IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 30 March 2025, accepted 6 May 2025, date of publication 22 May 2025, date of current version 4 June 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3572478

SURVEY

Drone-Based Sound Source Localization: A Systematic Literature Review

SÉRGIO F. CHEVTCHENKO[1], BELMAN JAHIR RODRÍGUEZ[1], RAFAELLA VALE[2], ABISHEK SOTI[1], YESHWANTH BETHI[1], NAQIB IBNUL[1], ALEXANDRE MARCIREAU[1], MOSTAFA RAHIMI AZGHADI[3], ANDREW WABNITZ[1], AND SAEED AFSHAR[1], (Member, IEEE)

[1]International Center for Neuromorphic Systems (ICNS), Western Sydney University, Penrith, NSW 2747, Australia
[2]Centro de Informática (CIn), Universidade Federal de Pernambuco (UFPE), Recife 50740-560, Brazil
[3]College of Science and Engineering, James Cook University, Townsville, QLD 4811, Australia

Corresponding author: Sérgio F. Chevtchenko (s.chevtchenko@westernsydney.edu.au)

ABSTRACT Sound source localization (SSL) using microphones mounted on uncrewed aerial vehicles (UAVs) holds significant potential for tasks ranging from search-and-rescue and gunshot detection to industrial inspection and wildlife monitoring, particularly in scenarios where camera-based sensing may be limited by poor visibility. Recent surveys take a broad view of SSL methods, with limited coverage of UAV-based approaches. This paper addresses this gap through a systematic review of UAV-based SSL, drawing on 49 studies published between 2014 and 2024. Four research questions are addressed: 1) What is the array configuration of the SSL platform? 2) What are the intended applications of SSL on UAV platforms? 3) What are the choices for the UAV platforms? 4) What are the solutions employed for performing the SSL task? The findings of this review indicate that 82% of the studies are intended for search-and-rescue contexts, 60% rely on arrays of eight microphones or fewer, and only three implement real-time onboard SSL. Notably, real-time processing of larger arrays and public availability of comprehensive datasets are identified among key obstacles to the field. This paper concludes by identifying research gaps and proposing future directions in multi-source localisation, real-time processing, and energy efficiency.

INDEX TERMS Sound source localisation, UAV, mapping study, systematic mapping.

I. INTRODUCTION

Sound source localisation is the ability to identify the origin of sounds in the environment. This capability, found in humans and animals, relies on auditory cues such as differences in arrival time and intensity of sounds at each ear. The brain interprets these differences to determine the direction of a sound, an essential skill that can often mean survival in the natural world. However, accurately localising sounds using robotics remains a significant challenge in complex, dynamic real-world environments. Although previous reviews have explored various aspects of sound source localisation (SSL), there remains a research gap regarding analyses specifically focused on drone-based SSL methods.

The recent expansion in drone technology has accelerated the development of acoustic localisation systems for uncrewed aerial vehicles (UAVs). Acoustically equipped UAVs can detect specific sound events in locations difficult or hazardous for humans to access, supporting applications such as search-and-rescue missions [1], [2], [3], [4], [5], [6], [7], [8], wildfire tracking [9], and gunshot localisation [10], [11].

However, deploying auditory sensors on UAVs introduces unique challenges. Capturing clear audio during flight is complicated by external noises, notably wind and internal (ego-noise) sources from the drone's propellers and motors, which can obscure important acoustic signals. To mitigate this, noise reduction techniques can be implemented as part of audio processing. Additionally, hardware considerations like microphone arrangement and choice of motor can be beneficial in reducing ego-noise. Operating power constraints must also be considered when working with UAV-based SSL.

The remainder of this paper is structured as follows: Section II reviews existing studies related to acoustics and UAVs, further clarifying the unique contributions of this

The associate editor coordinating the review of this manuscript and approving it for publication was Paulo Mendes.

systematic review. Section III outlines the research protocol. Section IV presents findings addressing the research questions, and Section V discusses open challenges. Limitations and mitigation strategies are detailed in Section VI. Finally, conclusions and recommendations for future research are presented in Section VII.

II. RELATED WORK

Other works have contributed to sound source localisation research by reviewing existing methods. Of note, a 2015 survey by Argentieri, Danès, and Souères [12] examines the available SSL methods in robotics, comparing binaural and array-based techniques. The unique challenges of the field, which can differ from the obstacles found in acoustics and signal processing, are identified, including embeddability of microphones, external and ego-noise, and real-time processing constraints. Their discussion recognises the need for more robust approaches due to the inherent complexities of real-world environments.

More recent investigations into this topic have also been presented. The 2020 review by Martinez-Carranza and Rascon [13] explores auditory perception techniques applied to UAVs, emphasising the various objectives, techniques, and challenges associated with the use of sound in UAV applications. The techniques are classified according to the origin of the sound and where the microphones are placed, whether air or land, and the objectives of the auditory perception process, which can include detection, classification, or localisation. The review analyses the evolution of microphone array configurations used in UAV-based auditory perception, pointing to a trend toward utilising a larger number of microphones, particularly in applications involving sound source localisation. It also discusses the challenges associated with UAV-based auditory perception, such as ego-noise generated by the UAV's motors and propellers, and the limited payload capacity that restricts the size and complexity of microphone arrays that can be mounted on UAVs.

In a 2022 review by Joshi et al. [14], an examination is made on aeroacoustics to characterise noise sources in aircraft, including the challenges involved. Acoustic-based detection techniques using these noise signatures are also studied, with a focus on beamforming and machine learning approaches. The paper highlights the importance of joining these two fields for the advancement of passive acoustic detection for applications ranging from aircraft collision avoidance to safe autonomous piloting systems.

Another 2022 review by Desai and Mehendale [15] provides a broad perspective on sound source localisation systems, focusing on techniques based on the human auditory system to determine the direction of a sound source. The review stresses SSL as an important capability for robotic ears. Different SSL techniques are categorised, including those based on conventional algorithms, relying on mathematical models like time difference of arrival, as well as methods utilising binaural signal processing and convolutional neural networks (CNNs). A discussion of various microphone array configurations and their impact on localisation accuracy is also presented.

Grumiaux et al. [16] present a survey on deep learning-based methods for SSL, covering papers published between 2011 and 2021. Their review focuses on single and multi-source localisation in indoor environments, and classifies existing approaches by neural network architecture, the type of input features, the output strategy (classification or regression), the types of data used for model training and evaluation, and the model training strategy. While this review provides detailed insights into deep learning applications for SSL, it does not specifically address challenges or adaptations related to UAVs.

A recent survey by Jekateryńczuk and Piotrowski [17] examines sound source localisation and detection methods, classifying them based on the number of microphones, spatial dimensions, and number of detectable sources. The survey analyses both classic methods and modern methods using artificial intelligence, highlighting the evolution of these techniques from classic approaches based on propagation models to more recent applications of machine learning and deep learning. The paper highlights the applications of these methods in military and civilian contexts and discusses the potential of reinforcement learning to improve accuracy and adaptability.

A comparative summary of related works is presented in Table 1. To the best of our knowledge, this paper presents the most recent research, in the form of a systematic review, into sound source localisation methods with emphasis on drone-based systems. While the 2020 work by Martinez-Carranza and Rascon [13] is of similar subject and also a review, it is not a systematic one and has a broader view of the landscape, such as different categories of auditory perception as well as the inclusion of UAVs as detection targets. In the present systematic review, the focus of contribution to the literature is research on sound source localisation methods utilising drones with mounted microphone arrays to detect and localise a given target.

III. SYSTEMATIC MAPPING STUDY PROCESS

A mapping study was conducted in accordance with the methodology laid out by Petersen et al. [18]. The structure afforded by this framework gave us the opportunity to methodically and comprehensively index related works in the field of sound source localisation using drone-mounted microphone arrays. Figure 1 describes the steps comprising the adopted methodology.

The objective of this systematic mapping study is to screen recent literature and identify gaps in the state of the art in the deployment and application of drone-mounted microphone arrays for sound source localisation. Specifically, drone parameters and SSL algorithms and their variations were noted and tabulated to analyse the efficacy of different strategies as they were adapted to their respective environments and challenges. Particular attention is given to

TABLE 1. Comparison of previous SSL review studies and the present work. The "SLR?" column identifies which approaches are systematic literature reviews. The "Review Range" column indicates the approximate coverage of primary research works analysed in each review.

Study	Year	Focus	UAV-Based?	SLR?	Review Range
[12]	2015	Robotics SSL (binaural and arrays)	Limited	No	–2014
[13]	2020	UAV auditory perception, mic configurations	Yes, but broad	No	2012–2020
[14]	2022	Aeroacoustics in aircraft (noise detection)	Partial	No	2010–2021
[15]	2022	SSL systems, traditional and CNN-based methods	Limited	No	–2022
[16]	2022	Indoor SSL with DL methods	No	No	2011–2021
[17]	2023	SSL/detection: classic vs. AI methods	Limited	No	2014–2023
Present Work	2025	UAV-mounted SSL methods	Yes	Yes	2014–2024

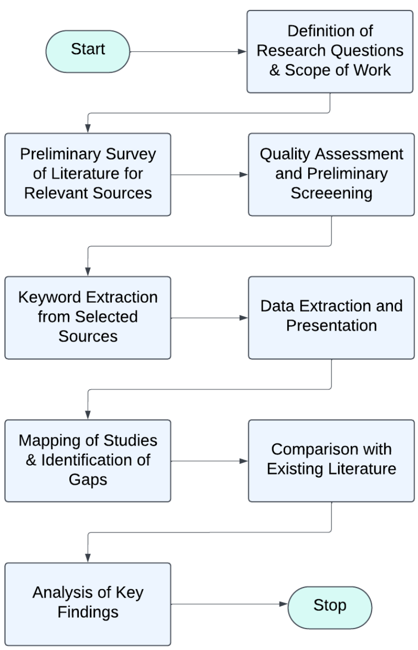

FIGURE 1. Flow chart depicting the methodology adopted.

whether SSL algorithms were executed in real time, whether onboard the drone or remotely, offline, or tested solely in simulation environments. Additionally, the study highlights the use of machine learning techniques, in particular deep neural networks, in the signal processing or localisation steps.

A. DEFINITION OF RESEARCH QUESTIONS

Initially, the state of the art in SSL using microphone arrays on drones was surveyed. The following research questions arose and were deemed suitable for the mapping study:

- RQ1: *What is the array configuration of the SSL platform?* The geometric configuration of the microphones with respect to the drone, the coordinate system used, the types of microphones, and the use case targeted in justification of the above are considered. For instance, an effort is made to determine which geometries (e.g., circular, linear, cubic, spherical) are most frequently chosen and why.
- RQ2: *What are the intended applications of SSL on UAV platforms?* The environmental factors, the breadth of sound sources, the impact of the ego-noise, and the reasoning for using a UAV as a platform are considered. Anticipating that more specialised applications would appear in only a few studies, emerging trends, such as drone-to-drone tracking and industrial inspections, are also looked for.
- RQ3: *What are the choices for the UAV platforms?* Factors such as the dimensions and the weight of the drone, the carrying capacity, the number of rotors, whether the vehicle is capable of powered flight or gliding, and which modes were used are considered. Justifications for the selection with respect to the targeted use case are noted. A strong correlation between UAV size and the chosen array design and data processing method was expected.
- RQ4: *What are the solutions employed for performing the SSL task?* The SSL algorithms used are identified. Specifically, an inquiry is made into whether any of the popular solutions, such as beamforming, MUltiple SIgnal Classification (MUSIC), or variations thereof, are used. Attention is given to whether the processing was conducted online or in post, in situ, or remotely, as well as to how the processing was implemented, including the choice of hardware such as digital signal processors (DSPs). Whether machine learning and neural networks were used in the processing is also taken into consideration. Deep learning and real-time tracking are expected to appear more

rarely, due to computational and power limitations on UAVs.

B. CONDUCT SEARCH FOR PRIMARY STUDIES

It was necessary to reduce the scope of the search for primary sources considering the breadth and volume of publications in the areas of applications for UAVs, SSLs, and machine learning. To this end, a search string to use across databases was created. This process is key in transitioning from step 1 to step 2 in the mapping study methodology. The string consists of keywords defined in abstracts in relevant papers on the topic and their synonyms and equivalent phrases which are delineated with the logical operators "AND" and "OR" and parentheses for syntax and the construction of the appropriate associations. This resulted in the following string: ("UAV" OR "Drone") AND (("Acoustic" OR "Sound") "Source Localization").

