyte gas sensor, reproduced with permission from 15. Electrochemical-based gas sensor translates gas concentration into an electrochemical signal via electrode reactions, but differ in electrolyte phase and transport pathway.

Figure 2.3 Cross sensitivity of MOX sensors comes from sensors' limited selectivity. It limits sensor's recognition capability in complex chemical backgrounds.

Figure 2.4 Surface-functionalization based sensor. (A). Receptor based gas sensor, reproduced with permission from 38. (B). Peptide based gas sensor, reproduced with permission from 39, The authors. (C). Live cell-based gas sensor, reproduced with permission from 40. (D). Molecular imprinting polymer-based gas sensor, reproduced with permission from 41. These approaches enhance sensitivity and selectivity by introducing delicate bio/chemically specific recognition elements at the sensing interface, illustrating a spectrum of trade-offs between molecular specificity, stability, and system complexity for real-world gas sensing applications.

Figure 2.5 Platforms of robotics gas perception. (A) Ground robotics, reproduced with permission from 68(B) Multirotor UAV, which induces significant downwash flow.

Reproduced with permission from 66 (C) Fixed-wing UAV, carrying a laser-based sensor, reproduced with permission from 67.

Figure 2.6 Statistical description of Gaussian plume model. Due to the existence of turbulence, the instantaneous concentration spatial and temporal distribution could be highly fluctuating, reproduced with permission from 76.

Figure 2.7 Reactive gas source localization approaches. Although with different strategies, these approaches rely on driving the robot towards potential gas sources, which is limited by robots' accessibility, reproduced with permission from8.

Figure 2.8 Statistical approach of gas source localization, reproduced with permission from 83. It estimates potential gas source location by Bayesian estimation with gas transmission model.

Figure 2.9 Gas source localization based on Deep Q Network (DQN), reproduced with permission from 68. DQN is a reinforcement learning algorithm that combines Q-learning with deep neural networks to learn optimal actions from high-dimensional input states.

Figure 2.10 Rule based gas distribution mapping, reproduced with permission from 101. This approach requires the robot to uniformly traverse the area and perform interpolation. The mapping quality is strongly coupled to coverage uniformity and sampling density, which can be difficult to guarantee in practical deployments with limited time, obstacles, or uneven exploration.

Figure 2.11 Input, prediction and error of different gas distribution mapping approach, reproduced with permission from 98. Gas Distribution Decoder (GDD) is a CNN based prediction model, which performs better accuracy, especially at edge fine structure compared to traditional method.

Figure 3.1 Working principle of PID sensors, reproduced with permission from 110. Gas

Figure 5.4 Performance of different approaches under different sampling balance. The Transformer-based approach shows significantly higher accuracy in both wind field and gas concentration distribution.

Figure 5.5 Evaluation with real-world data. (A) Models' performance on data from wind tunnel experiments. (B) Comparison of transformer's output of gas concentration with the real plume region.

Abbreviations

GSL	Gas Source Localization
GDM	Gas Distribution Mapping
MOX	Metal Oxide
PID	Photoionization Detector
SLAM	Simultaneous Localization and Mapping
TDLAS	Tunable Diode Laser Absorption Spectroscopy
FET	Field Effect Transistors
QCM	Quartz Crystal Microbalance
SPR	Surface Plasmon Resonance
SAW	Surface Acoustic Wave
DQN	Deep Q Network
LSTM	Long Short-Term Memory
CNN	Convolutional Neural Network
MFC	Mass Flow Controller
GPR	Gaussian Process Regression
CFD	Computational Fluid Dynamics
GAN	Generative Adversarial Network
PINN	Physics-Informed Neural Network
GDD	Gas Distribution Decoder
PI	Polyimide
PCB	Printed Circuit Board
MCU	Microcontroller Unit
ADC	Analog-to-Digital Converter
VOC	Volatile Organic Compound
LIDAR	Light Detection and Ranging
BCE	Binary Cross Entropy
RANS	Reynolds-Averaged Navier–Stokes Equations
UV	Ultraviolet

ppb	Parts Per Billion
ppm	Parts Per Million
ROC	Receiver Operating Characteristic
Kernel DM+V	Kernel Extrapolation Distribution Mapping + Variance
MSE	Mean Squared Error
SSIM	Structural Similarity Index Measure

Chapter 1

Introduction

Gaseous substances, while often invisible and hard to detect, are critical factors in various real-world contexts such as environmental monitoring, industrial safety, and disaster response. Conventional gas sensing systems fall short in dynamic or unknown environments. To address this limitation, the field of robotic spatial olfaction offers an active sensing paradigm, enabling mobile robots to detect, localize, and map gases in complex settings. This chapter introduces the two central tasks of robotic olfaction: gas source localization (GSL) in turbulent, obstacle-rich environments, and gas distribution mapping (GDM) in more stabilized dispersion scenarios, with challenges arising from the sparse, noisy, and spatially uneven nature of sensor data, as well as the complex fluid dynamics involved in gas propagation.

To overcome these challenges, this thesis proposes two novel approaches. First, for GSL, a SLAM (Simultaneous Localization and Mapping)-informed strategy is developed that leverages environmental layout and wind flow observations to better model the non-linear interactions between gas plumes and obstacles, enabling more accurate source inference. Second, for GDM, a Transformer-based model is introduced that uses positional encoding and attention mechanisms to reconstruct wind and gas concentration fields from sparse and biased measurements. These methods are supported by both simulation and real-world experiments, including the development of a mobile robot *platform, dataset collection under varying environmental conditions, and comparative model evaluations.*

1.1 Problem Statement

Among different industries, including environmental monitoring[1], industrial safety[2], and disaster rescue[3], gas is always an important but typically invisible environmental factor. Detecting, locating and mapping of gas in the working area is crucial for ensuring environmental compliance, preventing industrial accidents, and supporting rescue operations. Current gas monitoring systems primarily rely on fixed sensor networks[4–7], which can only perform passive monitoring and are inadequate for handling complex or unknown environments. To address this challenge, robotic spatial olfaction[8] aims to provide active and deployable gas sensing solutions capable of operating in complex or unknown environments, such as for inspection tasks or emergency response.

Depending on the specific scenario, robotic spatial olfaction for industry typically involves two primary tasks: gas source localization[9] (GSL) and gas distribution mapping[10] (GDM). These tasks reflect the dual nature of gas dispersion in real environments. When a gas source initially appears (e.g., a leakage occurs), the gas plume is dominated by turbulence, with high temporal and spatial fluctuations. At this early stage, a stable concentration gradient has yet to form, making direct inference of the source location from sparse measurements highly non-trivial. The GSL focuses on locating the source of a gas leak in this kind of unstable environment.

In contrast, at later stages of dispersion, particularly in enclosed or ventilated environments, the gas field tends to stabilize into a smoother concentration distribution governed by diffusion and convection. GDM addresses this regime by reconstructing the high-dimensional spatial distribution of gas concentrations across a region, often from a limited set of sensor readings. This task exemplifies the core challenge of robotic olfaction: inferring a complex global field from sparse, local measurements. Both GSL and GDM highlight the inverse nature of robotic olfactory inference, where perception is fundamentally about bridging the gap between limited, low-dimensional observations and the full, high-dimensional environmental state.

However, robotics spatial olfaction is always a challenging task, involving multiple sensing modals. First, as a form of matter dispersion, the gas propagation process involves complex interactions with the surrounding environment, including wind-induced convection, plume formation caused by turbulence at interfaces[11], and wind field distortion resulting from topography-induced effects. As a result, it requires not only perception of gas itself but also perception of physical environment. Besides, the chemical background encountered by robots is always highly complex. Industrial emissions, vehicle exhaust, and household waste can all release various types of gases. Robots need to be capable of selectively detecting harmful target gases within these complex chemical backgrounds. These cross-modal characteristics introduce significant complexity to robotic olfaction.

The primary challenge of gas source localization lies in the interaction between the gas plume and the surrounding topography. A robot for gas source localization typically uses a gas sensor to detect the gas from the source and an anemometer to estimate the direction of the source[12,13]. However, in real-world scenarios, observed gas concentrations and airflow vectors do not directly correlate with the source location due to complex wind-obstacle interactions. Early in gas dispersion, a continuous concentration gradient has not yet formed, and turbulence creates plumes with fluctuating concentrations. This means observed gas concentrations only indicate the presence of gas, not the distance to the source. Additionally, obstacles distort wind streamlines, making it difficult to infer source direction solely from an anemometer. As a result, gas source localization in complex environments with obstacles is challenging due to the limitations of sensory information.

As for gas distribution mapping, it presents three major challenges. First, the number and locations of sampling points are unfixed, making rigid network architectures such as multilayer perceptron unsuitable for this task. Second, the distribution of gas concentration within the region is highly sparse, with concentrations at locations outside the gas plume being nearly zero. This sparsity imposes higher demands on the model’s feature extraction capability, requiring it to effectively identify meaningful patterns from predominantly low-value or zero-concentration areas. Finally, due to the robot’s mobility, the sequence of inputs' sampling locations is not fixed. This necessitates a model capable of encoding the relative positional information of inputs to capture spatial correlations between samples, which is crucial for accurately reconstructing the fluid field.

In this thesis, two main problems are proposed and demonstrated successively:

For early stage GSL in complex obstacle-rich environment, to tackle the challenges posed by environmental obstacles, this thesis explores the use of spatial context, such as environmental layout and wind flow observations to model the complex and non-linear relationship between gas detections and the true source location. Information about the surrounding environment can be captured through Simultaneous Localization and Mapping (SLAM), a technique that enables robots to simultaneously build or update a map of an unfamiliar environment while localizing themselves within it. In the context of gas source localization, SLAM provides precise positions of gas encounters along with the corresponding environmental structures. Given that obstacle configurations can significantly influence gas dispersion dynamics, incorporating SLAM-derived spatial knowledge is expected to substantially improve the accuracy of source localization.

For stable stage GDM, this thesis proposes a Transformer-based approach for GDM from sparse and biased sensor inputs. The attention mechanism enables the model to capture long-range spatial correlations, while positional encoding allows the Transformer to effectively interpret the relative locations of sampling points, making Transformer a powerful model for fluid reconstruction[14–16]. Because of gravity and surface friction, near-ground airflow is predominantly horizontal. Therefore, GDM can be simplified as a mapping task within a two-dimensional plane, which typically corresponds to the robot's operational workspace. After applying positional encoding to the coordinates and feature encoding to the sensor readings, the inputs are fused and processed through an encoder-decoder architecture. The Transformer-based model could be able to predict the distribution of wind and gas concentration from sparse observation.

1.2 Objectives and Scope

In order to solve the problems stated above, the following objectives are to be accomplished by this thesis:

1. Design a novel gas source localization approach. By leveraging the spatial relationship between the gas plume encounter locations and the environmental map, this approach will be able to overcome the challenges of gas source localization in complex topography.

2. Design a Transformer-based gas distribution mapping approach. Reconstruct area gas and wind distribution from sparse and biased sensor sampling by spatial encoding and self-attention.

To achieve these objectives, the following experiments and analyses were conducted

1. Design and develop a robot platform with gas sensor, anemograph and LIDAR, to evaluate robotics olfaction strategy.

2. Collecting GSL dataset under varying topography and wind speed conditions, including environmental maps, wind direction, and gas encounter locations.

3. Design and train a U-net based GSL model in the dataset and evaluate it with new experiments.

4. Generation of GDM datasets with computational fluid dynamic method, including 2D gas dispersion with different wind direction, gas source location and intensity.

5. Design and train a Transformer-based GDM model based on the dataset, evaluate the performance with current baselines.

6. Evaluate the robustness and generalization ability of GDM model based on real-world

experiments.

1.3 Dissertation Overview

The thesis consists of the following 6 chapters:

Chapter 1 provides a rationale for the research and outlines the goals and scope. The thesis describes the development of robotic spatial olfaction methods aimed at addressing the challenges of gas source localization (GSL) and gas distribution mapping (GDM) in complex environments.

Chapter 2 provides a comprehensive overview of the current state of research in robotic spatial olfaction, including various gas sensing technologies along with their respective advantages and limitations, existing approaches for gas source localization and gas distribution mapping, as well as the key challenges that remain to be addressed.

Chapter 3 discusses the methodology of this thesis. It also covers the design of the sensor and robotic platform, the methodology for dataset generation, and the design and training procedures of the proposed models.

Chapter 4 elaborates the development of gas source localization. It covers the design and control of the robotic platform, the experimental setup for gas source localization, and the development and training of a U-Net-based localization model. This chapter demonstrates the effectiveness of the proposed approach in complex environments, including terrains with multiple obstacles and dynamic wind fields.

Chapter 5 elaborates the development of gas distribution mapping. It covers the generation of gas distribution dataset by CFD, and the training of Transformer-based model based on it. This chapter demonstrates the effectiveness of the proposed approach in reconstructing gas fields with sparse and unbalanced sensor sampling in real-world scenario.

Chapter 6 reviews the conclusions and implications of this thesis. The problems stated at the beginning of the thesis have been addressed. The implications of this work lay in extending the applicability of robotic spatial olfaction to complex and dynamic environments, including those with intricate chemical backgrounds and obstacle-rich spaces. This chapter also discusses the necessity of integrating additional sensing modalities in future work to further enhance system robustness and environmental adaptability.

1.4 Findings and Outcomes/Originality

The originality of this thesis lies in a core conceptual contribution: robotic gas sensing is framed as context-aware inverse inference from sparse, unstructured, and trajectory-dependent observations, where physical context is integrated as a structured prior to improve data efficiency and robustness under occlusions. Within this framing, gas source localization (GSL) is treated as inferring a source location without requiring contact, while gas distribution mapping (GDM) is posed as a set-to-field reconstruction problem that maps irregular samples to dense spatial fields. Building on this conceptual foundation, the thesis reports the following outcomes:

1. **Context-aware GSL under obstacle–wind interactions**. A topography-aware localization approach is developed in which environmental geometry is incorporated as an explicit observation modality, together with wind cues and plume encounter events. This design enables robust source inference in obstacle-rich environments where wind-obstacle interactions create complex, non-intuitive transport patterns, and it supports distance-based source prediction without continuous navigation-to-source tracing.

2. **Sparse-to-field GDM using attention-based modeling of unstructured observations**. A Transformer-based GDM framework is introduced to reconstruct gas concentration and wind fields from sparse and irregular sensor observations. By representing measurements as an unordered set of location–value tokens and using positional

encoding to capture sampling geometry, the attention mechanism models long-range spatial dependencies that are difficult to recover through local interpolation or grid-based CNN pipelines under sampling bias. Extensive evaluations on CFD-simulated datasets and real wind-tunnel measurements demonstrate robustness and practical utility in complex, partially observed sensing conditions.

Chapter 2

Literature Review

This chapter reviews existing gas sensing technologies and robotic olfaction strategies, highlighting key limitations in real-world applications. While electronic and bio-inspired sensors offer varying trade-offs between selectivity and robustness, most are unsuitable for complex, dynamic environments. Robotic platforms, ranging from ground vehicles to UAVs, face challenges in gas sampling fidelity due to terrain constraints or airflow disturbances.

For gas source localization (GSL), reactive and statistical methods struggle with turbulent dispersion and environmental occlusion, while emerging learning-based models show promise but lack spatial awareness. In gas distribution mapping (GDM), traditional interpolation and deep learning models often assume dense, regular sampling, limiting their generalization. These gaps underscore the need for novel GSL and GDM frameworks that can incorporate environmental context and handle sparse, irregular sensor data, goals that this thesis aims to address.

2.1 Current Gas Sensing Approaches

The challenge in current research of gas sensing stems from the complexity of chemical molecular information. Visual sensing fundamentally involves the detection of electromagnetic waves, while acoustic sensors merely need to detect vibrations in the air. In contrast to these two, gas sensing does not have a single, definitive physical quantity to target for detection. The olfactory system of mammals is based on the shape matching between olfactory receptors and gas molecules[17]. This recognition mechanism, combined with the hundreds of distinct olfactory receptors in mammals, ensures that biological olfactory systems maintain good specificity while also being able to produce consistent responses to previously unencountered gases. However, the fine-tuning of material structures, akin to olfactory receptors, significantly surpasses the current technical capabilities of material science. Currently, gas sensors for e-nose applications primarily operate based on redox reactions between gases and sensing materials (metal oxide sensors[18], electrochemical sensors[19]), or the selective adsorption of gas molecules on the sensor surface (Field Effect Transistor sensors[20]). The amount of information generated by these detection mechanisms is significantly less than that produced by biological olfactory systems. Due to the differences in sensing mechanisms, the information obtained from these detection methods cannot align with biological olfaction, making it difficult to accurately assess the characteristics of odors. Also, some studies attempt to create gas sensors based on olfactory receptors[21], engineered cells[22], or living organs[23]. Despite their superior selectivity, these biologically based sensors are costly, challenging to maintain, difficult to produce outside of a laboratory environment, and nearly impossible to fine-tune for customization, limiting their practical value for applications beyond the lab.

2.1.1 Electronic Based Sensors

In the electronic approach, sensor signals are generated on the surface of sensing electrodes. The types of sensors suitable for universal e-noses are relatively limited, typically utilizing redox reactions to achieve selectivity for gases. Based on the different initiation methods of redox reactions, sensors can be further divided into two categories: metal oxide gas sensors and electrochemical gas sensors.

Metal oxide (MOX) gas sensors are devices that detect gases through changes in the electrical conductivity of a metal oxide film (usually heated). When the sensor's surface, typically composed of transition metal oxide[24] like tin dioxide, titanium oxide or zinc oxide, is exposed to certain gases, the interaction between the target gas molecules and the adsorbed oxygen on the sensor's surface alters the concentration of charge carriers within the metal oxide. This change in carrier concentration affects the sensor's electrical resistance, which is measured and correlated to the gas concentration (Figure 2.1A). These sensors are widely appreciated for their sensitivity and relative low cost. Due to the varying properties of MOX sensors made from different materials, a common approach is to assemble an array of sensors made of different MOX materials. Since the sensors in the array exhibit differentiated responses to targets, functions such as gas identification[25] and health monitoring[26] can be achieved through machine learning or statistical methods.

The major limitation of MOX sensors is their requirement for high-temperature operation. Metal oxides typically exhibit poor gas sensing performance at room temperature and therefore need to be heated to an operating temperature, normally between 250°C and 500°C. In addition to increased energy consumption, this also results in a long initialization period after activation, often exceeding 12 hours, as the sensor requires time to reach thermal equilibrium and establish a stable depletion layer at the target operating temperature. Various approaches have been explored to reduce the operating temperature of metal oxide sensors[27,28], including metal doping[29], organic semiconductor functionalization[30,31] (Figure 2.1B), and the synergistic use of multiple metal oxides[32]. However, reports on room-temperature MOX sensors remain relatively scarce.

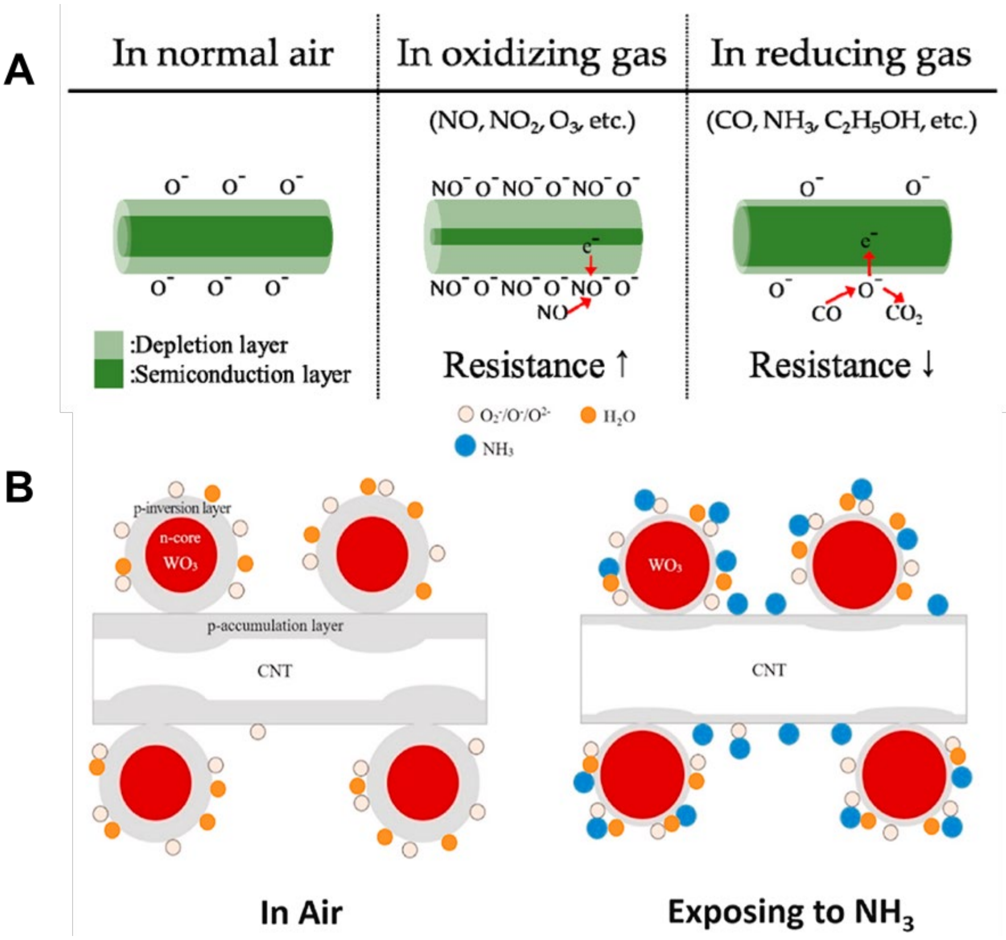

Figure 2.1 The MOX sensors' mechanism of gas sensing[30,33]. (A). Gas reacts with the surface oxygen layer, inducing change of carrier density, reproduced with permission from [33]. (B). Doping of factualization of MOX (e.g. with doping of carbon nanotube) can enhance sensors' sensitivity, reproduced with permission from [30]. MOX sensing is fundamentally governed by surface chemistry, and sensor performance can be substantially tuned by material engineering. Both of which are critical considerations for robust sensor response and related applications.

Electrochemical gas sensors operate based on the principle of redox reactions occurring at the electrode-electrolyte interface when gas molecules are present[34–38]. Because electrochemical sensors can detect specific redox potentials and generate characteristic cyclic voltammetry curves, they offer better selectivity compared to MOX sensors. These sensors can be broadly classified into two types based on the electrolyte medium: liquid electrolyte sensors[34] (Figure 2.2A) and solid-state electrolyte sensors[19] (Figure 2.2B). Liquid electrolyte sensors typically consist of a working electrode, a counter electrode, and a reference electrode immersed in an aqueous or gel-like ionic solution. When gas diffuses through a gas-permeable membrane into the electrolyte, it undergoes electrochemical reactions that generate a measurable current proportional to its concentration. Solid-state electrolyte sensors, on the other hand, replace the liquid medium with a solid ion-conducting material, such as a proton-conductive polymer or ceramic.

Measuring relative to the reference electrode, electrochemical gas sensors can provide a quantitative analysis of the gas presence. The advantage of electrochemical gas sensors is that their output contains more information from different testing methods and the corresponding current-voltage features. Voltammetry is the most commonly used electrochemical sensing method, where a potential is applied to an electrode and the resulting current is recorded. During the measurement, the potential is swept across a range, either linearly in linear sweep voltammetry or cyclically in cyclic voltammetry, while the current response of the electroactive species in solution is monitored. The resulting plot of current versus potential provides valuable information about the electrochemical processes occurring on the electrode. This type of data exhibits greater tolerance to noise, and through machine learning algorithms such as Support Vector Machines (SVMs)[39] and Random Forests[40], both qualitative and quantitative detection of gases can be achieved. Electrochemical gas sensors' selectivity can also be achieved through the choice of electrode materials and the potential applied across the electrodes. The electrode material is chosen for its ability to catalyze the redox reaction of the specific target gas while minimizing reactions with other substances[41]. By controlling the potential, it is possible to set an electrochemical window within which only the desired gas is oxidized or reduced, while other gases remain electrochemically inactive.

Despite their advantages, electrochemical sensors also face several limitations. In liquid electrolyte sensors, the electrolyte is affected by contamination or gradual depletion over time, which can lead to signal instability and reduced sensor lifespan. The need for careful sealing and maintenance also limits their robustness in long-term or field applications. Solid-state electrolyte sensors, while more durable, often suffer from low ion diffusion coefficients. This results in weak signal currents and high noise levels, reducing the signal-to-noise ratio and making it difficult to detect low concentrations of gas with high accuracy.

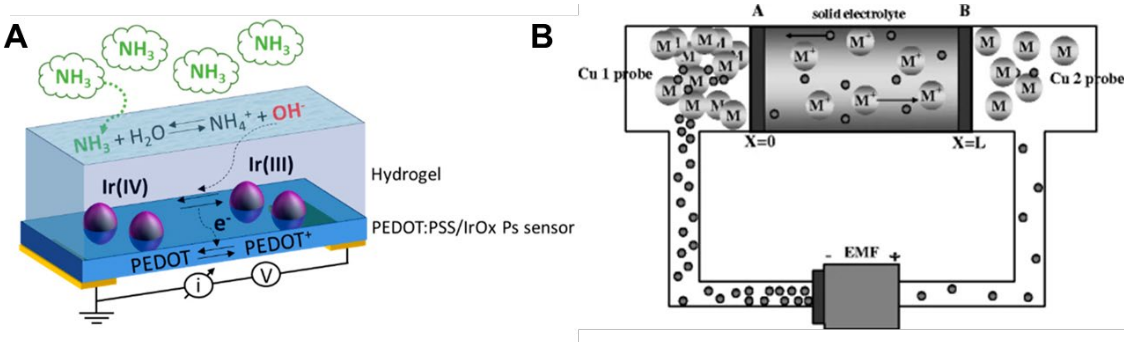

Figure 2.2 Electrochemical-based gas sensor. (A). Liquid electrolyte gas sensor, reproduced with permission from [34]. (B). Solid-state electrolyte gas sensor, reproduced with permission from [19]. Electrochemical-based gas sensor translates gas concentration into an electrochemical signal via electrode reactions, but differ in electrolyte phase and transport pathways.

Besides, all gas sensors based on redox reactions inevitably encounter issues of cross-sensitivity. Cross-sensitivity refers to a sensor's similar responses to different gases, where the sensor cannot distinctly identify individual gases due to overlapping selectivity (Figure 2.3). Actually, the redox capacity is a general indicator of chemical reactivity rather than a specific marker for the molecular shape, which is more directly related to the perception of odors. Since the shape of a molecule determines how it fits into olfactory receptors, creating artificial sensors that can discern this shape is a complex challenge that goes beyond simple redox reactions. This disconnect means that while redox-based sensors can detect the presence of certain gases, they are less adept at differentiating complex odor profiles that would be readily distinguishable by biological olfactory systems.

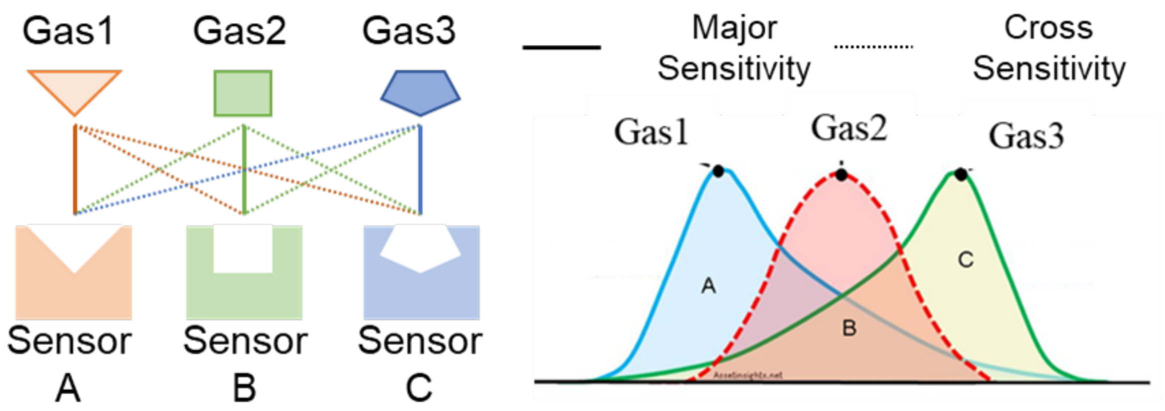

Figure 2.3 Cross sensitivity of MOX sensors comes from sensors' limited selectivity. It limits sensor's recognition capability in complex chemical backgrounds.

2.1.3 Surface-functionalization Based Sensors

Surface-functionalization based sensors represent a class of highly specific sensing technologies that rely on two main components: a functionalized recognition layer and a signal transduction layer. The surface of the sensor is modified with functional materials designed to selectively bind target molecules, such as proteins[42] (Figure 2.4A), peptides[43] (Figure 2.4B), living cells[44] (Figure 2.4C), or molecularly imprinted polymers[45] (Figure 2.4D). Once the target molecules bind to the functionalized surface, the state of surface is changed, which is then detected and converted into a measurable signal by the underlying transduction mechanism. Common signal transduction methods include field-effect transistors[46] (FET), quartz crystal microbalance[47,48] (QCM), surface plasmon resonance[49] (SPR), and surface acoustic wave[50] (SAW) devices. These sensors offer high selectivity and sensitivity, hardly affected by other components.

To bridge the gap between the readings from e-noses and the perception of biological olfaction, a clever approach is to directly utilize biomolecules related to olfaction as the sensing materials. Including Olfactory Receptor (OR)[21], Odorant Binding Proteins (OBP)[51] and peptides[52].

Olfactory receptors are the molecular basis of mammalian olfaction. Olfactory receptors are a class of transmembrane proteins distributed on the cell membranes of olfactory-related cells[53]. Upon binding with odor molecules, olfactory receptors undergo a change in shape. This conformational shift leads to a change in the cell membrane potential, which triggers downstream signaling pathways. When exposed to a mixture of gases, a multitude of different olfactory receptors produce varied responses, collectively forming an olfactory map that the brain uses to identify odors[54]. When such receptor proteins are modified on the surface of a semiconductor (typically a carbon-based field-effect semiconductor), the binding of the receptor to odor molecules alters the distance between the surface charges of the receptor and the semiconductor surface, affecting the carrier density of the semiconductor and thereby inducing a change in resistance. The potential of this method lies in its ability to measure odors in the same manner as the biological olfactory system. The challenge lies in the fact that olfactory receptors are transmembrane proteins with a hydrophobic, lipid-soluble region in their structure. This means they cannot stably exist in aqueous environments and cannot be expressed using conventional bioengineering methods[55]. To modify wild-type receptors on the surface of sensors, artificial cell membranes(Figure 2.4A) [42,44,56] or living cells (Figure 2.4C)[22,57,58] need to be used as carriers, which severely limits the development and application of such sensors. There are protein engineering techniques that can modify the sequence of olfactory receptors to make them water-soluble[59,60]. Apart from the issue of the universality of these techniques, the function of the receptor binding region is thought to be related to cell membrane tension[61], and it is still unclear whether simply modifying water solubility will affect the receptor's function. In addition, the specific correlations between odors and olfactory receptors, that is, the detailed olfactory map, are not yet clear. The challenges in expressing and purifying receptor proteins further complicate the investigation into their functionality.

Olfactory receptors are not the only biomolecules that can be utilized for e-noses. Odorant Binding Proteins (OBPs) are small, soluble proteins that are secreted into the mucus of the nasal epithelium in vertebrates. They play an essential role in the olfactory system by transporting volatile odor molecules through the aqueous mucus layer to olfactory receptors, which are located on the cilia of olfactory sensory neurons[62]. Due to their solubility in water, OBPs can be readily modified on sensor surfaces, enabling highly sensitive gas sensing[63]. However, while

OBPs function as universal carrier proteins, their specificity and diversity fall significantly short in comparison to olfactory receptors, which constrains their potential utility in e-nose applications. Utilizing peptide-modified sensors represents an alternative compromise solution. Based on wild-type receptor proteins, a peptide sequence can be designed to mimic the receptor's binding region while possessing improved stability and solubility in water[64,65]. However, there is currently no widely accepted method for designing receptor-like peptides, and the functional similarity between synthetic peptides and natural receptors remains questionable.