The databases considered for the search of the primary studies were the ACM Digital Library,[1] IEEE Xplore,[2] Web of Science,[3] Science Direct,[4] and Scopus.[5]

C. SCREENING OF PAPERS FOR INCLUSION AND EXCLUSION

Given the large number of articles unrelated to the present research topics, specific inclusion and exclusion criteria were established. These criteria are intended to narrow down the search and are a part of the preliminary screening process for quality assessment.

The inclusion criteria were:

- Publications that propose or deploy a drone-mounted microphone array for sound source localisation;
- Original works that were peer-reviewed;
- Papers that were published between 2014 and 2024.

The exclusion criteria were:

- Articles that do not implement SSL;
- Articles that did not use drones for mounting a microphone array;
- Articles in languages other than English;
- Articles in which the setup or findings were not clearly stated;
- Articles that were duplicates, found in multiple databases, or published in multiple sites;
- Works that were not original research, such as review papers. It is worth noting that the related review papers are still listed in Section II.

IV. RESULTS AND DISCUSSION

Tables 2 and 3 summarise the parameters and algorithms from the selected studies. Table 2 highlights UAV-related aspects such as microphone placement, array configuration, maximum tested distance, and the drone model or category (e.g., micro, very small, small). Table 3 focuses on the SSL algorithms used, the number of microphones, sampling rates, and considerations for multiple sound sources. The following sections discuss these findings in greater detail.

A. DESCRIPTIVE ANALYSIS

In May 2024, a thorough survey of the literature within the last decade was conducted. This search resulted in 296 papers, as illustrated in Figure 2. Of these, 99 were from Scopus, 86 from the ISI Web of Science, 78 from the IEEE Digital Library, 28 from the ACM Digital Library, and 5 from Science Direct. Out of this total, 108 were identified as duplicates. A further 126 were rejected as they did not address the research questions or were only relevant in some part but not all. In the end, 62 papers were selected for full-text reading.

During the full-text reading, a further 11 papers were rejected as they did not address the research questions directly, leaving 49 papers for the systematic review. A generally growing trend of publications can be observed in Figure 3.

Of the papers selected, 31 were published in an IEEE journal or conference; one in the Taylor & Francis Journal of Advanced Robotics; three in MDPI journals; three in journals by the Fuji Technology Press Ltd.; one in the Journal of Robotics and Mechatronics published by The Robotics Society of Japan; two by the Department of Mathematics, Computer Science and Physics at the University of Udine; one as a paper in the Automatic Target Recognition XXVIII proceedings by SPIE; one in a journal published by the Aeronautical and Astronautical Society of the Republic of China; one as a chapter in the book titled Unmanned Aerial Systems published by Elsevier; one as a journal article published by PLOS Computational Biology; two as articles in the EURASIP Journal on Audio, Speech, and Music Processing; one in the International Conference in Communications, Signal Processing, and Systems published by Springer; and one as a conference paper published by Universidade Federal do ABC.

B. WHAT IS THE ARRAY CONFIGURATION OF THE SSL PLATFORM?

1) ARRAY GEOMETRY AND SHAPE: INFLUENCE ON PERFORMANCE

The selection of the array geometry is essential in the SSL performance. It affects the spacial possibilities of localisation (azimuth and elevation), the frequency bands of the acoustics waves of the interest sound sources, the computational processing and the potential for longer-distance detection between the source and the microphones. For instance, in linear arrays, the array's frequency bands are determined by the time difference of arrival (TDOA), which depends on the shortest distance in the array (defining the highest frequency) and the longest distance between any two microphones in the array configuration (defining the lowest

[1]https://dl.acm.org
[2]IEEExplore.ieee.org/Xplore/home.jsp
[3]https://www.webofscience.com
[4]https://www.sciencedirect.com/
[5]https://www.scopus.com/

TABLE 2. Comparison of UAV-related parameters in the selected studies.

Ref.	Max Dist. (m)	Mic Placement Relative to UAV	Array Configurations	Drone Model(s)	Drone Category
[19]	35	Below	Circular	DJI Matrice 100	Small
[20]	6	Above	Circular	3DR IRIS	Very Small
[21]	2.4	Below	Cubic	MK-Quadro	Very Small
[22]	3	Above	Circular	3DR IRIS	Very Small
[23]	150	Around	Triangular, Tetrahedral	N/A	Micro, Very Small
[24]	200	Below	Circular	Custom Kiteplane	Small
[25]	N/A	Below	Cubic	N/A	Simulated
[5]	5	Above	Linear	Custom Built	Very Small
[26]	5	Below	Cylinder	N/A	Simulation
[10]	180	N/A	N/A	AR.Drone 2.0	Very Small
[27]	10	Around	Cylinder	enRoute PG-560	Very Small
[28]	151.5	Below	Circular	DJI Matrice 600	Small
[29]	6.4	Above, Below	Circular	N/A	N/A
[30]	22	Around	Sphere	N/A	N/A
[2]	30	Below	Rectangle	N/A	N/A
[7]	N/A	Below	Cubic	N/A	N/A
[31]	18	Around	Cylinder	ZION-PG560	Small
[32]	11	Around, Below	Circular	AscTech Pelican, enRoute Zion, Parrot Bebop	Very Small, Small
[33]	N/A	One Side	Sphere	N/A	Small
[34]	50	One Side	Sphere	DJI Inspire 2	Small
[35]	N/A	Around	Linear	AR.Drone 2.0	Micro
[36]	1	Above	Linear	Parrot Bebop	Very Small
[37]	60	N/A	Sphere	N/A	N/A
[38]	10	One Side	Sphere	MS-06LA	Small
[39]	20	Below	Linear	DJI Mavic Pro	Very Small
[40]	2	Above, One Side	Sphere	Matrice 100	Small
[41]	3	Above	Circular	3DR IRIS	Very Small
[42]	N/A	Around, One Side	Sphere	N/A	Small
[43]	6	Above	Circular	3DR IRIS	Very Small
[44]	N/A	N/A	Circular	N/A	N/A
[45]	N/A	Below	Cubic	N/A	N/A
[46]	20	One Side	Sphere	N/A	Small
[6]	1.6	One Side	Rectangle	Autel X-Star	Very Small
[47]	N/A	Below	Cubic	N/A	N/A
[1]	N/A	One Side	Sphere	MS-06LA	Small
[48]	N/A	Below	Sphere	N/A	N/A
[49]	0.8	One Side	Linear	Crazyflie 2.0	Micro
[50]	1	Around	Linear	AR.Drone 2.0	Micro
[51]	4	Above	Linear	Parrot Bebop	Very Small
[11]	100	Around	Rectangle	DJI Phantom 4	Very Small
[8]	N/A	Below	Cubic	N/A	N/A
[52]	5	Multiple Sides, One Side	Sphere	N/A	N/A
[53]	3	Around	Circular	AscTech Pelican	Very Small
[54]	30	N/A	Sphere	DJI S1000, DJI Inspire	Small
[55]	0.6	Around	Other	N/A	Micro
[56]	3	Around	Circular	AscTech Pelican	Very Small
[57]	3	Below	Circular	DJI M300 RTK	Small
[58]	N/A	Below	Cubic	N/A	N/A
[59]	2	Above	Hexagonal	N/A	N/A

frequency), as described by the following equations,

$$f_{\max} = \frac{c}{2d_{\min}} \tag{1}$$

$$f_{\min} = \frac{c}{2d_{\max}} \tag{2}$$

where

- $f_{\max}$ is the maximum frequency,
- $f_{\min}$ is the minimum frequency,
- c is the speed of sound in the medium (approximately 343 m/s in air),
- $d_{\min}$ is the minimum distance between two microphones in the array,
- $d_{\max}$ is the maximum distance between two microphones in the array.

The array configurations used in each of the reviewed works are presented in Table 2. According to Figure 4, the arrangement of distributed microphones in the arrays offers a variety of geometric shapes. Based on the collected data, the most common configurations are circular (14 instances) and spherical arrays (11 instances). Other authors use the DRone EGonoise and localizatiON (DREGON) dataset [21], which implements a cubic array with 7 instances, followed by linear arrays (7 instances). Other utilised geometries include cylindrical and rectangular arrays (3 instances). Finally, tetrahedral and hexagonal arrays are rarely used.

Configurations of array geometries can vary from simple 2D layouts to more complex 3D arrangements, with the specific arrangement affecting the localisation accuracy and the computational demands. The following analysis provides insights into the commonly used numbers of microphones and array configurations as observed across various studies.

Figure 5 displays the distribution of microphone array configurations according to the number of microphones used across different reviewed articles. Each colour in the chart represents a unique configuration, such as circular, cubic, cylinder, hexagon, linear, sphere, and rectangle among others, providing an overview of the preferred configurations and microphones in UAV-based direction of arrival (DOA) research.

A clear preference emerges for 8-microphone arrays, particularly in circular and cubic configurations. This indicates the popularity of the DREGON dataset [21], which is widely

TABLE 3. Comparison of SSL algorithms in the selected studies.

Ref.	# of Mics	SSL Baseline Algorithms	Sampling Rate (kHz)	Signal Source	Multiple Sources	Execution
[19]	8	MUSIC, SRP-PHAT	48	Human voice, Whistle, White noise		Offline
[20]	8	SRP-PHAT	N/A	Human voice		Simulation
[21]	8	SEVD-MUSIC, GEVD-MUSIC	44.1	Human voice, White noise		Offline
[22]	8	GEVD-MUSIC, SRP-PHAT	8	Human voice		Simulation
[23]	4	GCC-PHAT	40	Other, Drones	✓	Local
[24]	8	iGEVD-MUSIC	N/A	Human voice		Simulation
[25]	8	GCC-PHAT, GCC-NONLIN	44.1	Human voice	✓	Simulation
[5]	4	MUSIC	N/A	Human voice		Remote
[26]	6	MUSIC	16	Other		Simulation
[10]	5	GCC-PHAT	N/A	Gun shots		Simulation
[27]	16	GSVD-MUSIC	N/A	Human voice, Whistle	✓	Remote
[28]	32	DS-beamforming	25.6	Other		Offline
[29]	8	SRP-PHAT	8	Human voice, Other		Simulation
[30]	16	MUSIC	16	Other, Drones		Simulation
[2]	4	GCC	50	Whistle		Local
[7]	8	DS-beamforming, MVDR, GCC-PHAT, GCC-NONLIN	44.1	Human voice		Offline
[31]	16	GSVD-MUSIC	16	Human voice, Whistle	✓	Remote
[32]	8	iGEVD-MUSIC, iGSVD-MUSIC	16	Speech, Animal sounds, Vehicles, Other, Whistle	✓	Remote
[33]	12	MUSIC, SEVD-MUSIC, iGSVD-MUSIC	16	Whistle		Remote
[34]	16	MUSIC	16	Whistle		Simulation
[35]	2	iGEVD-MUSIC, DS-beamforming	N/A	Other		Offline
[36]	4	DU-beamforming, MUSIC	16	Human voice, Whistle		Simulation
[37]	16	MUSIC	16	Human voice	✓	Simulation
[38]	12	SEVD-MUSIC, iGSVD-MUSIC	16	Human voice, Whistle		Simulation
[39]	2	GCC-PHAT	N/A	Human voice		Offline
[40]	8	GCC-PHAT	8	Human voice		Simulation
[41]	8	MUSIC, GEVD-MUSIC, SRP-PHAT	8	Human voice		Simulation
[42]	16	MUSIC, GEVD-MUSIC	N/A	Human voice, Whistle	✓	Remote
[43]	8	DS-beamforming	8	Human voice		Simulation
[44]	21	DS-beamforming, MVDR, SRP-PHAT	N/A	Other	✓	Simulation
[45]	8	GCC-PHAT, SRP-PHAT, MVDR, DS-beamforming	44.1	Human voice, White noise		Simulation
[46]	16	MUSIC	N/A	Other (chirp)	✓	Remote
[6]	64	DS-beamforming	50	Human voice, Drones		Remote
[47]	8	N/A	16	Human voice		Simulation
[1]	12	SEVD-MUSIC, iGSVD-MUSIC	N/A	Human voice, Whistle		Remote
[48]	19	SH-DU-FSPT	48	Drones		Offline
[49]	2	N/A	144	Animal sounds		Local
[50]	2	DS-beamforming, iGEVD-MUSIC	N/A	Human voice		Offline
[51]	4	N/A	N/A	Human voice, Drones		Remote
[11]	4	GCC-PHAT	N/A	Gun shots		Offline
[8]	8	GCC-PHAT	44.1	Human voice		Simulation
[52]	16	MUSIC, SEVD-MUSIC	44.1	White noise		Simulation
[53]	16	SEVD-MUSIC, iGEVD-MUSIC, iGEVD-CMS, iGSVD-MUSIC	16	Human voice		Offline
[54]	19	MUSIC	N/A	Drones	✓	Simulation
[55]	15	MUSIC, GEVD-MUSIC, GCC-PHAT	48	Human voice	✓	Simulation
[56]	16	iGSVD-MUSIC	16	Speech, Animal sounds, Vehicles, Other	✓	Offline
[57]	N/A	DS-beamforming	N/A	Electrical discharge		Offline
[58]	8	MUSIC, GCC-PHAT, GCC-NONLIN	16	Human voice		Simulation
[59]	6	MUSIC, CSSM, SRP-PHAT, TOPS, WAVES	44.1	Other (chirp)		Offline

used in sound localisation studies due to its standardised 8-channel cubic setup as seen in Figure 6.