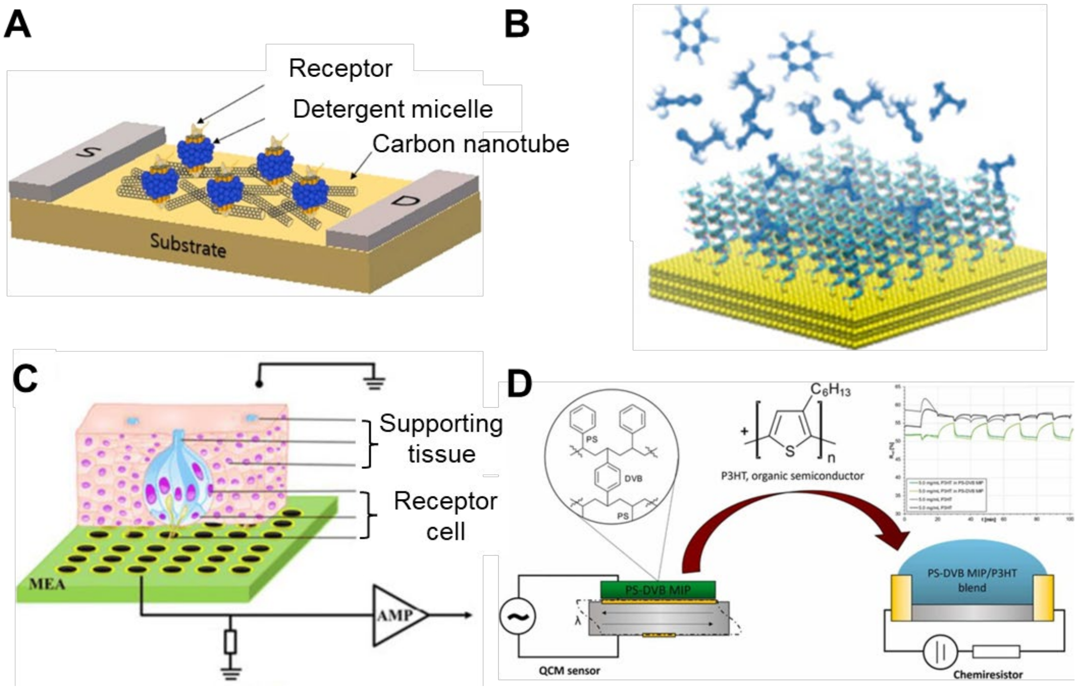

Figure 2.4 Surface-functionalization based sensor. (A). Receptor based gas sensor, reproduced with permission from [42]. (B). Peptide based gas sensor, reproduced with permission from [43], The authors. (C). Live cell-based gas sensor, reproduced with permission from [44]. (D). Molecular imprinting polymer-based gas sensor, reproduced with permission from [45]. These approaches enhance sensitivity and selectivity by introducing delicate bio/chemically specific recognition elements at the sensing interface, illustrating a spectrum of trade-offs between molecular specificity, stability, and system complexity for real-world gas sensing applications.

Despite their high selectivity, surface-functionalized sensors face several significant limitations. The fabrication of customized functional layers is often complex and labor-intensive, requiring precise control over molecular design and surface chemistry. Moreover, many of these functional layers are based on biological materials, such as proteins[42] or peptides[66], which are inherently fragile and have limited operational lifespans. These materials are easy to degrade under extreme conditions such as high temperature, intense light exposure, or fluctuations in pH, making them unsuitable for harsh industrial environments. As a result, surface-functionalized sensors are typically confined to highly controlled settings, such as biomedical diagnostics or laboratory research, and remain difficult to scale or deploy in real-world industrial applications.

2.1.4 Sensor Selection for Robot Deployment

To better contextualize gas sensor choices for robotic deployment, it is important to emphasize that mobility fundamentally changes what matters for robot gas sensing. Because a robot samples along a trajectory and must react online, the most deployment-critical metrics are response speed, recovery speed, and temporal consistency. Under these criteria, physically based PID sensors are among the most commonly adopted options in mobile robotic olfaction, primarily because they provide fast, repeatable responses that better match closed-loop navigation and rapid decision-making. In contrast, pattern-recognition approaches built on MOX sensor arrays often rely on chemical adsorption and surface reactions, which typically exhibit slower response and especially slower recovery, making them less compatible with fast-moving platforms that require frequent, reliable updates rather than long stabilization periods.

This robotics-oriented perspective also clarifies how sensor limitations translate into the algorithmic challenges addressed in later chapters. Slow recovery, baseline drift, and cross-sensitivity reduce the reliability of absolute concentration magnitudes and can create spurious detections along a moving trajectory, which motivates inference formulations that tolerate noise and exploit spatial-contextual cues. Meanwhile, spectroscopic approaches can offer fast response and species-specific signatures, but they typically require dedicated light sources and optical paths, increasing volume, power consumption, and integration complexity. As a result, they are usually practical only in specific scenarios where platform constraints and operating conditions permit. These trade-offs motivate the subsequent emphasis on robust inference under noisy, biased observations and on evaluation protocols that acknowledge the non-idealities of real sensors in mobile settings.

2.2 Current Platforms for Robotic Gas Perception

Robotic gas perception has been explored across various mobile platforms, each offering both advantages and trade-offs in terms of mobility, sampling resolution, and environmental interference. The most widely studied categories include ground robots, multirotor unmanned aerial vehicles (UAVs), and fixed-wing UAVs. These platforms differ not in their dynamic properties and how they interact with the surrounding gas field, which is a critical factor in gas sensing.

2.2.1 Ground Robots

Ground-based robots are commonly used in robotic olfaction due to their simple control schemes, high payload capacity, and low power consumption[13,67]. Their operation on relatively stable surfaces enables precise control over sampling locations and trajectories. Moreover, ground platforms can easily integrate with simultaneous localization and mapping (SLAM) technologies to incorporate spatial context into gas sensing. However, the mobility of ground robots is inherently constrained by terrain conditions. Their ability to access complex environments such as rubble, stairs, or vegetation is limited, which reduces their operational flexibility. Furthermore, ground robots typically sample near the surface, limiting their gas perception to near-ground distributions and potentially missing elevated or volumetric plume structures.

2.2.2 Multirotor UAVs

Multirotor drones, such as quadcopters, offer significantly enhanced spatial flexibility by enabling sampling in the vertical (z) direction. This capability allows robotic gas sensing systems to operate in three-dimensional environments, extending GDM beyond the two-dimensional ground plane. However, the strong downwash produced by multirotor propellers introduces substantial disturbance to the ambient gas field (Figure 2.5B). This airflow can dilute or redirect gas plumes, especially in turbulent or early-stage dispersion scenarios, thus contaminating the very signal that needs to be measured[68]. As a result, multirotor UAVs are best suited for use in environments where gas distributions have already stabilized, such as during stable-stage dispersion or indoor leakage scenarios with limited convection.

2.2.3 Fixed-wing UAVs

Fixed-wing UAVs present an appealing alternative for robotic gas sensing in open, large-scale environments (Figure 2.5C)[69]. Unlike multirotor, fixed-wing platforms generate minimal vertical turbulence during flight, thereby preserving the integrity of the surrounding gas field. Their higher flight speeds and longer operational ranges make them well-suited for covering wide areas and conducting aerial transects across suspected plume regions. However, fixed-wing UAVs cannot hover or maintain a stationary position, and their high-speed motion significantly reduces the interaction time between the sensor and the target gas. Consequently, only fast-response sensors, such as tunable diode laser absorption spectroscopy (TDLAS) or photo ionization sensors can be effectively deployed on fixed-wing platforms. The limited availability and higher cost of such sensors impose constraints on system design and restrict the types of gases that can be detected.

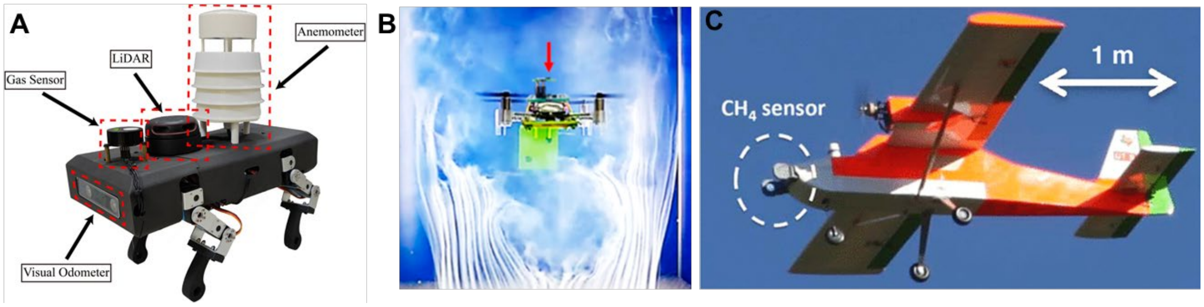

Figure 2.5 Platforms of robotics gas perception. (A) Ground robotics, reproduced with permission from [70](B) Multirotor UAV, which induces significant downwash flow. Reproduced with permission from [68] (C) Fixed-wing UAV, carrying a laser-based sensor, reproduced with permission from [69].

2.3 Properties of Gas Plume

In gas source localization, the source is typically modeled as a point emitter, with gas dispersion governed primarily by airflow rather than molecular diffusion. The resulting gas plume, described by Gaussian plume Law[71]:

$$C(x,y,z)=\frac{Q}{2\pi\sigma_y\sigma_z u}\exp\left(-\frac{y^2}{2\sigma_y^2}\right)\left[\exp\left(-\frac{(z-H)^2}{2\sigma_z^2}\right)+\exp\left(-\frac{(z+H)^2}{2\sigma_z^2}\right)\right] \qquad (1)$$

where $C(x,y,z)$ is the pollutant concentration, Q is the emission rate, u is the wind speed, H is the effective stack height, and σ_y,σ_z are the dispersion coefficients in the crosswind and vertical directions, respectively.

It exhibits several characteristic behaviors (Figure 2.6): 1) It originates from a point source and is transmitted downstream by wind. 2) Due to turbulence at plume-atmosphere interface, gas concentrations fluctuate over time and space, with the highest mean concentration typically found along the plume centerline. 3) Unlike diffusion-driven processes, the plume structure is highly intermittent and patchy, lacking a stable or monotonic concentration gradient. Consequently, the mere presence or detection of the plume rather than the gradient magnitude is often considered a more reliable cue for inferring the source location. Based on this, Some studies model gas plumes analytically[72,73] or statistically[74,75], while others focus on plume tracking strategies. Early plume tracking methods aim to keep robots in contact with the plume, such as adjusting speed or direction when losing contact[76,77]. These methods fall under reactive source localization, discussed in Chapter 2.2.2.

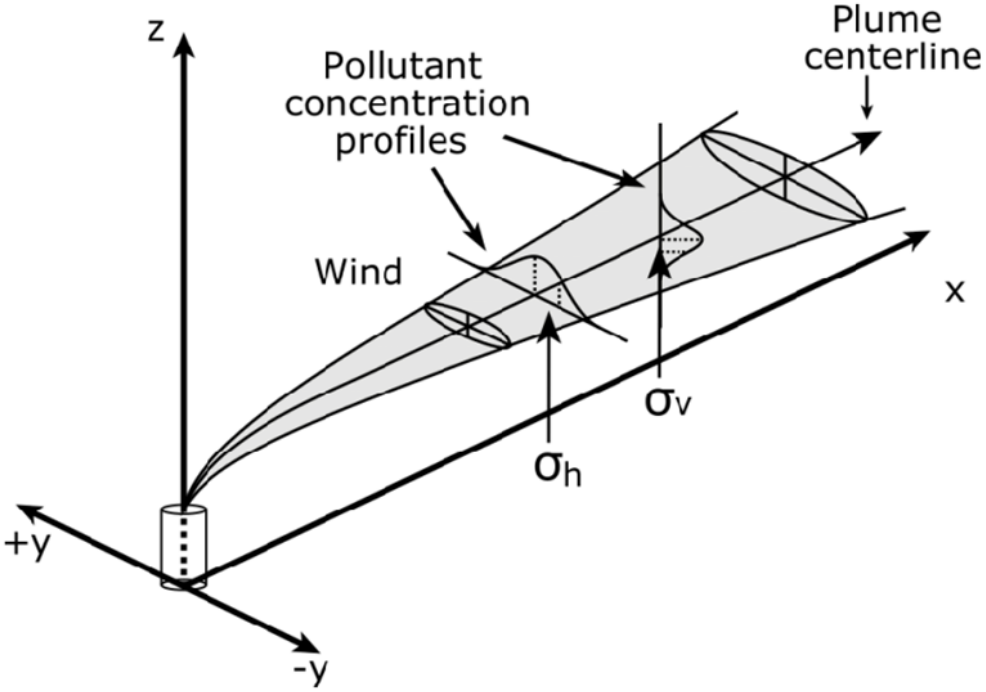

Figure 2.6 Statistical description of Gaussian plume model. Due to the existence of turbulence, the instantaneous concentration spatial and temporal distribution could be highly fluctuating, reproduced with permission from [78].

2.4 Current Approaches of Gas Source Localization

The localization of gas sources in complex environments has been a longstanding challenge in the field of mobile robot olfaction. Conventional approaches are unable to handle complex interaction patterns between dispersed gas and environmental topography. Recent advances in deep learning offer a promising solution by approximating these complex patterns and their relationship to the source location, motivating the development of novel learning-based localization approaches.

2.4.1 Reactive Gas Source Localization

Reactive methods of GSL rely on robots' direct responses to real-time sensor inputs, executing predefined behavior without maintaining internal memory or engaging in dynamic state estimation (Figure 2.7). These approaches are valued for their simplicity, low computational overhead, and ease of deployment across diverse robotic platforms[12]. A typical example is chemotaxis[79], in which a robot continuously moves toward regions of higher gas concentration. However, this strategy assumes the presence of a stable concentration gradient, which is often violated in turbulent dispersion environments where gas distribution is intermittent and patchy. To address this limitation, chemotaxis-anemotaxis[80] methods incorporate wind direction data alongside concentration readings, enabling robots to perform biologically inspired behaviors such as upwind surging and crosswind casting or zigzagging[81,82], which are patterns observed in moths and other insects during olfactory navigation.

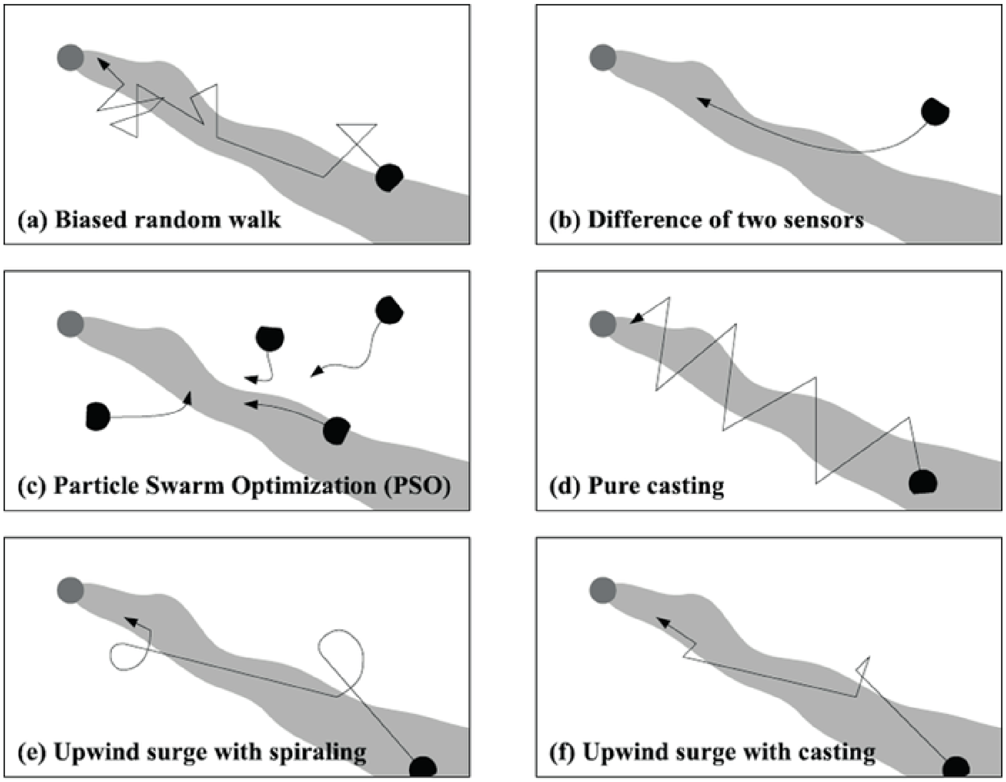

Figure 2.7 Reactive gas source localization approaches. Although with different strategies,

these approaches rely on driving the robot towards potential gas sources, which is limited by robots' accessibility, reproduced with permission from[9].

Despite their adaptability, reactive strategies generally suffer from limited flexibility and robustness in complex or dynamic environments. In disaster response scenarios, for example, obstacles and irregular terrain can disrupt simple tracking behaviors. Moreover, in multi-objective missions, where robots are required to follow predefined trajectories or conduct area surveillance persistent plume tracing becomes impractical. Under such constraints, the ability to infer or predict the gas source location from sparse or indirect observations during exploration becomes a more viable and efficient alternative, motivating the transition toward model-based and learning-driven approaches.

2.4.2 Statistical Gas Source Localization

To overcome the limitations of reactive approaches, statistical GSL methods employ probabilistic frameworks to estimate gas source locations by integrating sensor data, environmental knowledge, and prior assumptions[83–85] (Figure 2.8). These methods treat localization as a recursive state estimation problem, utilizing tools such as Bayesian filtering20, Kalman filtering[86], and particle filtering[84,87]to continuously update belief distributions as new measurements are acquired. A notable example is Infotaxis, a strategy that directs the robot's motion toward maximizing expected information gain, effectively balancing exploration and exploitation under uncertainty[88–90].

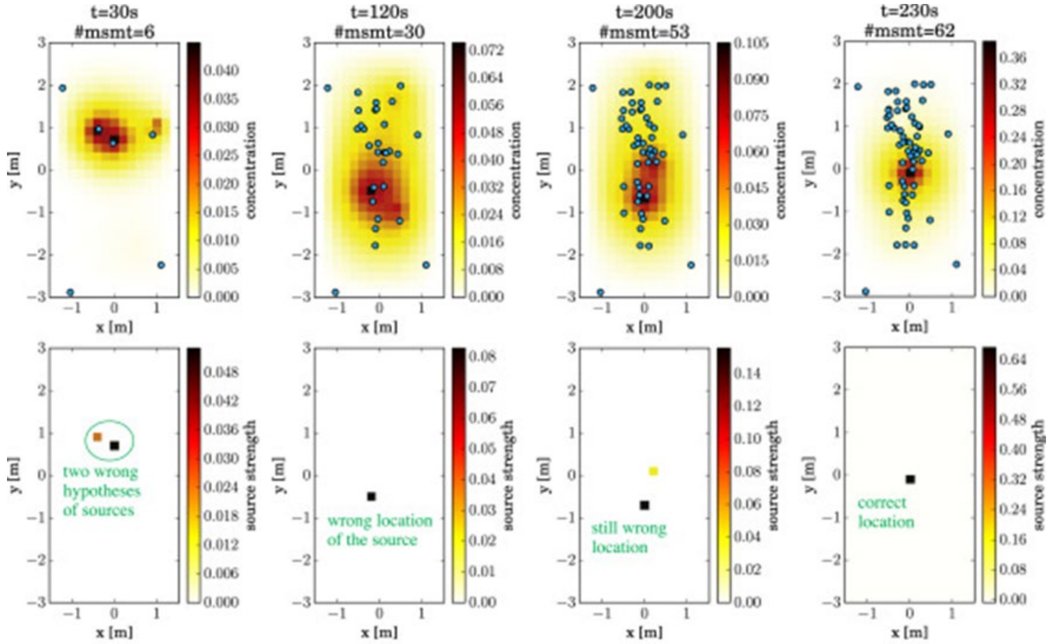

Figure 2.8 Statistical approach of gas source localization, reproduced with permission from [85]. It estimates potential gas source location by Bayesian estimation with gas transmission model.

Statistical methods offer several advantages: they are inherently capable of handling noisy and intermittent measurements, account for the stochastic nature of gas dispersion, and provide uncertainty quantification alongside location estimates. However, their performance highly depends on the validity of the underlying gas dispersion model[91]. Many implementations assume simplified environmental conditions such as a uniform wind field, where the wind vector is considered constant throughout the search space. While such assumptions may hold in controlled settings, they often break down in real-world scenarios featuring wind-topology interactions. This difference between model assumptions and environmental realities can severely affect localization accuracy, thus limiting the generalizability of statistical methods in practical deployments.

2.4.3 Learning-based Gas Source Localization

In recent years, deep learning has emerged as a powerful tool in robotics, offering strong decision-making capabilities under sparse or noisy input conditions[92]. In the context of gas source localization, deep reinforcement learning approaches, such as Deep Q Networks[67,93] (DQN) (Figure 2.9) and Long Short-Term Memory[94] (LSTM) based models have shown promising performance by learning effective navigation or inference policies directly from interaction with the environment. These methods overcome some of the modeling limitations inherent in traditional statistical techniques by capturing nonlinear dependencies between sensor readings, robot states, and environmental cues.

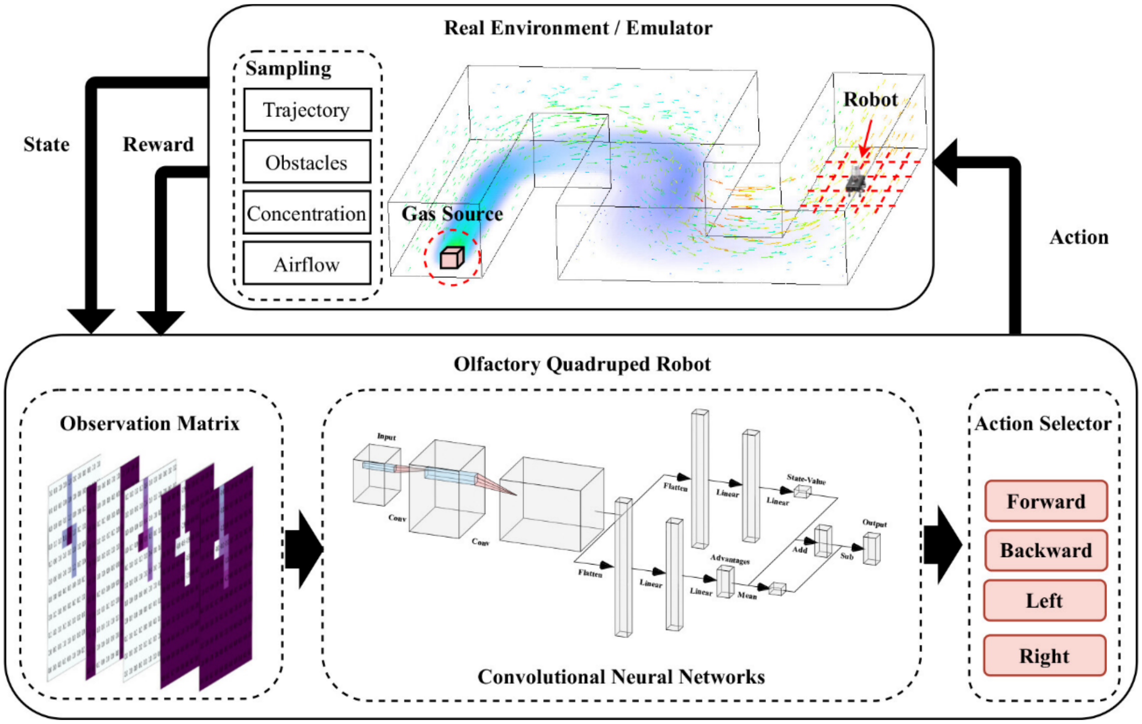

Figure 2.9 Gas source localization based on Deep Q Network (DQN), reproduced with permission from[70]. DQN is a reinforcement learning algorithm that combines Q-learning with deep neural networks to learn optimal actions from high-dimensional input states.

Furthermore, advances in computer vision have demonstrated the effectiveness of convolutional neural networks[95] (CNNs) and encoder-decoder architecture[96] in extracting spatial relationships from image-like data. Models such as U-Net[97], Mask R-CNN[98], and DeepLab[99] have achieved state-of-the-art performance in segmentation and spatial inference tasks by leveraging hierarchical feature representations. Drawing inspiration from these architectures, this thesis proposes to incorporate contextual spatial information embedded in grid maps to enhance gas source localization. By fusing in-situ sensor observations with environmental layouts, this approach aims to enable more accurate and robust source predictions, even in complex and dynamically changing environments.

2.5 Current Approaches Gas Distribution Mapping

Existing approaches to gas distribution mapping fall into two broad categories: rule-based methods and learning-based methods. Early rule-based techniques typically rely on interpolation schemes or probability estimates. While these methods are straightforward, they often assume dense, uniform sensor coverage. This inflexibility has motivated the development of learning-based approaches. Such data-driven methods like convolutional neural networks[7] (CNNs) or long short-term memory[100] (LSTM) can capture more complex plume patterns than rule-based ones, but they generally require large training datasets and may struggle to generalize when sample sizes are limited, or sensor readings are irregular and noisy. As a result, current approaches cannot fully address the challenges of varying sample densities, sparse and uneven sampling, or deployment in unstructured real-world environments.

2.5.1 Rule-based Gas Distribution Mapping

Early approaches to GDM predominantly relied on rule-based and interpolation-driven methods to estimate spatial gas concentration fields from sensor measurements. Techniques such as Kernel DM+V/W[74] and Gaussian Process Regression[101] (GPR) were widely adopted to construct continuous concentration maps from discrete sensor data. While these methods offer smooth interpolation and uncertainty quantification, they are inherently limited in modeling complex, nonlinear gas dispersion behaviors. Their performance is highly dependent on the density and uniformity of the sensor layout, often assuming regular sampling patterns and abundant data coverage, which are rarely met in practical deployments involving mobile sensors or sparse measurements (Figure 2.10).

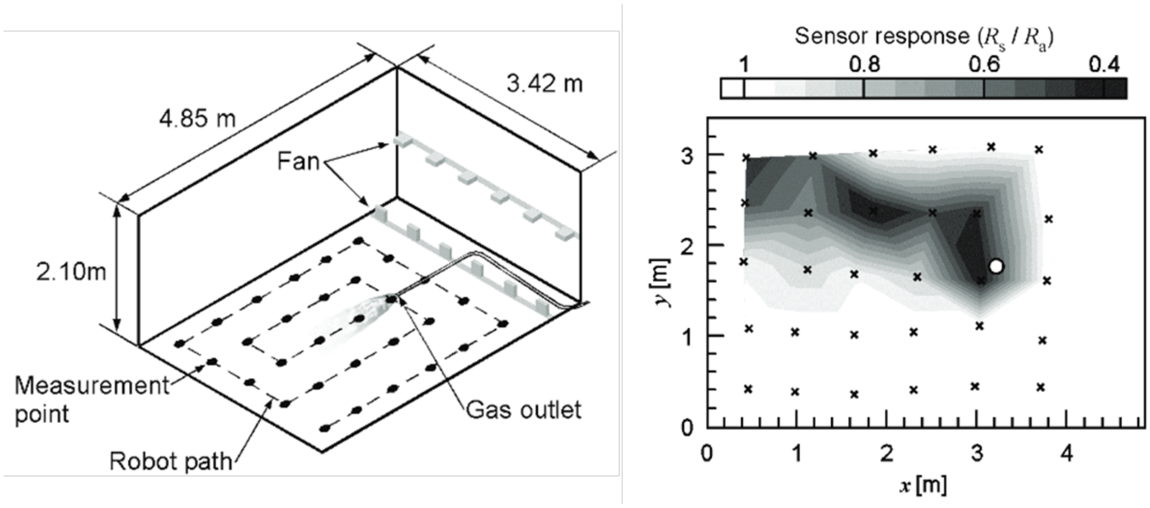

Figure 2.10 Rule based gas distribution mapping, reproduced with permission from [102]. This approach requires the robot to uniformly traverse the area and perform interpolation. The mapping quality is strongly coupled to coverage uniformity and sampling density, which can be difficult to guarantee in practical deployments with limited time, obstacles, or uneven exploration.

In parallel, physics-based approaches such as Computational Fluid Dynamics[103] (CFD) provide more reasonable modeling by numerically solving the Navier–Stokes and diffusion equations under specified boundary conditions. These methods can capture fine-grained flow-gas interactions and produce high-fidelity simulations of gas dispersion. However, the computational expense and sensitivity to initial and boundary conditions significantly constrain their applicability in dynamic or real-time settings, especially in large-scale or cluttered industrial environments where environmental parameters are not fully observable or change frequently. These limitations have motivated the exploration of learning-based GDM methods that aim to combine the predictive power of data-driven models with the physical interpretability of traditional approaches.

2.5.2 Learning-based Gas Distribution Mapping

In recent years, machine learning techniques, particularly deep learning models have gained increasing traction in GDM due to their adaptability, data efficiency, and capacity for real-time inference (Figure 2.11). Convolutional Neural Networks[7] (CNNs) and long short-term memory[100] (LSTM) have demonstrated promising performance in reconstructing spatial gas distributions from sparse sensor observations. These models can learn complex spatial patterns without explicit physical modeling. However, conventional CNN-based architectures often assume fixed-size grid inputs and rely on a large number of training samples to capture spatial dependencies effectively. Moreover, the limited receptive field of standard CNNs restricts their ability to model long-range spatial correlations and global distribution patterns, which are crucial for accurately reconstructing turbulent gas fields.

To address some of these challenges, Physics-Informed Neural Networks[104] (PINNs) have emerged as a compelling alternative by embedding the governing partial differential equations of fluid dynamics directly into the learning objective. By combining data-driven inference with physical constraints, PINNs improve the robustness and interpretability of predictions, particularly under data-scarce conditions. However, the highly nonlinear nature of fluid dynamics equations poses a significant optimization challenge[105]. The lack of large-scale, high-quality open-source datasets for GDM further enhance the difficulty, as small or noisy self-collected datasets can introduce inconsistencies between data loss and physics-based loss. This often leads to optimization failures or trivial solutions.