The frequent use of circular configurations and the cubic arrangement suggests a practical approach in planar setups that are simpler to implement and process. With 16 microphones, a wider range of configurations is observed, including spherical, circular and cylindrical shapes. The presence of 3D configurations highlights a move towards capturing a fuller acoustic environment, which can provide more precise localisation capabilities in azimuth and elevation angles.

In contrast, arrays with larger microphone counts, such as 32 and 64 microphones, are sparsely used and are primarily configured in simpler 2D layouts like circular (Figure 7) and rectangular (Figure 14).

This limited adoption of larger arrays may reflect practical constraints, as managing data from such arrays can be computationally demanding and may exceed the hardware limitations of UAV platforms, especially for real-time processing applications. For smaller arrays, typically those with two or four elements, 1D configurations, especially linear, are preferred. These simpler, planar setups provide a feasible solution for UAVs where space and computational resources are limited.

2) NUMBER OF MICROPHONES ON UAVS: KEY CONSIDERATIONS

In UAV-based sound source localisation, the number of microphones used typically ranges from 2, 4, 8 and 16, reflecting a symmetry in the array configurations and balance between achieving sufficient spatial resolution for accurate direction of arrival estimation and processing limitation of UAV platforms. The choice of the number of microphones is influenced by both the limited space available on UAVs and the hardware's processing capacity. Smaller arrays, such as those with four microphones, are often chosen for real-time

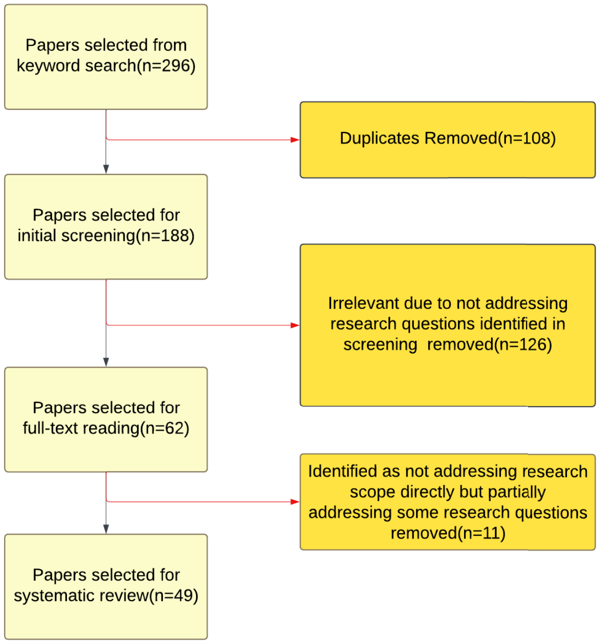

FIGURE 2. Flow chart of the paper selection process.

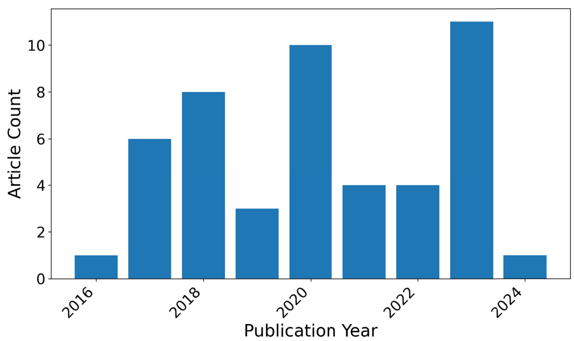

FIGURE 3. Number of primary studies by publication year.

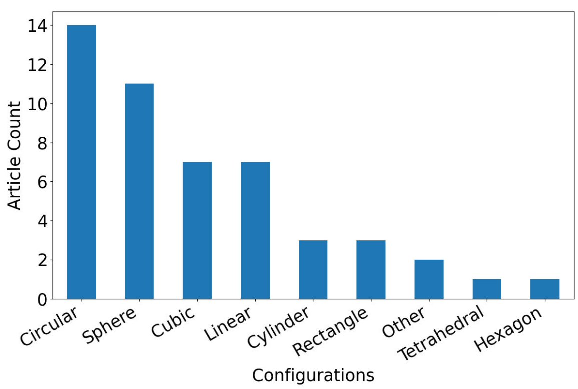

FIGURE 4. Different shapes of arrays.

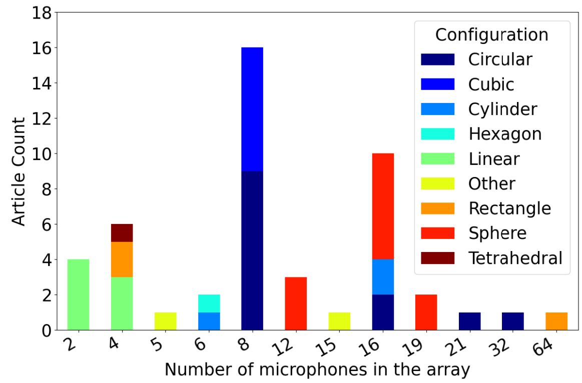

FIGURE 5. Array geometries and number of microphones.

FIGURE 6. The MiKroKopter quadrotor UAV used for the DREGON dataset, with 3D printed 8-channel microphone array mounted on the bottom. Green circles highlight two of the passive markers used for motion capture and yellow circles highlight two of the microphones (best seen in colours). Copyright ©2017 IEEE [21].

FIGURE 7. Four microphone modules, each with eight MEMS microphones, as shown in Figure 13 [28].

local processing directly on the UAV due to the limits of onboard computing power and energy efficiency. Larger arrays, like those with 8 or 16 microphones, are feasible for remote processing or offline analysis where data can be transmitted to external servers or processed post-flight. This distribution reflects the need to optimise performance within the unique hardware limitations of UAVs while maintaining sufficient localisation accuracy.

These data are summarised in Table 3, where the consulted references are listed, with the number of microphones under the column "# of Mics" and the DOA execution location under the column "Execution."

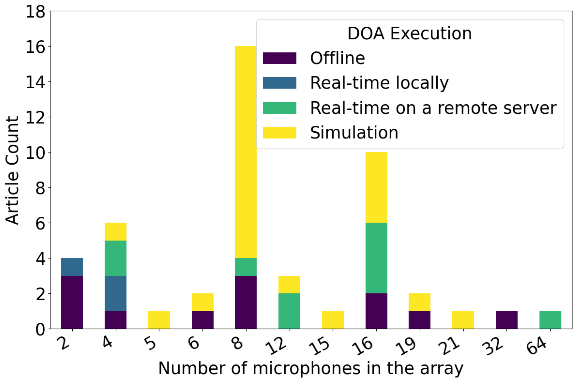

FIGURE 8. Distribution of DOA approaches vs. number of microphones.

Figure 8 demonstrates distinct trends in the use of microphone arrays for DOA tasks on UAVs, influenced by the execution location: **offline, real-time locally, real-time on a remote server, and simulation**.

For real-time local processing, smaller microphone arrays, typically with four or fewer microphones, are most common. This is because real-time local processing requires the computation to be performed directly on the UAV, which is limited by onboard processing power and energy resources. This point is further addressed in Section V. For instance, the use of two microphones in [49] is focused on mimicking the sound localisation and tracking capabilities of bats using a biomimetic miniature drone equipped with two microphones optimal for ultrasonic signals and an Atmel AVR32 micro-controller processing unit with 144 kHz sampling rate, to estimate the azimuth of the sound source in real time. The drone in Figure 9 successfully tracked the position of a moving sound-emitting target, mimicking a bat during search phase.

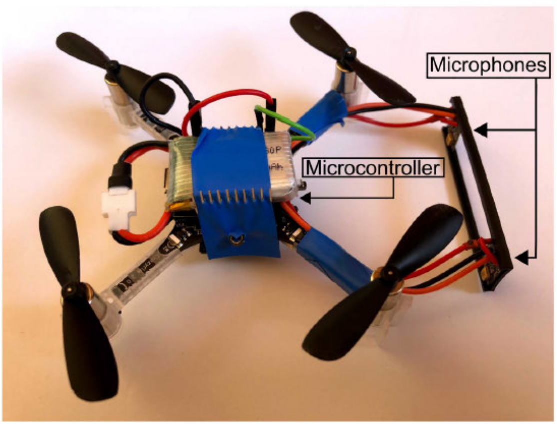

FIGURE 9. Image of the system with real-time local DOA execution. Drone with embedded electronics and two microphones for azimuth localisation [49].

Sibanyoni et al. [2] developed a 2D acoustic source localisation system for drones in search-and-rescue missions using an array of four microphones and a digital signal processor. The system estimated both the azimuth and the distance to the sound source and was able to localise a whistle blown at various distances. Using fewer microphones reduces both the computational load and power consumption, enabling real-time execution within the hardware limitations of the drone. More complex tasks, such as 2D sound localisation (azimuth & elevation), often utilise 4-microphone arrays for enhanced accuracy and resolution. A final instance of real-time onboard execution is found in Basiri et al. [23], with implementation of an onboard system for relative bearing estimation for teams of drones using sound. They tested both passive and active audio-based systems, with drones generating unique chirp sounds to aid in localisation. The drone utilised a 4-microphone array in a tetrahedral configuration, as illustrated in Figure 10. The active system achieved a detection in azimuth and elevation. The paper also demonstrated a fully autonomous leader-following behaviour.

FIGURE 10. Pictures of the two drones used for real-time onboard DOA. The microphones are indicated by red circles. (a) A small drone (440 grams) with a microphone array (21 grams) used for outdoor experiments. (b) A pocket-sized drone (31.4 grams) with a microphone array (3.5 grams) used for indoor experiments. Copyright ©2016 IEEE [23].

In the case of real-time remote processing, slightly larger arrays, such as with 8 or 16 microphones, are more feasible, as the computation is offloaded to a remote server. This approach bypasses the UAV's processing limitations, allowing for higher resolution in sound localisation. However, remote processing depends on reliable data transmission, which can introduce latency and may be sensitive to network conditions.

Offline processing appears across various microphone configurations, from smaller to mid-sized arrays (4, 8, and 16 microphones). Offline processing involves recording data during the UAV's flight and analysing it later on a more powerful, stationary system. This method is advantageous as it allows researchers to use larger arrays and more computationally intensive algorithms that would be impractical for real-time, onboard processing. Offline processing supports more detailed post-flight analysis, making it a preferred approach for testing and optimising algorithms without the constraints of real-time execution.

Finally, simulation is highly represented, particularly with 8-microphone arrays, primarily due to the availability of the

DREGON dataset. This dataset is widely used to simulate and evaluate sound localisation algorithms in a controlled setting, making it a standard tool for algorithm development. Simulations allow researchers to assess algorithm performance under controlled conditions before deploying them on actual UAVs, where hardware and environmental factors come into play.

3) MICROPHONE CHARACTERISATION: SAMPLING RATES AND CONSTRUCTION

Micro-electro-mechanical system (MEMS) microphones are known for their compact size, high signal-to-noise ratio (SNR), low power consumption, and high sensitivity [6]. These are commonly found in UAV-based acoustic applications. The combination of MEMS microphones with field-programmable gate array (FPGA)-based architectures is frequent in sound source localisation applications [60]. These features make them commercially advantageous for UAV integration, where size and weight constraints are crucial. Additionally, MEMS microphones offer a nearly flat frequency response, making them suitable for capturing a wide range of acoustic signals.

The integration of a microphone preamplifier and an analog-to-digital converter (ADC) in a single MEMS microphone chip further reduces costs and space, facilitating the development of systems with minimal components. One example is a 32-channel, time-synchronised MEMS microphone array mounted on a drone for acoustic source localisation [28]. In this work, the authors assembled four modules, each containing eight MEMS microphones, as illustrated in Figure 13 and Figure 7, positioned at four different locations on the drone.

Apart from custom-built arrays, some studies utilise commercially available microphone arrays for sound source localisation on UAVs. Two studies used the linear 4-microphone array from the PlayStation Eye. This array, integrated with a camera, was mounted on a UAV to detect humans. The PlayStation Eye's array microphone offered a practical and readily available solution for the research in [5] and [19], as illustrated in Figure 11.

Another example is research employing a commercially available microphone array designed for ambisonics audio-visual productions. The ZM-1 array uses a spherical harmonic multiple signal classification technique. Figure 12 illustrates this 19-channel system with MEMS microphones suspended below a drone.

Finally, one study implemented a 64-microphone array on a fixed structure instead of a drone, simulating the drone's location for sound source localisation. This study used a square array of 64 (8 × 8) MEMS microphones uniformly spaced on a rectangular printed circuit board as illustrated in Figure 14. This system employed a myRIO platform, and the setup allowed the researchers to evaluate the acoustic localisation algorithms in the presence of drone propeller noise without integrating the array onto the drone [6].

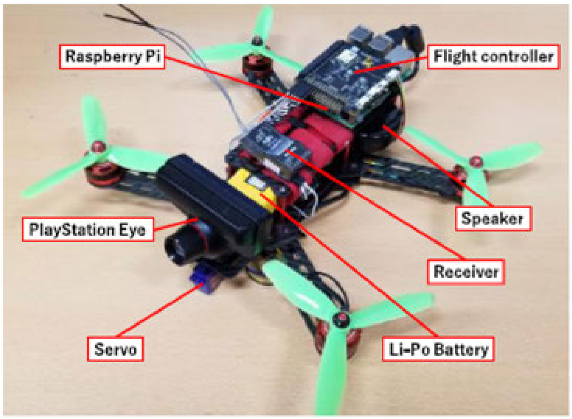

FIGURE 11. PlayStation Eye microphone array [5].

FIGURE 12. ZM-1 Ambisonics with 19 MEMS. The red arrow indicates the position of the microphone. Copyright ©2021 IEEE [48].

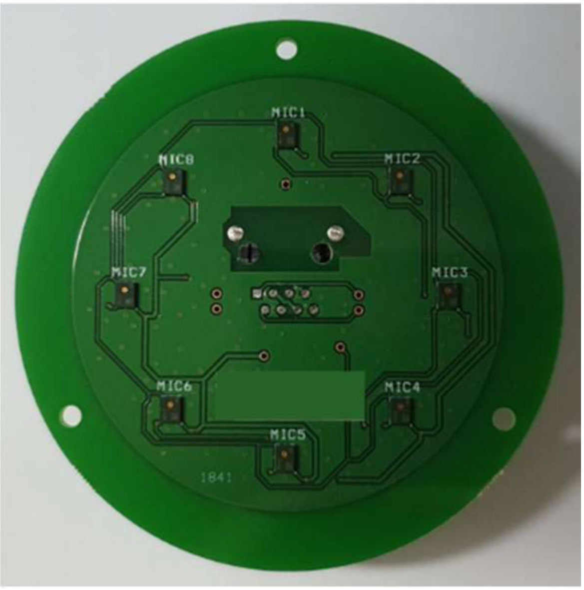

FIGURE 13. A module consisting of 8 MEMS microphones [28].

The sampling frequency is another crucial factor that directly relates to the acoustic nature of the sound source under investigation. A higher sampling rate helps capture

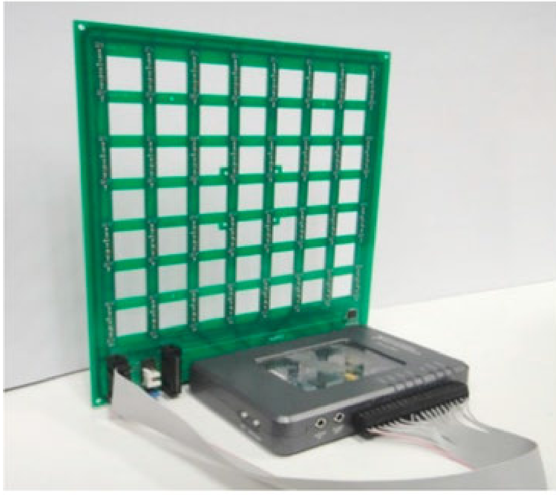

FIGURE 14. Array module with myRIO FPGA and 64 MEMS microphones board [6].

high-frequency details and avoids aliasing, which can distort the signal when the source has pronounced high-frequency components. Conversely, if the sound source is largely low-frequency, a lower sampling rate may be sufficient.

The studies reviewed most often adopt a sampling rate of 16 kHz, likely due to its balanced approach. It retains the key frequency range necessary for accurate localisation of human sounds (speech and whistle), while avoiding excessive data size and computational overhead. This balance is especially important for real-time or near-real-time implementations on hardware-limited UAV platforms. Figure 15 shows the distribution of sampling rates used across various studies for UAV-based sound source localisation.

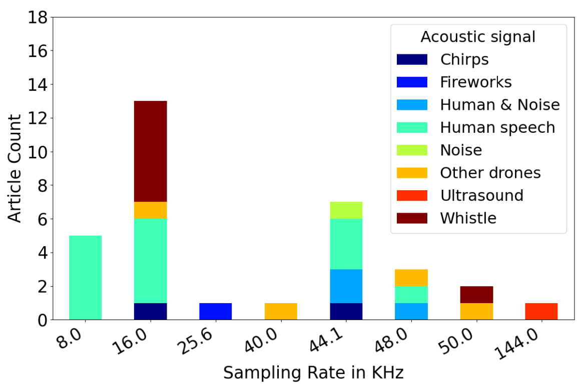

FIGURE 15. Audio sampling rate used in the reviewed works.

The 8 kHz sampling rate, also prominent, is primarily chosen in studies focused on human voice detection. Since human speech typically falls within the lower frequency range, 8 kHz is adequate for capturing the necessary information while minimising data size and computational demands—making it well-suited for UAVs with limited processing capabilities. The 44.1 kHz sampling rate is another widely used rate, strongly associated with the DREGON dataset. This CD quality rate provides high fidelity, allowing researchers to evaluate algorithms under optimal conditions and ensuring comprehensive coverage of the audio spectrum for more detailed analysis.

Less common sampling rates, such as 48 kHz and 50 kHz, appear in studies that may require higher resolution audio to capture finer spatial details in localisation tasks, provided the UAV platform can support the increased processing demands. Uncommon sampling rates of 40 kHz, 25.6 kHz, and 144 kHz are noted, with the latter employed in a biomimetic system on a drone for detecting ultrasonic signals [49].

4) MICROPHONE PLACEMENT STRATEGIES

One of the most significant challenges in sound source localisation with UAVs is dealing with low signal-to-noise ratios, as the microphone array is often near the drone's noise sources. Researchers have experimented with different microphone array placements on drone platforms to address this issue. Maintaining a balanced centre of gravity is essential for stability and effective manoeuvrability. The most common positions are below, above, and around the drone.

A study by Salvati et al. [19] utilised a uniform circular array (UCA) suspended one meter below a quadcopter by load-bearing ropes, as illustrated in Figure 16. This arrangement aimed to enhance the signal-to-propeller noise ratio by orienting the microphone array downward. This setup works well up to -19 dB of SNR, minimising the direct influence of propeller noise, but it affects manoeuvrability and SSL precision.

FIGURE 16. Circular microphone array mounted on the bottom of the Matrice 100 quadcopter, through a set of four nylon cords of 1 m lengtheach. Copyright ©2020 IEEE [19].

An alternative is to position the microphone array on one side of the drone, as shown in Figure 17. This setup aims to maximise distance from the drone's noise sources. However, it can lead to an unbalanced load, impacting stability and manoeuvrability. To address this, other researchers have employed counterbalance weights symmetrically placed around the drone as illustrated in Figure 18.

FIGURE 17. DJI Inspire 2 drone and 16-channel microphone array on one side of the drone. The distance between the drone centre and the microphone array is 50 centimetres [34].

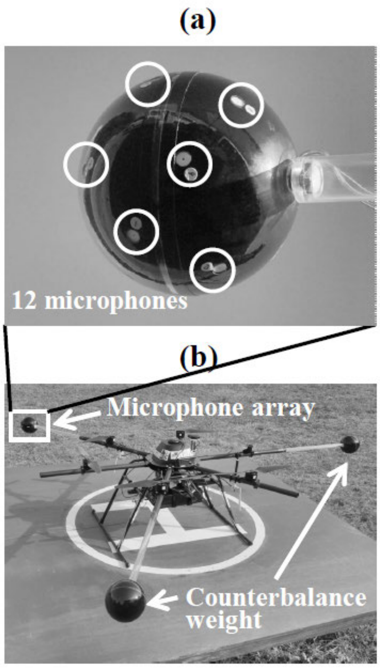

FIGURE 18. Microphone array with waterproof protection and counterbalance loads [38].

C. WHAT ARE THE INTENDED APPLICATIONS OF SSL ON UAV PLATFORMS?

As per Figure 19, the majority of the reviewed works focus on the localisation of human speech and whistle sounds, which suggests that search-and-rescue is the most common application for drone-based SSL systems. By using sound, drones can locate people in distress under adverse lighting conditions or in the presence of visual occlusion, such as in a forest [61].

Drone-based acoustic sensors have promising applications in industrial scenarios. For example, Li et al. [57] investigated the use of acoustic sensing to address the problem of electrical discharge location. Similarly, gas leakage detection and localisation represent another potential application [62]. However, this review did not identify any existing drone-based solutions for such tasks. While the above industrial applications remain underexplored, they often require the use of microphones in the ultrasonic range [63], which further increases the computational and power requirements due to higher sampling rates.

Notably, Hoshiba et al. [32] evaluated drone-based sound source localisation for 21 different sound types, including whistles, human speech, vehicle noise, crow calls, motorbike engines, horns, and alarms. Their system, employing a UAV-embedded microphone array, demonstrated localisation capabilities within a maximum range of 11 meters. A related work, also by Hoshiba et al. [33], introduces a spherical microphone array system designed for outdoor environments, illustrated in Figure 18. This system improves noise robustness and localisation accuracy, particularly for dynamic and noisy conditions. Despite these advancements, the range limitations of such systems remain a critical factor, especially in applications requiring localisation over greater distances.

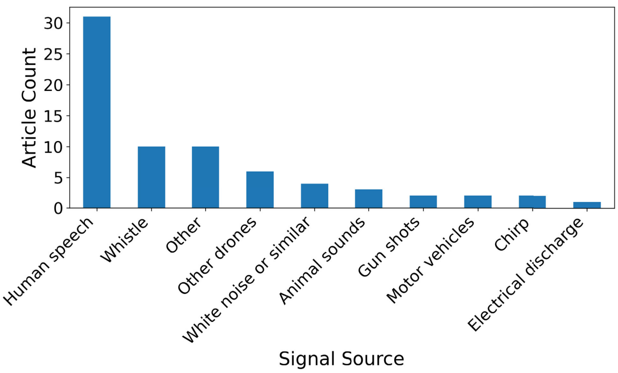

FIGURE 19. Variety of the reported sound sources.