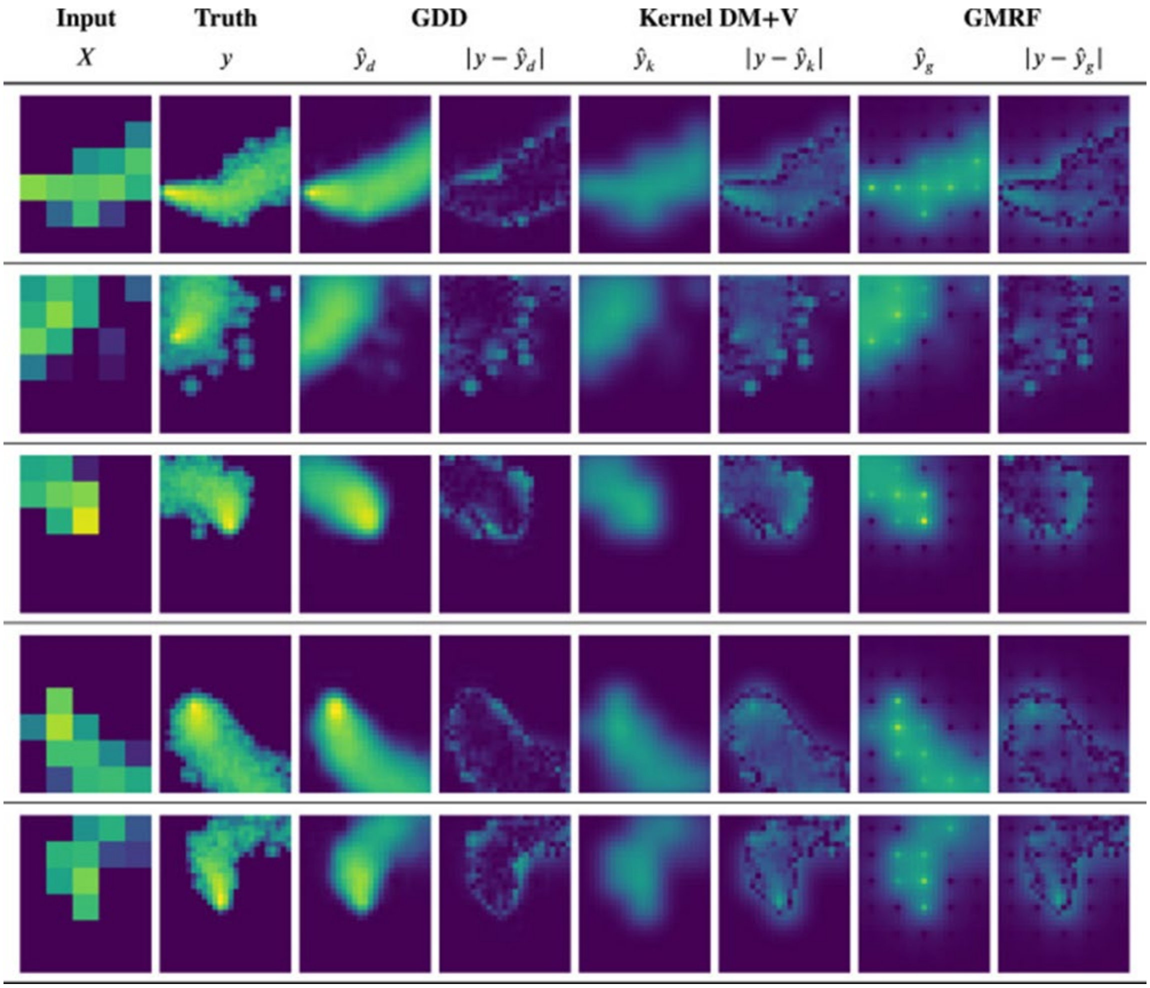

Figure 2.11 Input, prediction and error of different gas distribution mapping approach, reproduced with permission from [7]. Gas Distribution Decoder (GDD) is a CNN based prediction model, which performs better accuracy, especially at edge fine structure compared to traditional method.

Recently, Transformer architecture has shown remarkable potential in fluid-related learning tasks[14], including turbulence reconstruction[15] and temporal flow prediction[16]. Owing to its attention mechanism[106], Transformer models naturally handle variable-length and unordered inputs while capturing global spatial relationships through contextual embeddings. These capabilities make it especially suitable for sparse and irregularly sampled gas distribution mapping scenarios. Motivated by these strengths, we propose a Transformer-based framework tailored to reconstruct gas concentration fields from sparse sensor inputs, leveraging the model's capacity for long-range dependency modeling and flexible input representation. Beyond empirical performance, the architectural choice is primarily driven by the structure of robotic GDM inputs. Measurements arrive as an unordered set of samples with variable cardinality, each associated with a spatial coordinate and multi-modal attributes (e.g., concentration and wind). Rasterization-based networks like CNN pipelines require either dense gridding or ad-hoc interpolation to map such sets to fixed-size tensors, introducing discretization bias and coupling the representation to a particular sampling resolution. In contrast, attention-based set processing provides permutation invariance over sampled points, a global receptive field that scales naturally with sparsity, and adaptive, content-dependent aggregation that can be interpreted as a learned kernel operating over spatially embedded observations. This is particularly aligned with gas dispersion reconstruction, where long-range coherence of plumes and flow continuity can depend on non-local cues. Alternative deep models such as graph neural networks[107] can also represent irregular samples, but they typically require explicit graph construction (e.g., k-Nearest-Neighbor radius selection) that is brittle under highly varying sample densities and sampling bias. Therefore, the transformer is positioned in this thesis as a principled compromise: it retains flexibility for unstructured sensing while avoiding hand-crafted neighborhood definitions, and it supports cross-modal fusion through attention when jointly reconstructing wind and concentration fields.

2.6 Outstanding Questions

Currently, both gas source localization (GSL) and gas distribution mapping (GDM) remain open research challenges. Existing GSL methods often limit flexibility and have strong dependence on an accurate gas dispersion model, resulting in struggling in environments with complex topographies and turbulent plume dynamics like disaster sites. In real-world conditions, the relationship between observed gas concentrations, wind directions, and source location is highly nonlinear and environment-dependent, yet most existing approaches lack the capacity to incorporate spatial context such as obstacle layout or wind-field distortion caused by structures. As a result, localization accuracy remains limited in practical applications.

Similarly, while data-driven approaches have recently improved GDM performance, most models are still based on regular, grid-based architectures that assume fixed sensor layouts and uniform sampling. These assumptions do not hold in realistic robotic deployments, where sampling is inherently sparse, irregular, and biased due to motion constraints. Moreover, many existing GDM models fail to explicitly encode the spatial relationships between sampling points, making them sensitive to sampling errors and limiting their ability to reconstruct high-fidelity gas fields from minimal observations.

Considering these gaps, there is a pressing need for robotic olfaction systems that can (i) integrate environmental knowledge into the inference of gas source locations under turbulent and occluded conditions, and (ii) leverage flexible, spatially aware models for reconstructing gas distributions from sparse and unstructured sensor inputs. Addressing these challenges is essential for deploying robotic olfaction in real-world tasks such as industrial inspection, environmental hazard detection, and search-and-rescue missions. This thesis aims to contribute toward this goal by developing novel GSL and GDM frameworks that are both environment-aware and sampling-adaptive, advancing the robustness and applicability of robotic gas perception.

Chapter 3

Experimental Methodology

This chapter presents the methodology for developing robust robotic gas perception systems addressing two key tasks: gas source localization (GSL) and gas distribution mapping (GDM). For GSL, a learning-based inference framework is proposed to tackle the challenges of early-stage gas dispersion, where turbulent plumes and obstacle-induced wind interactions disrupt traditional tracking methods. A custom mobile platform equipped with Photoionization Detector (PID) gas sensors, ultrasonic anemometers, and LIDAR-enabled SLAM is used to collect spatially anchored gas and wind data. A U-Net-based model is trained using a dynamically weighted loss function that adapts to the number of plume encounters, enabling accurate source inference under uncertainty.

For GDM, a Transformer-based architecture is introduced to reconstruct gas and wind fields from sparse and irregular sensor observations. The model incorporates positional encoding and multi-head self-attention to capture long-range spatial dependencies and adapt to varying input geometries. Training is conducted on a synthetic CFD dataset generated via OpenFOAM simulations under randomized flow conditions, and model robustness is validated through real-world wind tunnel data. Curriculum learning and OneCycleLR scheduling are used to improve training efficiency and generalization. Together, these methods form an *integrated robotic olfaction framework capable of reliable operation in complex, real-world environments.*

3.1 Rationale for Selection

As two inverse problems arise at different stages of the underlying fluid-dynamic process, target variables of GSL and GDM are task-dependent. For gas source localization (GSL), the target is the source location $\mathrm{x}_s \in \Omega$, where Ωdenotes the 2D workspace domain and $\mathrm{x}_s$is the unknown source coordinate. In practice, GSL is often expressed as estimating a posterior distribution $p(\mathrm{x}_s \mid o_{1:t}^{\mathrm{GSL}})$, where $p(\cdot)$quantifies localization uncertainty and $o_{1:t}^{\mathrm{GSL}}$denotes the history of GSL observations up to time t. Since GSL must reason over spatial context and occlusions, its per-step observation is defined as $o_t^{\mathrm{GSL}} = (\mathcal{M}, \tilde{\mathrm{u}}(\mathrm{p}_t, t), e_t, \mathrm{p}_t)$, where $\mathcal{M}$is a geometric map encoding free space and obstacles, $\mathrm{p}_t \in \mathbb{R}^2$is the robot's planar position at time t, $\tilde{\mathrm{u}}(\mathrm{p}_t, t)$is the wind measurement acquired at p_t, and $e_t \in \{0,1\}$is a binarized gas-encounter indicator that equals 1 when the measured concentration exceeds a predefined threshold and 0 otherwise; this event-driven sensing can be interpreted as a sparse observation of the plume's reachable region or transport trajectory. In contrast, for gas distribution mapping (GDM), the target is the spatial concentration field $c(\mathrm{x}, t)$over $\mathrm{x} \in \Omega$, where $c(\mathrm{x}, t)$denotes the gas concentration at location $\mathrm{x}$and time t(and can be discretized as a grid field $\mathrm{C}_t$in implementation). The GDM observation at time $t$is a set of sparse measurements $o_t^{\mathrm{GDM}} = \{(\mathrm{p}_i, \tilde{c}_i, \tilde{\mathrm{u}}_i)\}_{i=1}^{N_t}$, where $\mathrm{p}_i$is the i-th sampling coordinate, $\tilde{c}_i \approx c(\mathrm{p}_i, t)$ is the measured concentration at p_i, $\tilde{\mathrm{u}}_i$ is the co-located wind measurement, and $N_t$is the number of available samples at time t. Finally, as both tasks are formulated for a mobile robot, we define the state in both problems as the robot's planar position $s_t \triangleq \mathrm{p}_t \in \mathbb{R}^2$; therefore, different encoding schemes must be selected to match the characteristics of each task.

In this thesis, the term physics-aware is used to describe a weak physical supervision approach tailored to robotic gas sensing, where the underlying dispersion process is physically well-defined yet practically complex and highly variable. Although gas transport is governed by conservation laws and advection-diffusion dynamics, acquiring large-scale, high-fidelity datasets remains difficult because training data must be obtained either from costly real-robot/wind-tunnel experiments or from computationally intensive CFD simulations. Under such data scarcity, purely data-driven models often suffer from limited data efficiency and may fail to generalize across environments and flow regimes. Incorporating physical priors can therefore improve sample efficiency; for example, physics-informed neural networks (PINNs) have shown promising performance on certain fluid problems by enforcing PDE residuals as training constraints[104,108]. However, direct equation enforcement is not always suitable for the GSL and GDM settings studied here. First, the PDE solution space can be highly nonlinear, making optimization sensitive and convergence difficult in practice. More importantly, the training observations of gas dispersion are not perfectly physics-consistent due to sensor noise, bias, discretization, and sim-to-real discrepancies, so a physics residual loss may induce gradient directions that conflict with the data-fitting loss. Such conflicts can push the optimization toward degenerate or trivial solutions (e.g., physically plausible outputs that nonetheless mismatch measurements). For these reasons, this thesis adopt a physics-aware approach that integrates physical context without directly enforcing governing equations. The work in this thesis injects physically relevant modalities, such as explicit obstacle information $\mathcal{M}$, the spatial encoding of sampling locations $\mathbf{p}$, and physically meaningful wind/encounter cues. As a result, the model is encouraged to learn patterns consistent with gas dispersion while being trained primarily under the data loss. In this sense, physical knowledge acts as a structured prior that guides representation learning, rather than as a hard PDE constraint.

For GSL, the early stage of gas dispersion is dominated by turbulent plumes and complex wind-obstacle interactions, which limits the formation of stable, continuous concentration gradients. Traditional tracking or gradient-based methods fail under such conditions due to low signal continuity and high uncertainty. To address this, a learning-based inference framework is proposed, which leverages spatial context, including wind direction and environmental layout obtained via SLAM to model the non-linear relationship between plume encounters and true source locations. A PID sensor is chosen for its fast response and stability, providing reliable detection events for learning-based inference. Furthermore, an ultrasonic anemometer and a LIDAR-based SLAM system are integrated to supply environmental structure and wind vectors, allowing gas perception to be anchored in physical space. The usage of U-Net architecture is based on its proven effectiveness in spatial inference tasks, and is further enhanced by a dynamic loss function tailored to the progressive accumulation of sensor data.

In contrast, GDM deals with the stable stage of dispersion. Here, the goal is to reconstruct a continuous spatial field from sparse and irregular sensor inputs. Traditional CNN architecture assumes grid-aligned inputs, which are ill-suited to real-world robotic sampling. To overcome this, a Transformer-based model is employed due to its ability to process set-structured inputs, capture long-range spatial correlations via self-attention, and flexibly handle variable sampling patterns. The model is trained on synthetic CFD data generated by OpenFOAM, which allows controlled exploration of a wide range of source and flow configurations. To ensure robustness and generalizability, curriculum learning is adopted to progressively increase task difficulty, and real-world validation is conducted using wind tunnel experiments. The combined use of positional encoding, feature fusion, and Transformer decoder queries enables the model to infer high-resolution gas and wind fields from minimal observations.

3.2 Gas Source Localization with Mobile Robot

3.2.1 Design of Robot Platform

As a easy access, nontoxic gas with density is close to that of air, ethanol was selected as the model gas to be emitted. Meanwhile, a photoionization detector[109] (PID) sensor were used as the sensing device, respectively, for the robotic spatial olfaction experiments.

PID is a type of gas sensor that detect volatile organic compounds (VOCs) and certain inorganic gases based on the principle of photoionization (Figure 3.1). In a PID sensor, gas molecules are ionized by high-energy ultraviolet (UV) light. When the energy of the UV photon exceeds the ionization potential of the target gas molecule, electrons are ejected, producing positively charged ions. These ions generate a measurable current that is proportional to the gas concentration. PID is characterized by its fast response time (often within seconds), high sensitivity, and high signal-to-noise ratio, making them ideal for real-time gas detection. In this thesis, the PID sensor is employed as a model gas sensor due to its rapid and stable response characteristics.

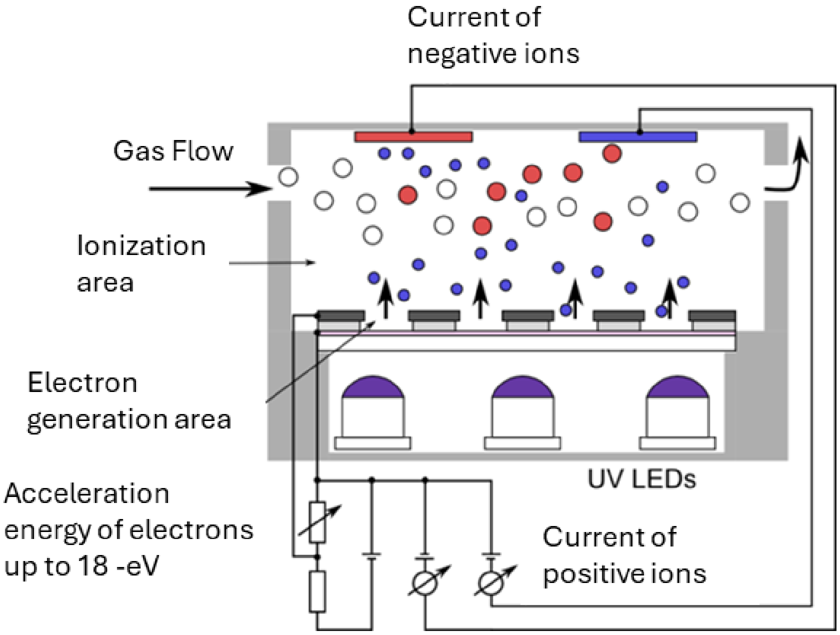

Figure 3.1 Working principle of PID sensors, reproduced with permission from [110]. Gas is detected by measuring the ions ionized from gas molecules by high-energy ultraviolet LED.

As an important variable in spatial olfaction tasks, environmental wind was measured using an ultrasonic anemometer. Ultrasonic anemometers[111] measure wind speed and direction based on the time-of-flight principle of ultrasonic sound waves (Figure 3.2). The device typically consists of multiple transducers arranged in pairs along orthogonal axes. These transducers alternately transmit and receive ultrasonic pulses. When wind is present, it alters the time it takes for the sound waves to travel between the transducers: sound waves traveling with the wind arrive faster, while those traveling against the wind arrive slower. By comparing these transit times in multiple directions, the anemometer calculates both the magnitude and direction of the wind vector. In robotic applications, their compact size, precision, and real-time capability make them ideal for monitoring airflow patterns.

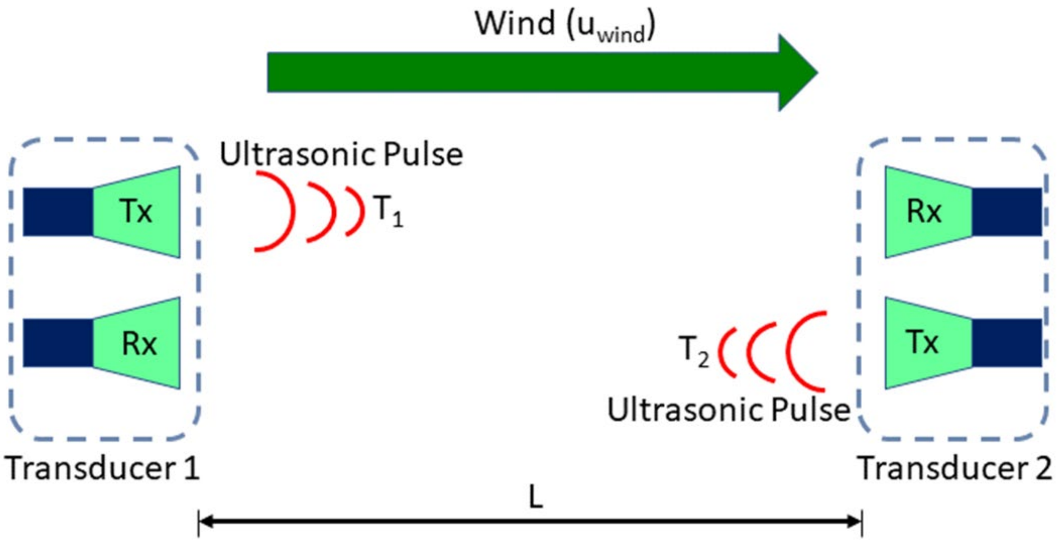

Figure 3.2 Working principle of ultrasonic anemometer, reproduced with permission from [112]. An ultrasonic anemometer measures wind speed and direction using the time-of-flight of ultrasonic sound waves between transducers.

The environmental map is scanned using Light Detection and Ranging[113] (LIDAR), which is a remote sensing technology that measures distance by emitting laser pulses and detecting the time it takes for the reflected light to return from surrounding objects (Figure 3.3). By scanning the environment with rapid, high-frequency laser pulses, LIDAR systems can generate precise two-dimensional maps of the surroundings, capturing structural details such as the shape, position, and size of objects. In robotics, LIDAR is widely used for simultaneous localization and mapping (SLAM), enabling a robot to build a map of an unknown environment while simultaneously determining its own position within that map. SLAM enables the robot to accurately associate gas measurements with spatial coordinates, which provides spatial context for gas source localization.

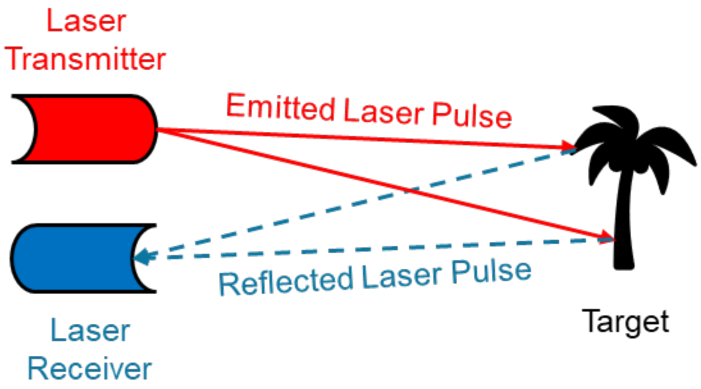

Figure 3.3 Working principle of LIDAR. LIDAR is a remote sensing technology that measures distances by illuminating a target with laser light and analyzing the reflected signal.

3.2.2 Experiment Setup and Dataset Collection

Experiments were conducted in an enclosed indoor environment measuring 10 meters by 10 meters, providing a controlled yet realistic setting for gas dispersion studies (Figure 3.4A). To simulate airflow, three electric fans were deployed to create an artificial wind field. A large central obstacle was placed within the space to significantly alter the global wind flow, while several smaller side obstacles were introduced to influence localized air flow and shape the resulting gas plume structure. All obstacles were intentionally designed to exceed the height of the fans, ensuring that airflow interacted with the obstacles rather than bypassing them.

The target gas used in this study was ethanol vapor, generated via an ultrasonic vaporizer (Figure 3.4B). Ethanol liquid was stored in a container where an absorbent material ensured continuous contact with an ultrasonic transducer. The ultrasonic produced ethanol vapor at room temperature without requiring thermal heating. Near the ground level, the vapor remained relatively stable in vertical position, reducing buoyancy-driven rise and allowing horizontal dispersion to be governed primarily by wind-induced advection rather than temperature gradients.

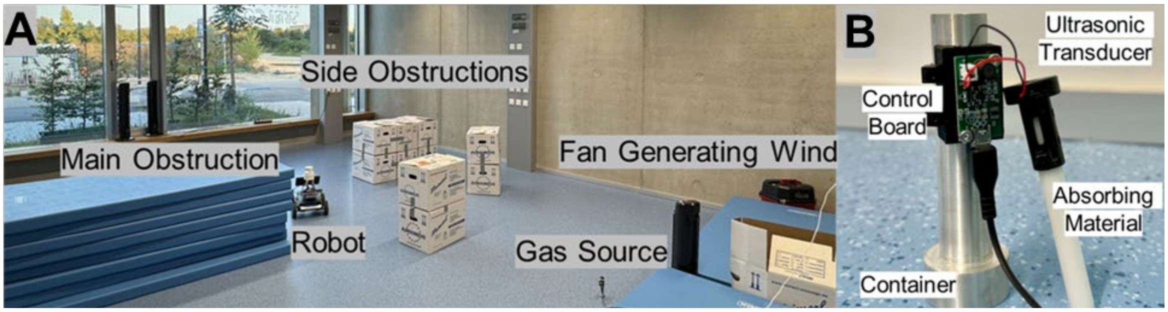

Figure 3.4 Setup of Gas Source Localization Experiments. (A) Setup of the experiment scenario. Different topographies are set by changing number and locations of obstructions. (B) The artificial gas source. Ethanol vapor is generated by an ultrasonic transducer.

The experimental trials were conducted under multiple configurations to evaluate the robustness and generalizability of the proposed method. Variables included different gas source locations, fan directions, obstacle arrangements, and initial robot deployment positions. This diversity of conditions aimed to simulate a range of environmental complexities and ensure comprehensive performance assessment.

The dataset was collected under a human-in-the-loop, event-based principle, collecting representative data that encodes the tempo-spatial pattern of plumes released by the gas source. This pattern is indirectly reflected by gas sensor responses, the corresponding position, and observed wind flow when the robot encounters a plume. To identify meaningful encounters with the plume, a two-stage detection and validation process was implemented. In the first stage, an event-based filter automatically flagged candidate encounters by monitoring rapid fluctuations in gas concentration signals, serving as a preliminary indicator of plume contact. In the second stage, a human operator reviewed the flagged events in real time to validate their significance and prevent redundant recordings, particularly those occurring in close spatial proximity to previously logged points. The event-based filter follows the approach proposed by Li et al[114], it binarizes the sensor readings, reducing data dimensionality and enhancing generalizability across different sensors (Figure 3.5).

$$C_{filtered} = \begin{cases} 0, if\ c_t \leq kc_{t-1} \\ 1,, if\ c_t > kc_{t-1} \end{cases} \tag{2}$$

$$\overline{c_t} = \gamma\overline{c_{t-1}} + (1-\gamma)c_t \tag{3}$$

where c_t is the sensor response at time t, k is a predefined noise threshold, and γ is a trade-off parameter controlling filter sensitivity. This filter converts continuous sensor readings into binary values indicating plume encounters. With an adaptive threshold $\overline{c_t}$, it detects plumes while mitigating the influence of accumulated gas in the environment. During the experiment, the gas sensor sampling frequency was set to 1Hz, with $\gamma = 0.2$ and $k = 1.04$, determined through preliminary experiments to minimize false positives. When the filter returns 1, the encounter location is recorded after operator confirmation, preventing excessive repeated triggers near the gas source.

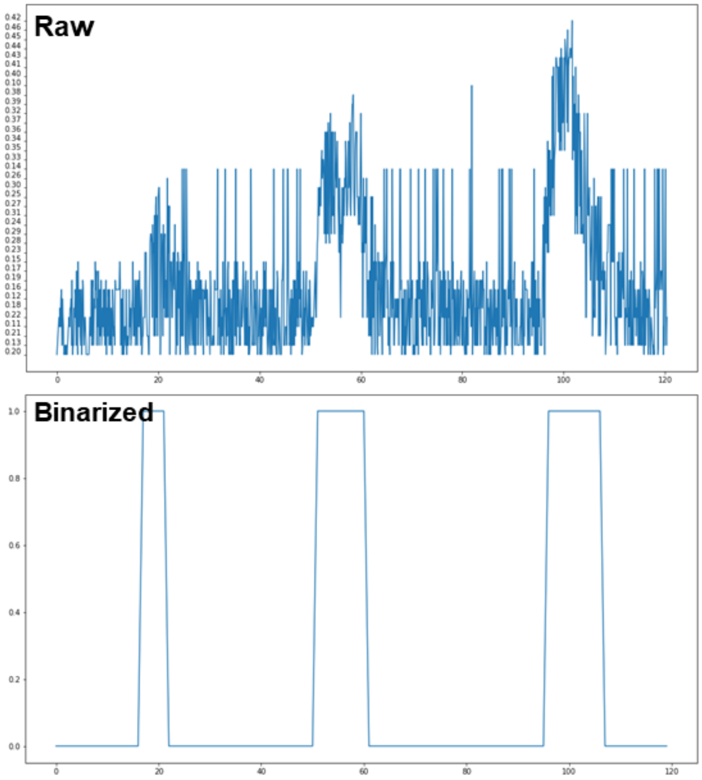

Figure 3.5 Demonstration of the event-based filter. Noisy, continuous sensor reading is transmitted to binary gas encounter event.

Training data collection requires multiple trials. A trial is defined as a mapping process from an initially empty map. When a plume encounter is confirmed, the current occupancy map, the robot's location, and anemometer data are saved. At the trial's end, the robot is guided to the gas source, and its final position on the map is recorded as the true gas source location (Figure 3.6D). Different trials vary in room topography, wind direction, gas source position, and initial robot location.

After initial collection, the dataset is then pre-processed for further training. The map is generated by cartographer, a widely used SLAM package. Maps are output in the form of occupancy grid map[115], which represents the environment as a two-dimensional grid of discrete cells. Each cell is assigned a probability value indicating whether it is occupied, free, or unknown based on sensor observations. In the grid map, unknown is marked as -1, free is 0, other grids are marked as an integer from 0 to 100, representing the probability of being occupied. Each recorded gas encounter is stored as a three-channel map tensor: map, gas encounter, and wind. The map channel records the occupancy grid map at the encounter moment, standardized by padding with -1. The encounter channel, matching the occupancy grid size, marks encounter coordinates with 1, while other grids remain at 0. All prior encounters are recorded cumulatively. During each trial, samples are sequentially generated in encounter order.

The wind channel plays a key role in gas source localization, assuming gas disperses from upwind to downwind. For a single gas source, the probability of its location in the downwind region relative to an encounter point is lower than in the upwind region. In our wind channel implementation, the anemometer provides a wind vector $\overrightarrow{V_{\text{wind}}}$, with wind field matrices matching the map's shape. Each grid cell's value W_i is determined by its orientation relative to the wind: upwind if $\vec{V_i} \cdot \overrightarrow{V_{\text{wind}}} \geq 0$, and downwind otherwise. The wind channel accumulates all encountered wind field matrices over time, offering a comprehensive view of wind conditions on the map.

$$W_i = \begin{cases} 1, if\ \vec{V_i} \cdot \overrightarrow{V_{\text{wind}}} \geq 0 \\ -1, if\ \vec{V_i} \cdot \overrightarrow{V_{\text{wind}}} < 0 \\ 0, if\ \overrightarrow{V_{\text{wind}}} = 0 \end{cases} \tag{4}$$

To enhance model robustness, all samples are rotated clockwise by $\frac{\pi}{2}, \pi, \frac{3\pi}{2}$ radians for data augmentation. Figure 3.6A, 3.6B, 3.6C are visualization of the map tensor sample corresponding to the 3rd encounter event in a trial.

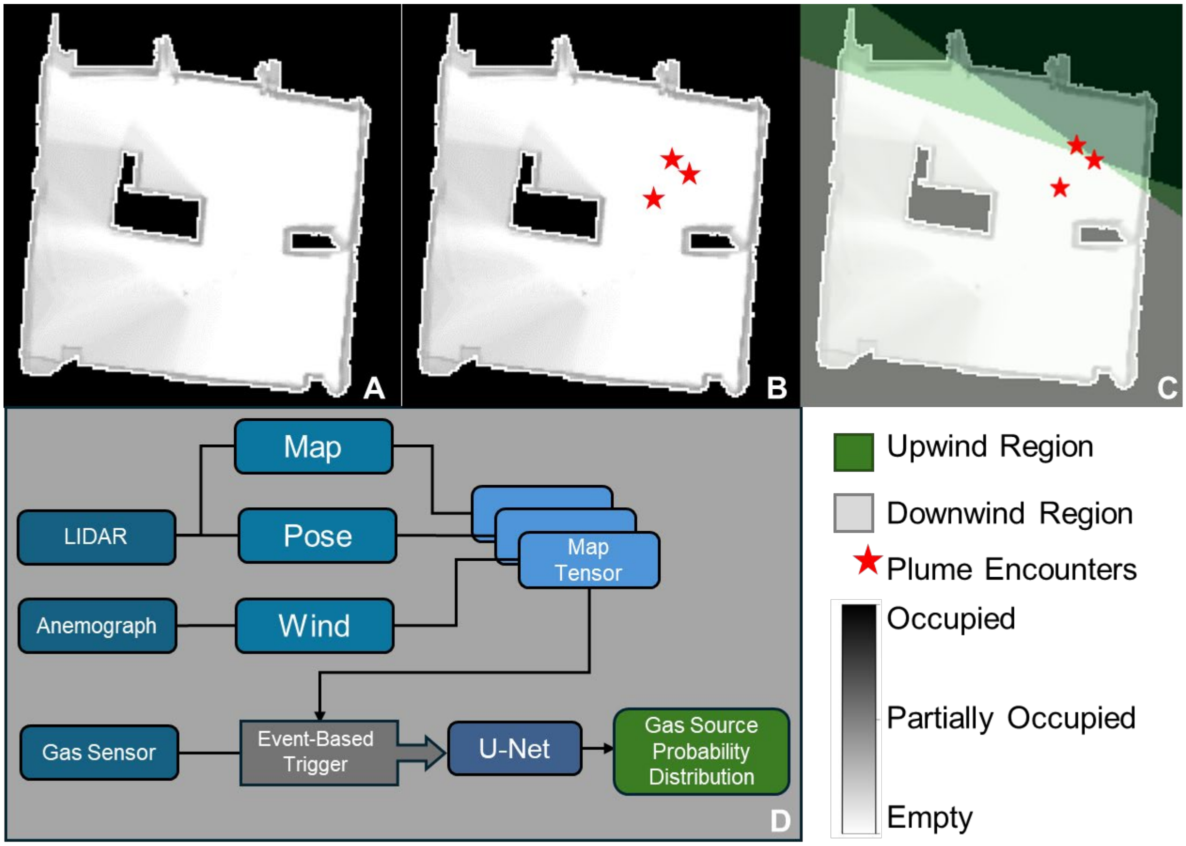

Figure 3.6 The pre-processed dataset and model diagram. Samples in dataset are 3-channel tensors (A) Occupancy grid map channel, where black area means unknown, marked as -1 in the matrix, white area means free, marked as 0, grey area is marked as integers from 0 to 100, representing the probability of being occupied. (B) Gas encounter locations channel, a matrix matching the occupancy grid size, marks encounter coordinates with 1, while other grids remain

0. (C) Wind field regions channel. Deeper green color means higher possibility of being upwind. (D) Diagram of data exploiting, three channels are from different sensor modalities and are combined to be proceeded by U-Net.