The maximum distance from the drone to the target, as reported in each of the surveyed works, is presented in Table 2 and summarised in Figure 20. It is worth noting that only five of these works report localisation at distances greater than 100 m [10], [11], [23], [24], [28]. However, these localisation distances are either reported in a simulated environment [24], with engines off [23], or for specific high-energy sounds, such as gun shots [10], [11] or firecrackers [28]. This narrow focus on shorter-range and single-source detection underscores a difficulty with more challenging and realistic tasks. This could be addressed by new public datasets and data collection in realistic outdoor scenarios, as described in Section V-A.

D. WHAT ARE THE CHOICES FOR THE UAV PLATFORMS?

The choice for UAV platforms is driven by several factors, including payload limitations, where smaller drones require lightweight setups and energy efficiency, as larger arrays increase power consumption. In the reviewed works,

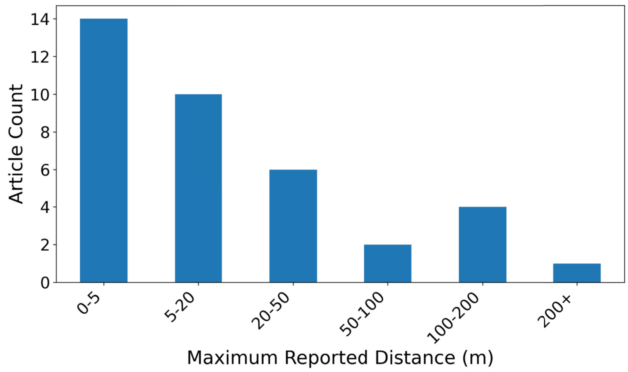

FIGURE 20. Maximum reported distances from the UAV to the sound source.

various applications of SSL corresponding to drone size are observed. Drones are categorised based on the following classifications: micro (250 g or less), very small (more than 250 g but not exceeding 2 kg), and small (more than 2 kg but not exceeding 25 kg). The drone size is not specified if it cannot be inferred from the information provided in the article. The distribution of drone sizes in the selected works is illustrated in Figure 21, with simulation as an additional category. The use of a simulated UAV allows for the exploration of high-density arrays without physical constraints. Real-world scenarios prioritise the balance between array size and flight performance, while simulations emphasise algorithm development under ideal conditions.

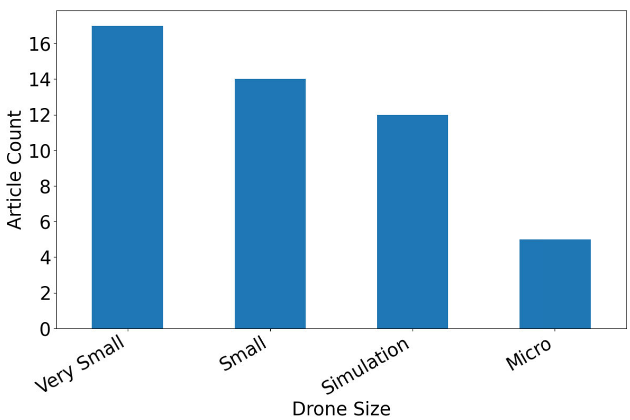

FIGURE 21. Drone category (Micro: 250 g or less; Very Small: more than 250 g but not more than 2 kg; Small: more than 2 kg but not more than 25 kg).

Figure 22 illustrates the distribution of microphone numbers in arrays mounted on drones of various sizes. Arrays with 2–5 microphones are mostly used on micro and very small drones due to weight and space limitations, whereas 8 microphones represent the most studied configuration, particularly in simulations and small drones, due to a balance between spatial resolution and payload capacity. Simulation represents setups using simulated environments such as

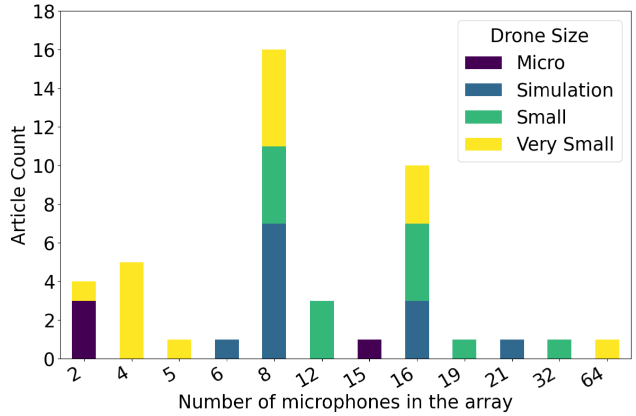

FIGURE 22. Number of microphones mounted on the drone, considering the drone size.

Pyroomacoustics or processing with third-party recorded datasets like DREGON. Larger arrays like 64-microphone arrays are a special case where the drone generates noise near an FPGA-based system (myRIO) rather than physically carrying the array.

The only case involving a significant number of microphones mounted on a micro category drone is the work proposed by Manamperi et al. [55]. In this study, the experimental setup features a drone equipped with 15 MEMS microphone modules (ICS-43432), as illustrated in Figure 23, distributed across various parts of its body, and a sound source placed on the ground. The experiments were conducted in a semi-anechoic chamber, where the drone was fixed on a rigid stand and connected to the processing system via cables.

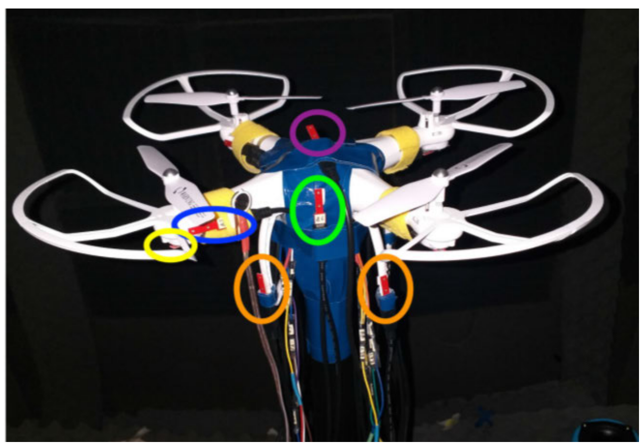

FIGURE 23. Micro drone with 15 MEMS microphones. Copyright ©2022 IEEE [55].

Although the microphones were mounted on the drone, the experimental conditions were not highly realistic, as the drone was stationary and held in place by a stand, rather than being operated in a free-flying or dynamic environment. This work aligns with the only array configuration that does not have a defined geometric shape, classified as "Other" in Figure 5.

Finally, a unique kiteplane system is presented in Kumon et al. [24], which enables longer and higher flights, thereby extending operational time for search-and-rescue missions. Its stable gliding motion significantly reduces ego-noise from rotors and airflow, which often complicates accurate sound source localisation in conventional drones. The navigation system alternates between rotor-driven flight and gliding to minimise ego-noise while maintaining altitude, as presented in Figure 24. While kiteplanes may be less practical than multi-rotors in terms of manoeuvrability and ease of deployment, their extended flight times and reduced rotor noise underscore the potential benefits of alternative UAV designs. Future research may benefit from vertical take-off and landing (VTOL) platforms that combine the agility of multi-rotors with the noise reduction and efficiency advantages of gliding flight.

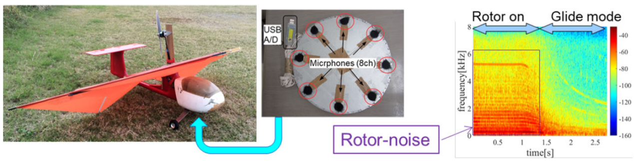

FIGURE 24. Kiteplane and Power Spectrogram in driving and gliding mode. Copyright ©2021 IEEE [24].

E. WHAT ARE THE SOLUTIONS EMPLOYED FOR PERFORMING THE SSL TASK?

This section provides an analysis of the most popular SSL algorithms. Their advantages and disadvantages are briefly covered, as well as their relation to other design choices, such as the number of microphones and the ego-noise mitigation strategy.

1) PREDOMINANT ALGORITHMS FOR SSL: BENEFITS AND APPLICATION-SPECIFIC CONSIDERATIONS

SSL approaches can be broadly classified into classic and AI-based methods [17]. Figure 25 shows the count of the most commonly used baselines among the classic algorithms. The most common are variations of MUltiple SIgnal Classification (MUSIC) [64]. It is widely used in applications with low signal-to-noise ratio (SNR) and multiple sources, due to its ability to estimate the noise from the observed signal and separate the signal and noise subspaces. Generalized eigenvalue decomposition MUSIC (GEVD-MUSIC) [65], relies on a previously recorded sound to estimate the noise correlation matrix, thus improving noise reduction. Generalized singular value decomposition MUSIC (GSVD-MUSIC) [66] is a later variation that improves the computational cost of the generalized eigenvalue decomposition. Incremental generalized singular value decomposition MUSIC (iGSVD-MUSIC) [67] is designed for real-time performance and adaptively estimates the noise correlation matrix from the observed signal, making it suitable for changing noise profiles.

The generalized cross-correlation with phase transform (GCC-PHAT) [68] is a well-established technique for estimating the time difference of arrival (TDOA) between pairs of microphones. By leveraging fast Fourier transforms (FFT) for correlation, it achieves computational efficiency and offers good performance in reverberant environments, such as indoors. However, it requires post-processing of the pairwise estimations for dealing with multiple microphones and is generally less immune to noise than MUSIC-based approaches. The steered-response power with phase transform (SRP-PHAT) [69] technique can be considered as a summation of GCC-PHAT calculations for each microphone pair in the array. It is reported to have similar, but generally lower localisation performance when compared to the MUSIC-based approaches [19], [22]. Delay-and-sum (DS), also known as classical beamforming is another popular and, due to its simplicity, hardware-friendly SSL technique. DS beamforming works by shifting in time and summing the signals received by each microphone in an array. This focuses the array's sensitivity toward a particular direction. Assuming a single source, the DOA can be estimated as the direction with the highest power.

Among AI-based approaches, an artificial neural network (ANN) is used for signal filtering in five of the reviewed works [29], [45], [47], [56], [57]. An end-to-end SSL solution based on an ANN is found in only three of the selected works [7], [39], [51]. It is worth noting that these end-to-end approaches are applied to a single sound source and to eight or fewer microphones, as per Table 3.

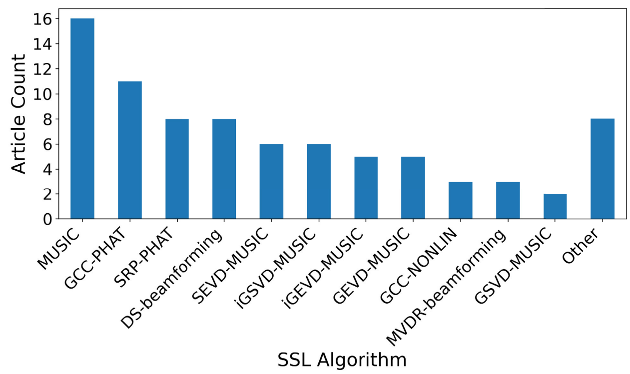

FIGURE 25. SSL algorithms implemented in the selected works.

2) ON-DEVICE VERSUS REMOTE EXECUTION AND REAL-TIME FEASIBILITY

As presented in Figure 26 and Table 3, most of the analysed studies opt for working in a simulated environment, where the ego-noise of the drone can be easily controlled in terms of both location and the SNR. For example, the widely used DREGON dataset also includes noise-free recordings that can be combined with ego-noise to create simulated noisy environments with different SNR levels [21]. This dataset also includes recordings of a stationary UAV with individual rotors turning at different speeds, which can be used to analyse the characteristics of rotor noise and generate recordings with synthetic noise.

Another approach is to record audio offline and test the algorithms on the recording. However, few of these works also report the execution time of these algorithms on embedded computers. This can become a significant limitation for future works aiming at implementing the proposed algorithms on autonomous drones. Some algorithms, such as MUSIC-based methods, tend to be more computationally intensive, making them unsuitable for real-time implementation on resource-constrained embedded platforms. For instance, Hoshiba et al. [33] find that while the iGSVD-MUSIC algorithm has high noise robustness, its high computational cost is a drawback for real-time applicability.

Transmitting the audio data to a ground station for processing allows for the use of more powerful algorithms and can improve the accuracy of the localisation. A drawback of this approach is the requirement of power for high speed transmission. A noteworthy approach is presented by Wakabayashi et al. [42], implementing a local extraction of the MUSIC spectrum, which is in turn transmitted to the ground station for further processing and localisation.

Finally, only three of the reviewed studies implement real-time onboard SSL [2], [23], [49]. Real-time onboard SSL is appealing for drone applications that require quick response, larger area cover and autonomous operation, not limited by line of sight or telemetry range. This gap underscores the need for higher-efficiency algorithms, as discussed in Section V. Zigelman et al. [49] propose a biomimetic miniature drone designed for real-time audio-based short-range tracking. The drone uses two microphones for angle estimation of bat chirping sounds and a microcontroller to perform onboard audio processing. Sibanyoni et al. [2] present a real-time 2D acoustic source localisation system designed for drones in search-and-rescue missions, focusing on safety whistle sounds. The system uses a TDOA algorithm for azimuth estimation, combined with triangulation from other drones to calculate distance. It incorporates four omnidirectional microphones arranged perpendicularly on a single drone and an STM32F4-discovery evaluation board equipped with a Cortex-M4 processor for real-time signal processing. Finally, Basiri et al. [23] propose an embedded system to locate other drones in real time. An Atmel AVR32 microcontroller is used on the drones for sound acquisition, processing, and localisation. However, the location system relies on active acoustic chirps, generated by onboard piezo-electric transducers. For instance, a simultaneous localisation of two drones is demonstrated, where one drone is producing upsweep chirps of 1.7 kHz to 4.7 kHz, and the other one produces downsweep chirps in the same frequency range.

3) EGO-NOISE MITIGATION STRATEGIES

A primary challenge for drone auditions is the presence of ego-noise, which can have a significant impact on localisation performance. Ego-noise in drones is primarily generated by the rotors and propellers, and is typically much louder than

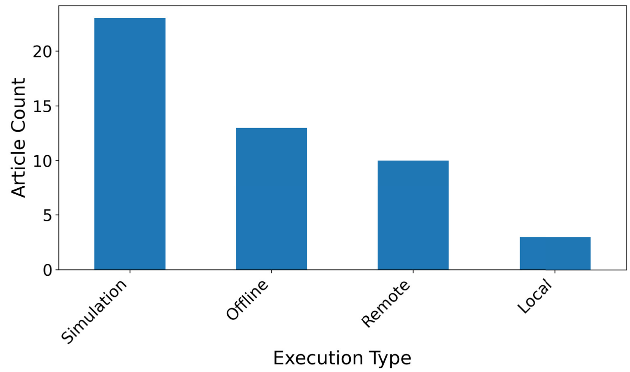

FIGURE 26. Execution of the implemented SSL algorithms.

the external signal. Furthermore, this noise is typically non-stationary, as each motor's speed is continuously modulated to achieve a stable and wind-resistant flight [61]. Several studies have investigated noise reduction techniques to improve the SNR of the audio signal. Some common methods include Wiener filtering [21], spectral subtraction [28] and mix source separation (MSS) [43]. MSS techniques offer the advantage of separating the target sound source from the noise without any prior knowledge of the sources, albeit at a typically higher computational cost.

Other studies propose leveraging the knowledge of the drone's ego-noise characteristics to improve localisation performance. For instance, the rotor speed information is used for speed correlated harmonics cancellation (SCHC) in [7]. Similarly, an SSL method for extremely low SNR conditions (lower than −20 dB) is proposed by Manamperi et al. [55], surpassing the performance of traditional methods like MUSIC, GEVD-MUSIC, and GCC-PHAT in terms of robustness and accuracy. This method employs a cross-correlation-based DOA estimation technique combined with noise angular spectrum subtraction. However, the method is demonstrated on a stationary drone (see Figure 23) and relies on pre-recorded drone-only noise characteristics to suppress self-generated noise.

Deep learning techniques have also been applied to this problem and offer promising results, albeit with a typically increased computational cost. Denoising autoencoders (DAE) are trained to reconstruct clean audio signals from noisy input, effectively reducing noise in Yen and Hioka [45]. A 1D convolutional architecture, DOANet, is used by Qayyum et al. [7] for an end-to-end SSL. DOANet directly estimates the azimuth and elevation angles of a sound source without requiring prior ego-noise reduction or hand-crafted feature extraction.

V. OPEN CHALLENGES

Despite significant recent progress in drone-based SSL, several challenges were identified during this systematic literature review, creating opportunities for future research. This section discusses key issues limiting the current state of the art and outlines potential research directions.

A. PUBLICLY AVAILABLE DATASETS

The low number of publicly available datasets is a significant challenge for drone-based SSL research. While some datasets have been released, they are often limited in terms of realistic applications. DREGON [21] is currently the most widely used dataset for UAV-embedded SSL, but it was recorded indoors and has limitations regarding the diversity of sound sources and noise conditions. It primarily focuses on single-source localisation in a controlled environment, which may not reflect real-world scenarios.

AIRA-UAS is another dataset, proposed by Carranza and Rascon [70], specifically designed for drone-to-drone localisation, utilising 8 microphones mounted on a DJI Matrice 100. The dataset includes audio recordings of two flying drones (3DR Solo and Parrot Bebop 2) at distances of up to 3 meters. While this dataset addresses drone-specific noise challenges, it remains constrained in distance to the sound source and in terms of the variety of environments and sound sources.

Wang et al. [20] provide a dataset for tracking moving sound sources, such as human speakers or emergency whistles, using an 8-microphone circular array mounted on a 3DR IRIS quadcopter. It features multiple scenarios with a loudspeaker as the moving sound source, which travels at a distance of up to 6 meters from the drone. The scenarios vary in noise conditions, including static and dynamic power levels for the drone's propellers.

Finally, many studies using real-world scenarios often do not report experimental parameters, such as the SNR and the distance between the drone and the sound source(s). Similarly, availability of the drone logs including speed, orientation, and location data would facilitate benchmarking of future works.

B. AI-BASED MULTI-SOURCE LOCALISATION

As discussed in Section IV-E, current research on end-to-end SSL approaches based on AI models remains limited, with only three studies focusing on single sound source localisation using convolutional neural networks or other deep learning architectures [7], [39], [51]. Thus, a promising future research direction could include localisation of multiple sound sources, particularly in realistic scenarios of range and noise levels. Additionally, state-of-the-art architectures, such as attention-based neural networks, could be used for localisation of multiple sound sources. The combination of CNNs for local feature extraction and Transformer-based global attention mechanisms has demonstrated promising results in multi-source localisation [71] and should be evaluated for drone-based applications as well.

C. REAL-TIME PROCESSING CONSTRAINTS

Real-time SSL remains a challenge, especially with eight or more microphones. While a larger number of microphones and a larger array aperture can improve performance, it increases the computational burden. Smaller arrays with two or four microphones are currently preferred for real-time processing directly on the UAV, due to onboard processing power and energy limitations. Furthermore, the use of ultrasonic microphones for applications such as electric discharge detection [57] may require higher sampling rates than conventional applications, aimed at human voice or whistle sounds. Thus, efficient implementation of deep learning and sound processing models on specialised hardware, such as FPGAs, is essential for achieving real-time performance for drone-based sound source localisation. Techniques that implement low-power SLL algorithms on FPGAs, such as those demonstrated by da Silva et al. [60], can enable drones to process audio signals at the edge, reducing latency and enhancing system performance.

D. ENERGY EFFICIENCY AND BIO-INSPIRED APPROACHES

The human ability to focus on a specific sound source amidst overlapping signals and noise is known as the "cocktail party problem" [72]. Unlike most SSL algorithms reviewed in this work, humans utilise a combination of interaural time difference (ITD) and interaural level difference (ILD) to localise sounds [73]. Incorporating similar mechanisms in drones, such as combining ITD and ILD measurements, can provide additional cues for localisation in noisy environments or across different frequency ranges. Animals can also improve their sound localisation ability by moving their head, and in some cases also their ears, which creates a different auditory perspective on a sound source [73]. Such approaches, although implemented in other domains [74], remain untested on UAVs.

Finally, neuromorphic approaches offer a promising path towards bio-inspired and energy efficient SSL [75]. These methods would be especially applicable to micro-drones weighing less than 250 g, where energy constraints and payload limitations are more severe. Despite their potential, micro-drones have been underexplored in SSL research, with existing works utilising up to four microphones for real-time applications. Spiking neural networks (SNNs) offer an alternative to traditional artificial neural networks (ANNs) with advantages in energy efficiency when run on neuromorphic hardware. Haghighatshoar and Muir [76] have demonstrated a low-power SNN-based audio source localisation method that uses a Hilbert transform spike encoding scheme, showing promise for energy-efficient implementations on drones.

VI. LIMITATIONS OF THIS REVIEW

Systematic mapping studies are subject to inherent limitations and risks [18]. Below, the primary limitations of this review and the corresponding mitigation strategies are outlined.

Research Question Formulation: The formulation of research questions (RQs) can inadvertently narrow the review scope, leading to the exclusion of relevant studies. To minimise this risk, the RQs were iteratively refined through discussions with co-authors to ensure they comprehensively captured the study goals.

Search Strategy: Despite employing a systematic search strategy, relevant studies may have been missed due to database coverage limitations, language restrictions, or exclusion during the initial selection. This was mitigated by testing and, when necessary, adapting the search strings across the databases. Furthermore, the search string is developed iteratively though initial manual search of relevant literature.

Selection and Publication Bias: The inclusion and exclusion criteria and preference for peer-reviewed studies may introduce bias, favouring published works that report significant results. To address this, objective inclusion and exclusion criteria were established, and efforts were made to ensure that the eligibility of the selected works was assessed by multiple reviewers during the initial reading of the abstracts and the subsequent in-depth analysis. The full list of gathered literature, including duplicated and rejected studies, is also made available.[6]

Data Extraction and Classification Accuracy: Differences in reviewer interpretation can result in misclassification or incomplete data extraction [77]. To mitigate this, double-checks were conducted by multiple reviewers, discrepancies were resolved through consensus, and a standardised extraction template was adhered to.

Generalisability of Findings: The focus on selected works and UAV-based sound source localisation limits generalisability. Results may not fully translate to other acoustic or sensor-based domains. However, the identified trends and gaps provide a foundation for further cross-domain research. It is strongly recommended that future research in UAV-based SSL also consider state-of-the-art developments in sound processing and localisation across other domains. Additionally, while deep neural networks are not commonly found in the selected works, readers may refer to the surveys by Jekateryńczuk and Piotrowski [17] and Grumiaux et al. [16] for an overview of more recent AI-based approaches.

Reporting Variability: Variations in how studies report experimental conditions (e.g., SNR, drone motion, and distances) made direct comparisons challenging. Future works should advocate for standardised metadata reporting, including SNR levels, drone and sound source position logs, and environmental conditions.

While these limitations are intrinsic to systematic reviews, the aim is to provide a comprehensive and unbiased evaluation of the literature within the scope of this review.

VII. CONCLUSION

This paper provides a comprehensive review of the state of the art in UAV-based sound source localisation. Through a systematic mapping study of 49 primary works published between 2014 and 2024, the main configurations, algorithms, and applications relevant to UAV-embedded microphone arrays were analysed. Below is the summary of the findings and recommendations for future research, based on the proposed research questions (RQs):

- Array Configurations (RQ1): Most SSL systems rely on relatively small, planar arrays of 2–8 microphones due to payload and energy constraints on UAVs. Larger arrays (up to 16 or more microphones) generally require either offline processing or remote computation, balancing the trade-off between accuracy and power and weight limitations. Future research should focus on the development of lightweight, power-efficient algorithms and hardware, capable of real-time processing with higher-channel count arrays. This would equip drones for autonomous search-and-rescue missions over large and remote areas.
- Intended Applications (RQ2): The predominant goal in existing works is human speech or whistle localisation for search-and-rescue operations. However, several promising use cases remain underexplored, such as industrial inspection, electrical discharge detection, drone-to-drone tracking and gunshot detection. Additionally, very few works demonstrate SSL at large distances (greater than 100 m). The development of long-range acoustic localisation for search-and-rescue is recommended in order to improve area coverage of such systems.
- UAV Platform Choices (RQ3): Drones are categorised into micro, very small, and small platforms, each demanding different trade-offs in payload, battery life, and stability. While smaller drones benefit from light microphone setups and require efficient, low-complexity algorithms, larger UAVs can carry heavier arrays and handle more computationally demanding methods. A glider-type or VTOL UAVs offer a promising research direction for extended flight times and lower ego-noise.
- SSL Solutions (RQ4): Classical methods like MUSIC, GCC-PHAT, and SRP-PHAT remain the foundation of drone-based SSL. Recently, AI-based approaches have begun to gain traction, although these are still limited to single-source localisation and few microphones. Thus, multi-source localisation and attention-based architectures remain unexplored for UAVs.