3.2.3 Model Architecture and Training

U-net was chosen as the backbone model. U-Net is a convolutional neural network architecture originally designed for image segmentation[97], but it has since been widely adopted for various dense prediction tasks requiring pixel or grid level outputs[116,117]. Its architecture is characterized by a symmetric encoder-decoder structure with skip connections, forming a distinctive "U" shape (Figure 3.7). The encoder path gradually reduces the spatial resolution of the input through a series of convolution and max-pooling layers, allowing the network to capture high-level features and global context. The decoder path uses up sampling to restore the original resolution, enabling precise localization. Skip connections between corresponding encoder and decoder layers allow the model to retain fine-grained spatial information lost during down sampling, improving the accuracy of the final output.

The processed dataset was split 4:1 into training and validation sets, ensuring robust evaluation while preventing data leakage by keeping samples from the same trial in the same set. The U-Net model features a lightweight symmetric encoder-decoder with 23 convolutional layers for deployment on the robot's processor. The encoder has four down sampling steps, each with two 3x3 convolutions, batch normalization, and ReLU activation, with 2x2 max-pooling halving the feature map size at each step. Starting with 64 channels, the count doubles per layer, reaching 1024 at the bottleneck. The decoder mirrors the encoder, using transposed convolutions for up sampling and concatenation with encoder features. A final 1x1 convolution outputs the class probabilities via a sigmoid function. The model converged stably after 100 epochs with a batch size of 8 and an initial learning rate of 1e-5. Learning rate halving and L2 regularization (weight decay: 0.01) were applied to prevent overfitting.

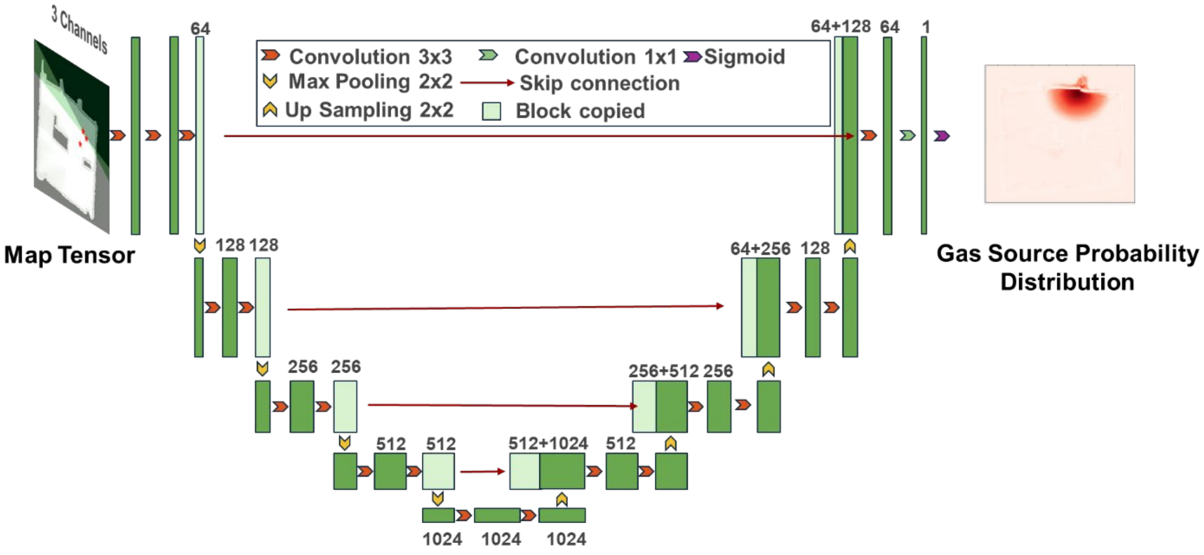

Figure 3.7 Architecture of the U-net based model. Input is 3-channel map tensor and output is source probability distribution map.

To enhance the stability and accuracy of gas source localization training, a specialized loss function incorporating a dynamic detection range (D_N) was designed. This approach allows the model to adaptively adjust its focus based on the number of plume encounters, thereby improving localization performance over time. The model outputs a posterior source map, a probabilistic grid where each cell represents the likelihood of containing the gas source. However, confidence in these predictions can vary depending on factors such as the number of plume encounters and the completeness of the SLAM-generated map. To accommodate this variability, the dynamic detection range D_N is defined as a function of the number of plume encounters N (Figure 3.8):

$$D_N = a \cdot e^{-k \cdot (N-1)} + c \qquad (5)$$

Here, parameter a is set to 3, and c is set to 1, meaning that the dynamic detecting range D_N starts at 4 meters after the first encounter and gradually converges to a 1 meter range as the number of encounters increases. k, representing the convergence rate, is set to 0.3. This adaptive range ensures that the model focuses on increasingly smaller and more precise regions as more data is collected, improving localization accuracy over time.

The hyperparameters are set following a design principle that balances early-stage reachability and late-stage precision. Specifically, c=1 defines the intended terminal tolerance: by the end of localization, the detection range is expected to shrink to approximately 1m, matching our target that the robot should confine the candidate region within a 1 m neighborhood of the source. In the experiments, the robot encounters the plume only a limited number of times before termination, the average number of plume encounters from start to finish is about four. We therefore calibrate the D_N schedule such that the learning objective encourages effective convergence within roughly four informative updates. The remaining parameters a and k control the initial scale and contraction rate and are chosen to stabilize the initial phase under occlusions. Because obstacles can block line-of-sight to the source, the robot often cannot directly observe the source location at initialization. If the initial D_N were entirely outside the robot's feasible scanning coverage, the policy would be forced into near-random proposals, leading to unnecessary training variance. Consequently, we select a and k so that the initial D_N is just large enough to reach the robot's effective scanning range at the beginning, while still contracting smoothly enough over subsequent steps.

The training loss is designed based on Binary Cross-Entropy[118] (BCE), a commonly used function for binary classification tasks. Instead of focusing on exact source coordinates, the model aims to catch the relevance between the source and the region inside D_N. Standard BCE, given by the following equation, is not well suited for the scenario due to the high variability in gas plume encounters and the need to handle spatial uncertainty.

$$\mathcal{L}_{\mathcal{BCE}}(y_i, \widehat{y_i}) = y_i \log(\widehat{y_i}) + (1 - y_i) \log(1 - \widehat{y_i}) \tag{6}$$

To address this, this thesis modifies the BCE by introducing a weight that reduces the influence of grid cells near the edges of the detection range, where predictions are less reliable. The proposed modification is critical since manually defined regions are not absolutely accurate, and the model may struggle to align precisely at the edges, leading to instability during training.

To ensure smoother training, it is necessary to reduce the weights at the edges of the predicted region. Each grid is assigned a weight according to:

$$y_{N,i} = \begin{cases} 0, & if\ d_i > D_N\ or\ m_i = -1 \\ \frac{D_N - d_i}{D_N} \cdot \frac{100 - m_i}{100}, & if\ d_i \leq D_N \end{cases} \tag{7}$$

in which d_i is the distance between grid i and the gas source location, m_i is the occupancy probability grid i. Grids outside D_N range of the gas source are labelled as 0, while grids inside D_N range are labelled as $\frac{(\mathrm{D_N - d_i})}{\mathrm{D_N}} \cdot \frac{(\mathrm{100 - m_i})}{100}$, considering both distance to gas source and occupancy probability of the map. The final training loss is defined by:

$$\mathcal{L} = -\frac{1}{n}\sum_{i=1}^{n} l(y_i, \widehat{y_i}) \tag{8}$$

$$l(y_i, \widehat{y_i}) = \begin{cases} \mathcal{L}_{\mathcal{BCE}}(y_i, 0), if \widehat{y_i} = 0 \\ \widehat{y_i}^{\frac{1}{5}} \cdot \mathcal{L}_{\mathcal{BCE}}(y_i, 1) \cdot W_{pos}, if \widehat{y_i} > 0 \end{cases} \tag{9}$$

The loss function $\mathcal{L}$ first computes the binary classification loss for the region in the range D_N, and then adjusts it using the decreased weight $\widehat{y_i}^{\frac{1}{5}}$ and a positive weighting factor W_{pos}, $\widehat{y_i}^{\frac{1}{5}}$ reduces the weight of predictions near the edges of the detection range, ensuring smoother training and preventing overfitting to noisy data. W_{pos} (set to 5.0) assigns a higher weight to the positive sample in case of imbalanced sample distribution, which ensures that the model pays sufficient attention to the rare but significant positive samples.

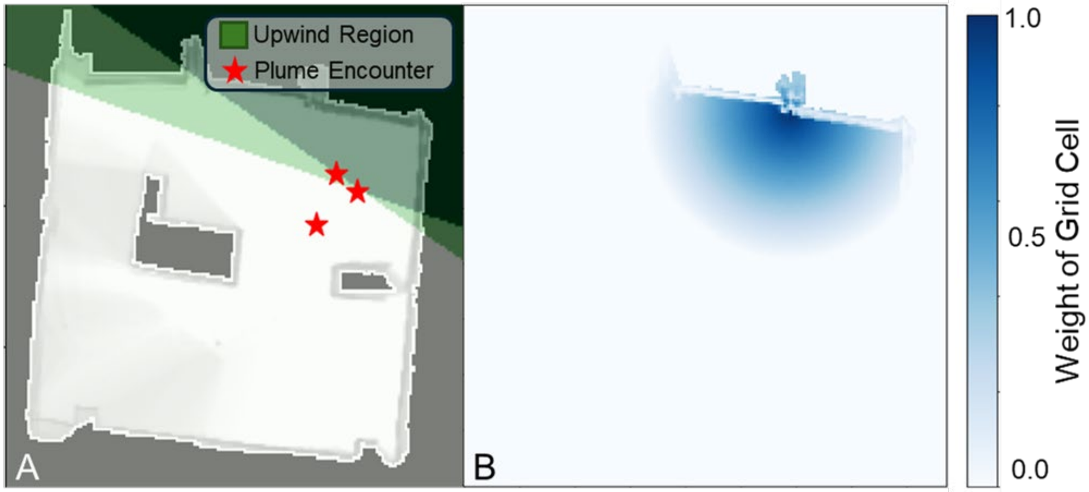

Figure 3.8 Dynamic detection region in loss function. Under the scenario (A): the detection region is shown in (B). The blue area is the region for loss calculation, which will shrink with the number of plume encounters. Grids near the edge are assigned smaller weights to avoid unstable loss gradient.

3.3 Gas Distribution Mapping

3.3.1 Dataset Generation and Selection

Computational Fluid Dynamics[119] (CFD) is a branch of fluid mechanics that uses numerical methods and algorithms to simulate and analyze fluid flow and related physical phenomena. The fundamental principle of CFD lies in solving the governing equations of fluid dynamics, which is the Navier-Stokes equations along with conservation laws for mass, momentum, and energy. These equations are discretized using methods such as finite volume, finite difference, or finite element methods, allowing the fluid domain to be represented in a computational mesh or grid. CFD simulations provide insights into velocity fields, pressure distributions, turbulence, heat transfer, and species transport, giving valuable simulation to fluid fields distribution.

OpenFOAM[120] is a widely used open-source CFD toolbox that provides a flexible and customizable platform for modeling fluid flow, turbulence, heat transfer, and chemical reactions.

OpenFOAM features a modular architecture that allows users to modify solvers, define custom boundary conditions, and couple multiple physical models. Its capabilities range from simple laminar flows to complex multiphase, reactive, and compressible flows. OpenFOAM is favored in both academia and industry due to its extensibility, active community support, and the absence of licensing costs.

In order to generate physically reliable and diverse training datasets for data-driven gas distribution reconstruction, a comprehensive series of two-dimensional CFD simulations were systematically performed. These simulations aim to replicate realistic turbulent airflow behavior and the subsequent transport and dispersion of passive scalars, specifically gas species, within a controlled computational domain. The primary objective is to capture the intricate interplay between turbulence-induced mixing and scalar transport processes, which are crucial for reconstructing accurate gas concentration fields from sparse sensing data.

The fluid flow is modeled as incompressible, which is a valid approximation under low-speed conditions typically encountered in indoor or urban dispersion scenarios[121]. To represent the turbulent nature of the flow, the Reynolds-averaged Navier-Stokes (RANS) equations are solved. This approach effectively separates the instantaneous velocity and pressure fields into mean and fluctuating components, enabling statistical treatment of turbulence effects. To close the system, the widely validated standard $k - \varepsilon$ turbulence model is employed, which provides transport equations for turbulent kinetic energy and its dissipation rate, thereby yielding an eddy viscosity that models the enhanced momentum transfer due to turbulence.

The transport of the gas concentration field is modeled as a passive scalar, meaning that it does not affect the flow field dynamically and is simply carried and diffused by the underlying velocity and turbulent eddies. The computational domain is a two-dimensional square region of 1 meter by 1 meter, discretized using a uniform mesh of 100 times 100 cells. This mesh resolution is chosen to balance computational efficiency and sufficient spatial detail to resolve key flow and scalar features.

The governing equations are the Reynolds-Averaged Navier-Stokes[122] (RANS) equations with the $k - \varepsilon$ turbulence model:

$$\nabla \cdot \vec{u} = 0 \tag{10}$$

$$\rho\left(\frac{\partial \vec{\mathrm{u}}}{\partial \mathrm{t}} + \vec{\mathrm{u}} \cdot \nabla \vec{\mathrm{u}}\right) = -\nabla \mathrm{p} + \mu \nabla^2 \vec{\mathrm{u}} \tag{11}$$

where $\vec{u} = (u, v)$ denotes the velocity vector field components in the Cartesian directions, p represents the fluid pressure, ρ is the fluid density (assumed constant for incompressibility), and μ is the dynamic viscosity. The continuity equation mathematically enforces the physical constraint that the fluid is incompressible, i.e., its volume does not change with time. The momentum equation balances the rate of change of momentum with forces arising from pressure gradients and viscous stresses. The nonlinear convective term $\vec{u} \cdot \nabla \vec{u}$ represents the advection of momentum by the flow itself, which is responsible for the complex behavior of fluid motion including turbulence generation.

In turbulent flows, velocity and pressure fields exhibit rapid fluctuations over time. To enable practical computation without resolving all turbulent scales, Reynolds decomposition is applied, decomposing instantaneous quantities into mean (time-averaged) and fluctuating components:

$$\vec{u} = \bar{\vec{u}} + \overrightarrow{u'} \tag{12}$$

$$p = \bar{p} + p' \tag{13}$$

where $\bar{\vec{u}}$ and $\bar{p}$ denote the mean velocity and pressure fields, respectively, and $\overrightarrow{u'}$ and p' are the zero-mean fluctuations.

Substituting these into the original Navier-Stokes equations and taking time-averages leads to the RANS equations:

$$\nabla \cdot \bar{\vec{u}} = 0 \tag{14}$$

$$\rho\left(\bar{\vec{u}} \cdot \nabla \bar{\vec{u}}\right) = -\nabla \bar{p} + \mu \nabla^2 \bar{\vec{u}} - \rho \nabla \cdot \overline{\vec{u'} \otimes \vec{u'}} \tag{15}$$

where the additional term $-\rho \nabla \cdot \overline{\vec{u'} \otimes \vec{u'}}$ represents the Reynolds stress tensor, embodying the momentum flux due to turbulent velocity fluctuations. This term is unknown a priori and necessitates turbulence modeling for closure.

Physically, Reynolds stresses act as an additional stress on the mean flow, enhancing mixing and momentum transport beyond molecular viscosity effects. To close the RANS system, the Boussinesq hypothesis[123] is introduced, which postulates that the Reynolds stresses can be modeled analogously to viscous stresses via an eddy viscosity μ_t. The expression is:

$$-\rho \overline{u_i' u_j'} = \mu_t \left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right) - \frac{2}{3} \rho k \delta_{ij} \tag{16}$$

where indices i, j correspond to spatial directions, k is the turbulent kinetic energy defined as $k = \frac{1}{2} \overline{u_i' u_i'}$, and δ_{ij} is the Kronecker delta symbol.

This approximation translates the unknown turbulent stresses into a scalar eddy viscosity multiplying the strain rate tensor, which significantly simplifies modeling while retaining the essential physics of enhanced momentum transfer caused by turbulence. Inserting this into RANS momentum yields the momentum equation in terms of the sum of molecular and turbulent viscosities:

$$\rho\left(\bar{\vec{u}} \cdot \nabla\bar{\vec{u}}\right) = -\nabla\bar{p} + \nabla \cdot \left[(\mu + \mu_t)\left(\nabla\bar{\vec{u}} + \left(\nabla\bar{\vec{u}}\right)^T\right)\right] - \frac{2}{3}\rho\nabla k \tag{17}$$

where the last term accounts for isotropic turbulent normal stresses.

The eddy viscosity μ_t is not directly known and must be computed from turbulence quantities. The standard $k - \varepsilon$model provides transport equations for k (turbulent kinetic energy) and ε (its dissipation rate), capturing production, diffusion, and destruction mechanisms of turbulence:

$$\mu_t = C_\mu \rho \frac{k^2}{\varepsilon} \tag{18}$$

where $C_\mu$is an empirical constant.

The governing transport equations for $k\ and\ \varepsilon$ are

$$\frac{\partial(\rho k)}{\partial t} + \nabla \cdot \left(\rho k \bar{\vec{u}}\right) = \nabla \cdot \left[\left(\mu + \frac{\mu_t}{\sigma_k}\right)\nabla k\right] + P_k - \rho \tag{19}$$

$$\frac{\partial(\rho\varepsilon)}{\partial t} + \nabla \cdot \left(\rho\varepsilon\bar{\vec{u}}\right) = \nabla \cdot \left[\left(\mu + \frac{\mu_t}{\sigma_\varepsilon}\right)\nabla\varepsilon\right] + C_{\varepsilon 1}\frac{\varepsilon}{k}P_k - C_{\varepsilon 2}\rho\frac{\varepsilon^2}{k} \tag{20}$$

where P_k is the production term of turbulent kinetic energy due to mean velocity gradients:

$$P_k = \mu_t\left(\nabla\bar{\vec{u}} + \left(\nabla\bar{\vec{u}}\right)^T\right) : \nabla\bar{\vec{u}} \tag{21}$$

with the colon operator denoting the double inner product of tensors.

The constants:

$$C_{\mu} = 0.09, \sigma_{k} = 1.0, \sigma_{\varepsilon} = 1.3, C_{\varepsilon 1} = 1.44, C_{\varepsilon 2} = 1.92 \tag{22}$$

are standard values obtained from extensive experimental validation. This turbulence model strikes a balance between accuracy and computational cost and is widely used for engineering flows, including indoor airflows and pollutant dispersion.

The dispersion of the gas is modeled as the transport of a passive scalar concentration c. Since the gas is assumed not to influence the flow field (no buoyancy or chemical reactions), its transport is governed by the convection-diffusion equation:

$$\frac{\partial c}{\partial t} + \nabla \cdot \left(c\bar{\vec{u}}\right) = \nabla \cdot (D\nabla c) \tag{23}$$

where D is the effective diffusion coefficient representing both molecular and turbulent diffusion mechanisms.

The turbulent diffusion is modeled by analogy to momentum diffusion, with the turbulent diffusivity proportional to the eddy viscosity and inversely proportional to the turbulent Schmidt number Sc_t:

$$D = D_m + \frac{\mu_t}{Sc_t \rho} \tag{24}$$

where D_m is the molecular diffusivity of the gas species in air, typically on the order of

$2.0 \times 10^{-5}\ \mathrm{m}^2/\mathrm{s}$. The turbulent Schmidt number Sc_t is an empirical constant set to 0.9, reflecting the ratio of momentum to mass diffusivity in turbulent flows. This formulation effectively captures the enhanced mixing and spreading of the gas plume due to turbulence, which is critical for generating realistic spatial concentration patterns for downstream data-driven analysis.

The finite volume method (FVM) discretizes the governing equations on the structured mesh. The following numerical schemes are employed:

Spatial discretization: Second-order central differencing for diffusion terms and a bounded second-order upwind scheme for convection terms to balance accuracy and stability.

Temporal discretization: Implicit backward Euler scheme for time integration, ensuring unconditional stability.

Pressure-velocity coupling: SIMPLE algorithm is used to iteratively solve the pressure and velocity fields until convergence at each time step.

Convergence criteria: Residuals for all equations are monitored, with a typical convergence tolerance set at 10^{-6}.

Grid independence studies were performed by comparing results on finer meshes $(e.g., 150\ \times 150\ and\ 200\ \times 200)$. No significant differences were observed in velocity profiles and concentration fields, confirming that the $100\ \times 100$ mesh is sufficient. Time step size was chosen as $\Delta\ t = 0.001$ s based on a Courant-Friedrichs-Lewy (CFL) number less than 1 to ensure numerical stability and temporal accuracy.

Two orthogonal velocity inlets are placed on the left and bottom boundaries of the domain. For each simulation case, inflow velocities at these inlets are sampled independently from a uniform distribution $\mathcal{U}(0.1,0.5)\sim\mathrm{m}/s$, introducing variability in overall flow direction and strength. The top and right domain boundaries are set as zero-gradient pressure outlets to allow natural outflow.

A single passive gas source is introduced within the central region of the domain. Its exact location is randomized for each case, mimicking scenarios where the gas leak origin is unknown. The gas is treated as a passive scalar and its emission rate is fixed across all simulations.

Table 3.1 Boundary conditions of CFD simulation.

Boundary	Type	Details
Left	Velocity Inlet	$u \sim U(0.1,\ 0.5)\ \mathrm{m/s}$
Bottom	Velocity Inlet	$u \sim U(0.1,\ 0.5)\ \mathrm{m/s}$
Right	Pressure Outlet	Zero gradient
Top	Pressure Outlet	Zero gradient
Interior	Gas Source	Random position & intensity

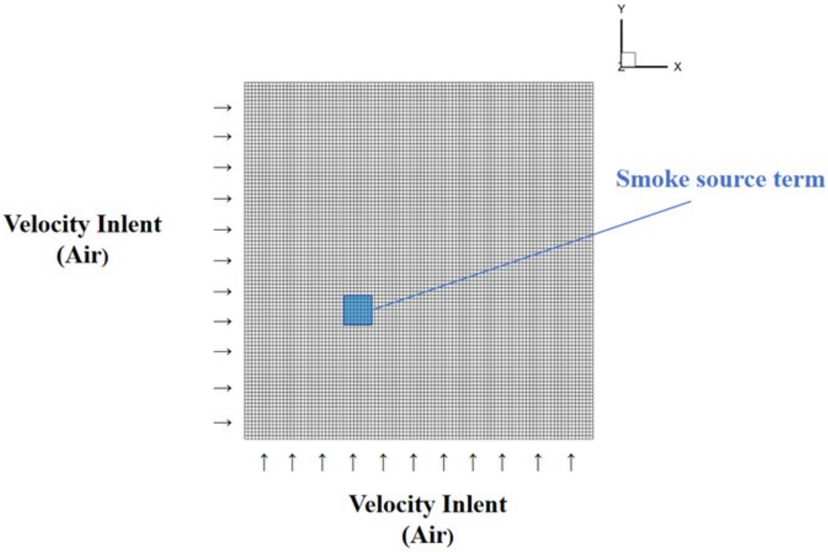

Figure 3.9 Setup and boundary conditions of CFD simulation.

Once steady-state is reached, each simulated sample is down sampled through interpolation into a uniform 64×64 grid, which serves as the ground truth (Figure 3.10). N sample points were randomly selected from the ground truth (N = 20, 30, 50) as sensor sampling. The sampled data includes (x, y) position, local gas concentration, and velocity components (u, v). The training set consists of 200 simulated cases. For each case, 32 random sampling trials are conducted to simulate different sensor input scenarios. During sampling, the gas concentration is used to weight the probability density, allowing the simulation of various unbalanced sampling conditions, such as higher, lower, or uniform sampling densities in the gas plume region. The test set is composed of 20 additional cases, which are down sampled using the same method as the training set.

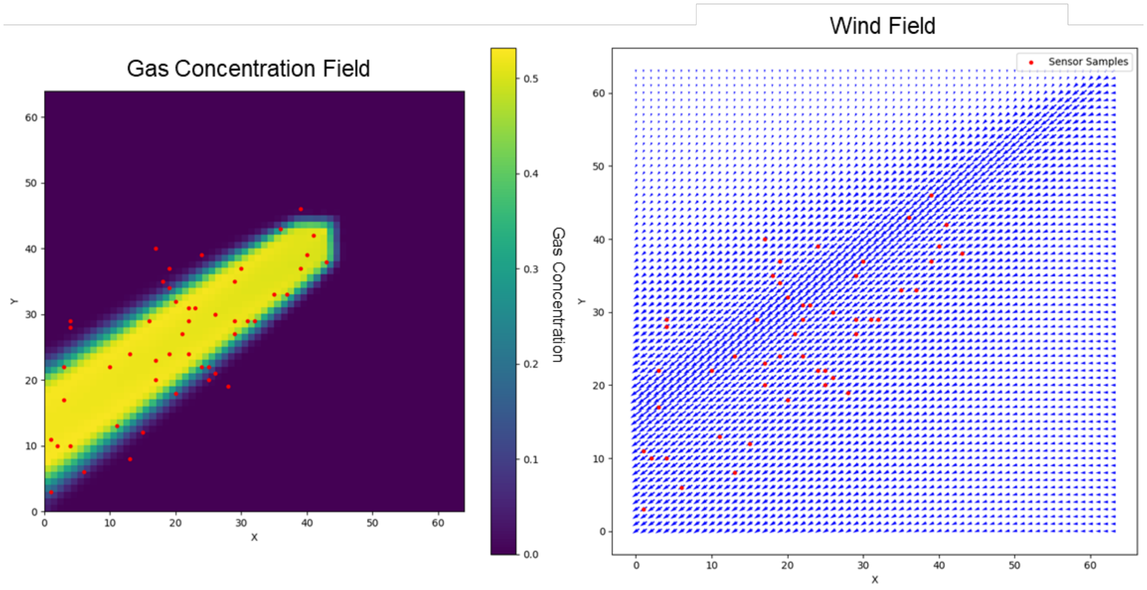

Figure 3.10 Sample of CFD simulation results. The simulation is under open environment, with a single gas source diffusing within a wind field, then down sampled to 64*64 grids. Sensor inputs is randomly sampled from the results.

The model's generalization ability and robustness are validated using real-world data. The validation of model is based on the dataset of ICASSP 2025 SP Grand Challenge: Gas source localization from real-world spatial in-situ concentration and wind measurements[124]. The dataset is collected from a low-speed wind tunnel within a test area measuring 2.40m x 2.40m x 0.88m (Figure 3.11A). The test area contains a factory model and a gas source releasing propylene glycol vapor. A robotic arm is used to measure horizontal wind speed and gas concentration across a total of 20 × 20 horizontal cells (in the XY-plane) and at 3 different vertical positions (Figure 3.11B). This dataset represents the spatial distribution of gas concentration in a realistic environment. The data is divided into different working planes based on vertical positions and interpolated into 64×64 distributions. Sensor inputs are then sampled following the same rules used in the training set. This setup is used to evaluate the extrapolation capability of our model in real-world scenarios (Figure 3.11C).

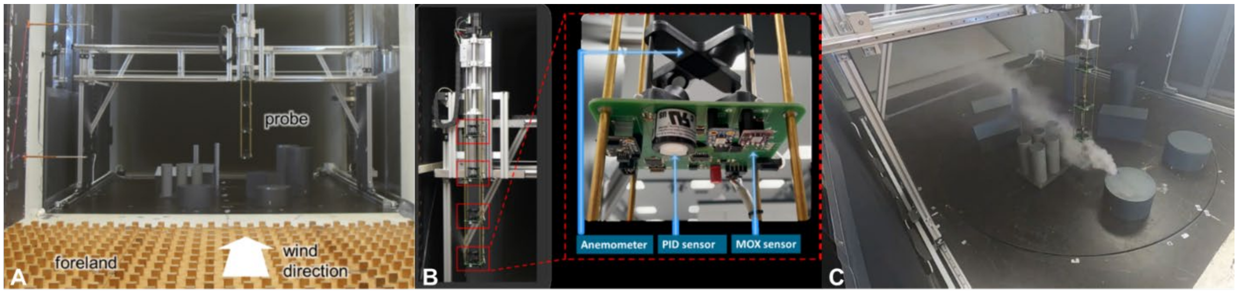

Figure 3.11 Setup of real-world validation experiment, reproduced with permission from [124]. (A) Testing area in the wind tunnel and the measurement probe. (B) Sensing probe with PID, MOX and wind sensors within each package. (C) A snapshot of the measurement process with an active gas source.

3.3.2 Model Architecture and Training

A Transformer-based model was designed to reconstruct gas concentration and wind fields from random sensor observations (Figure 3.12). Each input sample includes spatial coordinates (x, y), wind speed components (u, v), and gas concentration (c). The model uses two separate encoders: a Dynamic Positional Encoder and a Feature Encoder. The positional encoder applies Fourier-based encoding to spatial coordinates, followed by an MLP to generate 128-dimensional embeddings. This encoding captures high-frequency spatial patterns and enables generalization across different layouts. The feature encoder processes the (u, v, c) values through a two-layer MLP with GELU (Gaussian Error Linear Unit) activation and dropout, producing a feature embedding of the same dimension. The position and feature embeddings are projected into a shared latent space and fused using multi-head self-attention. A residual connection and LayerNorm are applied to stabilize training and preserve spatial structure. The fused embeddings are processed through a two-layer Transformer encoder. To predict the full grid, we construct a learnable decoder query by combining fixed positional encodings of the target grid points and a trainable latent query. This query is passed through a two-layer Transformer decoder, attending to the encoder memory to generate a global feature representation. The decoder output is fed into three parallel MLP heads to regress the u, v, and c fields over the entire grid. Each head independently predicts one of the target variables and reshapes the output into a 64×64 spatial map.

L2 regularization[125], which adds a penalty term proportional to the square of the model weights, is applied to prevent overfitting by discouraging overly complex models. This thesis adopts the OneCycleLR[126] strategy, a dynamic learning rate schedule that initially increases and then decreases the learning rate—to accelerate convergence and improve training stability. Additionally, curriculum learning[127] is employed to enhance training efficiency and robustness; this strategy, inspired by the way humans learn, involves presenting the model with easier tasks first before gradually introducing more difficult ones. Specifically, the model was initially trained on samples with 50 sensor inputs and then progressively adapted to more challenging cases with 30 and 20 inputs.

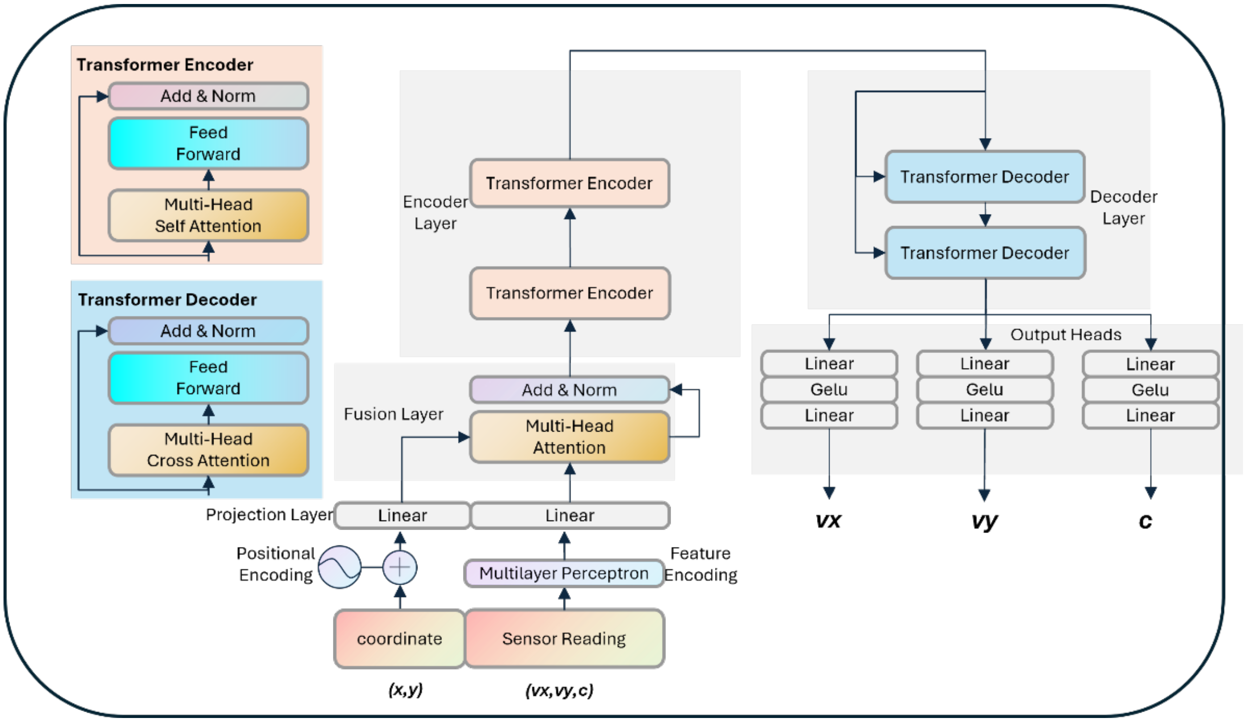

Figure 3.12 Architecture of the Transformer-based gas distribution mapping model.