REFERENCES

[1] K. Nakadai, M. Kumon, H. G. Okuno, K. Hoshiba, M. Wakabayashi, K. Washizaki, T. Ishiki, D. Gabriel, Y. Bando, T. Morito, R. Kojima, and O. Sugiyama, "Development of microphone-array-embedded UAV for search and rescue task," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2017, pp. 5985–5990.

[2] S. V. Sibanyoni, D. T. Ramotsoela, B. J. Silva, and G. P. Hancke, "A 2-D acoustic source localization system for drones in search and rescue missions," *IEEE Sensors J.*, vol. 19, no. 1, pp. 332–341, Jan. 2019.

[3] A. Deleforge, D. Di Carlo, M. Strauss, R. Serizel, and L. Marcenaro, "Audio-based search and rescue with a drone: Highlights from the IEEE signal processing cup 2019 Student competition [SP competitions]," *IEEE Signal Process. Mag.*, vol. 36, no. 5, pp. 138–144, Sep. 2019.

[6] www.osf.io/dekrs/?view_only=e8ed55e17df94137afb1c07b5ef99155

[4] B. P. A. Rohman, M. B. Andra, H. F. Putra, D. H. Fandiantoro, and M. Nishimoto, "Multisensory surveillance drone for survivor detection and geolocalization in complex post-disaster environment," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2019, pp. 9368–9371.

[5] Y. Yamazaki, C. Premachandra, and C. J. Perea, "Audio-Processing-Based human detection at disaster sites with unmanned aerial vehicle," *IEEE Access*, vol. 8, pp. 101398–101405, 2020.

[6] A. Izquierdo, L. del Val, J. J. Villacorta, W. Zhen, S. Scherer, and Z. Fang, "Feasibility of discriminating UAV propellers noise from distress signals to locate people in enclosed environments using MEMS microphone arrays," *Sensors*, vol. 20, no. 3, p. 597, Jan. 2020.

[7] A. B. A. Qayyum, K. M. N. Hassan, A. Anika, M. F. Shadiq, M. M. Rahman, M. T. Islam, S. A. Imran, S. Hossain, and M. A. Haque, "DOANet: A deep dilated convolutional neural network approach for search and rescue with drone-embedded sound source localization," *EURASIP J. Audio, Speech, Music Process.*, vol. 2020, no. 1, pp. 1–18, Dec. 2020.

[8] A. Banerjee, A. Nilhani, S. Dhabal, and P. Venkateswaran, "A novel sound source localization method using a global-best guided cuckoo search algorithm for drone-based search and rescue operations," in *Unmanned Aerial Systems*. Amsterdam, The Netherlands: Elsevier, 2021, pp. 375–415.

[9] S. D. Correia, J. Fé, S. Tomic, and M. Beko, "Drones as sound sensors for energy-based acoustic tracking on wildfire environments," in *Proc. IFIP Int. Internet Things Conf.* Cham, Switzerland: Springer, Jan. 2022, pp. 109–125.

[10] R. P. Fernandes, J. A. Apolinário, and A. L. L. Ramos, "Bearings-only aerial shooter localization using a microphone array mounted on a drone," in *Proc. IEEE 8th Latin Amer. Symp. Circuits Syst. (LASCAS)*, Feb. 2017, pp. 1–4.

[11] A. L. L. Ramos, J. G. C. Ribeiro, F. G. Serrenho, and J. A. Apolinário, "Effective direction of arrival estimation of gunshot signals from an in-flight unmanned aerial vehicle," *Proc. SPIE*, vol. 10648, pp. 89–99, Apr. 2018.

[12] S. Argentieri, P. Danès, and P. Souères, "A survey on sound source localization in robotics: From binaural to array processing methods," *Comput. Speech Lang.*, vol. 34, no. 1, pp. 87–112, Nov. 2015.

[13] J. Martinez-Carranza and C. Rascon, "A review on auditory perception for unmanned aerial vehicles," *Sensors*, vol. 20, no. 24, p. 7276, Dec. 2020.

[14] A. Joshi, M. M. Rahman, and J.-P. Hickey, "Recent advances in passive acoustic localization methods via aircraft and wake vortex aeroacoustics," *Fluids*, vol. 7, no. 7, p. 218, Jun. 2022.

[15] D. Desai and N. Mehendale, "A review on sound source localization systems," *Arch. Comput. Methods Eng.*, vol. 29, no. 7, pp. 4631–4642, Nov. 2022.

[16] P.-A. Grumiaux, S. Kitić, L. Girin, and A. Guérin, "A survey of sound source localization with deep learning methods," *J. Acoust. Soc. Amer.*, vol. 152, no. 1, pp. 107–151, Jul. 2022.

[17] G. Jekateryńczuk and Z. Piotrowski, "A survey of sound source localization and detection methods and their applications," *Sensors*, vol. 24, no. 1, p. 68, Dec. 2023.

[18] K. Petersen, S. Vakkalanka, and L. Kuzniarz, "Guidelines for conducting systematic mapping studies in software engineering: An update," *Inf. Softw. Technol.*, vol. 64, pp. 1–18, Aug. 2015.

[19] D. Salvati, C. Drioli, G. Ferrin, and G. L. Foresti, "Acoustic source localization from multirotor UAVs," *IEEE Trans. Ind. Electron.*, vol. 67, no. 10, pp. 8618–8628, Oct. 2020.

[20] L. Wang, R. Sanchez-Matilla, and A. Cavallaro, "Tracking a moving sound source from a multi-rotor drone," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2018, pp. 2511–2516.

[21] M. Strauss, P. Mordel, V. Miguet, and A. Deleforge, "DREGON: Dataset and methods for UAV-embedded sound source localization," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2018, pp. 1–8.

[22] L. Wang and A. Cavallaro, "Acoustic sensing from a multi-rotor drone," *IEEE Sensors J.*, vol. 18, no. 11, pp. 4570–4582, Jun. 2018.

[23] M. Basiri, F. Schill, P. Lima, and D. Floreano, "On-board relative bearing estimation for teams of drones using sound," *IEEE Robot. Autom. Lett.*, vol. 1, no. 2, pp. 820–827, Jul. 2016.

[24] M. Kumon, H. G. Okuno, and S. Tajima, "Alternating drive-and-glide flight navigation of a kiteplane for sound source position estimation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2021, pp. 2114–2120.

[25] A. B. A. Abdul Qayyum, A. Anika, Md. M. Monem Miah, Md. M. Rahman, K. M. Naimul Hasan, M. T. Islam, S. A. Imran Shouborno, Md. F. Shadiq, and M. A. Haque, "Direction of arrival estimation through noise supression: A novel approach using GSC beamforming and room acoustic simulation," in *Proc. IEEE Int. Conf. Signal Process., Inf., Commun. Syst. (SPICSCON)*, Nov. 2019, pp. 104–108.

[26] J. Li, F. Tong, Y. Zhou, Y. Yang, and Z. Hu, "Conformal cylindrical array sound source localization at the presence of shadowed elements," *IEEE Internet Things J.*, vol. 10, no. 20, pp. 17694–17704, Jun. 2023.

[27] M. Wakabayashi, H. G. Okuno, and M. Kumon, "Drone audition listening from the sky estimates multiple sound source positions by integrating sound source localization and data association," *Adv. Robot.*, vol. 34, no. 11, pp. 744–755, Jun. 2020.

[28] Y.-J. Go and J.-S. Choi, "An acoustic source localization method using a drone-mounted phased microphone array," *Drones*, vol. 5, no. 3, p. 75, Aug. 2021.

[29] L. Wang and A. Cavallaro, "Deep-learning-assisted sound source localization from a flying drone," *IEEE Sensors J.*, vol. 22, no. 21, pp. 20828–20838, Nov. 2022.

[30] T. Yamada, K. Itoyama, K. Nishida, and K. Nakadai, "Placement planning for sound source tracking in active drone audition," *Drones*, vol. 7, no. 7, p. 405, Jun. 2023.

[31] M. Wakabayashi, H. G. Okuno, and M. Kumon, "Multiple sound source position estimation by drone audition based on data association between sound source localization and identification," *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 782–789, Apr. 2020.

[32] K. Hoshiba, O. Sugiyama, A. Nagamine, R. Kojima, M. Kumon, and K. Nakadai, "Design and assessment of sound source localization system with a UAV-embedded microphone array," *J. Robot. Mechatronics*, vol. 29, no. 1, pp. 154–167, Feb. 2017.

[33] K. Hoshiba, K. Washizaki, M. Wakabayashi, T. Ishiki, M. Kumon, Y. Bando, D. Gabriel, K. Nakadai, and H. Okuno, "Design of UAV-embedded microphone array system for sound source localization in outdoor environments," *Sensors*, vol. 17, no. 11, p. 2535, Nov. 2017.

[34] T. Yamada, K. Itoyama, K. Nishida, and K. Nakadai, "Outdoor evaluation of sound source localization for drone groups using microphone arrays," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Honolulu, HI, USA, Oct. 2022, pp. 9296–9301.

[35] P. Misra, A. A. Kumar, P. Mohapatra, and P. Balamuralidhar, "DroneEARS: Robust acoustic source localization with aerial drones," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 80–85.

[36] D. Salvati, C. Drioli, G. Ferrin, and G. L. Foresti, "Beamforming-based acoustic source localization and enhancement for multirotor UAVs," in *Proc. 26th Eur. Signal Process. Conf. (EUSIPCO)*, Rome, Italy, 2018, pp. 987–991, doi: 10.23919/EUSIPCO.2018.8553514.

[37] T. Yamada, K. Itoyama, K. Nishida, and K. Nakadai, "Sound source tracking using integrated direction likelihood for drones with microphone arrays," in *Proc. IEEE/SICE Int. Symp. Syst. Integr. (SII)*, Jan. 2021, pp. 394–399.

[38] K. Hoshiba, K. Nakadai, M. Kumon, and H. G. Okuno, "Assessment of MUSIC-based noise-robust sound source localization with active frequency range filtering," *J. Robot. Mechatronics*, vol. 30, no. 3, pp. 426–435, Jun. 2018.

[39] J. Choi and J.-H. Chang, "Convolutional neural network-based direction-of-arrival estimation using stereo microphones for drone," in *Proc. Int. Conf. Electron., Inf., Commun. (ICEIC)*, Jan. 2020, pp. 1–5.

[40] M. Clayton, L. Wang, A. McPherson, and A. Cavallaro, "An embedded multichannel sound acquisition system for drone audition," *IEEE Sensors J.*, vol. 23, no. 12, pp. 13377–13386, Jun. 2023.

[41] L. Wang and A. Cavallaro, "Time-frequency processing for sound source localization from a micro aerial vehicle," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2017, pp. 496–500.

[42] M. Wakabayashi, K. Washizaka, K. Hoshiba, K. Nakadai, H. G. Okuno, and M. Kumon, "Design and implementation of real-time visualization of sound source positions by drone audition," in *Proc. IEEE/SICE Int. Symp. Syst. Integr. (SII)*, Jan. 2020, pp. 814–819.

[43] L. Wang and A. Cavallaro, "A blind source separation framework for ego-noise reduction on multi-rotor drones," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 28, pp. 2523–2537, 2020.

[44] Q. Guo, Z. Xin, T. Zhou, and S. Xu, "Off-grid space alternating sparse Bayesian learning," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–10, 2023.

[45] B. Yen and Y. Hioka, "Noise power spectral density scaled snr response estimation with restricted peak search for uav-based sound source localization," *EURASIP J. Audio, Speech, Music Process.*, vol. 2020, no. 1, p. 13, 2020.

[46] K. Yamada, M. Kumon, and T. Furukawa, "Belief-driven control policy of a drone with microphones for multiple sound source search," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Macau, China, 2019, pp. 5326–5332, doi: 10.1109/IROS40897.2019.8968119.

[47] T. Spadini, G. S. Imai Aldeia, G. Barreto, K. Alves, H. Ferreira, R. Suyama, and K. Nose-Filho, "On the application of SEGAN for the attenuation of the ego-noise in the speech sound source localization problem," in *Proc. Workshop Commun. Netw. Power Syst. (WCNPS)*, Oct. 2019, pp. 1–4.

[48] D. Salvati, C. Drioli, and G. L. Foresti, "Spherical harmonic diagonal unloading beamforming with ego-noise reduction for DOA estimation from autonomous systems," *IEEE Trans. Audio, Speech, Language Process.*, vol. 29, no. 10, pp. 216–220, Oct. 2021.

[49] R. Zigelman, O. Eitan, O. Mazar, A. Weiss, and Y. Yovel, "A bio-mimetic miniature drone for real-time audio based short-range tracking," *PLOS Comput. Biol.*, vol. 18, no. 3, Mar. 2022, Art. no. e1009936.

[50] P. Misra, A. A. Kumar, P. Mohapatra, and P. Balamuralidhar, "Aerial drones with location-sensitive ears," *IEEE Commun. Mag.*, vol. 56, no. 7, pp. 154–160, Jul. 2018.

[51] A. Toma, D. Salvati, C. Drioli, and G. L. Foresti, "CNN-based processing of acoustic and radio frequency signals for speaker localization from MAVs," in *Proc. Interspeech*, 2021, pp. 2147–2151.

[52] T. Yamada, K. Itoyama, K. Nishida, and K. Nakadai, "Sound source tracking by drones with microphone arrays," in *Proc. IEEE/SICE Int. Symp. Syst. Integr. (SII)*, Honolulu, HI, USA, Jan. 2020, pp. 796–801.

[53] T. Ohata, K. Nakamura, A. Nagamine, T. Mizumoto, T. Ishizaki, R. Kojima, O. Sugiyama, and K. Nakadai, "Outdoor sound source detection using a quadcopter with microphone array," *J. Robot. Mechatronics*, vol. 29, no. 1, pp. 177–187, Feb. 2017.

[54] A. Satish and A. Medda. (2023). *Acoustic Uav Detection Using Spherical Array Beamforming*. Georgia Tech Res. Inst. [Online]. Available: https://ieeexplore-ieee-org.ezproxy.uws.edu.au/document/10051923

[55] W. Manamperi, T. D. Abhayapala, J. Zhang, and P. N. Samarasinghe, "Drone audition: Sound source localization using on-board microphones," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 30, pp. 508–519, 2022.

[56] O. Sugiyama, S. Uemura, A. Nagamine, R. Kojima, K. Nakamura, and K. Nakadai, "Outdoor acoustic event identification with DNN using a quadrotor-embedded microphone array," *J. Robot. Mechatronics*, vol. 29, no. 1, pp. 188–197, Feb. 2017.

[57] Z. Li, Z. Chen, W. Wu, B. Chen, T. Han, and R. Wang, "Research and implementation of UAV partial discharge detection technology based on acoustic array," in *Proc. 3rd Int. Symp. Comput. Technol. Inf. Sci. (ISCTIS)*, Jul. 2023, pp. 108–113.

[58] S. Guan, R. Jia, L. Qiao, G. Gu, J. Kang, and Y. Song, "MUSIC-based sound source localization algorithm from UVA-embedded microphone array," in *Proc. Int. Conf. Commun., Signal Process., Syst.*, Jan. 2023, pp. 1–8.

[59] S. Azrad, A. Salman, and S. A. R. Al-Haddad, "Performance of doa estimation algorithms for acoustic localization of indoor flying drones using artificial sound source," *J. Aeronaut., Astronaut. Aviation*, vol. 56, no. 1, pp. 469–476, 2024.

[60] B. da Silva, L. Segers, A. Braeken, K. Steenhaut, and A. Touhafi, "Design exploration and performance strategies towards power-efficient FPGA-based architectures for sound source localization," *J. Sensors*, vol. 2019, pp. 1–27, Sep. 2019.

[61] L. Wang and A. Cavallaro, "Microphone-array ego-noise reduction algorithms for auditory micro aerial vehicles," *IEEE Sensors J.*, vol. 17, no. 8, pp. 2447–2455, Apr. 2017.

[62] J. Li, Y. Li, X. Huang, J. Ren, H. Feng, Y. Zhang, and X. Yang, "High-sensitivity gas leak detection sensor based on a compact microphone array," *Measurement*, vol. 174, Apr. 2021, Art. no. 109017.

[63] L. Yongfen, J. Shengchang, and L. Yanming, "Phased-ultrasonic receiving-planar array transducer for partial discharge location in transformer," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 53, no. 3, pp. 614–622, Mar. 2006.

[64] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. AP-34, no. 3, pp. 276–280, Mar. 1986.

[65] K. Nakamura, K. Nakadai, F. Asano, Y. Hasegawa, and H. Tsujino, "Intelligent sound source localization for dynamic environments," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2009, pp. 664–669.

[66] K. Nakamura, K. Nakadai, and G. Ince, "Real-time super-resolution sound source localization for robots," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2012, pp. 694–699.

[67] T. Ohata, K. Nakamura, T. Mizumoto, T. Taiki, and K. Nakadai, "Improvement in outdoor sound source detection using a quadrotor-embedded microphone array," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2014, pp. 1902–1907.

[68] C. Knapp and G. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-24, no. 4, pp. 320–327, Aug. 1976.

[69] J. H. DiBiase, "A high-accuracy, low-latency technique for talker localization reverberant environments using microphone arrays," Ph.D. dissertation, Brown Univ., 2000.

[70] O. Ruiz-Espitia, J. Martinez-Carranza, and C. Rascon, "AIRA-UAS: An evaluation corpus for audio processing in unmanned aerial system," in *Proc. Int. Conf. Unmanned Aircr. Syst. (ICUAS)*, Jun. 2018, pp. 836–845.

[71] D. Zhang, J. Chen, J. Bai, M. Wang, M. S. Ayub, Q. Yan, D. Shi, and W.-S. Gan, "Multiple sound sources localization using sub-band spatial features and attention mechanism," *Circuits, Syst., Signal Process.*, vol. 44, no. 4, pp. 2592–2620, Apr. 2025.

[72] S. Haykin and Z. Chen, "The cocktail party problem," *Neural Comput.*, vol. 17, no. 9, pp. 1875–1902, Sep. 2005.

[73] B. Grothe, M. Pecka, and D. McAlpine, "Mechanisms of sound localization in mammals," *Physiological Rev.*, vol. 90, no. 3, pp. 983–1012, Jul. 2010.

[74] M. Zohourian and R. Martin, "Binaural speaker localization and separation based on a joint ITD/ILD model and head movement tracking," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2016, pp. 430–434.

[75] J. Yang, R. Wang, Y. Ren, J. Mao, Z. Wang, Y. Zhou, and S. Han, "Neuromorphic engineering: From biological to spike-based hardware nervous systems," *Adv. Mater.*, vol. 32, no. 52, Nov. 2020, Art. no. 2003610.

[76] S. Haghighatshoar and D. R. Muir, "Low-power spiking neural network audio source localisation using a Hilbert transform audio event encoding scheme," *Commun. Eng.*, vol. 4, no. 1, p. 18, Feb. 2025.

[77] F. Q. B. da Silva, M. Suassuna, A. C. C. França, A. M. Grubb, T. B. Gouveia, C. V. F. Monteiro, and I. E. dos Santos, "Replication of empirical studies in software engineering research: A systematic mapping study," *Empirical Softw. Eng.*, vol. 19, pp. 501–557, Sep. 2012.

SÉRGIO F. CHEVTCHENKO received the bachelor's degree in electrical engineering and computer science and the M.Sc. degree in applied informatics, and the Ph.D. degree in computer science from the Universidade Federal de Pernambuco (UFPE), in 2023. He is currently a Postdoctoral Research Fellow with the MARCS Institute for Brain, Behaviour and Development. His research interests include neuromorphic engineering, reinforcement learning, and anomaly detection.

BELMAN JAHIR RODRÍGUEZ received the degree in sound engineering from the Universidad de San Buenaventura, Bogotá, and the master's degree in artificial intelligence from the Pontificia Universidad Javeriana de Colombia (PUJ). He is currently pursuing the Ph.D. degree with the MARCS Institute for Brain, Behaviour and Development, Western Sydney University. He is a Research Assistant with the International Centre for Neuromorphic Systems (ICNS). His research interests include audio processing, acoustics, and AI applications in sound engineering.

RAFAELLA VALE received the B.Sc. degree in computer science from the Universidade Federal Rural de Pernambuco (UFRPE), Brazil, in 2017, and the M.Sc. degree from the Universidade Federal de Pernambuco (UFPE), Brazil, in 2020, where she is currently pursuing the Ph.D. degree in computer science with the Centro de Informática. Her research interests include quantum circuits, quantum gate decomposition, and machine learning.

ABISHEK SOTI is currently pursuing the Bachelor of Engineering degree (Hons.) in electrical engineering with Western Sydney University. He is a Research Assistant with the International Centre for Neuromorphic Systems.

YESHWANTH BETHI received the B.Tech. degree in electrical engineering from the Indian Institute of Technology Bombay (IIT Bombay), India, in 2016, and the Ph.D. degree in neuromorphic engineering from Western Sydney University, Australia. He is currently a Postdoctoral Research Fellow with the International Center for Neuromorphic Systems (ICNS), Western Sydney University. His research focuses on developing novel computational architectures inspired by the principles of biological computation and signal processing.

NAQIB IBNUL received the Bachelor of Engineering (Hons) degree in electrical engineering from Western Sydney University. He is a Research Assistant at the International Centre for Neuromorphic Systems, currently working on Tool Condition Monitoring of Industrial Systems and Sound Source Localisation.

ALEXANDRE MARCIREAU received the M.Sc. (Eng.) degree in computer science from École Centrale, Paris, in 2015, and the Ph.D. degree in neuromorphic engineering from Sorbonne Université, Paris, in 2019. He is currently a Postdoctoral Research Fellow in neuromorphic engineering and space applications with the International Center for Neuromorphic Systems (ICNS), Western Sydney University. His research focuses on bioinspired computer vision, event vision sensors, event-based processing, and software development.

MOSTAFA RAHIMI AZGHADI is currently a Professor of electronics and computer engineering with James Cook University, Townsville, Australia. He is a recognized Leader in applied artificial intelligence and neuromorphic computing, with contributions spanning diverse fields, including agriculture, aquaculture, biosecurity, and medical applications. Over his career, he has successfully secured over U.S. $20 million in funding for a wide range of applied and fundamental research projects, reflecting his impactful contributions to advancing technology and research. He served as an Associate Editor for IEEE Access, from 2018 to 2024. He is an Associate Editor of *Frontiers in Neuromorphic Engineering* and *Computers*.

ANDREW WABNITZ received the B.E. degree (Hons.) in electrical engineering from The University of Adelaide, Australia, in 2005, and the Ph.D. degree from The University of Sydney, Australia, in 2013. He has over ten years' experience across industry, academia, and defense, working on projects involving embedded system design software development and FPGA design for applications in biomedical devices and cyber security and space. He is currently a Senior Researcher of cognitive technologies with the Defence Science and Technology Group, Department of Defence, Australia. His primary research interests include neuromorphic computing, bio-inspired algorithm design, and the application of this technology to real-world problems.

SAEED AFSHAR (Member, IEEE) received the B.Eng. degree in electrical engineering from the University of New South Wales, Sydney, NSW, Australia, and the M.Eng. degree in electrical engineering from Western Sydney University, Sydney. He is currently a Senior Lecturer with the International Centre for Neuromorphic Systems, Western Sydney University. His research interests include event-based vision and audio processing for neuromorphic hardware.