Chapter 4*

Topography Aware Gas Source Localization

Gas source localization in complex environments is critical for applications such as environmental monitoring, industrial safety, and disaster response. Traditional methods often struggle with the challenges posed by a lack of environmental topography integration, especially when interactions between wind and obstacles distort gas dispersion patterns. This chapter proposes a deep learning-based approach, which leverages spatial context and environmental mapping to enhance gas source localization. Based on PID and model gas equivalent to Chapter 4, by integrating Simultaneous Localization and Mapping (SLAM) with a U-Net-based model, the proposed method predicts the likelihood of gas source locations by analyzing gas sensor data, wind flow, and topography of the environment represented by a 2D occupancy map. The result demonstrates the efficacy of our approach using a wheeled robot equipped with a photoionization detector, a LIDAR, and an anemometer, in various scenarios with dynamic wind fields and multiple obstacles. The results show that this approach can robustly locate gas sources, even in challenging environments with fluctuating wind directions, outperforming conventional methods by utilizing topography contextual information. This study underscores the importance of topographical context in gas

---

*This Chapter is published as Changhao Tian, Annan Wang, Han Fan, Thomas Wiedemann, Yifei Luo, Le Yang, Weisi Lin, Achim J. Lilienthal, Xiaodong Chen. Deep Learning Based Topography Aware Gas Source Localization with Mobile Robot. International Conference on Robotics and Automation International Conference on Robotics and Automation (ICRA) 2025. IEEE, 2025: 4380-4386, no permission is required for thesis purpose.

source localization and offers a flexible and robust solution for real-world applications. Data and code are publicly available.

4.1 Introduction

In industrial and environmental contexts, gas leakage represents a significant yet often undetected threat. In different scenarios, from environmental monitoring, industrial safety, and disaster rescue, dispersed gas is always an important but typically invisible environmental factor. Detecting and locating potential gas sources is crucial for ensuring environmental compliance, preventing industrial accidents, and supporting rescue operations. Therefore, using mobile robots committing gas source localization in hazardous or hard-to-reach environments becomes appealing.

A typical robot for gas source localization typically uses a gas sensor to detect the gas from the source and an anemometer to estimate the direction of the source. However, the primary challenge arises from the complexity of gas dispersion processes. Firstly, turbulence can arise at the interface between the gas and the atmosphere due to viscosity. This unpredictable fluid dynamic effect disrupts the concentration gradient, rendering gas concentration an unreliable indicator for estimating the distance to the source. Besides, convective gas flow interacts with obstacles, causing the wind field to become distorted by the surrounding topography, making it difficult to infer source direction solely from an anemometer (Figure 4.1). As a result, gas source localization in the real-world is often a non-observable estimation problem, particularly in indoor environments where the presence of obstacles further complicates airflow patterns.

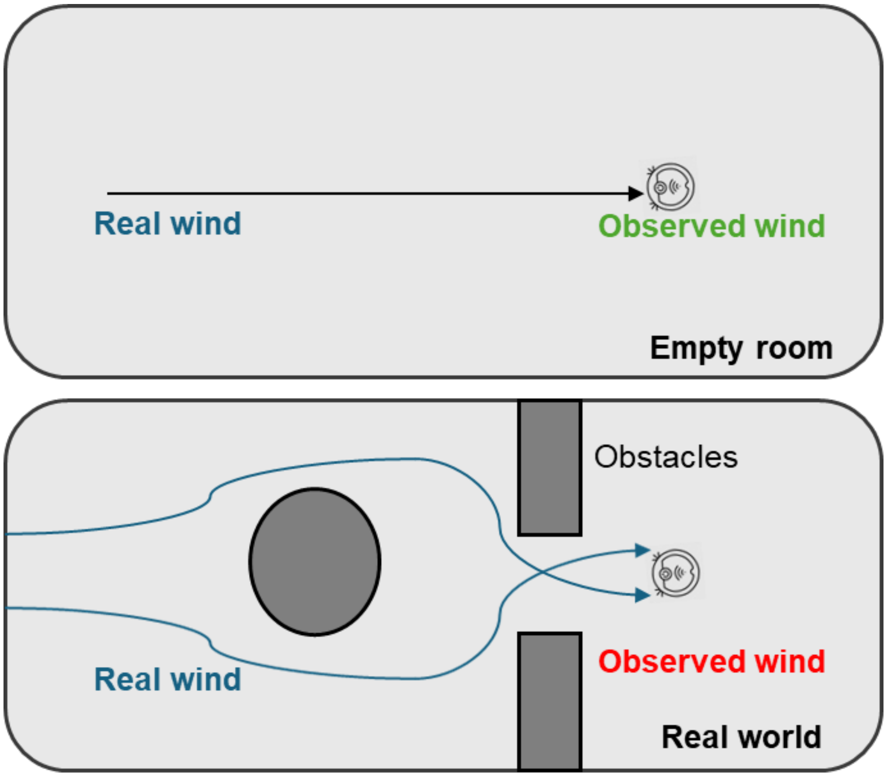

Figure 4.1 Schematic illustration of gas source localization. In real-world scenarios with obstacles, the interaction between airflow and physical obstructions further complicates the problem. Unlike empty room, in a real-world scenario, real wind cannot be inferred from observed wind.

To address this obstacle-induced challenge, this chapter leveraged spatial context, such as the environmental topography and observed wind flow, to learn the complex, non-analytical relationship between gas encounters and the actual gas source location. The contextual information in environmental topology can be perceived by SLAM, which allows robots to create or update maps of unknown environments while tracking their location. For gas source localization, SLAM can provide accurate locations where gas is detected and corresponding to surrounding topography[128]. Assuming that obstacle layout affects the gas dispersal process and, therefore, is crucial for predicting gas source locations, integrating SLAM could greatly enhance the localization process.

This chapter proposes a deep learning-based approach for gas source localization in complex environments. The goal of our work is to train a model that can predict the potential gas source location during the SLAM process when the robot encounters the plume, leveraging data from a LIDAR, an anemometer, and a gas sensor. The environment map, gas encounter locations and wind observations will be combined into a 3-channel tensor and put into a U-Net, which can capture the topography context related to gas dispersion. When the robot is deployed to an unknown environment, the model can predict the potential locations of gas sources despite a different topography. To implement and demonstrate our approach, a wheeled mobile robot equipped with a photoionization detector (PID), an anemometer, and a LIDAR is built and conducted real-world experiments. The results have shown that the method can locate the gas source effectively in an environment with multiple obstacles and complex gas plume distribution. Even under a changing wind field, the method can remain robust.

4.2 Experimental Method

4.2.1 Design of Robot Platform

The robotic platform is built on a Mecanum wheel chassis[129], shown by Figure 4.2A. It is equipped with a CZ0001 ultrasonic anemometer, an LDS-250D laser LIDAR, and a PID-AH2 gas sensor. The CZ0001 ultrasonic anemometer allows collection of wind data in the speed range of 0 to 40 m/s with an angular accuracy of ±2 degree, a speed accuracy of ±2% and a speed threshold of 0.05 m/s. The LDS-25D can generate point clouds of up to 14000 points per second with a 30 m range with an accuracy of ±0.01m at 14k Hz. The PID-AH2 is a highly sensitive photoionization detector (PID). A PID uses an ultraviolet (UV) lamp to ionize a detectable compound, including most Volatile Organic Compounds (VOCs) In our experiment, the PID-AH2 has a detection range from sub 1 parts per billion (ppb) to 50 parts per million (ppm). All the sensor payloads are controlled by an NVIDIA Jetson Nano processor. The robot's movement can be controlled via a remote controller. The SLAM functionality was implemented using the Cartographer ROS package[130].

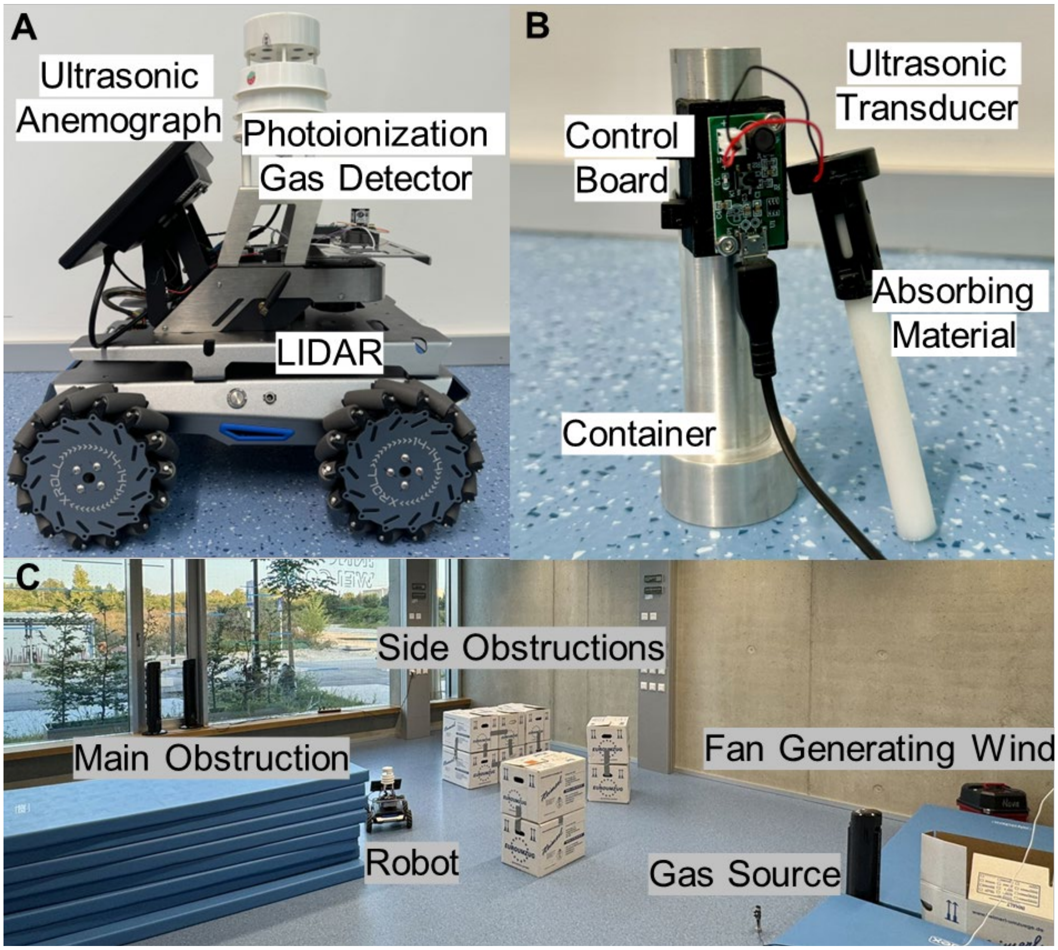

Figure 4.2 The experiment setup. (A) The Robot platform carrying sensor payloads. (B) The gas source generating ethanol vapor. (C) The experiment area with moveable obstructions and wind directions.

4.2.2 Target Gas Source and Area

The target gas is generated by an ultrasonic ethanol vaporizer (Figure 4.2B). Ethanol liquid in the container contacts an ultrasonic transducer through an absorbent material, producing ethanol vapor at room temperature. Near the ground surface, the gas can remain at a relatively constant altitude while being dispersed by wind currents without significant vertical floating via temperature gradient.

The experimental setup, depicted in (Figure 4.2C), consisted of an empty room measuring 10 meters by 10 meters. To facilitate gas dispersion, an artificial wind field was generated using three fans. A large central obstacle was introduced to alter the overall wind flow pattern, while smaller side obstacles were placed to influence the finer details of the resulting gas plume shape. All obstacles were higher than the fans to prevent airflow from bypassing them. The experiment was conducted under various conditions, including different fan configurations, gas source locations, obstacle arrangements, and initial robot positions

4.3 Principal Outcomes

4.3.1 Model Training

Based on dataset of 60 trials, each containing 4 to 6 samples, the model converged after 100 training epochs. The data were split into training and validation sets at a 4:1 ratio, with validation loss monitored to prevent overfitting. Specifically, each trial was treated as a unit during the split to ensure that samples from the same trial were not divided across different sets. This precaution was taken to avoid data leakage, where scenes encountered during training might reappear in the validation set. The loss curves (Figure 4.3) show that both the training and validation losses stabilized after 100 epochs, with no significant increase in validation loss. This indicates that the model converged well without exhibiting notable overfitting. The results also suggest that the customized loss function effectively reduced training instability.

Figure 4.3 Loss curve of the training process. The training converged after 100 epochs.

4.3.2 Gas Source Localization with Spatial Context

During testing, the proposed method achieved accurate gas source localization within as few as four plume encounters, as illustrated in Figure 4.4. This performance is primarily attributed to two critical factors: the availability of a more complete environmental map and the increased number of spatially distributed plume encounter locations.

Each time the robot encounters a gas plume, the model updates its inference by generating a probability distribution over the explored space. The shape of probability distribution map is aligned with the map from SLAM, the value of each grid means the grid's probability of containing gas source, indicating where the gas source is most likely to be located. As the robot continues to explore and gradually scans the surroundings of the gas source, progressing from Figure 4.4 A to Figure 4.4 C, the inferred probability distribution becomes more continuous and spatially coherent, reflecting a higher level of confidence in the source's estimated position.

For the first encounter (Figure 4.4A), the prediction (Figure 4.4B) is with significant difference from the real gas source location (green triangle in Figure 4.4G). This is because the spatial relationships between encounters have not yet been established, and the location of the gas source has not been identified by LiDAR, as is grey in Figure 4.4A. With each additional encounter, the distribution begins to converge, as the regions of high probability become increasingly focused and start to overlap around the actual source location. This convergence indicates that the model is accumulating sufficient spatial and sensory context to refine its prediction. Once a critical number of plume encounters has been reached, the aggregated information allows the model to accurately narrow down the probable area of the gas source, thus enabling reliable localization.

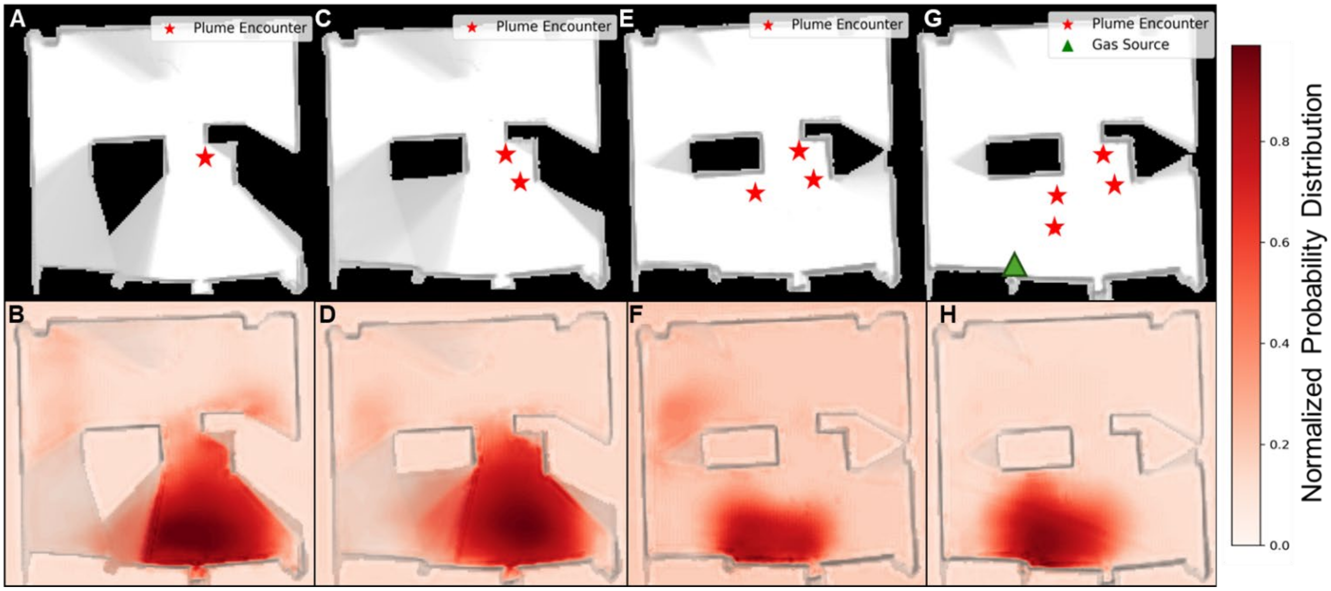

Figure 4.4 Model performance in one trial with increasing encounters and map completeness. The model's output demonstrates improved prediction accuracy and continuity as more gas encounters are acquired (A), (C), (E) and (G) and the coverage of the source posterior map increases (B), (D), (F) and (H). When 4 encounters were acquired (G), the grids assigned with the highest posteriors in (H) align closely with the actual source.

4.3.3 Quantified Evaluation of Model Performance

To evaluate the model's performance, how its output evolved with the number of plume encounters in the testing set is evaluated. Two quantitative indices were used, the ratio of positive grids and the distance from the probability distribution centroid to the actual gas source, reflecting the model's confidence and accuracy in localizing the gas source as new information is gathered.

The ratio of positive grids is defined as the proportion of grid cells whose predicted probability exceeds a threshold of 0.5, indicating it's likely to contain the gas source (Figure 4.5). This metric serves as a proxy for the model's output belief, a high ratio indicates that the model is broadly uncertain about the output, assigning elevated probabilities across a large area. Conversely, a lower ratio suggests that the model has become more selective and confident, focusing its probability mass on a smaller, more plausible region. As the number of plume encounters increases, this ratio decreases notably. This trend implies that the model progressively narrows its hypothesis space, reducing ambiguity and converging toward a smaller region that likely contains the true source.

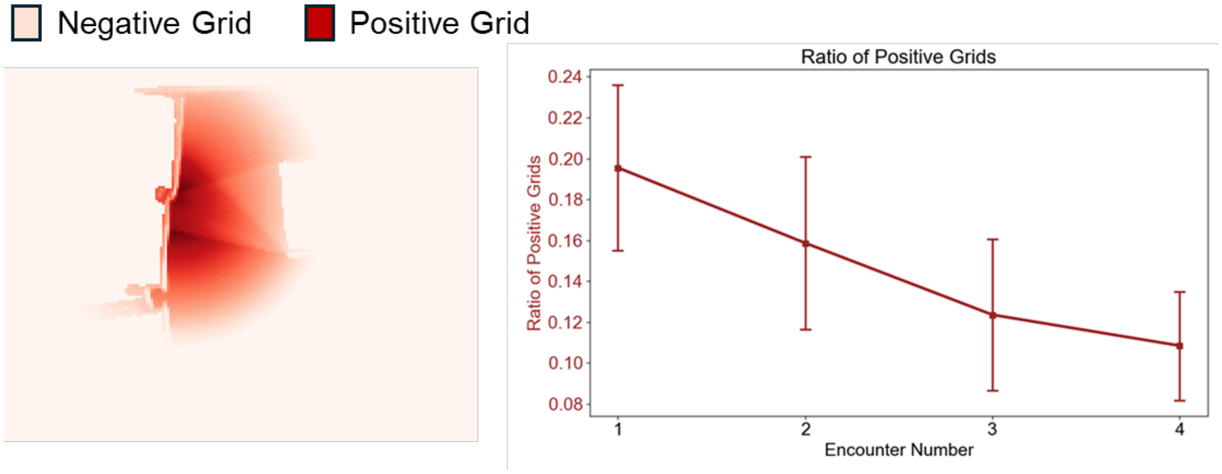

Figure 4.5 Schematic of the ratio of positive grids and its trend with respect to the number of encounters. As the number of encounters increases, the ratio gradually decreases, indicating an

increase in prediction belief.

The centroid-based distance measures the Euclidean distance between the predicted probability weighted centroid and the actual gas source location. This index evaluates localization accuracy, a large distance reflects a diffuse or misaligned prediction, while a smaller distance indicates that the model's probability mass is concentrated near the true source. As shown in Figure 4.6, after the first encounter, both the mean and variance of this distance are large, reflecting high uncertainty and low accuracy. However, with additional plume encounters, the centroid-based distance steadily declines, suggesting that the model integrates spatial and wind context effectively over time. This reduction underscores the role of contextual information in enhancing prediction precision, as the model is able to integrate sensory data with topographical and wind cues to localize the gas source more accurately.

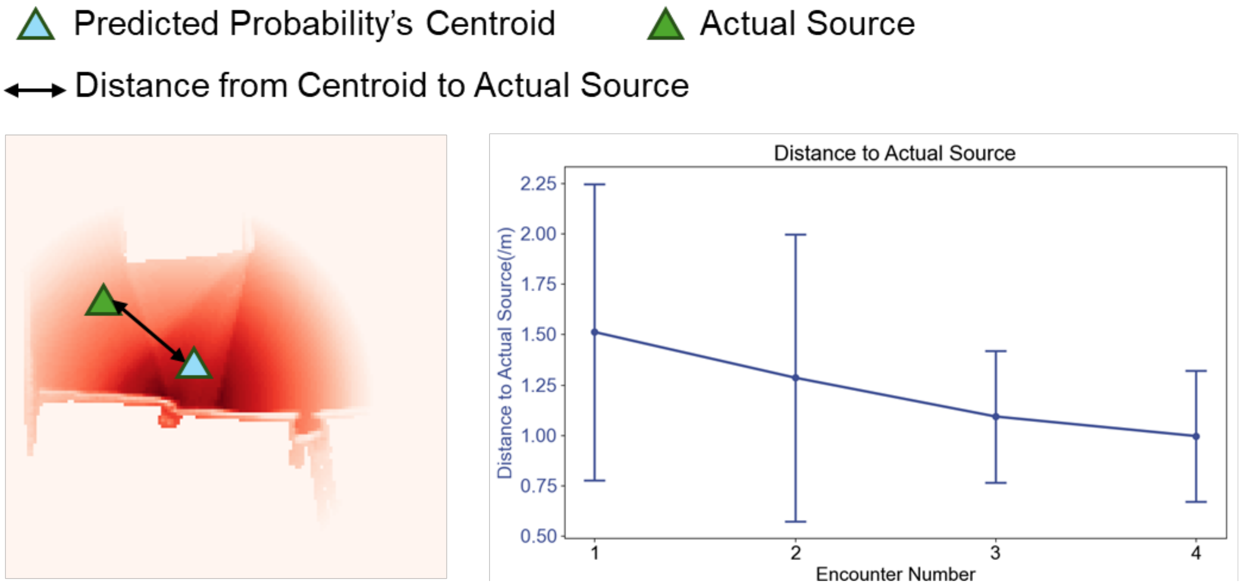

Figure 4.6 Schematic of the centroid-based distance and its trend with respect to the number of encounters. As the number of encounters increases, the ratio gradually decreases, indicating an increase in prediction accuracy.

It's worth noting that variances of both indices also decrease as the number of plume encounters increases. This indicates that the model's predictions become not only more accurate but also more consistent across different trials. The shrinking variance suggests that with more encounters, the model converges toward a stable inference pattern, reducing random fluctuations in its estimated source location. Such convergence reflects the model's ability to leverage accumulated spatial and wind-related context, rather than relying on isolated or noisy observations.

4.3.4 Robustness in Dynamic Environment

To further evaluate the robustness of our method, a set of tests under dynamic wind field conditions was conducted. In this setup, oscillating fans were used to generate a swinging wind pattern, introducing temporal variability and directional shifts in airflow that mimic more realistic environmental dynamics. As illustrated in Figure 4.7, the robot was able to accurately localize the gas source in both testing trials, despite the fluctuating wind directions.

It is important to note that the model was trained exclusively on data collected under static wind conditions, and that the onboard anemometer is limited to measuring local, instantaneous airflow at the robot's current position. This means the model had no prior exposure to dynamic airflow patterns and lacked direct access to a global wind field. Nevertheless, this approach maintained its ability to make reliable predictions based on the timing and location of plume encounters combined with the spatial context gathered through exploration.

This successful generalization to dynamic wind conditions highlights the central role of spatial context, such as obstacle layout, encounter locations, geometry among encounters, and accumulated sensory information in enabling robust gas source localization. Rather than relying on explicit wind field modeling, this method learns to infer the underlying structure of the environment through repeated plume interactions. The ability to operate effectively under such non-stationary conditions demonstrates the model's resilience and points to its potential for broader deployment in complex, uncontrolled real-world settings.

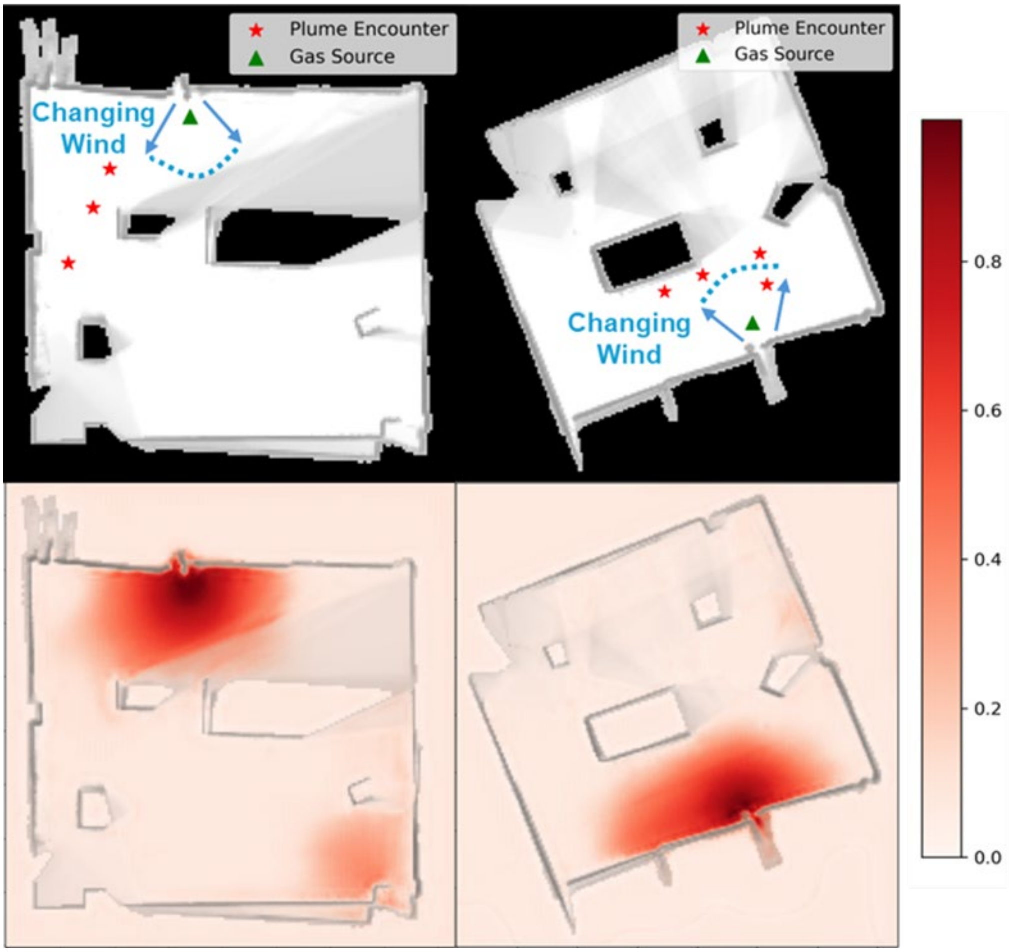

Figure 4.7 Demonstrations of test trials posed with changing wind field (shown by blue arrows). In dynamic environments that were never included in the training, the model remains robust in terms of prediction accuracy.

To complement the qualitative demonstrations under dynamic wind, a quantitative summary is provided by comparing static-wind and dynamic-wind evaluations under the same topography and experimental setting in the test set (Figure 4,8). Centroid-based distance and the ratio of positive grids are reported across different numbers of plume encounters, which are same metrics as analysis on the full testing set. The results indicate that dynamic wind leads to a slightly higher ratio of positive grids, suggesting that wind variability increases uncertainty in the inferred source region. However, the centroid-based distance exhibits a similar decreasing trend as encounter count increases, and the overall behavior remains aligned with the static-wind case. This quantitative evidence supports the robustness of GSL model under dynamic wind variations, indicating that the model learns interaction patterns between gas dispersion dynamics and terrain geometry, rather than merely memorizing the training data.

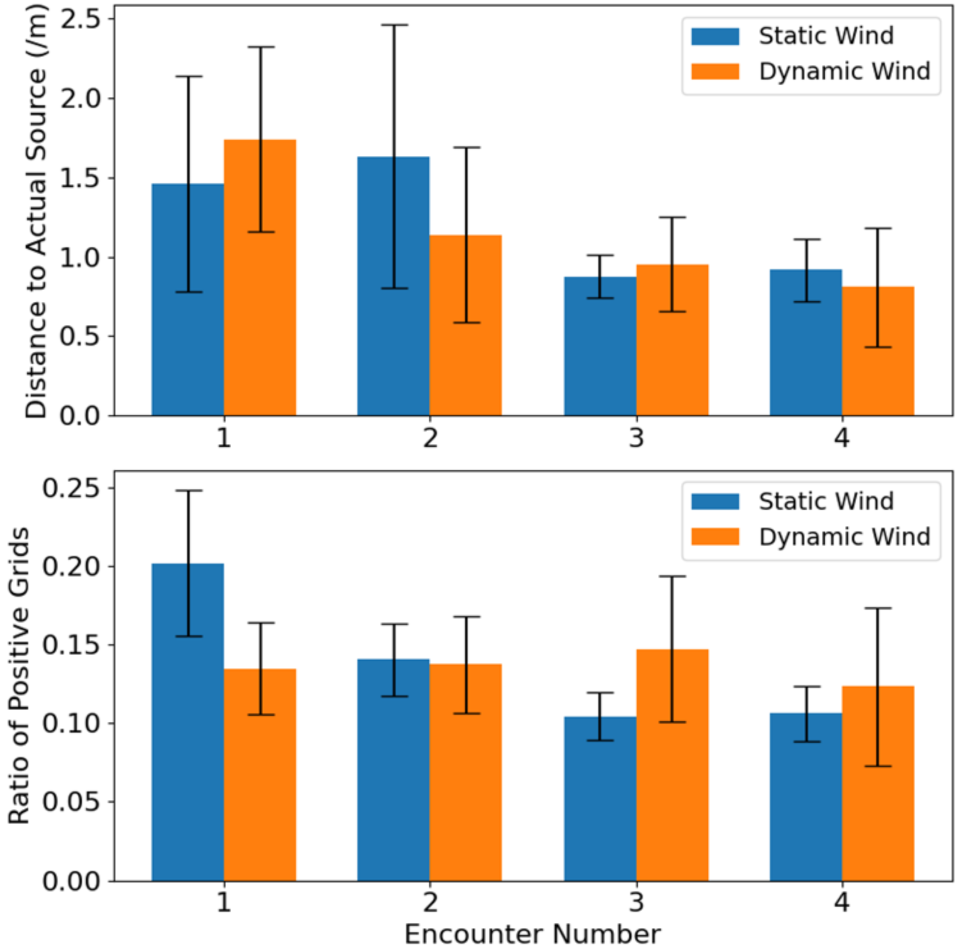

Figure 4.8 Quantitative comparison of GSL performance under static and dynamic wind conditions. The figure reports centroid-based distance between the predicted source distribution and the ground-truth source location and ratio of positive grids in the predicted result, evaluated on the test set under identical terrain and experimental settings. Results are grouped by the number of plume encounters. Dynamic wind yields a slightly higher ratio of positive grids, especially when prediction converges with multiple encounters, indicating a more diffuse support region (lower confidence), while the centroid-based distance and the overall convergence trend remain consistent with the static-wind case.

4.4 Conclusions

In this chapter, a topography aware gas source localization approach based on deep learning is proposed, which leverages environmental context to learn the spatial relationship between gas encounter events and the underlying source location. By incorporating topographical features perceived through SLAM and fusing them with sparse gas encounter signals, the proposed method enables probabilistic inference of the gas source location. Through real-world experiments conducted under both static and dynamic wind conditions, this chapter demonstrated that the proposed approach achieves robust and reliable performance. These findings suggest that environmental context-based prediction offers a promising direction for improving gas source localization in mobile robotics, particularly in complex or uncontrolled settings.

At present, the model utilizes a two-dimensional representation of the environment, which assumes that wind primarily flows around, rather than over, obstacles. While this simplification is reasonable in the most ground-level scenarios, it inherently limits the spatial resolution of gas dispersion modeling. To enable more comprehensive spatial perception of gas plumes, future work may incorporate 3D environmental representations using sensors such as 3D LiDAR or stereo cameras. These enhancements could allow the model to account for vertical airflows and volumetric plume structures, further improving localization accuracy in environments with significant vertical features.

Another important direction for future research is addressing the presence of multiple or dynamic gas sources. The current framework is developed and validated under the assumption of a single, stationary source. However, in many real-world applications, such as industrial leak detection or environmental monitoring, multiple or moving gas sources may coexist. Therefore, future work must explore strategies to disambiguate overlapping plumes, detect multiple sources, and adapt to changes in source position over time.

Systematically evaluate the generalization capability of the model across a wider range of conditions is also needed. This includes testing in environments of varying spatial configurations, such as open areas, narrow corridors, and cluttered spaces with irregular obstacle distributions. Furthermore, the future work should also expand the evaluation with more rigorous quantitative analyses, including Receiver Operating Characteristic (ROC) curves, precision-recall metrics, and ablation studies to isolate the contributions of key architectural components.

It is worth noting that existing GSL methods differ widely in both input format and evaluation criteria. For instance, reactive plume-tracking methods are often assessed based on the final distance between the robot and the gas source, whereas probabilistic or model-based approaches focus on distribution accuracy or prediction confidence[131]. A further and more fundamental reason why widely accepted baselines are difficult to adopt in this chapter is that the dominant formulation in the GSL literature is navigation-to-source: the robot is expected to physically reach (or nearly reach) the source, and performance is primarily reported in terms of arrival distance, time-to-reach, or success in contacting the source. Moreover, a large fraction of prior evaluations is conducted in relatively open arenas with limited or no obstacles, where classical plume-tracking behaviors and gradient/wind-cue exploitation are directly applicable. In contrast, the formulation considered in this thesis is prediction without contact in obstacle-rich environments. The objective is to infer the source location and to declare a source estimate without requiring the robot to physically arrive at the source. This mismatch in task definition and termination condition means that standard reach-the-source baselines are not directly comparable under the same evaluation endpoints, and re-purposing them would require non-trivial redesigning of stopping rules and objective functions, which would introduce additional degrees of freedom and potential fairness concerns. For these reasons, the evaluation protocol in this Chapter is designed around a source-declaration criterion and interface-consistent observations, enabling reproducible measurement of localization error, convergence speed, and robustness under occlusions, while aligning with the thesis objective of source prediction rather than source contact. Due to this lack of standardized benchmarks, direct comparisons across methods remain challenging. To address this gap, a well-defined gas source declaration protocol is also needed, enabling reproducible and quantitative evaluation of localization accuracy, convergence speed, and robustness under diverse environmental conditions.

Chapter 5

Gas Distribution Mapping from Unstructured Observation

This chapter presents a Transformer-based framework for gas distribution mapping (GDM) using sparse and unevenly distributed sensor observations. Traditional GDM methods, which rely on dense sensor networks or local interpolation, often fail under practical constraints such as limited sampling coverage and environmental complexity. Leveraging the Transformer's global attention mechanism, the proposed model effectively reconstructs both gas concentration and wind fields from limited and unbalanced measurements. To train and evaluate the model, a synthetic dataset was generated using OpenFOAM-based CFD simulations under varying wind and source conditions, followed by validation on real-world wind tunnel data. Quantitative comparisons against linear interpolation, Kernel DM+V, and CNN-based baselines demonstrate that the Transformer achieves significantly lower mean squared error and higher structural similarity across various levels of sparsity and sampling balance. Case studies further highlight the model's robustness to background drift and its ability to localize gas plumes under challenging conditions. The results show that although a noticeable simulation-to-real (sim-to-real) gap still exists, Transformer based approaches can be promising solutions for robotic spatial olfaction tasks requiring local field modeling from unstructured data.

5.1 Introduction

Gas emission is a very important factor in industrial manufacturing, environmental monitoring, and risk assignment. Accurate monitoring of gas concentration distribution is critical for ensuring workplace safety, optimizing industrial processes, and minimizing environmental impact. Traditional methods for Gas Distribution Mapping rely on dense sensor networks[132], which, while effective, are often expensive and difficult to deploy in complex environments. These limitations underscore the need for new approaches to gas concentration reconstruction that can achieve high spatial resolution with minimal sensor deployment.

Due to their mobility and flexible sampling capabilities, robots have become a highly popular solution for GDM[72,75]. However, traditional GDM methods often rely on interpolating data obtained from uniform sampling throughout the environment. In real-world scenarios, there are often regions that robots cannot easily access, leading to sparse and uneven sampling. This poses significant challenges to mapping accuracy and hinders practical applications. The limitations of traditional approaches highlight the need for methods that can effectively reconstruct gas distributions even with limited and irregularly distributed observations.

As a kind of matter transport, the distribution of gas concentration is closely correlated with the wind field within the region, which is the carrier of gas transport. Although methods such as Computational Fluid Dynamics (CFD) can solve fluid equations, their high computational cost makes them unsuitable for real-time applications. Additionally, as a typical model-driven approach, their strong reliance on predefined conditions limits their adaptability to unfamiliar environments. In contrast, AI-based approaches offer a promising alternative for reconstructing fluid fields by data-driven models. Machine learning techniques, like physics-informed neural networks[108,133] (PINNs), diffusion model[134], and generative adversarial model[135,136] (GAN), can learn the underlying relationships between sampling points and overall fluid fields, enabling point-to-point reconstruction with sparse and unbalanced observations.

Robotic gas distribution mapping presents three key challenges. First, the number and locations of sampling points vary with robot movement, making fixed-input models like multilayer perceptron unsuitable. Second, gas concentration is highly sparse, most areas have near-zero values, requiring the model to extract meaningful patterns from largely uninformative data. Third, input order is not fixed due to the robot's mobility, so the model must capture spatial relationships based on relative positions rather than sequence.

To address these challenges, a Transformer-based model for GDM is proposed in this chapter. The attention mechanism captures long-range spatial correlations, while positional encoding helps the model understand the relative layout of sampling points. Given that near-ground airflow is mainly horizontal, the mapping is performed in 2D, aligned with the robot's workspace. The model takes as input the sampling coordinates, gas concentration, and wind speed components (x and y). After encoding, the data is processed through an encoder–decoder architecture, with two output heads predicting the distributions of gas concentration and wind. This design enables accurate reconstruction of fluid fields from sparse, unordered measurements.

The GDM task in this thesis is formulated as learning a mapping from an unstructured observation set to dense 2D fields. Each observation is a tuple $(\mathbf{p}_i, \mathbf{x}_i)$, where $\mathbf{p}_i \in \mathbb{R}^2$denotes the sampling coordinate and $\mathbf{x}_i$contains multi-modal measurements (e.g., gas concentration and local wind). The number of observations Nvaries across trials and the sampling order is not meaningful, which makes the input naturally permutation-invariant. The adopted transformer treats $\{(\mathbf{p}_i, \mathbf{x}_i)\}_{i=1}^{N}$as a set and performs global message passing via self-attention, enabling every sample to condition on all others without relying on a pre-defined neighborhood graph or dense rasterization. Positional encoding $\phi(\mathbf{p}_i)$is used to embed spatial relations, and attention weights become content-adaptive operators that modulate long-range interactions, which is critical under extreme sparsity or sampling bias where local interpolation is under-determined. Compared with CNN-based encoder–decoders that require fixed grid inputs (often obtained through interpolation) and have limited effective receptive fields, the transformer provides a structurally consistent interface for irregular sensing and supports cross-modal fusion through attention when jointly reconstructing wind and concentration fields.

To generate training data, A CFD method was used to simulate gas dispersion under various scenarios. The dataset includes gas sources with different locations and intensities, as well as regional wind fields with varying directions. After training, the model was validated using real data collected from a low-speed wind tunnel. Results show that, compared with traditional statistical methods like Kernel DM+V[74] or typical neural network like CNN[7], this approach can more accurately reconstruct the gas concentration and wind field across the area, especially from sparse and unevenly distributed sensor samples.

5.2 Experimental Method

5.2.1 Training Data Generation by OpenFOAM

In order to generate physically reliable and diverse training datasets for data-driven gas distribution mapping, a comprehensive series of two-dimensional computational fluid dynamics simulations were systematically performed. These simulations aim to replicate realistic turbulent airflow behavior and the subsequent transport and dispersion of passive scalars, specifically gas species, within a controlled computational domain. The primary objective is to display the intricate interplay between turbulence-induced mixing and scalar transport processes, which are crucial for reconstructing accurate gas concentration fields from sparse sensing data.

The fluid flow is modeled as incompressible, which is a valid approximation under low-speed conditions typically encountered in indoor or urban dispersion scenarios. To represent the turbulent nature of the flow, the Reynolds-averaged Navier-Stokes[119] (RANS) equations are solved. This approach effectively separates the instantaneous velocity and pressure fields into mean and fluctuating components, enabling statistical treatment of turbulence effects. To close the system, the widely validated standard $k - \varepsilon$ turbulence model is employed, which provides transport equations for turbulent kinetic energy and its dissipation rate, thereby yielding an eddy viscosity that models the enhanced momentum transfer due to turbulence.

The transport of the gas concentration field is modeled as a passive scalar, meaning that it does not affect the flow field dynamically and is simply carried and diffused by the underlying velocity and turbulent eddies. The computational domain is a two-dimensional square region of 1m times 1m, discretized using a uniform mesh of 100 times 100 cells. This mesh resolution is chosen to balance computational efficiency and sufficient spatial detail to resolve key flow and scalar features. All simulations are done by OpenFOAM, an open-source CFD framework, the constants are standard values obtained from extensive past experimental validation.

To enable the model to learn a broader range of gas dispersion patterns and prevent overfitting, the dataset primarily consists of three types of cases, convection-dominated, diffusion-dominated and blank. In the convection case (Figure 5.1A), wind serves as the primary driving force for gas dispersion, resulting in a strong coupling between the gas concentration field and the wind field. In diffusion case (Figure 5.1B), the wind speed is relatively low compared to the source strength, so gas dispersion is dominated by diffusion. This results in a broader plume shape and smoother concentration gradients along the edges. In the blank case (Figure 5.1C), no gas source is present and the concentration distribution is entirely zero. This case is included to address sample imbalance and prevent the model from hallucinating gas presence in negative samples. 400 convection-dominated cases, 100 diffusion-dominated cases and 100 blank cases are included in the synthesized dataset.

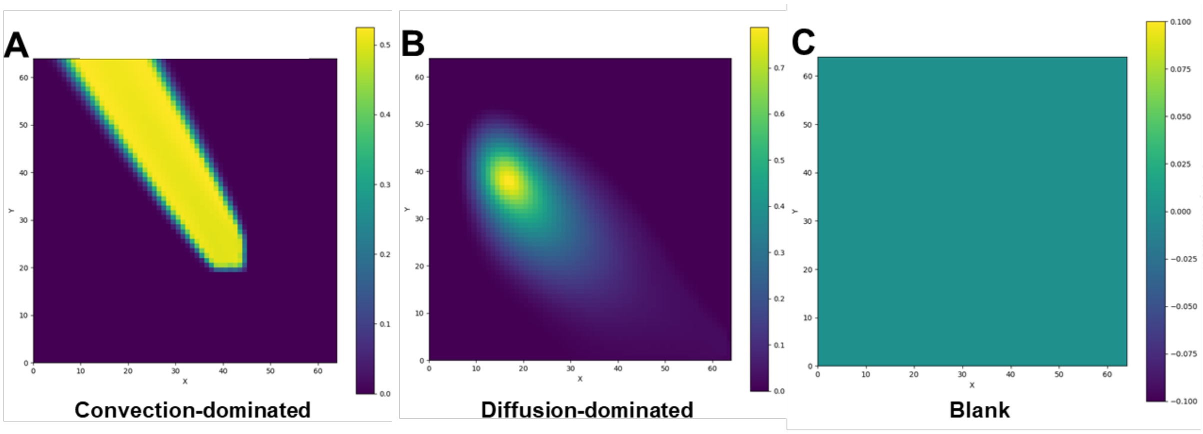

Figure 5.1 The composition of the dataset. (A) Convection-dominated cases. (B) Diffusion-dominated cases. (C) Blank cases.

5.2.2 Data Extraction and Sensor Down Sampling

The scalar concentration field is interpolated onto a 64 times 64 uniform grid to serve as the ground truth. To simulate real-world sensor limitations, a fixed number of sampling points (N) are drawn from this grid. The selection is randomized per case, and optionally biased toward high-concentration regions to mimic targeted deployment strategies. Multiple values of N are explored in this study (e.g., 20, 30, 50). For each simulation, 32 distinct sensor configurations are generated to ensure robustness. This results in a large and diverse dataset: 200 simulated cases for training and 20 cases for testing, each case will generate 32 different input-output samples.

5.2.3 Real-world Evaluation

The validation of our model is based on an open-source, real-world dataset collected in a low-speed wind tunnel[124]. The experimental area measures $2.40 \times 2.40 \times 0.88\ m$ and contains a scaled-down industrial site model with obstacles and a gas source emitting propylene glycol vapour (Figure 5.2A). As shown in Figure 5.2B, a robotic arm carries four gas sensing units positioned at different vertical heights (i.e., at Layer 1, Layer 2, Layer 3 and Layer 4), acquiring measurements across a 20 ×20 horizontal grid in the XY-plane. Each sensing unit comprises a MiCS5524 metal oxide (MOX) gas sensor, a PID-AH2 photo-ionization detector (PID) for measuring gas concentrations and a Trisonica Mini ultrasonic 3D anemometer for horizontal wind velocity. As one can observed in Figure 5.2C, compared to the PID, which provides calibrated and relatively sharp gas concentration peaks and fast recovery, the MOX sensor responses are characterized by slower recovery dynamics, susceptibility to baseline drift, and did not report calibrated concentration levels. This experimental setup yields a high-fidelity dataset that reflects realistic dispersion conditions with complex airflow structures and spatial obstructions.

For the evaluation of the learning models, the measurements collected from Layer 1 was selected, the plane closest to the gas source, where gas signal intensity and spatial variation are most prominent. Specifically, the MiCS5524 MOX readings from this layer are used as the input gas sensor signals for the learning models. To align with the spatial resolution used in our simulated training data, the raw measurements are interpolated onto a 64 ×64 grid, forming dense ground truth maps of gas concentration and wind fields. From these maps, sparse sensor inputs are generated using the same probabilistic sampling strategies defined in training, ensuring consistency in the sampling process between training and evaluation.

The ad hoc evaluation setup based on the real-world dataset allows a realistic and challenging benchmark to assess the considered approaches under practical conditions, featuring nonuniform sampling, real sensor noise, and complex airflow dynamics.

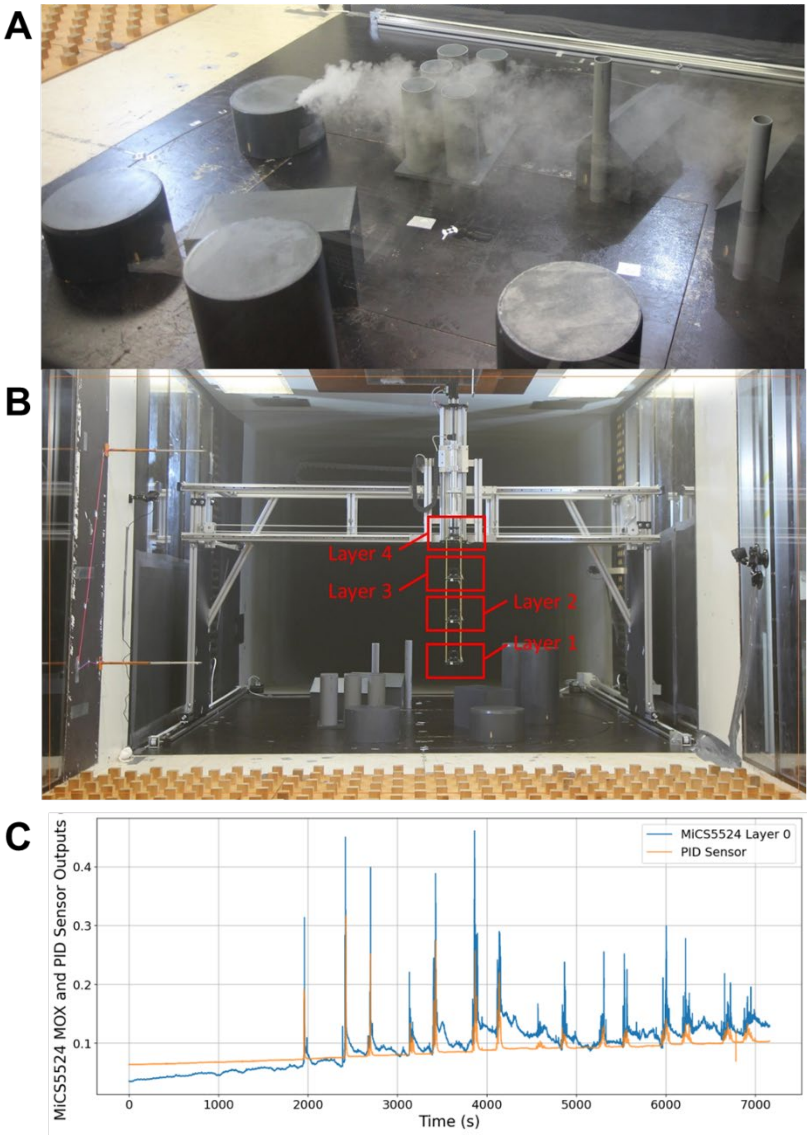

Figure 5.2 The experimental setup of the low-speed wind tunnel dataset[124]. (A) An example of the experimental environment with an active gas source (B) Vertical arrangement of 4 sensing units at different heights. (C) Time series of MOX and PID sensor responses at Layer 1 in the real-world dataset.

5.2.4 Model Evaluation

To assess the effectiveness of the proposed method, quantitative comparisons were conducted against alternative approaches. Linear interpolation, Kernel DM+V, and a CNN encoder–decoder architecture was selected as baseline methods for comparison.

Linear interpolation serves as a simple and widely used baseline, where unknown gas concentrations across the spatial domain are estimated by interpolating between known sensor measurements. Kernel Extrapolation Distribution Mapping + Variance (Kernel DM+V) is a widely used non-parametric statistical method that learns a probabilistic two-dimensional distribution model from a set of local gas sensor measurements. Unlike conventional interpolation techniques, Kernel DM+V not only estimates the mean of the gas distribution but also captures the observed variance. Incorporating predictive variance significantly enhances gas distribution modeling, as it enables the evaluation of model quality based on data likelihood and provides a principled way to learn meta-parameters. Lastly, a CNN encoder–decoder model is employed as a learning-based baseline. This architecture maps sparse sensor inputs to a dense concentration field through hierarchical feature extraction and spatial up sampling. Although purely data-driven, it provides a strong benchmark for learning-based reconstruction and has been effective in various inverse mapping tasks. This framework has achieved state-of-the-art performance in previous GDM super-resolution tasks. To ensure a fair comparison, a CNN encoder-decoder with a parameter count comparable to that of the Transformer model was used. It takes a 4-channel input of spatial resolution 64 times 64 and outputs a 3-channel prediction map of the same size. Since CNNs are not able to handle variable-length inputs, the sensor observations were first zero-padded based on their spatial coordinates to form a fixed-size 4-channel matrix of 64 times 64. The four channels represent gas concentration, wind velocity in the x and y direction, and a binary mask that explicitly indicates the locations of sensor observations. The encoder consists of three convolutional layers that progressively reduce spatial resolution (64-32-6) while increasing feature depth to capture high-level representations. The decoder then reconstructs the original resolution using two transposed convolutional layers followed by a final convolution that outputs the desired number of channels. LeakyReLU activations are used throughout the network to maintain non-linearity while mitigating the risk of dead neurons.

The evaluation of the models was primarily based on two metrics, mean squared error (MSE) and structural similarity index[137] (SSIM). MSE is a commonly used metric that measures the average squared difference between the predicted values and the ground truth. It is defined as:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y_i})^2 \tag{25}$$

where y_i is the true value and $\widehat{y_i}$ is the predicted value. A lower MSE indicates higher numerical accuracy and better overall fit. However, MSE does not account for perceptual similarity or structural features in spatial data. SSIM is a perceptual metric that evaluates the similarity between two spatial fields based on structural information. SSIM ranges from 0 to 1, with 1 indicating perfect structural similarity. SSIM captures spatial consistency, making it suitable for evaluating reconstructed fields such as gas concentration maps. SSIM is defined as:

$$\mathrm{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{26}$$

where μ_x and μ_y is the average value of field x and y, σ_x^2 and σ_y^2 is variance of field x and y, and σ_{xy} is covariance. C_1 and C_2 are constant used to avoid division by zero.

In evaluating the accuracy and robustness of the models, two categories of variables were controlled: sparsity and balance. Controlling sparsity was intended to evaluate the model's performance under varying numbers of sensor observations, specifically to assess whether it can reconstruct the spatial concentration field from sparse measurements, which is important for real-world applications. Models' performance was evaluated under sample numbers of 100, 50 and 20. The balance evaluation was designed to assess the model's stability when dealing with uneven and unstructured sampling patterns. This was intended to simulate real-world scenarios where a robot's mobility is constrained, leading to irregular and limited sampling coverage. Sampling balance was controlled by using concentration-weighted probabilistic sampling during sensor data extraction, thereby biasing sample selection based on local gas concentration values. The probability of selecting a given grid cell p_i was defined as:

$$p_i = \frac{(c_i + \varepsilon)^{\alpha}}{\sum_{j=1}^{N}\left(c_j + \varepsilon\right)^{\alpha}} \tag{27}$$

where c_i denotes the local gas concentration, ε is a small constant for numerical stability, which is assigned as 0.01, α is a tunable parameter that adjusts sampling preference: $\alpha>0$ means denser sampling in high-concentration regions, $\alpha<0$ means denser sampling in low-concentration areas, and $\alpha=0$ corresponds to uniform sampling. Models' performance was evaluated under α of 0.5, 0 and -0.5, with sampling number of 50.

5.3 Principal Outcomes

This part quantitatively evaluates all models using the held-out test set, with performance summarized in Table 5.1, 5.2, 5.3, and 5.4. Overall, across all sensor sparsity levels and sampling biases, the Transformer-based model consistently achieves the lowest mean squared error (MSE) and highest structural similarity (SSIM) for both gas concentration and wind field reconstruction. Compared to baseline methods, including linear interpolation, Kernel DM+V, and CNN, the Transformer demonstrates significantly higher robustness to reduced sample counts and biased sampling. The subsequent Section provides a more detailed breakdown of these results under varying sparsity levels and sampling biases, along with qualitative visual comparisons based on representative cases. Section 5.3.4 shows that the model produces a qualitatively realistic gas plume reconstruction with real wind-tunnel data. This result reflects this work's potential in sim-to-real generalization.

5.3.1 Model Training

Both the CNN encoder-decoder (Figure 5.1) and the Transformer models (Figure 5.2) were trained using a curriculum learning approach. The models were first trained on a dataset with 100 samples per instance and subsequently fine-tuned on a sparser dataset with only 50 samples per instance. At each stage, the learning rate was scheduled using the OneCycleLR method[126], which involves an initial warm-up phase followed by a gradual reduction in the learning rate. The learning rate and number of epochs for each stage were determined through preliminary experiments.

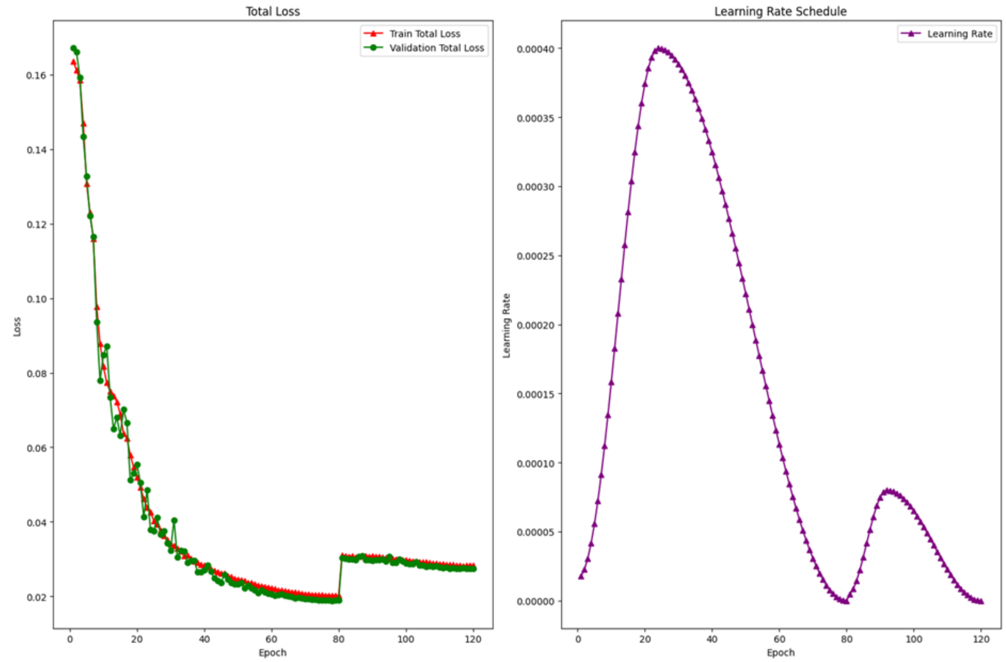

Figure 5.1 Loss curve and learning rate curve of CNN Encoder-decoder.

Throughout training, both models showed consistent convergence patterns. However, the Transformer-based architecture exhibited notably faster convergence compared to the CNN model, achieving lower validation loss within fewer epochs. During the second stage of training, the Transformer maintained stable performance with minimal loss fluctuations despite the increased sparsity of input data. In contrast, the CNN model experienced a temporary increase in validation loss during the stage transition, indicating higher sensitivity to changes in sample density.

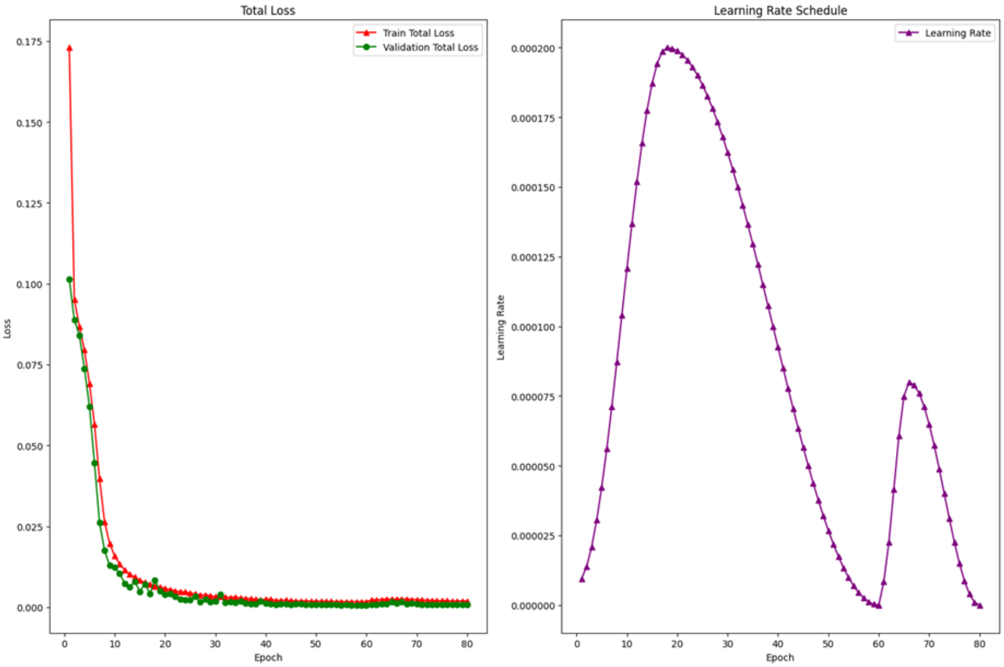

Figure 5.2 Loss curve and learning rate curve of Transformer.

These findings suggest that the Transformer model is better equipped to handle distributional shifts across stages, possibly due to its global receptive field and attention mechanism, which allow for more flexible integration of spatially sparse inputs.

5.3.2 Model Evaluation with Different Sparsity

As shown in Table 5.1, the results based on MSE reveal that the Transformer model consistently outperforms the other methods across all categories. Specifically, for both gas concentration and wind field reconstruction, the Transformer achieves MSE values that are approximately an order of magnitude lower than those of the baseline methods. Furthermore, the transformer maintains stable performance even as the number of sensor observations decreases, demonstrating strong robustness to sparsity. In contrast, the Kernel DM+V method exhibits more noticeable fluctuations in MSE as sampling becomes sparser. This performance drop can be attributed to the method's reliance on local kernel-based estimation, which is more sensitive to irregular or sparse sampling patterns. Unlike the Transformer, which can globally integrate spatial features, Kernel DM+V lacks the capacity to adaptively infer long-range dependencies from limited or unevenly distributed data.

Table 5.1 Mean Squared Error of Different Approaches under Different Sparsity.

MSE	Linear Interpolation	Kernel DM+V	CNN Encoder-decoder	Transformer
Gas 100 Samples	0.0400	0.0090	0.0095	0.0003
Wind 100 Samples	0.0098	0.0023	0.0011	0.0001
Gas 50 Samples	0.0386	0.0135	0.0125	0.0003
Wind 50 Samples	0.0115	0.0063	0.0019	0.0001
Gas 20 Samples	0.0382	0.0243	0.0202	0.0005
Wind 20 Samples	0.0165	0.0186	0.0076	0.0002

As shown in Table 5.2, the results based on SSIM further highlight the advantages of the Transformer model. Across all categories, the Transformer consistently achieves higher SSIM scores than the other methods, indicating better preservation of spatial structure and perceptual quality in the reconstructed fields. In contrast, the other methods, particularly linear interpolation and Kernel DM+V, show a marked decline in SSIM as the number of sensor observations decreases. This can be attributed to their limited ability to reconstruct fine-grained spatial patterns under sparse sampling. While interpolation methods tend to produce overly smoothed results, Kernel DM+V, though more adaptive, still struggles to infer structural coherence when local data support is insufficient. The Transformer, on the other hand, leverages global attention to capture both local details and long-range dependencies, enabling it to maintain structural fidelity even in low-observation scenarios.

Table 5.2 Structural Similarity Index of Different Approaches under Different Sparsity.

SSIM	Linear Interpolation	Kernel DM+V	CNN Encoder-decoder	Transformer
Gas 100 Samples	0.5308	0.6809	0.5560	0.9830
Wind 100 Samples	0.6186	0.8601	0.8276	0.9733
Gas 50 Samples	0.4932	0.6664	0.4851	0.9822
Wind 50 Samples	0.5710	0.7328	0.7928	0.9724
Gas 20 Samples	0.4669	0.6473	0.2875	0.9767
Wind 20 Samples	0.4609	0.4820	0.5945	0.9676

Figure 5.3 shows the case study of different approaches on a same sample, which further illustrates the behavior of different models under reduced sampling conditions. When the number of sensor observations is insufficient to evenly cover the entire plume region, linear interpolation tends to overestimate the values between positive samples. This results in unrealistic plume boundaries and expansion of the gas region. In contrast, the Kernel DM+V method underestimates concentrations in unsampled plume areas, often leading to hollow or fragmented reconstructions. The CNN encoder-decoder, while more robust, exhibits increasingly blurred predictions as the sample count decreases. This blurring effect likely reflects a bias toward the average patterns observed in the training data, causing the model to regress toward conservative, smoothed outputs in the absence of sufficient local information.

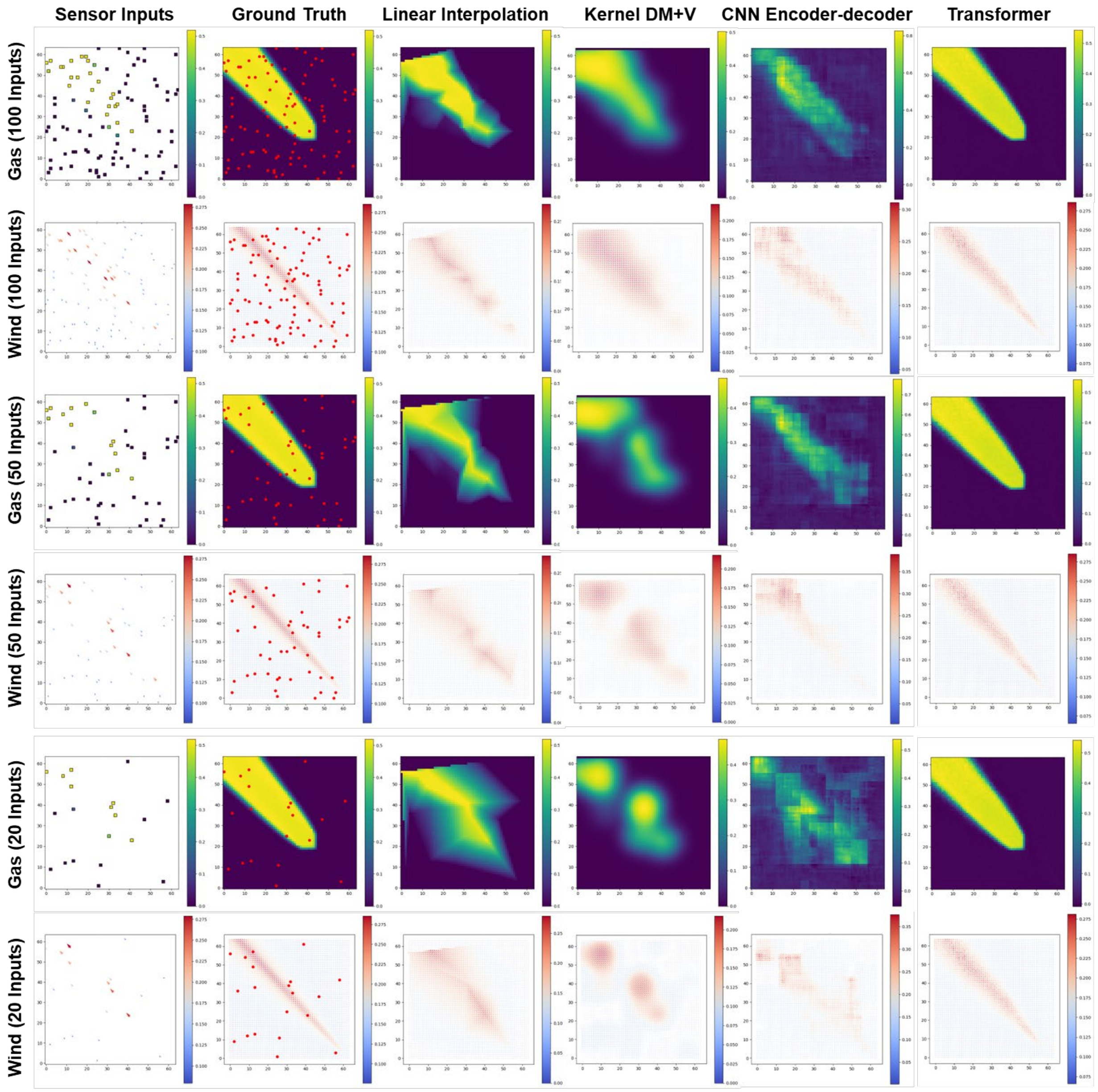

Figure 5.3 Performance of different approaches under different sampling sparsity. The Transformer-based approach shows significantly higher accuracy in both wind field and gas concentration distribution.

5.3.3 Model Evaluation with Different Sampling Balance

As shown by Table 5.3, the evaluation results under varying levels of sampling balance further demonstrate the robustness of the Transformer model in terms of MSE. Across all categories, the Transformer achieves MSE values that are consistently an order of magnitude lower than those of the other methods for both gas concentration and wind field prediction. Among the baseline methods, linear interpolation performs relatively well when sampling is dense within the plume region, as it can effectively capture local gradients. However, its performance degrades in areas with sparse observations or uneven coverage. Kernel DM+V and CNN models tend to perform better under uniformly distributed sampling, where spatial generalization is more reliable. When sampling becomes unbalanced, both methods exhibit noticeable increases in prediction error. While the Transformer also shows some increase in MSE when the plume region is sparsely sampled, it still outperforms all other methods.

Table 5.3 Structural Similarity Index of Different Approaches under Different Balance.

MSE	Linear Interpolation	Kernel DM+V	CNN Encoder-decoder	Transformer
Gas Alpha=0.5	0.0118	0.0464	0.0358	0.0004
Wind Alpha=0.5	0.0094	0.0137	0.0052	0.0003
Gas Alpha=0	0.0386	0.0135	0.0125	0.0003
Wind Alpha=0	0.0115	0.0063	0.0019	0.0001
Gas Alpha=-0.5	0.0327	0.0403	0.0522	0.0008
Wind Alpha=-0.5	0.0097	0.0151	0.0064	0.0002

The SSIM-based evaluation (Table 5.4) under different levels of sampling balance shows that, in every category, the Transformer consistently achieves higher SSIM scores than the other methods. As the sampling balance decreases, traditional methods show a significant drop in SSIM. Linear interpolation is affected by distorted spatial transitions in poorly sampled regions, leading to blocky or overly smooth reconstructions. Similarly, Kernel DM+V and CNN models experience a decline in structural similarity, as their ability to generalize across unsampled areas diminishes without evenly distributed training signals. In contrast, although the Transformer also experiences minor degradation under imbalanced sampling, it maintains a much higher level of structural fidelity.

Table 5.4 Structural Similarity Index of Different Approaches under Different Balance.

SSIM	Linear Interpolation	Kernel DM+V	CNN Encoder-decoder	Transformer
Gas Alpha=0.5	0.6090	0.4046	0.2119	0.9749
Wind Alpha=0.5	0.6418	0.5132	0.5989	0.9677
Gas Alpha=0	0.4932	0.6664	0.4851	0.9822
Wind Alpha=0	0.5710	0.7328	0.7928	0.9724
Gas Alpha=-0.5	0.6464	0.5844	0.2401	0.9675
Wind Alpha=-0.5	0.7045	0.5125	0.6514	0.9654

Figure 5.4 shows the case study of different approaches on a same sample on different sampling balance. When sensor observations are denser within the plume region, both linear interpolation and Kernel DM+V tend to overestimate the extent of the plume, resulting in some exaggeration of the gas region. Nonetheless, in these scenarios, they are still able to provide reasonable coverage of the overall concentration field. However, when the plume region is sparsely sampled, both methods show severe distortion. The dominance of negative samples causes them to under-represent or miss the plume. The CNN model, on the other hand, demonstrates a tendency to regress toward the mean prediction under both dense and sparse sampling conditions. This produces blurred and conservative outputs that fail to capture the sharp spatial structure of the plume. This behavior suggests that CNN may struggle to effectively model long-range spatial dependencies, making it less capable of adapting to irregular or unbalanced sensor layouts.

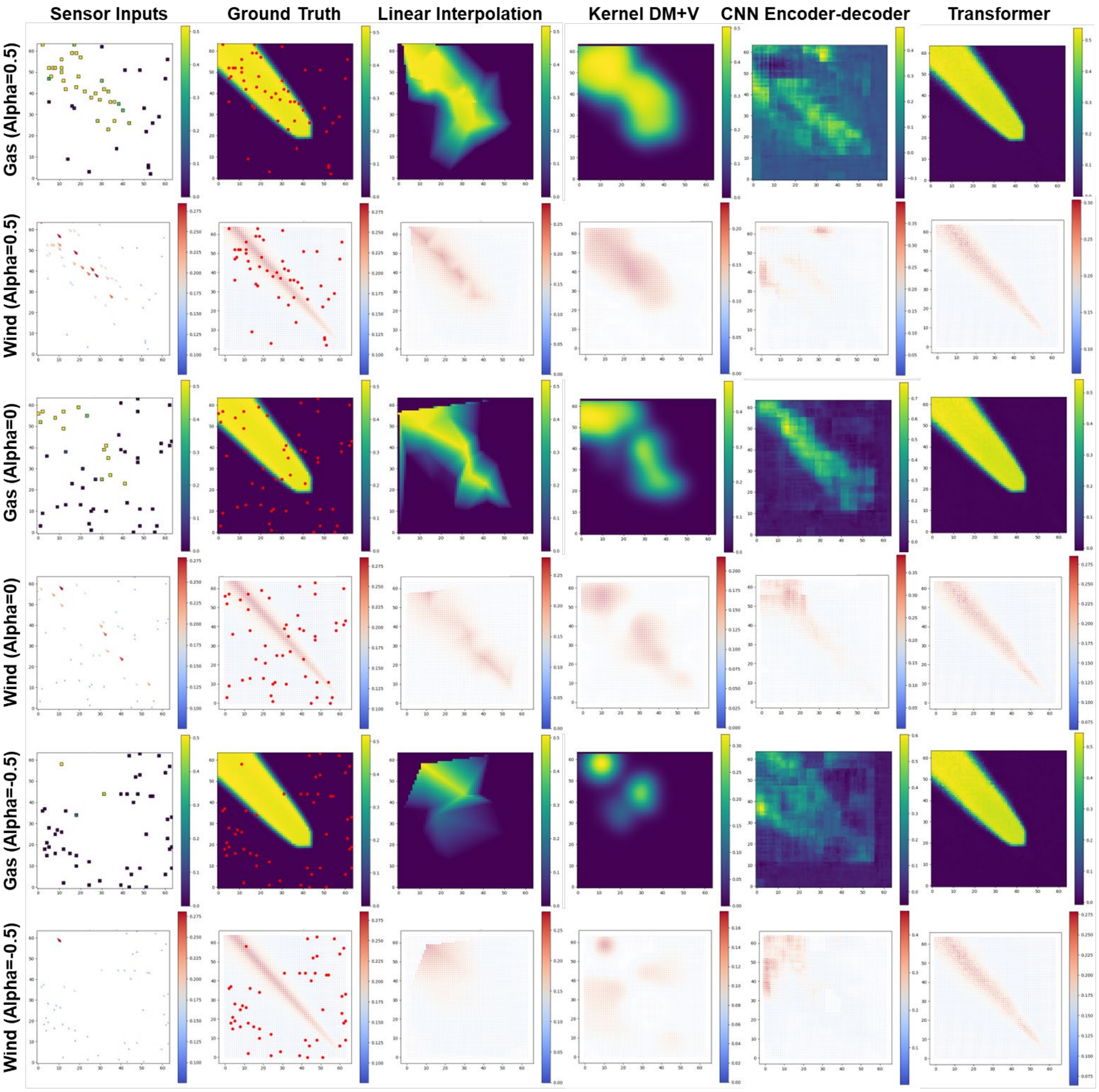

Figure 5.4 Performance of different approaches under different sampling balance. The Transformer-based approach shows significantly higher accuracy in both wind field and gas concentration distribution.

5.3.4 Case Study of Real-world Experiment

To evaluate the model's generalization ability and assess the sim-to-real gap, the performance of different models was tested on data collected from a real wind tunnel. Due to the absence of accurate ground truth (only noisy sensor observations were available), no quantitative comparison was conducted. Instead, a qualitative case study analysis was performed. To better simulate real-world conditions, no additional enhancement or filtering was applied to the results. The raw, uncalibrated sensor outputs were merely scaled to match the value range of the simulation results.

To evaluate the model's generalization ability and assess the sim-to-real gap, the performance of different models was tested on data collected from a real wind tunnel. Due to the absence of accurate ground truth (only noisy sensor observations were available), no quantitative comparison was conducted. Instead, a qualitative case study analysis was performed. In this setting, interpolated concentration map should not be interpreted as an exact ground-truth field. The sensor readings are subject to measurement noise, baseline drift, and response dynamics. Therefore, interpolation only yields an observed distribution that is consistent with the available measurements. In other words, the interpolated field reflects a measurement-driven representation rather than the true concentration field $c(\mathrm{x}, t)$. Under this representation, conventional pixel-wise error metrics would largely quantify agreement with the interpolation scheme and its smoothing bias, rather than physical reconstruction fidelity, and would be sensitive to the specific interpolation method and its hyperparameters. To better simulate real-world conditions, no additional enhancement or filtering was applied to the results. The raw, uncalibrated sensor outputs were merely scaled to match the value range of the simulation results. Nevertheless, the wind-tunnel data were collected via dense scanning using a robotic arm, producing high spatial coverage of measurements. Under such dense sampling, interpolation can reliably reveal the contrast between plume-affected regions and background levels and can provide a reasonable approximation of the plume footprint and contour. For this reason, the interpolated concentration maps are used only to support qualitative visualization and case-study interpretation.

As shown in Figure 5.5A, none of the tested models were able to accurately reconstruct the wind field distribution. This limitation is primarily attributed to the presence of physical obstacles within the test area, which introduced numerous low-speed zones and turbulent regions. These complex flow dynamics disrupt the spatial continuity of the wind field, making it difficult for numerical methods to interpolate or infer missing information from sparse observations. Moreover, in the absence of prior knowledge about the location or shape of obstacles, and without having encountered similar flow patterns during training, the models lacked the contextual information necessary to accurately predict the local disturbances caused by the obstacles.

Result of gas distribution mapping of the real wind tunnel experiment, shown in Figure 6.5B, illustrates the significant challenges posed by low signal-to-noise ratios in gas sampling. During the sequential sampling process from upwind to downwind, the sensor exhibited noticeable baseline drift, resulting in a systematically elevated background concentration in the upper half of the observed distribution. As a consequence, the actual gas plume, highlighted by the red box in Figure 6.5B, which corresponds to the visible smoke region in the physical experiment, was submerged within the background signal. Under these conditions, baseline-sensitive methods such as linear interpolation, Kernel DM+V, and CNN misinterpreted the influence range of the plume due to the background offset, leading to inaccurate reconstructions.

In contrast, the Transformer demonstrated a clear advantage by successfully reconstructing a complete plume region despite the strong background drift. Its global attention mechanism allowed it to distinguish the true signal pattern from the noise-dominated observation, effectively identifying the plume shape and orientation. However, the Transformer’s output was still affected by the noise during measurement, the reconstructed plume contours did not fully align with the ground truth, and two overlapping plume heads appeared in the upwind region. These artifacts likely stem from distorted wind direction observations, indicating that despite the model's robustness, a sim-to-real gap remains.

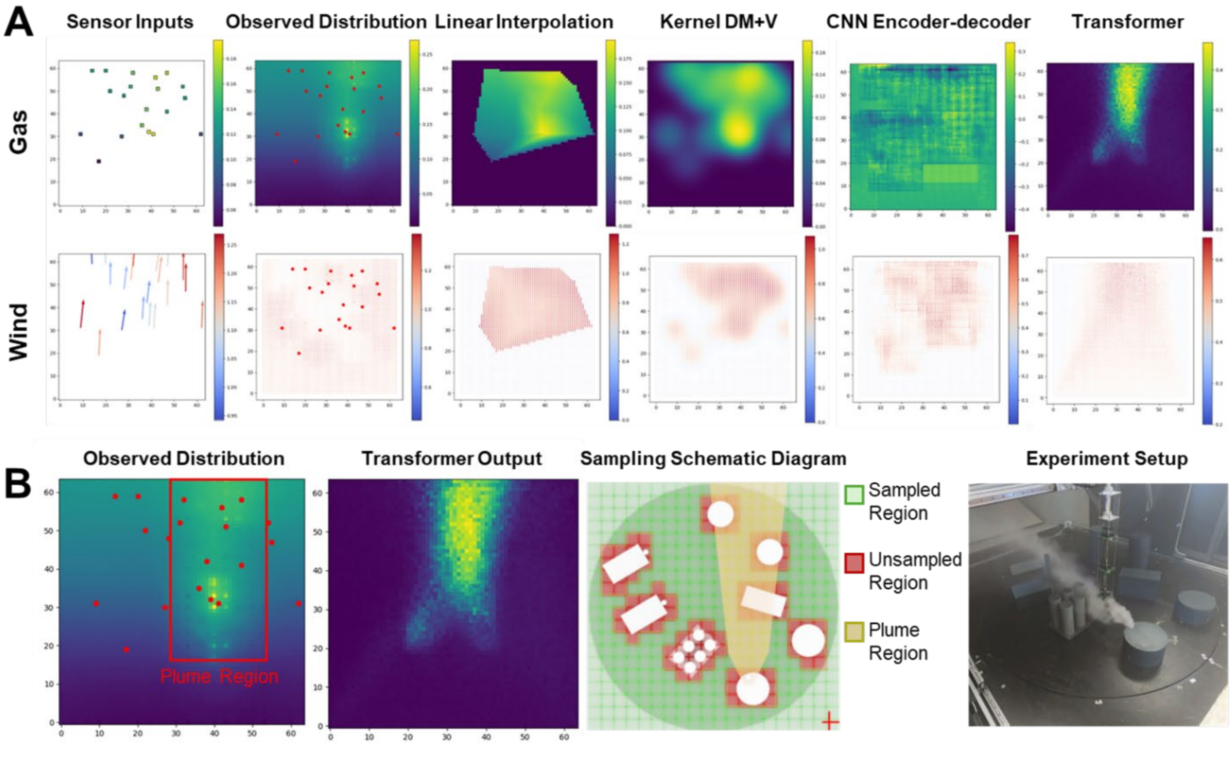

Figure 5.5 Evaluation with real-world data. (A) Models' performance on data from wind tunnel experiments. (B) Comparison of Transformer's output of gas concentration with the real plume region.

5.4 Conclusion

Gas distribution mapping (GDM) in mobile robotics has long faced the challenge of generating accurate spatial reconstructions from sparse and uneven sensor data. Traditional approaches often rely on dense sampling or assume idealized plume models, which rarely hold in complex real-world environments. Learning-based methods, especially convolutional networks, have improved reconstruction performance, but they typically struggle to handle highly irregular inputs or to capture long-range spatial dependencies. These gaps underscore the need for models that are both structurally flexible and data-efficient.

This chapter presents a Transformer-based model for GDM that directly addresses these limitations. By leveraging attention mechanisms and positional encoding, the model learns to globally integrate sparse sensor measurements and infer spatial structure from limited and biased samples. Quantitative evaluations on synthetic test cases demonstrate the robustness of the approach: under extremely sparse sampling (20 observations), the Transformer achieves an MSE of 0.0005 and SSIM above 0.96 for gas concentration prediction, substantially outperforming the baseline methods, linear interpolation, Kernel DM+V, and CNN-based models. Even under the sampling bias, where sensor inputs are intentionally skewed away from the plume, this work maintains strong reconstruction fidelity, while the compared baseline methods deteriorate significantly. In the real-world wind tunnel experiment, the Transformer model demonstrated an advantage in handling background noise caused by sensor baseline drifting. While other methods misinterpreted the elevated background as part of the plume, the Transformer-based approach effectively separated the true plume region from the noisy observations.

A potential concern for learning-based GDM is whether the model memorizes case-specific plume templates rather than learning transferable spatial relationships. The training and evaluation protocol in this chapter is designed to reduce this risk in two ways. First, during training, each CFD case is subjected to multiple rounds of random re-sampling of sparse observations. This prevents the model from associating a single fixed input pattern with a specific output field and forces reliance on geometry-conditioned relationships among samples. Second, during testing, evaluation is performed on previously unseen cases (with different source and flow realizations) and with multiple rounds of random re-sampling. Consistent performance across re-sampling trials indicates that predictions are not tied to a particular sampling template, but instead depend on the relative spatial arrangement and values of the observed samples. Together, the combination of case-level holdout and repeated stochastic re-sampling provides support that the attention mechanism captures transferable, geometry-aware aggregation rules under variable-length, irregular sensing, rather than memorizing a fixed map of input patterns to outputs.

These findings reinforce the emerging consensus in spatial modelling that Transformers offer appealing advantages over convolutional deep learning models, particularly for unstructured inputs and tasks involving long-range correlation. Whereas CNNs tend to regress toward averaged patterns seen during training, the Transformer is able to preserve both local plume features and broader spatial continuity, leading to more faithful reconstructions across a variety of input conditions. Compared to the data-driven Kernel DM+V, which assumes smooth spatial extrapolation and is sensitive to sampling bias, the Transformer is markedly more adaptive.

However, several limitations should be acknowledged. First, this work still exhibits a sim-to-real gap. The Transformer -based model is trained entirely on simulated data, and although it is evaluated on real wind tunnel measurements, the simulation-to-reality transfer remains imperfect. In real-world environments, obstacles introduce complex wind field distortions that are not represented in the training data. Simple noise injections during training are insufficient to capture such structural irregularities. Furthermore, due to the presence of cross-attention between wind and concentration channels, inaccuracies in wind prediction propagate to gas concentration estimates, causing distortions in the reconstructed plume shape and location.

A key challenge underlying this sim-to-real gap is the scarcity of high-quality real-world training data. In many wind-tunnel and field settings, dense ground-truth wind and concentration fields are difficult to obtain, which makes it necessary to rely heavily on simulation for supervised learning. Moving forward, one promising direction is to expand advanced dispersion imaging modalities, such as schlieren techniques[138] or multispectral remote sensing[139] to acquire higher-fidelity measurements that can support finer-grained supervision and more realistic domain coverage. In parallel, the pretraining pipeline can be strengthened by developing more physically reliable simulation procedures, particularly by modeling sensor noise with physics-grounded characteristics (e.g., drift dynamics, response latency, saturation, direction-dependent bias, and correlated disturbances) rather than using simple additive Gaussian perturbations. Such realism-oriented simulation and data acquisition efforts would directly reduce the mismatch between synthetic and real measurements and hence narrow the sim-to-real gap.

Extending the framework to full 3D reconstruction introduces additional constraints that are primarily data and compute driven. Generating and storing 3D CFD-based datasets significantly increases computational cost, and learning 3D fields typically requires larger-capacity models, which in turn demand more training data to avoid overfitting. Addressing this bottleneck will likely require community-level efforts to build and release high-quality open-source 3D benchmark datasets for robotic olfaction and turbulent dispersion. Complementarily, it is worthwhile exploring more data-efficient learning paradigms, such as physics-informed neural networks (PINNs) or neural operators, which can leverage governing-physics structure to improve sample efficiency and generalization, potentially reducing the dependence on large volumes of fully labeled 3D data.

Apart from the observed limitation in the wind field reconstruction, a potential concern is that the model might implicitly memorize typical plume patterns from the training data, and its success on the test set could partly stem from structural similarities between training and evaluation scenarios. Furthermore, the Transformer operates in a data-driven manner without explicitly enforcing physical laws such as mass conservation or fluid dynamics, which could limit its generalization under turbulent or multi-source gas dispersion.

The observed and potential limitations suggest that while the Transformer offers a powerful framework for data-driven GDM, future efforts are needed to address the generalization to real-world scenarios. Future work should improve the training data in terms of expanding the diversity of training scenarios, including more complex environmental geometries and variable gas dispersion patterns. In addition, integrating physics-informed training approach and domain-specific inductive biases could allow the model to be better aware of the underlying fluid dynamics. However, physics-informed training has higher requirement to data quality, as noise in training may cause gradient between physics loss and dataset loss, resulting in trivial solution. To improve robustness, especially to incorporate with widely used metal-oxide gas sensors, noise and sensor response models that more accurately reflect the behavior of real gas sensors should be considered. Constructing such training datasets will likely require a hybrid approach that combines high-fidelity simulations with real-world experimental data.

Additionally, deploying the model on a mobile robotic platform in outdoor or industrial environments would enable real-time evaluation under realistic conditions and provide insights into practical deployment, resilience to sensor drift, and adaptive planning for efficient sampling. Besides, extending the framework to three-dimensional environments is a natural but important future direction towards realistic, complex applications. In many practical scenarios, such as multi-floor buildings or underground tunnels, gas dispersion occurs in all three spatial dimensions. A 3D extension of GDM would enable more comprehensive monitoring and improved applicability to real-world deployments.

Chapter 6

Conclusions and Recommendations

This chapter concludes the thesis by summarizing its key contributions to robotic spatial gas perception, focusing on two core tasks: gas source localization (GSL) and gas distribution mapping (GDM). Through the design of task-specific learning architectures, sensor platforms, and experimental protocols, the thesis presents context-aware solutions that leverage environmental structure and wind information to infer high-dimensional gas patterns from sparse observations. The proposed U-Net-based GSL method integrates SLAM-derived maps and wind vectors to localize gas sources in complex, obstacle-rich environments, while the Transformer-based GDM model reconstructs gas and wind fields from unstructured sensor inputs with high spatial fidelity. Both methods demonstrate strong performance in simulation and real-world validation, advancing the robustness and applicability of robotic olfaction systems. The chapter also outlines directions for future research, including multi-modal sensing, sim-to-real transfer, and autonomous navigation strategies, to further enhance the adaptability and scalability of gas-sensing robots in real-world deployment scenarios.

6.1 Conclusions

This thesis primarily focuses on robotic spatial gas perception. Overall, the goal of robotic olfaction is to model high-dimensional gas distribution features based on limited observations. Achieving this goal is particularly challenging due to the inherent sparsity, irregularity, and uncertainty of gas sensing data, especially in dynamic or unstructured environments. By addressing two fundamental tasks at different stages of gas dispersion, gas source localization (GSL) in turbulent, obstacle-rich settings and gas distribution mapping (GDM) under sparse sampling constraints, this work proposes context-aware solutions that integrate environmental structure, wind flow, and sensor feedback. Through the design of specialized learning architectures, sensor platforms, and realistic experimental protocols, the thesis contributes toward building robust and generalizable gas perception systems suitable for real-world deployment.

Chapter 4 displayed a robotics olfaction application under the scenario of topography aware gas source localization, based on the characteristics of sensor developed in chapter 4. In this chapter, the proposed method integrates the physical map obtained from SLAM, the locations where the robot encounters gas within the map, and the corresponding wind velocity vectors into a unified tensor. A U-Net model is then trained to predict the probability density distribution of the gas source. Incorporating the physical map and using a U-Net to capture spatial context enables the model to predict gas source locations with improved accuracy even in environments with obstacles represented in the evaluation. This approach helps mitigate the effects of wind-topography interactions, which are considered one of the main challenges in gas source localization. The model's performance was evaluated using the predicted centroid location and the proportion of positive grid cells. As the number of recorded gas encounter positions increased, both accuracy and belief improved significantly, highlighting the critical role of spatial correlations among encounter locations and the physical map in effective gas source localization. Experiments conducted in dynamic wind fields demonstrated that the proposed method could maintain a consistent convergence trend under changing wind directions not seen during training, with a modest increase in uncertainty, indicating robustness within the tested dynamic-wind settings. The code and data in this chapter is open source.

Chapter 5 focuses on the problem of gas distribution mapping in robotic olfaction. GDM typically involves handling unstructured data, as observations are often sparse and unevenly distributed due to constraints such as the robot's accessibility. In this chapter, a Transformer-based model for GDM is proposed. With the spatial encoding and cross-attention mechanisms of the Transformer architecture, the model can effectively capture spatial relationships between observations and handle variable-length inputs, avoiding the loss of high-frequency structures that during padding or interpolation. Different approaches were evaluated using Mean Squared Error (MSE) and Structural Similarity Index Measure (SSIM). The results show that, compared to numerical methods such as linear interpolation, statistical methods like Kernel DM+V, and CNN-based models like CNN Encoder-decoder, the Transformer achieves superior performance in both accuracy and structural consistency, especially under conditions of sparse and imbalanced sampling. The model's generalization capability was validated through wind tunnel experiments. Even under conditions of significant baseline drift and low signal-to-noise ratio, the model was able to predict the approximate extent of the plume-affected region in qualitative case studies where dense physical ground truth is not available.

In summary, this thesis presents a comprehensive investigation into robotic spatial olfactory perception, encompassing sensor fusion, signal decoding, and spatial inference for gas-related tasks. The proposed solutions address critical challenges in real-world deployment, based on the properties of gas dispersion process, the need for robust modeling under sparse and uneven observations, and the complexity introduced by dynamic airflow and environmental obstacles. The contributions of this work form a task-driven pipeline that integrates sensor design, algorithmic modeling, and robotic evaluation, enabling robots to perceive gas-related features more reliably within the evaluated settings, while recognizing that broader deployment requires further validation across additional environments and longer-term sensing conditions.

6.2 Discussion and Future Work

6.2.1 Multi-modal Environmental Sampling in Robotic Olfaction

Gas dispersion is transport of matter, and its transmission process is subject to complex interactions with the physical environment. This determines that gas dispersion patterns are diverse and highly sensitive to the surrounding environment. Unfortunately, there is currently no and highly challenging to construct dataset that comprehensively captures all possible dispersion patterns. As a result, robotic spatial olfaction tasks often have to rely on small-scale datasets with limited representativeness. In such cases, robots need to actively perceive environmental features to help infer the underlying pattern of the olfactory task. The results in Chapter 4 demonstrated that incorporating physical maps can effectively address the observability issues inherent in gas source localization tasks. This confirms that expanding the sensing modalities is an effective approach to enhancing perceptual capability in small datasets.

When expanding sensing modalities, careful selection and preprocessing of data are crucial. Selecting sensing modalities based on prior knowledge and extracting task-relevant features can reduce prediction complexity and improve performance on small datasets. For example, in 2D GSL tasks, where the primary concern is the distribution of obstacles, processing LiDAR point cloud data into an occupancy grid map via SLAM can improve the signal-to-noise ratio and highlight the most relevant features. Future work could explore the integration of additional modalities such as thermal imaging, airflow, or obstacle height, depending on the environmental context and task demands. Developing modality selection strategies that dynamically adapt based on environmental cues would further enhance efficiency. Moreover, learning algorithms that explicitly model the relationships between modalities, such as cross-modal attention or graph-based fusion could help maximize the utility of each data source. Finally, creating benchmark datasets and simulation platforms with controllable modality availability would support systematic evaluation and accelerate progress in multi-modal robotic olfaction.

6.2.2 Overcoming Sim-to-real Gap in Robotic Olfaction

The Transformer-based model proposed in Chapter 6 demonstrates strong perceptual capabilities through its use of cross-attention mechanisms. This architecture proves particularly effective in handling unstructured inputs, making it a promising solution for physical field perception tasks where observations are sparse, irregular, and context-dependent. However, the results also reveal a persistent sim-to-real gap. Although the model benefits from noise-injected training data to improve generalization, real-world noise sources, such as baseline drift, sensor instability, and gas accumulation in confined corners do not follow simple Gaussian distributions. Instead, they exhibit spatial and temporal regularities driven by physical processes, which are not well captured in standard simulation pipelines.

Currently, large-scale gas dispersion datasets remain dependent on CFD simulations due to practical constraints in conducting real-world experiments. To overcome the sim-to-real gap[140,141], future datasets should incorporate noise patterns that align more closely with the underlying physical laws. This calls for more accurate multi-physics simulations that jointly model fluid dynamics, gas transport, and sensor behavior in realistic environmental conditions. For future works, firstly, attention should be given to physically realistic simulation. In order to enhance simulation fidelity by incorporating domain knowledge of gas dynamics and sensor mechanisms, allowing for the generation of synthetic data with physically meaningful noise and boundary effects. Furthermore, fine-tuning with real-world data is another promising approach to overcome sim-to-real gap. Pretraining on large-scale simulated data followed by fine-tuning on a limited amount of real experimental data can help adapt models to domain-specific deviations without requiring extensive field collection. Besides, Physics-Informed Neural Networks (PINNs) can impose physical constraints during model training to improve the realism of predicted outputs. However, due to the highly nonlinear solution space of governing equations such as the Navier-Stokes and advection-diffusion models, the loss landscape may become inconsistent with data-driven objectives. To address this, future research should explore techniques such as supervised pretraining, curriculum learning, and dynamic loss reweighting to avoid convergence to trivial or physically implausible solutions.

6.2.3 Navigation and Cooperation Strategies

The results in Chapters 5 and 6 have demonstrated that robots are capable of perceiving gas distributions, potential source and environmental structures using models. However, the robot's movement in these experiments was still either manually controlled or restricted to pre-defined paths. For practical deployment in unknown or hazardous environments, robotic systems are expected to exhibit a greater degree of autonomy. Building on the established perception models, future work should focus on developing strategies for autonomous navigation and multi-agent cooperation, enabling robots not only to sense but also to act purposefully in complex environments.

For navigation, the goal is to enable the robot to explore the environment in a manner that most efficiently reconstructs the complete gas distribution. Achieving this requires the integration of multiple information sources[92,142], including the physical structure derived from occupancy grid maps (such as obstacles and flow-blocking regions), the estimated gas distribution along with its associated uncertainty (from uncertainty-aware predictors), and the robot's own sensor coverage and motion constraints. To this end, future work may explore information-theoretic exploration strategies that leverage uncertainty maps[143] to prioritize regions with the highest expected information gain. Additionally, active sensing methods that incorporate fluid dynamics constraints can help avoid misleading measurements caused by phenomena such as gas accumulating around obstacles. Finally, navigation policies can benefit from graph-based planning within hybrid metric-topological representations, allowing the robot to balance global exploration with finer local sensing efficiency.

For cooperation, robotic olfaction tasks include multiple types of sensory information. While global modalities such as LiDAR-based occupancy maps can be shared among robots, local modalities, such as gas concentrations and wind vectors are inherently non-shareable and benefit from being measured independently by multiple agents. This motivates the use of heterogeneous multi-robot systems[144,145], where different robots carry different sensing payloads, such as gas sensors, anemometers, or thermal cameras, to enhance both efficiency and cost-effectiveness. Effective cooperation in such systems requires strategies for role allocation and load balancing, assigning robots to different tasks or regions based on their sensing capabilities. Additionally, distributed estimation and map fusion techniques are essential for asynchronously integrating local observations into a coherent global gas distribution model[146,147]. Dynamic task assignment mechanisms can further improve system robustness by allowing the team to adapt to changes in environmental conditions, such as shifting wind directions or emerging new gas sources.

To support the development of these complex decision-making strategies, reinforcement learning (RL) offers a promising framework for optimizing robot policies through experience[92,148,149]. However, standard RL suffers from low sample efficiency, especially in high-dimensional, stochastic environments like gas dispersion. A practical step forward is to build interactive simulators that support physically realistic gas distribution modeling, wind dynamics, and robot-environment interaction[131]. Such simulators can provide accelerable and controllable training environments for various robotic olfaction tasks, including source seeking, adaptive sampling, and collaborative mapping. The integration of perception, decision-making, and interaction within a unified robotic platform will also be essential for robust and scalable deployment of robotic olfaction systems in the field.

REFERENCES

References

1. Lewis, T. & Bhaganagar, K. A comprehensive review of plume source detection using unmanned vehicles for environmental sensing. Science of the Total Environment 762, 144029 (2021).

2. Kodali, R. K., Greeshma, R., Nimmanapalli, K. P. & Borra, Y. K. Y. IOT based industrial plant safety gas leakage detection system. in 2018 4th international conference on computing communication and automation (ICCCA) 1–5 (IEEE, 2018).

3. Fan, H., Hernandez Bennetts, V., Schaffernicht, E. & Lilienthal, A. J. Towards gas discrimination and mapping in emergency response scenarios using a mobile robot with an electronic nose. Sensors 19, 685 (2019).

4. Tsujita, W., Yoshino, A., Ishida, H. & Moriizumi, T. Gas sensor network for air-pollution monitoring. Sensors and Actuators B: Chemical 110, 304–311 (2005).

5. Song, Z. et al. Wireless Self-Powered High-Performance Integrated Nanostructured-Gas-Sensor Network for Future Smart Homes. ACS Nano 15, 7659–7667 (2021).

6. Somov, A. et al. Development of wireless sensor network for combustible gas monitoring. Sensors and Actuators A: Physical 171, 398–405 (2011).

7. Winkler, N. P. et al. Super-resolution for Gas Distribution Mapping. Sensors and Actuators B: Chemical 419, 136267 (2024).

8. Ishida, H., Wada, Y. & Matsukura, H. Chemical Sensing in Robotic Applications: A Review. IEEE Sensors Journal 12, 3163–3173 (2012).

9. Neumann, P. P. Gas Source Localization and Gas Distribution Mapping with a Micro-Drone. (Bundesanstalt für Materialforschung und-prüfung (BAM), 2013).

10. Loutfi, A., Coradeschi, S., Lilienthal, A. J. & Gonzalez, J. Gas distribution mapping of multiple odour sources using a mobile robot. Robotica 27, 311–319 (2009).

11. Brusca, S. et al. Theoretical and Experimental Study of Gaussian Plume Model in Small Scale System. Energy Procedia 101, 58–65 (2016).

12. Ferri, G. et al. A biologically-inspired algorithm implemented on a new highly flexible multi-agent platform for gas source localization. in The First IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics, 2006. BioRob 2006. 573–578 (IEEE, 2006).

13. Francis, A., Li, S., Griffiths, C. & Sienz, J. Gas source localization and mapping with mobile robots: A review. Journal of Field Robotics 39, 1341–1373 (2022).

14. Santos, J. E. et al. Development of the Senseiver for efficient field reconstruction from sparse observations. Nature Machine Intelligence 5, 1317–1325 (2023).

15. Xu, Q., Zhuang, Z., Pan, Y. & Wen, B. Super-resolution reconstruction of turbulent flows with a transformer-based deep learning framework. Physics of Fluids 35, (2023).

16. Drikakis, D., Kokkinakis, I. W., Fung, D. & Spottswood, S. M. Self-supervised transformers for turbulent flow time series. Physics of Fluids 36, (2024).

17. Gaillard, I., Rouquier, S. & Giorgi, D. Olfactory receptors. CMLS, Cell. Mol. Life Sci. 61, 456–469 (2004).

18. Rüffer, D., Hoehne, F. & Bühler, J. New Digital Metal-Oxide (MOx) Sensor Platform. Sensors 18, 1052 (2018).

19. Park, C. O., Fergus, J. W., Miura, N., Park, J. & Choi, A. Solid-state electrochemical gas sensors. Ionics 15, 261–284 (2009).

20. Hong, S. et al. FET-type gas sensors: A review. Sensors and Actuators B: Chemical 330, 129240 (2021).

21. Du, L., Wu, C., Liu, Q., Huang, L. & Wang, P. Recent advances in olfactory receptor-basedbiosensors. Biosensors and Bioelectronics 42, 570–580 (2013).

22. Bohrn, U. et al. Monitoring of irritant gas using a whole-cell-based sensor system. Sensors and Actuators B: Chemical 175, 208–217 (2012).

23. Schott, M., Wehrenfennig, C., Gasch, T. & Vilcinskas, A. Insect Antenna-Based Biosensors for In Situ Detection of Volatiles. in Yellow Biotechnology II: Insect Biotechnology in Plant Protection and Industry (ed. Vilcinskas, A.) 101–122 (Springer, Berlin, Heidelberg, 2013).

24. Khorramifar, A. et al. Environmental Engineering Applications of Electronic Nose Systems Based on MOX Gas Sensors. Sensors 23, 5716 (2023).

25. Torres-Tello, J., Guaman, A. V. & Ko, S.-B. Improving the Detection of Explosives in a MOX Chemical Sensors Array With LSTM Networks. IEEE Sensors Journal 20, 14302–14309 (2020).

26. Marín, D., Llano-Viles, J., Haddi, Z., Perera-Lluna, A. & Fonollosa, J. Home monitoring for older singles: A gas sensor array system. Sensors and Actuators B: Chemical 393, 134036 (2023).

27. Kumar, R., Liu, X., Zhang, J. & Kumar, M. Room-temperature gas sensors under photoactivation: From metal oxides to 2D materials. Nano-Micro Letters 12, 1–37 (2020).

28. Ou, L.-X., Liu, M.-Y., Zhu, L.-Y., Zhang, D. W. & Lu, H.-L. Recent progress on flexible room-temperature gas sensors based on metal oxide semiconductor. Nano-micro letters 14, 206 (2022).

29. Ganesh, R. S. et al. Low temperature ammonia gas sensor based on Mn-doped ZnO nanoparticle decorated microspheres. Journal of Alloys and Compounds 721, 182–190 (2017).

30. Duong, V. T., Nguyen, C. T., Luong, H. B., Nguyen, D. C. & Nguyen, H. L. Ultralow-detection limit ammonia gas sensors at room temperature based on MWCNT/WO3 nanocomposite and effect of humidity. Solid State Sciences 113, 106534 (2021).

31. Sun, D., Luo, Y., Debliquy, M. & Zhang, C. Graphene-enhanced metal oxide gas sensors at room temperature: A review. Beilstein journal of nanotechnology 9, 2832–2844 (2018).

32. Li, X. et al. Highly Sensitive Room-Temperature Detection of Ammonia in the Breath of Kidney Disease Patients Using Fe2Mo3O8/MoO2@ MoS2 Nanocomposite Gas Sensor. Advanced Science 11, 2405942 (2024).

33. Bak, S.-Y. et al. Sensitivity Improvement of Urchin-Like ZnO Nanostructures Using Two-Dimensional Electron Gas in MgZnO/ZnO. Sensors 19, 5195 (2019).

34. Serafini, M. et al. A Wearable Electrochemical Gas Sensor for Ammonia Detection. Sensors 21, 7905 (2021).

35. Singh, M. & Won, K. Diols as a novel signal enhancer for electrochemical ammonia gas sensing in ionic liquid electrolytes. Microchemical Journal 208, 112444 (2025).

36. Ji, X. et al. Electrochemical ammonia gas sensing in nonaqueous systems: A comparison of propylene carbonate with room temperature ionic liquids. Electroanalysis: An International Journal Devoted to Fundamental and Practical Aspects of Electroanalysis 19, 2194–2201 (2007).

37. Khan, M. A. et al. Recent trends in electrochemical detection of NH3, H2S and NOx gases. International Journal of Electrochemical Science 12, 1711–1733 (2017).

38. Sekhar, P. K. et al. Electrochemical gas sensor integrated with vanadium monoxide nanowires for monitoring low concentrations of ammonia emission. Journal of the Electrochemical Society 167, 027548 (2020).

39. Aliramezani, M., Norouzi, A. & Koch, C. R. A grey-box machine learning based model of an electrochemical gas sensor. Sensors and Actuators B: Chemical 321, 128414 (2020).

40. Wang, B. et al. Machine Learning-Assisted Volatile Organic Compound Gas Classification Based on Polarized Mixed-Potential Gas Sensors. ACS Appl. Mater. Interfaces 15, 6047–6057 (2023).

41. Fergus, J. W. Materials for high temperature electrochemical NOx gas sensors. Sensors and Actuators B: Chemical 121, 652–663 (2007).

42. Yoo, J. et al. Olfactory receptor-based CNT-FET sensor for the detection of DMMP as a simulant of sarin. Sensors and Actuators B: Chemical 354, 131188 (2022).

43. Wasilewski, T., Neubauer, D., Kamysz, W. & Gębicki, J. Recent progress in the development of peptide-based gas biosensors for environmental monitoring. Case Studies in Chemical and Environmental Engineering 5, 100197 (2022).

44. Wu, C. et al. Biomimetic Sensors for the Senses: Towards Better Understanding of Taste and Odor Sensation. Sensors 17, 2881 (2017).

45. Völkle, J., Kumpf, K., Feldner, A., Lieberzeit, P. & Fruhmann, P. Development of conductive molecularly imprinted polymers (cMIPs) for limonene to improve and interconnect QCM and chemiresistor sensing. Sensors and Actuators B: Chemical 356, 131293 (2022).

46. Zhang, C., Chen, P. & Hu, W. Organic field-effect transistor-based gas sensors. Chemical Society Reviews 44, 2087–2107 (2015).

47. Tang, Y.-L. et al. Ammonia gas sensors based on ZnO/SiO2 bi-layer nanofilms on ST-cut quartz surface acoustic wave devices. Sensors and Actuators B: Chemical 201, 114–121 (2014).

48. Liu, K. & Zhang, C. Volatile organic compounds gas sensor based on quartz crystal microbalance for fruit freshness detection: A review. Food Chemistry 334, 127615 (2021).

49. Mishra, S. K., Tripathi, S. N., Choudhary, V. & Gupta, B. D. SPR based fibre optic ammonia gas sensor utilizing nanocomposite film of PMMA/reduced graphene oxide prepared by in situ polymerization. Sensors and Actuators B: Chemical 199, 190–200 (2014).

50. Sun, X. et al. Enhanced sensitivity of SAW based ammonia sensor employing GO-SnO2 nanocomposites. Sensors and Actuators B: Chemical 375, 132884 (2023).

51. Choi, D. et al. Bioelectrical nose platform using odorant-binding protein as a molecular transporter mimicking human mucosa for direct gas sensing. ACS sensors 7, 3399–3408 (2022).

52. Compagnone, D. et al. Gold nanoparticles-peptide based gas sensor arrays for the detection of foodaromas. Biosensors and Bioelectronics 42, 618–625 (2013).

53. Spehr, M. & Munger, S. D. Olfactory receptors: G protein-coupled receptors and beyond. Journal of Neurochemistry 109, 1570–1583 (2009).

54. Saito, H., Chi, Q., Zhuang, H., Matsunami, H. & Mainland, J. D. Odor coding by a Mammalian receptor repertoire. Science signaling 2, ra9–ra9 (2009).

55. McCusker, E. C., Bane, S. E., O'Malley, M. A. & Robinson, A. S. Heterologous GPCR expression: a bottleneck to obtaining crystal structures. Biotechnology progress 23, 540–547 (2007).

56. Misawa, N. et al. Construction of a biohybrid odorant sensor using biological olfactory receptors embedded into bilayer lipid membrane on a chip. ACS sensors 4, 711–716 (2019).

57. Misawa, N., Mitsuno, H., Kanzaki, R. & Takeuchi, S. Highly sensitive and selective odorant sensor using living cells expressing insect olfactory receptors. Proceedings of the National Academy of Sciences 107, 15340–15344 (2010).

58. Wang, P., Liu, Q., Xu, Y., Cai, H. & Li, Y. Olfactory and taste cell sensor and its applications in biomedicine. Sensors and Actuators A: Physical 139, 131–138 (2007).

59. Zhang, S. et al. QTY code enables design of detergent-free chemokine receptors that retain ligand-binding activities. Proceedings of the National Academy of Sciences 115, E8652–E8659 (2018).

60. Looger, L. L., Dwyer, M. A., Smith, J. J. & Hellinga, H. W. Computational design of receptor and sensor proteins with novel functions. Nature 423, 185–190 (2003).

61. Latorraca, N. R., Venkatakrishnan, A. J. & Dror, R. O. GPCR dynamics: structures in motion. Chemical reviews 117, 139–155 (2017).

62. Heydel, J.-M. et al. Odorant-Binding Proteins and Xenobiotic Metabolizing Enzymes: Implications in Olfactory Perireceptor Events. The Anatomical Record 296, 1333–1345 (2013).

63. Lu, Y. et al. Olfactory biosensor using odorant-binding proteins from honeybee: Ligands of floral odors and pheromones detection by electrochemical impedance. Sensors and Actuators B: Chemical 193, 420–427 (2014).

64. Mascini, M. et al. Tailoring gas sensor arrays via the design of short peptides sequences as binding elements. Biosensors and Bioelectronics 93, 161–169 (2017).

65. Gaggiotti, S., Della Pelle, F., Mascini, M., Cichelli, A. & Compagnone, D. Peptides, DNA and MIPs in gas sensing. From the realization of the sensors to sample analysis. Sensors 20, 4433 (2020).

66. Bianchi, R. C., da Silva, E. R., Dall ‘Antonia, L. H., Ferreira, F. F. & Alves, W. A. A nonenzymatic biosensor based on gold electrodes modified with peptide self-assemblies for detecting ammonia and urea oxidation. Langmuir 30, 11464–11473 (2014).

67. Chen, X., Fu, C. & Huang, J. A Deep Q-Network for robotic odor/gas source localization: Modeling, measurement and comparative study. Measurement 183, 109725 (2021).

68. Burgués, J., Hernández, V., Lilienthal, A. J. & Marco, S. Smelling Nano Aerial Vehicle for Gas Source Localization and Mapping. Sensors 19, 478 (2019).

69. Nathan, B. J. et al. Near-Field Characterization of Methane Emission Variability from a Compressor Station Using a Model Aircraft. Environ. Sci. Technol. 49, 7896–7903 (2015).

70. He, Y. et al. Gas source localization using Dueling Deep Q-Network with an olfactory quadruped robot. International Journal of Advanced Robotic Systems 21, 17298806241255797 (2024).

71. Gifford Jr, F. Statistical properties of a fluctuating plume dispersion model. in Advances in geophysics vol. 6 117–137 (Elsevier, 1959).

72. Ishida, H., Nakamoto, T. & Moriizumi, T. Remote sensing of gas/odor source location and concentration distribution using mobile system. Sensors and Actuators B: Chemical 49, 52–57 (1998).

73. Marques, L., Nunes, U. & de Almeida, A. T. Olfaction-based mobile robot navigation. Thin solid films 418, 51–58 (2002).

74. Lilienthal, A. J., Reggente, M., Trincavelli, M., Blanco, J. L. & Gonzalez, J. A statistical approach to gas distribution modelling with mobile robots-the kernel dm+ v algorithm. in 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems 570–576 (IEEE, 2009).

75. Reggente, M. Statistical gas distribution modelling for mobile robot applications. (Örebro university, 2014).

76. Ishida, H., Nakayama, G., Nakamoto, T. & Moriizumi, T. Controlling a gas/odor plume-tracking robot based on transient responses of gas sensors. IEEE Sensors Journal 5, 537–545 (2005).

77. Lilienthal, A. J., Loutfi, A. & Duckett, T. Airborne chemical sensing with mobile robots. Sensors 6, 1616–1678 (2006).

78. Jauer, M.-L. Development and Implementation of an Ad-vanced Fuel Jettison Simulation and Comparison with the ICAO Standard Procedure. (2018).

79. Jatmiko, W., Sekiyama, K. & Fukuda, T. A pso-based mobile robot for odor source localization in dynamic advection-diffusion with obstacles environment: theory, simulation and measurement. IEEE Computational Intelligence Magazine 2, 37–51 (2007).

80. Gongora, A., Monroy, J. & Gonzalez-Jimenez, J. Gas source localization strategies for teleoperated mobile robots. An experimental analysis. in 2017 European Conference on Mobile Robots (ECMR) 1–6 (2017).

81. Kowadlo, G. & Russell, R. A. Robot odor localization: a taxonomy and survey. The International Journal of Robotics Research 27, 869–894 (2008).

82. Reddy, G., Murthy, V. N. & Vergassola, M. Olfactory sensing and navigation in turbulent environments. Annual Review of Condensed Matter Physics 13, 191–213 (2022).

83. Ojeda, P., Monroy, J. & Gonzalez-Jimenez, J. Robotic gas source localization with probabilistic mapping and online dispersion simulation. IEEE Transactions on Robotics (2024).

84. Neumann, P. P., Hernandez Bennetts, V., Lilienthal, A. J., Bartholmai, M. & Schiller, J. H. Gas source localization with a micro-drone using bio-inspired and particle filter-based algorithms. Advanced Robotics 27, 725–738 (2013).

85. Wiedemann, T., Shutin, D. & Lilienthal, A. J. Model-based gas source localization strategy for a cooperative multi-robot system—A probabilistic approach and experimental validation incorporating physical knowledge and model uncertainties. Robotics and Autonomous Systems 118, 66–79 (2019).

86. Blanco, J. L., Monroy, J. G., Lilienthal, A. & Gonzalez-Jimenez, J. A kalman filter based approach to probabilistic gas distribution mapping. in Proceedings of the 28th Annual ACM Symposium on Applied Computing 217–222 (2013).

87. Zhu, H., Wang, Y., Du, C., Zhang, Q. & Wang, W. A novel odor source localization system based on particle filtering and information entropy. Robotics and autonomous systems 132, 103619 (2020).

88. Chen, X. & Huang, J. Odor source localization algorithms on mobile robots: A review and future outlook. Robotics and Autonomous Systems 112, 123–136 (2019).

89. Jing, T., Meng, Q.-H. & Ishida, H. Recent Progress and Trend of Robot Odor Source Localization. IEEJ Transactions on Electrical and Electronic Engineering 16, 938–953 (2021).

90. Vergassola, M., Villermaux, E. & Shraiman, B. I. 'Infotaxis' as a strategy for searching without gradients. Nature 445, 406–409 (2007).

91. Wiedemann, T. Domain knowledge assisted robotic exploration and source localization. (Örebro University, 2020).

92. Wang, L., Pang, S. & Li, J. Olfactory-Based Navigation via Model-Based Reinforcement Learning and Fuzzy Inference Methods. IEEE Transactions on Fuzzy Systems 29, 3014–3027 (2021).

93. Wiedemann, T., Vlaicu, C., Josifovski, J. & Viseras, A. Robotic Information Gathering With Reinforcement Learning Assisted by Domain Knowledge: An Application to Gas Source Localization. IEEE Access 9, 13159–13172 (2021).

94. Bilgera, C., Yamamoto, A., Sawano, M., Matsukura, H. & Ishida, H. Application of convolutional long short-term memory neural networks to signals collected from a sensor network for autonomous gas source localization in outdoor environments. Sensors 18, 4484 (2018).

95. Alzubaidi, L. et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. Journal of big Data 8, 1–74 (2021).

96. Badrinarayanan, V., Kendall, A. & Cipolla, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE transactions on pattern analysis and machine intelligence 39, 2481–2495 (2017).

97. Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation. in Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18 234–241 (Springer, 2015).

98. He, K., Gkioxari, G., Dollár, P. & Girshick, R. Mask r-cnn. in Proceedings of the IEEE international conference on computer vision 2961–2969 (2017).

99. Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K. & Yuille, A. L. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. IEEE transactions on pattern analysis and machine intelligence 40, 834–848 (2017).

100. Winkler, N. P. et al. Learning from the past: Sequential deep learning for gas distribution mapping. in Iberian Robotics conference 178–188 (Springer, 2022).

101. Stachniss, C., Plagemann, C. & Lilienthal, A. J. Learning gas distribution models using sparse Gaussian process mixtures. Autonomous Robots 26, 187–202 (2009).

102. Wada, Y., Trincavelli, M., Fukazawa, Y. & Ishida, H. Collecting a database for studying gas distribution mapping and gas source localization with mobile robots. in The Abstracts of the international conference on advanced mechatronics: toward evolutionary fusion of IT and mechatronics: ICAM 2010.5 183–188 (The Japan Society of Mechanical Engineers, 2010).

103. Oberkampf, W. L. & Trucano, T. G. Verification and validation in computational fluid dynamics. Progress in aerospace sciences 38, 209–272 (2002).

104. Patel, Y., Mons, V., Marquet, O. & Rigas, G. Turbulence model augmented physics-informed neural networks for mean-flow reconstruction. Physical Review Fluids 9, 034605 (2024).

105. Zhang, D., Chen, Y. & Chen, S. Filtered partial differential equations: a robust surrogate constraint in physics-informed deep learning framework. Journal of Fluid Mechanics 999, A40 (2024).

106. Vaswani, A. et al. Attention is all you need. Advances in neural information processing systems 30, (2017).

107. Zhang, D., Liu, J., Jiang, C., Liu, A. & Xia, B. Quantitative detection of formaldehyde and ammonia gas via metal oxide-modified graphene-based sensor array combining with neural network model. Sensors and Actuators B: Chemical 240, 55–65 (2017).

108. Cai, S., Mao, Z., Wang, Z., Yin, M. & Karniadakis, G. E. Physics-informed neural networks (PINNs) for fluid mechanics: A review. Acta Mechanica Sinica 37, 1727–1738 (2021).

109. Freedman, A. N. The photoionization detector: Theory, performance and application as a low-level monitor of oil vapour. Journal of Chromatography A 190, 263–273 (1980).

110. Zimmer, C., Kallis, K. & Giebel, F. Micro-structured electron accelerator for the mobile gas ionization sensor technology. Journal of Sensors and Sensor Systems 4, 151–157 (2015).

111. Han, D., Kim, S. & Park, S. Two-dimensional ultrasonic anemometer using the directivity angle of an ultrasonic sensor. Microelectronics Journal 39, 1195–1199 (2008).

112. Abichandani, P., Lobo, D., Ford, G., Bucci, D. & Kam, M. Wind measurement and simulation techniques in multi-rotor small unmanned aerial vehicles. IEEE access 8, 54910–54927 (2020).

113. Reutebuch, S. E., Andersen, H.-E. & McGaughey, R. J. Light detection and ranging (LIDAR): an emerging tool for multiple resource inventory. Journal of forestry 103, 286–292 (2005).

114. Li, J.-G., Meng, Q.-H., Wang, Y. & Zeng, M. Odor source localization using a mobile robot in outdoor airflow environments with a particle filter algorithm. Auton Robot 30, 281–292 (2011).

115. Milstein, A. Occupancy grid maps for localization and mapping. Motion planning 381–408 (2008).

116. Siddique, N., Paheding, S., Elkin, C. P. & Devabhaktuni, V. U-net and its variants for medical image segmentation: A review of theory and applications. IEEE access 9, 82031–82057 (2021).

117. Wu, X., Hong, D. & Chanussot, J. UIU-Net: U-Net in U-Net for infrared small object detection. IEEE Transactions on Image Processing 32, 364–376 (2022).

118. Ruby, U., Yendapalli, V., & others. Binary cross entropy with deep learning technique for image classification. Int. J. Adv. Trends Comput. Sci. Eng 9, (2020).

119. Computational Fluid Dynamics. (Springer, Berlin, Heidelberg, 2009).

120. Jasak, H. OpenFOAM: Open source CFD in research and industry. International journal of naval architecture and ocean engineering 1, 89–94 (2009).

121. Woodward, P. & Colella, P. The numerical simulation of two-dimensional fluid flow with strong shocks. Journal of computational physics 54, 115–173 (1984).

122. Alfonsi, G. Reynolds-averaged Navier–Stokes equations for turbulence modeling. (2009).

123. Schmitt, F. G. About Boussinesq's turbulent viscosity hypothesis: historical remarks and a direct evaluation of its validity. Comptes Rendus Mécanique 335, 617–627 (2007).

124. Hinsen, P. et al. ICASSP 2025 SP Grand Challenge: Gas source localization from real-world spatial in-situ concentration and wind measurements. (2024).

125. Bilgic, B. et al. Fast image reconstruction with L2-regularization. Journal of magnetic resonance imaging 40, 181–191 (2014).

126. Al-Kababji, A., Bensaali, F. & Dakua, S. P. Scheduling techniques for liver segmentation: Reducelronplateau vs onecyclelr. in International conference on intelligent systems and pattern recognition 204–212 (Springer, 2022).

127. Bengio, Y., Louradour, J., Collobert, R. & Weston, J. Curriculum learning. in Proceedings of the 26th annual international conference on machine learning 41–48 (2009).

128. Aulinas, J., Petillot, Y., Salvi, J. & Lladó, X. The SLAM problem: a survey. Artificial Intelligence Research and Development 363–371 (2008).

129. Gfrerrer, A. Geometry and kinematics of the Mecanum wheel. Computer Aided Geometric Design 25, 784–791 (2008).

130. Hess, W., Kohler, D., Rapp, H. & Andor, D. Real-time loop closure in 2D LIDAR SLAM. in 2016 IEEE international conference on robotics and automation (ICRA) 1271–1278 (IEEE, 2016).

131. Erwich, H. H., Duisterhof, B. P. & de Croon, G. C. GSL-Bench: High Fidelity Gas Source Localization Benchmarking Tool. in 2024 IEEE International Conference on Robotics and Automation (ICRA) 17672–17678 (IEEE, 2024).

132. Burgués, J., Hernández, V., Lilienthal, A. J. & Marco, S. Gas distribution mapping and source localization using a 3D grid of metal oxide semiconductor sensors. Sensors and Actuators B: Chemical 304, 127309 (2020).

133. Mahmoudabadbozchelou, M., Karniadakis, G. E. & Jamali, S. nn-PINNs: Non-Newtonian physics-informed neural networks for complex fluid modeling. Soft Matter 18, 172–185 (2022).

134. Yang, G. & Sommer, S. A denoising diffusion model for fluid field prediction. arXiv preprint arXiv:2301.11661 (2023).

135. Werhahn, M., Xie, Y., Chu, M. & Thuerey, N. A multi-pass GAN for fluid flow super-resolution. Proceedings of the ACM on Computer Graphics and Interactive Techniques 2, 1–21 (2019).

136. Xie, Y., Franz, A., Chu, M. & Thuerey, N. tempoGAN: A temporally coherent, volumetric GAN for super-resolution fluid flow. ACM Transactions on Graphics (TOG) 37, 1–15 (2018).

137. Brunet, D., Vrscay, E. R. & Wang, Z. On the mathematical properties of the structural similarity index. IEEE Transactions on Image Processing 21, 1488–1499 (2011).

138. Ting, C.-C. & Chen, C.-C. Detection of gas leakage using microcolor schlieren technique. Measurement 46, 2467–2472 (2013).

139. Sabbah, S. et al. Remote sensing of gases by hyperspectral imaging: system performance and measurements. Optical Engineering 51, 111717–111717 (2012).

140. Peng, X. B., Andrychowicz, M., Zaremba, W. & Abbeel, P. Sim-to-Real Transfer of Robotic Control with Dynamics Randomization. in 2018 IEEE International Conference on Robotics and Automation (ICRA) 3803–3810 (2018).

141. Trentsios, P., Wolf, M. & Gerhard, D. Overcoming the sim-to-real gap in autonomous robots. Procedia CIRP 109, 287–292 (2022).

142. Gul, F., Rahiman, W. & Nazli Alhady, S. S. A comprehensive study for robot navigation techniques. Cogent Engineering 6, 1632046 (2019).

143. Roy, N., Burgard, W., Fox, D. & Thrun, S. Coastal navigation-mobile robot navigation with uncertainty in dynamic environments. in Proceedings 1999 IEEE international conference on robotics and automation (Cat. No. 99CH36288C) vol. 1 35–40 (IEEE, 1999).

144. Leitner, J. Multi-robot cooperation in space: A survey. 2009 advanced technologies for enhanced quality of life 144–151 (2009).

145. Yang, Q. & Parasuraman, R. Needs-driven heterogeneous multi-robot cooperation in rescue missions. in 2020 ieee international symposium on safety, security, and rescue robotics (ssrr) 252–259 (IEEE, 2020).

146. Alatise, M. B. & Hancke, G. P. A review on challenges of autonomous mobile robot and sensor fusion methods. IEEE access 8, 39830–39846 (2020).

147. Kam, M., Zhu, X. & Kalata, P. Sensor fusion for mobile robot navigation. Proceedings of the IEEE 85, 108–119 (1997).

148. Wiedemann, T., Vlaicu, C., Josifovski, J. & Viseras, A. Robotic information gathering with reinforcement learning assisted by domain knowledge: An application to gas source localization. IEEE Access 9, 13159–13172 (2021).

149. Kaelbling, L. P., Littman, M. L. & Moore, A. W. Reinforcement learning: A Survey. Journal of artificial intelligence research 4, 237–285 (1996).